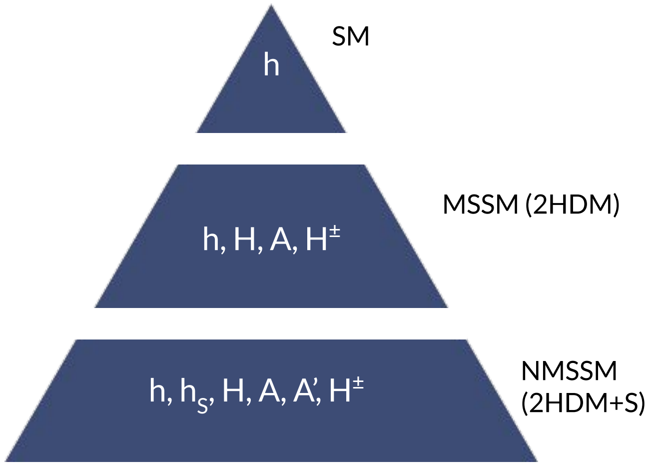
Studies of neural network architectures in the search for di-Higgs events in the context of NMSSM in the $\tau\tau$+bb final state

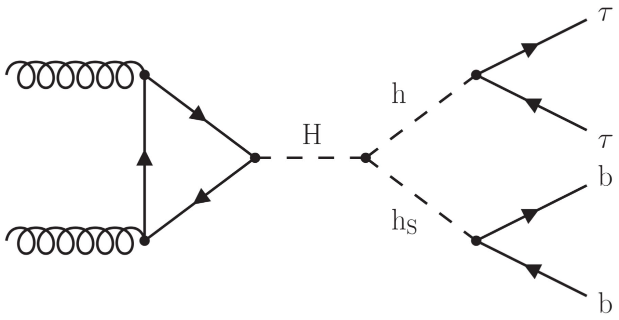
Ralf Schmieder

21.02.22

NMSSM phenomenology

- Decays of a heavy Higgs boson into two lighter bosons: Motivated e.g. by the next-to-minimal-supersymmetric standard model (NMSSM)
- Light Higgs boson h_S assumed to have significant admixture of the singlet field S
- In this case $\mathbf{H \rightarrow hh_S}$ is the dominant production process for $\mathbf{h_S}$
- $\mathbf{H \rightarrow h(\tau\tau)h_S(bb)}$ combines high branching ratio through bb decay with the lower bkg from the $\tau\tau$ decay

