SEARCHING FOR DARK MATTER IN FERMI-LAT UNIDENTIFIED SOURCES WITH MACHINE LEARNING

To be submitted, proceedings: PoS(ICRC2021)493 - SISSA

VIVIANA GAMMALDI

Departamento de Física Teórica, Universidad Autónoma de Madrid (UAM), Madrid, Spain
Instituto de Física Teórica (IFT UAM-CSIC), Madrid, Spain

IN COLLABORATION WITH

B. ZALDIVAR, J. CORONADO-BLAZQUÉZ, M. A. SÁNCHEZ-CONDE

19th MultiDark Consolider Workshop

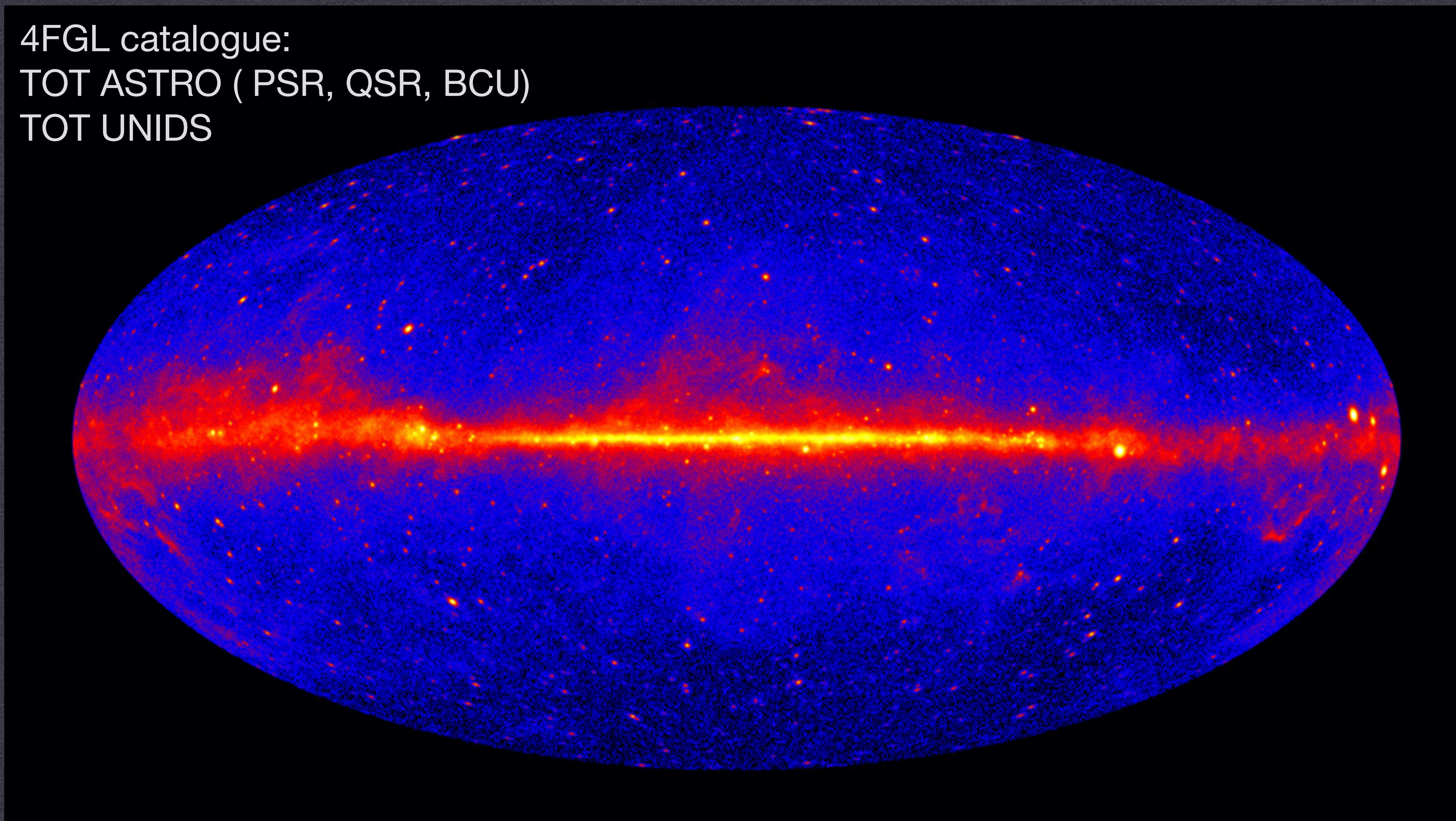
OUTLINE

- FERMI-LAT GAMMA-RAY DATA & BETA-PLOT
- DARK MATTER & BETA-PLOT
- “SYNTHETIC” FEATURES:
  DETECTION SIGNIFICANCE σ_d AND UNCERTAINTY ON β
- INTRODUCTION TO CLASSIFICATION IN MACHINE LEARNING
- PRELIMINARY RESULTS
- PRELIMINARY CONCLUSIONS

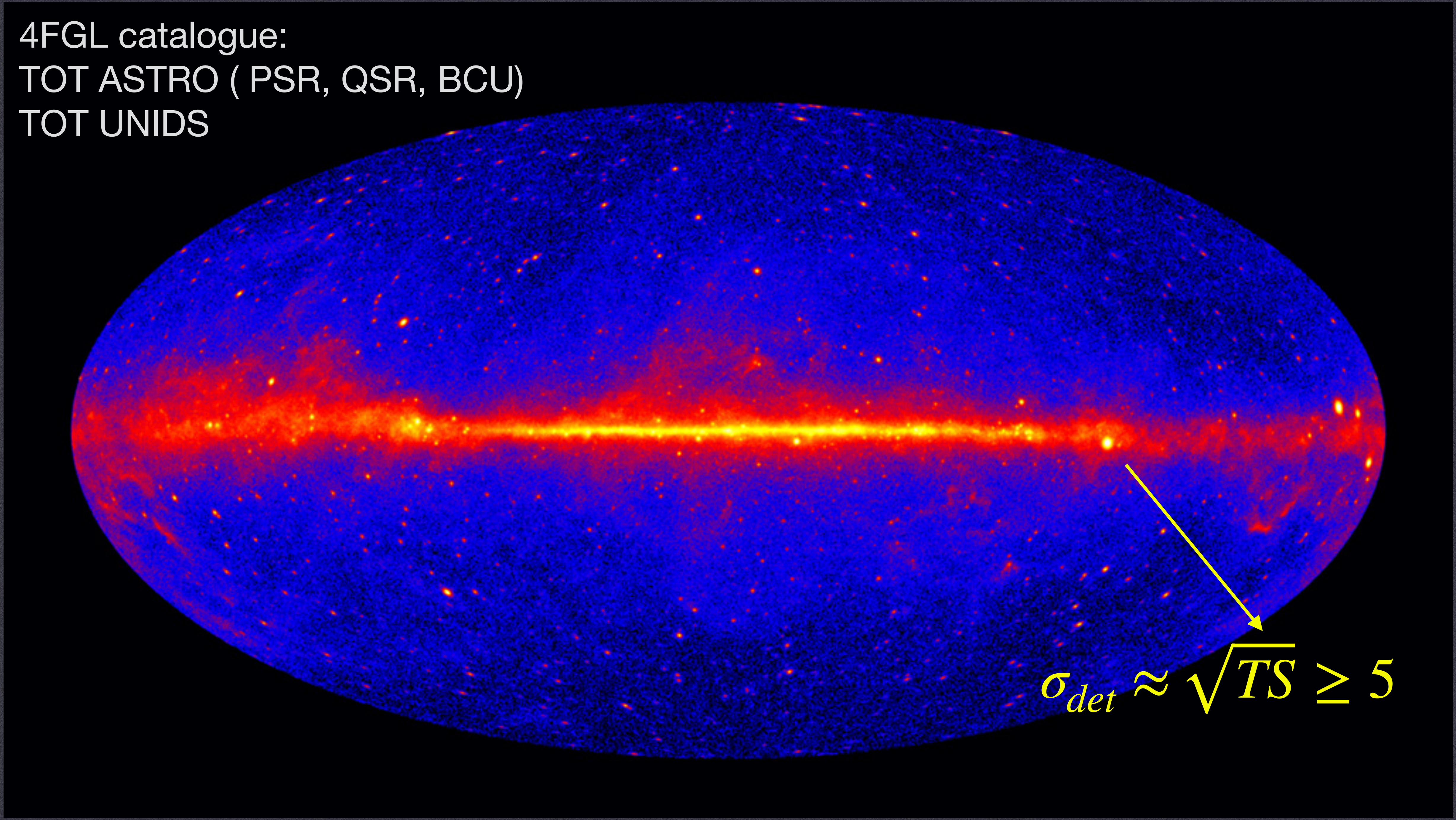
OUTLINE

- FERMI-LAT GAMMA-RAY DATA & BETA-PLOT
- DARK MATTER & BETA-PLOT
- “SYNTHETIC” FEATURES:
  DETECTION SIGNIFICANCE σ_d AND UNCERTAINTY ON β
- INTRODUCTION TO CLASSIFICATION IN MACHINE LEARNING
- PRELIMINARY RESULTS
- PRELIMINARY CONCLUSIONS

