A Search for Standard Model Higgs Boson Production in Association with Top-Quark Pairs At CMS

Dissertation Defense, April 22nd, 2015
Charlottesville, Virginia

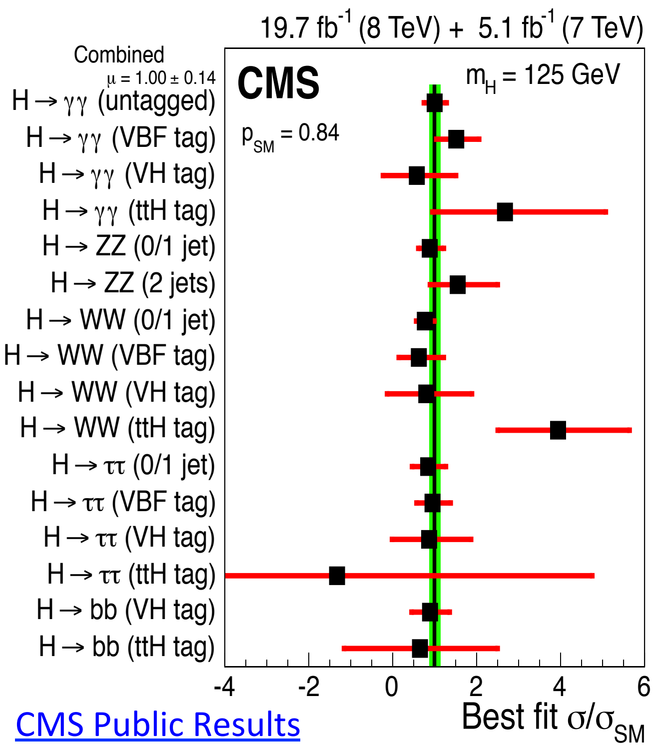
John Wood
University of Virginia
CMS Collaboration

Higgs Discovery at the LHC

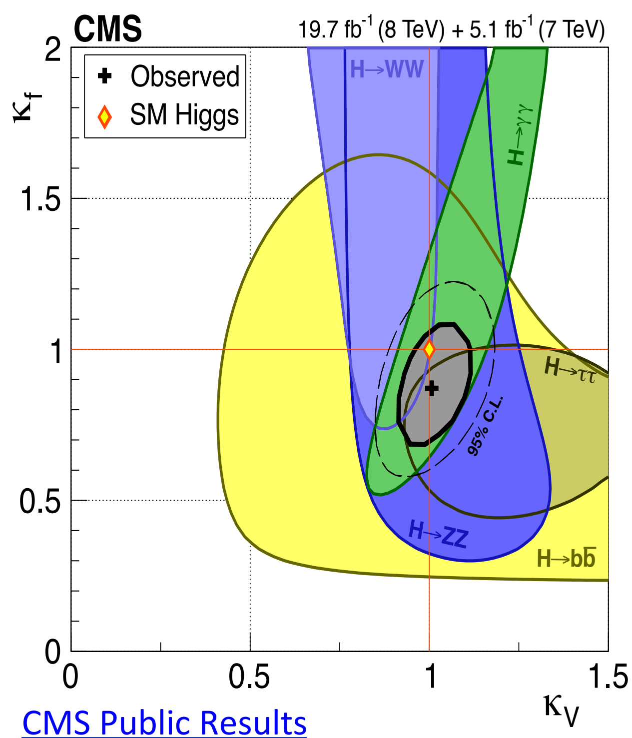
Latest Higgs Results

- Newly Discovered Boson
  - ***Mass of ~125 GeV***, consistent with precision EWK measurements
  - ***Spin 0, positive parity***, consistent with SM Higgs predicted properties
  - ***Coupling to SM leptons and bosons***, consistent with SM Higgs predicted properties

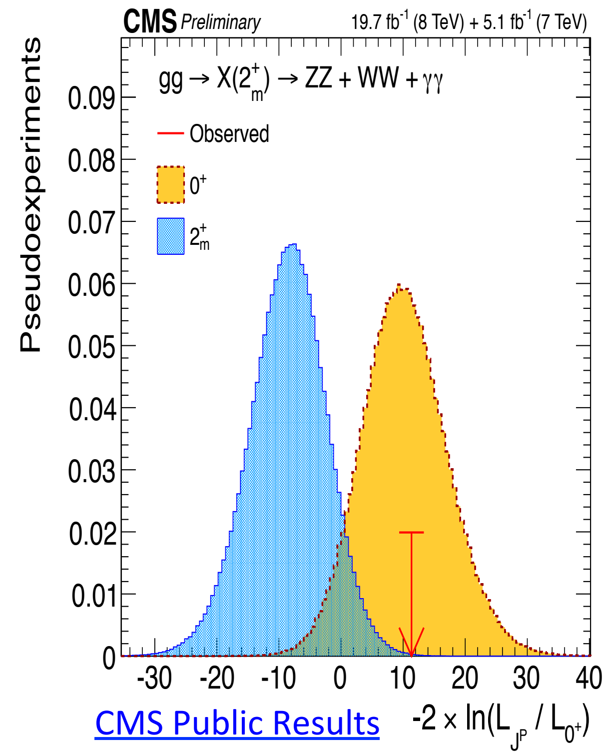
- Significant progress has been made towards the characterization of this new particle, but have yet to measure the coupling to the top-quark



CMS Public Results

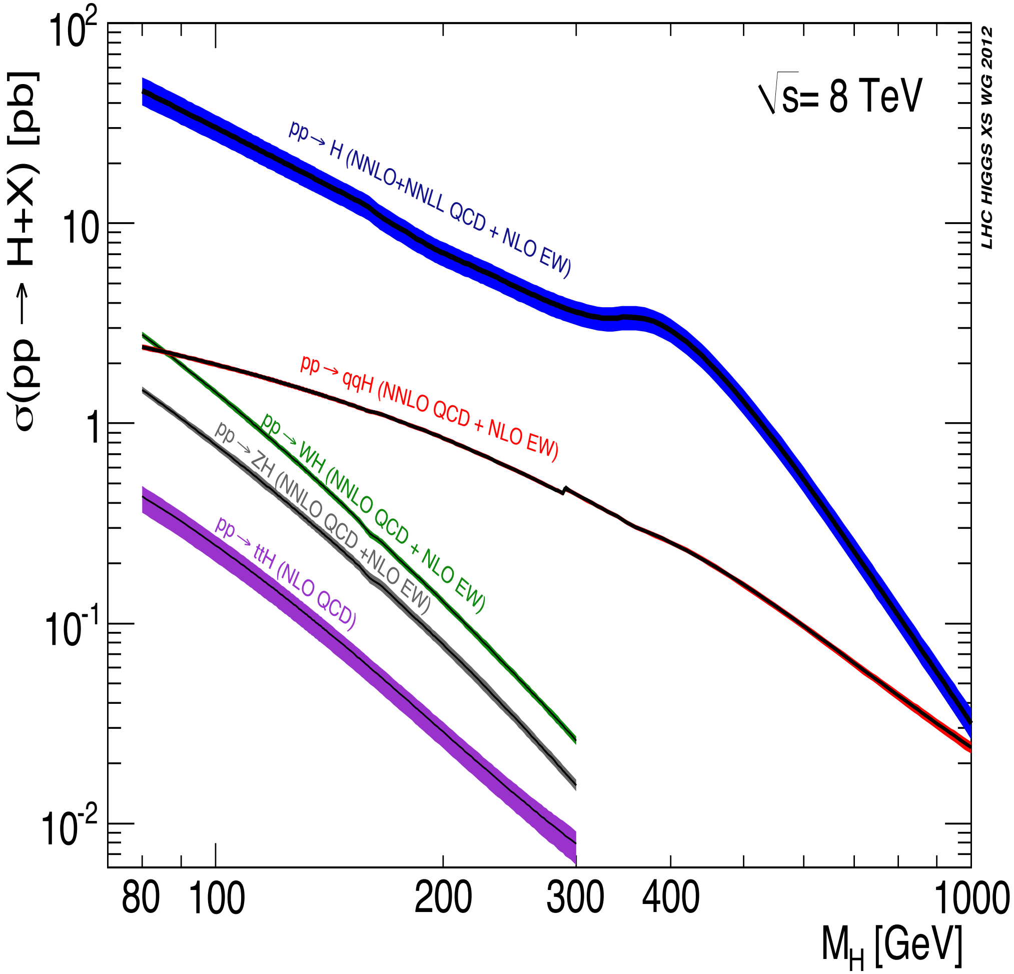


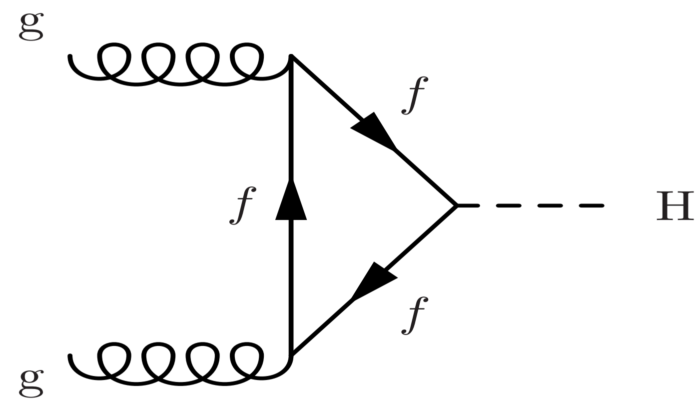
CMS Public Results

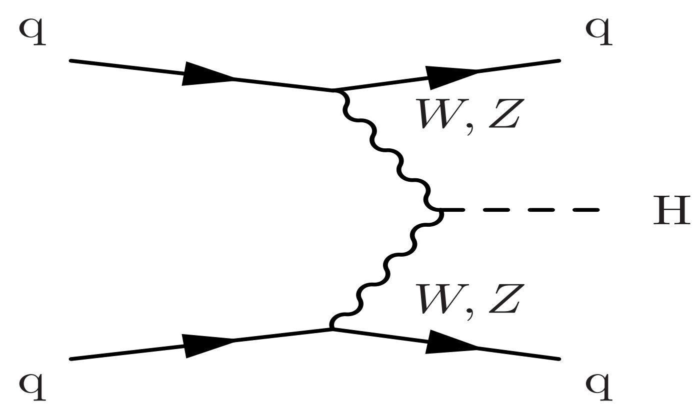


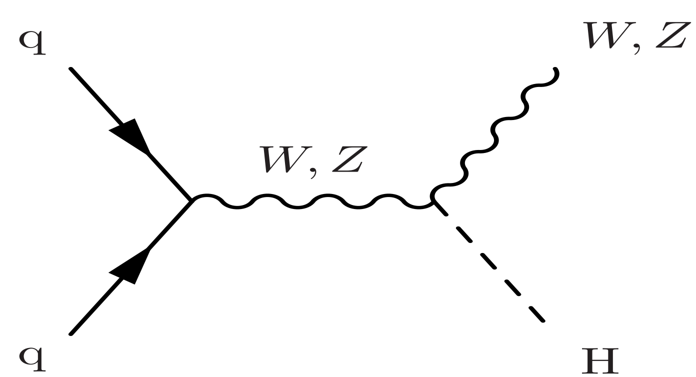
CMS Public Results

Higgs Production at the LHC



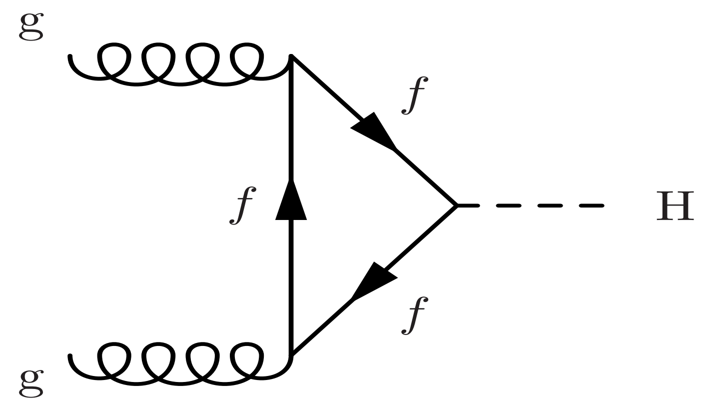


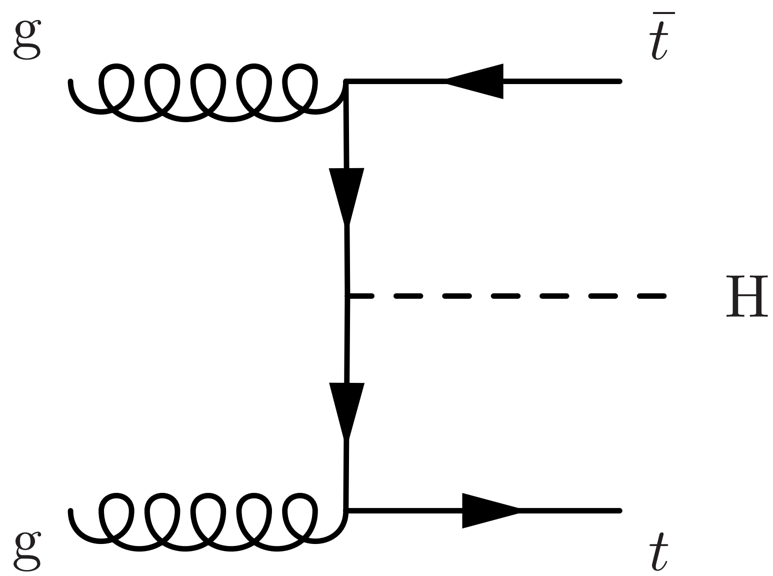




Top-Higgs Interplay

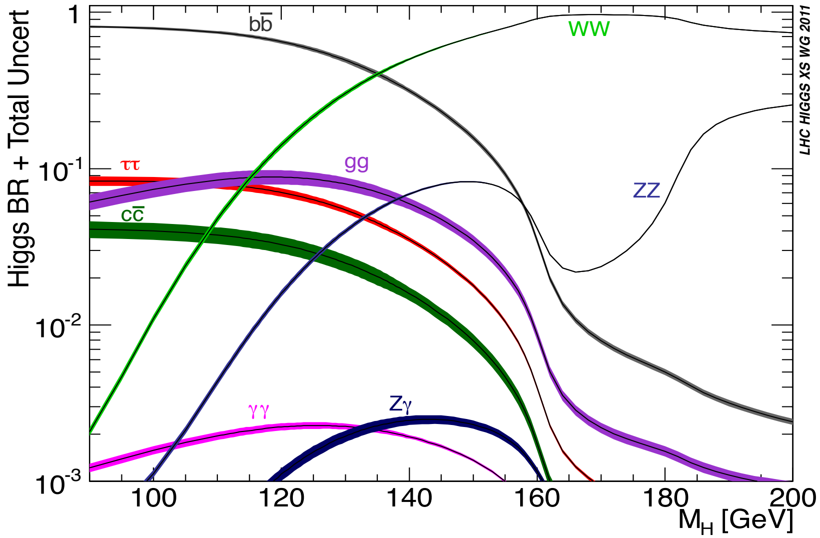
- **Top quark is essential for characterizing the new, recently observed boson**
  - If it is the SM Higgs: need to see coupling to all fermions
  - Large top-quark mass means large coupling to SM Higgs
  - ttH: best production mode **directly** sensitive to top-Higgs Yukawa coupling
  - Gluon-Gluon fusion: triangle loop to produce Higgs has contributions for other SM and potentially BSM particles



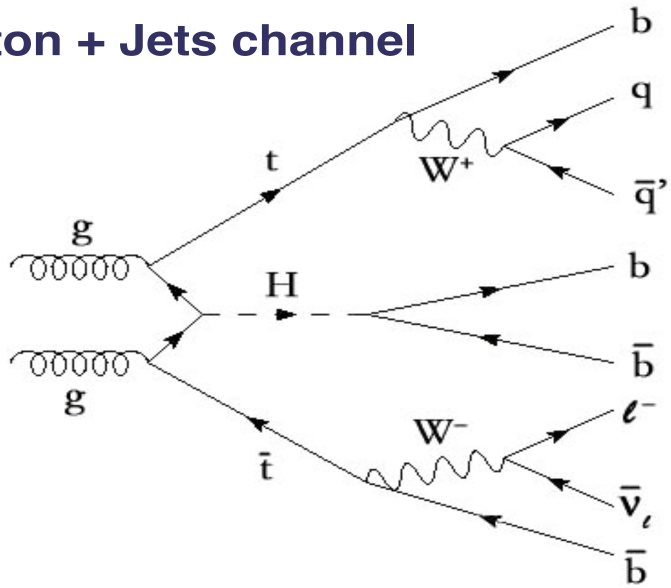


Direct Search for t̄tH

- **Virtues**
  - Access to all Higgs decay modes
    - bb was the obvious place to start
    - WW, ZZ, ττ, γγ are additional possibilities
  - Top quarks have distinctive, characteristic signature
    - Focus initially on final states where one or both top-quarks decay leptonically

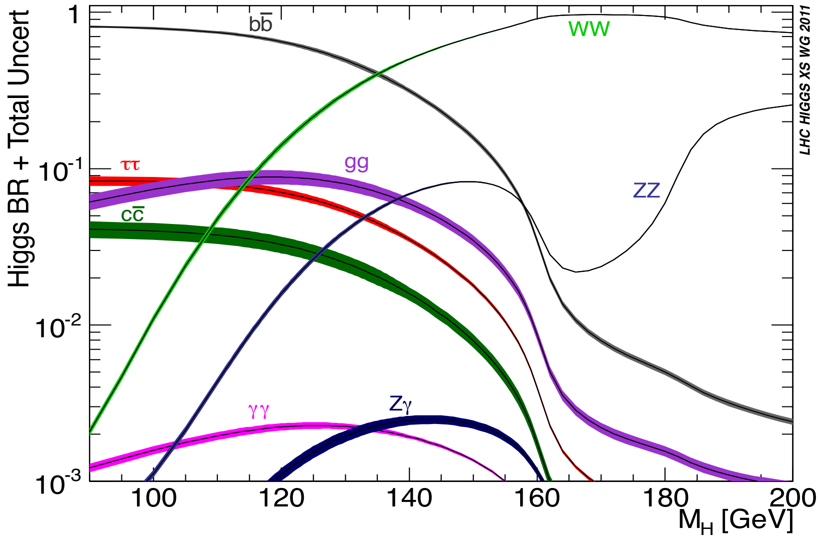


Lepton + Jets channel

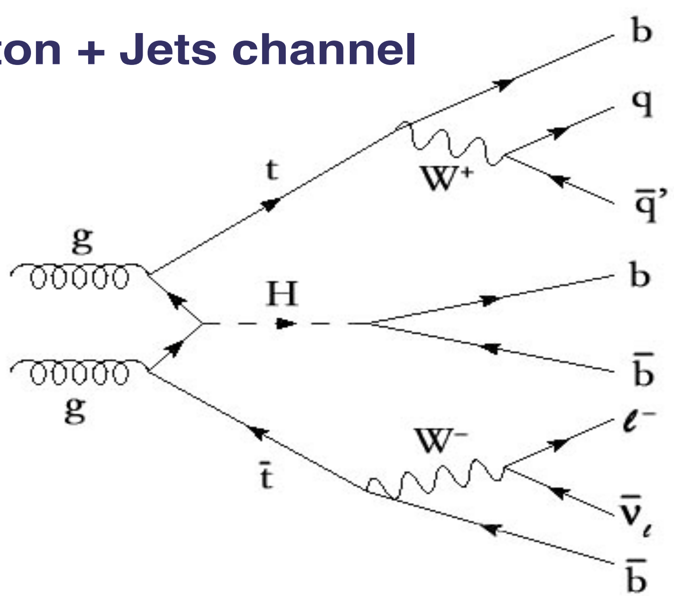


Direct Search for ttH

- **Virtues**
  - Access to all Higgs decay modes
    - bb was the obvious place to start
    - WW, ZZ, ττ, γγ are additional possibilities
  - Top quarks have distinctive, characteristic signature
    - Focus initially on final states where one or both top-quarks decay leptonically
- **Challenges**
  - Low production cross section
    - Signal phase space suffers low statistics
  - Swamped with tt+jets background
  - tt+bb difficult to discriminate
    - Irreducible, poorly known
  - Very busy events
  - Recoil of ttbar system against Higgs
    - Final state is more complex than typical tt+jets system + typical Higgs system