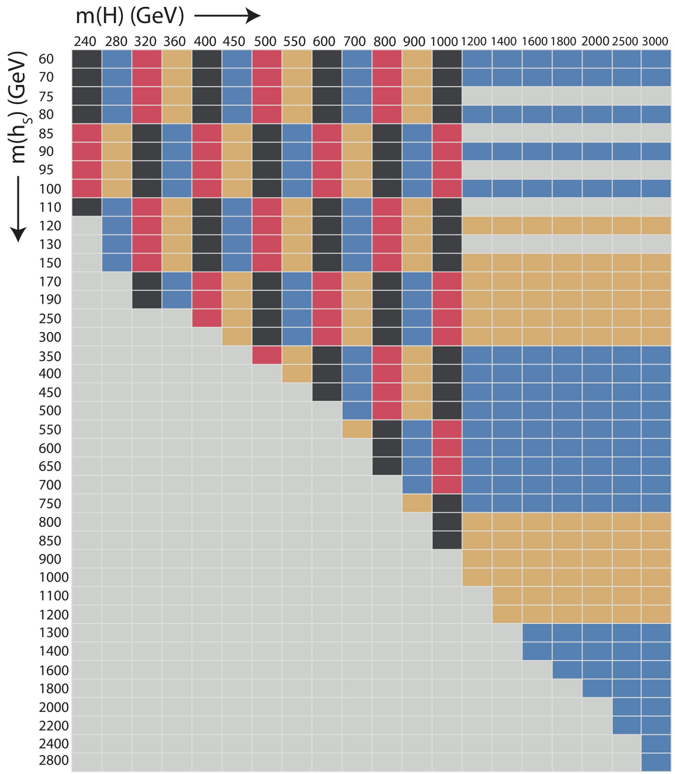




2HDM+S = 2 Higgs doublet model + Singlet

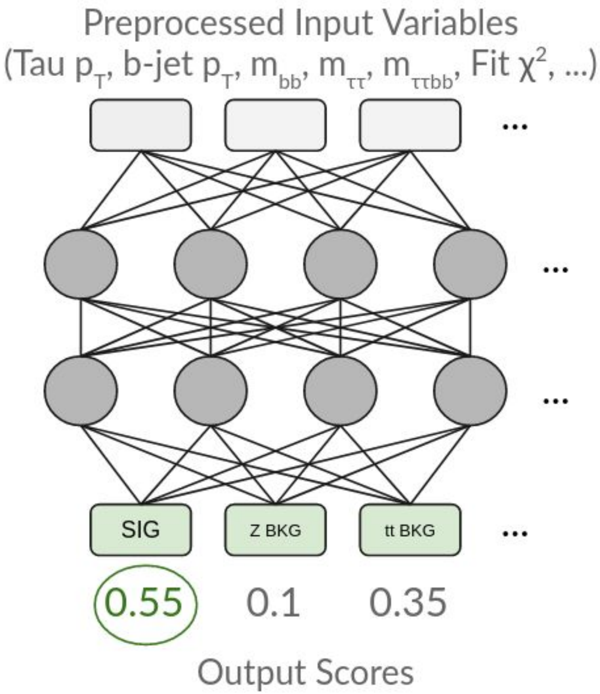
Analysis strategy in a nutshell

- In **H →hh$_S$** m(H) and m(h$_S$) are unknown
- A grid of 420 mass hypotheses are simulated with up to 500k events each between m(H) = 240 GeV and 3000 GeV
- Event categorization with the help of neural network (NN) multiclassification
- Grouping of mass hypotheses in individual trainings (color code on right figure) → resulting in **68 trainings per $\tau\tau$ final state**
- Depending on the $\tau\tau$ **final state, ~45-95% of backgrounds estimated from data**



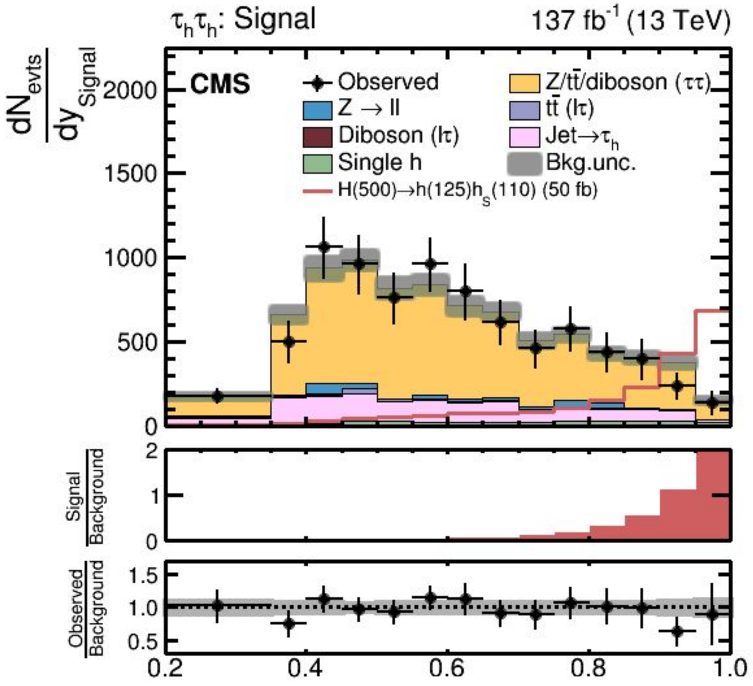
Event Categorisation

- Multiclassification based on neural networks (**NNs**) with one signal and four background categories
    - NN returns probability-like score for each category. The event is assigned to the category with the highest NN score
    - NN score used as final discriminator for signal extraction
- Mass hypotheses have different kinematic properties → NN training in 68 groups of signals



Example distribution of NN score

- Training group comprising the signal samples with m(H)=500 GeV, m(h_S)=[110, 120, 130, 150] GeV
- Signal with m(H)=500 GeV, m(h_S)=150 GeV indicated by red unstacked histogram
- Total of 45 such histograms enter 420 combined maximum likelihood fits, one for each mass hypothesis, for signal extraction



NN Score signal class

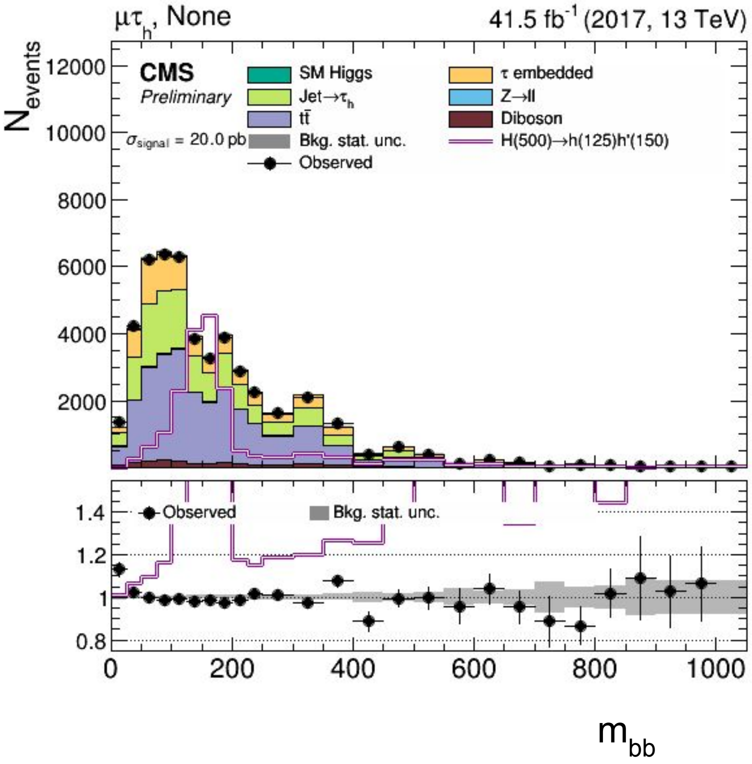
Analysis of the invariant bb mass as input to the NN

