Search for H → WW* → lνlν
Based on Boosted Decision Trees

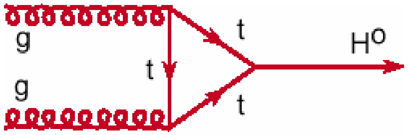
Hai-Jun Yang
University of Michigan

LHC Physics Signature Workshop
January 5-11, 2008

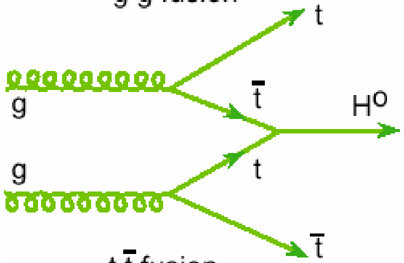
Outline

- H→WW* – a possible early discovery channel
- Brief Introduction of Boosted-Decision-Trees
- H → WW* → lνlν analysis based on BDT
- ATLAS Sensitivity of H → WW* → lνlν
- Summary and Outlook

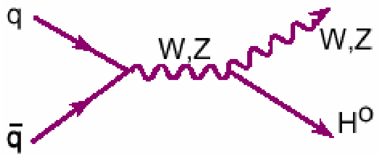
Higgs Production at LHC



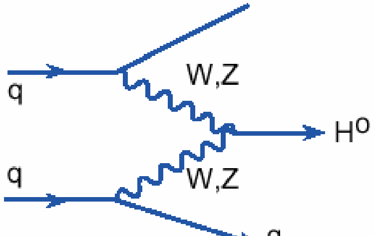
g g fusion



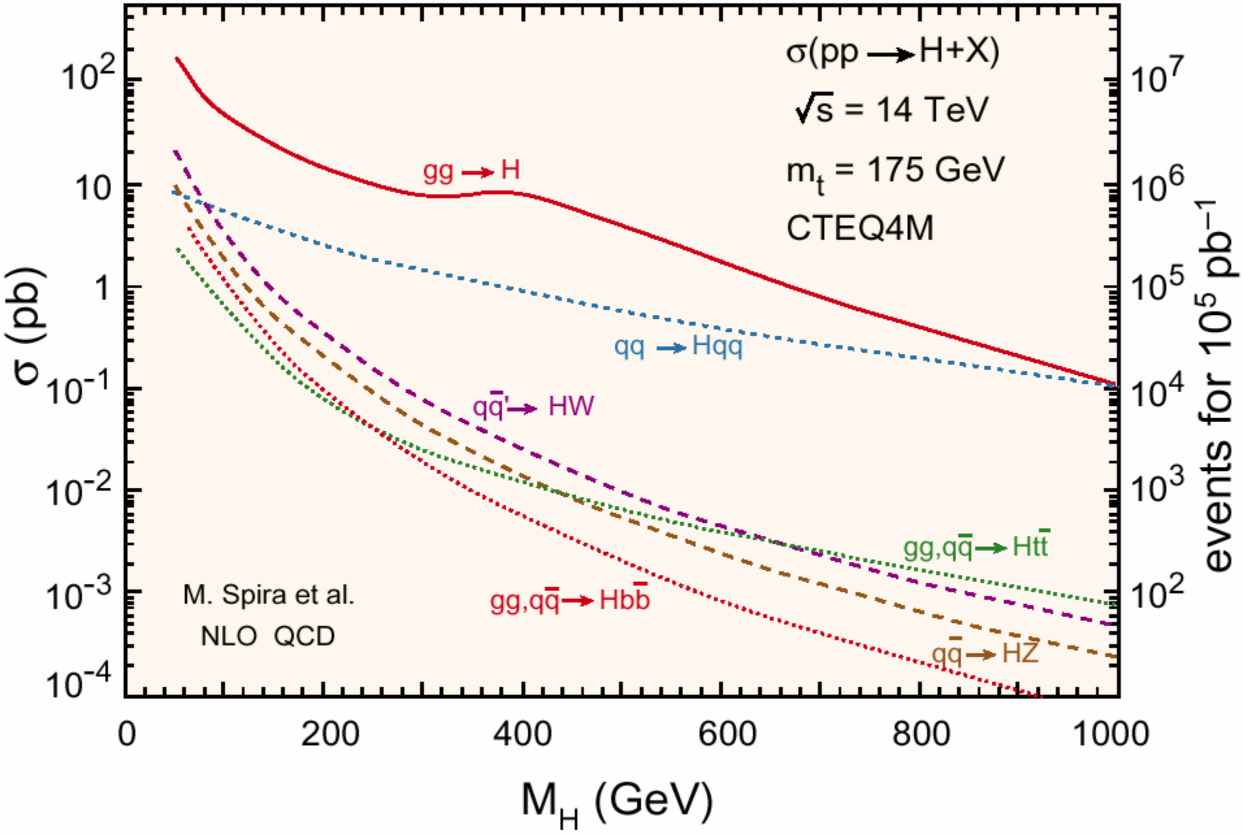
t $\bar{t}$ fusion



W, Z bremsstrahlung



WW, ZZ fusion



➔ Gluon-gluon fusion and WW/ZZ fusion are two dominant Higgs production mechanism.

