The Quest for the Higgs

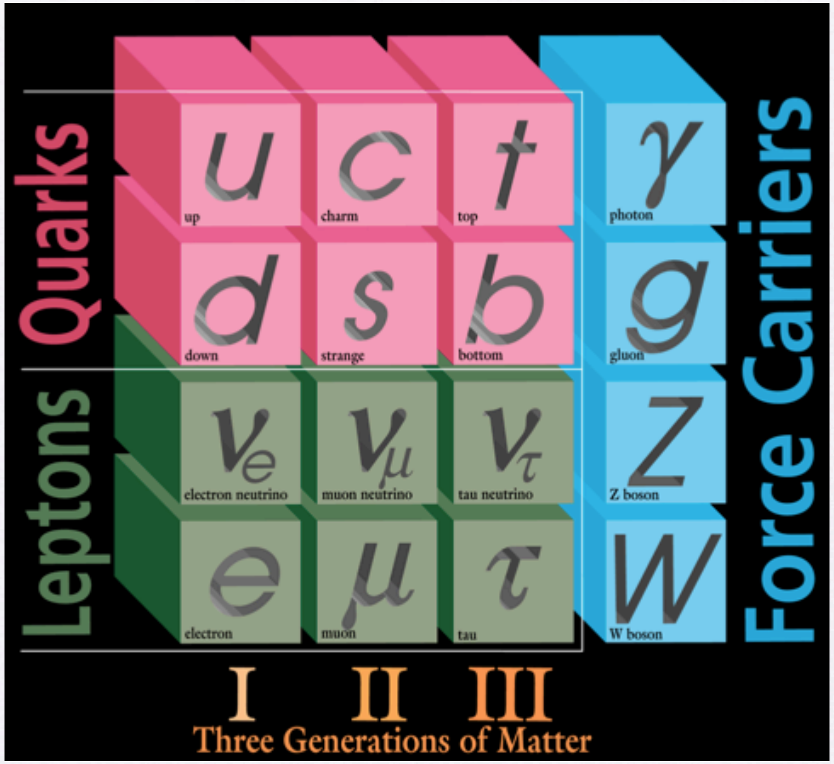
Sabine Lammers
Columbia University

Outline

- Introduction to SM Higgs Physics
- Experimental Apparatus
- Higgs Searches at the Tevatron
- WH and Combination
- Higgs Searches at the LHC
- Vector Boson Fusion (VBF) Higgs
- VBF at the Tevatron

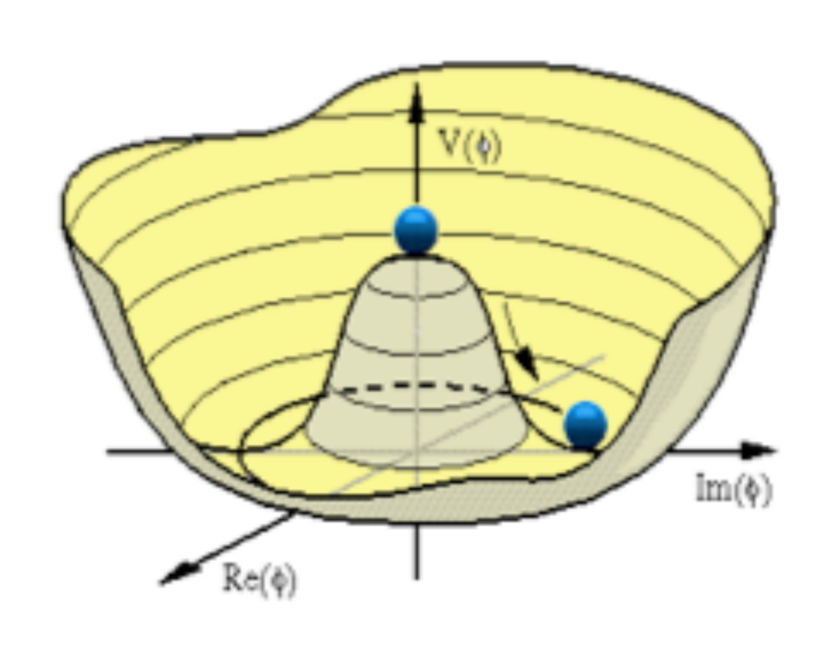
Standard Model

- The SM is a Quantum Field Theory: fusion of Special Relativity and Quantum Mechanics
- There are 3 main ingredients:
  - Forces
    - Electromagnetism(γ), Weak($W^{\pm}$,Z), Strong(g)
  - Matter
    - 6 quarks, 6 leptons in 3 generations
  - Spontaneous Symmetry Breaking
    - Higgs Mechanism



Higgs Phenomenology

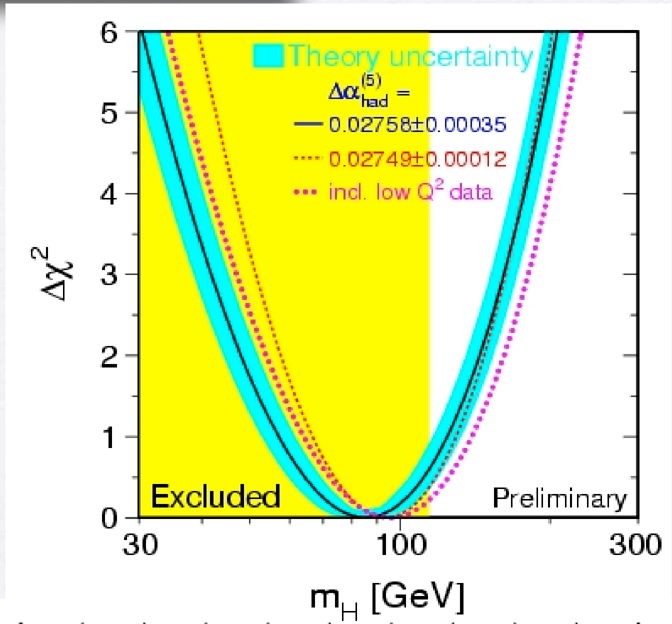
- Higgs field is a complex scalar field introduced to break electroweak symmetry and to introduce mass terms in the Standard Model (SM) Lagrangian
- Neutral, spin 0 Higgs Boson must be found to complete SM picture
- Higgs mass is a parameter of the theory

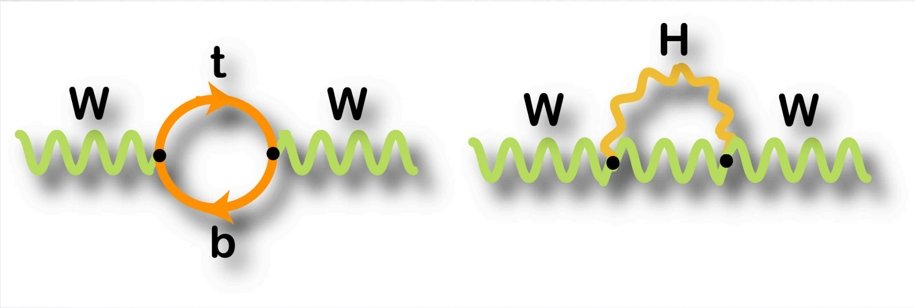


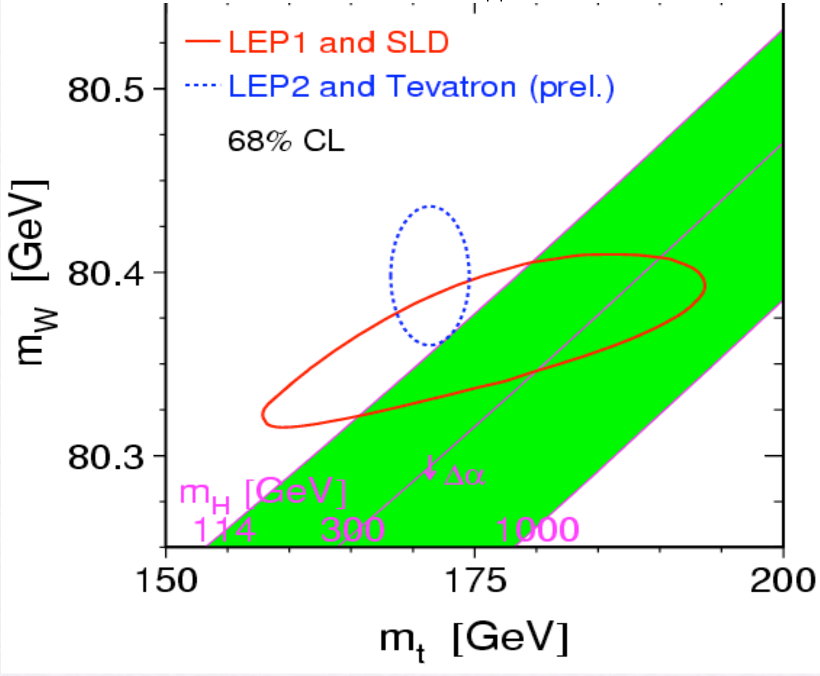
Indirect Constraints on Higgs mass

- Precision Fit of electroweak precision data, including top quark and W masses
- best fit Higgs mass = 76 + 33 - 24 GeV

➡ **light Higgs is preferred**

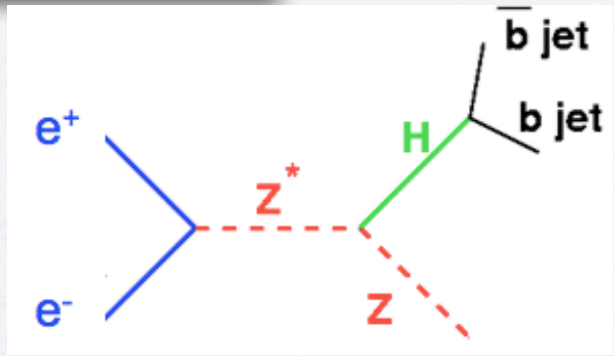






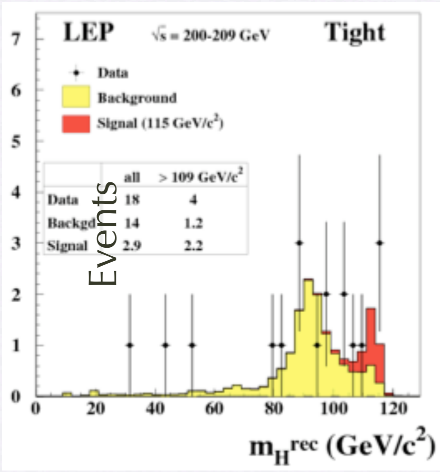
LEP Direct Searches

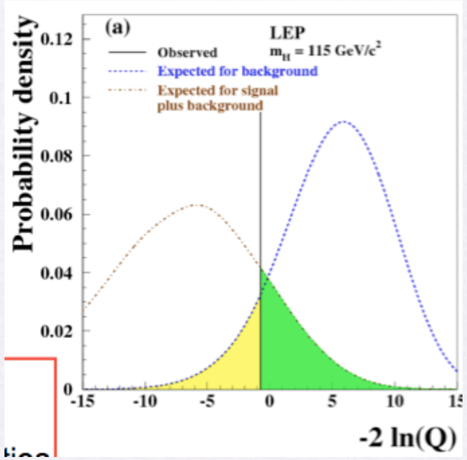
- LEP **direct search** result : combination from four experiments found hint of a signal at m_H ~118 GeV, but could be fluctuation
- LEP technique for deriving limits
  - Ratio of Poisson Likelihoods
  - Comparison of signal+background and background only hypotheses to data
  - Probability densities determined using toy MC experiments whose event makeup vary according to statistical and systematic uncertainties



$$\sqrt{s} - M_Z = 206.7 - 91.2 = 115.5\text{GeV}$$

$m_H \geq 114.4$ GeV @ 95% CL

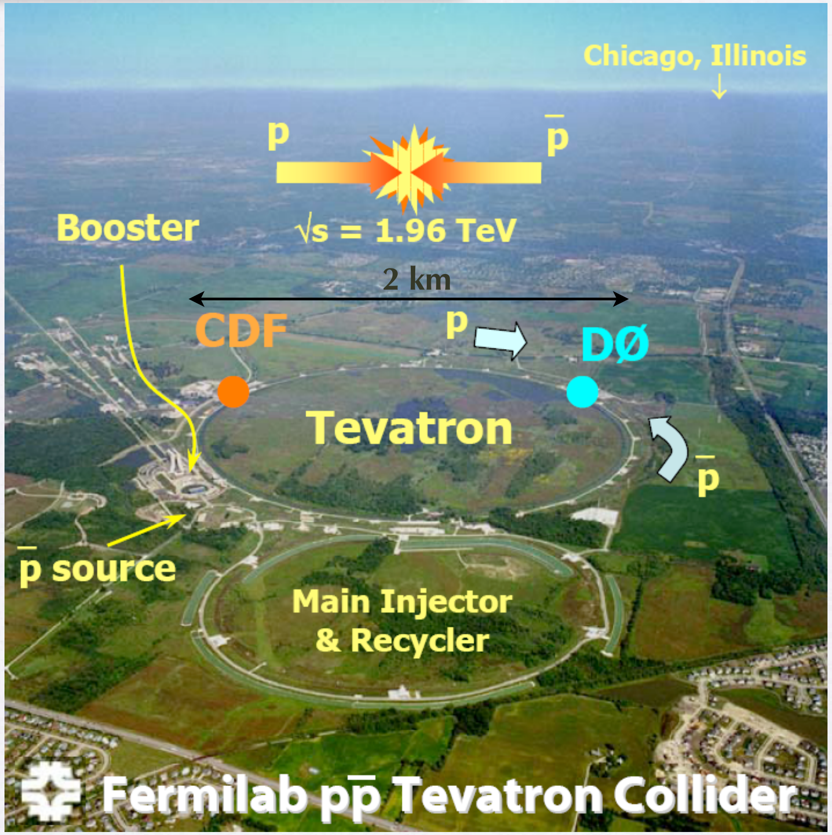


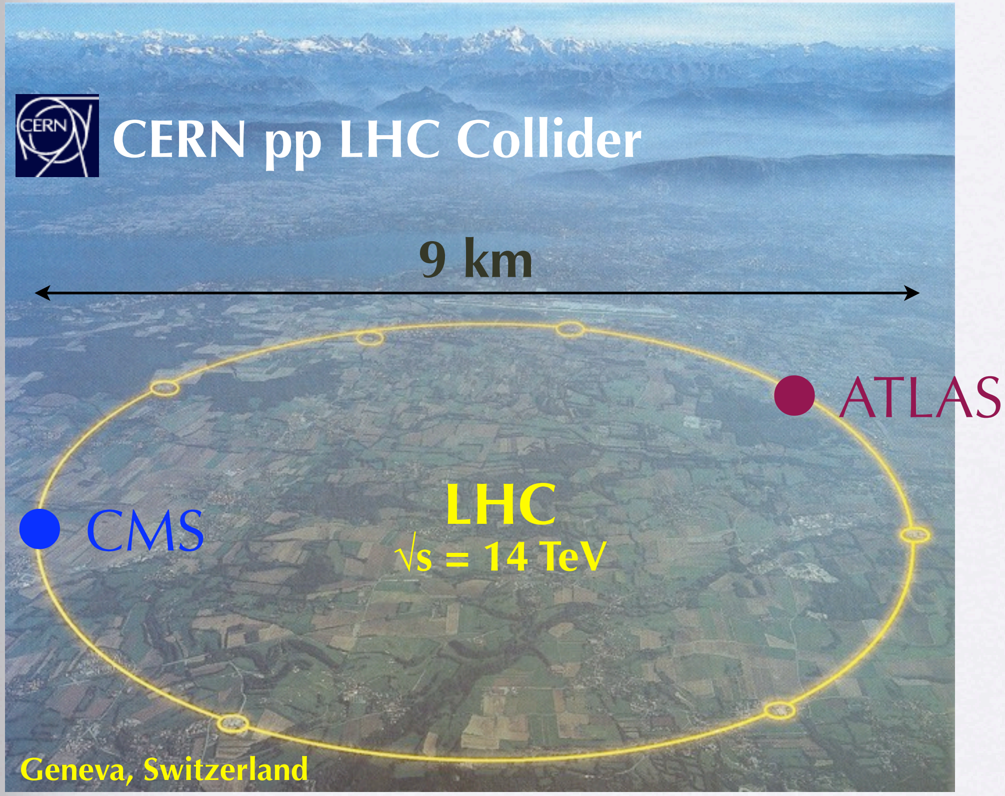


Experiments

Tevatron and LHC

- Tevatron - energy frontier accelerator for nearly 2 decades
  - $p\bar{p}$ collisions at $\sqrt{s}$ = 1.96 TeV



