Study of the Higgs boson in the WW channel

FCPPL team(2013)

LAL	Nanjing U. / USTC
Z. Zhang	S. Chen
S. Binet	Y. Zhu
Y. Li (Joint PhD)	

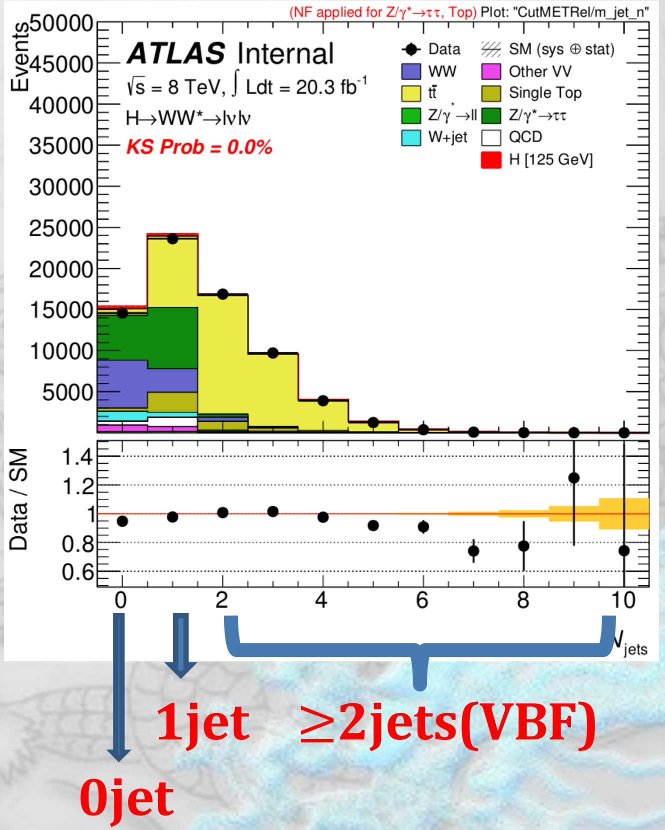
Outline

- Rate analysis in HWW:
  - top background study in 0jet channel
- Spin analysis in HWW
  - MVA optimization in 1jet channel
  - top background study 1jet channel
- Muon fake rates study
  - Fake muon suppression using BDT

Rate Analysis in HWW Channel

- HWW channel: Higgs → $WW^{(*)}$→ lvlv
- HWW channel is divided into 3 sub-channels according to the jet multiplicity in the final states (right plot):
    - 0jet, 1jet and VBF channels

- Rate analysis is to measure H production rate and test if it agrees with SM prediction by a likelihood fit
- Main backgrounds: di-boson, top, Z+jets and W+jets



Top Background in 0jet Channel of Rate Analysis

- Top process is one of the main backgrounds in 0jet channel

Process	Signal(MC)	Di-boson(MC)	Top(MC)	Z+jets(MC)	W+jets&QCD(data-driven)	data
0jet channel	303	7961	1216	6627	1682	14340

- It is important to normalize its MC prediction with data-driven method
  - In principle, there are always 2 bjets in top final states. But we only count jet whose pT is above 25 GeV, if both bjets are below this threshold, top event fall into 0jet channel.
- It includes several sub-processes:
  - ttbar and Wtop(main)
  - Single top of s/t channel
- There are several data-driven methods to determine the normalization of top quark background:
  - Jet Veto Survival Probability(JVSP) method(baseline)
  - Template method and its simplified version
  - In-situ b-tagging efficiency based method(IBEB)
- These methods, together with their performance in HWW+0jet analysis will be discussed and compared in the following

