Search for additional Higgs bosons in $H \rightarrow h(\tau\tau)h_s(bb)$ decays with CMS Run 2 data

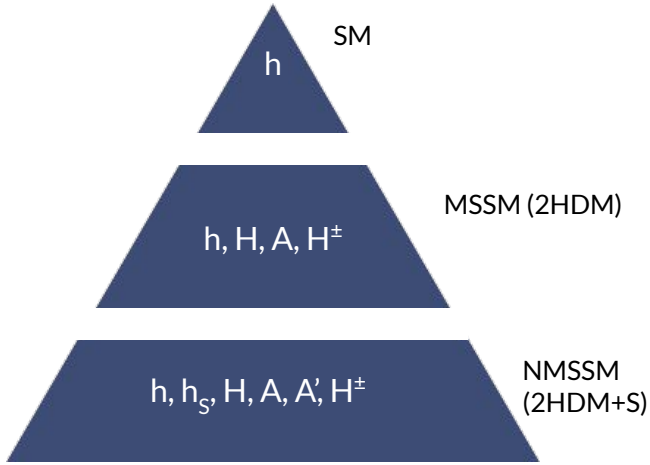
Ralf Schmieder
On behalf of the CMS Collaboration

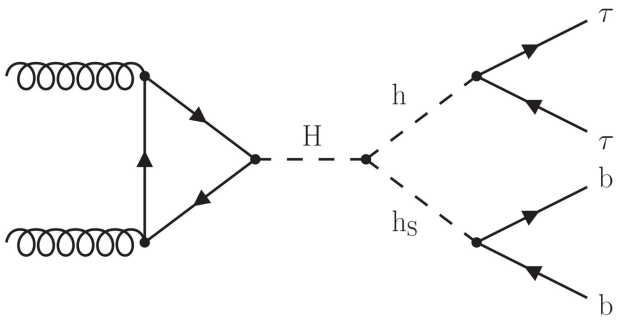
20.09.21

HIG-20-014 (submitted to JHEP)

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 $\mathbf{H \rightarrow h h_S}$ 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

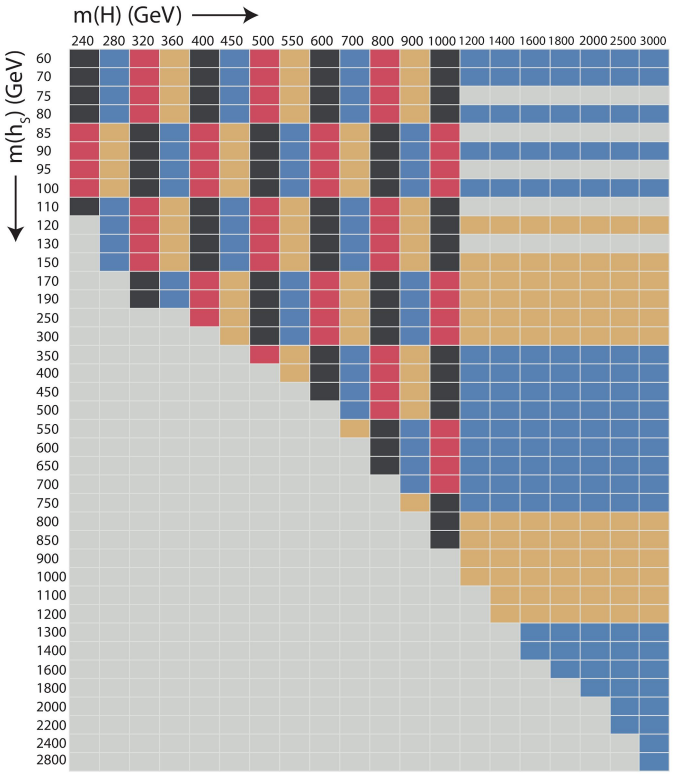




2HDM+S = 2 Higgs doublet model + Singlet

Analysis strategy in a nutshell

- In $\mathbf{H \rightarrow hh_S}$ 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**