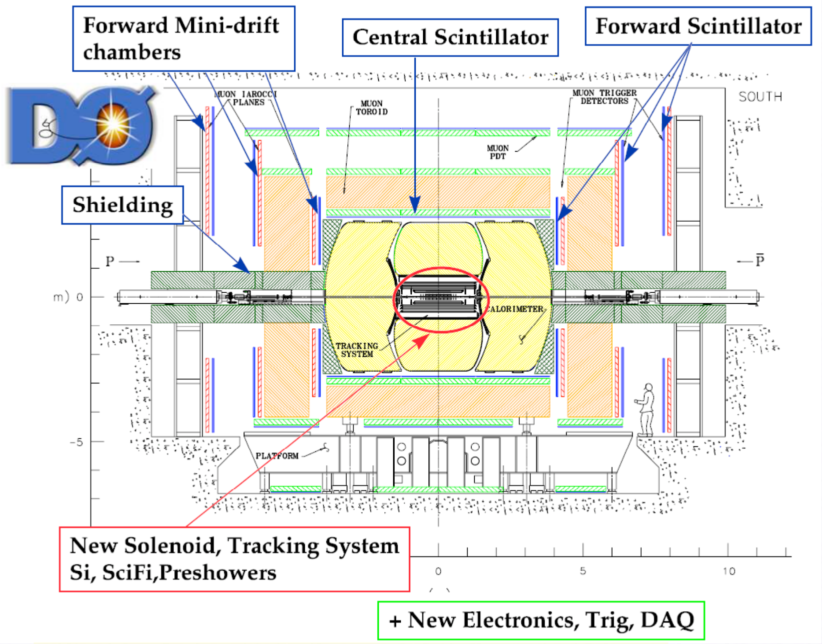


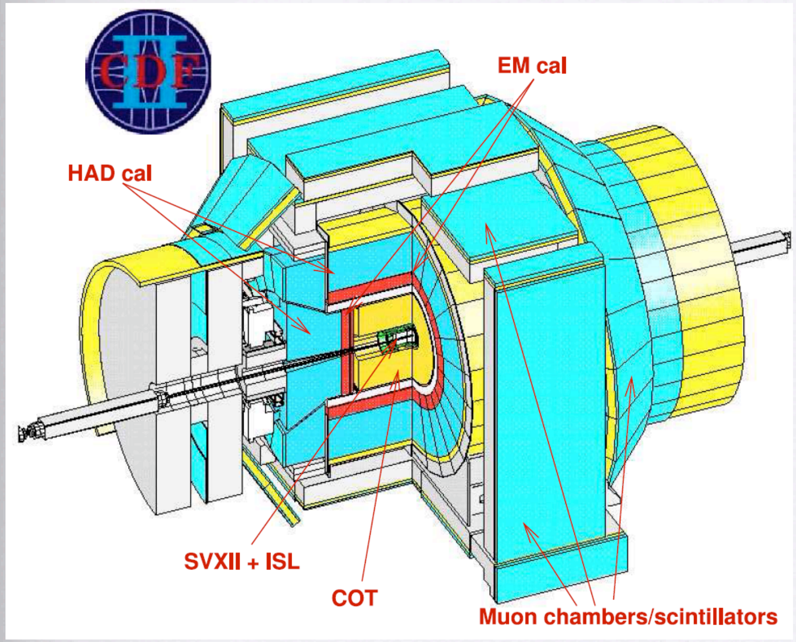
- LHC - will probe Terascale phenomena as energy frontier machine for the next decades
  - pp collisions at $\sqrt{s}$ = 14 TeV

Tevatron Detectors: DØ and CDF

- DØ - Liquid Argon and Uranium Scintillator sampling calorimeter
- Silicon Microstrip and Fiber tracking
- Good muon coverage $|\eta| < 2$
- 2T magnetic field

$\eta = -\ln(\tan\theta/2)$



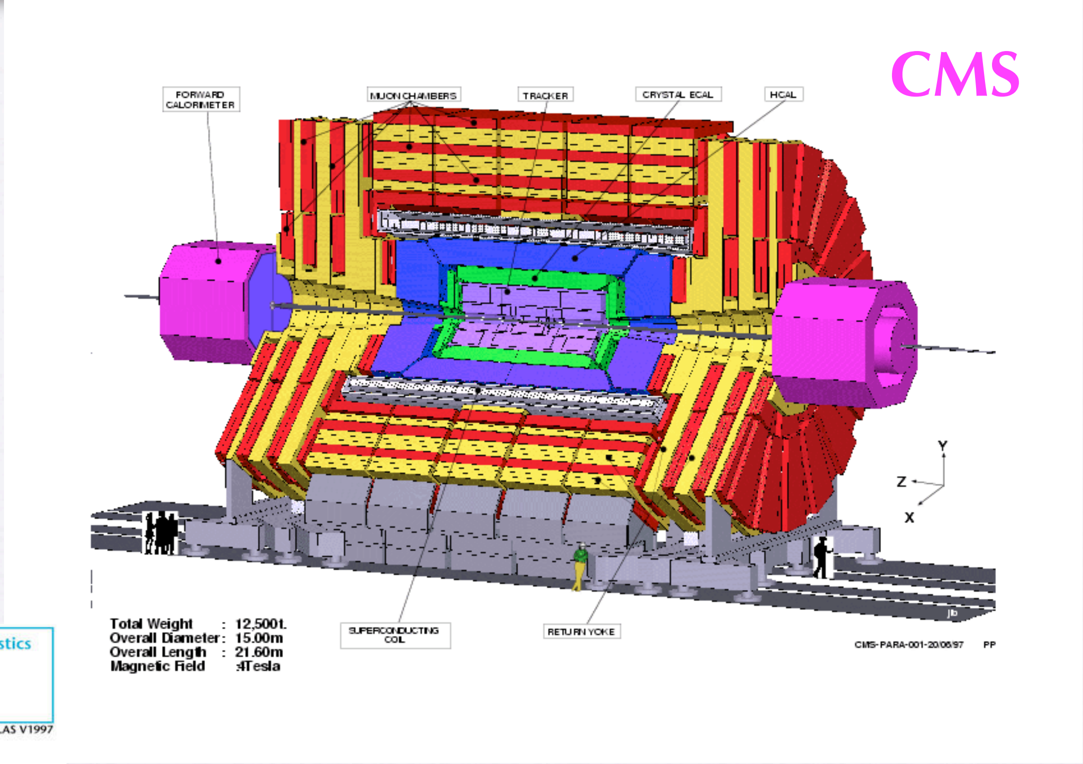


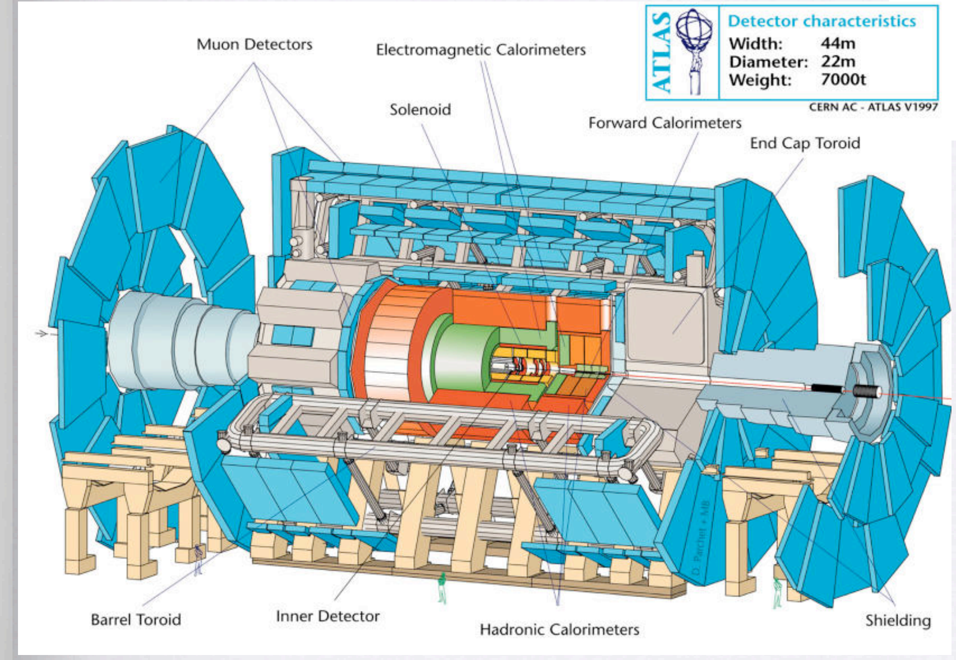
- CDF - Lead Scintillator sampling calorimeter
- Large tracking volume + silicon
- Muon coverage $|\eta| < 1.5$
- 1.5 T magnetic field

LHC Detectors: CMS and ATLAS

CMS

- Lead Tungstate crystal EM calorimeter
- full silicon tracking



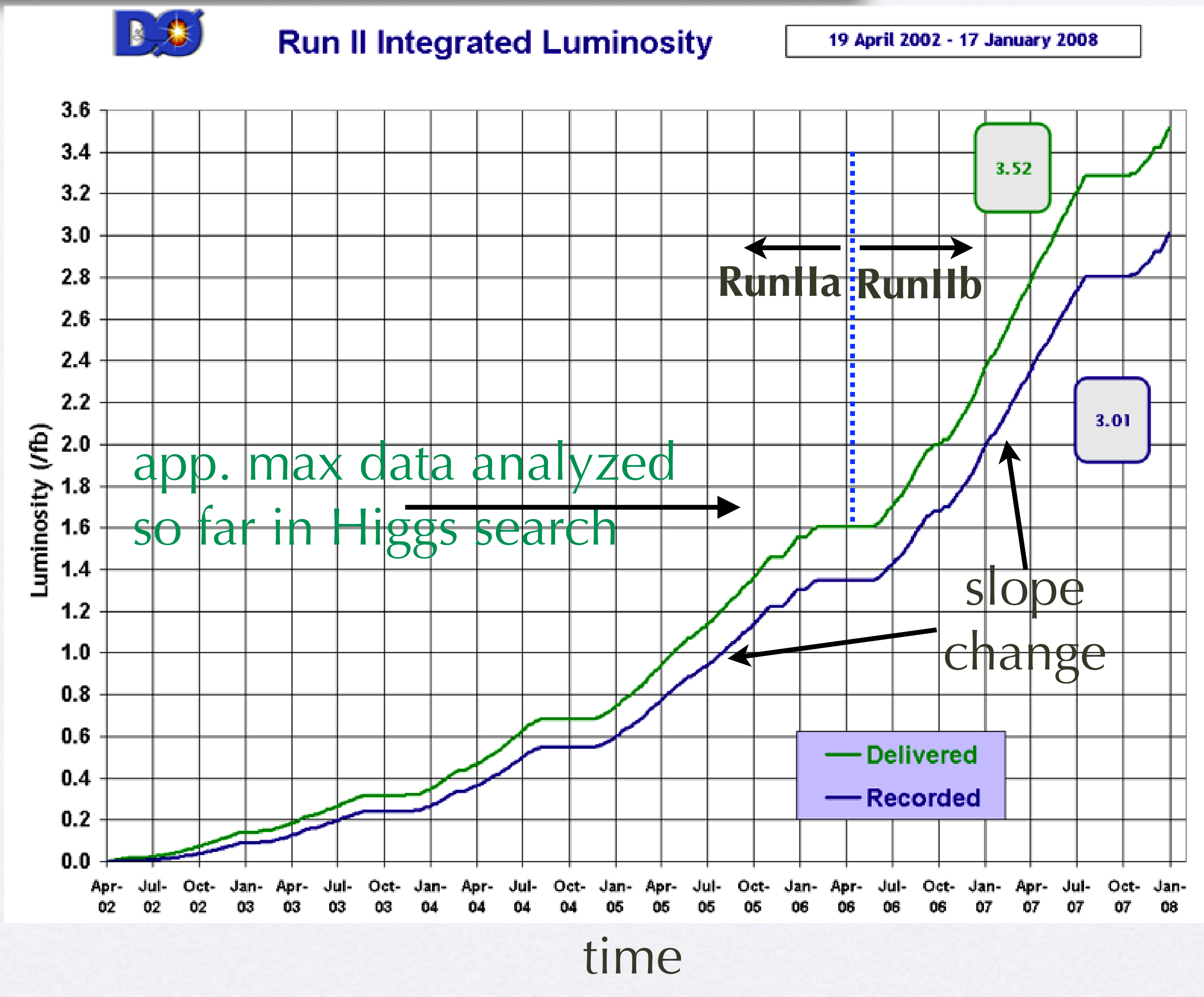


ATLAS

- liquid Argon calorimeter
- muon coverage to $|\eta|<2.5$

Tevatron Performance

- DZero RunIIB upgrades: L1Cal/ L1CalTrack trigger and new silicon layer added to inner tracking detector



time

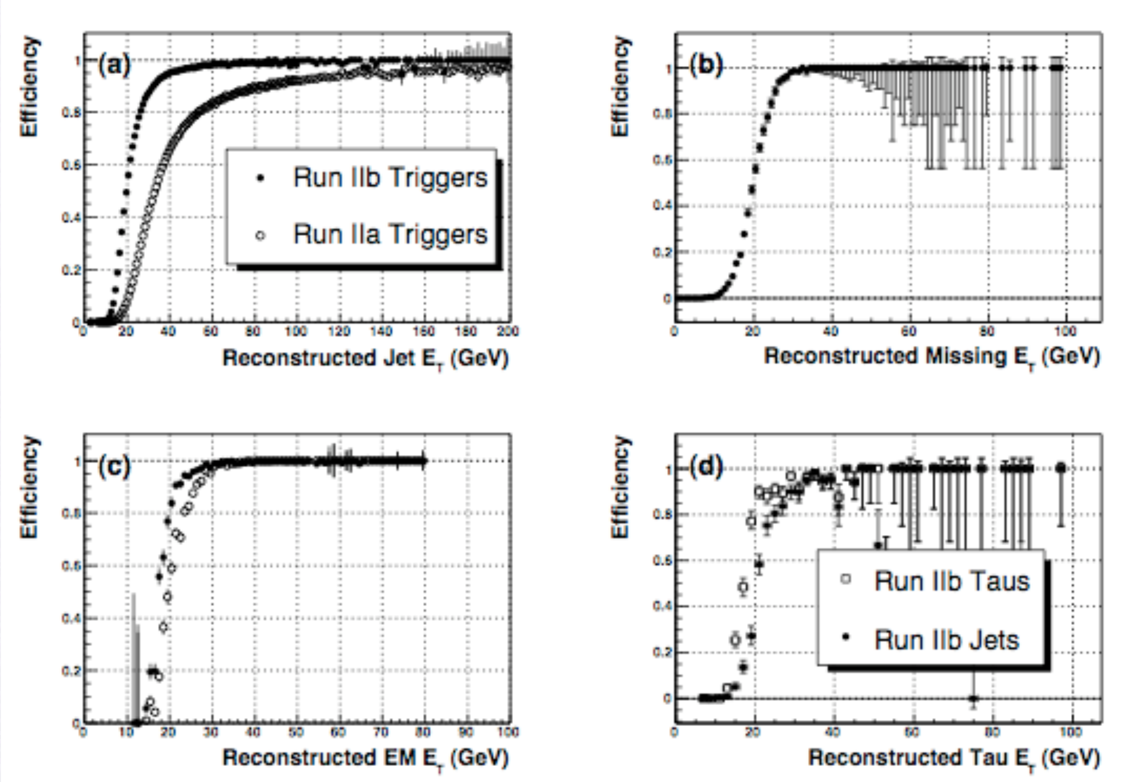
Trigger upgrades ensure high trigger efficiency at high instantaneous lumi
Silicon upgrade provides better b-tagging

L1Cal2b Upgrade

- Upgraded trigger electronics provide better digitization and allows for sophisticated hardware (sliding window) algorithms including clustering at Level 1.

- New features include triggers for jets, taus, isolated electrons, missing E_T, and topological triggers, e.g. acoplanar jets or back-to-back electrons

Improved L1Cal2b algorithms allows us to run at higher instantaneous luminosity with no degradation (enhancement in some cases) in trigger efficiency

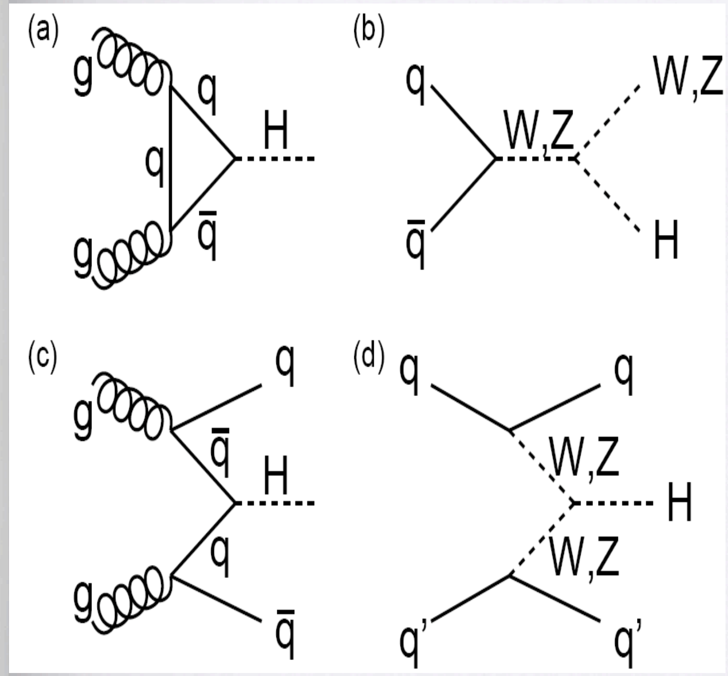


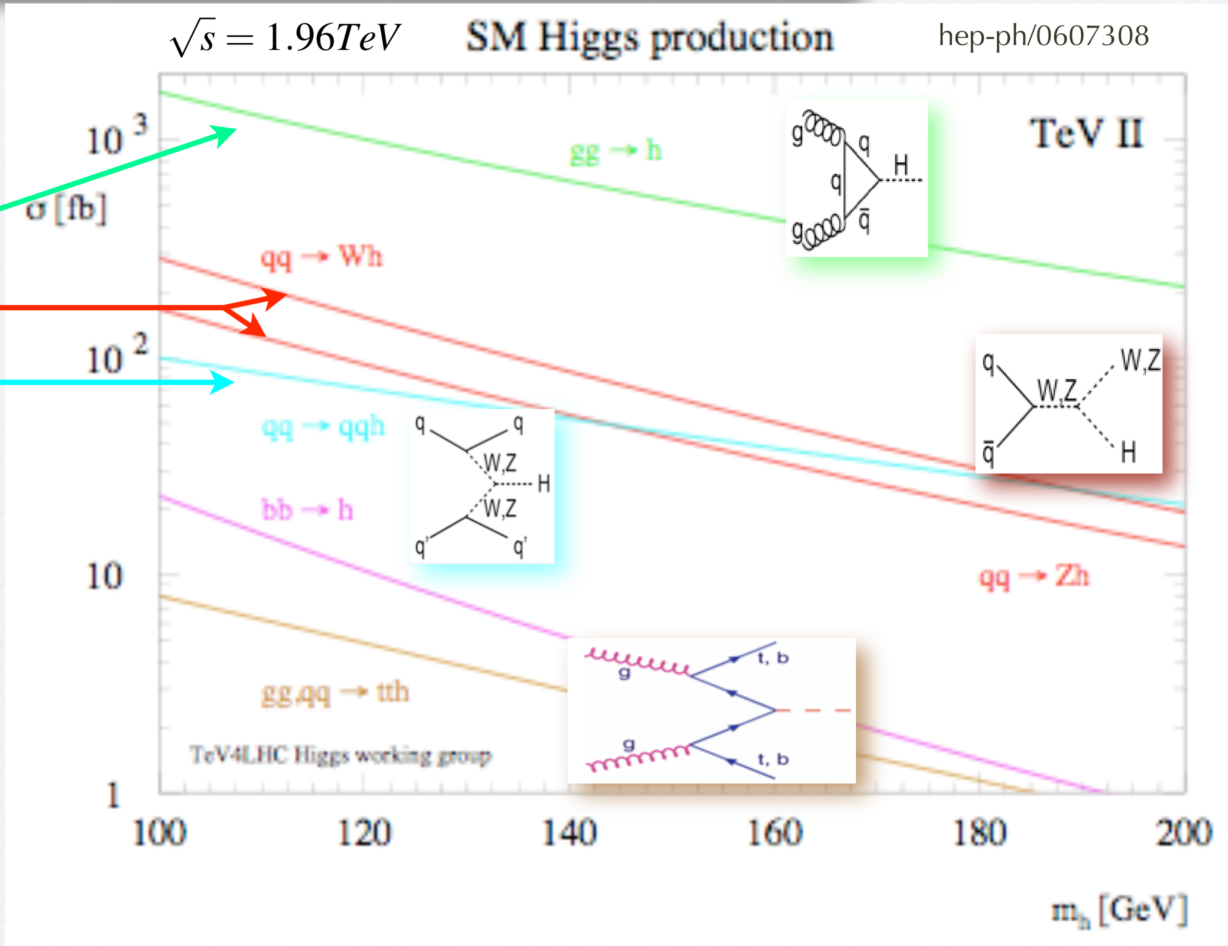
Nucl. Instrum. and Methods, A 584/1, 75-97 (2007)

