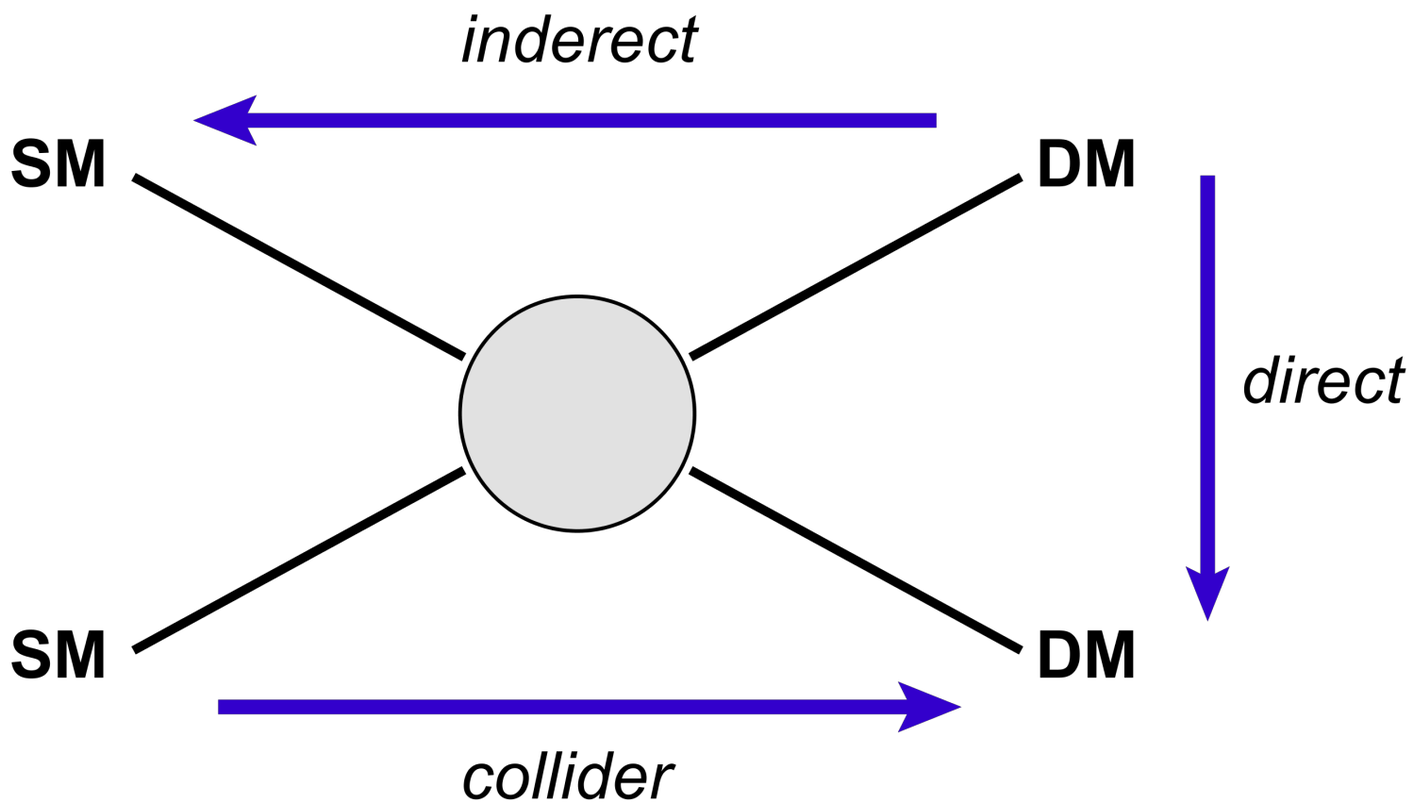
Collider searches for Dark Matter using Machine Learning

Charanjit Kaur Khosa and Michael Soughton
University of Sussex

Work in progress

In collaboration with Veronica Sanz

Introduction



DM searches channels at collider : Mono jet/W/top/photon ← Caterina Doglioni's talk

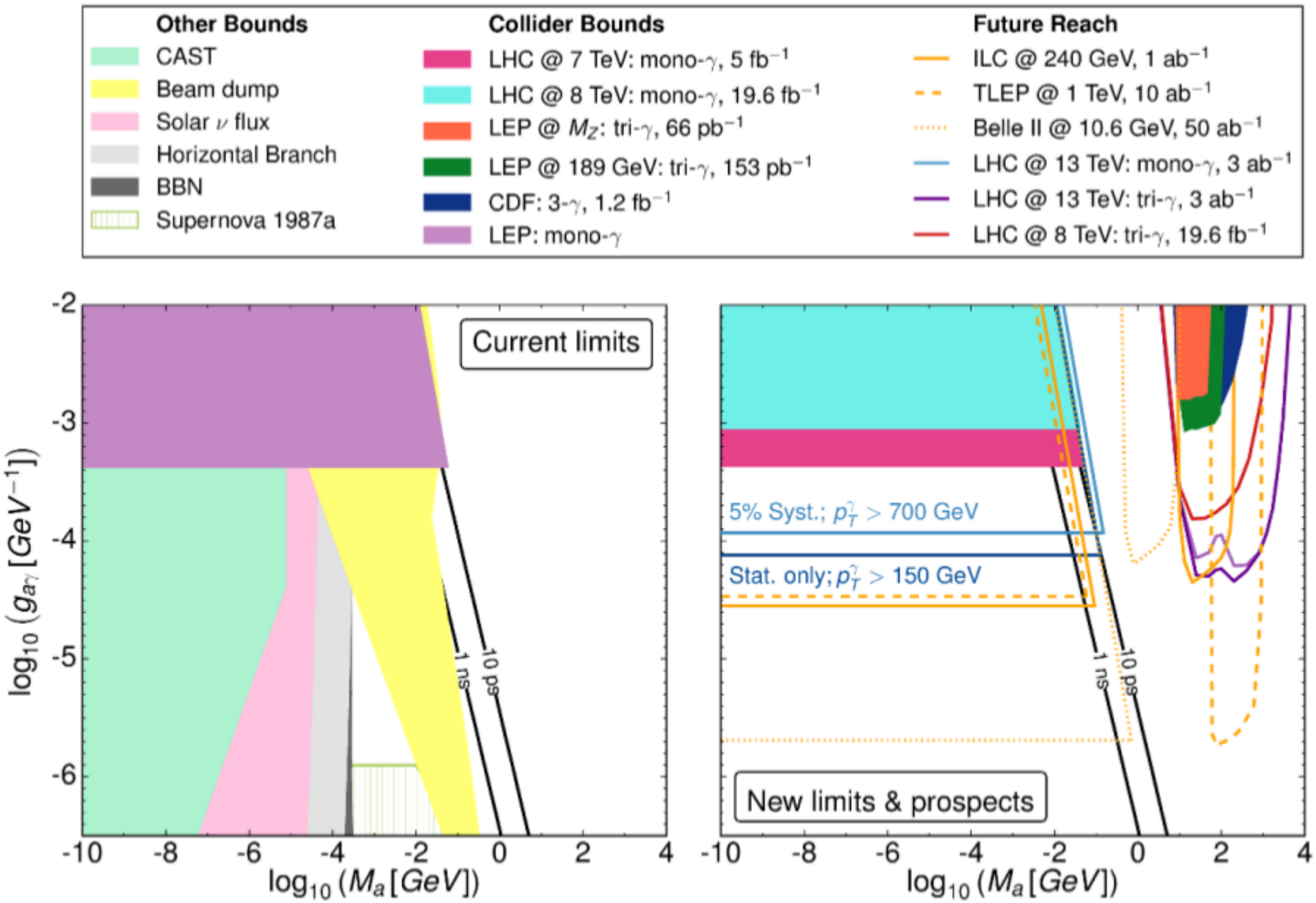
Signal : Monojet + MET **Model : Axion-like particles (ALPs)/WIMPs**

ALPs could look like MET if they decay outside the detector volume

Plan

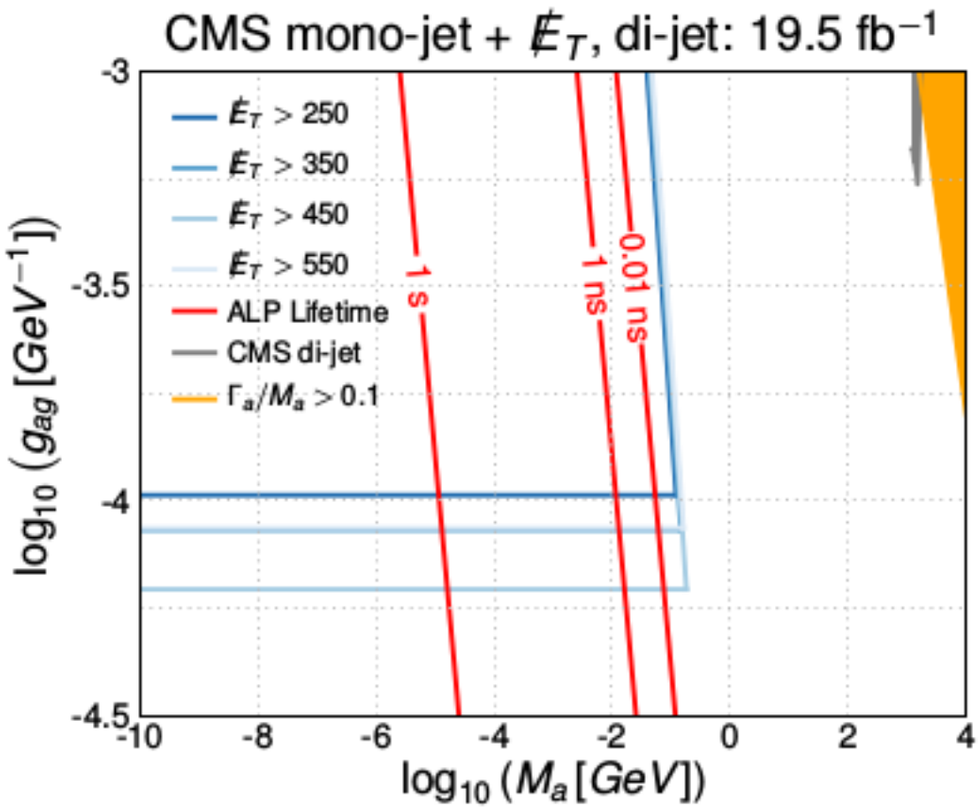
- Collider stable particles (Long lived particles)
- **First comparison:** Axions could be very light (collider stable) whereas typical DMs are not **(we expect distinction)**
- Multi-dimensional density information?
- Supervised to Unsupervised algorithms
- New physics?
- Anomaly detection