JVSP Method

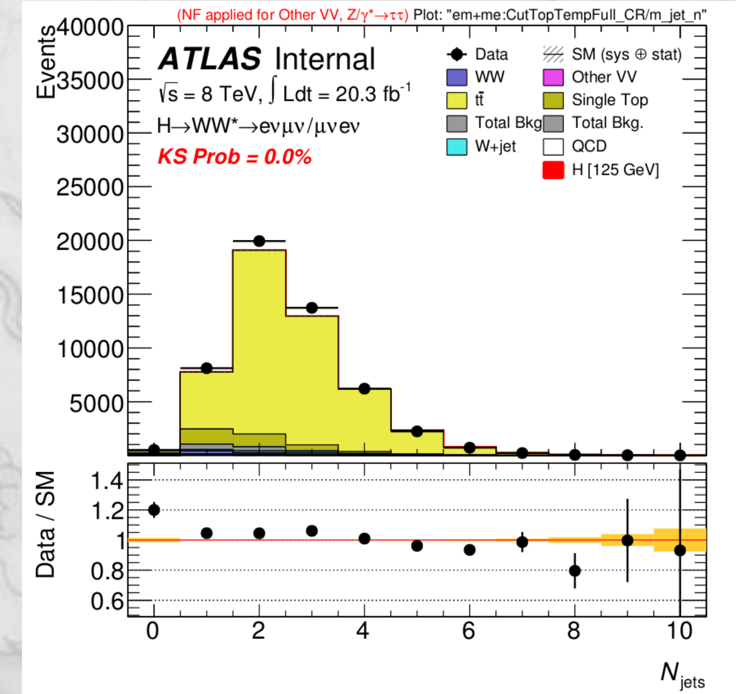
- Based on the simple formula below:
  - $N_{0\text{jet}} = N_{all} \times P_{0jet}$ (1)
  - N_{all} is the number of top events with all possible jet multiplicity
  - P_{0jet} is the probability of top events having 0 jet
- Principle: using data to estimate quantities at the right side of formula (1)

- For $\mathbf{N_{all}}$, we replace it with the prediction from (data – NonTop MC)
  - $N_{all} \rightarrow N_{all}^{\text{estimated}} = N_{all}^{data} - N_{all}^{NonTop,MC}$
- For P_{0jet}, we do a data-driven correction on MC predicted P_{0jet}^{MC}
  - $P_{0jet} \rightarrow P_{0jet}^{estimated} = P_{0jet}^{MC} \times (\frac{P_{1,CR}^{data}}{P_{1,CR}^{MC}})^2$, $\frac{P_{1,CR}^{data}}{P_{1,CR}^{MC}}$ is a per-bjet correction(since there are 2, it is squared) calculated in a very pure top CR
- Finally, we get the data-driven version of formula (1) :
  - $N_{0\text{jet}}^{estimated} = N_{all}^{estimated} \times P_{0jet}^{estimated}$

Template Method

- The goal is to get a data-driven template of the jet multiplicity distribution of top bkg in the signal region → its contribution in each sub-channel can be predicted just by counting events in the corresponding bin.
- First, a data-driven top jet multiplicity distribution is extracted in top CR by using data – NonTop MC (plot below)
  - $T^{CR}_{top,estimated} = T^{CR}_{data} - f \times T^{CR}_{Nontop}$
  - f is the normalization of non-top MC
- Then extrapolate this template from CR to SR by top MC predicted extrapolation factors:
  - $T^{SR}_{top,estimated} = \frac{T^{SR}_{top,MC}}{T^{CR}_{top,MC}} \times T^{CR}_{top,estimated}$
- The nomalization factor "f" is determined by a likelihood fit in SR using the above template.
- Finally, data-driven prediction of top events in 0jet channel is the 0jet bin of the template:

$$N^{estimated}_{top,0jet} = T^{SR}_{top,estimated}(0jet)$$



Simplified Template Method

- Instead of extracting template from top CR and doing a fit, we draw normalization factor for 0jet bin directly from a smaller top CR:
  - Defined inside 0jet sub-channel by doing a 20~25GeV btagging.
  - For 0jet channel, there is no jet with >25GeV, but 20~25GeV is possible.

Process	Signal(MC)	Di-boson(MC)	Top(MC)	Z+jets(MC)	W+jets&QCD(data-driven)	data
0jet & btagged	3	109	297	127	53	687

- We use the ratio $N = \frac{N^{CR}_{data-nontop,MC}}{N^{CR}_{top,MC}}$ to normalize MC predicted top events in 0jet bin $N^{MC}_{top,0jet} \rightarrow N^{estimated}_{top,0jet} = N^{MC}_{top,SR} \times N$

