Search for $H^{\pm}$ boson

DESY Summer Student Programme 2019

Artem Nigoian

Supervisors: Dirk Kruecker, David Brunner

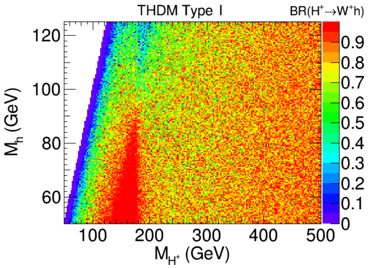
Deutsches Elektronen-Synchrotron

7.8.2019

Two higgs doublet model

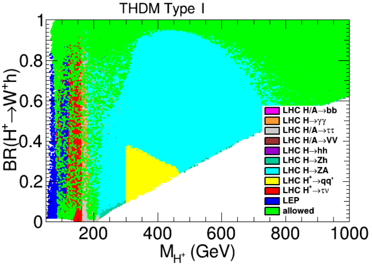
Model specification:

- Type I ($\Phi_1 \to -\Phi_1$) (no FCNC)
- $m_H = 125.09$ GeV
- $\sin(\alpha - \beta) = 0 \to m_h < m_H$
- $m_{H^\pm} = m_A, \tan(\beta) = 3$

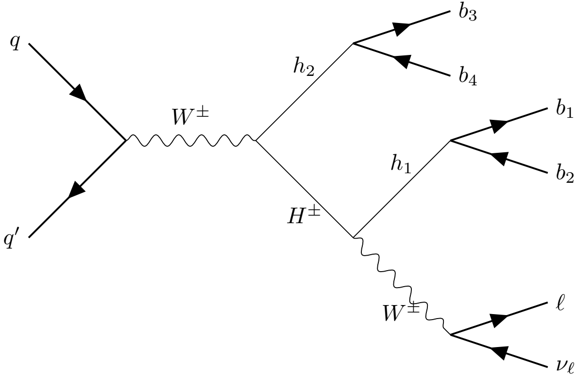


Features of this specification:

- High BR($H^\pm \to W^\pm h$) (first plot)
- h decays to fermions (b, τ)
- Large unexcluded region for high BR($H^\pm \to W^\pm h$) (second plot)

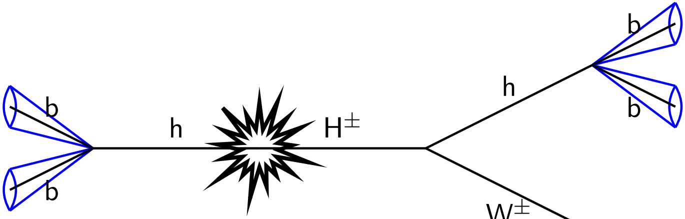


Feynman diagram of analysis process

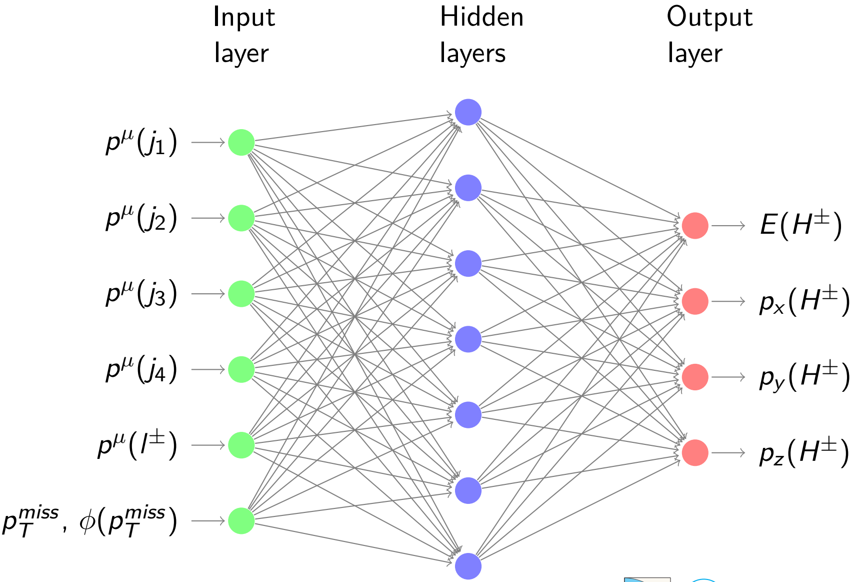


Signal jet kinematic

Non-boosted topology:



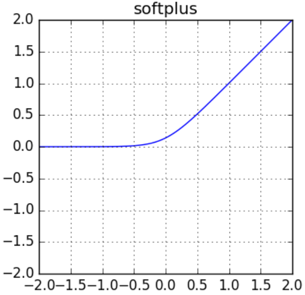
Neural network for $m(H^{\pm})$ reconstruction



Parameters of the neural network

During the network's hyperparameters tuning (random search in the grid) the following set of parameters turned out to be optimal:

→ Number of hidden layers: 5
→ Nodes in each hidden layer: 200
→ Batch size: 100
→ Activation function for nodes in hidden layers: softplus

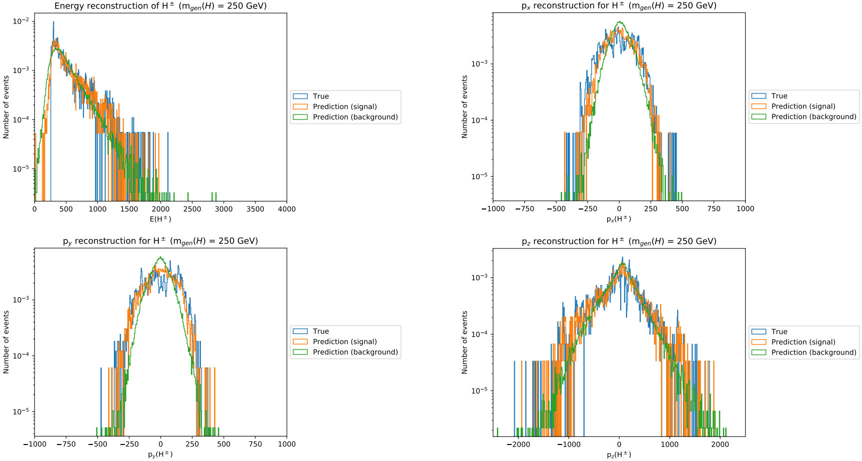


Data for training

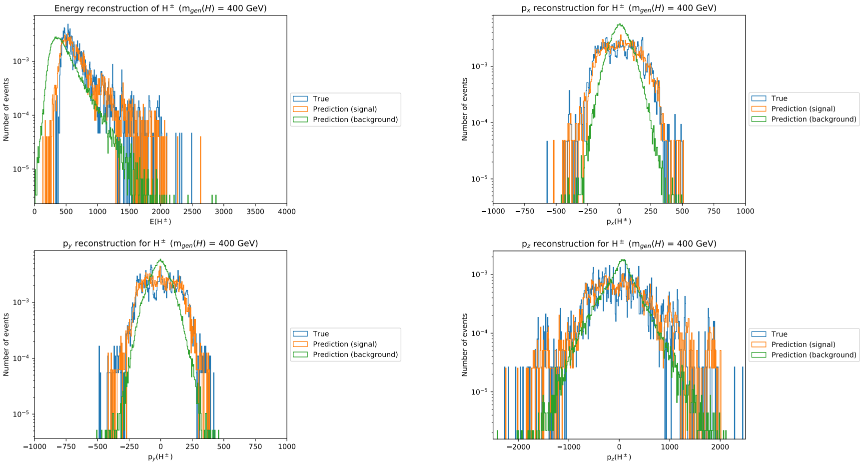
MC-generated events with $m(h_1) = 100$ GeV and $m(H^{\pm}) = 200, 250, ..., 600$ GeV were used for training

- $\rightarrow$ $\sim 3 \cdot 10^5$ events
- $\rightarrow$ unbalanced data (from $\sim 36 \cdot 10^3$ events for $m(H^{\pm}) = 200$ GeV to $\sim 14 \cdot 10^3$ events for $m(H^{\pm}) = 600$ GeV) $\rightarrow$ using *weighted* MSE as a loss function
- $\rightarrow$ training : test : validation $= 81 : 10 : 9$

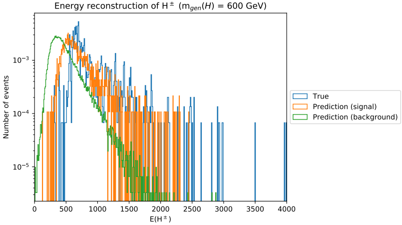
Results (4-momenta, $m^{gen}(H^{\pm}) = 250$ GeV)

