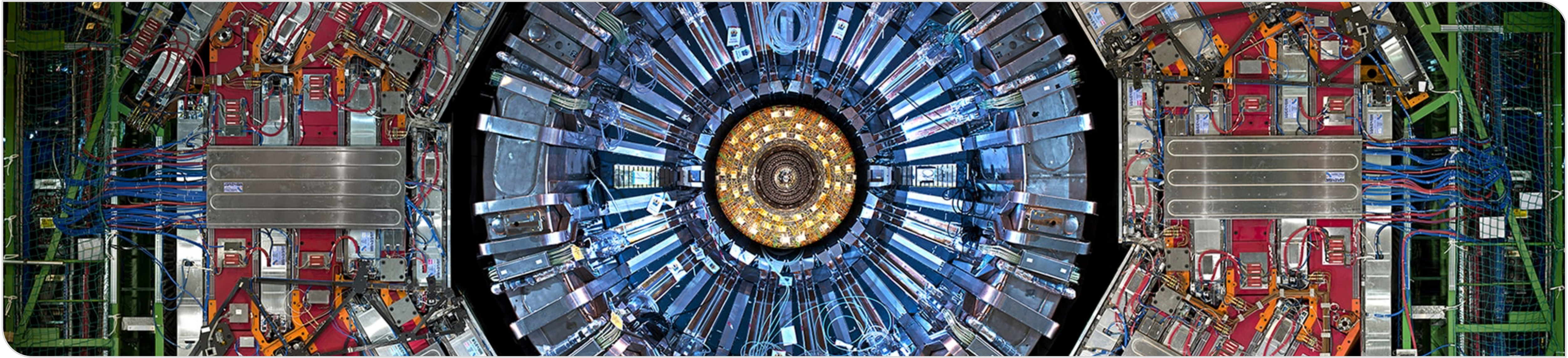
Karlsruhe Institute of Technology

Institute of Experimental Particle Physics

Probing multivariate methods to study $\mathrm{H} \rightarrow \mu\mu$ events in the VH and VBF channels

Erik Hettwer Benitez

erik.hettwer@gmail.com

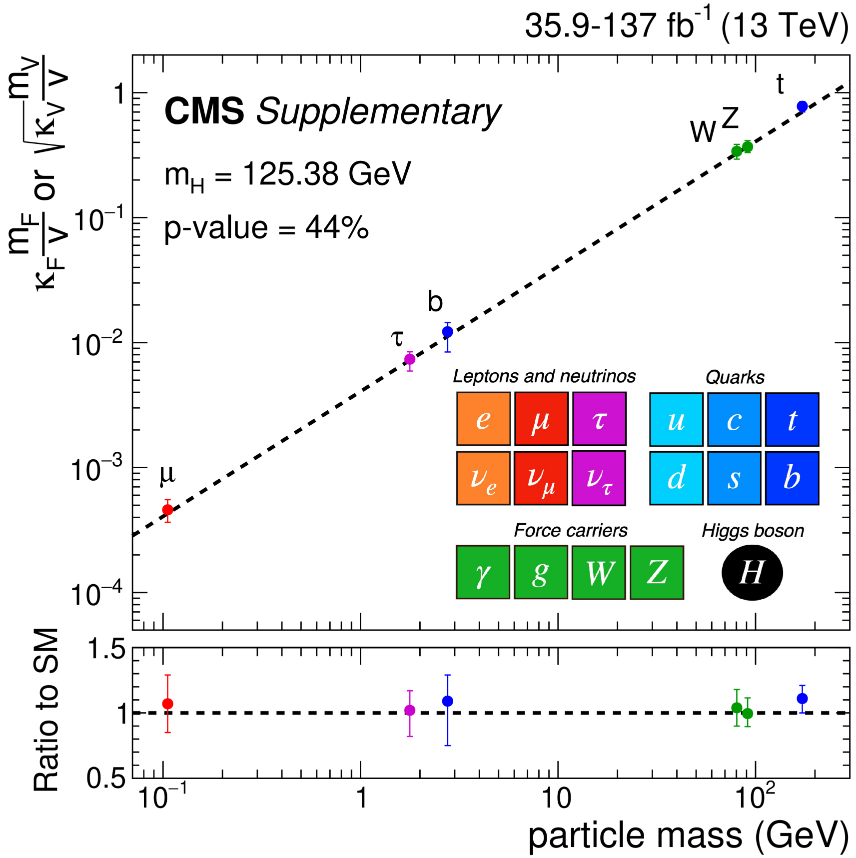


Motivation

„Why does mass appear? We don't know. If we understood that, we'd know a lot more about the whole world.“

R. Feynman

- The Yukawa coupling of the Higgs boson to second generation fermions is of great interest
- Clean final state with high sensitivity due to well understood muons
- Rare decay due to the small muon mass
- High luminosity enables this search



Higgs coupling to fermions

Source: https://ep-news.web.cern.ch/

Overview

- Introduction
- Run 2 analysis strategy
- Decay signature of $H \rightarrow \mu\mu$ and background processes
- Event Selection
- Invariant dimuon mass spectrum for WH and VBF
- Neural Networks
- $H \rightarrow \mu\mu$ with NN for WH and VBF Higgs boson production
- Summary of expected upper limits on the signal strength parameter
- Conclusion

Previous Searches

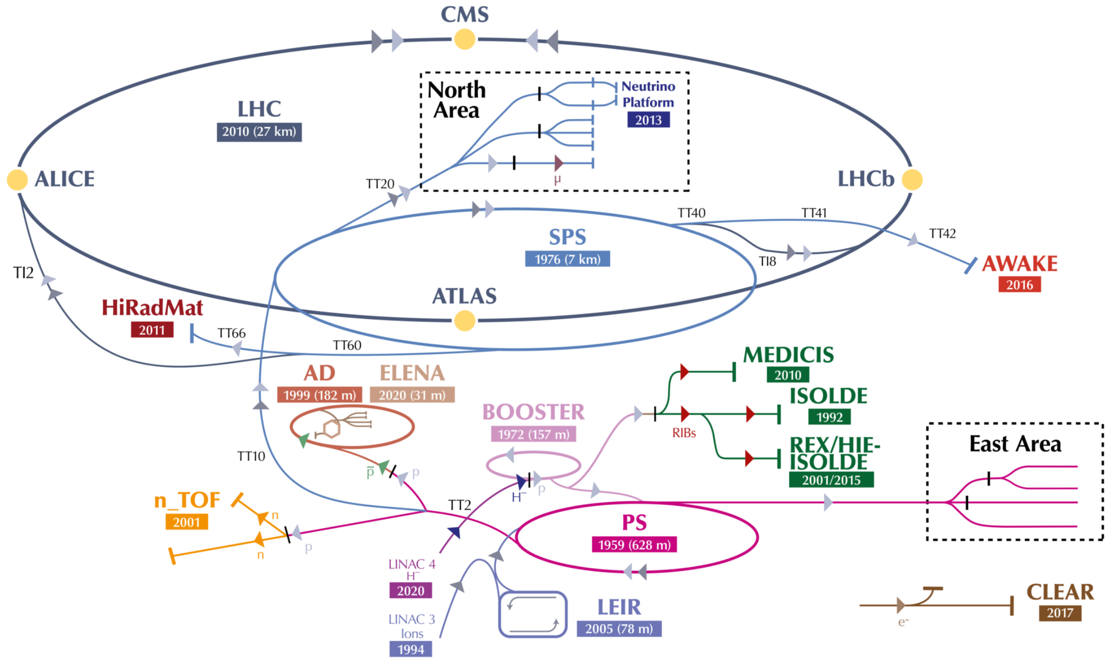
Institute of Experimental Particle Physics

