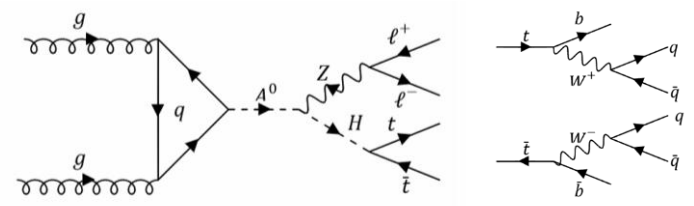
Improving a search for heavy neutral Higgs bosons in the $t\bar{t}Z$ final state at CMS using parameterized neural networks

DPG Frühjahrstagung 2025, Göttingen

Bianca Weidner, Matteo Bonanomi, Lukas Ebeling, Yannick Fischer, Johannes Haller, Daniel Hundhausen, Matthias Schröder

Overview

- Search targets $A \rightarrow ZH$ with decay of $H \rightarrow t\bar{t}$ (see previous talk T 25.1 by Yannick Fischer)
- Focus on hadronic decay of $t\bar{t}$
- CMS result with 138/fb of Run 2 data at 13 TeV published [1]
- Now: work on Run 3 data
- Here: study of neural network to separate signal of heavy Higgs boson events from background
- Considered mass range: $430\ \text{GeV} \leq m_A \leq 950\ \text{GeV}$, $330\ \text{GeV} \leq m_H \leq 850\ \text{GeV}$ with $\Delta m \geq 100\ \text{GeV}$



Limit plot Yannick

[1] [arXiv: 2412.00570 (subm. to PLB)]

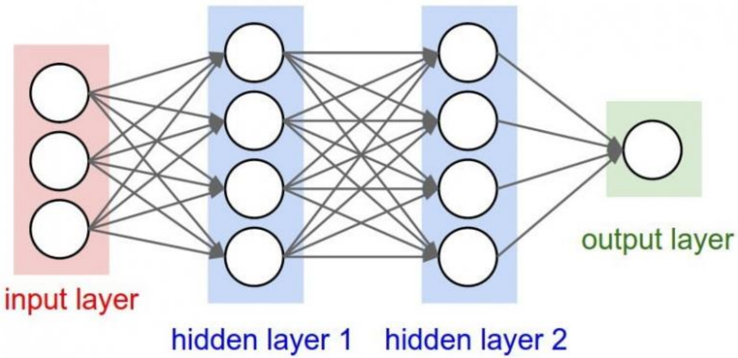
Two Higgs Doublet Models (2HDM)

- Second Higgs doublet → 5 different Higgs bosons (h , H, A, $H^{\pm}$)
- Free parameters
  - Masses of the Higgs bosons (m_h, m_H, m_A, $m_{H^{\pm}}$)
  - Ratio of the vacuum expectation values: $\tan(\beta) = \frac{v_1}{v_2}$
  - Mixing angle α between the CP-even Higgs bosons
- “Alignment limit”: $\cos(\beta - \alpha) \to 0$
  - Beyond the standard model h boson couples like the standard model Higgs boson

