Predicting the Subhalo Mass Functions in Simulations from Galaxy Images

Andreas Filipp [1 2 3] **Tri Nguyen** [4 5] **Laurence Perreault-Levasseur** [1 2 3 6 7 8] **Jonah Rose** [9] **Chris Lovell** [10]
Nicloas Payot [1 2 3] **Francisco Villaescusa-Navarro** [11 12] **Yashar Hezaveh** [1 2 3 6 8]

Abstract

Strong gravitational lensing provides a powerful tool to directly infer the dark matter (DM) subhalo mass function (SHMF) in lens galaxies. However, comparing observationally inferred SHMFs to theoretical predictions remains challenging, as the predicted SHMF can vary significantly between galaxies — even within the same cosmological model — due to differences in the properties and environment of individual galaxies. We present a machine learning framework to infer the galaxy-specific SHMF from galaxy images, conditioned on the assumed inverse warm DM particle mass M_{DM}^{-1}. To train the model, we use 1024 high-resolution hydrodynamical zoom-in simulations from the DREAMS suite. Mock observations are generated using `Synthesizer`, excluding gas particle contributions, and SHMFs are computed with the `Rockstar` halo finder. Our neural network takes as input both the galaxy images and the inverse DM mass. This method enables scalable, image-based predictions for the theoretical DM SHMFs of individual galaxies, facilitating a direct comparisons with observational measurements.

1. Introduction

One of the most striking open questions in modern astrophysics is the nature of dark matter (DM), which constitutes approximately 80% of the universe's matter content (e.g., Hinshaw et al., 2013; Planck Collaboration et al., 2020). While its presence is inferred from gravitational phenomena across a wide range of cosmic scales, from galaxy clusters to large-scale structure, DM has yet to be detected through any non-gravitational interactions, and its fundamental properties remain unknown.

Different models predict distinct clustering behaviors for DM, especially on small, sub-galactic scales. On these scales, the distribution of dark matter — quantified by the subhalo mass function (SHMF) — is highly sensitive to its particle nature, making it a powerful discriminator between DM models (e.g., Ferreira, 2021). In warm DM (WDM) scenarios, for example, smaller particle masses correspond to higher thermal velocities, which suppress the formation of low-mass halos below the free-streaming scale (e.g., Colín et al., 2000; Gilman et al., 2020; Loudas et al., 2022).

Strong gravitational lensing offers a unique way to probe the distribution of matter on these small scales. Unlike methods that rely on luminous tracers, lensing is sensitive to all matter — luminous or dark — making it a powerful observational tool to constrain the SHMF. Traditional analyses infer the presence of individual subhalos by evaluating whether introducing localized perturbers to a smooth lens model leads to a statistically significant improvement in the fit to the observed lensed images (e.g., Vegetti et al., 2010; Hezaveh et al., 2016b). More recently, simulation-based studies have demonstrated that machine learning approaches could enable population-level inference of the SHMF by combining data across ensembles of lensing systems (e.g., Brehmer et al., 2020; 2019; Coogan et al., 2022; Zhang et al., 2022; Wagner-Carena et al., 2023; 2024; Zhang et al., 2024; Filipp et al., 2024).

However, connecting these observational constraints to theoretical predictions remains challenging. Even within a fixed cosmology, the SHMF depends on properties of individual galaxies, including, for example, total mass, morphology, merger history, and local environment (e.g., Hezaveh et al., 2016a), leading to significant system-to-system variation.

In this work, we introduce a machine learning framework to predict theoretical SHMFs directly from galaxy images, conditioned on an assumed WDM mass. This framework is used to place constraints on the possibilities of theoretically predicted SHMFs and aims to constrain theoretical predic-

[1]Department of Physics, University of Montreal, Montreal, Canada [2]MILA Quebec AI Institute, Montreal, Canada [3]CIELA Institute, Montreal Institute for Astrophysics and Machine Learning, Montreal, Canada [4]Center for Interdisciplinary Exploration and Research in Astrophysics, Northwestern University, Evanston, USA [5]The NSF-Simons AI Institute for the Sky, Chicago, USA [6]Center for Computational Astrophysics, Flatiron Institute, New York, USA [7]Perimeter Institute for Theoretical Physics, Waterloo, Canada [8]Trottier Space Institute, McGill University, Montreal, Canada [9]Center for Computational Astrophysics, New York, USA [10]University of Portsmouth, Portsmouth, UK [11]Princeton University, Princeton, USA [12]Simons Foundation, New York, USA. Correspondence to: Andreas Filipp <andreas.filipp@umontreal.ca>.

ML4Astro 2025, Vancouver, CA.

tions specific to galaxies that have observational constraints on the SHMF from strong gravitational lensing. This is not a stand-alone method, but needs constrains through other methods, since we predict the galaxy-specific SHMF given a specific WDM model. To make these predictions, we use the DREAMS simulation suite (Rose et al., 2025), as well as `Synthesizer` (Vijayan et al., 2020) to create realistic galaxy images. Our method accounts for inter-galaxy variability and enables scalable, image-based inference of theoretical predictions. This approach provides a new pathway to compare dark matter models with forthcoming lensing observations, enabling more precise, per-galaxy tests of DM's small-scale gravitational effects.

The paper is structured as follows: In Section 2, we describe the hydrodynamical simulation suite used in this work. Section 3 begins with an overview of how we created the galaxy images from the hydrodynamical simulations, then provides the assumed SHMF profile and the correlation with different DM models, as well as the neural network architecture used to make predictions. Section 4 presents our results, and we conclude in Section 5.