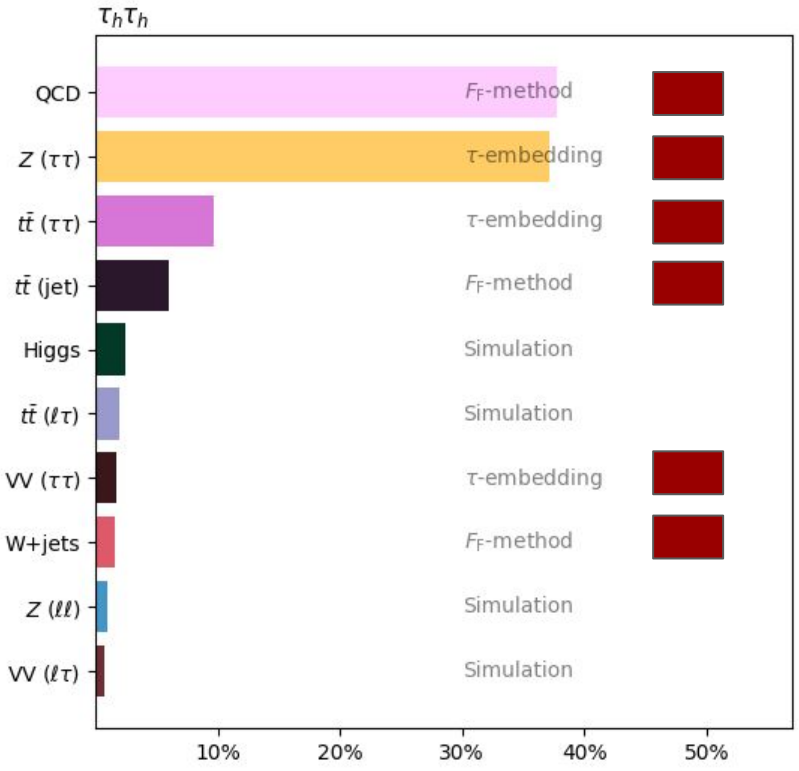
Background (BG) Estimation

3 methods used:

1. **τ-embedding** (JINST 14 (2019) P06032) for events decaying into two genuine tau leptons
2. **F_F method** (EPJ C 78 (2018) 708, JHEP 09 (2018) 007) for events entering the analysis due to jet $\rightarrow \tau_h$ misidentification
3. **Monte Carlo simulation** for everything else

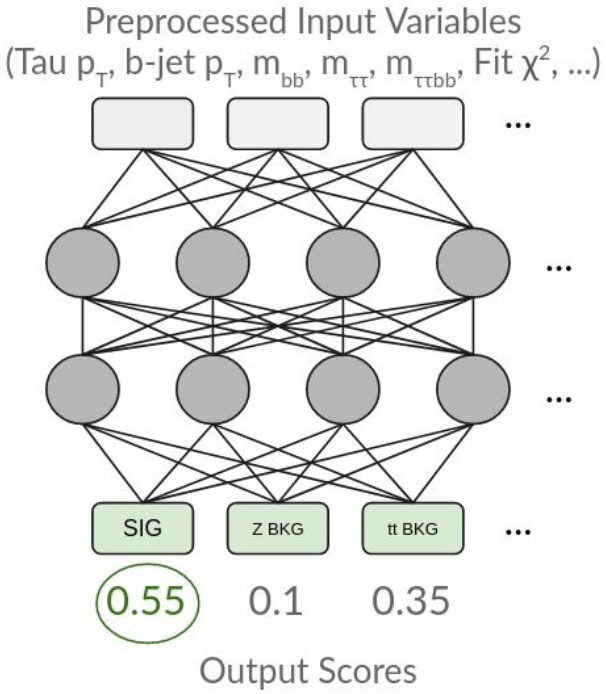
$e(\mu)\tau_h$: 42 % of BGs estimated from data