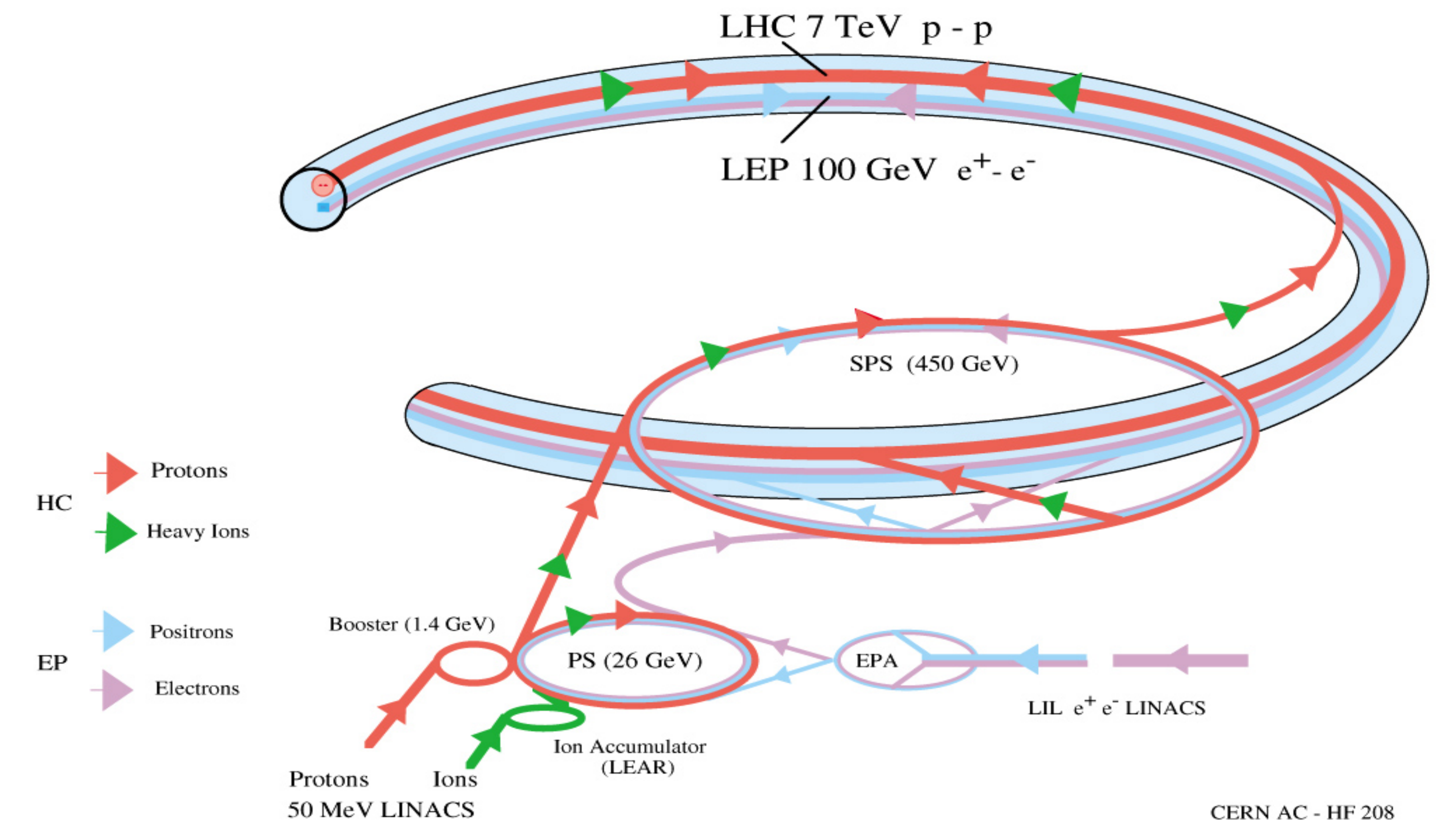
Lepton + Jets channel



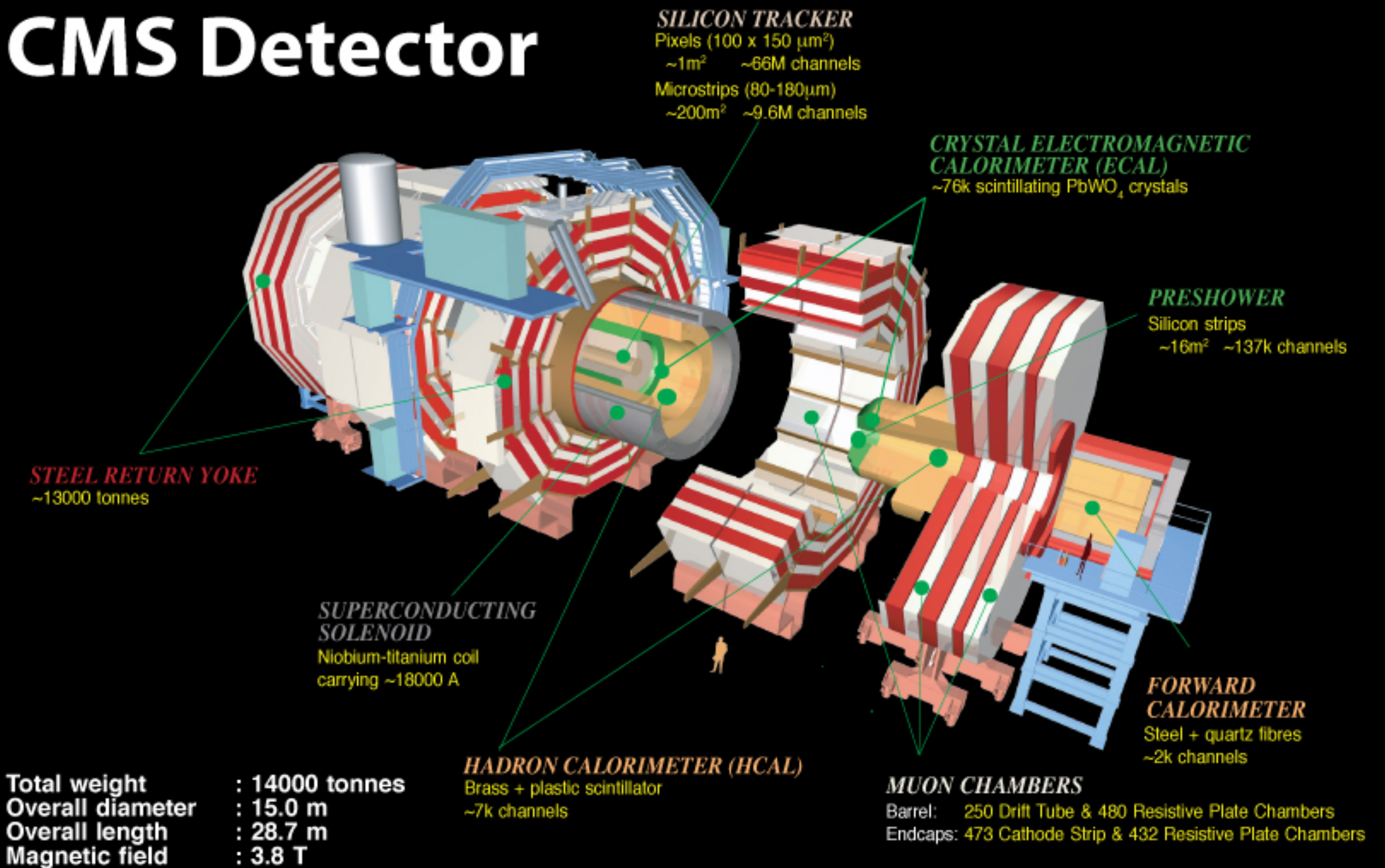
LHC Overview

LHC Overview

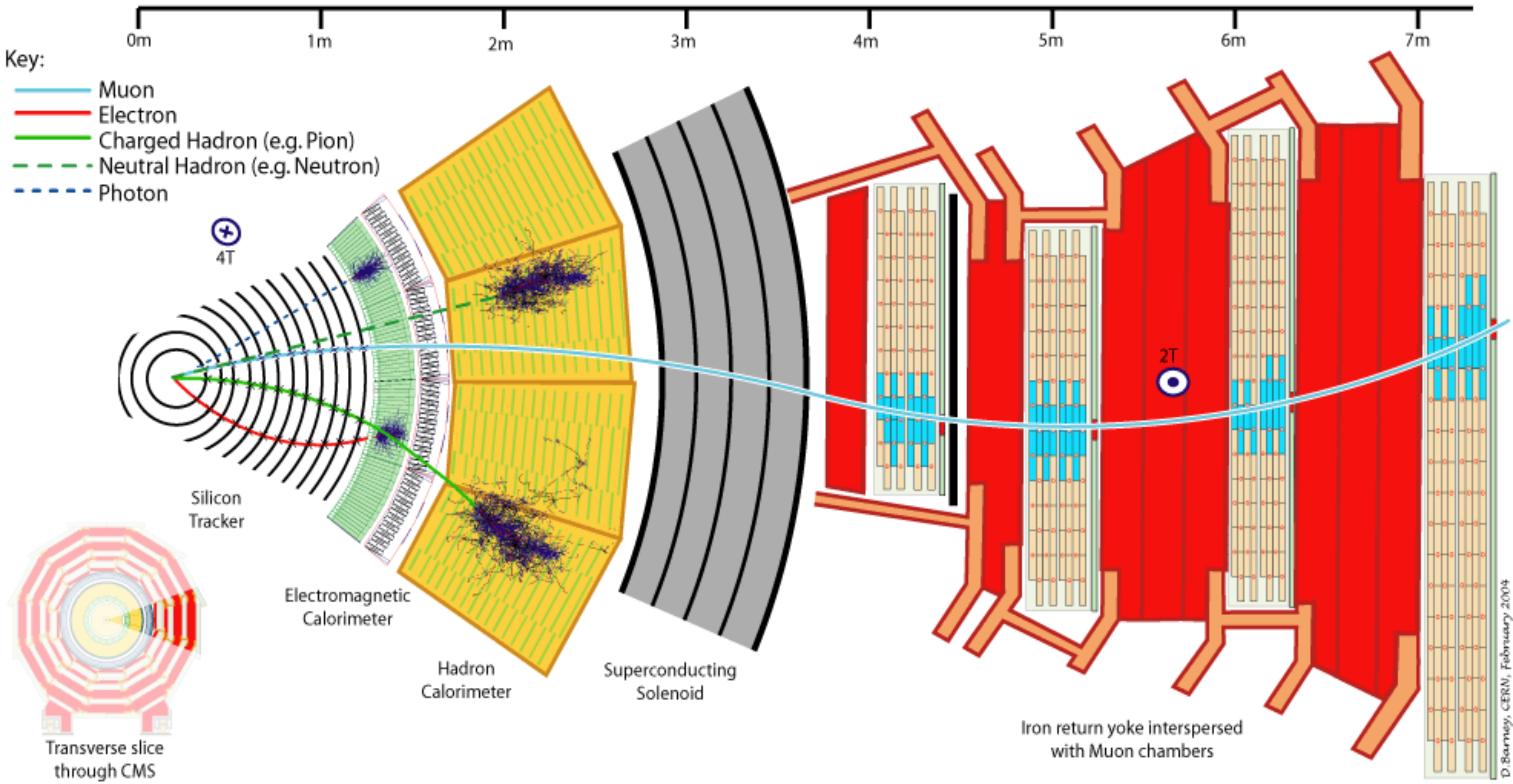
The LHC injection complex



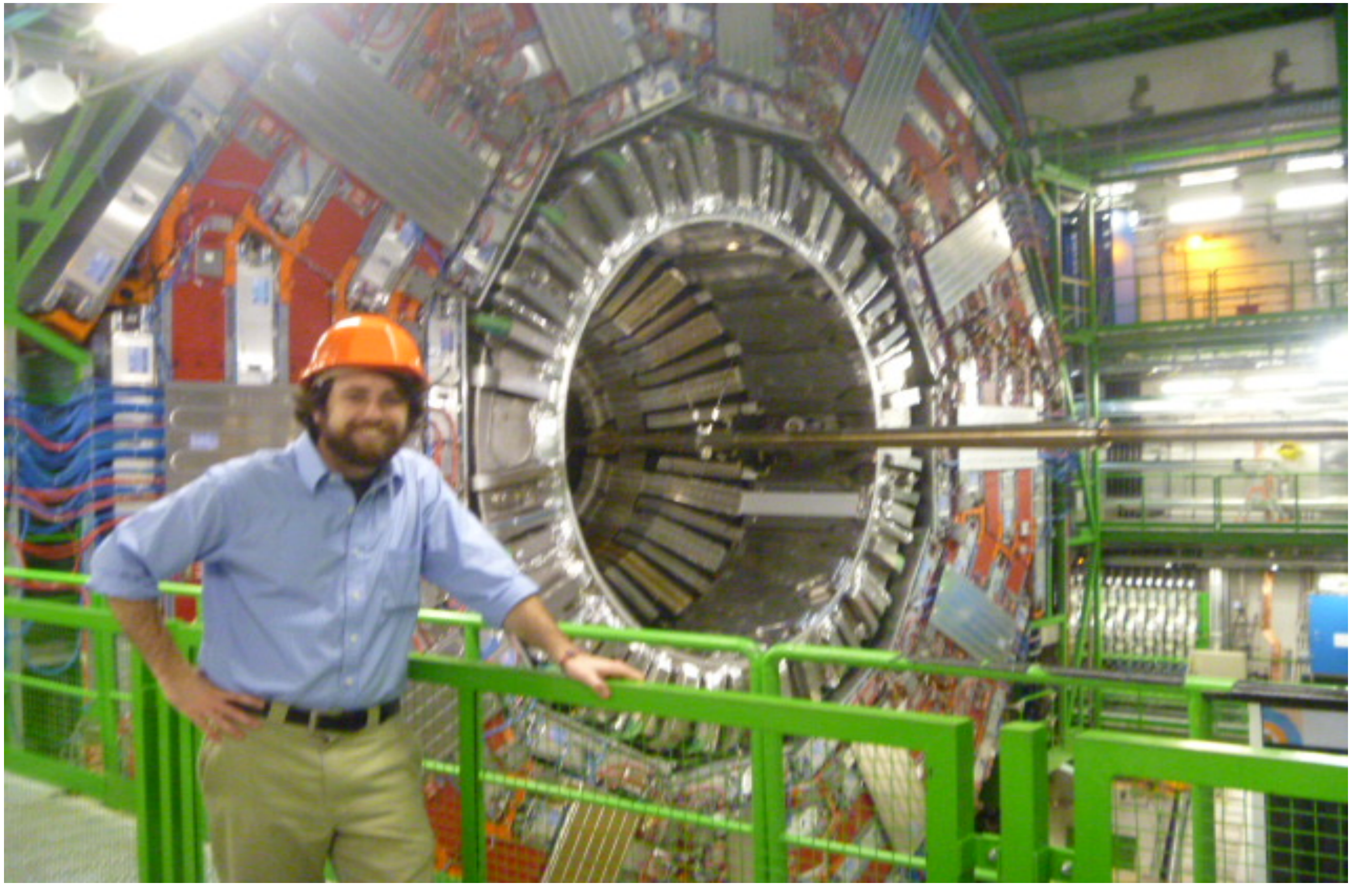
CMS Detector Overview



CMS Detector Overview

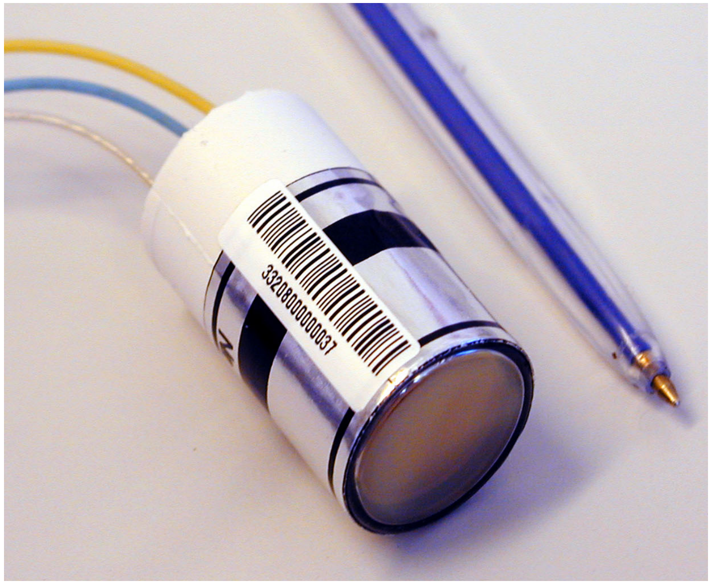


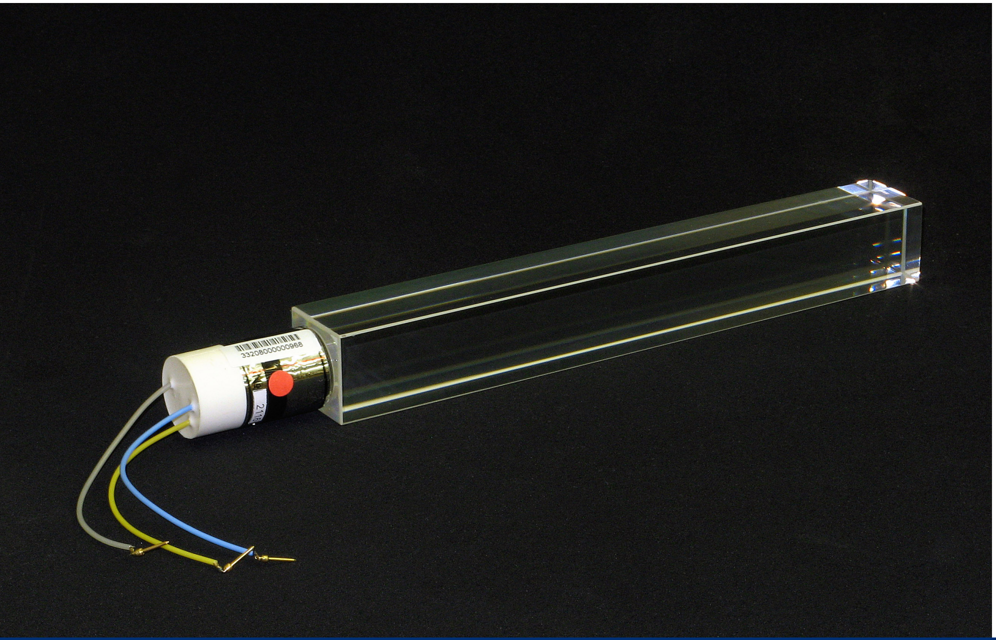
CMS Detector Overview

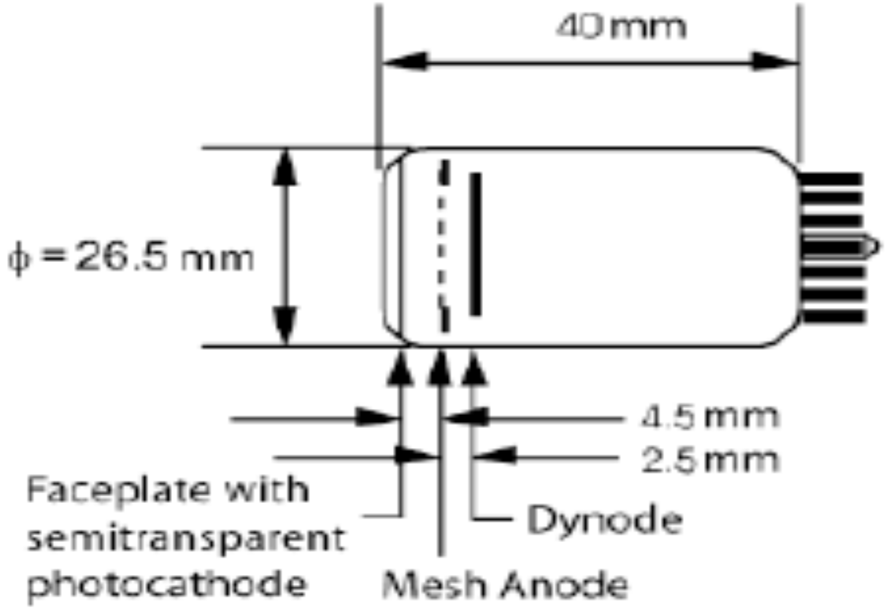


Vacuum Photo-Triodes (VPTs) and the ECAL Endcap

- Vacuum Photo-Triode – single stage photomultiplier device, adhered to back of lead tungstate crystals – measures scintillation light
- Cesium cathode for photoelectron production
- Anode is a mesh grid, captures ~50% of photo-electrons passing through it, 800V
- Dynode is behind anode, 600V, pushes electrons back towards anode for collection







ECAL related sidebar - VPT Test Stand - Description

- VPTs perform well in magnetic fields, and are radiation hard – great for ECAL endcap environment

- Over time VPT anode response decreases, described well by a function of two exponentials with different time constants, modeling the two primary effects of VPT aging
  - Ion damage to cathode due to loss of vacuum and large electric potentials
  - Depletion of photo-electrons from the cathode

- In-situ at CMS, this effect is convoluted with light loss of crystals from radiation damage – necessary to study VPT lifetime performance effects independently

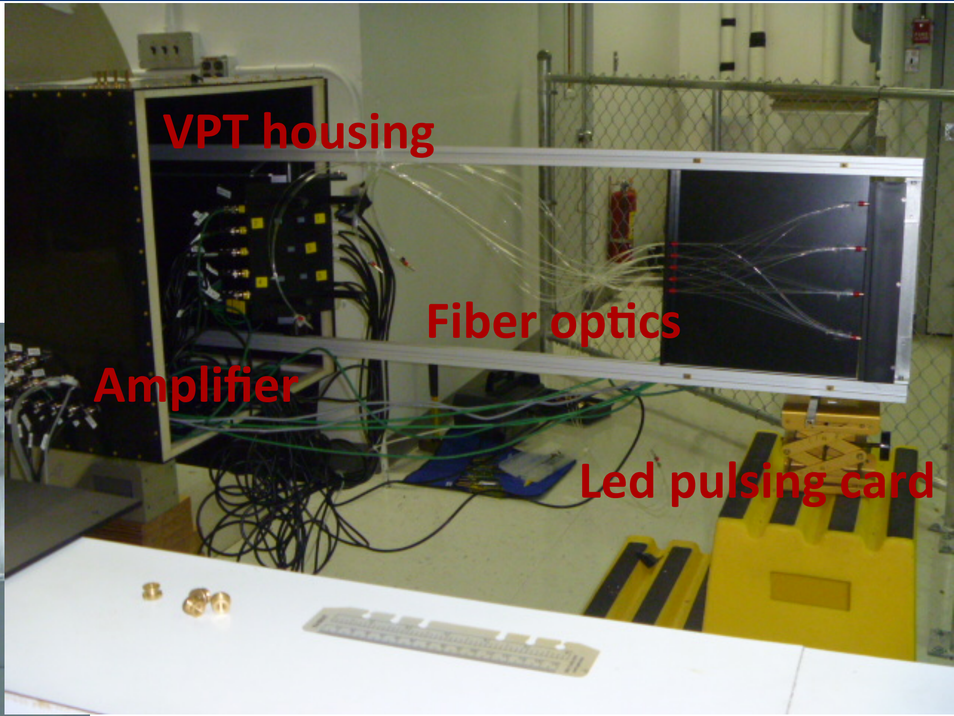
- University of Virginia HEP has access to a large-bore 3.8T solenoid magnet
  - Cryogen cooled – required $LN_{2,}$ LHe fills once a week, and month respectively
  - Test-stand – 5 VPTs tested at a time
    - PIN diodes used as a control for measuring LED source light, signal amplification in housing next to VPTs to prevent noise pickup on small signal output

