Dark matter constraints from dwarf galaxies: a data-driven analysis

Bryan Zaldívar [a1]

[1]Institute of Corpuscular Physics, University of Valencia and CSIC, Spain

The dwarf satellite galaxies of the Milky Way are ideal targets to search for non-gravitational interactions of dark matter using gamma-ray data. A new data-driven method is proposed to bypass the currently limited knowledge of the physics behind standard astrophysical emissions.

Dark matter (DM) constitutes about 27% of the matter-energy content of our Universe and discovering its nature is one of the biggest challenges of contemporary physics. In the leading paradigm, DM is a new particle, beyond the framework of the standard model for particle physics. Several candidates in different ranges of mass and interaction strength exist and can account for the purely gravitational DM observational evidence. One of the most promising class of candidates are weakly interacting massive particles (WIMPs), which can naturally account for the observed DM cosmological relic density with annihilation cross section $\langle \sigma v \rangle$ around the benchmark value of 3×10^{-26} cm^3/s. Via their residual annihilations today, these DM particle candidates can produce stable particles such as photons and charged cosmic rays as a consequence of decay and hadronisation of the final states (quarks, gauge bosons, leptons, higgses), see Fig. 1. The ensuing signal can, in principle, be detected by space- and ground-based telescopes. However, this possibility is challenged by the typically dominant astrophysical emissions, which one has to disentangle a WIMP signal from. Among the best channels to look for DM is the high energy gamma-rays, from hundreds of MeV up to TeV energies. One of the most promising targets are dwarf spheroidal galaxies (dSphs), satellites of the Milky Way located within a few hundreds kiloparsec from the Galactic centre and whose mass is dominated by DM, as inferred from the kinematics of their stars. With little intrinsic astrophysical background emission expected, dSphs could be shining in gamma rays mostly in reason of their DM content, through annihilation (or decay) of WIMP particles.

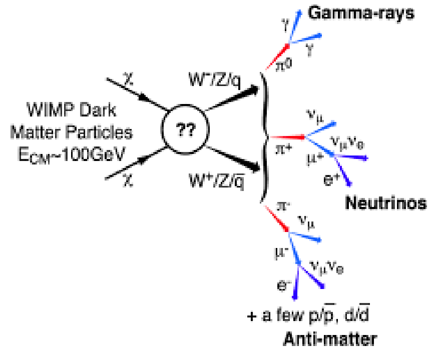

Figure 1: Pair annihilation of DM particles χ into different possible Standard Model particles. Fig. taken from the Fermi-LAT collaboration, fermi.gsfc.nasa.gov.

The leading experiment measuring GeV-to-TeV gamma-ray data from the whole sky, including dSphs, is the Large Area Telescope, aboard the Fermi satellite (hereafter, Fermi-LAT) from the NASA, see Fig. 2. The Fermi collaboration models the standard astrophysical diffuse emission of gamma-rays as a linear combination of spatial "templates" including different physical components, as the neutral hydrogen as well as CO interstellar gas, inverse compton scattering, plus a modelling of the emission from large-scale structures as the so-called "Fermi-bubbles", extragalactic isotropic emission, and finally cataloged point-like sources. Such modelling has several limitations though, one of the most important ones being its incapability to account for a spatially-dependent unresolved population of unknown astrophysical objects. In addition to that, it is unclear how to determine in a sound way the uncertainties associated to the expected astrophysical emission from these templates.

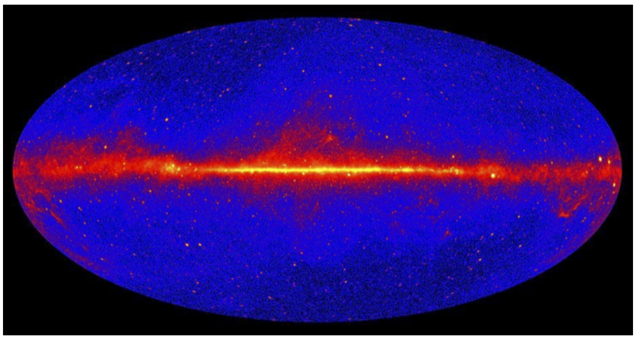

Figure 2: All-sky gamma-ray map produced by Fermi-LAT using 9 years of data collected from 2008 to 2017. See fermi.gsfc.nasa.gov.

In the context of DM searches, the above astrophysical emissions constitute the background against which to extract a potential DM signal, the latter representing the physics of interest. In this sense, the background emission can be thought (in a statistical jargon) as a "nuisance contribution". The idea, first presented in Ref. [1], is then to model the background emission using Machine Learning (ML) techniques, in a purely data-driven way, agnostic to the physics gov-