FERMI-LAT GAMMA-RAY DATA & BETA-PLOT

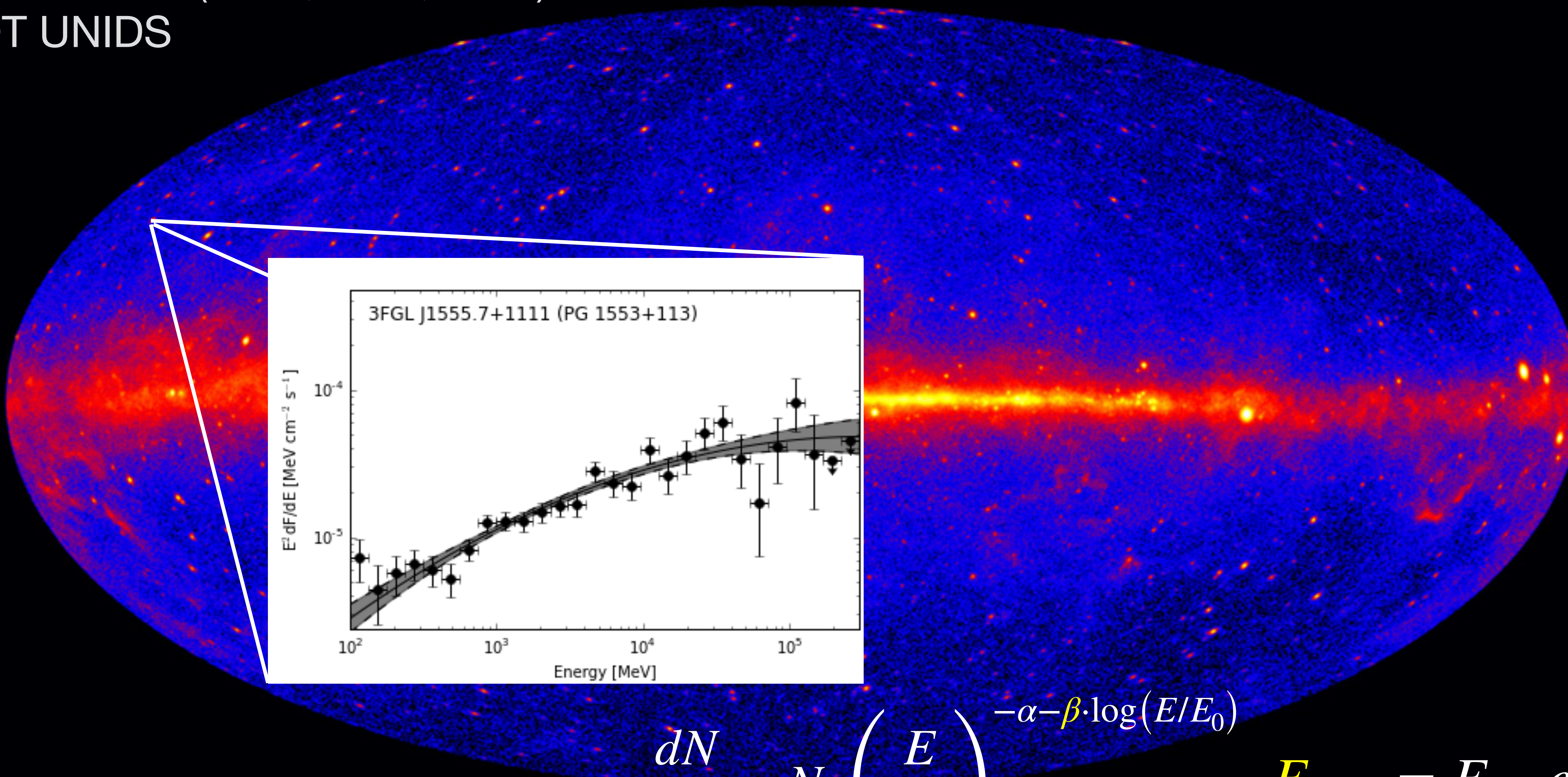


FERMI-LAT GAMMA-RAY DATA & BETA-PLOT



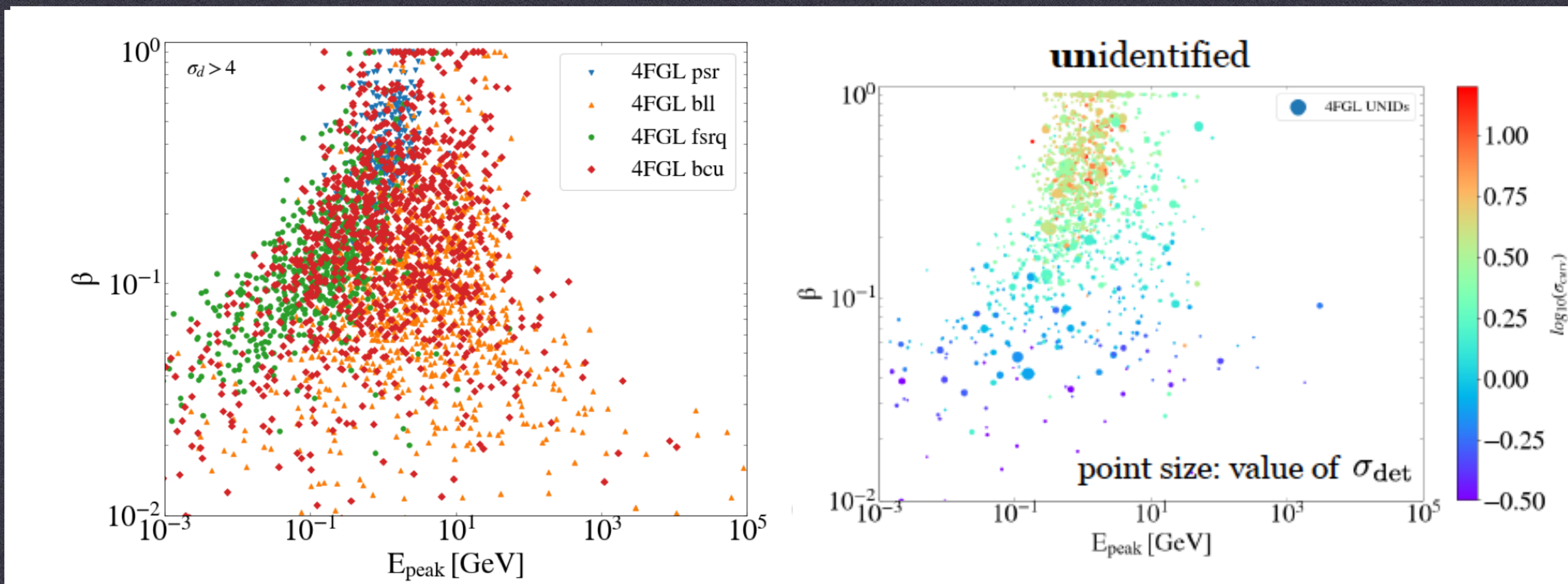
FERMI-LAT GAMMA-RAY DATA & BETA-PLOT

4FGL catalogue:
TOT ASTRO (PSR, QSR, BCU)
TOT UNIDS



Log-Parabola: $$\frac{dN}{dE} = N_0 \left(\frac{E}{E_0}\right)^{-\alpha-\beta\cdot\log(E/E_0)} , E_{peak} = E_0 \cdot e^{\frac{2-\alpha}{2\beta}}$$

FERMI-LAT GAMMA-RAY DATA & BETA-PLOT



$$\frac{dN}{dE} = N_0 \left(\frac{E}{E_0}\right)^{-\alpha-\beta\cdot\log(E/E_0)} , \qquad E_{peak} = E_0 \cdot e^{\frac{2-\alpha}{2\beta}}$$

PREVIOUS WORKS

3FGLzoo. Classifying 3FGL Unassociated Fermi-LAT Gamma-ray Sources by Artificial Neural Networks

D. Salvetti[1⋆], G. Chiaro[1,2†], G. La Mura[2], and D. J. Thompson[3]

Artificial Neural Network Classification of 4FGL Sources

S. Germani,[1⋆] G. Tosti,[1] P. Lubrano,[2] S. Cutini,[2] I. Mereu,[2] A. Berretta[1]
[1]*Dipartimento di Fisica e Geologia, Univ. degli Studi di Perugia, Via A. Pascoli snc, I-06123 Perugia, Italy*
[2]*INFN – Istituto Nazionale di Fisica Nucleare Sez. Perugia, Via A. Pascoli snc, I-06123 Perugia, Italy*

Searches for Pulsar-like Candidates from Unidentified Objects in the Third Catalog of Hard *Fermi*-LAT (3FHL) sources with Machine Learning Techniques

C. Y. Hui,[1⋆] Jongsu Lee,[2] K.L. Li,[1,3,4] Sangin Kim,[2] Kwangmin Oh,[2] Shengda Luo,[5] Alex P. Leung,[5] A. K. H. Kong,[4] J. Takata,[6] K. S. Cheng[7]

Machine learning application to Fermi-LAT data: sharpening all-sky map and emphasizing variable sources

Shogo Sato, Jun Kataoka, Soichiro Ito, Jun'ichi Kotoku, Masato Taki, Asuka Oyama, Takaya Toyoda, Yuki Nakamura, Marino Yamamoto

PREVIOUS WORKS

Spectral and spatial analysis of the dark matter subhalo candidates among *Fermi* Large Area Telescope unidentified sources

Javier Coronado-Blázquez[a,b] Miguel A. Sánchez-Conde[a,b] Mattia Di Mauro[c,d] Alejandra Aguirre-Santaella[a,b] Ioana Ciucă[e] Alberto Domínguez[f] Daisuke Kawata[e] Néstor Mirabal[c,g]

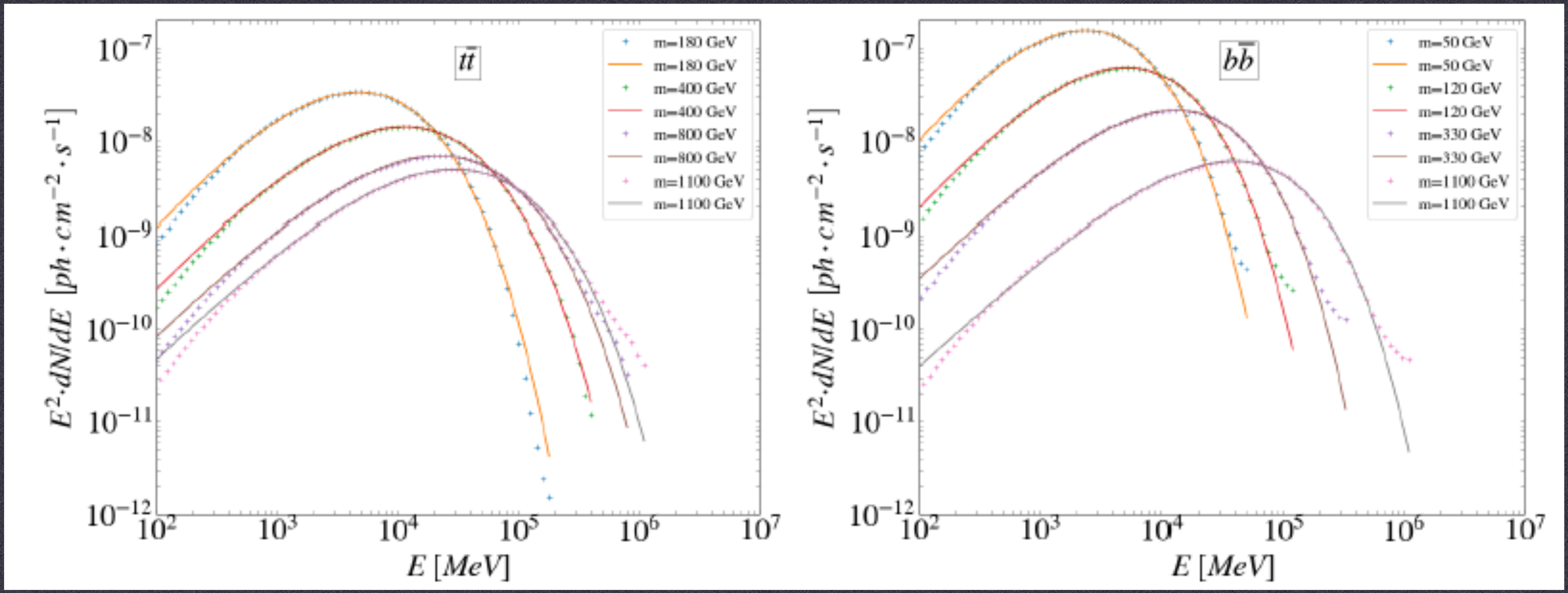
Unidentified Gamma-ray Sources as Targets for Indirect Dark Matter Detection with the Fermi-Large Area Telescope

Javier Coronado-Blazquez, Miguel A. Sanchez-Conde, Alberto Dominguez, Alejandra Aguirre-Santaella, Mattia Di Mauro, Nestor Mirabal, Daniel Nieto, Eric Charles

OUTLINE

- FERMI-LAT GAMMA-RAY DATA & BETA-PLOT
- DARK MATTER & BETA-PLOT
- "SYNTHETIC" FEATURES:
  DETECTION SIGNIFICANCE σ_d AND UNCERTAINTY ON β
- INTRODUCTION TO CLASSIFICATION IN MACHINE LEARNING
- PRELIMINARY RESULTS
- PRELIMINARY CONCLUSIONS

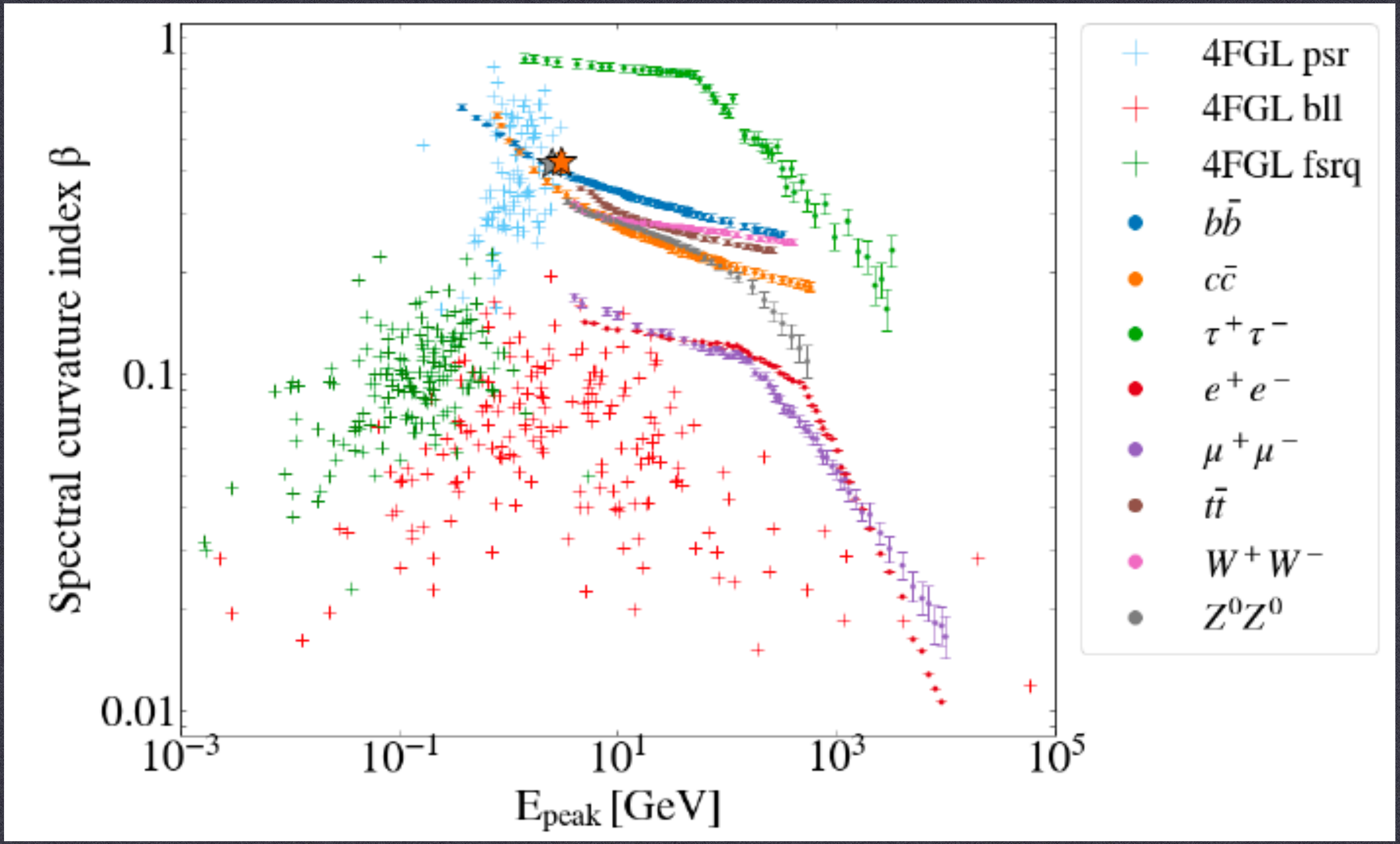
DARK MATTER & BETA-PLOT



J.Coronado-Blazquez et al. JCAP07(2019)020

$$\frac{dN}{dE} = N_0 \left(\frac{E}{E_0}\right)^{-\alpha-\beta\cdot\log(E/E_0)}, \qquad E_{peak} = E_0 \cdot e^{\frac{2-\alpha}{2\beta}}$$