Higgs Decay Branching Ratio and Discovery Channels

Low mass region: m(H) < 2 m_Z

$H \to \gamma\gamma$

$H \to bb$

$H \to \tau\tau$

$H \to ZZ^* \to 4\ell$

$H \to WW^* \to \ell\nu\ell\nu$ or $\ell\nu jj$

m(H) > 2 m_Z

$H \to ZZ \to 4\ell$

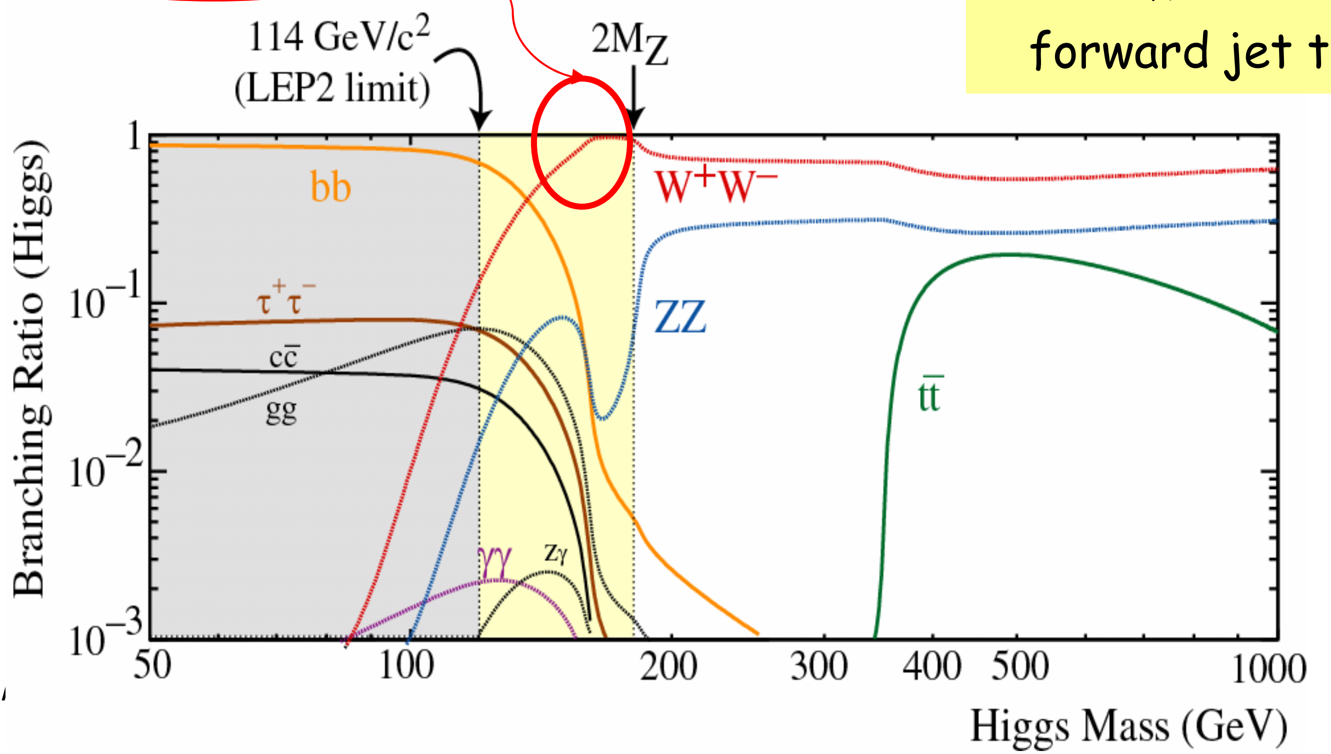
$qqH \to ZZ \to \ell\ell\ \nu\nu$ *

$qqH \to ZZ \to \ell\ell\ jj$ *

$qqH \to WW \to \ell\nu jj$ *

* for $m_H > 300$ GeV

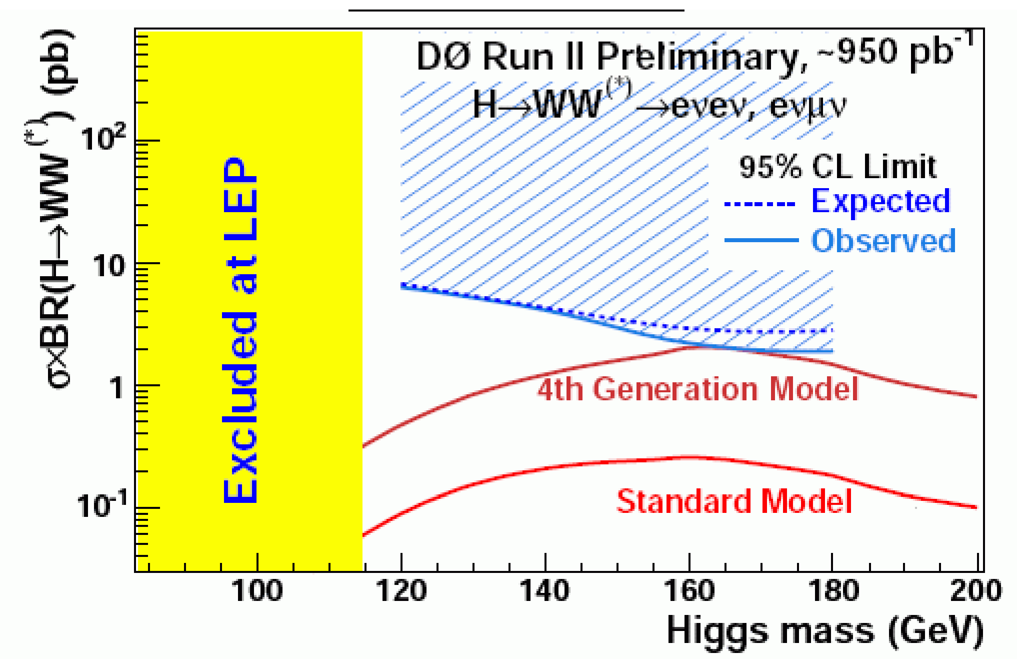
forward jet tag

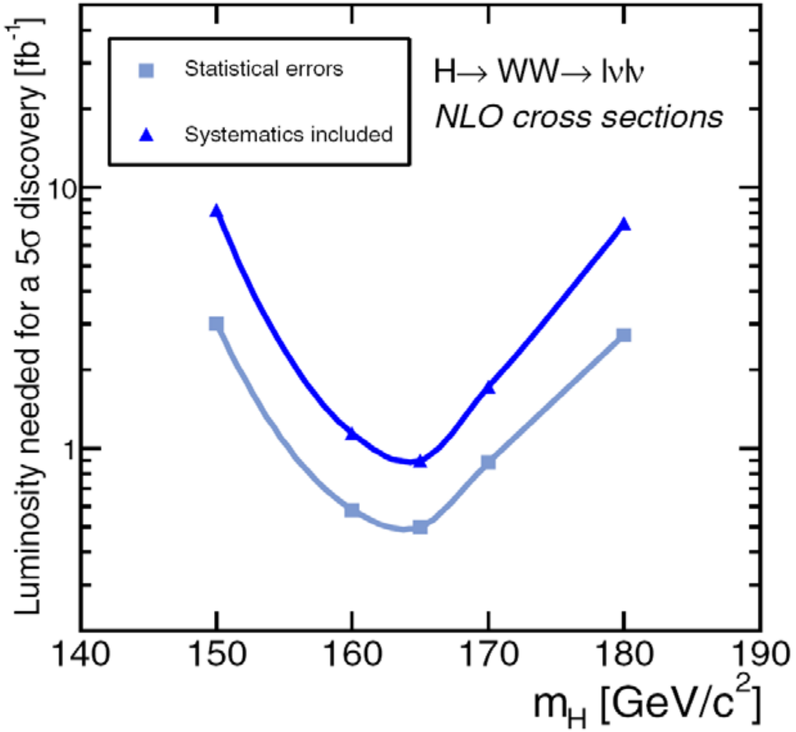


H → WW* → lνlν
Current limit and discovery potential at LHC

Excluded cross section times Branching Ratio at 95% C.L.

CMS Phys. TDR 2006



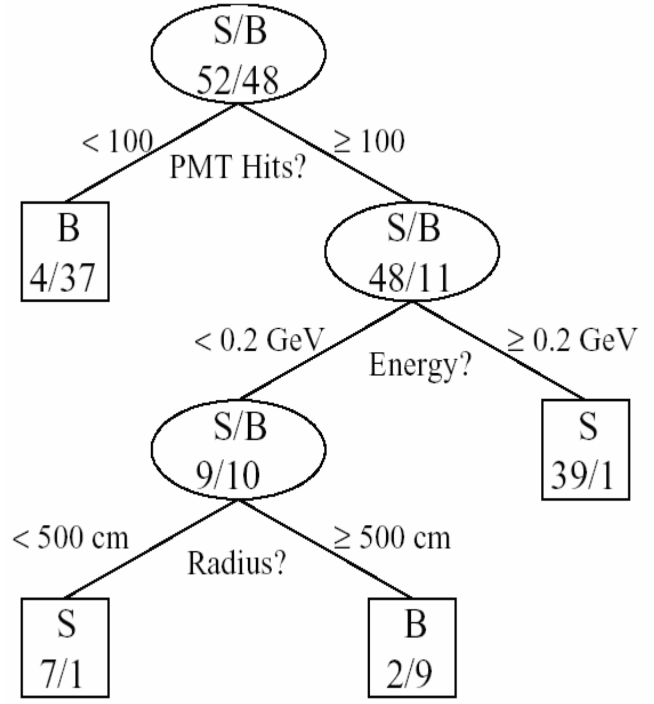


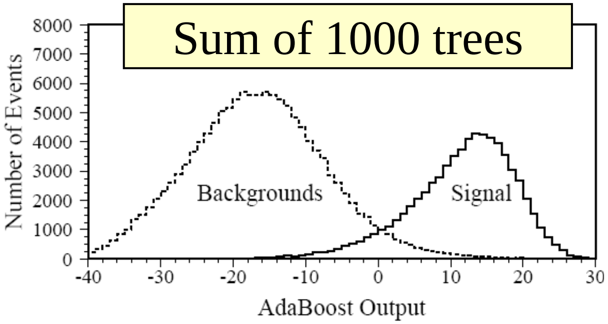
ATLAS Physics ‘Commissioning’

- Study the new physics discovery potential with CSC (computing system commissioning) program (started from summer of 2006)
- Physics ‘TDR’ will be updated soon with ATLAS CSC note using many 10th of Million fully simulated CSC MC data sets and with advanced analysis tools.
- We have developed and applied the BDT technique in diboson physics and Higgs discovery studies with the ATLAS CSC program.

Boosted Decision Trees (BDT)

➔Relative new in HEP – MiniBooNE, BaBar, D0(single top discovery), ATLAS
➔Advantages: robust, understand ‘powerful’ variables, ‘not a black box’, …



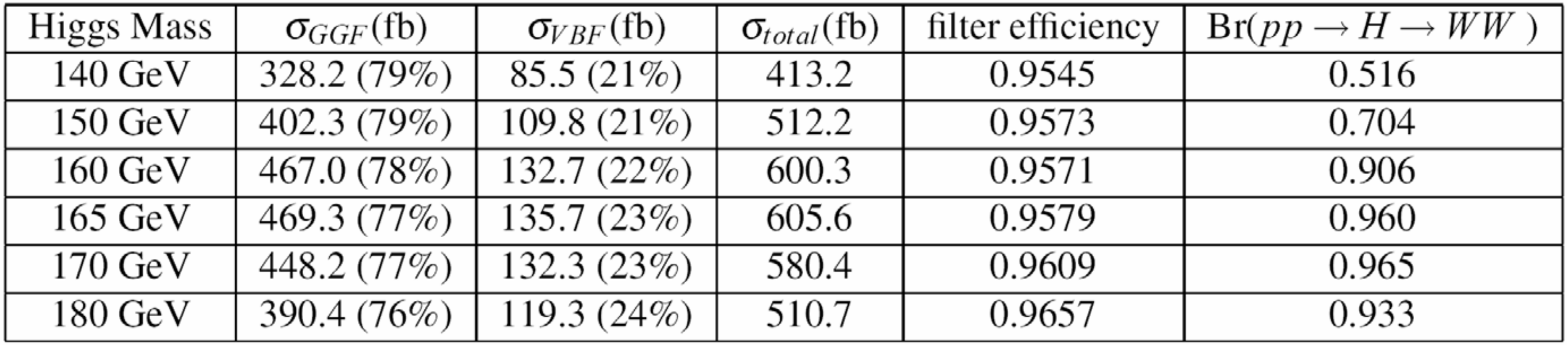


- Split data recursively based on input variables until a stopping criterion is reached (e.g. purity, too few events)
- Every event ends up in a “signal” or a “background” leaf
- Misclassified events will be given larger weight in the next tree (boosting)
- For a given event, if it lands on the signal leaf in one tree, it is given a score of 1, otherwise, -1. The sum of scores from all trees is the final score of the event.

B.P. Roe, H.J. Yang, et.al., physics/0408124, NIM A543 (2005) 577
H.J. Yang, B.P. Roe, et.al., physics/0610276, NIM A574 (2007) 342

H → WW* → lνlν (l = e, μ)

- Cross sections of H → WW* → lνlν (GGF & VBF) at LO (Pythia), K-factor ~ 1.9



Higgs Mass	σ_{GGF}(fb)	σ_{VBF}(fb)	σ_{total}(fb)	filter efficiency	Br($pp \to H \to WW$)
140 GeV	328.2 (79%)	85.5 (21%)	413.2	0.9545	0.516
150 GeV	402.3 (79%)	109.8 (21%)	512.2	0.9573	0.704
160 GeV	467.0 (78%)	132.7 (22%)	600.3	0.9571	0.906
165 GeV	469.3 (77%)	135.7 (23%)	605.6	0.9579	0.960
170 GeV	448.2 (77%)	132.3 (23%)	580.4	0.9609	0.965
180 GeV	390.4 (76%)	119.3 (24%)	510.7	0.9657	0.933

H → WW signal and background simulations used
ATLAS software release v12 (for CSC note)
Full ATLAS detector simulation and reconstruction

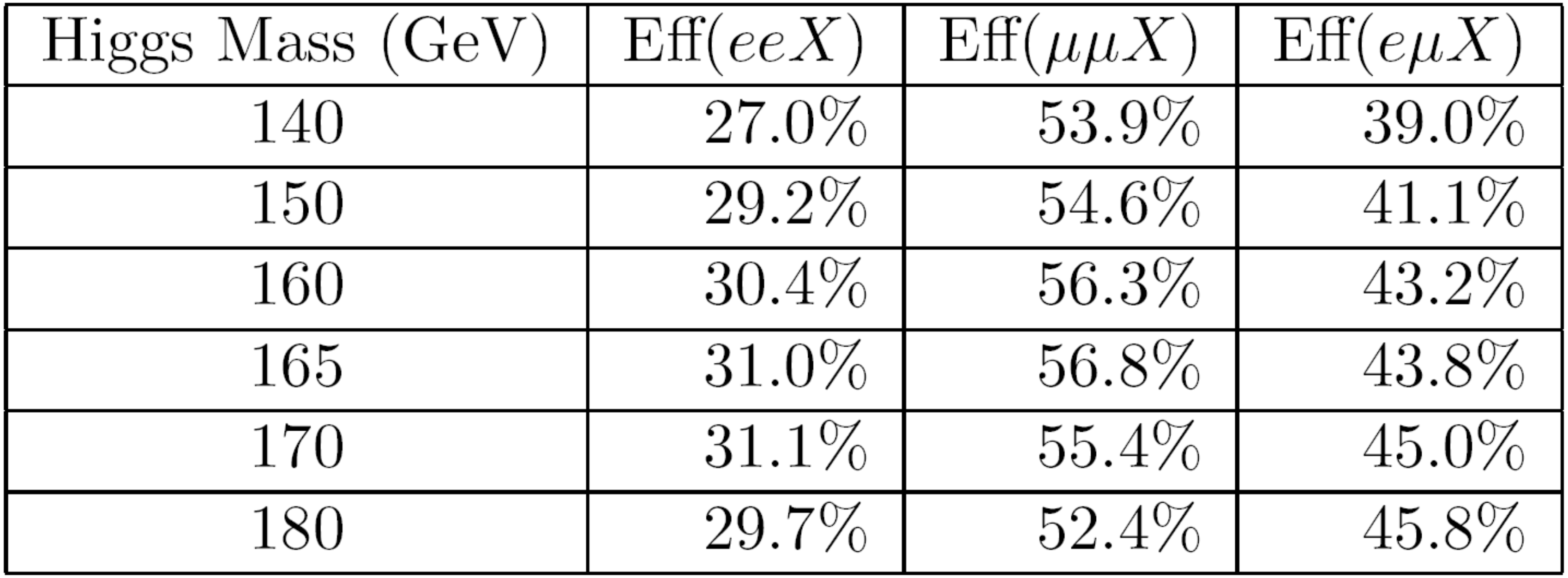
Backgrounds

Process	MC sample	cross-section
WW → lvlv (l=e,μ,τ)	372.5K,	11.72 pb
gg2WW → lvlv (l=e,μ,τ)	209.1K,	0.54 pb
ttbar →l + X	584.1K,	450.0 pb
WZ → lvll (l=e,μ)	281.4K,	0.7 pb
Z → ll (l=e,μ,τ)	1.15 M,	4.6 nb

- W/Z + Jets are potential background, using 1.1M fully simulated MC events (Alpgen generator), no event is selected in our final sample.
- Background estimate uncertainty ~ 15 – 20 %.

