Inferring dark matter halo properties for H I-selected galaxies

Tariq Yasin [ORCID],[1]★ Harry Desmond [ORCID],[1,2,3] Julien Devriendt[1] and Adrianne Slyz[1]

[1]*Astrophysics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, UK*
[2]*McWilliams Center for Cosmology, Department of Physics, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA 15213, USA*
[3]*Institute of Cosmology & Gravitation, University of Portsmouth, Dennis Sciama Building, Portsmouth PO1 3FX, UK*

ABSTRACT

We set constraints on the dark matter halo mass and concentration of $\sim$22 000 individual galaxies visible both in H I (from the ALFALFA survey) and optical light (from the Sloan Digital Sky Survey). This is achieved by combining two Bayesian models, one for the H I line width as a function of the stellar and neutral hydrogen mass distributions in a galaxy using kinematic modelling, and the other for the galaxy's total baryonic mass using the technique of inverse subhalo abundance matching. We hence quantify the constraining power on halo properties of spectroscopic and photometric observations, and assess their consistency. We find good agreement between the two sets of posteriors, although there is a sizeable population of low-line width galaxies that favour significantly smaller dynamical masses than expected from abundance matching (especially for cuspy halo profiles). Abundance matching provides significantly more stringent bounds on halo properties than the H I line width, even with a mass–concentration prior included, although combining the two provides a mean gain of 40 per cent for the sample when fitting an NFW profile. We also use our kinematic posteriors to construct a baryonic mass–halo mass relation, which we find to be near power law, and with a somewhat shallower slope than expected from abundance matching. Our method demonstrates the potential of combining photometric and spectroscopic observations to precisely map out the dark matter distribution at the galaxy scale using upcoming H I surveys such as the SKA.

Key words: dark matter – galaxies: formation – galaxies: kinematics and dynamics – galaxies: statistics.

1 INTRODUCTION

In the Lambda cold dark matter (ΛCDM) paradigm, galaxies form from the condensation of baryons in the centre of dark matter (DM) haloes, virialized overdensities seeded by primordial perturbations in the density field at the end of the inflationary epoch (Mo, van den Bosch & White 2010). Elucidating the 'galaxy–halo connection,' the set of correlations between the properties of galaxies and those of their host DM haloes, is a key endeavour in modern astrophysics, vital both to construct theoretical models of galaxy formation within ΛCDM and to enable tests of the model itself that rely on the DM distribution on small or large scales.

The first evidence for DM came from the dynamics of the Coma cluster (Zwicky 1933), and the kinematics of astrophysical objects remains one of our most powerful probes of its distribution. The most detailed modern studies of this type use radio interferometry to measure the rotation curves (RCs) of galaxies, the rotational velocity of the gas as a function of radius (e.g. Walter et al. 2008; Lelli, McGaugh & Schombert 2016). This has allowed detailed comparison between different halo models (Katz et al. 2017; Li et al. 2020), provided insights into the scaling relationships between galaxies (e.g. Desmond et al. 2019; Lelli et al. 2019) and afforded tests of modified gravity (Burrage, Copeland & Millington 2017; Naik et al. 2019; Chae et al. 2021). However, due to the observational expense, even the largest collations of RCs only contain hundreds of galaxies,

with ongoing surveys such as MIGHTEE-H I (Maddox et al. 2021) and APERTIF (Oosterloo, Verheijen & van Cappellen 2010) aiming to extend this to thousands.

In contrast, the width of the global H I 21-cm emission line of a galaxy (henceforth W_{50}) is a readily observable kinematic tracer and has already been derived for tens of thousands of galaxies. The current state of the art is the ALFALFA survey (Haynes et al. 2018), which has produced a catalogue of W_{50} for $\sim$30 000 galaxies out to $z \sim 0.06$ using a blind drift scan with the ARECIBO telescope. Future single dish and interferometric surveys such as WALLABY (5×10^5 galaxies; Koribalski et al. 2020) and the SKA ($\sim$10^9 galaxies out to $z \sim 2$; Yahya et al. 2015) will increase the number of W_{50} observations by orders of magnitudes while extending to significantly higher redshift.

The drawback of using W_{50} rather than resolved RCs to localize DM is that it lacks spatial information within the galaxy, instead providing a weighted average across the RC. In particular, W_{50} is dependant on both the spatial distribution of H I gas in the galaxy and the velocity profile that traces the total gravitational potential. If the H I gas disc extends into the flat part of the RC with speed V_{flat}, the H I spectral profile takes the form of the classic symmetric two-horned shape, due to flux building up at the wavelengths equivalent to the Doppler shifting of the H I line by $\pm V_{\text{flat}}$. In this case, the inclination-corrected line width $W_{50}/\sin i$ is approximately $2V_{\text{flat}}$, which is typically dominated by the gravity of the DM halo. However, depending on the H I distribution, W_{50} can either be weighted towards the inner parts of the galaxy where the baryonic contribution is important, causing it to be closer to twice the peak velocity V_{max}, or

★ E-mail: tariq.yasin@physics.ox.ac.uk

to lower values if the rotation curve is still rising or observed close to edge-on, causing a single peaked profile. W_{50} can also be affected by morphological asymmetries caused by environmental interactions or internal processes (Bok et al. 2019; Reynolds et al. 2020; Watts et al. 2020). Despite these complexities W_{50} contains a large amount of potential information about the dynamical mass of a galaxy. Lelli et al. (2019) find that for the SPARC sample W_{50} from archival single-dish data give a tighter baryonic Tully–Fisher relation (BTFR) than all velocity summary statistics (V_{max}, V_{2R_e}, $V_{2.2}$) other than V_{flat}.

In this study, we extract constraints on halo properties for individual galaxies with W_{50} measurements. To do this, we develop a Bayesian forward-modelling framework: we first construct models for the full H I line profile of a galaxy for given baryonic and DM mass distribution and then compress this to the W_{50} statistic. Unlike models that simply equate W_{50} to some summary statistic of the rotation curve (typically V_{flat} or V_{max}), possibly with empirical corrections for the effects described above, our method naturally takes full account of the H I distribution and shape of the rotation curve based on the available observational data. We test the accuracy of our method, which uses survey data and empirical relationships to construct the distribution of baryons, using the resolved spectroscopy and photometry of the SPARC data set.

The properties of DM haloes can also be inferred from photometric data using empirical models that connect the properties of haloes in dark matter-only (DMO) cosmological simulations to large-scale galaxy surveys. These models assign galaxy properties to simulated haloes by making simple parametrized assumptions, which are then tested and the parameters constrained using observations such as galaxy clustering (see Wechsler & Tinker 2018 for a review). This avoids invoking a full galaxy formation model as in hydrodynamical simulations or semi-analytical models, the complex baryonic physics of which are not yet well understood. We will use a particular empirical model called subhalo abundance matching (SHAM) to provide further constraints on the DM distribution of our observational sample, and to test for consistency between inferences made from photometric and spectroscopic data.

SHAM assigns galaxy masses based on the virial mass or rotation velocity of their halo, with some scatter to allow for stochasticity in the galaxy–halo connection (Kravtsov et al. 2004; Conroy, Wechsler & Kravtsov 2006; Behroozi, Conroy & Wechsler 2010; Guo et al. 2010; Moster et al. 2010). The ALFALFA survey has revealed that H I-selected samples are much more weakly clustered than optically selected (Li et al. 2012; Martin et al. 2012; Papastergis et al. 2013), and recent work has shown that reproducing their clustering signal requires performing SHAM on a subset of the simulated halo population selected on properties that have a similar clustering bias, such as formation time (Guo et al. 2017; Stiskalek et al. 2021).

SHAM, along with the similar empirical technique of Halo Occupation Distribution modelling, are commonly combined with the halo model or DMO simulations to construct mock catalogues for the massive volumes of the sky probed by cosmological surveys. They can also be applied in the inverse direction: using photometric observations of galaxies to map out the DM distribution (e.g. Desmond et al. 2018). Therefore, it is vital that these empirical methods are validated by independent tests, especially at low masses where clustering constraints become weak or unavailable.

In this work, we use the SHAM prescription of Stiskalek et al. (2021, henceforth ST21), which is specifically tailored to the H I-selected galaxies of the ALFALFA sample, to construct an inverse SHAM method that produces a Bayesian posterior on the halo properties of a galaxy given its baryonic mass as observational input. Comparing the independent constraints on halo properties

from H I kinematics and SHAM will allow us to answer the following questions:

(i) What is the DM content of H I-selected galaxies and how does this correlate with galaxy variables at both the population and individual galaxy levels?

(ii) To what extent are the halo properties implied by kinematics and photometry consistent?

(iii) How much information on the DM distribution is contained in cheaper photometric observations of a galaxy, versus the more expensive W_{50}?

(iv) Does combining kinematic and abundance matching posteriors give us tighter constraints on halo properties than either one alone?

The paper is structured as follows. Section 2 describes the observed and simulated data we use. Section 3 describes our Bayesian models and inference for W_{50} and SHAM, and the metrics we use to characterize the posteriors of individual galaxies. In Section 4, we present the constraints on halo properties obtained by the two models for our sample of H I-selected galaxies, their summary metrics and their correlation with galaxy properties. We expand on the implications of our results, discuss potential systematic errors and compare to the literature in Section 5, and conclude in Section 6. We assume $H_0 = 70$ km s^{-1}, and all logarithms are base-10.

2 OBSERVED AND SIMULATED DATA

2.1 ALFALFA

We take H I line widths, H I masses, and distances from the AL-FALFA[1] (Haynes et al. 2018) data set. This contains H I line sources in 7000 deg^2 of high Galactic latitude out to a redshift of 0.06 generated from a blind survey using the Arecibo telescope. The complete α.100 catalogue contains $\sim$31 500 extragalactic sources with H I masses ranging from 10^6 to 10^{11} M$_\odot$. We use the W_{50} line width measurement, which is the velocity width measured at 50 per cent of the peak flux. This is calculated by identifying the two peaks of the classic twin horned H I line profile, then fitting polynomials to the data between the peak and zero flux on each side. W_{50} is corrected for instrumental broadening, as described in Springob et al. (2005), but not for turbulent motion, disc inclination or cosmological effects. The catalogue also contains W_{20}, the velocity width measured at 20 per cent of the peak flux. Theoretically, this is expected to be a better tracer of the DM-dominated outskirts of the galaxy. However, Haynes et al. (2018) find the measurement less robust, and its errors harder to quantify, and so only an error on W_{50} is provided.

The catalogue also contains distances. For objects with $cz > 6000$ km s^{-1} this is cz, where z is the redshift in the CMB frame measured from the centroid of the H I line. For objects closer than this, a distance is calculated using the local peculiar velocity model of Masters (2005). Primary distances are used if they are available in the literature. The H I mass $M_{H I}$ is provided, calculated from the total integrated H I flux and the assumed distance. A code indicates the reliability of each detection based on the signal-to-noise ratio (SNR) and other observational criteria. Code 1 sources (of which there are 25 434) are the highest quality observations, while code 2 sources may have lower SNR, but have been successfully cross-matched with

[1]http://egg.astro.cornell.edu/alfalfa/data/index.php

known optical counterparts at the same redshift as the H I line. We use both in our fiducial analysis.

Yu et al. (2022) perform a reanalysis of the ALFALFA raw spectra, including calculating a different, integrated definition of the line width using the 'curve of growth' method of Yu, Ho & Wang (2020). Their measurement (henceforth W_{Yu85}) is defined as the velocity width enclosing 85 per cent of the total flux. We use W_{Yu85} in a non-fiducial model to test for possible systematics due to the ALFALFA data reduction.

2.2 NASA-Sloan Atlas

We use the NASA-Sloan Atlas (NSA) to obtain stellar masses and optical axis ratios (used to derive inclinations) for the ALFALFA galaxies. The NSA consists of images and parameters of local galaxies derived from Sloan Digital Sky Survey (SDSS) imaging data, with the addition of *Galaxy Evolution Explorer* (*GALEX*; Martin et al. 2005) data. We use NSA v1_0_1[2] which is based on SDSS DR13 (Albareti et al. 2017) and contains ∼640 000 galaxies out to redshift $z = 0.15$. The image analysis pipeline utilizes enhanced object detection, deblending and other improvements over standard SDSS processing, which improves its performance for larger and brighter galaxies (Blanton et al. 2011). The catalogue contains both elliptical Petrosian and Sérsic aperture photometry fits, with the former considered more reliable. The fits are K-corrected to $z = 0.0$ using the `kcorrect code v4_2` (Blanton & Roweis 2007), which also estimates the stellar mass. A Chabrier initial mass function (IMF) is assumed and spectral energy distributions (SEDs) are fitted to the broad-band optical SDSS fluxes as well as ultraviolet fluxes from *GALEX* when available.

To cross-match the NSA catalogue with ALFALFA, we follow ST21 by requiring a 5 arcsec skymatch between NSA galaxies and the ALFALFA × SDSS optical counterparts of Durbala et al. (2020), and a maximum line-of-sight distance of 10 Mpc. These criteria should yield a low number of mismatches. There are 21 776 galaxies in our final ALFALFA×NSA catalogue. We note that the ALFALFA × SDSS catalogue of Durbala et al. (2020) contains estimates of stellar mass for ∼30 000 galaxies, but we prefer to use NSA's improved image analysis pipeline, which also improves consistency with ST21.

2.3 SPARC

SPARC[3] (Lelli et al. 2016) is a data base of 175 late-type galaxies for which both high quality H I rotation curves and near-infrared *Spitzer* photometry are available. We use SPARC's publicly available resolved rotation curves, detailed mass models and W_{50} measurements, as well as the H I surface density as a function of radius (F. Lelli, private communication) in our analysis. The high-resolution SPARC data are used as a truth against which we test the approximations that allow us to model the much larger ALFALFA data set.

The sample spans a large range in luminosity (10^7 to 10^{12} $L_\odot$), surface brightness (∼5 to ∼5000 $L_\odot$ pc^{-2}), gas mass (∼10^7 to $10^{10.6}$ $M_\odot$), and morphology (S0 to Im/BCD). The advantage of *Spitzer* photometry is that at 3.6 μm the mass-to-light ratio of galaxies is relatively constant, which helps to break the degeneracy between the velocity contribution from the stars and the DM. We perform a cross-match between ALFALFA and SPARC using the same procedure as for the NSA, which yields 45 matches.

Lelli et al. (2019) compile H I line widths from various sources in the literature and for various definitions for the SPARC galaxies. The closest line width definition they include to ALFALFA's W_{50} is $W_{50\text{Mc}}$: the width at 50 per cent flux of the mean flux, where the mean flux is taken across the whole line width, with corrections for instrumental resolution and relativistic broadening. As we find the two measures are very similar for galaxies that have both, we test our W_{50} model using the 125 galaxies in SPARC for which there is either an ALFALFA W_{50} or a SPARC $W_{50\text{Mc}}$.

2.4 Simulation data

As input to our SHAM model we use the 140 Mpc h^{-1} 'Shin-Uchuu' box of the Uchuu suite of cosmological *N*-body simulations[4] (Ishiyama et al. 2021). Shin-Uchuu contains 6400^3 particles, with a particle mass of 8.97×10^5 $M_\odot$ h^{-1} and force softening length of 0.4 h^{-1} kpc. The simulation uses the GreeM *N*-body code (Ishiyama, Fukushige & Makino 2009; Ishiyama, Nitadori & Makino 2012) and the 2018 Planck flat ΛCDM cosmology (Planck Collaboration VI 2020): $H_0 = 67.74$ km s^{-1} Mpc^{-1}; $\Omega_0 = 0.3089$; $\lambda_0 = 0.6911$, scalar spectral index $n_s = 0.9667$; root-mean-square matter fluctuation on 8 Mpc h^{-1} scales $\sigma_8 = 0.8159$. Haloes and subhaloes were identified using the Rockstar halo finder (Behroozi, Wechsler & Wu 2013a) and the Consistent Trees Merger Tree Code (Behroozi et al. 2013b). The halo finder calculates the halo mass in a profile-independent manner by simply summing the mass of the particles belonging to each halo using the overdensity condition $\Delta_{\text{vir}} = 178$ of Bryan & Norman (1998). We use the trimmed Shin-Uchuu catalogue of haloes with $M_{\text{vir}} > 10^9$ $M_\odot$, so each halo has a minimum of 1100 particles and is therefore well resolved. We use the same overdensity condition to define virial quantities for the haloes in our kinematic model.

3 METHODOLOGY

3.1 Overview and verification with SPARC

For each galaxy in our sample, we have two separate Bayesian models: the kinematic model, for which the observable is the ALFALFA W_{50}, and the SHAM model, for which the observable is the galaxy's baryonic mass M_{bar}. Our kinematic model also contains nuisance parameters describing the galaxy, the priors of which are informed by the NSA and ALFALFA data. The free parameters and their priors for each model are listed in Table 1.

DMO simulations (on which our SHAM model is based) implicitly assume that baryons and DM behave identically and hence are perfectly mixed. The virial halo masses listed in their halo catalogues (which we label M_{vir}) are therefore different to the masses inferred from fitting parametrized halo profiles to kinematic data (which we label M_{halo}) according to $M_{\text{vir}} = M_{\text{halo}} + M_{\text{bar}}$, where M_{bar} is the mass of *observed* stars and cold gas. M_{halo} therefore includes DM as well as any unobserved baryons, which are inferred indiscriminately in the kinematic analysis. We assume that the sum of these components follows a standard density profile (Section 3.2.1), e.g. because they have the same distribution. In reality, the distribution of hot gas around galaxies and its association with the DM is poorly known. Hafen et al. (2019) found that f_b can vary substantially between galaxies, as part of the circumgalactic medium can be accreted from the intergalactic medium and other galaxies. However, this potential systematic uncertainty in M_{vir} of up to ∼ 15 per cent (for a halo completely stripped

---

[2]https://www.sdss.org/dr13/manga/manga-target-selection/nsa/
[3]http://astroweb.cwru.edu/SPARC/

[4]http://skiesanduniverses.org/Simulations/Uchuu/

Table 1. The free parameters in our abundance matching and kinematic models, their physical definitions and their Bayesian priors. All parameters are sampled in logarithmic space except inclination and distance.