$\tau_h\tau_h$: 94 % of BGs estimated from data



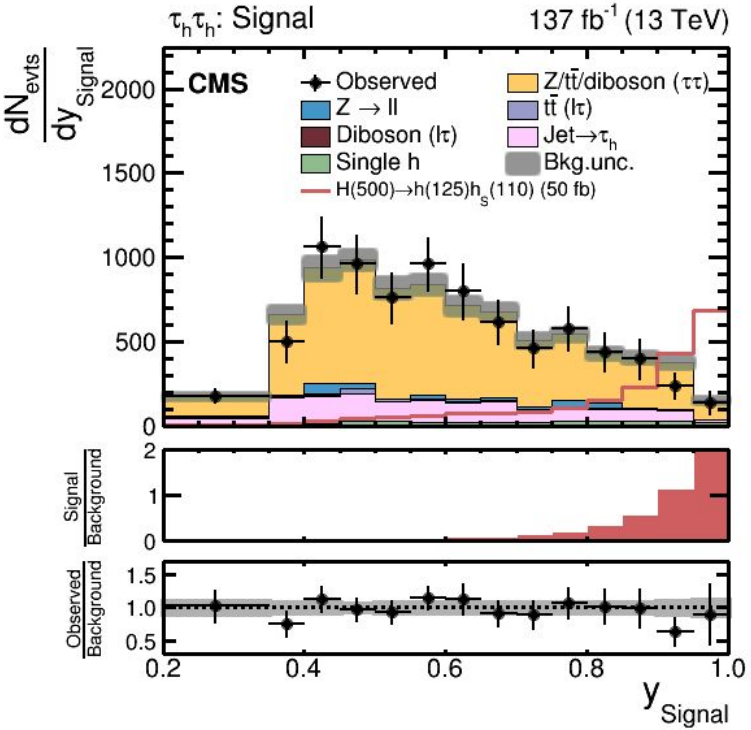
Event Categorisation

- Multiclassification based on **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



