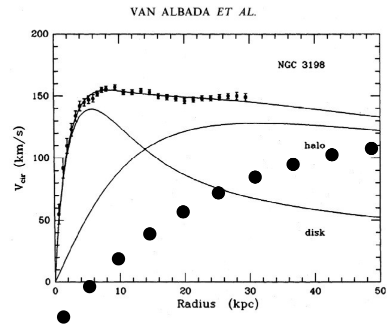
WIMPs or else?
Using Machine Learning for Dark Matter Detection

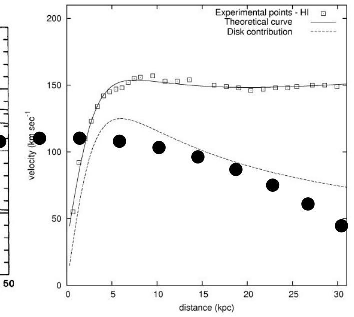
Charanjit Khosa
University of Sussex

UNIVERSITY OF SUSSEX

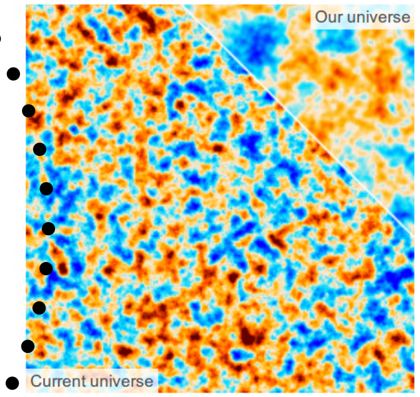
NExT Meeting, University of Sussex, 20th November 2019

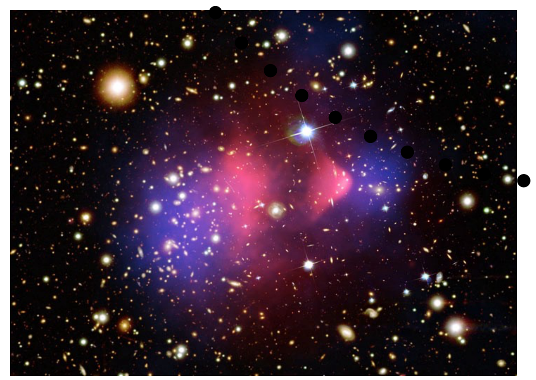
Based on: 1. CKK, Veronica Sanz, Michael Soughton, arXiv: 1910.06058 [hep-ph]
2. CKK, Lucy Mars, Joel Richards, Veronica Sanz [in prep.]

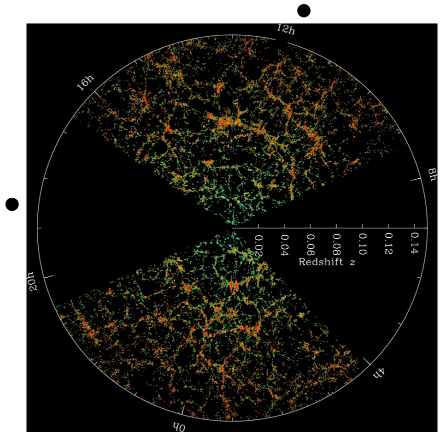


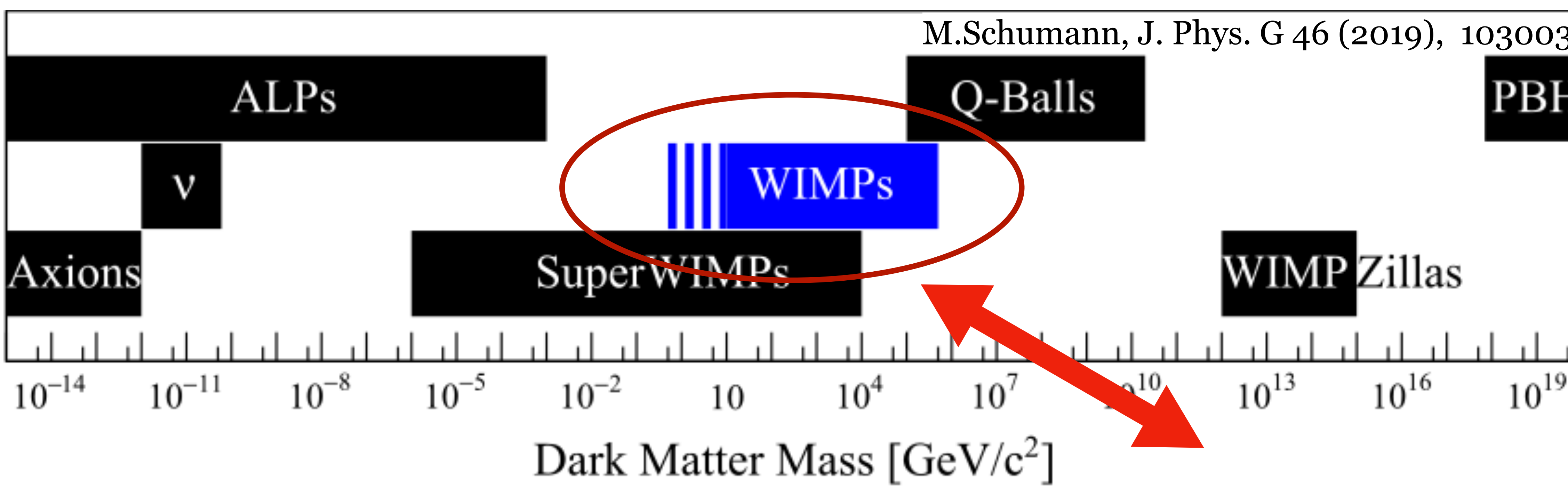


Dark Matter Exists





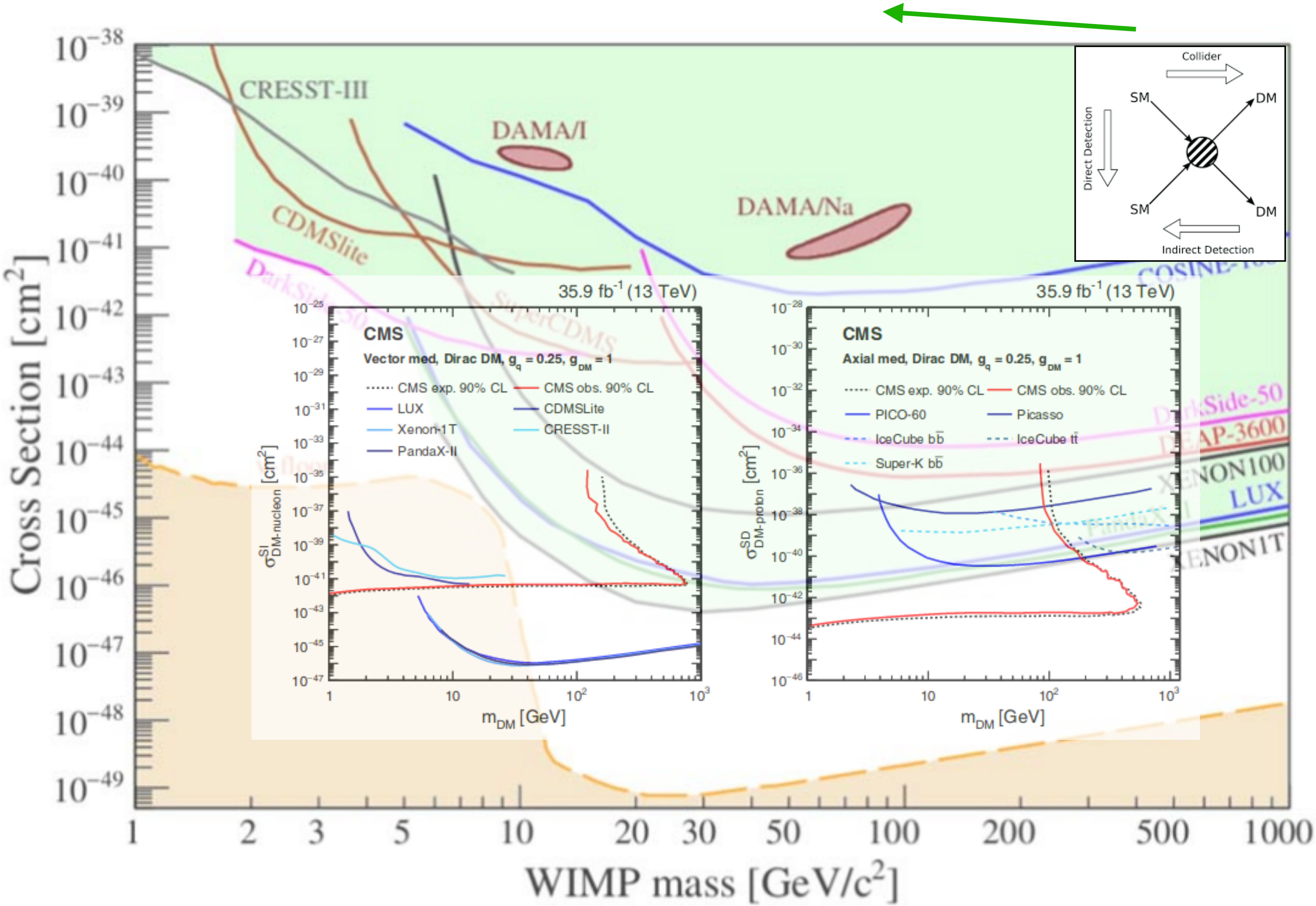




Beyond the Standard Model

Dark Matter Detection: Current Constraints

MET+X (jet, photon, W/Z (jets/ll),H(b/tau)+HF (b/t) pair+single top)



CMS Collaboration, Phys. Rev. D 97 (2018), 092005 M.Schumann, J. Phys. G 46 (2019), 103003

REVIEW

https://doi.org/10.1038/s41586-018-0361-2

Machine learning at the energy and intensity frontiers of particle physics

Alexander Radovic[1]*, Mike Williams[2]*, David Rousseau[3], Michael Kagan[4], Daniele Bonacorsi[5,6], Alexander Himmel[7], Adam Aurisano[8], Kazuhiro Terao[4] & Taritree Wongjirad[9]

Our knowledge of the fundamental particles of nature and their interactions is summarized by the standard model of particle physics. Advancing our understanding in this field has required experiments that operate at ever higher energies and intensities, which produce extremely large and information-rich data samples. The use of machine-learning techniques is revolutionizing how we interpret these data samples, greatly increasing the discovery potential of present and future experiments. Here we summarize the challenges and opportunities that come with the use of machine learning at the frontiers of particle physics.

Table 1 | Effect of machine learning on the discovery and study of the Higgs boson

Analysis	Years of data collection	Sensitivity without machine learning	Sensitivity with machine learning	Ratio of P values	Additional data required
CMS[24] $H \to \gamma\gamma$	2011–2012	2.2σ, $P = 0.014$	2.7σ, $P = 0.0035$	4.0	51%
ATLAS[43] $H \to \tau^+\tau^-$	2011–2012	2.5σ, $P = 0.0062$	3.4σ, $P = 0.00034$	18	85%
ATLAS[99] $VH \to bb$	2011–2012	1.9σ, $P = 0.029$	2.5σ, $P = 0.0062$	4.7	73%
ATLAS[41] $VH \to bb$	2015–2016	2.8σ, $P = 0.0026$	3.0σ, $P = 0.00135$	1.9	15%
CMS[100] $VH \to bb$	2011–2012	1.4σ, $P = 0.081$	2.1σ, $P = 0.018$	4.5	125%

A. Radovic et al., Nature 560(2018) no. 7716,41

https://doi.org/10.1038/s41586-018-0361-2

Machine learning at the energy and intensity frontiers of particle physics

Alexander Radovic[1]*, Mike Williams[2]*, David Rousseau[3], Michael Kagan[4], Daniele Bonacorsi[5,6], Alexander Himmel[7], Adam Aurisano[8], Kazuhiro Terao[4] & Taritree Wongjirad[9]

Our know[illegible] particles of nature and their interactions is summarized by the standard model of particle [illegible] er energies and intensities, which produce extremely large and information-rich data samples. The use of machine-learning techniques is revolutio[illegible] ential of present and future experiments [illegible] achine learning at the frontiers of particle physics.