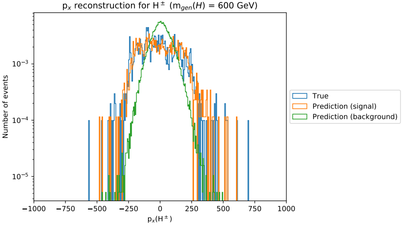


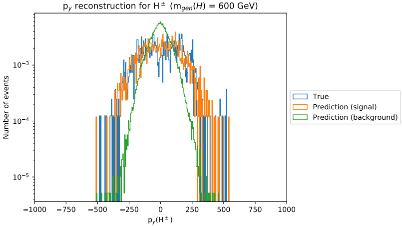
Results (4-momenta, $m^{gen}(H^{\pm}) = 400$ GeV)

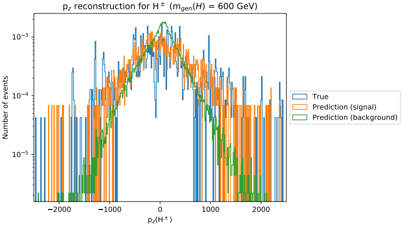


Results (4-momenta, $m^{gen}(H^{\pm}) = 600$ GeV)

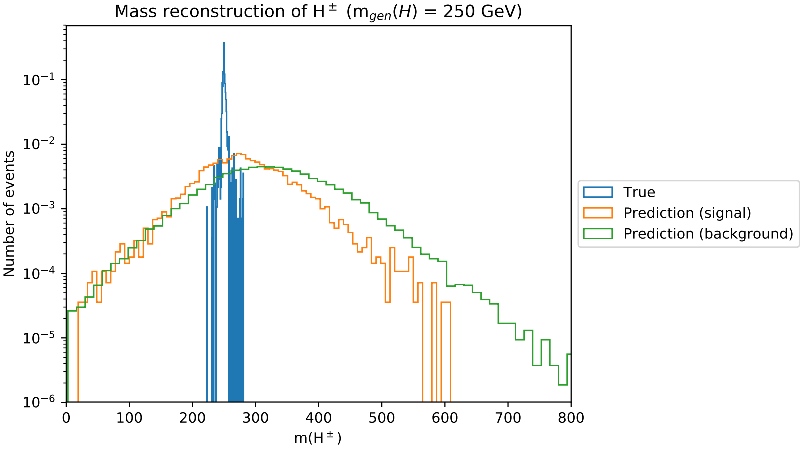




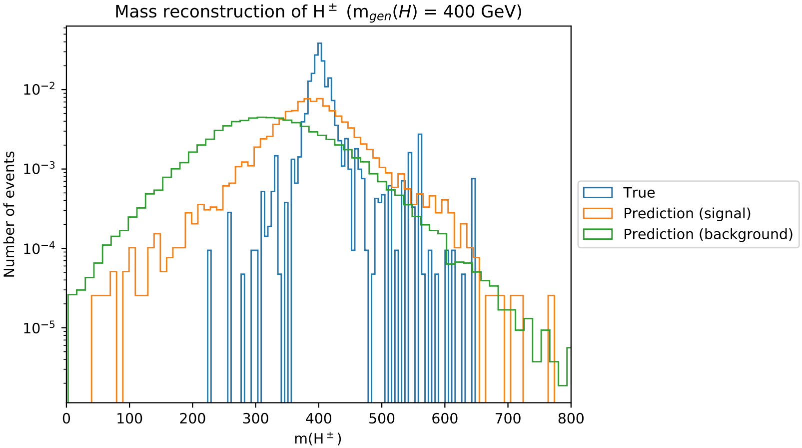




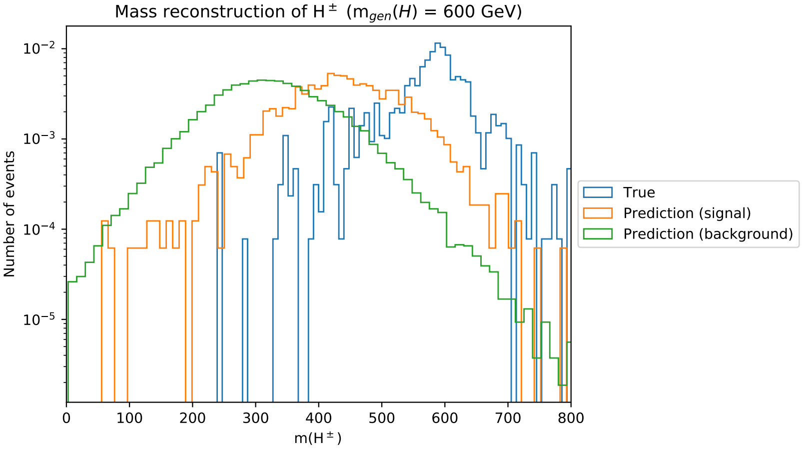
Results ($m^{pred}(H^{\pm})$ for $m^{gen}(H^{\pm}) = 250$ GeV)



