H$\rightarrow \tau\tau$ measurements using neural networks

with 2016+2017 data at CMS

Teresa Lenz (DESY)
on behalf of the CMS Collaboration

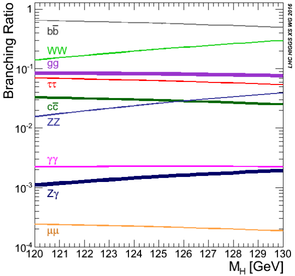
HELMHOLTZ RESEARCH FOR GRAND CHALLENGES

Terascale 13th Annual Meeting (DESY)

November 27th, 2019

Introduction

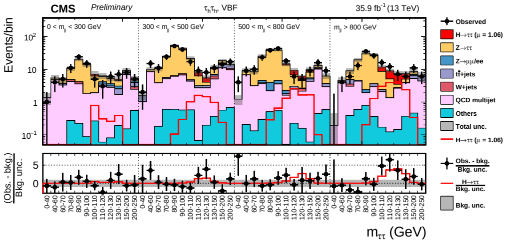
- Precision measurements of Yukawa couplings important test of SM.
- $H \to \tau\tau$ second largest branching ratio to fermions, lower background than $H \to bb$.



Observation of $H \to \tau\tau$ with $> 5\sigma$ (2016 data combined with 8 TeV result)

Phys. Lett. B 779 (2018) 283

- Cut-based analysis
- Three categories (VBF/qqH, ggH, background)
- 2-d discriminators



New results using machine learning techniques

CMS-PAS-18-032

Measurement of $H \to \tau\tau$ cross sections:

- Inclusively
- Split by production mode (ggH and qqH)
- In different kinematic regions (Higgs p_T, n_{Jets},...) $\to$ STXS stage-1

New compared to previous analysis:

- Analyzed 2016+2017 data (77 fb^{-1}).
- Multi-class neural network approach.
- STXS stage-1 (oriented) cross section measurement.
- Data-driven approaches for 90% of the background contributions:
  - $Z \to \tau\tau$: τ-embedding
  - Jet $\to \tau_h$: fake factor method

Analysis strategy

Before the neural net enters the stage:

- **Four** channels are separately analysed:

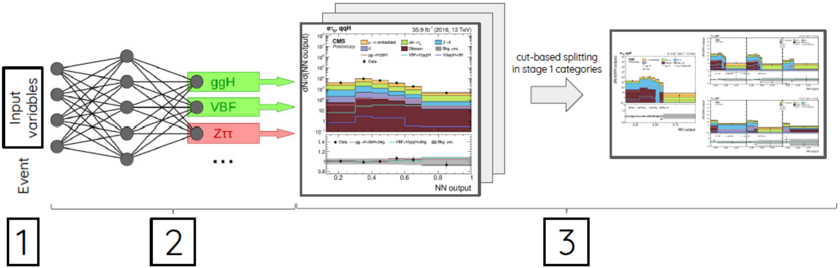
$$\tau_h\tau_h\ (\tau\tau),\ \tau_\mu\tau_h\ (\mu\tau),\ \tau_e\tau_h\ (e\tau),\ \tau_e\tau_\mu\ (e\mu)$$

- **Loose** preselection:
  - Oppositly charged leptons.
  - Minimal p_{T} thresholds required (see table).
  - Background suppressing cuts: e.g. b-jet veto in $e\mu$.