Machine Learning techniques increase the discovery potential of the experiments

Table 1 | Effect of machine learning on the discovery and study of the Higgs boson

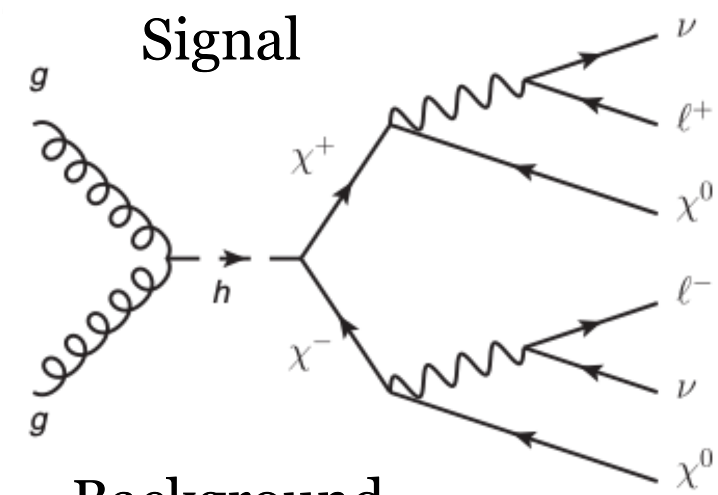
Analysis	Years of data collection	Sensitivity without machine learning	Sensitivity with machine learning	Ratio of P values	Additional data required
CMS[24] $H \to \gamma\gamma$	2011–2012	2.2σ, $P = 0.014$	2.7σ, $P = 0.0035$	4.0	51%
ATLAS[43] $H \to \tau^+\tau^-$	2011–2012	2.5σ, $P = 0.0062$	3.4σ, $P = 0.00034$	18	85%
ATLAS[99] $VH \to bb$	2011–2012	1.9σ, $P = 0.029$	2.5σ, $P = 0.0062$	4.7	73%
ATLAS[41] $VH \to bb$	2015–2016	2.8σ, $P = 0.0026$	3.0σ, $P = 0.00135$	1.9	15%
CMS[100] $VH \to bb$	2011–2012	1.4σ, $P = 0.081$	2.1σ, $P = 0.018$	4.5	125%

A. Radovic et al., Nature 560(2018) no. 7716,41

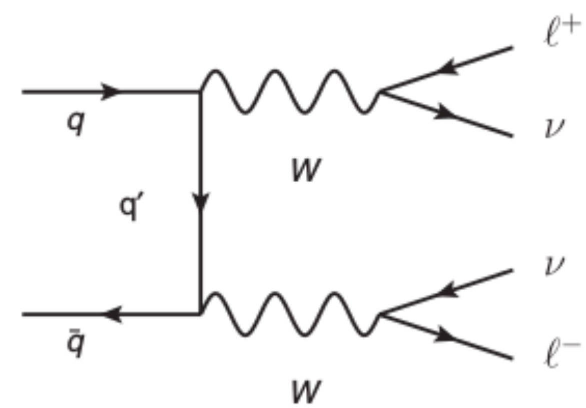
nature COMMUNICATIONS

r

Signal



Background



ARTICLE

Received 19 Feb 2014 | Accepted 4 Jun 2014 | Published 2 Jul 2014 DOI: 10.1038/ncomms5308

Searching for exotic particles in high-energy physics with deep learning

P. Baldi[1], P. Sadowski[1] & D. Whiteson[2]

Collisions at high-energy particle colliders are a traditionally fruitful source of exotic particle discoveries. Finding these rare particles requires solving difficult signal-versus-background classification problems, hence machine-learning approaches are often used. Standard approaches have relied on 'shallow' machine-learning models that have a limited capacity to learn complex nonlinear functions of the inputs, and rely on a painstaking search through manually constructed nonlinear features. Progress on this problem has slowed, as a variety of techniques have shown equivalent performance. Recent advances in the field of deep learning make it possible to learn more complex functions and better discriminate between signal and background classes. Here, using benchmark data sets, we show that deep-learning methods need no manually constructed inputs and yet improve the classification metric by as much as 8% over the best current approaches. This demonstrates that deep-learning approaches can improve the power of collider searches for exotic particles.

Low-level features: $p_T^{l_1}, p_T^{l_2}, \sum p_T^j, MET, N_j$

High-level features: **Axial MET**, M_{T_2}, **razor quantities**

SUSY benchmark:
chargino production (lepton+MET final state)

Technique	Low-level	High-level	Complete
AUC			
BDT	0.850 (0.003)	0.835 (0.003)	0.863 (0.003)
NN	0.867 (0.002)	0.863 (0.001)	0.875 (<0.001)
$NN_{dropout}$	0.856 (<0.001)	0.859 (<0.001)	0.873 (<0.001)
DN	0.872 (0.001)	0.865 (0.001)	0.876 (<0.001)
$DN_{dropout}$	0.876 (<0.001)	0.869 (<0.001)	0.879 (<0.001)
Discovery significance			
NN	6.5σ	6.2σ	6.9σ
DN	7.5σ	7.3σ	7.6σ

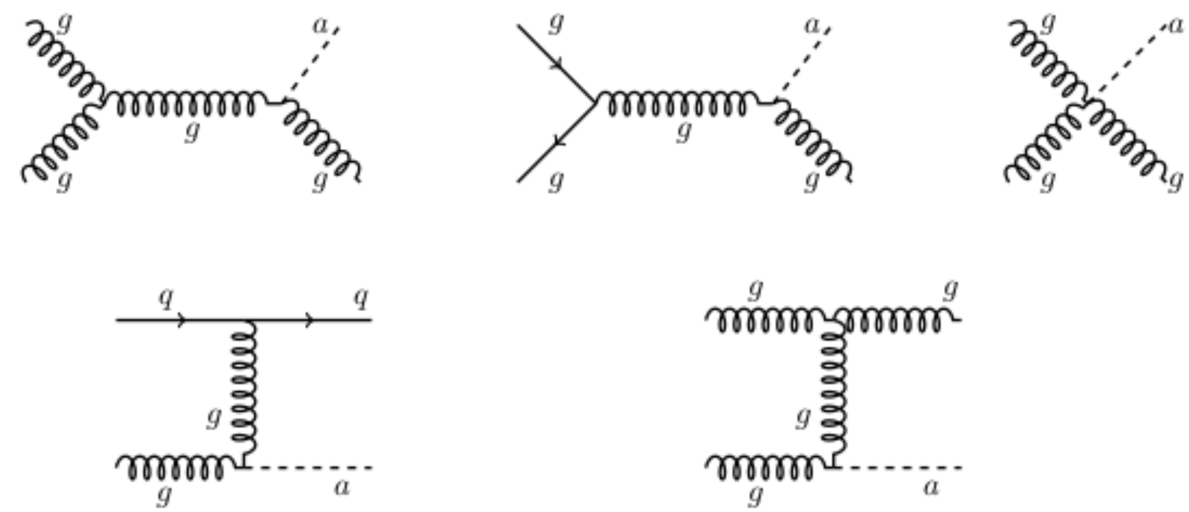
BDT, boosted decision tree; DN, deep neural network; NN, shallow neural network; SUSY, supersymmetry particle.

Deep Learning methods improve the reach of the collider searches for new physics searches

- **SMEFT (new physics deformations)** — J. Brehmer, K. Cranmer, G. Louppe and J Pavez, Phys. Rev. Lett. 121 (2018) 111801, Phys. Rev. D 98 (2018) 052004; Felipe F. Freitas, CKK, Veronica Sanz, arXiv: 1902.05803 [hep-ph]
- **Top tagging** — G.Kasieczka, T.Plehn, M.Russell and T.Schell, JHEP 1705 (2017) 006
- **Novelty detection** — J. Hajer, Y.Y. Li, T. Liu and H. Wang, arXiv: 1807.10261[hep-ph]
- **Anomaly detection** — M. Farina, Y. Nakai and D. Shih, arXiv: 1808.08992[hep-ph]
- **Cosmological phase transitions** — M.L.Piscopo, M.Spannowsky and P.Waite, arXiv:1902.05563 [hep-ph]
- **DijetGAN** — R.Di Sipio, M.Faucci Giannelli et al, arXiv:1903.02433 [hep-ex]
- **many more..**

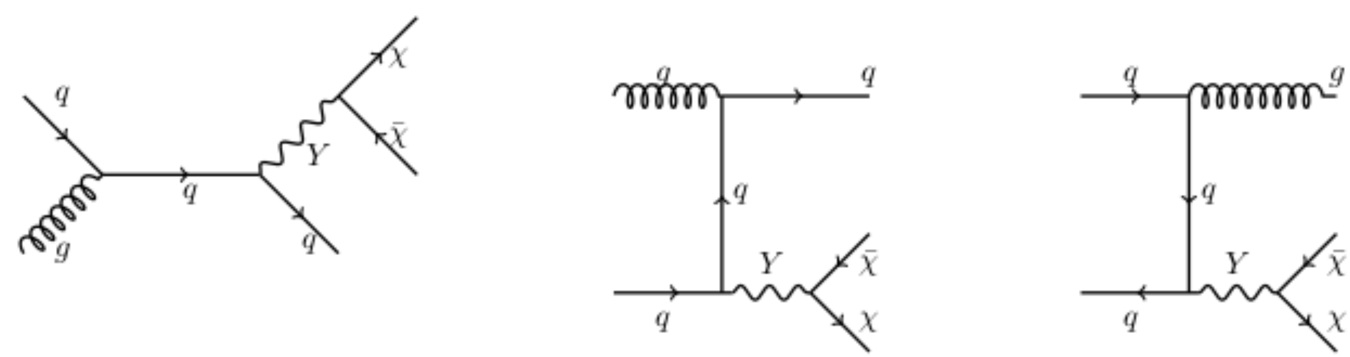
DM at LHC: Monojet Channel

Model	Mass	Type of coupling
SUSY1	$m_{\tilde{\chi}^0} = 100$ GeV	Bino-like
SUSY2	$m_{\tilde{\chi}^0} = 200$ GeV	Bino-like
SUSY3	$m_{\tilde{\chi}^0} = 300$ GeV	Bino-like
ALP	negligible	gluon-ALP
EFT	negligible	4-fermion


(a) Monojet process in linear ALPs

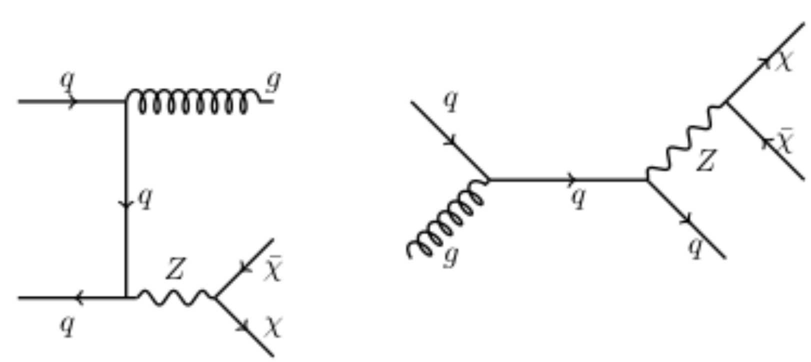
Axion-Like particles (ALPs) could decay after being produced but not inside the detector

$$\mathscr{L}_a \supset -\frac{g_{agg}}{2} a\,\mathbf{Tr}\left[G_{\mu\nu}\tilde{G}^{\mu\nu}\right]$$


(b) Monojet process in case of spin 1 mediator

Simplified models: spin-1 mediator

$$\mathscr{L}_Y = \bar{\chi}\gamma_\mu g_\chi^V \chi Y^\mu$$


(c) Monojet process in MSSM

CKK, Veronica Sanz, Michael Soughton, arXiv: 1910.06058 [hep-ph]

Analysis set-up

LO, parton level analysis

$p_T^j(MET), \eta^j, \phi^j$ $\qquad p_T^j > 130\text{GeV}$

$\sqrt{s} =$ 14 TeV, 400K events
using MC@NLO Madgraph

SUSY-WIMP $\qquad pp \to \tilde{\chi}_1^0 \tilde{\chi}_1^0 j$

ALPs $\qquad pp \to aj$

EFT, spin-1 mediator $\qquad pp \to \chi\bar{\chi}j$

Feynrules Model :
MSSM-SLHA2,ALPsEFT, DMSimp

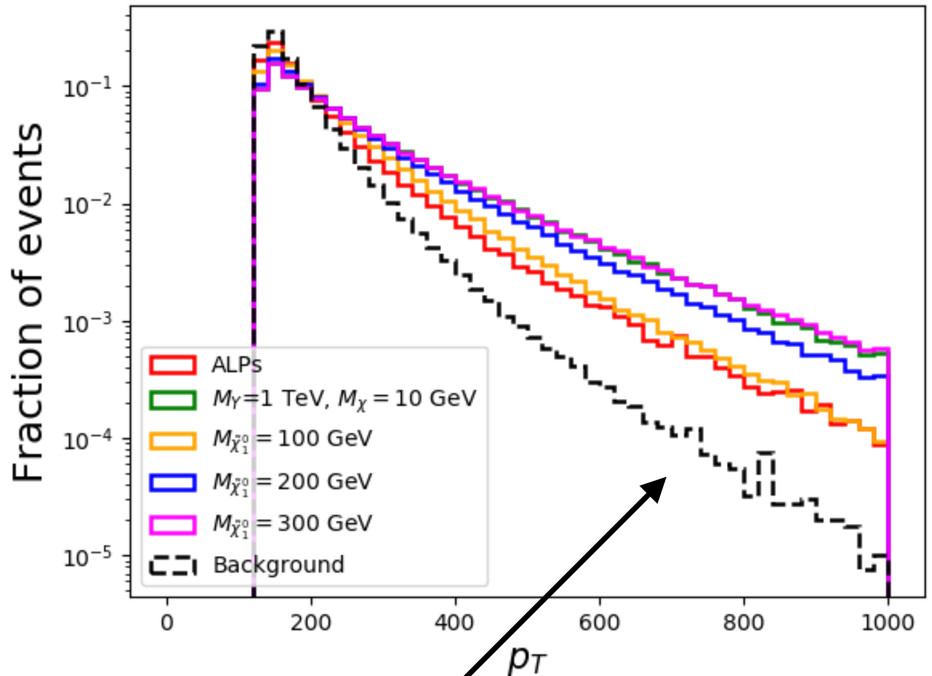
NLO, detector level analysis

200K events
DELPHES (ATLAS default run card)

$p_T^{j_1}, p_T^{j_2}, \eta^{j_1}, \eta^{j_2}, \text{MET}, \Delta\phi_{j_1 j_2}, \Delta\phi_{MET}^{j_1}, \Delta\phi_{MET}^{j_2}$ $\qquad p_T^{j_1} > 130\text{GeV}, p_T^{j_2} > 25\text{GeV}$