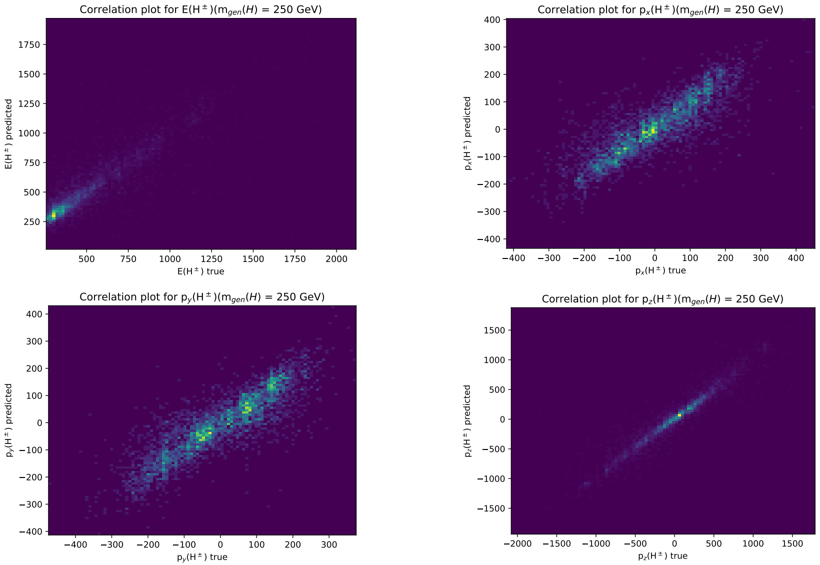
Results ($m^{pred}(H^{\pm})$ for $m^{gen}(H^{\pm}) = 400$ GeV)



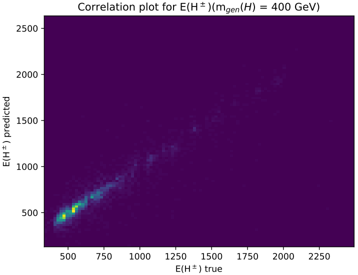
Results ($m^{pred}(H^{\pm})$ for $m^{gen}(H^{\pm}) = 600$ GeV)

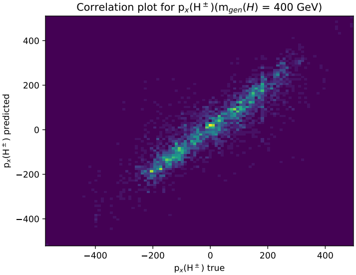


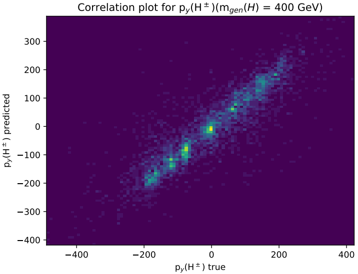
Correlation plots ($m^{gen}(H^{\pm}) = 250$ GeV)

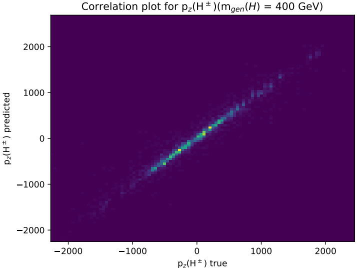


Correlation plots ($m^{gen}(H^{\pm}) = 400$ GeV)