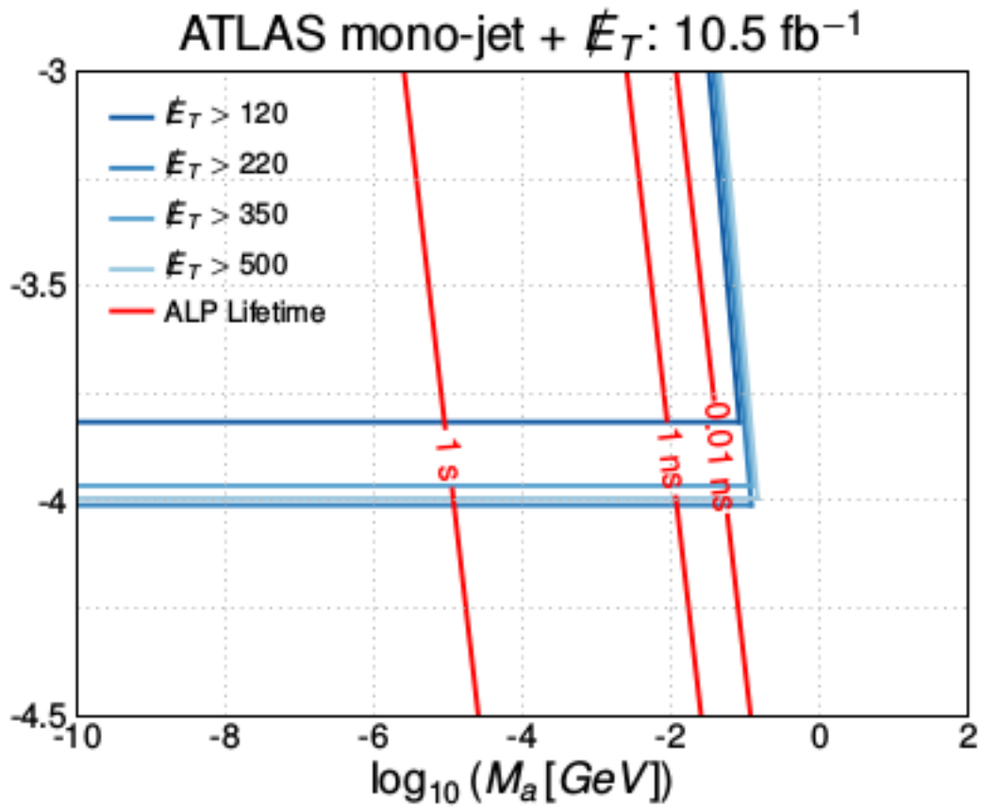
ALPs-Photon coupling



K. Mimasu and V. Sanz, JHEP06(2015)173

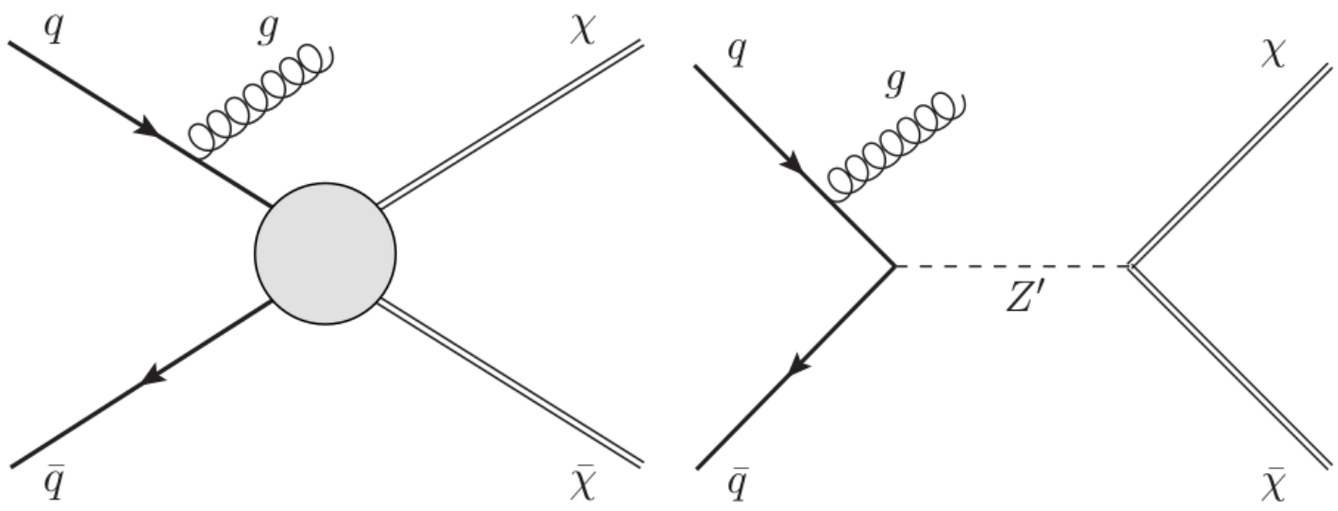
ALPs gluon coupling





K. Mimasu and V. Sanz, JHEP06(2015)173

Effective interactions: WIMPs with SM



$$\frac{m_q}{M^2}\chi^\dagger\chi\bar{q}q$$

See, for example : ATLAS analysis, Eur.Phys.J. C75 (2015) 299

Analysis set-up : mono-jet

BP1: ALPs $pp \to aj$

Madgraph_amc@NLO, FeynRules Models

$$\mathscr{L}_a = \frac{1}{2}\partial_\mu a\,\partial^\mu a - \frac{1}{2}M_a^2 a^2 - \frac{g_{a\gamma}}{4} a\,F_{\mu\nu}\tilde{F}^{\mu\nu} - \frac{g_{agg}}{2} a\,\mathbf{Tr}\left[G_{\mu\nu}\tilde{G}^{\mu\nu}\right] + \sum_\psi g_a^\psi\, m_\psi\, a\bar{\psi}\gamma^5\psi$$

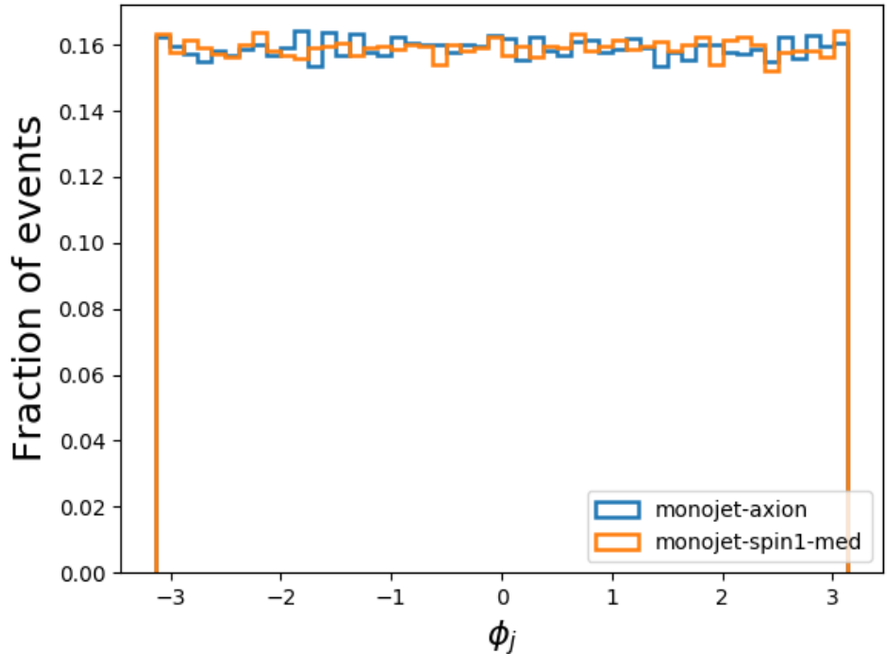
BP2 : Simplified dark matter models : heavy mediator decays to DM

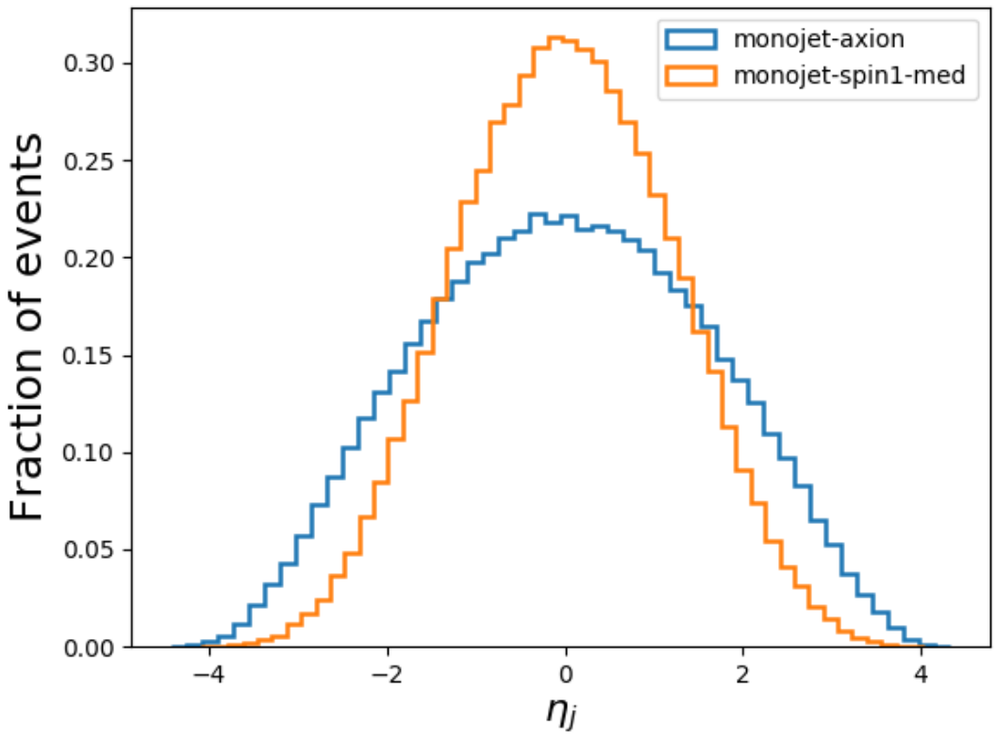
$pp \to \chi_d \chi_d j$ $\qquad$ $\frac{m_q}{M^2}\chi^\dagger\chi\bar{q}q$