Final state	First object	Second object
$\mathrm{e}\mu^{\dagger}$	$p_{\mathrm{T}}^{\mathrm{e}} > 13\,\mathrm{GeV}$, $\lvert\eta^{\mathrm{e}}\rvert < 2.5$	$p_{\mathrm{T}}^{\mu} > 10\,\mathrm{GeV}$, $\lvert\eta^{\mu}\rvert < 2.4$
$\mathrm{e}\tau_{\mathrm{h}}$	$p_{\mathrm{T}}^{\mathrm{e}} > {26^{\P} \atop 28(25)^{\S}}\,\mathrm{GeV}$, $\lvert\eta^{\mathrm{e}}\rvert < 2.1$	$p_{\mathrm{T}}^{\tau_{\mathrm{h}}} > 30\,\mathrm{GeV}$, $\lvert\eta^{\tau_{\mathrm{h}}}\rvert < 2.3$
$\mu\tau_{\mathrm{h}}$	$p_{\mathrm{T}}^{\mu} > {23(20)^{\P} \atop 25(21)^{\S}}\,\mathrm{GeV}$, $\lvert\eta^{\mu}\rvert < 2.1$	$p_{\mathrm{T}}^{\tau_{\mathrm{h}}} > 30\,\mathrm{GeV}$, $\lvert\eta^{\tau_{\mathrm{h}}}\rvert < 2.3$
$\tau_{\mathrm{h}}\tau_{\mathrm{h}}$	$p_{\mathrm{T}}^{\tau_{\mathrm{h}}} > 40\,\mathrm{GeV}$, $\lvert\eta^{\tau_{\mathrm{h}}}\rvert < 2.1$	

Analysis strategy

1. Selection + validation of input variables.
2. Classification into 2 signal classes (ggH, qqH) and several background classes (e.g. $Z \to \tau\tau$, $t\bar{t}$) $\to$ used as control regions.
3. Cut-based splitting into stage-1 STXS categories.



1.) Selection + validation of input variables

Modelling of input variables checked with Goodness-of-fit tests.

1. Start from the most general set of sensitive variables.
2. Goodness-of-fit tests based on saturated model (including all systematic uncertainties).
3. Check p-value: below 5% $\rightarrow$ discard.

Test modelling of correlations

1. 2-d distributions are "unrolled" to 1-d distributions.
2. Look individually to all p-values.
3. Severe mismodelling of correlations $\rightarrow$ discard variable.

1.) Selection + validation of input variables

- **17-22 variables selected (channel dependent).**

used for 2016

$m_{\tau\tau}$	✓✓	✓✓	✓✓	✓✓

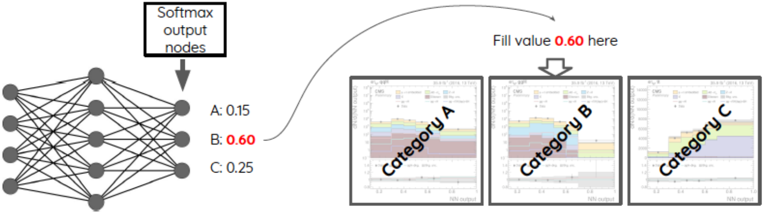
used for 2017

Variable	$e\mu$	$e\tau_h$	$\mu\tau_h$	$\tau_h\tau_h$
$m_{\tau\tau}^{SV}$	✓✓	✓✓	✓✓	✓✓
$m_{T\tau\tau}^{SV}$	✓✓	−−	−−	−−
$p_{T\tau\tau}^{SV}$	✓✓	−−	−−	−−
m_{vis}	✓−	✓−	✓−	✓✓
p_T^{vis}	✓✓	✓✓	✓−	✓−
$p_T^{\tau_1}$	−−	−−	✓−	✓✓
$p_T^{\tau_2}$	✓−	✓✓	✓✓	✓−
$\Delta R^{e\mu}$	✓✓	−−	−−	−−
$p_T(\text{jet}_1)$	✓✓	✓✓	✓✓	✓−
$\eta\ (\text{jet}_1)$	✓−	−−	−−	−−
$p_T(\text{jet}_2)$	✓✓	✓✓	✓✓	✓✓
$\eta\ (\text{jet}_2)$	✓−	−−	−−	−−
m_{jj}	✓✓	✓✓	✓✓	✓✓
$\Delta\eta_{jj}$	✓✓	✓✓	✓✓	✓✓