	Parameter	Units	Definition	Prior
AM	$M_{\rm vir}$	$M_\odot$	Virial mass	2D halo probability density from Uchuu simulation
Kinematics	$c_{0.1}$ $M_{\rm vir}$ $c_{0.1}$	$M_\odot$	Custom halo concentration, defined in equation (11) Virial mass $M_{\rm vir} = M_{\rm halo} + M_{\rm bar}$ (see Section 3.1) Custom halo concentration, defined in equation (11)	Flat in range $\log(M_{\rm bar}/M_\odot) < \log(M_{\rm vir}/M_\odot) < 15.5$ Flat in range $0.5 < \log c_{0.1} < 2$, or halo mass–concentration relationship
	$M_{\rm H\,I}$	$M_\odot$	H I mass	Gaussian prior from ALFALFA observed value and its uncertainty
	M_*	$M_\odot$	Stellar mass	Gaussian prior from NSA observed value with adopted 0.2 dex uncertainty
	$r_{\rm H\,I}$	kpc	Scale length of exponential gas disc	Gaussian prior from Wang et al. (2016) mass–size relationship
	$r_{\rm disc}$	kpc	Scale length of exponential stellar disc	Gaussian prior from empirical relationship based on M_* (Dutton et al. 2011)
	$r_{\rm bulge}$	kpc	Half light radius of stellar bulge	Gaussian prior from empirical relationship based on M_* (Dutton et al. 2011)
	D	Mpc	Physical distance to galaxy	Gaussian prior from ALFALFA value and its uncertainty
	i	deg	Inclination ($0°$ face on; $90°$ edge on)	Equation (13) with NSA $b/a \pm 10$ per cent, and a flat prior on q in the range 0.15 to b/a.

of baryons) is still small compared to the typical uncertainties on our derived halo masses. We do not modify halo concentrations to account for the fact that a fraction of $M_{\rm halo}$ is in baryons (see also Section 5.2.1).

In our kinematic model, we sample $\log M_{\rm vir} = \log(M_{\rm halo} + M_{\rm bar})$ rather than $M_{\rm halo}$ itself, setting the lower bound on its flat prior to be $\log M_{\rm bar}$. We find some galaxies to have non-zero posterior probability at $M_{\rm halo} = 0$, so this parametrization aids sampling by raising the lower limit of the posterior to a finite value. We could perform the SHAM inference with the same parametrization as the kinematics, but it is unnecessary. Our SHAM posterior does not approach $M_{\rm halo} = 0$, so we can simply convert to or from $M_{\rm vir}$ in post-processing.

In Fig. 1, we show relationship between our observed and simulated input quantities and our Bayesian models, as well as the verification of the W_{50} model using the SPARC data. In Section 3.2, we describe our kinematic model and its verification using SPARC, and in Section 3.3 our SHAM model.

3.2 Kinematic model

The kinematic model puts constraints on halo parameters by combining a parametrized form for the halo density profile with a baryonic mass distribution to forward-model W_{50} for comparison with the ALFALFA data. The forward-model consists of three steps:

(1) Calculate the model rotation curve $V_c(r)$ and H I surface density $\Sigma_{\rm HI}(r)$ for the galaxy for a given set of model free parameters.

(2) Construct the model H I line profile from $\Sigma_{\rm HI}(r)$ and $V_c(r)$.

(3) Calculate W_{50} from the line profile and compare to the observed value.

As SPARC has mass models for the H I and stars of each galaxy, as well as the observed RC, we can use it to carry out checks on our model:

(i) Are the observed W_{50} of real galaxies well modelled as a product of their azimuthally averaged RC and H I surface density?

(ii) How much do our simple mass models based on empirical relations and non-resolved observations bias line widths compared to the detailed mass models in SPARC?

(iii) How accurate are optical inclinations from SDSS photometry compared to those derived from tilted-ring fits to spatially resolved velocity fields?

The H I spectrum is a spatially integrated quantity, so to calculate W_{50} we must first model the full RC. We assume the rotational speed is equal to the circular speed (we discuss the validity of this in Section 5.2.2). The total circular speed, $V_c(r)$, is the sum in quadrature of the circular speed due to each of the galaxy's components (DM, stellar bulge, stellar disc, gas disc),

$$V_c^2(r) = V_{\rm DM}|V_{\rm DM}| + V_{\rm bulge}|V_{\rm bulge}| + V_{\rm disc}|V_{\rm disc}| + V_{\rm gas}|V_{\rm gas}|, \quad (1)$$

where V_x is also a function of r. To construct the H I spectrum from $V_c(r)$ and $\Sigma_{\rm HI}(r)$, we follow the method of Obreschkow et al. (2009), by considering the gas disc as a series of flat rings each with a constant circular speed V_c. The flux of the H I spectrum at a given wavelength λ corresponds to H I gas with radial velocity V_λ, which we define relative to the galaxy's systemic velocity so the mid-point of the spectrum is at $V_\lambda = 0$. The flux at wavelength λ due to a single, infinitely thin ring of gas is given by

$$\tilde{\psi}(V_\lambda, V_c) = \begin{cases} \frac{1}{\pi\sqrt{V_c^2 - V_\lambda^2}} & \text{if } |V_\lambda| < V_c \\ 0, & \text{otherwise .} \end{cases} \quad (2)$$

The singularity is smoothed by introducing a velocity dispersion $\sigma_{\rm HI}$ for the gas, which models its random motion. We fix the dispersion to be 10 km s^{-1}, based on observations of nearby galaxies (Leroy et al. 2008; Mogotsi et al. 2016). This broadens the flux from the ring:

$$\psi(V_\lambda, V_c) = \frac{\sigma_{\rm HI}^{-1}}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} dV \exp\left[\frac{(V_\lambda - V)^2}{-2\sigma_{\rm HI}^2}\right] \tilde{\psi}(V, V_c). \quad (3)$$

The total flux observed at λ is then obtained by integrating over the galaxy:

$$\Psi_{\rm HI}(V_\lambda) = \frac{2\pi}{M_{\rm HI}} \int_0^\infty dr\, r\, \Sigma_{\rm HI}(r) \psi(V_\lambda, V_c(r)). \quad (4)$$

As our model profiles are symmetrical, W_{50} is calculated by identifying the peak flux and then finding the outermost point where the flux is 50 per cent of the maximum.

To check the accuracy of modelling W_{50} in this way, we calculate a 'SPARC model W_{50}' by applying equation (4) to $V_{c,\rm obs}$, the observed
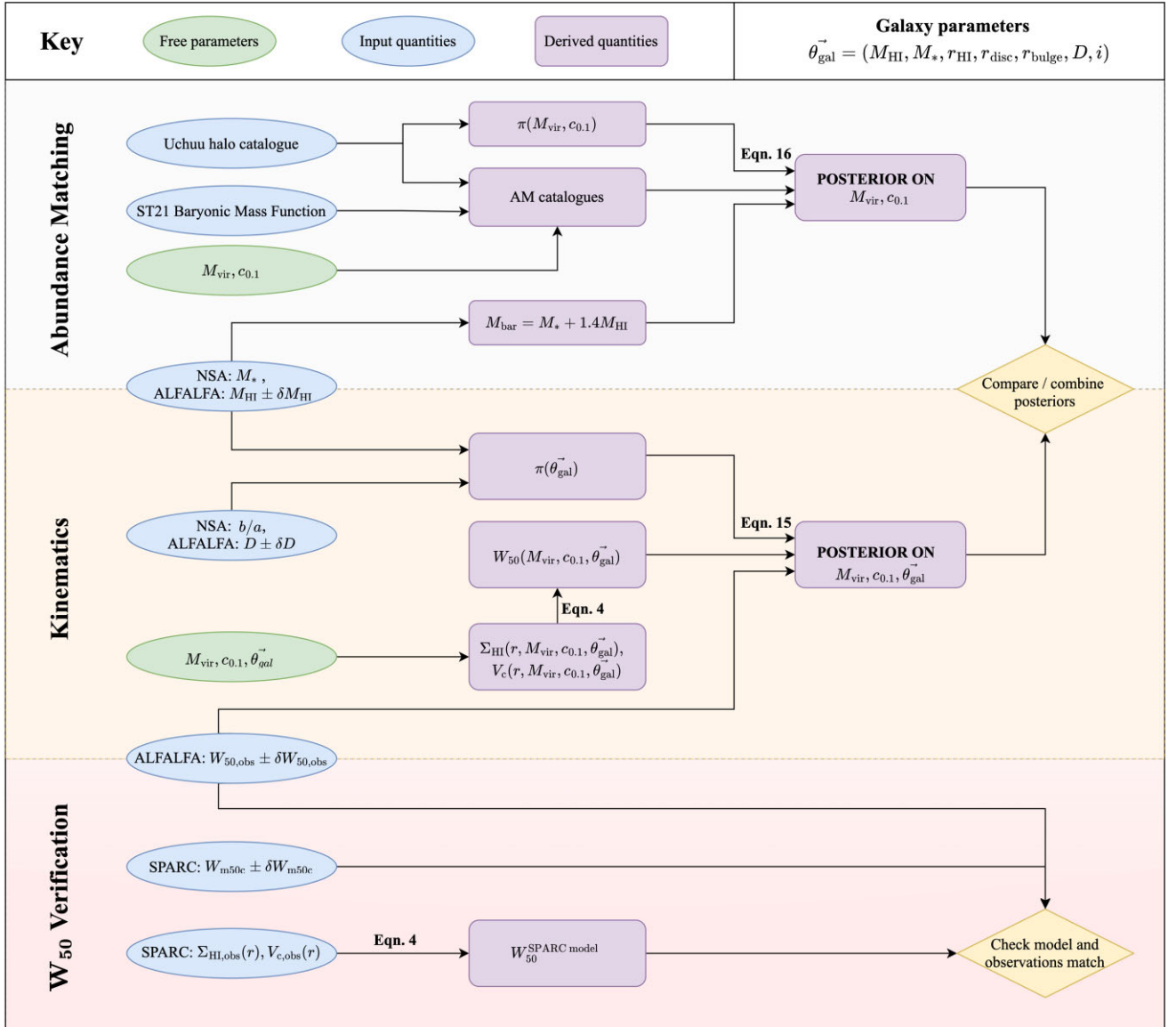
Figure 1. Schematic of our workflow for constraining halo mass and concentration from abundance matching and kinematics. The kinematics observable is the ALFALFA W_{50}. Observations of galaxy parameters from the NSA and ALFALFA inform the priors on the free parameters in the kinematic inference. The observed $M_{\rm H\,I}$ and M_* are also used to calculate the galaxy's baryonic mass $M_{\rm bar}$, which is the observable in the SHAM model. The resolved rotation curves ($V_{\rm c,obs}$) and H I surface density observations ($\Sigma_{\rm H\,I,obs}$) from the SPARC data set are used to verify that the model for calculating W_{50} (equation 4) can match observations.

inclination-corrected RC, and $\Sigma_{\rm H\,I,obs}$, the observed H I surface density profile, of the SPARC galaxies. For some SPARC galaxies, the H I surface density observations extend beyond the final data point of the RC. We extrapolate the RC as $V_{\rm flat}$ if it is defined for the galaxy (Lelli, McGaugh & Schombert 2015), or linearly extrapolate the RC if it is not. We show this procedure for a single galaxy in Fig. 2. To approximate uncertainties on the SPARC model W_{50}, we use the maximum and minimum W_{50} generated by combinations of the observational errors on the RC and $V_{\rm flat}$, and extrapolating linearly or with $V_{\rm flat}$.

In Fig. 3, we compare the SPARC model W_{50} to the observed W_{50} for the 125 galaxies in the SPARC data set with either an ALFALFA W_{50} or the very similar SPARC $W_{\rm 50Mc}$ (described in Section 2.3). Although the uncertainties on the SPARC model W_{50} are crude, there are only three galaxies for which there is $>3\sigma$

tension between the model and observed value. It is interesting that the model W_{50} on average slightly underpredicts the observed value across a large range, suggesting it cannot be due to a constant effect such as instrumental broadening. A possible cause is non-circular motions such as outflows. We conclude that the model works well for $W_{50} \gtrsim 200\,{\rm km\,s^{-1}}$. The scatter between model and observations increases at lower W_{50}, but this can be explained by the increased uncertainty in the extrapolation of the RC. To apply the W_{50} model to the ALFALFA galaxies, we now need to construct model $\Sigma_{\rm H\,I}(r)$ and $V_{\rm c}(r)$ for the ALFALFA sample.

3.2.1 DM distribution

The circular speed contribution from the DM is calculated by assuming a halo profile, with the halo mass $M_{\rm halo}$ and its concen-
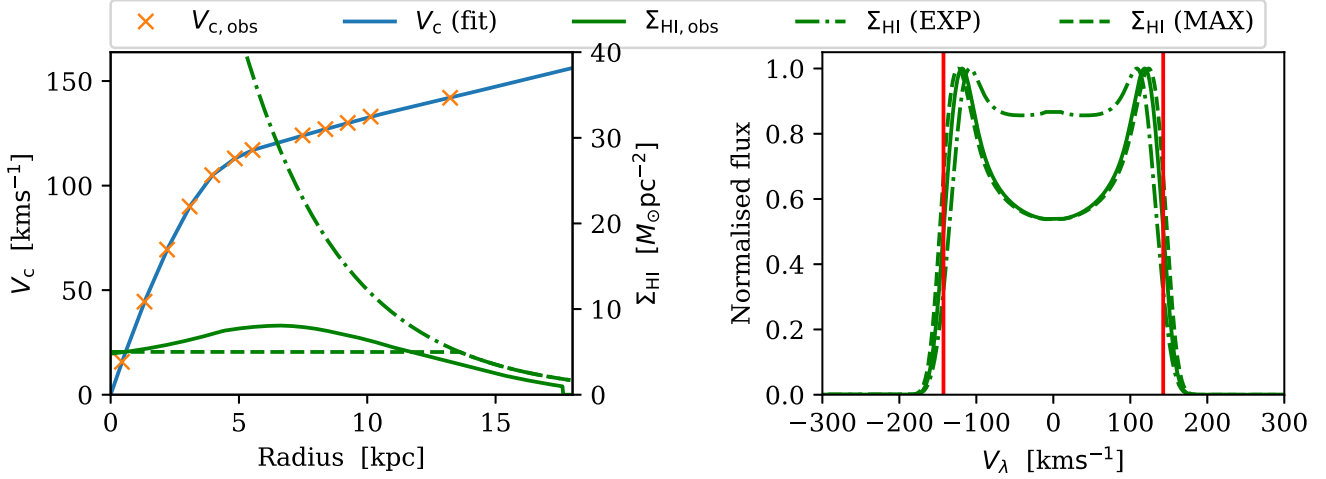
Figure 2. *Left-hand panel*: For the low surface brightness galaxy F568-1, we plot $V_{c,obs}$ from SPARC (orange) and its fit (blue) with linear extrapolation. The observed $\Sigma_{H I,obs}$ from SPARC is shown, as well as the untruncated exponential model (EXP) and the truncated model (MAX) with $\Sigma_{MAX} = 5\,M_\odot\,pc^{-2}$ (see Section 3.2.2). The model $\Sigma_{H I}$ are based on the mass–size relationship of Wang et al. (2016). *Right-hand panel*: The H I spectra produced by applying equation (4) to V_c (fit) and the three $\Sigma_{H I}$. V_λ is the observed radial velocity relative to the galaxy's systemic velocity. The resulting W_{50} are: 285.4 km s^{-1} for $\Sigma_{HI, obs}$ (shown in red); 271.2 km s^{-1} for EXP; 287.3 km s^{-1} for MAX. The untruncated model produces lower values of W_{50} generally. This galaxy is an example where even though the MAX model does not provide a perfect fit to $\Sigma_{H I}$, it is still capable of reproducing W_{50}.
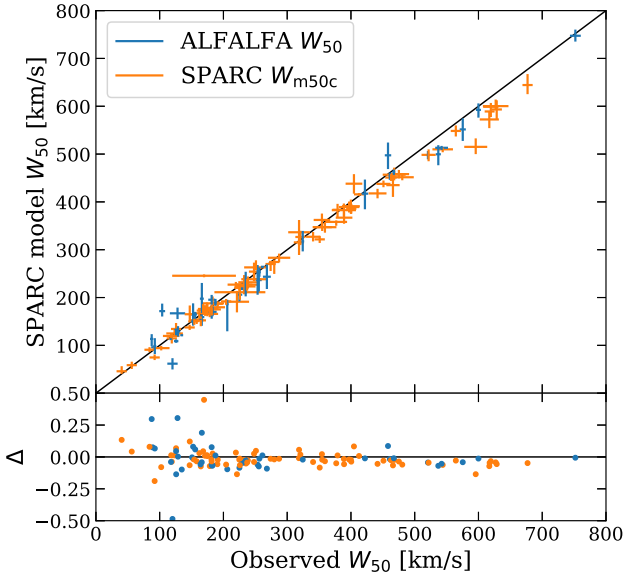

Figure 3. The SPARC model W_{50} (calculated from the observed SPARC H I surface density and RC using equation (4)) plotted against the observed W_{50}, for the 125 galaxies in SPARC with either an ALFALFA W_{50} or SPARC W_{50Mc} (see Section 2.3). The subplot shows the relative differences $\Delta = \frac{Model-Observed}{Observed}$. The model and observations are in good agreement, although there is increasing scatter in the residuals at lower values. Both quantities are calculated/corrected using the inclinations from the SPARC kinematic fitting.

tration as free parameters. There are a large number of different halo profiles in literature (see e.g. Li et al. 2020). We perform our analysis for the cuspy Navarro–Frenk–White (NFW; Navarro, Frenk & White 1997) and the cored Burkert (Burkert 1995) profiles as representatives of cusped and cored profiles, respectively. We caution that neither of these is likely to be entirely accurate after accounting for baryonic feedback, as we discuss further in Section 5.2.

3.2.1.1 NFW NFW found that the haloes in cosmological DMO simulations are well fit by a universal density profile

$$\rho_{NFW}(r) = \frac{\rho_s}{\left(\frac{r}{r_s}\right)\left[1 + \left(\frac{r}{r_s}\right)\right]^2}, \tag{5}$$

where r_s is a scale radius and ρ_s a characteristic density. The enclosed mass at radius r is

$$M_{NFW}(r) = 4\pi\rho_s r_s^3 \left[\ln(1+x) - \frac{x}{1+x}\right], \tag{6}$$

where $x \equiv r/r_s$.

