Characterizing γ-rays maps of the Galactic Center with neural density estimation

Anonymous Author(s)
Affiliation
Address
`email`

Abstract

Machine learning methods have enabled new ways of performing inference on high-dimensional datasets modeled using complex simulations. We leverage recent advancements in simulation-based inference in order to characterize the contribution of various modeled components to γ-ray data of the Galactic Center recorded by the *Fermi* satellite. A specific goal here is to differentiate "smooth" emission, as expected for a dark matter origin, from more "clumpy" emission expected for a population of relatively bright, unresolved astrophysical point sources. Compared to traditional techniques based on the statistical distribution of photon counts, our method based on density estimation using normalizing flows is able to utilize more of the information contained in a given model of the Galactic Center emission, and in particular can perform posterior parameter estimation while accounting for pixel-to-pixel spatial correlations in the γ-ray map.

1 Introduction

Dark matter (DM) represents one of the major unsolved problems in particle physics and cosmology today. The traditional Weakly-Interacting Massive Particle (WIMP) paradigm envisions production of dark matter in the early Universe through freeze-out of dark sector particles weakly coupled to the Standard Model (SM) sector. In this scenario, one of the most promising avenues of detecting a dark matter signal is through an observation of excess γ-ray photons at $\sim \mathrm{GeV}$ energies from DM-rich regions of the sky. The *Fermi* γ-ray Galactic Center Excess (GCE), first identified over a decade ago [1–4] using data from the *Fermi* Large Area Telescope (LAT) [5], is an excess of photons in the Galactic Center with properties—such as energy spectrum and spatial morphology—broadly compatible with expectation due to annihilating DM [6, 7].

The high dimensionality of γ-ray data has traditionally necessitated a description of the photon map in terms of hand-crafted summary quantities *e.g.*, the probability distribution of photon counts [8, 9] or a wavelet decomposition of the photon map [10–13], in order to enable computationally tractable analyses. While effective, this reduced description necessarily involves loss of information compared to that contained in the original γ-ray map. On the other hand, recent developments in machine learning have enabled analysis techniques that can extract more information from high-dimensional datasets. Machine learning methods have recently shown promise for analyzing γ-ray data [14] and specifically for understanding the nature of the *Fermi* GCE [15–17].

Here, we showcase a complementary approach that leverages recent developments in simulation-based inference (SBI, also referred to as likelihood-free inference; see, *e.g.*, Ref. [18] for a recent review) in order to weigh in on the nature of the GCE. In particular, we use conditional density estimation techniques based on normalizing flows [19, 20] to characterize the contributions of various modeled components, including "clumpy" PS-like and "smooth" DM-like emission spatially tracing the GCE, to the γ-ray photon sky at $\sim \mathrm{GeV}$ energies in the Galactic Center region. Rather than using

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

hand-crafted summary statistics, we employ a graph-based convolutional neural network architecture (previously utilized in Refs. [15, 16]) in order to extract summary statistics from γ-ray maps optimized for the downstream task of estimating the distribution of parameters characterizing the contribution of modeled components to the GCE. Unlike traditional approaches based on the statistics of photon counts, this approach lets us capture more of the information contained in a model of the Galactic Center emission, and in particular implicitly uses the distribution of correlations between pixels as an additional discriminating handle. As we show in our extended paper [21] alongside more details and validation tests on the analysis, this fact makes our method more resilient to certain systematic uncertainties associated with model misspecification in real *Fermi* data.

2 Model and inference

The forward model We use the datasets and spatial templates from Refs. [22, 23] to create simulated maps of *Fermi* data in the Galactic Center region. The maps are spatially binned using the `HEALPix` [24] pixelization scheme with resolution parameter `nside=128`, roughly corresponding to pixel area $\sim 0.5\,\mathrm{deg}^2$. The inner region of the Galactic plane, where the observed emission is especially difficult to model, is masked at latitudes $|b| < 2^\circ$, and a radial cut $r < 25^\circ$ defines our region of interest (ROI).

