Deep learning the astrometric signature of dark matter substructure

Kyriakos Vattis, Michael W. Toomey, and Savvas M. Koushiappas

Phys. Rev. D 104, 123541 — Published 20 December 2021
DOI: 10.1103/PhysRevD.104.123541
Deep learning the astrometric signature of dark matter substructure

Kyriakos Vattis,† Michael W. Toomey,‡ and Savvas M. Koushiappas†

Department of Physics, Brown University, Providence, RI 02912-1843, USA and
Brown Theoretical Physics Center, Brown University, Providence, RI 02912-1843, USA

(Dated: October 21, 2021)

We study the application of machine learning techniques for the detection of the astrometric signature of dark matter substructure. In this proof of principle a population of dark matter subhalos in the Milky Way will act as lenses for sources of extragalactic origin such as quasars. We train ResNet-18, a state-of-the-art convolutional neural network to classify angular velocity maps of a population of quasars into lensed and no lensed classes. We show that an SKA -like survey with extended operational baseline can be used to probe the substructure content of the Milky Way, and demonstrate how axiomatic attribution can be used to localize substructures in lensing maps.

PACS numbers:

I. INTRODUCTION

The Standard Cosmological Model, ΛCDM consisting of 70% dark energy in the form of a cosmological constant, 25% of Cold Dark Matter and 5% of baryonic matter has been quite successfully tested in recent years from galactic scales to cosmological scales [1–5]. However, one of its main components, dark matter, has only been observed through its gravitational effects despite all the efforts through direct and indirect detection [6–19] and colliders like the Large Hadron Collider [20, 21] to observe Weakly Interacting Massive Particles (WIMP) or other popular candidates like axions [22–25]. While the efforts continue, the null results highlight the necessity of finding alternative gravitational signatures that would shed light on the elusive nature of dark matter.

A very promising way to probe various dark matter models gravitationally is to study their effects on the distribution of dark matter on subgalactic scales, an idea that has been studied thoroughly in the literature. Several methods have been proposed including the utilization of tidal streams [26–31], astrometric observations [32–35], pulsar timing array observations [36], stellar wakes [37] and images of strongly-lensed background galaxies or quasars [38–52] looking for dark subhalos. For most of the aforementioned methods the expected signal is small and difficult to observe.

With the increase of computational power and the improvement of machine learning algorithms there have been multiple attempts to involve neural network–based approaches for such analyses. In the last few years machine learning has been used in an array of problems in cosmology [53] as well as in other problems in the physical sciences [54]. Examples include applications in large-scale structure [55], the Cosmic Microwave Background [56–58], the cosmological 21 cm signal [59] and lensing studies [60–68]. In the search for dark matter, supervised machine learning algorithms have been trained to identify substructure properties with simulated galaxy-galaxy strong lensing images. More recently, convolutional neural networks (hereafter, CNN) have been used for classification of different subhalo mass cut-offs [69], classifying DM halos with and without substructure [70, 71], to distinguish dark matter models with disparate substructure morphology [70], and also inference of population level properties for DM substructure [72].

In this work we investigate the possibility of applying machine learning techniques and in particular a CNN, to detect the presence of lensing effects from dark matter substructure on a population of quasars in astrometric data from future surveys based on simulated data as a proof of principle. We also demonstrate how axiomatic attribution can be used to localize substructures in lensing maps. In Sec. II we outline the formalism followed in order to simulate the lensing signal followed by a description of creating simulated data maps. We discuss the implementation of the CNN in Sec. III, our results in Sec. IV and our conclusions are shown in Sec. V.

II. ASTROMETRIC LENSING

A. Formalism

We begin by considering the lensing effect of a dark matter halo on the astrometric parameters of a background source (position, velocity and acceleration), similar in spirit to [33, 34]. We utilize a modified version of the code provided in [34] to extract the astrometric signal in a suite of models that are used to create the simulated data set used in deep learning model training.