3.2.1.2 Burkert The Burkert profile was proposed as a modification to the PISO profile in order to improve the fit to observations of dwarf spheroids at larger radii. The density profile is

$$\rho_{Burkert}(r) = \frac{\rho_s}{\left(1 + \frac{r}{r_s}\right)\left[1 + \left(\frac{r}{r_s}\right)^2\right]}, \tag{7}$$

and the enclosed mass is given by

$$M_{Burkert}(r) = 2\pi\rho_s r_c^3 \left[\frac{1}{2}\ln\left(1+x^2\right) + \ln\left(1+x\right) - \arctan(x)\right]. \tag{8}$$

It should be noted that while differing at small radii, at large radii both NFW and Burkert profiles have the same slope $\rho \propto 1/r^3$. Therefore, the large extrapolation necessary to calculate the halo's total mass (the virial radius lies far outside the radius probed by W_{50}) is similar for both. The contribution of the halo to the circular speed at radius r is $V_{DM} = \sqrt{GM_{DM}(r)/r}$. It is most convenient to calculate relative to the virial quantities

$$\frac{V_{DM}(r)}{V_{halo}} = \sqrt{\frac{M_{DM}(r)}{M_{halo}}\frac{R_{halo}}{r}}, \tag{9}$$

where $V_{halo} \equiv \sqrt{GM_{halo}/R_{halo}}$ is the circular speed at the virial radius R_{halo} (which is inferred from M_{halo}).
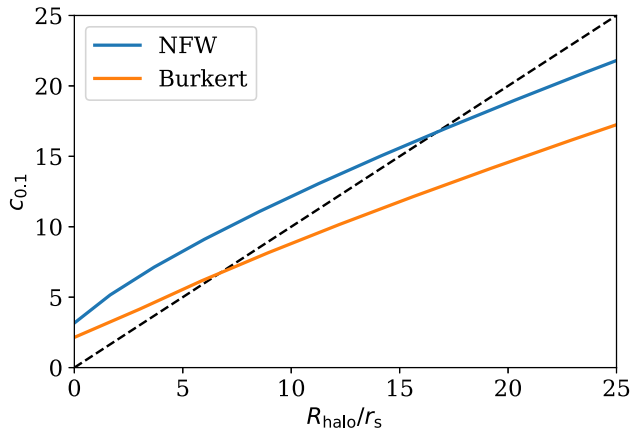
Figure 4. The mapping between $R_{\rm halo}/r_s$ (where $R_{\rm halo}$ is the virial radius) and our new concentration definition $c_{0.1}$ (equation 11). The scale length r_s is defined separately for the NFW (equation 5) and Burkert (equation 7) profiles. For the NFW profile $r_s = r_{-2}$.

Traditionally, concentration is defined as

$$c_{\rm halo} = \frac{R_{\rm halo}}{r_{-2}},\qquad(10)$$

where r_{-2} is the radius at which the slope of the logarithmic density profile is equal to -2 (for NFW $r_s = r_{-2}$). We instead use a new concentration definition

$$c_{0.1} = \frac{R_{\rm halo}}{r_{0.1}},\qquad(11)$$

where $r_{0.1}$ is the radius enclosing 10 per cent of the halo mass. This new definition has three advantages: (1) unlike r_{-2}, $r_{0.1}$ is defined for all halo profiles; (2) $r_{0.1}$ can be calculated from haloes in simulations without assuming a profile but just counting particles in spheres grown from the centre; (3) it is easier to interpret, as it does not depend on the profile shape. In Fig. 4, we show the mapping between $c_{0.1}$ and $R_{\rm halo}/r_s$, which we calculate numerically.

3.2.2 Baryon distribution

We set the prior on M_* to be a Gaussian with mean equal to the NSA M_* and a scatter of 0.2 dex for all galaxies, as a representative value of uncertainty on the mass-to-light ratio. We compare the NSA M_* to the GALEX-SDSS-WISE Legacy Catalog (Salim, Boquien & Lee 2018) values; this is a smaller sample than NSA, but all galaxies have SDSS spectra and WISE photometry which are used in their alternative SED fitting pipeline. For the ~11 000 galaxies in both catalogues and ALFALFA, the mean difference in M_* is 0.04 dex, with a standard deviation of 0.19 dex. With 0.2 dex uncertainty, the NSA M_* are also consistent with those from SPARC (adopting $\Upsilon_* = 0.5\,{\rm M_\odot/L_\odot}$ as per Lelli et al. 2016). We scale the NSA M_* according to the distances listed in ALFALFA.

We use $M_{\rm H\,I}$ and its uncertainty from ALFALFA. McGaugh, Lelli & Schombert (2020) find that the molecular gas mass for late-type galaxies is around 7 per cent of the stellar mass. This correction is minor for most of our sample. Therefore to be consistent with ST21 we set the total baryonic mass to be simply $M_{\rm bar} = M_* + 1.4 M_{\rm H\,I}$, where the factor of 1.4 accounts for cosmological helium and metals.

To estimate the spatial distribution of the stars we use the empirical relationships of Dutton et al. (2011), who use the GIM2D software (Simard et al. 2002) to perform two-component bulge and disc fits to r-band and g-band images and hence derive structural properties

of ~200 000 late and early type galaxies from SDSS DR7. They find empirical relationships for $M_*-r_{\rm disc}$ and $M_*-r_{\rm bulge}$, and the bulge fraction. We use their relationships for late-type galaxies, as these comprise the bulk of our sample. We fix the bulge fraction to the mean relationship and set the Bayesian priors on $r_{\rm disc}$ and $r_{\rm bulge}$ to be a Gaussian with mean given by Dutton et al. (2011) and a scatter of 0.5 dex (estimated from their fig. 4).

For the model H I surface density $\Sigma_{\rm H\,I}(r)$ we use the empirical relationship of Wang et al. (2016, henceforth W16), based on a sample of 500 mainly late-type galaxies with spatially resolved H I observations. They find

$$\log(D_{\rm H\,I}/{\rm kpc}) = 0.51 \log(M_{\rm H\,I}/{\rm M_\odot}) - 3.29,\qquad(12)$$

with 0.06 dex of scatter, where $D_{\rm H\,I}$ is the diameter of the $1\,{\rm M_\odot\,pc^{-2}}$ isophote of $\Sigma_{\rm H\,I}$. W16 also find that for late-type galaxies, there is a homogeneous exponential profile in the outer parts with scale length $r_{\rm H\,I} = 0.1 D_{\rm HI}$. Adopting this and equation (12) specifies the full $\Sigma_{\rm H\,I}(r)$. We also fit exponential profiles to the outer radii of the SPARC galaxies, and find a similar relationship.

Observed H I discs are not exponential all the way to the centre (Leroy et al. 2008; Wang et al. 2014, 2016). Many galaxies have H I cores or holes as the high density H I converts to H_2. The conversion is also dependant on metallicity and temperature, so varies between galaxies (Bigiel & Blitz 2012). Therefore, we truncate $\Sigma_{\rm H\,I}(r)$ to have a maximum value $\Sigma_{\rm max}$, which is set by the requirement to reproduce the observed $M_{\rm H\,I}$. For the W16 model, this gives $\Sigma_{\rm max} \approx 5\,{\rm M_\odot}$, which is consistent with observations. This truncated model is better at handling slowly rising RCs, where an unrealistic central peak in H I could overweight the velocities at low radius and thus cause W_{50} to be underestimated. However, it is not very sensitive to the precise value of $\Sigma_{\rm max}$. Fig. 2 compares the model $\Sigma_{\rm H\,I}(r)$ to SPARC observations for a particular galaxy. Although the model is not a precise match, we see that this does not cause a large shift in W_{50} for a galaxy with a slowly rising RC. This is because the H I density towards the centre changes the flux and shape, but does not shift the position of W_{50}, which is mainly set by the outer, flatter part of the RC that forms the horns.

3.2.3 Inclination

The model circular speed $V_c(r)$ must be corrected for inclination ($i = 0°$ face on; $i = 90°$ edge on) by a factor of $\sin(i)$. As ALFALFA does not resolve the H I disc, inclination is derived from optical photometry. Assuming axisymmetry, the relationship between intrinsic relative thickness q, observed axis ratio b/a and inclination is

$$\cos^2 i = \frac{(b/a)^2 - q^2}{1 - q^2}.\qquad(13)$$

Typically b/a is obtained from infrared photometry, due to the lower extinction. However, this light predominantly comes from older stars that do not reside in a thin disc. There is much discussion on q in the literature, and its variation with galaxy type. It is common to simply assume $q = 0.2$ for line width surveys (Zwaan, Meyer & Staveley-Smith 2010), based on population studies of late-type galaxies (e.g. Unterborn & Ryden 2008). Based on colour, bulge fractions, and Sérsic index, the majority of the ALFALFA sample are late-type. However, as baryonic mass increases (above ~$10^{10}\,{\rm M_\odot}$), there is an increasing population of galaxies in ALFALFA with properties more consistent with early-type galaxies. Dwarf galaxies are also observed to have thicker discs (Méndez-Abreu, Sánchez-Janssen & Aguerri 2010).
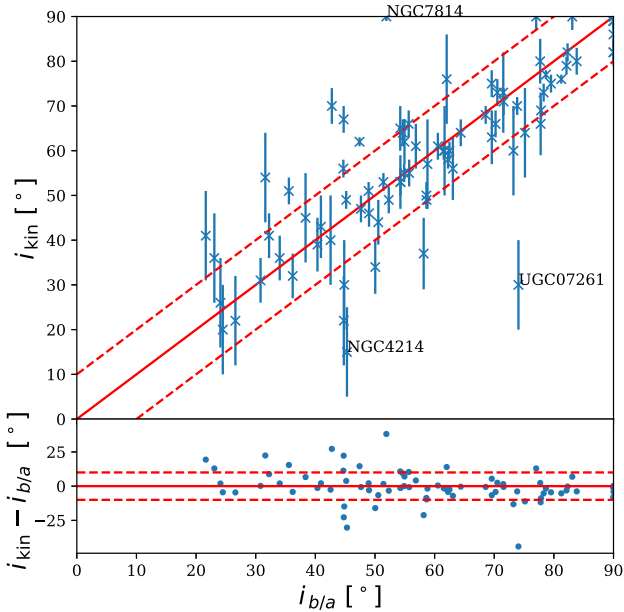
Figure 5. The inclinations from NSA *r*-band optical axis ratios assuming $q = 0.2$ ($i_{b/a}$) plotted against inclinations obtained from tilted ring fits to the resolved H I velocity field (i_{kin}) for the SPARC sample. The solid red line shows equality, and the dashed red lines $\pm 10°$ disagreement. The subplot shows their residuals. Our fiducial model has a flat prior on q, so galaxies above the solid red line are accounted for in the scatter of the posterior probability on halo properties. We account for galaxies below the red line by adopting a 10 per cent uncertainty on *b/a*.

In view of the importance of, and uncertainty in, inclinations we consider three different models for it, ordered from least to most conservative:

(1) Assume q = 0.2 and derive *i* from equation (13) using the observed *b/a* and its uncertainty $\sigma_{b/a}$.

(2) Give *q* a flat prior between 0.15 (the lowest *b/a* in the NSA) and the observed *b/a* of the galaxy. Then the prior on inclination is $\pi(i|b/a \pm \sigma_{b/a})$, as calculated from equation (13). We use this prior in our fiducial analysis as it is more conservative. However, assuming the true distribution of *q* for our sample is peaked at lower values, it will bias our results towards higher *i* and hence lower dynamical mass. This is because $b/a \leq 1$, so changing *q* has a bigger effect on the numerator of equation (13) than the denominator.

(3) Simply assume the galaxies are randomly oriented. The angle between two random vectors in 3D is distributed as $\sin(i)$, so the prior on inclination is then $\pi(i) = \frac{1}{2}\sin(i)$.

For *b/a*, we use the SERSIC_BA value from the NSA, which is calculated from single-component two-dimensional Sérsic fits to *r*-band photometry. In Fig. 5, we plot the SPARC kinematic inclinations against the NSA optical inclinations (with $q = 0.2$), for the 84 galaxies in both SPARC and the NSA. The kinematic inclinations are obtained from tilted-ring fits to the velocity fields, and so are expected to be more accurate.

There are a couple of extreme outliers. The optical inclination of UGC 07261 severely overpredicts the kinematic inclination. On inspection, its SDSS image shows it to be heavily barred, causing it to have a far lower *b/a* that does not reflect the stellar disc itself, which appears to be nearly face on. This could be mitigated by calculating *b/a* using outer isophotes only. However, we find all other measures of *b/a* in the NSA, such as those based on Petrosian fits at radii

containing 50 per cent and 90 per cent of the total light, produce less plausible inclinations across the whole sample.

The optical inclination of NGC 7814 is a severe underestimate of the kinematic value. The SDSS image reveals it to be an edge-on thin disc with an extremely prominent bulge, resulting in a high *b/a*. The flat prior on *q* in our fiducial model accounts for this, as $q = 1$ is appropriate for a bulge. In fact, our fiducial model is flexible enough that it can account for any case where the optical inclination is an underprediction of the true value.

NGC 4214 is the most severe example of a galaxy where the inclination is overpredicted by the optical *b/a* due to the irregularity of the light distribution in its SDSS image. Irregularity has a tendency to make circular objects appear to have a lower *b/a*. To account for these galaxies, we adopt an uncertainty of 10 per cent on *b/a*, which makes the SPARC kinematic and optical inclinations consistent within 2σ for all galaxies where the optical *b/a* overpredicts inclination, except for UGC 07261.

3.2.4 Likelihood

We compare our model line width to the observed line width by using Bayes' theorem to calculate the probability for the parameters θ given model $\mathcal{M}$ and data D,

$$\mathcal{P}(\theta|D, \mathcal{M}) = \frac{\mathcal{P}(D|\theta, \mathcal{M})\mathcal{P}(\theta|\mathcal{M})}{\mathcal{P}(D|\mathcal{M})}. \quad (14)$$

We choose the likelihood to be

$$\mathcal{P}(W_{50,\text{obs}}|\theta, \mathcal{M}) = \frac{\exp\{-(W_{50,\text{obs}} - W_{50})^2/(2\delta W_{50,\text{obs}}^2)\}}{\sqrt{2\pi}\,\delta W_{50,\text{obs}}}. \quad (15)$$

We also repeat the same inference, but imposing the mass–concentration relation (converted to $c_{0.1}$) derived from the Uchuu simulations (Ishiyama et al. 2021, equation 2) as a prior, with 0.11 dex log-normal scatter. This breaks the degeneracy between mass and concentration.

3.3 Abundance matching

In its simplest form, SHAM posits that the most massive (or brightest) galaxy forms in the most massive halo or subhalo, the second most massive galaxy in the second most massive halo and so on (Kravtsov et al. 2004). When applied to a galaxy survey and an equivalently sized simulation box this yields a monotonic relationship between halo mass and galaxy mass. Conroy et al. (2006) showed that this simple non-parametric model could produce an excellent fit to galaxy clustering from the present day up to $z = 5$. The model has been extended to allow stochasticity through the SHAM scatter parameter σ_{AM}, which models both intrinsic scatter in the galaxy–halo connection and scatter from observational uncertainties in the galaxy mass or luminosity (Behroozi et al. 2010, henceforth BCW10). Halo assembly bias can be included in the SHAM framework by allowing secondary halo parameters such as concentration to affect the order in which galaxies are assigned to haloes (Reddick et al. 2013; Chaves-Montero et al. 2016; Lehmann et al. 2016; Stiskalek et al. 2021).

In SHAM, the property used to rank galaxies is traditionally luminosity or stellar mass, as these quantities are readily available for the large samples required to calculate accurate abundance and correlation functions. In this work we use a SHAM model from ST21 (Section 4.2) that is based on a sample of H I-selected galaxies, and instead ranks galaxies using their baryonic mass. This is possible as the galaxy sample is a cross-match of NSA and ALFALFA, and hence contains both a stellar mass and H I mass for each galaxy.
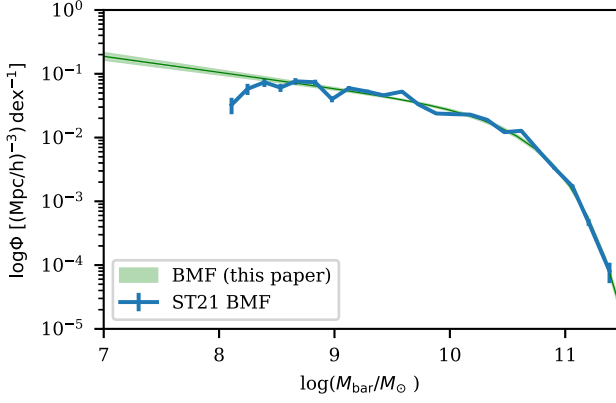
Figure 6. The BMF used in the abundance matching analysis (green line), derived by fitting a Schechter function to the ST21 BMF (blue). The four lowest mass data points are not included in the fit, as the faint end is potentially biased by incomplete treatment of selection effects (ST21). The shaded band shows the 1σ uncertainty.

Baryonic mass is expected to be more fundamentally related to halo mass than stellar mass as star formation has only an indirect effect on galaxies' baryonic mass fractions. This is implied empirically by the Tully–Fisher relation and radial acceleration relation, both of which are tighter and more regular when plotted in terms of total cold baryonic mass rather than stellar mass.

To construct the galaxy baryonic mass function (BMF), we fit a Schechter function (Schechter 1976)

$$\phi(M_{\rm bar}) = \ln(10)\phi_* \left(\frac{M_{\rm bar}}{M_*}\right)^{\alpha+1} e^{-\left(\frac{M_{\rm bar}}{M_*}\right)}, \qquad (16)$$

where the fit parameters are M_*, ϕ_*, and α, to the ALFALFA $\times$ NSA BMF of ST21 (plotted in their fig. 2). We remove the four lowest mass data points, as ST21 suggest that the faint end is biased by incomplete treatment of selections effects. The derived BMF, which we plot in Fig. 6, has parameters $\alpha = -1.24 \pm 0.02$, $\log(M_* h^2/M_\odot) = 10.20 \pm 0.02$ and $\phi^* = 3.3 \pm 0.2 \times 10^{-3}\ h^3$ Mpc^{-3} dex^{-1} (with α and ϕ_* anticorrelated with M_*). The uncertainty on the BMF (which is dominated by α) does not significantly impact the assignment of galaxies to haloes in the mass range probed by ALFALFA, and so is not propagated through our analysis. We find the fitted value of α is in good agreement with that of the ALFALFA H IMF derived by Jones et al. (2018), which is unhampered by selection effects induced by the ALFALFA $\times$ NSA cross-match.

For the SHAM proxy, haloes are first selected that have a peak-mass redshift below a certain value $z_{\rm cut}$, and then ranked by their present-day $M_{\rm vir}$. The $z_{\rm cut}$ parameter allows the model to reproduce the weaker clustering of H I-selected samples with lower $\sigma_{\rm AM}$. ST21 found $z_{\rm cut} = 0.22^{+0.4}_{-0.2}$ and $\sigma_{\rm AM} = 0.42^{+0.8}_{-0.2}$ dex from clustering constraints. We use these maximum likelihood values in our fiducial model. These uncertainties are also not propagated, as discussed further in Section 5.3.

The basic abundance matching procedure links each halo in the catalogue to a galaxy baryonic mass as follows:

(1) Deconvolve the galaxy mass function with the chosen SHAM scatter (BCW10).

(2) Remove haloes with a peak formation time before the redshift $z_{\rm cut}$.

(3) Rank haloes by the proxy (in our case $M_{\rm vir}$).