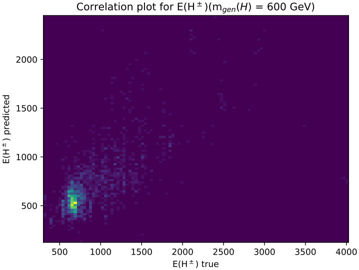


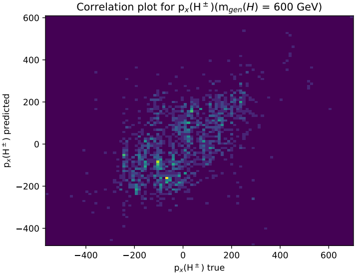


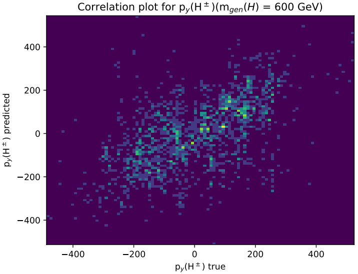


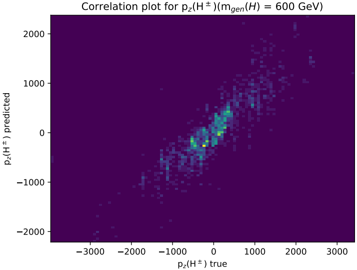


Correlation plots ($m^{gen}(H^{\pm}) = 600$ GeV)









Analysis status (31.07)

What is achieved so far:

→ Neural network is implemented

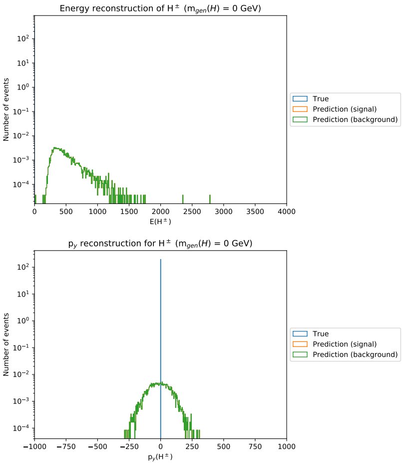
→ Training process is optimised (GPUs are used)

→ Trained neural network is able to predict $m(H^{\pm})$ for low masses and gives shifted results for higher masses

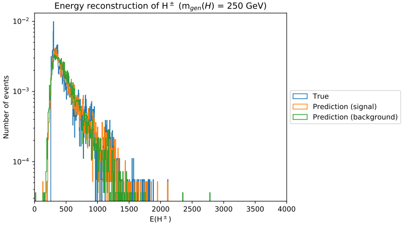
Plans:

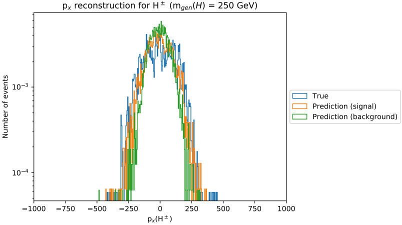
→ Include background events into dataset for training

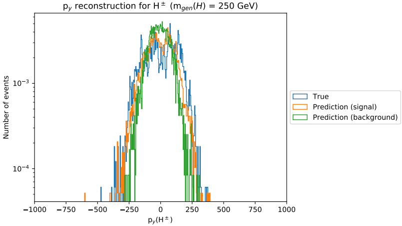
Results (4-momenta, $m^{gen}(H^{\pm}) = 0$ GeV, bkg$\rightarrow 0$)

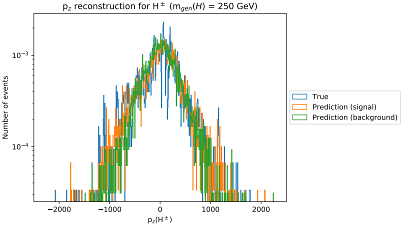


Results (4-momenta, $m^{gen}(H^{\pm}) = 250$ GeV, bkg$\to 0$)

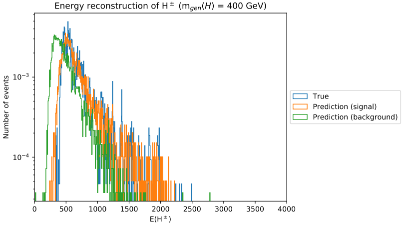


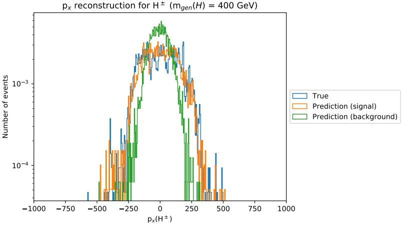


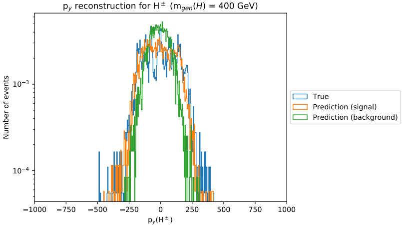


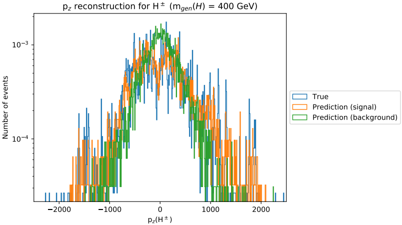


Results (4-momenta, $m^{gen}(H^{\pm}) = 400$ GeV, bkg$\to 0$)