IBEB Method

- Based on the formula below:
  - $N_{top,0jet} = \frac{N_{top,tagged}}{\epsilon_{top,tag}}$ (2)
  - $N_{top,tagged}$ is the number of top events in 0jet channel being tagged to have bjet
  - $\epsilon_{top,tag}$ is the top tag efficiency
- Same principle as JVSP method
- For $N_{top,tagged}$, we estimated it using (data – NonTop MC)
  - $N_{top,tagged} \rightarrow N^{estimated}_{top,tagged} = N^{data}_{top,tagged} - N^{NonTop,MC}_{top,tagged}$
- For $\epsilon_{top,tag}$, we do a data-driven correction on MC predicted $\epsilon^{MC}_{top,tag}$
  - $\epsilon^{MC}_{top,tag} \rightarrow \epsilon^{estimated}_{top,tag} = \epsilon^{MC}_{top,tag} \times \frac{\epsilon^{data}_{top,tag,CR}}{\epsilon^{MC}_{top,tag,CR}}$, $\frac{\epsilon^{data}_{top,tag,CR}}{\epsilon^{MC}_{top,tag,CR}}$ is a correction factor derived in top CR
- Finally, we get the data-driven version of formula (2) :
  - $N^{estimated}_{top,0jet} = \frac{N^{estimated}_{top,tagged}}{\epsilon^{estimated}_{top,tag}}$

Comparison between Each Method

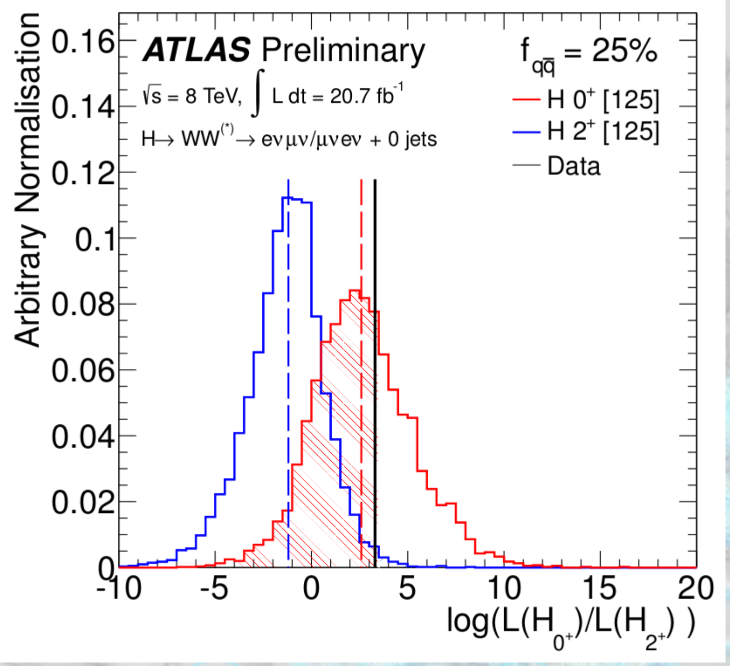
- Here are the results of each method and their corresponding sys errs

Mothod	JVSP	Template	Template(Simp)	IBEB
Stat	1.8%	7.6%	6.5%	6.6%
Exp.	4.8%	6.5%	6.5%	1.5%
Theo.	3.0%	3.5%	3.5%	1.8%
NonTop	2.7%	7.1%	7.1%	7.1%
Total sys err	6.3%	10.2%	10.2%	7.5%
Total sys + stat err	6.5%	12.8%	12.1%	10.0%
NF	1.11 ± 0.07	1.32 ± 0.17	1.38 ± 0.17	1.37 ± 0.14
Comment	Most precision Limited by exp err	Except for theo. err All others are large	Except for theo. err All others are large	Least affected by exp. and theo. err

- There results are preliminary(approximations are used)
- JVSP is the best (so it's our baseline method)