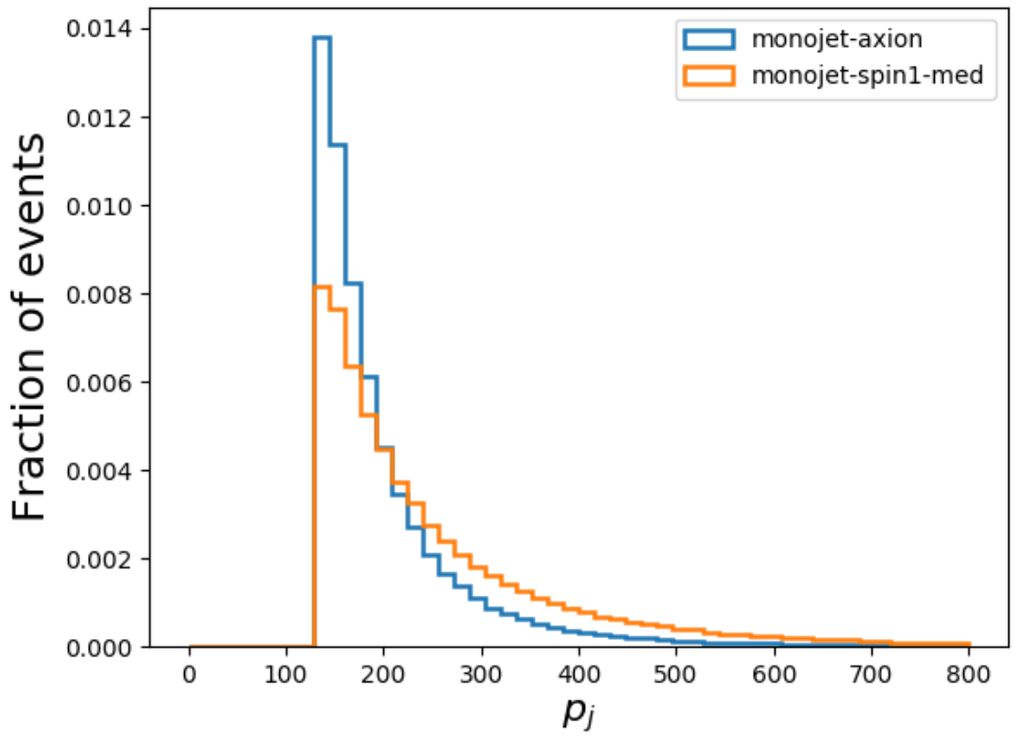
$p_T^j > 130\,GeV$

Assumption : very well understanding of the background $\qquad$ $p_T^j(MET), \eta_j, \phi_j$

1D distributions







Preliminary Results

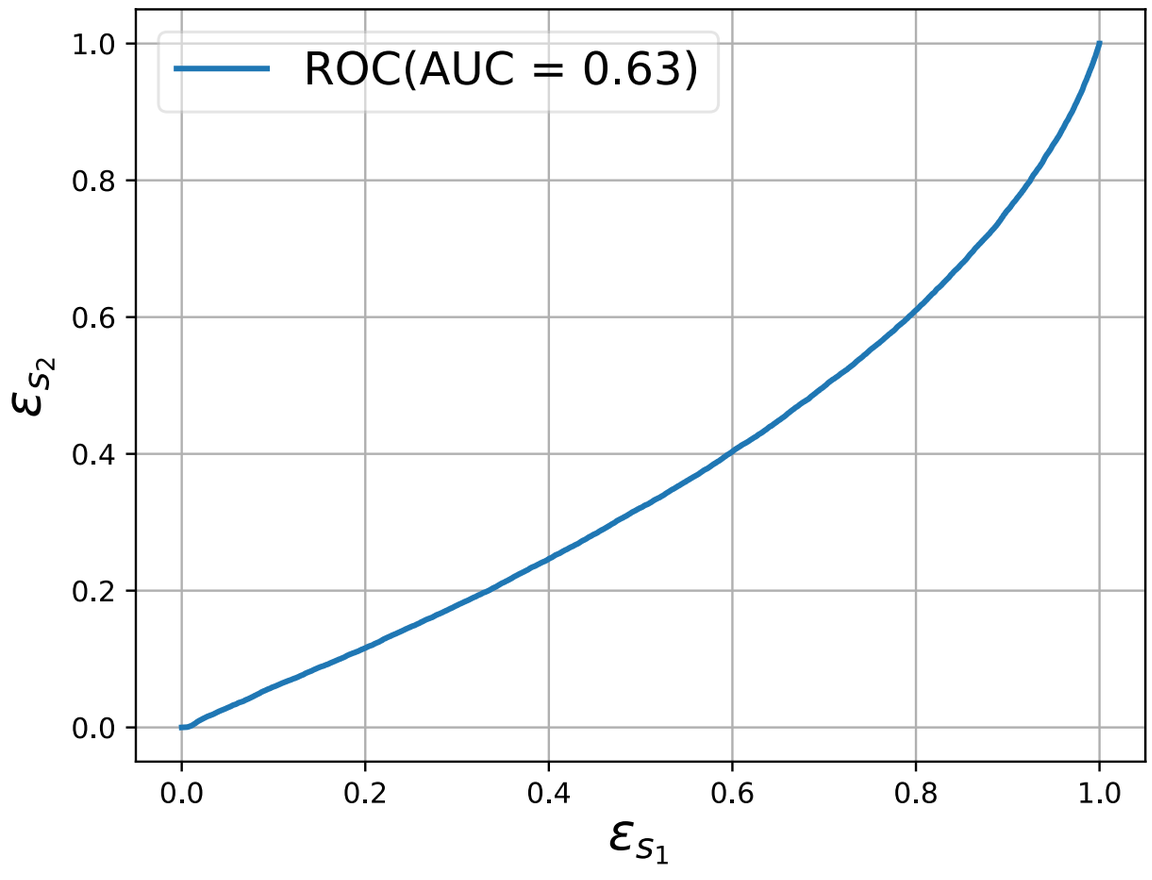
Supervised Learning (at the moment) before trying the semi-supervised/unsupervised algorithms

Logistic regression

Features : p_T^j, η_j, ϕ_j

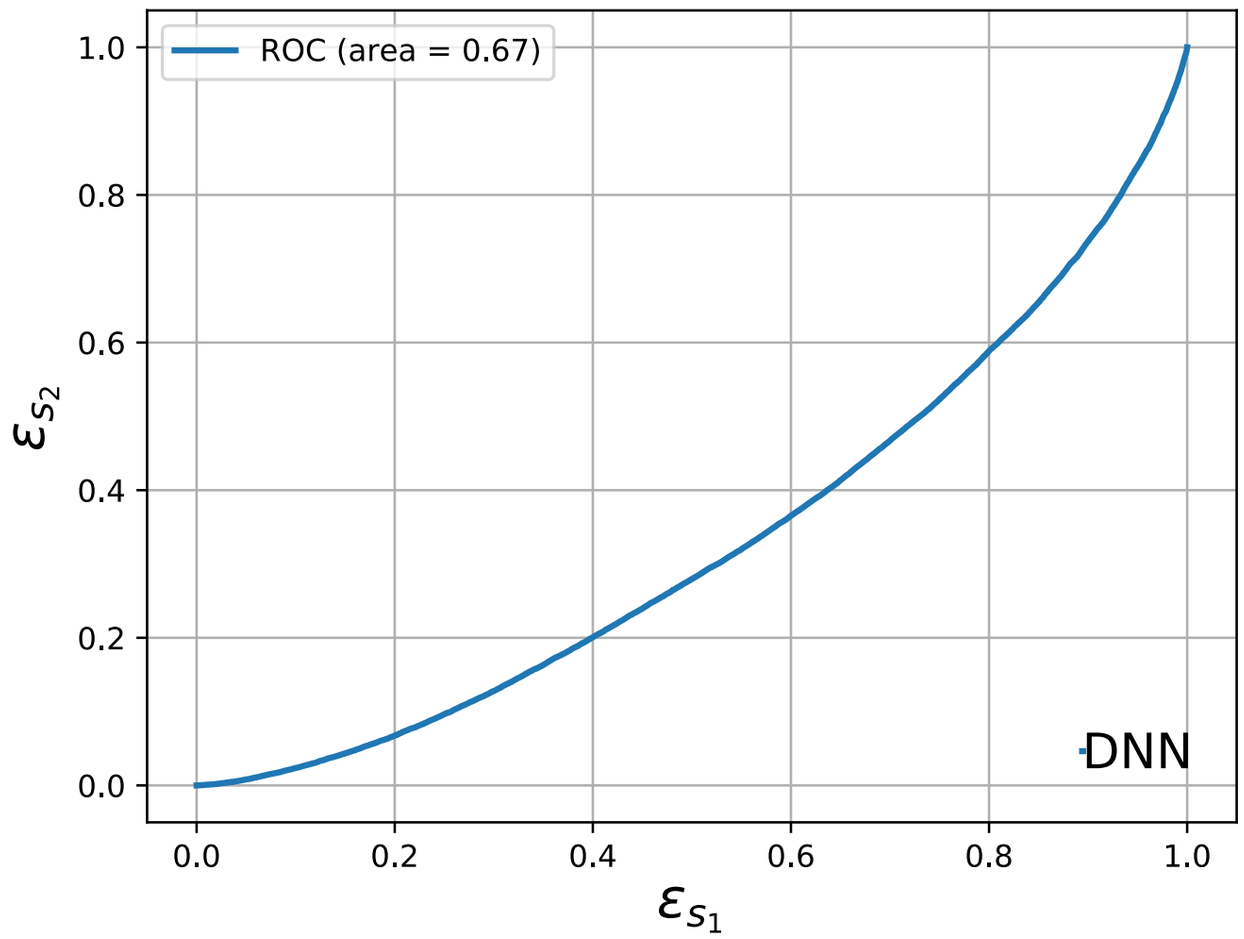
ϵ_{S_1} **: Correct identification of ALPs**

ϵ_{S_2} **: BP2 events labelled as ALPs**



DNN

- Hidden Layers:(after 5, same performance)
- Epochs : (>100, same)
- Activation function : ReLU
- Loss function : Binary cross-entropy