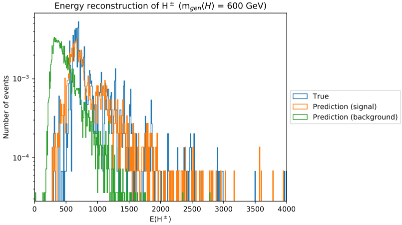


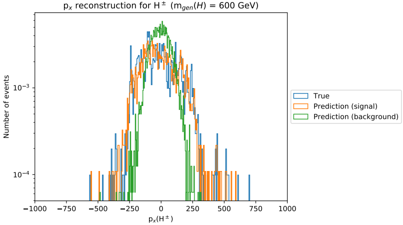


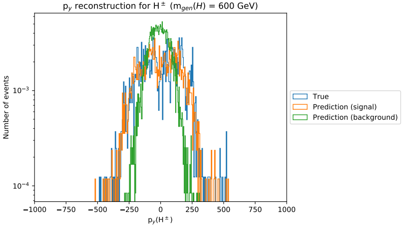


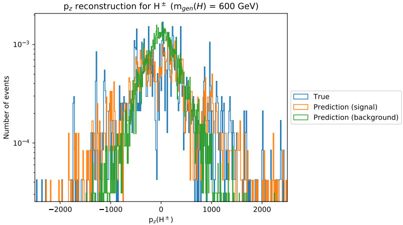


Results (4-momenta, $m^{gen}(H^{\pm}) = 600$ GeV, bkg$\rightarrow 0$)

