LABORATÓRIO DE INSTRUMENTAÇÃO E FÍSICA EXPERIMENTAL DE PARTÍCULAS
partículas e tecnologia

Fundação para a Ciência e a Tecnologia
MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E ENSINO SUPERIOR

Measurement of $H \rightarrow \tau\tau$ with a multivariate analysis tool

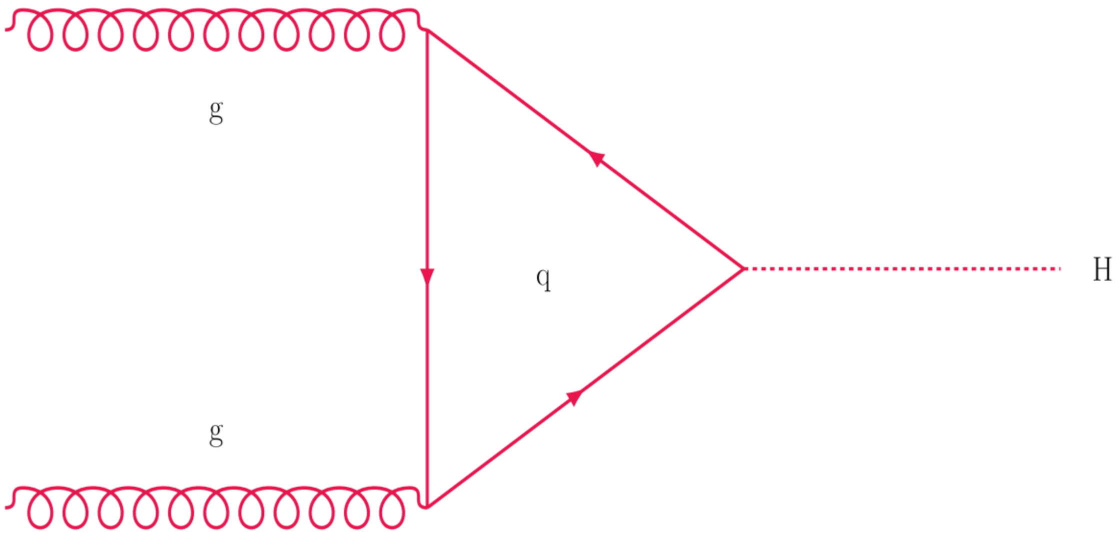
Dr. Pedrame Bargassa

Luís Sintra, Ricardo Cipriano and Tomás Alvim

Outline

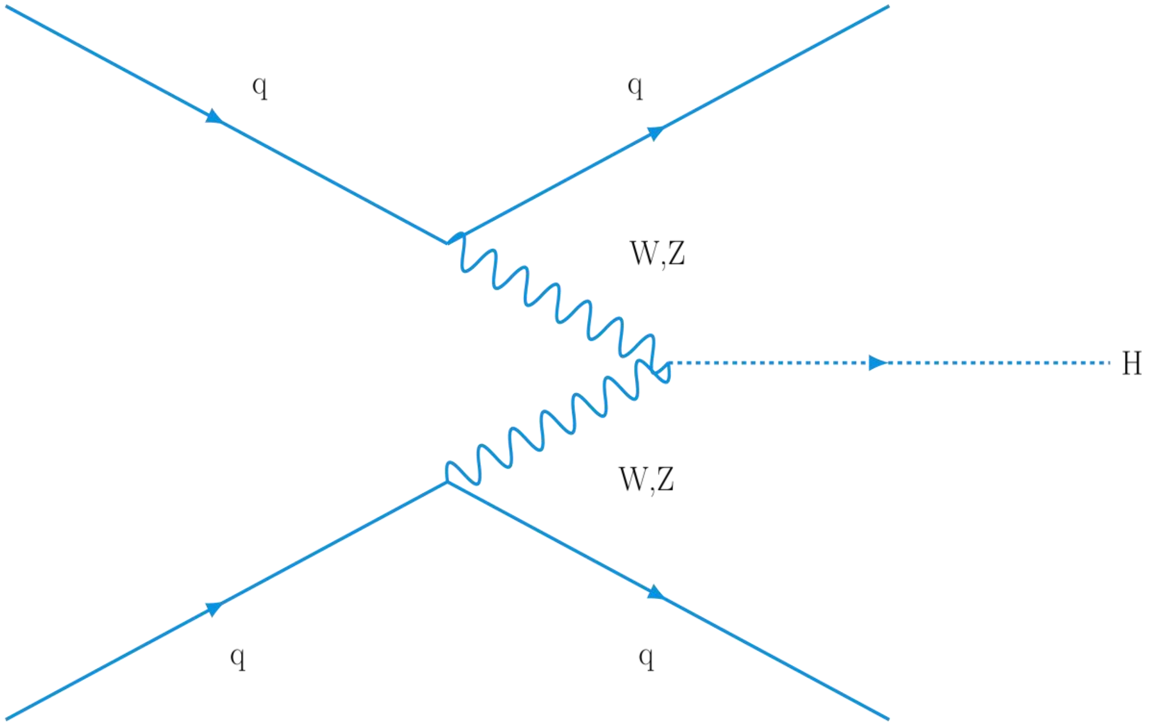
I. Higgs production & main background
II. Neural networks and categorization of events
III. Preselection cuts
IV. Confirmation of previous results
V. 1st new approach: N(btag)=0
VI. 2nd new approach: new NN
VII. Conclusions

Higgs production processes



gluon gluon fusion (ggH)

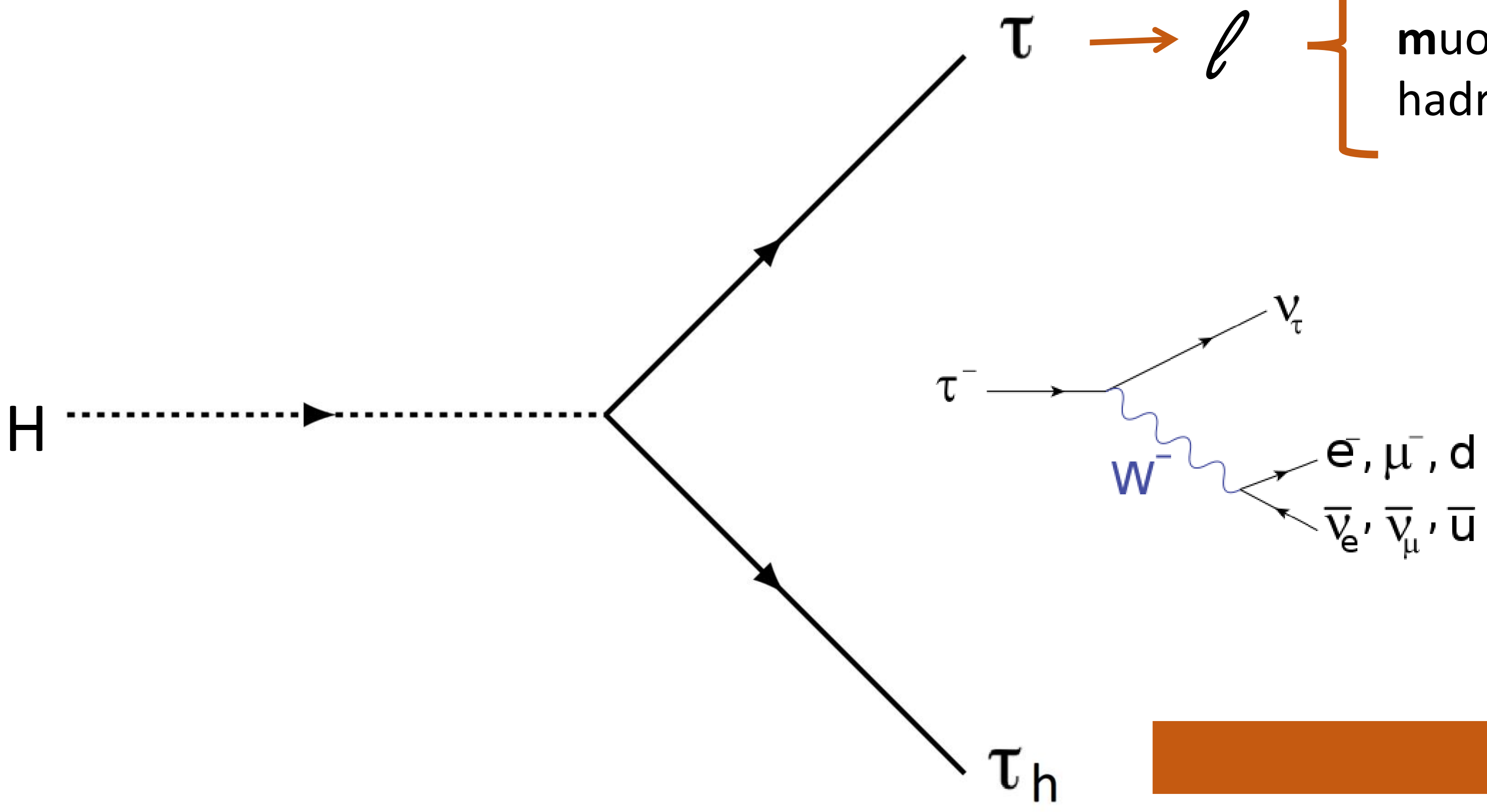
Cross section: 12.6 pb



vector boson fusion (qqH/VBF)

Cross section: 0.5 pb

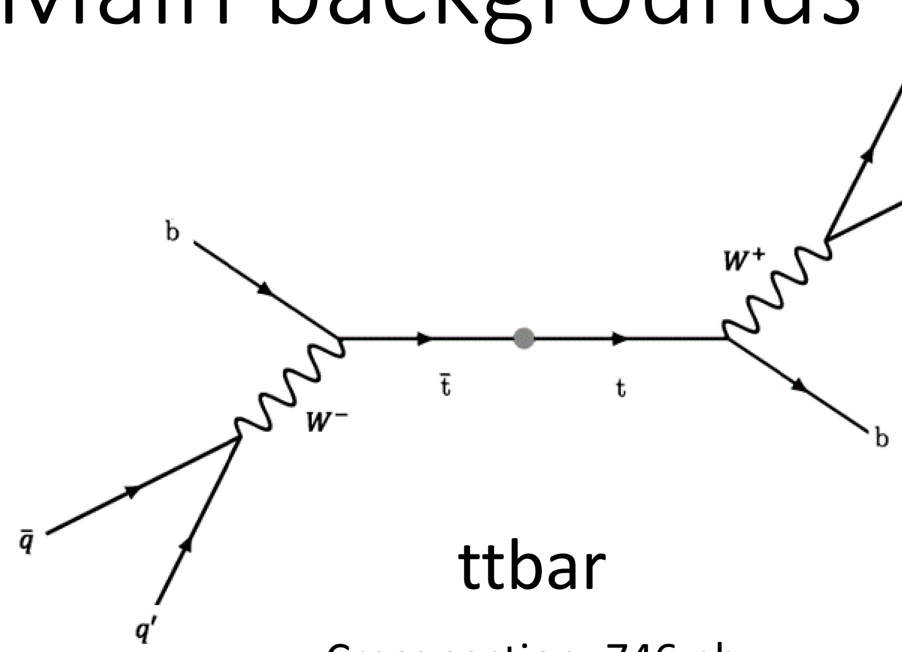
Higgs decay channels



electron
muon
hadronic **t**au

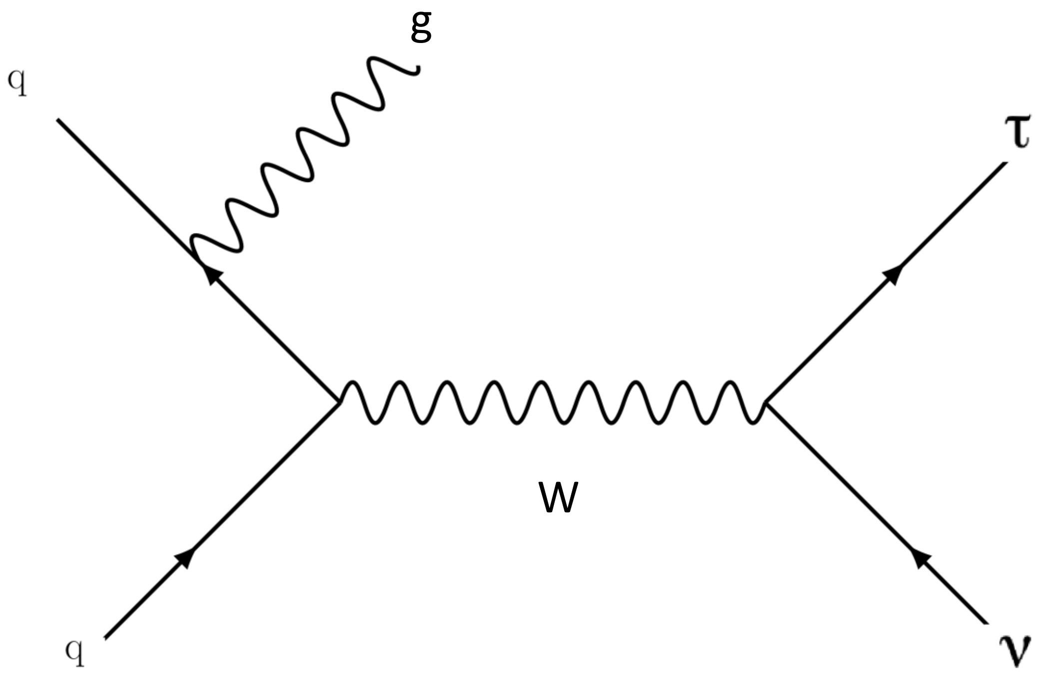
The channels are:
- electron – tau (et)
- muon – tau (mt)
- tau – tau (tt)

Main backgrounds

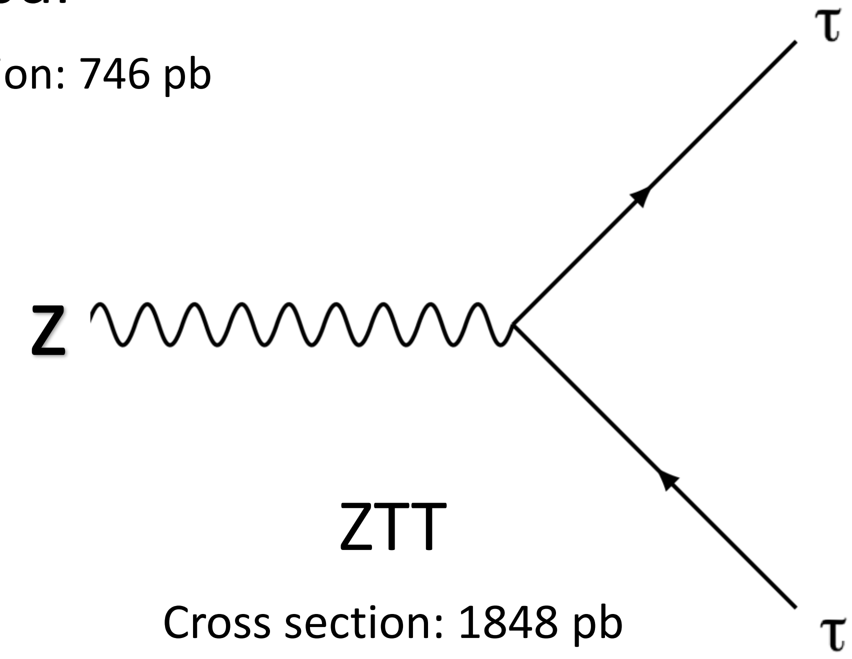


ttbar

Cross section: 746 pb



W+jet



ZTT

Cross section: 1848 pb

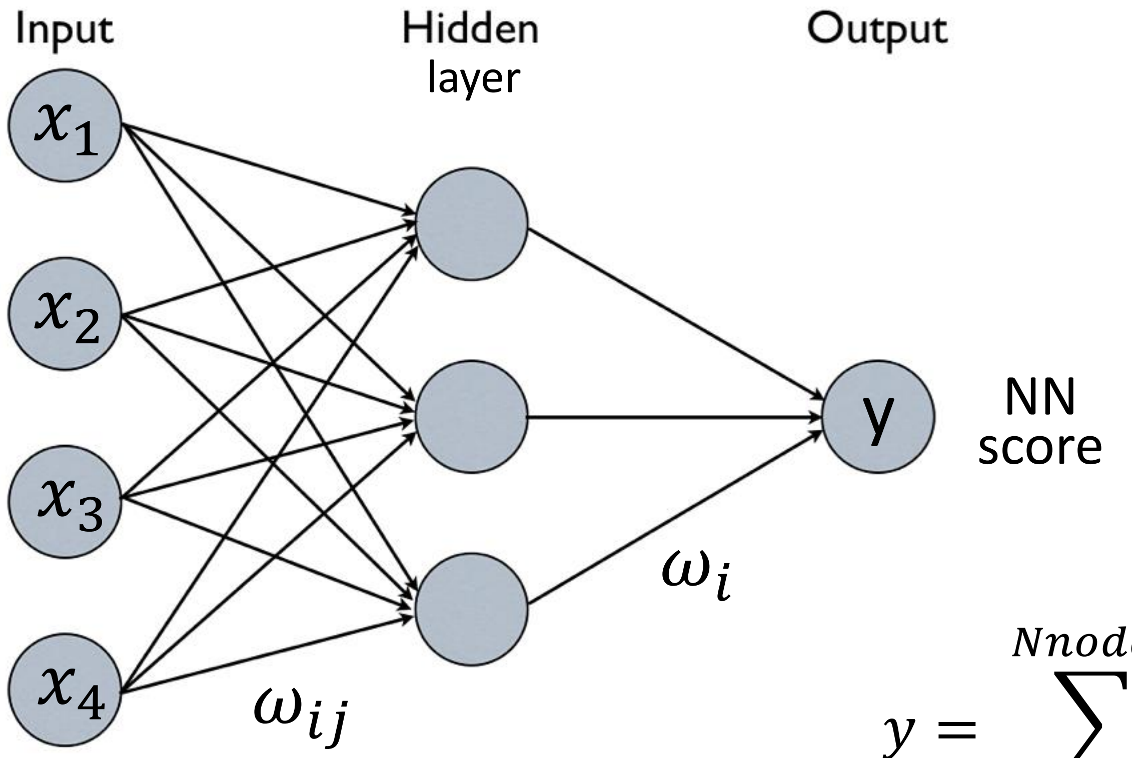
Pre-selection cuts

The pre-selection cuts applied on events that are going to be analyzed by the NN are meant to diminish as much background as possible, while leaving enough efficiency for $H \rightarrow \tau\tau$

