Inferring dark matter substructure with global astrometry beyond the power spectrum

Anonymous Author(s)
Affiliation
Address
email

Abstract

Astrometric lensing has recently emerged as a promising avenue for characterizing the population of dark matter clumps—subhalos—in our Galaxy. Leveraging recent advances in simulation-based inference and neural network architectures, we introduce a novel method to look for global dark matter-induced lensing signatures in astrometric datasets. Our method shows significantly greater sensitivity to a cold dark matter population compared to existing approaches, establishing machine learning as a powerful tool for characterizing dark matter using astrometric data.

1 Introduction and background

Although there exists plenty of evidence for dark matter (DM) on galactic scales and above, the distribution of DM clumps—subhalos—on sub-galactic scales is less well-understood and remains an active area of cosmological study. This distribution additionally correlates with and may provide clues as to the underlying particle physics nature of dark matter, highlighting its relevance across multiple domains. While larger subhalos can be detected and characterized through their association with luminous tracers such as bound stellar populations, subhalos with smaller masses $\lesssim 10^9\,\mathrm{M}_\odot$ are not generally associated with luminous matter [1, 2], rendering their characterization challenging. Gravitational effects provide one of the few avenues to probe the distribution of these otherwise-invisible subhalos. Gravitational lensing—the bending of light from a background source due to a foreground mass—is one such effect and has been proposed in various incarnations as a probe of dark subhalos.

Astrometry refers to the precise measurement of the positions and motions of luminous objects like stars and other galaxies. Gravitational lensing of these background objects by a moving foreground mass, such as a dark matter subhalo, can imprint a characteristic pattern of motions on the measured kinematics (angular velocities and/or accelerations) of these objects. Refs. [3, 4] introduced several methods for extracting this signature with the aim of characterizing the subhalo population in our Galaxy, including those based on computing convolutions of the expected lensing signal and detecting local kinematic outliers. Ref. [5] further proposed using the angular power spectrum of the astrometric field as an observable to infer the global properties of a dark matter population.

Astrometric datasets are inherently high-dimensional, consisting of positions and kinematics of potentially millions of objects. Especially when the signal consists of the collective imprint of a large number of subhalos, characterizing their population properties involves marginalizing over all possible configurations of subhalos, rendering the likelihood intractable and necessitating a reduction to data summaries like the power spectrum. While shown to be effective, such simplification can result in significant loss of information when the signal is non-Gaussian in nature.

The dawn of the era of precision astrometry, with the *Gaia* satellite having delivered the most precise astrometric dataset to-date [6–8] and surveys including the Square Kilometer Array (SKA) [9, 10]

Submitted to 35th Conference on Neural Information Processing Systems (NeurIPS 2021). Do not distribute.

and Roman Space Telescope [11] set to achieve further leaps in sensitivity, calls for methods that can extract more information from these datasets than is possible using existing techniques. In this paper we introduce a method that leverages recent advances in simulation-based inference in order to characterized the subhalo population in our Galaxy using astrometric data.

2 Model and inference

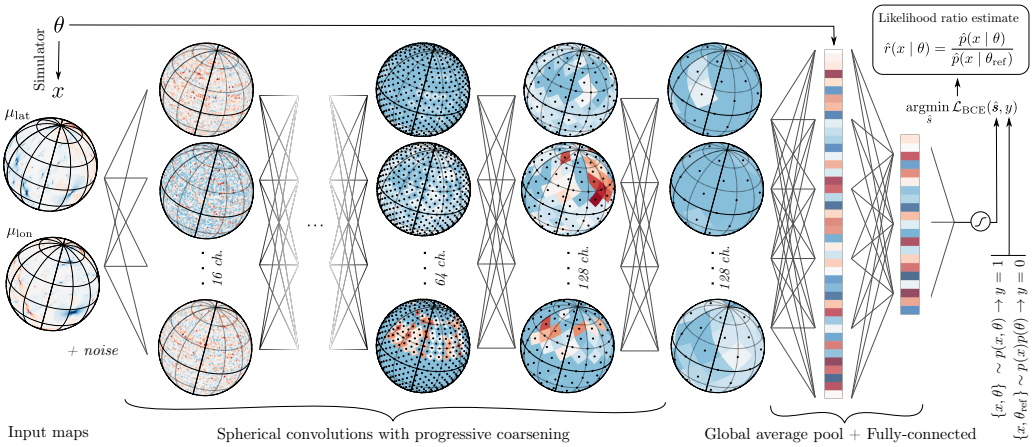

Figure 1: An illustration of the method and neural network architecture used in this work.