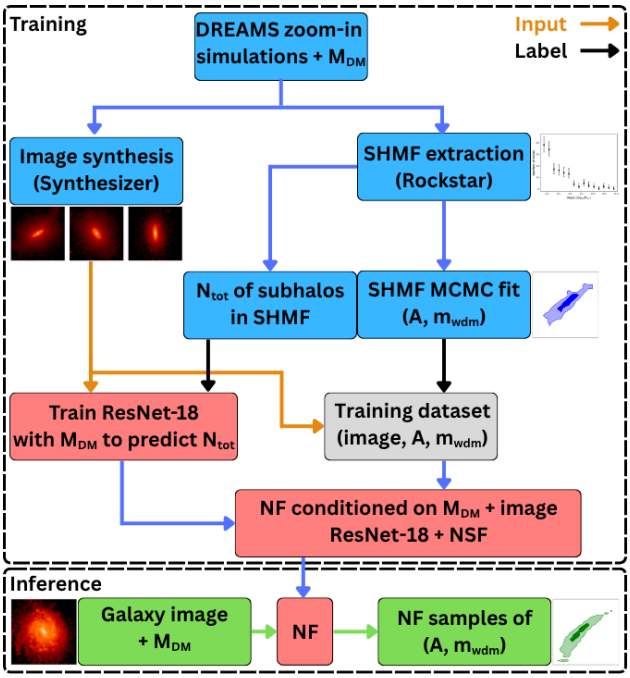

Figure 1. **A flowchart showing the training procedure, label generation, and inference.**

2. DREAMS Simulations

The DREAMS simulation suite contains, among other products, a set of high-resolution cosmological zoom-in hydrodynamic simulations designed to explore galaxy formation under varying DM and baryonic physics. Each zoom-in simulation is run using the `AREPO` code (Springel, 2010; Springel et al., 2019; Weinberger et al., 2020), enabling accurate modeling of complex baryonic processes such as gas cooling, star formation, and feedback. The initial conditions for each zoom-in are constructed by selecting a random, isolated Milky Way–mass halo from a low-resolution volume, then iteratively refining its Lagrangian region using intermediate- and high-resolution particle resampling to define the zoom-in domain (see Rose et al., 2025).

These zoom-in simulations are performed across a range of different WDM models in the range $M_{\rm DM} \in [1.8, 30.3]\,{\rm keV}$, sampled uniformly in the inverse $M_{\rm DM}^{-1}$. Further, the supernova (SN) wind, SN energy, and active galactic nuclei (AGN) parameters vary between each zoom-in simulation (Rose et al., 2025). The simulations are modeled using the IllustrisTNG baryonic physics prescriptions.

By varying both the DM physics and feedback parameters, the DREAMS simulations allow us to marginalize over baryonic uncertainties and learn a mapping from observable galaxy properties and WDM masses to the underlying SHMF. This marginalization ensures that our model generalizes across different astrophysical scenarios.

The DM sub-halos in the WDM zoom-in simulations are identified using the `Rockstar` halo finder (Behroozi et al., 2013). At the time of submitting this work, `Rockstar` halo catalogs were made available for 815 of the 1024 DREAMS zoom-in simulations. These constitute the labeled dataset used for supervised training of the SHMF inference model.

3. Methods

3.1. Image Generation with Synthesizer

Creating realistic galaxy images from hydrodynamic simulations typically requires computationally expensive radiative transfer simulations. As an efficient alternative, we use `Synthesizer` (Wilkins et al., 2020; Vijayan et al., 2020) to generate realistic observational mock images from simulated galaxies.

For each of the 815 zoom-in galaxies with available `Rockstar`-catalog, we generate 12 different projections by varying the line-of-sight orientation of the particles. `Synthesizer` produces galaxy images by generating spatially resolved spectral energy distributions (SEDs) and applying instrument-specific wavelength filters from the Spanish Virtual Observatory (SVO) filter service[1] (Rodrigo et al., 2012; Rodrigo & Solano, 2020; Rodrigo et al., 2024).

A detailed description of the image creation process can be found in Appendix A. Examples of the generated galaxy images are shown in Figure 2, alongside their corresponding galaxy-specific SHMFs.

[1] https://svo2.cab.inta-csic.es/theory/fps/

3.2. Dark Matter Subhalo Mass Function

We assume the functional form of the DM SHMFs to be a power-law with a slope of -0.9, as an approximation to theoretical predictions of cold DM (CDM) (e.g., Kuhlen et al., 2007; Diemand et al., 2007; Springel et al., 2008; Hiroshima et al., 2018). To model the subhalo mass function of WDM, we add to the power-law form of CDM a low-mass cutoff (e.g., Gilman et al., 2020):

$$\begin{aligned} \frac{dN}{dm}|_{\rm wdm} &= \frac{dN}{dm}|_{\rm CDM} \cdot \left(1 + \left(\frac{m_{\rm wdm}}{m}\right)\right)^{-1.3} \\ &= A \cdot m^{-0.9} \cdot \left(1 + \left(\frac{m_{\rm wdm}}{m}\right)\right)^{-1.3} \end{aligned} \quad (1)$$

where $m_{\rm wdm}$ is the cutoff mass, a characteristic of WDM scenarios, and the parameter A is the normalization.

We fit the parameterized WDM SHMF form to the data of the simulated galaxies, with the subhalos of the individual galaxies identified by `Rockstar`. We bin the identified subhalos in 15 mass bins, linearly spaced in logarithmic 10 base from $\log_{10}(M/M_\odot) = 7.75$ to 11. The uncertainties in the observed subhalo counts $\hat{n}_i$ are dominated by Poisson noise, see Appendix B. To fit the parameters of 1, we use the likelihood defined in eqn (9) in Appendix B:

$$\ln \mathcal{L}(\hat{n} \mid \theta) = -\frac{1}{2} \sum_i \frac{\left(\hat{n}_i - n_i(\theta)\right)^2}{V_i \; - \; V_i' \left(\hat{n}_i - n_i(\theta)\right)} \,, \quad (2)$$

where the index i runs over the mass bins, $n_i(\theta)$ is the model prediction of eqn (1), given the parameters $\theta = (A, m_{\rm wdm})$. We use the `emcee` package (Foreman-Mackey et al., 2013) to perform Markov Chain Monte Carlo (MCMC) sampling over this likelihood. This yields posterior samples for the amplitude A and the WDM cutoff mass $m_{\rm wdm}$. We then use the samples $(A, m_{\rm wdm})$ obtained in this way to train a normalizing flow (NF), which allows us to account for correlations between A and $m_{\rm wdm}$. We have access to the MCMC samples, because we know the underlying truth from the DREAMS simulations; performing the MCMC fits from observational data is not possible.

For each zoom-in galaxy, we use 100 posterior samples from the fitted distribution of $(A, m_{\rm wdm})$ as training labels for the NF. This ensures that the NF learns the continuous distribution of plausible parameters for A and $m_{\rm wdm}$, rather than a single point estimate. As the flowchart of in Figure 1 illustrates, the MCMC samples are only used during training to help the NF learn the continuous distributions of A and $m_{\rm wdm}$ conditioned on the galaxy morphologies and $M_{\rm DM}$. The framework is not aimed to be used on survey wide data, but to compare observational determined SHMFs from strong lenses with theoretical predictions of the SHMF of the lens galaxy.