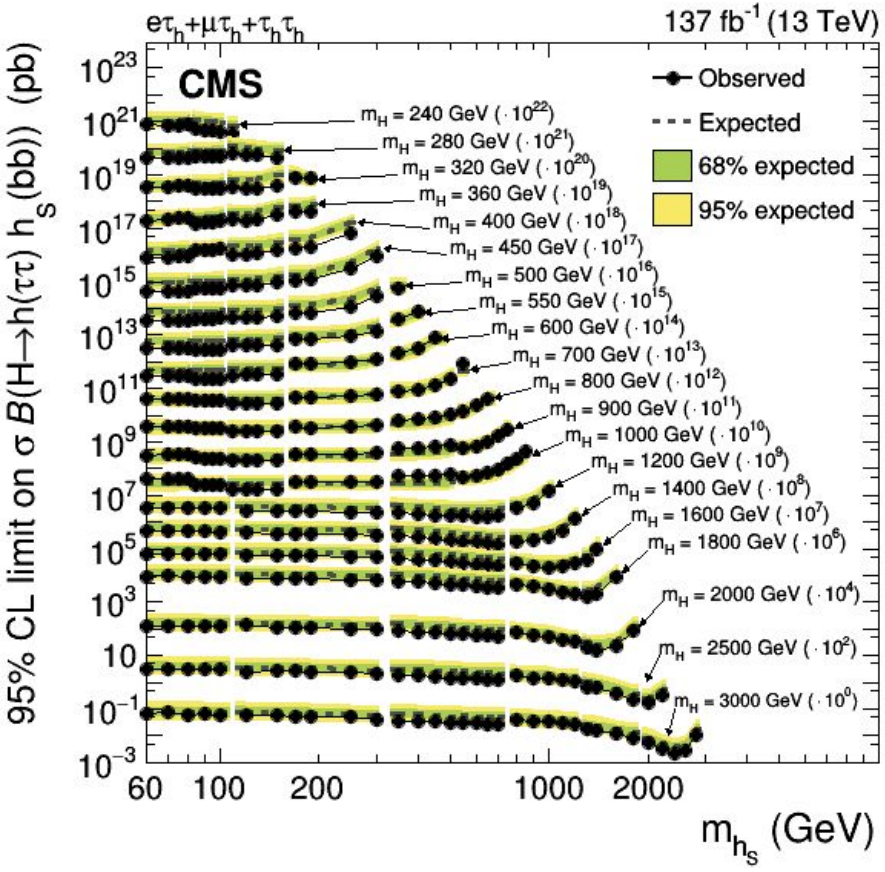
Example distribution of NN score

- 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



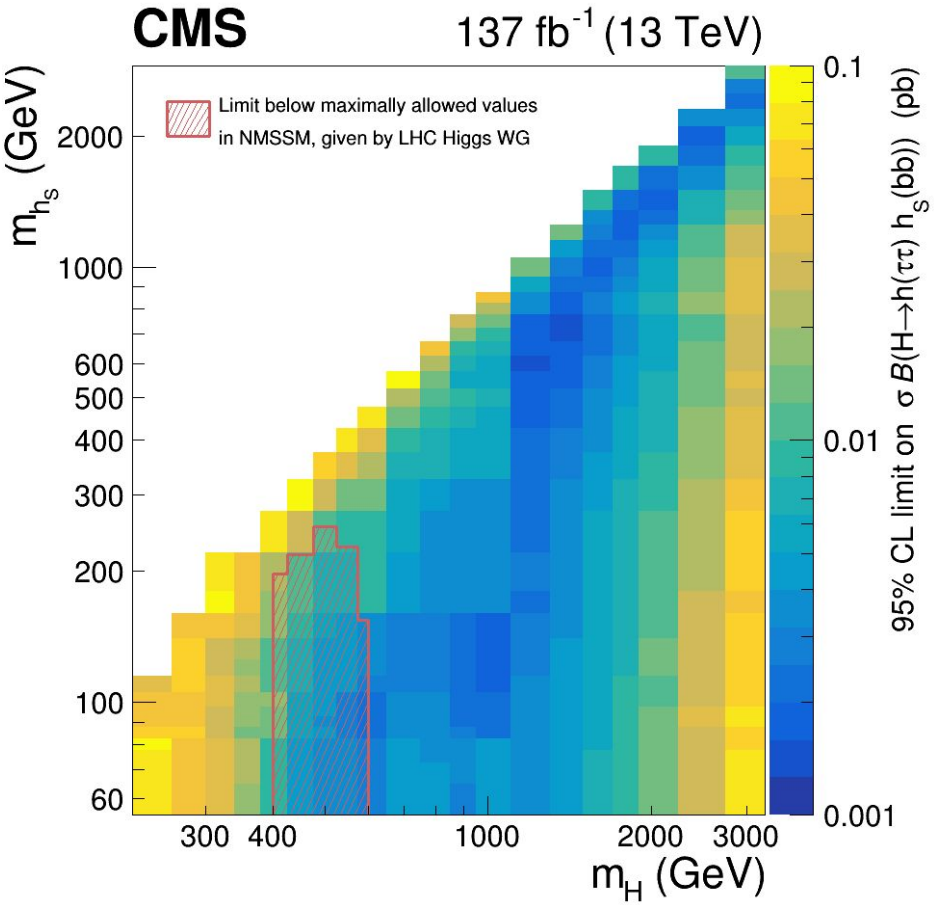
Exclusion limits

- No significant signal excess found
- Expected and observed 95% CL upper limits calculated
- For better visibility the limits of each corresponding mass value have been scaled by orders of 10 as indicated in the figure
- Numerical results of the search can be found in: https://www.hepdata.net/record/ins1869505



Interpretation in NMSSM

- Comparison of observed upper limits with maximally allowed σ x BRs for the $\mathbf{H \rightarrow h h_s}$ process in the NMSSM
- Red hatched region in the m(H)-m(h_s) plane can be further constrained by these data
- **First search for such a process at the LHC**