Higgs Searches at the Tevatron

Higgs Production

Dominant production by (a,c) gluon fusion; also (b)Higgsstrahlung, (d) Vector Boson Fusion

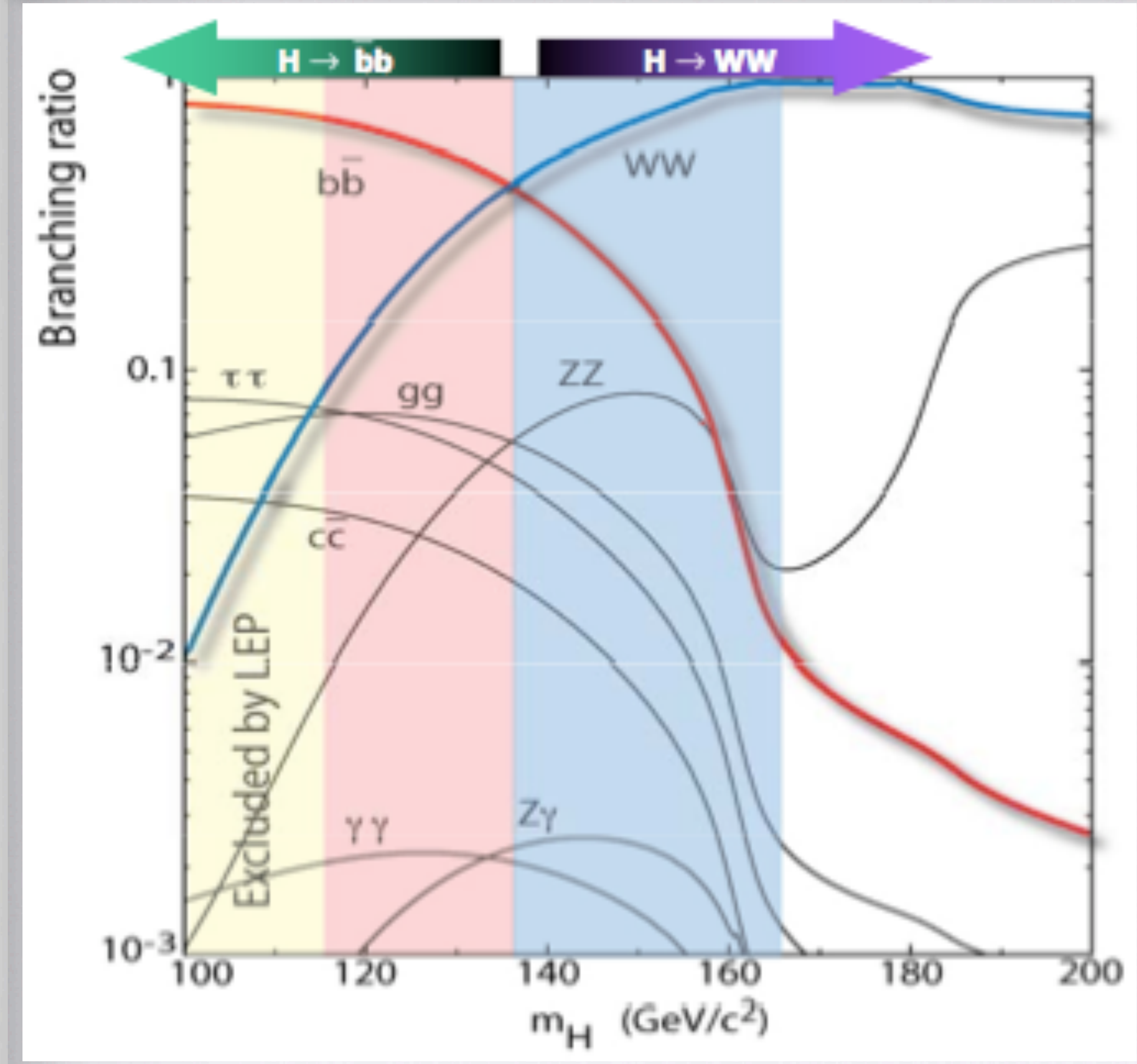




For maximal signal significance:

- Higgsstrahlung or "associated production" searches at low mass
- gluon fusion searches at high mass

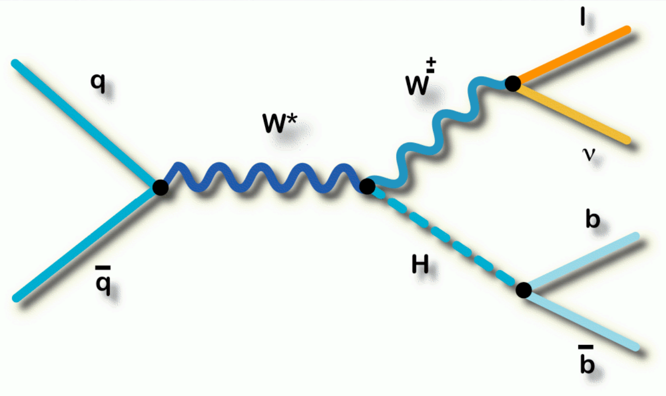
Higgs Decay



- Higgs decays to pairs of fermions or bosons
- Couplings to fermions are proportional to the masses
- Selection depends on available phase space to produce real particles
- Dominant decay
  - b-quark pairs when $m_H < 135$ GeV
  - W pairs when $m_H > 135$ GeV

WH Channel

$$WH \rightarrow l\nu b\bar{b},\ l = e, \mu$$

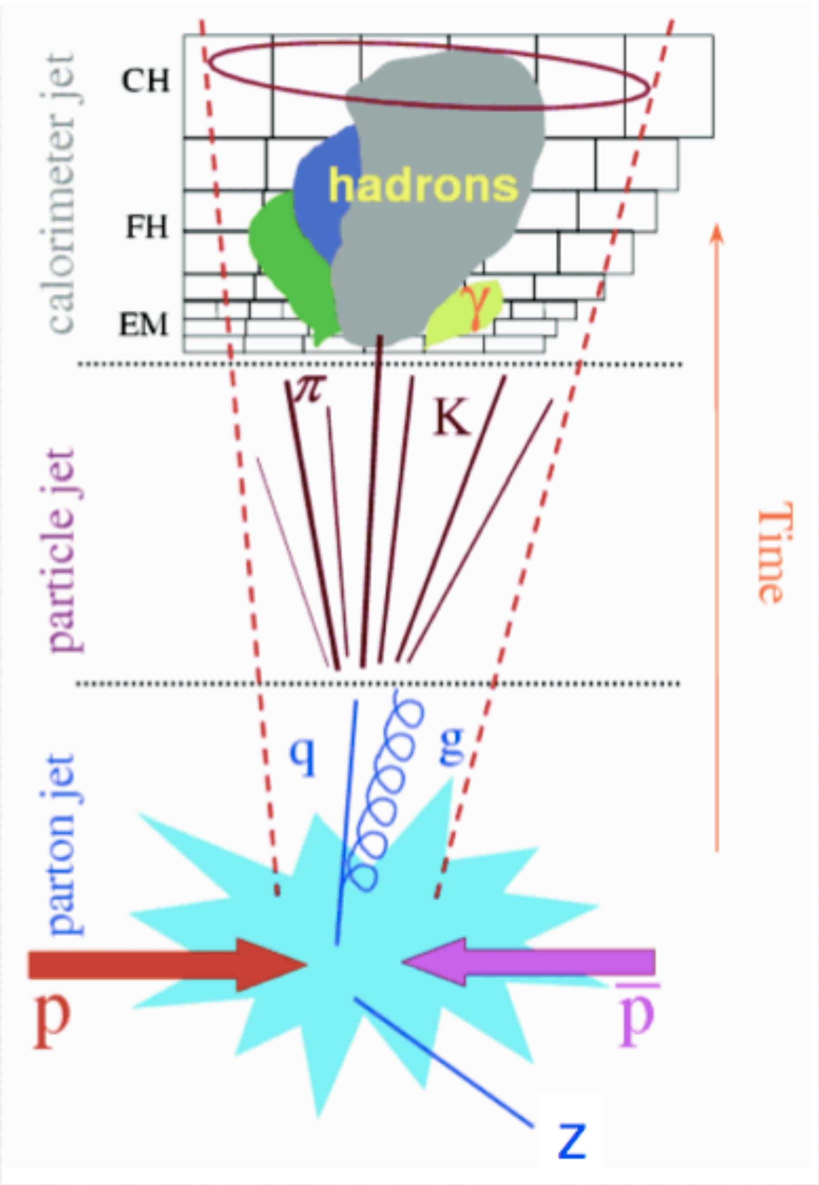


Analysis Ingredients

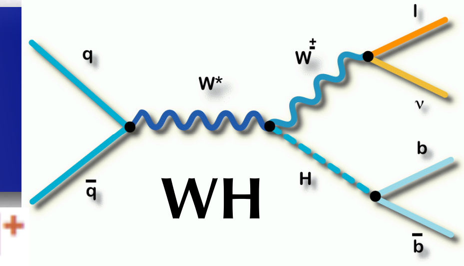
- Selection of phase space
  - want high acceptance, reconstruction and trigger efficiency for Higgs events
- Reconstruction of final state particles
- Simulation of background processes
- Normalizing the backgrounds and K-factors
  - good modeling of the data needed for further analysis
- Analyzing the data with multivariate techniques
- In the absence of signal, extracting limits

Phase Space and Reconstruction

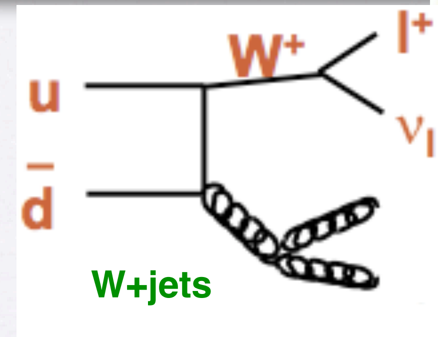
- Event Selection: $WH \rightarrow l\nu b\bar{b},\ l = e, \mu$
  - electron, muon p_T > 15 GeV
  - electron, muon $|\eta|$ < 1.1, 2 $\eta = -\ln(\tan\theta/2)$
  - missing E_T > 20 GeV
  - scalar sum of jet energies > 60 GeV
  - 2 jets with p_T> 20 GeV, $|\eta|$ < 2.5
  - 1 jet with p_T> 25 GeV, $|\eta|$ < 2.5
  - single or double b-tagging
- Electron Reconstruction
  - EM fraction > 0.9
  - shower shape requirement
  - cone isolation requirements
  - EM deposit matched to 5 GeV track
  - likelihood requirement
- Muon Reconstruction
  - hits in all layers of muon system
  - scintillator hits
  - track matching between central tracking and muon systems
  - isolation requirements
- B Jet Tagging
  - NN algorithm based on 7 lifetime observables

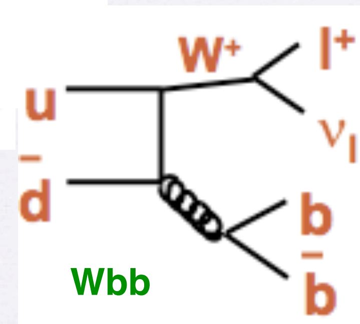


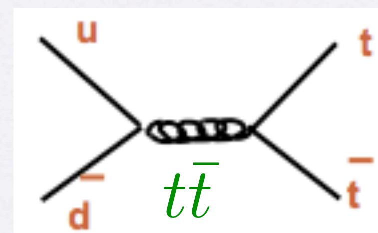
Backgrounds

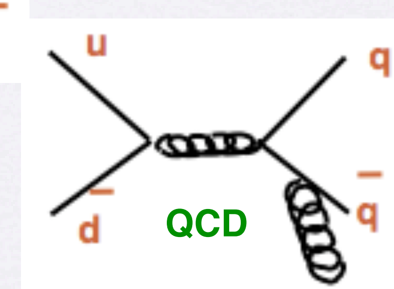


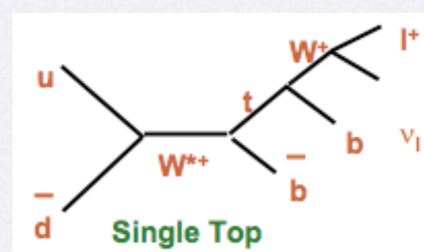
- W+jets - any process that produces a W and light flavor jets
    - dominant background before requiring b-tagging
- Wbb, Wcc - production of W and the heavier charm and bottom jets
    - dominant background after b-tagging
- tt - direct production of top pairs which decay to Wb
    - dominant background at high dijet mass
- QCD - pure jet events in which one jet mimics lepton signature
- additional contributions from single top, diboson and others

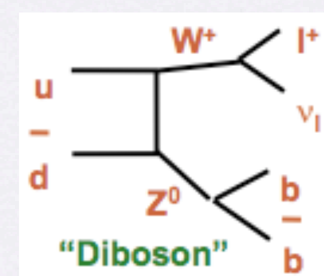










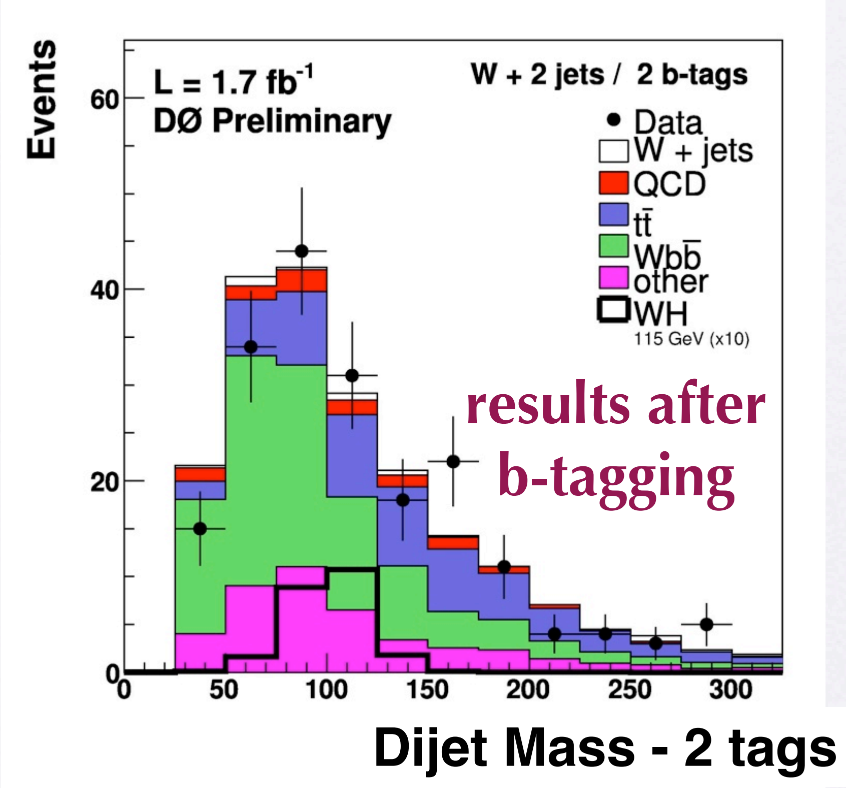


QCD instrumental backgrounds

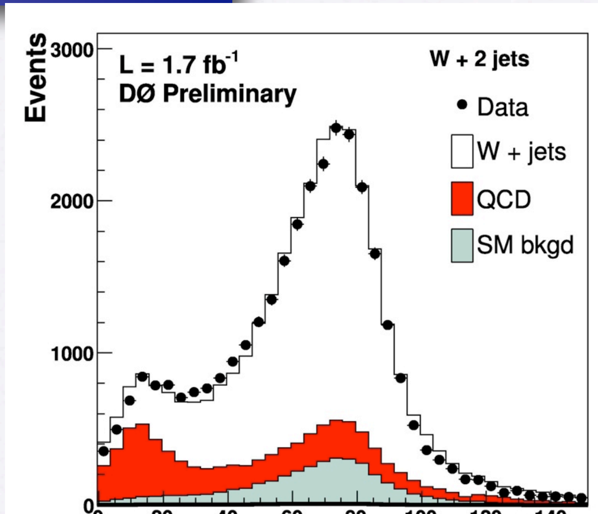
- Jets mimic signature of electrons or muons
  - electrons: jet has high electromagnetic fraction
  - muon: semi-leptonic quark decay is mis-identified as being isolation
- Fake jet contribution can be reduced by requiring “lepton” to be well separated from other jets in the event $\Delta R_{lepton\text{-}jet} > 0.5$
- Independent analysis performed on QCD-enriched data sample to determine probability that jets pass lepton identification criteria.
- QCD shape estimated separately for each distribution