The system we consider is shown in Fig. (1); a source i moving with velocity \( \mathbf{v}_i \) is lensed by a lens \( l \) with velocity \( \mathbf{v}_l \) as seen by an observer with velocity \( \mathbf{v}_o \). The distance from observer to a lens is \( D_l \) and from observer to a source is \( D_i \). Both are assumed to be constant and thus all vectors involved in the calculation are two dimensional laying on the celestial sphere.
For such a system the deflection angle $\Delta \theta_{il}$ between the real position and the position of the image of the observer. The time derivative of the magnitude of the impact factor is given by

$$|b_{il}'| = \hat{b}_{il} \cdot \mathbf{v}_{il}.$$  

(5)

Similarly we take the time derivative of Eq. (3) to calculate the lensing induced acceleration as

$$a = \Delta \theta_{il}' = - \left( 1 - \frac{D_i}{D_l} \right) \frac{4G_N M(b_{il})}{c^2 b_{il}} \hat{b}_{il}$$

$$+ \frac{M'(b_{il})}{b_{il}^2} \left[ 2|b_{il}| \mathbf{v}_{il} + |b_{il}| \mathbf{v}_{il}' - 4|b_{il}^3| \hat{b}_{il} \right]$$

$$- \frac{2M(b_{il})}{b_{il}^2} \left[ 2|b_{il}| \mathbf{v}_{il} + |b_{il}| \mathbf{v}_{il}' + 3|b_{il}|^2 \hat{b}_{il} \right],$$  

(6)

with $M''$ the second derivative of the enclosed mass with respect to the impact factor and

$$|b_{il}''| = \frac{\mathbf{v}_{il} \cdot (\mathbf{b}_{il} \mathbf{v}_{il} - |b_{il}| \mathbf{b}_{il})}{b_{il}^2}.$$  

(7)

In order to compute the deflection angle, velocity and acceleration (equations 1, 3 and 6) we need to specify the distribution of matter in the lens, i.e., lens density profile $\rho(r)$. We assume that the density profile of the halos is described by the Navarro–Frenk–White (NFW) profile $[73]$

$$\rho(r) = \frac{\rho_s}{\left( \frac{r}{r_s} \right) \left( 1 + \frac{r}{r_s} \right)^2}.$$  

(8)

where $r_s$ is the characteristic scale and $\rho_s$ the characteristic density of the profile.

In order to demonstrate the aforementioned effects, Figure 2 shows an example of induced deflections, velocities and accelerations in arbitrary units (see also [33]). The lens is located at the origin moving towards the $+\hat{y}$ direction in all three plots. The left panel depicts the angular deflection. The maximum angular deflection occurs for impact parameters approximately equal to the characteristic scale, as for $b_{il} < r_s$ the enclosed mass $M(< b_{il})$ decreases rapidly. On the other hand, if $b_{il} \gg r_s$ the enclosed mass approaches a constant and therefore the deflection decreases as $\sim 1/b_{il}$. The middle panel in Figure 2 shows the corresponding induced velocity; the largest amplitude in the velocity vector occurs near the centre and in a direction opposite to the direction of motion of the lens with a dipole-like behavior in the far-field limit. Finally, the right panel of Figure 2 shows the acceleration pattern which exhibits a quadrupole-like behavior in the far-field.  

1 All these characteristics are in agreement to what has been demonstrated in of [33].
B. Population of lenses

We now proceed to simulate the lensing effect on a population of sources by a population of lenses. Following [34], we choose to simulate the lens population as NFW subhalos in the Milky Way that follow a mass function of the form

\[
\frac{dN}{dM} = A_0 M^{-1.9}
\]  

(9)

normalized so that there are 150 subhalos in the mass range \(10^8 - 10^{10} M_\odot\) [74] consistent with results from recent hydrodynamical simulations [75, 76].