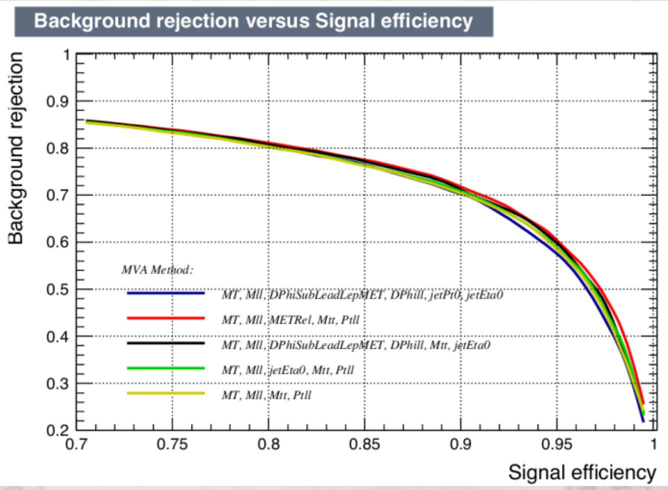
Spin Analysis in HWW Channel

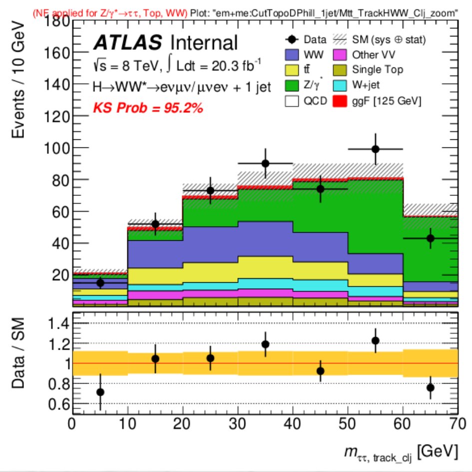
- Two spin models are considered: spin0, spin2
- Method to test which model is preferable:
  - Different H spin -> different topo structure of final state(angles between objects)
  - Use quantities describing this structure to train BDT: 2 BDTs are trained, one from spin0 vs All bkg, the other from spin2 vs All bkg.
  - Construct a 2-D BDT and fit it to data.
- An example fit shown at right plot.
- Different training variables are used in 0jet and 1jet channels
- Only 0jet channel was used previously

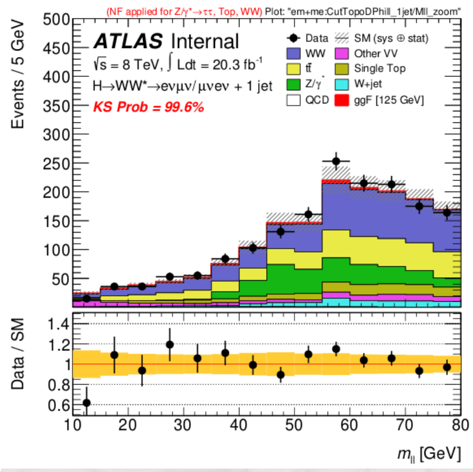


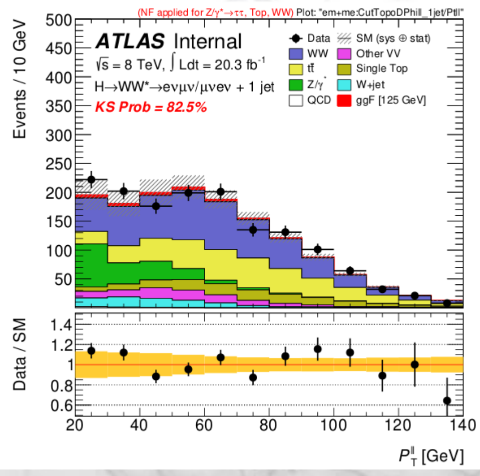
BDT optimization in 1jet Channel of Spin Analysis

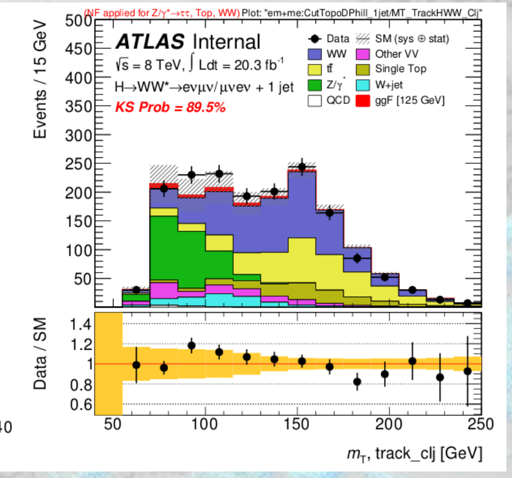
- To include 1jet bin as well, need optimization
- Tried various combination of input variables and get several sets having similar performance(right plot) → choose the simplest one, which are shown below:



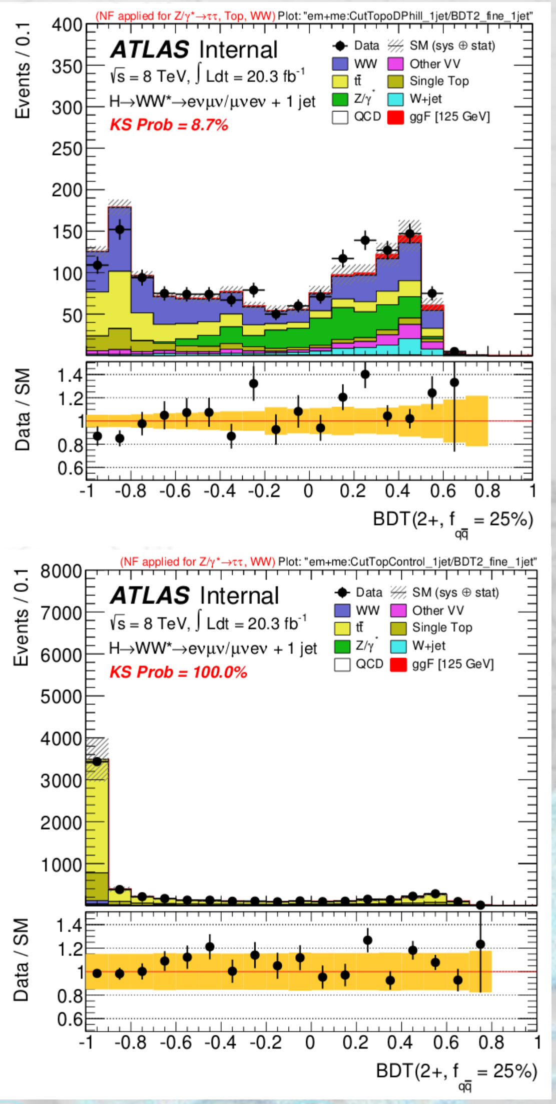








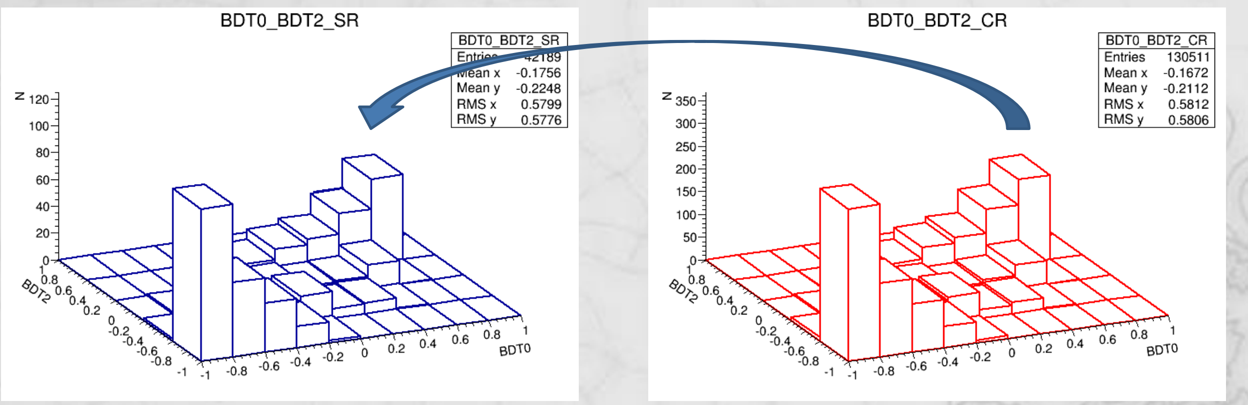
Top Background in 1jet Channel of Spin Analysis



- Top is one of the main bkg in 1jet channel(right upper plot)
- We need to constrain top bkg's shape in SR