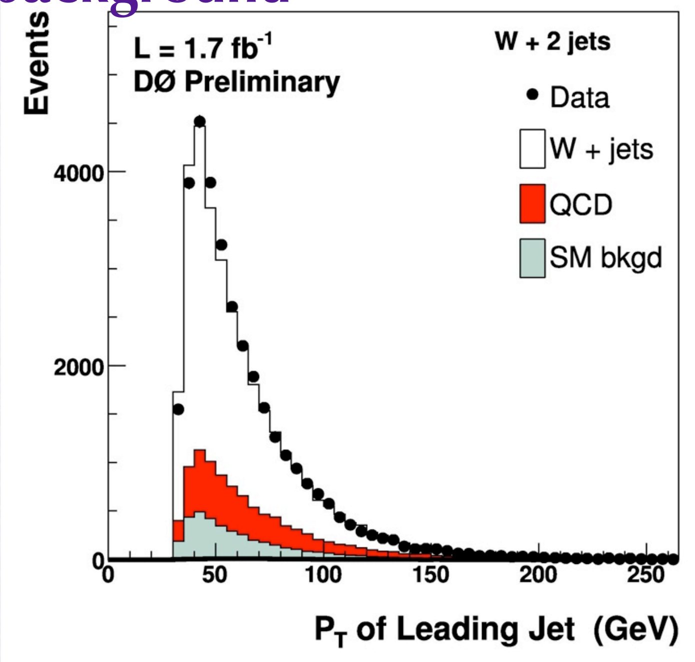
Results

- Backgrounds modeled by MC, normalisation scaled to NLO cross sections when available; W+jets normalized to data; QCD background determined with data
- Data binned by number of jets, number of b-tags. Most sensitive in 2 jet, 2 b-tag sample
- Signal to background ratio still very small after all cuts



good modeling of background





Event Yields

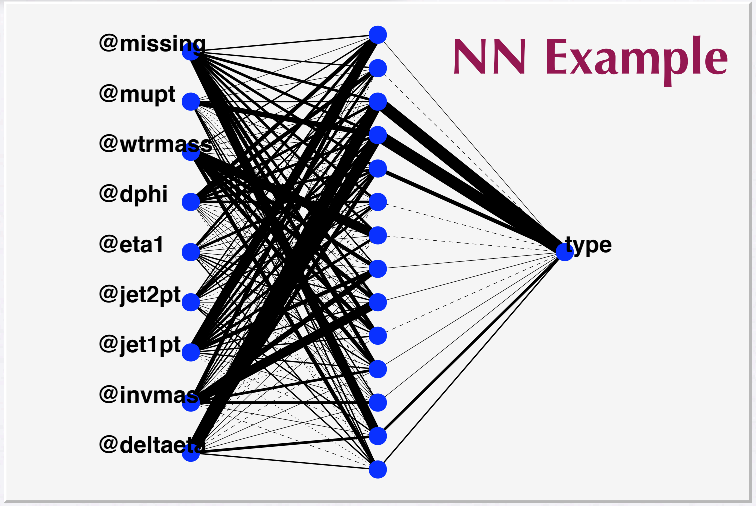
	W + 2 jets	W + 2 jets (1 b jet)	W + 2 jets (2 b jets)	W + 3 jets	W + 3 jets (1 b jet)	W + 3 jets (2 b jets)
WH	9.92 ± 1.44	3.94 ± 0.63	2.32 ± 0.44	2.43 ± 0.42	0.95 ± 0.18	0.59 ± 0.12
WZ	645 ± 90	38 ± 6	7.6 ± 1.34	153 ± 24	10 ± 2	2.4 ± 0.5
$Wb\bar{b}$	1352 ± 346	441 ± 117	91.7 ± 26.0	433 ± 118	137 ± 39	33.9 ± 10.0
$t\bar{t}$	348 ± 83	139 ± 34	53.8 ± 14.3	596 ± 152	238 ± 63	122.4 ± 34.3
Single top	189 ± 37	78 ± 16	19.4 ± 4.4	62 ± 13	25 ± 6	10.1 ± 2.5
QCD Multijet	2908 ± 436	193 ± 36	10.8 ± 3.3	1051 ± 158	87 ± 16	12.2 ± 4.7
W+ jets (light,c)	28013 ± 3181	470 ± 137	20.9 ± 6.9	5332 ± 836	132 ± 41	11.5 ± 4.0
Total expectation	33458 (n.t.d.)	1360 ± 187	204.1 ± 31.0	7627 (n.t.d.)	630 ± 86	192.5 ± 36.3
Observed Events	33458	1403	193	7627	570	173

- Table summarizes the number of expected events for Higgs events and all background processes given all cuts in different jet/btag bins
- W+2jets samples used for analysis, W+3jets samples used for control
- In most sensitive bin, $S/\sqrt{(S+B)} = 2.32 / \sqrt{204.1} = .162$

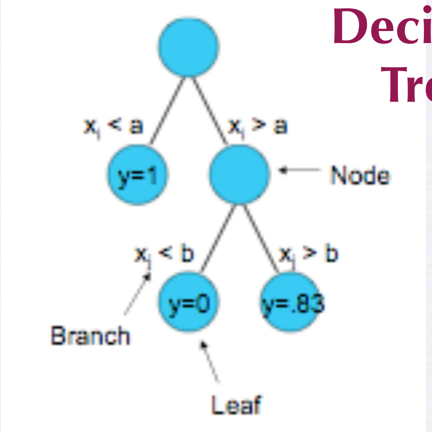
Multivariate Techniques

- Neural Networks
  - exploits correlations between kinematic properties of event objects
  - “trained” on reconstructed variables in signal and background MC samples to find correlations
  - run on data to identify events with high signal probability
- Matrix Element Discriminant
  - use LO matrix elements to calculate event probabilities
  - for each event and process, integrate ME over phase space including efficiency and resolution functions
- Decision Trees
  - similar to neural networks, classifies events as more signal-like or background like

NN Example



Decision Tree

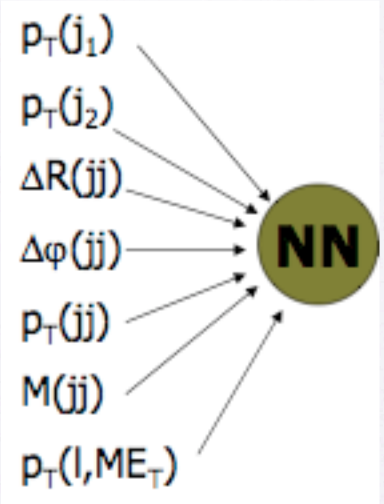


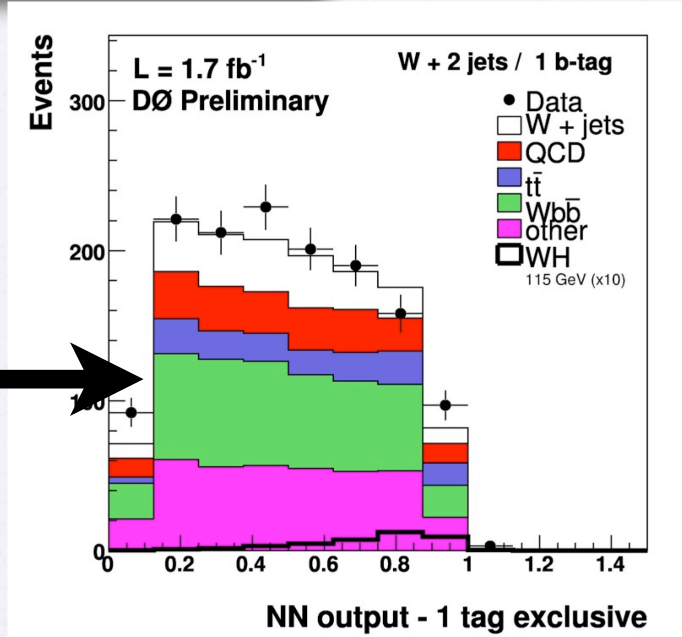
Different techniques usually give comparable improvement in sensitivity

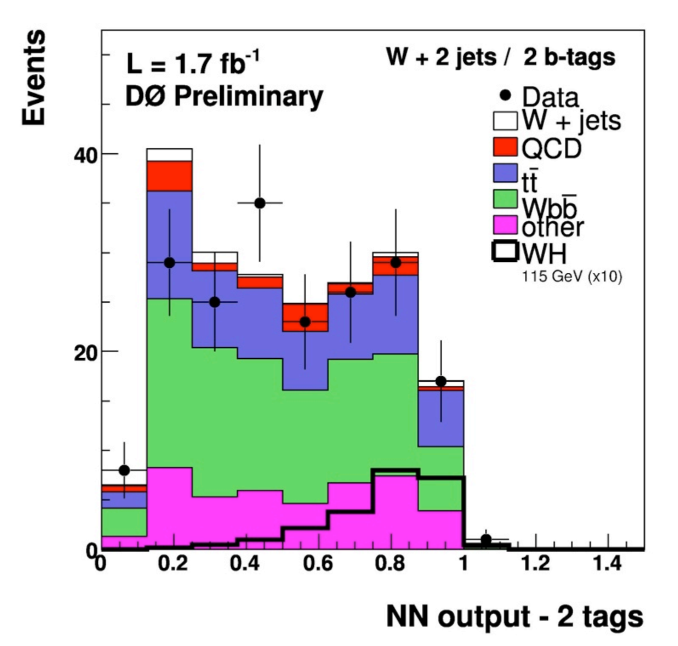
NN applied to WH

$WH \to l\nu b\bar{b},\ l = e, \mu$

- Neural net discriminant tuned to further enhance signal and background separation
  - Event variables are inputs:
  - NN trained on subset of background samples, but run on all backgrounds



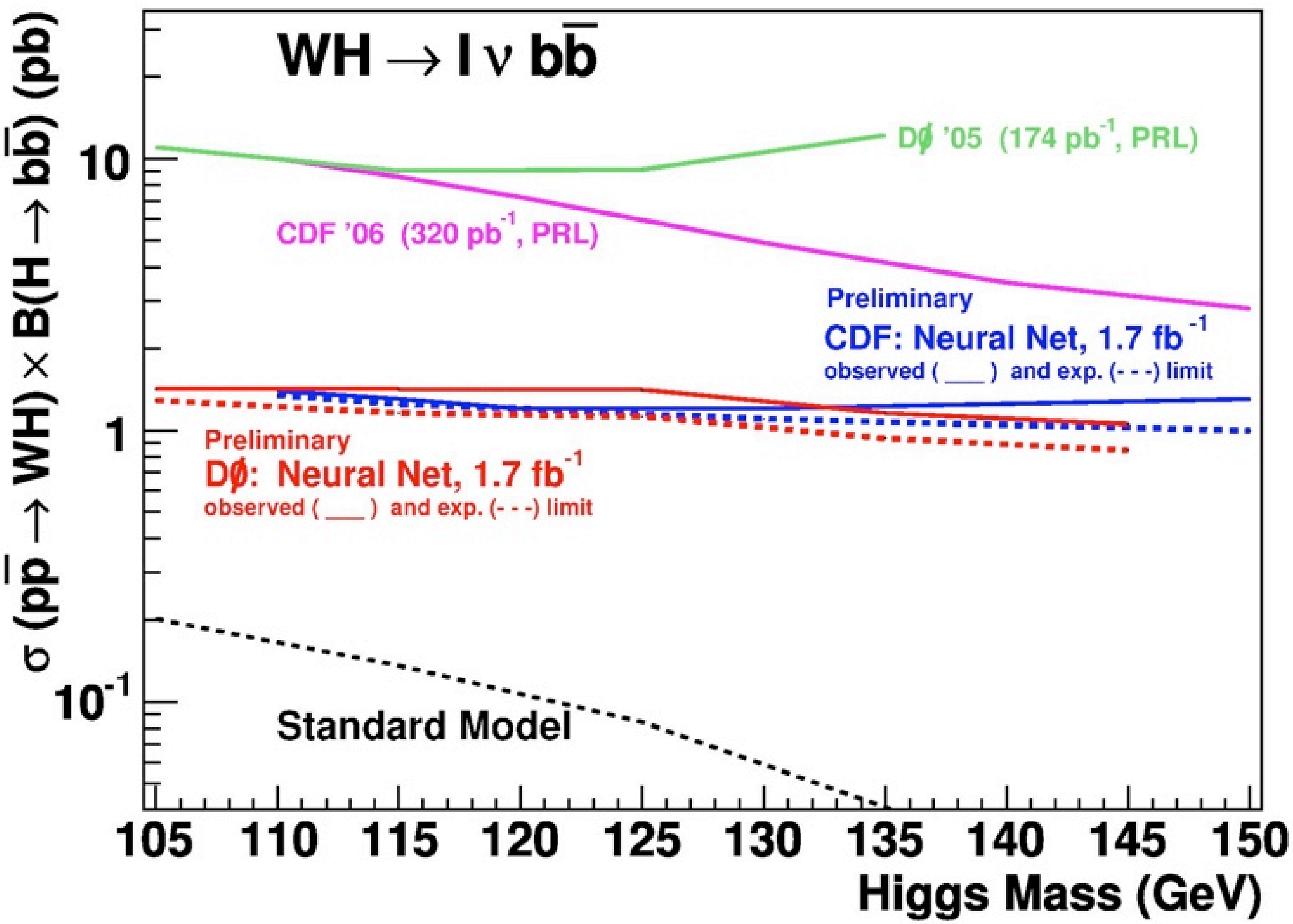




- Systematics
  - luminosity and normalisation
  - QCD background estimation
  - input background cross-sections
  - jet energy scale, dijet mass resolution
  - b-tagging, lepton-id
- Final result determined from fit to NN output

Final Result

$WH \rightarrow l\nu b\bar{b},\ l = e, \mu$



Result: Limit/SM expectation ~10 for m_H = 115 GeV

Summary of all Modes

Channel	Lumi /Technique	Final state	#chan.
$WH \rightarrow l\nu\ bb$	1.7 fb^{-1} / NN	e/μ, 1b/2b	2*(2+2)
$ZH \rightarrow ll\ bb$	1.1 fb^{-1} / NN	e/μ, 1b/2b	2+2
$ZH \rightarrow \nu\nu\ bb$	0.9 fb^{-1} / NN	$Z \rightarrow \nu\nu$, $W \rightarrow \ell\nu$ (2b)	2
$H \rightarrow WW^*$	1.7 fb^{-1} /NN	ee, $e\mu$, $\mu\mu$	2*3
$WH \rightarrow WWW^*$	1 fb^{-1} / 2D LHood	ee, $e\mu$, $\mu\mu$	3

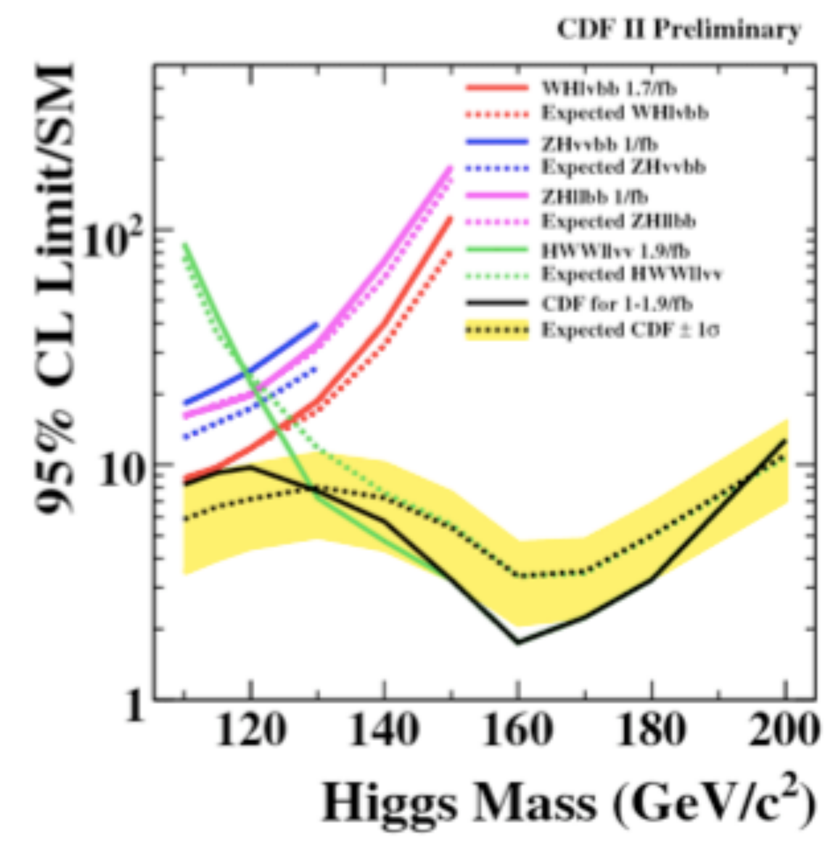
Total of 23 DØ channels combined (tau-channels not included yet)