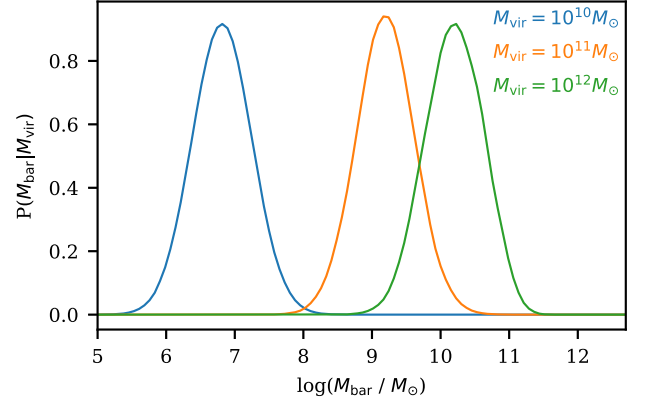(4) Assign a baryonic mass to each halo by matching abundances as described above.

Figure 7. Probability distributions $P(M_{\rm bar}|M_{\rm vir})$ for different values of $M_{\rm vir}$, calculated from an ensemble of abundance matching catalogues. These form the likelihood in our Bayesian inverse abundance matching method (Section 3.3).

(5) Add the SHAM scatter according to the prescription of BCW10.

We repeat this process 500 times, thereby assigning 500 $M_{\rm bar}$ to each halo in the catalogue. We bin the haloes onto a grid of $M_{\rm vir}$ and $c_{0.1}$, and use kernel density estimation to calculate the 1D probability distribution $P(M_{\rm bar}|M_{\rm vir}, c_{0.1})$ for each bin. The probability distributions are then interpolated over the entire space of $\{M_{\rm vir}, c\}$. As our proxy is simply $M_{\rm vir}$ after $z_{\rm cut}$ pre-selection, the probability is actually independent of $c_{0.1}$ here. We show examples in Fig. 7. The posterior probability that the halo of a galaxy has parameters $\{M_{\rm vir}, c_{0.1}\}$, given that it is observed to have baryonic mass $M_{\rm bar,obs} \pm \delta M_{\rm bar,obs}$, is

$$P(M_{\rm vir}, c_{0.1}|M_{\rm bar,obs}) = \frac{P(M_{\rm bar,obs}|M_{\rm vir}, c_{0.1})P(M_{\rm vir}, c_{0.1})}{P(M_{\rm bar,obs})}. \qquad (17)$$

The likelihood for the observed baryonic mass is connected to the likelihood for the 'true' $M_{\rm bar}$ from abundance matching by

$$P(M_{\rm bar,obs}|M_{\rm vir}, c_{0.1}) = \int_0^\infty P(M_{\rm bar,obs}|M_{\rm bar})P(M_{\rm bar}|M_{\rm vir}, c_{0.1})\, dM_{\rm bar}, \qquad (18)$$

where

$$P(M_{\rm bar,obs}|M_{\rm bar}) = \frac{1}{\sqrt{2\pi}(\delta M_{\rm bar,obs})} e^{-\frac{(M_{\rm bar,obs} - M_{\rm bar})^2}{2(\delta M_{\rm bar,obs})^2}}. \qquad (19)$$

The prior $P(M_{\rm vir}, c_{0.1})$ is the 2D probability density function in $\{M_{\rm vir}, c_{0.1}\}$ for all haloes in the simulation.

3.4 Inference methods and analysis of posteriors

We generate the posterior probability distributions using the emcee ensemble sampler (Foreman-Mackey et al. 2013). We initiate the sampler with 200 walkers and use the stretch move $a = 2$. We run a small sample of galaxies until the strict $\tau > 50$ autocorrelation time convergence condition is reached. For the whole sample, we use only 10 000 steps, which for most galaxies does not reach $\tau = 50$, but we check our posteriors are unaffected using the galaxies with fully converged chains. We remove the first 5000 steps as burn-in. The acceptance fraction is ~ 0.15. We rerun a subsample of galaxies with Multinest to check our results are not dependant on the choice of sampler. We place a flat prior on $\log(M_{\rm vir}/M_\odot)$ between $\log(M_{\rm bar})$

Table 2. The statistics we use to characterize and compare the 2D posteriors in halo mass and concentration from kinematics and AM.

Metric	Definition	Interpretation
$\mathcal{F}$	The fraction of the marginalised 1D $\log(M_{vir})$ posterior probability from kinematics that lies in the region $\log(M_{vir}) < \log(M_{bar}) + 0.2$.	The extent to which a galaxy is compatible with having $M_{vir} = M_{bar}$ (i.e. $M_{halo} = 0$). Galaxies with large observational uncertainties will have higher $\mathcal{F}$. We interpret $\mathcal{F} > 0.01$ as $M_{halo} = 0$ not being excluded by W_{50}.
$\mathcal{O}$	The fraction of the marginalized 2D $\{M_{vir}, c_{0.1}\}$ posterior probability from kinematics that lies inside the 2σ contour of the SHAM posterior.	The level of agreement between SHAM and kinematics. We interpret $\mathcal{O} < 0.01$ as the two models being in tension.
$\mathcal{I}$	$\left[\dfrac{\sigma_{AM}}{\sigma_{AM+KIN}} - 1\right]$, where σ_{AM} is the size of the 1σ contour of the marginalized 1D M_{vir} posterior for AM, and σ_{AM+KIN} is the same quantity for the combined posterior of SHAM and kinematics.	The improvement in the constraint on M_{vir} obtained when combining kinematics with SHAM, compared to SHAM alone. $\mathcal{I} = 0$ corresponds to no improvement and $\mathcal{I} = 1$ to a constraint that is twice as tight.

and 15.5, corresponding to an extremely massive cluster. We sample in $\log(c_{0.1})$, with a flat prior between a lower bound corresponding to the smallest value $c_{0.1}$ can take for an NFW halo, and an upper bound of 2, as there are extremely few haloes in the Uchuu simulation with higher concentration.

In Table 2, we define a number of metrics to summarize and compare the posteriors produced by SHAM and kinematics. We quantify the extent to which the posterior is compatible with $M_{halo} = 0$ using the $\mathcal{F}$ metric. It is defined as the fraction of the posterior for which $\log(M_{vir}) < \log(M_{bar}) + 0.2$, with 0.2 chosen because it is the uncertainty on M_*, which is typically much larger than the uncertainties on M_{HI} and so is roughly the maximum uncertainty on M_{bar}.

The tension between the SHAM and kinematics posteriors is quantified by the overlap metric $\mathcal{O}$, the fraction of the kinematics posterior that lies inside the 2σ contour of the SHAM posterior. We interpret two posteriors as being in tension if $\mathcal{O} < 0.01$, roughly corresponding to a 2σ disagreement. We choose 0.01 (rather than 0.05) so small posteriors that overlap with much larger posteriors are not considered to be in tension. This statistic is easy to calculate from the posteriors and can easily handle a variety of shapes of the kinematics posterior, at the price of a nontrivial interpretation. Like other assessments of tension such as the Bayesian evidence, this method is sensitive to the choice of prior. Two posteriors that overlap perfectly will also be considered more in tension if one of them is very wide. In this sense $\mathcal{O}$ measures the similarity of the two posteriors along the lines of the Bhattacharyya distance (Bhattacharyya 1946). We avoid using Bayes factors to evaluate tension, because commonly used samplers may not evaluate the evidence correctly for posteriors with plateaus and shallow likelihoods, which are common for our kinematic model (Fowlie, Handley & Su 2020; Schittenhelm & Wacker 2021; although see Fowlie, Handley & Su 2021 for a solution). To ensure $\mathcal{O}$ is calculated accurately we check that the 2σ contours have converged with respect to chain length, and we use a large number of samples (1 000 000) so the tails of the distributions are well resolved.

The metric $\mathcal{I}$ quantifies the tightening of the constraint on M_{vir} from combining kinematics with SHAM. It is defined relative to the SHAM constraint alone because this tends to be significantly tighter than the kinematic constraint, and because it is more common to have photometric than spectroscopic measurements. We compare in M_{vir} rather than M_{vir}–$c_{0.1}$, because M_{vir} and $c_{0.1}$ are highly degenerate in the kinematic inference. We check that using a different contour, e.g. 2σ, or using the Kullback–Leibler divergence as a measure of information content, leads to the same conclusions. We redefine the SHAM likelihood to include the prior, so it can have the same flat prior on M_{vir} and $c_{0.1}$ as the kinematic model. Then the combined

posterior can be obtained by direct multiplication of the two separate posteriors, due to the assumption of independence.

4 RESULTS

4.1 Individual posteriors

4.1.1 Idealized DMO inference

To understand the posterior shapes for the kinematic model, it is instructive to consider first the case of a single galaxy where we assume the potential is sourced purely by the DM, and the galaxy parameters (i, M_*, r_{HI}, etc.) are perfectly known. This leaves M_{vir} and $c_{0.1}$ as the only free parameters and δW_{50} the only source of uncertainty. We apply this to AGC 742 791 and show the posterior in Fig. 8 for different assumed values of r_{HI} and i. The elongated posterior is due to the perfect degeneracy between mass and concentration; the same line width can be sourced by a high mass halo with a low concentration or vice versa. What matters to first order is the dynamical mass contained within the gas disc that contributes to the rotational velocity. As there are no additional parameters to marginalize over, the posteriors are the locus in the mass–concentration plane required to produce the observed W_{50}, broadened by the observational error δW_{50}.

The naive model $W_{50} = 2V_{max}$, where V_{max} is the maximum circular speed of the halo, is only equal to our full model for values of $c_{0.1}$ where V_{max} occurs at a similar radius to r_{HI}. When V_{max} occurs at a larger radius than r_{HI} more halo mass is required to produce the same W_{50}, so our model posterior curves to the right of the $2V_{max}$ posterior. This effect is more pronounced at low values of r_{HI} and $c_{0.1}$, where V_{max} occurs at a larger radius. Conversely, when V_{max} occurs at a much smaller radius than r_{HI} there is a large contribution to the H I spectral line from the gradually decreasing part of the rotation curve beyond V_{max}, which biases W_{50} to lower values. More mass is required at a given concentration to produce the observed W_{50}, so once again our posterior curves to the right of the $2V_{max}$ case, with the effect more pronounced for higher values of r_{HI} and $c_{0.1}$ (the effect is only significant at extreme $c_{0.1}$). Adding baryons to the idealised model would shift the posteriors to the left, as less DM is required to generate the observed W_{50}.

The right-hand panel of Fig. 8 shows the posterior for our idealized model at different assumed inclinations. At high inclination the $\sin i$ correction to W_{50} is slowly changing, so as inclination decreases the posterior slowly sweeps to higher mass/concentration. As i decreases the posterior sweeps to higher mass and concentration at an increasing rate. Therefore, our 'randomly distributed' inclination model (3) would have a posterior with have a sharp peak at the $i =$
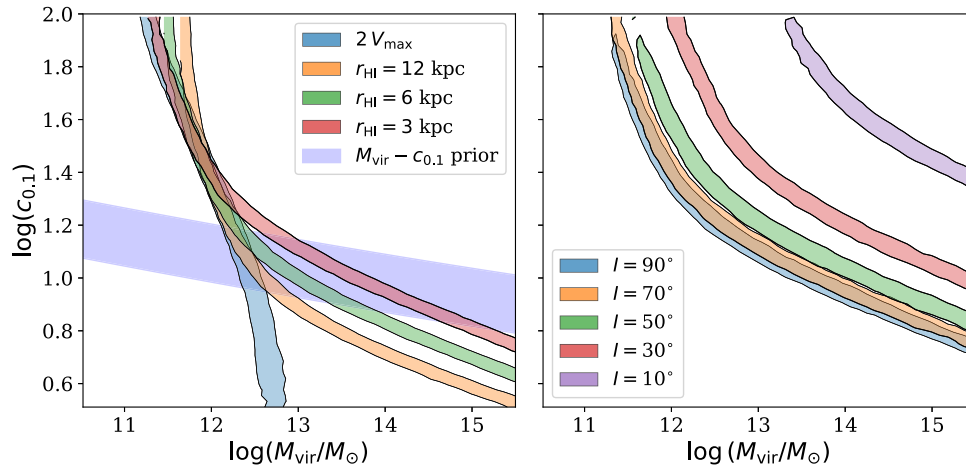
Figure 8. The posterior for AGC 742 791 ($W_{50} = 428 \pm 25$ km s^{-1}), obtained by assuming the line width is only sourced by the DM, and with no uncertainty on inclination or the gas disc size r_{HI}. The contours show the 1σ region, which has width only due to the uncertainty on the line width. In the left-hand panel, r_{HI} is varied. We also show the simple model where $W_{50} = 2 V_{max}$, where V_{max} is the maximum circular velocity of the DM halo, and the mass–concentration prior. As the size of the gas disc decreases and probes a smaller radius, the posterior deviates further from the $2 V_{max}$ model. In the right-hand panel, r_{HI} is fixed and the inclination i is varied.

90° region, with a long tail all the way to the high M_{vir} and $c_{0.1}$ prior boundary at low values of i. For our fiducial inclination model (2), with a flat prior on intrinsic relative thickness q, the minimum inclination occurs at $q = 0.15$, where the inclination probability also peaks. Hence, there is a competing effect between the peak in inclination probability at the lowest value of inclination, and the peak in the 2D posterior corresponding to $i = 90°$ described above. For the simple $q = 0.2$ model (1), the uncertainty on inclination is only due to the uncertainty on b/a, so the posterior is a band.

4.1.2 The full model

The output of our model for a single galaxy is the posterior probability distribution on the two halo parameters M_{vir} and $c_{0.1}$ for both the SHAM and kinematic model, as well as the six galaxy parameters for kinematics (listed in Table 1). We also calculate the joint posterior of both models. Our kinematic model puts Gaussian priors on all galaxy parameters (except inclination), with the width set by the observational or empirical model uncertainty. The line width does not further constrain these galaxy parameters beyond their priors in most cases, so only the 2D and 1D posteriors on M_{vir} and $c_{0.1}$ are of primary interest. Inclination i is sometimes prevented from taking on very low values (face-on) if this would cause too high an intrinsic line width, requiring a DM halo with M_{vir} or $c_{0.1}$ beyond the prior boundaries. It can also be prevented from taking on high values (edge-on), as for some galaxies this would make the intrinsic line width so low that the circular velocity from the baryons alone exceeds it. However, the galaxy parameters are still important, as the uncertainty on them broadens the posteriors on halo properties, often dramatically. For example the high-end tail of the prior on M_* can cause the M_{vir} posterior to have a significant tail to low values.

When we apply the real kinematic model to the whole ALFALFA sample, the resulting 2D posteriors on $M_{vir} - c_{0.1}$ vary greatly between galaxies. In Fig. 9, we show the posteriors for a selection of galaxies to illustrate this diversity, as well as their $\mathcal{F}$, O and $\mathcal{I}$ metrics (see Table 2 for definitions). The kinematics posterior of AGC 724 223 (first panel) is most similar to the theoretical example above. Its W_{50} is large enough relative to M_{bar} that the baryonic contribution

to the dynamical mass within the gas disc is subdominant to DM across the whole range of galaxy priors. Its b/a is small, so there is a narrow prior on inclination that does not greatly broaden the posterior. It has $\mathcal{F} = 0$, because $M_{halo} = 0$ is strongly disfavoured. AGC 742 791 (middle right-hand panel) is the opposite case – the observed W_{50} can be explained with negligible DM within the priors on the baryonic component; it can only have a DM halo of significant mass if $c_{0.1}$ is so low that most of the halo mass sits outside of the gas disc. This results in much higher values of $\mathcal{F}$. AGC 742 791 is an example between these two extremes. Some region of the priors on galaxy parameters are compatible with having little DM within the gas disc, as shown by the 2σ contour extending down to the baryonic mass of the galaxy. The weak constraints are in large part driven by its high b/a, which results in a broad prior on inclination. AGC 200 359 (far right panel) also has a high b/a, but its W_{50} is very large relative to M_{bar}, so the kinematic constraints are still relatively tight. The posterior strongly disfavours $M_{halo} = 0$, so $\mathcal{F} = 0$.

Fig. 9 also shows the difference between the NFW and Burkert profiles. A higher halo mass is required for the cored Burkert profile to generate the same W_{50} as the cuspy NFW profile. The bend in the banana shape is much more pronounced for the Burkert profile, going almost completely horizontal below a certain concentration for each galaxy, meaning a large range of halo masses generates the same line width at a fixed concentration. The cored Burkert profile has a much more slowly rising circular velocity profile than NFW, so for typical concentrations the larger halo mass does not become apparent until radii larger than the extent of the gas disc. This effect is illustrated in Fig. 10.