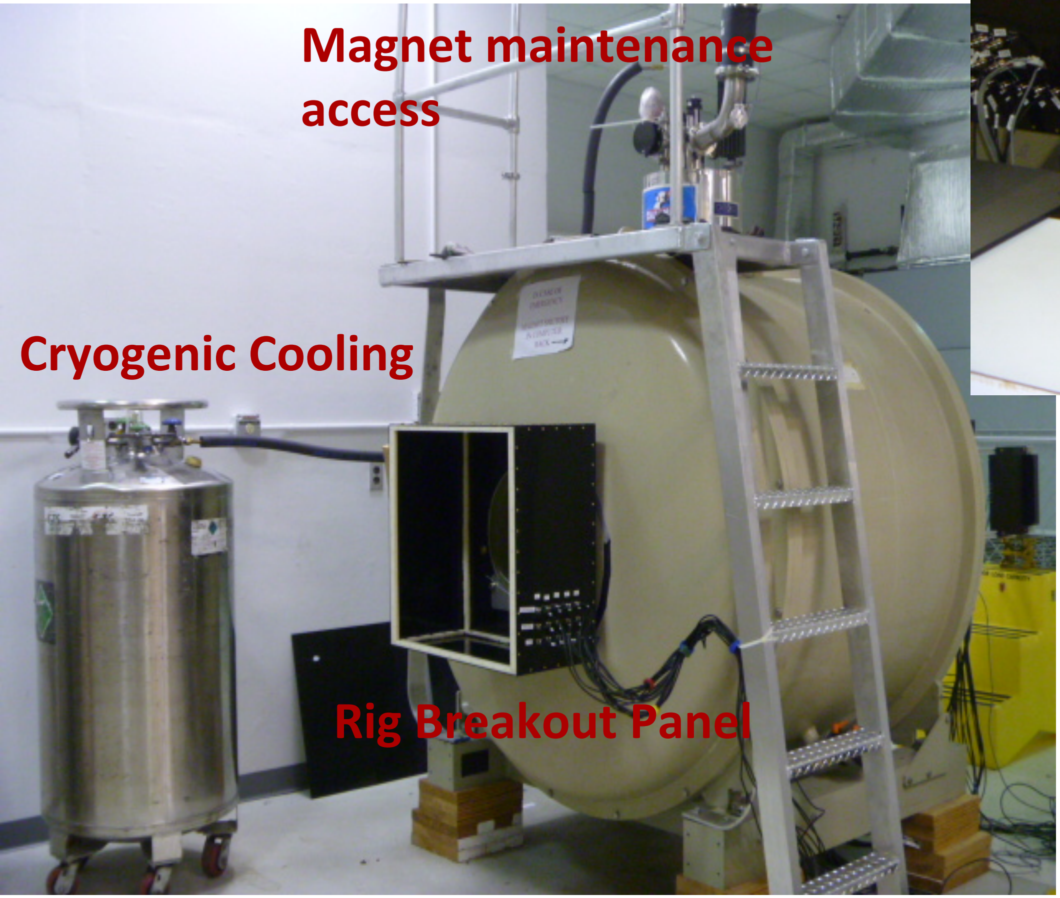
- Three LEDs are used to pulse VPTs
  - Blue Load LED – simulates beam cycle and light exposure from crystal
  - Blue & Orange Reference LEDs – measurement LEDs for VPT performance

- Load light illumination and measurement cycle continue for ~3-4 months per run
  - 6 runs completed at UVa

Magnet and Test Stand

- Below:
  - 3.8T superconducting solenoid magnet
  - Inner bore houses VPTs and amplifier equipment

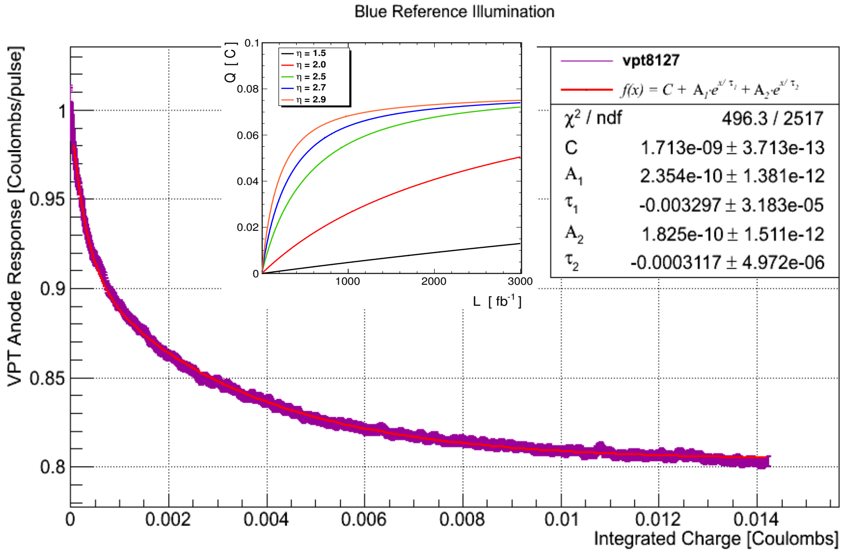


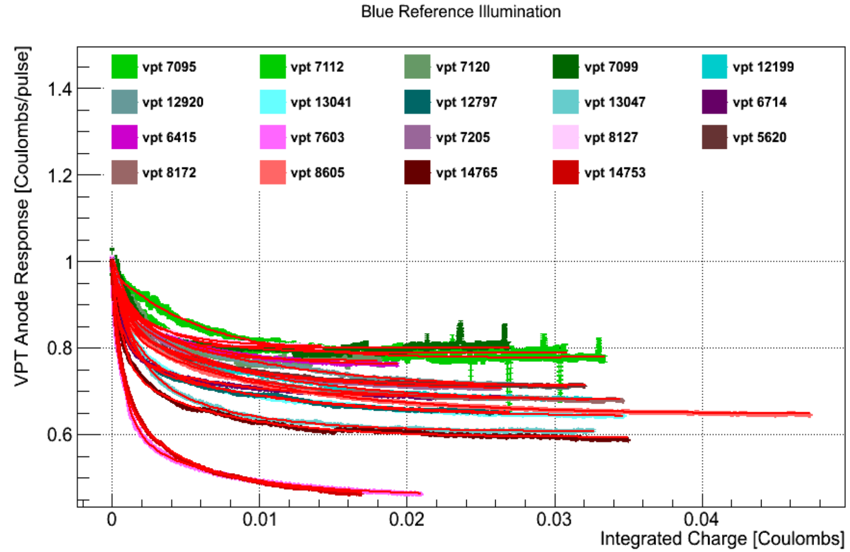


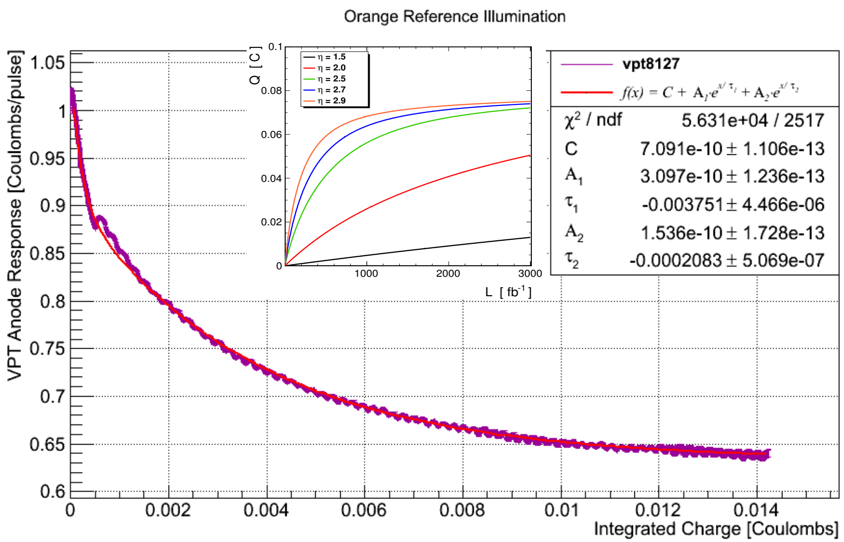
- Above:
  - VPT, amplifier housing, fiber optics, and LED pulsing card, mounted on test stand railing system
  - In picture – maintenance position, outside of bore

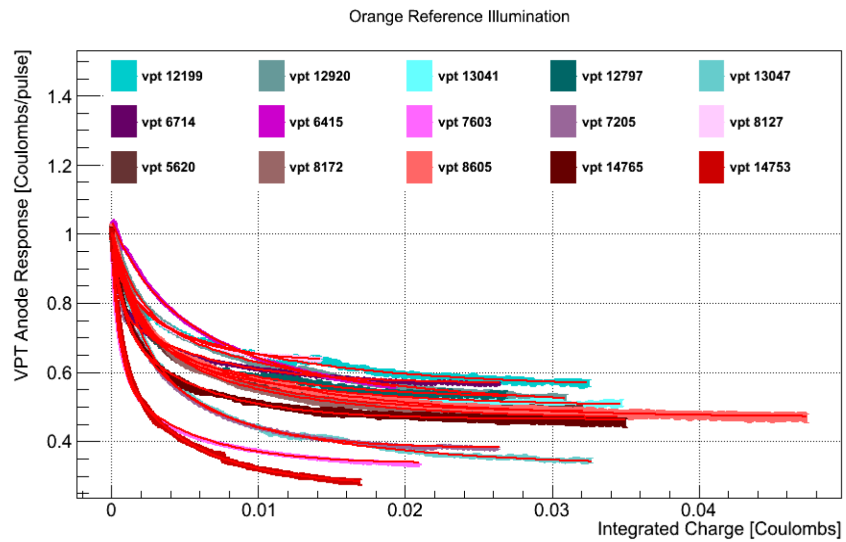
Blue and Orange Reference LED – Fits Summary

❖ NIM paper expected to be published later this year from these results



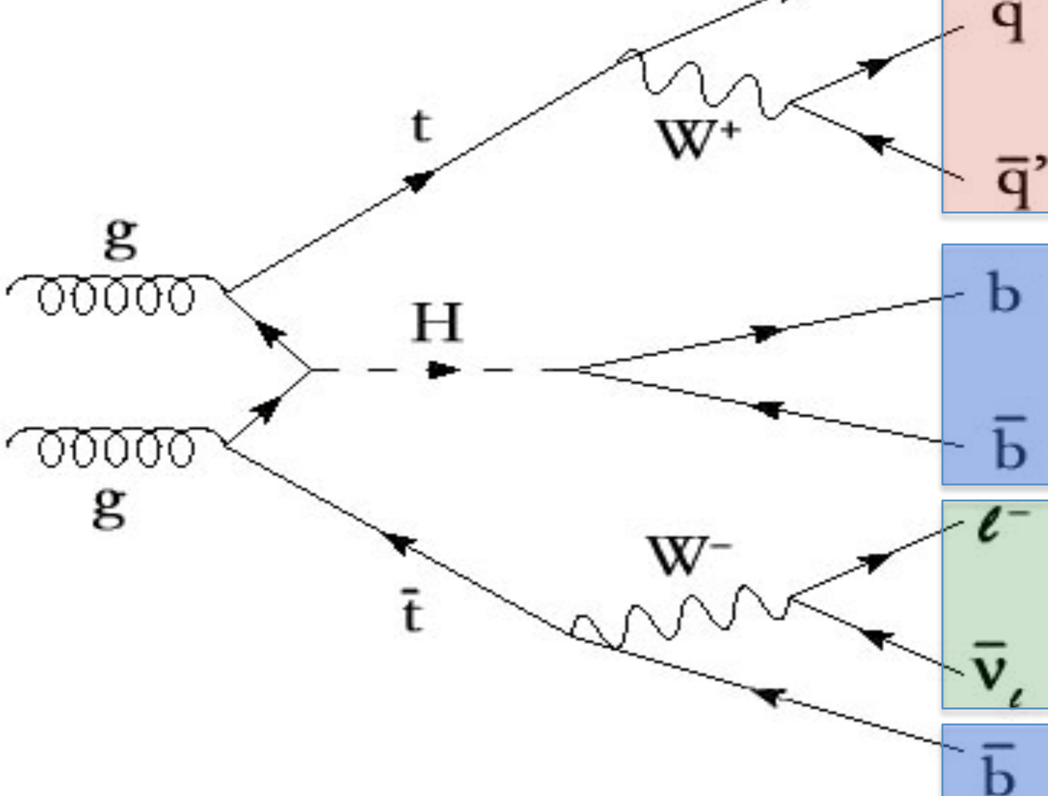






ttH Signature

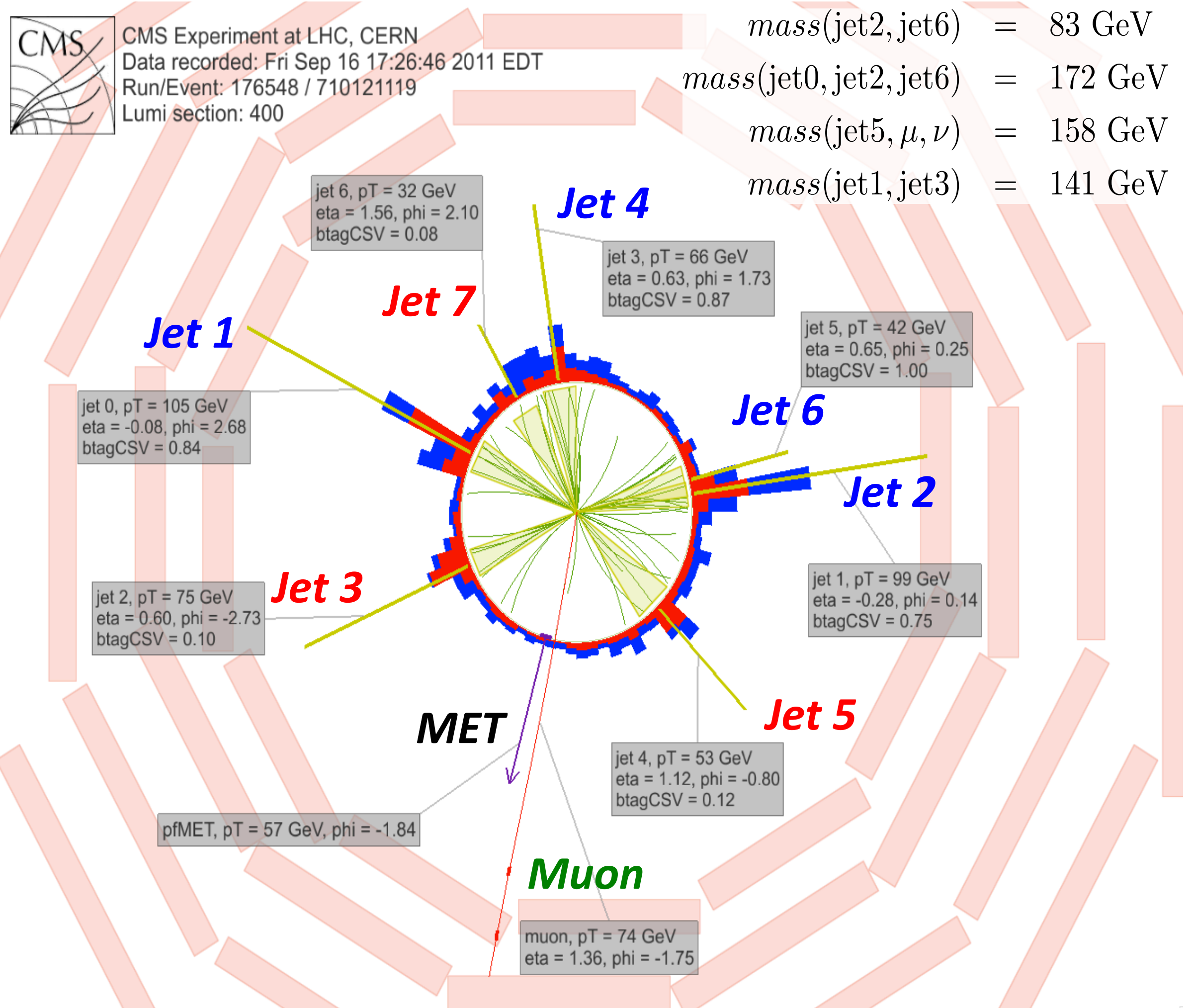
Lepton + Jets channel



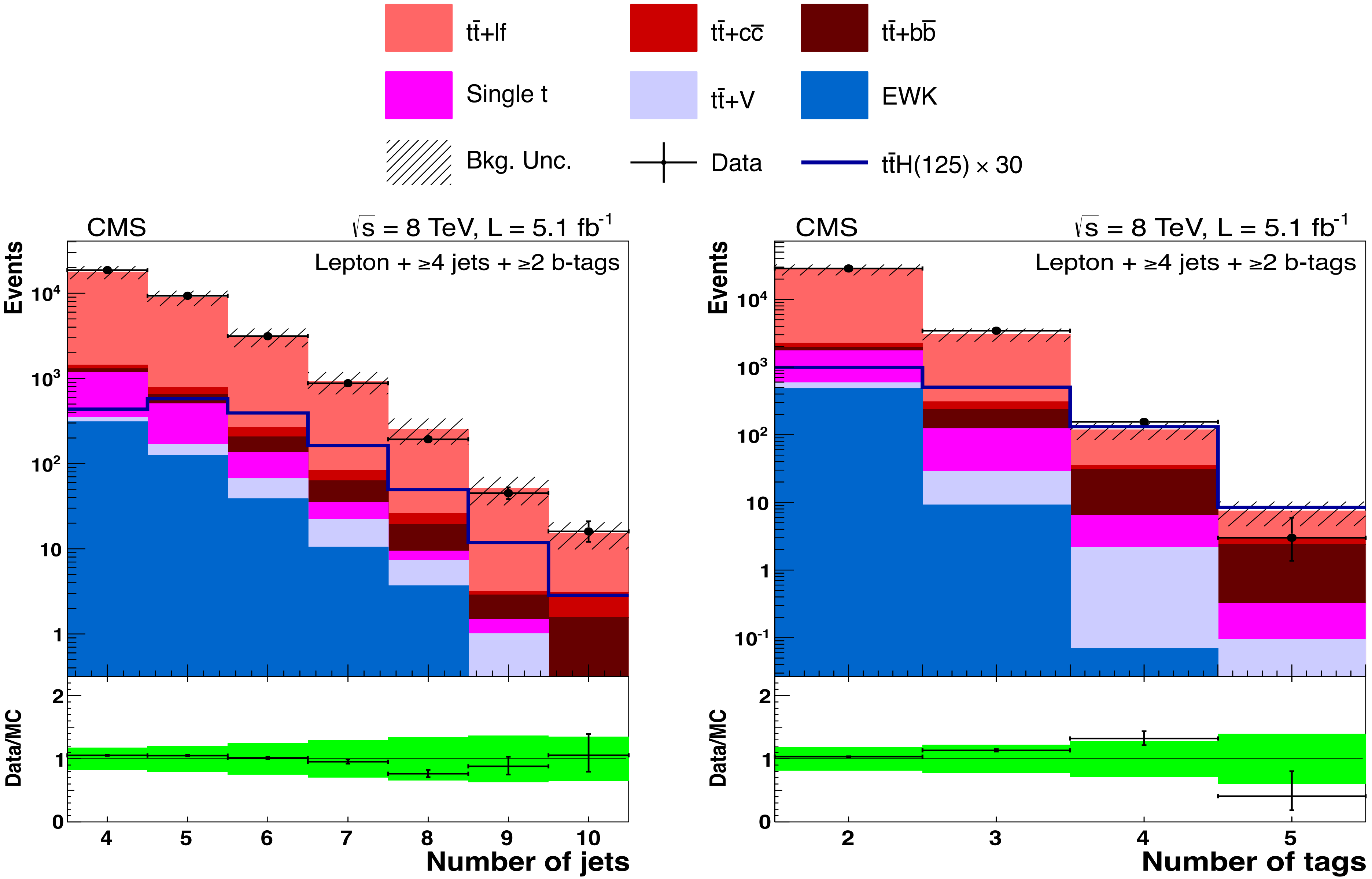
- ❖ b quark from top
- ❖ quarks from hadronic W
- ❖ b quarks from Higgs
- ❖ Lepton and MET from W
- ❖ anti-b quark from anti-top

Event Selection

- **8TeV Data - 2012, 5.1 fb^{-1}**
- **Start from tt-enriched sample:**
  - Event passes single muon/ electron+jets trigger
  - 1 isolated lepton (30 GeV)
  - At least 3 jets (40GeV) plus 1 jet (30GeV)
  - At least 2 b-tagged jets (CSVM)
  - Categorize events into bins of nJets and nTags
    - Signal will occupy higher jet/tag multiplicity bins
    - Background will occupy lower bins



Number of Jets/Tags Distributions



Categorization by Number of Jets/Tags

- S/Sqrt(B) for each jet/tag category
  - Highlighted green region is signal-enriched
- **Lepton+Jets, 7 categories:**
  - 4 jets (3, ≥ 4 tags)
  - 5 jets (3, ≥ 4 tags)
  - >=6 jets (2, 3, ≥ 4 tags)

Number of Jets

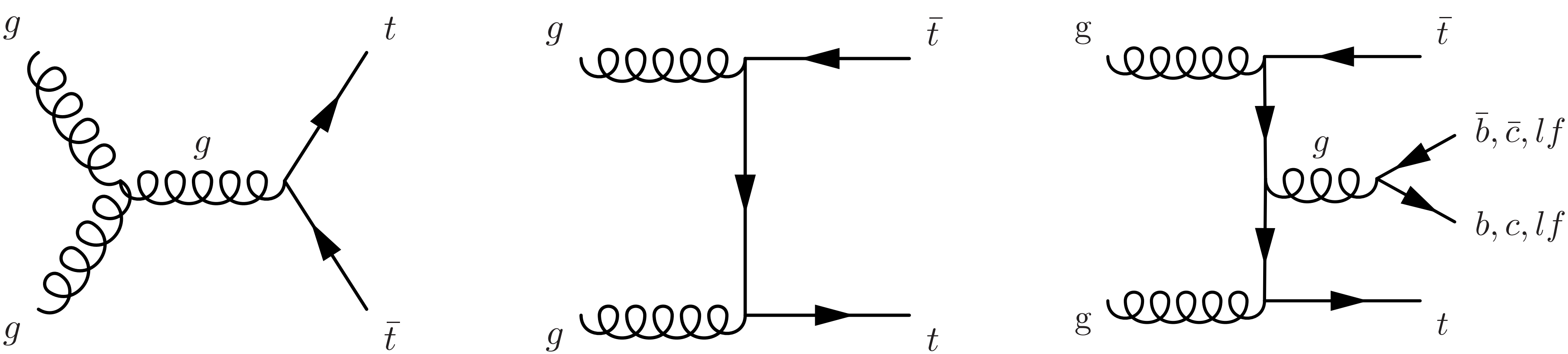
Number of Tags

	4	5	>=6
2	-	-	0.19
3	0.10	0.20	0.27
>=4	0.13	0.23	0.31

Background Processes

- **ttbar** – top-quark pair production, with extra jets coming from ISR/FSR or higher-order processes