8

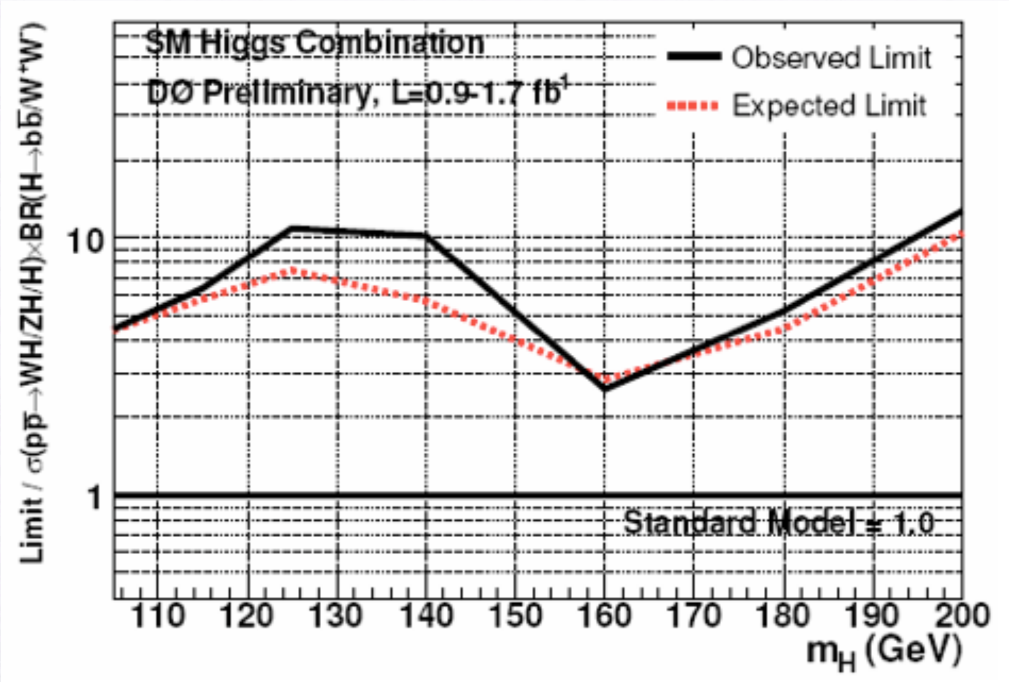
Combinations & Higgs Sensitivity

- Current state-of-the-art limits on Higgs production for $m_H < 200$ GeV per experiment

CDF

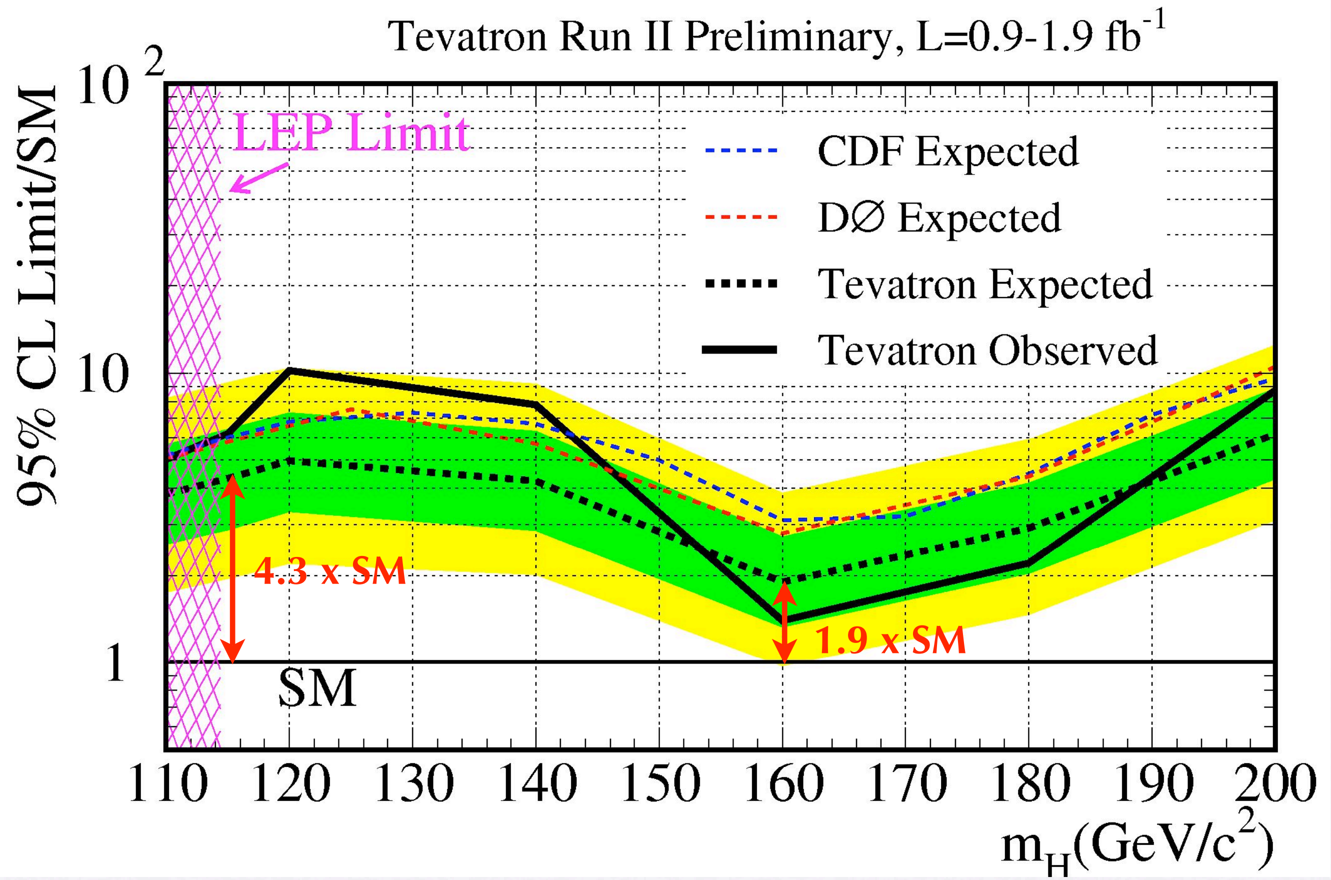


DØ



Latest Higgs Results from Tevatron

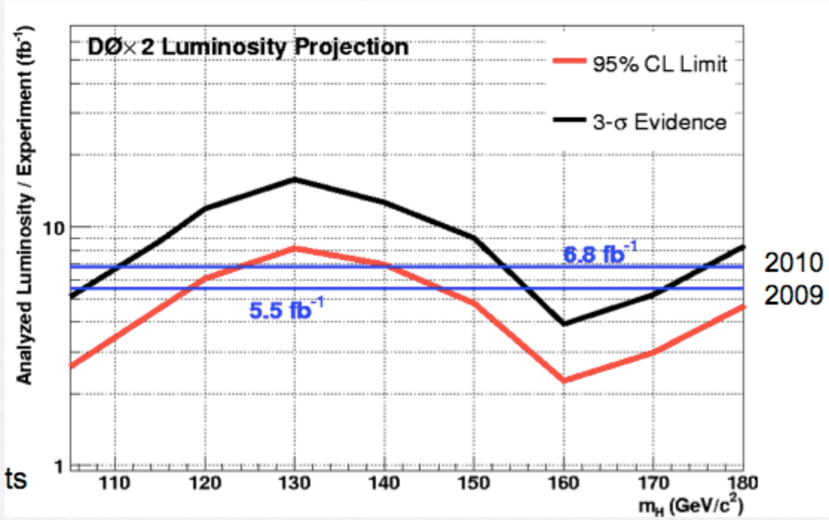
- Nearly at required sensitivity for m_H = 160 GeV! Look for tantalizing results at upcoming conferences (maybe even Moriond ‘08).
- D0 and CDF sensitivities are largely similar, differences can appear as each experiment updates their analyses

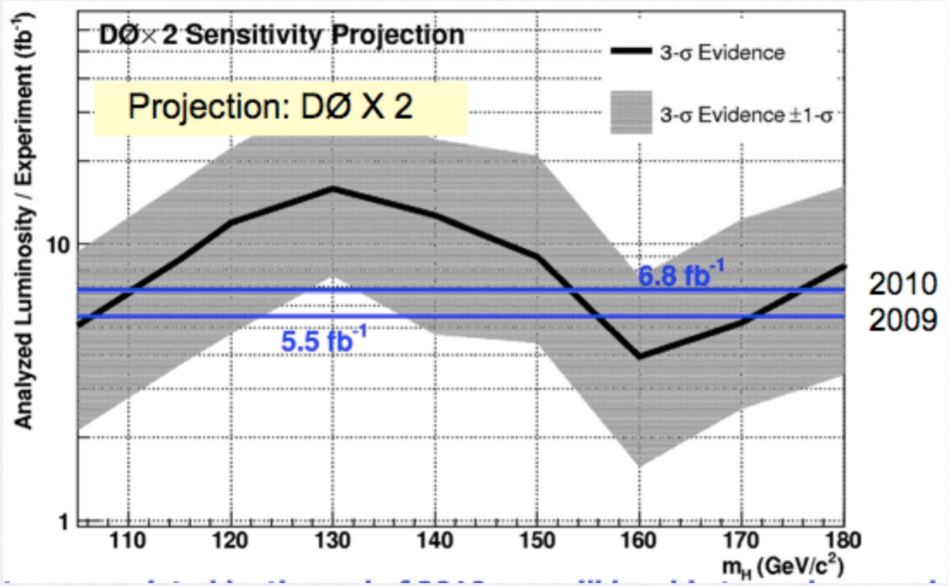


Observed limit @ m_H=160 GeV - 1.4 times SM expectation

Tevatron Projections

- Including data taking efficiency, projected full data set will be
  - 5.5 fb-1 by end of 2009
  - 6.8 fb-1 by end of 2010
- Assumption: projected sensitivity for m_H = 115 GeV will be factor x2 higher than current for full dataset
  - Improvement from 2005 -> 2007 was factor 1.7
  - Several possibilities for improvement:
    - Better b-tagging with Layer 0
    - dedicated group studying dijet mass resolution
    - many gains to be made in acceptance
    - implementation of multivariate techniques

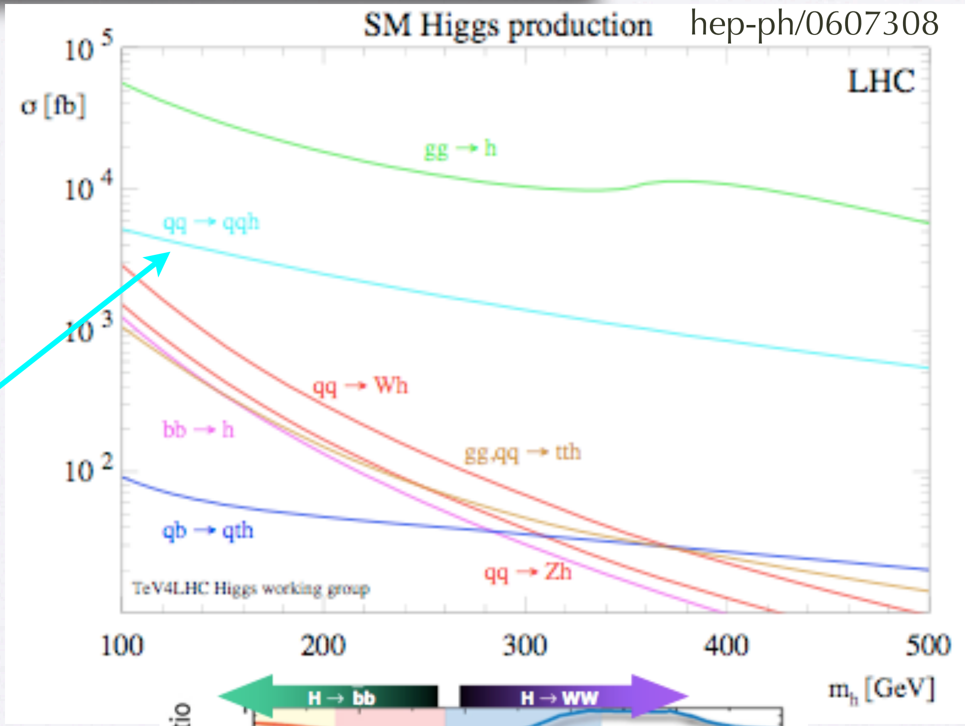


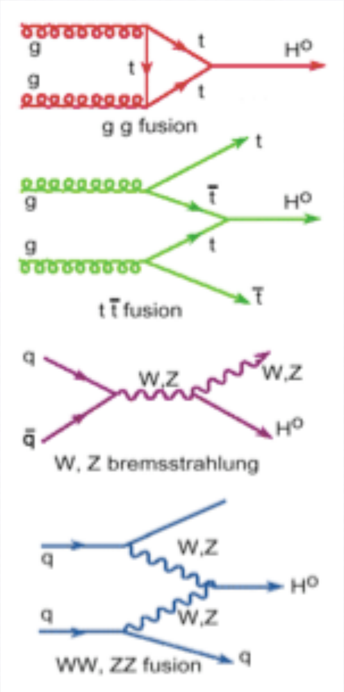


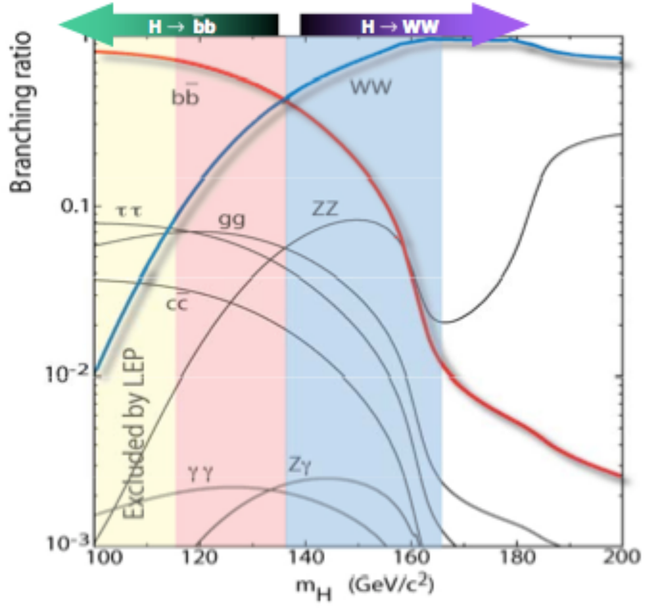
Higgs Searches at the LHC

Higgs Production at the LHC

- All cross sections go up by 1-2 orders of magnitude (backgrounds go up as well)
- Still dominated by gluon fusion
- Relative Vector Boson Fusion rate much higher than at the Tevatron

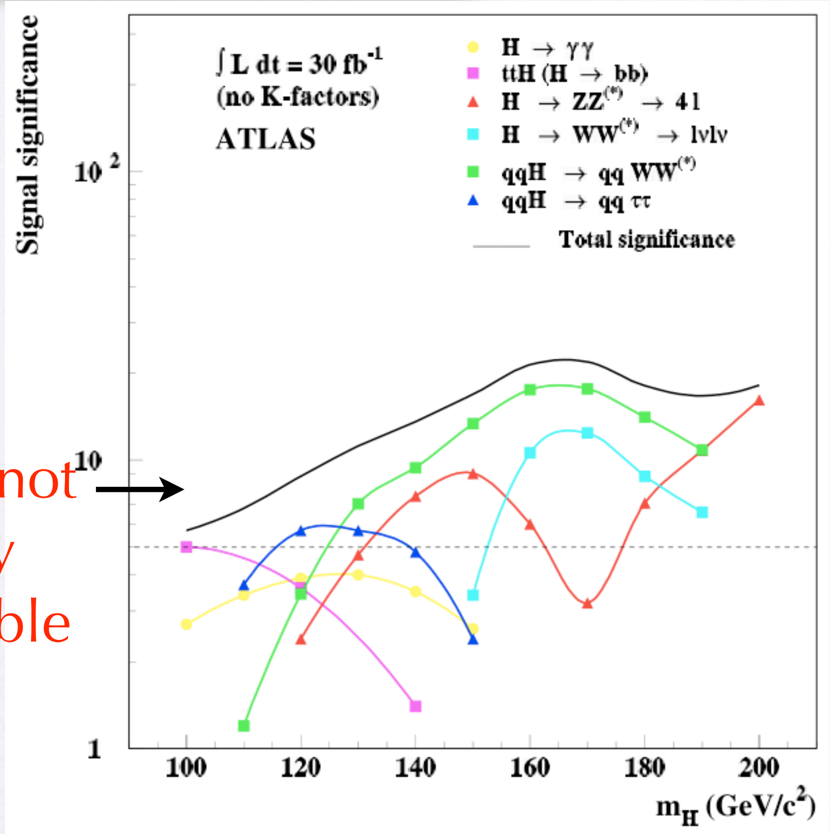




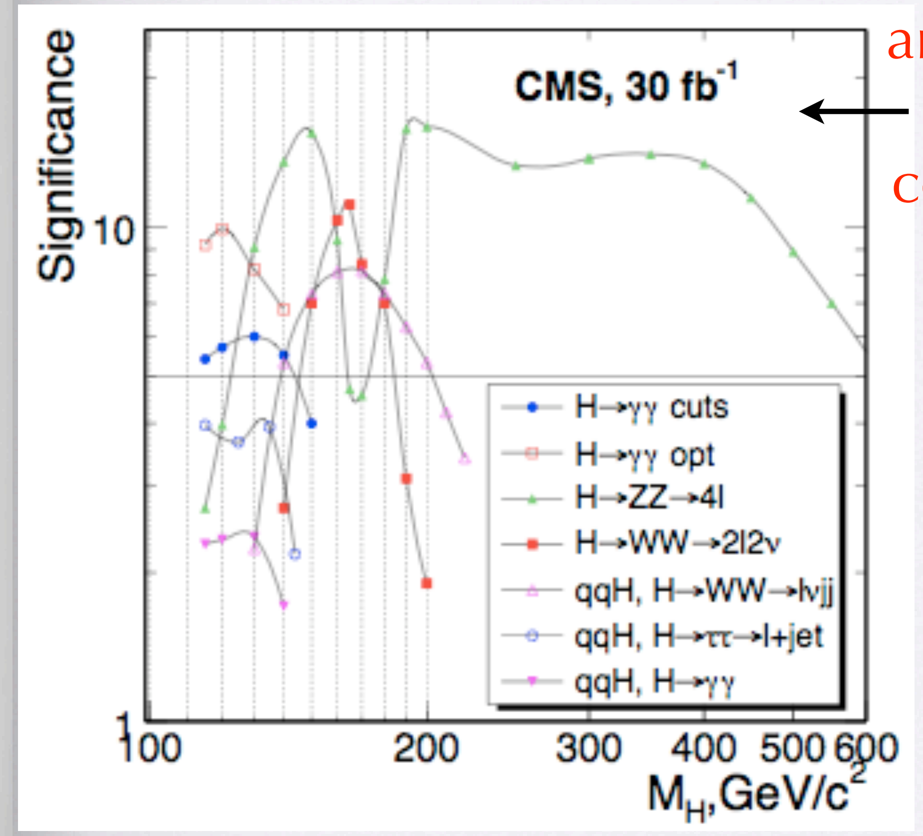


Higgs Sensitivity at LHC

- Highest sensitivity at ATLAS is for VBF production in the tau channel for $m_H < 130$ GeV and in the W channel for $m_H > 130$ GeV.