We then define the spatial distribution of the subhalos, taking into account the tidal disruption due to the gradient of the Galactic potential towards the Galactic centre that depletes the fraction of mass bound in substructures in this region, by assuming an Einasto profile fitted to the results of the Aquarius simulation [77]

\[
\rho_g(r) = \exp \left\{ -\frac{2}{\gamma} \left[ \left( \frac{r}{r_\epsilon} \right)^{\gamma} - 1 \right] \right\}
\]  

(10)

with \(r_\epsilon = 199\text{kpc}\) and \(\gamma_\epsilon = 0.678\). We use this distribution to sample the radii at which the subhalos lay.

The next ingredient for defining a lens population is the dark matter velocity distribution. We assume that in the Galactic frame and far away from the Sun’s gravitational potential the dark matter velocity distribution is described by the Standard Halo Model (SHM) as

\[
f_\infty(v) = N \left( \frac{1}{\pi v_0^3} \right)^{3/2} e^{-v^2/v_0^2}
\]  

(11)

for \(|v| < v_{\text{esc}}\) and 0 otherwise (however see also [78]).

\(N\) is a normalization factor to account for the truncation at \(v_{\text{esc}}, v_0 = 220\text{km/s}\) [79] and the escape velocity is \(v_{\text{esc}} = 550\text{km/s}\) [80]. We transform this velocity distribution at the position of Earth (as we are interested in measurements at the Earth’s frame) by applying a Galilean transformation from the Galactic frame to the Earth frame

\[
f_\odot(v) \approx f_\infty(v + v_\odot)
\]  

(12)

where \(v_\odot = \{11, 232, 7\} \text{km/s}\) is the velocity of the Sun in \(\{U_\odot, V_\odot, W_\odot\}\) coordinates [81]. We ignore the annual modulation due to the motion of the Earth around the Sun as it is a factor of ten smaller compared to the velocity components.

We normalize the density profile of each subhalo in the following way. The NFW profile is defined by two parameters, \(\rho_s\) and \(r_s\), or alternatively, the mass of the halo, \(M_{200}\) and its concentration \(c_{200}\). The concentration is a measure of the compactness of the halos, defined as

\[
c_{200} = R_{200}/r_s
\]

where \(R_{200}\) is the radius that encloses \(M_{200}\) whose density is 200 times the critical density of the universe \(\rho_c = 3H_0^2/8\pi G_N\). With this definition, the characteristic density \(\rho_s\) and scale radius \(r_s\) are obtained from

\[
M_{200} = \int_0^{R_{200}} dr' 4\pi r'^2 \rho(r') = 4\pi \rho_s r_s^3 f(c_{200})
\]  

(13)

where

\[
f(x) = \ln(1 + x) - \frac{x}{1 + x}
\]
Figure 3: The induced astrometric lensing effect on a population of quasar sources by a population of NFW subhalos distributed in the Milky Way, under the assumption of pure signal and no instrumental noise. We truncate the effect within 20 degrees of the centre of each subhalo. The calculation was done on a HEALPIX map with \textit{nside} = 128 and we simulate \( \sim 1200 \) subhalos with masses between \( 10^7 \) and \( 10^{10} M_\odot \). The first row shows the angular displacement \( \Delta \theta_{il} \), the second row the induced velocity \( u \) and finally the third row the induced acceleration \( a \).

We generate a population of sources with \( D_i = \infty \) and \( v_i = 0 \) at the center of each pixel of a HEALPIX map with \textit{nside} = 128, and we simulate \( \sim 1200 \) subhalos between \( 10^7 \) and \( 10^{10} M_\odot \). This approach is similar to future observations of quasars with the Square Kilometer Array (SKA), or other instruments with similar capabilities. In the absence of instrumental noise, the resulting signal within 20 degrees from the center of each halo is shown in shown in Fig. (3).