H→ WW Pre-selection

- At least one lepton pair (ee, μμ, eμ) with $P_T > 10$ GeV, $|\eta|<2.5$
- Missing $E_T > 15$ GeV
- $|M_{ee} – M_z| > 10$ GeV, $|M_{\mu\mu} – M_z| > 15$ GeV to suppress background from Z → ee, μμ



Higgs Mass (GeV)	Eff(eeX)	Eff($\mu\mu X$)	Eff($e\mu X$)
140	27.0%	53.9%	39.0%
150	29.2%	54.6%	41.1%
160	30.4%	56.3%	43.2%
165	31.0%	56.8%	43.8%
170	31.1%	55.4%	45.0%
180	29.7%	52.4%	45.8%

- **IsEM & 0x7FF == 0 (tight electron id cuts)**
- **Staco-muon id**

BDT Training with pre-selected events

Input physics variables to BDT program (1)

- Energy and Momentum

  - $p_T(\ell)$, $p_T(\ell,\ell)$
  - MET, total recoil E_T
  - scalar $\sum E_T(jet)$, vector $\sum E_T(\ell, MET)$

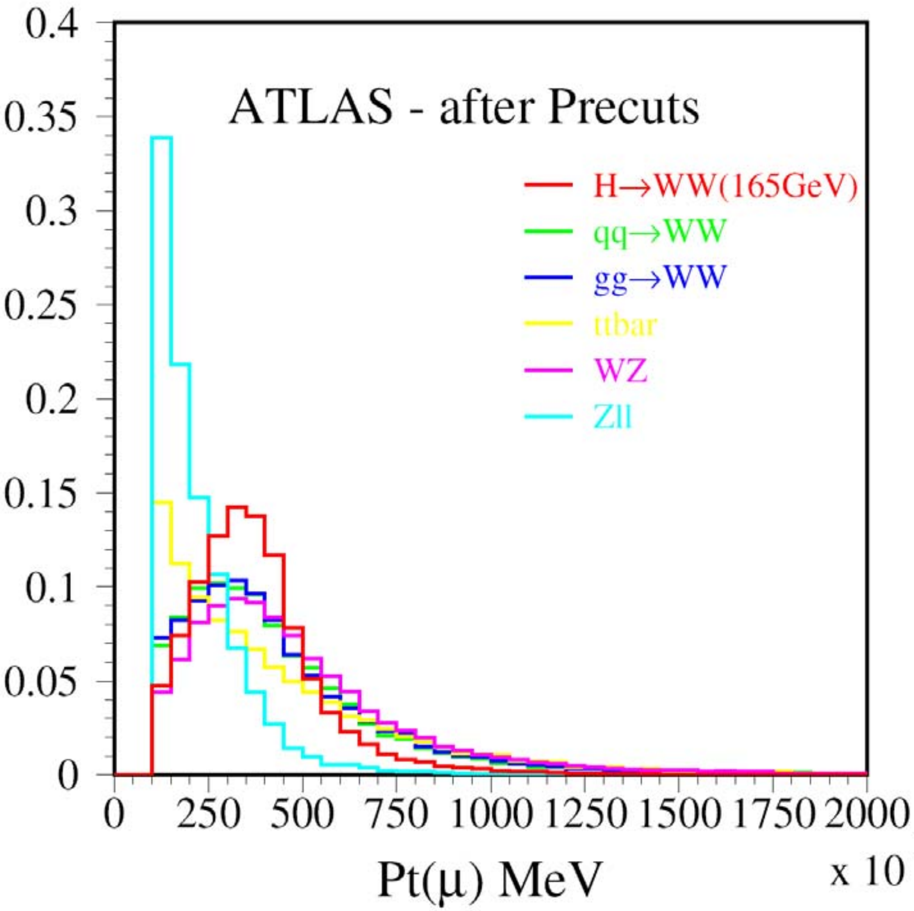
- Lepton Isolation

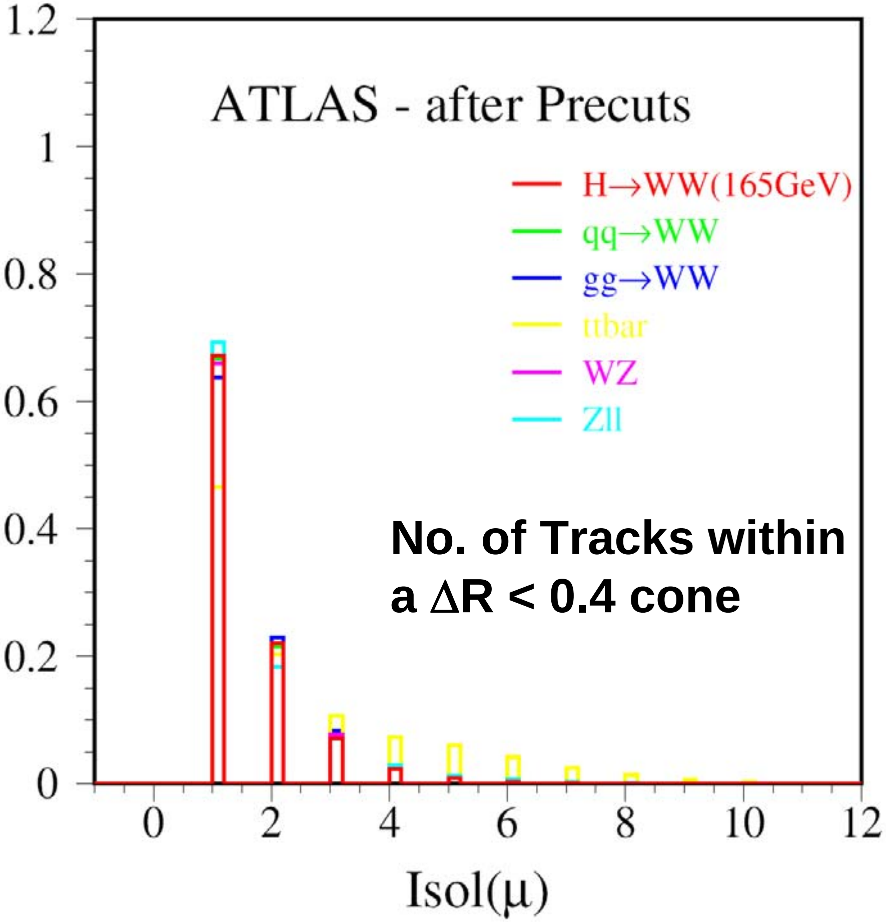
  - Number of tracks in $\Delta R < 0.4$ cone around ℓ
  - Sum of track p_T in $\Delta R < 0.4$ cone around ℓ
  - Sum of jet E_T in $\Delta R < 0.4$ cone around ℓ

Input physics variables to BDT program (2)

- Event Topology
  - Number of Jets with E_T >30 GeV
  - E(ℓ)/P(ℓ)
  - A0 (impact parameter) of ℓ, $\Delta A0(\ell,\ell)$, $\Delta Z(\ell,\ell)$
  - $\Delta R(\ell,\ell)$, $\Delta\phi(\ell,\ell)$, $\Delta\phi(\ell,MET)$
  - $\Delta\Omega(\ell,\ell)$ - opening angle of two leptons
- Mass Information
  - Invariant mass(ℓ,ℓ)
  - Transverse mass($\ell\ell$,MET)
  - Transverse mass(ℓ,MET)