- Pt (muon) > 20
- Pt (electron) > 25
- Pt (tau): > 20 for et & mt channels; > 40 for tt channel
- Min (mT(l1,MET) , mT(l2,MET)) < 50
- 2 opposite charge leptons

Neural networks: Concept

The measurement of the H → ττ decay is being made with a neural network (NN)



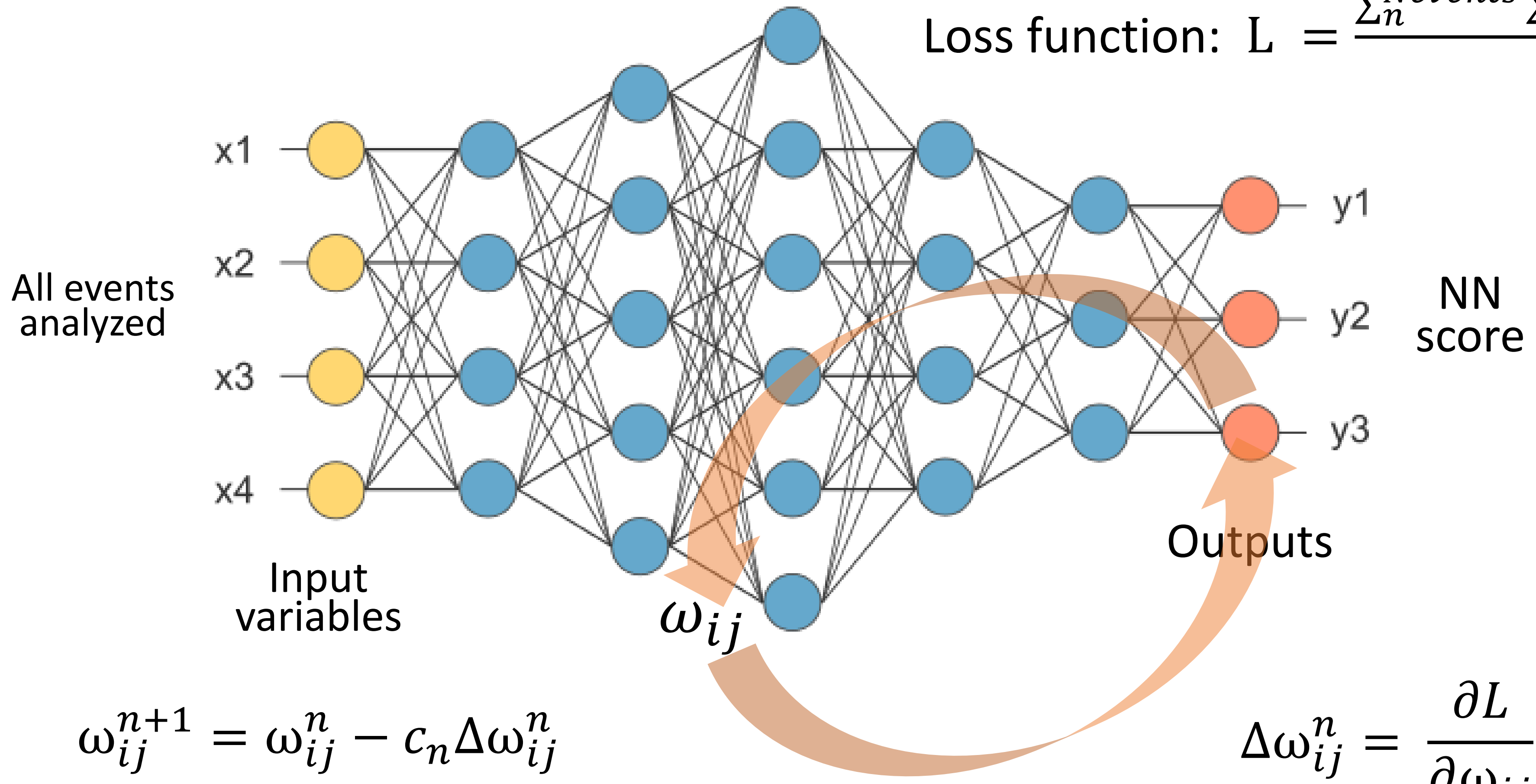
$$y = \sum_{j}^{Nnodes} \omega_j \, g\left(\sum_{i}^{Ninputs} \omega_{ij} x_i \right)$$

Neural networks: Training

NN response

Loss function: $L = \frac{\sum_n^{Nevents} \sum_i^{Noutputs} (E_i^n - y_i^n)^2}{Nevents}$

Correct response



$$\omega_{ij}^{n+1} = \omega_{ij}^{n} - c_n \Delta\omega_{ij}^{n}$$

$$\Delta\omega_{ij}^{n} = \frac{\partial L}{\partial \omega_{ij}}$$

Process categorization

Our multi-class NN is trained to distinguish signal from background and categorizes them in different classes.

Background

Signal

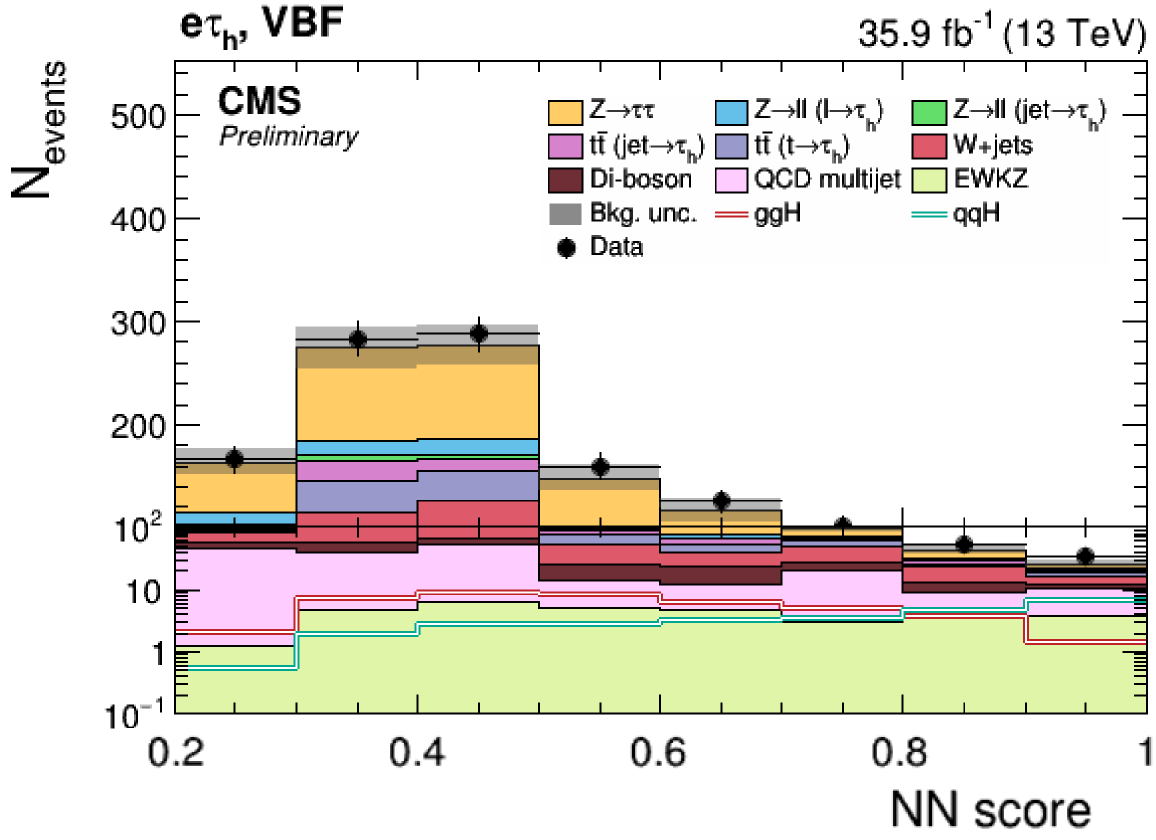
et/mt	
Classes	Processes
misc	EWK VV
w	W
tt	TTJ TTT
zll	ZL ZJ
ss	QCD
ztt	ZTT
qqH	qqh
ggH	ggh

tt	
Classes	Processes
misc	EWK VV W TTT ZL ZJ
noniso	QCD
ztt	ZTT
qqH	qqh
ggH	ggh

Event categorization

The multi-class NN attributes a score for each class which represents the estimated probability of that event to belong in that given class.

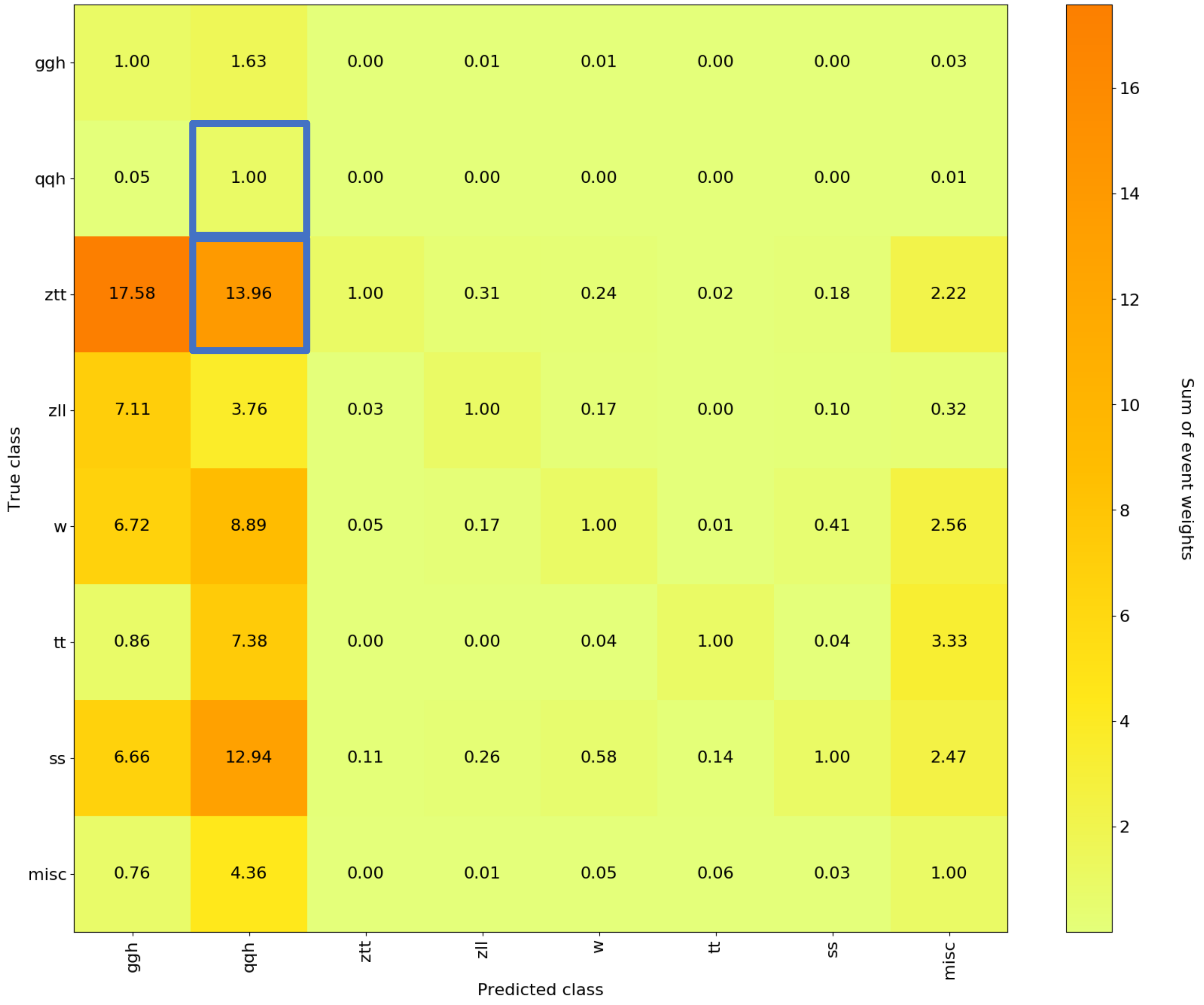
For each event, we define the class with highest NN score to be the predicted class, which might be different from the true class.



Confusion matrix

A confusion matrix tells how much the multi-class NN is confused when classifying a true process Y in the predicted process X.

This is an example of a purity matrix.



Replication of Pedrame's results

Before we tried to improve the L2N200 NN, we needed to see if we could reproduce the basic results.

Pedrame's results:

	Uncertainty on signal strength
et	± 0.44
mt	± 0.23
tt	± 0.33

Combined uncertainty on signal strength	± 0.18

Confidence Limit: 68%

Our results:

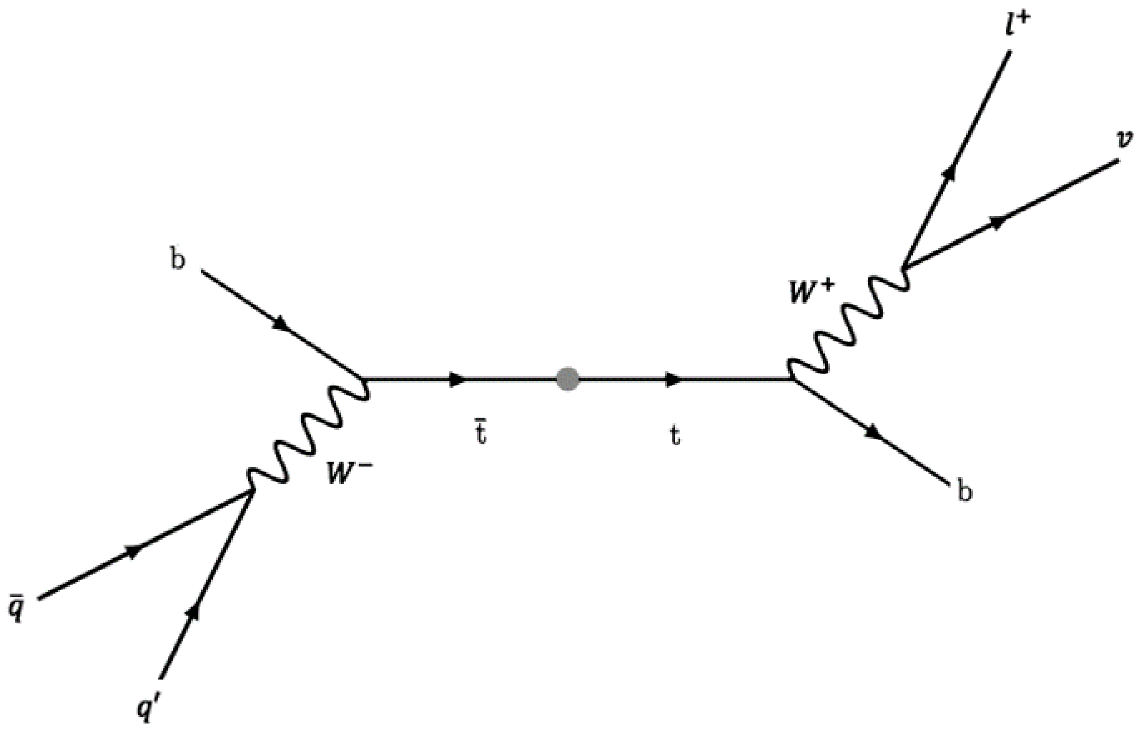
	Uncertainty on signal strength
et	± 0.44
mt	± 0.23
tt	± 0.33

Combined uncertainty on signal strength	± 0.18

Nbtag = 0 cut

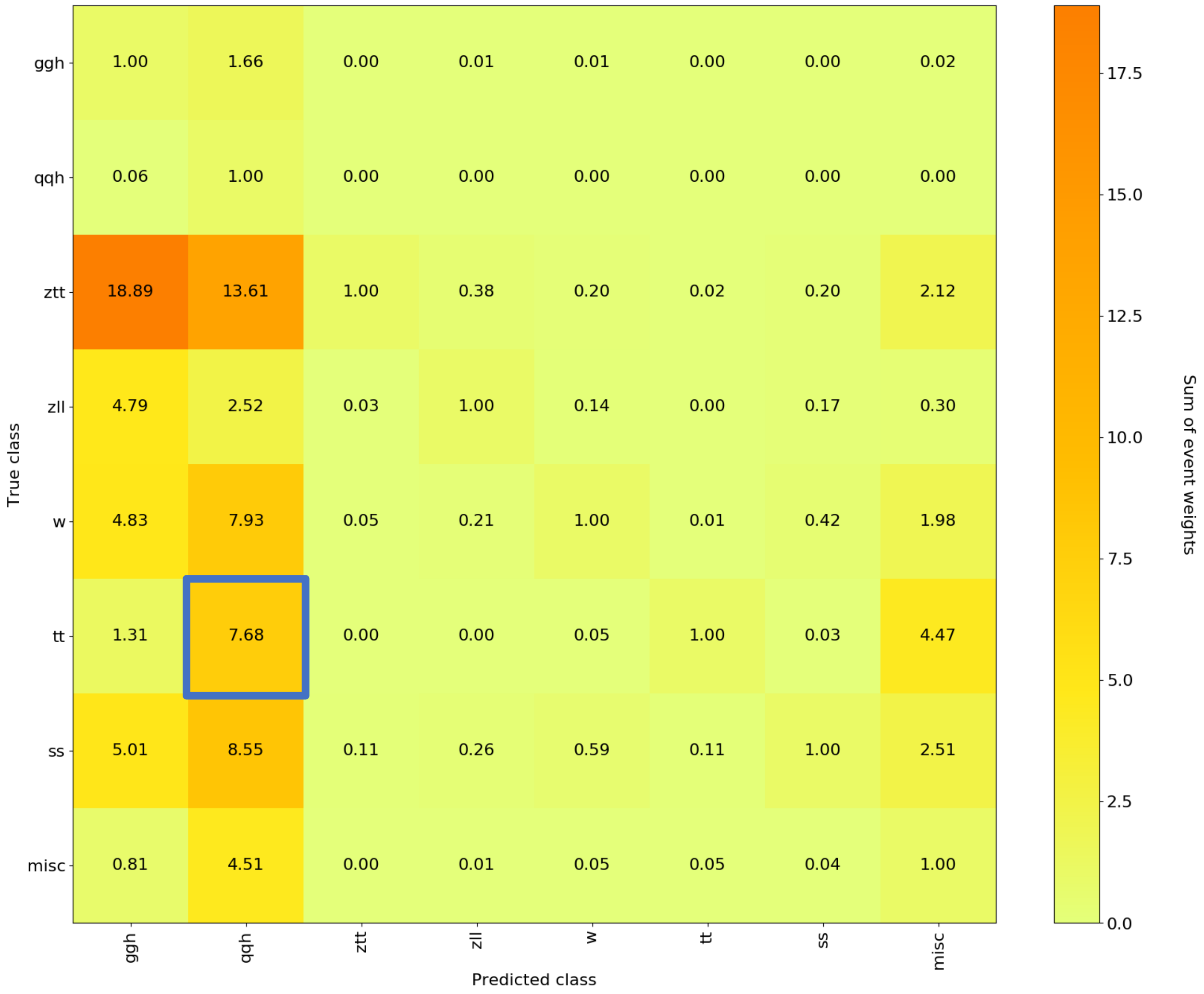
ttbar involves production of b quarks, while H → ττ (most of time) doesn't.