3.3. Network Architecture and Training

Our architecture consists of two main components: a convolutional neural network (CNN) to process the image of the galaxy and a conditional NF.

Within the same cosmological model the predicted SHMF depends on the properties of each individual galaxy. Our architecture infers the relation between the observable properties of the galaxies and their SHMF, given the WDM mass. Figure 1 shows a flowchart of the training and inference scheme we used. We first pre-train a ResNet-18 with a two layer multi-layer perceptron (MLP) to predict the number of subhalos in the galaxy images, with the inverse WDM mass $M_{\rm DM}^{-1}$ concatenated to the first of the two MLP layers, using a mean squared error (MSE) loss. Our pretraining ensures a meaningful embedding space for the images, before using the output of the pretrained CNN as a condition for the NF.

The embedded image and the inverse of the WDM mass $M_{\rm DM}^{-1}$ are used as condition for a neural spline flow (NSF) to predict the parameters A and $m_{\rm WDM}$. We use the `zuko` library to implement the NSF (Rozet et al., 2022). We use 730 of the zoom-in galaxies as training data and keep 85 as validation data, for both - the pretraining of the CNN and the training of the NSF - the same sets.

To train the full model all the weights of the ResNet are allowed to update, enabling end-to-end optimization of the feature extractor and density estimator. This allows us to efficiently model the posterior over the SHMF fit parameters conditioned on both the image and the WDM mass.

4. Results

Figure 2 shows galaxy-specific SHMFs with the corresponding galaxy image. We compare the NF conditioned on the galaxy images and $M_{\rm DM}$, and a NF conditioned only on $M_{\rm DM}$. This gives an estimate how including the morphological information of the galaxy improves the SHMF prediction in comparison pure theoretical predictions. The network is used for galaxy-specific predictions of the SHMF under the assumption of different WDM masses, therefore we only show galaxy specific results. For each shown galaxy we only have access to the SHMF under one specific WDM temperature, therefore we cannot probe the same galaxy under different WDM temperatures against the ground truth. We achieve an improvement in constraints on the SHMF by including information about the galaxy morphology. The displayed galaxies come from three different WDM mass regions. The examples on the left come from higher WDM masses, the one in the middle from medium masses, and the right side from low WDM masses. The black dots are the subhalo counts from the DREAMS simulations with Poisson uncertainties. The red and blue solid lines shows the 50th percentile fit of the NF samples conditioned on the

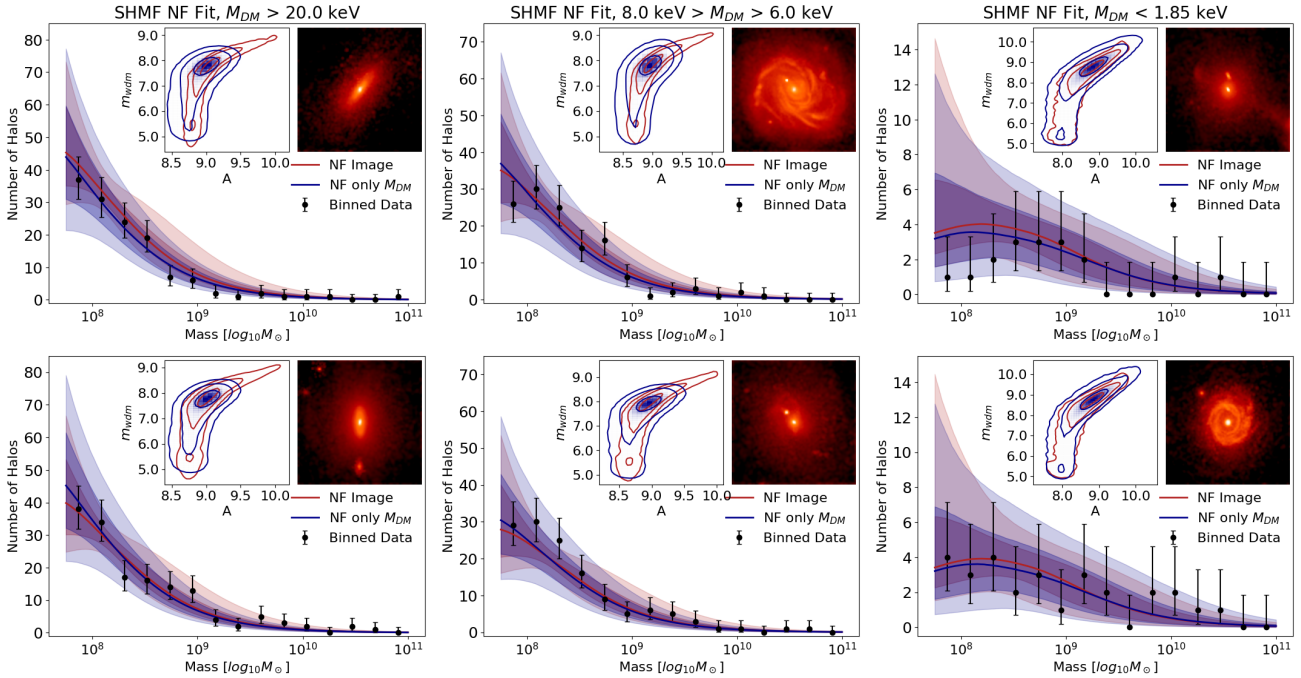

Figure 2. **The galaxy specific halo mass functions under different cosmologies.** The figure shows the galaxy specific halo mass functions for different WDM masses, with the corresponding galaxy image in the plot. The displayed galaxies come from three different WDM mass regions. The black dots are the counts of subhalos from the DREAMS simulation with Poisson uncertainties. The blue contours show the samples of the NF conditioned only on $M_{\rm DM}$, and the red contours the samples of the NF conditioned on both - $M_{\rm DM}$ and the galaxy image. The NF conditioned only on $M_{\rm DM}$ shows a broader posterior range and bigger uncertainties on the fit.

Table 1. **PQMass-χ^2 and RSME values**. The table shows the mean χ^2 values obtained with PQMass, as well as the median RSME, for the NF conditioned on images and $M_{\rm DM}$, and a NF conditioned only on $M_{\rm DM}$ for the entire validation set. The comparison is made between the MCMC samples and the different NF samples for A and $m_{\rm wdm}$.