DARK MATTER & BETA-PLOT

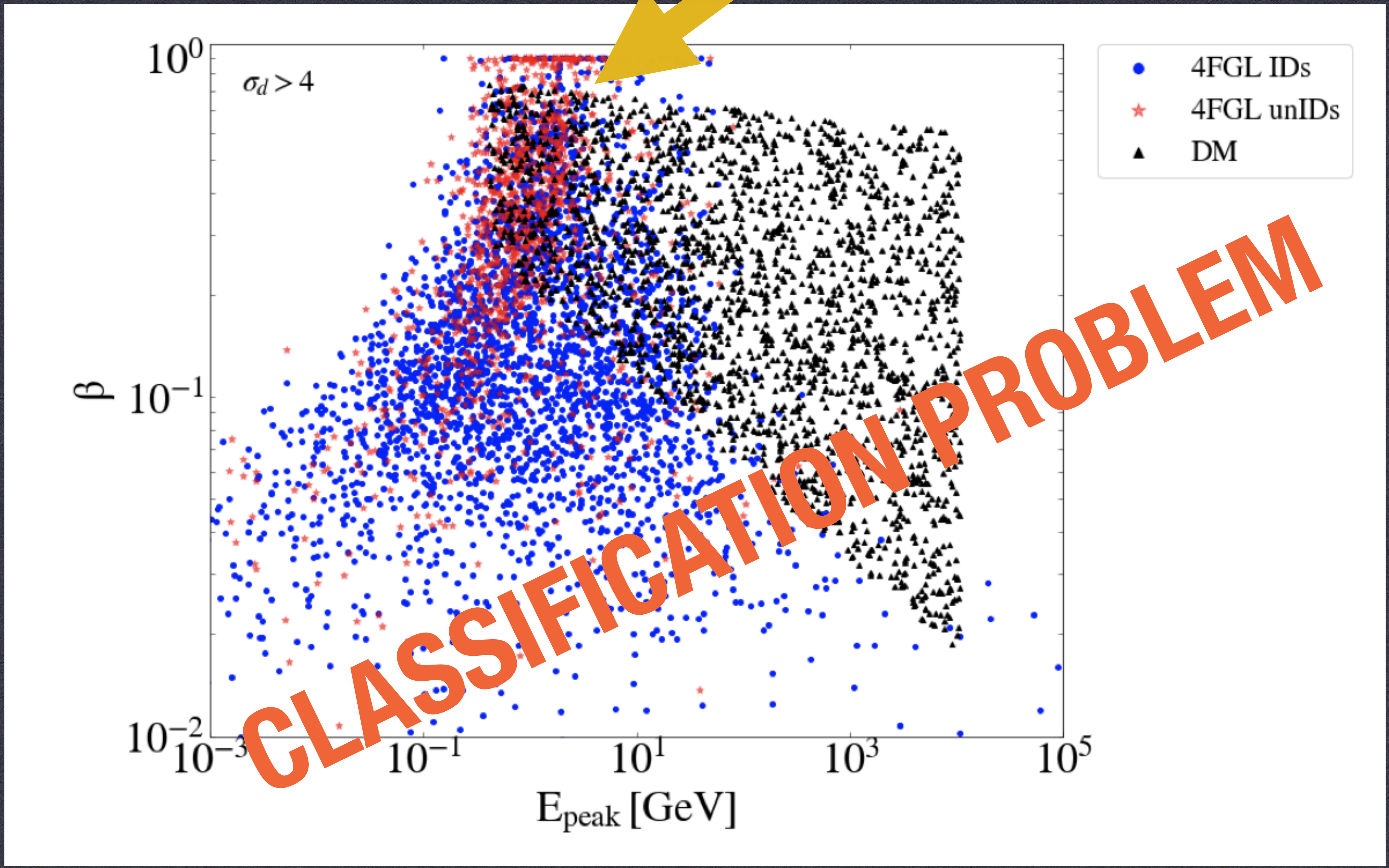


J. Coronado-Blázquez et al., JCAP11(2019)045

DARK MATTER & BETA-PLOT

$$\frac{dN}{dE} = B_r \left(\frac{dN}{dE}\right)_{C_1} + (1 - B_r)\left(\frac{dN}{dE}\right)_{C_2}$$

Degeneracy of pulsar and DM signal



DARK MATTER & BETA-PLOT

Our strategy:

1. The classification algorithm is trained on a sample of Astrophysical (Astro) and Dark Matter (DM) sources. The classification accuracy is tested on a subsample of data;

2. The "machine" has learned the classification problem and it is applied to the unIDs dataset: we expect the algorithm telling us if any unIDS could be a DM source with a given probability.

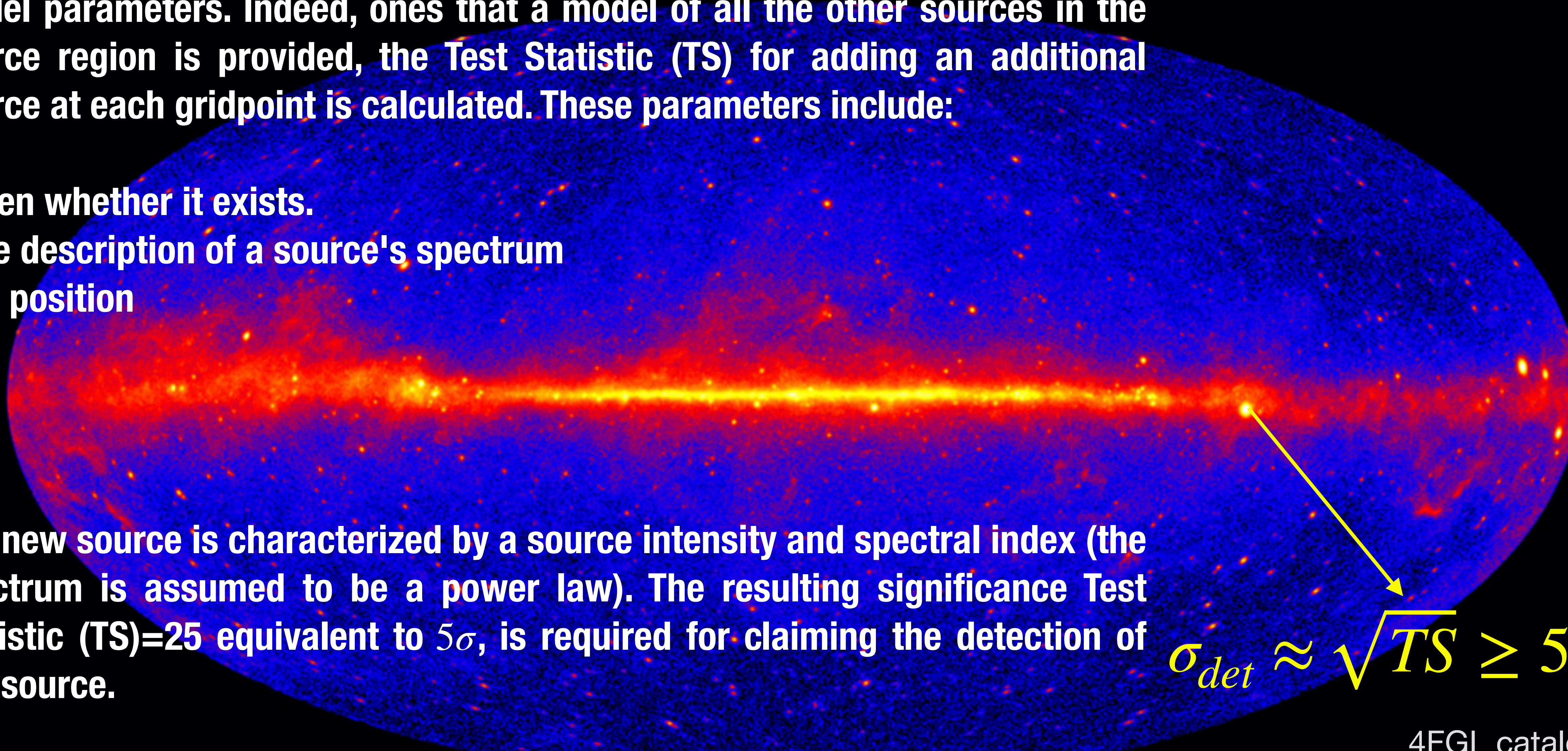
OUTLINE

- FERMI-LAT GAMMA-RAY DATA & BETA-PLOT
- DARK MATTER & BETA-PLOT
- “SYNTHETIC” FEATURES:
  DETECTION SIGNIFICANCE σ_d AND UNCERTAINTY ON β
- INTRODUCTION TO CLASSIFICATION IN MACHINE LEARNING
- PRELIMINARY RESULTS
- PRELIMINARY CONCLUSIONS

DETECTION SIGNIFICANCE

To analyze LAT data, the collaboration tools construct the likelihood that is applicable to the LAT data, and then use this likelihood to find the best fit model parameters. Indeed, ones that a model of all the other sources in the source region is provided, the Test Statistic (TS) for adding an additional source at each gridpoint is calculated. These parameters include:

- even whether it exists.
- the description of a source's spectrum
- its position



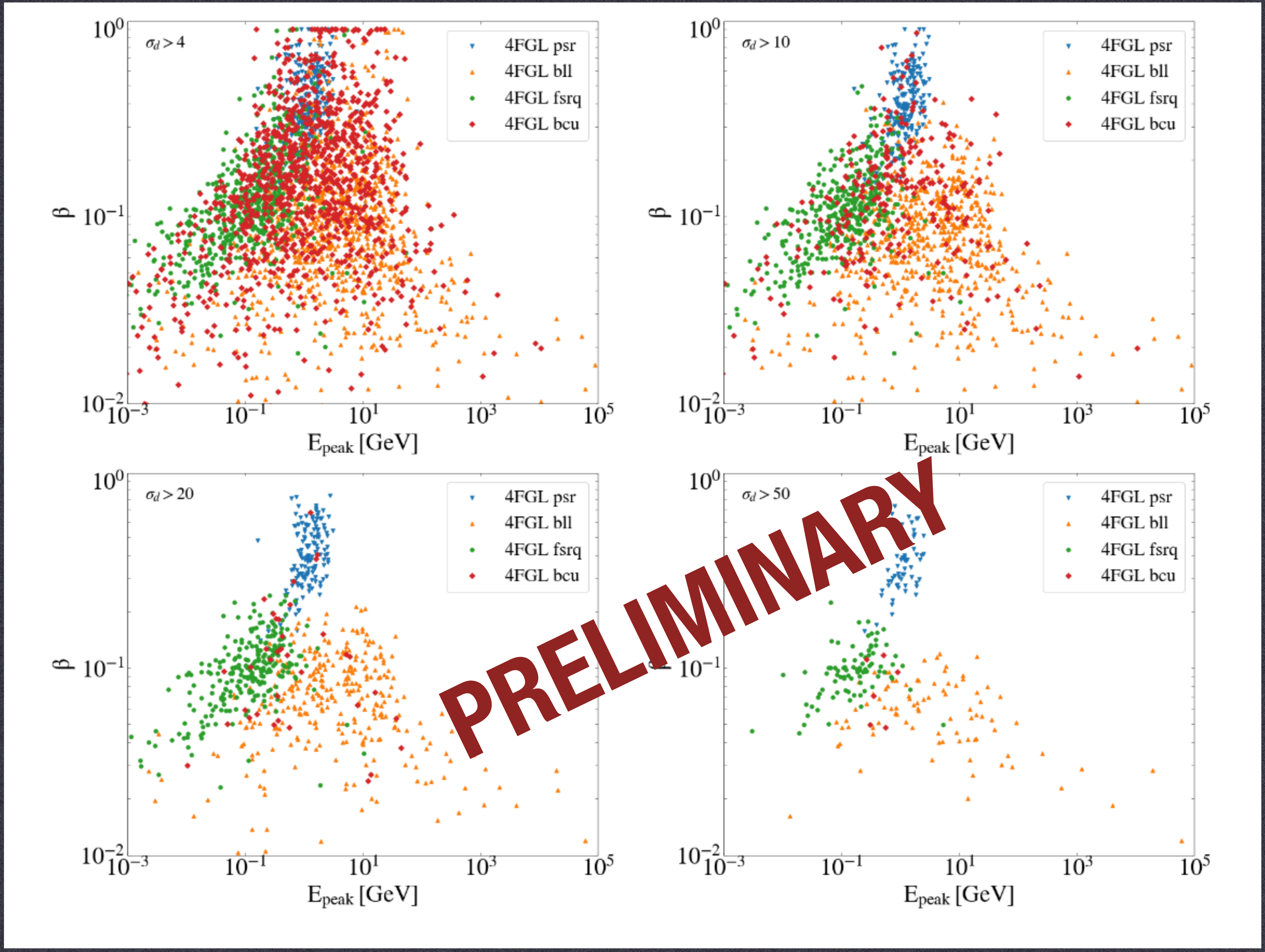
The new source is characterized by a source intensity and spectral index (the spectrum is assumed to be a power law). The resulting significance Test Statistic (TS)=25 equivalent to 5σ, is required for claiming the detection of any source.

$$\sigma_{det} \approx \sqrt{TS} \geq 5$$

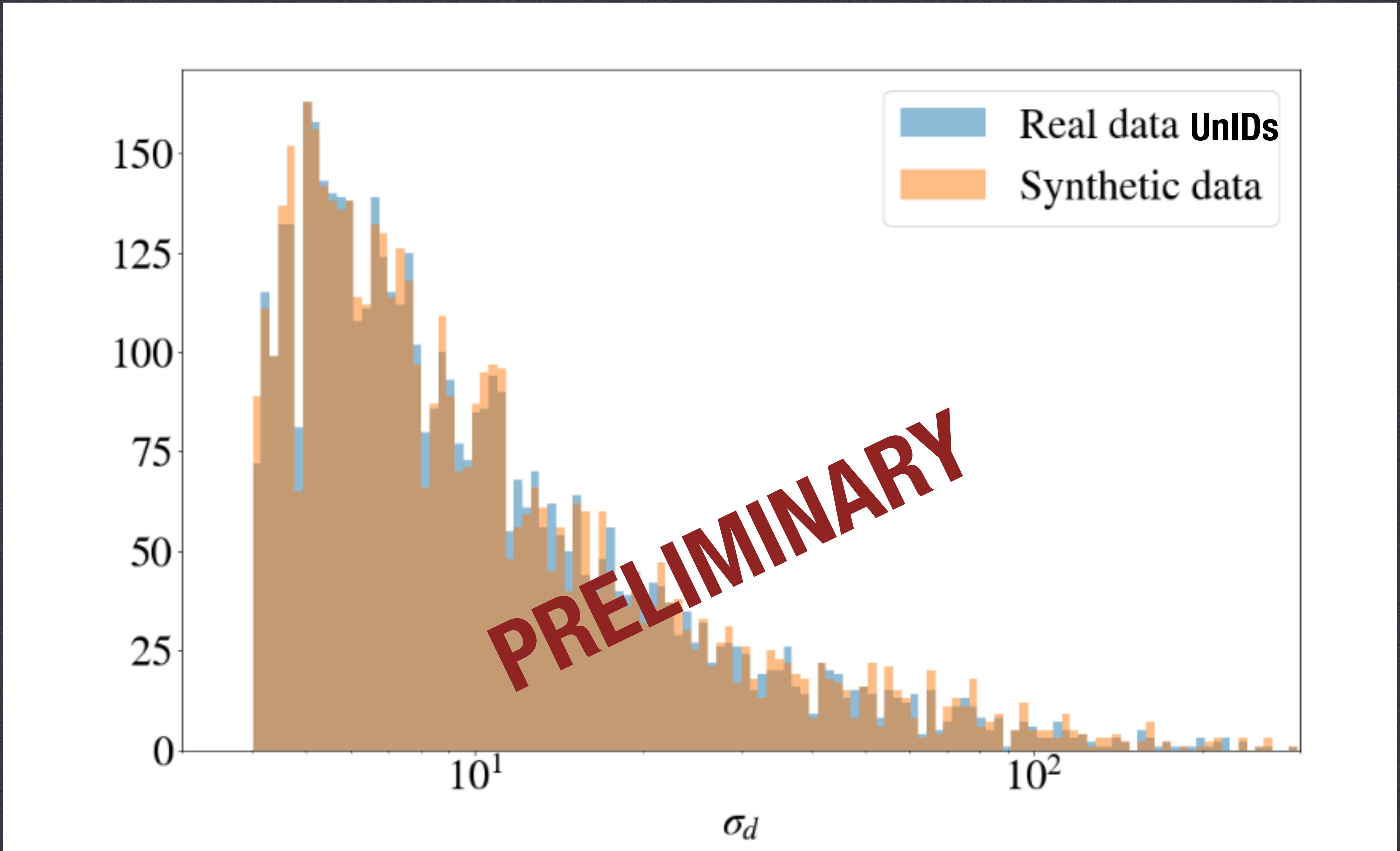
Hereafter, we will use the so-defined detection significance σ_d as a feature of our classification problem.