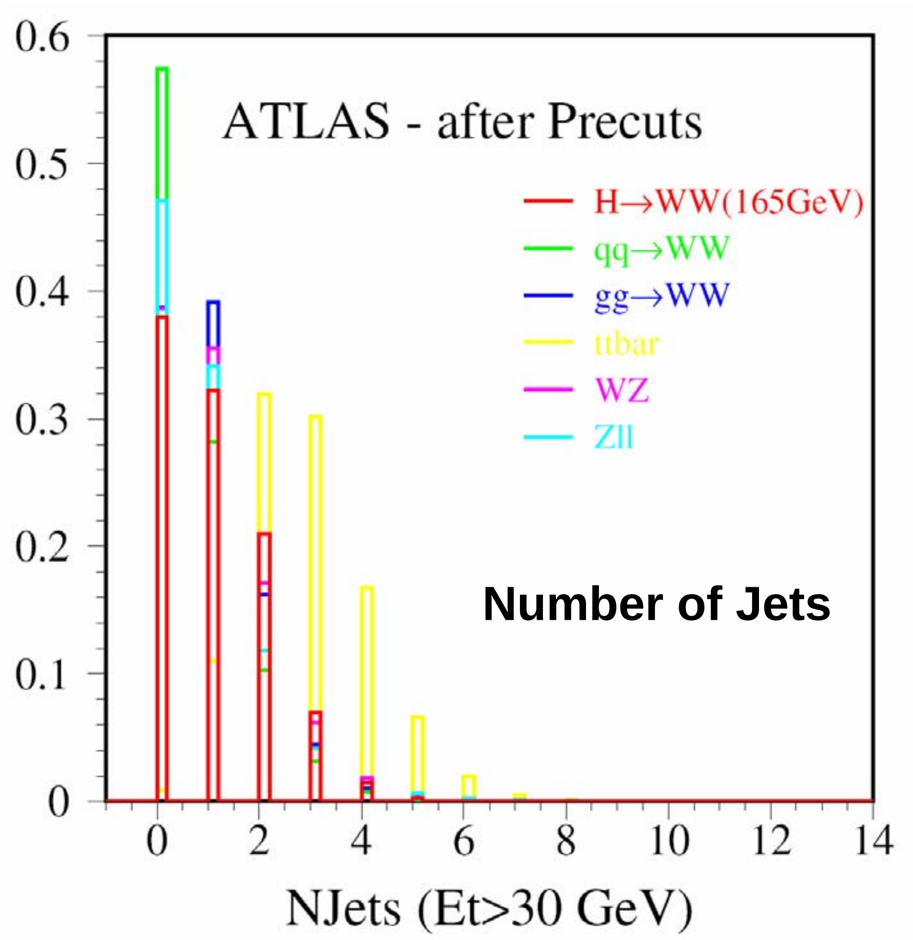
Some Training Variable Distributions



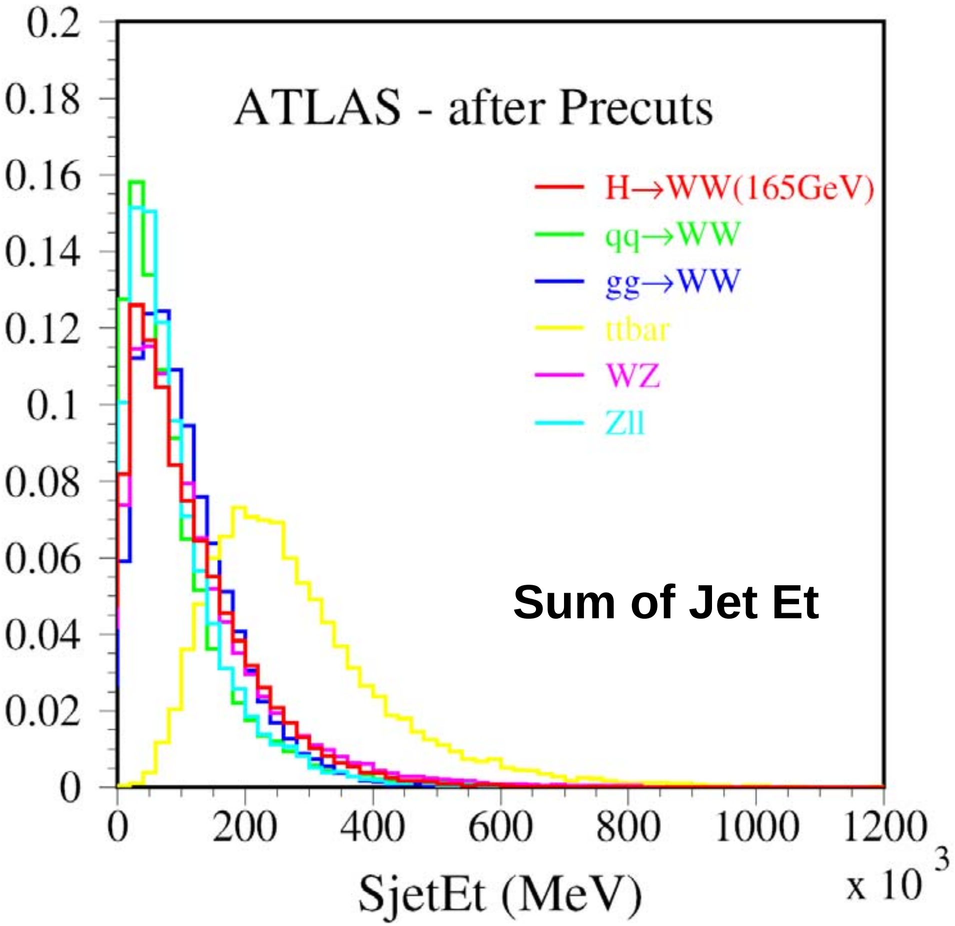


No. of Tracks within a $\Delta R < 0.4$ cone

Some Training Variables

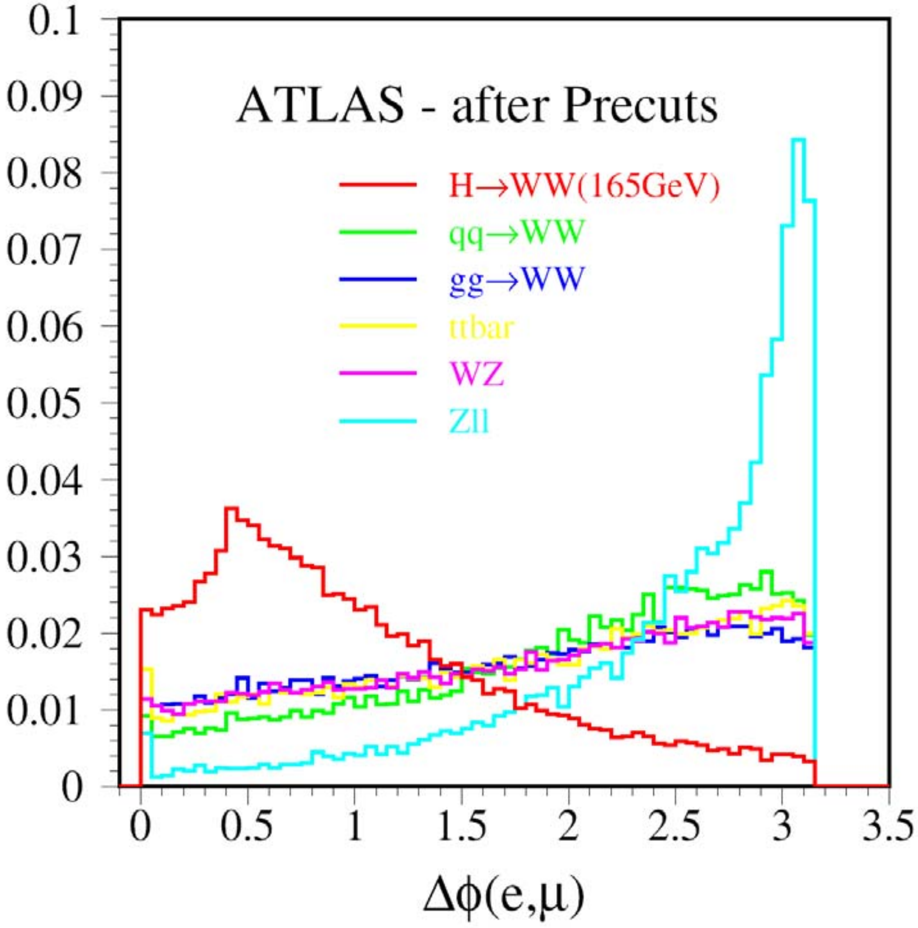


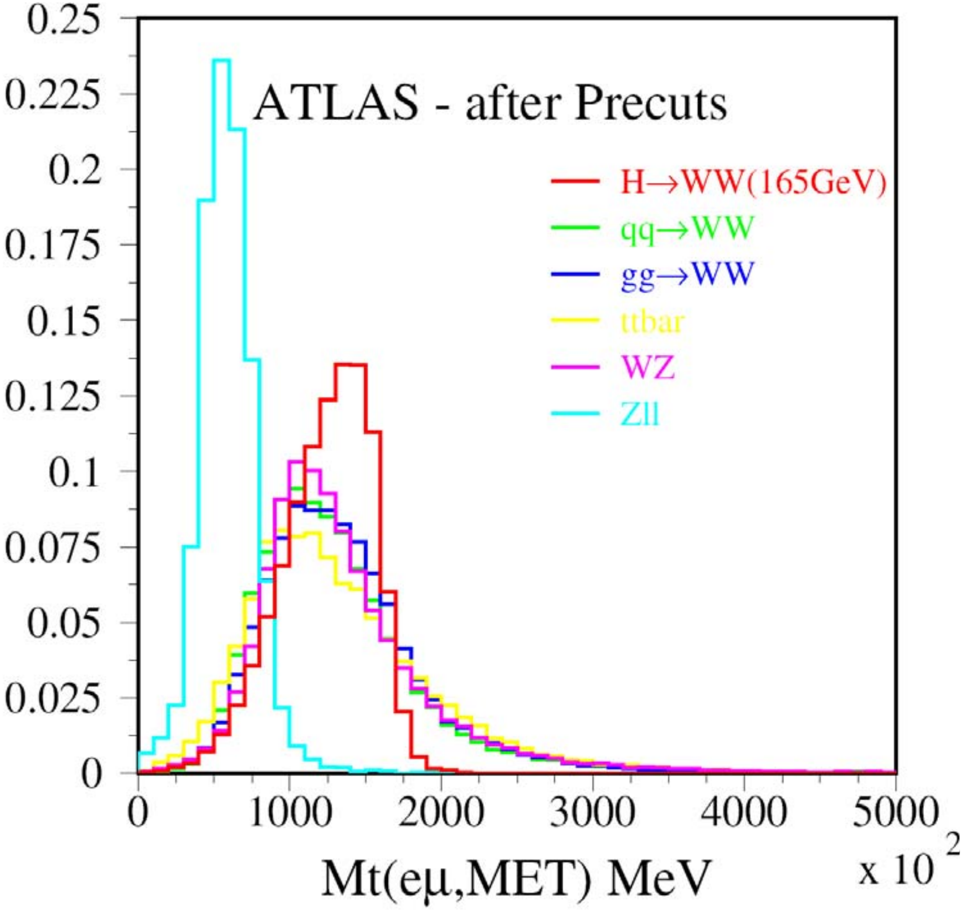
Number of Jets



Sum of Jet Et