	NF with image	NF without image
PQM χ^2	343.7 ± 92.5	385.8 ± 121.7
Median RMSE	0.26 ± 0.20	0.55 ± 0.17

image and the samples of the NF only conditioned on $M_{\rm DM}$ respectively. The shaded regions show the 1σ and 2σ fits of the SHMF from eqn 1 using the sampled parameters. The displayed contours show the 1σ, 2σ, and 3σ contours of the parameter samples $(A, m_{\rm wdm})$. The titles indicate the WDM mass range, which is passed along with the image to the neural network. For each range of $M_{\rm DM}$ we show two examples in the same column. The NF conditioned only on $M_{\rm DM}$ shows broader posteriors and bigger uncertainties on the parameters of the SHMF. Including the image information allows tighter constraints on both parameters, A and $m_{\rm wdm}$, and highlights their correlations, resulting in tighter constraints of the SHMF. This supports the conclusion that incorporating galaxy images enables more precise predictions of the SHMF fit parameters than using the WDM mass $M_{\rm DM}$ alone.

Further, we test the NF conditioned on the galaxy images and $M_{\rm DM}$ against the NF only conditioned on $M_{\rm DM}$ qualitatively, to show that the inclusion of the galaxy images leads to a better performance of the NF. In Table 1 we report the mean and standard deviation of the PQMass-χ^2 values (Lemos et al., 2024) for the entire validation set for the NF conditioned on images and $M_{\rm DM}$, and the NF only conditioned on $M_{\rm DM}$. We also report the median root mean squared error (RMSE) and its standard deviation of the entire validation set for both network cases. For that we use for each condition the MCMC samples of the SHMF fit as target distribution and compare those against the NF samples of A and $m_{\rm wdm}$. The closer the distributions are, the lower is the PQMass-χ^2 value. We do not expect the NF samples to fit the MCMC fits perfectly, because each individual galaxy samples a distribution of possible SHMF for a given WDM model, making the inference under-specified without additional information. Providing additionally galaxy images leads to a more informative inference of the possible SHMF, since they provide more constraining information. The lower RMSE value for the normalizing flow with access to the morphological information shows that the performance of the architecture improves and has tighter constraints on the SHMF. Additionally, we compare the MCMC fits of the SHMFs and the NF samples of the NF with image conditioning in Figure 3 in Appendix C. We do not compare the NF samples here with the MCMC samples, because the goal is to show that the NF with access to the image data does outperform the pure theoretical NF predictions based on the WDM only without image access.

The current sample size of the simulation is too small to have a separate test set, therefore we show the performance for the validation set. The provided examples show that the network is able to learn the galaxy specific fit parameters within the predicted uncertainties.

5. Discussion

Future iterations of this work will focus on increasing the realism of the input images and expanding the training dataset. The current images do not include the effects of dust attenuation, assume only Gaussian noise, and use a simplified Gaussian point spread function (PSF). Additionally, the images are expressed in flux units $(\mathrm{erg}, \mathrm{s}^{-1}, \mathrm{Hz}^{-1})$ rather than in instrument-specific units. We plan to convert them to the native units of the Hubble Space Telescope (HST), i.e., $\mathrm{e}^{-}, \mathrm{s}^{-1}$, and to incorporate non-Gaussian, instrument-specific noise using SLIC (Legin et al., 2023), along with more realistic PSF models. These improvements will be accompanied by the inclusion of the full set of `Rockstar` halo catalogs, including those for the remaining 209 DREAMS zoom-in simulations, to expand the training and validation sets and enable evaluation on an independent test set.

The results demonstrate strong performance and highlight the potential of applying this architecture to more realistic simulations. However, there are important limitations to note regarding the training data. All galaxies used during training and validation are Milky Way-mass halos in isolated systems. This constraint is due to the availability of high-resolution DM zoom-in simulations, which currently exist only for these galaxy types, since running such detailed zoom-ins on cluster galaxies is computationally too expensive. Whilst the masses of the galaxies are for all galaxies in the order of the Milky Way, they are approximately equally distributed between spiral and elliptical galaxies. In contrast, strong gravitational lenses are more commonly observed to be massive galaxies in dense environments - often the central galaxies of clusters. These systems are not represented in the necessary mass resolutions in current hydrodynamical simulations. For these purposes the subhalos should be resolved down to masses of $\log_{10}(M/M_{\odot}) = 7.75$ and below to be able to resolve small subhalos and confidentially infer the SHMF.

We expect that future high-resolution hydrodynamical simulations will model the relevant environments in greater detail, allowing us to extend the training set and improve generalization to observed lensing galaxies. Until then, our results should be interpreted within this limitation.

Acknowledgements

This work is partially supported by Schmidt Futures, a philanthropic initiative founded by Eric and Wendy Schmidt as part of the Virtual Institute for Astrophysics (VIA). The work is in part supported by computational resources provided by Calcul Quebec and the Digital Research Alliance of Canada. Y.H. and L.P. acknowledge support from the Canada Research Chairs Program, the National Sciences and Engineering Council of Canada through grants RGPIN-2020-05073 and 05102. A.F. acknowledges the support from the Bourse J. Armand Bombardier and UdeMs final year scholarship. TN is supported by the CIERA Postdoctoral Fellowship.

References

Behroozi, P. S., Wechsler, R. H., and Wu, H.-Y. The ROCKSTAR Phase-space Temporal Halo Finder and the Velocity Offsets of Cluster Cores. , 762(2):109, January 2013. doi: 10.1088/0004-637X/762/2/109.

Brehmer, J., Mishra-Sharma, S., Hermans, J., Louppe, G., and Cranmer, K. Mining for Dark Matter Substructure: Inferring Subhalo Population Properties from Strong Lenses with Machine Learning. , 886(1):49, November 2019. doi: 10.3847/1538-4357/ab4c41.

Brehmer, J., Louppe, G., Pavez, J., and Cranmer, K. Mining gold from implicit models to improve likelihood-free inference. *Proceedings of the National Academy of Sciences*, 117(10):5242–5249, 2020. doi: 10.1073/pnas.1915980117. URL https://www.pnas.org/doi/abs/10.1073/pnas.1915980117.

Chang, L. J. and Necib, L. Dark matter density profiles in dwarf galaxies: linking Jeans modelling systematics and observation. , 507(4):4715–4733, November 2021. doi: 10.1093/mnras/stab2440.

Colín, P., Avila-Reese, V., and Valenzuela, O. Substructure and Halo Density Profiles in a Warm Dark Matter Cosmology. , 542(2):622–630, October 2000. doi: 10.1086/317057.