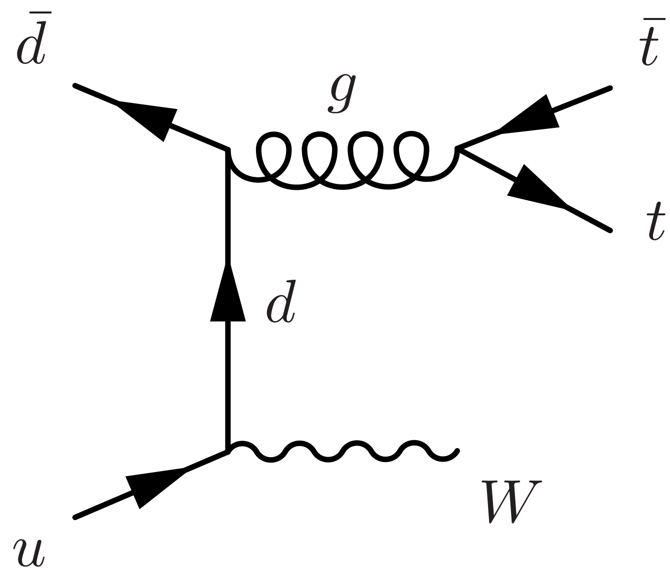
	≥6 jets 2 b-tags	4 jets 3 b-tags	5 jets 3 b-tags	≥6 jets 3 b-tags	4 jets 4 b-tags	5 jets ≥4 b-tags	≥6 jets ≥4 b-tags
$t\bar{t}H(125)$	11.7 ± 1.9	3.9 ± 1.8	6.1 ± 2.8	6.9 ± 3.1	0.6 ± 0.3	1.5 ± 0.7	2.5 ± 1.2
$t\bar{t}$+lf	3460 ± 940	1320 ± 280	870 ± 210	570 ± 170	18.0 ± 5.1	27.6 ± 8.6	41 ± 15
$t\bar{t} + b\bar{b}$	61 ± 34	35 ± 19	43 ± 24	35 ± 20	2.5 ± 1.7	8.4 ± 5.3	15.4 ± 9.4
$t\bar{t} + c\bar{c}$	62 ± 17	19.6 ± 5.1	25.0 ± 6.9	25.9 ± 7.7	0.6 ± 0.4	0.8 ± 0.9	3.7 ± 1.8
tt V	35.7 ± 7.5	4.5 ± 1.1	6.1 ± 1.4	8.6 ± 2.1	0.1 ± 0.1	0.7 ± 0.2	1.5 ± 0.4
Single t	79 ± 18	56 ± 11	25.6 ± 6.2	10.3 ± 2.9	0.3 ± 0.6	3.1 ± 2.2	1.0 ± 0.6
V+jets	53 ± 40	5.9 ± 6.0	0.8 ± 0.9	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Diboson	1.2 ± 0.4	1.8 ± 0.6	0.5 ± 0.2	0.2 ± 0.1	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Total bkg	3760 ± 980	1440 ± 300	970 ± 230	650 ± 190	21.5 ± 6.1	41 ± 12	63 ± 21
Data	3503	1646	1116	686	28	56	74

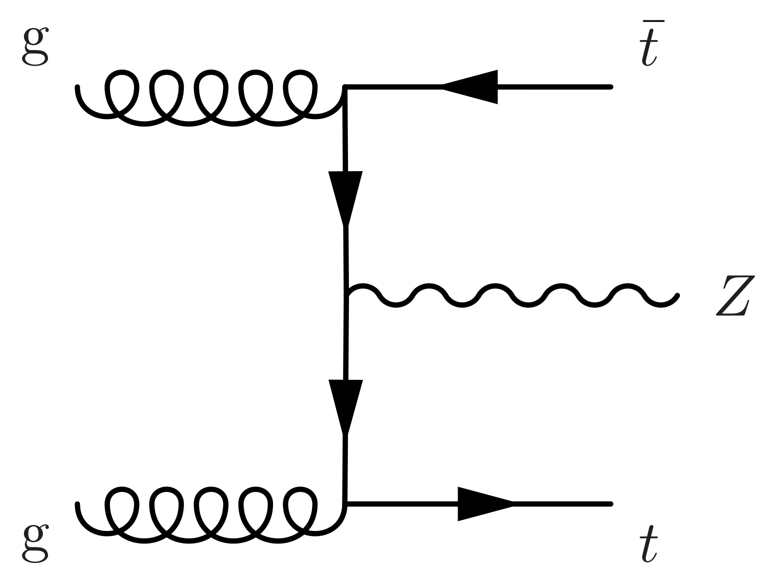


Background Processes

- **ttbar** – top-quark pair production, with extra jets coming from ISR/ FSR or higher-order processes
- **ttbar+V** – top-quark pair-production, with an additional W/Z boson

	≥6 jets 2 b-tags	4 jets 3 b-tags	5 jets 3 b-tags	≥6 jets 3 b-tags	4 jets 4 b-tags	5 jets ≥4 b-tags	≥6 jets ≥4 b-tags
$t\bar{t}H(125)$	11.7 ± 1.9	3.9 ± 1.8	6.1 ± 2.8	6.9 ± 3.1	0.6 ± 0.3	1.5 ± 0.7	2.5 ± 1.2
$t\bar{t}$+lf	3460 ± 940	1320 ± 280	870 ± 210	570 ± 170	18.0 ± 5.1	27.6 ± 8.6	41 ± 15
$t\bar{t} + b\bar{b}$	61 ± 34	35 ± 19	43 ± 24	35 ± 20	2.5 ± 1.7	8.4 ± 5.3	15.4 ± 9.4
$t\bar{t} + c\bar{c}$	62 ± 17	19.6 ± 5.1	25.0 ± 6.9	25.9 ± 7.7	0.6 ± 0.4	0.8 ± 0.9	3.7 ± 1.8
$t\bar{t}V$	35.7 ± 7.5	4.5 ± 1.1	6.1 ± 1.4	8.6 ± 2.1	0.1 ± 0.1	0.7 ± 0.2	1.5 ± 0.4
Single t	79 ± 18	56 ± 11	25.6 ± 6.2	10.3 ± 2.9	0.3 ± 0.6	3.1 ± 2.2	1.0 ± 0.6
V+jets	53 ± 40	5.9 ± 6.0	0.8 ± 0.9	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Diboson	1.2 ± 0.4	1.8 ± 0.6	0.5 ± 0.2	0.2 ± 0.1	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Total bkg	3760 ± 980	1440 ± 300	970 ± 230	650 ± 190	21.5 ± 6.1	41 ± 12	63 ± 21
Data	3503	1646	1116	686	28	56	74

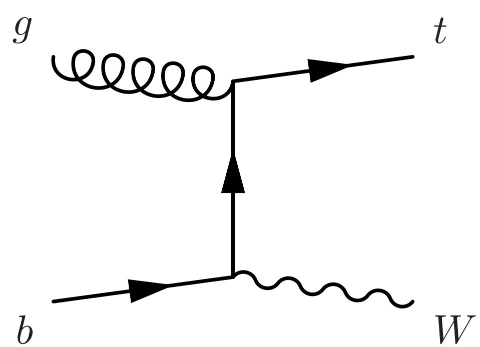


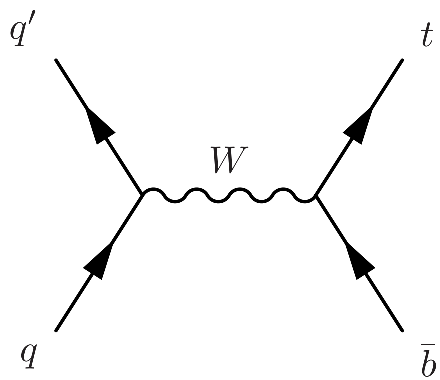


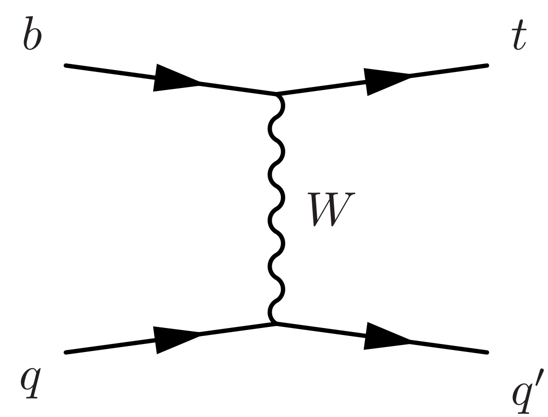
Background Processes

- **ttbar** – top-quark pair production, with extra jets coming from ISR/FSR or higher-order processes
- **ttbar+V** – top-quark pair-production, with an additional W/Z boson
- **Single t** – top quark produced in association of extra jets, or W boson

	$\geq$6 jets 2 b-tags	4 jets 3 b-tags	5 jets 3 b-tags	$\geq$6 jets 3 b-tags	4 jets 4 b-tags	5 jets $\geq$4 b-tags	$\geq$6 jets $\geq$4 b-tags
$t\bar{t}H(125)$	11.7 ± 1.9	3.9 ± 1.8	6.1 ± 2.8	6.9 ± 3.1	0.6 ± 0.3	1.5 ± 0.7	2.5 ± 1.2
$t\bar{t}$+lf	3460 ± 940	1320 ± 280	870 ± 210	570 ± 170	18.0 ± 5.1	27.6 ± 8.6	41 ± 15
$t\bar{t} + b\bar{b}$	61 ± 34	35 ± 19	43 ± 24	35 ± 20	2.5 ± 1.7	8.4 ± 5.3	15.4 ± 9.4
$t\bar{t} + c\bar{c}$	62 ± 17	19.6 ± 5.1	25.0 ± 6.9	25.9 ± 7.7	0.6 ± 0.4	0.8 ± 0.9	3.7 ± 1.8
$t\bar{t}\,V$	35.7 ± 7.5	4.5 ± 1.1	6.1 ± 1.4	8.6 ± 2.1	0.1 ± 0.1	0.7 ± 0.2	1.5 ± 0.4
Single t	79 ± 18	56 ± 11	25.6 ± 6.2	10.3 ± 2.9	0.3 ± 0.6	3.1 ± 2.2	1.0 ± 0.6
V+jets	53 ± 40	5.9 ± 6.0	0.8 ± 0.9	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Diboson	1.2 ± 0.4	1.8 ± 0.6	0.5 ± 0.2	0.2 ± 0.1	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Total bkg	3760 ± 980	1440 ± 300	970 ± 230	650 ± 190	21.5 ± 6.1	41 ± 12	63 ± 21
Data	3503	1646	1116	686	28	56	74



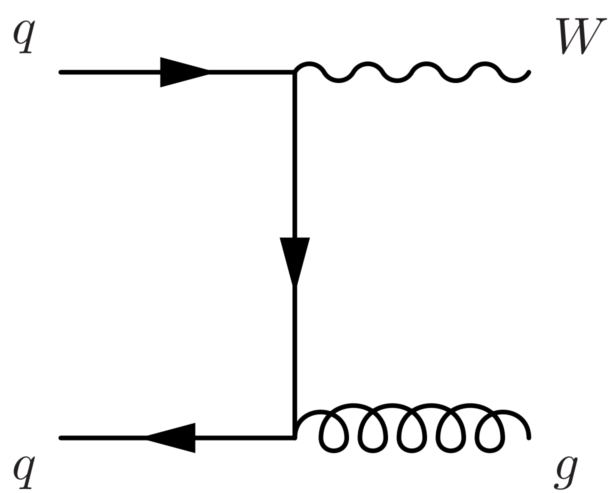


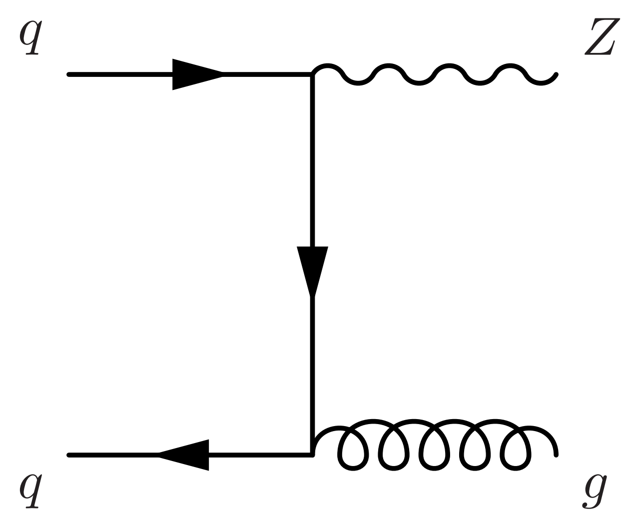


Background Processes

- **ttbar** – top-quark pair production, with extra jets coming from ISR/FSR or higher-order processes
- **ttbar+V** – top-quark pair-production, with an additional W/Z boson
- **Single t** – top quark produced in association of extra jets, or W boson
- **V+Jets** – EWK production of W/Z bosons with extra jets

	≥6 jets 2 b-tags	4 jets 3 b-tags	5 jets 3 b-tags	≥6 jets 3 b-tags	4 jets 4 b-tags	5 jets ≥4 b-tags	≥6 jets ≥4 b-tags
$t\bar{t}H(125)$	11.7 ± 1.9	3.9 ± 1.8	6.1 ± 2.8	6.9 ± 3.1	0.6 ± 0.3	1.5 ± 0.7	2.5 ± 1.2
$t\bar{t}$+lf	3460 ± 940	1320 ± 280	870 ± 210	570 ± 170	18.0 ± 5.1	27.6 ± 8.6	41 ± 15
$t\bar{t} + b\bar{b}$	61 ± 34	35 ± 19	43 ± 24	35 ± 20	2.5 ± 1.7	8.4 ± 5.3	15.4 ± 9.4
$t\bar{t} + c\bar{c}$	62 ± 17	19.6 ± 5.1	25.0 ± 6.9	25.9 ± 7.7	0.6 ± 0.4	0.8 ± 0.9	3.7 ± 1.8
$t\bar{t}$ V	35.7 ± 7.5	4.5 ± 1.1	6.1 ± 1.4	8.6 ± 2.1	0.1 ± 0.1	0.7 ± 0.2	1.5 ± 0.4
Single t	79 ± 18	56 ± 11	25.6 ± 6.2	10.3 ± 2.9	0.3 ± 0.6	3.1 ± 2.2	1.0 ± 0.6
V+jets	53 ± 40	5.9 ± 6.0	0.8 ± 0.9	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Diboson	1.2 ± 0.4	1.8 ± 0.6	0.5 ± 0.2	0.2 ± 0.1	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Total bkg	3760 ± 980	1440 ± 300	970 ± 230	650 ± 190	21.5 ± 6.1	41 ± 12	63 ± 21
Data	3503	1646	1116	686	28	56	74

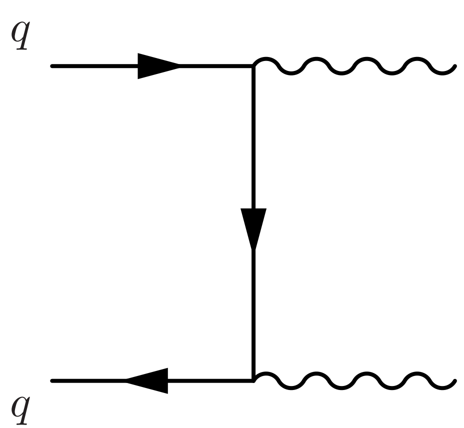


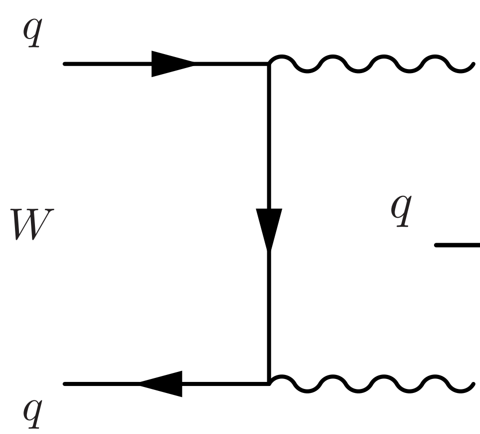


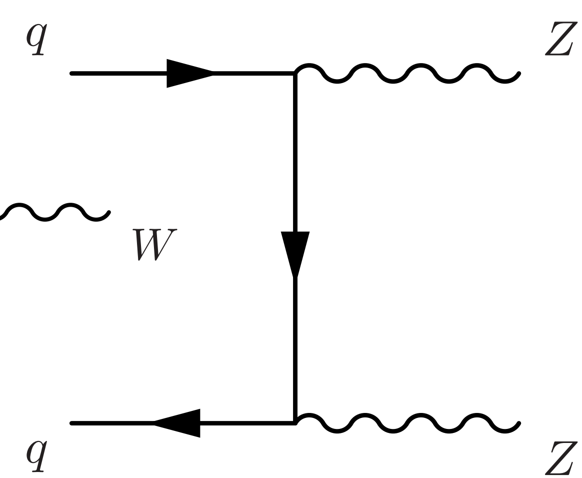
Background Processes

- **ttbar** – top-quark pair production, with extra jets coming from ISR/FSR or higher-order processes
- **ttbar+V** – top-quark pair-production, with an additional W/Z boson
- **Single t** – top quark produced in association of extra jets, or W boson
- **V+Jets** – EWK production of W/Z bosons with extra jets
- **diBoson** – EWK production of WW/WZ/ZZ pairs

	$\geq$6 jets 2 b-tags	4 jets 3 b-tags	5 jets 3 b-tags	$\geq$6 jets 3 b-tags	4 jets 4 b-tags	5 jets $\geq$4 b-tags	$\geq$6 jets $\geq$4 b-tags
$t\bar{t}H(125)$	11.7 ± 1.9	3.9 ± 1.8	6.1 ± 2.8	6.9 ± 3.1	0.6 ± 0.3	1.5 ± 0.7	2.5 ± 1.2
$t\bar{t}$+lf	3460 ± 940	1320 ± 280	870 ± 210	570 ± 170	18.0 ± 5.1	27.6 ± 8.6	41 ± 15
$t\bar{t} + b\bar{b}$	61 ± 34	35 ± 19	43 ± 24	35 ± 20	2.5 ± 1.7	8.4 ± 5.3	15.4 ± 9.4
$t\bar{t} + c\bar{c}$	62 ± 17	19.6 ± 5.1	25.0 ± 6.9	25.9 ± 7.7	0.6 ± 0.4	0.8 ± 0.9	3.7 ± 1.8
$t\bar{t}\,V$	35.7 ± 7.5	4.5 ± 1.1	6.1 ± 1.4	8.6 ± 2.1	0.1 ± 0.1	0.7 ± 0.2	1.5 ± 0.4
Single t	79 ± 18	56 ± 11	25.6 ± 6.2	10.3 ± 2.9	0.3 ± 0.6	3.1 ± 2.2	1.0 ± 0.6
V+jets	53 ± 40	5.9 ± 6.0	0.8 ± 0.9	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Diboson	1.2 ± 0.4	1.8 ± 0.6	0.5 ± 0.2	0.2 ± 0.1	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Total bkg	3760 ± 980	1440 ± 300	970 ± 230	650 ± 190	21.5 ± 6.1	41 ± 12	63 ± 21
Data	3503	1646	1116	686	28	56	74

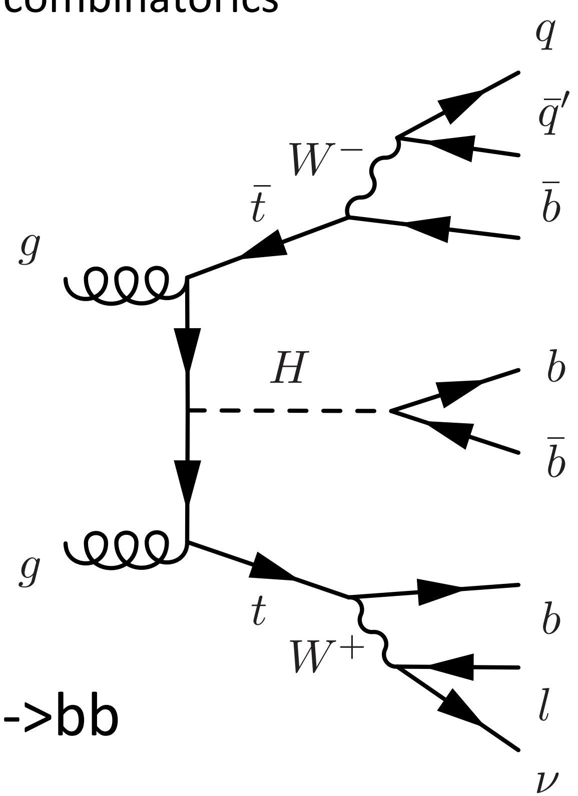




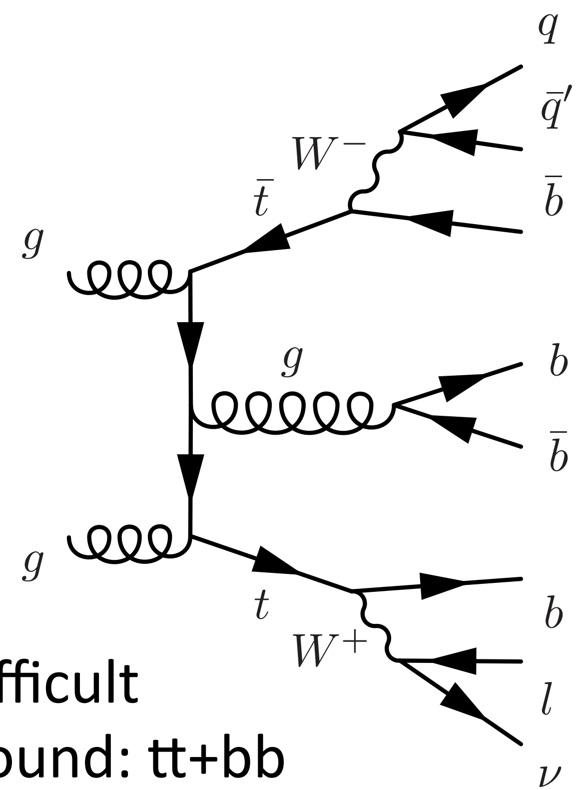


Background Processes

- **Biggest Challenges**
  - Overwehelmingly large background contributions from tt+jets
  - tt+bb background, since if resembles our signal very closely
  - Large number of objects in event also make it difficult to distinguish which jet came from which decay, ie finding the two bTagged jets associated with Higgs
  - Attempts to form invariant mass of of Higgs from two bTagged jets gets smeared out from combinatorics



