XXXIII Canary Islands Winter School of Astrophysics

Overlaps at the Frontiers of Astrophysics, Cosmology and Particle Physics

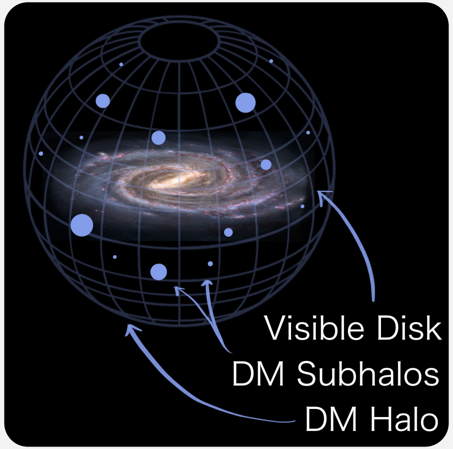
Nov 28th 2022

A Machine Learning Approach to Searching Dark Matter Subhalos in Fermi-LAT Sources

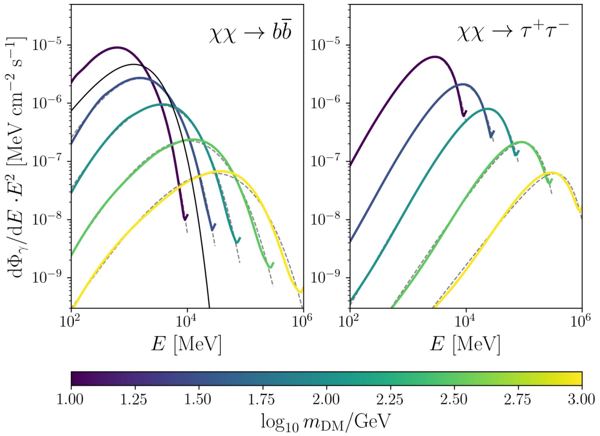
Anja Butter [1], *Michael Krämer* [2], *Silvia Manconi* [2], ***Kathrin Nippel*** [2]

[1] *ITP, U. Heidelberg* [2] *TTK RWTH Aachen*

Physics Motivation



- Galaxy populated by clumps of dark matter
  ➜ N-body simulations*
- Assuming WIMP dark matter:
  $\chi\chi \rightarrow$ SM SM ($\rightarrow \gamma$)
  ➜ A signal like this could already be detected among Fermi-LAT sources**



- The Fermi-LAT 4FGL source catalog can help constrain the properties of dark matter
  1. Create realistic set of subhalo simulations
  2. Assess detectability
  3. Look for subhalo-like spectra among unclassified sources
- Machine Learning is a powerful tool for classification tasks***
  ➜ We employ a neural network to effectively classify DM subhalos

This talk

* see e.g. Zavala, Frenk (2019) 1907.1175
Springel et al. (2008) 0809.0898

** see e.g. Hooper, Witte (2017) 1610.07587
Coronado-Blásquez et al (2019) 1910.14429
Calore et al. (2019) 1910.13722
Di Mauro et al. (2020) 2007.08535
Gammaldi et al. (2022) 2207.09307
...

*** see e.g. Finke et al. (2021) 2012.05251
Butter et al. (2022) 2112.01403

Simulations
Subhalo Population

PPPC 4 DM ID: *Cirelli et al. (2012)*

DM annihilation spectra for each mass, and primary annihilation channel, assuming WIMPs

CLUMPY V3: *Hütten et al. (2018)*

J-factor and sky position of galactic subhalos

https://clumpy.gitlab.io/CLUMPY/

fermipy: *Wood et al. (Fermi-LAT collaboration, 2017)*

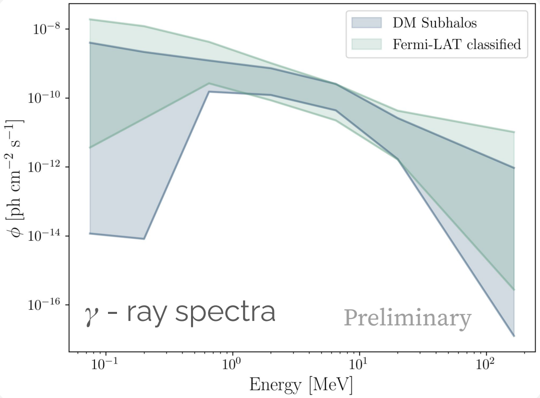
Simulate detector effects

DM model dependent: Prefactor, PPPC 4 DM ID

$$\phi = \underbrace{\frac{\langle \sigma v \rangle}{8 \cdot \pi \cdot m_{\mathrm{DM}}}}_{\text{Prefactor}} \cdot \mathcal{J} \cdot \underbrace{\frac{\mathrm{d}N}{\mathrm{d}E}}_{\text{PPPC 4 DM ID}}$$