Coogan, A., Montel, N. A., Karchev, K., Grootes, M. W., Nattino, F., and Weniger, C. One never walks alone: the effect of the perturber population on subhalo measurements in strong gravitational lenses, 2022.

Diemand, J., Kuhlen, M., and Madau, P. Formation and Evolution of Galaxy Dark Matter Halos and Their Substructure. , 667(2):859–877, October 2007. doi: 10.1086/520573.

Ferreira, E. G. M. Ultra-light dark matter. , 29(1):7, December 2021. doi: 10.1007/s00159-021-00135-6.

Filipp, A., Hezaveh, Y., and Perreault-Levasseur, L. Robustness of Neural Ratio and Posterior Estimators to Distributional Shifts for Population-Level Dark Matter Analysis in Strong Gravitational Lensing. *arXiv e-prints*, art. arXiv:2411.05905, November 2024. doi: 10.48550/arXiv.2411.05905.

Foreman-Mackey, D., Hogg, D. W., Lang, D., and Goodman, J. emcee: The MCMC Hammer. , 125(925):306, March 2013. doi: 10.1086/670067.

Gilman, D., Birrer, S., Nierenberg, A., Treu, T., Du, X., and Benson, A. Warm dark matter chills out: constraints on the halo mass function and the free-streaming length of dark matter with eight quadruple-image strong gravitational lenses. , 491(4):6077–6101, February 2020. doi: 10.1093/mnras/stz3480.

Hezaveh, Y., Dalal, N., Holder, G., Kisner, T., Kuhlen, M., and Perreault Levasseur, L. Measuring the power spectrum of dark matter substructure using strong gravitational lensing. , 2016(11):048, November 2016a. doi: 10.1088/1475-7516/2016/11/048.

Hezaveh, Y. D., Dalal, N., Marrone, D. P., Mao, Y.-Y., Morningstar, W., Wen, D., Blandford, R. D., Carlstrom, J. E., Fassnacht, C. D., Holder, G. P., Kemball, A., Marshall, P. J., Murray, N., Perreault Levasseur, L., Vieira, J. D., and Wechsler, R. H. Detection of Lensing Substructure Using ALMA Observations of the Dusty Galaxy SDP.81. , 823 (1):37, May 2016b. doi: 10.3847/0004-637X/823/1/37.

Hinshaw, G., Larson, D., Komatsu, E., Spergel, D. N., Bennett, C. L., Dunkley, J., Nolta, M. R., Halpern, M., Hill, R. S., Odegard, N., Page, L., Smith, K. M., Weiland, J. L., Gold, B., Jarosik, N., Kogut, A., Limon, M., Meyer, S. S., Tucker, G. S., Wollack, E., and Wright, E. L. Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Parameter Results. , 208(2):19, October 2013. doi: 10.1088/0067-0049/208/2/19.

Hiroshima, N., Ando, S., and Ishiyama, T. Modeling evolution of dark matter substructure and annihilation boost. , 97(12):123002, June 2018. doi: 10.1103/PhysRevD.97.123002.

Kuhlen, M., Diemand, J., and Madau, P. The Shapes, Orientation, and Alignment of Galactic Dark Matter Subhalos. , 671(2):1135–1146, December 2007. doi: 10.1086/522878.

Legin, R., Adam, A., Hezaveh, Y., and Perreault-Levasseur, L. Beyond Gaussian Noise: A Generalized Approach to Likelihood Analysis with Non-Gaussian Noise. , 949(2): L41, June 2023. doi: 10.3847/2041-8213/acd645.

Lemos, P., Sharief, S., Malkin, N., Salhi, S., Stone, C., Perreault-Levasseur, L., and Hezaveh, Y. PQMass: Probabilistic Assessment of the Quality of Generative Models using Probability Mass Estimation. *arXiv e-prints*, art. arXiv:2402.04355, February 2024. doi: 10.48550/arXiv.2402.04355.

Loudas, N., Pavlidou, V., Casadio, C., and Tassis, K. Discriminating power of milli-lensing observations for dark matter models. , 668:A166, December 2022. doi: 10.1051/0004-6361/202244978.

Planck Collaboration, Aghanim, N., Akrami, Y., Ashdown, M., Aumont, J., Baccigalupi, C., Ballardini, M., Banday, A. J., Barreiro, R. B., Bartolo, N., Basak, S., Battye, R., Benabed, K., Bernard, J. P., Bersanelli, M., Bielewicz, P., Bock, J. J., Bond, J. R., Borrill, J., Bouchet, F. R., Boulanger, F., Bucher, M., Burigana, C., Butler, R. C., Calabrese, E., Cardoso, J. F., Carron, J., Challinor, A., Chiang, H. C., Chluba, J., Colombo, L. P. L., Combet, C., Contreras, D., Crill, B. P., Cuttaia, F., de Bernardis, P., de Zotti, G., Delabrouille, J., Delouis, J. M., Di Valentino, E., Diego, J. M., Doré, O., Douspis, M., Ducout, A., Dupac, X., Dusini, S., Efstathiou, G., Elsner, F., Enßlin, T. A., Eriksen, H. K., Fantaye, Y., Farhang, M., Fergusson, J., Fernandez-Cobos, R., Finelli, F., Forastieri, F., Frailis, M., Fraisse, A. A., Franceschi, E., Frolov, A., Galeotta, S., Galli, S., Ganga, K., Génova-Santos, R. T., Gerbino, M., Ghosh, T., González-Nuevo, J., Górski, K. M., Gratton, S., Gruppuso, A., Gudmundsson, J. E., Hamann, J., Handley, W., Hansen, F. K., Herranz, D., Hildebrandt, S. R., Hivon, E., Huang, Z., Jaffe, A. H., Jones, W. C., Karakci, A., Keihänen, E., Keskitalo, R., Kiiveri, K., Kim, J., Kisner, T. S., Knox, L., Krachmalnicoff, N., Kunz, M., Kurki-Suonio, H., Lagache, G., Lamarre, J. M., Lasenby, A., Lattanzi, M., Lawrence, C. R., Le Jeune, M., Lemos, P., Lesgourgues, J., Levrier, F., Lewis, A., Liguori, M., Lilje, P. B., Lilley, M., Lindholm, V., López-Caniego, M., Lubin, P. M., Ma, Y. Z., Macías-Pérez, J. F., Maggio, G., Maino, D., Mandolesi, N., Mangilli, A., Marcos-Caballero, A., Maris, M., Martin, P. G., Martinelli, M., Martínez-González, E., Matarrese, S., Mauri, N., McEwen, J. D., Meinhold, P. R., Melchiorri, A., Mennella, A., Migliaccio, M., Millea, M., Mitra, S., Miville-Deschênes, M. A., Molinari, D., Montier, L., Morgante, G., Moss, A., Natoli, P., Nørgaard-Nielsen, H. U., Pagano, L., Paoletti, D., Partridge, B., Patanchon, G., Peiris, H. V., Perrotta, F., Pettorino, V., Piacentini, F., Polastri, L., Polenta, G., Puget, J. L., Rachen, J. P., Reinecke, M., Remazeilles, M., Renzi, A., Rocha, G., Rosset, C., Roudier, G., Rubiño-Martín, J. A., Ruiz-Granados, B., Salvati, L., Sandri, M., Savelainen, M., Scott, D., Shellard, E. P. S., Sirignano, C., Sirri, G., Spencer, L. D., Sunyaev, R., Suur-Uski, A. S., Tauber, J. A., Tavagnacco, D., Tenti, M., Toffolatti, L., Tomasi, M., Trombetti, T., Valenziano,