Signal:
ttH, H->bb

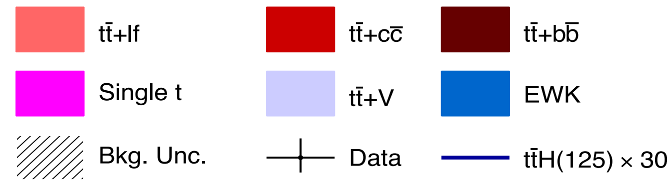


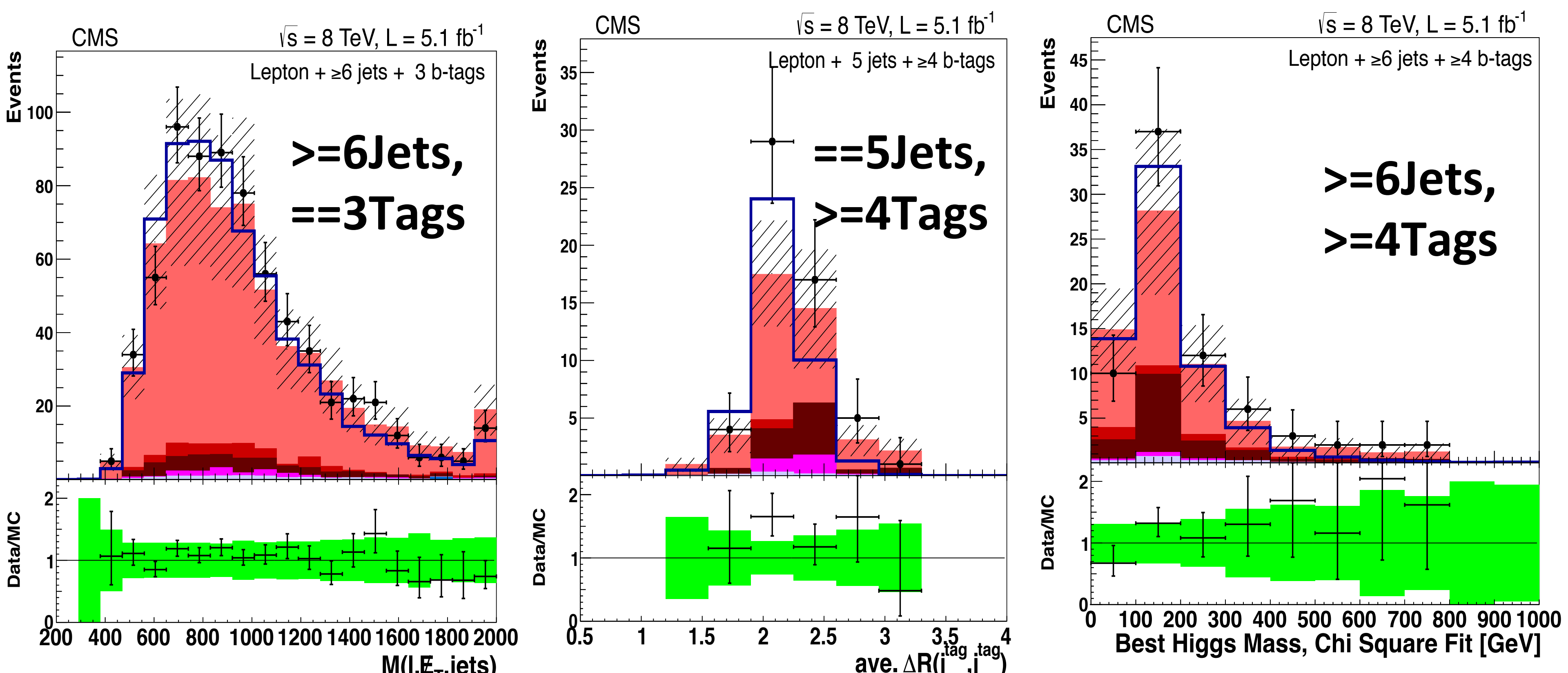
Most difficult
background: tt+bb

Searching for Discriminating Variables

- High Jet multiplicity gives many possible combinations to form the invariant mass of the Higgs system, smears mass resolution
  - no single variable with excellent discrimination against background

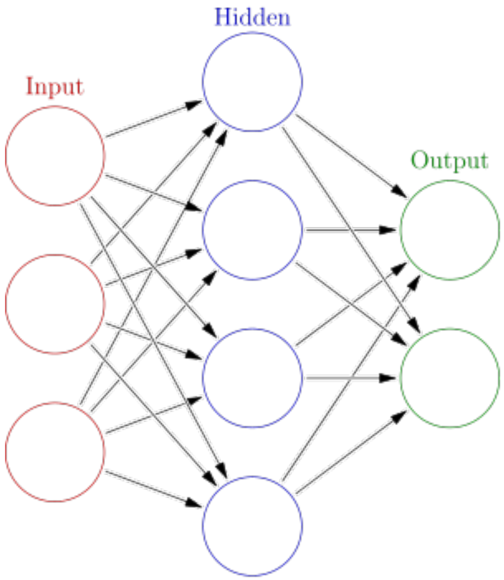
- Signal normalized to sum of bkg





Multi-Variate Analysis Techniques

- **Artificial Neural Network (ANN)** – specific implementation of an MVA.
  - The values of the input variables are assigned to a "node"
  - **Nodes** are independent functions that map an input value onto an output value, which is determined by an **Activation Function** – a weighted response to the input value, also known as the **Synapse Response**
  - Nodes are arranged in layers, and each node in a layer only connects to the nodes in the layer in front of it, this is known as a **Multi-Layer Perceptron**
  - **Training events** are given to the ANN from a known signal and background source, ANN output is compared to desired result (0=background, 1=signal) and adjusts weights to try and maximize output towards signal/background-like response



"Activation Function"

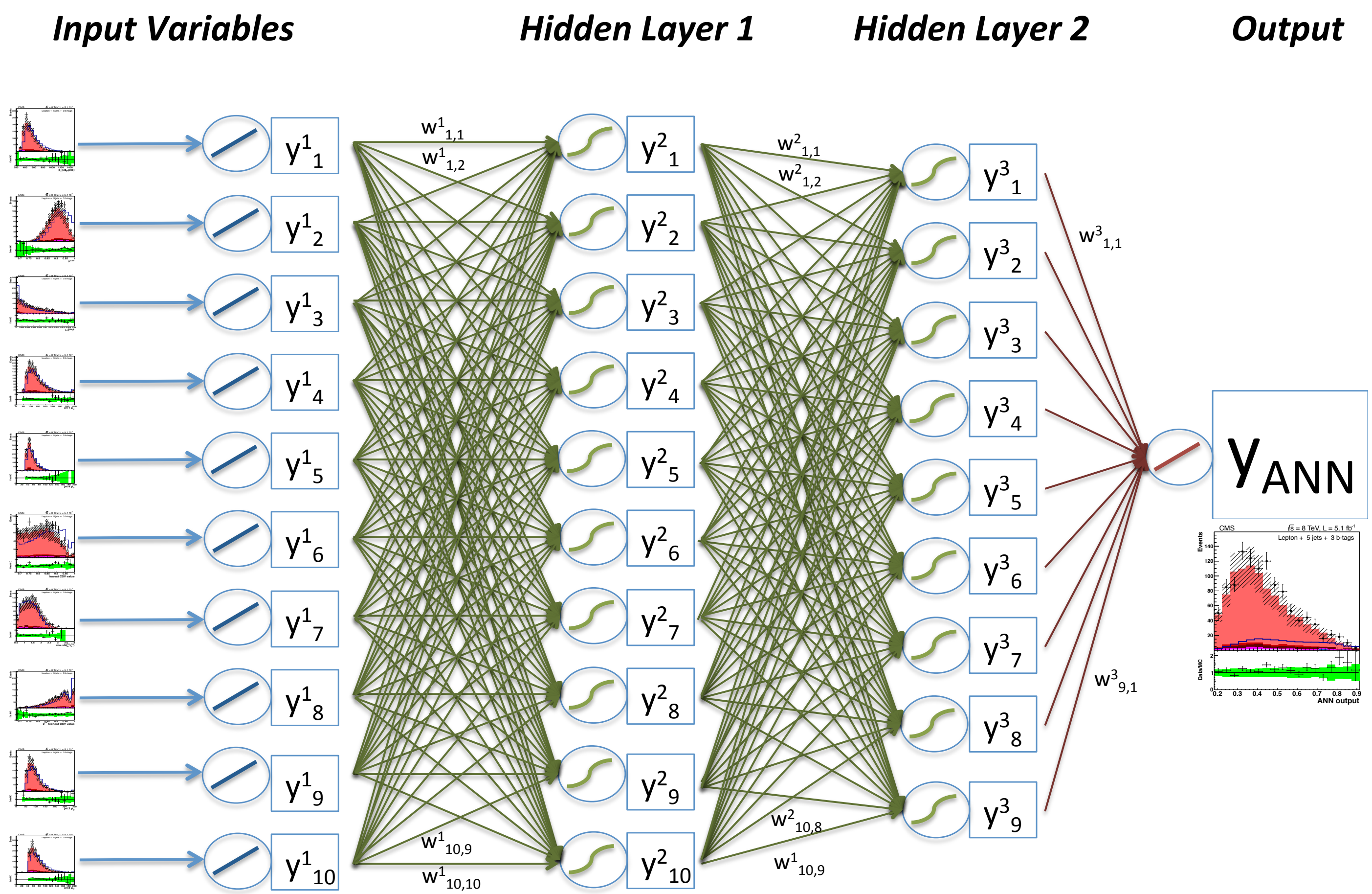
$$A(\alpha) = \frac{1}{1+e^{-x}}$$

"Synapse Response"

$$\alpha = w_{0j}^{(l)} + \sum_{i=1}^{n} y_i^{(l)} w_{ij}^{(l)}$$

Multi-Variate Analysis Techniques

- **Clermond-Ferrand Multi Layer Percepton (CFMlpANN)** - specific implementation of an ANN. Ours uses 10 input variables, and two hidden layers with 10, 9 nodes respectively



$$y_{ANN} = \sum_{k=1}^{n-1} y_k^{(3)} w_{k1}^{(3)} = \sum_{k=1}^{n-1} A\left(\sum_{j=1}^{n} y_j^{(2)} w_{jk}^{(2)}\right) w_{k1}^{(3)} = \sum_{k=1}^{n-1} A\left(\sum_{j=1}^{n} A\left(\sum_{i=1}^{n} x_i w_{ij}^{(1)}\right) w_{jk}^{(2)}\right) w_{k1}^{(3)}$$

Input Variables, all Categories

- **Artificial Neural Network (ANN)** – specific implementation of an MVA. This uses 10 input variables for each jet/tag category

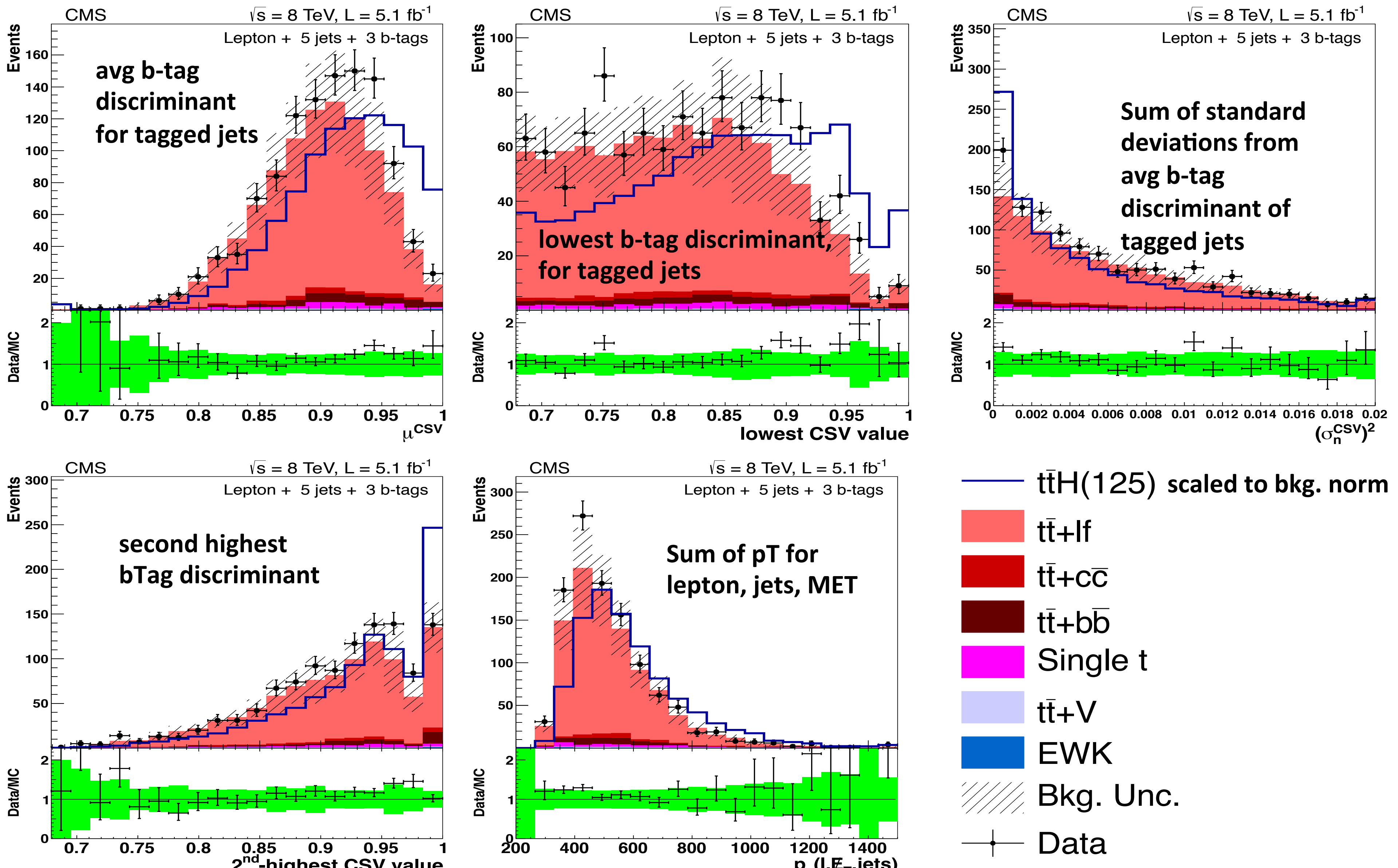
- Variables with a Star have the highest discriminating power

- Variables are generally classified as being related to:
  - Single jet kinematics
  - Dijet/MultiJet kinematics
  - Angular distributions
  - bTag discriminant

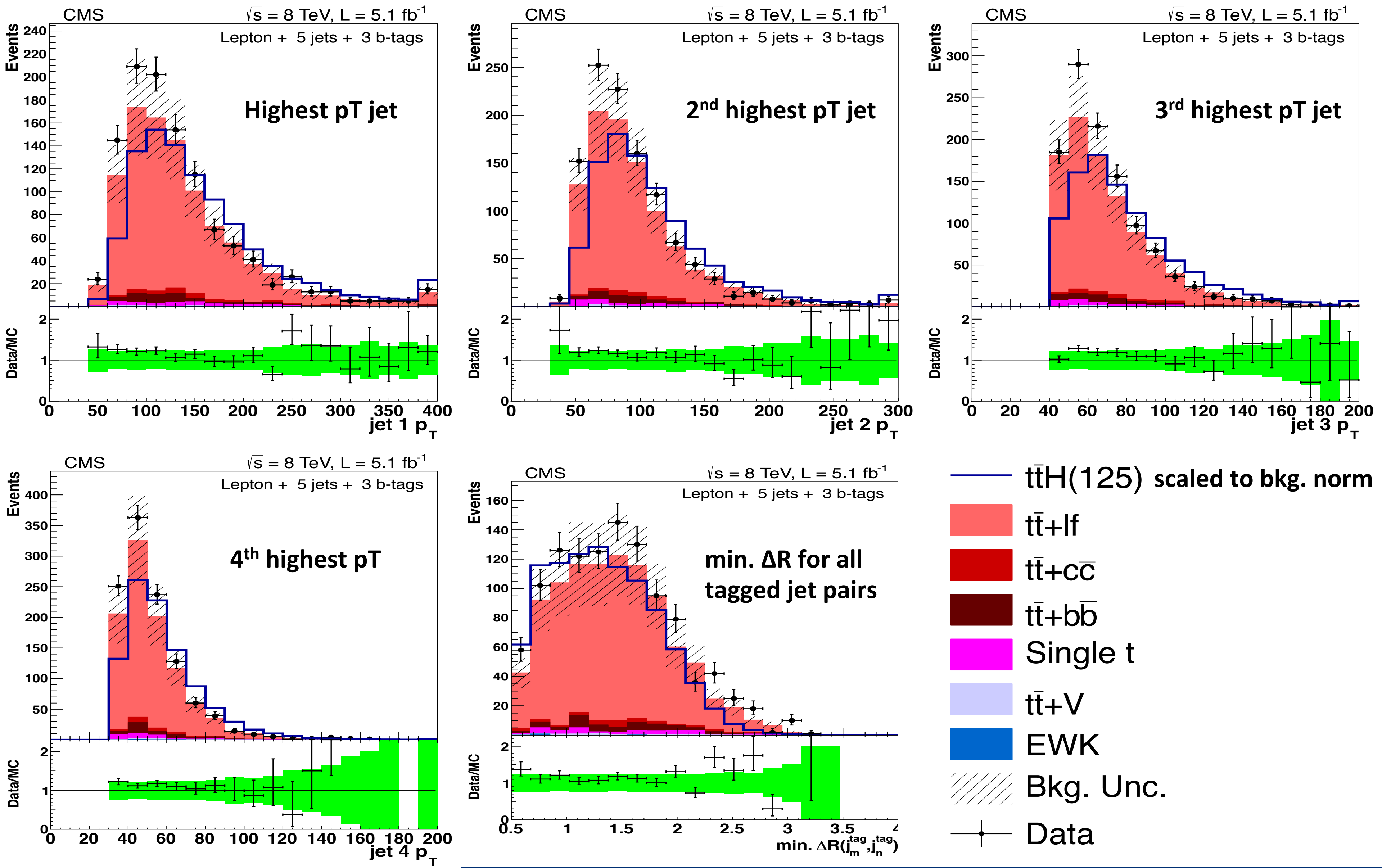
- bTag discriminant variables are present in almost every jet/tag category
  - Average bTag discriminant of all jets tends to be the single most powerful variable

	Lepton+Jets						
Jets Tags	≥ 6 2	4 3	5 3	≥ 6 3	4 4	5 ≥ 4	≥ 6 ≥ 4
Jet 1 p_T		✓	✓		✓		
Jet 2 p_T		✓	✓				
Jet 3 p_T	✓	✓	✓			✓	
Jet 4 p_T	✓	✓	✓			✓	
N_{jets}							
$p_T(\ell, E_T^{\text{miss}}, \text{jets})$		★	✓		✓	✓	
$M(\ell, E_T^{\text{miss}}, \text{jets})$	✓	✓		✓	✓		✓
Average $M((j_m^{\text{untag}}, j_n^{\text{untag}}))$	✓			✓			
$M((j_m^{\text{tag}}, j_n^{\text{tag}})_{\text{closest}})$							✓
$M((j_m^{\text{tag}}, j_n^{\text{tag}})_{\text{best}})$							✓
Average $\Delta R(j_m^{\text{tag}}, j_n^{\text{tag}})$				✓	✓	✓	✓
Minimum $\Delta R(j_m^{\text{tag}}, j_n^{\text{tag}})$			✓				
$\Delta R(\ell, j_{\text{closest}})$						✓	✓
Sphericity	✓			✓			✓
Aplanarity	✓				✓		
H_0	✓						
H_1	✓				✓		
H_2				✓			✓
H_3	★			✓			✓
μ^{CSV}	✓	✓	★	★	★	★	★
$(\sigma_n^{\text{CSV}})^2$		✓	✓	✓	✓	✓	
Highest CSV value						✓	
2^{nd}-highest CSV value		✓	✓	✓	✓	✓	✓
Lowest CSV value		✓	✓	✓	✓	✓	✓

Artificial Neural Network – Input, 5J 3T

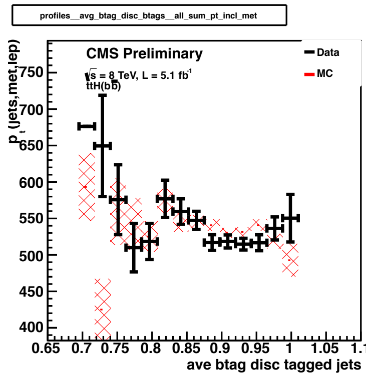


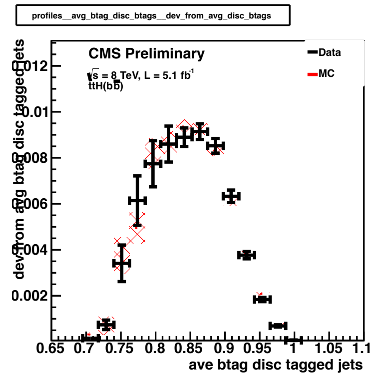
Artificial Neural Network – Input, 5J 3T

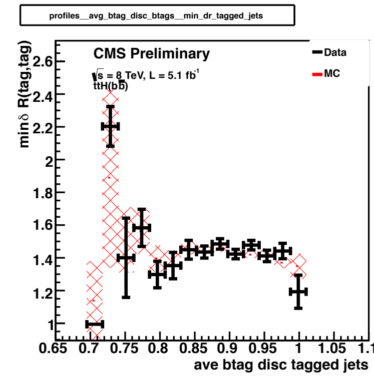


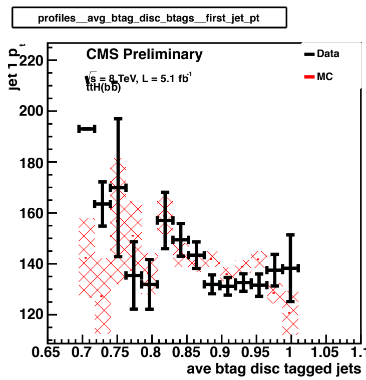
Artificial Neural Network – Input, 5J 3T

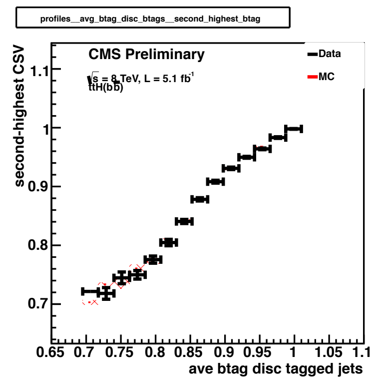
- Data to Monte Carlo comparisons of the correlations between the Average bTag Discriminant and all other variables in the 5j3t category
- These plots are the “profiles” of the 2D correlations – ie for each bin of the Average bTag Discriminant on the X-axis, the average value of the other input variable for events that would fall into that x-axis bin, is plotted, with the std. deviation of that average reported as the error
- The good data/MC agreement shows that the correlations between variables are also well-modeled