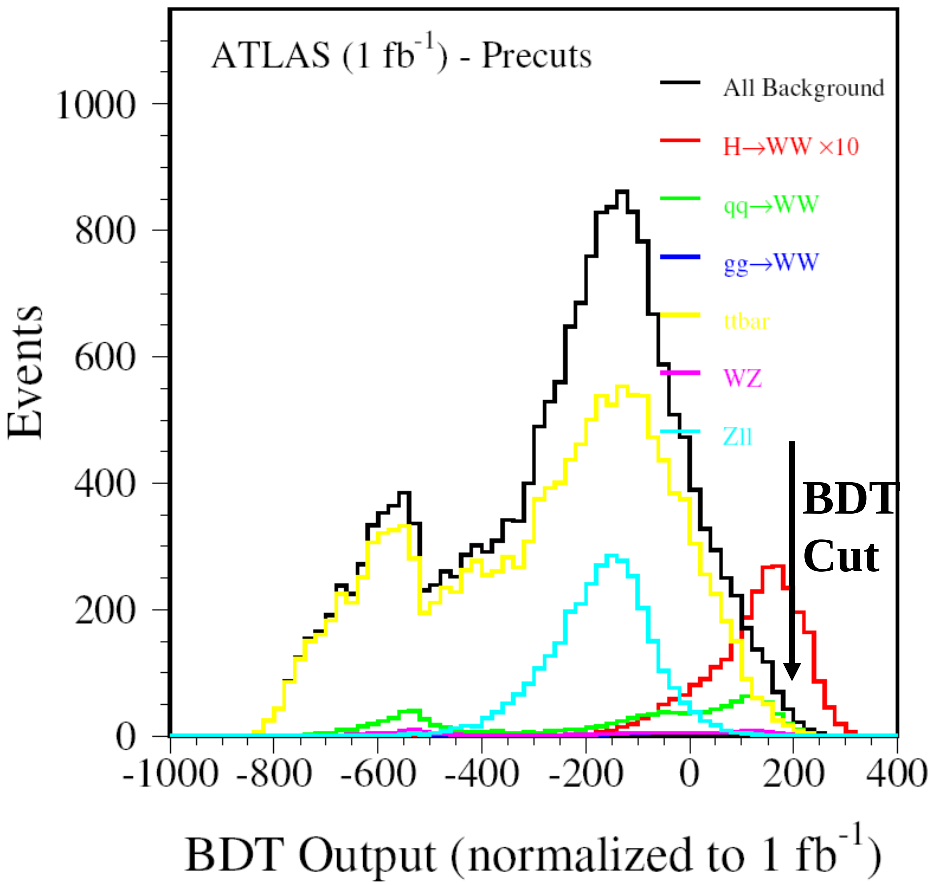
Some Training Variables

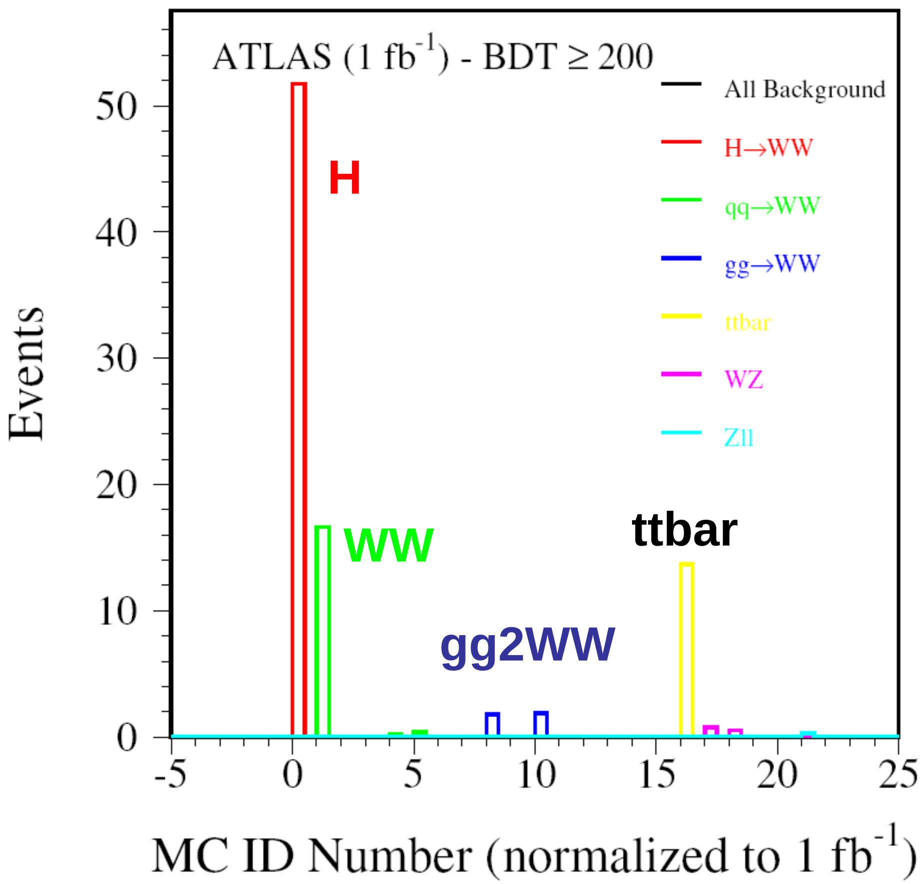




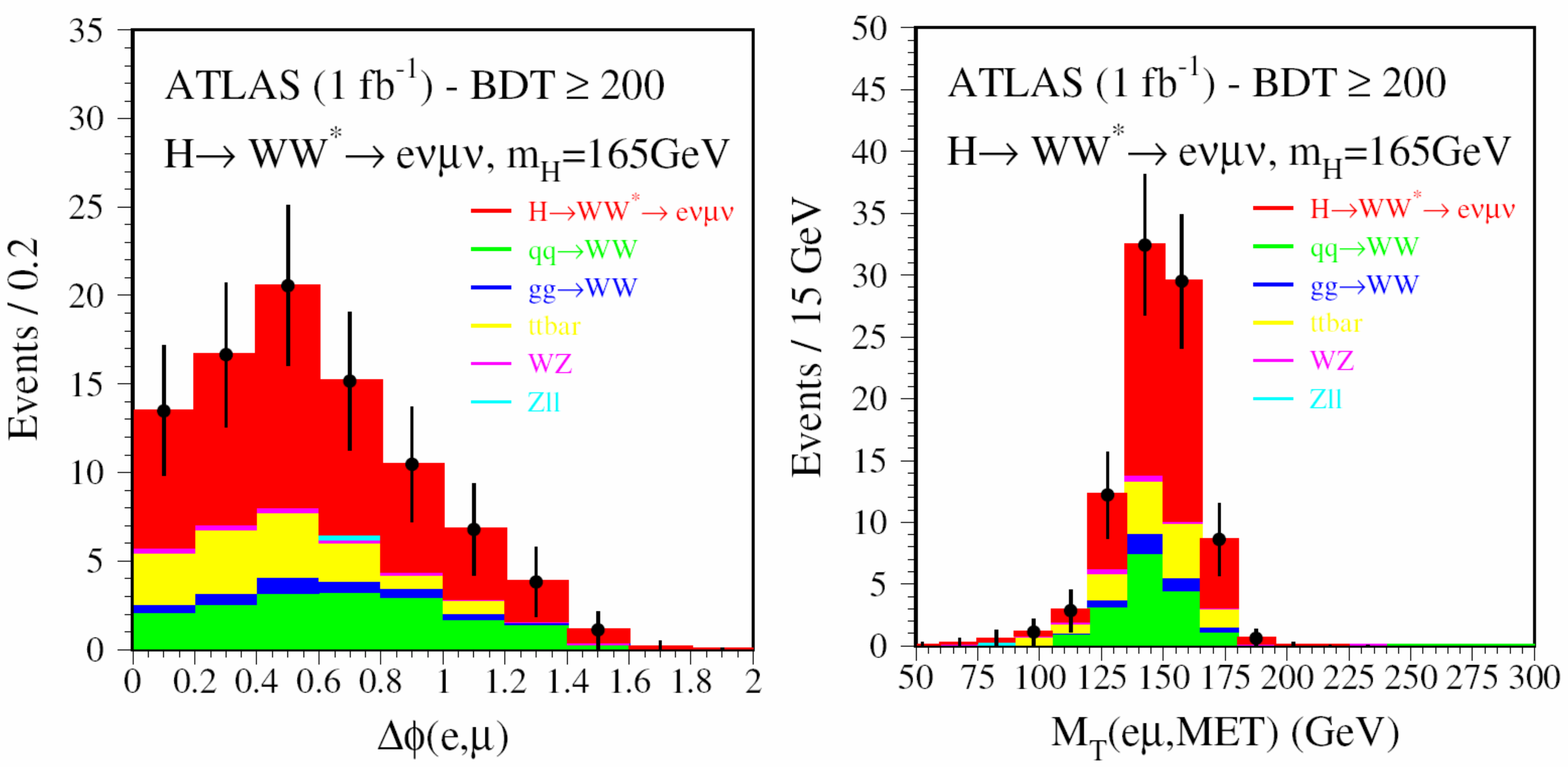
H→WW*→eνμν (165 GeV)

BDT output spectrum and selected signal & background events for $1fb^{-1}$

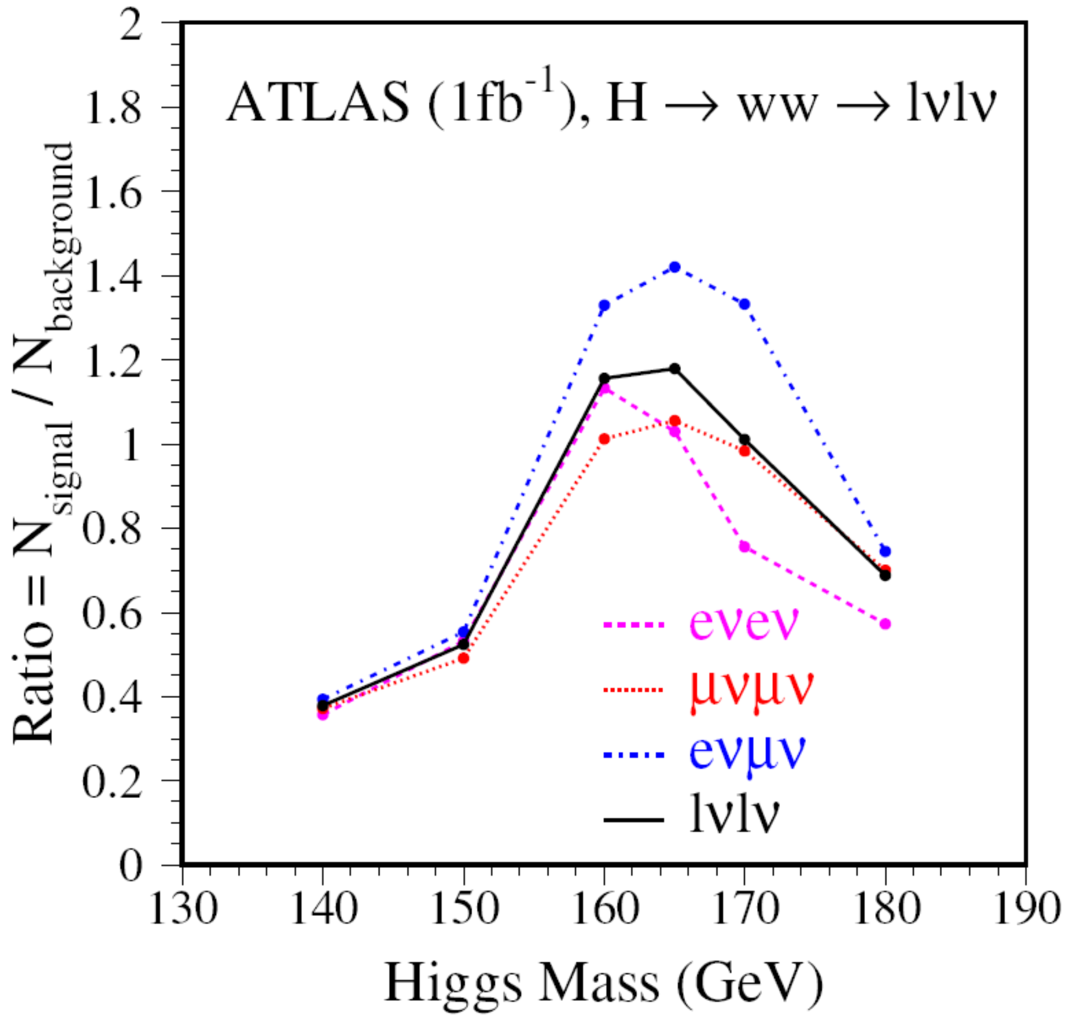




After BDT Selection (H→WW*→eνμν)