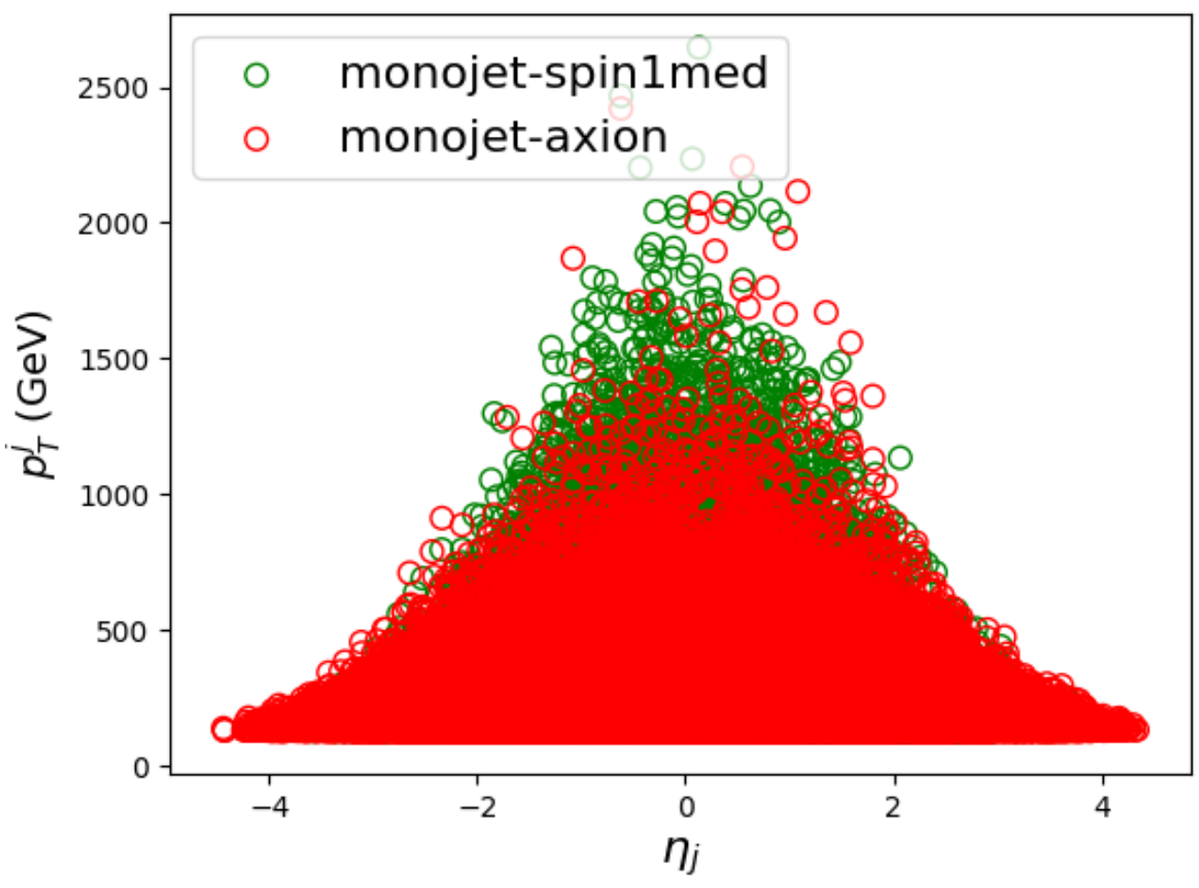
Multi-dimensional probability information

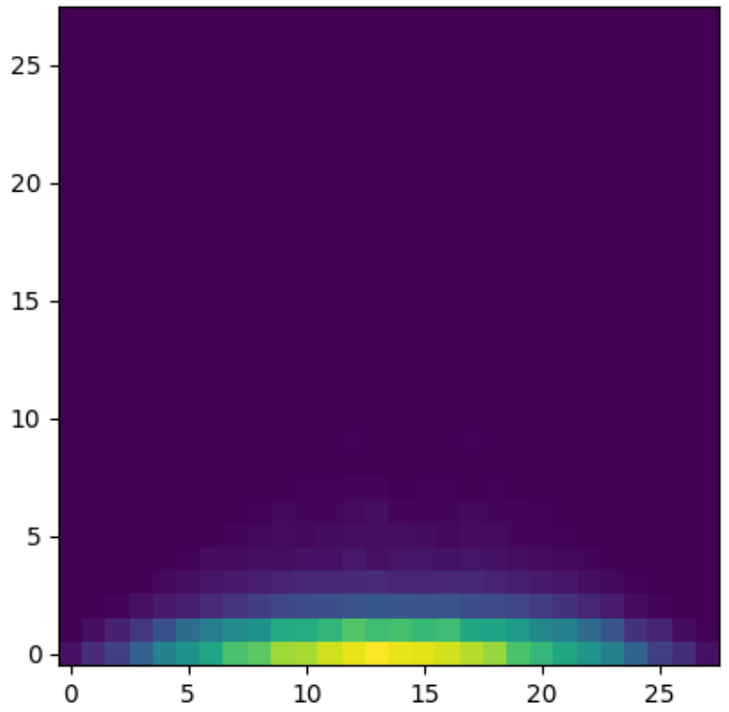
2D distributions : mono-jet

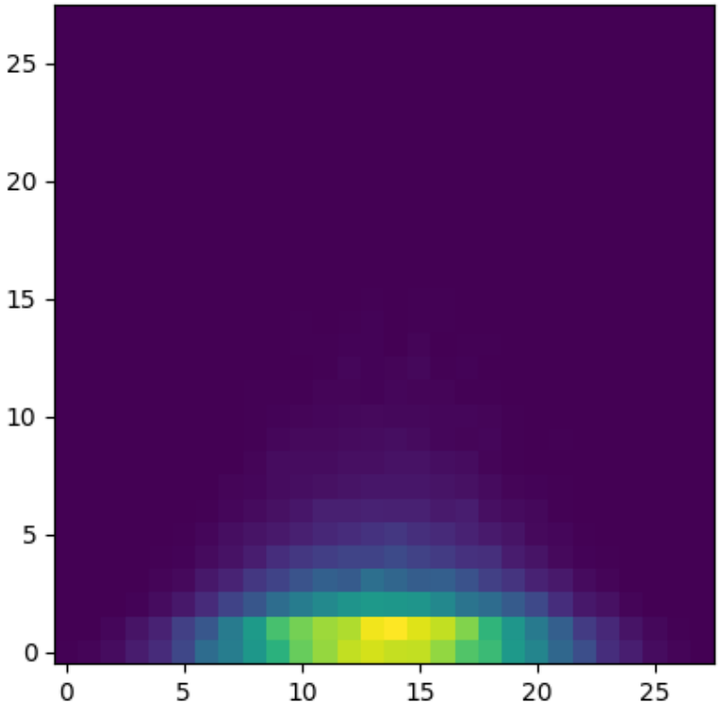
Binning: 29×29 (50K events per image)