The simulated maps are a combination of diffuse (alternatively referred to as smooth or Poissonian) and PS contributions. The smooth contributions include *(i)* the Galactic diffuse foreground emission [25], *(ii)* spatially isotropic emission accounting for, *e.g.*, uniform emission from unresolved sources of extragalactic origin, *(iii)* emission from resolved PSs included in the *Fermi* 3FGL catalog [26], and *(iv)* lobe-like emission associated with the *Fermi* bubbles [27]. Finally, *(v)* Smooth DM-like emission is modeled using a line-of-sight integral of the (squared) generalized Navarro-Frenk-White (NFW) [28, 29] profile, $\rho_{\mathrm{gNFW}}(r) \propto (r/r_\mathrm{s})^{-\gamma}\,(1 + r/r_\mathrm{s})^{-3+\gamma}$ with inner slope $\gamma = 1.2$ motivated by previous GCE analyses [30, 6, 31]. The total smooth component is obtained as a Poisson realization of a linear combination of these spatial templates.

Assuming the locations of individual PSs are not known a-priori, the statistics of multiple PS populations can be completely specified through *(i)* their spatial distribution, described by templates T^p discretized over pixels p, *(ii)* the distribution of expected photon counts S contributed by each PS, $p(S)$, and *(iii)* the distribution of the number of PSs for each population. Additionally, the modeled instrumental point-spread function quantifies the spatial distribution of photon counts sourced by individual PS around its location due to the finite angular resolution of the instrument. Here, we parameterize the distributions of photon counts S contributed by each PS through a doubly-broken power law specified by the break locations $\{S_{\mathrm{b},1}, S_{\mathrm{b},2}\}$, spectral indices (slopes) $\{n_1, n_2, n_3\}$, and appropriately normalized to unity. Together, we denote these parameters by θ_{PS}.

The PS components of the simulated *Fermi* map are created as follows, practically implemented using the code package `NPTFit-Sim` [32]. The total number of PSs to be simulated is drawn as $n \sim \mathrm{Pois}(n \mid n_{\mathrm{pix}}\lambda)$, where n_{pix} is the number of pixels in the ROI. The sample of PS angular positions is drawn from a PDF constructed by linearly interpolating the relevant pixel-wise spatial template T^p; $\{r_n\} \sim p(r) \propto T(r)$. The expected number of photons emitted by each PS, indexed by i, is drawn by sampling from the mean source-count distribution, $S \sim p\,(S \mid \theta_{\mathrm{PS}})$, and scaling to correct for non-uniform exposure of the satellite. The actual sample of photon counts emitted by the simulated PSs, $\{x_n\}$, is taken to be a Poisson realization of this expectation. The procedure is repeated for each PS population, and the final simulated PS map is constructed by binning the sampled photon positions within the ROI according to the pixelization scheme used. The total map is obtained by combining the simulated diffuse and PS components. The inclusion of PSs in the forward model introduced a large number of latent variable—the positions and fluxes associated with each PS—and renders the full likelihood of the model intractable.

Modeled PS populations are often compactly described through the so-called source-count distribution (SCD) $\mathrm{d}N/\mathrm{d}S$, which quantifies the differential number density of sources per unit angular area (more formally $\mathrm{d}^2N/\mathrm{d}S\mathrm{d}\Omega$, although we leave the area dependence implicit) emitting S photons in expectation. The source-count distribution jointly describes the distribution of photon counts from individual PSs $p(S \mid \theta_{\mathrm{PS}})$ and their mean per-pixel abundance λ, and is related to these as $\mathrm{d}N/x\mathrm{d}S = \lambda\, p(S \mid \theta_{\mathrm{PS}})/\Omega_{\mathrm{pix}}$ where the the pixel area Ω_{pix} is used to convert the per-pixel source count to per-area, agnostic to pixel size. We will present our results in terms of the source fluxes

$\mathrm{d}N/\mathrm{d}F$, with the conversion $S = \langle\epsilon\rangle F$ where $\langle\epsilon\rangle$ is the mean exposure in the region considered. Two PS populations are modeled—*(i)* those correlated with the GCE, following an NFW profile, and *ii* those tracing the Galactic disk, spatially modeled using a doubly exponential profile.

The forward model is thus specified by a total of 18 parameters—6 for the overall normalizations of the Poissonian templates, and 6×2 parameters modeling the source-count distributions associated with GCE-correlated and disk-correlated PS populations $\{\langle S^{\mathrm{PS}}\rangle, n_1, n_2, n_3, S_{\mathrm{b},1}, S_{\mathrm{b},2}\}$. $\langle S^{\mathrm{PS}}\rangle$ denotes the mean per-pixel counts contributed by a given PS population, and parameterizes their overall abundance.