$\mathcal{J}$: CLUMPY – Halo model dependent

Initial / Benchmark Setup	
Halo model	DM only
m_{DM}	80 GeV
$\langle \sigma v \rangle$	10^{-23} cm^3 s^{-1}
Final state	$b\bar{b}$



➜ Benchmark classification training set for comparing subhalos with 4FGL catalog
- Realistic scenario with simulations as close as possible to real sources
- Number of detectable subhalos sufficient for ML approach

Simulations

Detector Effects

Use **fermipy** for simulating 12 years of Fermi-LAT data

Input: Individual subhalo with given position in sky & flux fitted with 'PLSuperExpCutoff'*

$$\phi = \phi_0 \left(\frac{E}{E_0}\right)^{\gamma} \exp\left(-\left(\frac{E}{E_0}\right)^{\beta}\right)$$

↓

Define ROI around subhalo

↓

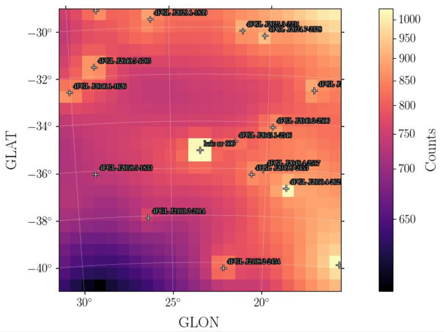
Fit source among background (diffuse + isotropic) & point sources (4FGL-DR3)

↓

Detection threshold

$$\mathrm{TS} = 2\log\left(\mathcal{L}/\mathcal{L}_0\right) \overset{!}{\geq} 25$$

ROI counts map

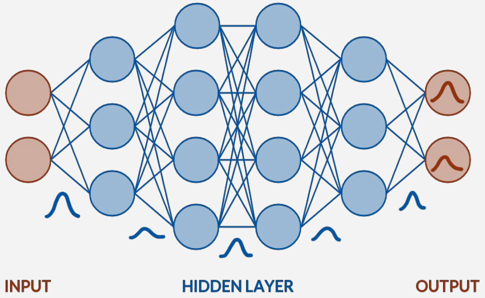


* see also Calore et al. (2017) 1611.03503

➜ Benchmark classification training set for comparing subhalos with 4FGL catalog

- Realistic scenario with simulations as close as possible to real sources
- Number of detectable subhalos sufficient for ML approach

Machine Learning Approach
Bayesian Neural Network Classification



- Replace individual weight of Dense NN with weight distributions
  - Shape of distribution allows for uncertainty estimation of outputs
  - BNN learns posterior distribution $p\left(w|D\right)$ by approximating variational weight distribution $q_\theta\left(w\right)$ using the KL-divergence

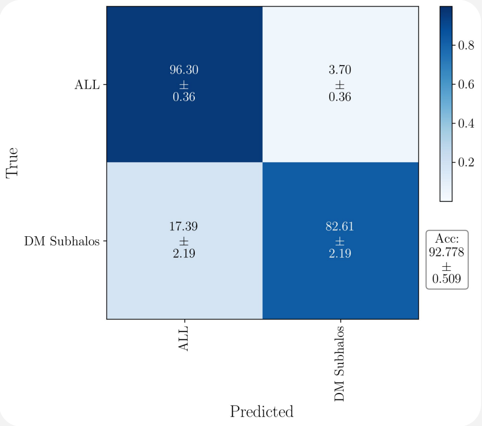
$$
\begin{aligned}
\mathrm{KL}(q(w)||p(w|D)) &= \int dw\; q(w) \log \frac{q(w)}{p(w|D)} \\
&= \int dw\; q(w) \log \frac{q(w)}{p(D|w)p(w)} + \mathrm{const} \\
&= \mathrm{KL}(q(w)||p(w)) - \int dw\;\; q(w) \log(p(D|w)) + \mathrm{const} \\
&= \sum_i \log \frac{\sigma_{p,i}}{\sigma_{q,i}} + \frac{\sigma_{q,i}^2 + (\mu_{p,i} - \mu_{q,i})^2}{2\sigma_{p,i}^2} - \frac{1}{2}
\end{aligned}
$$