For all galaxies, the SHAM posterior occupies a constrained region with a defined peak. Towards low mass there is a sharp drop to zero probability, with longer tails to high M_{vir}, high $c_{0.1}$ and low $c_{0.1}$. The SHAM posteriors vary smoothly with M_{bar}, as it is the only galaxy property that the model depends upon (the additional scatter on halo parameters due to the observational uncertainty on M_{bar} is subdominant to the SHAM model's intrinsic scatter). In particular, the posterior smoothly decreases in size in the M_{vir} dimension as M_{bar} decreases (compare the third and fourth columns of Fig. 9). This is due to the differing slopes of the BMF and HMF, meaning
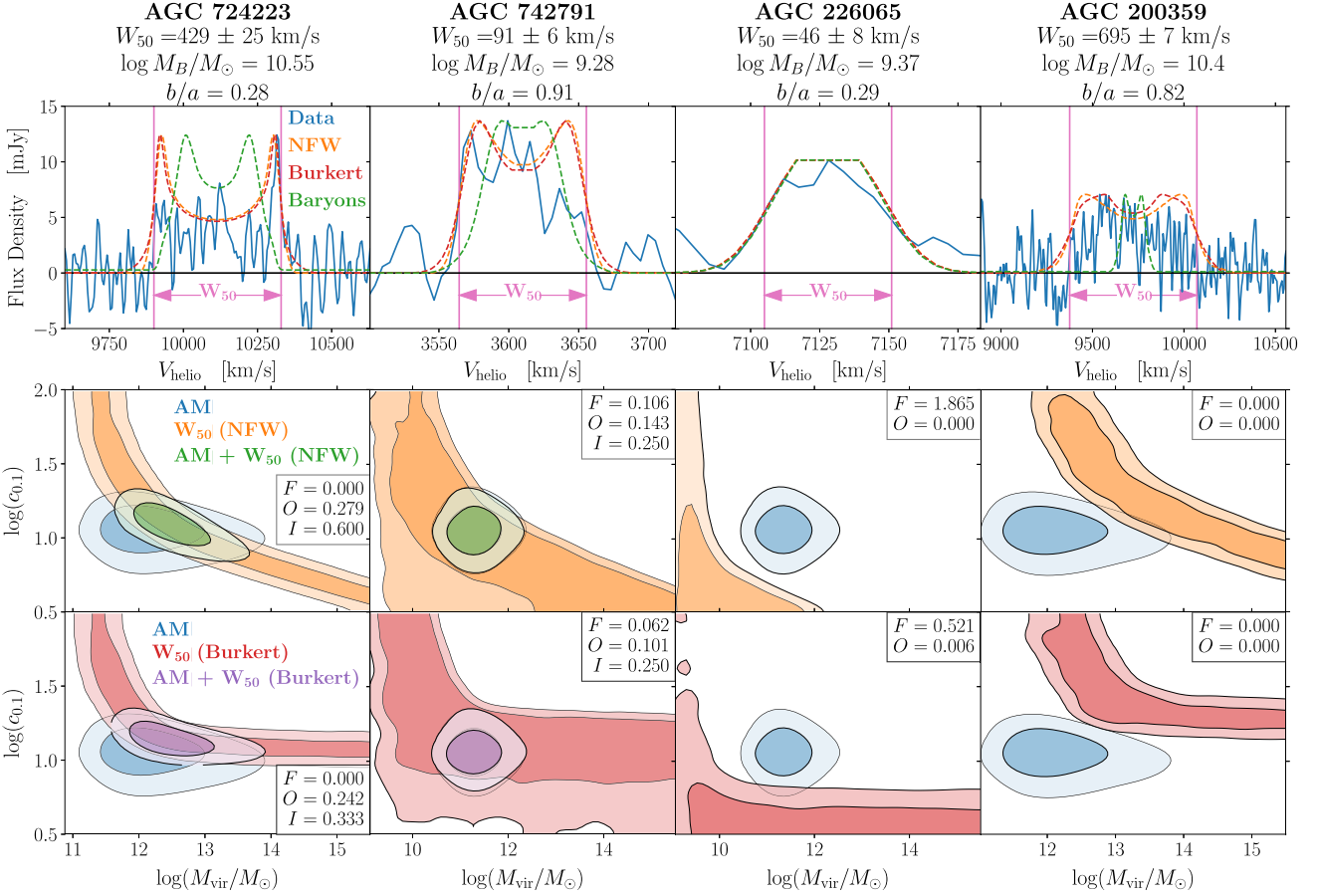
Figure 9. Results for four representative galaxies. From left to right: (i) one with a well-constrained DM distribution from kinematics consistent with SHAM, (ii) a poorly constrained distribution consistent with SHAM, (iii) an apparently DM-free galaxy, and (iv) a galaxy with a more massive halo predicted from W_{50} than expected from SHAM. *Top:* The raw ALFALFA spectra, with our best-fitting model using an NFW and Burkert halo profile. We also show the line profile generated by the baryons only, which is calculated using the mean of the priors on the galaxy parameters. As we are only fitting to W_{50}, the best-fitting model is not necessarily a good match to the spectrum, often due to asymmetries. We leave fitting the full spectrum to future work. *Bottom:* Posteriors on halo mass and concentration from SHAM, kinematics assuming NFW (upper) or Burkert (lower) profiles, and their combination (1 and 2σ isoprobability contours shown). The inset lists the $\mathcal{F}$, $\mathcal{O}$ and $\mathcal{I}$ summary statistics (see Table 2). We only combine the kinematics and SHAM posteriors in the first two case, where they are not in tension.
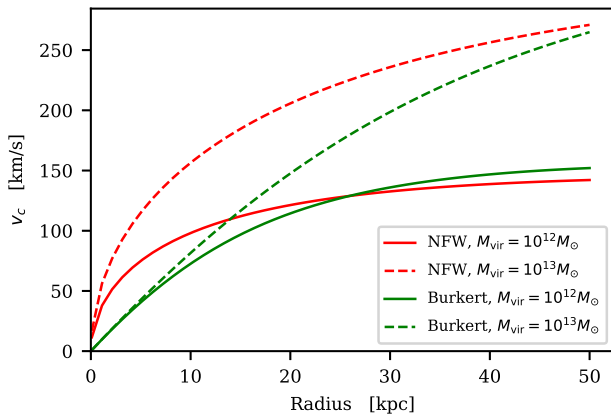

Figure 10. The circular velocity of a DM halo with $c_{0.1} = 10$ for difference masses and halo profiles. For the Burkert profile, increasing the halo mass by a factor of 10 results in very little change in circular speed within 20 kpc. For NFW, the difference is much more pronounced. This explains why increasing the mass of Burkert haloes that are below a certain concentration only slowly increases the predicted line width, and hence why Burkert masses are poorly constrained.

that the scatter $\sigma(M_{\rm vir}|M_{\rm bar})$ increases towards higher mass, even though the SHAM scatter parameter $\sigma_{\rm AM} = (M_{\rm bar}|M_{\rm vir})$ is constant. This also results in an asymmetric posterior: a galaxy of a given $M_{\rm bar}$ is assigned to a long tail of high-mass haloes. We discuss the effect of SHAM scatter further in Section 5.3. The distribution of $c_{0.1}$ in the SHAM posterior corresponds to the mass–concentration relationship of haloes that survive the formation time cut, which have a lower mean concentration at a given mass than the whole catalogue.

The galaxies in Fig. 9 are ordered from least in tension to most in tension. For the first two panels SHAM and kinematics are in good agreement, so $\mathcal{O}$ is high. Therefore, we combine the SHAM and kinematic posteriors, which leads to a significant tightening in the constraints for AGC 724 223 due to its tighter kinematic constraint (and wider SHAM constraint due to the higher $M_{\rm bar}$), whereas there is no improvement for AGC 742791. For AGC 226 065, a very low dynamical mass is inferred from W_{50}, causing tension with AM. The tension is less severe with the Burkert profile, as it is possible to have a larger $M_{\rm vir}$, while keeping the dynamical mass inside the gas disc the same. For cases where SHAM and kinematics are in tension ($\mathcal{O} < 0.01$), we do not combine the posteriors.
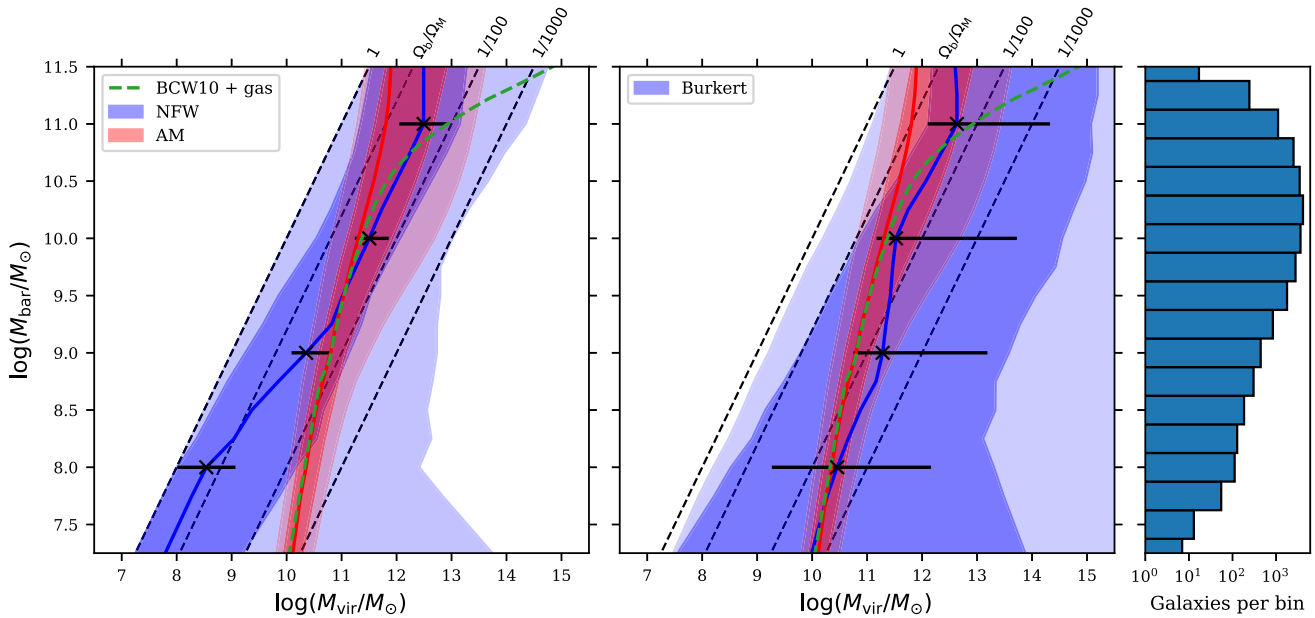
Figure 11. The baryonic mass–halo mass relation formed by stacking the 1D marginalized posteriors on M_{vir} in 0.25 dex bins of baryonic mass for kinematics (blue) and SHAM (red). For kinematics, the mass–concentration prior is applied. The solid line shows the mode of the stacked distribution in each bin, and the shaded regions the 1 and 2σ isoprobability contours. 'BCW10 + gas' shows the best-fitting SHAM relationship from Behroozi et al. (2010), which is derived for optically selected galaxies, converted to baryonic mass using the mean gas fraction in each bin for the ALFALFA galaxies. Black points show the average per-galaxy uncertainties M_{vir} at given M_{bar}; that these errorbars are much smaller than the width of the blue band shows that this is driven mainly by variations between galaxies in a bin. The diagonal dashed lines show constant M_{bar}/M_{vir} as indicated at the top.

4.2 Population behaviour

4.2.1 $M_{bar}-M_{vir}$ relationship

We now derive a $M_{bar}-M_{vir}$ relationship from our modelling of the H I line width and AM. We do this by stacking the 1D posteriors on M_{vir} in bins of M_{bar} with 0.25 dex width, applying the mass–concentration prior to break the degeneracy of those parameters. The result is shown in Fig. 11. The previously discussed gradual variation in the size and asymmetry of the AM posterior in M_{vir} as M_{bar} changes is clearly visible. The kinematic relations have high scatter due to the significant fraction of galaxies for which there is very little constraint on M_{vir} (e.g. the second panel of Fig. 9). For the Burkert profile, the stacked posteriors extend all the way to the upper limit of the prior, due to the cored profile causing a long tail to high mass in most galaxies (e.g. the right-hand panel of Fig. 9), which overlaps with the mass–concentration prior.

Despite the large scatter, the mode of the posterior contains important information about the $M_{bar}-M_{vir}$ relationship implied by H I kinematics. It is stable to random resampling for bins with more than ~150 galaxies. Above $M_{bar} = 10^{9.5}$ M$_{\odot}$, the NFW and Burkert relationships modes are similar, while towards lower mass the NFW mode diverges from the SHAM mode to a much lower M_{vir}. The Burkert bends away from the SHAM mode in the other direction, but becomes extremely close to it again for the lowest M_{bar} bins. For NFW, the slope of the $M_{bar}-M_{vir}$ relationship is much shallower than for AM, and is even compatible with a constant M_{bar}/M_{vir}. For NFW, the mean and median of the distribution in each bin are similar to the mode. However due to the long tail to high M_{vir} for Burkert, the mean and median are different to the mode, and are dependant on the upper bound of the prior. This makes the mode the more robust statistic.

The stacked intervals are very wide for Burkert kinematics, as many individual galaxies have M_{vir} that are poorly constrained by

W_{50}. The divergence between NFW kinematics and SHAM at low baryonic mass is caused by the substantial population of galaxies with low line widths given their baryonic mass, which seem to have little if any DM within the radius probed by their gas disc (e.g. 3rd column of Fig. 9).

The large scatter of the kinematic relations derive from a combination of per-galaxy uncertainty on M_{vir} and an offset of the M_{vir} posteriors of different galaxies in a given M_{bar} bin. To investigate this, we show in black in Fig. 11 the per-galaxy uncertainties for four M_{bar} bins: the points are at the stacked modal M_{vir} and the errorbars show the average over all galaxies in the bin of the distance between the mode and the lower/upper 1σ interval of each individual galaxy posterior. Their size is a much weaker function of M_{bar} than the stacked intervals, showing that the flaring in the blue band at low M_{bar} is driven by increased scatter in M_{vir} *between* galaxies at fixed M_{bar}, rather than increased uncertainties for particular galaxies.

To see how much the relationship is tightened when the galaxies with the weakest constraints are removed, and to check its robustness, we consider a quality cut to the data (Code 1 galaxies only; SNR > 10; $\delta W_{50}/W_{50} < 0.1$; $b/a < 0.7$, which corresponds to $i \approx 45$ for $q = 0.2$), and recalculate $M_{bar}-M_{vir}$ using the remaining 4232 galaxies. The mode and shape of the confidence intervals are unchanged, except for bins with few galaxies after the cut ($M_{bar} < 10^{8.25}$ M$_{\odot}$). The mean width of the 1σ region is 30 per cent smaller for NFW after the cut, but the posterior still reaches the prior bound for Burkert.

4.2.2 Compatibility with $M_{vir} = 0$

Sixty per cent of galaxies in the sample have $\mathcal{F} < 0.01$ for both Burkert and NFW; the other 40 per cent have posteriors that are compatible with $M_{halo} = 0$. To investigate which galaxy properties this depends upon, we trained a Random Forest Classifier (Pedregosa et al. 2011)
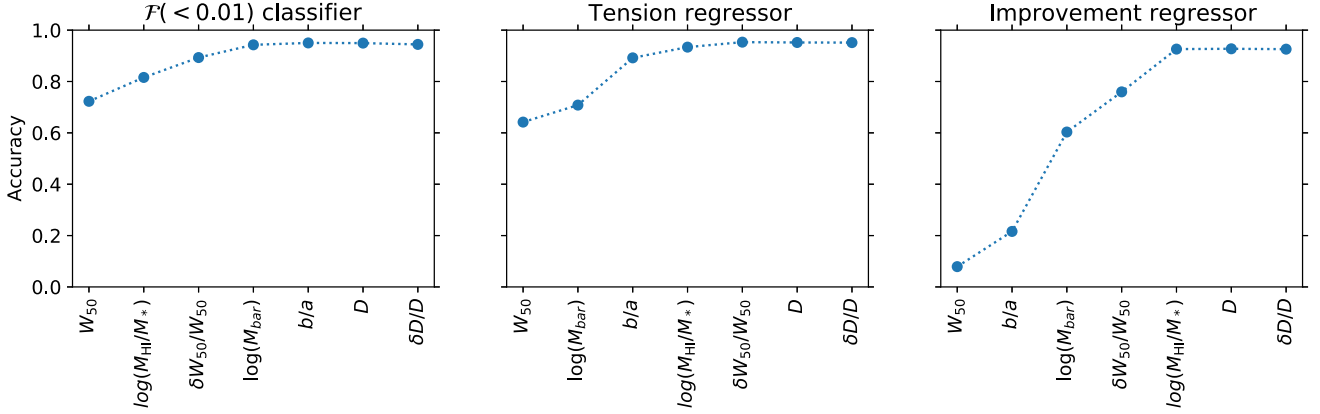
Figure 12. The test set accuracy of the random forest as a function of the cumulative number of features used to predict the posterior metrics $\mathcal{F}$ (left column), tension $\mathcal{O}$ (middle), and improvement $\mathcal{I}$ (right column) (see Table 2 for definitions). We use random forest regressors for $\mathcal{O}$ and $\mathcal{I}$, and a binary classifier for $\mathcal{F}$ on the condition $\mathcal{F}(< 0.01)$, because its distribution is sharply peaked at $\mathcal{F} = 0$. Features are added from left to right in the order that maximizes the increment in accuracy, as described in Section 4.2.2. Accuracy is measured by the fraction of correctly predicted labels for the classifier, and by R^2 (equation 20) for the regressors. We see that W_{50} is the most predictive galaxy property for $\mathcal{F}$ and $\mathcal{O}$, giving reasonable accuracy even when it alone is used to the train the random forest. No galaxy property is predictive on its own for $\mathcal{I}$, but together W_{50}, b/a, M_{bar}, $\delta W_{50}/W_{50}$, and $M_{H\,I}/M_*$ yield good accuracy. D and its fractional error add no new information in any case, as would be expected. The results shown are for the NFW profile, but very similar results are obtained with the Burkert profile.

on the sample for the binary classification $\{\mathcal{F} < 0.01, \mathcal{F} > 0.01\}$, optimizing its hyperparameters using threefold cross-validation. As our labels are roughly balanced, we assess accuracy using the fraction of correctly predicted labels. To select important features we use the method of Stiskalek et al. (2022, Section 3.6), in which galaxy properties are added sequentially to the set of features used to train the random forest. At each increment, the next feature added is the one that generates the greatest improvement in accuracy when combined with the current set of features. This produces a list of features, ordered from most to least important, and the new accuracy after their inclusion. This method of identifying important features avoids ambiguities associated with highly correlated features because the improvement is conditioned on some set of features already being used. The result is shown in the left-hand panel of Fig. 12. For the $\mathcal{F}$ statistic, the most important galaxy property is W_{50}, with a much smaller dependence on $\delta W_{50}/W_{50}$, log $(M_{H\,I}/M_*)$, and M_{bar}.

The top panel of Fig. 13 correlates $\mathcal{F}$ with various galaxy properties, and the bottom panel correlates these properties with W_{50}. $\mathcal{F}$ is a strong function of W_{50}, with $\mathcal{F} > 0.01$ for nearly all galaxies in the lowest W_{50} bin. We use the reduced residual method to remove the dependence on W_{50}, and find that the residuals of $\delta W_{50}/W_{50}$, M_{bar}, b/a, n correlate positively with $P(\mathcal{F} < 0.01)$, and the residuals of log (M_{HI}/M_*) anti-correlate. These are all as expected: higher values of $\delta W_{50}/W_{50}$ and b/a give weaker constraints and therefore higher $\mathcal{F}$; higher gas fractions result in tighter constraints and hence lower $\mathcal{F}$, as a larger $M_{H\,I}$ means a larger $r_{H\,I}$, so W_{50} probes further into the halo at fixed M_{bar}; higher M_{bar} galaxies have a lower mean gas fraction and higher n, leading to weaker constraints.

4.2.3 Agreement between SHAM and kinematics

In Fig. 14, we show the distribution of our tension metric $\mathcal{O}$ (described in Table 2). Around 1000 galaxies are in tension ($\mathcal{O} < 0.01$) for both halo profiles. The second peak in $\mathcal{O}$ corresponds to galaxies where the SHAM posterior lands exactly on the kinematics posterior, such as the left-hand panel of Fig. 9. Most cases of tension are due to kinematics predicting too little dynamical mass relative to AM, as in the third column of Fig. 9. Nearly all these galaxies are also compatible with $M_{halo} = 0$ according to our $\mathcal{F}$ statistic. There is a

smaller population of galaxies where the dynamical mass inferred from the kinematic model is higher than SHAM (as in the right column of Fig. 9).

We correlate tension with galaxy properties in Fig. 13. The too low/too high dynamical mass populations are visible in the two peaks in the correlation with W_{50}, with a higher peak at low W_{50} than high W_{50}. There is higher tension for NFW than Burkert at low W_{50}, and also at low M_{bar}, which is also seen in the stacked $M_{vir} - M_{bar}$ relationship. There are slightly more galaxies for Burkert that overpredict the dynamical mass than for NFW. It is harder for kinematics to be in tension with SHAM from predicting too high a M_{vir} due to the SHAM posterior's long tail to high M_{vir}, especially at high M_{bar}.