L., Valiviita, J., Van Tent, B., Vibert, L., Vielva, P., Villa, F., Vittorio, N., Wandelt, B. D., Wehus, I. K., White, M., White, S. D. M., Zacchei, A., and Zonca, A. Planck 2018 results. VI. Cosmological parameters. , 641:A6, September 2020. doi: 10.1051/0004-6361/201833910.

Rodrigo, C. and Solano, E. The SVO Filter Profile Service. In *XIV.0 Scientific Meeting (virtual) of the Spanish Astronomical Society*, pp. 182, July 2020.

Rodrigo, C., Solano, E., and Bayo, A. SVO Filter Profile Service Version 1.0. IVOA Working Draft 15 October 2012, October 2012.

Rodrigo, C., Cruz, P., Aguilar, J. F., Aller, A., Solano, E., Gálvez-Ortiz, M. C., Jiménez-Esteban, F., Mas-Buitrago, P., Bayo, A., Cortés-Contreras, M., Murillo-Ojeda, R., Bonoli, S., Cenarro, J., Dupke, R., López-Sanjuan, C., Marín-Franch, A., de Oliveira, C. M., Moles, M., Taylor, K., Varela, J., and Ramió, H. V. Photometric segregation of dwarf and giant FGK stars using the SVO Filter Profile Service and photometric tools. , 689:A93, September 2024. doi: 10.1051/0004-6361/202449998.

Rose, J. C., Torrey, P., Villaescusa-Navarro, F., Lisanti, M., Nguyen, T., Roy, S., Kollmann, K. E., Vogelsberger, M., Cyr-Racine, F.-Y., Medvedev, M. V., Genel, S., Anglés-Alcázar, D., Kallivayalil, N., Wang, B. Y., Costanza, B., O'Neil, S., Roche, C., Karmakar, S., Garcia, A. M., Low, R., Lin, S., Mostow, O., Cruz, A., Caputo, A., Farahi, A., Muñoz, J. B., Necib, L., Teyssier, R., Dalcanton, J. J., and Spergel, D. Introducing the DREAMS Project: DaRk mattEr and Astrophysics with Machine Learning and Simulations. , 982(2):68, April 2025. doi: 10.3847/1538-4357/adb8e5.

Rozet, F. et al. Zuko: Normalizing flows in pytorch, 2022. URL https://pypi.org/project/zuko.

Springel, V. E pur si muove: Galilean-invariant cosmological hydrodynamical simulations on a moving mesh. , 401 (2):791–851, January 2010. doi: 10.1111/j.1365-2966.2009.15715.x.

Springel, V., Wang, J., Vogelsberger, M., Ludlow, A., Jenkins, A., Helmi, A., Navarro, J. F., Frenk, C. S., and White, S. D. M. The Aquarius Project: the subhaloes of galactic haloes. , 391(4):1685–1711, December 2008. doi: 10.1111/j.1365-2966.2008.14066.x.

Springel, V., Pakmor, R., and Weinberger, R. AREPO: Cosmological magnetohydrodynamical moving-mesh simulation code. Astrophysics Source Code Library, record ascl:1909.010, September 2019.

Tanabashi, M., Hagiwara, K., Hikasa, K., Nakamura, K., Sumino, Y., Takahashi, F., Tanaka, J., Agashe, K., Aielli, G., Amsler, C., Antonelli, M., Asner, D. M., Baer, H., Banerjee, S., Barnett, R. M., Basaglia, T., Bauer, C. W., Beatty, J. J., Belousov, V. I., Beringer, J., Bethke, S., Bettini, A., Bichsel, H., Biebel, O., Black, K. M., Blucher, E., Buchmuller, O., Burkert, V., Bychkov, M. A., Cahn, R. N., Carena, M., Ceccucci, A., Cerri, A., Chakraborty, D., Chen, M. C., Chivukula, R. S., Cowan, G., Dahl, O., D'Ambrosio, G., Damour, T., de Florian, D., de Gouvêa, A., DeGrand, T., de Jong, P., Dissertori, G., Dobrescu, B. A., D'Onofrio, M., Doser, M., Drees, M., Dreiner, H. K., Dwyer, D. A., Eerola, P., Eidelman, S., Ellis, J., Erler, J., Ezhela, V. V., Fetscher, W., Fields, B. D., Firestone, R., Foster, B., Freitas, A., Gallagher, H., Garren, L., Gerber, H. J., Gerbier, G., Gershon, T., Gershtein, Y., Gherghetta, T., Godizov, A. A., Goodman, M., Grab, C., Gritsan, A. V., Grojean, C., Groom, D. E., Grünewald, M., Gurtu, A., Gutsche, T., Haber, H. E., Hanhart, C., Hashimoto, S., Hayato, Y., Hayes, K. G., Hebecker, A., Heinemeyer, S., Heltsley, B., Hernández-Rey, J. J., Hisano, J., Höcker, A., Holder, J., Holtkamp, A., Hyodo, T., Irwin, K. D., Johnson, K. F., Kado, M., Karliner, M., Katz, U. F., Klein, S. R., Klempt, E., Kowalewski, R. V., Krauss, F., Kreps, M., Krusche, B., Kuyanov, Y. V., Kwon, Y., Lahav, O., Laiho, J., Lesgourgues, J., Liddle, A., Ligeti, Z., Lin, C. J., Lippmann, C., Liss, T. M., Littenberg, L., Lugovsky, K. S., Lugovsky, S. B., Lusiani, A., Makida, Y., Maltoni, F., Mannel, T., Manohar, A. V., Marciano, W. J., Martin, A. D., Masoni, A., Matthews, J., Meißner, U. G., Milstead, D., Mitchell, R. E., Mönig, K., Molaro, P., Moortgat, F., Moskovic, M., Murayama, H., Narain, M., Nason, P., Navas, S., Neubert, M., Nevski, P., Nir, Y., Olive, K. A., Pagan Griso, S., Parsons, J., Patrignani, C., Peacock, J. A., Pennington, M., Petcov, S. T., Petrov, V. A., Pianori, E., Piepke, A., Pomarol, A., Quadt, A., Rademacker, J., Raffelt, G., Ratcliff, B. N., Richardson, P., Ringwald, A., Roesler, S., Rolli, S., Romaniouk, A., Rosenberg, L. J., Rosner, J. L., Rybka, G., Ryutin, R. A., Sachrajda, C. T., Sakai, Y., Salam, G. P., Sarkar, S., Sauli, F., Schneider, O., Scholberg, K., Schwartz, A. J., Scott, D., Sharma, V., Sharpe, S. R., Shutt, T., Silari, M., Sjöstrand, T., Skands, P., Skwarnicki, T., Smith, J. G., Smoot, G. F., Spanier, S., Spieler, H., Spiering, C., Stahl, A., Stone, S. L., Sumiyoshi, T., Syphers, M. J., Terashi, K., Terning, J., Thoma, U., Thorne, R. S., Tiator, L., Titov, M., Tkachenko, N. P., Törnqvist, N. A., Tovey, D. R., Valencia, G., Van de Water, R., Varelas, N., and Venanzoni, G. Review of Particle Physics*. , 98(3):030001, August 2018. doi: 10.1103/PhysRevD.98.030001.