4FGL catalogue:
TOT ASTRO (PSR, QSR, BCU)
TOT UNIDS

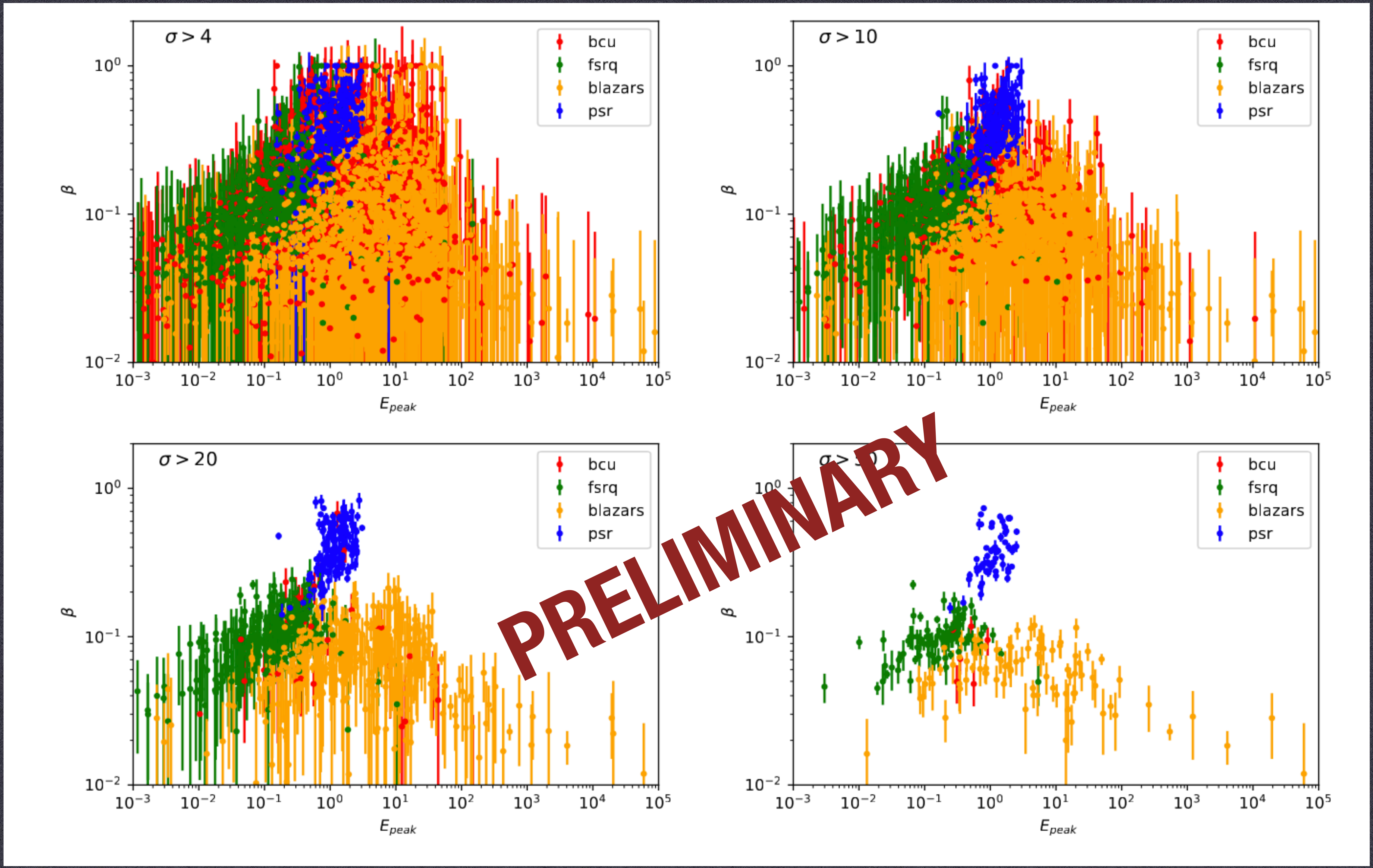
DETECTION SIGNIFICANCE



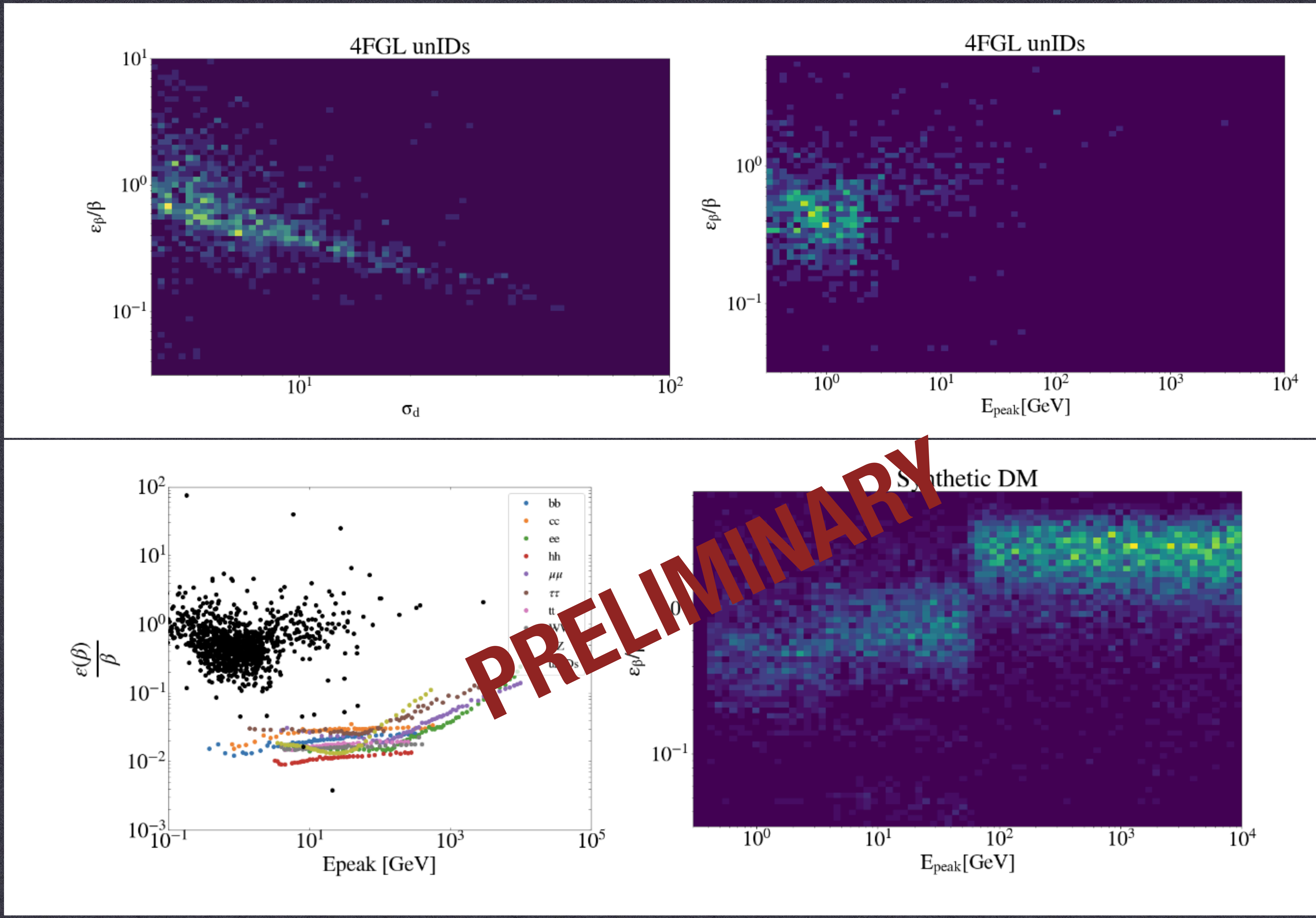
DETECTION SIGNIFICANCE



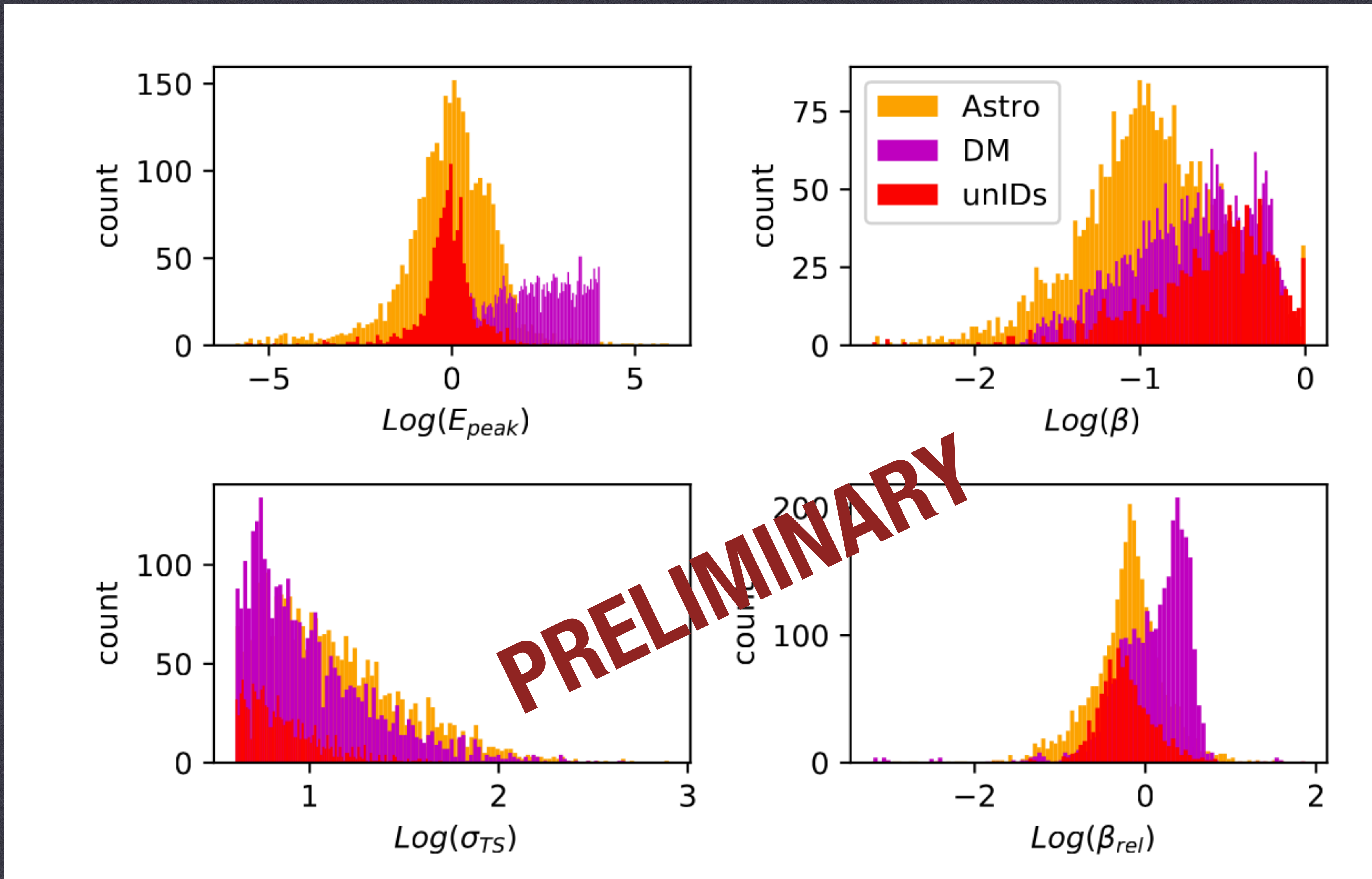
UNCERTAINTY ON β



UNCERTAINTY ON β

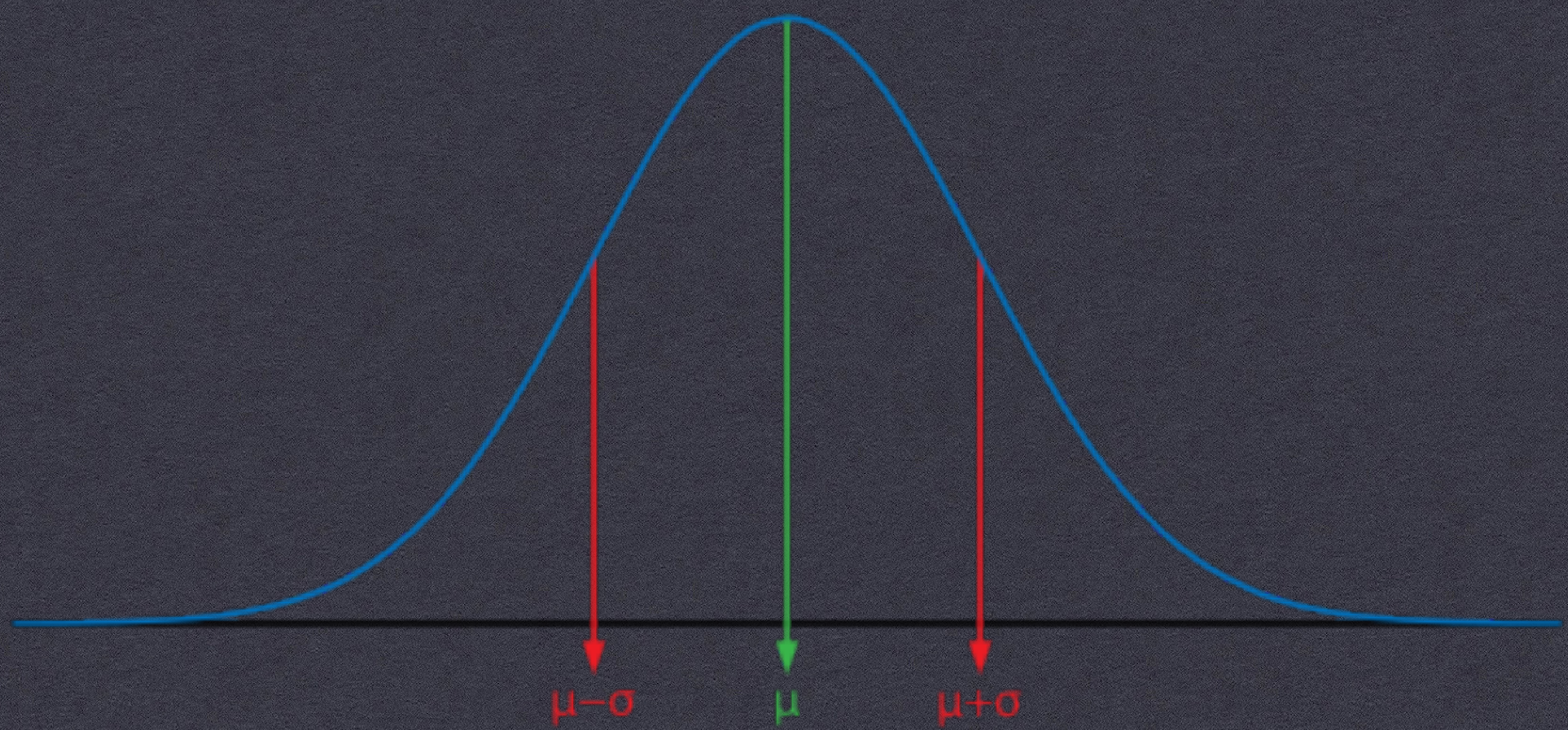


4 FEATURES DISTRIBUTIONS



GAUSSIAN SAMPLING OF β UNCERTAINTY

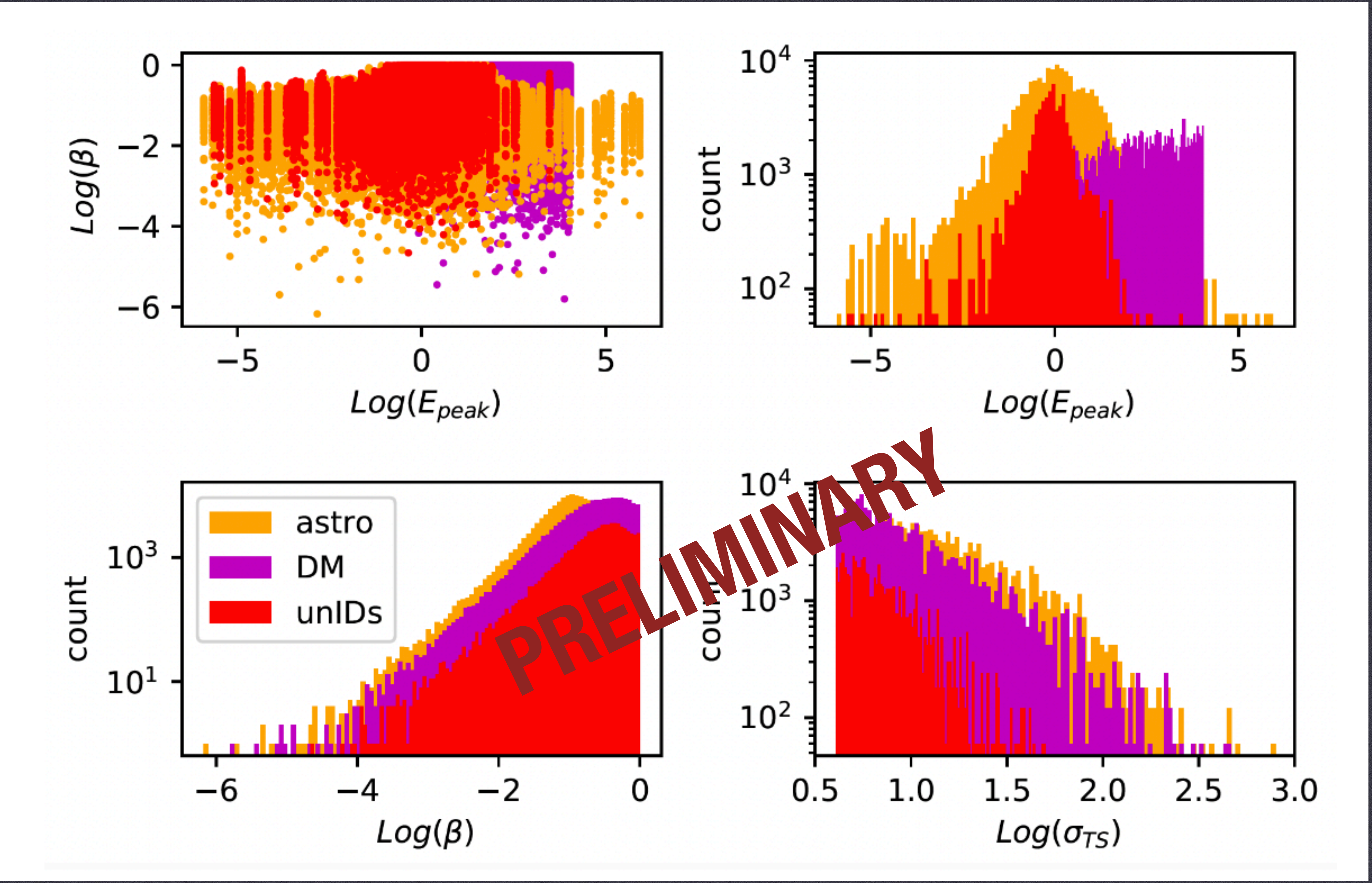
$M = 60$



$\beta - \epsilon_\beta$ β $\beta + \epsilon_\beta$

$0 < \beta \leq 1$ **Is required if β is small and ϵ_β is big**

GAUSSIAN SAMPLING OF β UNCERTAINTY



GAUSSIAN SAMPLING OF β UNCERTAINTY

Related issues:

- **Increasing the number of data** from N (Astro+DM datasets) to MxN makes the learning process slower;

- After the learning step and in order to classify the unIDs, the method would also require the **sample of the unIDs uncertainty**, that is useless for the classification intent itself.

OUTLINE

- FERMI-LAT GAMMA-RAY DATA & BETA-PLOT
- DARK MATTER & BETA-PLOT
- “SYNTHETIC” FEATURES:
  DETECTION SIGNIFICANCE σ_d AND UNCERTAINTY ON β
- <u>INTRODUCTION TO CLASSIFICATION IN MACHINE LEARNING</u>