We apply a random forest regressor to our $\mathcal{O}$ metric to assess its dependence on galaxy properties, once again optimizing hyperparameters using threefold cross-validation. We use the same importance ranking procedure as before, but this time assessing accuracy using the coefficient of determination

$$R^2 = 1 - \frac{\sum_i (y_{i,\text{test}} - y_{i,\text{pred}})^2}{\sum_i (y_{i,\text{test}} - \hat{y}_{\text{test}})^2}, \quad (20)$$

where $y_{i,\text{test}}$ is the test set value, $y_{i,\text{pred}}$ is the corresponding prediction, and $\hat{y}_{\text{test}}$ is the mean test set value. $R = 1$ is perfect prediction, and $R = 0$ is given by a model that always predict $\hat{y}_{\text{test}}$, disregarding the input data. We find O is most dependant on W_{50}, and to a lesser extent on M_{bar} and b/a (Fig. 12, middle panel).

In Fig. 15, we plot the 2D $W_{50} - M_{bar}$ distribution of the ALFALFA sample, which shows clearly that tension primarily occurs when W_{50} is low for a given M_{bar}. The split is very clean in the log $W_{50} - \log M_{bar}$ plane and the $M_{bar} - W_{50}$ relationship from the SPARC sample lies within it. The scatter in the ALFALFA data is much greater than in SPARC. We see from the relatively similar scatter of the ALFALFA and SPARC data describing the same galaxies (red and white dots) that the ALFALFA data is reasonably accurate for the SPARC sample. Therefore, the increased scatter is due to systematics in the ALFALFA × NSA data that are not present for galaxies that are also in SPARC and/or because the ALFALFA × NSA sample has a different intrinsic log $W_{50} - \log M_{bar}$ distribution to SPARC; e.g. we do not expect the entire sample to be rotationally supported

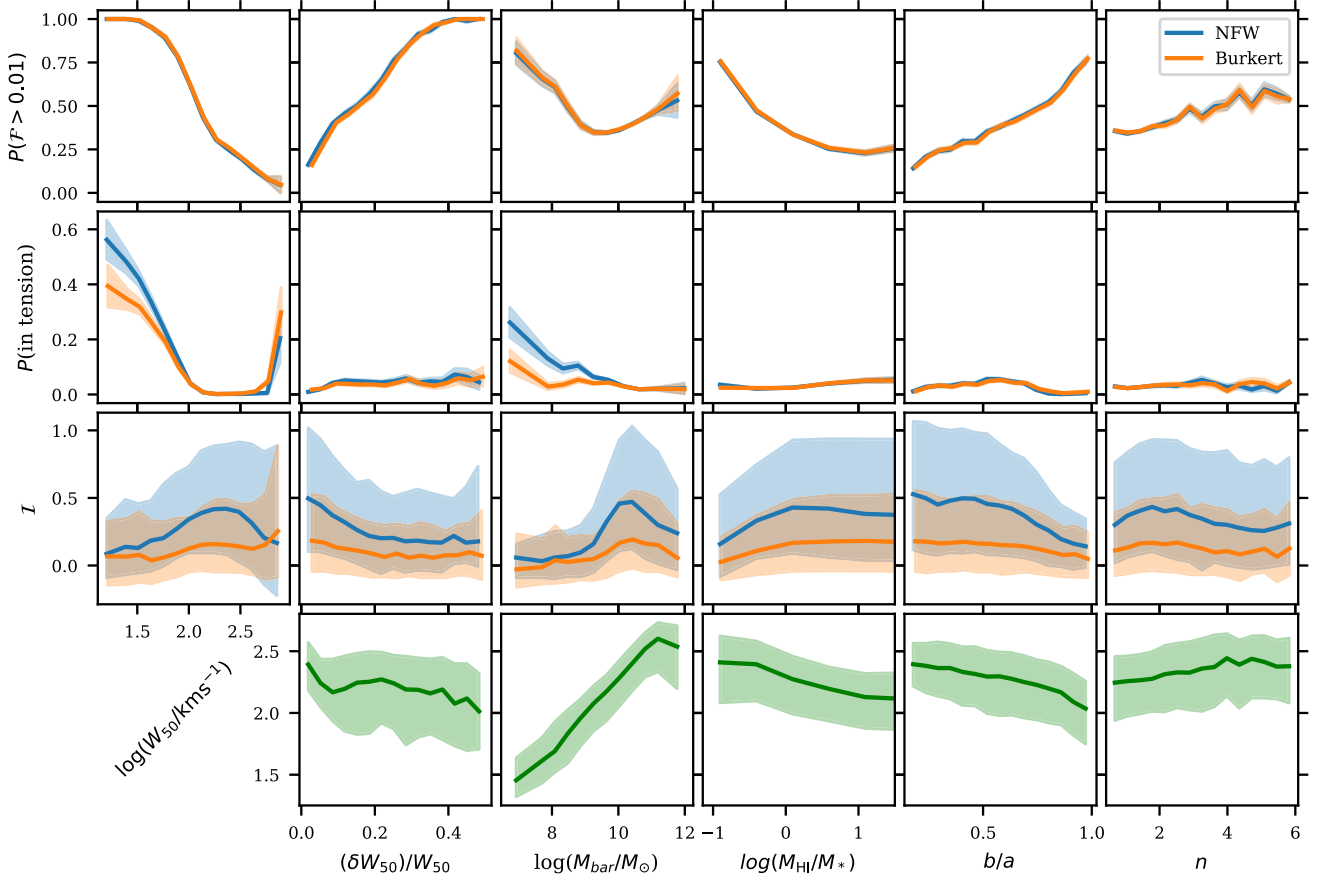**Figure 13.** The correlation of galaxy parameters with the $\mathcal{F}$, $\mathcal{O}$, and $\mathcal{I}$ metrics, and with the ALFALFA line width W_{50}. The top two panels show the fraction of galaxies in each bin for which $M_{\mathrm{halo}} = 0$ is not excluded by kinematics ($\mathcal{F} > 0.01$, top panel) or exhibit tension between SHAM and kinematics ($\mathcal{O} < 0.01$, second panel). The shaded bands show the 1σ uncertainties calculated by bootstrap resampling each bin. The bottom two panels show the median (solid line) of $\mathcal{I}$ and $\log W_{50}$, with the shaded region showing the 1σ variation between the galaxies in each bin. We initially set 15 bins in each plot, then merge the outermost bins if there are fewer than 10 galaxies in one of them. We show the correlation of each quantity with W_{50} because we saw in Fig. 12 that this was the most important quantity for determining $\mathcal{F}$, $\mathcal{O}$, and $\mathcal{I}$. Galaxy parameter values that give tighter constraints (low δW_{50}, high gas fraction, low b/a, high W_{50}) produce low values of $\mathcal{F}$ and high values of $\mathcal{I}$. Tension is most prevalent at low W_{50} and low M_{bar}, and is higher for NFW than Burkert.
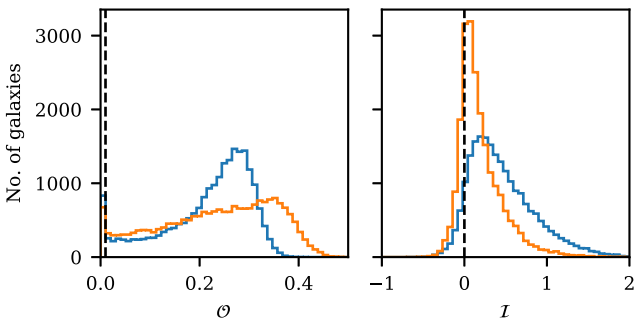

Figure 14. Distribution of the overlap metric $\mathcal{O}$ and improvement metric $\mathcal{I}$ over all galaxies in the sample (see Table 2). SHAM and kinematics are in tension for galaxies with $\mathcal{O}$ to the left of the vertical dashed line. For those with $\mathcal{I}$ to the right of the vertical dashed line, the constraint on M_{vir} from SHAM is tightened when combined with the constraint from kinematics. For $I = 1$ the constraint on M_{vir} is twice as tight.

and in equilibrium. We discuss potential systematics extensively in Section 5.2.

We repeat the random forest regressor procedure for the improvement statistic $\mathcal{I}$, and find that it is dependant on a greater number of

features than $\mathcal{F}$ and $\mathcal{O}$ (Fig. 12, right-hand panel). The improvement $\mathcal{I}$ is much greater for NFW than Burkert due to the Burkert posterior's long tail to high mass, which for many galaxies aligns with the high mass of tail of the SHAM posterior. We correlate $\mathcal{I}$ with galaxy properties in Fig. 13. The positive correlation of $\mathcal{I}$ with M_{bar} is because the SHAM constraint is much weaker at high M_{bar} due to the increasing scatter $\sigma(M_{\mathrm{vir}}|M_{\mathrm{bar}})$. The correlations with other galaxy parameters are for the same reasons as discussed for the $\mathcal{F}$ statistic, as I is strongly correlated with the tightness of the kinematic constraint.

5 DISCUSSION

5.1 Interpretation of the results

5.1.1 Abundance matching versus kinematic constraints

Our W_{50}-based kinematic model and our abundance matching (AM) model produce Bayesian posteriors on the mass and concentration of individual galaxies. The constraints offered on halo mass by kinematics are reasonably strong (<1 dex with the mass–concentration prior applied) for around 40 per cent of galaxies when assuming a cuspy profile. When assuming a Burkert profile the constraints are much weaker, often only providing a lower limit on mass, as the
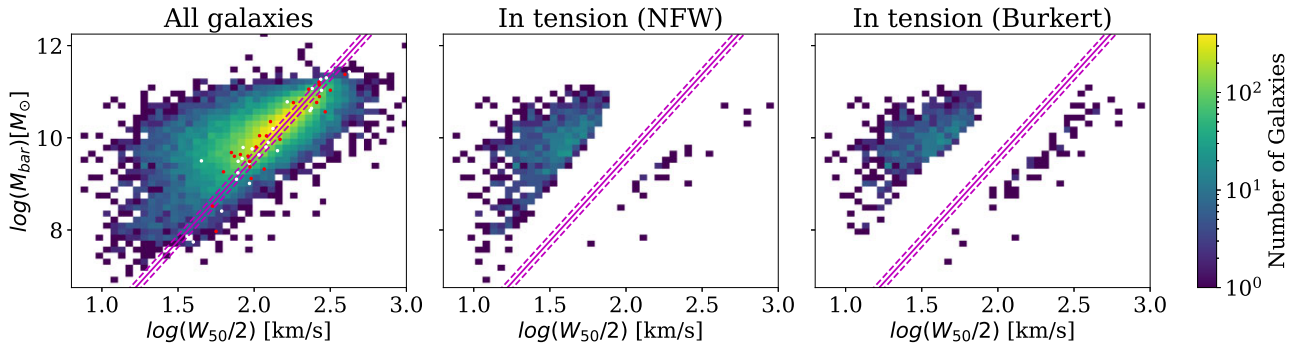
Figure 15. A heat map of the distribution of W_{50} and M_{bar} for the whole ALFALFA sample, and for galaxies where SHAM and kinematics are in tension. The W_{50} have been corrected for inclination assuming the intrinsic relative thickness $q = 0.2$. The $M_{bar}-W_{50}$ relationship from SPARC (Lelli et al. 2019) is shown (magenta line) with its 3σ intrinsic scatter (dashed lines). The 25 galaxies that are in SPARC × ALFALFA × NSA are plotted both using their ALFALFA data (red dots) and their SPARC data (white dots) in the left-hand panel. For the latter we use the W_{m50} line width measure and kinematic inclinations, as these were used to derive the SPARC $M_{bar}-W_{50}$ relationship.

more slowly rising circular velocity profiles due to the core mean that haloes with vastly different masses can have relatively minor differences in rotational velocity within the gas disc. The model is flexible enough that the mode of the posterior produces a W_{50} within 2σ of the observed value for all galaxies. Thus, it is not possible for the constraints to be artificially tightened by a poor fit, which is a concern when fitting resolved rotation curves.

In general, the constraints offered by SHAM are tighter than for kinematics. When assuming an NFW profile, combining SHAM with kinematics yields stronger constraints on halo mass for a majority of galaxies, with a 1σ constraint that is twice as tight for around 1/5 of the sample. The improvement is more pronounced at the high-mass end, where the SHAM constraints are weaker. Our analysis therefore shows that SHAM can be augmented with the abundant H I line widths of future large-scale surveys to constrain better the distribution of DM, especially for haloes with cuspy profiles.

The improvement from combining probes may in fact be even stronger, as the tightness of our SHAM constraints are likely overestimated because our analysis does not account for the uncertainties on the SHAM model itself (e.g. in the proxy and scatter). This is especially true at the low-mass end, where there is no clustering data to constrain the SHAM parameters. We discuss the uncertainty on the SHAM model further in Section 5.3.

5.1.2 Tension

Overall, we do not find significant tension between kinematics and SHAM when comparing the posteriors of individual galaxies, apart from a population of galaxies for which the line width predicts too small a dynamical mass relative to AM, and a smaller population of galaxies for which the inverse is true. The lack of tension is to some extent unsurprising given the weak constraints on halo properties from kinematics for many galaxies. For our SHAM model, the mean M_{vir} changes by only 2 dex over 4 dex in baryonic mass, which is comparable to the scatter in M_{vir} from kinematics for a substantial fraction of our sample even with the mass–concentration prior applied.

Although more individual galaxies are in tension between SHAM and kinematics for both profiles at lower baryonic mass (in part due to the tighter constraints from SHAM at low mass), fewer are in tension with Burkert than NFW. This is also clearly visible in the stacked $M_{bar}-M_{vir}$ relationship: the mode of the stacked posterior lies at much

lower M_{vir} when assuming an NFW profile, compared to both SHAM and the Burkert profile, which are in good agreement with each other. Therefore both individual and stacked posteriors suggest that at low mass NFW underpredicts halo mass relative to AM/Burkert. This is expected when fitting a cuspy profile to a cored halo (Trujillo-Gomez, Kruijssen & Reina-Campos 2022). Therefore, the disagreement may imply that core formation is a decreasing function of mass, and/or that SHAM needs further refining for H I-selected samples.

There is much literature, both theoretical and observational, on DM core formation (see Del Popolo & Le Delliou 2021 for a review). A popular hypothesis is that core formation proceeds through the kinematic heating of the DM in galactic centres by supernova driven cycles of gas inflows and outflows (Read & Gilmore 2005; Pontzen & Governato 2014). Some simulations (e.g. Di Cintio et al. 2014) and analytic calculations (Peñarrubia et al. 2012) have found core formation to be a function of M_*/M_{vir}, with M_* a proxy for the energy available to drive gas outflows with supernovae, and M_{vir} related to the amount of DM that needs to be removed and the depth of the potential well that it needs to be removed from. Read, Walker & Steger (2019) find observational evidence for this in local dwarf galaxies. From our stacked M_*-M_{vir} relationship (which is qualitatively very similar to our $M_{bar}-M_{vir}$ relationship), we see no such correlation with M_*/M_{vir}: SHAM is most in agreement with cored halos relative to cusps at low M_*, rather than at some value of M_*/M_{vir}, as suggested by the above models. However our results in this regard are weak, and do not rule out any scenario. The disagreement between kinematics with NFW and SHAM at low mass could also be resolved by haloes expanding to lower concentrations than in DMO simulations. Semi-analytical studies based on SHAM have found some evidence that this is necessary to explain the Tully–Fisher (Desmond & Wechsler 2015) and radial acceleration (Desmond 2017a) relations. Both these papers use a relaxation model in which the halo as a whole can expand or contract relative to its pristine NFW profile (Dutton et al. 2007), finding mild evidence for expansion. However, applying a different relaxation model to the radial acceleration relation, Paranjape & Sheth (2021) found evidence for expansion only in the outer halo.

The $M_{bar}-W_{50}$ relationship for the SPARC galaxy sample (Lelli et al. 2019) is extremely tight (see Fig. 15). Papastergis, Adams & van der Hulst (2016) take an extremely restricted sample of only 90 ALFALFA galaxies and construct a similarly tight relationship: their criteria are edge-on, high gas fraction and $W_{50} > 100\,\mathrm{km\,s^{-1}}$, and they inspect each spectrum for residual noise. They find a correlation between increasing offset of a galaxy from their mean relationship

to lower W_{50} and the degree to which the H I spectrum only has a single peak (measured by kurtosis). The offset also increases at lower baryonic mass. Our full ALFALFA $\times$ NSA sample is a very different population of galaxies compared to these two studies, so it is expected that the $M_{\rm bar}-W_{50}$ distribution will differ. However with such a large sample based on unresolved observations, some of our galaxies will inevitably have biased halo properties due to systematic uncertainties. Galaxies that lie far from the literature relations fall under particular suspicion. We consider potential systematics in the kinematic modelling in detail in Section 5.2, but we note here that our results are robust to basic quality cuts on SNR and b/a, such as the one used to test the $M_{\rm bar}-M_{\rm vir}$ relationship in Section 4.2.1.

5.1.3 $M_{bar}-M_{vir}$ relationship

A key result is the $M_{\rm bar}-M_{\rm vir}$ relationship derived from H I kinematics (Fig. 11), which we find to be roughly linear in log-space. This is reminiscent of the BTFR between $M_{\rm bar}$ and $V_{\rm rot}$, which is linear over six decades in mass. Canonical optically selected SHAM relationships such as that of BCW10, however, display a break in the stellar mass–halo mass relation (SHMR) around the Milky Way mass. Generating a linear BTFR from such SHAM relationships is non-trivial (Desmond 2017b). Observations of massive spirals have also failed to detect the SHMR break expected from SHAM (Li et al. 2020; Posti & Fall 2021; Mancera Piña et al. 2022), and McGaugh & van Dokkum (2021) pointed out that the halo mass predictions for the Milky Way and Andromeda from kinematics lie well below the SHAM prediction. Neither the $M_*-M_{\rm vir}$ or $M_{\rm bar}-M_{\rm vir}$ relationship from our kinematic model display a break, but continue approximately linearly with a mode close to the cosmic baryon fraction at the highest masses. This result is robust to quality cuts in the data. Although for the Burkert profile the relationship is very wide, the modal relation is still linear.

Posti & Fall (2021), also applying the mass–concentration relationship as a prior in their kinematic analysis, argue for a linear SHMR relationship for late-type galaxies, with elliptical galaxies displaying the expected break. Unlike Posti & Fall (2021), we do not have a clean sample of late-type galaxies. When assessed by standard cuts on colour and Sérsic index (n), it appears the majority of galaxies in our sample above $10^{10.5}\,{\rm M}_\odot$ are early type. When we split our sample using a crude cut on morphological type ($n < 2$ for late-type, $n > 3$ for early-type), we do not find differing relationships in $M_{\rm bar}-M_{\rm vir}$. Multiple empirical studies using different methods have attempted to measure the SHMR for passive and star-forming galaxies separately, but no consensus has yet been reached on whether they differ (Wechsler & Tinker 2018).