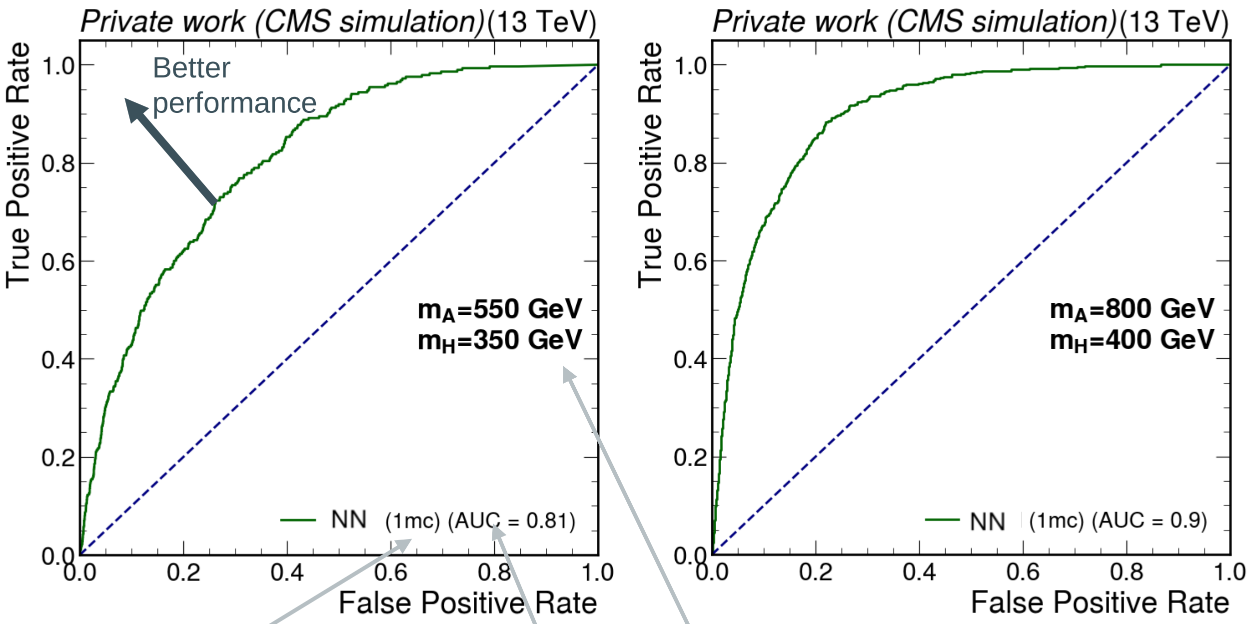
Deep neural networks

- Parameterized neural networks (PNN): true masses (m_A, m_H) used for generation of the Monte Carlo simulated signal events are used as an additional input feature
  - Background: gets random values for the true masses out of the list used for the true masses of the signals

- Implementation of a tuner to find optimal hyperparameters (Keras Hyperband Optimizer)



Dedicated neural network per mass combination



NN (1 mc): Not parameterized network (trained with 1 mass combination)

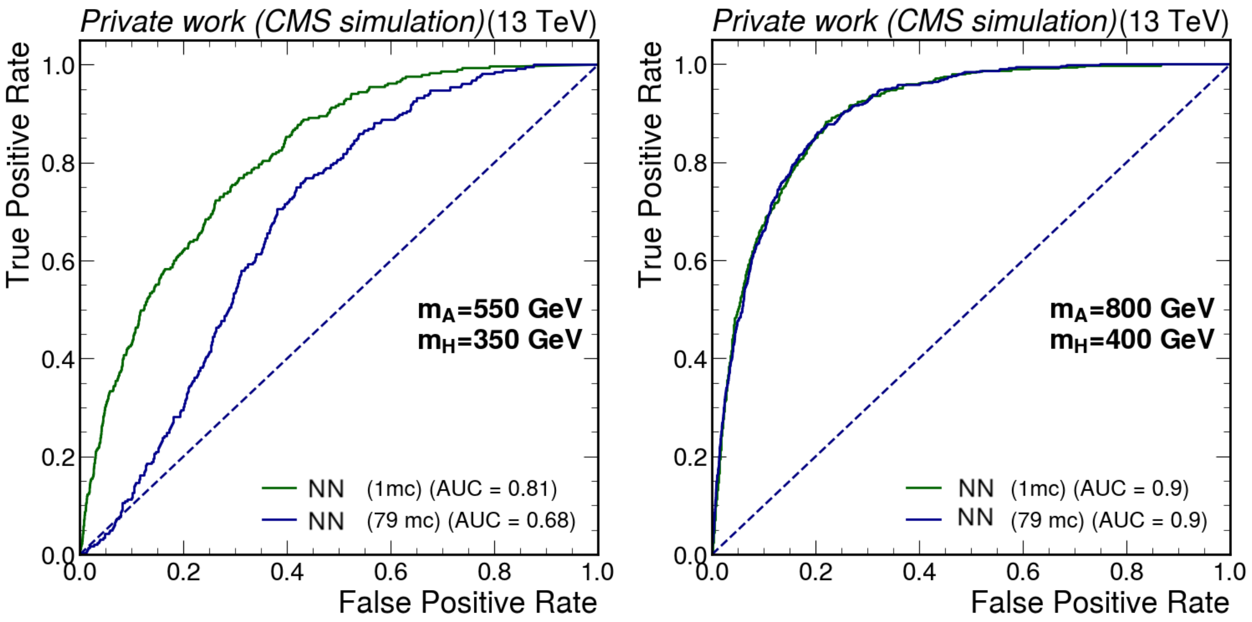
→ Performance is good, but it is impractical
→ For each mass combination a separate network is necessary