- PRELIMINARY RESULTS
- PRELIMINARY CONCLUSIONS

CLASSIFICATION ALGORITHMS

- LOGISTIC REGRESSION (LR) (SCIKITS-LEARN)
- ARTIFICIAL NEURAL NETWORK (NN) (SCIKITS-LEARN)
- NAIVES BAYES (NB) (TENSOR FLOW)
- GAUSSIAN PROCESS (GP) (TENSOR FLOW)

CLASSIFICATION ALGORITHMS

- **LOGISTIC REGRESSION (LR):** PROBABILISTIC DISCRIMINATIVE MODEL. DESPITE ITS NAME, IS A CLASSIFICATION MODEL RATHER THAN REGRESSION MODEL.
- **NEURAL NETWORK (NN):** PROBABILISTIC DISCRIMINATIVE MODEL. ARE A NON-LINEAR STATISTICAL DATA MODELING TOOL COMPOSED OF HIGHLY INTERCONNECTED NODES THAT CAN MODEL COMPLEX RELATIONSHIPS BETWEEN INPUTS AND OUTPUTS.
- **NAIVE BAYES (NB):** GENERATIVE MODEL. A PROBABILISTIC CLASSIFIER BASED ON BAYES' THEOREM, WHICH ASSUMES THAT EACH FEATURE MAKES AN INDEPENDENT AND EQUAL CONTRIBUTION TO THE TARGET CLASS.
- **GAUSSIAN PROCESS (GP):** NON-PARAMETRIC MODEL. IT IS A STOCHASTIC PROCESS, I.E. A COLLECTION OF RANDOM VARIABLES, SUCH THAT EVERY FINITE LINEAR COMBINATION OF THEM IS NORMALLY DISTRIBUTED. THE DISTRIBUTION OF A GP IS THE JOINT DISTRIBUTION OF ALL THOSE RANDOM VARIABLES.