Inference with likelihoods based on simplified data representations The 1-point PDF (probability distribution function) framework, first introduced in the context of γ-ray analyses in Ref. [33] and extended in Refs. [8, 9] under the name of non-Poissonian template fitting (NPTF), considers a simplification of the problem by computing the pixel-wise likelihood assuming each pixel to be statistically independent (*1-point* then referring to values over individual, independent spatial positions in the sky). This significantly reduces the latent space dimensionality by eliminating the positions of individual PSs as latent variables, localizing them within a pixel and modulating their expected abundance by the modeled spatial template (*e.g.*, GCE-correlated or disk-correlated in our case). Here, we use this method as a comparison point, and sample the posterior associated with parameters of interest with `dynesty` [34] using the likelihood from `NPTFit` [23].

Extracting representative features from γ-ray maps Rather than relying on hand-crafted data summaries, a neural network is used to extract representative features $s_\varphi(x)$ of the data x optimized for the downstream density estimation task. We use the `DeepSphere` architecture [35–37] with a configuration similar to and inspired by that employed in Ref. [15]. `DeepSphere` is a graph-based spherical convolutional neural network architecture tailored to data sampled on a sphere, and in particular is able to leverage the hierarchical structure of data in the `HEALPix` representation. This makes it well-suited for our purposes. The architecture consists of graph convolutional layers which, following a ReLU nonlinearity, coarsens the pixel representation by a factor of 4 with max pooling while doubling the number of feature dimensions until a maximum of 256. The output of the final convolution layer is passed through a fully connected neural network with 1024 hidden units before outputting 128 summaries.

Simulation-based inference with normalizing flows Simulation-based inference (SBI) refers to a class of methods for performing inference when the data-generating process does not have a tractable likelihood. This is the case for the forward model used here, where the presence of a large number of PSs leads to a large latent space. We approximate the joint posterior over the parameters of interest θ given a γ-ray map x through a distribution $\hat{p}_\phi(\theta \mid s_\varphi(x))$ conditioned on summaries $s_\varphi(x)$ from simulated samples $\{x\}$. The conditioned posterior distribution is parameterized by ϕ and defined via normalizing flows [19, 20], which are a class of models that provide an efficient way of modeling high-dimensional probability distributions. Specifically, we use Masked Autoregressive Flows (MAFs) [38] to define the flow transformation. We use 8 MAF transformations, each made up of a 2-layer masked autoregressive neural network [39] with 128 hidden units and $\tanh$ activations. Each transformation is conditioned on summaries $s_\varphi(x)$ by including these as additional inputs into the transformation blocks.

Normalizing flows allow for tractable density evaluation, and $\log \hat{p}_\phi(\theta \mid s_\varphi(x))$ is used as the training objective to simultaneously optimize parameters $\{\phi, \varphi\}$ associated with the convolution and flow neural networks, respectively. 10^6 samples from the forward model are produced, with 15% of these held out for validation. The model is trained for up to 30 epochs with early stopping, using a batch size of 256. The `AdamW` optimizer [40, 41] is used with initial learning rate 10^{-3} and weight decay 10^{-5}, with learning rate decayed through cosine annealing.

3 Application to *Fermi* data

We apply our neural simulation-based inference pipeline to the real *Fermi* dataset. As a point of comparison, we also run the NPTF method on the data using the same spatial templates and prior assumptions as those used in the corresponding SBI analyses. The results of the NPTF analysis are shown in the bottom panel of Fig. 1. The left column shows the median (solid lines) as well