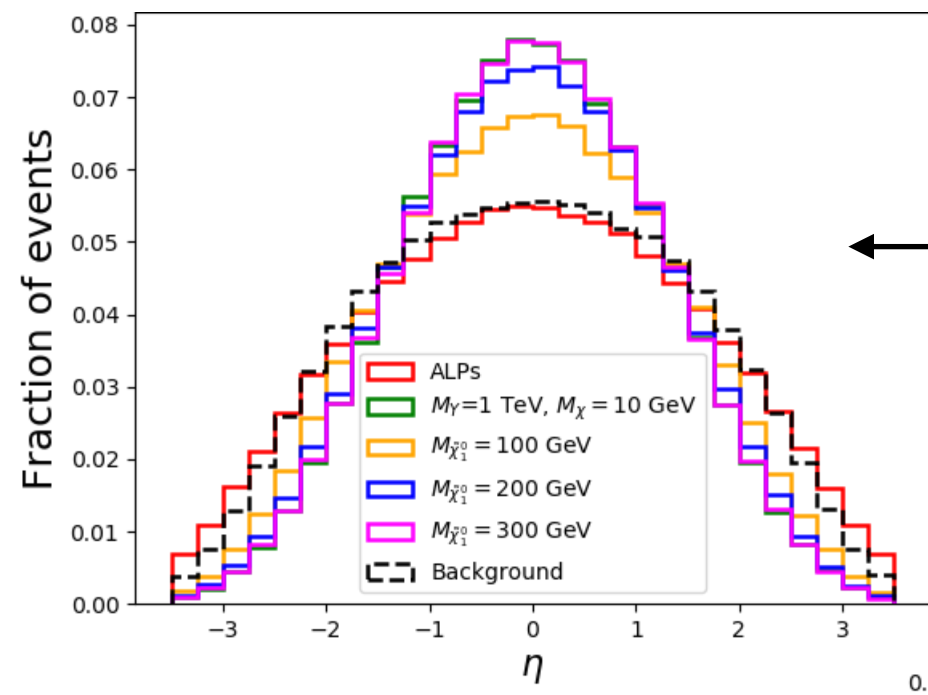
CKK, Veronica Sanz, Michael Soughton, arXiv: 1910.06058 [hep-ph]

Kinematic Distributions

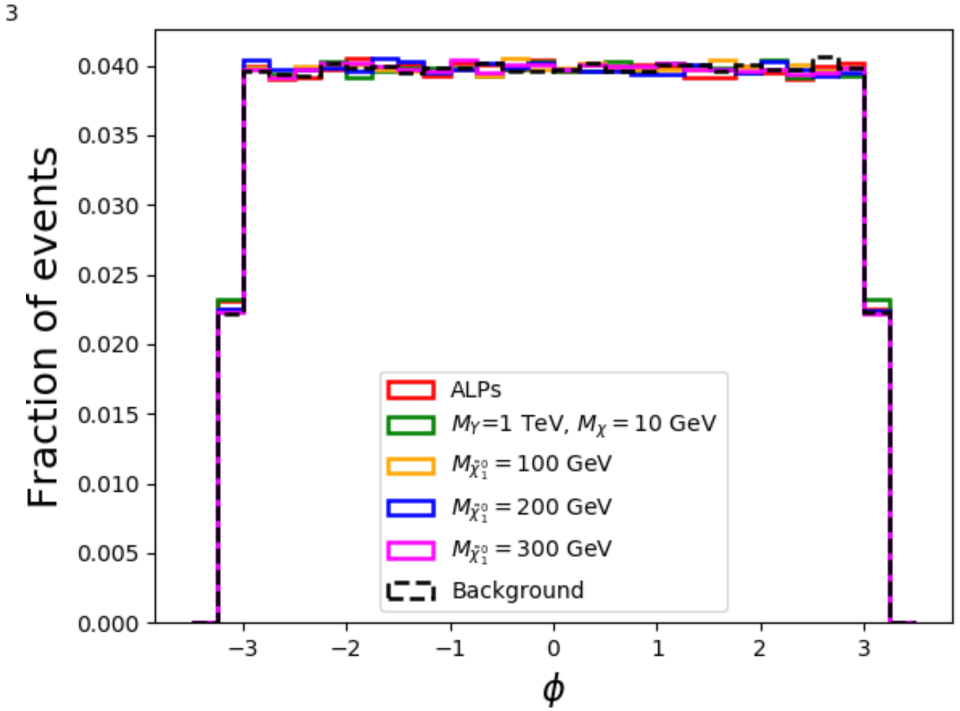


Transverse momentum of the jet

$$pp \to Z(\to \nu\bar{\nu})j$$



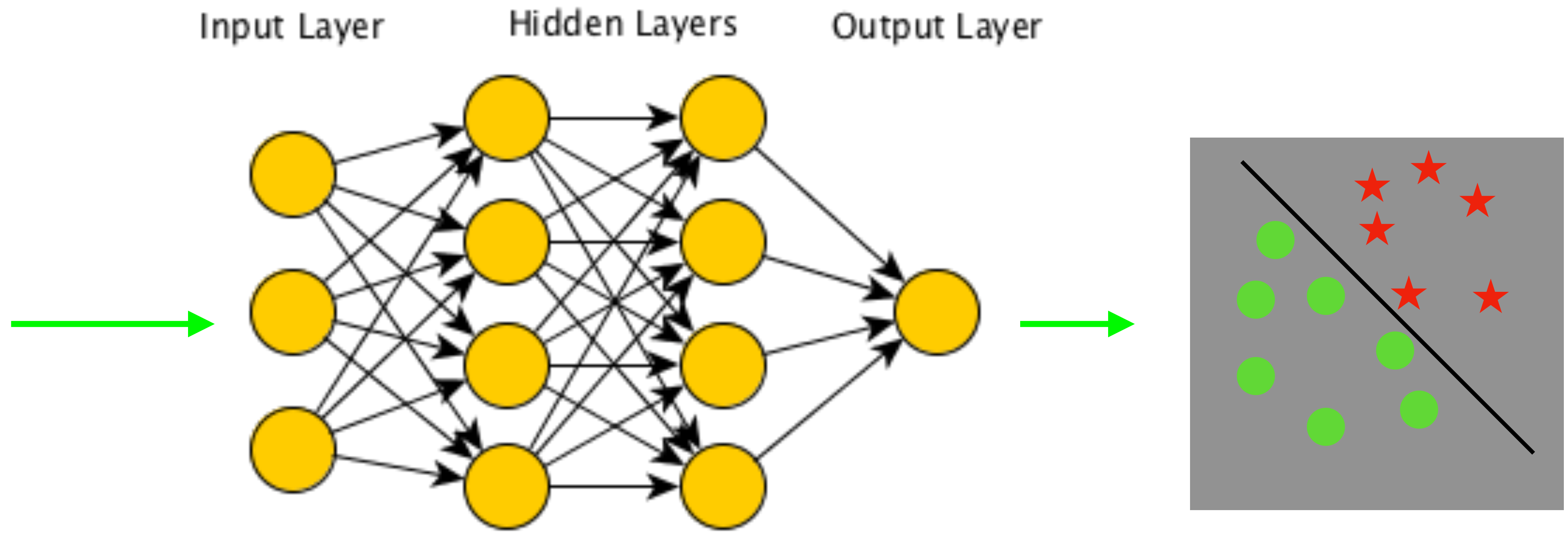
ALP has similar distribution as SM background



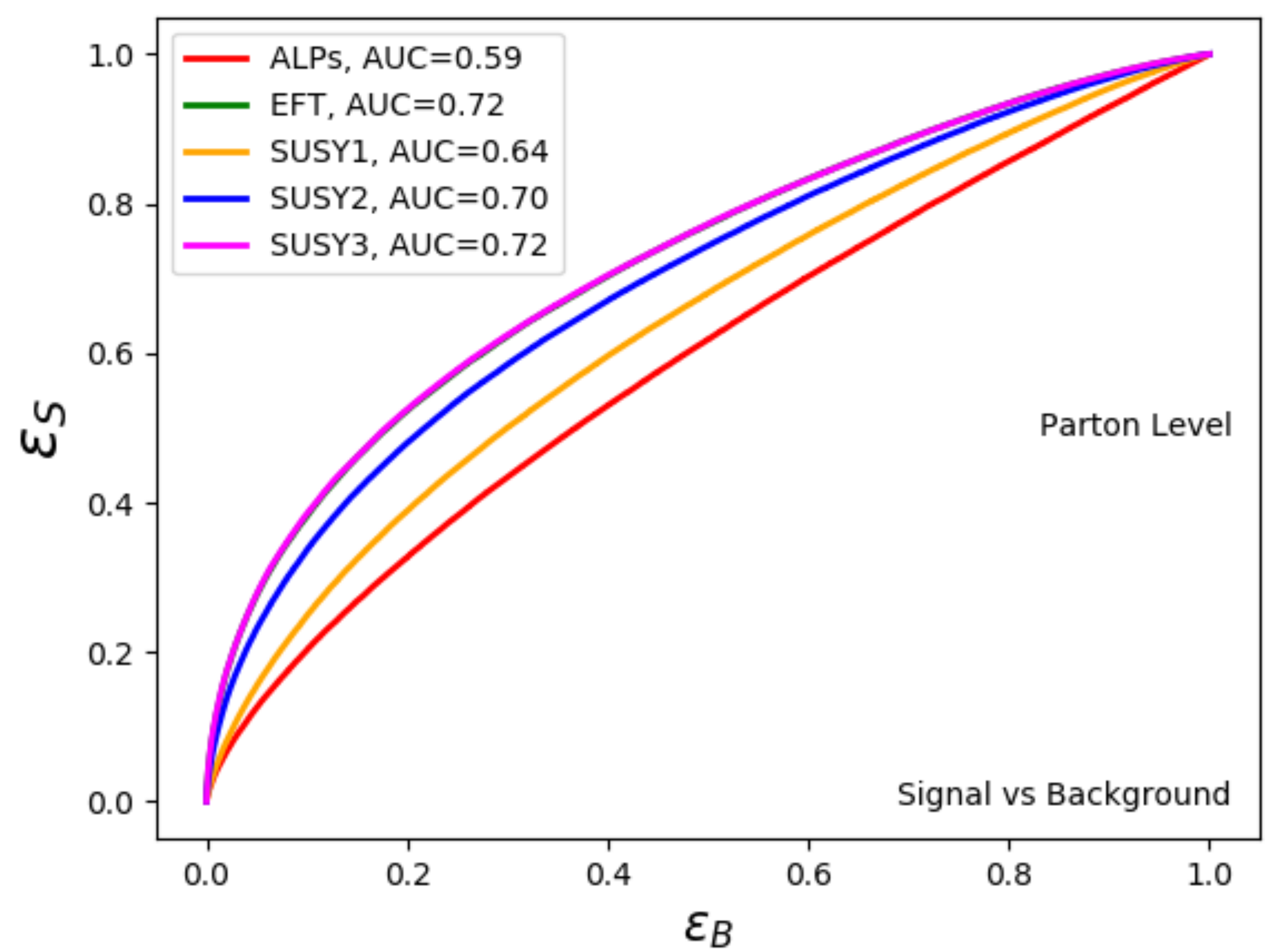
Azimuthal angular distribution,
No preferred direction for the WIMP