- Measurements have been performed at CMS for LHC Run 2 data
- Branching Ratio was measured to be within $0.7 \times 10^{-4} < \mathscr{BR}(\mathrm{H} \rightarrow \mu\mu) < 4.5 \times 10^{-4}$ at 95% CL
- There is evidence at 3 σ for this decay, but no observation can be claimed so far
- Analysis methods in the analysis from 2021 vary between Higgs production channels

- Different methods have been probed in this work to explore possible pathways for a Run 3 analysis
- No systematic uncertainties have been taken into account in this work, so further studies are needed

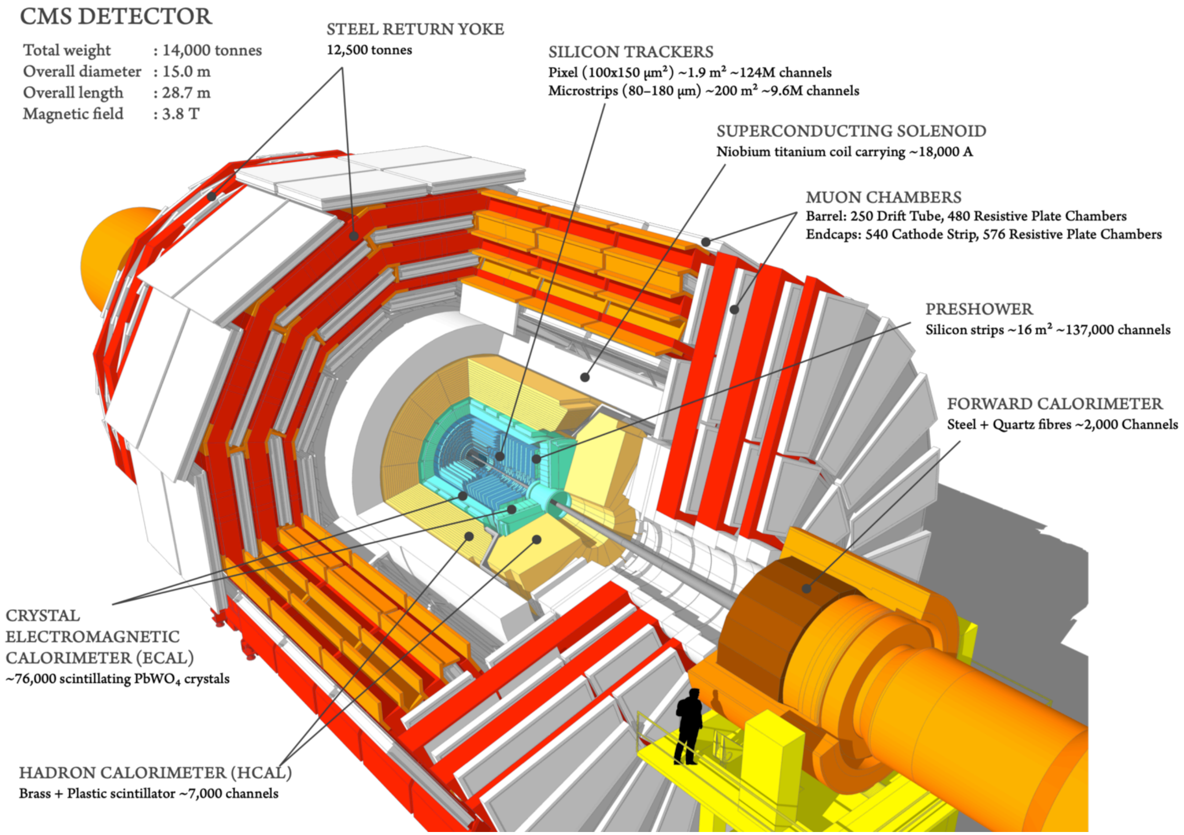
The Large Hadron Collider (LHC)

- High energy particle accelerator that collides protons at 13 TeV
- Four main experiments (ATLAS, CMS, LHCb, ALICE)
- Integrated Luminosity in 2018 (year of subject to this study): $L = 67.86\,\text{fb}^{-1}$
- High energy collisions produce particles of interest like the Higgs boson
- These collisions are studied at the individual experiments



The CMS Experiment

- Hermetic nearly 4π multi purpose particle detector
- Components (inner to outer layers):
  - Silicon Pixel Detector
  - Silicon Strip Detector
  - Electromagnetic Calorimeter
  - Hardon Calorimeter
  - Solenoid Magnet
  - Muon Chambers

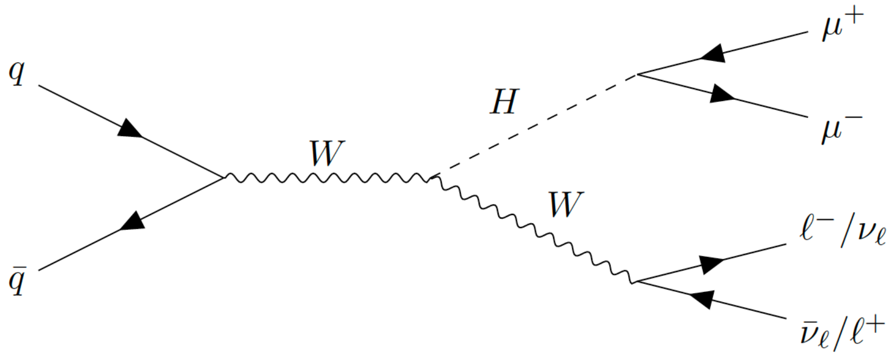


Analysis Strategy from 2021

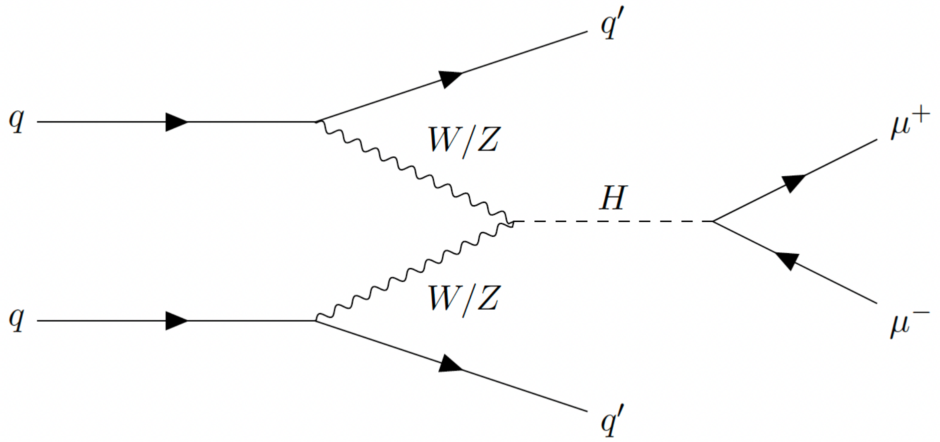
- **ggH, VH, $t\bar{t}$ - Higgs production:**
  - Phase Space selection
  - Training of multiple BDTs on variables
  - Selection on classifier output to define samples of different signal purity
  - Dimuon mass ML fit on those samples
  - Calculation of CLs for each individual sample and combination
- **VBF Higgs production:**
  - Phase space selection
  - Training of deep neural networks on variables including the dimuon mass
  - ML fit of the classifier output
  - Calculation of CLs

Decay signature of H → $\mu\mu$ and background processes

- WH Higgs production (lept. W decay):
  - Two opposite-sign muons
  - Invariant muon mass expected close to Higgs mass
  - One Additional visible lepton