The forward model We consider a population of Navarro-Frenk-White (NFW) [12] subhalos following a power-law mass function, $\mathrm{d}n/\mathrm{d}m \propto m^{\alpha}$, with slope $\alpha = -1.9$ as expected if the population is sourced from nearly scale-invariant primordial fluctuations in the canonical Λ Cold Dark Matter (ΛCDM) scenario. The subhalo fraction $f_{\rm sub}$, quantifying the expected fraction of the mass the Milky Way mass contributed by subhalos in the range 10^{-6}–$10^{10}\,\mathrm{M}_{\odot}$, is taken to be the parameter of interest. The spatial distribution of subhalos is modeled using results from the Aquarius simulation, following Refs. [13, 14], which also motivates the fiducial test parameter point containing 150 subhalos in expectation between 10^{8}–$10^{10}\,\mathrm{M}_{\odot}$, corresponding to $f_{\rm sub} \simeq 0.2$.

Our datasets consist of the 2-dimensional angular velocity map on the celestial sphere of quasi-stellar objects (QSOs), also known as quasars which, owing to their large distances from us, are expected to have small intrinsic angular velocities. Given a subhalo lens with velocity $\mathbf{v}_l$, the expected induced velocity for a quasar at impact parameter $\mathbf{b}$ is given by [3, 5]

$$\boldsymbol{\mu}(\mathbf{b}) = 4G_{\rm N}\left\{\frac{M(b)}{b^2}\left[2\hat{\mathbf{b}}\left(\hat{\mathbf{b}}\cdot\mathbf{v}_l\right) - \mathbf{v}_l\right] - \frac{M'(b)}{b}\hat{\mathbf{b}}\left(\hat{\mathbf{b}}\cdot\mathbf{v}_l\right)\right\} \tag{1}$$

where $M(b)$ and $M'(b)$ are the projected mass and its gradient. An example of the induced velocity signal on part of the celestial sphere, projected along the latitudinal and longitudinal directions and exhibiting dipole-like structures, is shown in the leftmost column of Fig. 1.

In order to enable comparison with traditional approaches—which are generally not expected to be sensitive to a CDM subhalo population with next-generation astrometric surveys [3, 5]—we benchmark using an optimistic observational configuration corresponding to measuring the proper motions of 10^{8} quasars with noise $\sigma_{\mu} = 0.1\,\mu\mathrm{as\,yr}^{-1}$.

The power spectrum approach Ref. [5] introduced an approach for extracting the astrometric signal due to a dark matter subhalo population by decomposing the observed map into its angular (vector) power spectrum. The power spectrum is a summary statistic ubiquitous in astrophysics and cosmology and quantifies the amount of correlation contained at different spatial scales. In the case of data on a sphere, the basis of spherical harmonics is often used, and the power spectrum then encodes the correlation structure on different multipoles ℓ. The power spectrum effectively captures the linear component of the signal and, when the underlying signal is a Gaussian random field, captures *all* of the relevant information contained in the map(s) [15].

The expected signal and corresponding sensitivity in the power spectrum domain can be computed analytically using the formalism described in Ref. [5], which we use here as a comparison point. While effective, reduction of the full astrometric map to its power spectrum results in loss of information; this can be seen from the fact that the signal on the leftmost column of Fig. 1 is far from Gaussian. Furthermore, the existence of systematic correlations on large angular scales due to *e.g.*, biases in calibration of celestial reference frames [16] introduces degeneracies with a putative signal and precludes their usage in the present context. This is especially true in the case of *relative* astrometric observations, where systematic variations among observed patches of the sky are present. For these reason multipoles $\ell < 10$ were discarded in Ref. [5], degrading the projected sensitivity.