Variable	$e\mu$	$e\tau_h$	$\mu\tau_h$	$\tau_h\tau_h$
p_T^{jj}	✓✓	✓✓	✓✓	✓✓
$p_T(\text{b jet}_1)$	−−	✓✓	✓✓	✓✓
$p_T(\text{b jet}_2)$	−−	✓✓	✓✓	✓✓
p_T^{miss}	−✓	✓✓	✓−	✓−
D_ζ	✓✓	−−	−−	−−
m_T^{e}	−−	✓✓	−−	−−
m_T^{μ}	✓✓	−−	✓✓	−−
$m_T^{e+\mu}$	✓−	−−	−−	−−
$\max(m_T^{\mu}, m_T^{e})$	✓✓	−−	−−	−−
$m_T^{\tau_h}$	−−	✓✓	✓−	✓✓
$p_T^{\tau\tau+miss}$	✓✓	✓✓	✓✓	✓✓
$p_T^{\tau\tau jj+miss}$	−✓	−−	−−	−−
$N_{b\ jet}$	−−	✓✓	✓✓	✓✓
N_{jet}	✓✓	✓✓	✓✓	−✓

2.) Multi-class neural network for event classification

- **Two step approach:**
  1. Training: supervised learning with simulated samples.
  2. Prediction: 90% of background processes are data-driven.
- **Targets:** signal and background classes/categories.
- **Output:** probability that an event belongs to a certain class.



Event is classified according to the highest probability

2.) Multi-class neural network for event classification

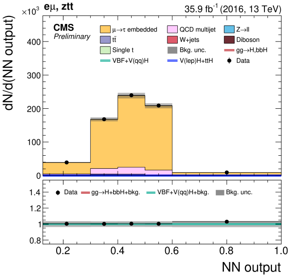
- Training of neural networks per year and per channel.
- To exploit full statistical power of MC samples $\rightarrow$ two folds are trained.
- All processes weighted equally in the loss function.

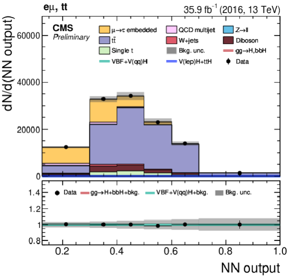
<table>
<tr><th></th><th colspan="4">Classes/Categories per final state</th></tr>
<tr><th>Process</th><th>$e\mu$</th><th>$e\tau_h$</th><th>$\mu\tau_h$</th><th>$\tau_h\tau_h$</th></tr>
<tr><td>$gg \rightarrow H$</td><td>ggH (0.20)</td><td>ggH (0.23)</td><td>ggH (0.27)</td><td>ggH (0.54)</td></tr>
<tr><td>VBF</td><td>qqH (0.74)</td><td>qqH (0.72)</td><td>qqH (0.72)</td><td>qqH (0.57)</td></tr>
<tr><td>$Z \rightarrow \tau\tau$</td><td>ztt (0.52)</td><td>ztt (0.66)</td><td>ztt (0.63)</td><td>ztt (0.62)</td></tr>
<tr><td>QCD</td><td>qcd (0.45)</td><td>qcd (0.21)</td><td>qcd (0.17)</td><td>qcd (0.48)</td></tr>
<tr><td>$t\bar{t}$</td><td>tt (0.55)</td><td>tt (0.79)</td><td>tt (0.75)</td><td rowspan="5">misc (0.45)</td></tr>
<tr><td>$Z \rightarrow \ell\ell$</td><td rowspan="2">misc (0.24)</td><td>zll (0.55)</td><td>zll (0.53)</td></tr>
<tr><td>W+jets</td><td>wj (0.43)</td><td>wj (0.51)</td></tr>
<tr><td>Diboson</td><td>db (0.46)</td><td rowspan="2">misc (0.21)</td><td rowspan="2">misc (0.28)</td></tr>
<tr><td>Single t</td><td>st (0.30)</td></tr>
</table>

2.) Multi-class neural network for event classification

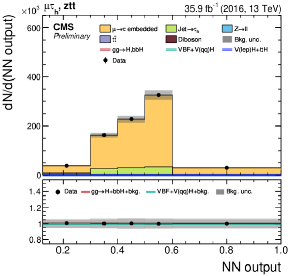
2016 exemplary background classes:

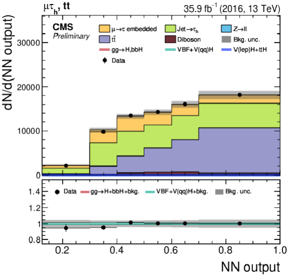
***eμ*-channel**





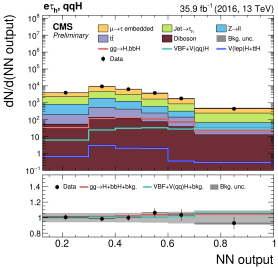
***μτ*-channel**

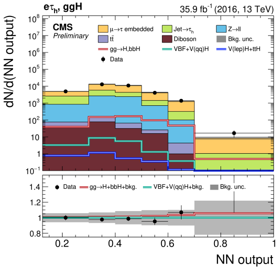




2.) Multi-class neural network for event classification

2016 $e\tau_h$ signal classes:





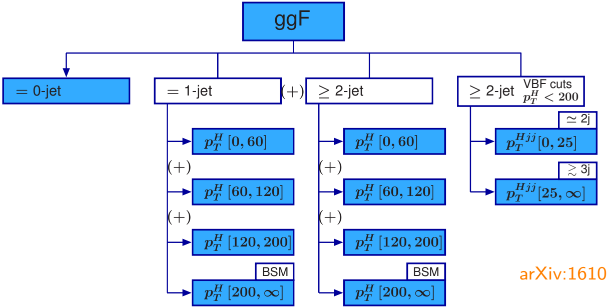
To further improve sensitivity:

→ Cut-based splitting of inclusive classes.

→ Cuts (on reconstructed variables) aligned with the generator-level STXS stage-1 categories.

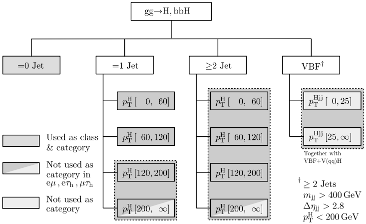
Simplified template cross sections (STXS)

- Defined in the “LHC Higgs Cross Section Working Group“.
- Intention: easier combination and BSM interpretation.
- Stage-1 categories used (got updated in the meanwhile stage 1.1).



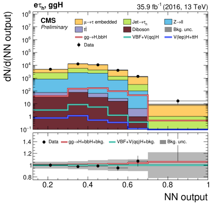
arXiv:1610.07922

3.) Cut-based splitting into STXS stage-1 categories

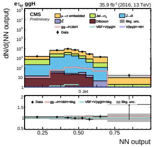


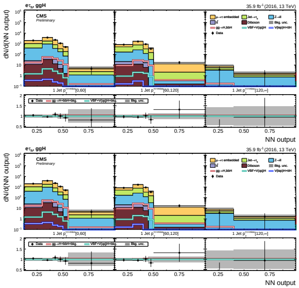
One inclusive category

Seven STXS stage-1 motivated categories



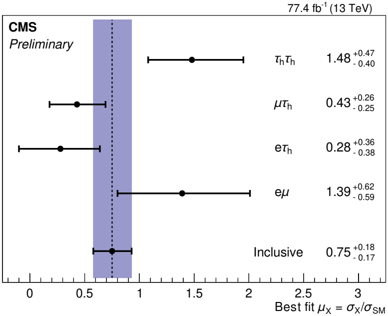
→



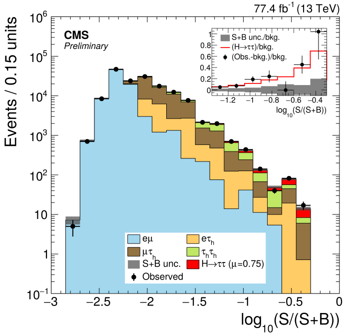


Results: Inclusive

- Background + cut-based splitted signal categories → statistical inference.