(Deep) Neural Network

$p_T^j(MET), \eta^j, \phi^j$

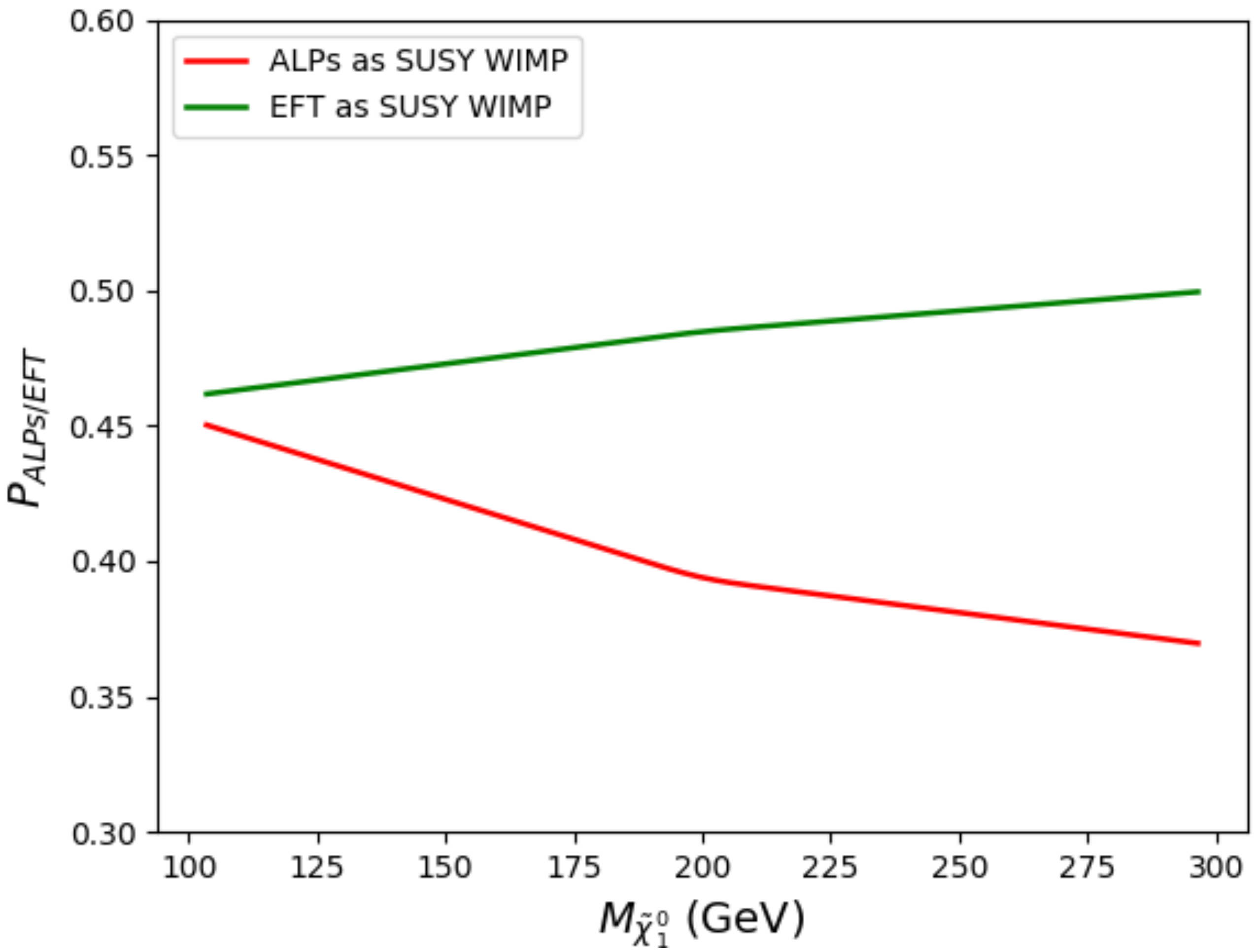


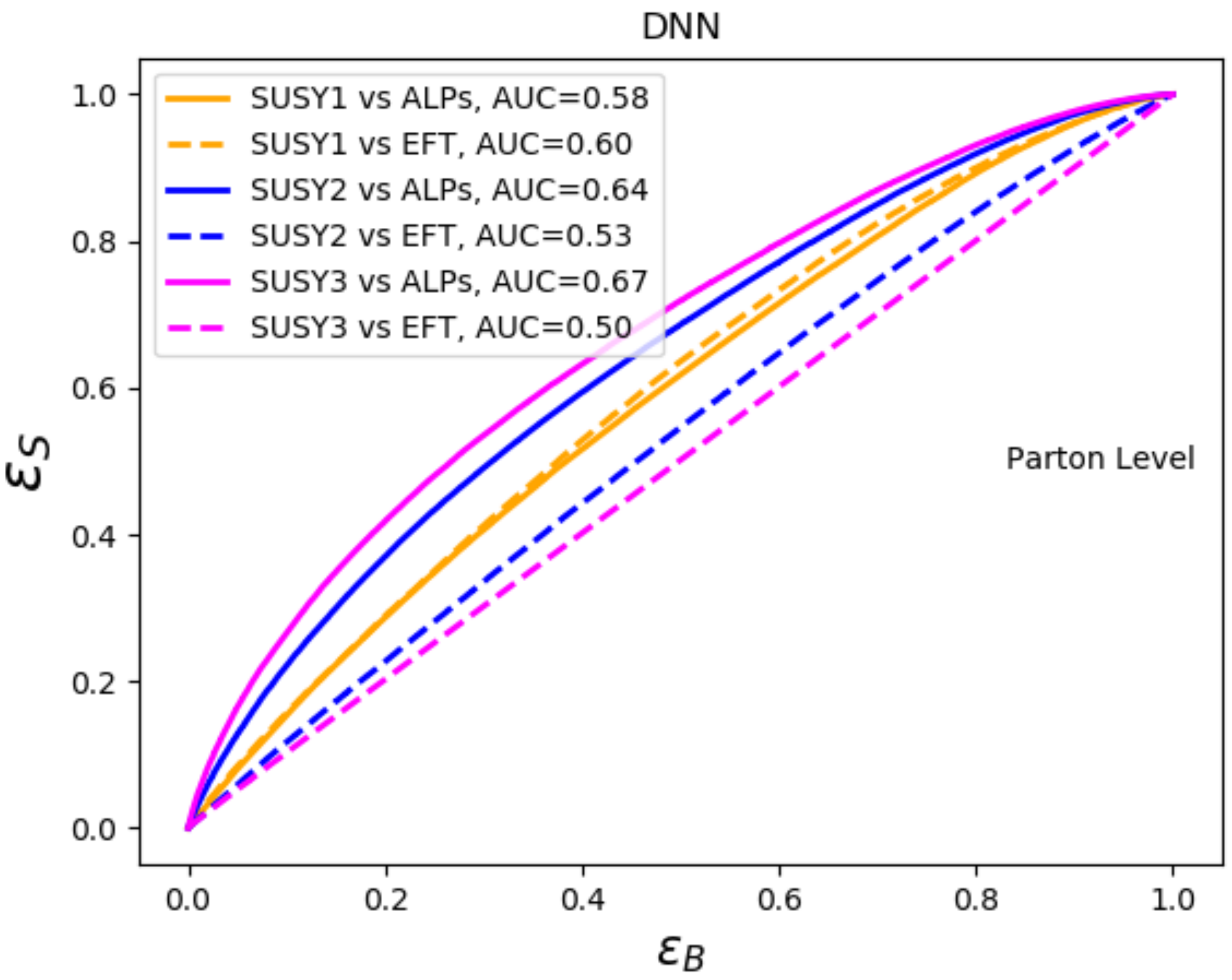
- **Training set: Test set = 70 %: 30% (data scaling)**
- **Hidden layers: 5 (optimised)**
- **Activation function: ReLu**
- **Dropouts: 0.2**
- **Loss function: Binary Cross-entropy**



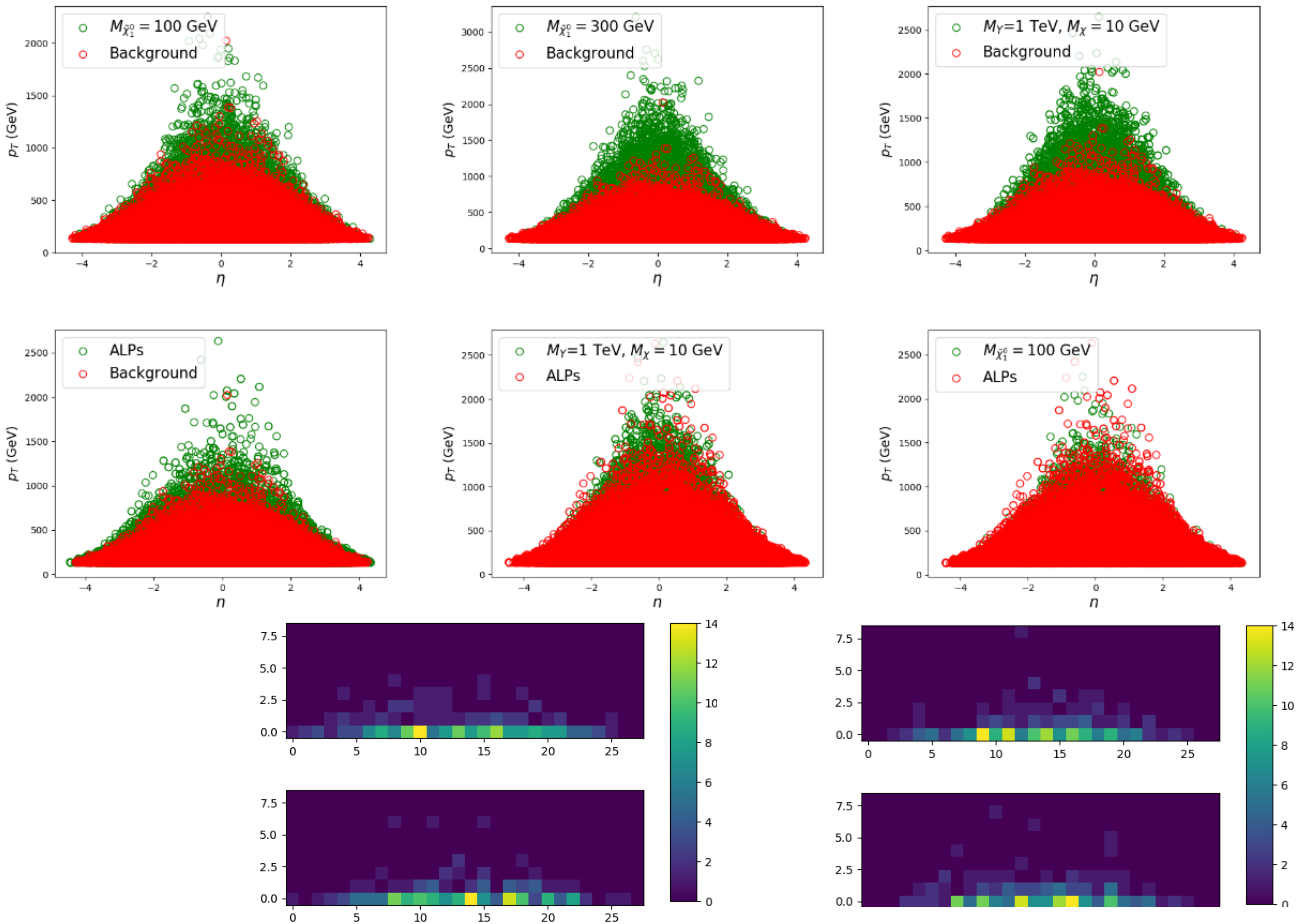
ϵ_S : **Signal Selection**
ϵ_B : 1− **Background Rejection**
AUC: Area under the ROC curve

DM Characterization

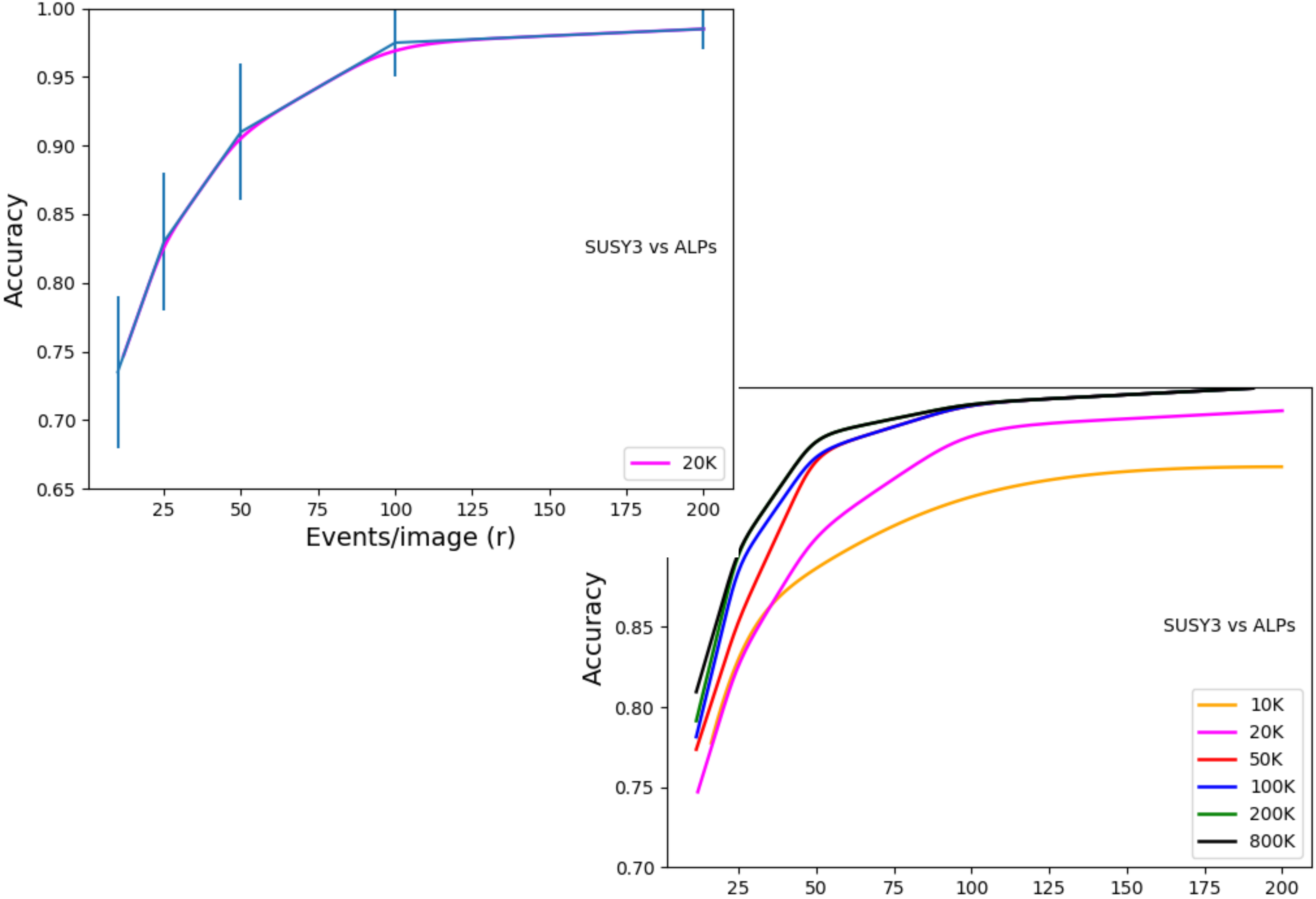




2D Distributions (LO Parton Level)