Simulation-based inference with parameterized classifiers Recent advances in machine learning have enabled methods that aim to directly extract information from models defined through high-dimensional simulations; see Ref. [17] for a recent review. Here, we make use of parameterized classifiers [18–23] (previously used in Refs. [24, 25] in the context of inferring DM substructure) in order to directly approximate the likelihood ratios associated with all-sky astrometric maps containing signatures of dark matter. Given a classifier that can distinguish between samples $\{x\} \sim p(x \mid \theta)$ and those from a fixed reference hypothesis $\{x\} \sim p(x \mid \theta_{\rm ref})$, the decision function output by the optimal classifier $s(x, \theta) = p(x \mid \theta)/(p(x \mid \theta) + p(x \mid \theta_{\rm ref}))$ is one-to-one with the likelihood ratio, $r(x \mid \theta) \equiv p(x \mid \theta)/p(x \mid \theta_{\rm ref}) = s(x, \theta)/(1 - s(x, \theta))$, a fact appreciated as the likelihood-ratio trick [18, 26]. The classifier $s(x, \theta)$ in this case is a neural network that can work directly on the high-dimensional data x, and is parameterized by θ by having it included as a feature. In order to improve numerical stability and reduce dependence on the fixed reference hypothesis $\theta_{\rm ref}$, we follow Ref. [23] and train a classifier to distinguish between data-sample pairs $\{x, \theta\} \sim p(x, \theta)$ and those from a product of marginal distributions $\{x, \theta\} \sim p(x)p(\theta)$ (obtained by shuffling samples within a batch) using the binary cross-entropy loss as the optimization objective.

Extracting information from high-dimensional datasets We bin the the velocity maps on a `HEALPix` grid with resolution parameter `nside=64`, the value in each pixel then quantifying the average velocity of quasars within that pixel. All inputs are normalized to zero mean and unit standard deviation across the training sample. We use `DeepSphere` [27, 28], a graph-based convolutional neural network tailored to data sampled on a sphere and able to leverage the hierarchical structure of the `HEALPix` representation. In particular, `DeepSphere` efficiently performs convolutions in the spectral domain, using a basis of Chebychev polynomials as convolutional kernels [29]. Starting with 2 input channels representing the two orthogonal components of the velocity vector, we perform a graph convolution operation, increasing the channel dimension to 16, followed by a batch normalization, ReLU nonlinearity, and coarsening of the representation to resolution `nside=32` with max pooling (downsampling the representation by a factor of 4). Four more such layers are employed, increasing the channel dimension by a factor of 2 at each step until a maximum of 128, with the final map having `nside=2` corresponding to 48 pixels. At this stage, we average over the spatial dimension in order to encourage rotational invariance [30], outputting 128 features onto which the parameter of interest $f_{\rm sub}$ is appended. This is passed through a fully-connected network with (1024, 128) hidden units outputting the classifier decision by applying a sigmoidal projection.

10^5 samples from the forward model were produced. The estimator is trained for up to 50 epochs with early stopping, with a batch size of 64. The `AdamW` optimizer [31, 32] is used with initial learning rate 10^{-3} and weight decay 10^{-5}, with learning rate decayed through cosine annealing. Figure 1 summarizes the neural network architecture and method used in this work.

3 Experiments on simulated data

The left panel of Fig. 2 shows the expected test statistic (TS), defined as $\mathrm{TS} \equiv -2\ln(\hat{r}/\hat{r}_{\rm max})$, as a function of substructure fraction $f_{\rm sub}$ evaluated as an expectation over 50 test maps of the fiducial signal with $f_{\rm sub} \simeq 0.2$. Corresponding curves using the power spectrum approach are shown in blue, using minimum multipole thresholds of $\ell \geq 5$ (dashed) and $\ell \geq 10$ (solid). Thresholds corresponding to 1- and 2-σ significance are shown as the horizontal grey lines. We see that sensitivity gains of over a factor of ~ 2 can be expected for this benchmark when using the machine learning approach presented here when compared to the traditional power spectrum approach. No significant systematic bias on the central value of the inferred DM abundance is observed.

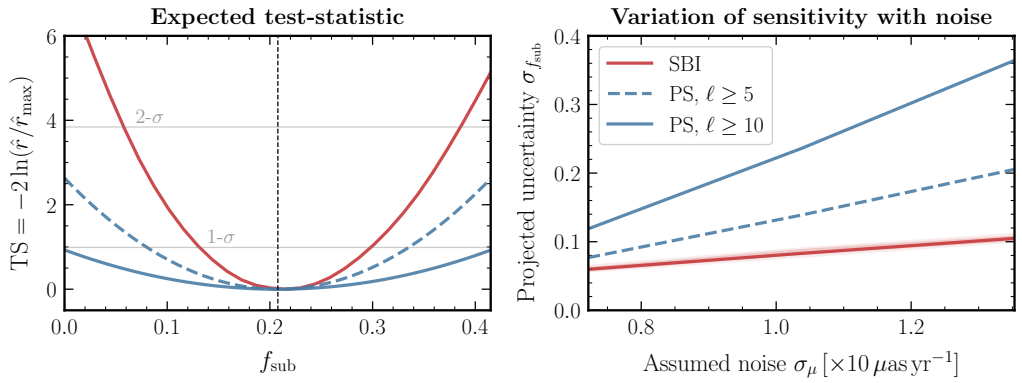

Figure 2: *(Left)* The expected test statistic using the machine learning-based method introduced in this work (red line) compared with existing approaches using power spectrum summaries with different multipole thresholds (blue lines). *(Right)* Scaling of the expected sensitivities with instrumental noise.