Conclusion

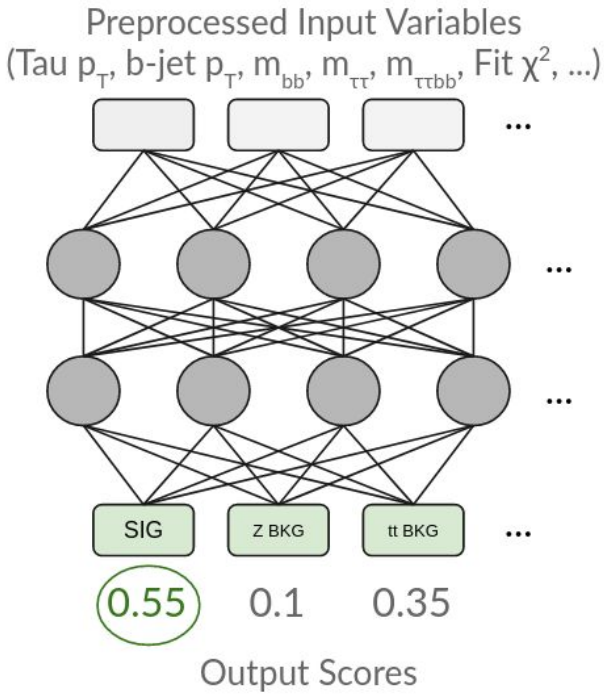
- A novel search for $\mathbf{H \rightarrow hh_s}$ decays in the context of the NMSSM was presented using the full CMS Run 2 Data
- ~45-95% of all backgrounds are estimated from data
- Event categorisation is performed using multiclass NNs
- No significant excess was found
- This is the first search for such a process at the LHC

Thank you for your attention!

Backup

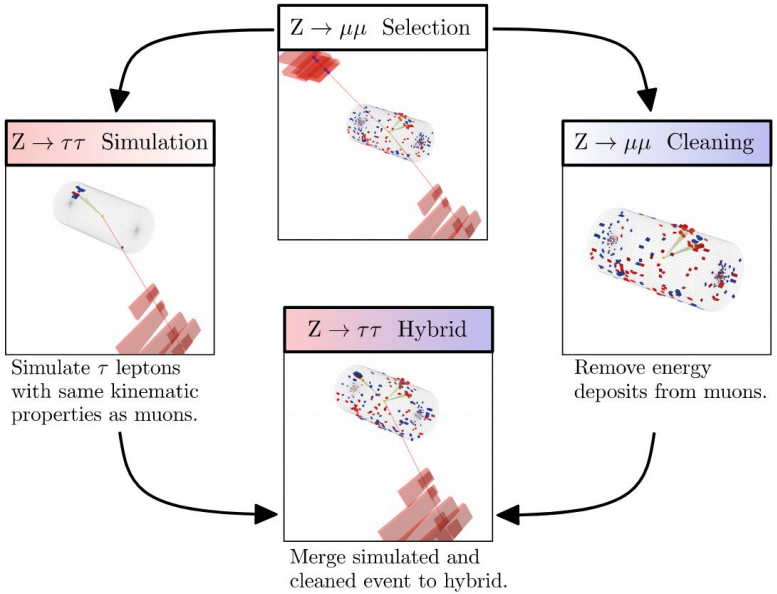
Event Categorisation

- 4 **background** categories, 1 **signal** category
  - Bkg genuine $\tau\tau$ pairs →background
  - jet→τ_h misidentified→background
  - Top quark pairs→background
  - Low cross section processes, like h($\tau\tau$) →background
  - NMSSM processes → signal