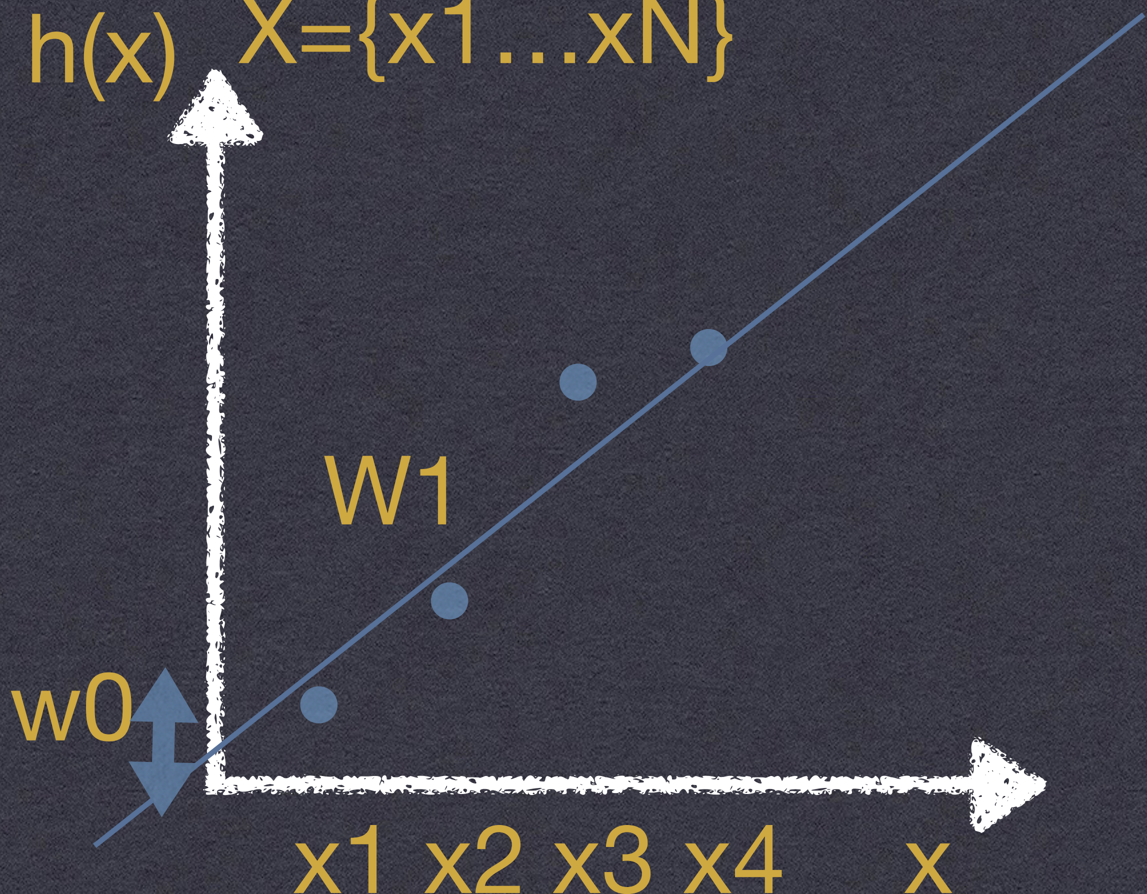
CLASSIFICATION ALGORITHMS

LINEAR REGRESSION

1-FEATURE

h(x)=w0+w1x

X={x1…xN}



P-features

h(x)=w0+w1x…+wpx=WTX

[X]=NxP N data-points
P features

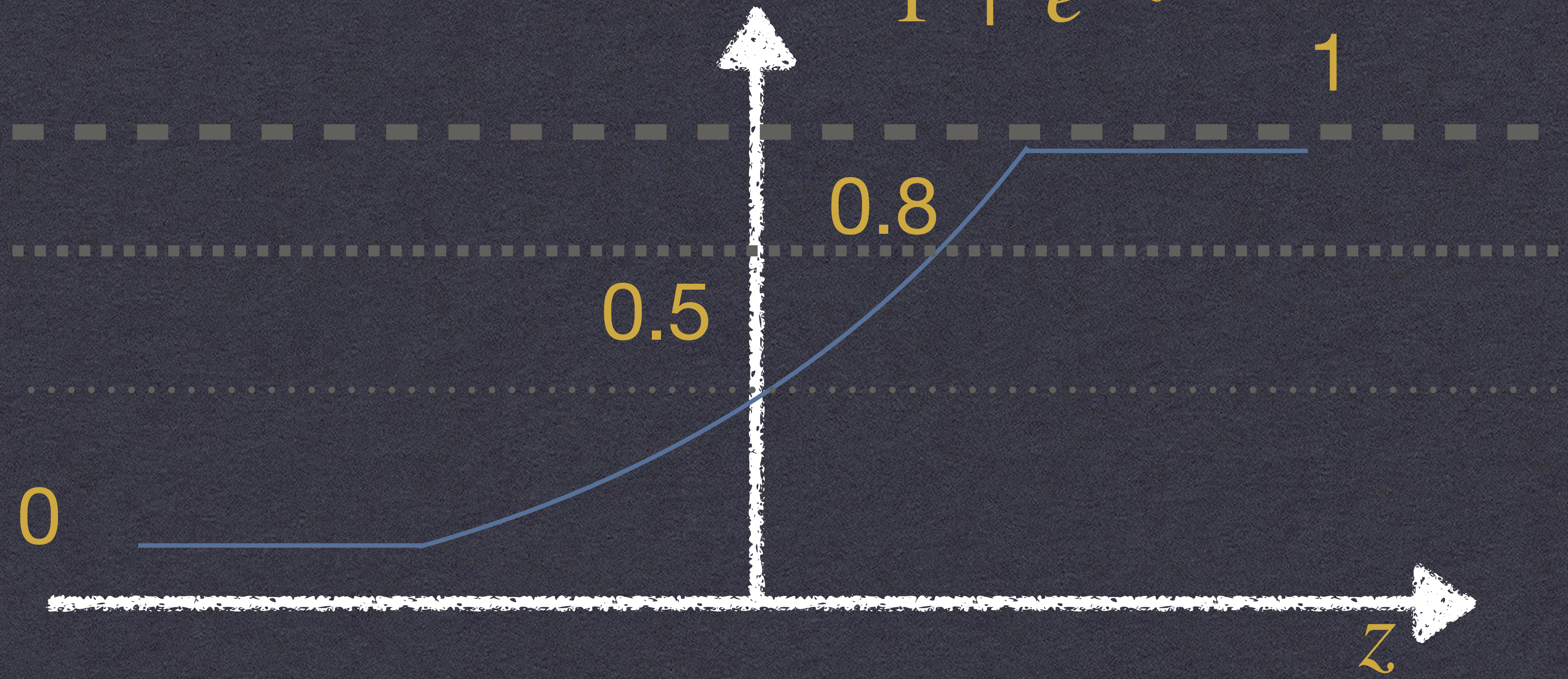
W=(w0…wp)
coefficients of the fits

$$J(\mathbf{W}) = \frac{1}{2}(h(x) - \mathbf{Y})^2 = \frac{1}{2}\sum_{i=1}^{N}((\mathbf{W}^T\mathbf{X})_i - Y_i^2)$$

CLASSIFICATION ALGORITHMS

LOGISTIC REGRESSION

Activation function: $h(x) = g(z) = \dfrac{1}{1+e^{-z}}$



1-F

$g(z) \leq 0.5 \rightarrow Y = 0$ (e.g. Astro) $g(z) > 0.5 \rightarrow Y = 1$ (e.g. DM)

LINEAR REGRESSION

$$\mathrm{J}(\mathbf{W}) = \frac{1}{2}(h(x) - \mathbf{Y})^2 = \frac{1}{2}\sum_{i=1}^{N}((\mathbf{W}^T\mathbf{X})_i - Y_i^2)$$

LOGISTIC REGRESSION

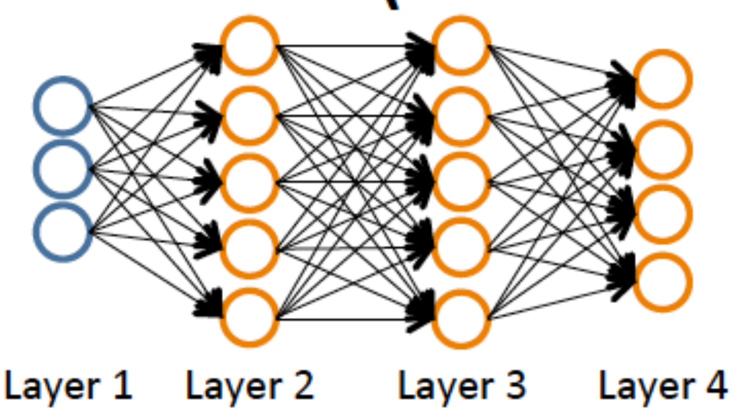
$$J(W) = -\frac{1}{N}[\sum_{i=1}^{N} y^{(i)} \log(h_w(x^{(i)})) + (1 - y^{(i)}) \log(1 - h_w(x^i))))]$$

CLASSIFICATION ALGORITHMS

ARTIFICIAL NEURAL NETWORK

Neural Network (Classification)

$$\{(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \ldots, (x^{(m)}, y^{(m)})\}$$



$L =$ total no. of layers in network

$s_l =$ no. of units (not counting bias unit) in layer l

Binary classification

$y = 0 \text{ or } 1$

1 output unit

Multi-class classification (K classes)

$y \in \mathbb{R}^K$ E.g. $\begin{bmatrix}1\\0\\0\\0\end{bmatrix}$, $\begin{bmatrix}0\\1\\0\\0\end{bmatrix}$, $\begin{bmatrix}0\\0\\1\\0\end{bmatrix}$, $\begin{bmatrix}0\\0\\0\\1\end{bmatrix}$

pedestrian car motorcycle truck

K output units

CLASSIFICATION ALGORITHMS

ARTIFICIAL NEURAL NETWORK

Cost function

Logistic regression:

$$J(\theta) = -\frac{1}{m}\left[\sum_{i=1}^{m} y^{(i)} \log h_\theta(x^{(i)}) + (1 - y^{(i)}) \log(1 - h_\theta(x^{(i)}))\right] + \frac{\lambda}{2m}\sum_{j=1}^{n}\theta_j^2$$

Neural network:

$$h_\Theta(x) \in \mathbb{R}^K \quad (h_\Theta(x))_i = i^{th} \text{ output}$$

$$J(\Theta) = -\frac{1}{m}\left[\sum_{i=1}^{m}\sum_{k=1}^{K} y_k^{(i)} \log(h_\Theta(x^{(i)}))_k + (1 - y_k^{(i)}) \log(1 - (h_\Theta(x^{(i)}))_k)\right]$$

$$+ \frac{\lambda}{2m}\sum_{l=1}^{L-1}\sum_{i=1}^{s_l}\sum_{j=1}^{s_{l+1}}(\Theta_{ji}^{(l)})^2$$

CLASSIFICATION ALGORITHMS NAIVE BAYES

the appropriate variables. We can now interpret $p(\mathcal{C}_k)$ as the prior probability for the class $\mathcal{C}_k$, and $p(\mathcal{C}_k|\mathbf{x})$ as the corresponding posterior probability. Thus $p(\mathcal{C}_1)$ represents the probability that a person has cancer, before we take the X-ray measurement. Similarly, $p(\mathcal{C}_1|\mathbf{x})$ is the corresponding probability, revised using Bayes' theorem in light of the information contained in the X-ray. If our aim is to minimize the chance of assigning $\mathbf{x}$ to the wrong class, then intuitively we would choose the class having the higher posterior probability. We now show that this intuition is correct, and we also discuss more general criteria for making decisions.

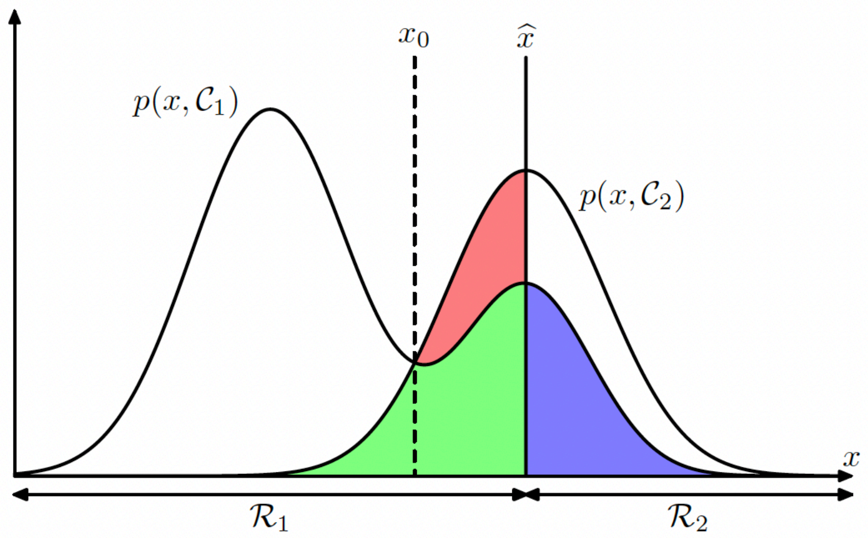


Figure 1.24 Schematic illustration of the joint probabilities $p(x, \mathcal{C}_k)$ for each of two classes plotted against x, together with the decision boundary $x = \widehat{x}$. Values of $x \geqslant \widehat{x}$ are classified as class $\mathcal{C}_2$ and hence belong to decision region $\mathcal{R}_2$, whereas points $x < \widehat{x}$ are classified as $\mathcal{C}_1$ and belong to $\mathcal{R}_1$. Errors arise from the blue, green, and red regions, so that for $x < \widehat{x}$ the errors are due to points from class $\mathcal{C}_2$ being misclassified as $\mathcal{C}_1$ (represented by the sum of the red and green regions), and conversely for points in the region $x \geqslant \widehat{x}$ the errors are due to points from class $\mathcal{C}_1$ being misclassified as $\mathcal{C}_2$ (represented by the blue region). As we vary the location $\widehat{x}$ of the decision boundary, the combined areas of the blue and green regions remains constant, whereas the size of the red region varies. The optimal choice for $\widehat{x}$ is where the curves for $p(x, \mathcal{C}_1)$ and $p(x, \mathcal{C}_2)$ cross, corresponding to $\widehat{x} = x_0$, because in this case the red region disappears. This is equivalent to the minimum misclassification rate decision rule, which assigns each value of x to the class having the higher posterior probability $p(\mathcal{C}_k|x)$.

Christopher M. Bishop, Springer 2006.

CLASSIFICATION ALGORITHMS
NAIVE BAYES

$$p(\mathbf{x}_\mathrm{I}, \mathbf{x}_\mathrm{B}|\mathcal{C}_k) = p(\mathbf{x}_\mathrm{I}|\mathcal{C}_k)p(\mathbf{x}_\mathrm{B}|\mathcal{C}_k).$$

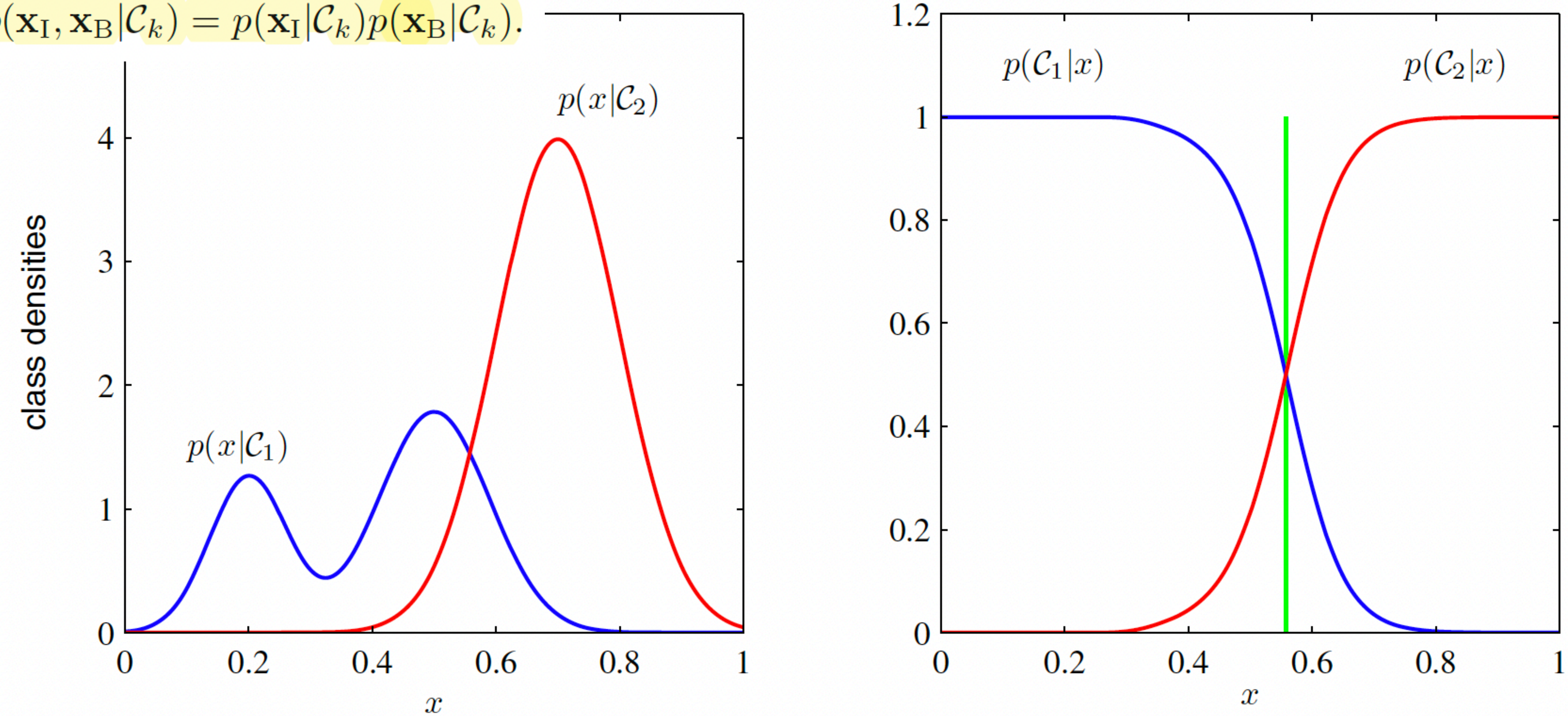


Figure 1.27 Example of the class-conditional densities for two classes having a single input variable x (left plot) together with the corresponding posterior probabilities (right plot). Note that the left-hand mode of the class-conditional density $p(\mathbf{x}|\mathcal{C}_1)$, shown in blue on the left plot, has no effect on the posterior probabilities. The vertical green line in the right plot shows the decision boundary in x that gives the minimum misclassification rate.