The right panel of Fig. 2 shows the scaling of expected sensitivity on substructure fraction $f_{\rm sub}$ with assumed noise per quasar, keeping the number of quasars fixed (red band, showing 1-σ variation over test datasets) compared to the power spectrum approach (blue lines). A far more favorable scaling of the machine learning approach is seen compared to the power spectrum approach, suggesting that it is disproportionately advantageous in low signal-to-noise regimes that are generally most relevant for dark matter searches.

4 Conclusions and outlook

We have introduced a method to analyze astrometric datasets over large regions of the sky using techniques based on machine learning with the aim of inferring the lensing signature of a dark matter subhalo population. We have shown our method to be significantly more sensitive a CDM subhalo population compared to established methods based on global summary statistics, with more favorable scaling with instrumental noise. Since data collection and reduction is an expensive endeavor, the use of methods that can take advantage of more of the available information can be equated to potentially years of observations, underscoring their importance. Additionally, unlike the power spectrum approach, the current method does not require the construction of a numerically-expensive estimator to account for non-uniform exposure and instrumental noise—these, as well as any other observational effects can be incorporated directly into the forward model. The use of architectures that explicitly take the vector nature of the input into account [33] can further improve the fidelity of the analysis.

We have focused in this work on assessing sensitivity to a CDM-like subhalo population with quasar velocity astrometry, which is within the scope of upcoming radio telescopes like the SKA. Our method can also be applied in a straightforward manner to look for the *acceleration* lensing signal imprinted on Milky Way stars, in particular sourced by a population of more compact subhalos than those expected in CDM. Both of these features are expected to introduce a larger degree of non-Gaussianity than in the signal explored here (as can be seen, *e.g.*, in Fig. 1 of Ref. [5]). Such an analysis is within purview of the upcoming Roman exoplanet microlensing survey [34] as well as future *Gaia* data releases, and a machine learning approach could provide significant sensitivity gains over existing methods.

Astrometric lensing has been established as a promising way to characterize the Galactic dark matter population, with theoretical progress in recent years going in step with advances on the observational front. While this work is a first attempt at bringing principled machine learning techniques to this field, with the availability of increasingly complex datasets we expect machine learning to be an important general-purpose tool for future astrometric dark matter searches.

Code used for reproducing the results presented in this paper is available at `https://github.com/xxx/yyy`.

References

[1] A. Fitts *et al.*, *"FIRE in the Field: Simulating the Threshold of Galaxy Formation,"* Mon. Not. Roy. Astron. Soc. **471**, 3547 (2017), arXiv:1611.02281 [astro-ph.GA].

[2] J. I. Read, G. Iorio, O. Agertz, and F. Fraternali, *"The stellar mass-halo mass relation of isolated field dwarfs: a critical test of ΛCDM at the edge of galaxy formation,"* Mon. Not. Roy. Astron. Soc. **467**, 2019 (2017), arXiv:1607.03127 [astro-ph.GA].

[3] K. Van Tilburg, A.-M. Taki, and N. Weiner, *"Halometry from Astrometry,"* JCAP **07**, 041 (2018), arXiv:1804.01991 [astro-ph.CO].

[4] C. Mondino, A.-M. Taki, K. Van Tilburg, and N. Weiner, *"First Results on Dark Matter Substructure from Astrometric Weak Lensing,"* Phys. Rev. Lett. **125**, 111101 (2020), arXiv:2002.01938 [astro-ph.CO].

[5] S. Mishra-Sharma, K. Van Tilburg, and N. Weiner, *"Power of halometry,"* Phys. Rev. D **102**, 023026 (2020), arXiv:2003.02264 [astro-ph.CO].

[6] Gaia Collaboration, *"The Gaia mission,"* Astron. Astrophys. **595**, A1 (2016), arXiv:1609.04153 [astro-ph.IM].