In practice the sensitivity and precision of the instrument plays a significant role. For this proof of principle example, we assume an SKA-like sensitivity\cite{83}. Given the current forecasts \cite{83} after 10 years of operation SKA will be able to observe up to \( N_q \sim 10^8 \) quasars with a peculiar velocity uncertainty of \( \sigma_u \approx 1 \, \mu\text{as/yr} \) which we label scenario A. Additionally we consider four more sets of parameters (scenarios B, C, D and E) as shown in Ta-
As we will see the weakness of the signal makes the significance of detection highly dependent on these two parameters $N_q$ and $\sigma_u$.

### Table I: Scenarios of astrometric parameters $N_q$ and $\sigma_u$ used in this work.

| Scenario | $N_q$  | $\sigma_u$ [\(\mu\text{as/yr}\)] |
|----------|--------|----------------------------------|
| A        | $10^8$ | 1                                |
| B        | $10^8$ | 0.1                              |
| C        | $10^9$ | 1                                |
| D        | $10^9$ | 0.1                              |
| E        | $3 \times 10^9$ | 1  |

III. THE ASTROMETRIC LENSING SIGNAL

A. Characteristics of the signal in simulated data sets

In Fig. (3) we show the expected astrometric signal for the ideal case with no instrumental noise. We observe qualitative characteristics similar to Fig. (2), such as the gradient of the magnitude of the displacement field as a function of distance is small compared to the gradient in the velocity and acceleration fields (due to the $1/b_{il}$ factors – see footnote 1). In addition, the colorbar depicts the magnitude of the expected effect. For the displacement field the median amplitude is 1 mas, while for the velocity and acceleration fields the median values are $10^{-2}$ nas yr$^{-1}$, $5 \times 10^{-10}$ nas yr$^{-2}$ respectively.

For the purposes of this work, the velocity vector $\mathbf{u}$ is an adequate probe of substructure lensing as quasars are stationary in the galactic frame. However note that the acceleration can also be used for some halo mass and concentration functions as was shown in [33].

In order to apply machine learning methods in our efforts to extract dark matter substructure from the astrometric data we first need to create a library of fake data sets which we will use in order to test the method outlined above. While Fig. (3) shows the pure signal effect we must take into account instrumental noise. To create simulated data sets we distribute the $N_q$ quasars on a HEALPIX grid with $nside = 32$ corresponding to 12,288 equal-area pixels. We assume quasars follow a Poisson distribution among pixels with a mean value given by the total number of quasars divided by the total number of pixels.

At each pixel we sample a longitudinal and latitudinal velocity component from a Normal distribution centered at the expected value in the presence of lensing and with a standard deviation $\sigma_u$. With all quasars being assigned a measured velocity under the assumption that all quasars in a pixel will experience the same lensing effect, we average all quasars and assign a single velocity to each pixel. For the case of only instrumental noise, i.e., no signal, we follow the same procedure with the normal distribution always centered at 0. We repeat the process for all five scenarios of Tab. I.

Fig. (4) shows an example of the simulated data after taking into account the sensitivity of the instrument for parameter set C. The upper two projections contain the lensing signal while the bottom two projections show the no-signal, only instrumental noise case.
Next, we will use this method to generate big data sets for machine learning training and validation data sets. The hope is that by training a model to distinguish between a Gaussian noise from the non-Gaussian signal maps it will be possible to extract information about the underlying population of lenses in future surveys.

B. Machine learning implementation

The structure of the data leads naturally to an implementation of a CNN. One can appreciate this from Fig. (4) where it is clear that correlations among pixels will likely serve a key role in helping to distinguish between the signal and no signal case. Given this intuition, we trained state-of-the-art architecture ResNet-18 [84] as a classifier between these two classes. We selected ResNet-18 to train across all our sets after testing other algorithms (VGG [85] and AlexNet [86]), as it consistently achieved noticeably better performance - see Fig. (5).

ResNet-18 differs from a normal CNN in that it has residual blocks which are characterized by their use of skip connections, helping to alleviate the problem of vanishing gradients. One problem, however, is that our simulations are based on data in HEALPIX format. While there are architectures that are designed to be trained on this data structure, notably graph based CNN DeepSphere [87], we instead opt to project our maps to Cartesian coordinates allowing us to train on a greater diversity of established architectures like ResNet.