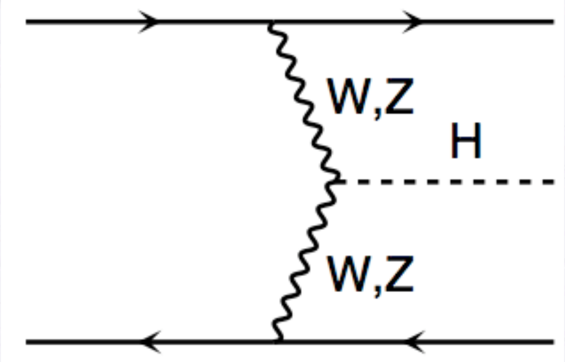
analyses not directly comparable

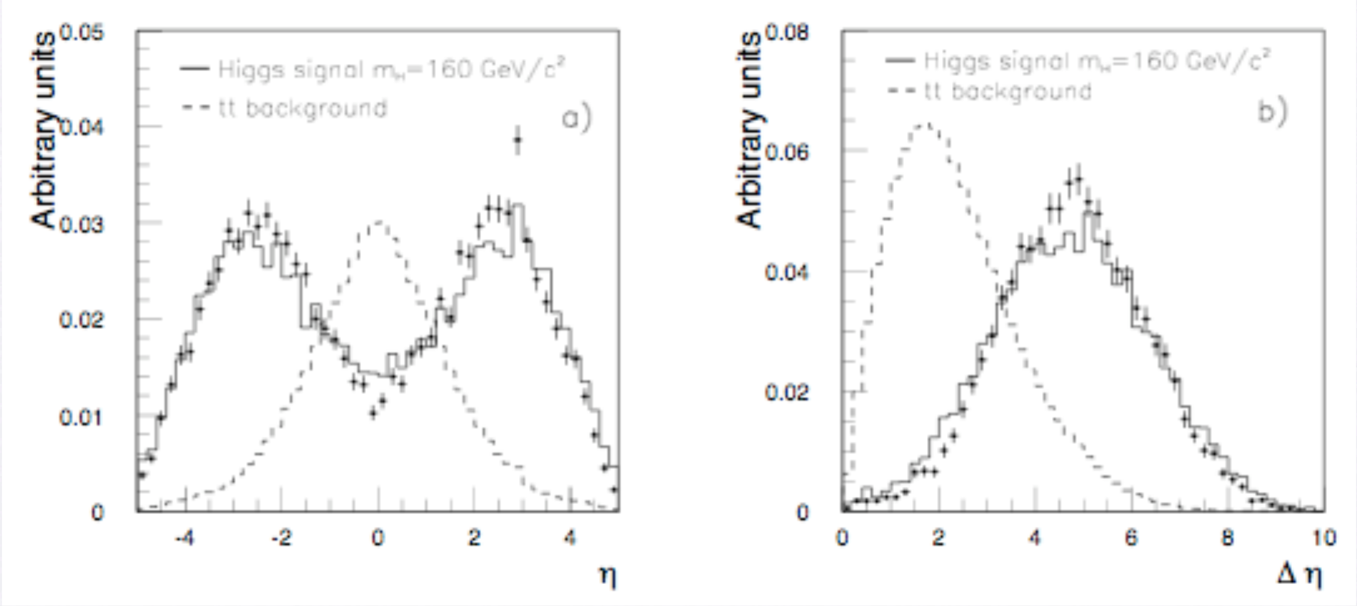


- CMS has better sensitivity in H->γγ, and excellent sensitivity in H->ZZ except for the window with $2m_W < m_H < 2m_Z$.

Vector Boson Fusion at LHC

- W or Z radiated from each of the incoming quarks, produces a central Higgs
- Topological feature of quark jets produced very forward increases sensitivity
- Introduces possibility to veto on jet activity in the central region to reduce backgrounds

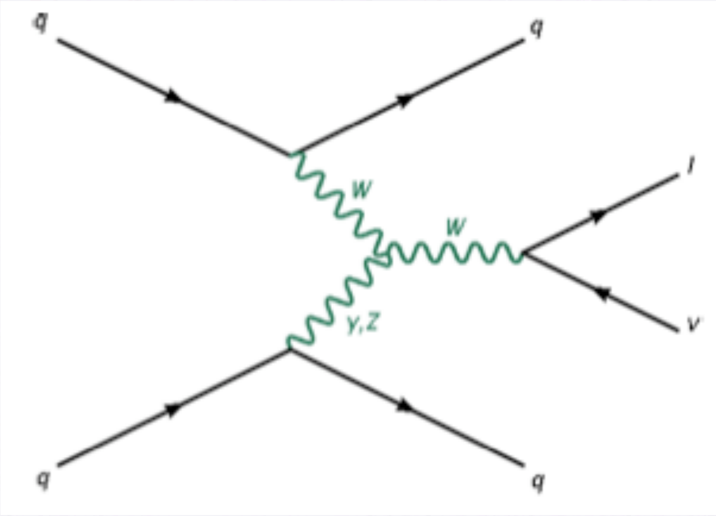




Vector Boson Fusion at the Tevatron

VBF at the Tevatron

- Production of W,Z or Higgs by weak boson fusion process (i.e. not gluon induced)
- Testbed for VBF search methodology being employed by LHC searches
- Validation of VBF-produced W, Z standard model cross sections
- Currently studying W production, where dominant process is W-γ fusion.
- Event signature is similar to t-channel single top or WH production, but no b-tagging.



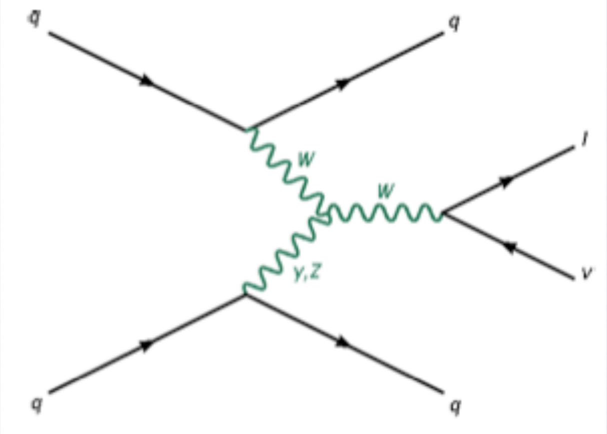
new analysis at Tevatron

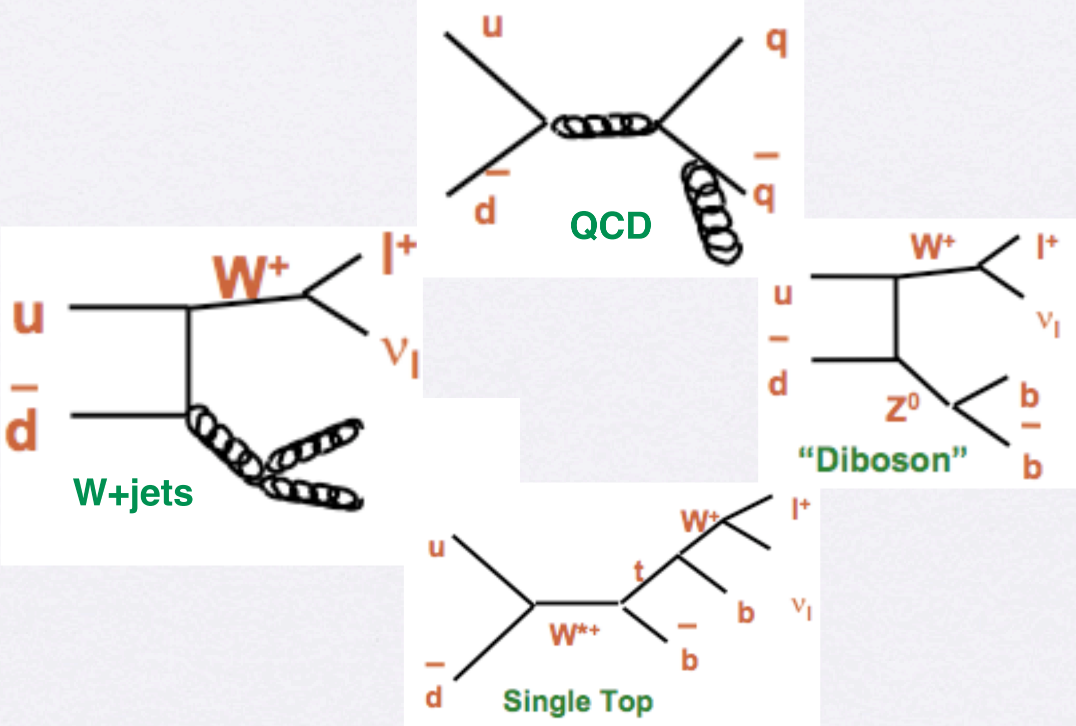
Analysis Ingredients

- Phase Space Selection
  - muon $p_T > 15$ GeV
  - muon $|\eta| < 2$
  - missing $E_T > 20$ GeV
  - 2 jets with $p_T > 20$ GeV
- Similar backgrounds to WH

W+jets - dominant
Wbb, Wcc
tt and single top
Z->μμ
Diboson
QCD

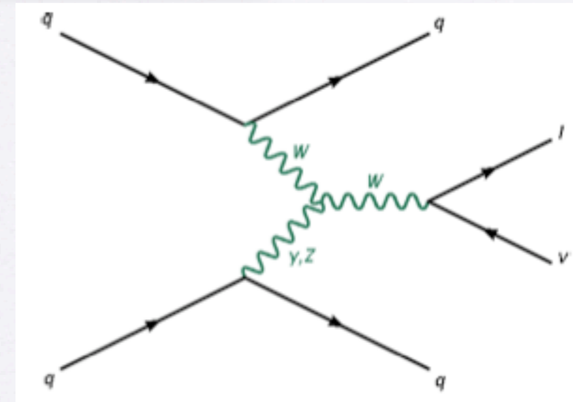
muon channel
only so far

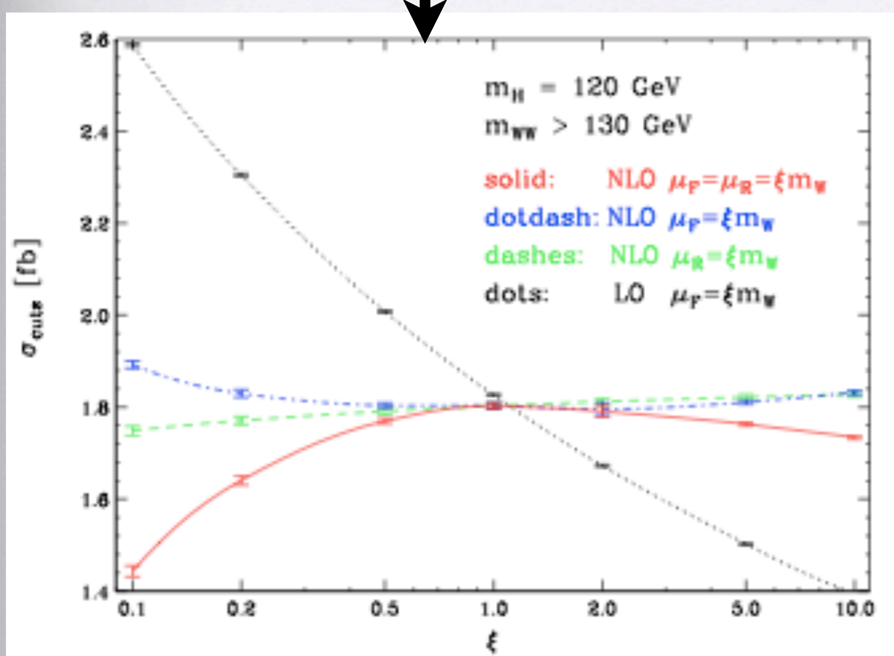


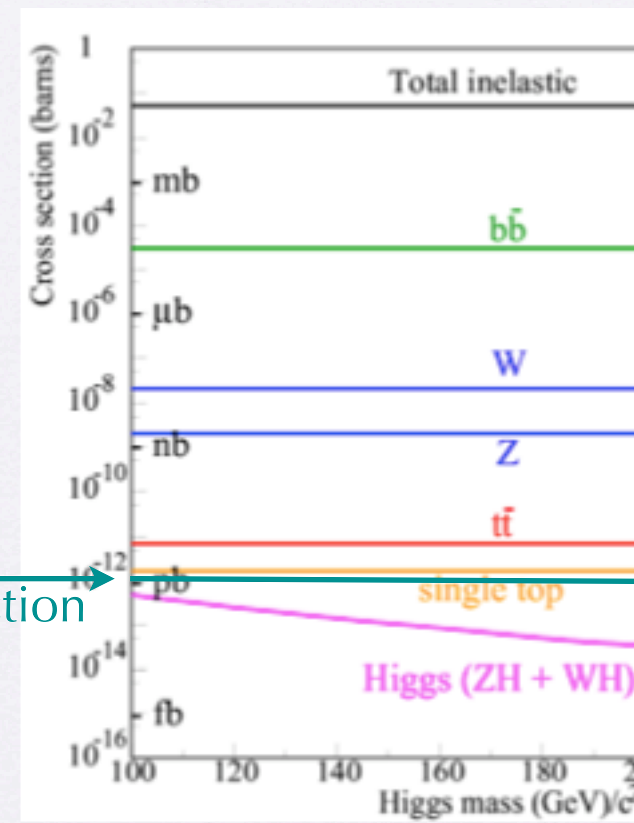


Signal Monte Carlo

- NLO production cross section ~2.4 pb
- VBF can be simulated with Herwig, Sherpa, VBFNLO
- VBFNLO fully flexible MC that generates event files at LO, cross sections at NLO. Has very good control over theoretical uncertainties.