[7] Gaia Collaboration, *"Gaia Data Release 2. Summary of the contents and survey properties,"* Astron. Astrophys. **616**, A1 (2018), arXiv:1804.09365 [astro-ph.GA].

[8] L. Lindegren *et al.*, *"Gaia Data Release 2. The astrometric solution,"* Astron. Astrophys. **616**, A2 (2018), arXiv:1804.09366 [astro-ph.IM].

[9] E. B. Fomalont and M. Reid, *"Microarcsecond astrometry using the SKA,"* New Astron. Rev. **48**, 1473 (2004), arXiv:astro-ph/0409611.

[10] M. J. Jarvis, D. Bacon, C. Blake, M. L. Brown, S. N. Lindsay, A. Raccanelli, M. Santos, and D. Schwarz, *"Cosmology with SKA Radio Continuum Surveys,"* (2015), arXiv:1501.03825 [astro-ph.CO].

[11] WFIRST Astrometry Working Group, *"Astrometry with the Wide-Field Infrared Space Telescope,"* Journal of Astronomical Telescopes, Instruments, and Systems **5**, 044005 (2019), arXiv:1712.05420 [astro-ph.IM].

[12] J. F. Navarro, C. S. Frenk, and S. D. M. White, *"The Structure of cold dark matter halos,"* Astrophys. J. **462**, 563 (1996), arXiv:astro-ph/9508025.

[13] M. Hütten, C. Combet, G. Maier, and D. Maurin, *"Dark matter substructure modelling and sensitivity of the Cherenkov Telescope Array to Galactic dark halos,"* JCAP **09**, 047 (2016), arXiv:1606.04898 [astro-ph.HE].

[14] V. Springel, J. Wang, M. Vogelsberger, A. Ludlow, A. Jenkins, A. Helmi, J. F. Navarro, C. S. Frenk, and S. D. M. White, *"The Aquarius Project: the subhalos of galactic halos,"* Mon. Not. Roy. Astron. Soc. **391**, 1685 (2008), arXiv:0809.0898 [astro-ph].

[15] M. Tegmark, *"How to measure CMB power spectra without losing information,"* Phys. Rev. D **55**, 5895 (1997), arXiv:astro-ph/9611174.

[16] Gaia Collaboration, *"Gaia Data Release 2. The celestial reference frame (Gaia-CRF2),"* Astron. Astrophys. **616**, A14 (2018), arXiv:1804.09377 [astro-ph.GA].

[17] K. Cranmer, J. Brehmer, and G. Louppe, *"The frontier of simulation-based inference,"* Proc. Nat. Acad. Sci. **117**, 30055 (2020), arXiv:1911.01429 [stat.ML].

[18] K. Cranmer, J. Pavez, and G. Louppe, *"Approximating Likelihood Ratios with Calibrated Discriminative Classifiers,"* (2015), arXiv:1506.02169 [stat.AP].

[19] P. Baldi, K. Cranmer, T. Faucett, P. Sadowski, and D. Whiteson, *"Parameterized neural networks for high-energy physics,"* Eur. Phys. J. C **76**, 235 (2016), arXiv:1601.07913 [hep-ex].

[20] J. Brehmer, K. Cranmer, G. Louppe, and J. Pavez, *"A Guide to Constraining Effective Field Theories with Machine Learning,"* Phys. Rev. D **98**, 052004 (2018), arXiv:1805.00020 [hep-ph].

[21] J. Brehmer, G. Louppe, J. Pavez, and K. Cranmer, *"Mining gold from implicit models to improve likelihood-free inference,"* Proc. Nat. Acad. Sci. **117**, 5242 (2020), arXiv:1805.12244 [stat.ML].

[22] J. Brehmer, K. Cranmer, G. Louppe, and J. Pavez, *"Constraining Effective Field Theories with Machine Learning,"* Phys. Rev. Lett. **121**, 111801 (2018), arXiv:1805.00013 [hep-ph].

[23] J. Hermans, V. Begy, and G. Louppe, *"Likelihood-free MCMC with Amortized Approximate Ratio Estimators,"* (2019), arXiv:1903.04057 [stat.ML].

[24] J. Brehmer, S. Mishra-Sharma, J. Hermans, G. Louppe, and K. Cranmer, *"Mining for Dark Matter Substructure: Inferring subhalo population properties from strong lenses with machine learning,"* Astrophys. J. **886**, 49 (2019), arXiv:1909.02005 [astro-ph.CO].