Motivation for selecting events with no btag before training: seriously bring down the level of one of the main backgrounds (ttbar) as to confuse less the multi-class NN.

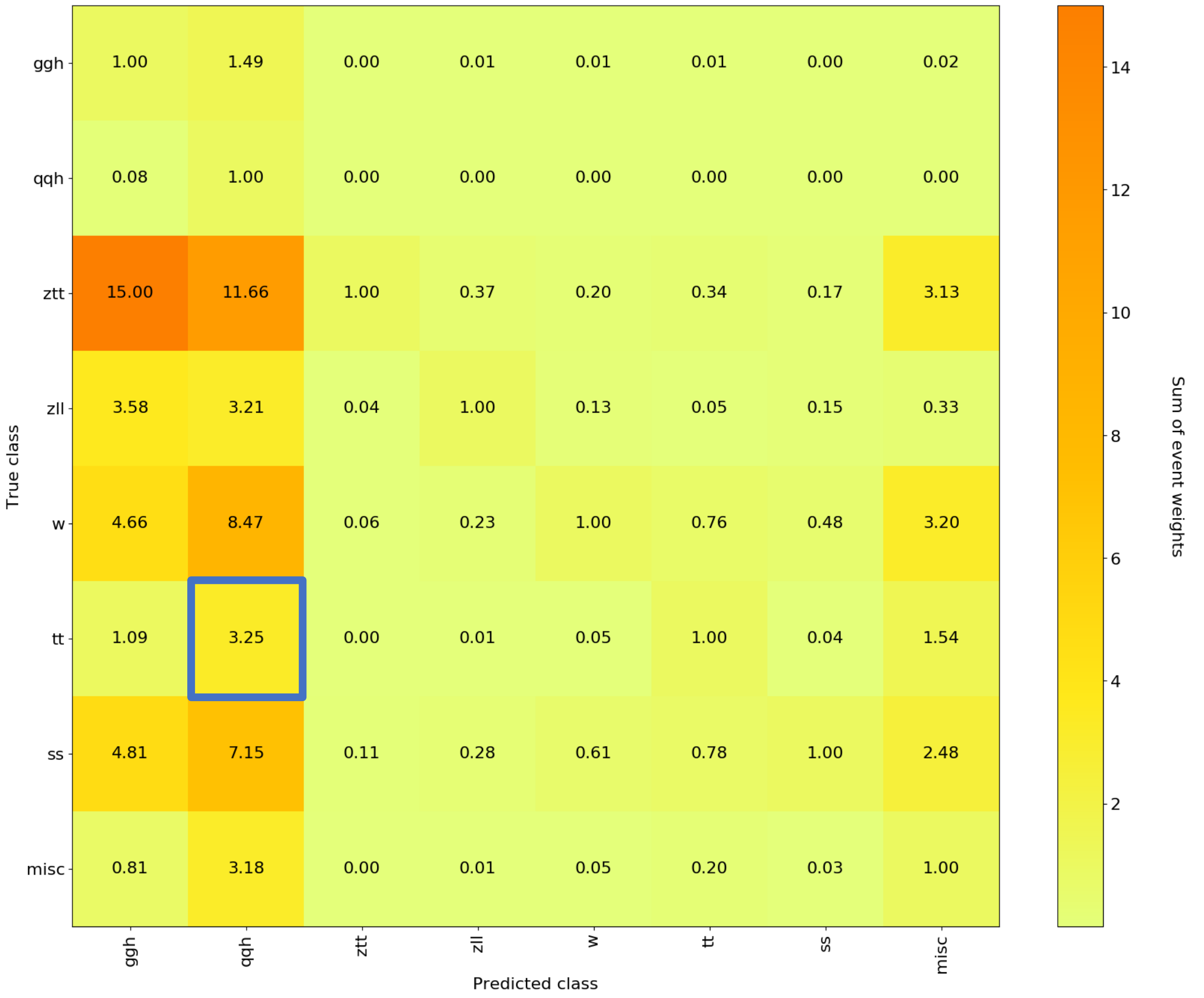


Nbtag = 0 cut efficiency	et	mt	tt
ttbar	0.2609 ± 0.0021	0.2602 ± 0.0014	-
ggH	0.9737 ± 0.0002	0.9747 ± 0.0001	0.9327 ± 0.0008
qqH	0.9575 ± 0.0002	0.9567 ± 0.0002	0.9542 ± 0.0003

Confusion matrix comparison for the et channel

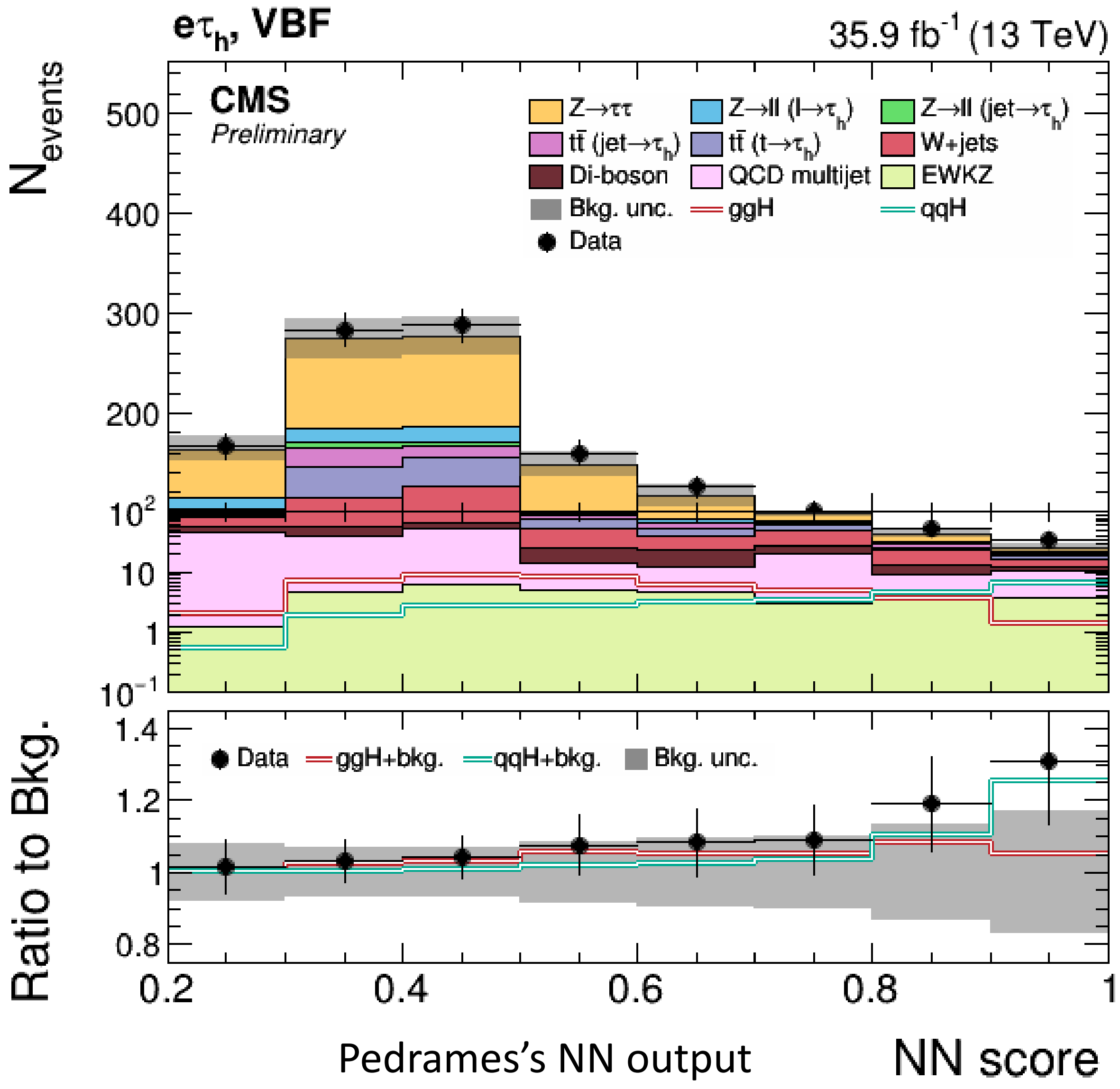


Pedrame's confusion matrix

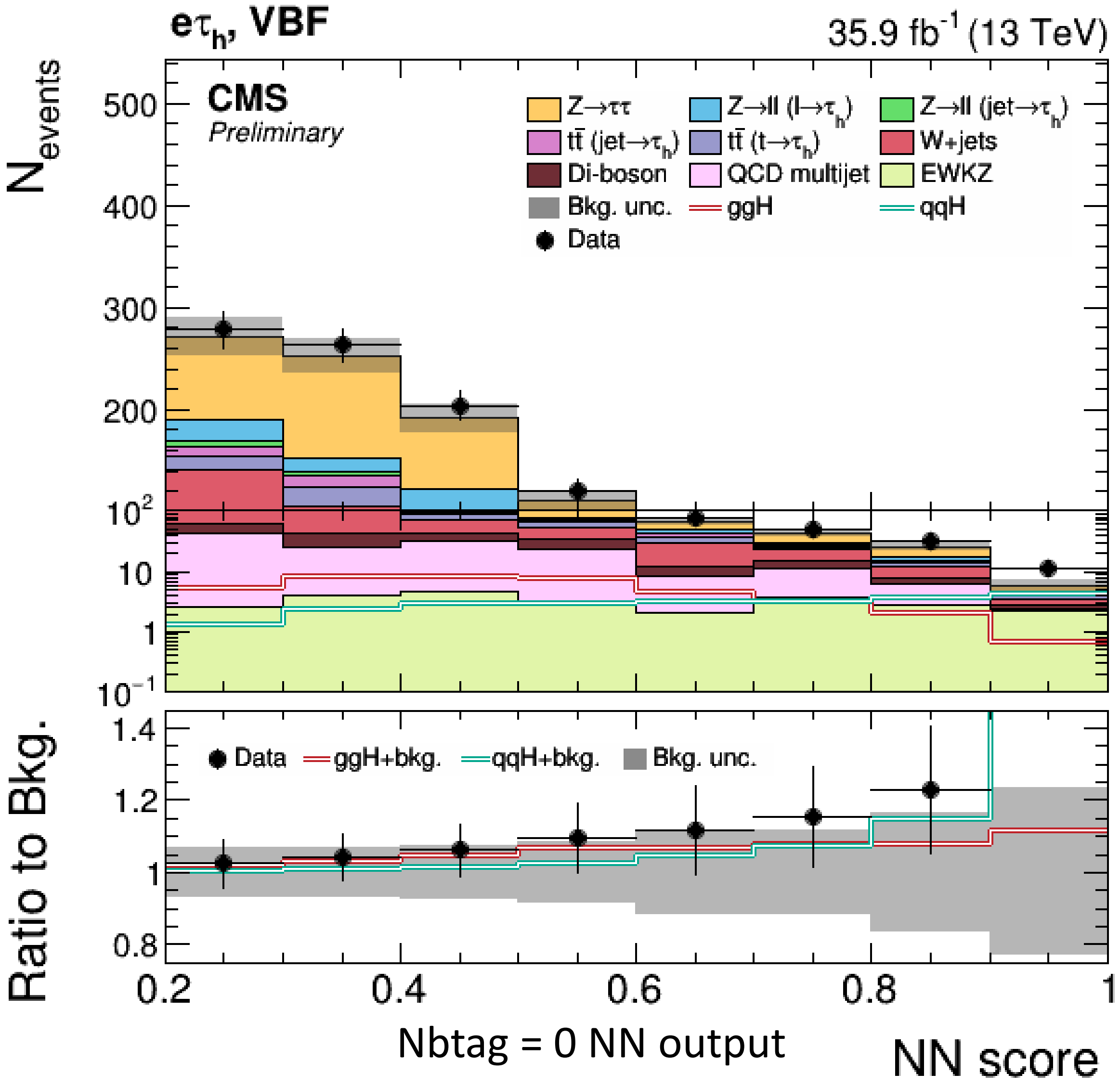


Nbtag = 0 confusion matrix

Pedrame's and Nbtag = 0: qqH NN scores for et



Pedrames's NN output



Nbtag = 0 NN output

Comparison of signal strength precision

Pedrame's results:

	Uncertainty on signal strength
et	± 0.44
mt	± 0.23
tt	± 0.33

Combined uncertainty on signal strength	± 0.18

Nbtag = 0 results:

	Uncertainty on signal strength
et	± 0.39
mt	± 0.22
tt	± 0.31

Combined uncertainty on signal strength	± 0.16

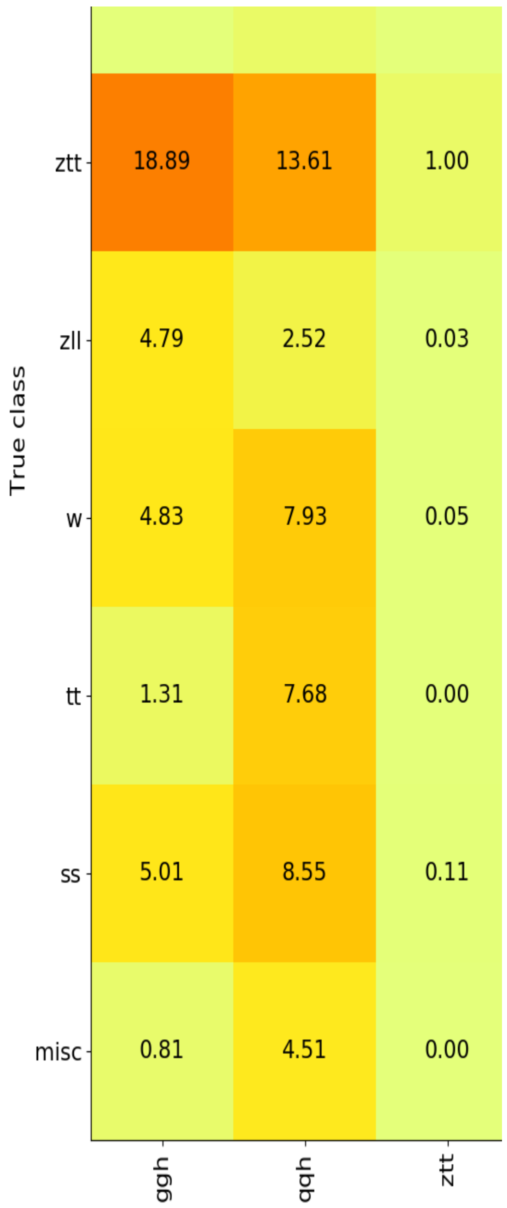
- For et the cut improved the signal strength precision in 11.4%
- For mt the cut improved the signal strength precision in 4.3%
- For tt the cut improved the signal strength precision in 6.1%
- For CMB the cut improved the signal strength precision in 11.1%

NN score cut

After the Nbtag = 0 proved successful, we decided to try a different cut, based on a NN output.