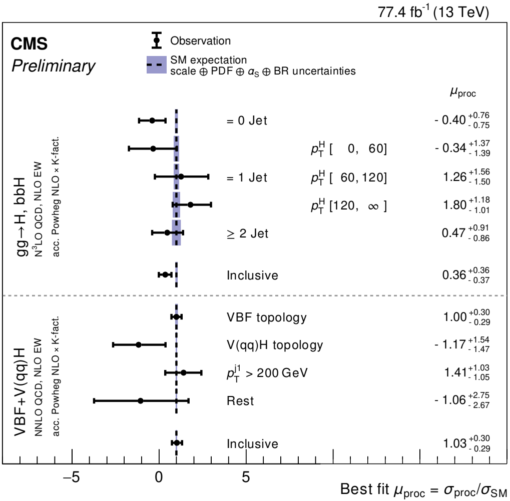


- **Visualization of excess.**



- **Signal strength** $\mu = 0.75^{+0.18}_{-0.17}$.

Results: stage-0 and stage-1



- Due to missing sensitivity: Several STXS stage-1 categories merged.

Conclusion

- First results in $H \to \tau\tau$ with the use of neural networks.
  - CMS-PAS-HIG-18-032
- First measurement of stage-1 STXS.
- Further improvements under study for full Run-II result.

Thank you.

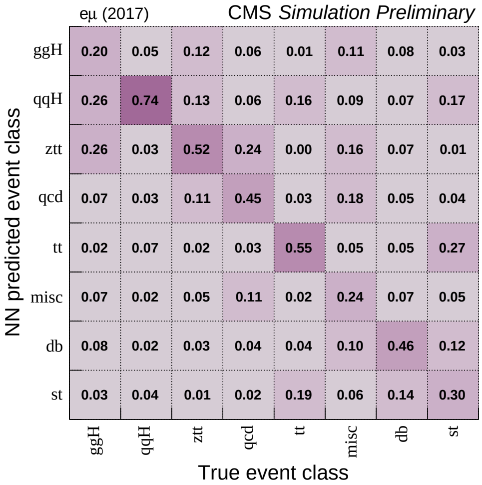
BACKUP

2.) Multi-class neural network for event classification

Confusion matrix ($e\mu$):

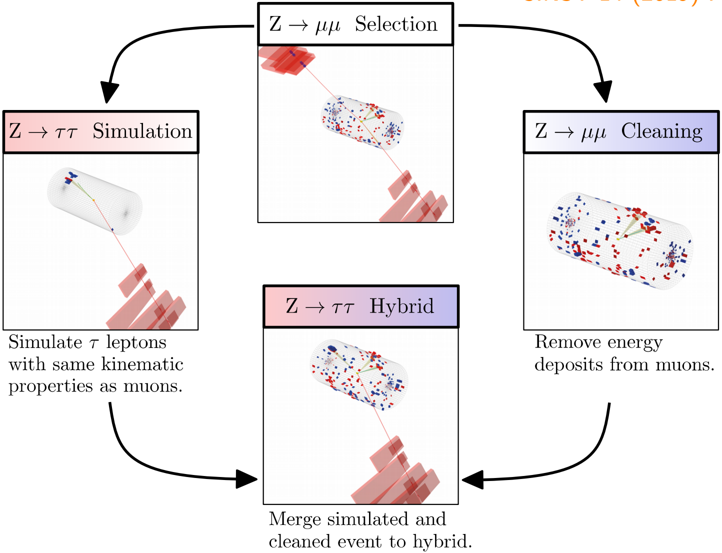
- Efficiency matrix (columns normalized to one).

→ More efficient categorization of qqh than ggh events.