mc: number of mass combinations included in the training

Mass combination used for evaluation of the network

AUC: area under the ROC curve

Neural network trained with all mass combinations

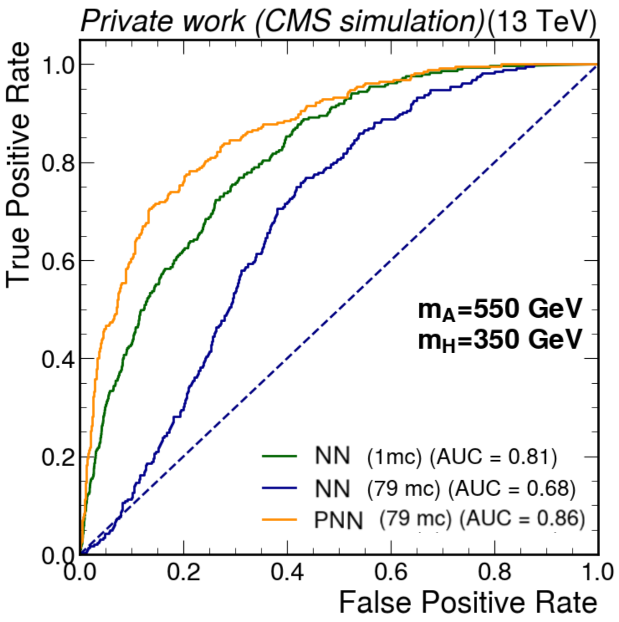


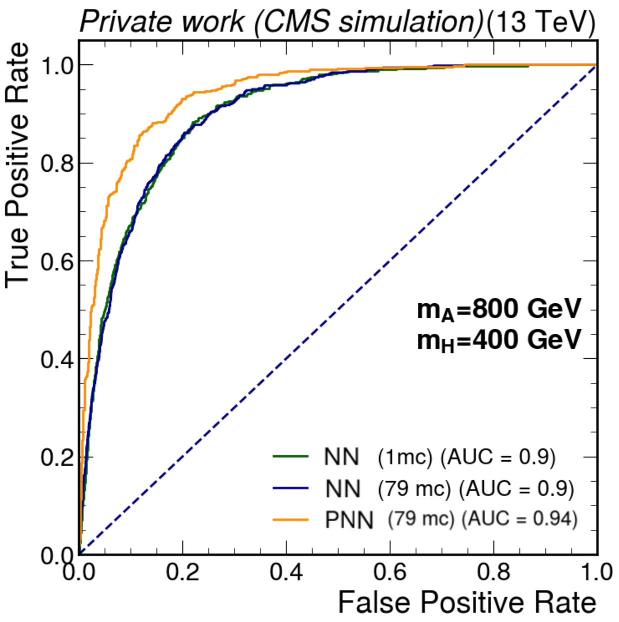
NN (1 mc): Not parameterized network (trained with 1 mass combination)

NN (79 mc): Not parameterized network (trained with 79 different mass combinations)

→ The inclusion of all signal events without parameterization leads to a significant decrease in performance for some mass combinations

Parameterized neural network trained with all mass combinations





NPN (1 mc): Not parameterized network (trained with 1 mass combination)

NPN (79 mc): Not parameterized network (trained with 79 different mass combinations)