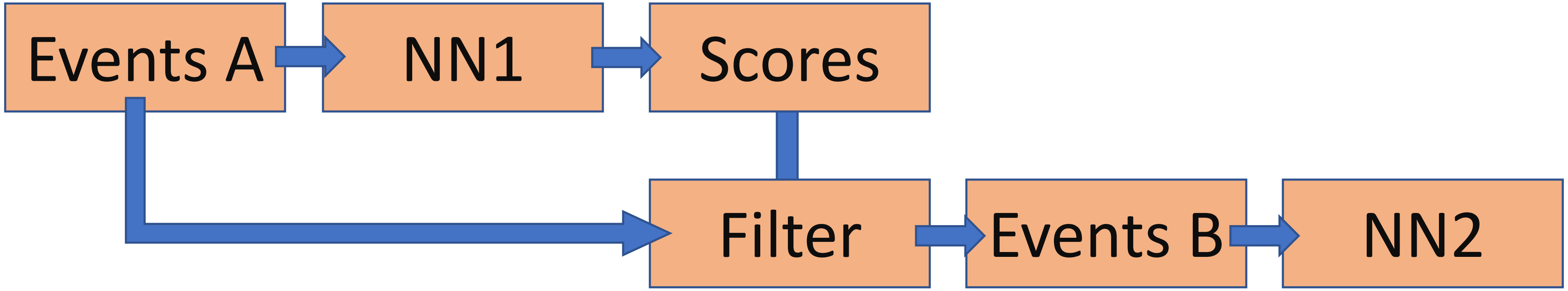
In this approach, we used two NN's, one of which serves to filter events to be fed to the other one, the latter being used as before.

With the 1st NN, we want to diminish the contribution of processes other than ztt, so that when we run the NN it has less backgrounds and can hopefully better distinguish between signal & ztt.



Pedrame's Purity 1 matrix for et

Architecture of the NN score cut approach

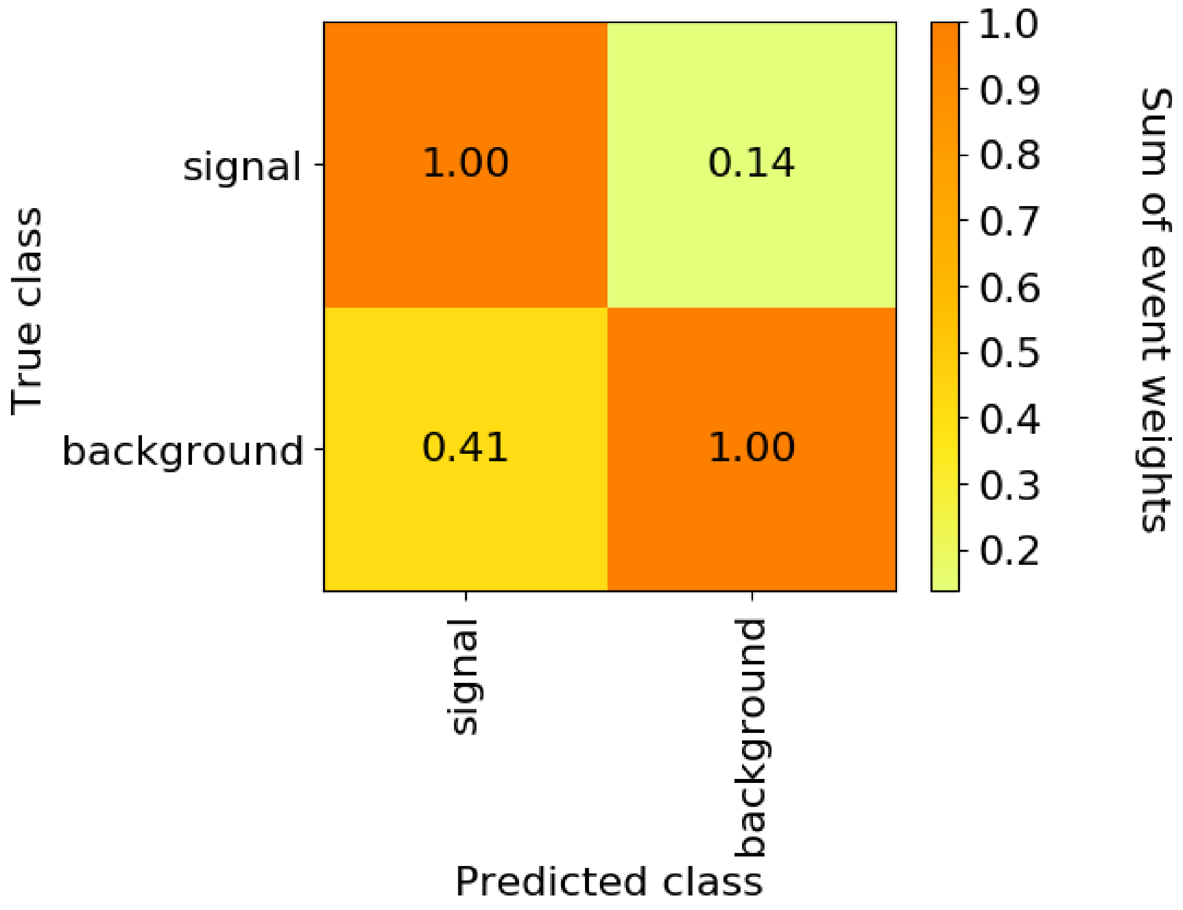


et/mt		
Classes NN1	Classes NN2	Processes
Background	misc	EWK VV
	w	W
	tt	TTJ TTT
	zll	ZL ZJ
	ss	QCD
Signal	ztt	ZTT
	qqH	qqh
	ggH	ggh

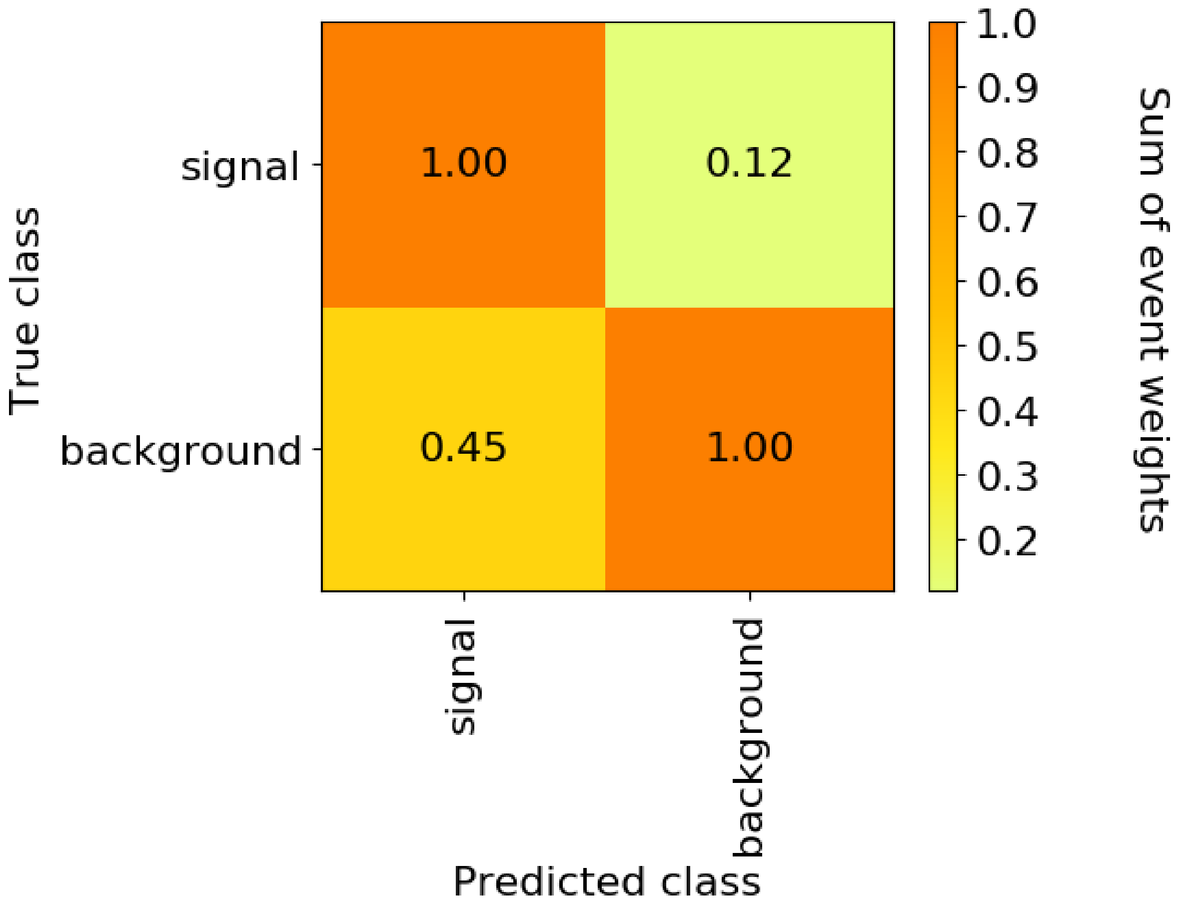
tt		
Classes NN1	Classes NN2	Processes
Background	misc	EWK VV W TTT ZL ZJ
	noniso	QCD
Signal	ztt	ZTT
	qqH	qqh
	ggH	ggh

NN2 has the same categorization as NN0(Pedrame's) but trains with filtered events

Confusion matrices for NN1



et channel confusion matrix



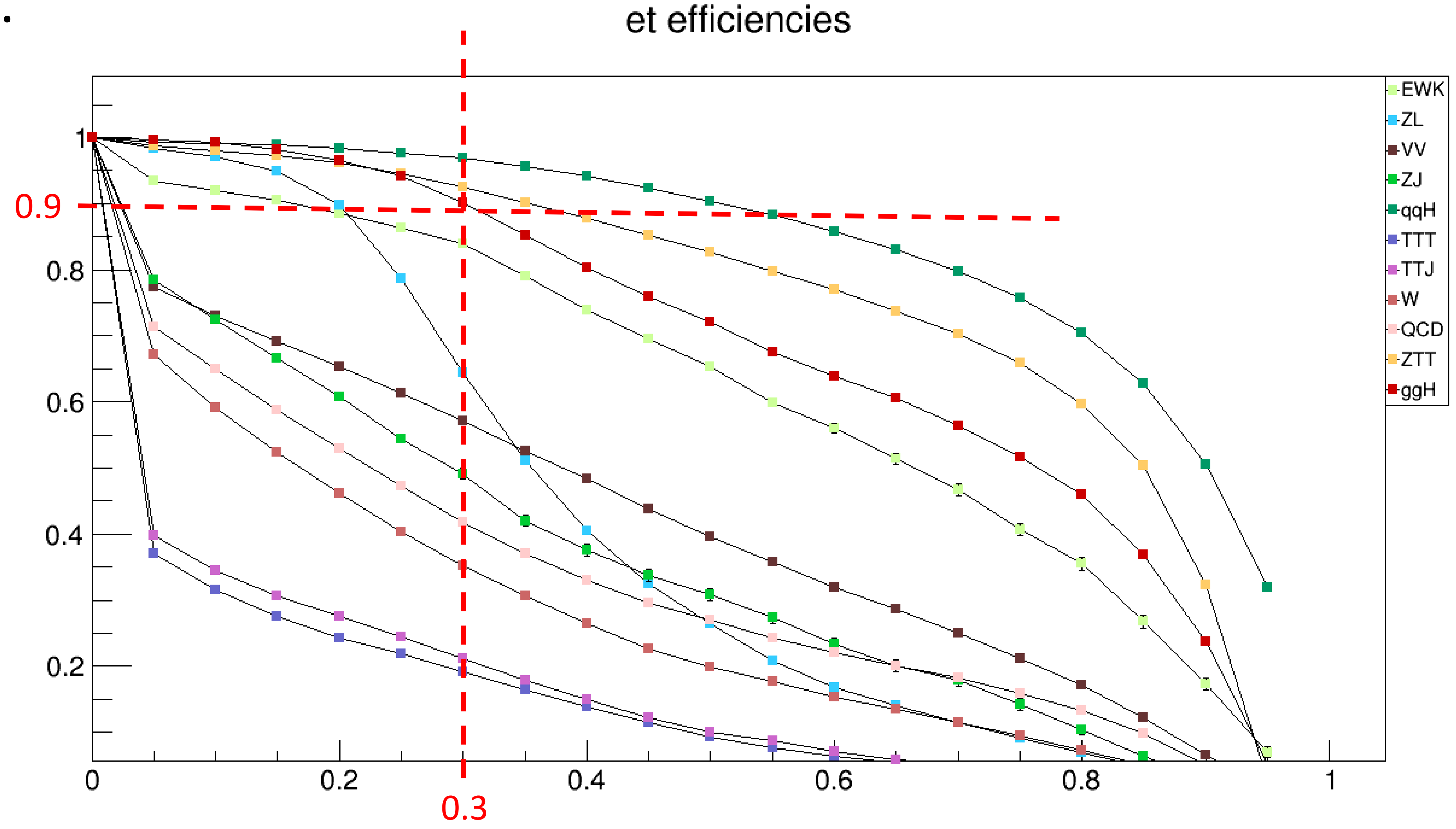
mt channel confusion matrix

Selection of events B (for NN2)

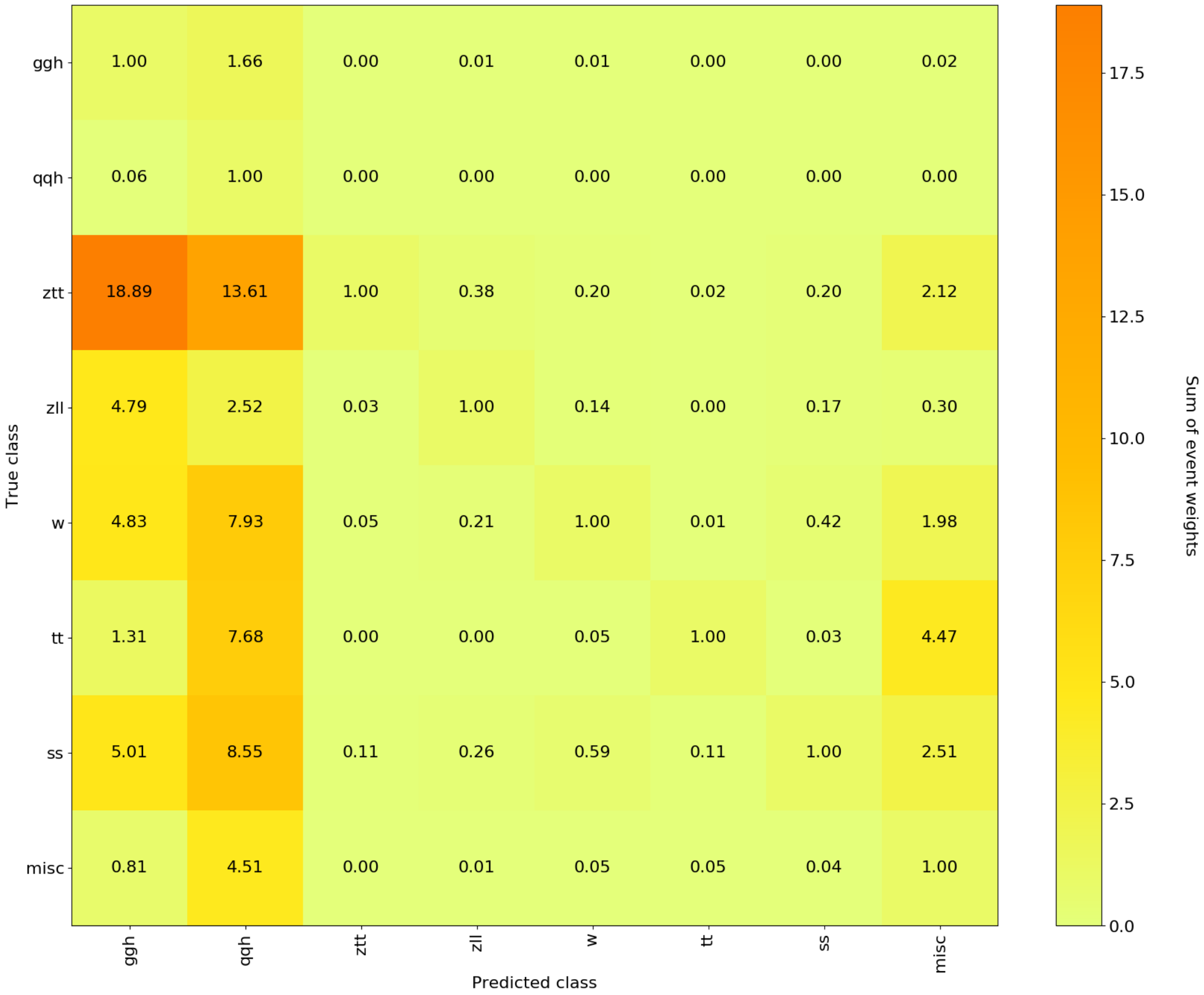
We chose the place of the cut where the efficiency is at least 90% for both qqH and ggH.