- VBF Higgs production:
  - Two opposite-sign muons
  - Invariant muon mass expected close to Higgs mass
  - Two jets from quark interaction



Decay signature of $\mathrm{H} \to \mu\mu$ and background processes

- Any process that produces the same final state is considered a background process
- The primary event selection should reduce these backgrounds as much as possible while preserving the signal
- Cuts on kinematic variables are used to reduce the total phase space

Event Selection Criteria:

- HLT_IsoMu24 trigger
- Loose muon isolation
- $(p_\mathrm{T})_1 > 26\,\mathrm{GeV}$ (offline)
- $(p_\mathrm{T})_{2/3} > 20\,\mathrm{GeV}$

WH leptonic phase space:

- Sum of lepton charges is uneven
- $m_{\mu\mu,\mathrm{cand.Higgs}} \in [110{,}150]\,\mathrm{GeV}$
- If $m_{\mu\mu} \in [81{,}101]\,\mathrm{GeV}$ -> discard
- No medium b-tagged jets, max two loose b-tagged jets

VBF phase space:

- Sum of lepton charges is even
- At least two jets (sorted by p_T)
- $(p_\mathrm{T})_{\mathrm{j}1} > 35\,\mathrm{GeV}$
- $(p_\mathrm{T})_{\mathrm{j}2} > 25\,\mathrm{GeV}$
- $m_{\mathrm{jj}} > 400\,\mathrm{GeV}$
- $|\eta_{\mathrm{jj}}| > 2.5$
- No medium b-tagged jets, max one loose b-tagged jet

Decay signature of H $\rightarrow \mu\mu$ and background processes

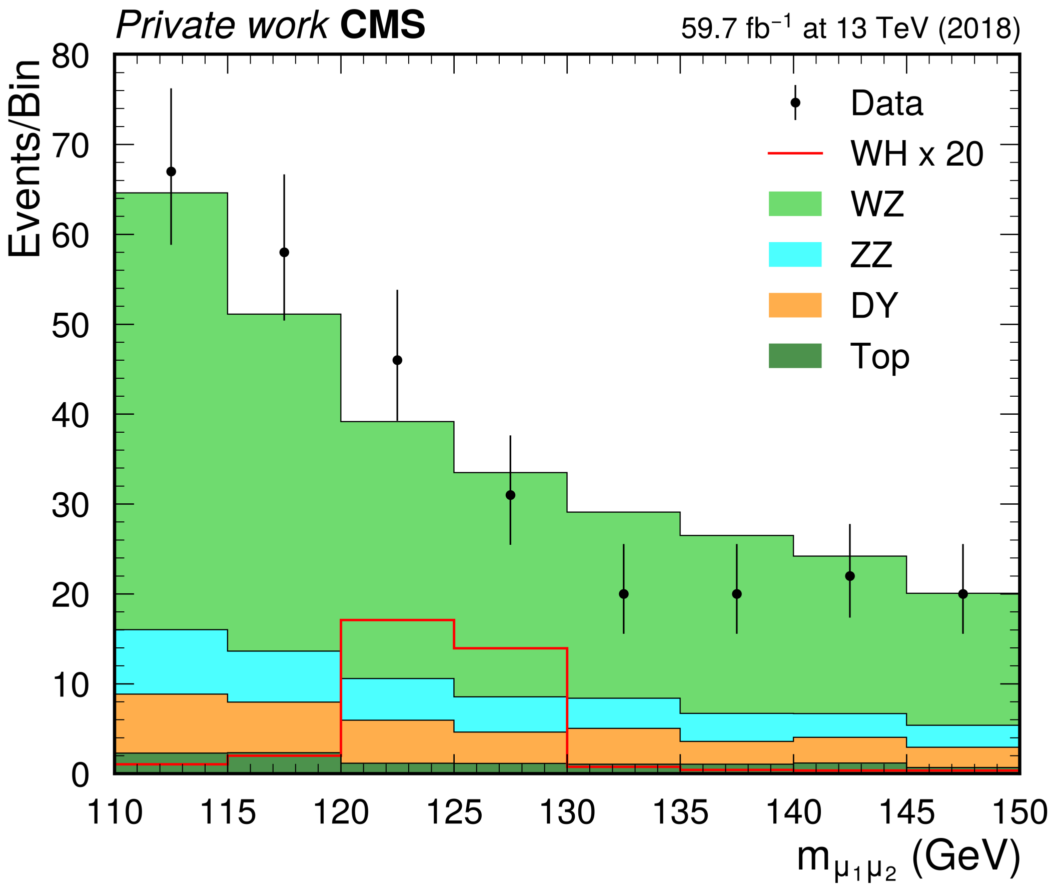
- Any process that produces the same final state is considered a background process
- The primary event selection should reduce these backgrounds as much as possible while preserving the signal
- Cuts on kinematic variables are used to reduce the total phase space

- Backgrounds WH (leptonic):
  - $WZ \rightarrow 3\ell + \nu$
  - $ZZ \rightarrow 4\ell$
  - Drell Yan
  - $t\bar{t}$
- Backgrounds VBF:
  - Drell Yan
  - $t\bar{t}$

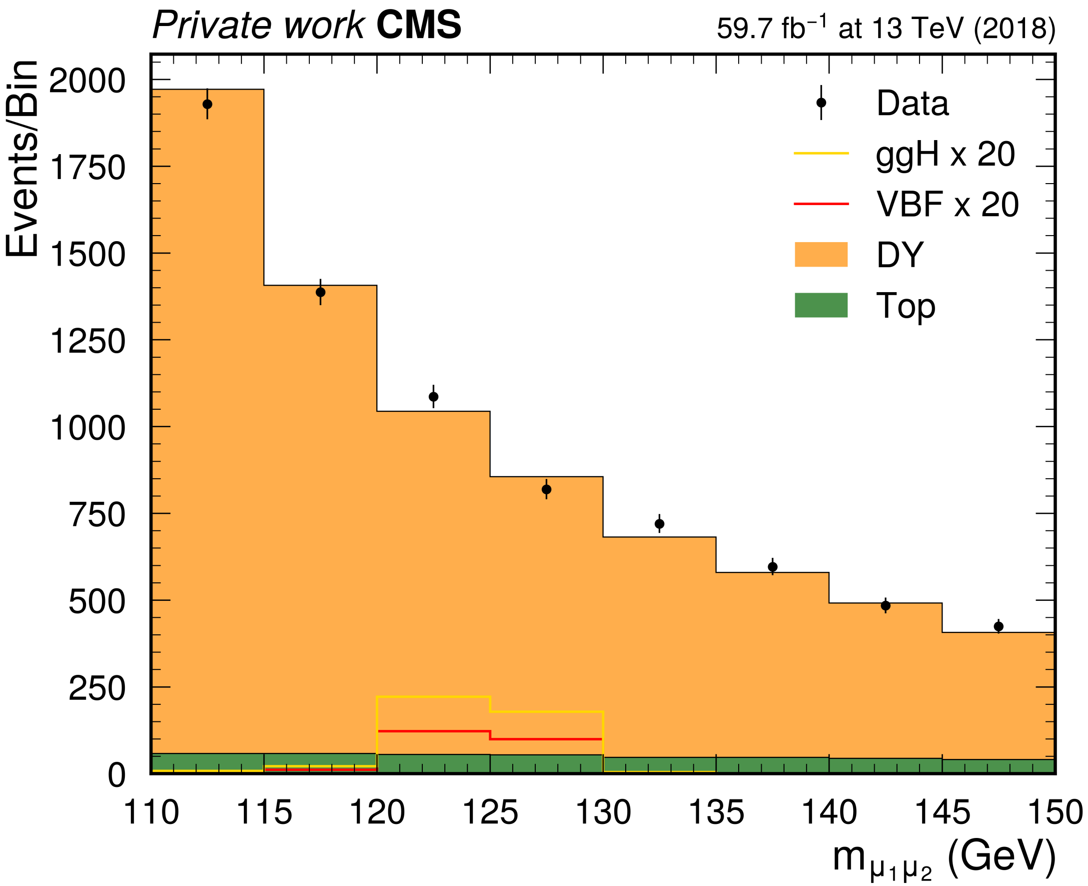
Invariant Dimuon Mass Spectrum

Institute of Experimental Particle Physics

WH leptonic phase space:



VBF phase space:



ML Fit of the Dimuon Mass Spectrum

- Goal: Calculate expected upper limits on the signal strength parameter

$$\mu = \frac{\sigma_{\text{observed}}}{\sigma_{\text{SM}}}$$

- It Compares the observed signal to the Standard Model, with $\mu = 1$ matching the prediction, $\mu > 1$ indicating excess, and $\mu < 1$ showing a deficit.
- The signal confidence level is defined as:

$$\text{CL}_\text{s} = \frac{\text{CL}_\text{s+b}}{\text{CL}_\text{b}} \qquad \text{with } \text{CL}_\text{s+b} = 1 - p_{s+b} \text{ and } \text{CL}_\text{b} = 1 - p_b$$

- Avoids overly conservative results by including the possibility of background fluctuations

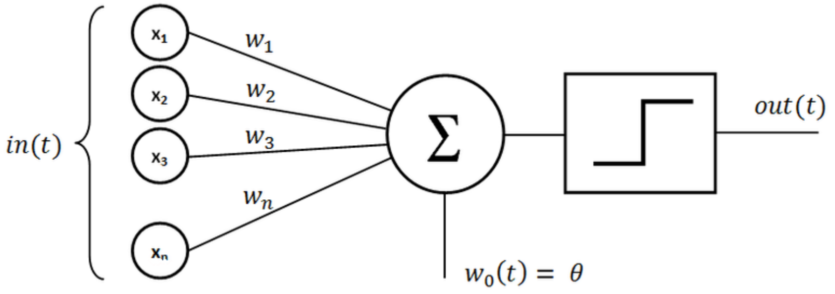
Fit Results Dimuon Mass Fit

- The resulting distributions have been fit using COMBINE under a background only hypothesis
- An expected 95% CLs upper limit is calculated on the signal strength parameter μ

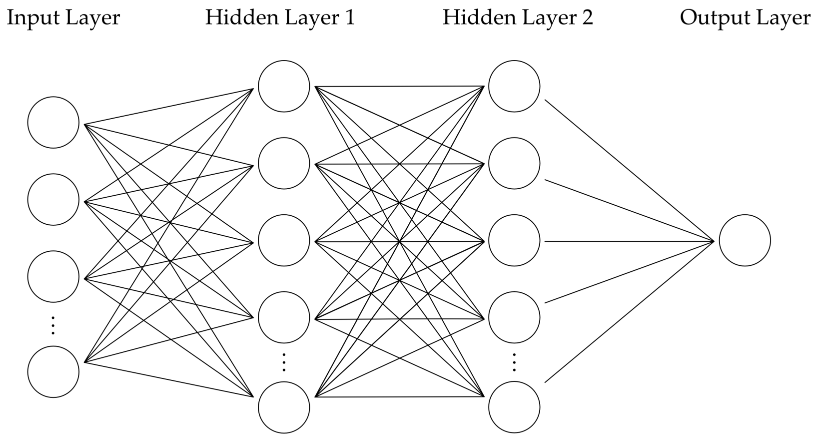
	WH leptonic	VBF
Expected upper Limit	$\mu = 10.19^{+4.47}_{-3.01}$	$\mu = 3.64^{+1.45}_{-1.04}$
Expected Significance	$0.21\,\sigma$	$0.55\,\sigma$

Interlude: Neural Networks

- **Perceptron:** A basic unit in neural networks that computes a weighted sum of inputs.
- **Activation:** An activation function determines the output.
- **Linear Limit:** A single perceptron solves linearly separable problems.
- **MLP:** A multi-layer perceptron connects neurons to solve non-linear problems.
- **Hidden Layers:** Layers between input and output that learn abstract features.
- **Weights & Biases:** Values that adjust how inputs influence the output during training.

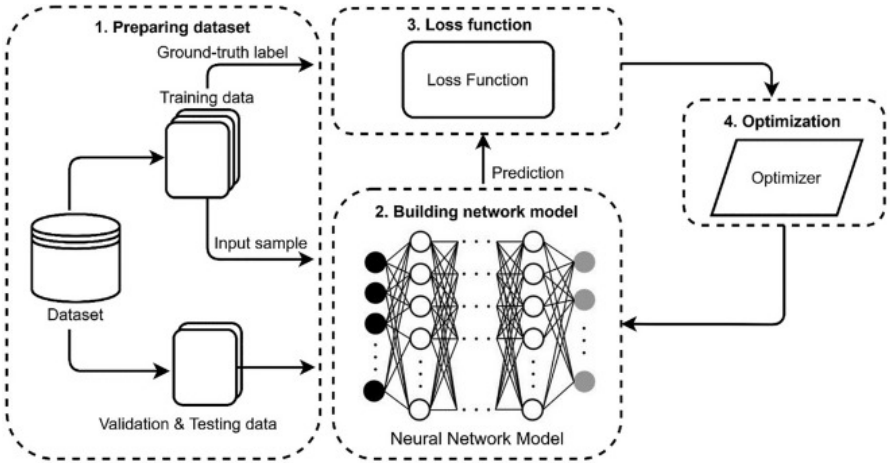


Source: https://upload.wikimedia.org/.../Perceptron_moj.png



Interlude: Neural Networks

- **Training:**
  - Feed data through the network (feedforward).
  - Calculate the error using a loss function.
  - Use back-propagation to update weights and biases.
  - Repeat until the model converges or reaches a stopping criterion.
- **Inference:**
  - Feed new data into the trained network.
  - Generate predictions using the learned weights.



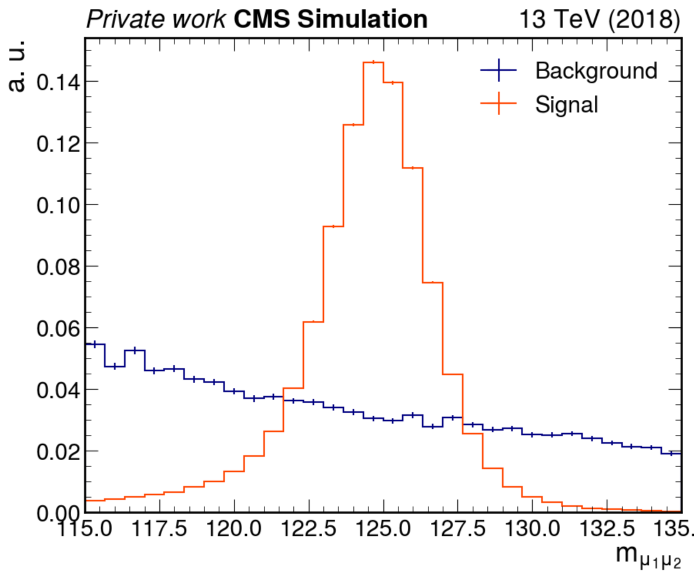
Source: https://www.sciencedirect.com/.../neural-network-training