For each scenario of Table 1, our training set consists of 25,000 training and 2,500 validation images per class (signal and no-signal cases). With a large number of simulated data, cross-validation is unnecessary as is the case for a lot of deep-learning applications. The cross-entropy loss was minimized over at most 50 epochs in batches of 32 with the Adam optimizer where the learning rate was initialized with a value of $1 \times 10^{-1}$ and decayed by a factor of 10 if the validation loss was not improved after 5 epochs.

We implement ResNet-18 with the PyTorch package and run on four NVIDIA Tesla V100 GPUs at the Brown Center for Computation and Visualization. As a metric for classifier performance we utilize the well-established Area Under the receiver operating characteristic (ROC) Curve (AUC).
the desired input,

\[
IG_1(x, x') := (x_i - x'_i) \int_{0}^{1} \frac{\partial F(x' + \alpha(x - x'))}{\partial x_i} d\alpha.
\]  

Here \( F \) corresponds to a trained model (perhaps a classifier between two types of dark matter) and \( x \) an input (perhaps an astrometric lensing image velocity map). For a 2-dimensional image input, one can construct an assignment map where each pixel is assigned an attribution score. A positive (negative) attribution score contributes favorably (negatively) to the final network prediction whereas pixels with no attribution do not contribute. Inputs with the largest magnitude attribution score are the most influential in the networks final decision.

IV. RESULTS

In Fig. (6) we present the ROC curves and their corresponding AUC values for each set of parameters of Tab. 1. This is the main result of this work. The red lines correspond to \( \sigma_u = 1 \mu\text{as/yr} \) while the blue lines are for \( \sigma_u = 0.1 \mu\text{as/yr} \) and the total number of quasars \( N_q \) is designated by the line’s width. As expected larger \( N_q \) and smaller \( \sigma_u \) lead to higher AUC values and higher detection significance since the noise level of each pixel scales as \( \sim \sigma_u / \sqrt{n_q} \) where \( n_q \) is the number of quasars per pixel. Additionally we can see that a decrease in the uncertainty \( \sigma_u \) is more impactful than increasing the number of quasars \( N_q \) and thus a more precise survey will be better suited to detect the signal than a more sensitive one.

Our model trained on an SKA-like survey, realized here as parameter scenario A, is not sensitive to the signal induced by substructure after achieving an AUC of \( \sim 0.5 \). In this configuration the data is likely too noise-dominated even for a CNN. However, a more optimistic survey with ten times more quasars does manage to achieve limited power at discerning between signal and noise as evidenced by a marginal AUC score of 0.723 for scenario C, but this is still not very well equipped to detect dark matter substructure based on a 10 year operational baseline.

In order to achieve a more robust performance, corresponding to AUC values \( \gtrsim 0.8 \), we need either the number of quasars to increase even more by a factor of at least 3 compared to scenario C, and/or \( \sigma_u \) to decrease by \( \sim \) a factor of 10, or some combination of the above. Encouragingly, it is very promising that the detection significance improves drastically with parameter scenario D where \( N_q = 10^8 \) and \( \sigma_u = 0.1 \mu\text{as/yr} \), allowing for almost perfect classification using the ResNet-18 architecture.

In other words, if SKA runs for longer than a decade, it is possible that the proposed technique can be successfully applied in the future to detect Milky Way substructure. For example, doubling its operational time reduces the uncertainty to \( \sigma_u = 0.5 \mu\text{as/yr} \), and with the currently planned \( N_q = 10^8 \) quasars it is possible to obtain an AUC of \( \gtrsim 0.8 \), and thus reliably probe the substructure content of the Milky Way.