$\eta_j : [-4,4]$

$p_T^j : [130,2000] GeV$







Data processing : events to images

- Information of correlation between variables exists in 2D histograms

- Little point (yet) in using 3D, 4D,..histograms

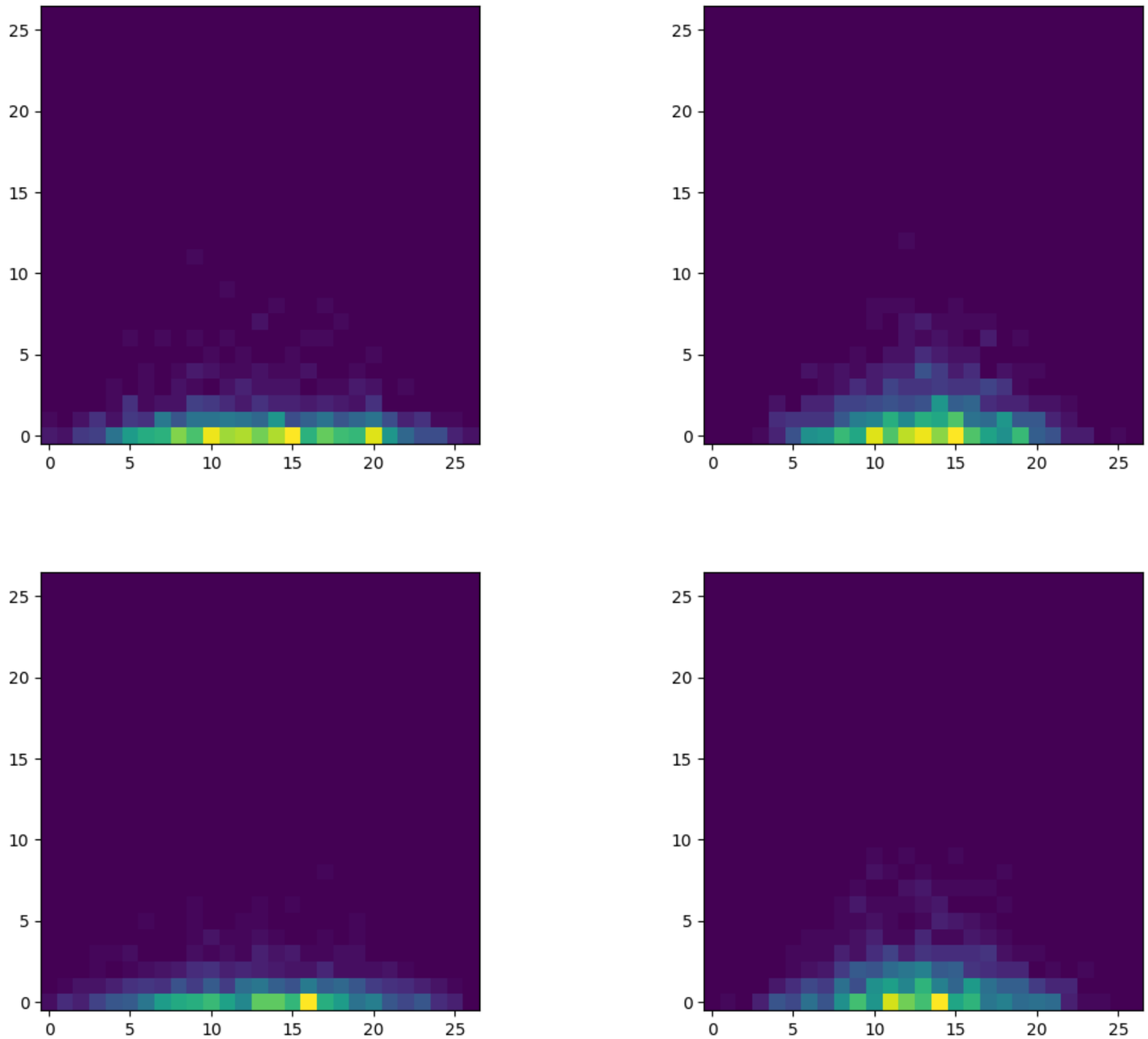
- Able to run CNN

- N=1000 no of events per image

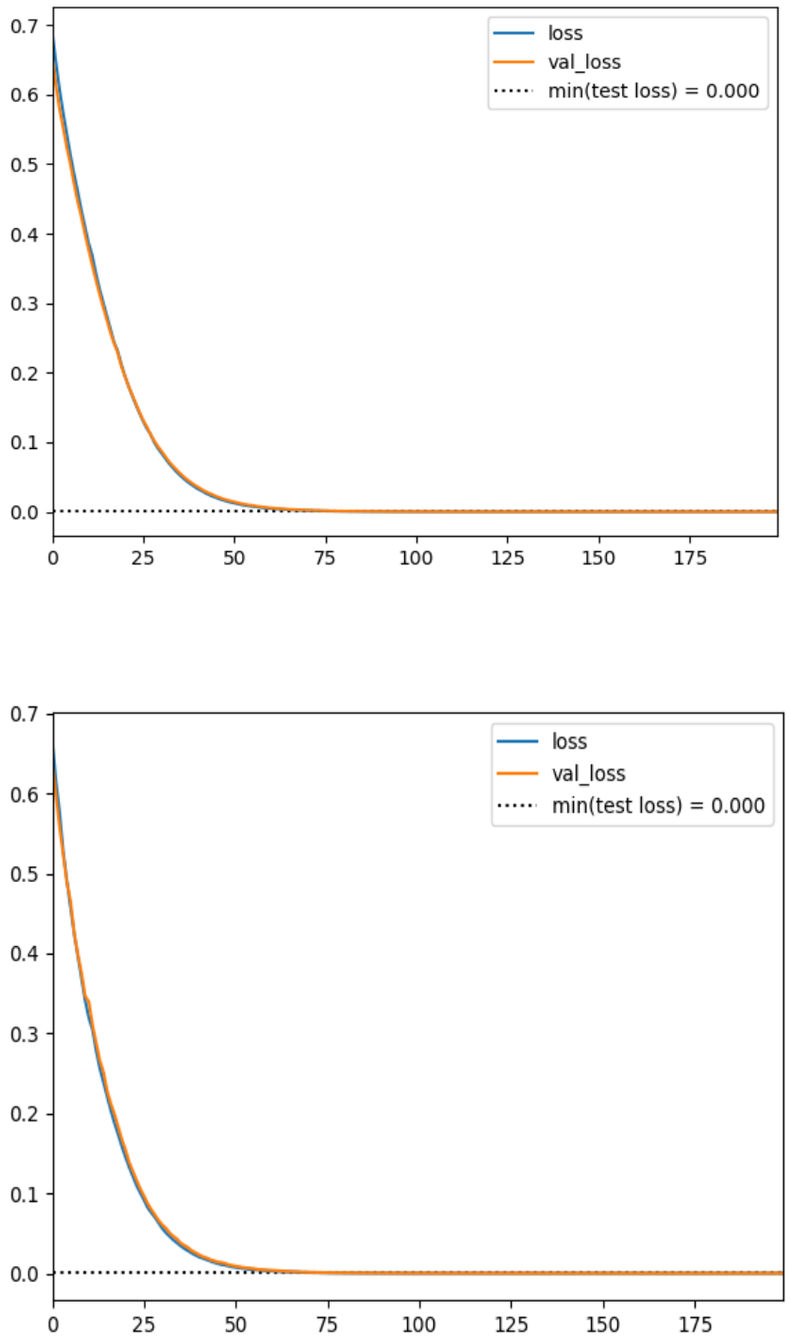
- Training : Validation : Test Sample = 320:40:40 images

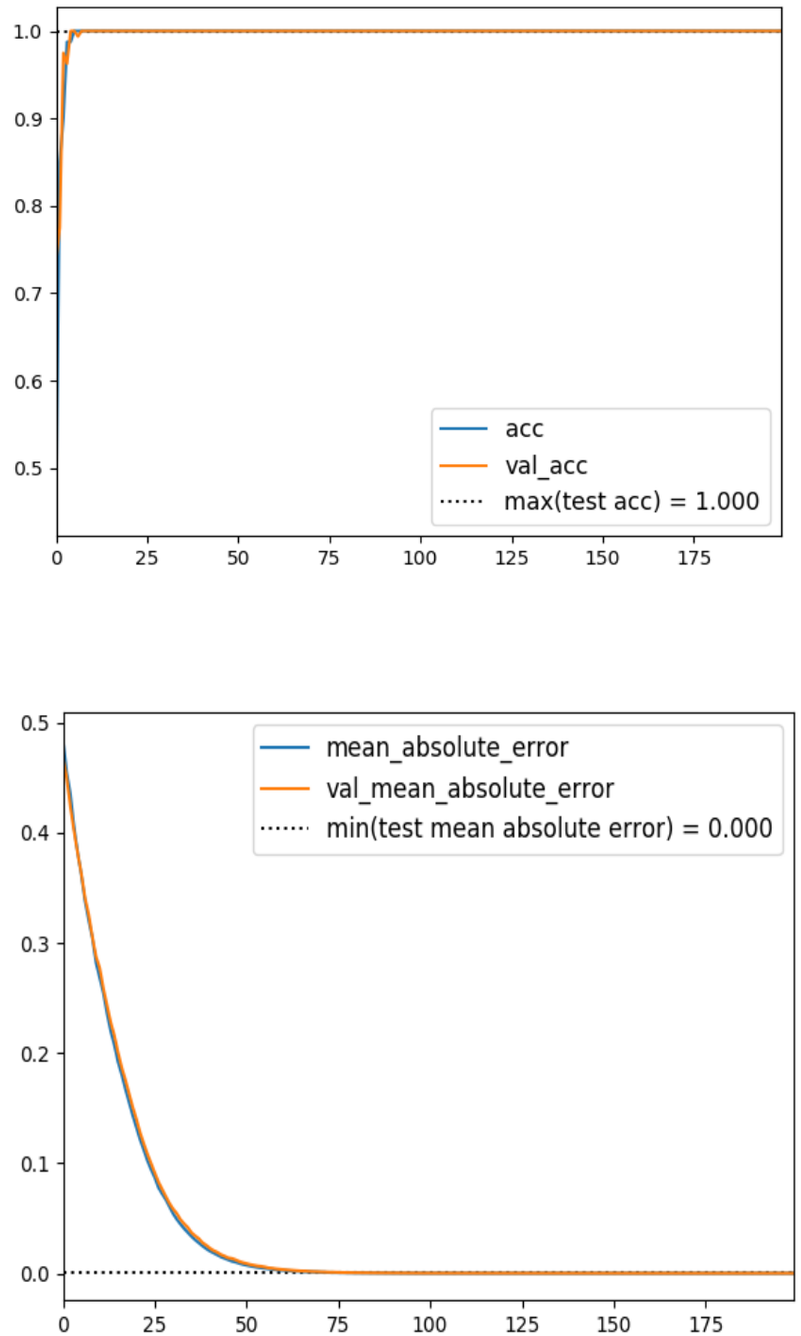
- Hidden layers: 2, Loss function: Binary cross-section, Activation functions : ReLU

Training the DNN (p_T:[130,1000]GeV)

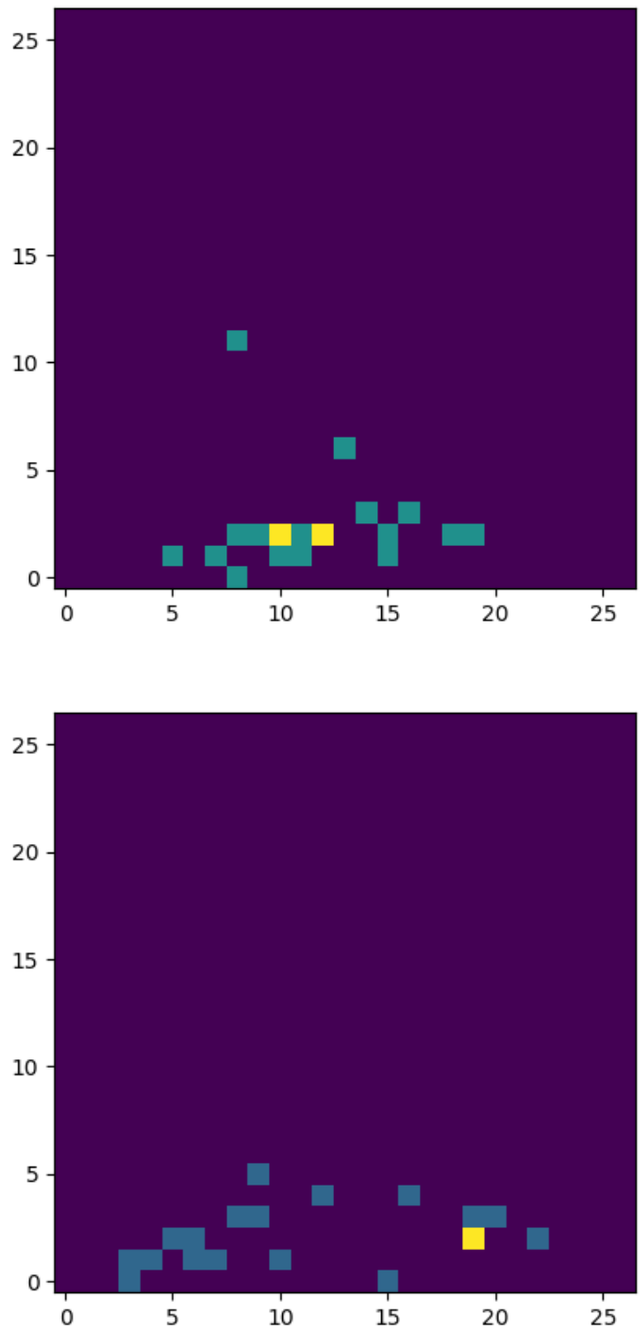
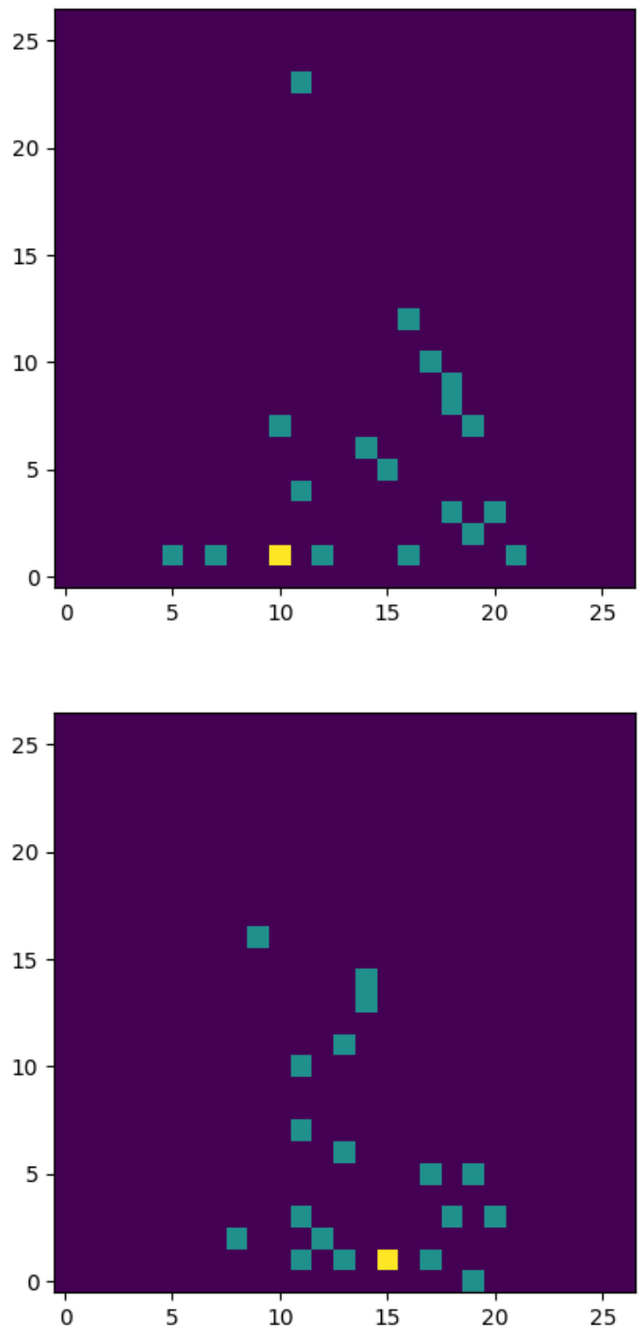