Our fiducial SHAM model, based on H I-selected galaxies, has a scatter of $\sigma_{\rm AM} = 0.42$ dex, whereas the scatter for optically selected galaxies is well constrained to $\sim$0.2 dex (at least under the assumption that $\sigma_{\rm AM}$ is not a function of mass; e.g. Reddick et al. 2013). As $\sigma_{\rm AM}$ increases the highest mass galaxies are increasingly assigned to the much more numerous lower mass haloes, washing out the characteristic break in the mean $M_{\rm bar}-M_{\rm vir}$ relationship, as can be seen in Fig. 11. However, there is still an increasing scatter $\sigma(M_{\rm vir}|M_{\rm bar})$ towards high mass, with the posterior extending to high $M_{\rm vir}$. There is also still a break in galaxy formation efficiency as a function of halo mass. The fiducial model is in better agreement with our kinematic relationship than BCW10, but we caution that $\sigma_{\rm AM}$ is very poorly constrained for H I-selected galaxies. The model also assigns around a third of the most massive galaxies into haloes such that $M_{\rm bar}/M_{\rm vir} > \Omega_{\rm b}/\Omega_{\rm M}$, suggesting SHAM models with scatter

this high may not produce realistic galaxy populations, even if they reproduce the clustering signal.

5.2 Kinematic modelling caveats

5.2.1 Baryonic effects on halo density profiles

We have studied the NFW and Burkert profiles as representative cases of a cusped and cored profile, respectively. However, previous studies have shown that no halo profile is a good fit to all observations (Katz et al. 2017; Li et al. 2020). This may be explained by the difficulty in calculating baryon-induced modifications to haloes due to the uncertainty in the numerical implementation of baryonic physics and resolution constraints. Initially, haloes are expected to contract adiabatically from baryonic infall as galaxies form (Li et al. 2022 incorporate this effect into their halo fitting procedure). The subsequent expulsion of gas from galaxies by star formation is then expected to expand the halo. Recently, Velmani & Paranjape (2023) studied halo relaxation in large volume hydrodynamical simulations and found it to vary substantially with halo mass and concentration, and star formation rate. Paranjape et al. (2021) demonstrated that changes to halo relaxation physics can significant alter W_{50}.

Any baryonic effect that modifies halo density profiles would be expected to alter our mass and concentration constraints, and thus our analysis would need to be repeated for halo profiles inspired by hydrodynamical simulations (e.g. Di Cintio et al. 2014). As there is currently no convergence in the precise effects of baryons on DM across a range of galaxy scales, this is left for future work. Ultimately, theoretical and observational progress in understanding the net effect of feedback processes is required before the modification of haloes by baryons can be robustly accounted for in kinematic analyses.

5.2.2 Modelling the line width

Our model assumes the line width is the product of a galaxy's azimuthally averaged rotation curve and H I surface density, with a H I velocity dispersion of 10 km s^{-1}. We tested this using the SPARC sample and found good agreement. However very few of these galaxies had observed line widths < 100 km s^{-1}, which is the region in which most of the tension is found. Furthermore, we see a trend in Fig. 3 where towards lower line width there is weaker agreement between model and observations, although this could be explained by uncertainty in the extrapolation of RCs.

Pressure support has the effect of reducing the rotational velocity below the circular velocity, causing the dynamical mass to be underestimated (Bureau & Carignan 2002; Oh et al. 2015; Iorio et al. 2016). This 'asymmetric drift' correction becomes important in galaxies where $V_{\rm rot}$ is comparable to the gas dispersion, and becomes more important at larger radii. Our models for $\sigma_{\rm H I}$ and $\Sigma_{\rm H I}$ do not capture the galaxy-to-galaxy variation required to sensibly apply the correction, which is challenging even with resolved data. As a simple test of the sensitivity of our results to asymmetric drift, we rerun our inference replacing the assumption that $V_{\rm rot}(r) = V_{\rm c}(r)$ with $V_{\rm rot}^2(r) = V_{\rm c}^2(r) - \sigma_{\rm HI}^2$, with $\sigma_{\rm H I} = 10$ km s^{-1}. We find that for our stacked $M_{\rm bar}-M_{\rm vir}$ relationship the mode of $M_{\rm vir}$ is increased by 0.5 dex for NFW and 0.15 dex for Burkert in the lowest mass bin ($M_{\rm bar} = 10^{7.25}\,{\rm M}_\odot$), where the effect is greatest. At $M_{\rm bar} = 10^{8.5}\,{\rm M}_\odot$, the difference with our fiducial model is less than 0.15 dex for both haloes. We conclude that, although there is an increasing number of galaxies towards lower mass and W_{50} that are affected by asymmetric drift, our overall conclusions are likely robust to it.

Non-equilibrium motions are also expected to be increasingly important towards lower mass, due to supernovas driving gas out of the plane of the disc, and radial outflows (e.g. Verbeke et al. 2017). Asymmetries due to the increasing irregularity of galaxies may also cause W_{50} not to reflect the dynamical mass (Reynolds et al. 2020). High velocity clouds in the observed galaxy can also create high velocity wings in the flux profile, leading to W_{50} overpredicting the rotational velocity (Schulman, Bregman & Roberts 1994).

5.2.3 Measuring the line width

Extracting the line width from often noisy spectra is a difficult process. We investigate the difference between the base ALFALFA catalogue W_{50} line width measurement and the W_{Yu85} measurement from the Yu et al. (2022) reanalysis. W_{Yu85} implies a slightly lower mass for most galaxies, but there is a significant population for which it infers a much higher mass. The strength and trends of tension with galaxy properties are the same for both line width measures, although there is not good agreement on which specific galaxies are in tension. W_{Yu85} produces a very similar stacked $M_{bar}-M_{vir}$ relationship to W_{50}, with the mode not different by more than 0.4 dex in any bin.

Haynes et al. (2018) caution that at low SNR, radio frequency interference can cause the line width to be underestimated. This could potentially explain the population of galaxies with extremely low W_{50} relative to their baryonic mass. However, we find that galaxies with H I code 2 are not overrepresented among in-tension galaxies, and we do not find a trend between SNR and tension above $SNR = 6$, suggesting this cannot be the sole cause. Yu et al. (2022) provide a different cut on galaxies more vulnerable to RFI. Again we found these galaxies were not overrepresented among the in-tension galaxies. Another potential source of error is confusion, where the separation between two galaxies is smaller than the beam width. Jones et al. (2016) show that the impact of confusion for the catalogue as a whole is not significant, although they say it is easy to identify specific examples. Using the flag for crowding in the reanalysis of Yu et al. (2022), we found that crowded galaxies are not over-represented among the in-tension galaxies.

Ball et al. (2022) present an ALFALFA reanalysis that utilizes a similar curve-of-growth based algorithm and crowding analysis to Yu et al. (2022), with the aim of minimizing the BTFR scatter by removing outliers. Future work may benefit from testing their methods and sample selections in the mass-modelling context.

5.2.4 Inclination

As discussed in Section 3.2.3, the gas disc inclination calculated from the observed optical axis ratio can be highly inaccurate. The most troubling potential inclination systematic is when the measured b/a is biased low, leading to too high (edge-on) an inclination, too low an inferred intrinsic W_{50} and hence too low a dynamical mass within the gas disc. The majority of our in-tension galaxies are of the type where the dynamical mass inferred from W_{50} is lower than the SHAM expectations, with the frequency of tension increasing at lower M_{bar}. In our SPARC sample, we find a single severe underestimation of b/a in the optical, for a heavily barred face-on galaxy. As bars are most prevalent in galaxies with stellar masses $10^9 < M_*/M_\odot < 10^{11}$ (Méndez-Abreu et al. 2010), this is unlikely to explain the increasing tension that we observed towards even lower M_{bar}.

Several irregular galaxies in the SPARC sample had a NSA b/a that was somewhat too low (given both their kinematic inclination and visual appearance), which we accounted for by adopting a 10 per cent error on b/a. This sample was too small to look for a correlation between mass or flux and disagreement with kinematic inclination, to see if the effect increases towards lower mass, as does the tension. In general, we expect irregularity to increase towards low M_{bar}, as the weaker self-gravity of the system makes it more susceptible to internal and environmental effects.

Applying a flat prior on q, when the true distribution for the sample is likely peaked at $q \approx 0.2$, will cause the inferred masses to be biased low. We calculated the $M_{bar}-M_{vir}$ relationship for the $q = 0.2$ inclination model, and found it made little difference in the lowest and highest mass bins, where it raises M_{vir} somewhat. However, it is in these bins that we expect q to be higher than 0.2.

Finally, Almeida & Filho (2019) have argued that low dynamical masses may be caused by ignoring triaxiality when calculating inclination using the optical b/a. We also neglect triaxiality, so our results are subject to the same potential bias. A more sophisticated analysis could use the expected population distribution of inclination to infer the distribution of axial ratios as a function of galaxy properties (Putko et al. 2019). However this is complicated for ALFALFA, as the selection function of blind spectroscopic H I surveys is dependant on inclination (Lang et al. 2003).

5.2.5 Baryonic mass

Using an erroneously high M_{bar} will cause M_{vir} to be underestimated. The uncertainties on $M_{H I}$ are smaller than those on M_*. We verified the NSA M_* in two ways. Firstly, we compared it to the GALEX survey and found consistency within the uncertainties. Secondly, we compared the NSA to SPARC, and also found good agreement on the whole, but with increasing disagreement towards low mass ($< 10^9 M_\odot$). However, low-mass galaxies tend to be more gas-dominated, so the bias on M_{vir} from M_* will be less important, suggesting it is unlikely to be the cause of the observed trend of tension with mass. Ball et al. (2022) also identified foreground stars as a potential source of overestimated stellar masses, finding this pathology in 11 per cent of the most extreme BTFR outliers in their data.

H I self-absorption has not been accounted for in H I masses. Depending on the model used, the correction can range from insignificant except for the most edge-on galaxies, to a 30 per cent correction for all galaxies (see Jones et al. 2018 for a thorough discussion). Underestimating the H I mass causes the halo circular velocity to be overestimated, and the size of the H I disc to be underestimated. Both of these result in too high a halo mass being inferred. Applying the inclination-based correction from Jones et al. (2018) ($\Delta \log M_{H I} = 0.13 \log(b/a)$) results in insignificant differences to our results above $M_{bar} = 10^9 M_\odot$. However at lower mass, where galaxies are more gas-rich, the effect is significant, with the mode of the stacked M_{vir} distribution 0.4 dex lower at $M_{bar} = 10^{7.5} M_\odot$ for NFW, putting more galaxies in tension. As the Jones et al. (2018) model is based on thin disc galaxies, the true effect may be larger for thicker dwarf galaxies. Future work may benefit from improved modelling of H I self-absorption.

5.2.6 Gas distribution

The adopted gas model from Wang et al. (2016) is based on a sample of 500 dwarf and spiral galaxies, with masses down to $M_{H I} = 10^7 M_\odot$. They find that early-type galaxies, although still lying on the $M_{H I}-D_{H I}$ relation, tend to have flatter gas profiles, with a larger fraction of their gas lying outside of $D_{H I}$. Adopting

too low an r_{HI} causes M_{vir} to be overestimated. On the other hand, explaining the tension between SHAM and kinematics at the low-mass end would require that the true size of the gas disc be smaller than in the model. The $M_{HI}-r_{HI}$ relationship of W16 shows no evolution towards lower mass. The most in-tension galaxies in this regime, for which the baryonic mass alone is enough to generate the observed line width, are little sensitive to changes in gas distribution as both the stars and gas are contained within the gas disc.

5.3 Abundance matching caveats

The abundance matching relationship for gas-selected galaxies is constrained by clustering only for $\log M_B > 9.4$, and even for those ranges it is poorly constrained (see fig. 14 of ST21) due to the weaker clustering of H I-selected samples and the comparatively small sample size of H I surveys.

Rather than sample from the entire posterior of ST21, which includes regions with extremely high σ_{AM} for which the galaxy–halo connection is essentially fully randomized, we adopted their maximum likelihood point under the assumption that the SHAM parameters are independent of mass (although this was strongly ruled out for optically selected samples by ST21). As discussed previously, the high-mass end is particularly sensitive to σ_{AM}. We find far weaker dependence on the z_{cut} parameter.

The extrapolation of the baryonic mass function becomes important for galaxies with $\log M_{bar} < 10^{8.5} \, M_\odot$. However, the trends in our results are not noticeably different in the extrapolated regime. The only way to increase the baryon fraction for SHAM at the low-mass end (hence reducing tension with kinematics) is to strongly increase the steepness of the baryonic mass function.

Finally, at the faint end it is possible that a qualitatively different SHAM prescription is required to deal with the different formation scenarios of lower mass galaxies. ST21 showed that fainter optically selected samples require higher scatter, and argue that low-mass galaxies may require a different set of SHAM assumptions, such as increased galaxy formation bias or a difference between satellite and central galaxies. On the other hand, Nadler et al. (2020) find an upper limit $\sigma_{AM} = 0.2$ for Milky Way satellites.

5.4 Comparison to literature

Previous studies, most notably of the SPARC galaxies, have used resolved rotation curves to infer halo properties (Katz et al. 2017; Li et al. 2020). The advantage of resolved rotation curves is that they are able to provide much more information on density profiles than a summary statistic that effectively samples the RC at a single radius. Li et al. (2020) tested a wide variety of rotation curves, and found that in general cored profiles such as Burkert provided better fits than cuspy profiles such as NFW, with many galaxies favouring cores even at higher mass. Even with the full rotation curve, the constraints on halo properties are weak in some cases. For example, Li et al. (2020) find many galaxies for which the 1σ halo mass constraint spans over two orders of magnitude. However for many of the galaxies they recover a good constraint without any prior applied, as the shape of the rotation curves is enough to break the degeneracy between mass and concentration. Our approach is complementary, sacrificing precision of the dynamical measurements for a much larger sample size and thus trading potential systematic error in relating the small SPARC sample to the entire halo population for weaker galaxy-by-galaxy constraints. For us, robust conclusions are available only statistically across the full sample.

Our analysis sheds light on the 'small-scale problems' of ΛCDM (Bullock & Boylan-Kolchin 2017). The Too-Big-to-Fail problem (TBTF) is the observation that the kinematically inferred halo masses of the Milky Way's satellites are much lower than the masses of the largest subhaloes of Milky Way-sized haloes in N-body simulations (Boylan-Kolchin, Bullock & Kaplinghat 2011). This was later generalised to populations of field galaxies that were found to have kinematics that implied a lower halo mass than predicted by abundance matching (Ferrero et al. 2012). Papastergis et al. (2015) studied the problem in a sample of ALFALFA isolated dwarfs for which resolved rotation curves also exist, using a SHAM procedure that ranks galaxies by their line width. Fitting an NFW profile, they found that haloes with resolved outer rotational velocities of less than 25 km s^{-1} are incompatible with the haloes implied by AM. They find that fitting a halo profile with a mass-dependant core reduces the tension, but does not fully alleviate it. Although we do not separate satellite and field galaxies, or restrict our sample to galaxies with resolved rotation curves (and hence more robust kinematic halo masses), our results are similar to the above studies: we also find a population of dwarf galaxies for which the halo mass inferred from kinematics is significantly below the SHAM prediction, such that the two measurements are in statistical tension. The disagreement is partially alleviated by fitting a Burkert profile instead of NFW. There is much literature on proposed solutions to TBTF for both satellite and field galaxies, including modelling and observational uncertainties, and new DM physics (see Papastergis & Shankar 2016 for a review).

A related small-scale problem is the observed diversity of rotation curves at fixed galaxy mass, which does not appear in ΛCDM simulations (Oman et al. 2015). Baryonic models that solve small-scale problems such as TBTF through core-formation create cores too uniformly, and therefore fail to generate this rotation curve diversity (Sales, Wetzel & Fattahi 2022). In our sample, we observe great diversity in W_{50} at fixed M_{bar} at lower mass, where the gas disc does not probe so far into the halo. This could be indicative of different arrangements of the baryons and/or different DM central densities at fixed halo mass, assuming the abundance matching relation does not flare at low mass. However, as we do not have accurate baryonic distributions we cannot provide more concrete results.

We find a significant number of ALFALFA galaxies are DM-deficient according to our model, as their line width is completely explained by the baryons alone. The existence of apparently DM-deficient galaxies has been previously noted for the ALFALFA sample by Guo et al. (2020), using a simple method where the dynamical mass is estimated from the gas disc scale length and the observed line width without full modelling. They apply quality cuts to the sample and find 19 dark-matter deficient galaxies (14 of which are isolated) out of a sample of 324 (although Almeida & Filho 2019 argue this may due to neglected triaxiality, see Section 5.2.4).

Mancera Piña et al. (2019) study a sample of six ALFALFA galaxies with low linewidths for their baryon masses using HI interferometric data, deriving resolved RCs (with 2–3 resolution elements per galaxy side) which support the galaxies being baryon dominated. Furthermore, using higher resolution observations, Mancera Piña et al. (2021) identified an apparently 'DM-free' isolated galaxy in the sample, although the inclination is still a potentially significant systematic uncertainty and its stability in the absence of DM has been contested (Sellwood & Sanders 2022). Two DM-free dwarf galaxies have also been identified using globular cluster dynamics (van Dokkum et al. 2018, 2019, although see Saifollahi et al. 2021) speculated to have formed from gas stripped in a galaxy–galaxy collision (van Dokkum et al. 2022).

In general, claimed observational detections of dark-matter deficient galaxies tend to be controversial due to modelling uncertainties, even with far better data than our unresolved observations. It is interesting to speculate, however, whether the galaxies we identify as being plausibly DM-deficient would remain so given more precise measurements. Jackson et al. (2021) and Moreno et al. (2022) find DM-free galaxies produced in tidal interactions in simulations, the latter predicting that 30 per cent of central galaxies host at least one DM-free satellite.

Another approach to comparing H I kinematics with ΛCDM expectations is to forward model the line width velocity function using either semi-analytical models (Chauhan et al. 2019; Paranjape et al. 2021) or hydrodynamical simulations (Dutton, Obreja & Macciò 2018; El-Badry et al. 2018). This does not require the inclinations of individual galaxies, avoiding a major source of uncertainty. Dutton et al. (2018) find that the velocity function for dwarf galaxies in the hydrodynamical NIHAO simulations are in good agreement with ALFALFA line widths in the range $10 < W_{50}/2 < 80 \, \mathrm{km \, s^{-1}}$. They identify turbulent motions, projection effects due to intrinsic H I disc thickness and flattened DM distributions as important factors in lowering the observed line widths relative to expectations. We do not properly account for turbulent motions (although we test our model sensitivity to them) and we assume an infinitely thin H I disc. We find less tension at low line width when fitting a Burkert profile, but our inferred core-formation dependence is in disagreement with the M_*/M_{vir} dependence seen in their simulation. It is plausible that the turbulent motions, H I disc thickness and feedback physics in their simulation account for the differences with our results.