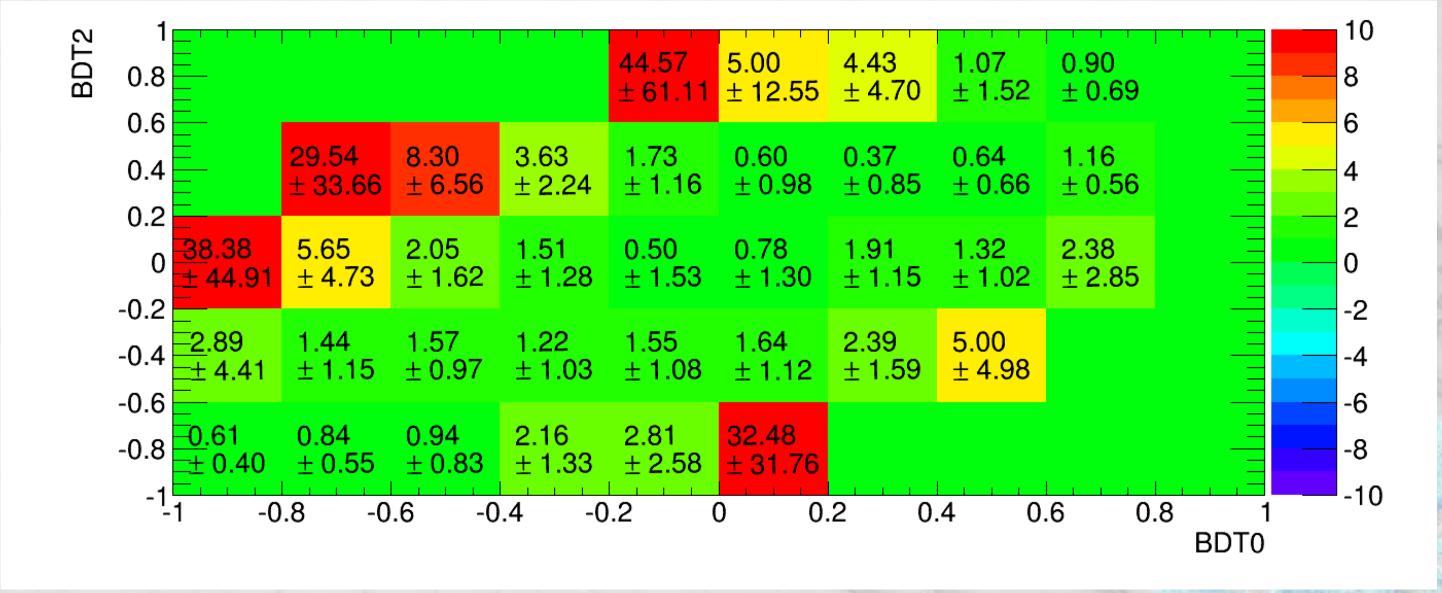
  This can be done by extrapolating the shape from CR to SR

  The CR is defined by having non-zero bjet(right bottom plot)
- Then we need to evaluate sys errs for the extrapolation factor α = N_SR/N_CR in each bin. (graphics below)



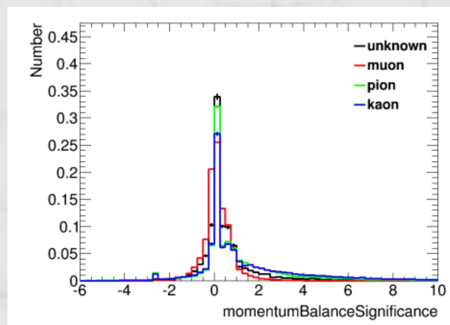
Theoretical Uncertainties of Top Shape Extrapolation

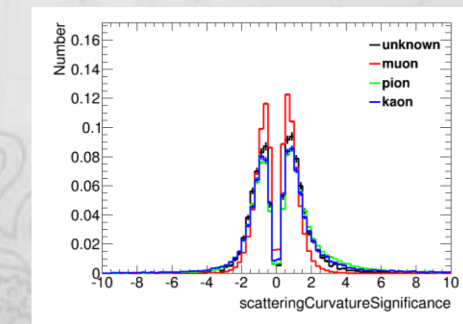
- Theoretical uncertainties for the extrapolation are studied using the same group of samples in top 0jet case(privately produced, large stat ~ 10000 fb^{-1})
- Take the QCD scale uncertainty as an example here
- Large stat sample → stat uncertainty ~ 1%
- Stat distribute around diagonal → errs on boarder bins can't be estimated precisely → but they contributes little to the fit