[25] J. Hermans, N. Banik, C. Weniger, G. Bertone, and G. Louppe, *"Towards constraining warm dark matter with stellar streams through neural simulation-based inference,"* (2020), arXiv:2011.14923 [astro-ph.GA].

[26] S. Mohamed and B. Lakshminarayanan, "Learning in implicit generative models," (2017), arXiv:1610.03483 [stat.ML].

[27] M. Defferrard, M. Milani, F. Gusset, and N. Perraudin, *"DeepSphere: a graph-based spherical CNN,"* arXiv e-prints , arXiv:2012.15000 (2020), arXiv:2012.15000 [cs.LG].

[28] N. Perraudin, M. Defferrard, T. Kacprzak, and R. Sgier, *"DeepSphere: Efficient spherical Convolutional Neural Network with HEALPix sampling for cosmological applications,"* Astron. Comput. **27**, 130 (2019), arXiv:1810.12186 [astro-ph.CO].

[29] M. Defferrard, X. Bresson, and P. Vandergheynst, *"Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering,"* arXiv e-prints , arXiv:1606.09375 (2016), arXiv:1606.09375 [cs.LG].

[30] M. Lin, Q. Chen, and S. Yan, "Network in network," (2014), arXiv:1312.4400 [cs.NE].

[31] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," (2017), arXiv:1412.6980 [cs.LG].

[32] I. Loshchilov and F. Hutter, "Decoupled weight decay regularization," (2019), arXiv:1711.05101 [cs.LG].

[33] C. Esteves, A. Makadia, and K. Daniilidis, "Spin-weighted spherical cnns," (2020), arXiv:2006.10731 [cs.CV].

[34] K. Pardo and O. Doré, *"Detecting dark matter subhalos with the Nancy Grace Roman Space Telescope,"* (2021), arXiv:2108.10886 [astro-ph.CO].

Checklist

1. For all authors...
   (a) Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope? [Yes]
   (b) Did you describe the limitations of your work? [Yes] See Sec. 4
   (c) Did you discuss any potential negative societal impacts of your work? [N/A] Potential negative societal impacts were considered, and we believe this work does not present any issues in this regard.
   (d) Have you read the ethics review guidelines and ensured that your paper conforms to them? [Yes]
2. If you are including theoretical results...
   (a) Did you state the full set of assumptions of all theoretical results? [N/A] No theoretical results were obtained in this work.
   (b) Did you include complete proofs of all theoretical results? [N/A]
3. If you ran experiments...
   (a) Did you include the code, data, and instructions needed to reproduce the main experimental results (either in the supplemental material or as a URL)? [Yes] The code repository associated with this paper and needed to reproduce all the results will be linked in Sec. 4. At review stage, it is redacted in order to maintain the required anonymity.
   (b) Did you specify all the training details (e.g., data splits, hyperparameters, how they were chosen)? [Yes] These are described in Sec. 3.
   (c) Did you report error bars (e.g., with respect to the random seed after running experiments multiple times)? [Yes] The main results in Fig. 2 show the expectation evaluated over multiple test datasets. The right panel specifically shows the error bars associated with running over different trials.

(d) Did you include the total amount of compute and the type of resources used (e.g., type of GPUs, internal cluster, or cloud provider)? [No] Due to space constraints limiting the total length of the extended abstract to 4 pages, this information will be included in the camera-ready version of the paper.

4. If you are using existing assets (e.g., code, data, models) or curating/releasing new assets...
   (a) If your work uses existing assets, did you cite the creators? [Yes] All code used for this project is cited in the Acknowledgments section, which is redacted during blind review. Code citations will be reinstated for the camera-ready version of the paper.
   (b) Did you mention the license of the assets? [N/A] Licenses are mentioned in the links associated with individual code packages.
   (c) Did you include any new assets either in the supplemental material or as a URL? [N/A] No new assets (excluding the code used to reproduced the experiments) were produced in this work.
   (d) Did you discuss whether and how consent was obtained from people whose data you're using/curating? [N/A]
   (e) Did you discuss whether the data you are using/curating contains personally identifiable information or offensive content? [N/A] No personal information is included in the assets utilized in this paper.
5. If you used crowdsourcing or conducted research with human subjects...
   (a) Did you include the full text of instructions given to participants and screenshots, if applicable? [N/A]
   (b) Did you describe any potential participant risks, with links to Institutional Review Board (IRB) approvals, if applicable? [N/A]
   (c) Did you include the estimated hourly wage paid to participants and the total amount spent on participant compensation? [N/A]