Dutta, Khandai & Rana (2022) use an abundance matching-based method to extract $M_{\mathrm{HI}}–M_{\mathrm{vir}}–V_{\mathrm{rot}}–W_{50}$ scaling relations for the ALFALFA sample. They use group finder-based halo masses to obtain a reduced halo mass function corresponding to the ALFALFA sample, which they abundance match to the ALFALFA H I MF. The resulting H I-selected H I-to-halo mass relationship (their fig. 7) is similar to the mean of our stacked abundance matching H I-to-halo mass relationship at low mass, but at high mass has a much stronger break. This is largely driven by the lack of AM scatter in their model (compared to 0.42 dex in ours), which prevents an apples-to-apples comparison.

5.5 Future work

This paper presents a first attempt to compare the halo properties inferred from abundance matching and H I line widths for an entire population of H I-selected galaxies. Future H I surveys will improve the constraints on SHAM models for H I-selected galaxies by reducing the uncertainty on galaxy clustering and extending it to lower masses, allowing a more robust assessment of the agreement between the two methods and their relative constraining power.

The increased precision and reduction in systematics on W_{50} of future surveys should also improve the constraints from kinematics. The increasing number of observed line widths will also increase the statistical power at the low and high mass ends. Future surveys will also allow both H I line width and abundance matching studies to be extended to higher redshifts. Ponomareva et al. (2021) have already used line widths from MeerKAT to study the BTFR out to $z = 0.081$. Glowacki, Elson & Davé (2021) predict evolution in the BTFR with redshift from the SIMBA simulation.

More information is contained in the H I flux profile, of which W_{50} is a summary statistic. Exploiting this would increase the precision of halo parameter inference. Using a similar model as this work, Paranjape et al. (2021) showed this by performing a full

spectrum fitting for some nearby galaxies with very well resolved ALFALFA spectra. This method is potentially very powerful if it can be applied to whole populations of galaxies. A potential problem is the uncertainty in the detailed H I distribution, which may have cores, holes, or asymmetries. These may bias the inferred halo properties if not adequately modelled.

Finally, other methods of inferring the properties of DM haloes could potentially be combined, which could probe the DM distribution at different radial distances from the halo centre. The H I line width probes the central region of the halo, but weak lensing measures the acceleration towards the outskirts of stacked galaxies. The velocity dispersions of stars in early-type galaxies (or of galaxies in groups or clusters) could also be used. For example, Schulz, Mandelbaum & Padmanabhan (2010) used weak lensing to measure the DM halo profile in the outskirts of massive elliptical galaxies, extrapolated it to the centre assuming an NFW halo and then compared the resulting central dynamic mass to the SDSS velocity dispersion, finding evidence for halo contraction.

6 CONCLUSIONS

We have compared the constraints on halo mass and concentration inferred from the kinematic modelling of the H I line width with those inferred from an (inverse) abundance matching model specifically tailored to H I-selected galaxies, for the ~22 000 galaxies in the ALFALFA × NSA data set. Our conclusions are as follows:

(i) The two methods produce consistent halo constraints galaxy by galaxy in most cases, with the kinematics posterior broader and requiring a mass–concentration prior for bounded constraints on either quantity.

(ii) The halo posteriors of SHAM can be augmented with information from the H I line width to produce tighter constraints on the DM distributions of individual galaxies. The gains are greater when assuming a cuspier halo profile.

(iii) Towards low baryonic mass there is an increasing population of galaxies with smaller line widths than expected from abundance matching. For some galaxies, this implies a dynamically insignificant amount of DM within their gas disc, leading to extremely high baryon fractions when the halo is extrapolated to the virial radius. The disagreement with abundance matching is more severe when fitting an NFW halo than Burkert, which we interpret as weak evidence for a cored central DM density at low baryonic mass. There is a smaller population of galaxies for which SHAM and kinematics disagree because the dynamical mass inferred from kinematics is higher than from AM.

(iv) The $M_{\mathrm{bar}}–M_{\mathrm{vir}}$ relation reconstructed from H I kinematics is in statistical agreement with that from SHAM (Fig. 11). It is however closer to a power law, with a deviation (especially assuming an NFW profile) towards lower M_{vir} at fixed M_{bar} at the faint end. When assuming a Burkert profile there is less information to be gleaned on M_{vir} from the line width, resulting in a very uncertain relation.

(v) We formulate statistics to quantify whether a galaxy (i) exhibits tension between its kinematics and SHAM modelling results, (ii) affords a strong improvement in halo constraints by combining the two methods, and (iii) has $M_{\mathrm{halo}} = 0$ excluded by the kinematic modelling. We also develop a machine learning-based method for assessing the extent to which these statistics correlate with various galaxy properties, finding line width to be the most important feature in each case.

Our analysis demonstrates the potential for combined photometric and spectroscopic constraints on the galaxy–halo connection,

even when using low-resolution spectroscopic products such as H I line widths. With future surveys set to improve dramatically our knowledge of the H I universe, we anticipate that our framework will be useful for inferring DM distributions, constraining kinematic and empirical models, and advancing understanding of the physical processes that underlie galaxy formation.

ACKNOWLEDGEMENTS

DATA AVAILABILITY

Data from the ALFALFA and NSA surveys are publicly available at http://egg.astro.cornell.edu/alfalfa/data/index.php and https://www.sdss.org/dr13/manga/manga-target-selection/nsa/. The SPARC database is located at http://astroweb.cwru.edu/SPARC/. The data of the Yu et al. (2022) reanalysis of ALFALFA will be publicly released with their paper. Data from the Uchuu simulation suite is available at http://skiesanduniverses.org/Simulations/Uchuu/. Other data will be made available on reasonable request to the corresponding author.

REFERENCES

Albareti F. D., et al., 2017, ApJS, 233, 25
Almeida J. S., Filho M., 2019, RNAAS, 3, 191
Ball C. J., Haynes M. P., Jones M. G., Peng B., Durbala A., Koopmann R. A., Ribaudo J., O'Donoghue A., 2022, ApJ, 950, 87
Behroozi P. S., Conroy C., Wechsler R. H., 2010, ApJ, 717, 379
Behroozi P. S., Wechsler R. H., Wu H.-Y., 2013a, ApJ, 762, 109
Behroozi P. S., Wechsler R. H., Wu H.-Y., Busha M. T., Klypin A. A., Primack J. R., 2013b, ApJ, 763, 18
Bhattacharyya A., 1946, Sankhyā: The Indian Journal of Statistics (1933-1960), 7, 401
Bigiel F., Blitz L., 2012, ApJ, 756, 183
Blanton M. R., Kazin E., Muna D., Weaver B. A., Price-Whelan A., 2011, AJ, 142, 31
Blanton M. R., Roweis S., 2007, AJ, 133, 734
Bok J., Blyth S.-L., Gilbank D. G., Elson E. C., 2019, MNRAS, 484, 582
Boylan-Kolchin M., Bullock J. S., Kaplinghat M., 2011, MNRAS, 415, L40
Bryan G. L., Norman M. L., 1998, ApJ, 495, 80
Bullock J. S., Boylan-Kolchin M., 2017, ARA&A, 55, 343
Bureau M., Carignan C., 2002, AJ, 123, 1316
Burkert A., 1995, ApJ, 447, L25
Burrage C., Copeland E. J., Millington P., 2017, PRD, 95, 064050
Chae K.-H., Desmond H., Lelli F., McGaugh S. S., Schombert J. M., 2021, ApJ, 921, 104
Chauhan G., Lagos C. d. P., Obreschkow D., Power C., Oman K., Elahi P. J., 2019, MNRAS, 488, 5898
Chaves-Montero J., Angulo R. E., Schaye J., Schaller M., Crain R. A., Furlong M., Theuns T., 2016, MNRAS, 460, 3100
Conroy C., Wechsler R. H., Kravtsov A. V., 2006, ApJ, 647, 201
Del Popolo A., Le Delliou M., 2021, Galaxies, 9, 123
Desmond H., 2017a, MNRAS, 464, 4160

Desmond H., 2017b, MNRAS, 472, L35
Desmond H., Ferreira P. G., Lavaux G., Jasche J., 2018, MNRAS, 474, 3152
Desmond H., Katz H., Lelli F., McGaugh S., 2019, MNRAS, 484, 239
Desmond H., Wechsler R. H., 2015, MNRAS, 454, 322
Di Cintio A., Brook C. B., Macciò A. V., Stinson G. S., Knebe A., Dutton A. A., Wadsley J., 2014, MNRAS, 437, 415
Durbala A., Finn R. A., Odekon M. C., Haynes M. P., Koopmann R. A., O'Donoghue A. A., 2020, AJ, 160, 271
Dutta S., Khandai N., Rana S., 2022, MNRAS, 511, 2585
Dutton A. A. et al., 2011, MNRAS, 416, 322
Dutton A. A., Obreja A., Macciò A. V., 2018, MNRAS, 482, 5606
Dutton A. A., van den Bosch F. C., Dekel A., Courteau S., 2007, ApJ, 654, 27
El-Badry K. et al., 2018, MNRAS, 477, 1536
Ferrero I., Abadi M. G., Navarro J. F., Sales L. V., Gurovich S., 2012, MNRAS, 425, 2817
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Fowlie A., Handley W., Su L., 2020, MNRAS, 497, 5256
Fowlie A., Handley W., Su L., 2021, MNRAS, 503, 1199
Glowacki M., Elson E., Davé R., 2021, MNRAS, 507, 3267
Guo H., Li C., Zheng Z., Mo H. J., Jing Y. P., Zu Y., Lim S. H., Xu H., 2017, ApJ, 846, 61
Guo Q. et al., 2020, Nature Astron., 4, 246
Guo Q., White S., Li C., Boylan-Kolchin M., 2010, MNRAS, 404, 1111
Hafen Z. et al., 2019, MNRAS, 488, 1248
Haynes M. P. et al., 2018, ApJ, 861, 49
Iorio G., Fraternali F., Nipoti C., Di Teodoro E., Read J. I., Battaglia G., 2016, MNRAS
Ishiyama T. et al., 2021, MNRAS, 506, 4210
Ishiyama T., Fukushige T., Makino J., 2009, PASJ, 61, 1319
Ishiyama T., Nitadori K., Makino J., 2012, preprint (arXiv:1211.4406)
Jackson R. A. et al., 2021, MNRAS, 502, 1785
Jones M. G., Haynes M. P., Giovanelli R., Moorman C., 2018, MNRAS, 477, 2
Jones M. G., Haynes M. P., Giovanelli R., Papastergis E., 2016, MNRAS, 455, 1574
Katz H., Lelli F., McGaugh S. S., Di Cintio A., Brook C. B., Schombert J. M., 2017, MNRAS, 466, 1648
Koribalski B. S. et al., 2020, Ap&SS, 365, 118
Kravtsov A. V., Berlind A. A., Wechsler R. H., Klypin A. A., Gottlöber S., Allgood B., Primack J. R., 2004, ApJ, 609, 35
Lang R. H. et al., 2003, MNRAS, 342, 738
Lehmann B. V., Mao Y.-Y., Becker M. R., Skillman S. W., Wechsler R. H., 2016, ApJ, 834, 37
Lelli F., McGaugh S. S., Schombert J. M., 2015, ApJ, 816, L14
Lelli F., McGaugh S. S., Schombert J. M., 2016, AJ, 152, 157
Lelli F., McGaugh S. S., Schombert J. M., Desmond H., Katz H., 2019, MNRAS, 484, 3267
Leroy A. K., Walter F., Brinks E., Bigiel F., de Blok W. J. G., Madore B., Thornley M. D., 2008, AJ, 136, 2782
Li C., Kauffmann G., Fu J., Wang J., Catinella B., Fabello S., Schiminovich D., Zhang W., 2012, MNRAS, 424, 1471
Li P., Lelli F., McGaugh S., Schombert J., 2020, ApJS, 247, 31
Li P., McGaugh S. S., Lelli F., Schombert J. M., Pawlowski M. S., 2022, A&A, 665, A143
Maddox N. et al., 2021, A&A, 646, A35
Mancera Piña P. E. et al., 2019, ApJ, 883, L33
Mancera Piña P. E., Fraternali F., Oosterloo T., Adams E. A. K., di Teodoro E., Bacchini C., Iorio G., 2022, MNRAS, 514, 3329
Mancera Piña P. E., Fraternali F., Oosterloo T., Adams E. A. K., Oman K. A., Leisman L., 2021, MNRAS, 512, 3230
Martin A. M., Giovanelli R., Haynes M. P., Guzzo L., 2012, ApJ, 750, 38
Martin D. C. et al., 2005, ApJ, 619, L1
Masters K. L., 2005, Proceedings of the 22nd Texas Symposium on Relativistic Astrophysics . p. 374
McGaugh S. S., Lelli F., Schombert J. M., 2020, RNAAS, 4, 45
McGaugh S. S., van Dokkum P., 2021, RNAAS, 5, 23

Méndez-Abreu J., Sánchez-Janssen R., Aguerri J. A. L., 2010, ApJ, 711, L61

Mo H., van den Bosch F. C., White S., 2010, Galaxy Formation and Evolution. Cambridge University Press, Cambridge

Mogotsi K. M., de Blok W. J. G., Caldu-Primo A., Walter F., Ianjamasimanana R., Leroy A. K., 2016, AJ, 151, 15

Moreno J. et al., 2022, Nat. Astron., 6, 496

Moster B. P., Somerville R. S., Maulbetsch C., van den Bosch F. C., Macciò A. V., Naab T., Oser L., 2010, ApJ, 710, 903

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

Naik A. P., Puchwein E., Davis A.-C., Sijacki D., Desmond H., 2019, MNRAS, 489, 771

Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493

Obreschkow D., Croton D., De Lucia G., Khochfar S., Rawlings S., 2009, ApJ, 698, 1467

Oh S.-H. et al., 2015, AJ, 149, 180

Oman K. A. et al., 2015, MNRAS, 452, 3650

Oosterloo T., Verheijen M., van Cappellen W., 2010, Proceedings of the ISKAF2010 Science Meeting. preprint (arXiv:1007.5141)

Papastergis E., Adams E. A. K., van der Hulst J. M., 2016, A&A, 593, A39

Papastergis E., Giovanelli R., Haynes M. P., Rodríguez-Puebla A., Jones M. G., 2013, ApJ, 776, 43

Papastergis E., Giovanelli R., Haynes M. P., Shankar F., 2015, A&A, 574, A113

Papastergis E., Shankar F., 2016, A&A, 591, A58

Paranjape A., Sheth R. K., 2021, MNRAS, 507, 632

Paranjape A., Srianand R., Choudhury T. R., Sheth R. K., 2021, preprint(astro-ph/2105.04570)

Pedregosa F. et al., 2011, J. Mach. Learn. Res., 12, 2825

Peñarrubia J., Pontzen A., Walker M. G., Koposov S. E., 2012, ApJ, 759, L42

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

Ponomareva A. A. et al., 2021, MNRAS, 508, 1195

Pontzen A., Governato F., 2014, Nature, 506, 171

Posti L., Fall S. M., 2021, A&A, 649, A119

Putko J., Almeida J. S., Munoz-Tunon C., Ramos A. A., Elmegreen B. G., Elmegreen D. M., 2019, ApJ, 883, 10

Read J. I., Gilmore G., 2005, MNRAS, 356, 107

Read J. I., Walker M. G., Steger P., 2019, MNRAS, 484, 1401

Reddick R. M., Wechsler R. H., Tinker J. L., Behroozi P. S., 2013, ApJ, 771, 30

Reynolds T. N., Westmeier T., Staveley-Smith L., Chauhan G., Lagos C. D. P., 2020, MNRAS, 493, 5089

Saifollahi T., Trujillo I., Beasley M. A., Peletier R. F., Knapen J. H., 2021, MNRAS, 502, 5921

Sales L. V., Wetzel A., Fattahi A., 2022, Nat. Astron., 6, 897

Salim S., Boquien M., Lee J. C., 2018, ApJ, 859, 11

Schechter P., 1976, ApJ, 203, 297

Schittenhelm D., Wacker P., 2021, preprint (arXiv:2005.08602)

Schulman E., Bregman J. N., Roberts M. S., 1994, ApJ, 423, 180

Schulz A. E., Mandelbaum R., Padmanabhan N., 2010, MNRAS, 408, 1463

Sellwood J. A., Sanders R. H., 2022, MNRAS, 514, 4008

Simard L. et al., 2002, ApJS, 142, 1

Springob C. M., Haynes M. P., Giovanelli R., Kent B. R., 2005, ApJS, 160, 149

Stiskalek R., Bartlett D. J., Desmond H., Anbajagane D., 2022, MNRAS, 514, 4026

Stiskalek R., Desmond H., Holvey T., Jones M. G., 2021, MNRAS, 506, 3205 (ST21)

Trujillo-Gomez S., Kruijssen J. M. D., Reina-Campos M., 2022, MNRAS, 510, 3356

Unterborn C. T., Ryden B. S., 2008, ApJ, 687, 976

van Dokkum P. et al., 2018, Nature, 555, 629

van Dokkum P. et al., 2022, Nature, 605, 435

van Dokkum P., Danieli S., Abraham R., Conroy C., Romanowsky A. J., 2019, ApJ, 874, L5

Velmani P., Paranjape A., 2023, MNRAS, 520, 2867

Verbeke R., Papastergis E., Ponomareva A. A., Rathi S., Rijcke S. D., 2017, A&A, 607, A13

Walter F., Brinks E., de Blok W. J. G., Bigiel F., Kennicutt J., Thornley M. D., Leroy A. K., 2008, AJ, 136, 2563

Wang J. et al., 2014, MNRAS, 441, 2159

Wang J., Koribalski B. S., Serra P., van der Hulst T., Roychowdhury S., Kamphuis P., Chengalur J. N., 2016, MNRAS, 460, 2143 (W16)

Watts A. B., Power C., Catinella B., Cortese L., Stevens A. R. H., 2020, MNRAS, 499, 5205

Wechsler R. H., Tinker J. L., 2018, ARA&A, 56, 435

Yahya S., Bull P., Santos M. G., Silva M., Maartens R., Okouma P., Bassett B., 2015, MNRAS, 450, 2251

Yu N., Ho L. C., Wang J., 2020, ApJ, 898, 102

Yu N., Ho L. C., Wang J., Li H., 2022, ApJ SupplementSeries, 261, 21

Zwaan M. A., Meyer M. J., Staveley-Smith L., 2010, MNRAS, 403, 1969

Zwicky F., 1933, Helvetica Phys. Acta, 6, 110

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