- Best possible setup: individual training per mass hypothesis
- Study of possible loss of sensitivity due to the grouping of hypotheses
    - Most groups contain four signal hypotheses
- Compare the results of individual training and grouped training
- Study of impact of mass estimator of di-b-system to the NN categorization

Analysis of the invariant bb mass as input to the NN

- Two informations from bb mass estimator:
    - Signal is peaking
    - Signal peak is fixed
- Hypothesis:

  In groups of 4 or more hypothesis this character is washed out



NN - Taylor coefficient analysis

- Designed to get a hint on the impact of the input features **x** (Tau p_T , $m_{\tau\tau}$, ...) on the NN output function $f(\mathbf{x})$
- Expand $f(\mathbf{x})$ in its input features x_i up to the second order
- For each event **a** with input feature values a_i, $f(\mathbf{x})$ is expanded around **a**
- Taylor coefficients t_i

$\rightarrow$ **Study impact of the input feature m_{bb}**

$$f(\mathbf{x}) \approx f(\mathbf{a}) + (x_1 - a_1)t_{x_1} + (x_2 - a_2)t_{x_2} + (x_1 - a_1)^2 t_{x_1 x_1} + (x_1 - a_1)(x_2 - a_2)t_{x_1 x_2} + \ldots$$

NN - Taylor coefficient analysis

Example: $x_1 = m_{\tau\tau}$, $x_2 = m_{bb}$

Impact of the correlation of $m_{\tau\tau}$ and m_{bb} on the NN output function

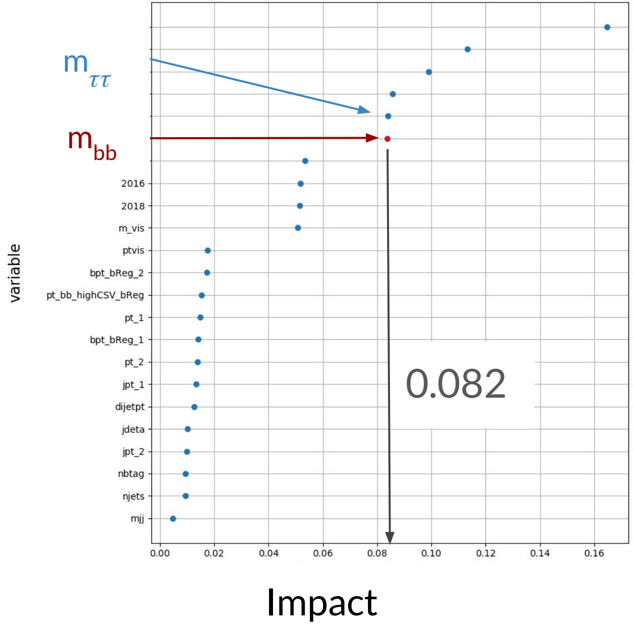
Impact of m_{bb} on the NN output function

Impact of $m_{\tau\tau}$ on the NN output function

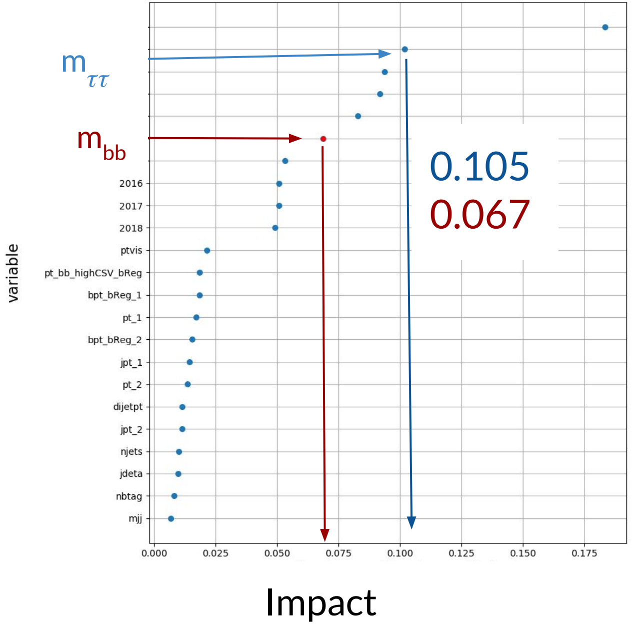
$$f(\mathbf{x}) \approx f(\mathbf{a}) + (x_1 - a_1)\boxed{t_{x_1}} + (x_2 - a_2)\boxed{t_{x_2}} + (x_1 - a_1)^2 t_{x_1 x_1} + (x_1 - a_1)(x_2 - a_2)\boxed{t_{x_1 x_2}} + \dots$$