Cuts:

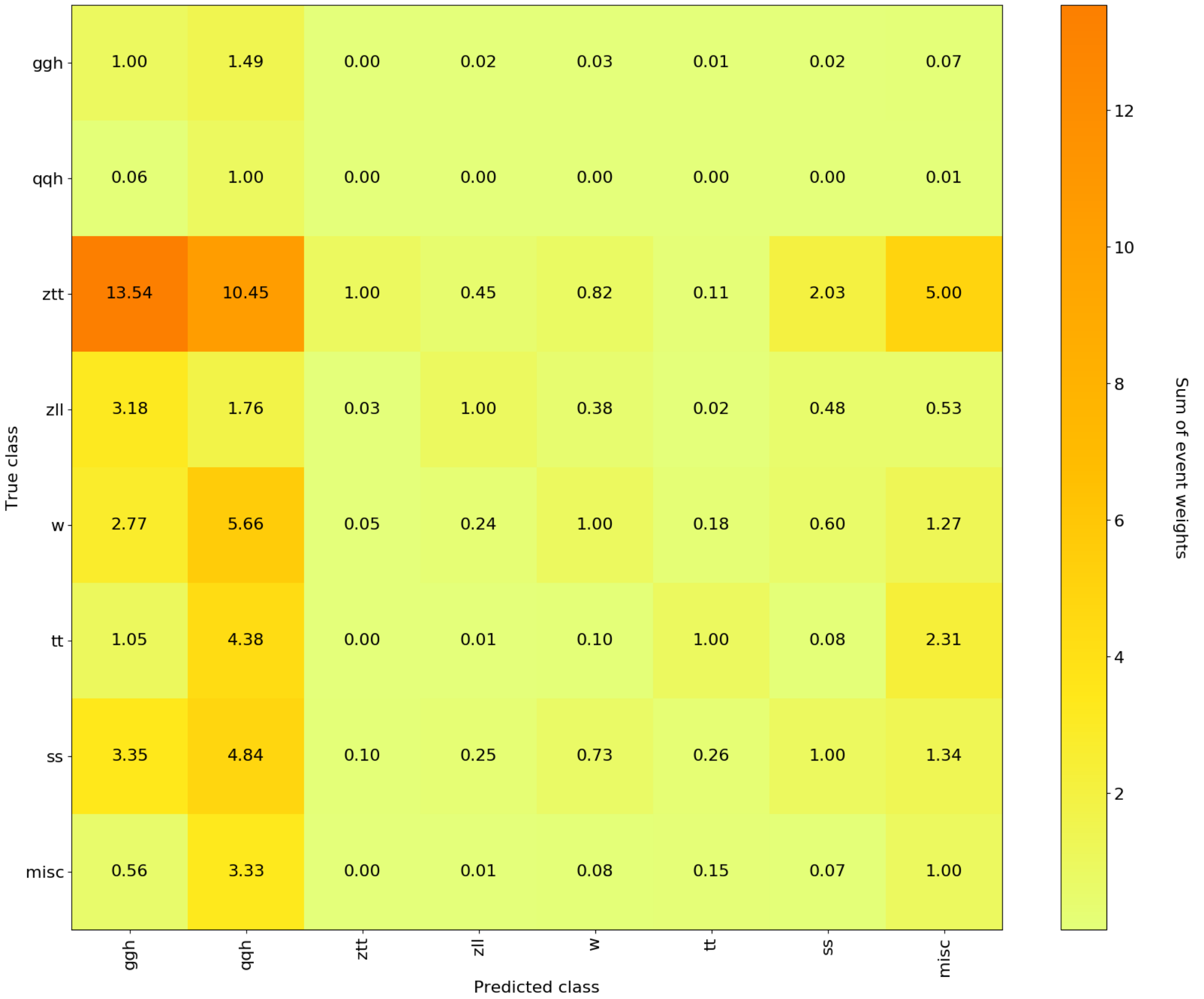
- et: 0.3
- mt: 0.3
- tt: 0.4



Confusion matrix comparison for et channel on NN2

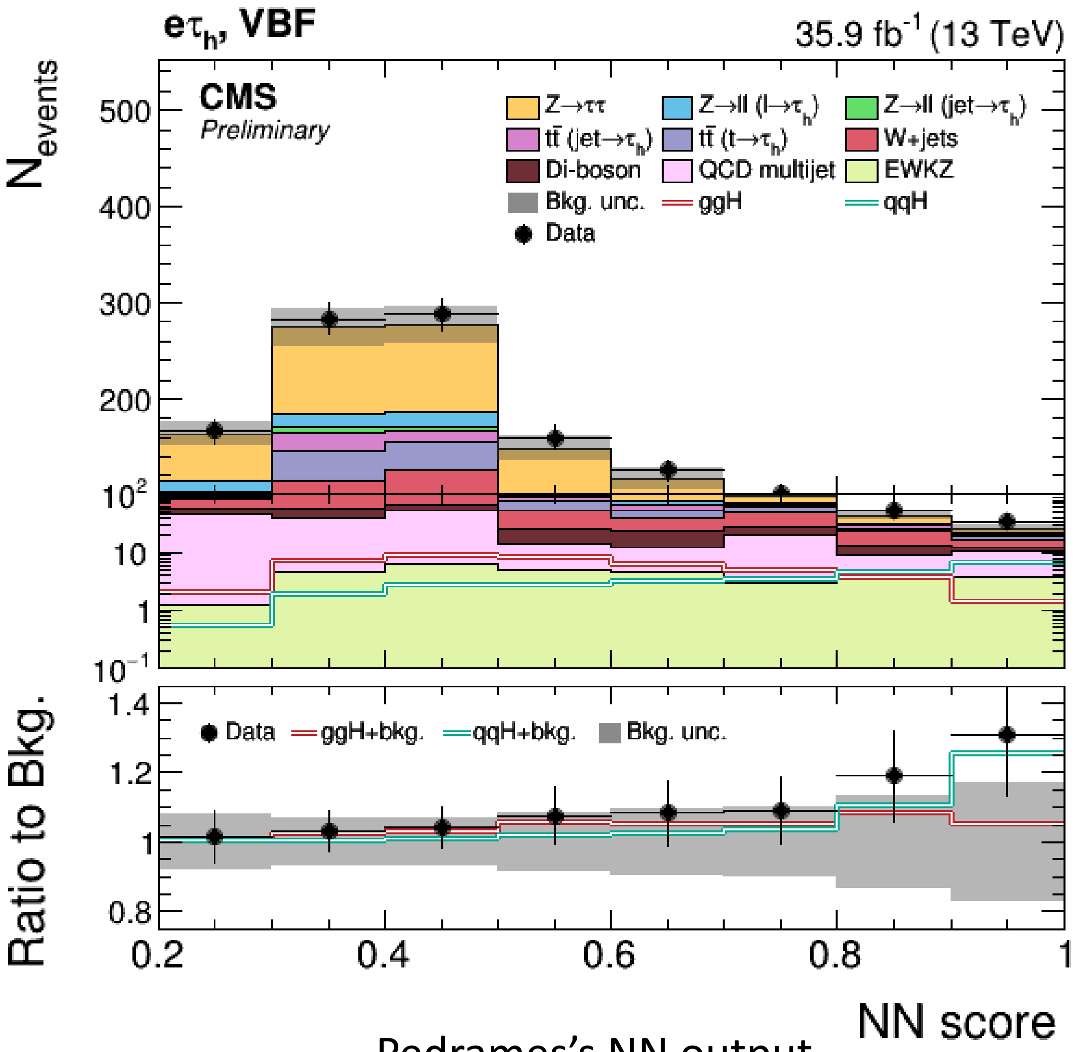


Pedrame’s confusion matrix

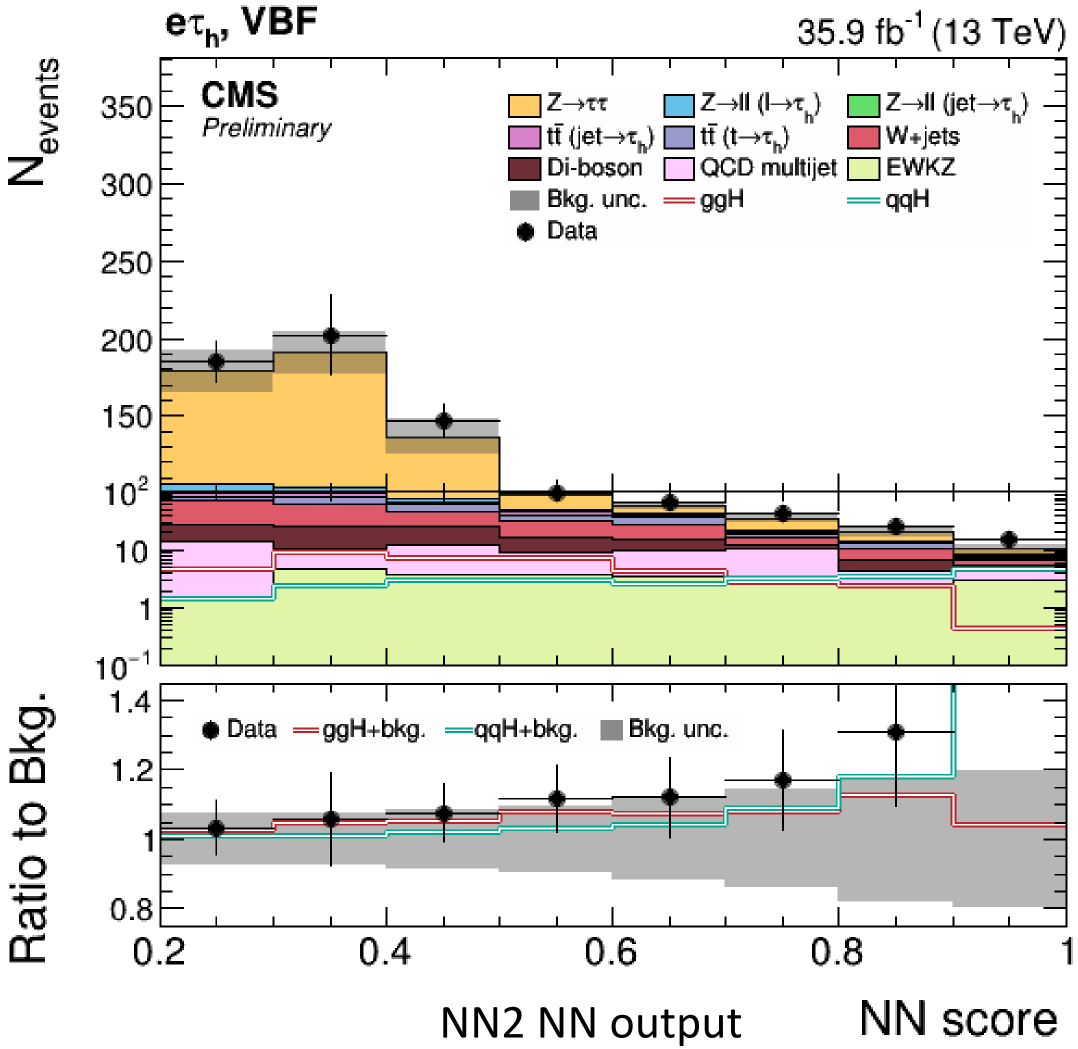


NN2 confusion matrix

Pedrame’s and NN2: qqH scores for et



Pedrames’s NN output



NN2 NN output

Comparison of signal strength precision

Pedrame's results:

	Uncertainty on signal strength
et	± 0.45
mt	± 0.25
tt	± 0.33

Combined uncertainty on signal strength	± 0.18

New approach results:

	Uncertanty on signal strength
et	± 0.39
mt	± 0.25
tt	-

Combined uncertainty on signal strength	-

- For et the cut improved the signal strength precision in 13%
- For mt the cut didn't change the signal strength precision.

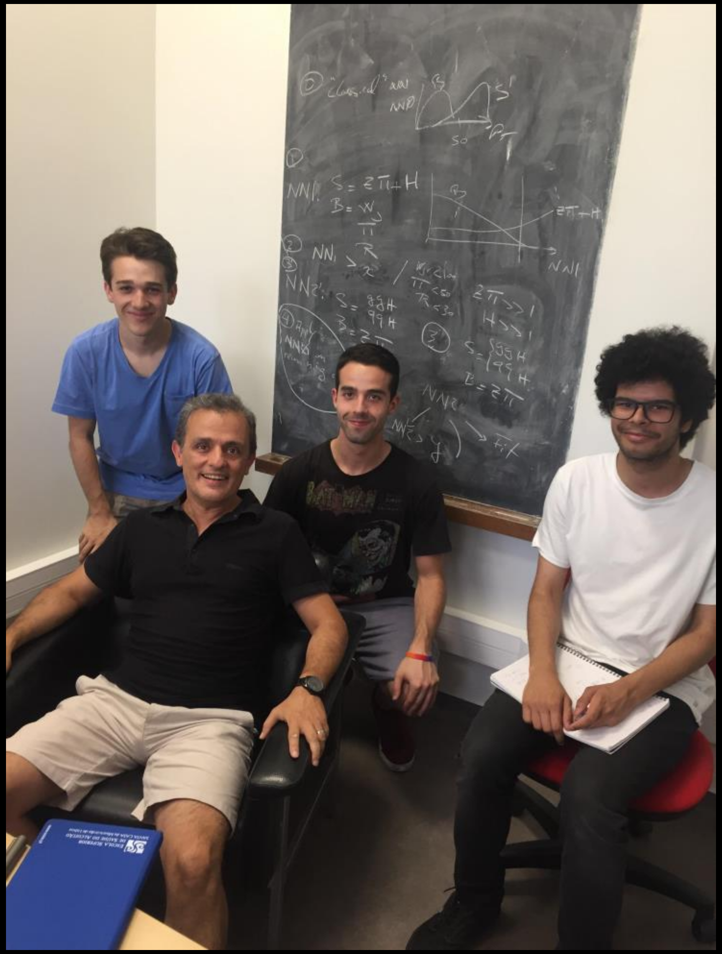
Conclusions

- We acquainted ourselves with the challenges of Higgs search where the signal is orders of magnitude smaller than the stardand model background.
- We learned to work with a multi class NN
- We tried two new approaches that are improving the performance of the normal approach.
    1. One with nbtag = 0 (bringing down level of ttbar background)
    2. A new one (bringing down the level of standard model background by doing a two level neural net)

Further work

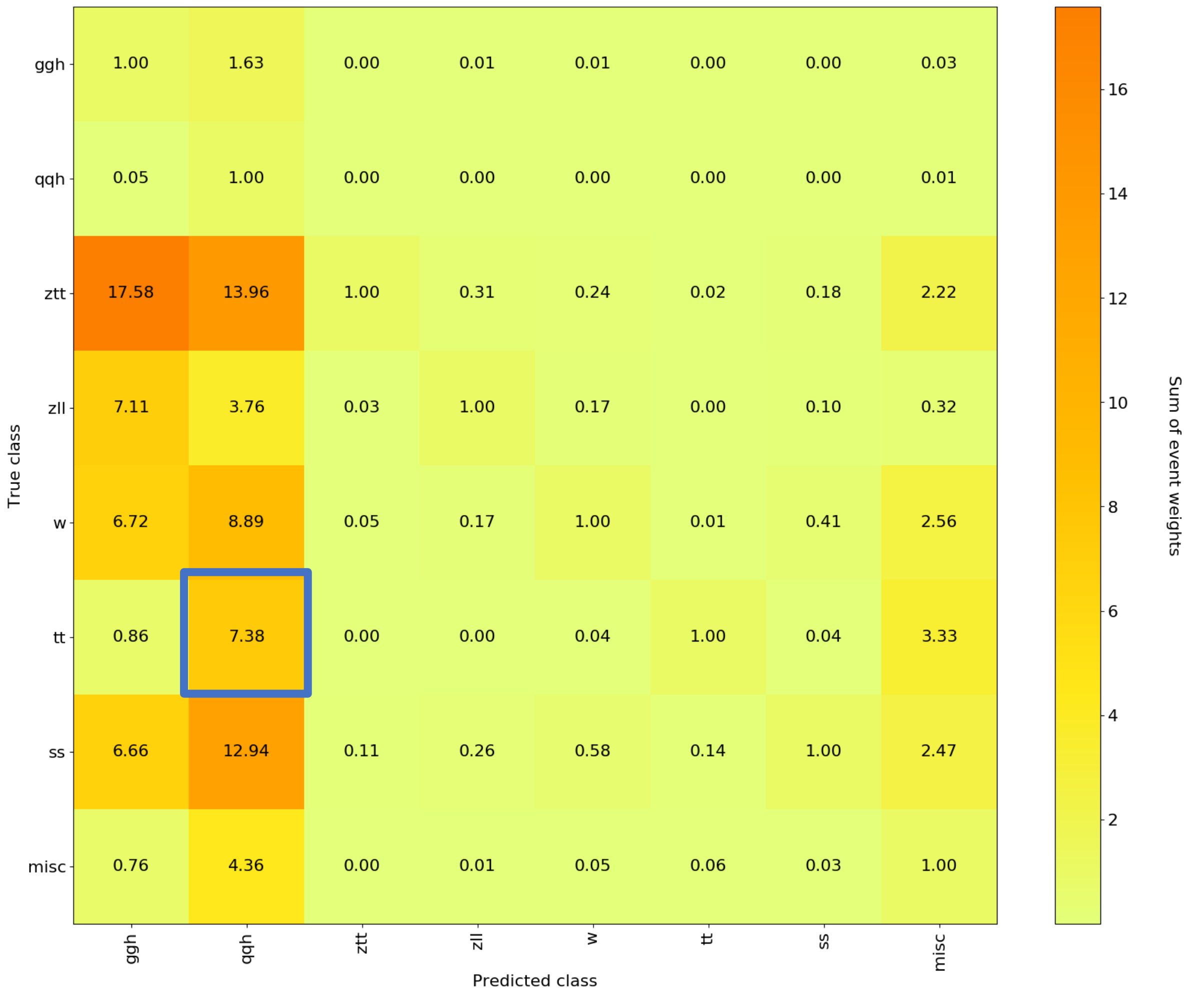
- Re-checking analysis part of the used framework which fails under some conditions.
- Try different NN architectures

Thanks!