Christopher M. Bishop, Springer 2006.

CLASSIFICATION ALGORITHMS
GAUSSIAN PROCESS

Multi-class Gaussian Process Classification with Noisy Inputs

Autor (es): Villacampa-Calvo, Carlos; Zaldívar, Bryan; Garrido-Merchán, Eduardo C.; Hernández Lobato, Daniel

Entidad: UAM. Departamento de Ingeniería Informática

Editor: Microtome Publishing

Fecha de edición: 2021-01

Cita: Journal Of Machine Learning Research 22.36 (2021): 1−52

ISSN: 1532-4435 (print); 1533-7928 (online)

SETUPS

- **2-FEATURES (2F) CLASSIFICATION (LR, NN, NB):** INCLUDES THE 2-FEATURES INTRODUCED SO FAR, INDEED (E_{peak}, β)

- **4-FEATURES (4F) CLASSIFICATION (LR, NN, NB):** INCLUDES THE SYSTEMATICS UNCERTAINTY, BY INCLUDING TWO MORE FEATURES, THAT ARE: $(E_{\text{peak}}, \beta, \sigma_d, \beta_{\text{rel}})$ WHERE $\beta_{\text{rel}} = \epsilon_\beta / \beta$

- **3-FEATURES AUGMENTED (3F-A) (LR, NN, NB):** AN AUGMENTED DATASET CONTAINING THREE FEATURES: $(E_{\text{peak}}, \beta, \sigma_d)$ INSTEAD OF INCORPORATING THE UNCERTAINTY β_{rel} AS AN EXTRA FEATURE, THE STRATEGY HERE IS TO AUGMENT THE DATASET BY THE FOLLOWING PROCEDURE: FOR EACH OBSERVATION, WE ASSUME THAT THE VARIABLE β FOLLOWS A TRUNCATED GAUSSIAN DISTRIBUTION, WHOSE MEAN IS PRECISELY THE OBSERVED VALUE, AND THE STANDARD DEVIATION IS PRECISELY THE OBSERVED UNCERTAINTY ϵ_β, BUT TRUNCATED SUCH THAT $0 < \beta \leq 1$.

- **3F-B (GP):** A DATASET CONTAINING THE THREE SAME FEATURES AS ABOVE, I.E. $(E_{\text{peak}}, \beta, \sigma_d)$. HOWEVER, NOW THE UNCERTAINTIES ϵ_β ARE INCLUDED IN THE STATISTICAL MODEL. CONCRETELY, THIS SETUP WILL CONCERN EXCLUSIVELY THE NIMGP MODEL MENTIONED ABOVE.

OUTLINE

- FERMI-LAT GAMMA-RAY DATA & BETA-PLOT
- DARK MATTER & BETA-PLOT
- “SYNTHETIC” FEATURES:
  DETECTION SIGNIFICANCE σ_d AND UNCERTAINTY ON β
- INTRODUCTION TO CLASSIFICATION IN MACHINE LEARNING
- <u>PRELIMINARY RESULTS</u>
- PRELIMINARY CONCLUSIONS

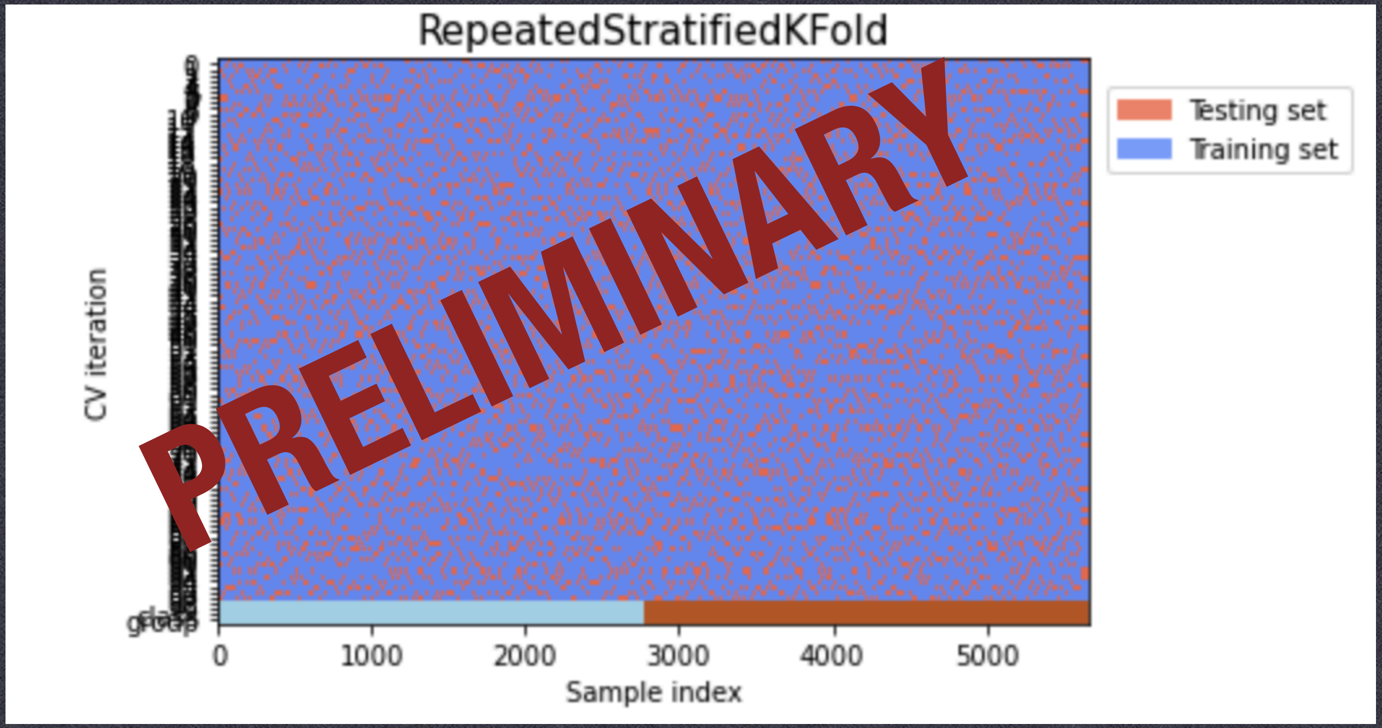
DATA PRE-PROCESSING

1. 10^(-3)GeV < E_peak < 10^6 GeV , reliable range of the Fermi-LAT sensitivity in energy
2. Balanced data: same number of DM and Astro
3. Log scale classification
4. Standardised data: each feature is normalised with respect to their medium values.
5. Training/Testing data set split:

RepeatedStratifiedKFold(n_splits=N_splits, n_repeats=N_Repeats)
Number of folds, N_splits=5 -> Train set = 4530 (80%) data Test set=1132 (20%)
Number of times cross-validator needs to be repeated, N_Repeats=20
N_class=N_splits x N_Repeats= 100

Stratified: The split into N_folds preserve the percentage of samples for each class and without repeated data in different folds.

Repeated: the cross-validation is repeated a number of times with different random seed

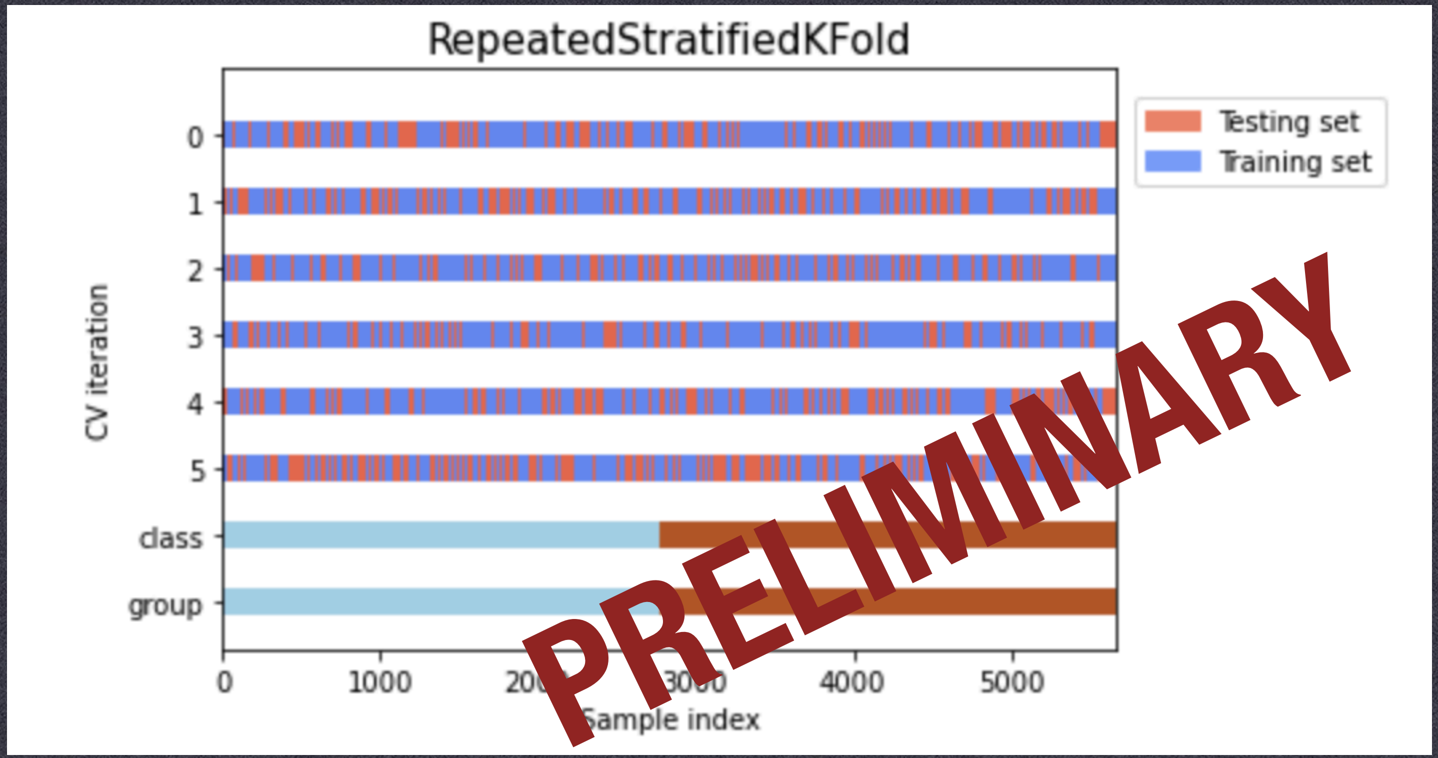


DATA PRE-PROCESSING: CHECK

RepeatedStratifiedKFold(n_splits=N_splits, n_repeats=N_Repeats)
Number of folds, N_splits=3 -> Train set = 3774 (80%) data Test set=1888 (33%)
Number of times cross-validator needs to be repeated, N_Repeats=2
N_class=N_splits x N_Repeats= 6

Stratified: The split into N_folds preserve the percentage of samples for each class and without repeated data in different folds.

Repeated: the cross-validation is repeated a number of times with different random seed



PRELIMINARY CLASSIFICATION RESULTS

$$\text{Overall accuracy (OA)}(y, \hat{y}) = \frac{1}{n_{\text{samples}}} \sum_{i=0}^{n_{\text{samples}}-1} 1(\hat{y}_i = y_i)$$

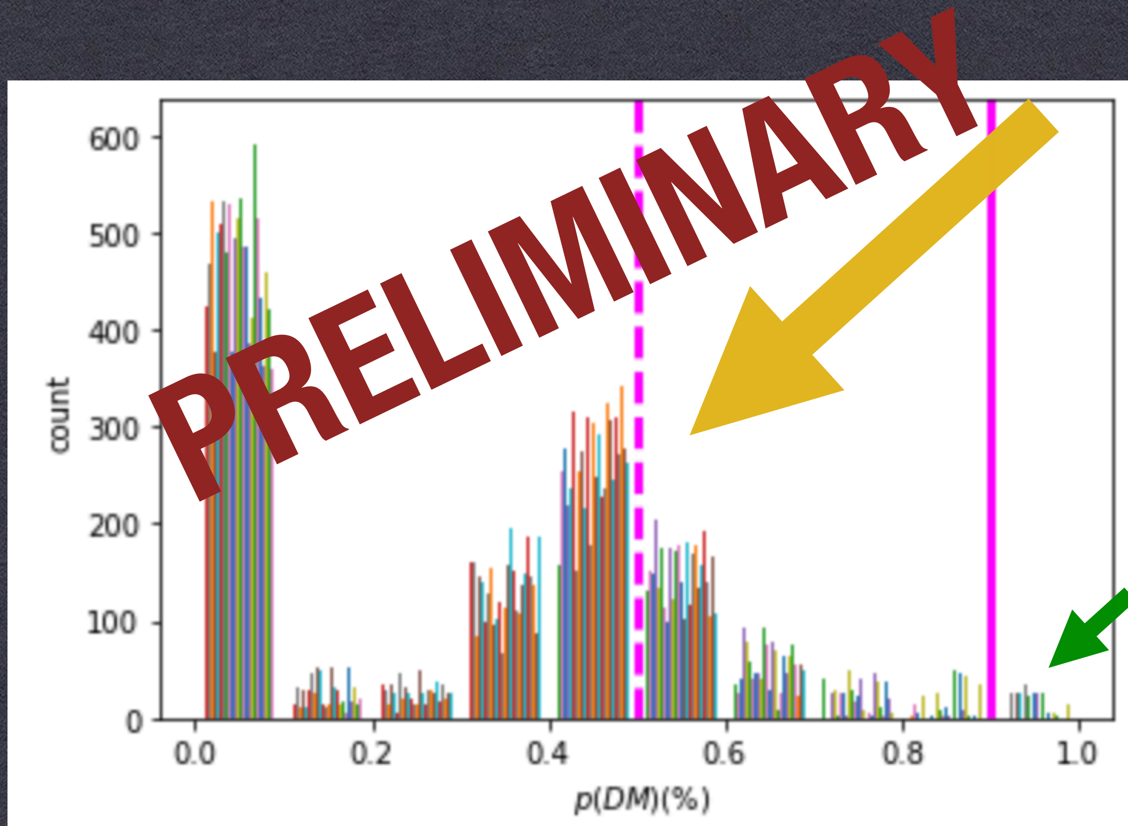
- **TRUE NEGATIVE:** PERCENTAGE OF WELL CLASSIFIED ASTRO SOURCES (NORMALISED TO THE TOTAL NUMBER OF ASTRO SOURCES)
- **TRUE POSITIVE:** PERCENTAGE OF WELL CLASSIFIED DARK MATTER SOURCES (NORMALISED TO THE TOTAL NUMBER OF DM SOURCES)