Vegetti, S., Koopmans, L. V. E., Bolton, A., Treu, T., and Gavazzi, R. Detection of a dark substructure through gravitational imaging. , 408(4):1969–1981, November 2010. doi: 10.1111/j.1365-2966.2010.16865.x.

Vijayan, A. P., Lovell, C. C., Wilkins, S. M., Thomas,

P. A., Barnes, D. J., Irodotou, D., Kuusisto, J., and Roper, W. J. First Light And Reionization Epoch Simulations (FLARES) – II: The photometric properties of high-redshift galaxies. *Monthly Notices of the Royal Astronomical Society*, 501(3):3289–3308, 11 2020. ISSN 0035-8711. doi: 10.1093/mnras/staa3715. URL https://doi.org/10.1093/mnras/staa3715.

Wagner-Carena, S., Aalbers, J., Birrer, S., Nadler, E. O., Darragh-Ford, E., Marshall, P. J., and Wechsler, R. H. From Images to Dark Matter: End-to-end Inference of Substructure from Hundreds of Strong Gravitational Lenses. , 942(2):75, January 2023. doi: 10.3847/1538-4357/aca525.

Wagner-Carena, S., Lee, J., Pennington, J., Aalbers, J., Birrer, S., and Wechsler, R. H. A Strong Gravitational Lens Is Worth a Thousand Dark Matter Halos: Inference on Small-Scale Structure Using Sequential Methods. *arXiv e-prints*, art. arXiv:2404.14487, April 2024. doi: 10.48550/arXiv.2404.14487.

Weinberger, R., Springel, V., and Pakmor, R. The AREPO Public Code Release. , 248(2):32, June 2020. doi: 10.3847/1538-4365/ab908c.

Wilkins, S. M., Lovell, C. C., Fairhurst, C., Feng, Y., Matteo, T. D., Croft, R., Kuusisto, J., Vijayan, A. P., and Thomas, P. Nebular-line emission during the Epoch of Reionization. *Monthly Notices of the Royal Astronomical Society*, 493(4):6079–6094, 03 2020. ISSN 0035-8711. doi: 10.1093/mnras/staa649. URL https://doi.org/10.1093/mnras/staa649.

Zhang, G., Mishra-Sharma, S., and Dvorkin, C. Inferring subhalo effective density slopes from strong lensing observations with neural likelihood-ratio estimation. , 517(3):4317–4326, December 2022. doi: 10.1093/mnras/stac3014.

Zhang, G., Şengül, A. Ç., and Dvorkin, C. Subhalo effective density slope measurements from HST strong lensing data with neural likelihood-ratio estimation. , 527(2):4183–4192, January 2024. doi: 10.1093/mnras/stad3521.

A. Image Generation with Synthesizer

`Synthesizer` produces galaxy images by generating spatially resolved spectral energy distributions (SEDs) and applying instrument-specific wavelength filters from the Spanish Virtual Observatory (SVO) filter service[2] (Rodrigo et al., 2012; Rodrigo & Solano, 2020; Rodrigo et al., 2024). We use the Hubble Space Telescope Wide Field Camera (HST/WFC) F105W filter to simulate realistic near-infrared observations.

The synthetic images are constructed with only taking stellar particles into account. We neglect the contributions of dust and gas particles, which are expected to have a relatively minor impact on the morphological features relevant to our analysis. In future steps, we will include the effects of dust and interstellar gas. As first step of the image creation we use an incident emission model and place the galaxies at redshift $z = 0$.

The particles are convolved with a smoothing filter for optical more appealing visualization. The smoothing length of the individual stellar particles are taken for each particle from the simulation data directly, which means that any artifacts of isolated particles in the imaging are originating from the simulation. The smoothing length is the co-moving radius of the sphere centered on the particle enclosing the 32 ± 1 nearest particles of this same type.

To ensure image fidelity, we first generate high-resolution images and convolve them at the high resolution with a Gaussian point spread function (PSF) to approximate observational effects. The PSF has a full width half maximum of 3 pixels. The resulting images are then downsampled to a target resolution of 0.78125 kpc/pixel. We add gaussian noise of $10^{23} \frac{\text{erg}}{\text{s}\cdot\text{Hz}}$. We do not yet include instrument specific observational noise or Poisson noise. Examples of the generated galaxy images are shown in Figure 2, alongside the corresponding galaxy-specific SHMFs.