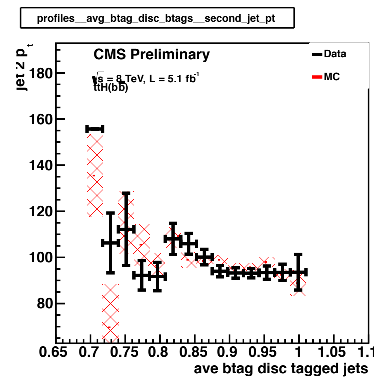


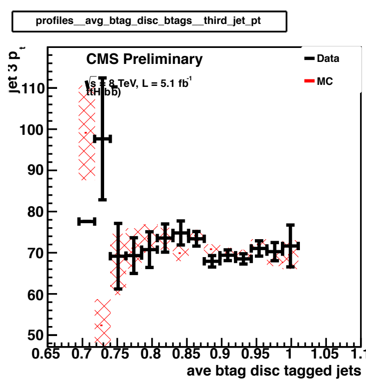


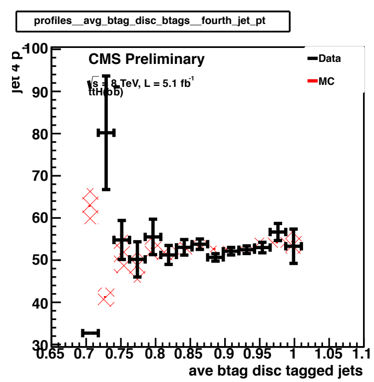


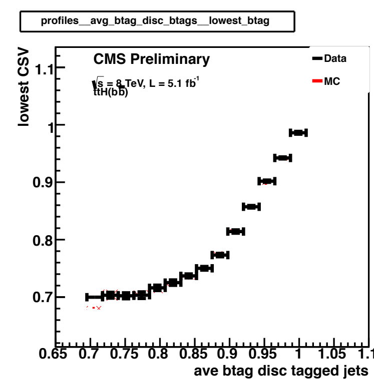






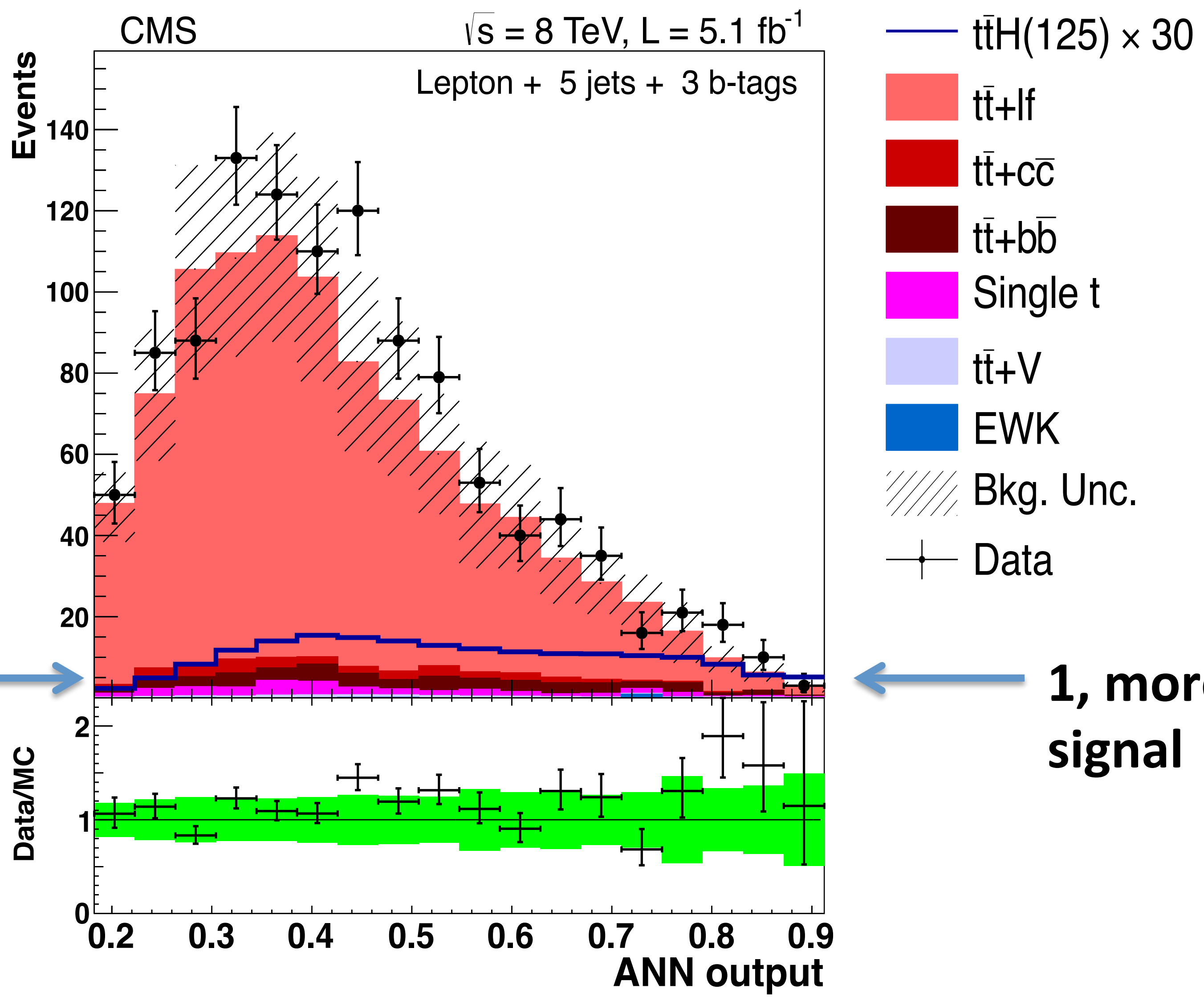






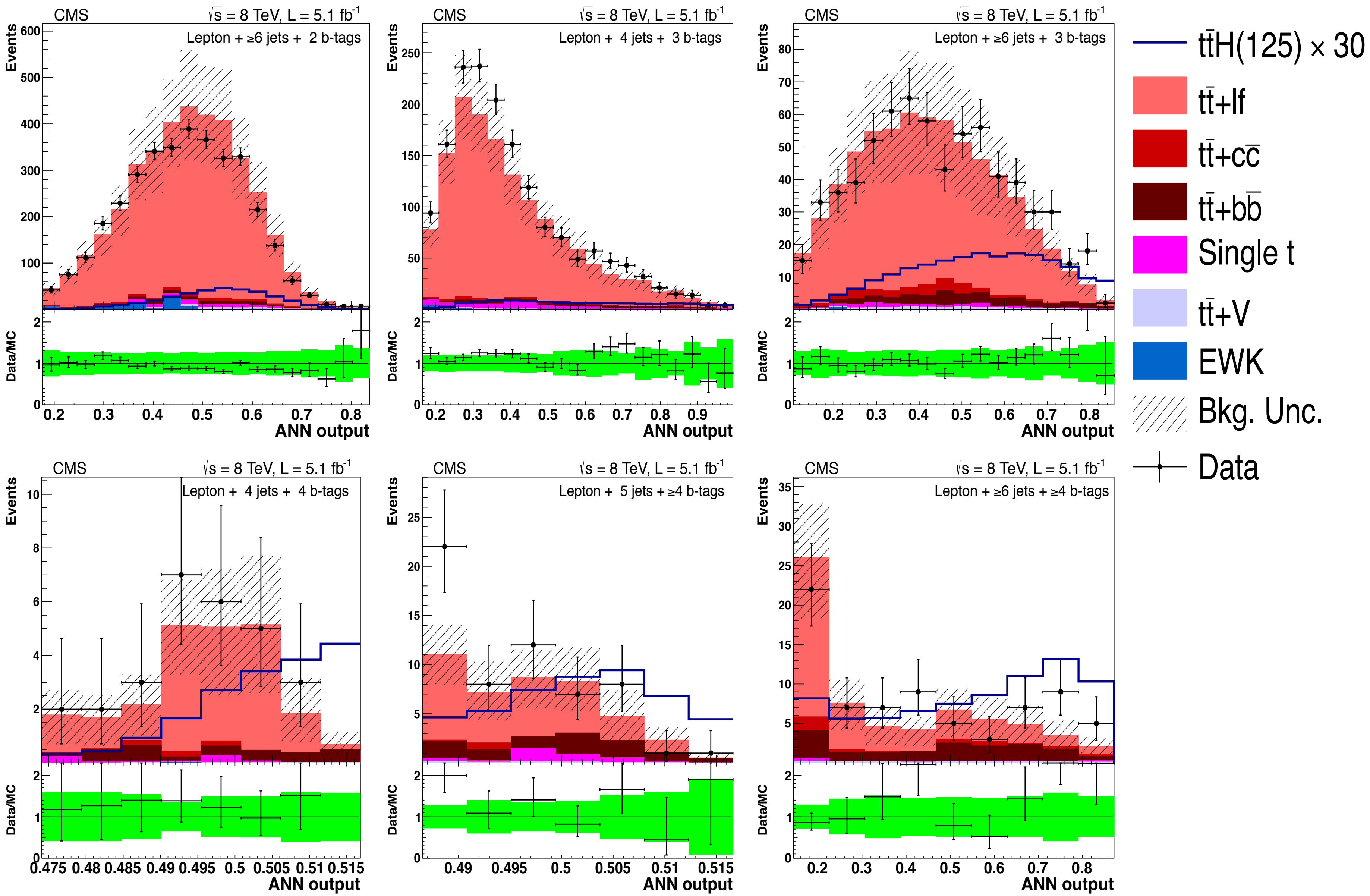
Artificial Neural Network – Output

0, more like background



1, more like signal

Artificial Neural Network – Output



Systematic Uncertainties

❖ *Below:* Table of complete list of uncertainties, % change to normalization, and whether or not uncertainty is used in shape comparison

Source	Rate Uncertainty	Shape	Remarks
Luminosity (7 TeV)	2.2%	No	All signal and backgrounds
Luminosity (8 TeV)	4.4%	No	All signal and backgrounds
Lepton ID/Trig	4%	No	All signal and backgrounds
Pileup	1%	No	All signal and backgrounds
Additional Pileup Corr.	–	Yes	All signal and backgrounds
Jet Energy Resolution	1.5%	No	All signal and backgrounds
Jet Energy Scale	0–60%	Yes	All signal and backgrounds
b-Tag SF (b/c)	0–33.6%	Yes	All signal and backgrounds
b-Tag SF (mistag)	0–23.5%	Yes	All signal and backgrounds
MC Statistics	–	Yes	All backgrounds
PDF (gg)	9%	No	For gg initiated processes ($t\bar{t}$, $t\bar{t}Z$, $t\bar{t}H$)
PDF ($q\bar{q}$)	4.2–7%	No	For $q\bar{q}$ initiated processes ($t\bar{t}W$, W, Z).
PDF (qg)	4.6%	No	For qg initiated processes (single top)
QCD Scale ($t\bar{t}H$)	15%	No	For NLO $t\bar{t}H$ prediction
QCD Scale ($t\bar{t}$)	2–12%	No	For NLO $t\bar{t}$ and single top predictions
QCD Scale (V)	1.2–1.3%	No	For NNLO W and Z prediction
QCD Scale (VV)	3.5%	No	For NLO diboson prediction
Madgraph Scale ($t\bar{t}$)	0–20%	Yes	$t\bar{t}$ + jets/$b\bar{b}$/$c\bar{c}$ uncorrelated. Varies by jet bin.
Madgraph Scale (V)	20–60%	No	Varies by jet bin.
$t\bar{t} + b\bar{b}$	50%	No	Only $t\bar{t} + b\bar{b}$.

Systematic Uncertainties

❖ *Below:* Table of complete list of uncertainties, % change to normalization, and whether or not uncertainty is used in shape comparison

Source	Rate Uncertainty	Shape	Remarks
Luminosity (7 TeV)	2.2%	No	All signal and backgrounds
Luminosity (8 TeV)	4.4%	No	All signal and backgrounds
Lepton ID/Trig	4%	No	All signal and backgrounds
Pileup	1%	No	All signal and backgrounds
Additional Pileup Corr.	–	Yes	All signal and backgrounds
Jet Energy Resolution	1.5%	No	All signal and backgrounds
Jet Energy Scale	0–60%	Yes	All signal and backgrounds
b-Tag SF (b/c)	0–33.6%	Yes	All signal and backgrounds
b-Tag SF (mistag)	0–23.5%	Yes	All signal and backgrounds
MC Statistics	–	Yes	All backgrounds
PDF (gg)	9%	No	For gg initiated processes ($t\bar{t}$, $t\bar{t}Z$, $t\bar{t}H$)
PDF ($q\bar{q}$)	4.2–7%	No	For $q\bar{q}$ initiated processes ($t\bar{t}W$, W, Z).
PDF (qg)	4.6%	No	For qg initiated processes (single top)
QCD Scale ($t\bar{t}H$)	15%	No	For NLO $t\bar{t}H$ prediction
QCD Scale ($t\bar{t}$)	2–12%	No	For NLO $t\bar{t}$ and single top predictions
QCD Scale (V)	1.2–1.3%	No	For NNLO W and Z prediction
QCD Scale (VV)	3.5%	No	For NLO diboson prediction
Madgraph Scale ($t\bar{t}$)	0–20%	Yes	$t\bar{t}$ + jets/$b\bar{b}$/$c\bar{c}$ uncorrelated. Varies by jet bin.
Madgraph Scale (V)	20–60%	No	Varies by jet bin.
$t\bar{t} + b\bar{b}$	50%	No	Only $t\bar{t} + b\bar{b}$.

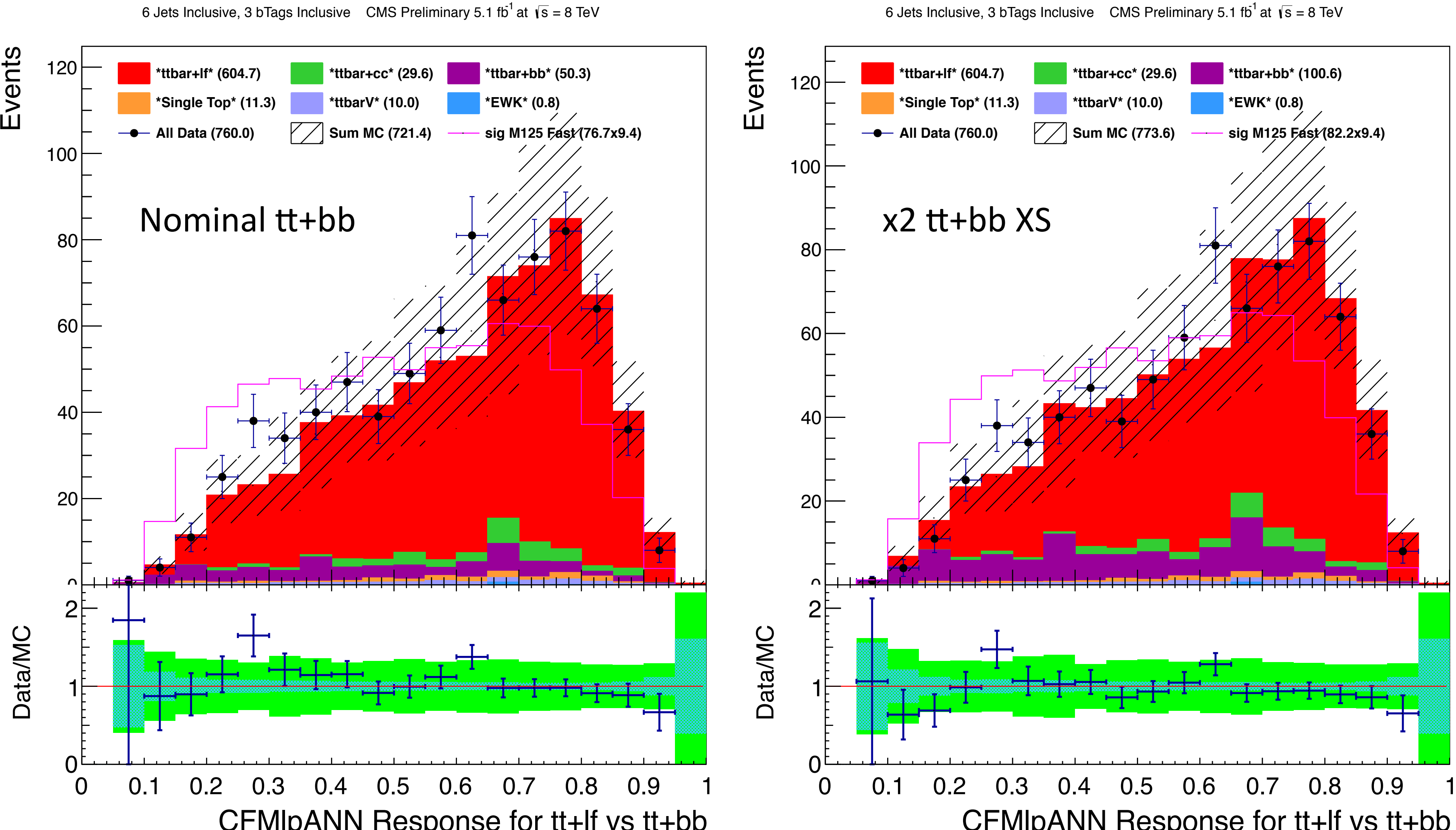
Systematic Uncertainties

- *Below:* Table of complete list of uncertainties, % change to normalization, and whether or not uncertainty is used in shape comparison

Source	Rate Uncertainty	Shape	Remarks
Luminosity (7 TeV)	2.2%	No	All signal and backgrounds
Luminosity (8 TeV)	4.4%	No	All signal and backgrounds
Lepton ID/Trig	4%	No	All signal and backgrounds
Pileup	1%	No	All signal and backgrounds
Additional Pileup Corr.	–	Yes	All signal and backgrounds
Jet Energy Resolution	1.5%	No	All signal and backgrounds
Jet Energy Scale	0–60%	Yes	All signal and backgrounds
b-Tag SF (b/c)	0–33.6%	Yes	All signal and backgrounds
b-Tag SF (mistag)	0–23.5%	Yes	All signal and backgrounds
MC Statistics	–	Yes	All backgrounds
PDF (gg)	9%	No	For gg initiated processes ($t\bar{t}$, $t\bar{t}Z$, $t\bar{t}H$)
PDF ($q\bar{q}$)	4.2–7%	No	For $q\bar{q}$ initiated processes ($t\bar{t}W$, W, Z).
PDF (qg)	4.6%	No	For qg initiated processes (single top)
QCD Scale ($t\bar{t}H$)	15%	No	For NLO $t\bar{t}H$ prediction
QCD Scale ($t\bar{t}$)	2–12%	No	For NLO $t\bar{t}$ and single top predictions
QCD Scale (V)	1.2–1.3%	No	For NNLO W and Z prediction
QCD Scale (VV)	3.5%	No	For NLO diboson prediction
Madgraph Scale ($t\bar{t}$)	0–20%	Yes	$t\bar{t}$ + jets/$b\bar{b}$/$c\bar{c}$ uncorrelated. Varies by jet bin.
Madgraph Scale (V)	20–60%	No	Varies by jet bin.
$t\bar{t} + b\bar{b}$	50%	No	Only $t\bar{t} + b\bar{b}$.

tt+bb vs tt+lf discriminant– 8TeV L+J, mu+ele, >=6J, >=3T

- Here, in the most sensitive category for tt+bb, nominal tt+bb XS shows more consistent levels of agreement than the doubled XS



Limit Setting Technique

- In the case of no clear signal, ask what is the upper limit on ttH production possible that can still be explained by this dataset

- CLs is modified frequentist technique (~Bayesian, relies on background prediction)

- Begins by forming a likelihood function assuming a Poisson distribution for the number of events observed in a given bin
  - For a given signal strength multiplier, mu, the likelihood is maximized if the prediction for the number of events in the bin (u*s+b) is comparable to the number observed, n_i
  - Systematic uncertainties are accounted for through nuisance parameters, using log-normal distributions for rate uncertainties, and template fitting for shape uncertainties

$$\begin{aligned}\mathcal{L}(\mathrm{data}|\mu,\theta) &= Poisson(\mathrm{data}|\mu\cdot s(\theta)+b(\theta))\cdot p(\tilde{\theta}|\theta)\\ &= \prod_i \frac{(\mu s_i+b_i)^{n_i}}{n_i!}e^{-(\mu s_i+b_i)}\cdot p(\tilde{\theta}|\theta)\end{aligned}$$

Limit Setting Technique

- A test statistic is formed for two hypotheses – one testing a particular value of mu, the other comparing the best fit of mu within the range 0 < mu_best < mu_tested
  - This assures that an upward fluctuation of background can't be used to exclude a signal – ie no lower bound on the value of mu
- The statistic q_u is formed for the signal+background hypothesis (mu != 0), and the background only hypothesis (mu=0)

$$\tilde{q}_\mu = -2\ln\frac{\mathcal{L}(\text{data}|\mu,\hat{\theta}_\mu)}{\mathcal{L}(\text{data}|\hat{\mu},\hat{\theta})} \qquad , 0 \leq \hat{\mu} \leq \mu$$

- PDFs of q_u are formed by varying the signal and background templates, constrained by the range allowed by the systematic uncertainties, and p-values calculated for the signal +background and background-only hypotheses

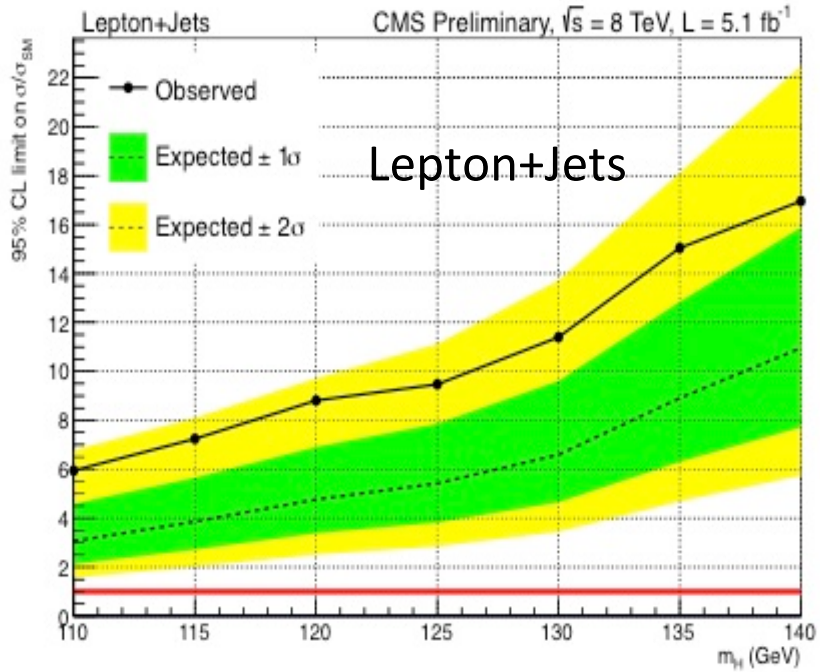
$$p_\mu = P(\tilde{q}_\mu \geq \tilde{q}_\mu^{obs}|signal + background) = \int_{\tilde{q}_\mu^{obs}}^{\text{inf}} f(\tilde{q}_\mu|\mu,\hat{\theta}_\mu^{obs})d\tilde{q}_\mu$$