S/B Ratio of H → WW* → lνlν



1 fb^{-1}	eνeν	μνμν	eνμν
BG140	36.0	83.9	94.5
H140	12.8	31.2	37.2
BG150	26.6	85.6	81.0
H150	14.1	42.1	44.9
BG160	15.5	39.7	35.2
H160	17.5	40.2	46.8
BG165	22.2	43.5	36.4
H165	22.9	45.9	51.8
BG170	37.2	36.1	32.1
H170	28.2	35.5	42.8
BG180	34.6	36.8	59.4
H180	19.8	25.8	44.3

Discovery Confidence Level Calculation

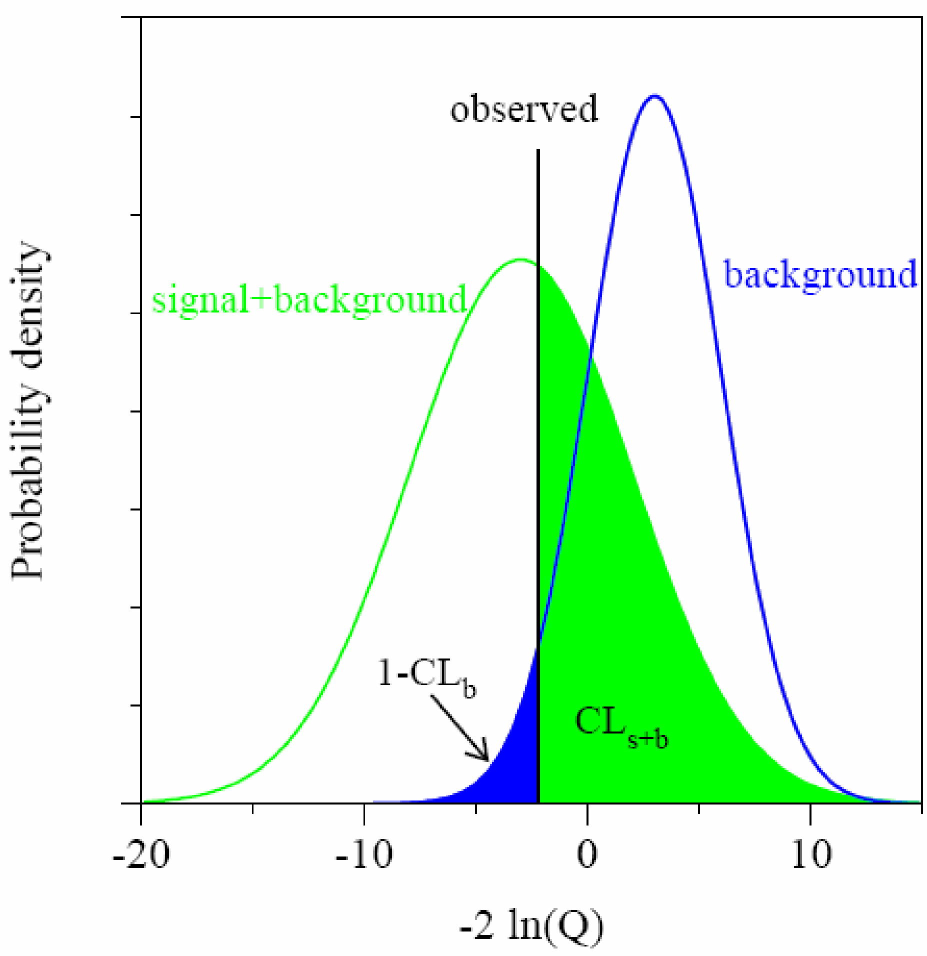
- ➔ Log-likelihood ratio test-statistics by using BDT bins and 3 Higgs decay channels

$$Q = \frac{L(s+b)}{L(b)}$$

- ➔ MC experiments are based on Poisson statistics

- ➔ CL_b represents C.L. to exclude “background only” hypothesis

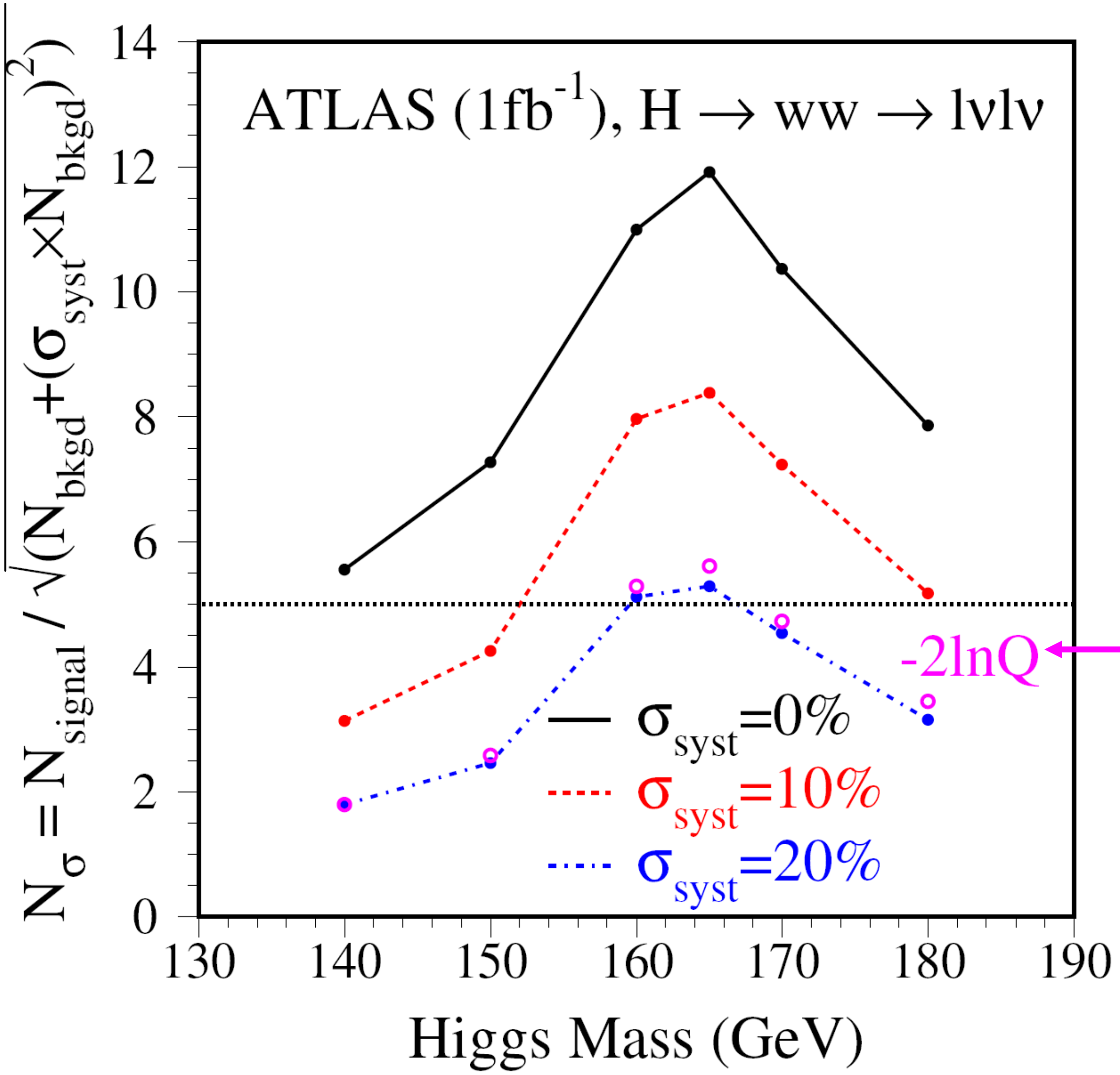
(used for LEP Higgs Search)



Results ($H \rightarrow WW^* \rightarrow l\nu l\nu$, for $1fb^{-1}$)

M_{Higgs} (GeV)	$Eff_s = Eff_{pre} \times Eff_{bdt}$	N_s	N_{bg}	N_σ (stat)	N_σ(10 % syst)	N_σ(20 % syst)	$N_{\sigma 20}$ (-2lnQ)
140	9.7%	81.2	214.4	5.5	3.1	1.8	1.8
150	9.9%	101.1	193.2	7.3	4.2	2.5	2.6
160	9.0%	104.6	90.4	11.0	8.0	5.1	5.3
165	10.3%	120.5	102.2	11.9	8.4	5.3	5.6
170	9.6%	106.5	105.4	10.4	7.2	4.5	4.7
180	9.2%	89.9	130.8	7.9	5.2	3.1	3.4