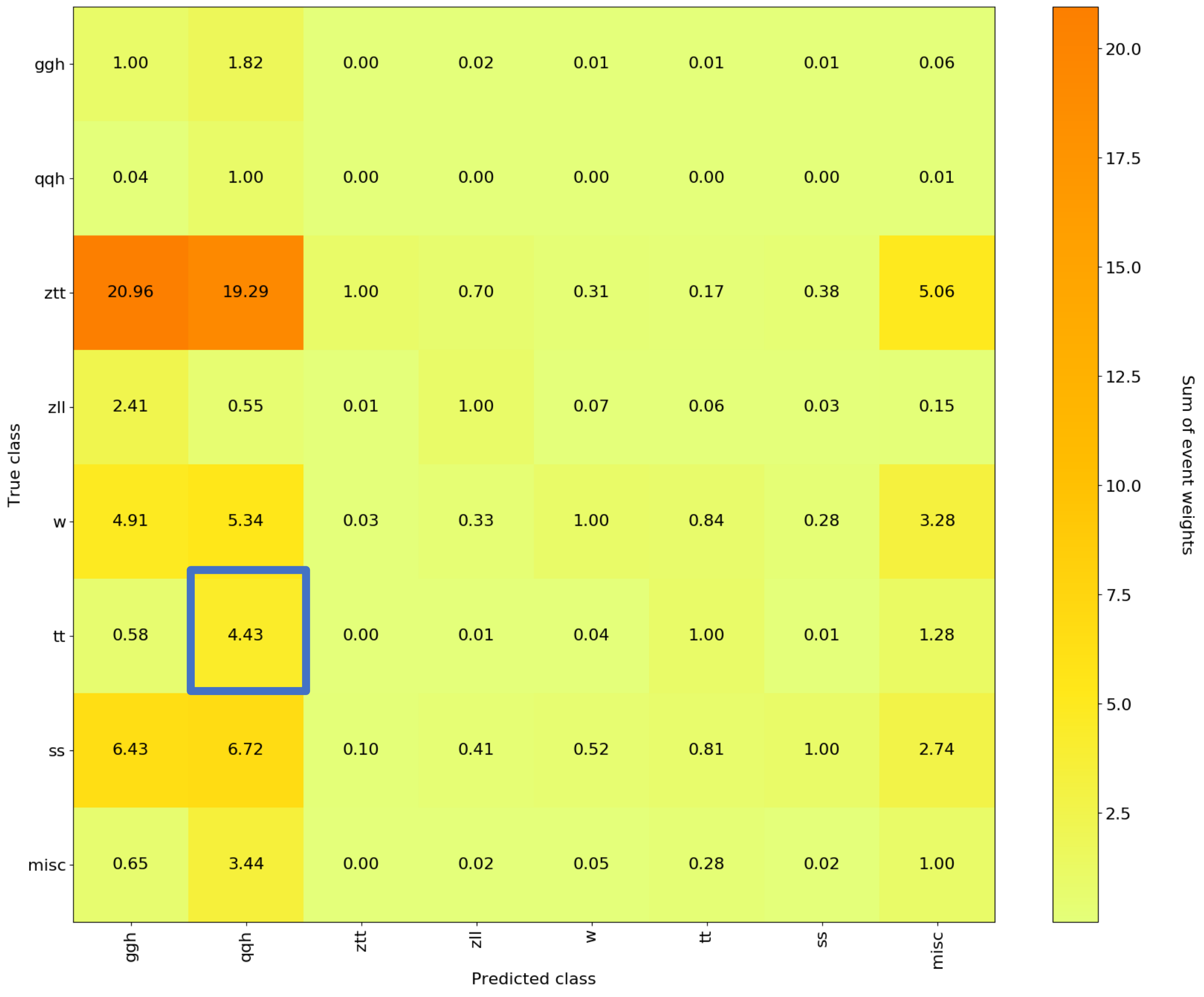


Additional information slides:

Confusion matrix comparison for mt channel

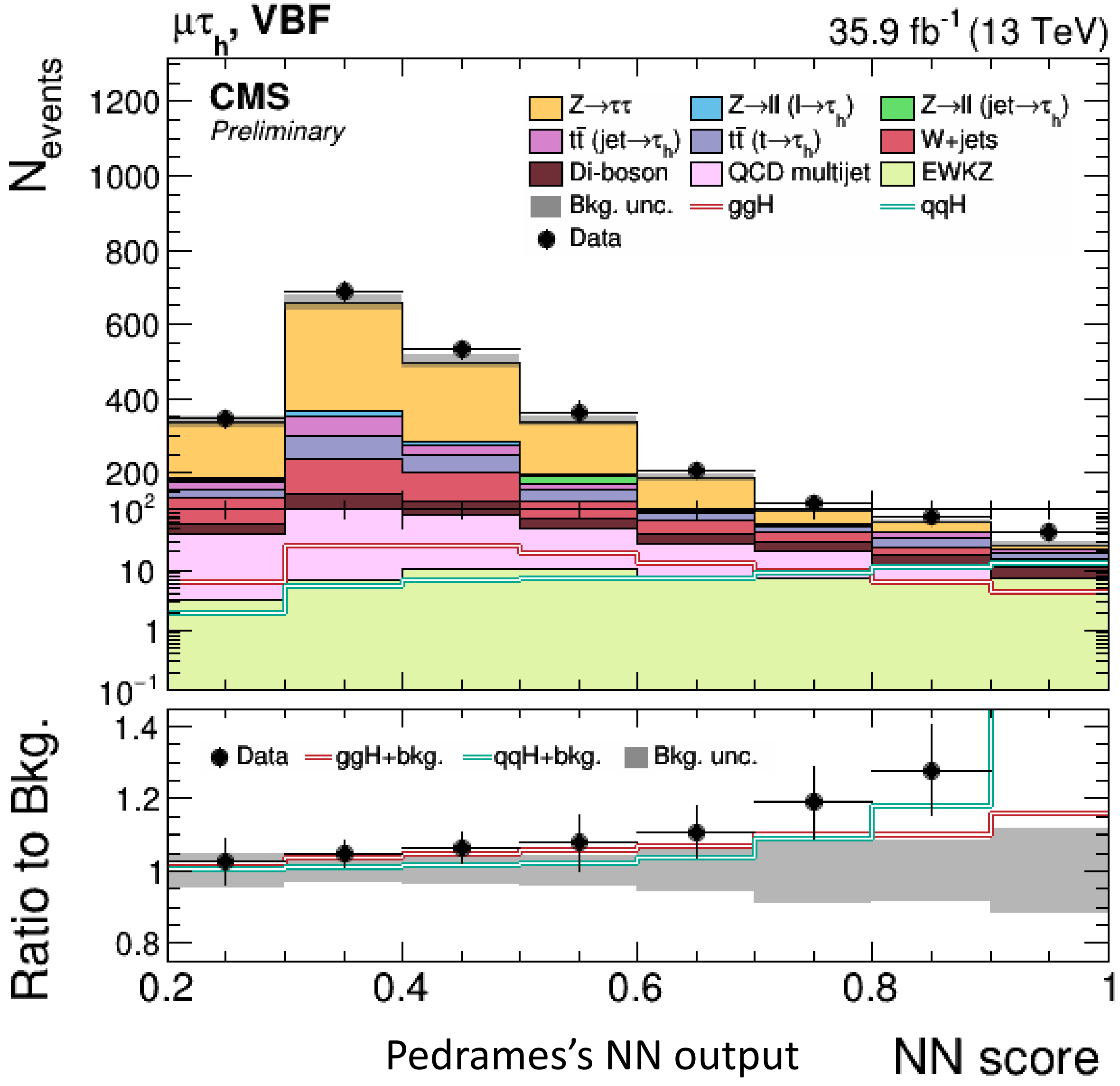


Pedrame’s confusion matrix

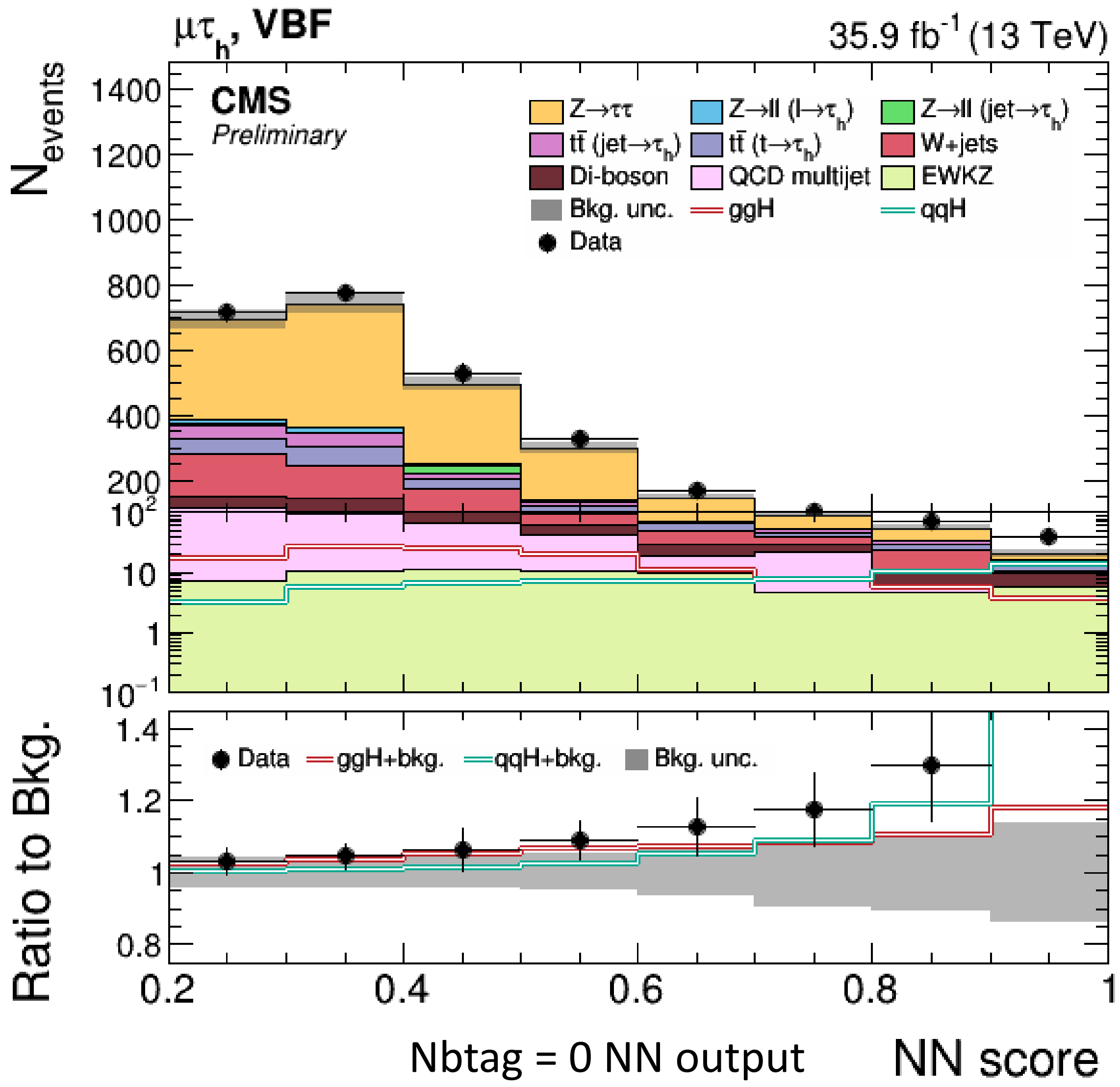


Nbtag = 0 confusion matrix

Pedrame’s and Nbtag = 0: qqH NN scores for mt

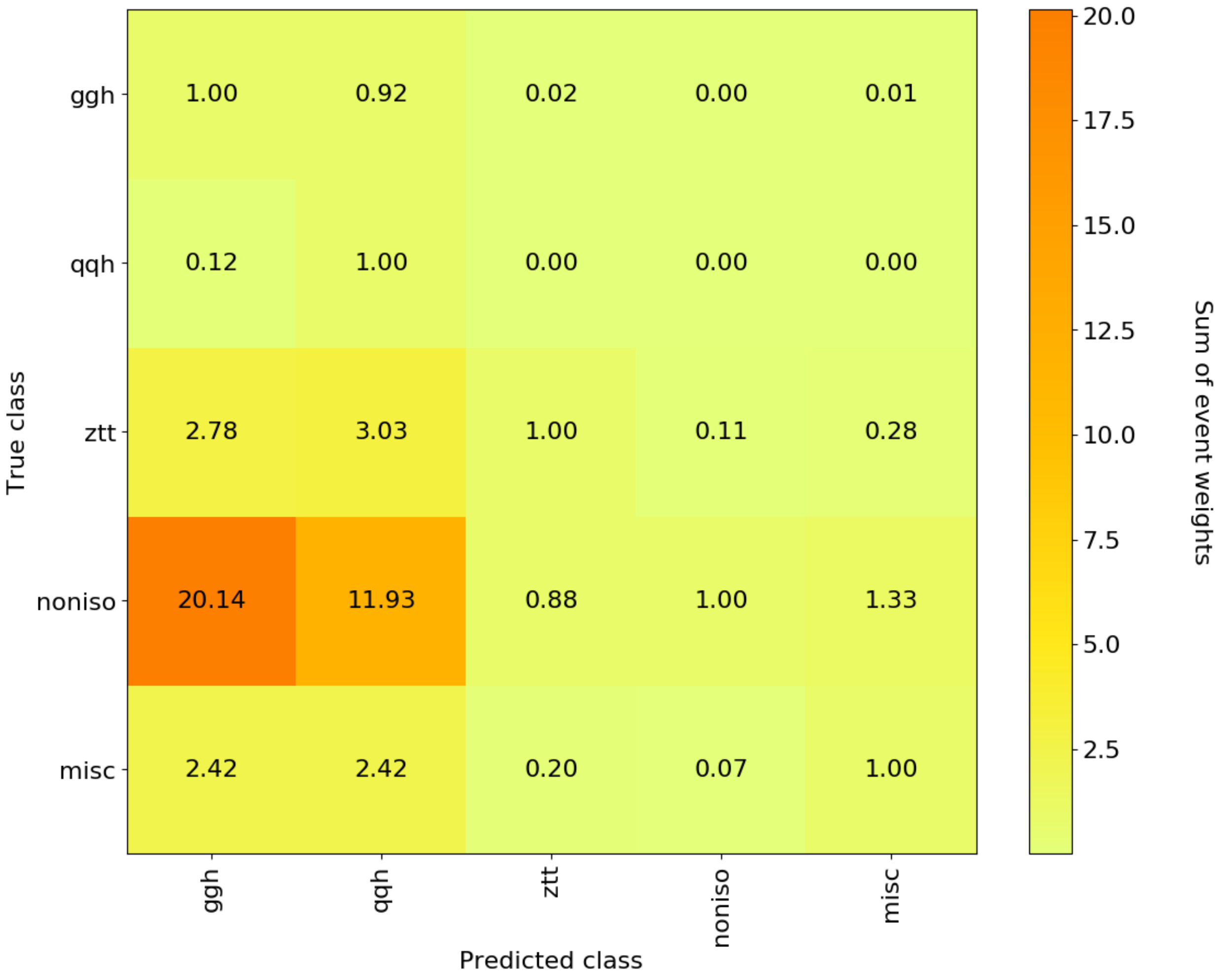


Pedrames’s NN output

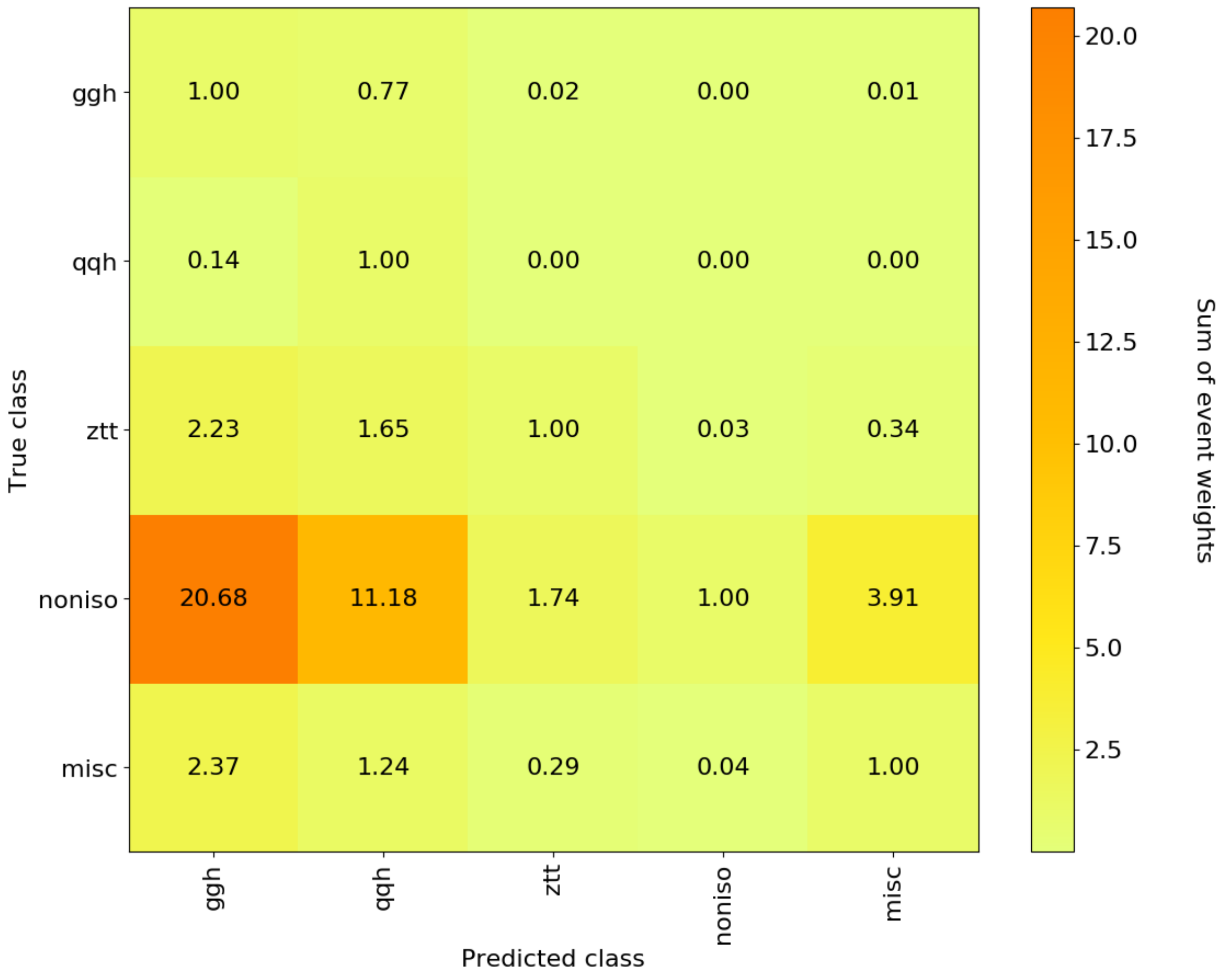


Nbtag = 0 NN output

Confusion matrix comparison for tt.