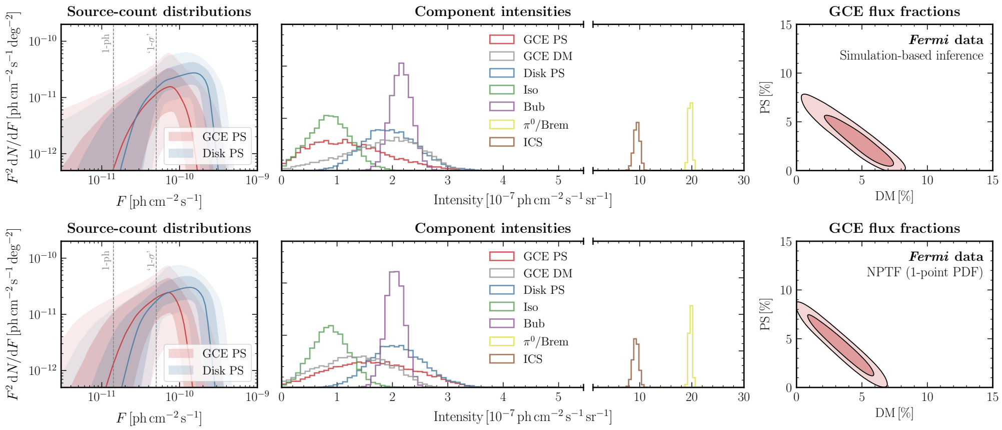

Figure 1: Results of the baseline analysis on real *Fermi* data. *(Top row)* Analysis using neural simulation-based inference with normalizing flows, and *(bottom row)* using the 1-point PDF likelihood implemented in the non-Poissonian template fitting (NPTF) framework. While moderate preference for a PS-like origin of the GCE is seen in the case of the NPTF analysis (bottom), the simulation-based inference analysis attributes a smaller fraction of the GCE to PS-like emission (top).

as middle-68/95% containment (dark/light shaded regions) of the posteriors on the source-count distributions $F^2 \mathrm{d}N/\mathrm{d}F$ of GCE-correlated (red) and disk-correlated (blue) PSs, evaluated point-wise in flux F. The dashed grey vertical lines correspond to the flux associated with a single expected photon count per source (below which Poissonian and PS-like emission is expected to be perfectly degenerate) and the approximate 1-σ threshold for detecting individual sources (below which the degeneracy is often observed in practice [42, 25]). The middle column shows the posteriors on various modeled emission components. The right column shows the joint posterior on the fraction of DM- and PS-like emission in proportion to the total inferred flux in the ROI.

Consistent with previous 1-point PDF studies using a similar configuration, a significant fraction of the GCE—$55.0^{+8.8}_{-22.9}\%$–is attributed to PS-like emission. The top panel of Fig. 1 shows results using the neural simulation-based analysis pipeline introduced in this paper. Although posteriors for the astrophysical background templates are seen to be broadly consistent with those inferred in the NPTF analysis, the preference for PSs is somewhat reduced in this case, with $38.0^{+9.0}_{-19.4}\%$ of the GCE emission being PS-like.

4 Discussion

We have leveraged recent advances in neural simulation-based inference in order to jointly characterize a putative DM-like signal and PS population associated with the observed *Fermi* Galactic Center Excess. While broadly consistent with results of the traditional method, our method shows a reduced preference for PS-like emission correlated with the GCE. In our extended work [21], we present additional details of our analysis, including a validation of our pipeline on simulated data as well as a discussion of the impact of model misspecification within our framework. We show there that, owing to the fact that it can extract more information from the forward model, our method is less sensitive to certain forms of model misspecification than traditional approaches.

As in any Galactic Center γ-ray analysis, we caution of the potential of unknown systematics, such as mismodeling on the scale of the size of the LAT point-spread function, to bias the results and conclusions of our analysis. Although machine learning-based analyses can utilize more of the information encoded in the forward model, and in particular in the present case can take advantage of pixel-to-pixel correlations, this can also make them more susceptible to specific modeled features compared to traditional techniques based on data reduction to hand-crafted data summaries. We leave a more detailed investigation of the impact of these effects to future work.

Code used for reproducing the results presented in this paper is available at https://github.com/smsharma/fermi-gce-flows.

References

[1] L. Goodenough and D. Hooper, *"Possible Evidence For Dark Matter Annihilation In The Inner Milky Way From The Fermi Gamma Ray Space Telescope,"* (2009), arXiv:0910.2998 [hep-ph].

[2] D. Hooper and L. Goodenough, *"Dark Matter Annihilation in The Galactic Center As Seen by the Fermi Gamma Ray Space Telescope,"* Phys. Lett. B **697**, 412 (2011), arXiv:1010.2752 [hep-ph].

[3] A. Boyarsky, D. Malyshev, and O. Ruchayskiy, *"A comment on the emission from the Galactic Center as seen by the Fermi telescope,"* Phys. Lett. B **705**, 165 (2011), arXiv:1012.5839 [hep-ph].

[4] D. Hooper and T. Linden, *"On The Origin Of The Gamma Rays From The Galactic Center,"* Phys. Rev. D **84**, 123005 (2011), arXiv:1110.0006 [astro-ph.HE].