Impact of mass estimators to the signal class

Single mass hypothesis

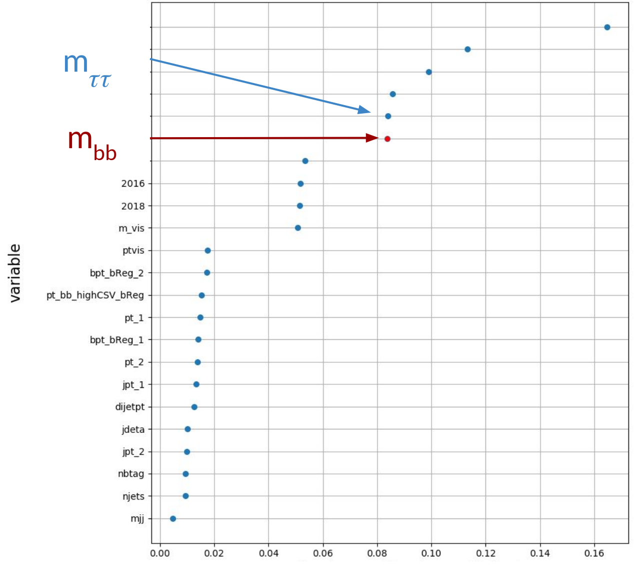


Four mass hypotheses



Impact of mass estimators to the signal class

Single mass hypothesis



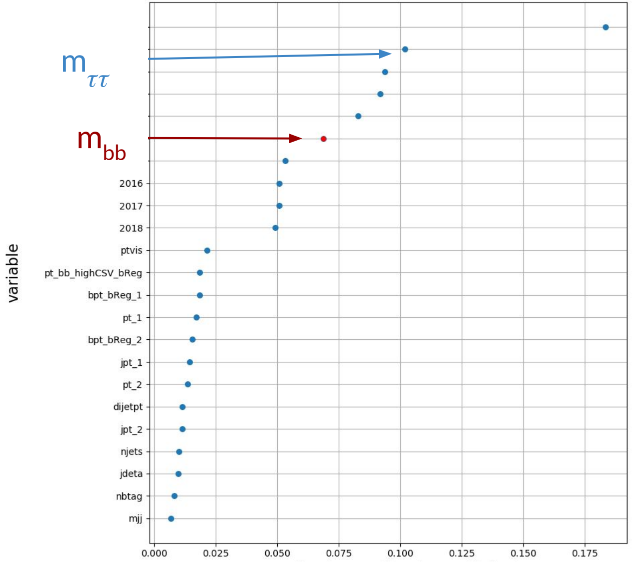
Impact

Impact

$m_{\tau\tau}$ ↑

m_{bb} ↓

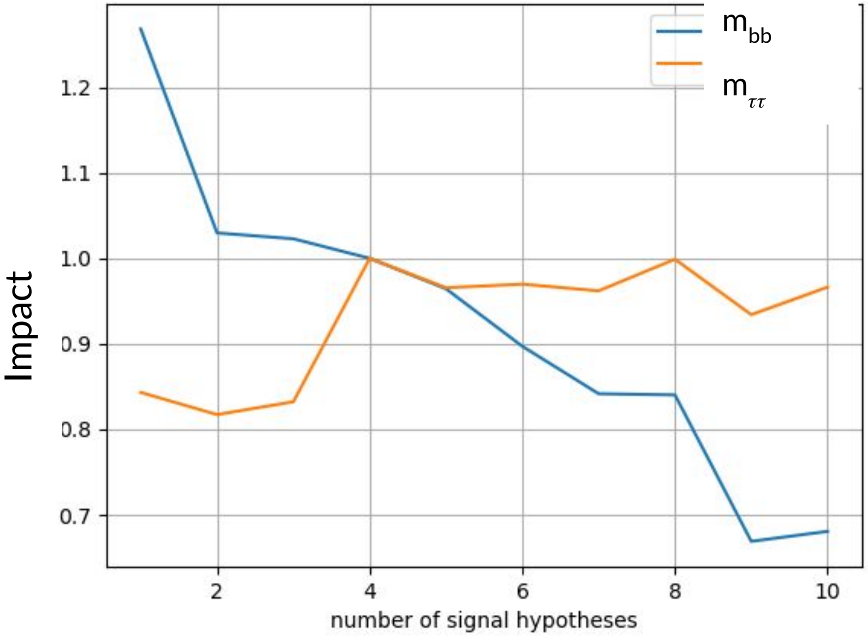
Four mass hypotheses



Impact

Impact of mass estimators to the signal class

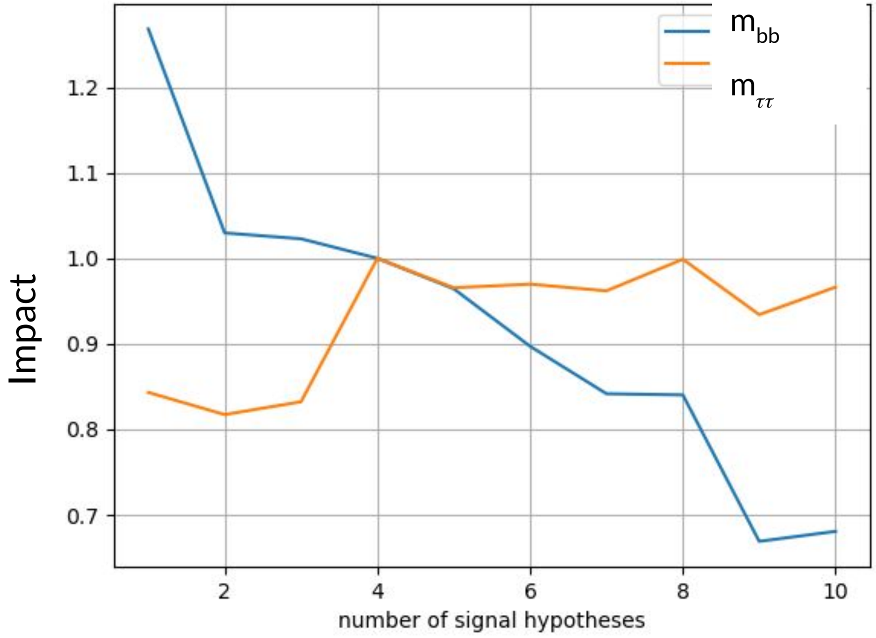
- Investigate inverse trends of impact of m_{bb} and $m_{\tau\tau}$
- vary number of signal hypotheses that form signal class
- Impact normalized to the impact with 4 signal hypotheses