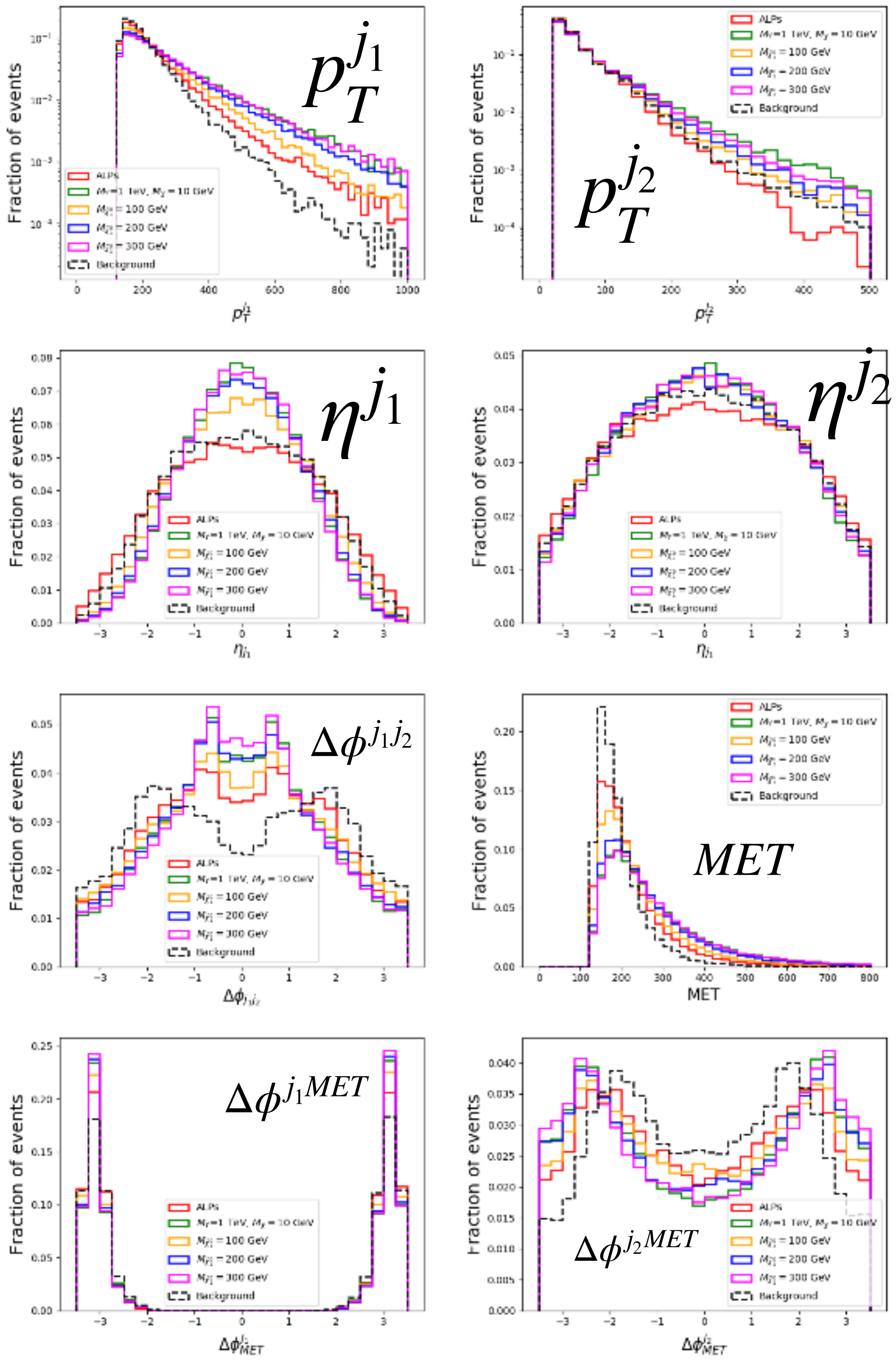


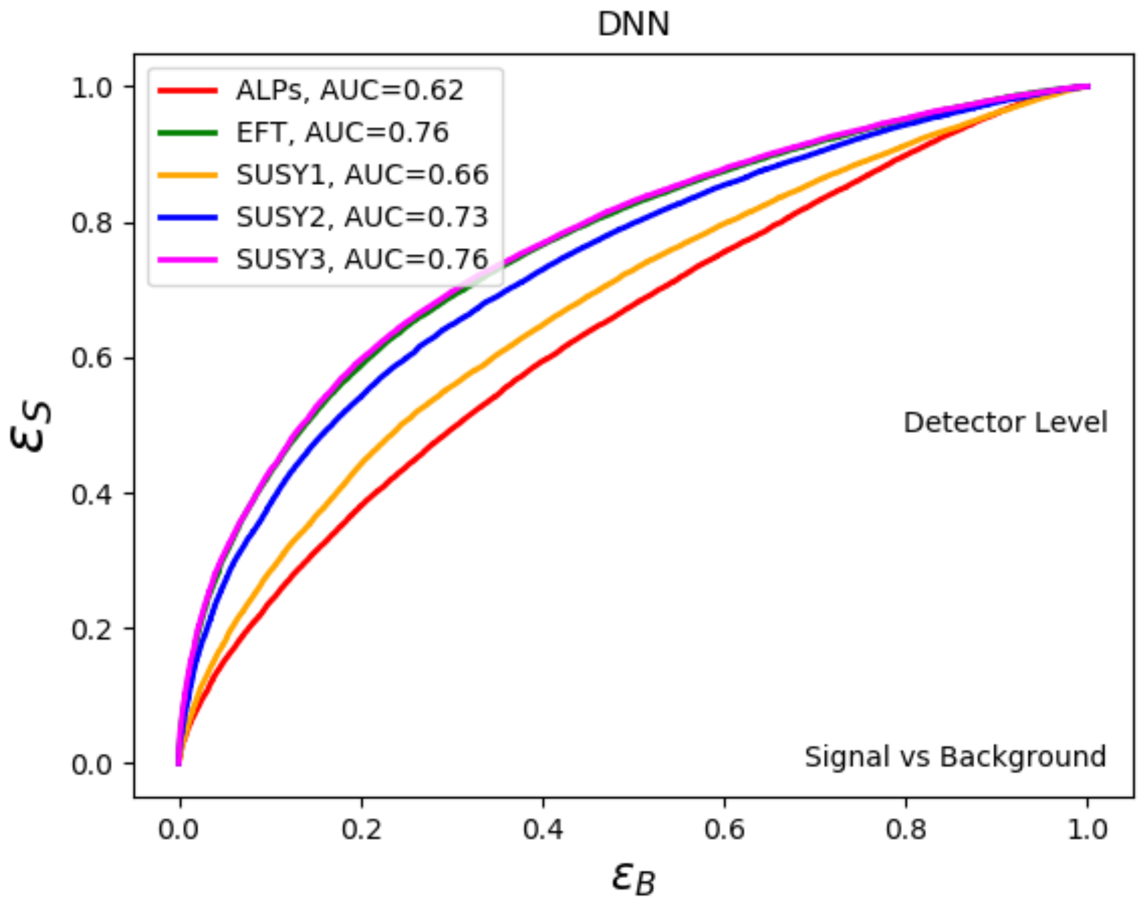
DNN for 2D Histograms



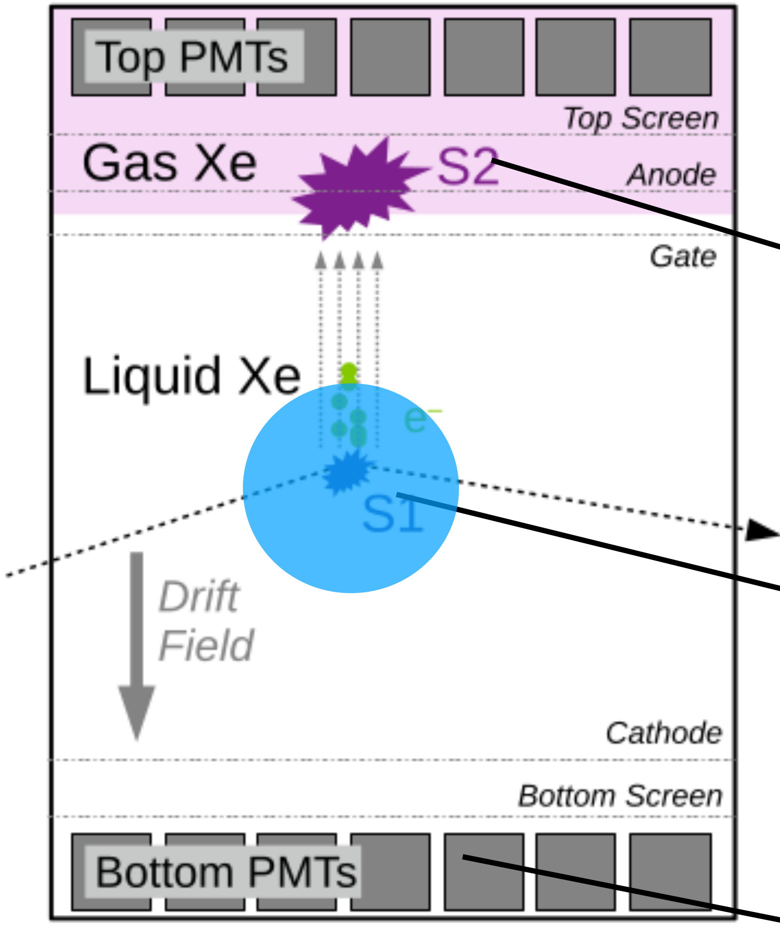
Dijet Case

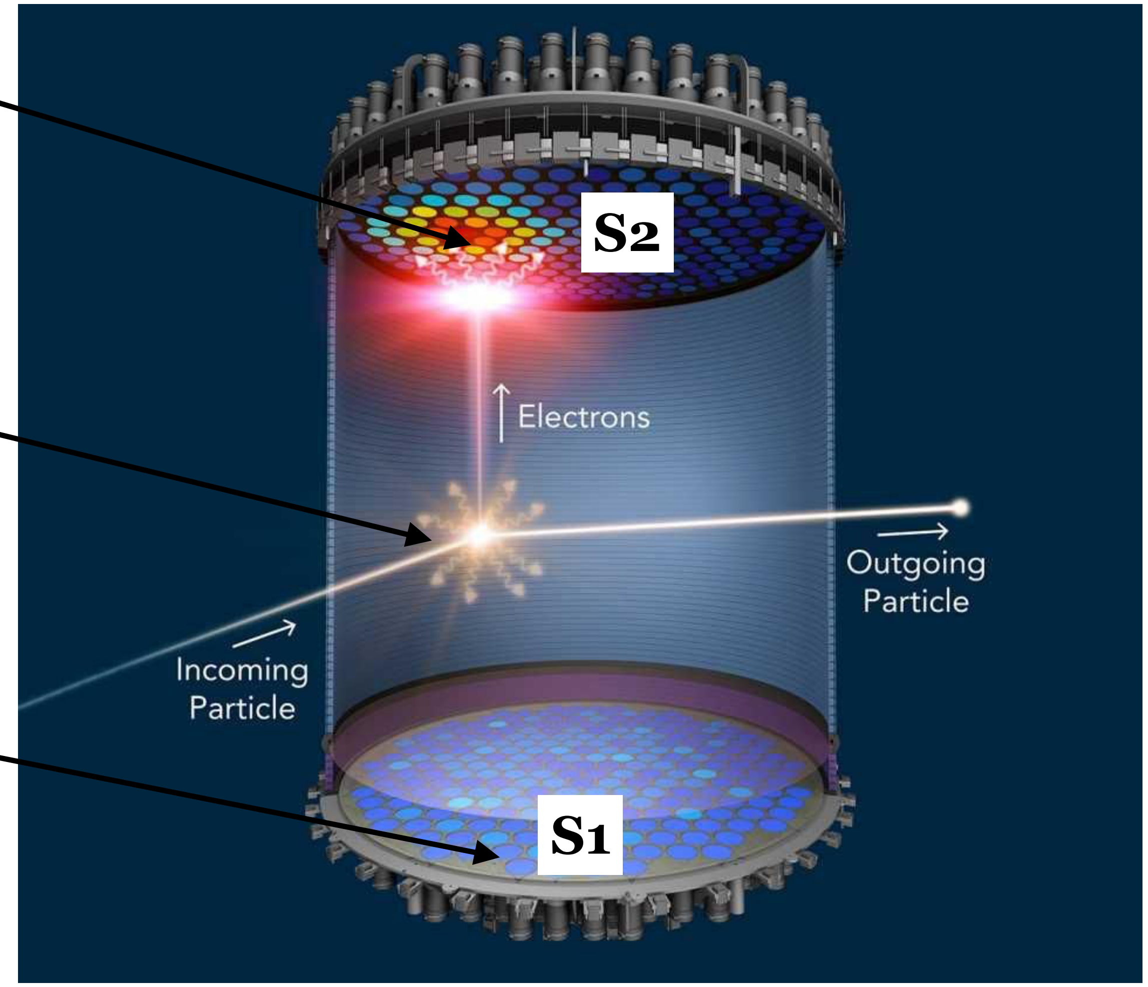
Same NN Architecture

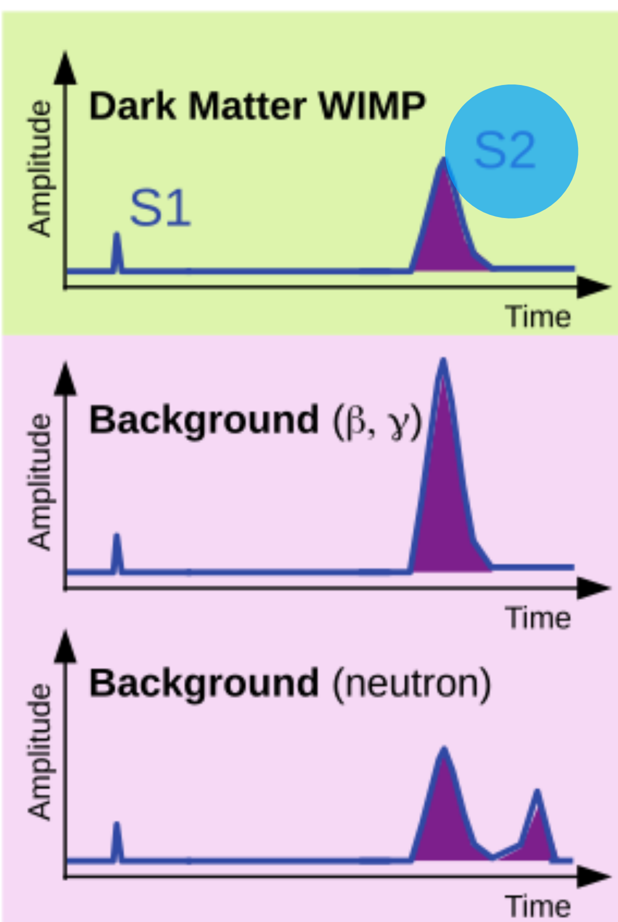




TPC for Xenon1T Experiment







Eur. Phys. J. C (2017) 77:881

Dual-Phase TPC:

- **Particle interacts with LXe**
- **Prompt scintillation, S1**
- **Ionisation electrons drift upwards due to E_{drift}**
- **Extracted using $E_{extraction}$ at GXe**
- **Delayed signal, S2**

TPC output as images

Name	Expected number of events
Electronic recoils (ER)	61.879487
CNNS (ν)	0.000901
Radiogenic neutrons	0.058570
Accidental coincidences (acc)	0.220000
Wall leakage (wall)	0.520000
Anomalous (anom)	0.090004
500 GeV/c^2, 10^{-45} cm^2 WIMP	35.029005

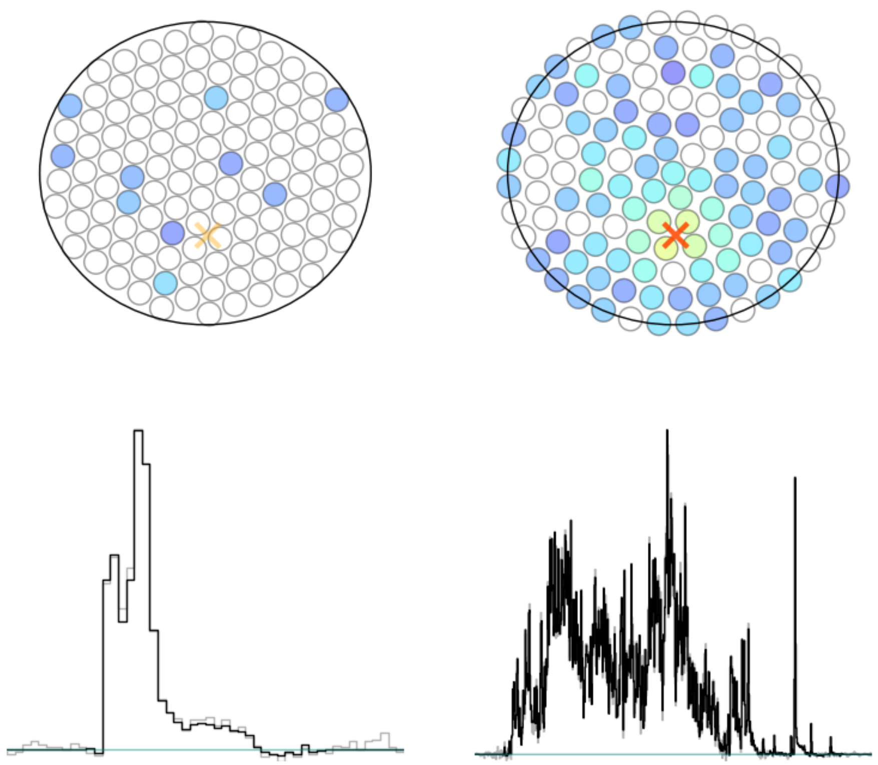
TABLE I. Expected number of events for each type of background within the fiducial mass and a 500 GeV/c^2, 10^{-45} cm^2 WIMP (Generated using Laidbax).

Simulation Tools for XENON1T

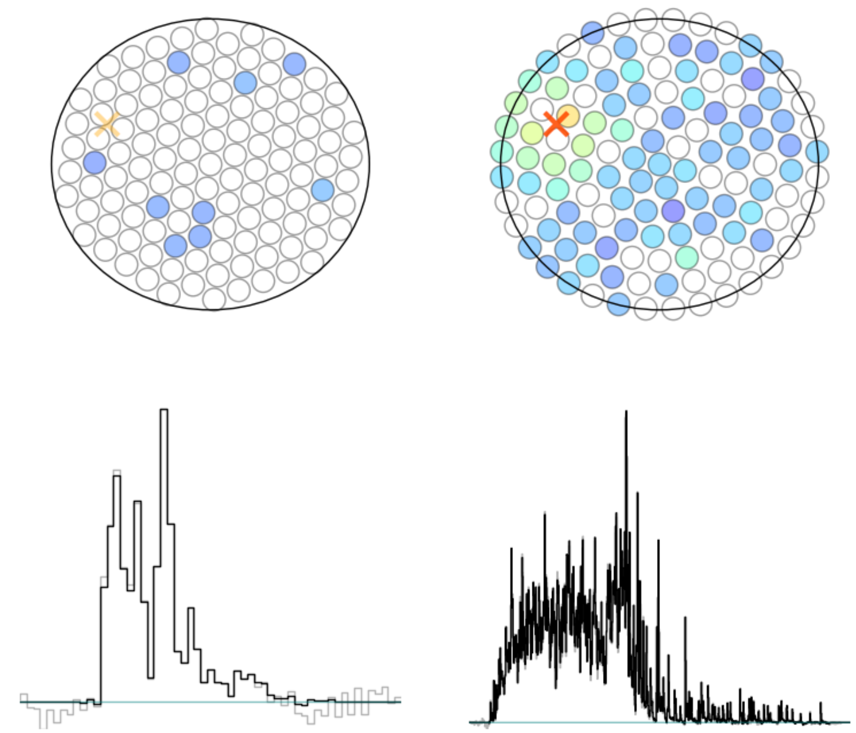
Laidbax
Wimprates
Blueice

→ **FAX** → **PAX**

WIMP event

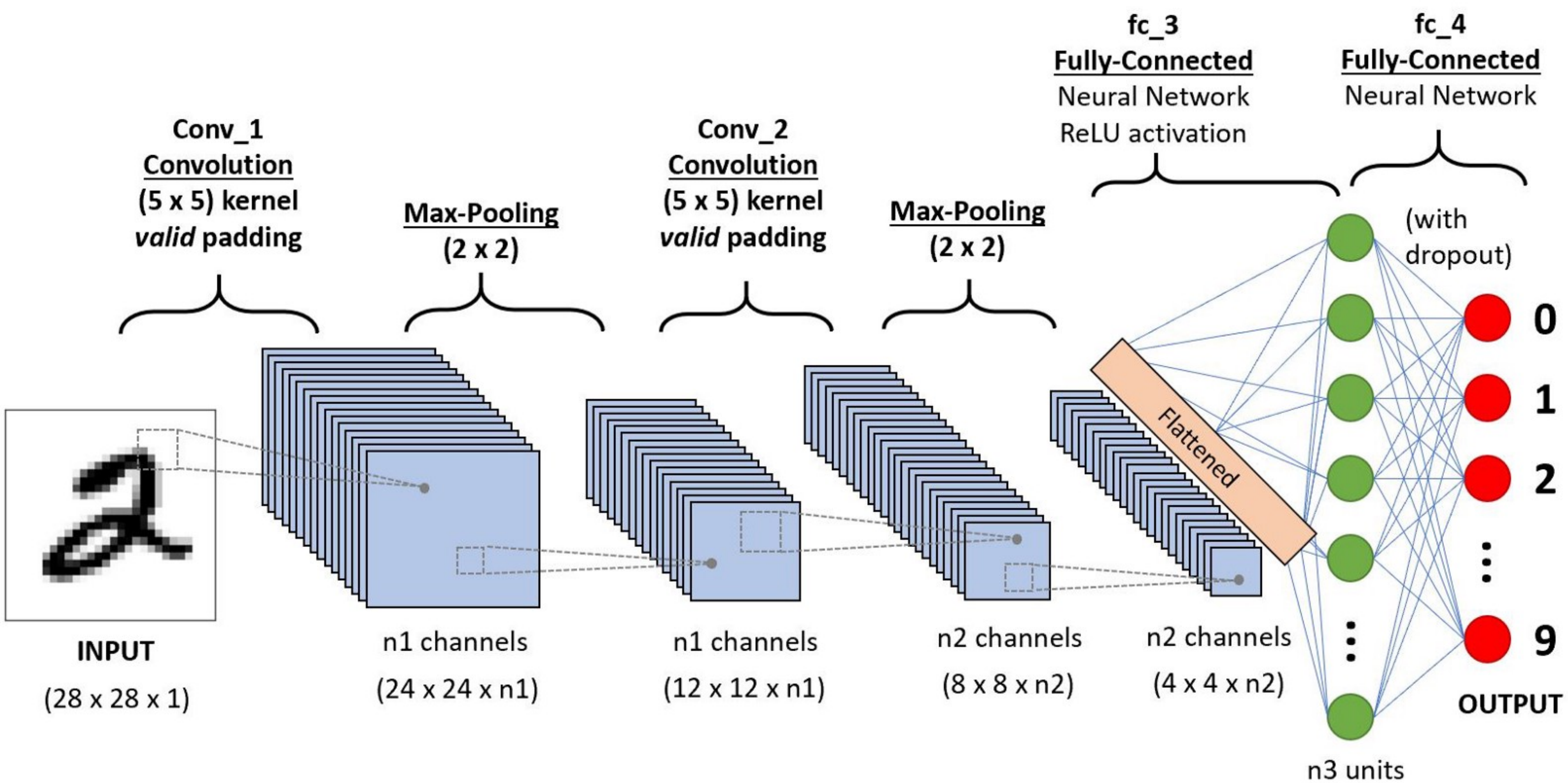


ER background



https://github.com/XENON1T/pax
https://github.com/XENON1T/laidbax
Aalbers, J. (2018) *Dark Matter Search with XENON1T* (Doctoral dissertation, University of Amsterdam)

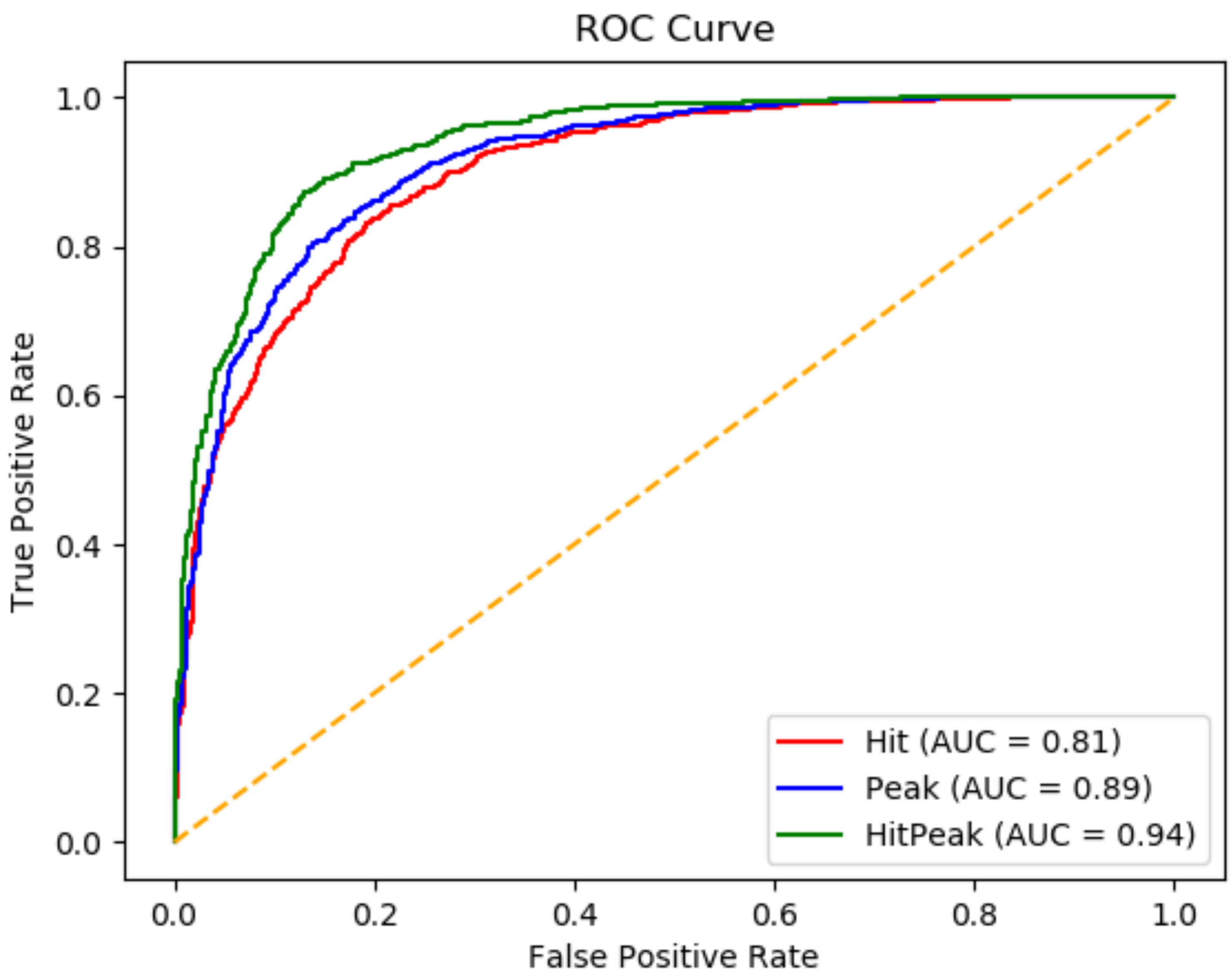
Convolutional Neural Network



- **Convolutional layer: detect edges**
- **Pooling: compress the information**
- **Classification: (fully) connected hidden layers**