We observe that for the current best velocity sensitivity estimates for a 10 year run of an SKA-like survey \( \sigma_u \sim 1 \mu\text{as/yr} \), a total number of \( N_q \sim 3 \times 10^9 \) quasars would be required for the classifier to detect the signal with an AUC value \( > 0.8 \). Much better performance is achieved for \( N_q \sim 10^8 \) but much smaller velocity uncertainty \( \sigma_u \sim 0.1 \mu\text{as/yr} \) with AUC value of \( \sim 0.96 \). We see that a smaller uncertainty has much higher impact on the ability of the classifier to detect the dark matter substructure and it can be achieved by expanding the duration of the survey to more than a decade. Our results are in qualitative agreement with the results of [34] though we cannot quantitatively compare since the AUC...
values used here don’t directly map to the discovery significance shown in Fig. 6 of that work.

We have further shown a method for localizing subhalos using integrated gradients, a method of axiomatic attribution. Concretely, we find that the map for positive attributions, pixels which voted for the substructure class, are found to be consistent with the location of known strong lenses in our simulations. While we don’t investigate in detail here, the negative attributions are likely to encode crucial information about various noise and backgrounds present in the data set.

While it may be possible that there is an architecture that could be better suited to tackle the complexity of the data set at the level of expectations for SKA, it should be pointed out that the results presented here are based upon an assumed dark matter model defined by the mass and distribution of its substructure. It may be the case that dark matter is something other than the CDM picture presented here and, thus, prospects for detection may differ. Similar to the spirit of [90] where anomaly detection, a form of unsupervised machine learning, was used to detect the presence of dark matter substructure in strong lensing images, it seems natural to extend this work to the unsupervised scenario where an autoencoder is trained on sets of images corresponding to SKA measurements with no substructure and qualifying the constraining power of identifying any DM substructure in SKA data. In this way one is taking a theory agnostic approach to the detection of dark matter – i.e. our classifier performance will not depend on an arbitrary choice of model and its various parameters. Additionally the inclusion of acceleration data on top of the velocity maps can further strengthen the detection of the signal. We leave this for future work.

Acknowledgments

We acknowledge useful conversations with Stephon Alexander, Jatan Buch, and John Leung. This research was conducted using computational resources and services at the Center for Computation and Visualization, at Brown University. SMK is partially supported by NSF PHY-2014052. We gratefully acknowledge the support of Brown University.

[1] Planck Collaboration, N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, R. B. Barreiro, N. Bartolo, et al., ArXiv e-prints (2018), 1807.06209.
[2] T. M. C. Abbott, F. B. Abdalla, A. Alarcon, S. Alam, F. Andrade-Oliveira, J. Annis, S. Avila, M. Banerji, N. Banik, K. Bechtol, et al., MNRAS 477, 1604 (2018), 1801.02689.
[3] H. Gil-Marín, J. Guy, P. Zarrour, E. Burtin, C.-H. Chuang, W. J. Percival, A. J. Ross, R. Ruggeri, R. Tojerio, G.-B. Zhao, et al., MNRAS 477, 1604 (2018), 1801.02689.
[4] E. Corbelli and P. Salucci, MNRAS 311, 441 (2000), astro-ph/9909252.
[5] S. W. Randall, M. Markevitch, D. Clowe, A. H. Gonzalez, and M. Bradac, ApJ 679, 1173 (2008), 0704.0261.
[6] A. K. Drukier, K. Freese, and D. N. Spergel, Phys. Rev. D33, 3495 (1986).
063514 (2019), 1906.03156.

[64] Y. D. Hezaveh, L. Perreault Levasseur, and P. J. Marshall, Nature 548, 555 (2017), 1708.08842.

[65] L. Perreault Levasseur, Y. D. Hezaveh, and R. H. Wechsler, Astrophys. J. 850, L7 (2017), 1708.08843.

[66] W. R. Morningstar, Y. D. Hezaveh, L. Perreault Levasseur, R. D. Blandford, P. J. Marshall, P. Putzky, and R. H. Wechsler (2018), 1808.00011.