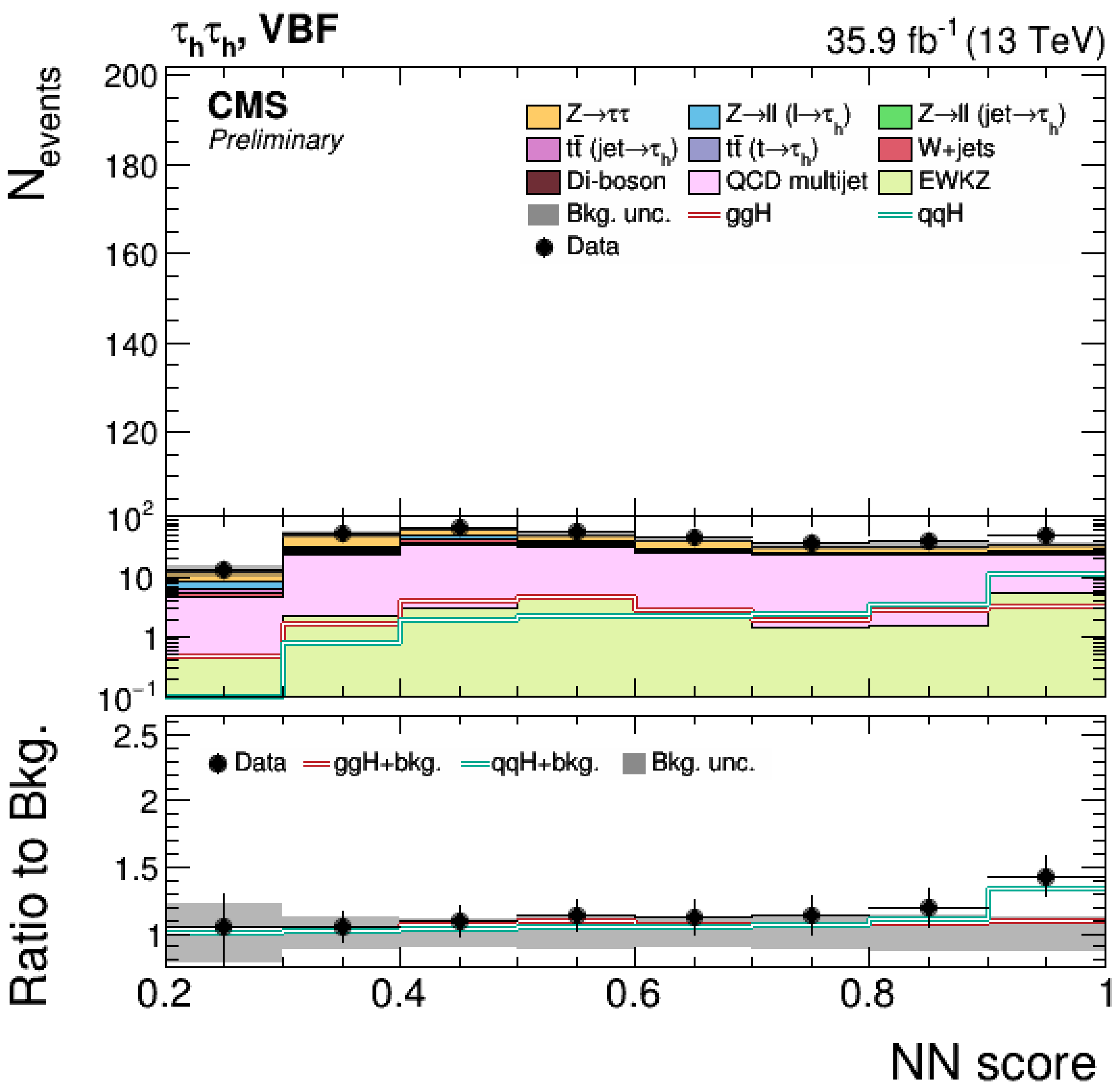


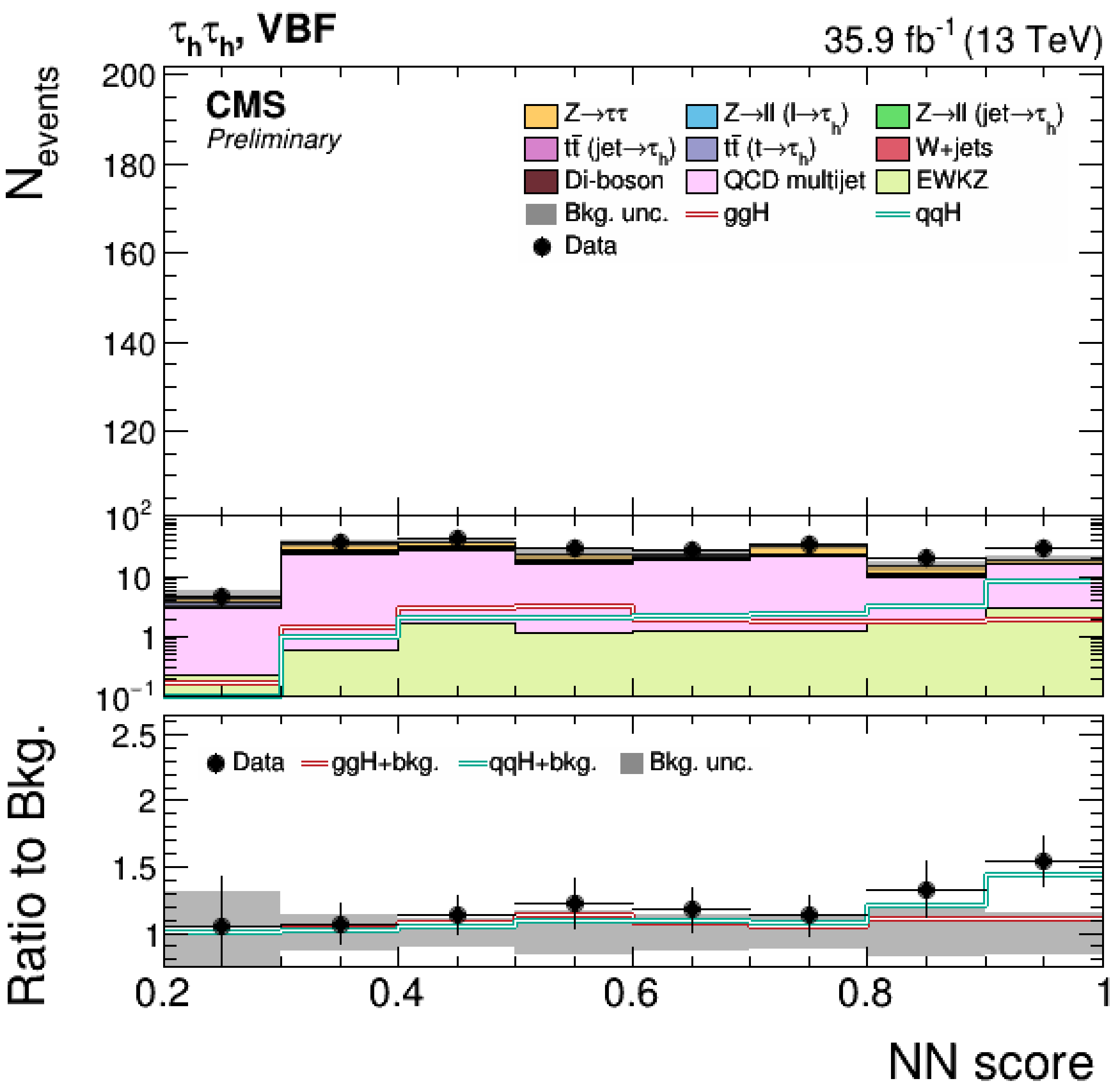
Pedrame's confusion matrix.



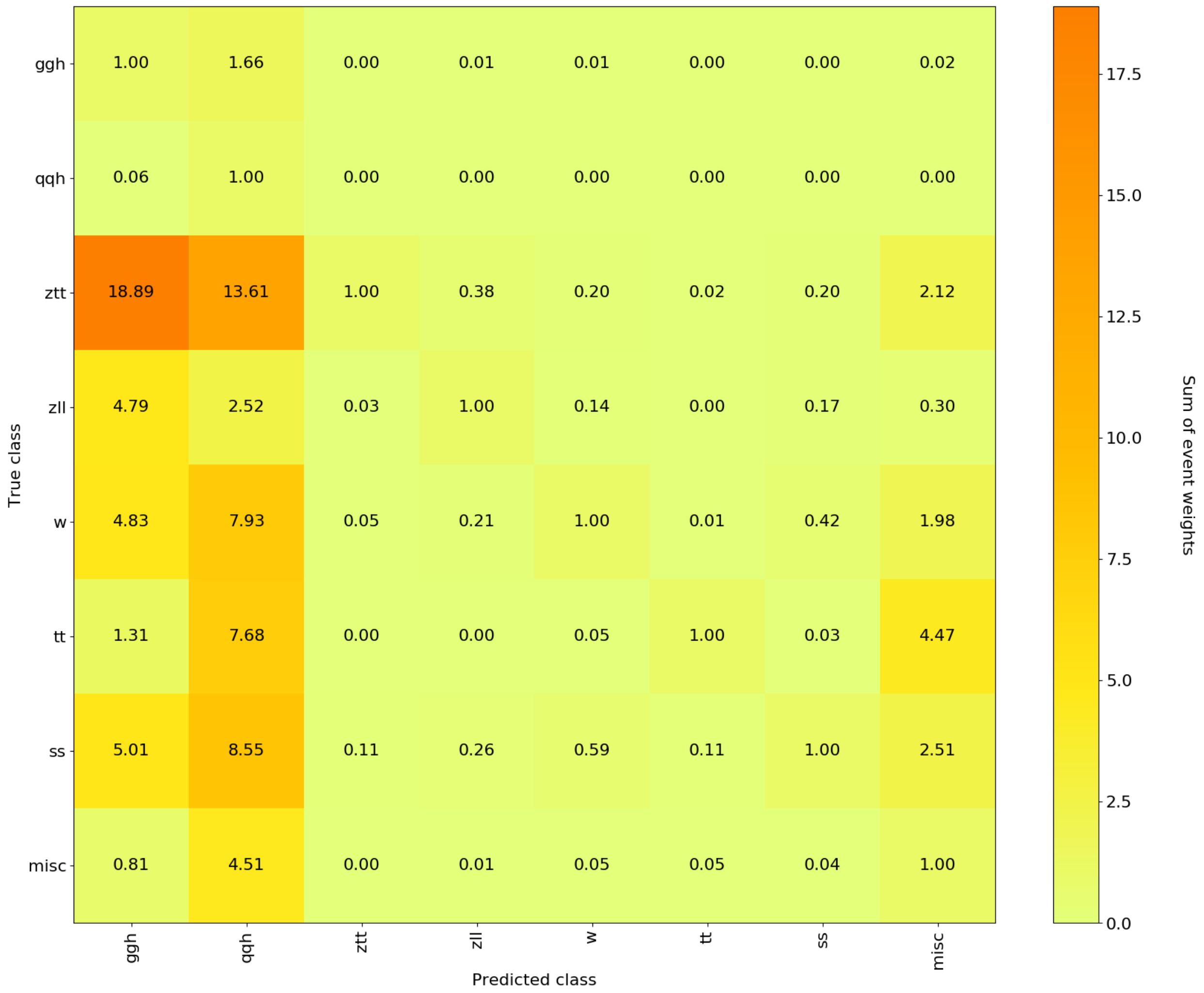
Nbtag == 0 confusion matrix.