DNN training rates

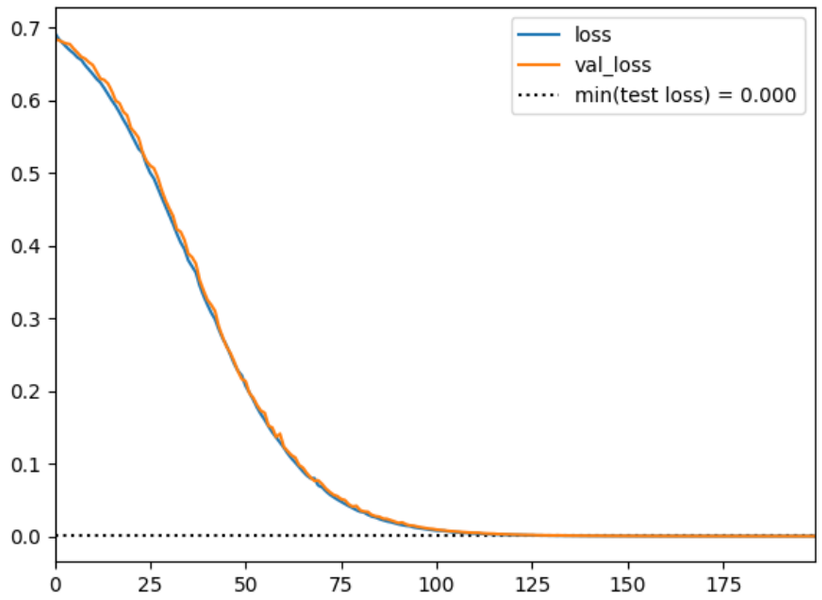


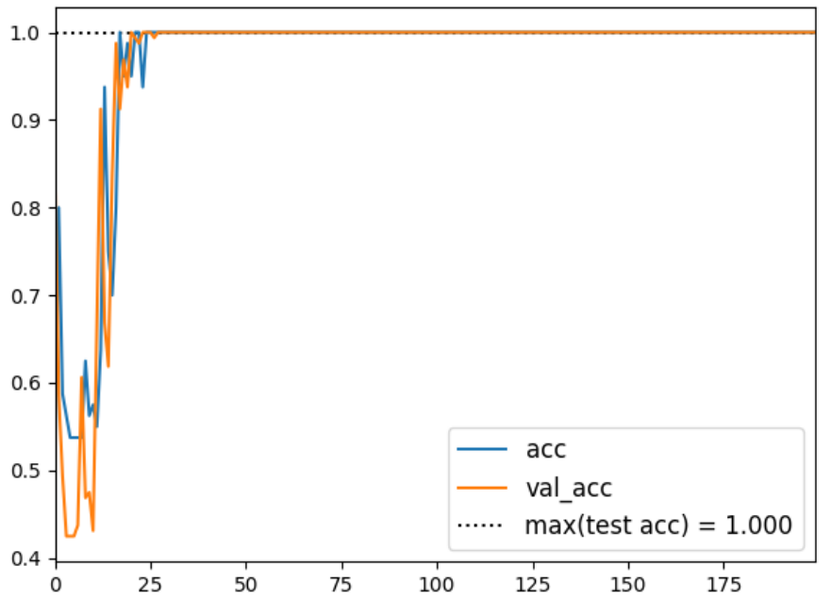


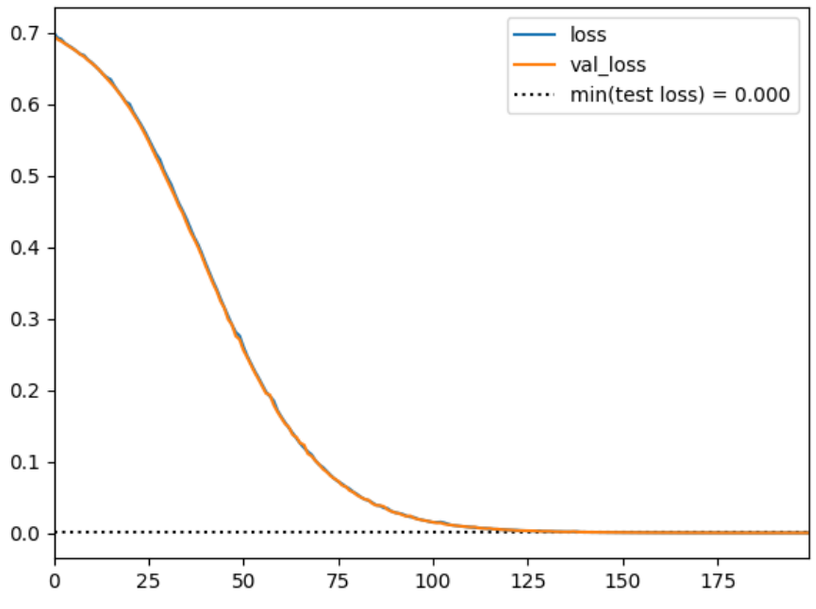
Testing against events with 20 events/image

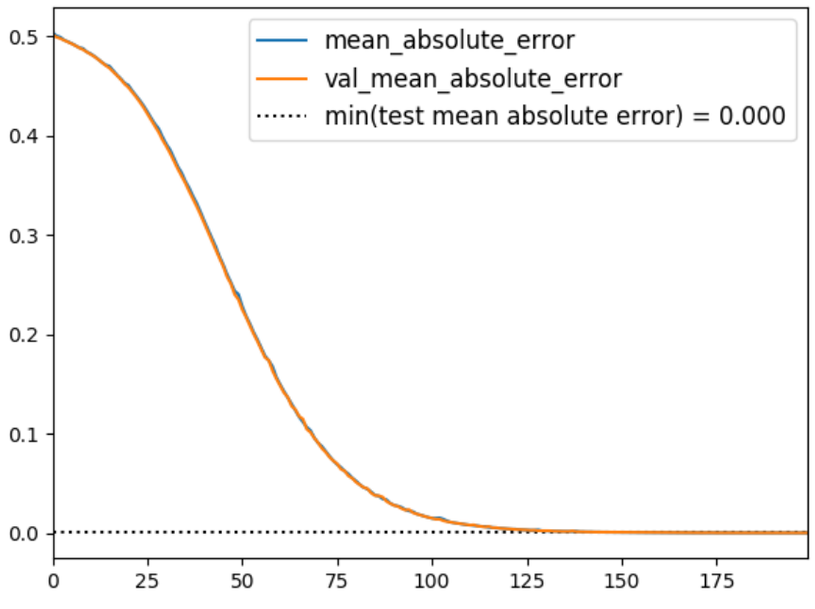



CNN training rates









Comparison of scores for density plots

- DNN accuracies:
  - Training: ~100%
  - Test: ~100%
  - Against data of 20 events/image: ~89.6%
- DNN mean absolute errors:
  - Training: ~2.11×10^{-5}
  - Test: ~1.36×10^{-5}
  - Against data of 20 events/image: ~0.165
- CNN accuracies:
  - Training: ~100%
  - Test: ~100%
  - Against data of 20 events/image: ~83.7%
- CNN mean absolute errors:
  - Training: ~2.14×10^{-5}
  - Test: ~3.29×10^{-5}
  - Against data of 20 events/image: ~0.332