B. Poisson Noise from Samples

To compute the Poisson noise of the counts of each bin of the SHMF we get using the `Rockstar` catalogs, we follow the prescription in Tanabashi et al. (2018); Chang & Necib (2021). We compute the confidence intervals using the inverse cumulative distribution function of the χ^2 distribution:

$$\mu_{\text{lower}} = \frac{1}{2} F^{-1}_{\chi^2}\Big(\tfrac{\alpha}{2};\, 2\hat{n}\Big) \tag{3}$$

$$\mu_{\text{higher}} = \frac{1}{2} F^{-1}_{\chi^2}\Big(1 - \tfrac{\alpha}{2};\, 2(\hat{n}+1)\Big) \tag{4}$$

with $F^{-1}_{\chi^2}$ the inverse of the χ^2 cumulative distribution and $\hat{n}$ the number of counts. The confidence level is given by $100(1-\alpha)$. From these intervals, we define the asymmetric error bars and variances with:

$$\sigma_{\text{lower},i} = \hat{n}_i \;-\; \mu_{\text{lower},i} \tag{5}$$

$$\sigma_{\text{higher},i} = \mu_{\text{higher},i} \;-\; \hat{n}_i \tag{6}$$

$$V_i = \sigma_{\text{lower},i}\, \sigma_{\text{higher},i} \tag{7}$$

$$V_i' = \sigma_{\text{higher},i} \;-\; \sigma_{\text{lower},i} \tag{8}$$

We use the variances and uncertainties on the counts per bin to compute the likelihood of our fits to the data.

The likelihood function to fit the SHMF (eqn 1) is defined by:

$$\ln \mathcal{L}(\hat{n} \mid \theta) = -\frac{1}{2} \sum_i \frac{\big(\hat{n}_i - n_i(\theta)\big)^2}{V_i \;-\; V_i' \big(\hat{n}_i - n_i(\theta)\big)} \tag{9}$$

where $n_i(\theta)$ is the model prediction of eqn (1), given the parameters $\theta = (A, m_{\text{wdm}})$. We use the likelihood function to fit the `Rockstar`-catalog data with an MCMC.

C. Comparison of Normalizing Flow Samples with MCMC Fits

In Figure 3 we show galaxy-specific SHMFs obtained by NF samples and MCMC fits along with the corresponding galaxy image. The figure is structured as Figure 2. We use the same galaxies as before, but under different projections. We do not

[2] https://svo2.cab.inta-csic.es/theory/fps/

expect to perfectly recreate the MCMC fits. Our method serves as a tool to get the theoretical predictions on the SHMF given a WDM theory. This does not serve as constraints on the SHMF, but is rather to be used to compare theoretical predictions with observational constraints that can be achieved with strong lensing results. This is not a stand alone method to learn about dark matter in observational surveys, but needs a comparison counterpart that constrained the SHMF. The blue and red solid contours show the samples of the MCMC fit and of the NF samples respectively. The dark and light shaded areas are the 1σ and 2σ regions of the fits. The displayed contours show the 1σ, 2σ, and 3σ contours of the MCMC fits and NF samples of $(A, m_{\rm wdm})$, defining the SHMF.

The sampled parameters of the NF represent the fits obtained directly through the likelihood and cover the full range of the parameters sampled with MCMC.

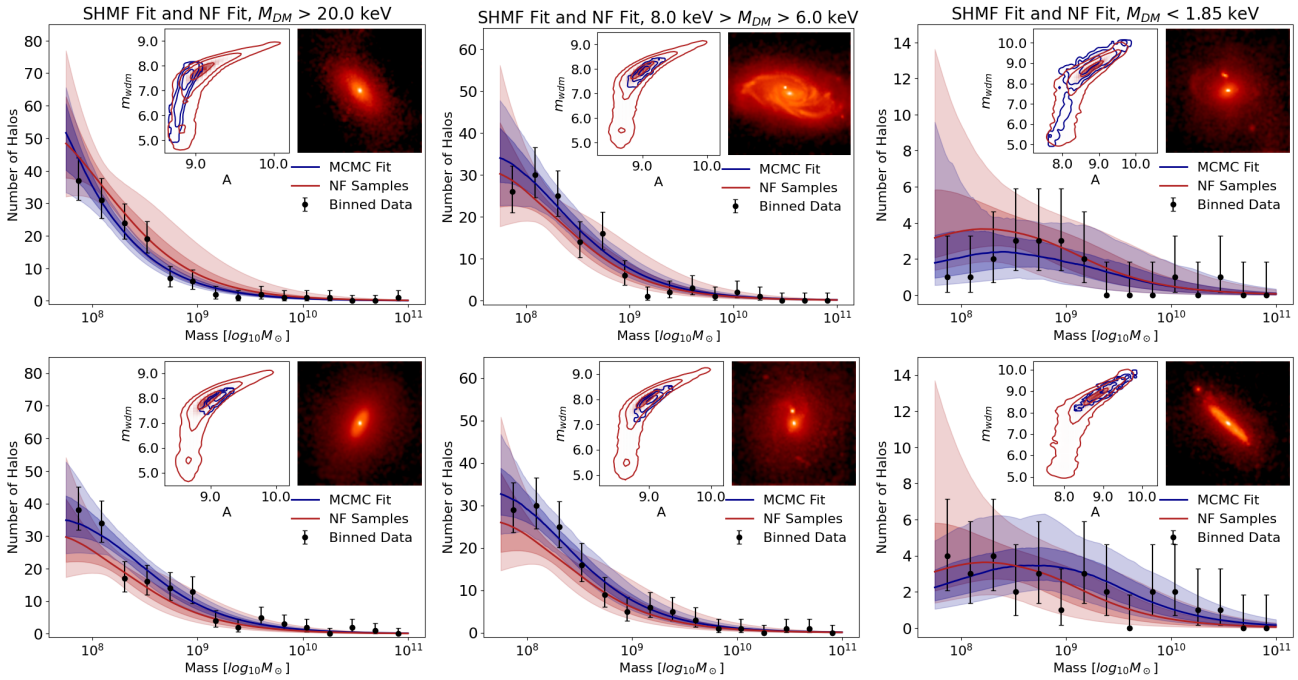

Figure 3. **Similar to Figure 2.** The shown galaxies are the same as in Figure 3, but the orientation of the galaxy images is different. The blue contours show the samples of the MCMC fit, and the red contours the samples of the NF conditioned on $M_{\rm DM}$ and the galaxy image. Overall, the sampled parameters of the NF represent the fits obtained directly through the likelihood and cover the full range of MCMC sampled parameters.