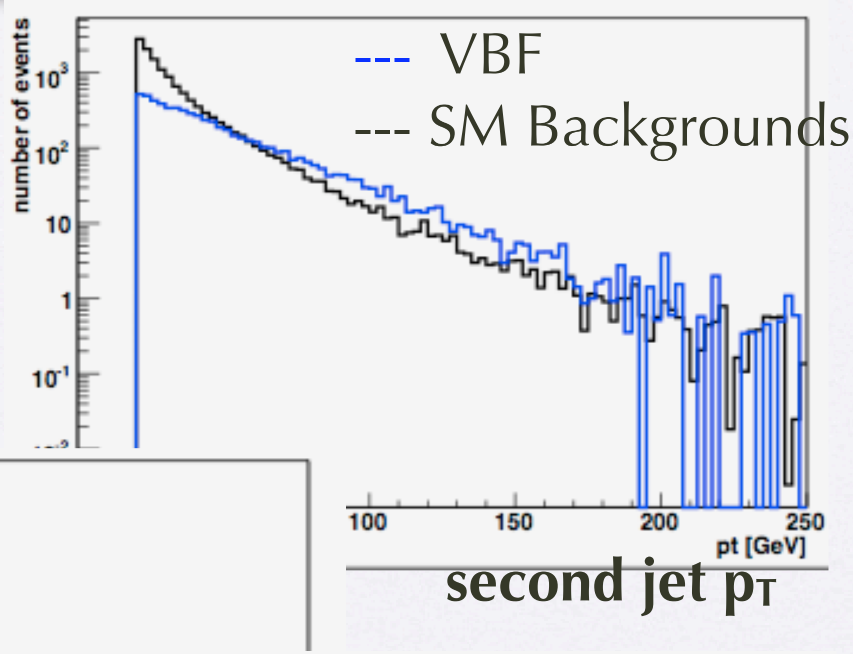


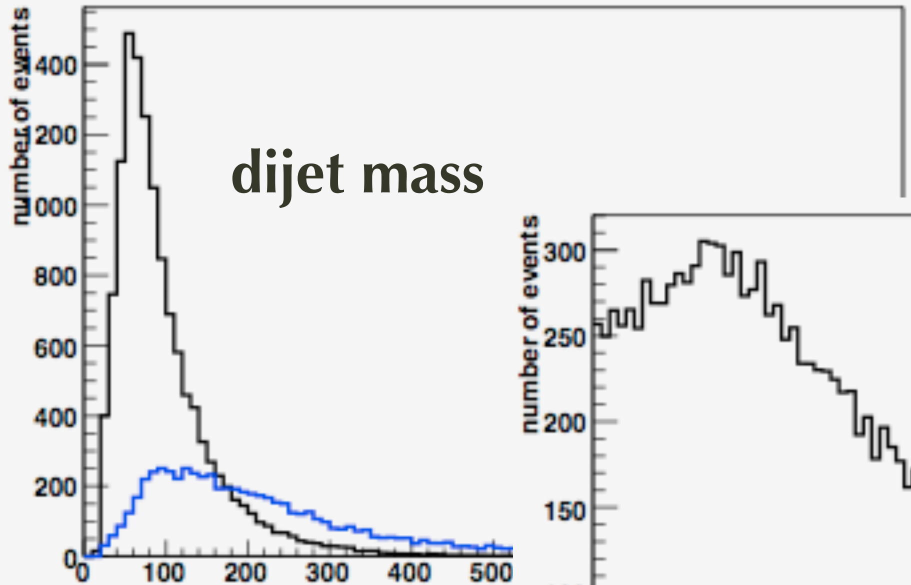


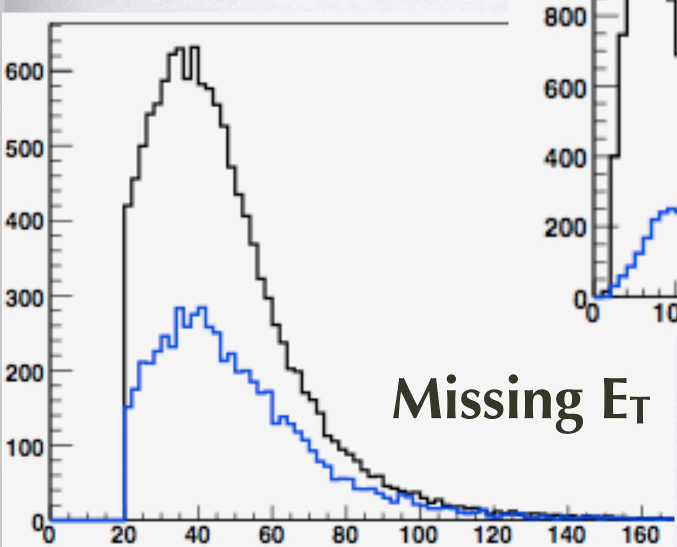
VBF W production

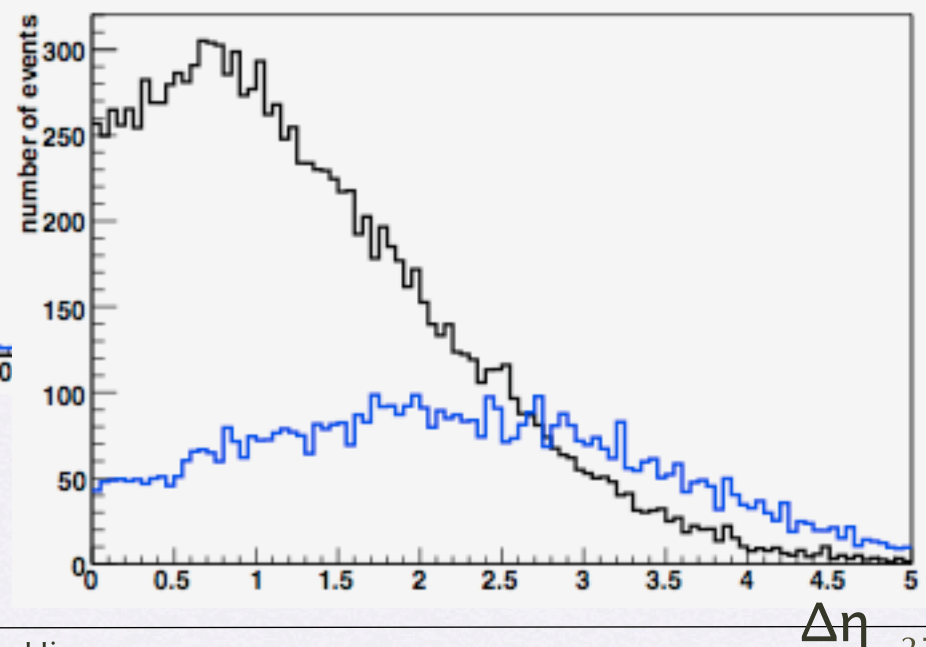
VBF - W production

- Several topological features which distinguish VBF from SM backgrounds:
  - Forward jets => large $\Delta\eta$
  - Forward jets => large dijet mass
  - harder jet p_T spectra







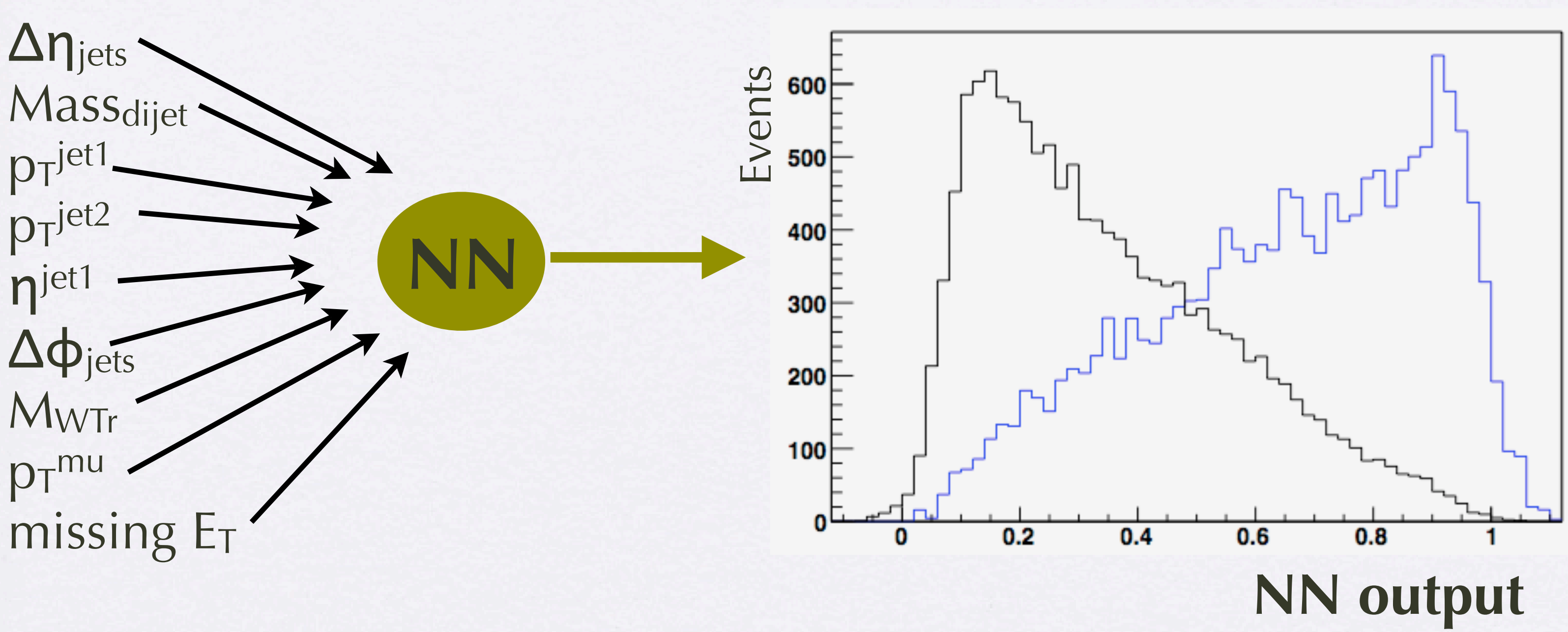


Neural Network Discriminant

- Shape spectra differences between WBF and other SM processes make VBF ideally suited for multivariate approach
- NN trained on all simulated background samples and VBFNLO signal sample
- Currently being studied/optimized

--- VBF

--- SM Backgrounds



Generator issues

VBF very sensitive to correct modeling of jet emission angle

Different generators show non-negligible variation in jet η.

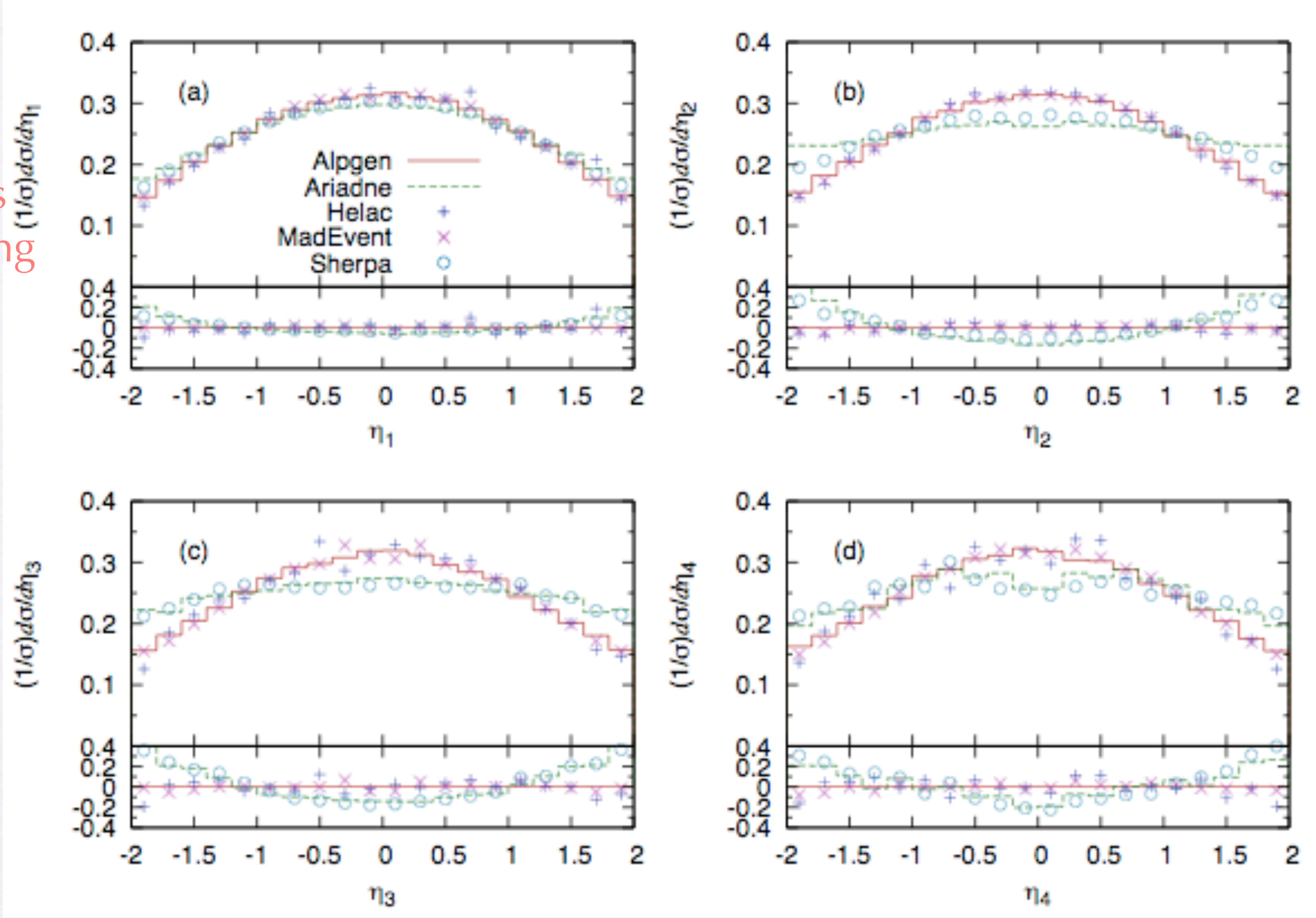
arXiv:0706.2569

Alpgen

- most common MC used at TeV for modeling backgrounds
- employs MLM matching

Sherpa

- can simulate VBF
- uses CKKW matching
- being studied at LHC



Crucial to understand for VBF measurement!

Ongoing investigations

- Neural network optimization for separating signal and background
  - adding observables
  - optimizing background samples for training
- Forward jet description difficult for LO generators
  - Correct inclusion of all diagrams by chosen generator?
  - Color connection between forward jets and proton remnant can get hairy
  - Different parton showering schemes
  - Scheme chosen for matching matrix elements and parton showers
- Validation of forward jet tagging and mini-jet veto. VBF Higgs searches at LHC rely heavily on this method.

Summary and Conclusion

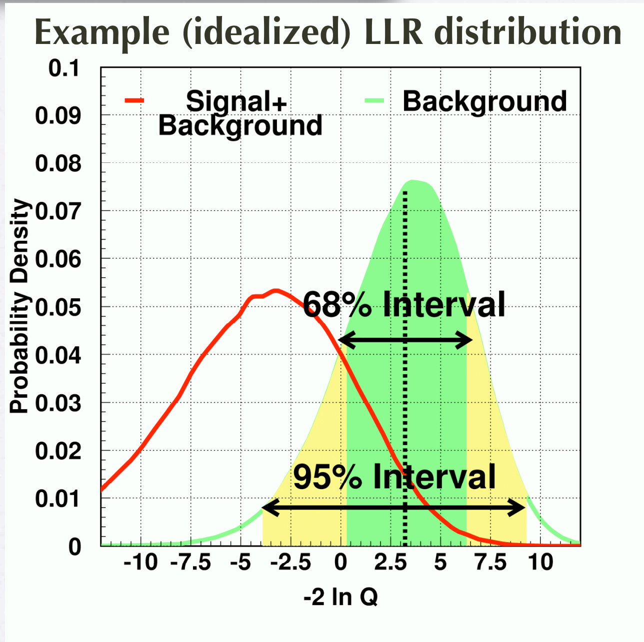
- Higgs Searches at the Tevatron and LHC are among the most exciting work being done in HEP today.
- Tevatron sensitivity is improving faster than the increase in luminosity due to intensified efforts in improving reconstruction efficiencies, triggering, jet resolutions, b-tagging algorithms, and more.
- LHC is the only place that will unequivocally discover the SM Higgs Boson (if it is there!), but the Tevatron may get a glimpse of it first.
- Tevatron is a good testbed for search techniques employed at LHC.
- Many important lessons to take from the Tevatron
  - QCD can be very difficult to model accurately
  - Multivariate techniques for object ID (like b-tagging) and event selection perform extremely well.
  - Choose multivariate techniques that are complementary
- Understanding VBF physics at the Tevatron will be useful for validating LHC Higgs searches.

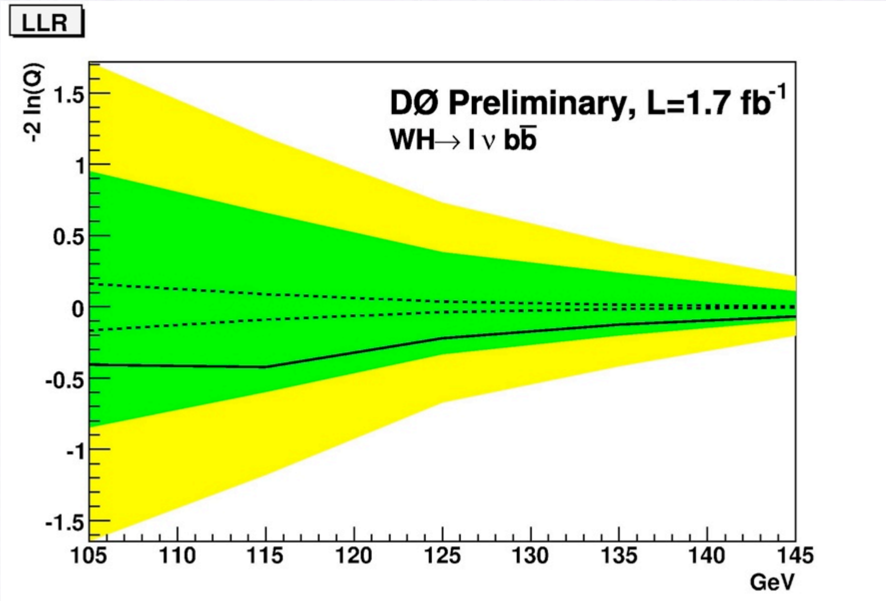
Backup

Deriving Limits

$WH \rightarrow l\nu b\bar{b},\ l = e, \mu$

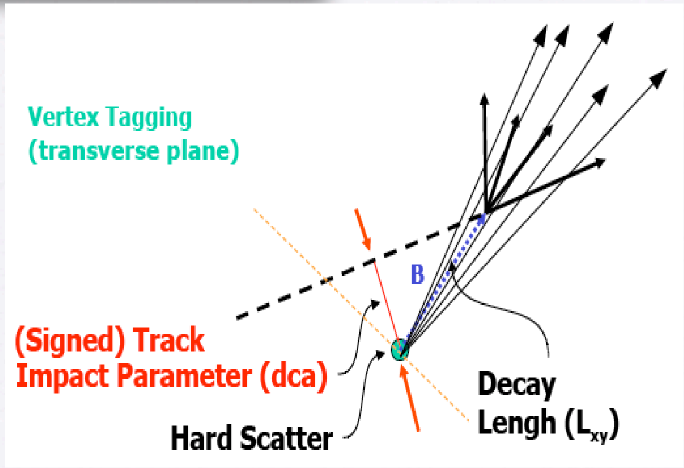
- Limits derived using semi-frequentist CL_s method where test statistic is LLR = -2LogQ = -2Log[P(s+b)/P(b)]
  - P are probability distribution functions for the signal+background and background only hypotheses
  - P are populated via random Poisson trials with mean values given by the expected number of events in each hypothesis.
  - Systematic uncertainties are incorporated by varying the expected number of events in each hypothesis according to the size and correlations of the uncertainties