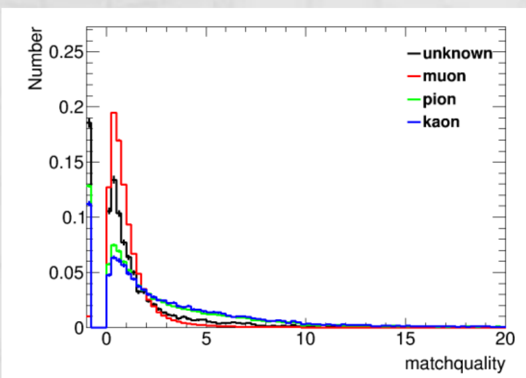


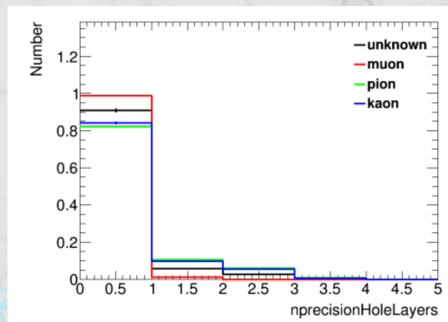
Other activities: Fake Muon Study

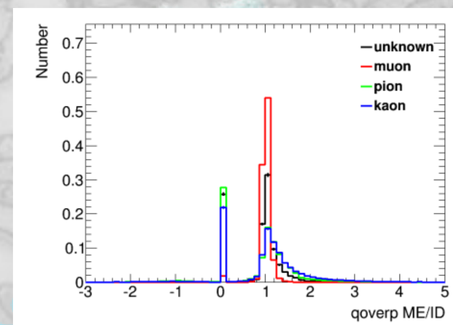
- This work could contribute to HWW analysis indirectly: since an quality-improved muon collections will reduce some main background, such as Wjets. Of course other analysis will also benefit(Bs to mumu rare decay search)
- The motivation is the existing working point for 3rd chain muon(the future muon collection that will be used) is not tight enough(too many fake muon), so we tried to train a BDT to suppress the fakes
- 5 variables that are chosen:









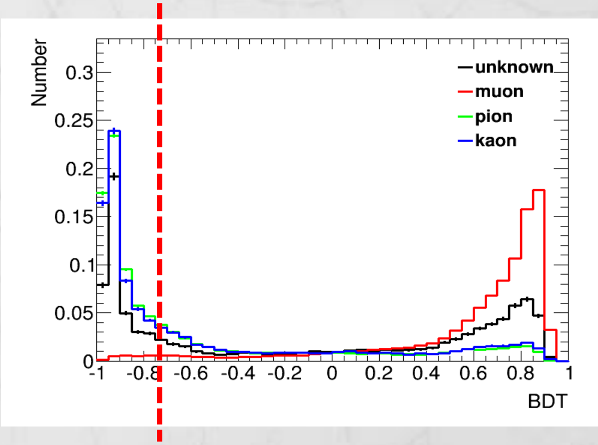


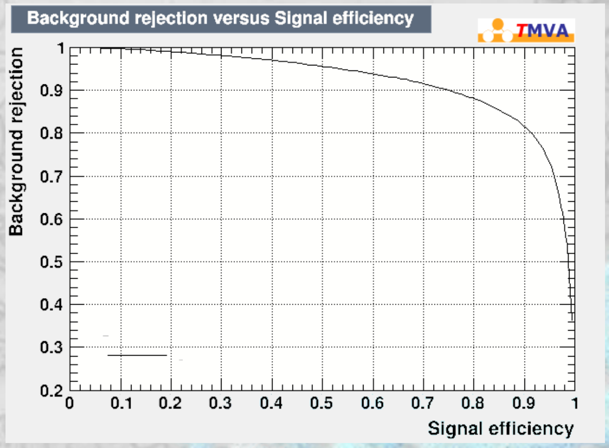
Expected Performance of the BDT

- Output BDT shapes and its ROC are shown in right plots.
- Choose the cut on BDT at which we'll have around the same true muon efficiency as the existing working point
- Comparison with the existing wokring point in terms of true muon efficiency / fake muon reduction

Effi/Redu.(%)	Muon	Pion	Kaon	Proton	Others	unknown
Medium+	97.87	-37.22	-37.02	-60.07	-55.74	-28.35
BDT	97.83	-59.60	-57.25	-80.78	-82.64	-37.12
Improvement	-0.04	60.12	54.66	34.49	48.25	30.92

- When having same true muon efficiency, BDT has almost 2 times stronger fake reduction than medium+ does
- Next step → validation of input variable shapes using data in Z->mumu control region
  - Some discrepancies are observed → to be understood and fixed.





Summary

- In HWW+0jet rate analysis, we normalized the top bkg using data-driven method(JVSP) and compared other optional methods with it:
  - JVSP method is the best.
- In HWW+1jet spin analysis, we optimized the BDT training and constraint top's shape as well as the theoretical uncertainties on this constraint.
- We have a promising fake muon suppression method which needs further study.

- We will continue the project with an extended term in 2014:
- LAL: Z. Zhang, D. Rousseau
- Nanjing U. / Shandong U. / USTC
- S. Chen, L. Ma, Y. Zhu
- Y. Li (Joint PhD student)