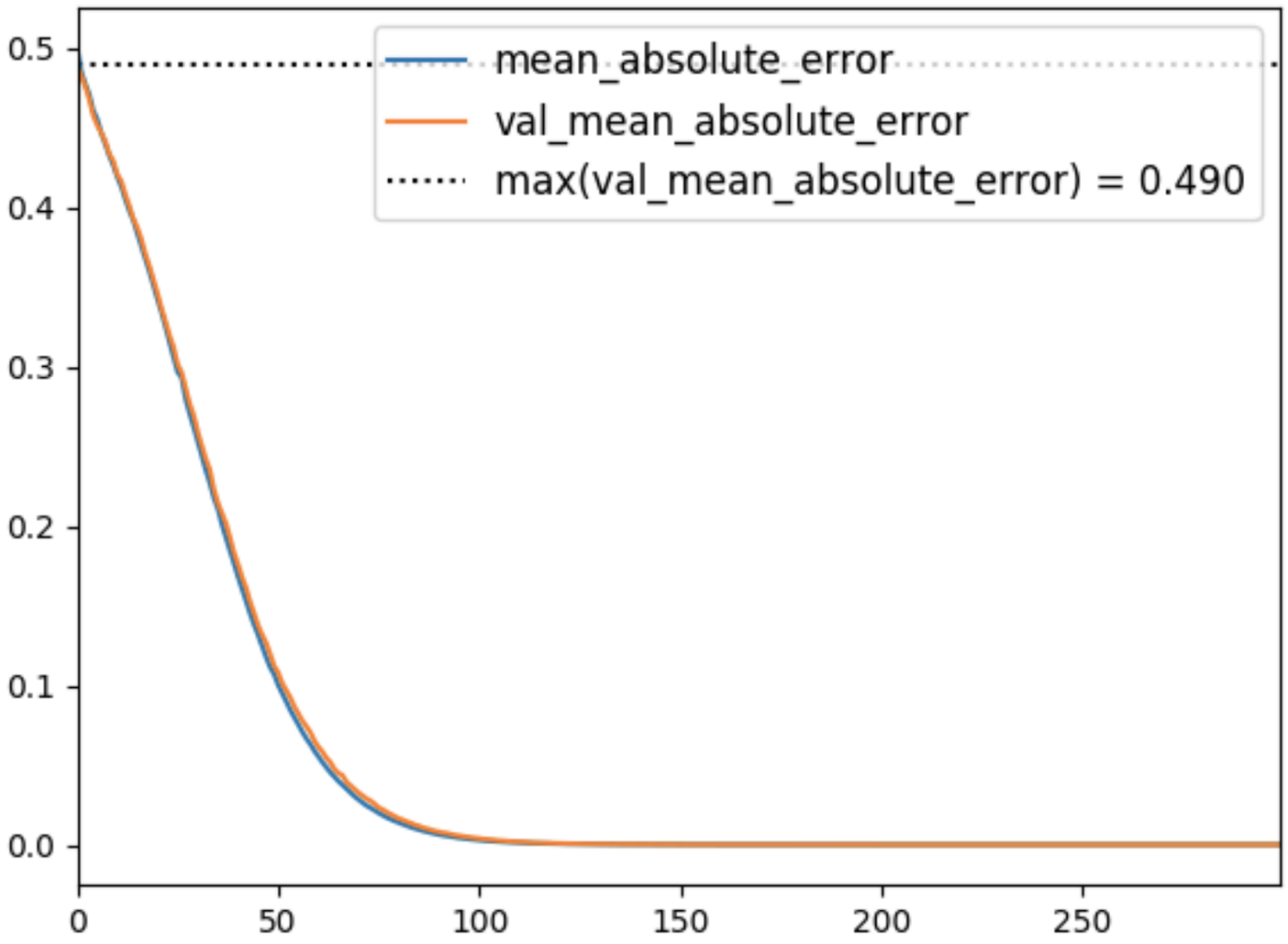
Summary and next steps

- **Parton level event generation for ALPs and simplified models**
- **Able to distinguish between signals when only a small number of events are present**
- **Looking to explore further the relations of events/image between datasets: is it a correct approach ?**
- **Compare to expected number of events/timeframe from LHC – requires knowledge of cross sections and model parameters**
- **Analysis with different DM/collider stable models and incorporate showering/hadronisation, detector effects, NLO effects & SM background.** $(p_T^{j_1}, p_T^{j_2}, \eta_{j_1}, \eta_{j_2}, MET, \Delta\phi_{jj}, \Delta\phi_{METj_1}, \Delta\phi_{METj_2})$
- **Move towards unsupervised learning algorithms**

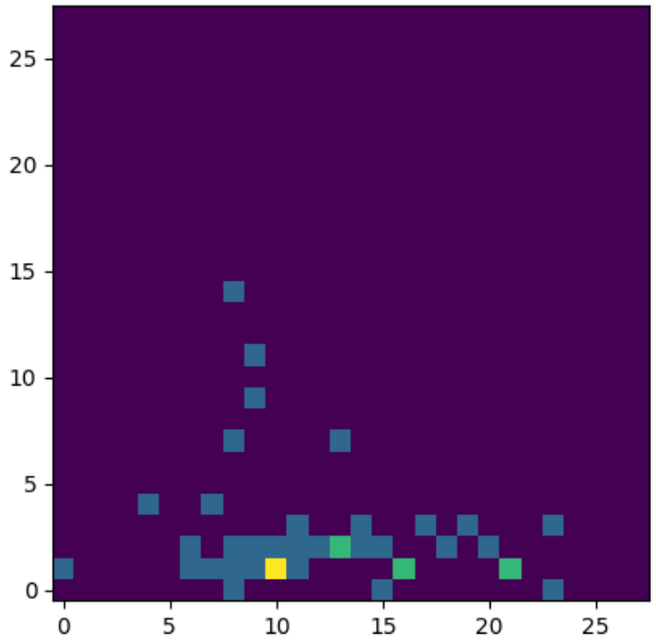
Thanks

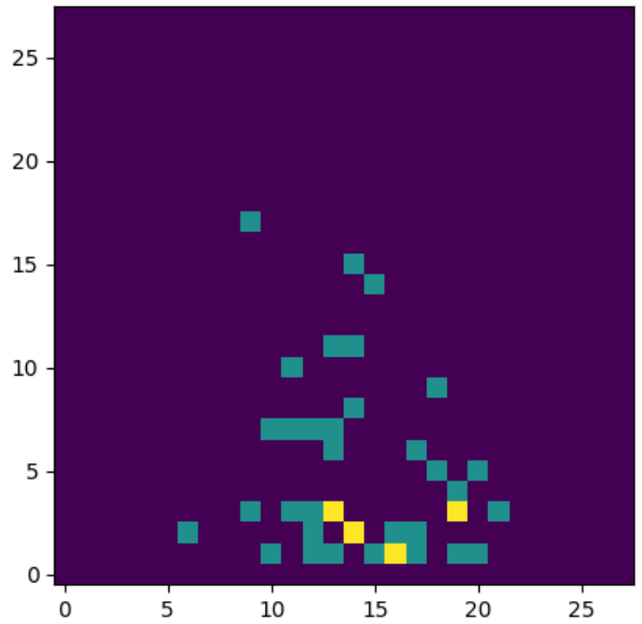
Suggestions/Questions ?

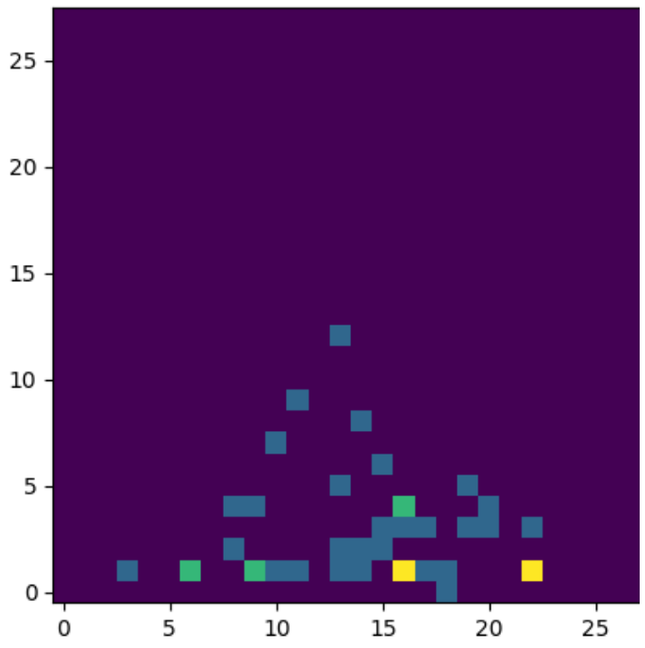
Small batch size

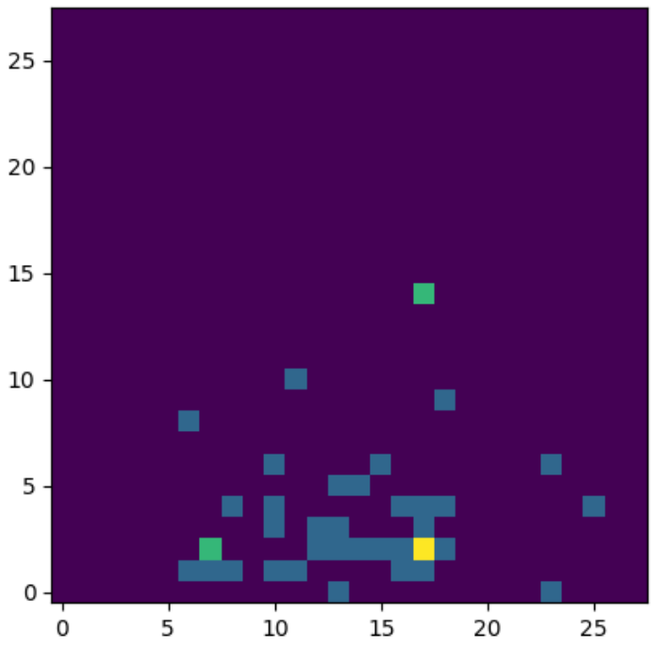


Data Sample 1 (p_T:[130,1000]GeV)