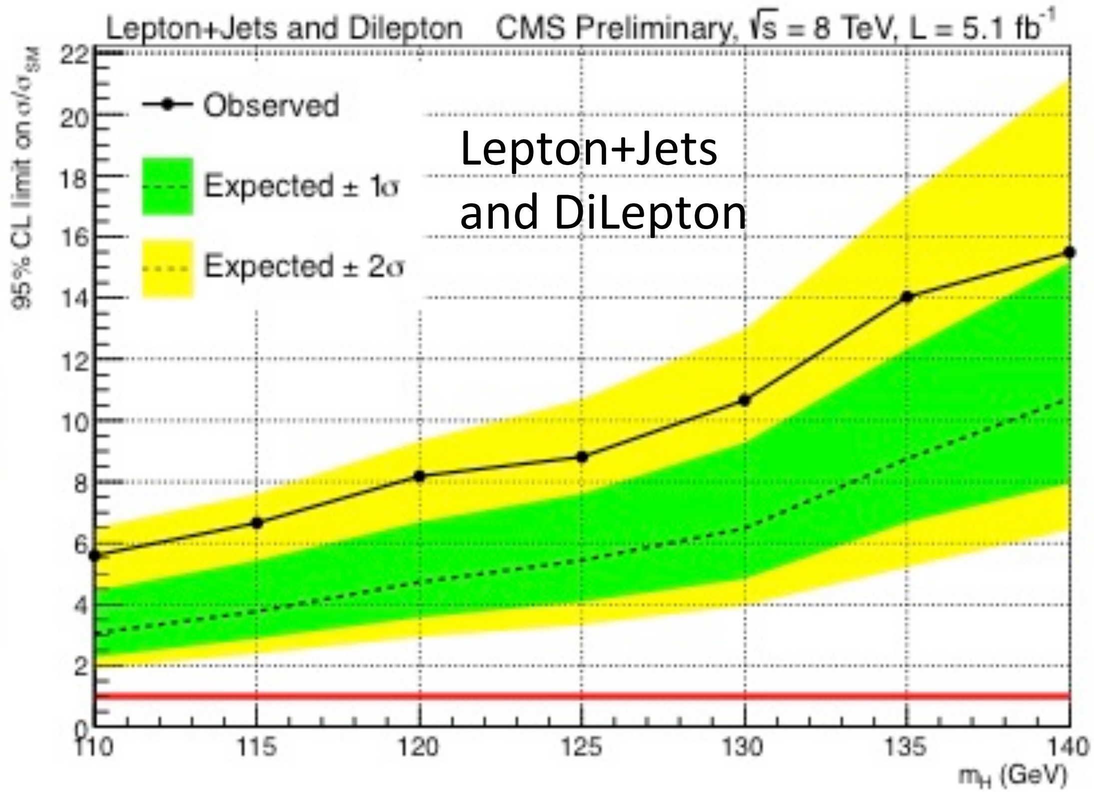
$$1 - p_0 = P(\tilde{q}_\mu \geq \tilde{q}_\mu^{obs}|background - only) = \int_{\tilde{q}_0^{obs}}^{\text{inf}} f(\tilde{q}_\mu|0,\hat{\theta}_0^{obs})d\tilde{q}_\mu$$

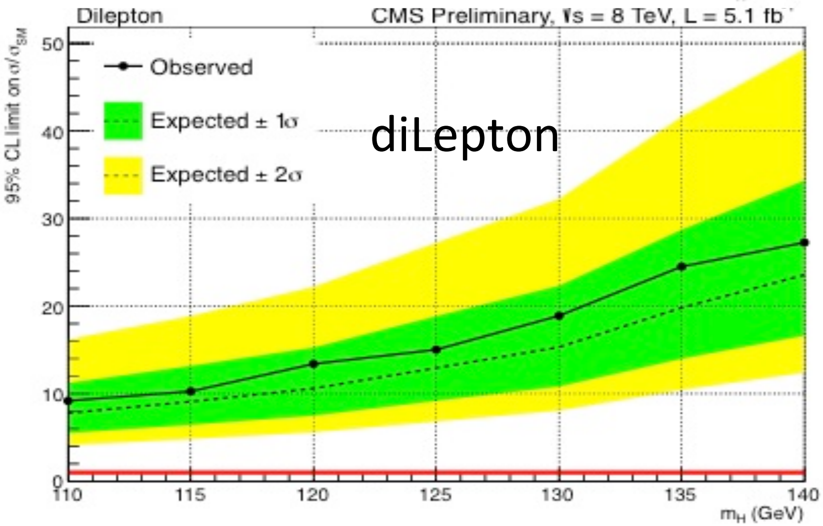
- The 95% CLs upper limit is determined by the value of mu that gives a ratio of 0.05

$$CL_s(\mu) = \frac{p_\mu}{1-p_0}$$

Limit Setting – 8TeV Only







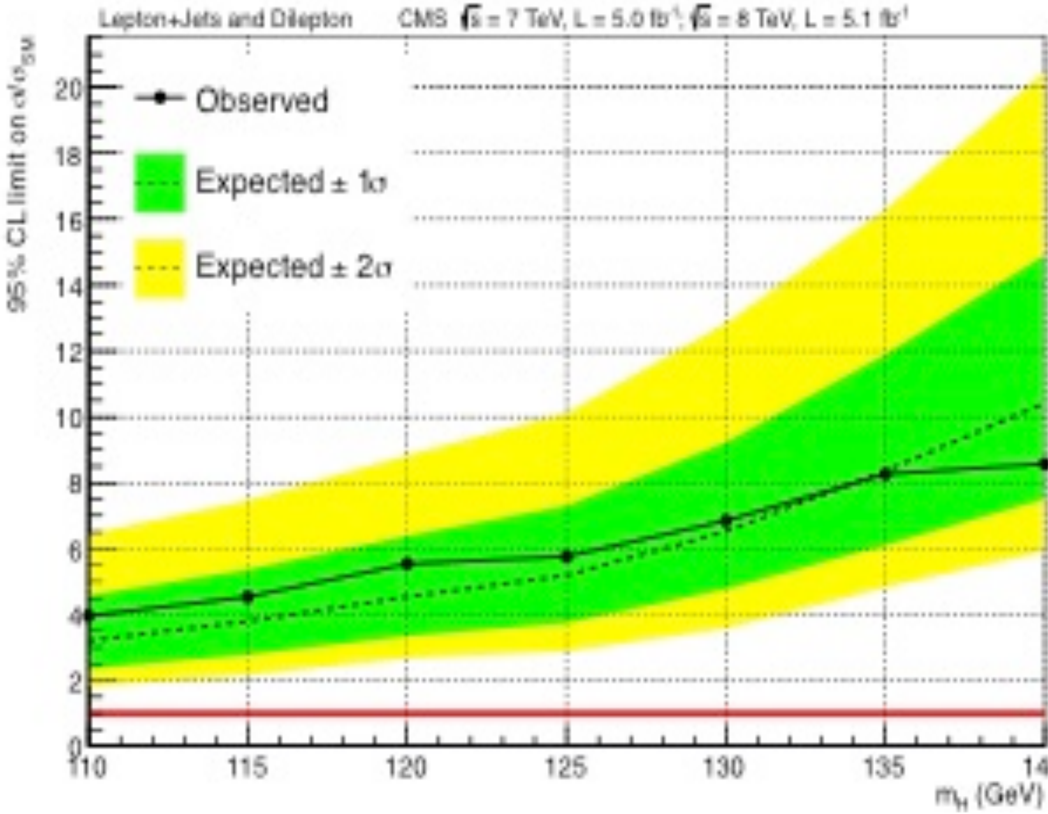
- Ideally, extract and measure Higgs signal – but no large excess in SM predictions seen
- Proceed by setting upper limit on Higgs production
- At Higgs mass 125 GeV, Lepton+Jets Channel
  - ***expect to set a limit of 5.4x σ_{SM}***
  - ***observed upper limit: 8.8 x σ_{SM}***

Summary of Initial 7+8TeV Results

- A similar analysis also exists for the 7TeV data, with approximately equal statistics

- Combine 7+8TeV results to improve sensitivity to ttH production

- ***At M125***
  - ***Expected limit = 5.2 x σ_{SM}***
  - ***Observed limit = 5.8 x σ_{SM}***



7 TeV + 8 TeV Lepton+Jets and Dilepton combined				
		Expected		
m_H	Observed	Median	68% CL Range	95% CL Range
110 GeV	4.0	3.2	[2.4, 4.6]	[1.8, 6.5]
115 GeV	4.5	3.8	[2.8, 5.4]	[2.2, 7.5]
120 GeV	5.5	4.5	[3.3, 6.4]	[2.7, 8.9]
125 GeV	5.8	5.2	[3.7, 7.3]	[2.9, 10.1]
130 GeV	6.8	6.5	[4.8, 9.2]	[3.6, 12.9]
135 GeV	8.3	8.4	[6.1, 11.9]	[4.8, 16.3]
140 GeV	8.6	10.4	[7.5, 14.9]	[5.9, 20.5]

Summary of Initial 7+8TeV Results

- A similar analysis also exists for the 7TeV data, with approximately equal statistics
- Combine 7+8TeV results to improve sensitivity to ttH production
- ***At M125***
  - ***Expected limit = 5.2 x σ_{SM}***
  - ***Observed limit = 5.8 x σ_{SM}***
- **Publication –** 5 fb^{-1} 7TeV data, first 5 fb^{-1} 8TeV data
  - Lepton+Jets, diLepton Channels
  - JHEP, May 2013, [JHEP paper link]
  - Arxiv: http://arxiv.org/pdf/1303.0763v1.pdf

EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH (CERN)

CERN-PH-EP/2013-027
2013/03/05

CMS-HIG-12-035

Search for the standard model Higgs boson produced in association with a top-quark pair in pp collisions at the LHC

The CMS Collaboration*

Abstract

A search for the standard model Higgs boson produced in association with a top-quark pair is presented using data samples corresponding to an integrated luminosity of 5.0 fb^{-1} (5.1 fb^{-1}) collected in pp collisions at the center-of-mass energy of 7 TeV (8 TeV). Events are considered where the top-quark pair decays to either one lepton+jets ($t\bar{t} \to \ell\nu q\bar{q}'b\bar{b}$) or dileptons ($t\bar{t} \to \ell^+\nu\ell^-\bar{\nu}b\bar{b}$), ℓ being an electron or a muon. The search is optimized for the decay mode $H \to b\bar{b}$. The largest background to the $t\bar{t}H$ signal is top-quark pair production with additional jets. Artificial neural networks are used to discriminate between signal and background events. Combining the results from the 7 TeV and 8 TeV samples, the observed (expected) limit on the cross section for Higgs boson production in association with top-quark pairs for a Higgs boson mass of 125 GeV is 5.8 (5.2) times the standard model expectation.

Submitted to the Journal of High Energy Physics

*See Appendix A for the list of collaboration members

2nd Publication, 19.5 fb^{-1} 8 TeV

- **Update to Full 19.5 fb^{-1} of 8TeV data**
  - Improved data statistics – x4 stats wrt 2011
  - Higher statistics Monte Carlo
  - Derived custom b-tag Scale Factors
  - Improved MVA structure, input variables
  - Additional Search Channels for Combination
    - ttH, H->ττ
    - ttH, H->WW,ZZ, τ's-> 3, 4, leptons
    - ttH, H->ZZ -> Same-sign diLepton
    - ttH, H->γγ
  - Public PAS Hig-13-019

- **Publication –** 5 fb^{-1} 7TeV data, 19.5 fb^{-1} 8TeV data
  - JHEP, September 2014, JHEP paper link
  - Arxiv: http://arxiv.org/pdf/1408.1682v1.pdf

- **For the rest of the talk**
  - Updated Lepton+Jets analysis in the H->bb decay mode – full 8TeV data
  - Initial ttH combined channel results

EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH (CERN)

CMS-HIG-13-029

CERN-PH-EP/2014-189
2014/08/08

Search for the associated production of the Higgs boson with a top-quark pair

The CMS Collaboration*

Abstract

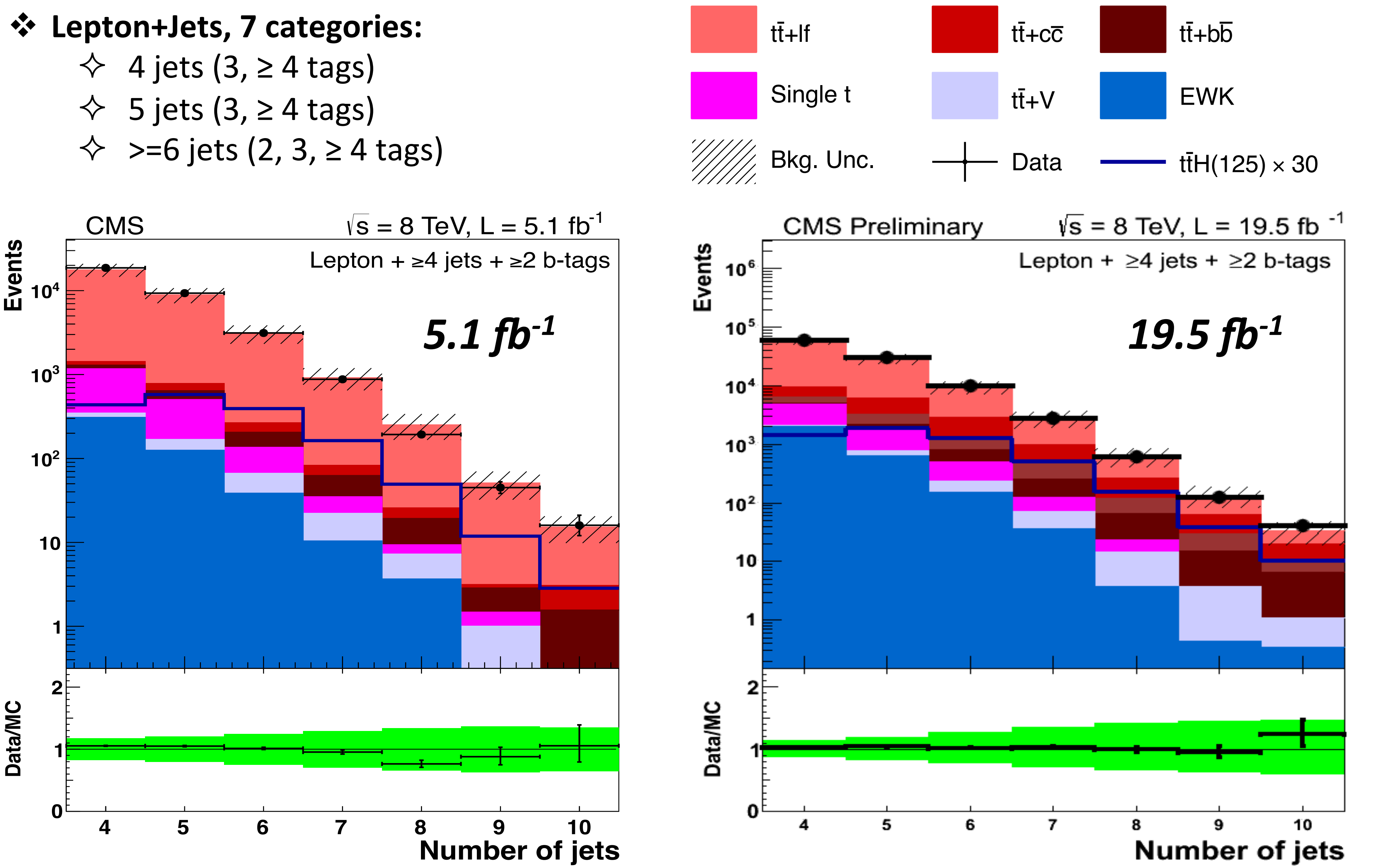
A search for the standard model Higgs boson produced in association with a top-quark pair ($t\bar{t}H$) is presented, using data samples corresponding to integrated luminosities of up to 5.1 fb^{-1} and 19.7 fb^{-1} collected in pp collisions at center-of-mass energies of 7 TeV and 8 TeV respectively. The search is based on the following signatures of the Higgs boson decay: H → hadrons, H → photons, and H → leptons. The results are characterized by an observed $t\bar{t}H$ signal strength relative to the standard model cross section, $\mu = \sigma/\sigma_{SM}$, under the assumption that the Higgs boson decays as expected in the standard model. The best fit value is $\mu = 2.8 \pm 1.0$ for a Higgs boson mass of 125.6 GeV.

Submitted to the Journal of High Energy Physics

*See Appendix A for the list of collaboration members

Number of Jets/Tags Distributions

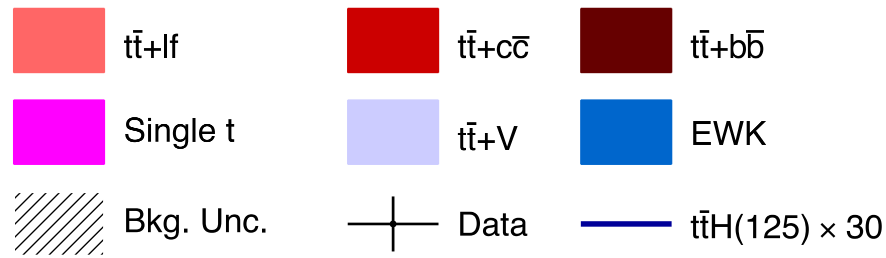
- **Lepton+Jets, 7 categories:**
  - 4 jets (3, ≥ 4 tags)
  - 5 jets (3, ≥ 4 tags)
  - >=6 jets (2, 3, ≥ 4 tags)

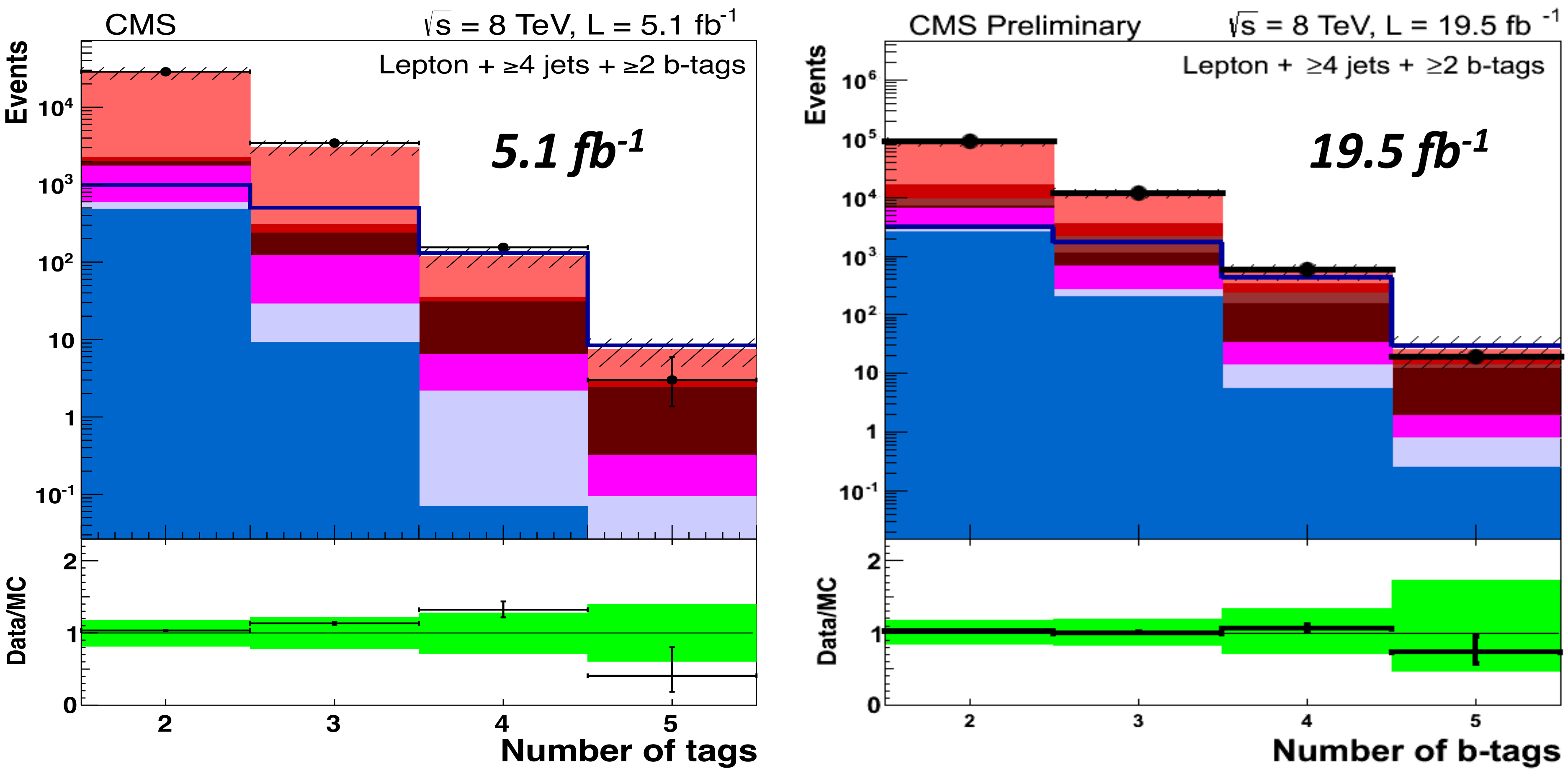


Number of Jets/Tags Distributions

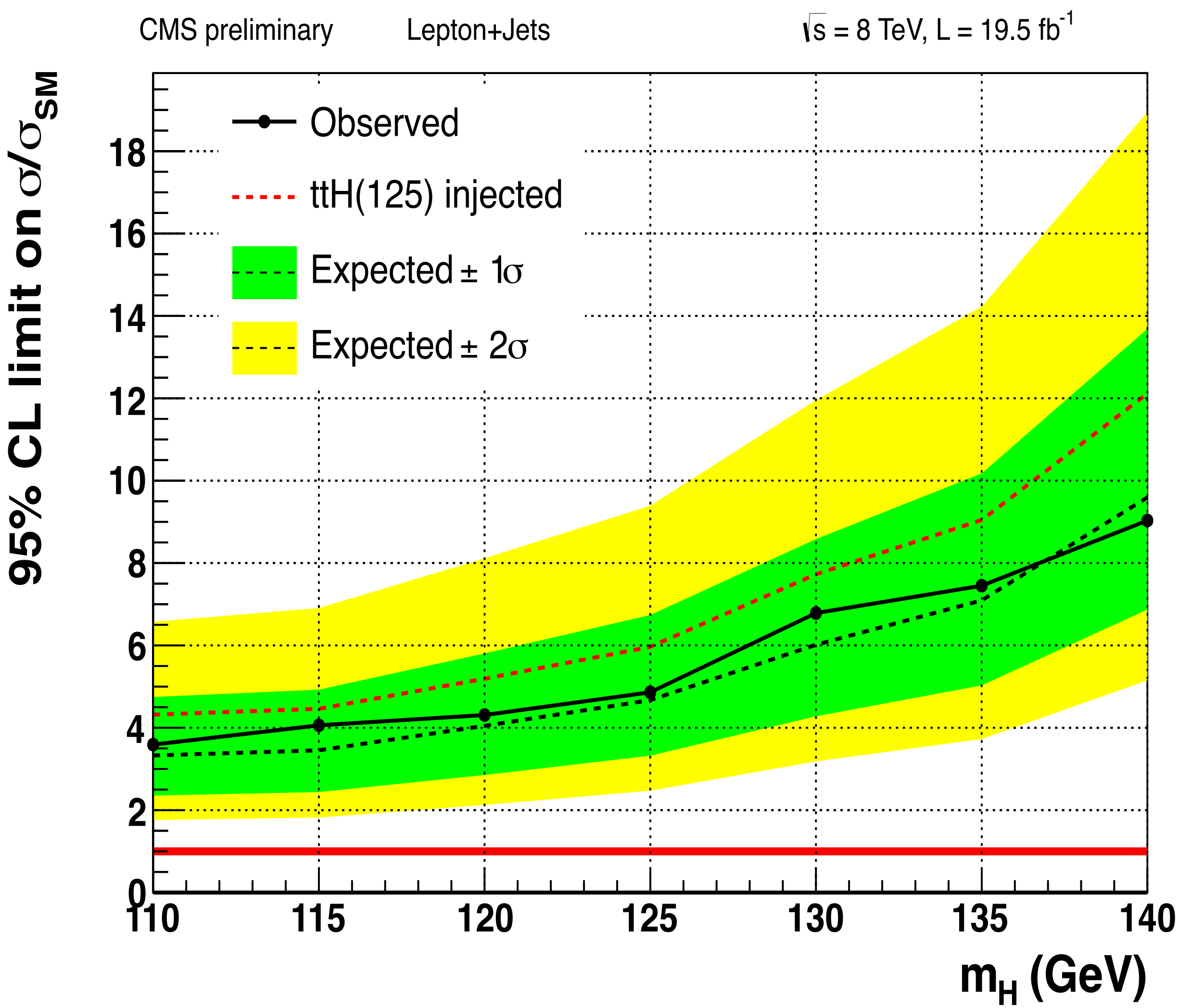
❖ **Lepton+Jets, 7 categories:**

- ✧ 4 jets (3, ≥ 4 tags)
- ✧ 5 jets (3, ≥ 4 tags)
- ✧ >=6 jets (2, 3, ≥ 4 tags)