ATLAS Sensitivity of H → WW* → lνlν



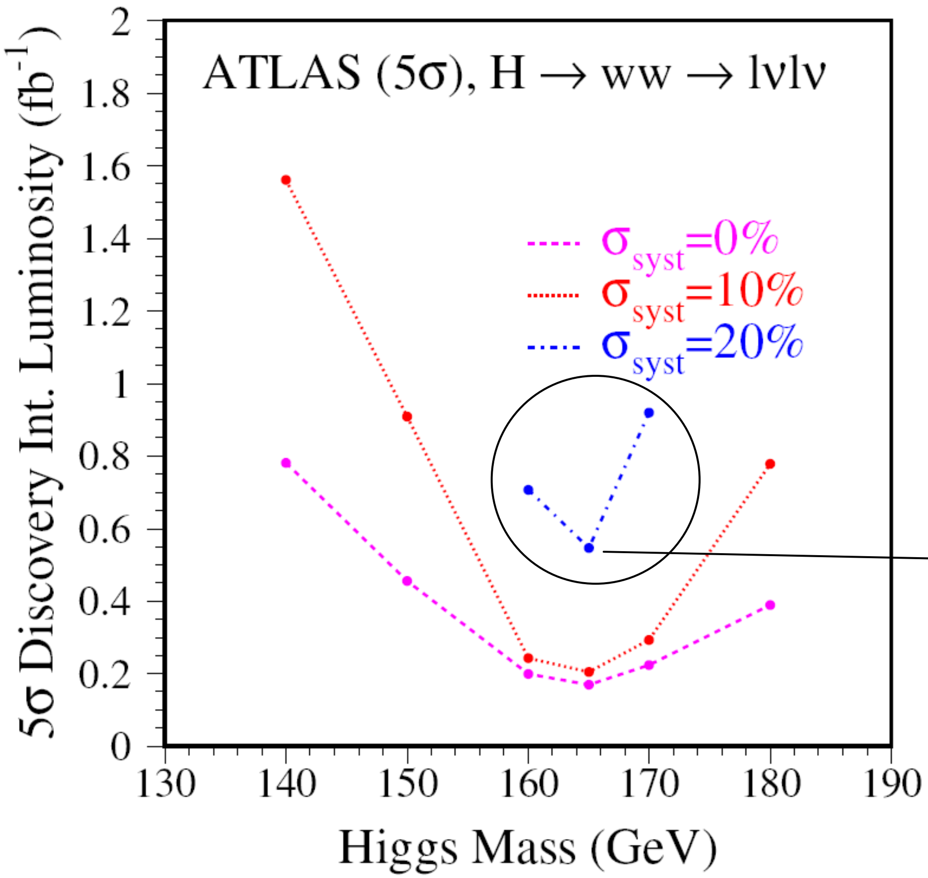
$$Q = \frac{L(s+b)}{L(b)}$$

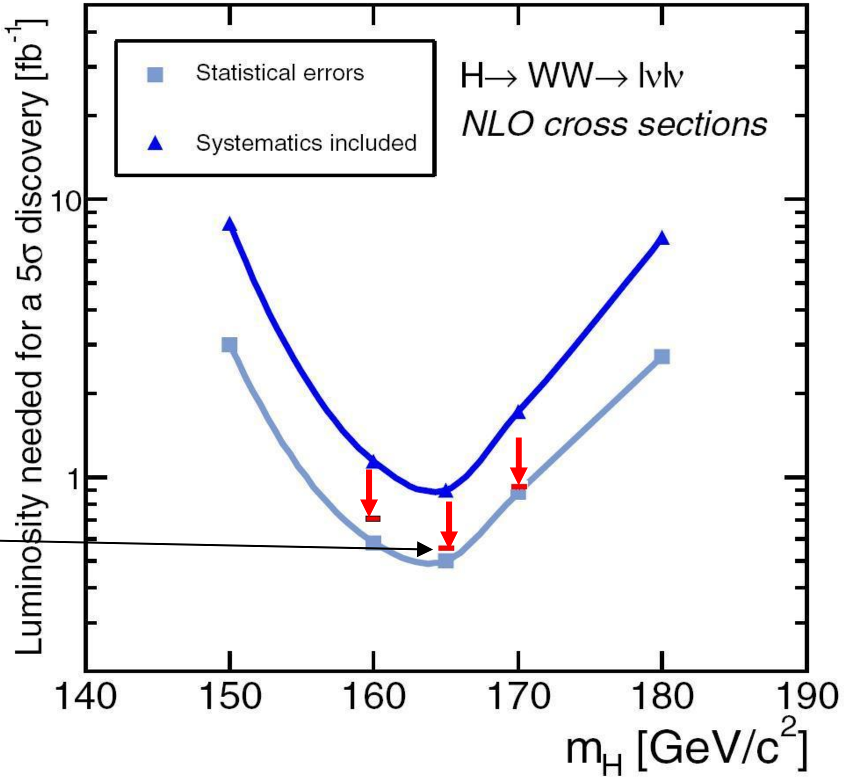
Log-likelihood Ratio with 20% syst. error

Required Int. Lumi for 5σ Discovery

BDT Analysis, H → WW* → lνlν (l=e,μ)