Impact of mass estimators to the signal class

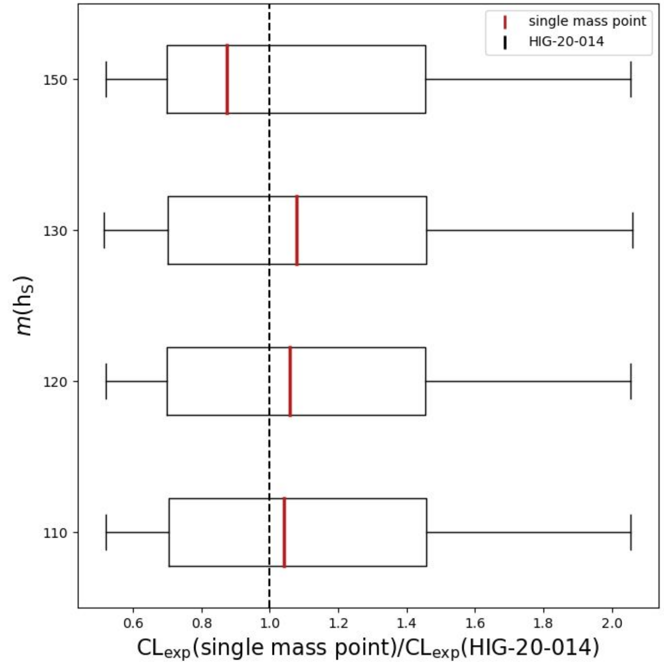
Impact of m_{bb} increases with the number of signals

Impact of $m_{\tau\tau}$ stagnates after 4 signals



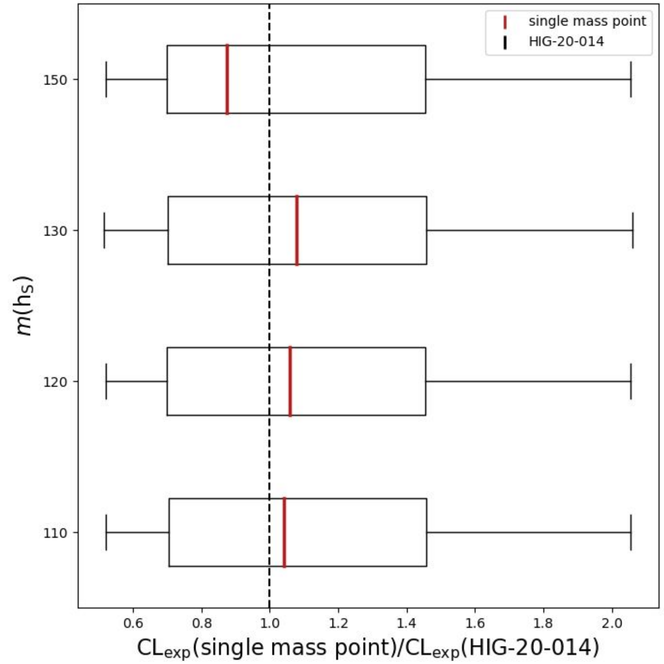
Analysis of the invariant bb mass as input to the NN

- No signal excess -> result is an exclusion limit
- Compare two results -> compare two exclusion limits
- Ratio plot:
    - Expected limits normalized to the limits of the base analysis
    - Boxes represent the 1σ band of the base analysis
    - Whiskers represent the 2σ band of the base analysis



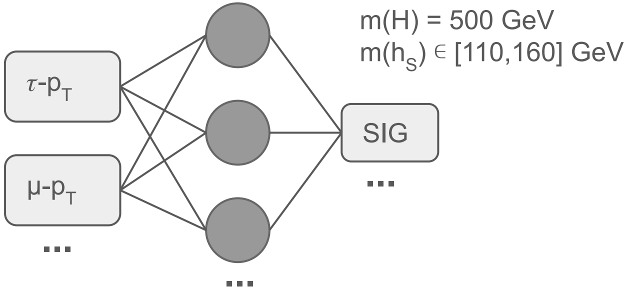
Analysis of the invariant bb mass as input to the NN