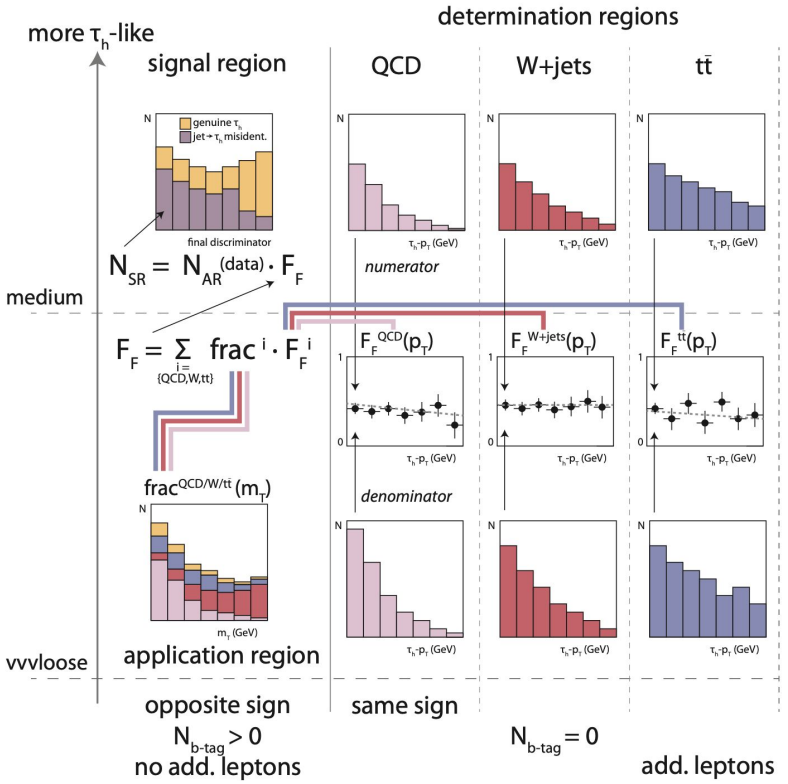
Background Estimation - Tau-Embedding

2019 JINST 14 P06032



Background Estimation - F_F Method

- Background processes are selected by process specific properties
- Sideband region with a high number of misidentified jets is defined
- Multivariate transfer function F_F is calculated and applied to signal region



NN input variables

Label	Description
pt_1	p_T of the muon, electron or p_T-leading τ_h
pt_2	p_T of the τ_h (p_T-subleading in $\tau_h\tau_h$)
m_vis	Visible mass of the $\tau\tau$ system
ptvis	Visible p_T of the $\tau\tau$ system
m_sv_puppi	SVFit mass of the $\tau\tau$ system
nbtag	Number of b-tagged jets
bpt_1	p_T of p_T-leading b-tagged jet
bpt_2	p_T of p_T-subleading b-tagged jet
mbb	Invariant mass of the two b-tagged jets
ptbb	p_T of the two b-tagged jets
njets	Number of non-b-tagged jets
jpt_1	p_T of the p_T-leading non-b-tagged jet
jpt_2	p_T of the p_T-subleading non-b-tagged jet
jdeta	$\Delta\eta$ between the two p_T-leading non-b-tagged jets
mjj	Invariant mass of the two p_T-leading non-b-tagged jets
dijetpt	p_T of the two p_T-leading non-b-tagged jets
m_ttvisbb	Invariant mass of the visible $\tau\tau$+bb system
kinfit_mH	m_H estimator derived by the kinematic fit
kinfit_mh2	Discrete $m(h_S)$-value selected for the minimal χ^2-value of the kinematic fit
kinfit_chi2	Minimal χ^2-value of the kinematic fit
2016	True if the event was recorded in the 2016 run period, false otherwise
2017	True if the event was recorded in the 2017 run period, false otherwise
2018	True if the event was recorded in the 2018 run period, false otherwise
bm_1†	Mass of the p_T-leading b-jet
bm_2†	Mass of the p_T-subleading b-jet
bcsv_1†	b-jet discriminator score of the p_T-leading b-jet
bcsv_2†	b-jet discriminator score of the p_T-subleading b-jet
jetCSV†	In case of only one jet passing the `medium` b-discriminator working point, b-jet discriminator score of non-b-tagged jet used for the $b\bar{b}$ system

Exclusion limits

- In absence of a signal excess, computation of exclusion limits with the CLs method
- Scan with a fixed light mass
- No significant deviation from Standard Model observed