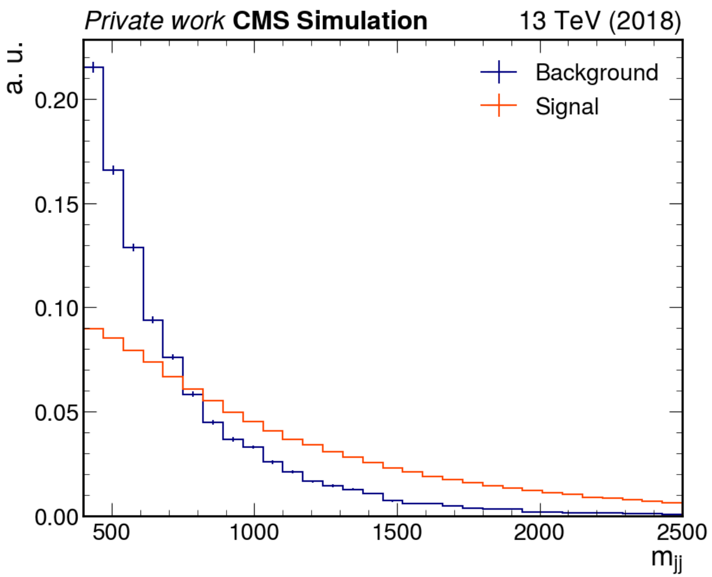
Event Categorization using Neural Networks

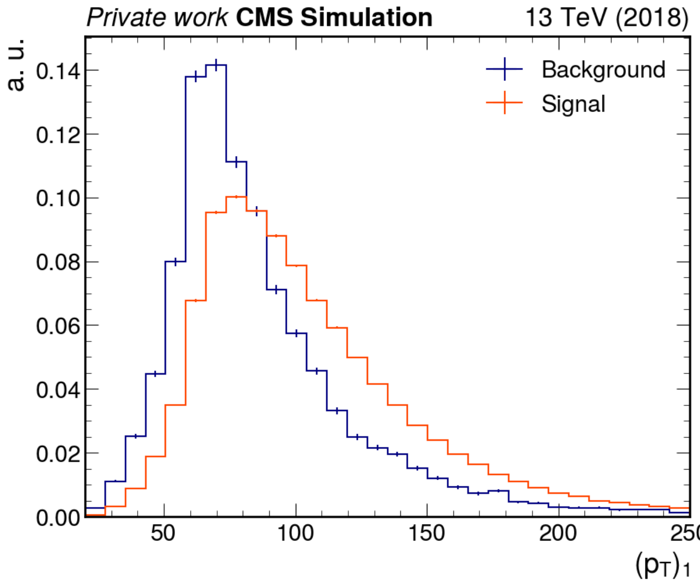
- The main discriminating variable in this problem is the invariant dimuon mass
- Other variables do also possess discriminating power
- Utilizing more variables may lead to better signal sensitivity

Event Categorization using Neural Networks

- The main discriminating variable in this problem is the invariant dimuon mass
- Other variables do also possess discriminating power
- Utilizing more variables may lead to better signal sensitivity





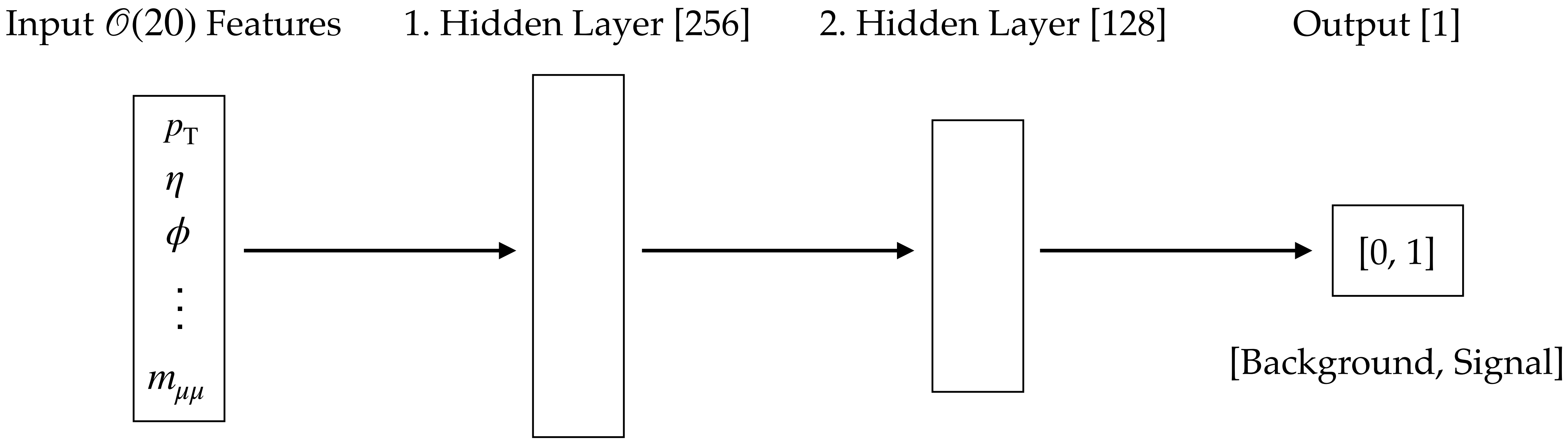


VBF: dimuon mass | dijet mass | leading muon transverse momentum

Event Categorization using Neural Networks

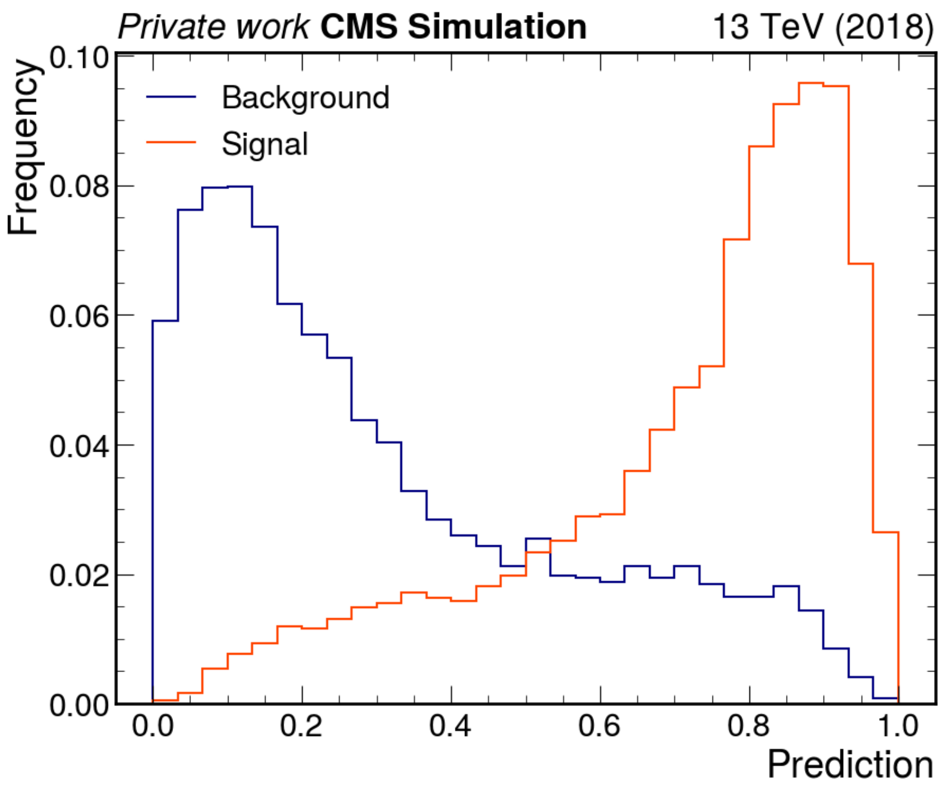
Institute of Experimental Particle Physics