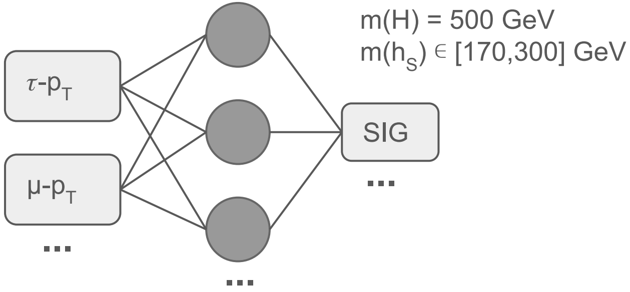
Grouping of four different but similar signal signatures to a signal class has no significant effect to the limits



Parameterized NN

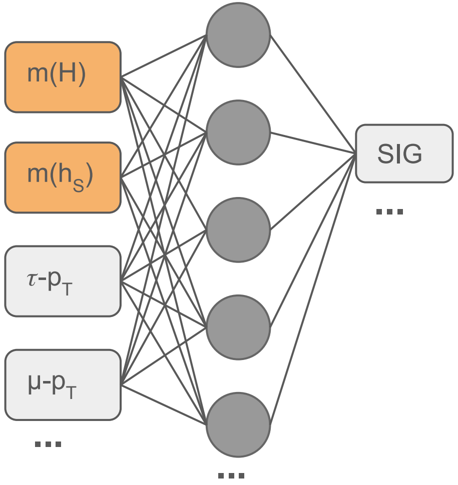
- Goal: replace 68 trainings per final state with a single training
- Idea: expand the input vector by parameters → m(H), m(h_S)
- Replace the NNs for different masspoints by a single NN, which trains on the same samples but has the parameters m(H), m(h_S) as input





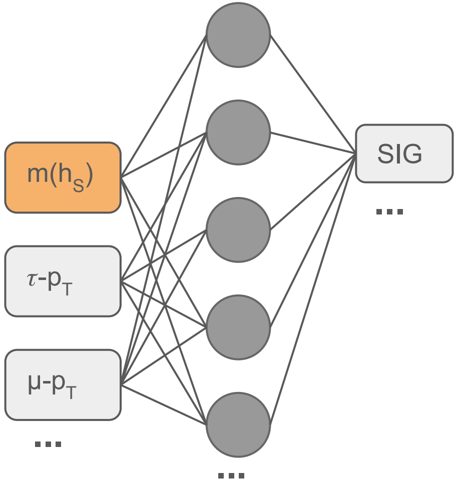
Parameterized NN

- Goal: replace 68 trainings per final state with a single training
- Idea: expand the input vector by parameters → m(H), m(h_S)
- Replace the NNs for different masspoints by a single NN, which trains on the same samples but has the parameters m(H), m(h_S) as input



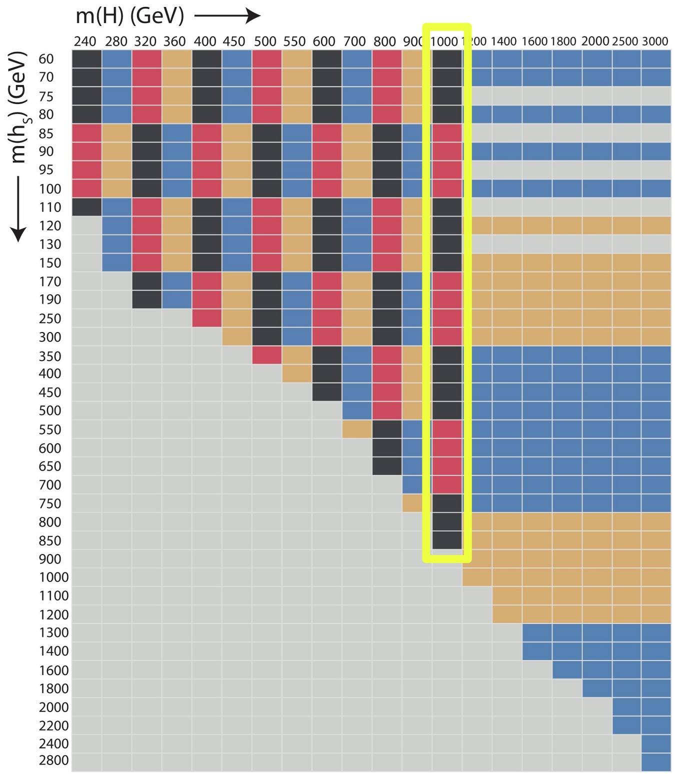
Parameterized NN

- First test: fix heavy Higgs boson mass m(H) to m(H)=1000 GeV and pass only $m(h_S)$
- Pass random value to backgrounds and true value to signal event
- Compare exclusion limits with the limits of the presented analysis