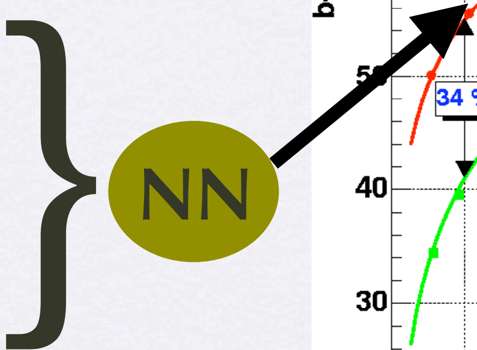
B-Tagging

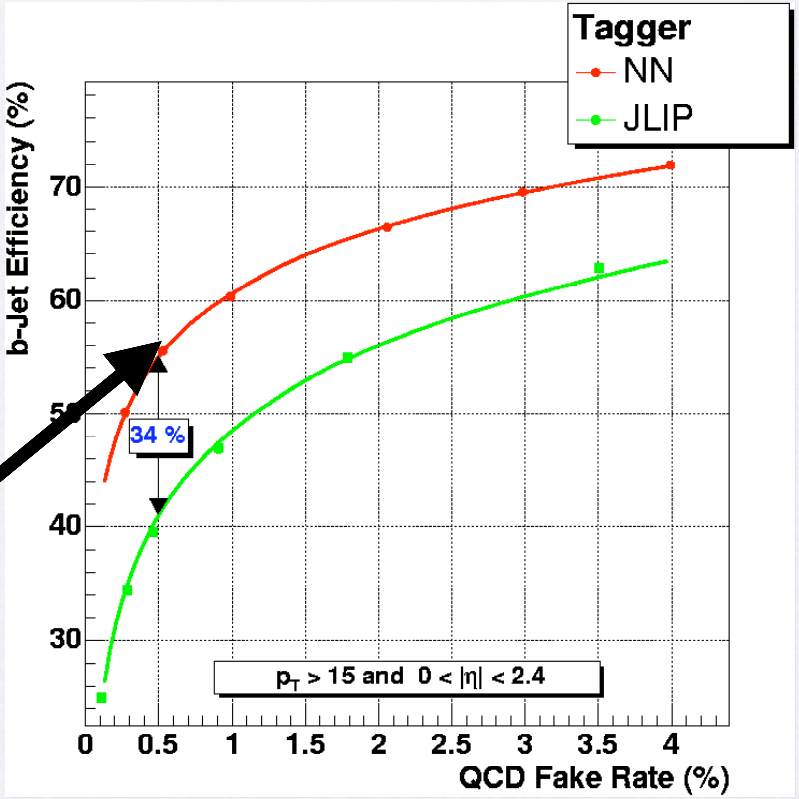
- Several approaches:
  - soft lepton tag
  - IP based tagging
  - secondary vertex reconstruction
- Most D0 analyses now use neural network discrimination for b-quarks
  - large improvement over individual taggers
  - Loose 70% eff, 4.5% mistag
  - Tight 50% eff, 0.3% mistag
  - WH: Tight,Loose operating points



Combine in Neural Network:

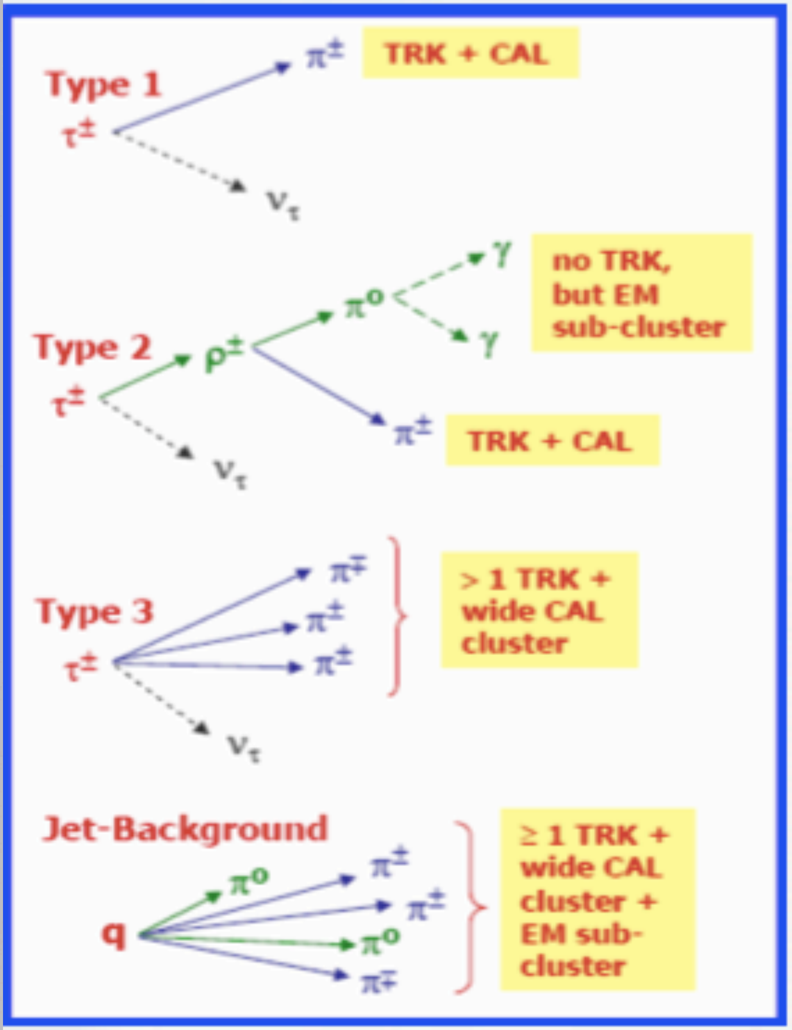
- vertex mass
- vertex number of tracks
- vertex decay length significance
- chi2/DOF of vertex
- number of vertices
- two methods of combined track impact parameter significances



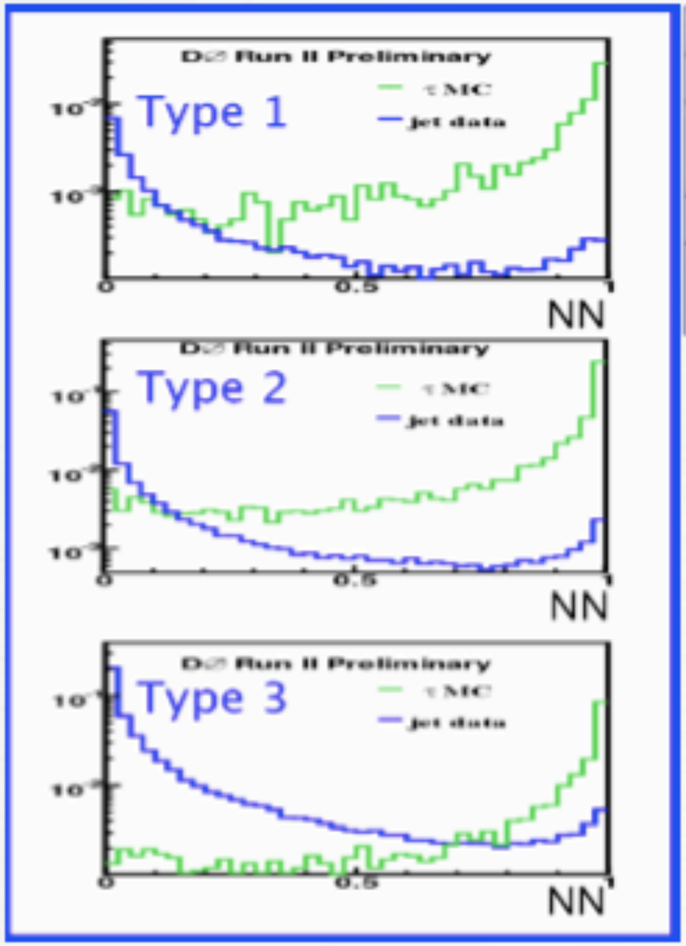


Tau identification

- Neural-net based ID
- 3 NN's for three distinct τ types:



- Performance (for $p_T > 15$ GeV):



Tau Type	1	2	3
Reconstruction			
Jets	1.5	10	38
Taus	9.1	50	20
NN > 0.9			
Jets	0.04	0.2	0.8
Taus	5.8	37	13

- Validated with the Z's (the first Tevatron Run II Z cross section measurement)

LHC Jet Efficiencies

hep-ph/0402254

- Studied in MC for LHC VBF Higgs production
- Never studied with real data!

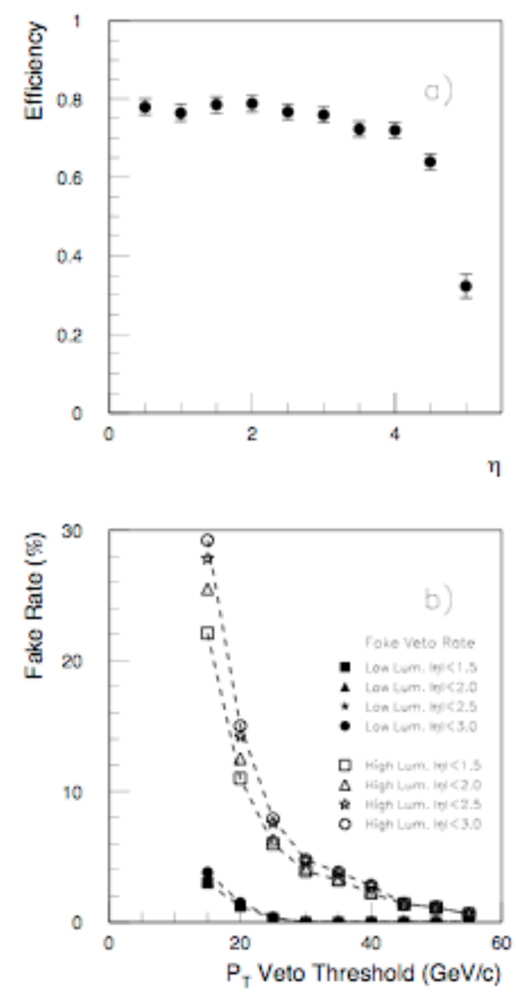
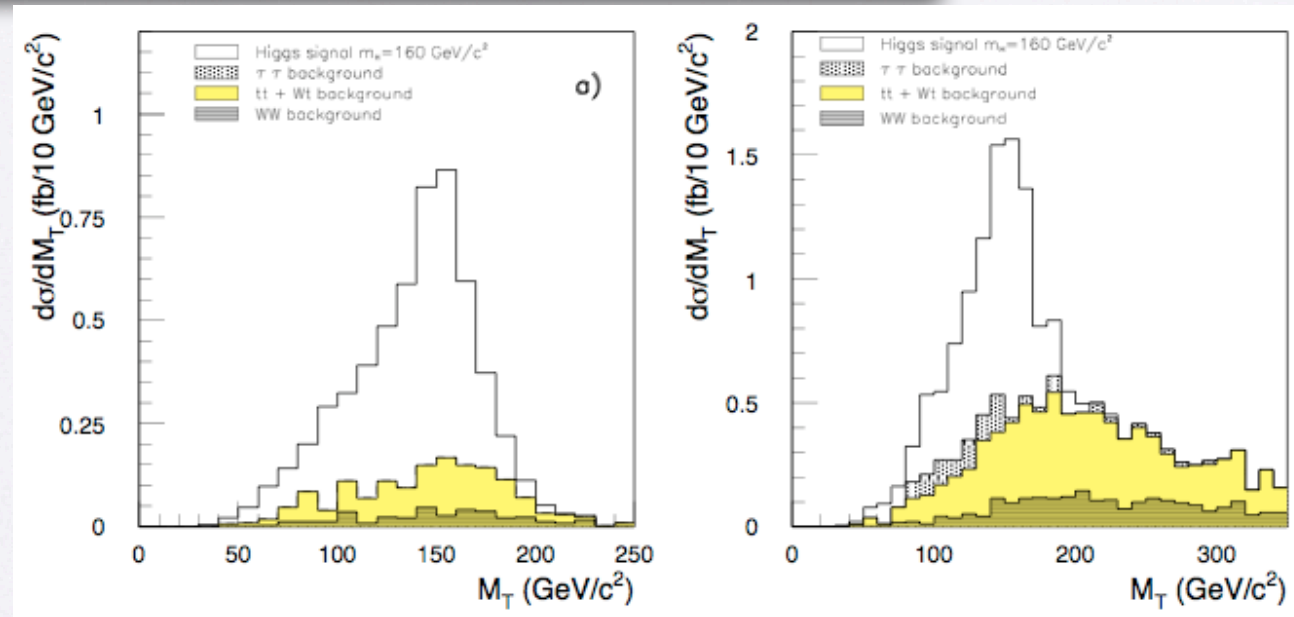
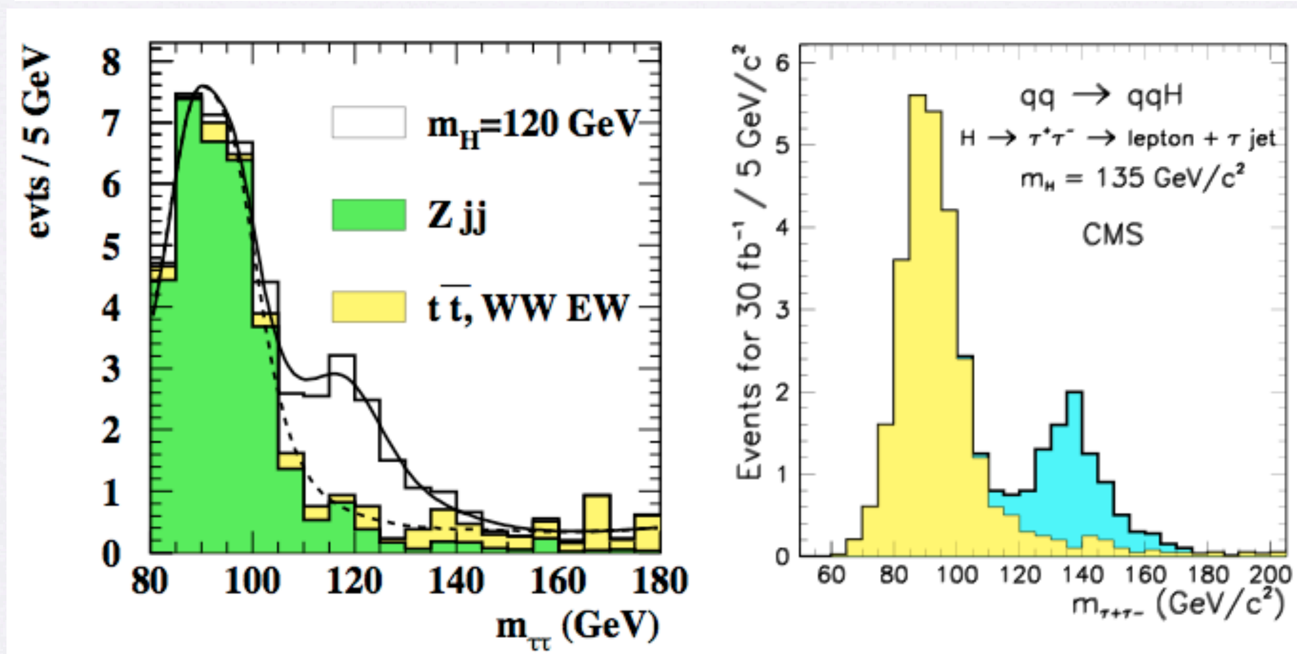


Fig. 2. *a) Efficiency for reconstructing a tag jet with $P_T >$ 20 GeV/c which originates from a parton with $P_T >$ 20 GeV/c as a function of pseudorapidity η of the parton. b) Probability for finding at least one jet from pileup events in central rapidity intervals in the ATLAS detector as a function of the P_T value used in the jet definition. The dashed curves connect the points for pseudorapidity itervals $|\eta| < 1.5$ and $|\eta| < 3.0$ for low and high LHC luminosities.*

VBF H->tautau -> emu

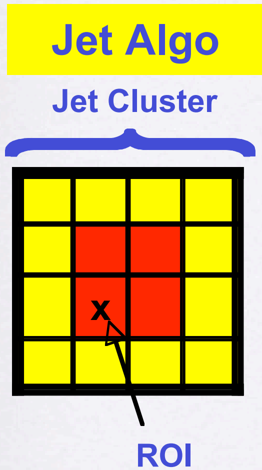
- Projected Higgs signal at m_H = 160 GeV for tight and loose electron criteria





L1Cal Algorithms

Introduce clustering algorithms at L1: high efficiency, low latency

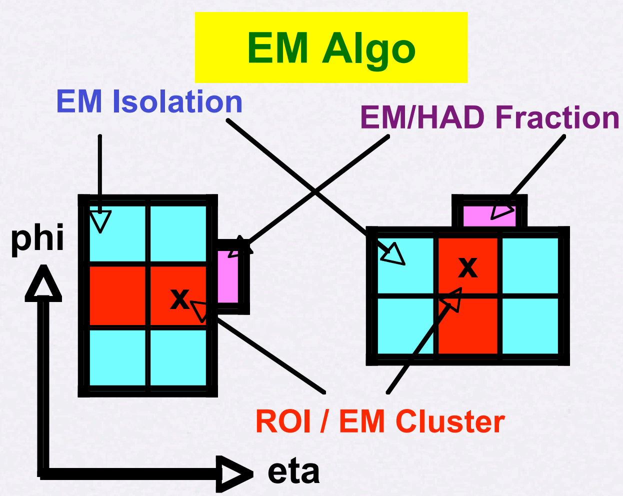


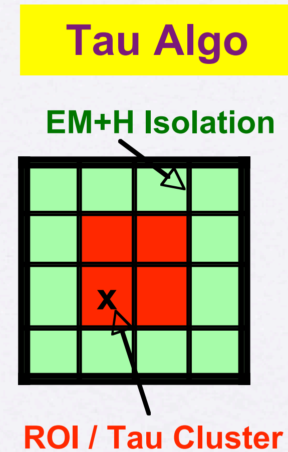
Jets:

- EM+HAD trigger sums
- 2x2 LM
- 4x4 TT geometry

EM Objects:

- 2x1 EM TT sums
- Isolation regions defined by adjacent towers





Taus:

- narrow jet
- ratio of core 2x2 to 4x4 sum
- EM+HAD energies

L1Cal Algorithms

Topological Terms:

- Back-to-Back EM objects
- Acoplanar Jets
- Jet Free Regions

Also missing E_T , scalar E_T
inclusion of ICR in algos