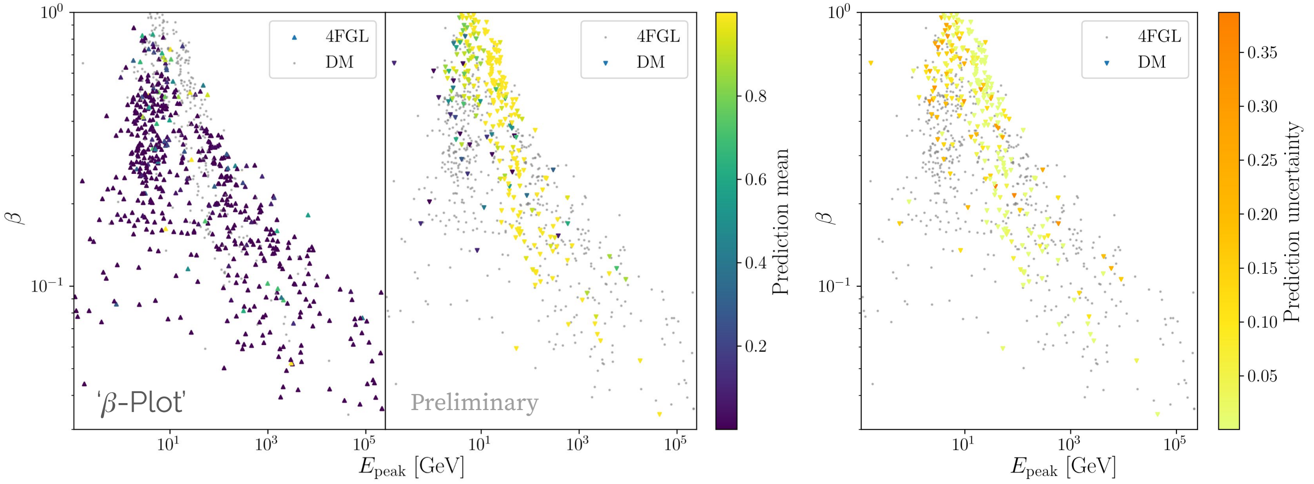
Assuming multivariate, diagonal Gaussians

- In practice: Use the Flipout estimator (Wen et al., 2018)
  - Performs a Monte Carlo approximation of the distribution integrating over the weight and bias to minimize the KL-divergence

Preliminary Results

Subhalo vs 4FGL Prediction Uncertainty

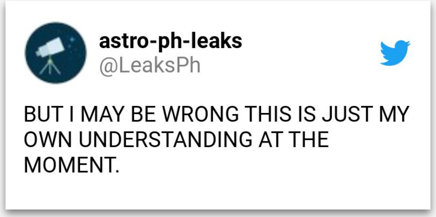


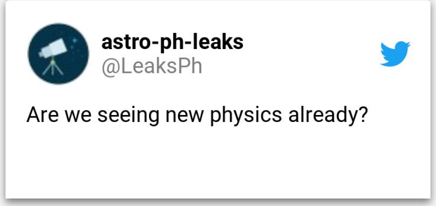


LogParabola fit: $\phi = \phi_0 \left(\frac{E}{E_0}\right)^{-\alpha-\beta\cdot\log(E/E_0)}$ $\quad E_{\mathrm{peak}} = E_0 \cdot e^{\frac{2-\alpha}{2\beta}}$

- ➜ Trained network can give reliable estimate on which unclassified sources in 4FGL are compatible with DM subhalo model at hand
  - ➜ Limits of accuracy: Inherent statistics in data and simulation

Conclusions & Outlook





- Using **CLUMPY**, PPPC 4 DM ID and **fermipy**, we have constructed a set of realistic DM subhalo simulations for a given model
- We have carefully evaluated the detectability using complete simulations of 12 years of Fermi-LAT data and used this to compare to the 4FGL-DR3 source catalog
- We use a Bayesian Neural Network classification approach to
  - Estimate the uncertainty of γ-ray classifier predictions
  - Apply classification to unclassified 4FGL sources to gauge a number of DM subhalo candidates
- This approach can be extended to any DM model