[5] W. B. Atwood *et al.* (Fermi-LAT), *"The Large Area Telescope on the Fermi Gamma-ray Space Telescope Mission,"* Astrophys. J. **697**, 1071 (2009), arXiv:0902.1089 [astro-ph.IM].

[6] T. Daylan, D. P. Finkbeiner, D. Hooper, T. Linden, S. K. N. Portillo, N. L. Rodd, and T. R. Slatyer, *"The characterization of the gamma-ray signal from the central Milky Way: A case for annihilating dark matter,"* Phys. Dark Univ. **12**, 1 (2016), arXiv:1402.6703 [astro-ph.HE].

[7] F. Calore, I. Cholis, and C. Weniger, *"Background Model Systematics for the Fermi GeV Excess,"* JCAP **03**, 038 (2015), arXiv:1409.0042 [astro-ph.CO].

[8] S. K. Lee, M. Lisanti, and B. R. Safdi, *"Distinguishing Dark Matter from Unresolved Point Sources in the Inner Galaxy with Photon Statistics,"* JCAP **05**, 056 (2015), arXiv:1412.6099 [astro-ph.CO].

[9] S. K. Lee, M. Lisanti, B. R. Safdi, T. R. Slatyer, and W. Xue, *"Evidence for Unresolved γ-Ray Point Sources in the Inner Galaxy,"* Phys. Rev. Lett. **116**, 051103 (2016), arXiv:1506.05124 [astro-ph.HE].

[10] R. Bartels, S. Krishnamurthy, and C. Weniger, *"Strong support for the millisecond pulsar origin of the Galactic center GeV excess,"* Phys. Rev. Lett. **116**, 051102 (2016), arXiv:1506.05104 [astro-ph.HE].

[11] B. Balaji, I. Cholis, P. J. Fox, and S. D. McDermott, *"Analyzing the Gamma-Ray Sky with Wavelets,"* Phys. Rev. D **98**, 043009 (2018), arXiv:1803.01952 [astro-ph.HE].

[12] S. D. McDermott, P. J. Fox, I. Cholis, and S. K. Lee, *"Wavelet-Based Techniques for the Gamma-Ray Sky,"* JCAP **07**, 045 (2016), arXiv:1512.00012 [astro-ph.HE].

[13] Y.-M. Zhong, S. D. McDermott, I. Cholis, and P. J. Fox, *"Testing the Sensitivity of the Galactic Center Excess to the Point Source Mask,"* Phys. Rev. Lett. **124**, 231103 (2020), arXiv:1911.12369 [astro-ph.HE].

[14] S. Caron, K. Dijkstra, C. Eckner, L. Hendriks, G. Jóhannesson, B. Panes, R. Ruiz De Austri, and G. Zaharijas, *"Identification of point sources in gamma rays using U-shaped convolutional neural networks and a data challenge,"* (2021), arXiv:2103.11068 [astro-ph.HE].

[15] F. List, N. L. Rodd, G. F. Lewis, and I. Bhat, *"The GCE in a New Light: Disentangling the γ-ray Sky with Bayesian Graph Convolutional Neural Networks,"* Phys. Rev. Lett. **125**, 241102 (2020), arXiv:2006.12504 [astro-ph.HE].

[16] F. List, N. L. Rodd, and G. F. Lewis, *"Dim but not entirely dark: Extracting the Galactic Center Excess' source-count distribution with neural nets,"* (2021), arXiv:2107.09070 [astro-ph.HE].

[17] S. Caron, G. A. Gómez-Vargas, L. Hendriks, and R. Ruiz de Austri, *"Analyzing γ-rays of the Galactic Center with Deep Learning,"* JCAP **05**, 058 (2018), arXiv:1708.06706 [astro-ph.HE].

[18] K. Cranmer, J. Brehmer, and G. Louppe, *"The frontier of simulation-based inference,"* Proceedings of the National Academy of Sciences **117**, 30055 (2020).

[19] G. Papamakarios, E. Nalisnick, D. J. Rezende, S. Mohamed, and B. Lakshminarayanan, *"Normalizing flows for probabilistic modeling and inference,"* arXiv preprint arXiv:1912.02762 (2019).

[20] D. Rezende and S. Mohamed, *"Variational inference with normalizing flows,"* in *International Conference on Machine Learning* (PMLR, 2015) pp. 1530–1538.

[21] S. Mishra-Sharma and K. Cranmer, *"A neural simulation-based inference approach for characterizing the Galactic Center γ-ray excess,"* In preparation .