WIMP versus ER background

- **CNN:**
  - **2 Convolution layers + ReLU activation function**
  - **20 Neurons per layer**
  - **Pooling layer:**
    - **Max Pooling**
    - **Flatten**
    - **ReLU activation function**
  - **Output layer:**
    - **Sigmoid activation function**



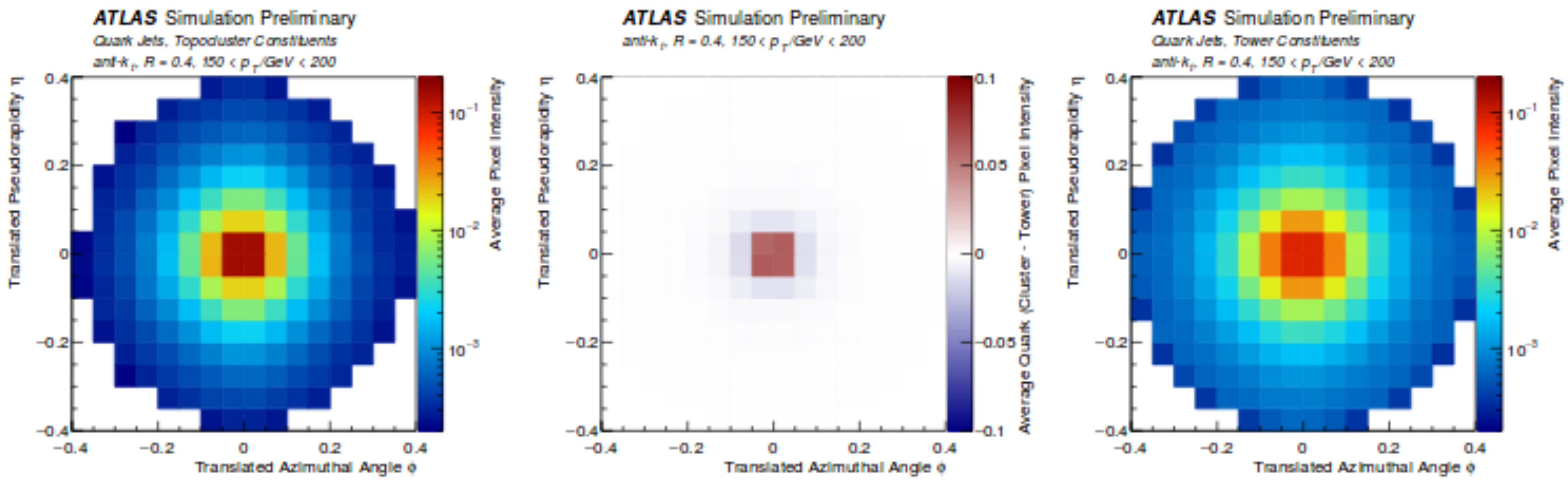
Conclusions and Outlook

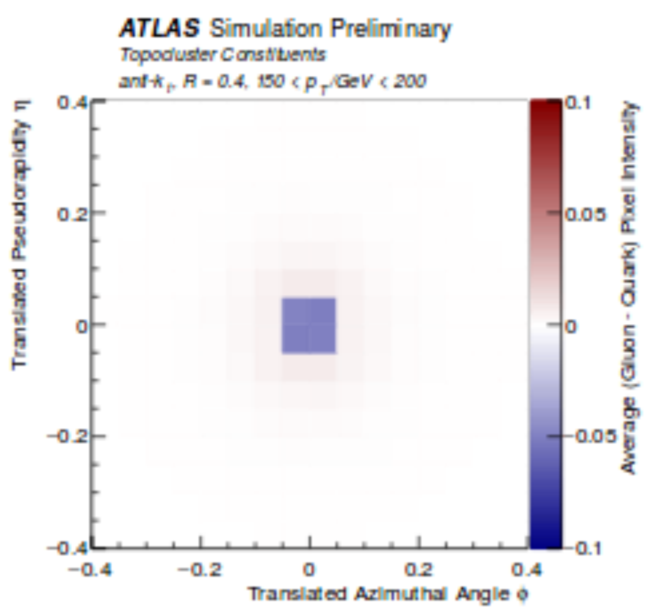
- We demonstrate the possibility of using ML for the Dark Matter Searches

- It shows the promising reach for disentangling among collider Dark Matter searches as well as for the direct detection experiments

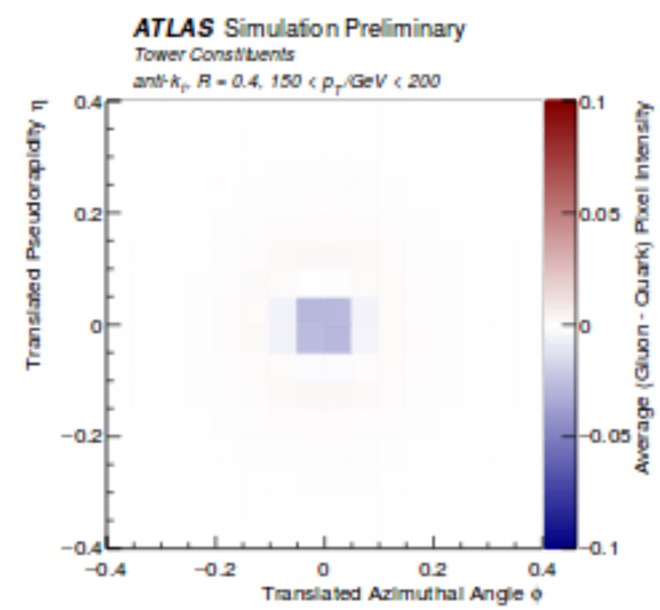
- These works set the benchmark for the unsupervised anomaly detection methods

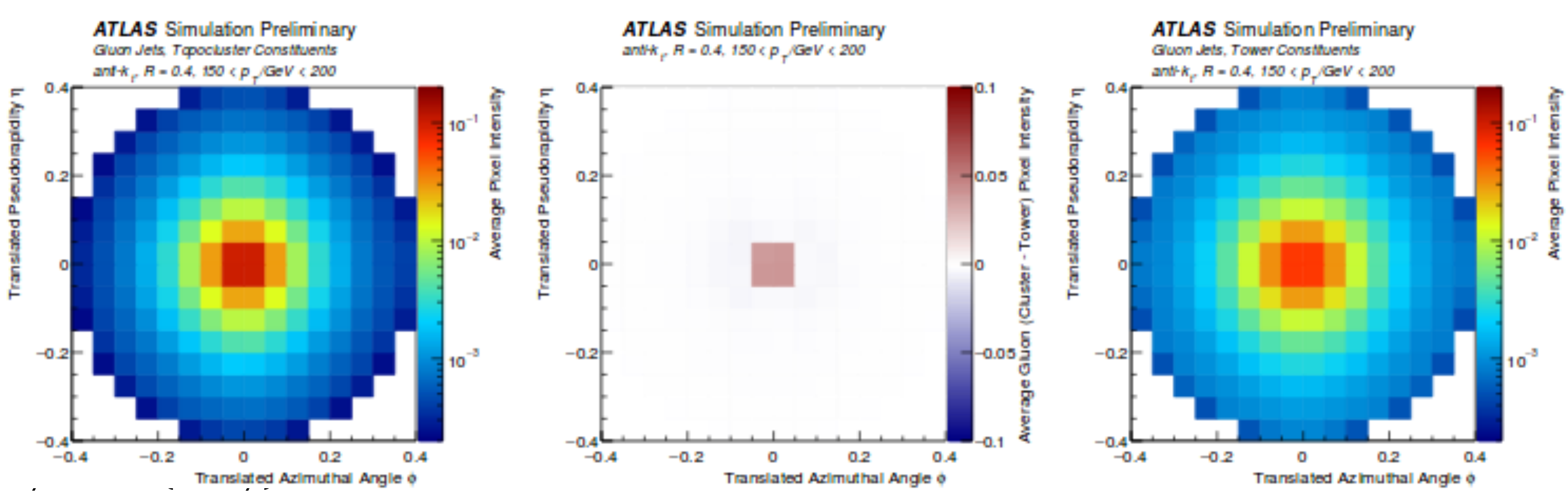
Detector output can be viewed as images





The ATLAS Collaboration. Quark versus Gluon Jet Tagging using Jet Images with the ATLAS Detector. Report No. ATL-PHYS-PUB-2017-017, https://cds.cern.ch/record/2275641 (CERN, 2017)
CMS Collaboration. New Developments for Jet Substructure Reconstruction in CMS. Report No. CMS-DP-2017-027, https://cds.cern.ch/record/2275226(CERN, 2017)