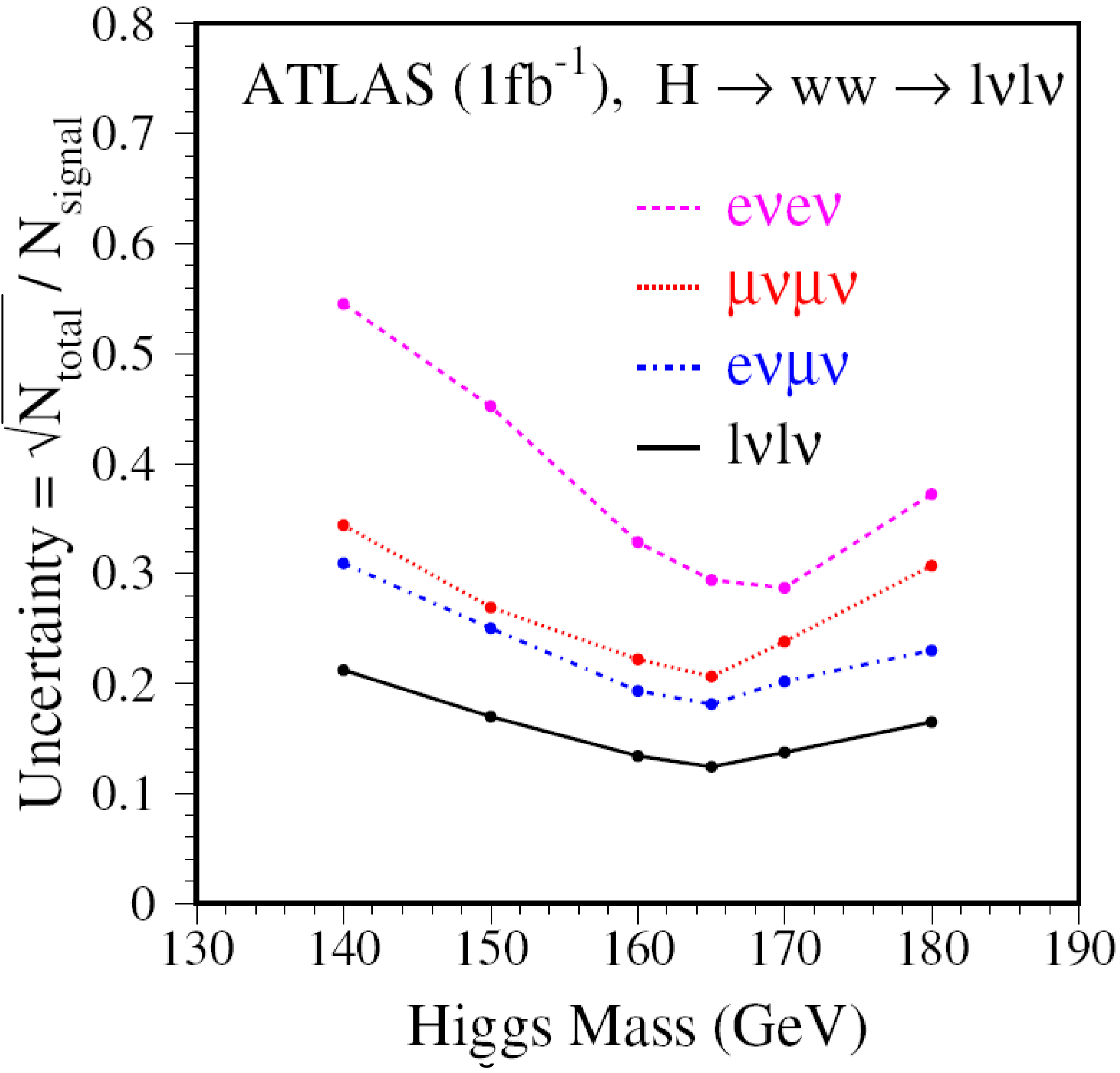
CMS Phys. TDR 2006





σ_{syst} = 19%, 16%, 11% for 1, 2, 10 fb^{-1}

Cross Section Uncertainty of $H \rightarrow WW^* \rightarrow l\nu l\nu$



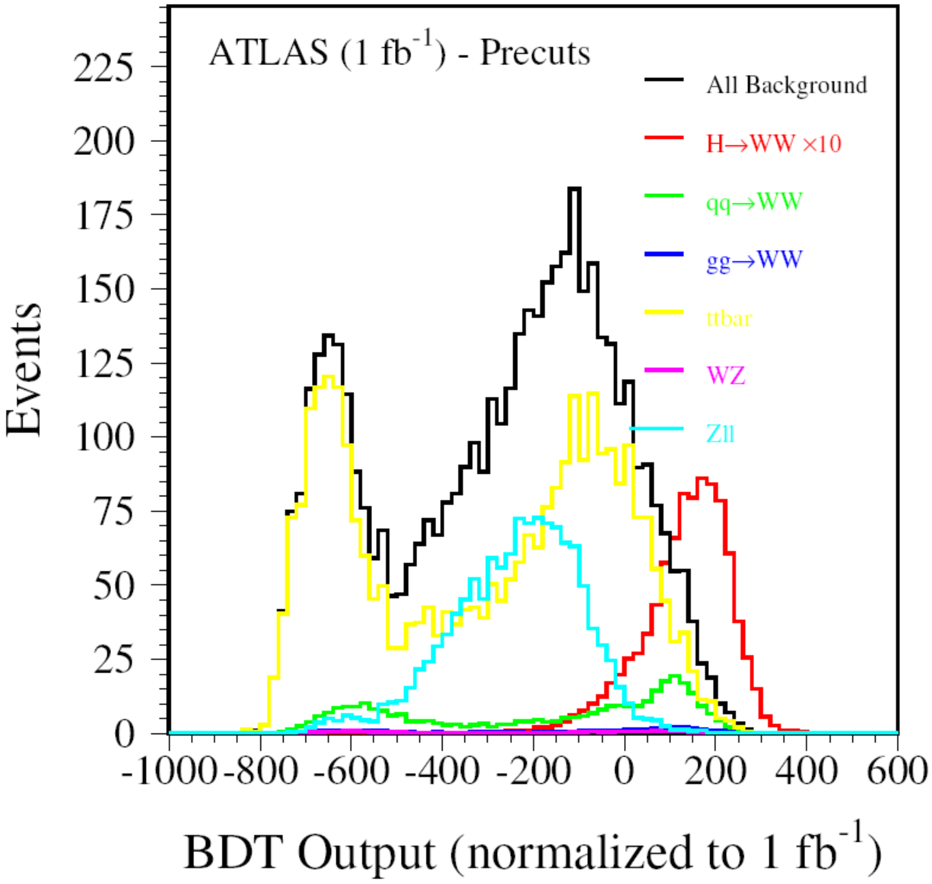
Summary and Outlook

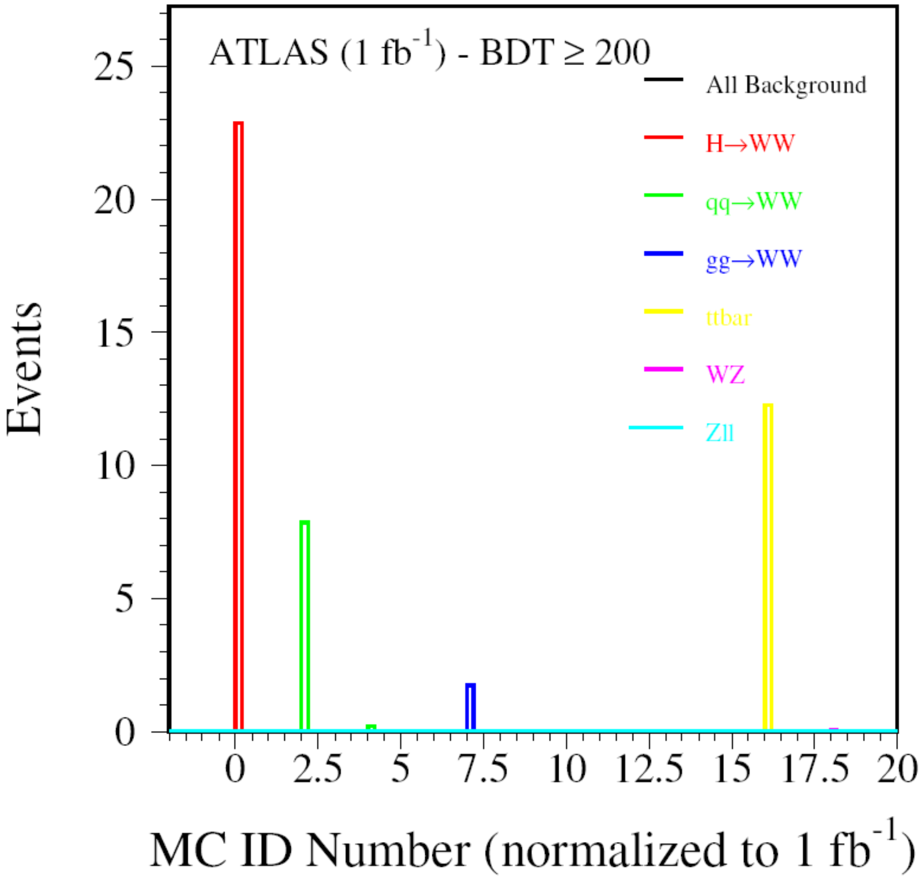
- H → WW* → lνlν analysis based on BDT has significant impact on early discovery potential.
- For 140-180 GeV SM Higgs 5σ discovery only needs a few fb^{-1} integrated luminosity.
- Major backgrounds for H → WW searches come from WW(50-60%) and ttbar(30-40%).

➔BDT is anticipated to have wide application in LHC physics analysis, especially for particle searches.

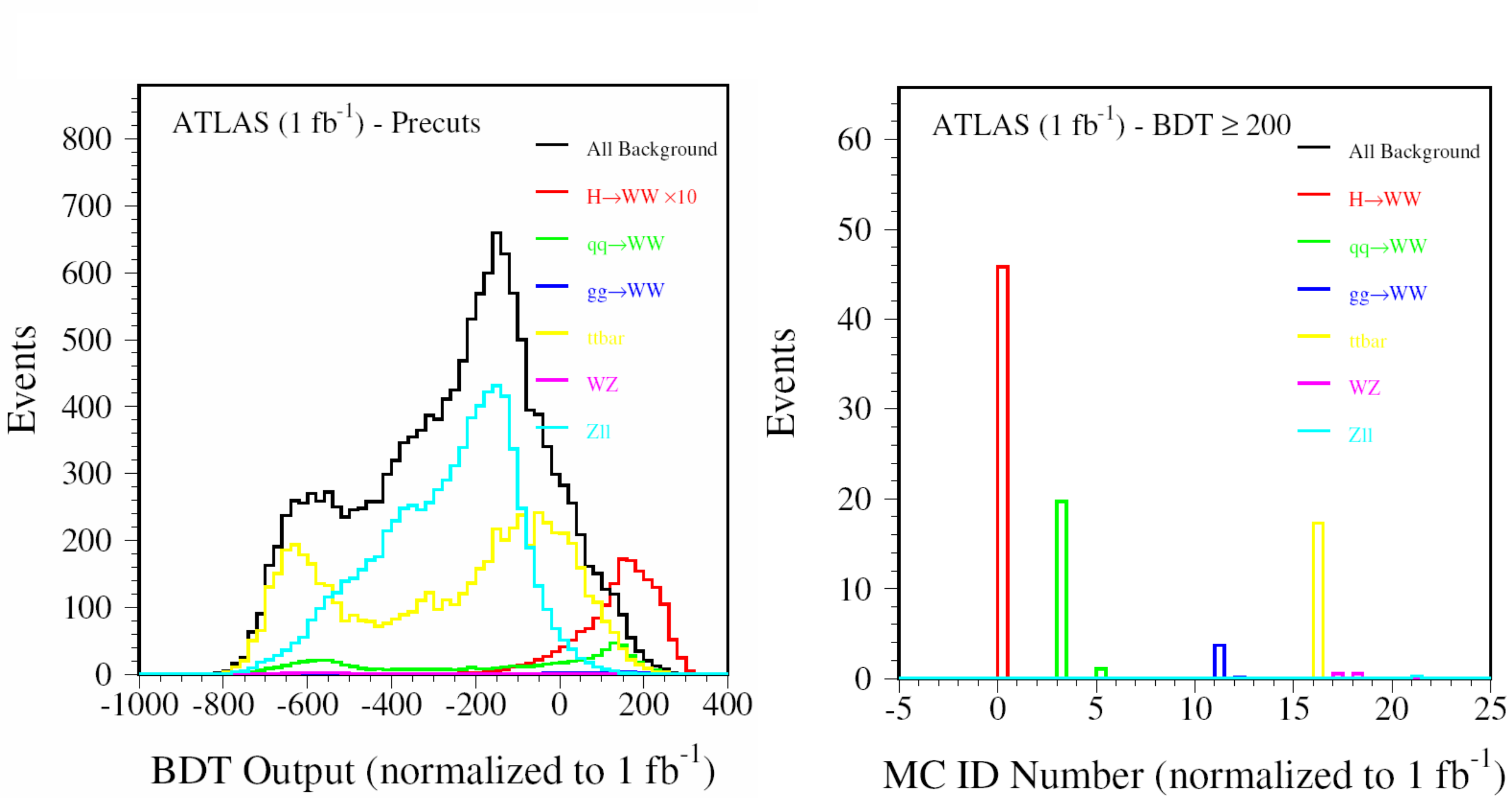
Backup Slides

H→WW*→eνeν (165 GeV)

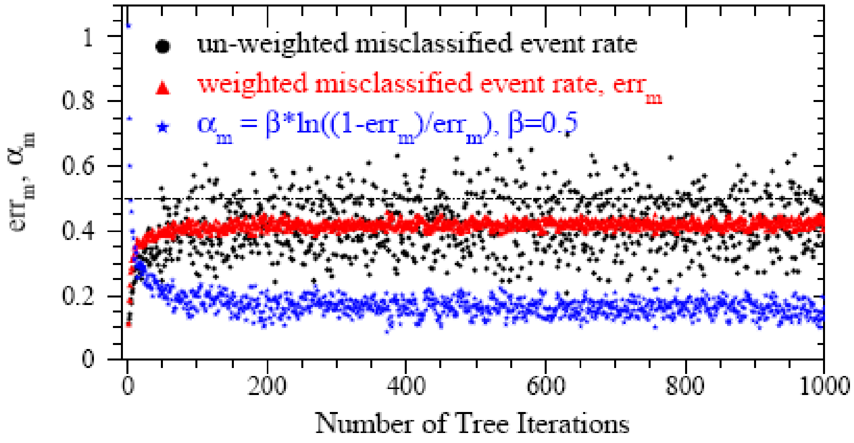




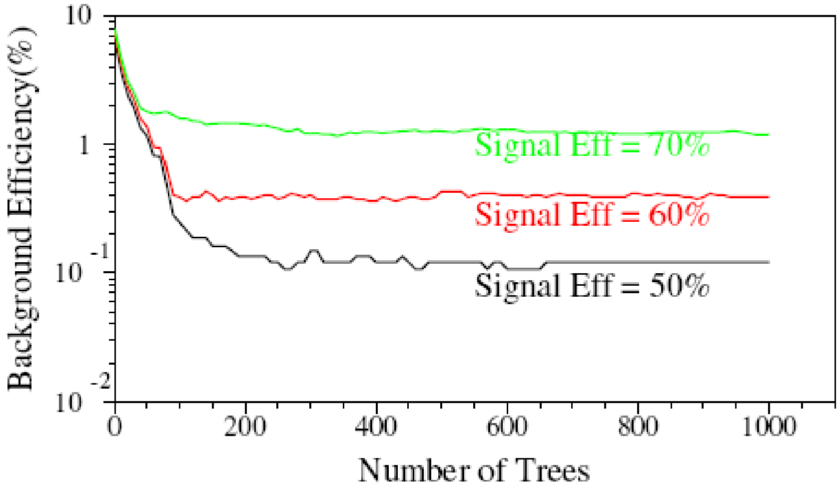
H→WW*→μνμν (165 GeV)



Weak → Powerful Classifier



➔The advantage of using boosted decision trees is that it combines many decision trees, "weak" classifiers, to make a powerful classifier. The performance of boosted decision trees is stable after a few hundred tree iterations.



➔ Boosted decision trees focus on the misclassified events which usually have high weights after hundreds of tree iterations. An individual tree has a very weak discriminating power; the weighted misclassified event rate err_m is about 0.4-0.45.

Ref1: H.J.Yang, B.P. Roe, J. Zhu, "*Studies of Boosted Decision Trees for MiniBooNE Particle Identification*", physics/0508045, Nucl. Instum. & Meth. A 555(2005) 370-385.

Ref2: H.J. Yang, B. P. Roe, J. Zhu, " *Studies of Stability and Robustness for Artificial Neural Networks and Boosted Decision Trees* ", physics/0610276, Nucl. Instrum. & Meth. A574 (2007) 342-349.

BDT Training with Event Reweighting

- In the original BDT training program, all training events are set to have same weights in the beginning (the first tree). It works fine if all MC processes are produced based on their production rates.
- Our MCs are produced separately, the event weights vary from various backgrounds. e.g. 1 fb^{-1} ,wt (ww)=0.07, wt (ttbar)=0.72
- If we treat all training events with different weights equally using “standard” training algorithm, ANN/BDT tend to pay more attention to events with lower weights (high stat.) and introduce training prejudice.

- Ref: http://arxiv.org/abs/0708.3635, Hai-Jun Yang, Tiesheng Dai, Alan Wilson, Zhengguo Zhao, Bing Zhou, ”A Multivariate Training Technique with Event Reweighting”