[67] W. R. Morningstar, L. Perreault Levasseur, Y. D. Hezaveh, R. Blandford, P. Marshall, P. Putzky, T. D. Rueter, R. Wechsler, and M. Welling (2019), 1901.01359.

[68] R. Canameras, S. Schuldt, S. H. Suyu, S. Taubenberger, T. Meinhardt, L. Leal-Taixe, C. Lemon, K. Rojas, and E. Savary (2020), 2004.13048.

[69] S. Varma, M. Fairbairn, and J. Figueroa (2020), 2005.05353.

[70] S. Alexander, S. Gleyzer, E. McDonough, M. W. Toomey, and E. Usai, Astrophys. J. 893, 15 (2020), 1909.07346.

[71] J. Brehmer, S. Mishra-Sharma, J. Hermans, G. Louppe, and K. Cranmer, Astrophys. J. 886, 49 (2019), ISSN 1538-4357, 1909.02005, URL http://dx.doi.org/10.3847/1538-4357/ab4c41.

[72] J. F. Navarro, C. S. Frenk, and S. D. M. White, ApJ 462, 563 (1996), astro-ph/9508025.

[73] M. Hütten, C. Combet, G. Maier, and D. Maurin, J. Cosmology Astropart. Phys. 2016, 047 (2016), 1606.04898.

[74] P. Mollitor, E. Nezri, and R. Teyssier, MNRAS 447, 1353 (2015), 1405.4318.

[75] J. F. Navarro, C. S. Frenk, A. Fattahi, J. F. Navarro, R. G. Bower, R. A. Crain, C. Dalla Vecchia, M. Furlong, J. C. Helly, A. Jenkins, et al., MNRAS 457, 1931 (2016), 1511.01098.

[76] M. Roos, Journal of Modern Physics 3, 1152 (2012), 1208.3662.

[77] K. Vattis and S. M. Koushiappas, Phys. Rev. D 97, 103003 (2018), 1801.06556.

[78] J. F. Kerr and D. Lynden-Bell, MNRAS 221, 1023 (1986).

[79] M. Schaller, C. Combet, G. Maier, and D. Maurin, J. Cosmology Astropart. Phys. 2016, 047 (2016), 1606.04898.

[80] P. Mollitor, E. Nezri, and R. Teyssier, MNRAS 447, 1353 (2015), 1405.4318.

[81] T. Sawala, C. S. Frenk, A. Fattahi, J. F. Navarro, R. G. Bower, R. A. Crain, C. Dalla Vecchia, M. Furlong, J. C. Helly, A. Jenkins, et al., MNRAS 457, 1931 (2016), 1511.01098.

[82] R. Schönrich, J. Binney, and W. Dehnen, MNRAS 403, 1829 (2010), 0912.3693.

[83] M. A. Sánchez-Conde and F. Prada, MNRAS 442, 2271 (2014), 1312.1729.

[84] E. Fomalont and M. Reid, New Astronomy Reviews 48, 1473 (2004).

[85] K. He, X. Zhang, S. Ren, and J. Sun, arXiv e-prints arXiv:1512.03385 (2015), 1512.03385.

[86] K. Simonyan and A. Zisserman, arXiv e-prints arXiv:1409.1556 (2014), 1409.1556.

[87] A. Krizhevsky, I. Sutskever, and G. E. Hinton, Commun. ACM 60, 84–90 (2017), ISSN 0001-0782, URL https://doi.org/10.1145/3065386.

[88] N. Perraudin, M. Defferrard, T. Kacprzak, and R. Sgier, Astron. Comput. 27, 130 (2019), 1810.12186.

[89] M. Sundararajan, A. Taly, and Q. Yan, arXiv e-prints arXiv:1703.01365 (2017), 1703.01365.

[90] B. Ostdieck, A. Diaz Rivero, and C. Dvorkin (2020), 2009.06663.

[91] S. Alexander, S. Gleyzer, H. Parul, P. Reddy, M. W. Toomey, E. Usai, and R. Von Klar (2020), 2008.12731.