	OA(%)	TN (%)	TP (%)
LR			
2F	84.9 ± 0.6	85.4 ± 1.3	84.4 ± 1.0
4F	86.0 ± 0.5	86.8 ± 1.2	85.6 ± 0.7
3F-A	82.9 ± 0.1	84.9 ± 0.2	80.9 ± .0.1
NN			
2F	86.8 ± 0.3	86.4 ± 2.4	87.2 ± 2.3
4F	**93.1 ± 0.4**	**94.7 ± 1.1**	**91.4 ± 1.0**
3F-A	85.0 ± 0.1	88.7 ± 0.8	81.3 ± 1.1
NB			
2F	82.0 ± 1.3	80.4 ± 2.7	83.8 ± 2.1
4F	83.7 ± 0.9	81.1 ± 1.9	86.4 ± 0.5
3F-A	82.6 ± 0.1	83.4 ± 0.2	81.3 ± 0.1
GP			
3F-B	87.0±0.1	84.5±0.2	89.4±0.2

PRELIMINARY

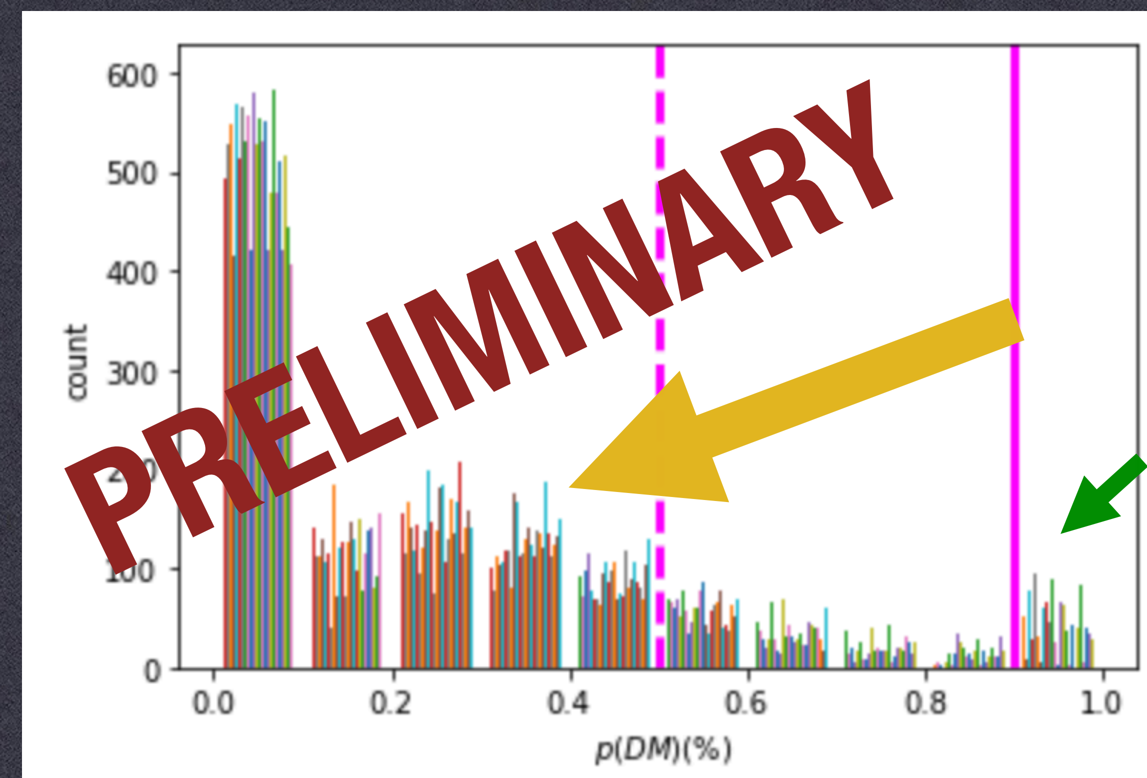
UNIDS CLASSIFICATION WITH NN

- 2-FEATURES (2F) CLASSIFICATION



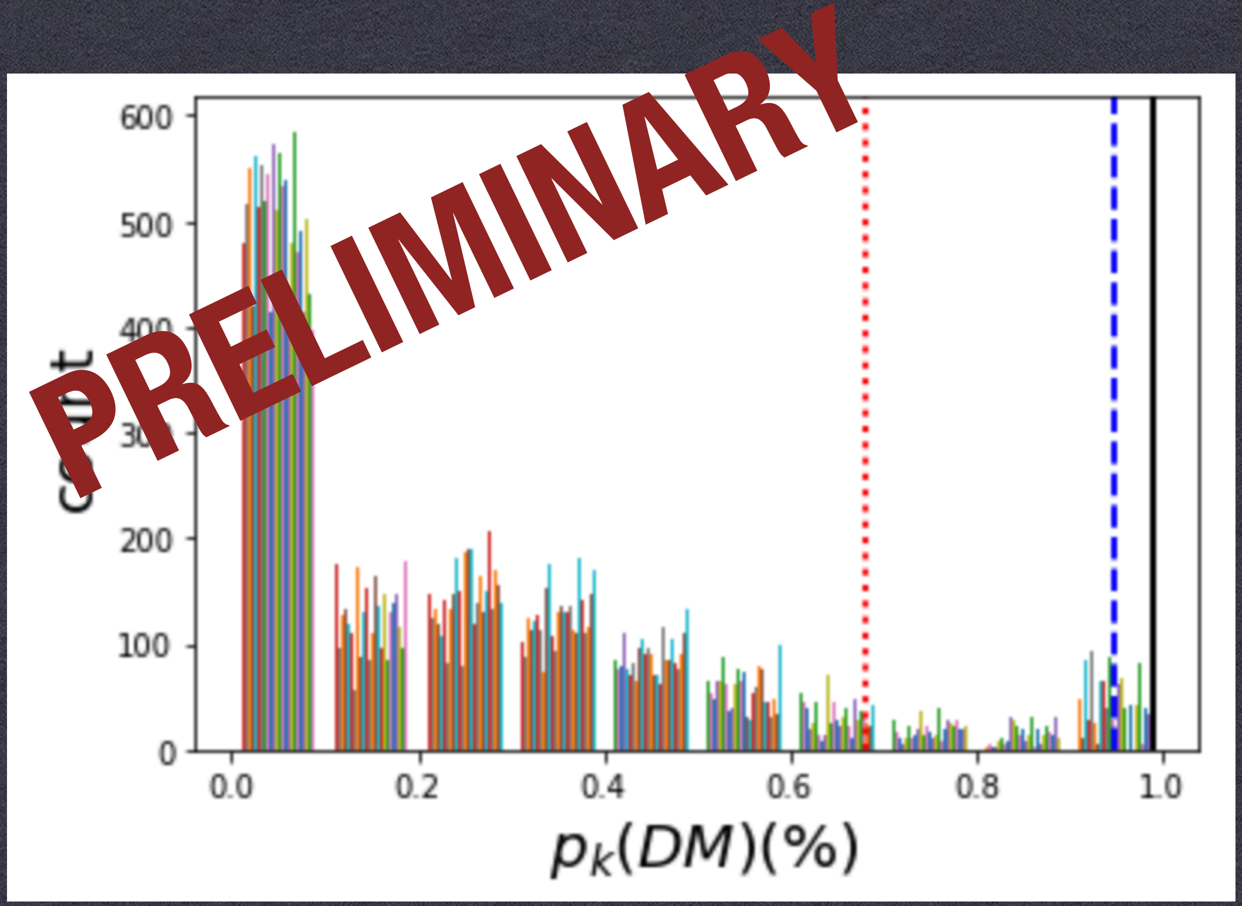
- 6 ± 10 UNIDS CLASSIFIED AS DM WITH $p_k > 90\,\%$ (ERROR DEFINED ON 100 CLASSIFICATION)
- 0 UNIDS WITH $\bar{p} > 90\,\%\,(50\%)(40\%)$

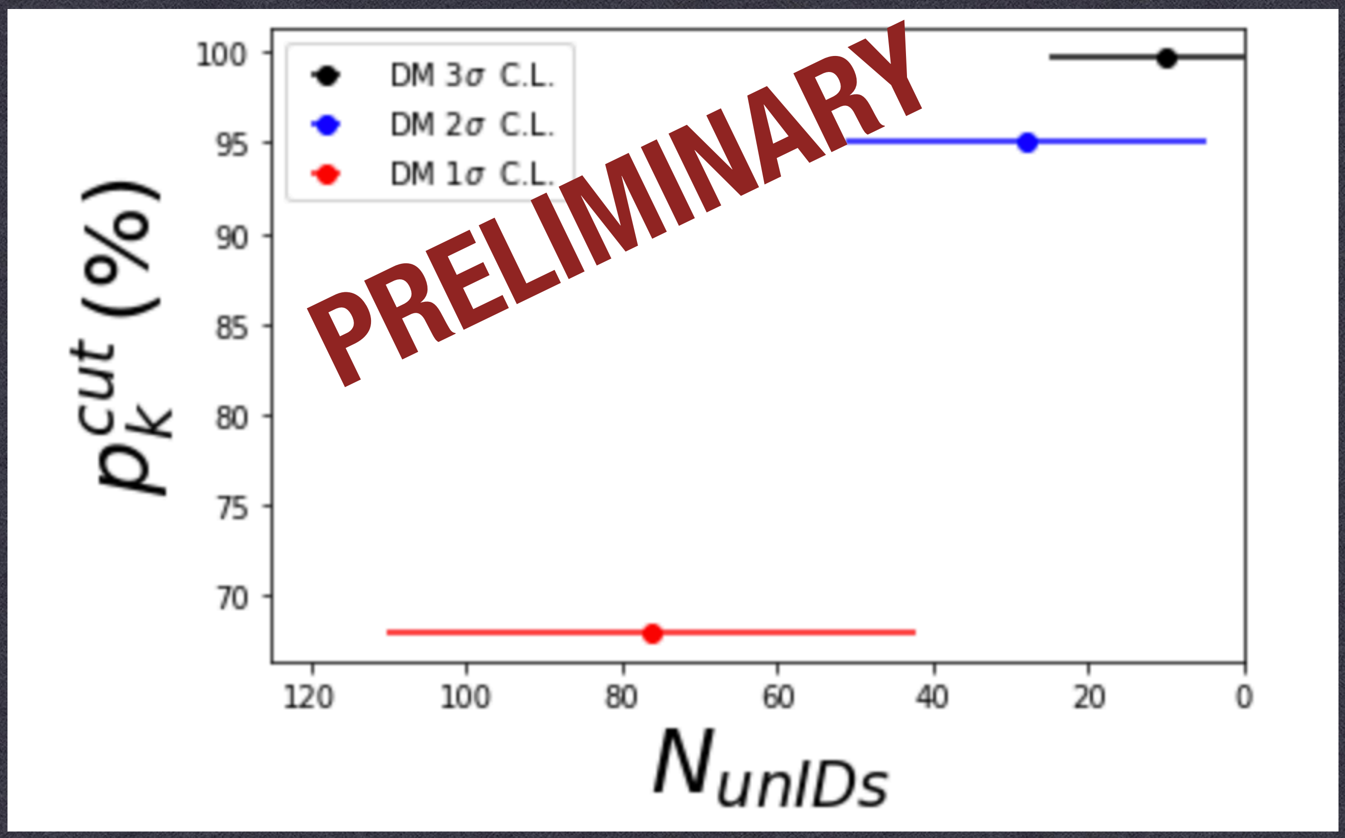
- 4-FEATURES (4F) CLASSIFICATION



- 36 ± 26 UNIDS CLASSIFIED AS DM WITH $p_k > 90\,\%$ (ERROR DEFINED ON 100 CLASSIFICATION)
- 0 UNIDS WITH $\bar{p} > 90\,\%\,(50\%)$
- FEW UNIDS WITH $\bar{p} > 40\,\%$

UNIDS CLASSIFICATION WITH NN





PRELIMINARY CONCLUSIONS

- WE TRAINED FOUR DIFFERENT MACHINES ON A SAMPLE OF BOTH EXPERIMENTAL AND THEORETICAL DATA
- WE INTRODUCED THE SYNTHETIC FEATURES AND FOUR DIFFERENT SET-UPS
- WE PROPOSED A METHODOLOGY TO INCLUDE SYSTEMATIC UNCERTAINTY IN CLASSIFICATION PROBLEMS, IMPROVING THE OVERALL CLASSIFICATION ACCURACY FOR ALL THE TRAINED ALGORITHMS.
- THE NN IS THE BEST CLASSIFIER AMONG OUR SELECTION OF DIFFERENT ML ALGORITHMS.
- THE NN IN THE 4-FEATURES SETUP IMPROVES THE DEGENERACY OF PULSARS AND DM SIGNAL
- THE RESULTS ARE IN STATISTICAL AGREEMENT WITHIN DIFFERENT RANDOM SEEDS
- NO UNIDS ARE CLASSIFIED AS DM IN AGREEMENT WITH PREVIOUS WORKS.
- THE PROPOSED METHODOLOGY COULD BE APPLIED TO DIFFERENT SCIENTIFIC CASES

THANK YOU
FOR YOUR ATTENTION