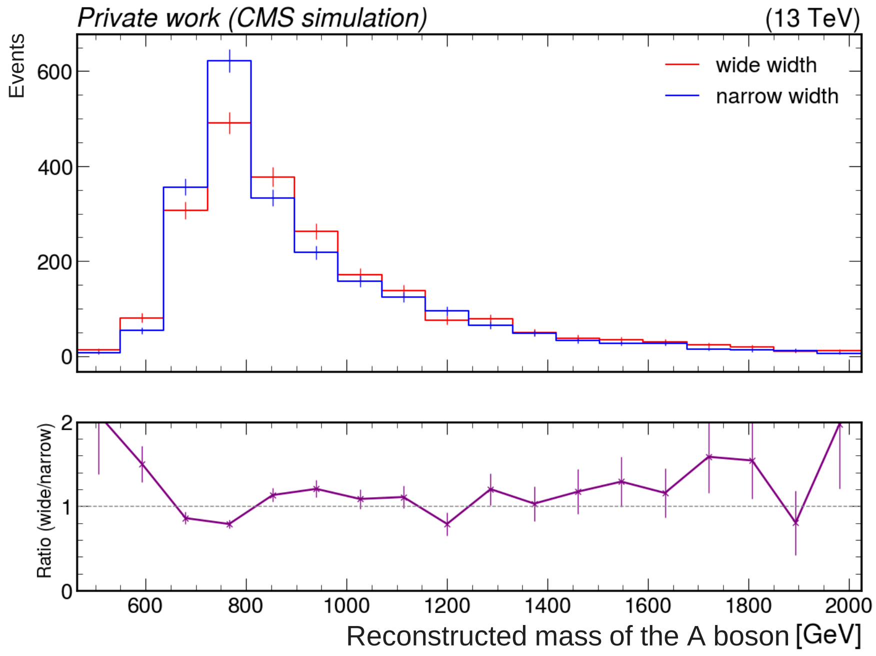
PNN: Parameterized network (trained with 79 different mass combinations)

→ The inclusion of all signal events with parameterization leads to a significant increase in performance

Advantage of parameterized networks

- One network can be used to evaluate all mass combinations
- Enhanced capability to interpolate to mass combinations not included in training compared to NN
- Better performance compared to networks trained with one mass combination is presumable due to more statistics included in the training
  - Statistics is for different mass combination
  - More statistics leads only to increase in case of parametrisation

Impact of changes of the decay width



Narrow width:

$$\frac{\Gamma_{A/H}}{m_{A/H}} = 0.03$$

Wide width:

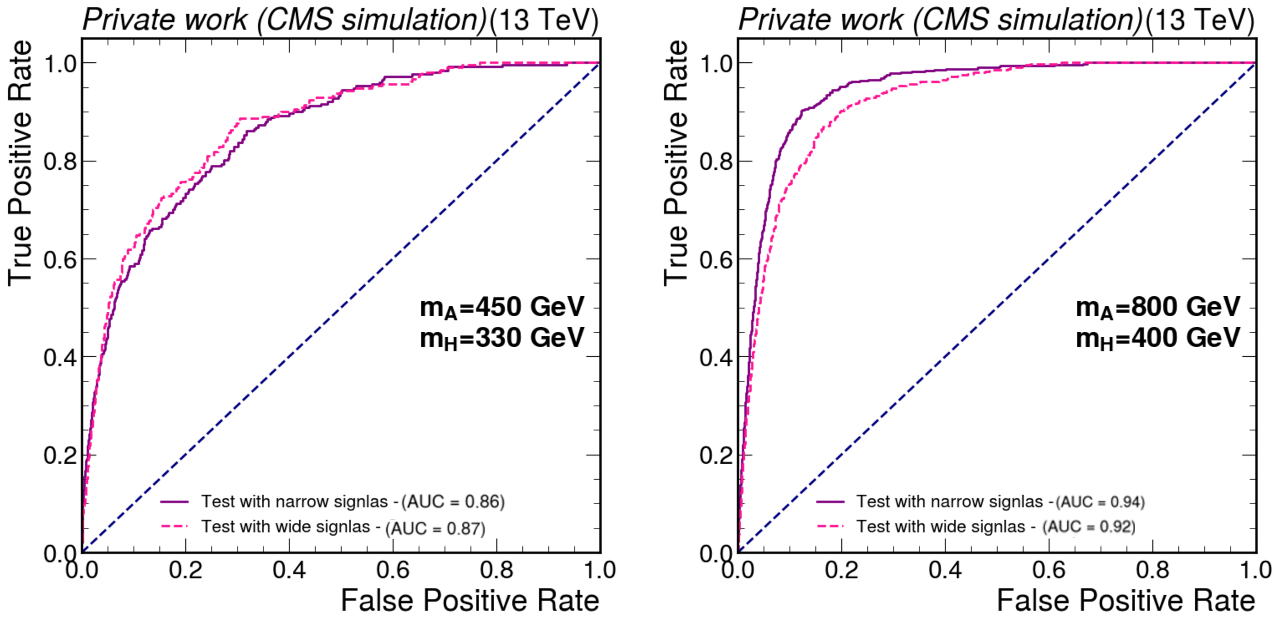
As predicted in 2HDM e.g.