[22] S. Mishra-Sharma, N. L. Rodd, and B. R. Safdi, "Supplementary material for NPTFit," (2016).

[23] S. Mishra-Sharma, N. L. Rodd, and B. R. Safdi, *"NPTFit: A code package for Non-Poissonian Template Fitting,"* Astron. J. **153**, 253 (2017), arXiv:1612.03173 [astro-ph.HE].

[24] K. M. Gorski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelman, *"HEALPix - A Framework for high resolution discretization, and fast analysis of data distributed on the sphere,"* Astrophys. J. **622**, 759 (2005), arXiv:astro-ph/0409513.

[25] M. Buschmann, N. L. Rodd, B. R. Safdi, L. J. Chang, S. Mishra-Sharma, M. Lisanti, and O. Macias, *"Foreground Mismodeling and the Point Source Explanation of the Fermi Galactic Center Excess,"* Phys. Rev. D **102**, 023023 (2020), arXiv:2002.12373 [astro-ph.HE].

[26] F. Acero *et al.* (Fermi-LAT), *"Fermi Large Area Telescope Third Source Catalog,"* Astrophys. J. Suppl. **218**, 23 (2015), arXiv:1501.02003 [astro-ph.HE].

[27] M. Su, T. R. Slatyer, and D. P. Finkbeiner, *"Giant Gamma-ray Bubbles from Fermi-LAT: AGN Activity or Bipolar Galactic Wind?"* Astrophys. J. **724**, 1044 (2010), arXiv:1005.5480 [astro-ph.HE].

[28] 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 [astro-ph].

[29] J. F. Navarro, C. S. Frenk, and S. D. White, *"A Universal density profile from hierarchical clustering,"* Astrophys.J. **490**, 493 (1997), arXiv:astro-ph/9611107 [astro-ph].

[30] C. Gordon and O. Macias, *"Dark Matter and Pulsar Model Constraints from Galactic Center Fermi-LAT Gamma Ray Observations,"* Phys. Rev. D **88**, 083521 (2013), [Erratum: Phys.Rev.D 89, 049901 (2014)], arXiv:1306.5725 [astro-ph.HE].

[31] B. Zhou, Y.-F. Liang, X. Huang, X. Li, Y.-Z. Fan, L. Feng, and J. Chang, *"GeV excess in the Milky Way: The role of diffuse galactic gamma-ray emission templates,"* Phys. Rev. D **91**, 123010 (2015), arXiv:1406.6948 [astro-ph.HE].

[32] N. L. Rodd and M. W. Toomey, *NPTFit-Sim* (2017).

[33] D. Malyshev and D. W. Hogg, *"Statistics of gamma-ray point sources below the Fermi detection limit,"* Astrophys. J. **738**, 181 (2011), arXiv:1104.0010 [astro-ph.CO].

[34] J. S. Speagle, *"dynesty: a dynamic nested sampling package for estimating bayesian posteriors and evidences,"* Monthly Notices of the Royal Astronomical Society **493**, 3132 (2020).

[35] M. Defferrard, M. Milani, F. Gusset, and N. Perraudin, *"Deepsphere: a graph-based spherical cnn,"* arXiv preprint arXiv:2012.15000 (2020).

[36] 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].

[37] M. Defferrard, N. Perraudin, T. Kacprzak, and R. Sgier, *"DeepSphere: towards an equivariant graph-based spherical CNN,"* in *ICLR Workshop on Representation Learning on Graphs and Manifolds* (2019) arXiv:1904.05146.

[38] G. Papamakarios, T. Pavlakou, and I. Murray, *"Masked autoregressive flow for density estimation,"* in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, NIPS'17 (Curran Associates Inc., Red Hook, NY, USA, 2017) pp. 2335–2344.

[39] M. Germain, K. Gregor, I. Murray, and H. Larochelle, *"Made: Masked autoencoder for distribution estimation,"* in *International Conference on Machine Learning* (PMLR, 2015) pp. 881–889.

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

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

[42] L. J. Chang, S. Mishra-Sharma, M. Lisanti, M. Buschmann, N. L. Rodd, and B. R. Safdi, *"Characterizing the nature of the unresolved point sources in the Galactic Center: An assessment of systematic uncertainties,"* Phys. Rev. D **101**, 023014 (2020), arXiv:1908.10874 [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 is linked in Sec. 4.
   (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]
   (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]