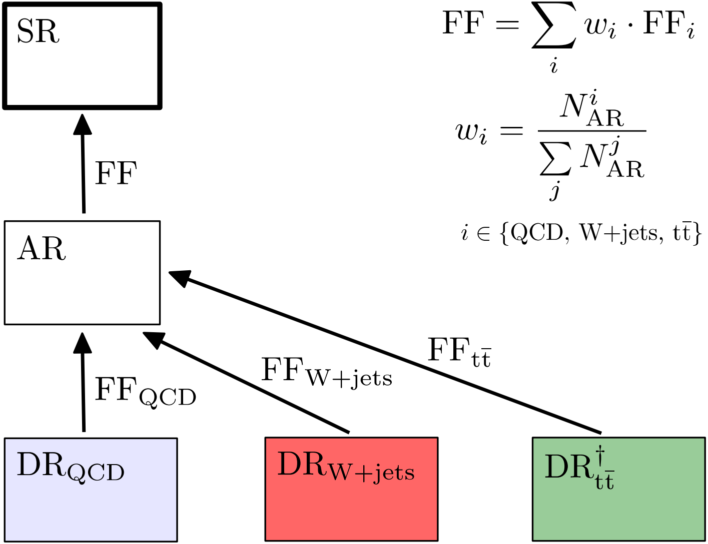
τ-Embedding

JINST 14 (2019) P06032



Fake factor method: Jet $\rightarrow \tau_h$

JHEP 09 (2018)007



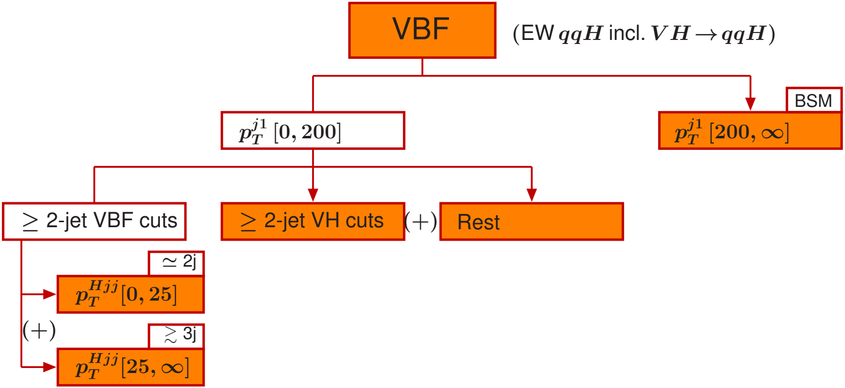
$$\mathrm{FF} = \sum_i w_i \cdot \mathrm{FF}_i$$

$$w_i = \frac{N^i_{\mathrm{AR}}}{\sum_j N^j_{\mathrm{AR}}}$$

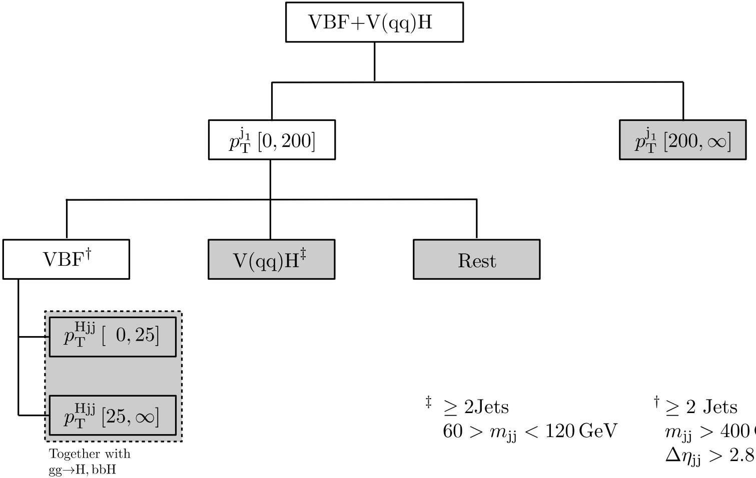
$i \in \{\mathrm{QCD,\ W{+}jets,\ t\bar{t}}\}$

†: taken from simulation

STXS categories: qqH



STXS categories for $H \rightarrow \tau\tau$ analysis: qqH



‡ $\geq$ 2Jets
$60 > m_{jj} < 120\,\mathrm{GeV}$

† $\geq$ 2 Jets
$m_{jj} > 400\,\mathrm{GeV}$
$\Delta\eta_{jj} > 2.8$