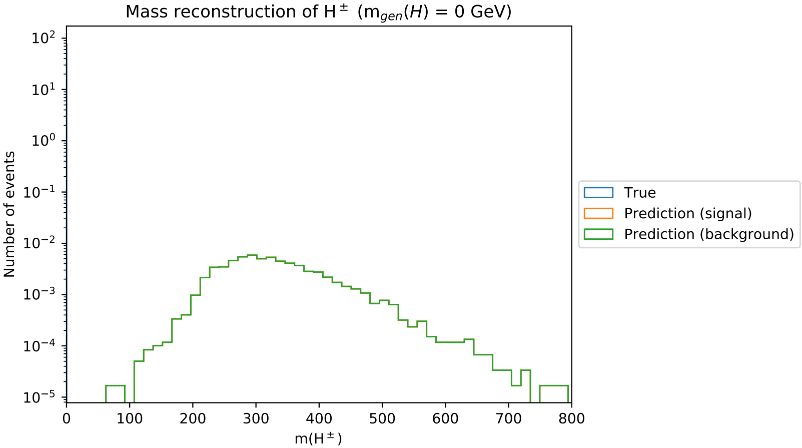




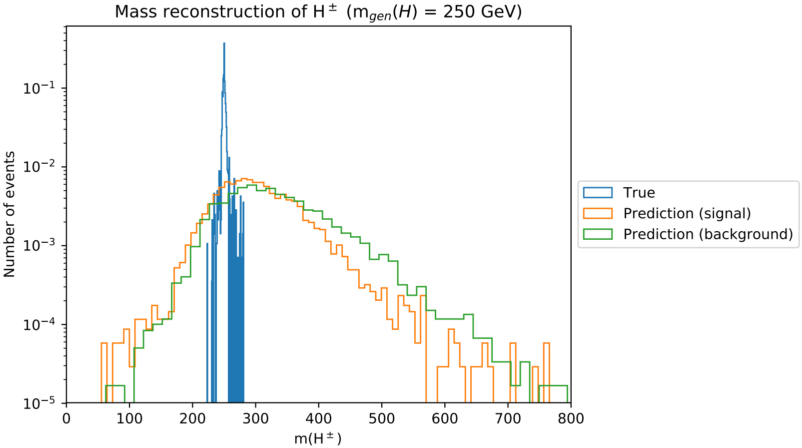




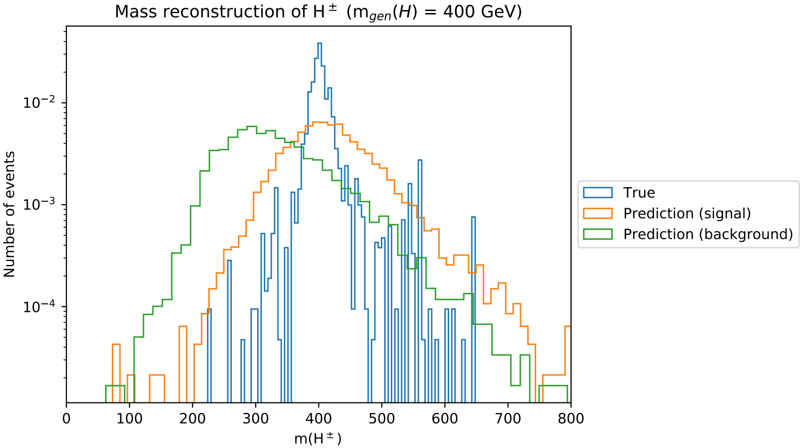
Results ($m^{pred}(H^{\pm})$ for $m^{gen}(H^{\pm}) = 0$ GeV, bkg$\rightarrow 0$)



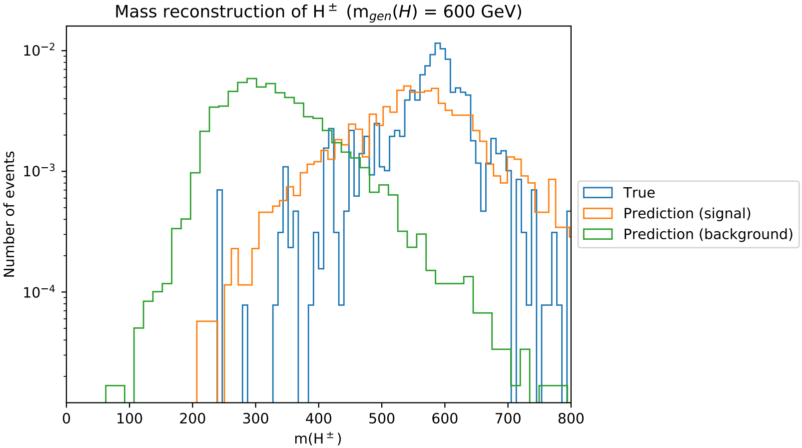
Results ($m^{pred}(H^{\pm})$ for $m^{gen}(H^{\pm}) = 250$ GeV, bkg$\rightarrow 0$)



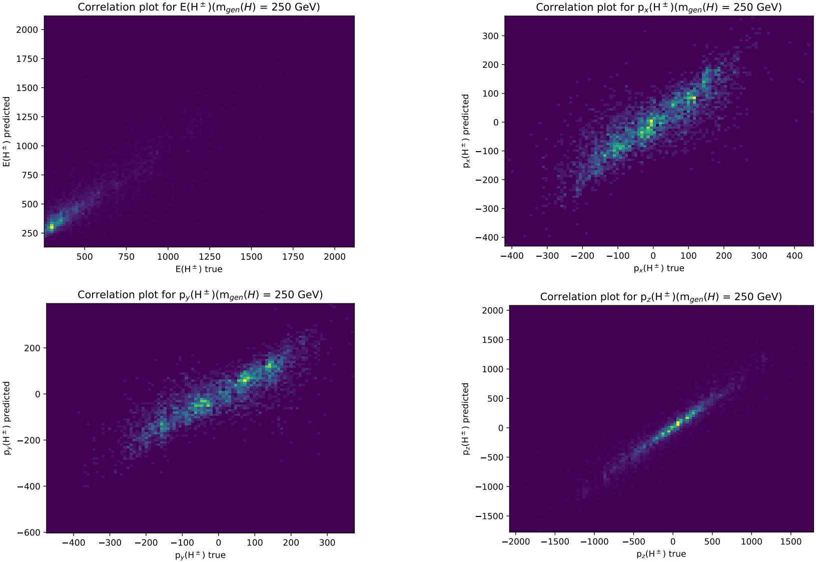
Results ($m^{pred}(H^{\pm})$ for $m^{gen}(H^{\pm}) = 400$ GeV, bkg$\to 0$)