Pedrame's vs Nbtag == 0: qqH NN scores for tt

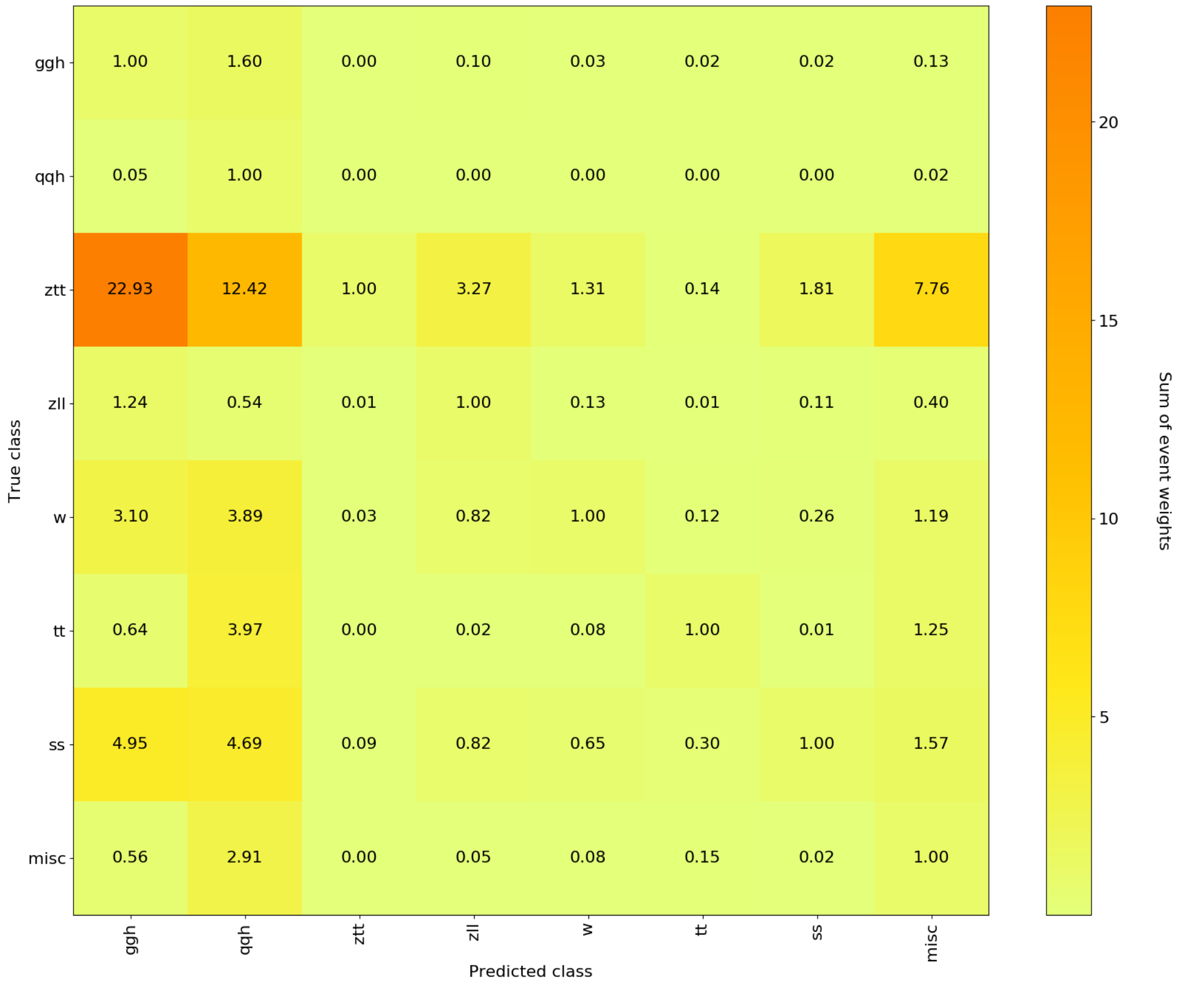




Confusion matrix comparison for mt channel on NN2

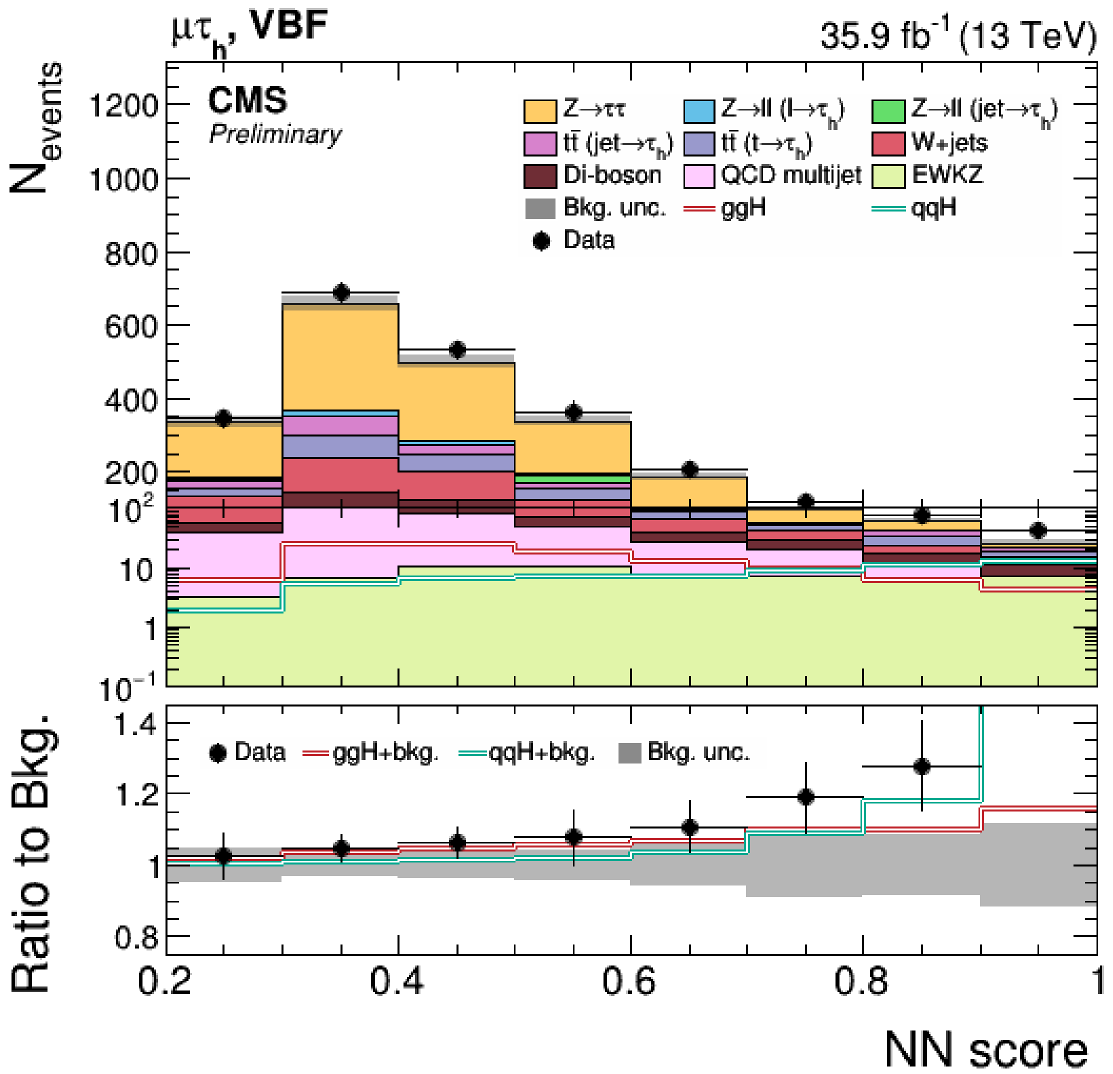


Pedrame's confusion matrix

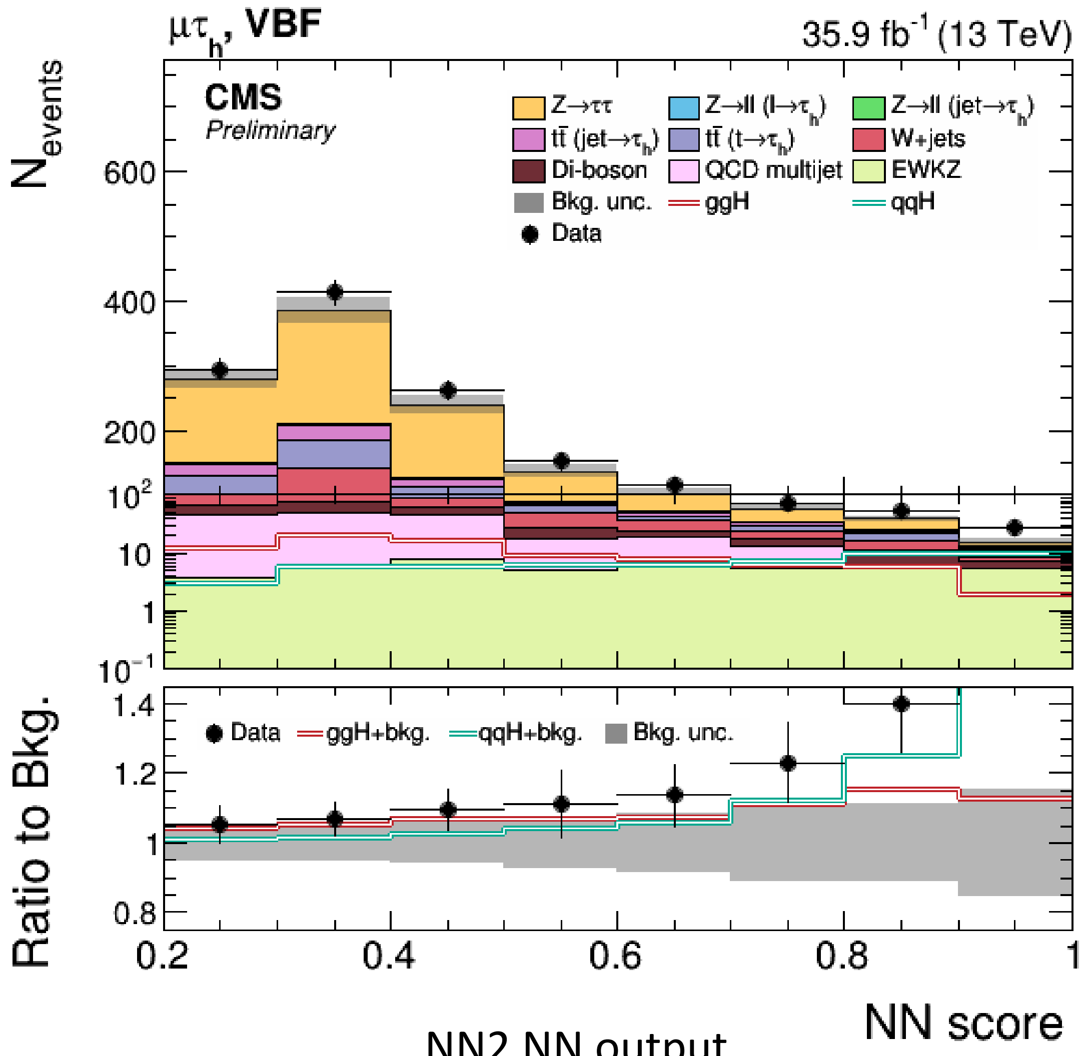


NN2 confusion matrix

Pedrame’s and new NN: qqH NN scores for mt

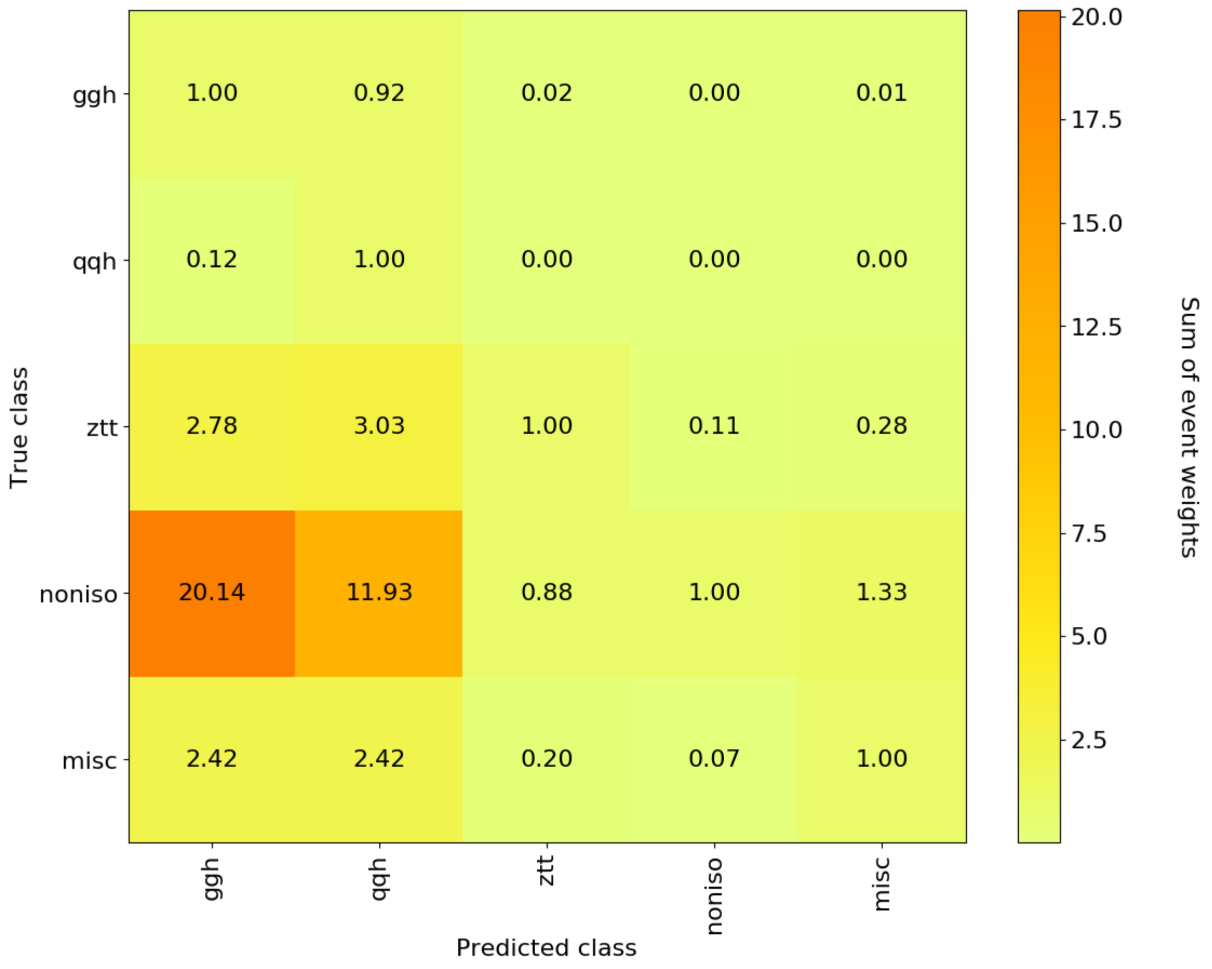


Pedrames’s NN output

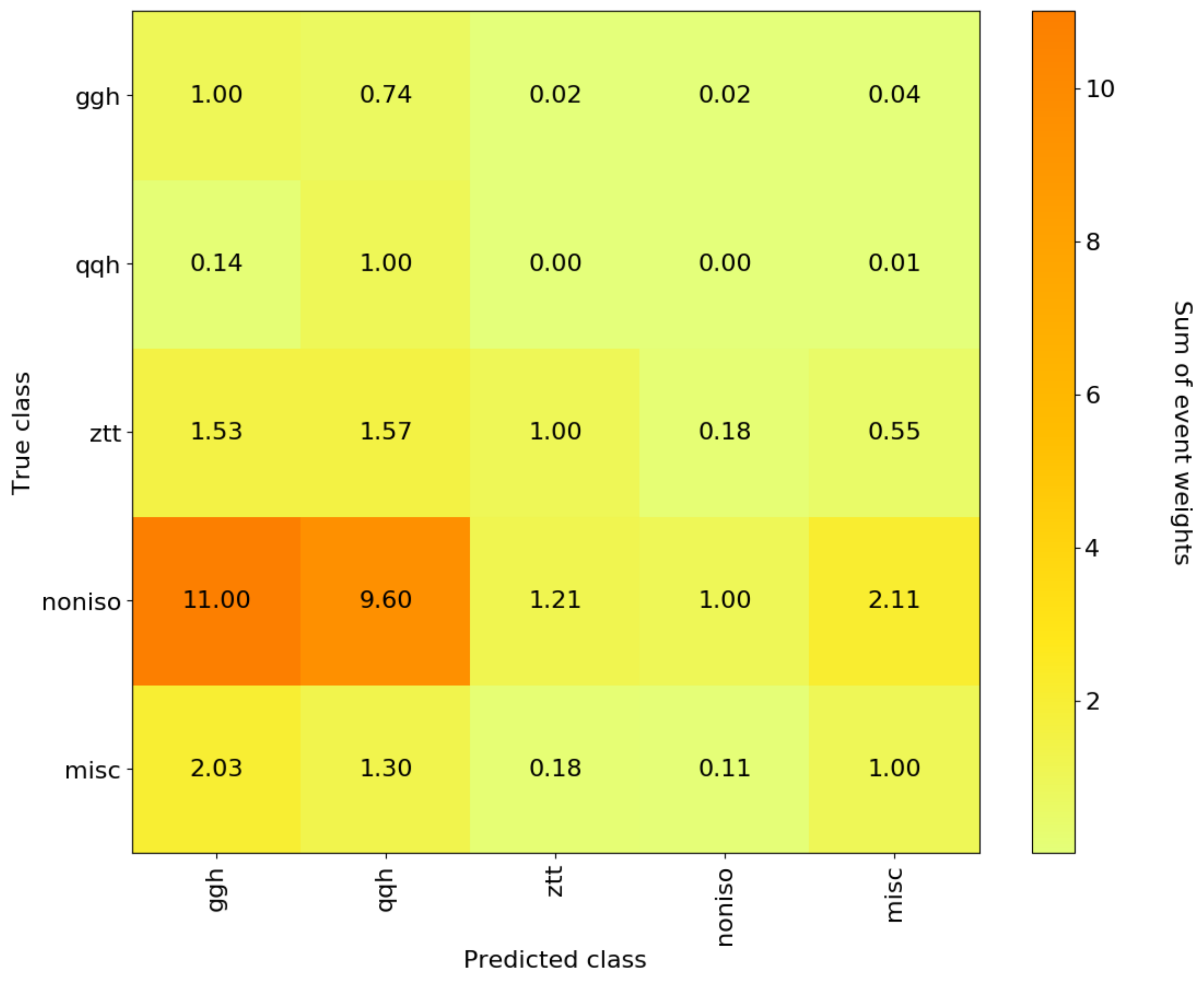


NN2 NN output

Confusion matrix comparison for tt channel on NN2

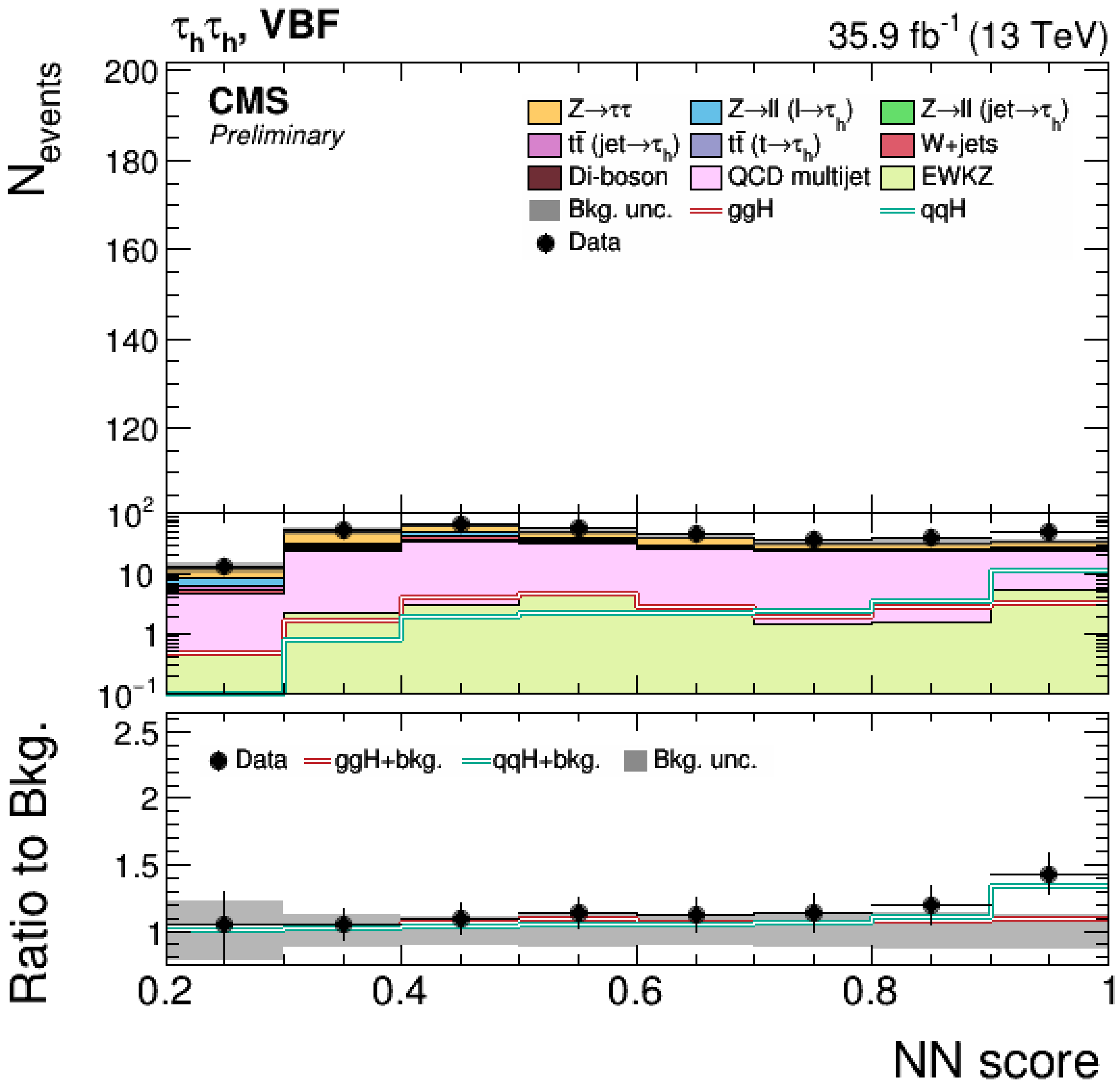


Pedrame's confusion matrix



NN2 confusion matrix

Pedrame's and NN2: qqH NN scores for tt



Waiting for results