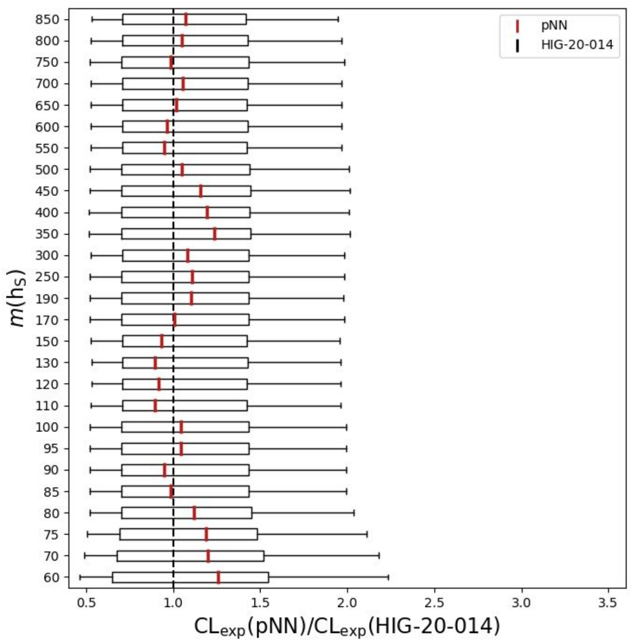
Parameterized NN

- First test: fix heavy Higgs boson mass m(H) to m(H)=1000 GeV and pass only m(h_s) as an input parameter
- 27 mass hypotheses tested in a single NN
  → replace 7 individual NN



Parameterized NN

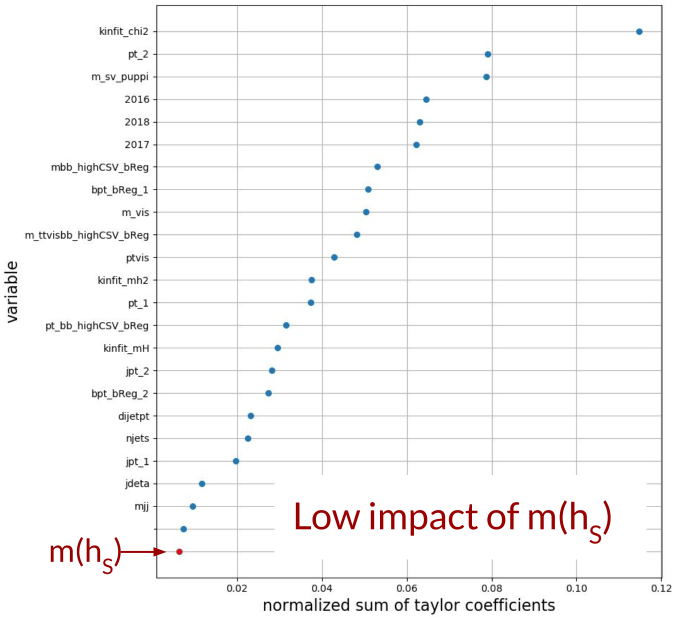
- Ratio plot:
  - Expected limits normalized to the limits of the base analysis
  - Boxes represent the 1σ band of the base analysis
  - Whiskers represent the 2σ band of the base analysis



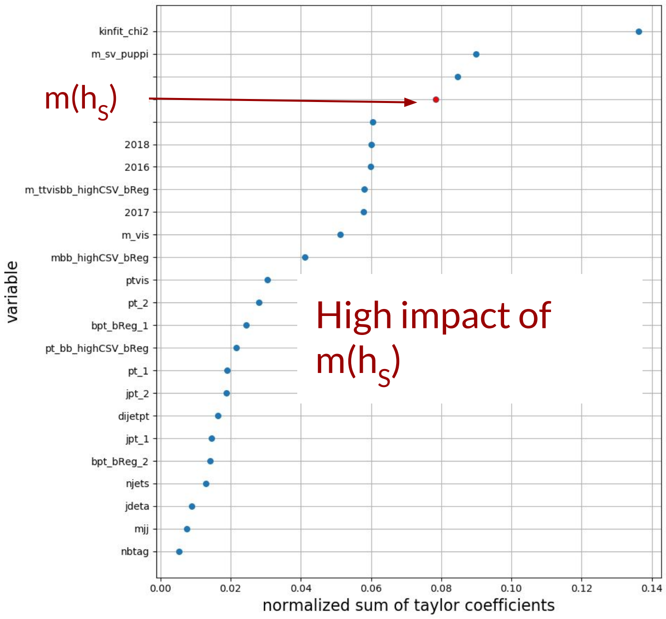
All tested mass hypotheses are within the 1σ band of the base analysis (HIG-20-014)