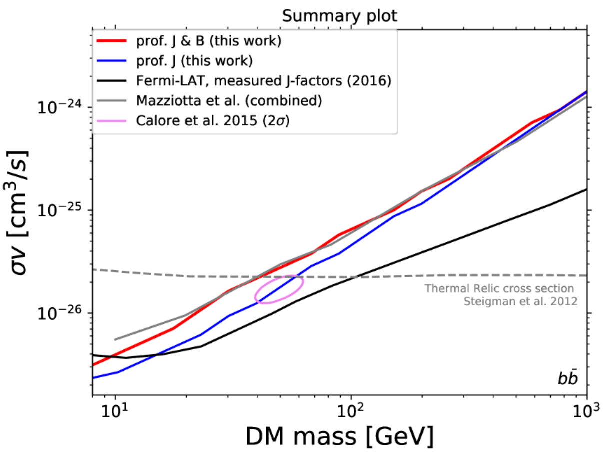

Figure 3: Comparison of current limits on the DM parameters: the main result of this work when profiling over J-factor (blue line) and also background uncertainties (red line), Fermi-LAT (black line), and other data-driven approaches (grey line). The purple contour represents the 2σ best-fit region.

erning the background. Once we have estimated the expected number of gamma-ray counts from the background in this way, a traditional procedure is followed to compare it with the observations at the different dSph galaxies, in order to make a statistical statement about the DM emission.

A data-driven model of the background emission

Traditional Machine Learning (ML) models assume a parametric probability distribution (likelihood) of the observables y. For example, in the case of real variables, a common assumption is a Gaussian likelihood, $y(\mathbf{x}) \sim \mathcal{N}(\mu(\mathbf{x};\theta), \sigma)$, with mean $\mu(\mathbf{x};\theta)$ given by some non-linear function (e.g. a neural network) with parameters θ, and a d-dimensional input $\mathbf{x}$. However, in many situations the Gaussian assumption is not justified, and may give poor predictive performances. We follow instead another ML approach, concerning the so-called " non-parametric" models, giving a likelihood function whose expressivity increases with the number of data points considered. Concretely, we write the joint likelihood $p(\mathbf{x}, y)$ as:

$$p(\mathbf{x}, y) = \frac{1}{N}\sum_{i=1}^{N} k_\sigma(\mathbf{x}-\mathbf{x}_i) g_\varsigma(y-y_i) \ , \qquad (1)$$

where the index i runs over all the N available data points, while $k_\sigma(\cdot)$ and $g_\varsigma(\cdot)$ are kernel functions with parameters σ and ς respectively. For our particular problem, $\mathbf{x}$ represent the vector of galactic longitude and latitude, while the observable y represents the number of photons measured by Fermi-LAT. Under very weak hypotheses of continuity and smoothness, Eq. (1) is guaranteed to provide an unbiased estimator for the true (but unknown) underlying likelihood. The above joint likelihood $p(\mathbf{x}, y)$ provide us a way to estimate the expected value $\lambda_{\rm bckg}(\mathbf{x})$ of the observable as:

$$\lambda_{\rm bckg}(\mathbf{x}) \equiv \mathbb{E}_{p(y|\mathbf{x})}[y] = \frac{\sum_{i=1}^{N} k_\sigma(\mathbf{x}-\mathbf{x}_i) y_i}{\sum_{i=1}^{N} k_\sigma(\mathbf{x}-\mathbf{x}_i)} \ . \qquad (2)$$

In order to make predictions, we first need to fix the model parameters σ, ς. In this work we follow a Maximum Likelihood Estimate procedure, where we choose the parameters that maximise the global log-likelihood of the whole data.

Bounds on the dark matter interactions

The expected background emission predicted with the above procedure should be added to the expected emission coming from the (hypothetical) DM interactions with the photons. We choose a typical WIMP dark matter, and for concreteness, focus on the case where these particles self-annihilate directly to bottom quarks, which subsequently decay and produce a spectrum of gamma-ray photons. The expected number of photons coming from DM annihilation at a given dwarf galaxy d can be expressed as:

$$\lambda_{\rm DM} = J_d \langle\sigma v\rangle f(m_{\rm DM}) \ , \qquad (3)$$

where J_d encodes the DM content (actually, its integral along the line-of-sight) at the dSph d, $\langle\sigma v\rangle$ is the velocity-averaged DM annihilation cross section, and the function $f(\cdot)$, depending on the DM mass $m_{\rm DM}$, encodes the spectral information of the DM annihilation.

As customary in the literature, we are interested in extracting bounds on the DM annihilation, for which we follow a frequentist approach. Fig. 3 shows a summary of our results. The solid lines represent the upper bound on $\langle\sigma v\rangle$ as a function of the DM mass. The most statistically complete result corresponds to the red curve, where we have in a way marginalised over the background contributions, something that was not possible so far in the literature. Comparing with the official results from Fermi-LAT (see black curve) we realise that our bounds on DM turn out to be more conservative, which is the result of taking into account the background uncertainties (unlike the Fermi result).

Concluding remarks

A proper data-driven inclusion of background uncertainties in the derivation of DM upper limits is not only possible, but very much needed for a fair comparison with limits from other targets, where theoretical background uncertainties are traditionally more explored than in dSphs analyses.

Notes

a. Email: b.zaldivar.m@csic.es

References

[1] F. Calore, P. D. Serpico and B. Zaldivar, *JCAP* ***10*** *(2018), 029 [arXiv:1803.05508 [astro-ph.HE]].*