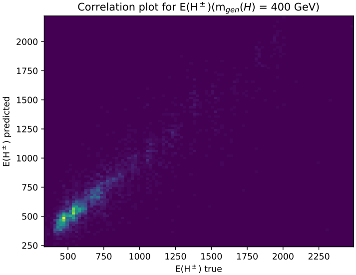
Results ($m^{pred}(H^{\pm})$ for $m^{gen}(H^{\pm}) = 600$ GeV, bkg$\rightarrow 0$)

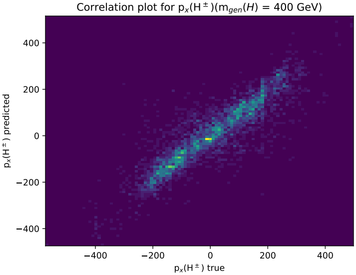


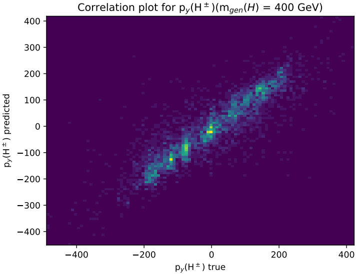
Correlation plots ($m^{gen}(H^{\pm}) = 250$ GeV, bkg$\rightarrow 0$)

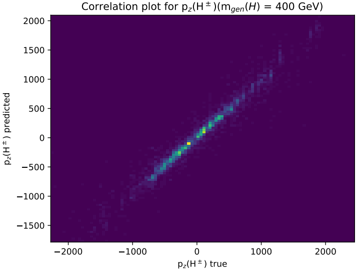


Correlation plots ($m^{gen}(H^{\pm}) = 400$ GeV, bkg$\rightarrow 0$)

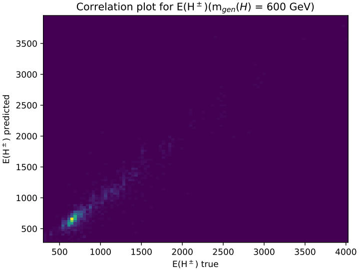


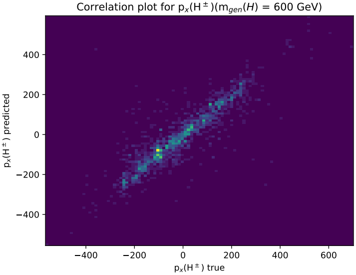


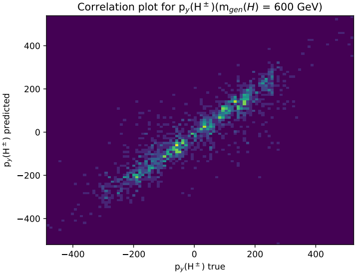


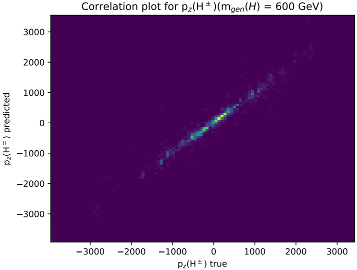


Correlation plots ($m^{gen}(H^{\pm}) = 600$ GeV, bkg$\rightarrow 0$)









Comments on other approaches of including bkg int

We have tried these approaches:

- → Predict for background events 4-momentum with high mass (in our case: 10^3 GeV and 10^4 GeV) and $\vec{p} = \vec{0}$ GeV
- → Predict for background events 4-momentum randomly: $E \sim U[0\ GeV, 500\ GeV]$, $p_x, p_y, p_z \sim U[-500\ GeV, 500\ GeV]$

Results: worse than for bkg $\rightarrow 0$
Decision: remove (temporarily?) background events from training

New architecture?

New proposal: we decided to use more complicated architecture of the neural network. Specificially, based on some positive results (see "ML applications to jet tagging in CMS", Mauro Verzetti) we added convolutional *input layers* before the Dense Layers.
We form a table 6×4 with the inputs (particle $\times$ momentum).
Particles were ordered by p_t.
Hyperparameters tuning gives us this set:

- → Number of hidden layers: 2
- → Nodes in each hidden layer: 50
- → Batch size: 175
- → Activation function for nodes in hidden layers: softplus
- → Convolutional layer: 1
- → Number of kernels: 4
- → Size of kernel: 3×4 (particle $\times$ momentum)

Analysis status (07.08)

What is achieved for this week:

- → Bkg was included into training (and then removed)
- → Some new architectures have been tested

Plans:

- → More architectures?
- → ? Consider events with more than 4 jets ?