Parameterized NN - Taylor coefficient analysis

Background classes



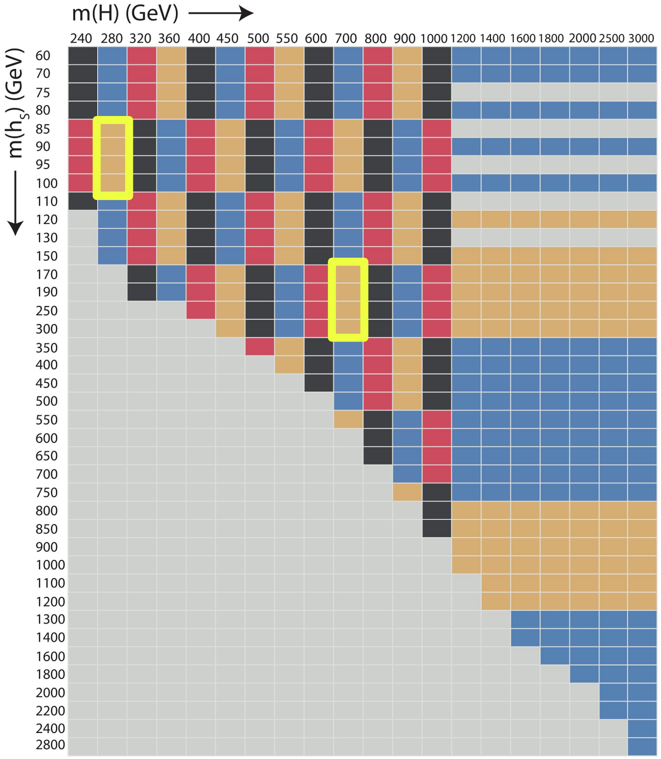
Signal class



Parameterized NN with sub NNs

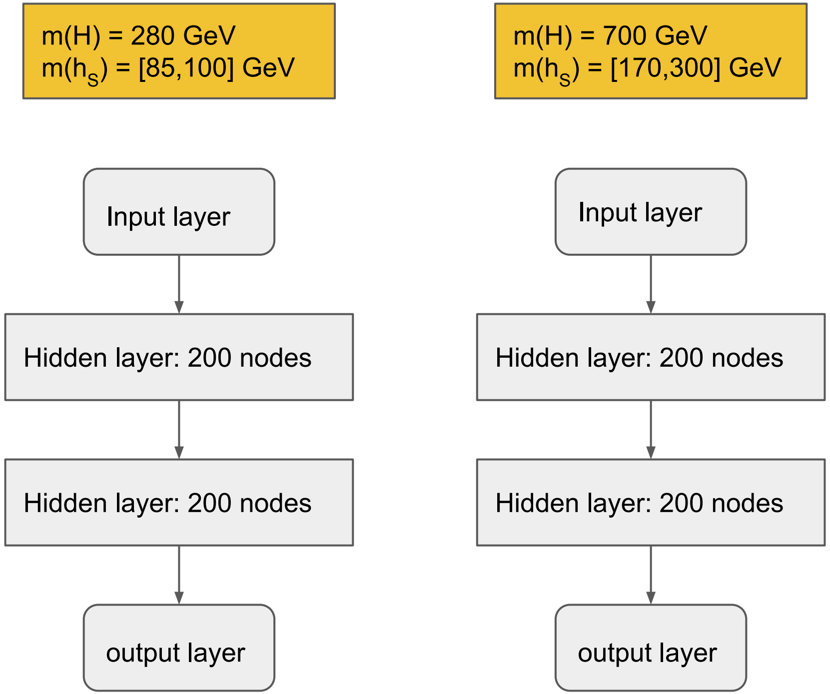
- Goal: single NN with a several sub NNs

- **First test**: pick two batches from the 68, implement a functional NN and compare the limits with the base analysis
    - m(H) = 280 GeV, batch 2
    - m(H) = 700 GeV, batch 4



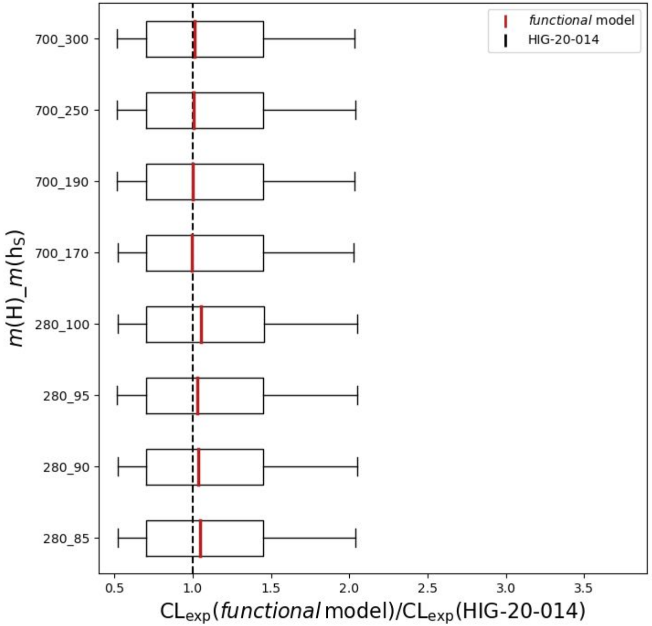
Parameterized NN with sub NNs

- Build a single NN with multiple input and multiple output tensors, each for one batch
- Each sub NN has the same architecture as the base analysis
- Sub NNs are independent → no shared layers



Parameterized NN with sub NNs

- Build a single NN with multiple input and multiple output tensors, each for one batch
- Each sub NN has the same architecture as the base analysis
- Sub NNs are independent → no shared layers



Both NNs lead to the same exclusion limits

Summary

- Base analysis: search for $\mathbf{H} \rightarrow \mathbf{hh_s}$ decays in the context of the NMSSM
- No significant loss of discrimination power with grouping signals hypotheses
- PNN as a single NN to replace 68 individual NNs
- PNN reaches same sensitivity as base analysis setup in tested parameter space and is able to interpolate between mass hypotheses
- PNN with sub NNs leads to the same exclusion limits as base analysis