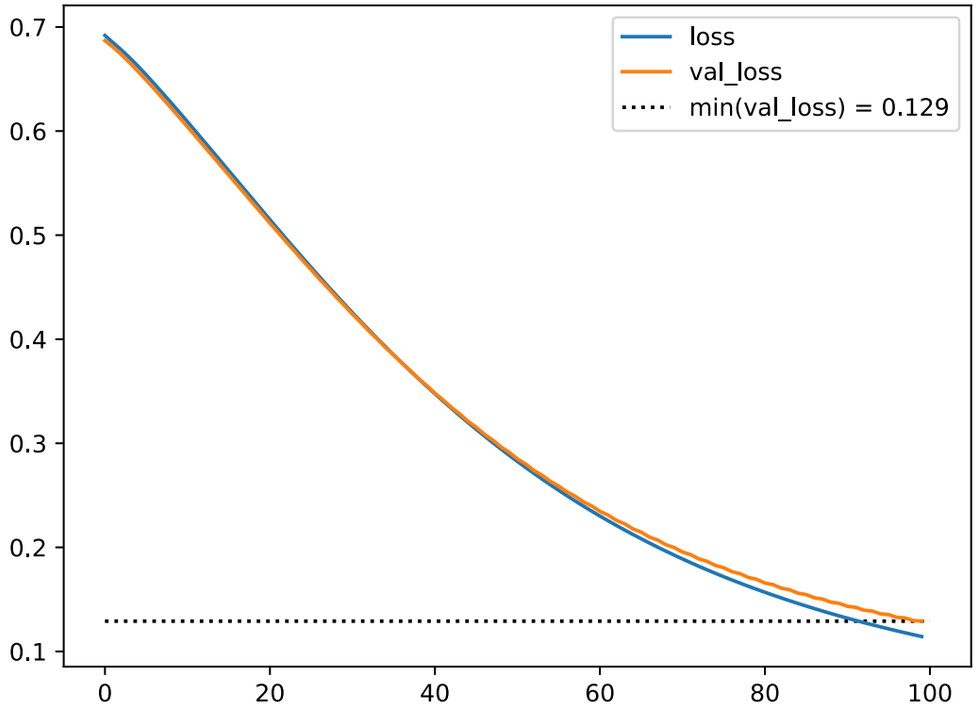


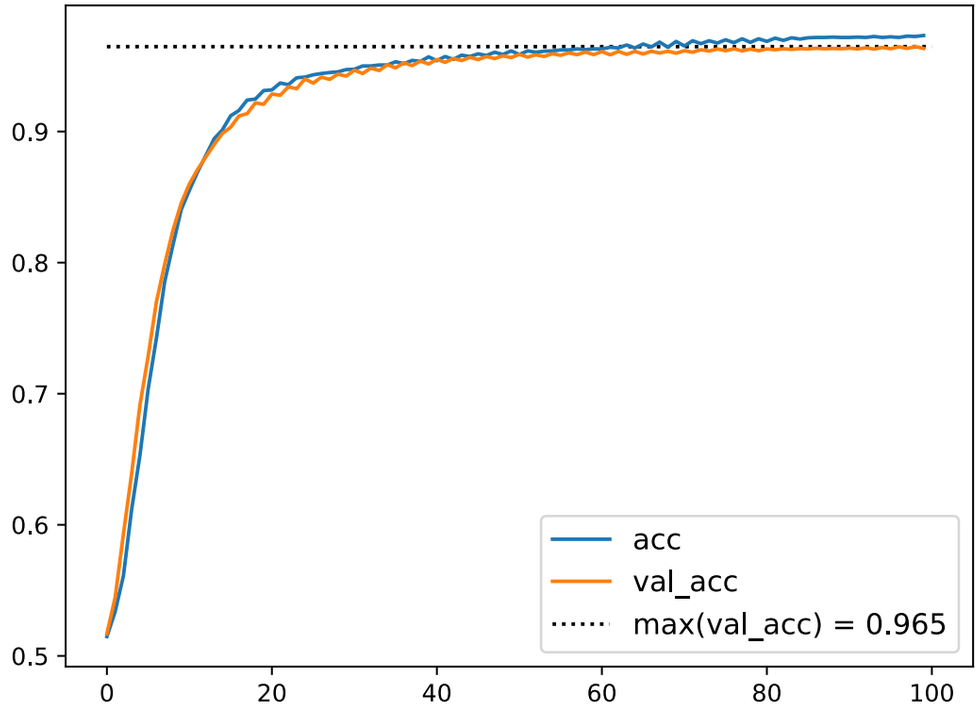


DNN (image recognition)

2D images : information of density is 1D array inout for NN

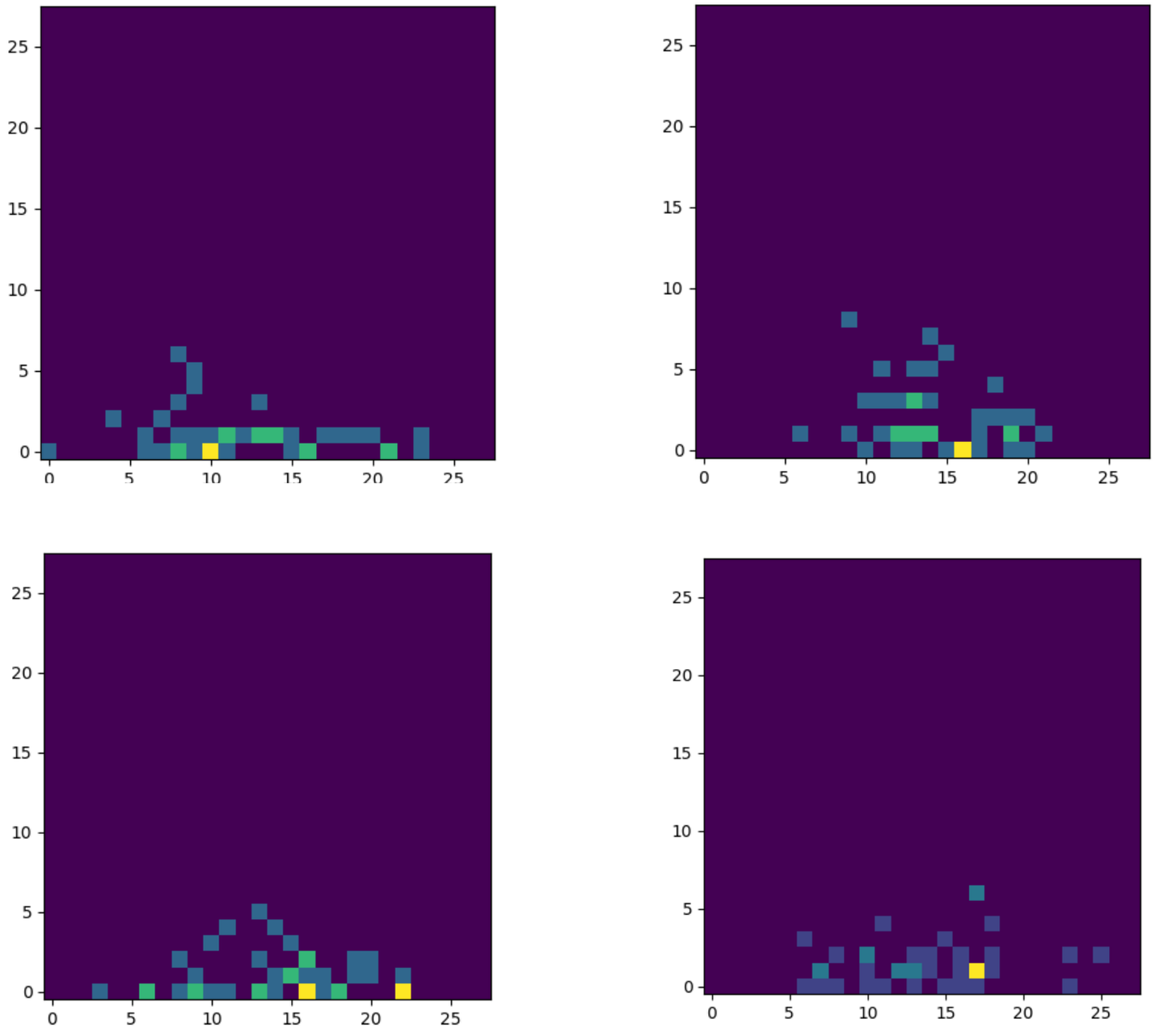
40 events per image





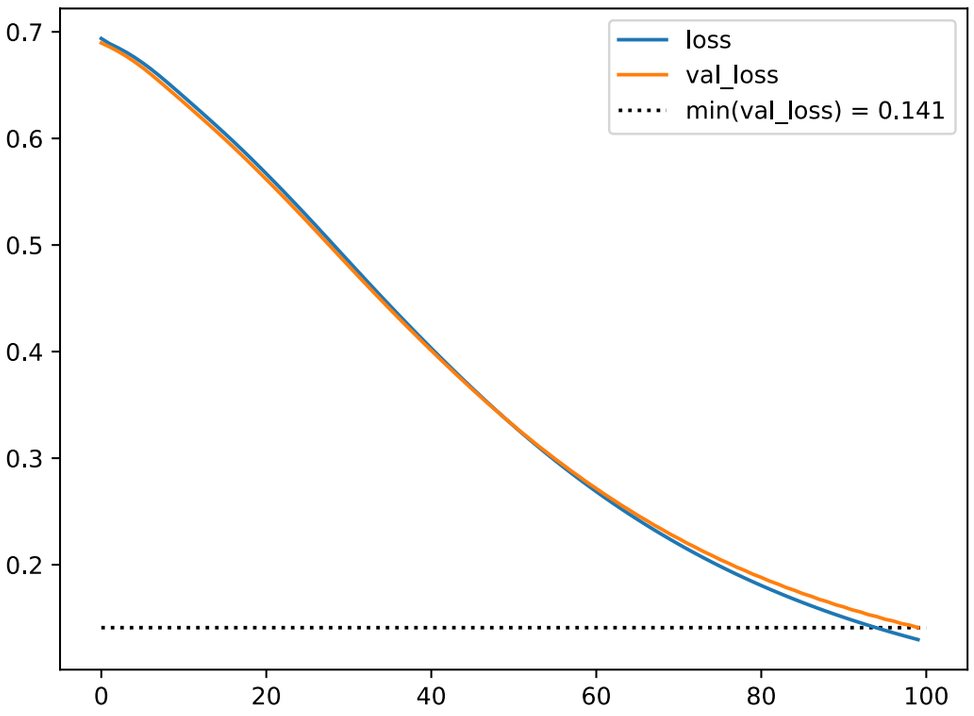
Hidden Layers : 2, Loss function : Binary cross-entropy, Activation function : ReLU

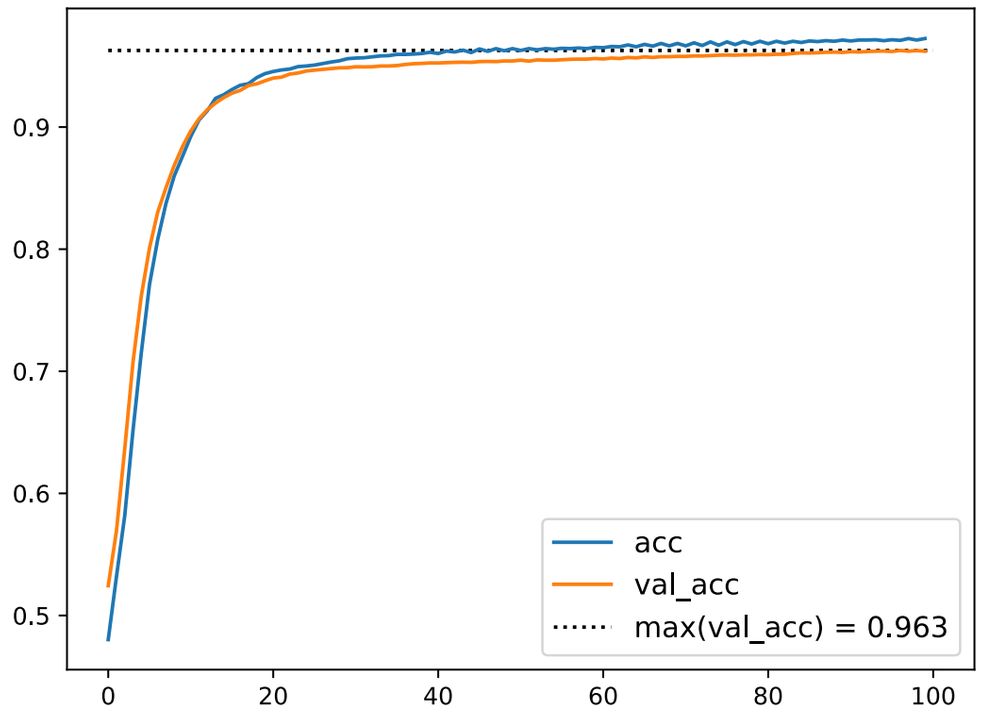
Data Sample 2 (P_T :[130,2000]GeV)



Algorithm has an information of rare events also

Same Performance ...





Dropouts ?