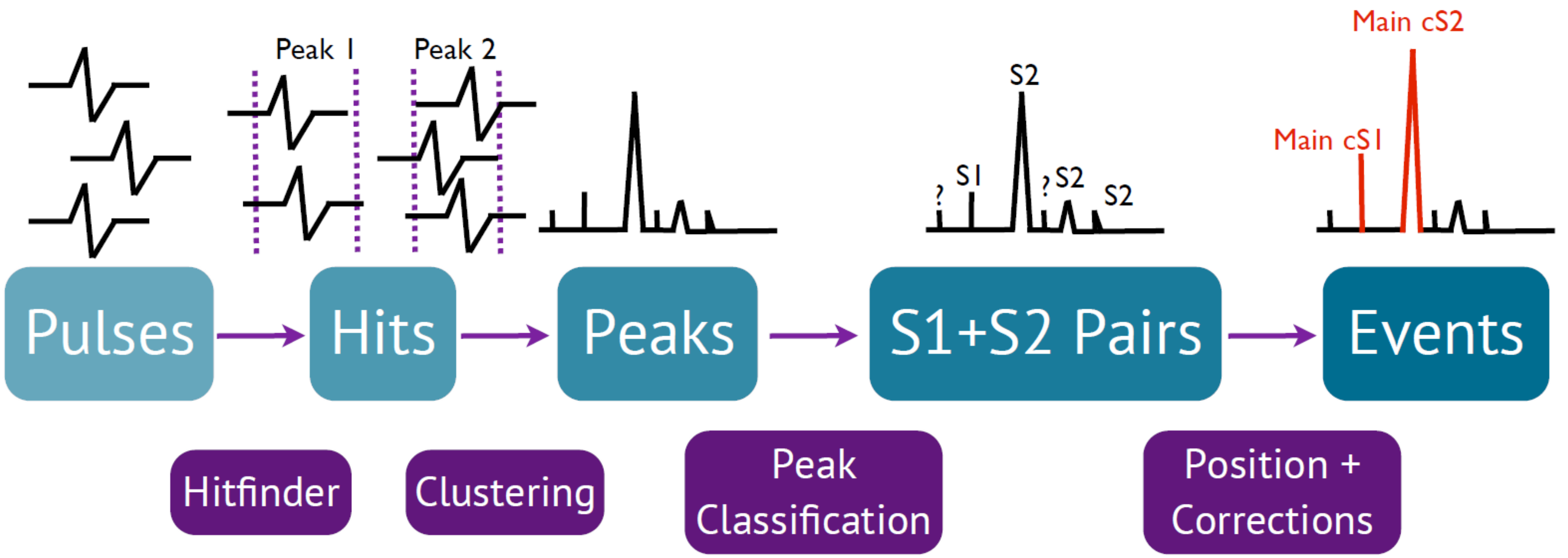


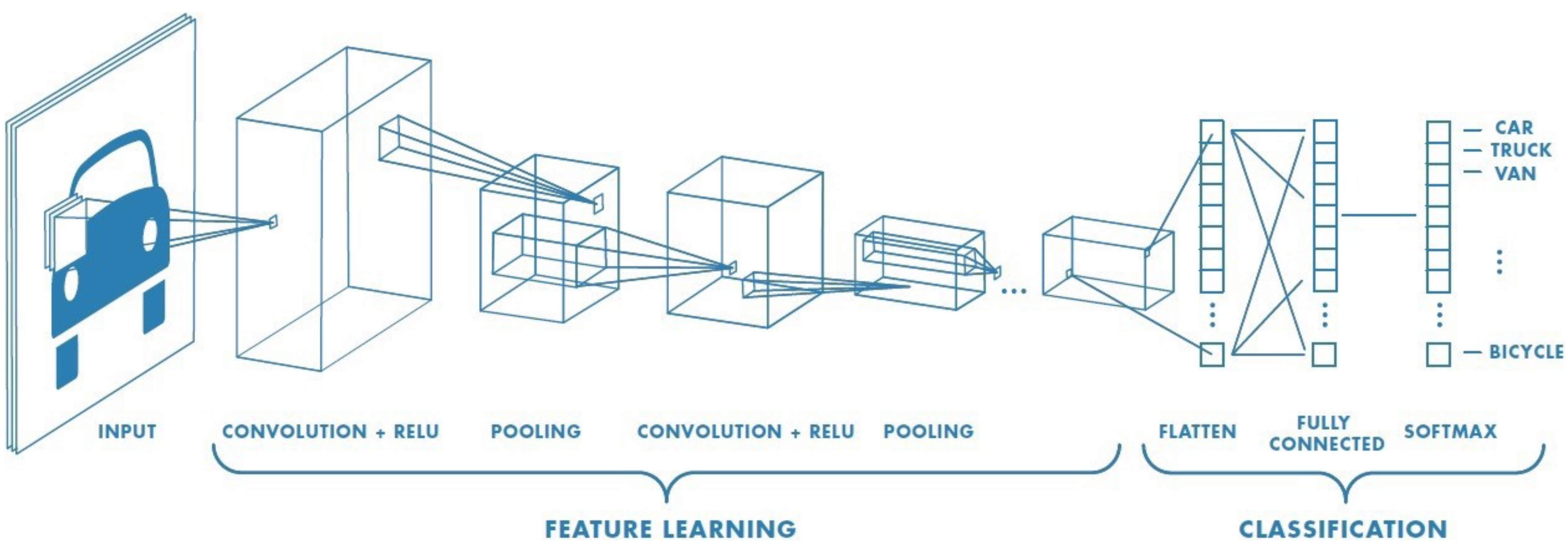


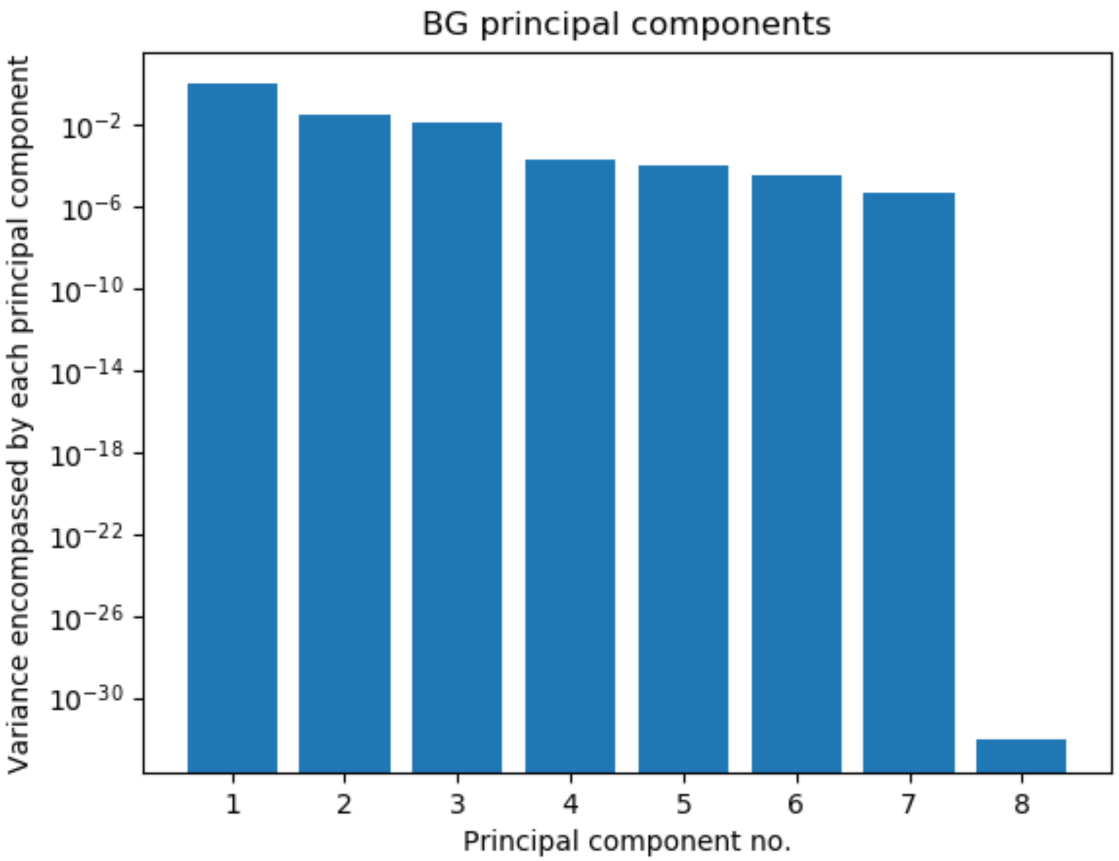
Neutrino Physics:

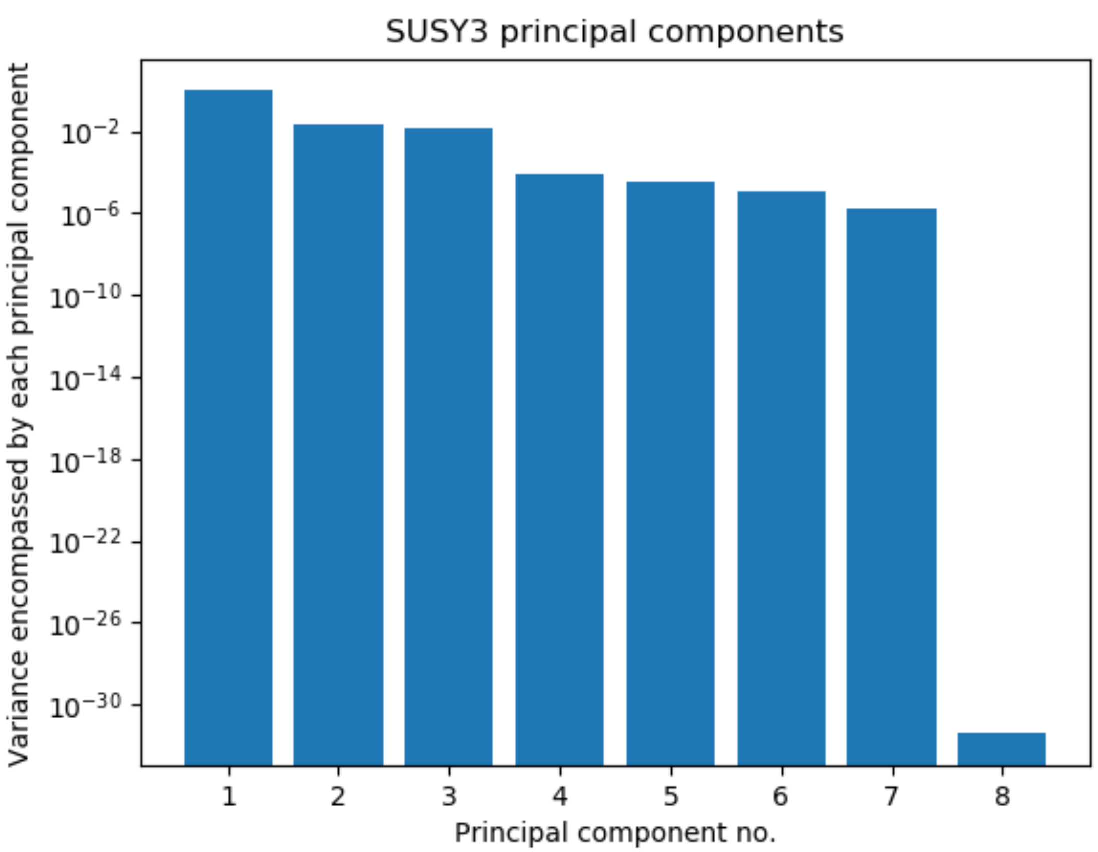
See e.g. Adamson, P. et al. in NOvA. Phys. Rev. D 96,

PAX (Processor for Analysing Xenon)









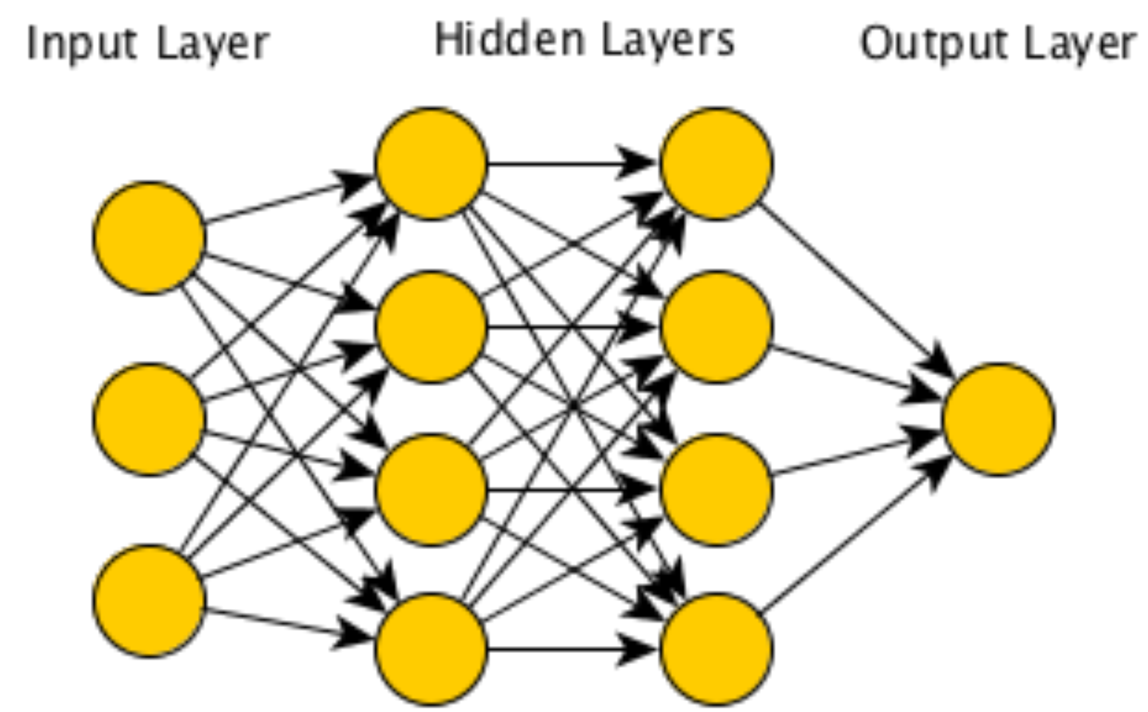
Background PCA correlations								
	p_T^{j1}	p_T^{j2}	η_{j_1}	η_{j_2}	$\Delta\phi_{j_1 j_2}$	MET	$\Delta\phi^{j1}_{\mathrm{MET}}$	$\Delta\phi^{j2}_{\mathrm{MET}}$
PC-1	0.67	0.41	-0.01	-0.01	-0.01	0.62	0.00	-0.00
PC-2	0.00	0.00	0.01	0.00	0.45	-0.00	-0.77	-0.45
PC-3	-0.00	-0.00	0.09	0.10	-0.70	-0.00	0.00	-0.70
PC-4	-0.01	0.00	-0.70	-0.70	-0.09	-0.01	-0.01	-0.10
PC-5	-0.12	0.89	-0.01	0.02	-0.00	-0.45	0.00	0.00
PC-6	-0.01	0.01	0.71	-0.71	-0.01	-0.01	0.01	-0.00
PC-7	0.73	-0.23	0.00	0.00	0.00	-0.65	0.00	0.00
PC-8	0.00	0.00	-0.00	-0.00	0.54	-0.00	0.64	-0.54

SUSY3, $M_{\tilde{\chi}_1^0} = 300$ GeV PCA correlations								
	p_T^{j1}	p_T^{j2}	η_{j_1}	η_{j_2}	$\Delta\phi_{j_1 j_2}$	MET	$\Delta\phi^{j1}_{\mathrm{MET}}$	$\Delta\phi^{j2}_{\mathrm{MET}}$
PC-1	0.67	0.32	0.00	0.01	0.00	0.67	-0.00	-0.00
PC-2	-0.00	-0.01	0.01	-0.01	0.32	-0.00	-0.76	-0.57
PC-3	0.00	-0.00	-0.07	-0.06	-0.80	0.00	0.11	-0.58
PC-4	-0.01	0.02	0.70	0.70	-0.07	-0.01	0.01	-0.05
PC-5	-0.22	0.95	-0.00	-0.02	0.00	-0.23	-0.01	-0.00
PC-6	-0.00	0.01	-0.71	0.71	0.01	-0.00	-0.01	-0.00
PC-7	0.71	-0.01	-0.00	0.00	-0.00	-0.71	0.00	0.00
PC-8	0.00	0.00	0.00	-0.00	0.50	0.00	0.64	-0.58

Neural Network (NN): Basic Structure

Training sample, validation sample, test sample

- **Input layer nodes: set of observables(kinematical features)/images**
- **Number of hidden layers (shallow or deep NN)**
- **Output layer: predictions**



Train the network using training sample and make predictions for the test set