- The main discriminating variable in this problem is the invariant dimuon mass
- Other variables do also possess discriminating power
- Utilizing more variables may lead to better signal sensitivity



Training Process

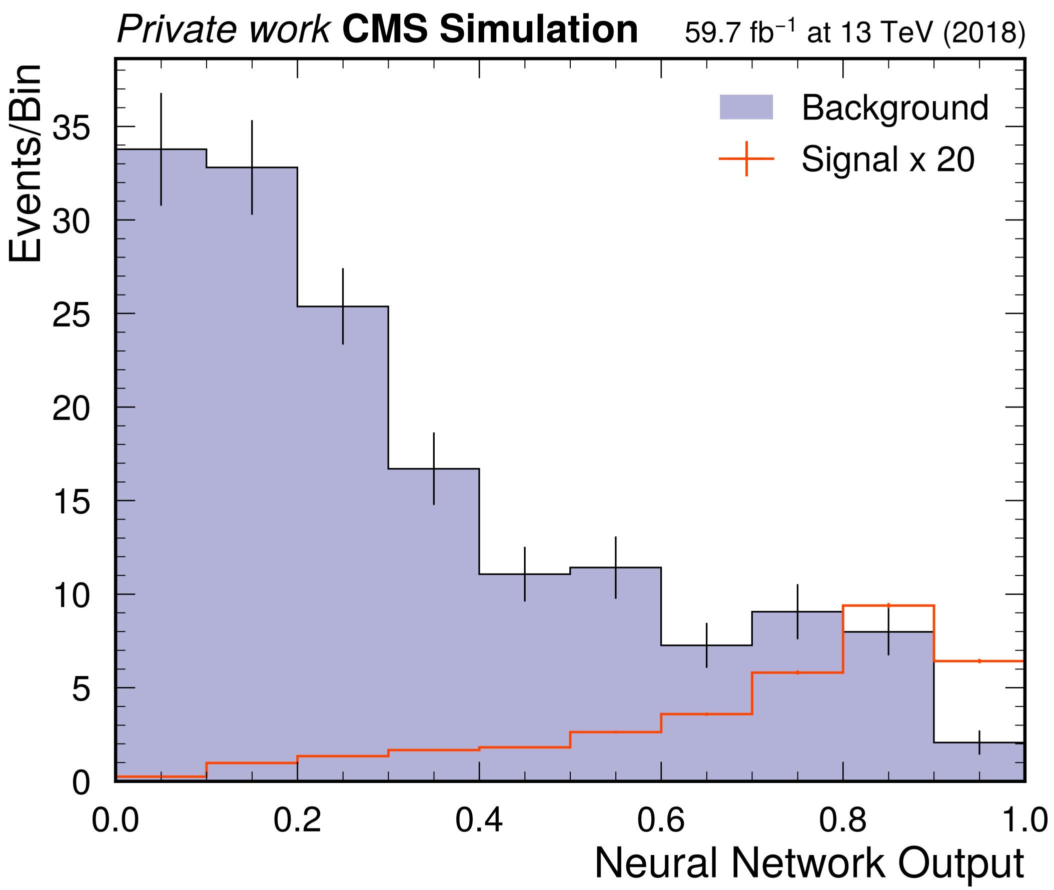
- Neural Network is trained on Monte Carlo (MC) simulated events
- Training/Validation dataset split of 4/1
- Uses batch normalisation as wells as Dropout of 0.5 to avoid over-fitting
- ReLU activation between layers + sigmoid activation for output layer
- Bayesian optimisation used to find optimal hyper-parameters



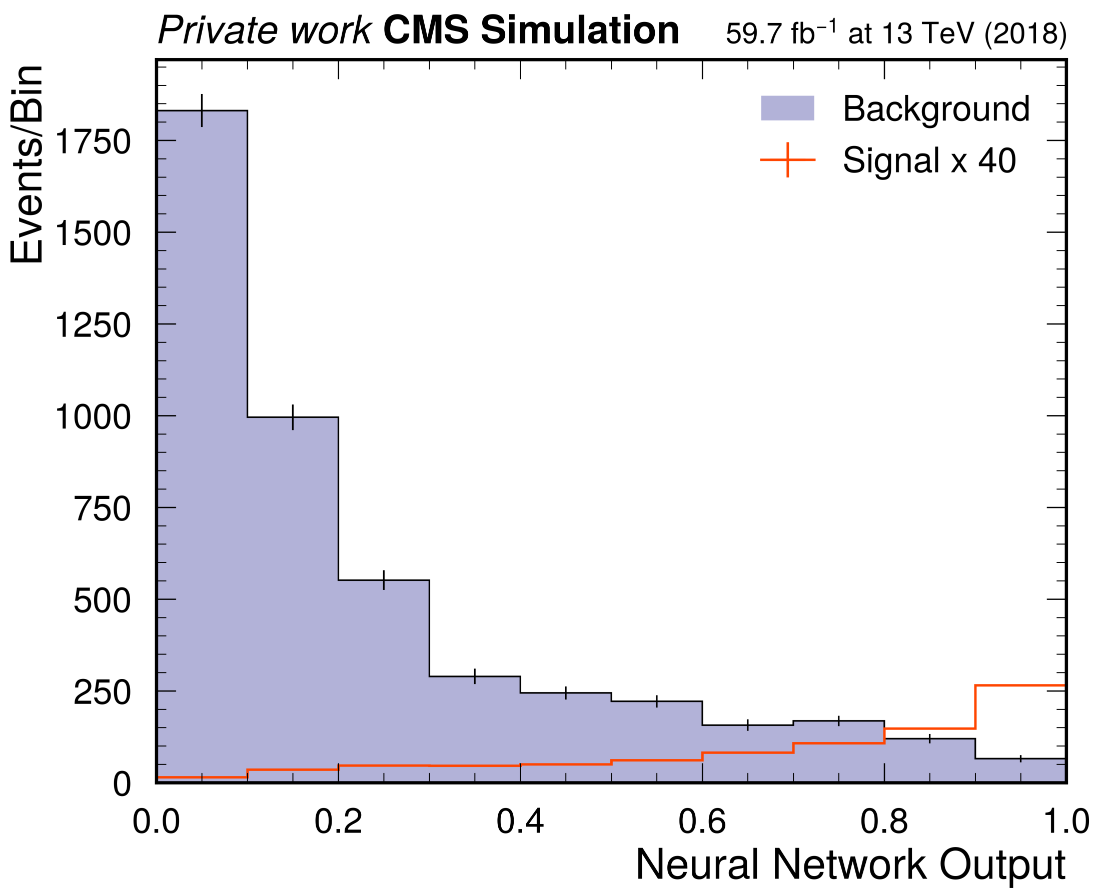
WH leptonic NN output inclusive $m_{\mu\mu}$

Neural Network Output

WH leptonic phase space:



VBF phase space:



Fit results NN Output

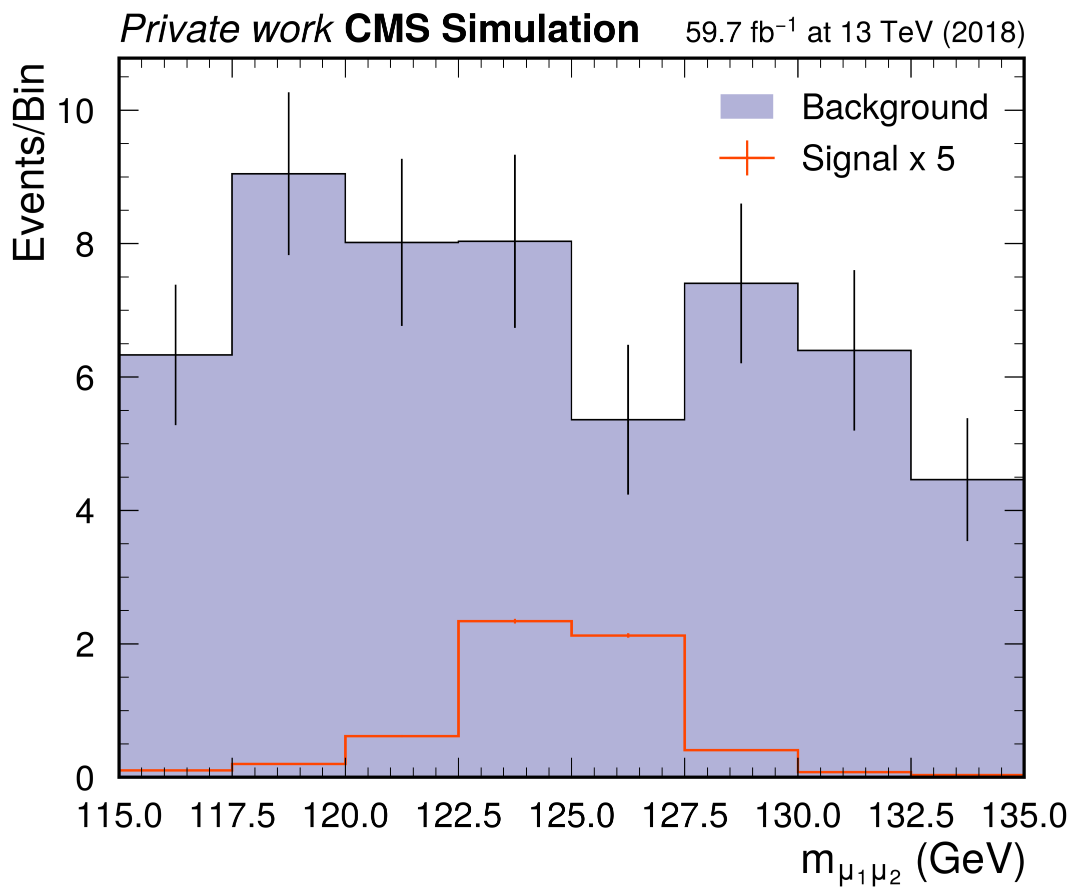
Institute of Experimental Particle Physics

- The resulting distributions have been fit using COMBINE under a background only hypothesis
- An expected 95% CLs upper limit is calculated on the signal strength parameter μ

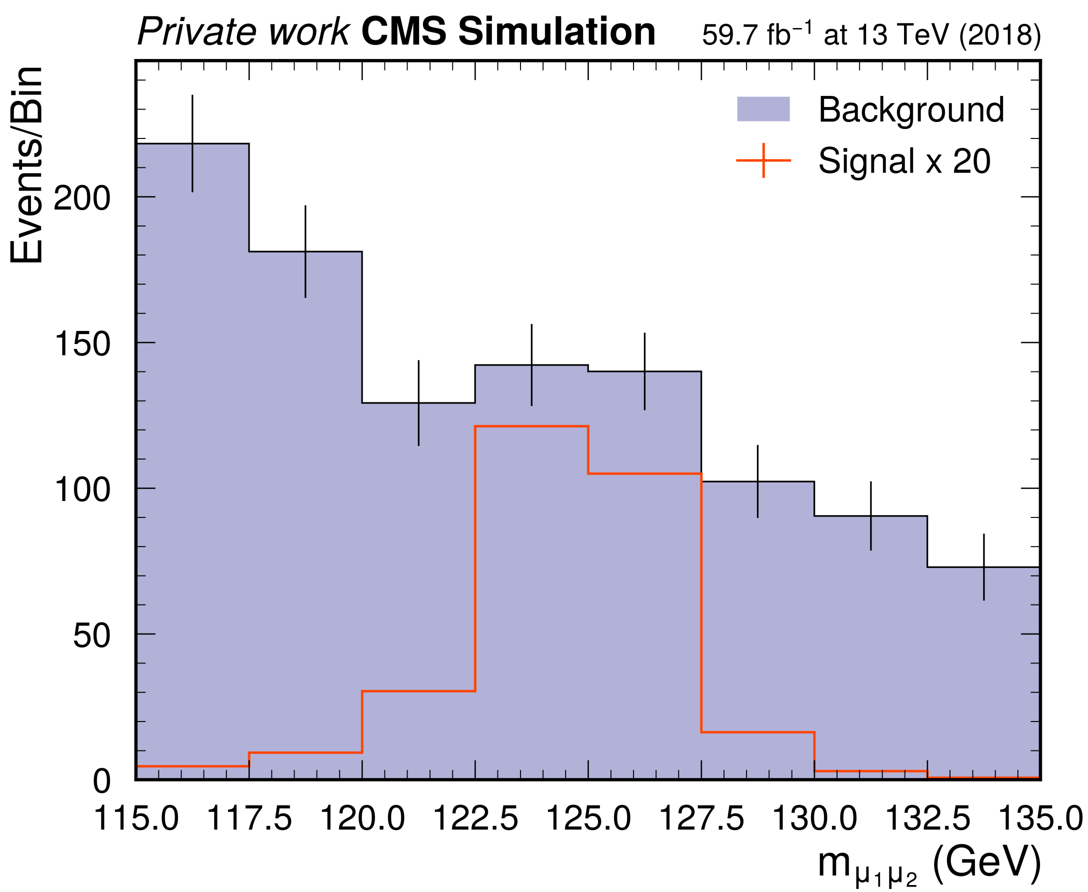
	WH leptonic	VBF
Expected upper Limit	$\mu = 7.72^{+3.66}_{-2.42}$	$\mu = 2.21^{+0.92}_{-0.64}$
Expected Significance	$0.3\,\sigma$	$0.93\,\sigma$

Neural Network Output Cut + Dimuon Mass Fit

WH leptonic phase space:



VBF phase space:

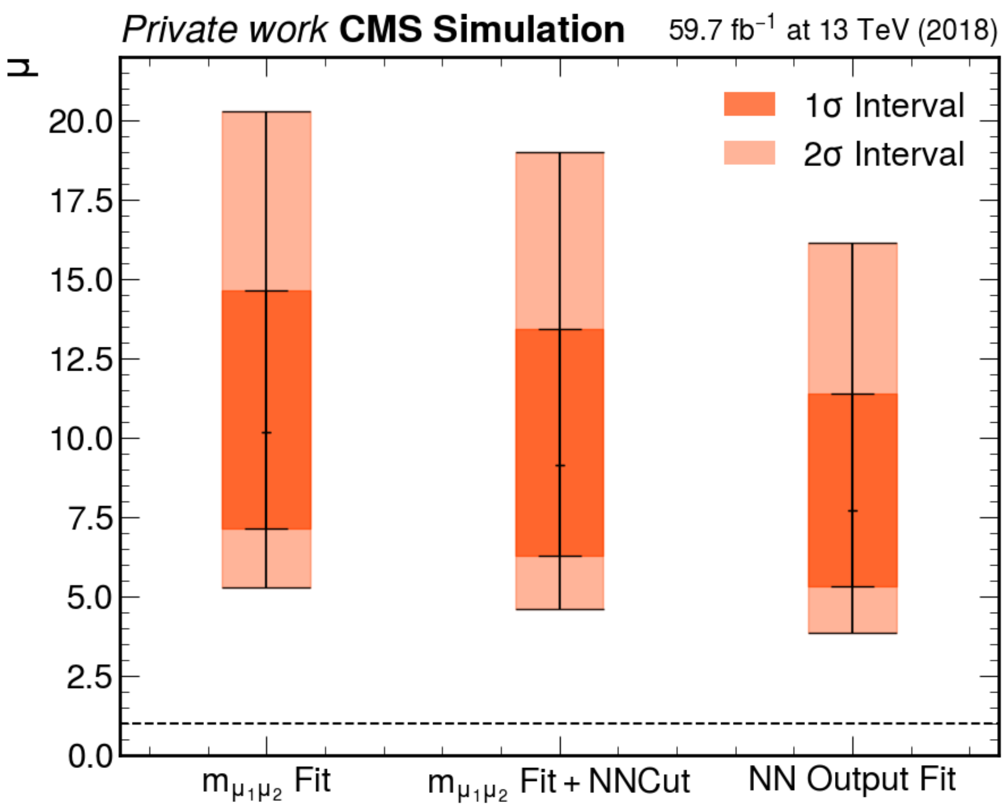


Results NN Output Cut + Dimuon Mass Fit

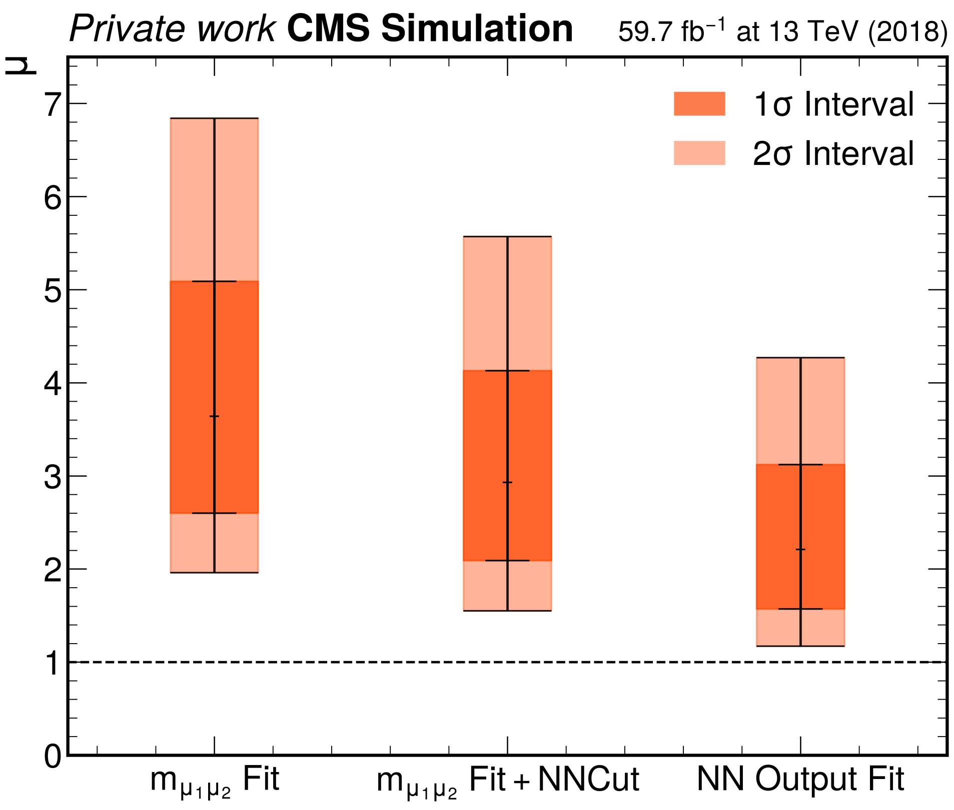
- The resulting distributions have been fit using COMBINE under a background only hypothesis
- An expected 95% CLs upper limit is calculated on the signal strength parameter μ

	WH leptonic	**VBF**
Expected upper Limit	$\mu = 9.13^{+4.29}_{-2.83}$	$\mu = 2.93^{+1.20}_{-0.84}$
Expected Significance	$0.25\,\sigma$	$0.69\,\sigma$

Summary of Expected upper limits though different methods



WH leptonic



VBF

Conclusion

- Expected upper limits on the signal strength parameter where set in two Higgs production channels using three different methods:
  - Fitting the invariant dimuon mass spectrum
  - Fitting the classifier output of a deep neural network
  - Fitting the dimuon mass spectrum after cutting on the classifier
- The best result is obtained when fitting the classifier output directly for both channels
- Care must be taken when using this approach to model all contributions correctly!
- The neural network works as a de facto black box which carries an inherent risk
- Both multivariate methods carry this risk but one seems to perform better

Training Variables

WH leptonic phase space

Variable Name	Description
$p_T(\mu_1)$	transverse momentum of the leading muon
$p_T(\mu_2)$	transverse momentum of the sub-leading muon
$p_T(\ell_3)$	transverse momentum of the remaining lepton
$m_{\mu_1\mu_2}$	dimuon mass of Higgs boson candidate
$p_T(\mu_1\mu_2)$	transverse momentum of Higgs candidate
$\eta(\mu_1\mu_2)$	pseudorapidity of Higgs candidate
$\phi(\mu_1\mu_2)$	azimuth angle of Higgs candidate
$\Delta\eta_{13}$	difference in pseudorapidity between μ_1 and ℓ_3
$\Delta\eta_{23}$	difference in pseudorapidity between μ_2 and ℓ_3
$\Delta\eta_{12;3}$	difference in pseudorapidity between Higgs cand. and ℓ_3
$\Delta\phi_{13}$	difference in azimuth angle between μ_1 and ℓ_3
$\Delta\phi_{23}$	difference in azimuth angle between μ_2 and ℓ_3
$\Delta\phi_{12;3}$	difference in azimuth angle between Higgs cand. and ℓ_3
ΔR_{12}	measure of separation between μ_1 and μ_2
ΔR_{13}	measure of separation between μ_1 and ℓ_3
ΔR_{23}	measure of separation between μ_2 and ℓ_3
$\cos\theta^*_{12}$	opening angle between μ_1 and μ_2 in the μ_1 rest frame
$\cos\theta^*_{13}$	opening angle between μ_1 and ℓ_3 in the μ_1 rest frame
$\cos\theta^*_{23}$	opening angle between μ_2 and ℓ_3 in the μ_2 rest frame
η_1	pseudorapidity of μ_1
η_2	pseudorapidity of μ_2
q_n	charge of the lepton n

VBF phase space

Variable Name	Description
$p_T(\mu_1)$	transverse momentum of the leading muon
$p_T(\mu_2)$	transverse momentum of the sub-leading muon
ϕ_1	azimuth angle of μ_1
ϕ_2	azimuth angle of μ_2
η_1	pseudorapidity of μ_1
η_2	pseudorapidity of μ_2
$(p_T)_{j1}$	transverse momentum of the leading jet
$(p_T)_{j2}$	transverse momentum of the sub-leading jet
ϕ_{j1}	azimuth angle of leading jet
ϕ_{j2}	azimuth angle of sub-leading jet
η_{j1}	pseudorapidity of leading jet
η_{j1}	pseudorapidity of sub-leading jet
$(p_T)_{jj}$	di-jet transverse momentum
$(p_T)_{\mu\mu}$	dimuon transverse momentum
m_{jj}	di-jet mass
$m_{\mu\mu}$	dimuon mass
n_{jets}	number of jets
n_μ	number of muons