$m_A = 800\ \mathrm{GeV}, \Gamma_A = 74.75\ \mathrm{GeV}$

(for tan(β)=2)

→ Decay width of the A boson is no longer neglectable compared to detector resolution

Impact of changes of the decay width



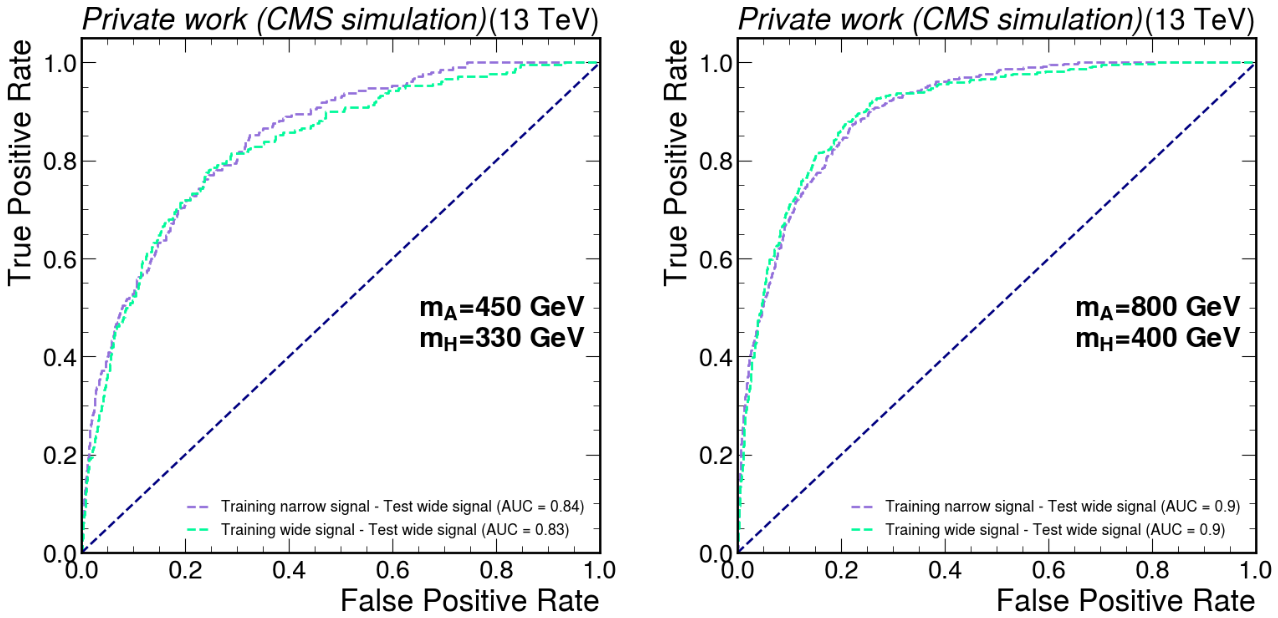
Training: narrow signal events

Testing: narrow and wide signal events in comparison

→ Even if we train the network with narrow signal events, the network performs nearly similar for testing with wide signal events

→ Neural network performs robust when the signal width is changed

Impact of changes of the decay width



Training: narrow and wide signal events in comparison

Testing: wide signal events

→ Using wide signal events for training does not lead to an increase in performance compared to training with narrow signals when testing with wide signal events

Summary

- Search for heavy Higgs bosons ($A \rightarrow ZH \rightarrow l\bar{l}\, t\bar{t}$)
- Studied neutral networks for signal/background classification
- Network performance measured using ROC and AUC score
- Parameterized network trained on several mass points performs better than several individual networks
- Parameterized neural network is robust to the effects of change in the width of the A boson
  - Only slightly decrease is observable for some mass combinations