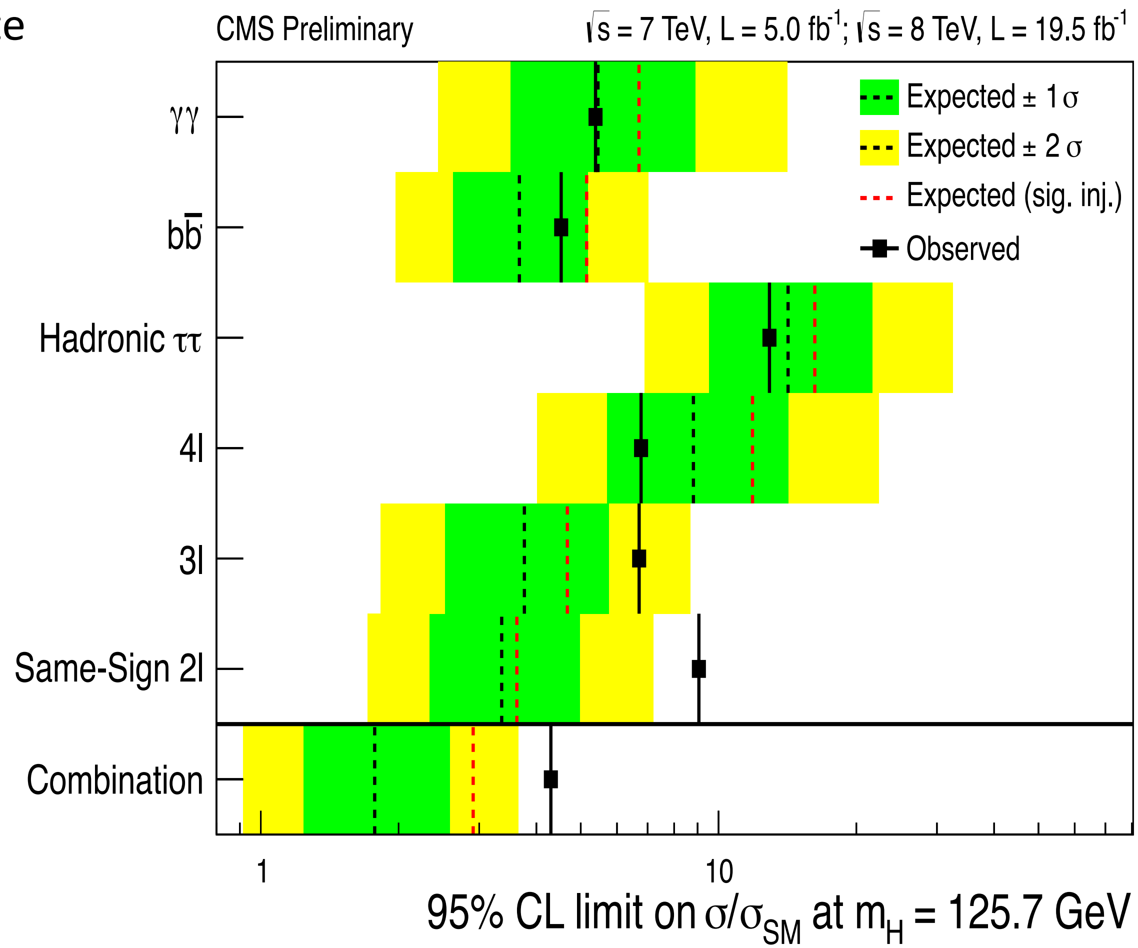
Signal Extraction – LJ, 8TeV



- Ideally, extract and measure Higgs signal – but no large excess in SM predictions seen
- Proceed by setting upper limit on Higgs production
- At Higgs mass 125 GeV
  - ***Expected limit: 4.7xσ_{SM}***
  - ***Observed limit: 4.9xσ_{SM}***
  - ***ttH injected: 6.0xσ_{SM}***
- Compared to 5fb^{-1} 8 TeV
  - ***Expected limit: 5.4xσ_{SM}***
  - ***Observed limit: 8.8σ_{SM}***
- Improvement less than expected from increasing lumi alone due to a different and more conservative estimate of systematic uncertainties

Summary of Initial 7+8TeV Results

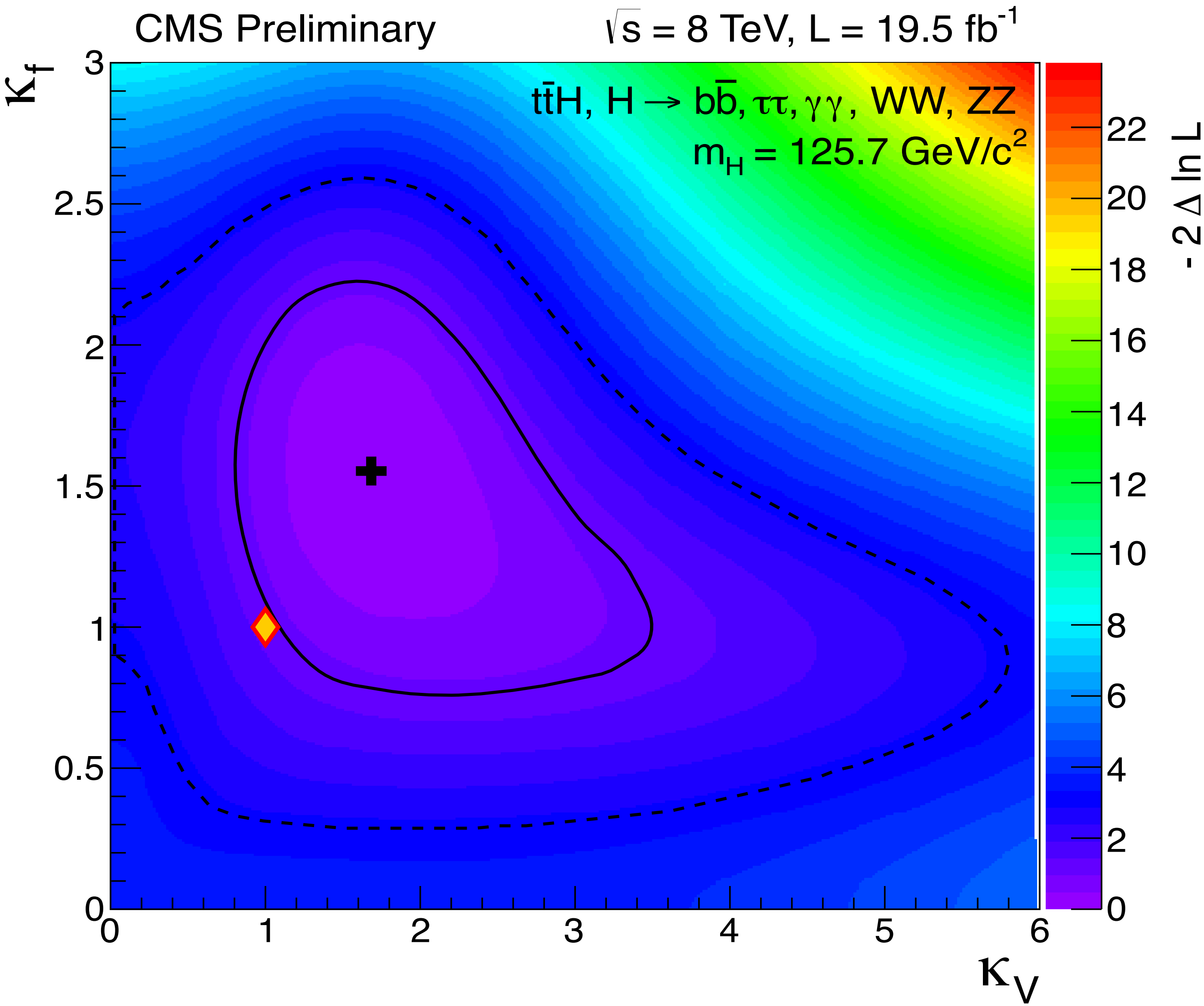
- Several Analysis Groups contribute to ttH searches at CMS
- Can access all decay modes with one production mechanism!
  - ttH, H->bb, L+J
  - ttH, H->bb, diLepton
  - ttH, H->ττ
  - ttH, H->γγ
  - ttH, H->WW,ZZ, τ's-> 4 Leptons
  - ttH, H->WW,ZZ, τ's-> 3 Leptons
  - ttH, H->ZZ -> Same-Sign dilepton



- ***At M125.7***
  - ***Expected: 1.8x σ_{SM}***
  - ***Observed: 4.3x σ_{SM}***
  - ***Signal Injected: 2.9x σ_{SM}***

Summary of Initial 7+8TeV Results

- Measurement of coupling constant to fermions (k_f), and bosons (k_V)
- SM is red/yellow diamond at (1,1)
- Solid line is 68% CL, Dashed Line is 95%



Looking towards Run 2

- This analysis will continue towards its goal of observing the Higgs produced in association with top-quark pairs during the next run of data collection at the LHC

- Additional improvements to the analysis are also underway
  - Investigations into NLO Monte Carlo
    - NLO tt+0/1/2 jets aMC@NLO + Pythia8, Powheg
    - NLO tt+bb aMC@NLO + Pythia, Powheg, Sherpa
    - NLO ttH+0/1 jets aMC@NLO + Pythia8, Powheg
  - Improved theoretical cross section measurements
    - Theoretical uncertainty on tt+jets cross section is one of the largest uncertainties for background normalization in >=6Jets >=4tags
  - Improved jet association algorithms
    - Correctly identifying which jets came from ttbar versus Higgs system will improve kinematic handle on ttH events
    - Exploitation of spin correlated variables of decay products
    - Improved X^2 minimization, BDT driven algorithm
  - Optimized BDT techniques
    - Tiered BDT ie MVA discriminant for each major tt+X background

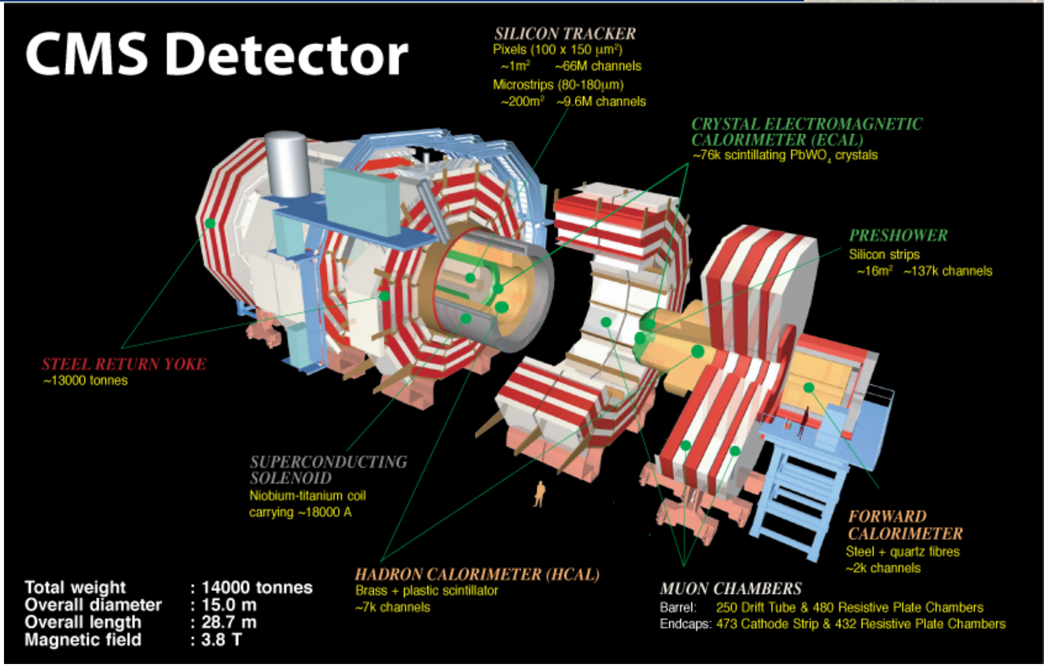
Summary

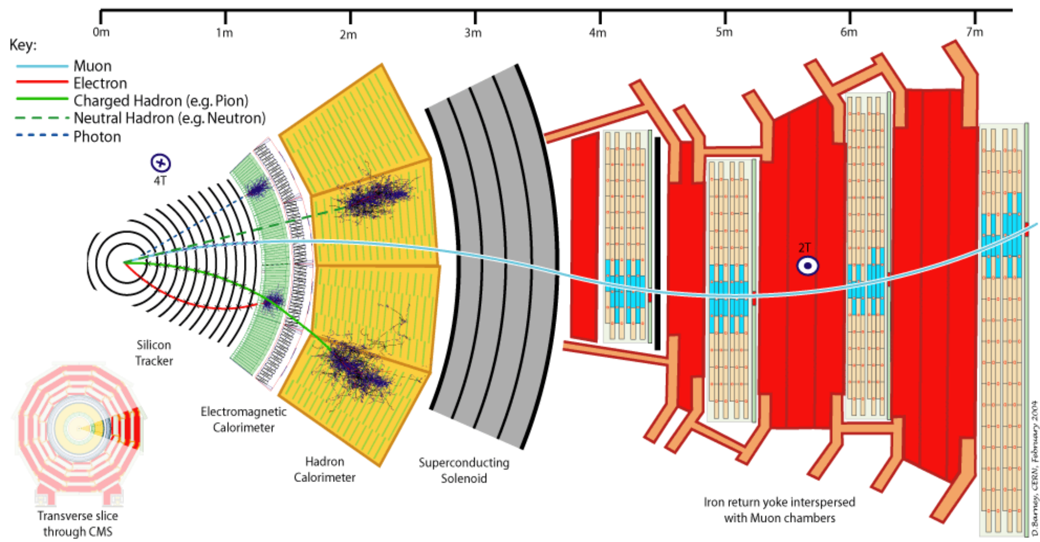
- This analysis is closing in on SM Higgs sensitivity!
- Adding in full 8TeV data, 20 fb^{-1} (x4 more data), plus different set of decay modes has given greater sensitivity to ttH production
  - Expected limit of the combination at ~1.8 * SM production sensitivity
- A measurement of the top-Higgs coupling is essential for the Run2 program at the LHC
- Exciting results are coming!!!

BACKUP

CMS Detector Overview

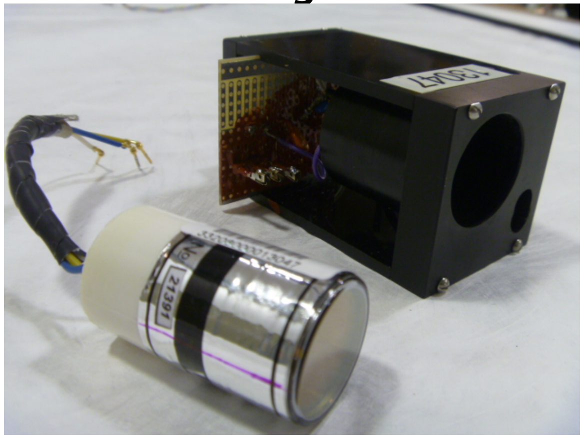
- **CMS detector sub-systems**
- *Tracker*
    - 66 (9.3) million channels in silicon pixels (strips) – high precision detection of ionizing radiation from charged particles
- *ECAL*
    - 60k(14k) $PbWO_4$ crystals - Absorbing/Scintillating medium from electromagnetic particles – e^- and γ deposit the most
- *HCAL*
    - 17 layers of plastic scintillator interspaced in brass (steel for outer/innermost layers) absorbers- initiates cascade & samples energy
- *Muon System*
    - 3 types of detectors- DTs, CSCs, RPCs, each specializing for use in tracking, spatial, and timing resolution



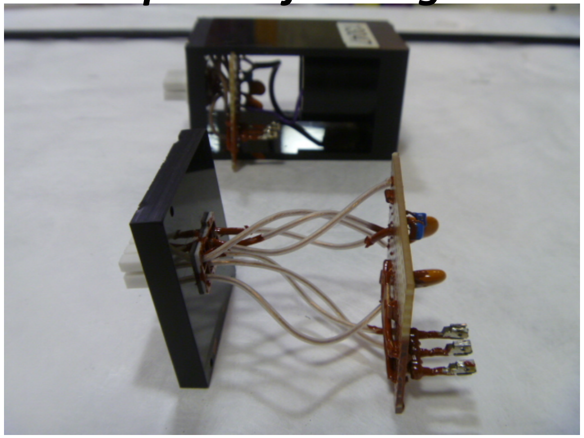


- Connection Scheme

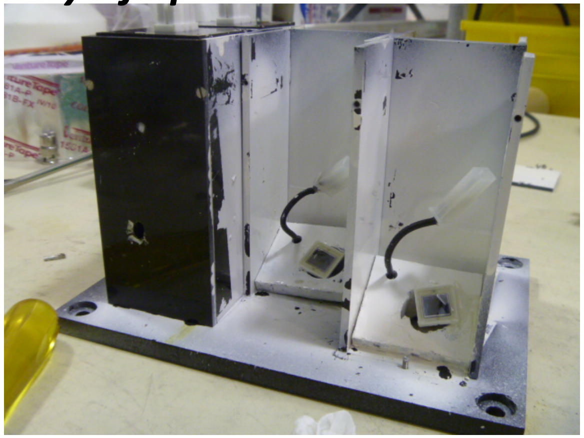
VPT & Housing

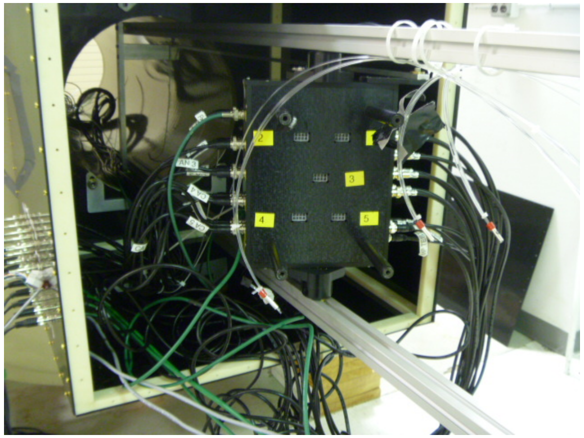


Simple HV filtering

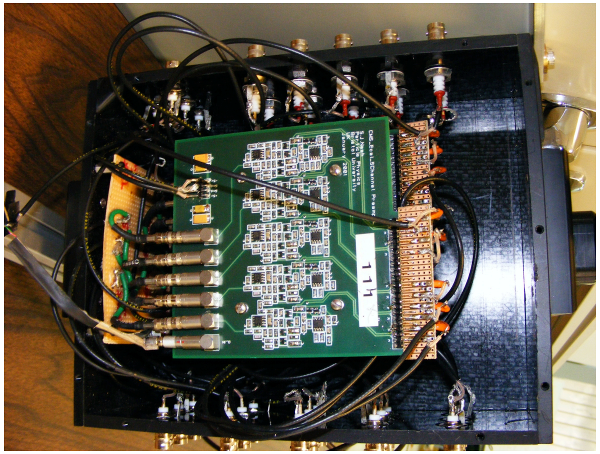


Array of vpt & PIN chambers

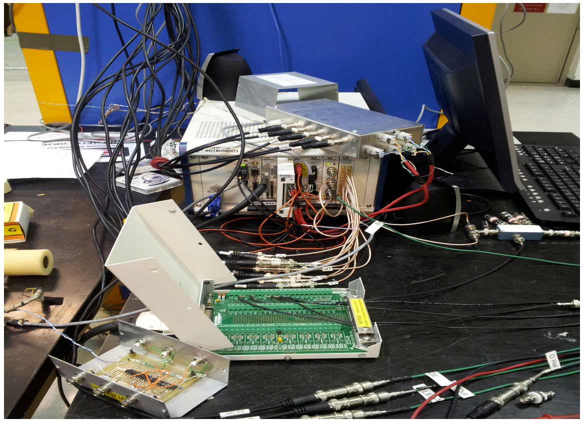




Connection of chambers to amplifier



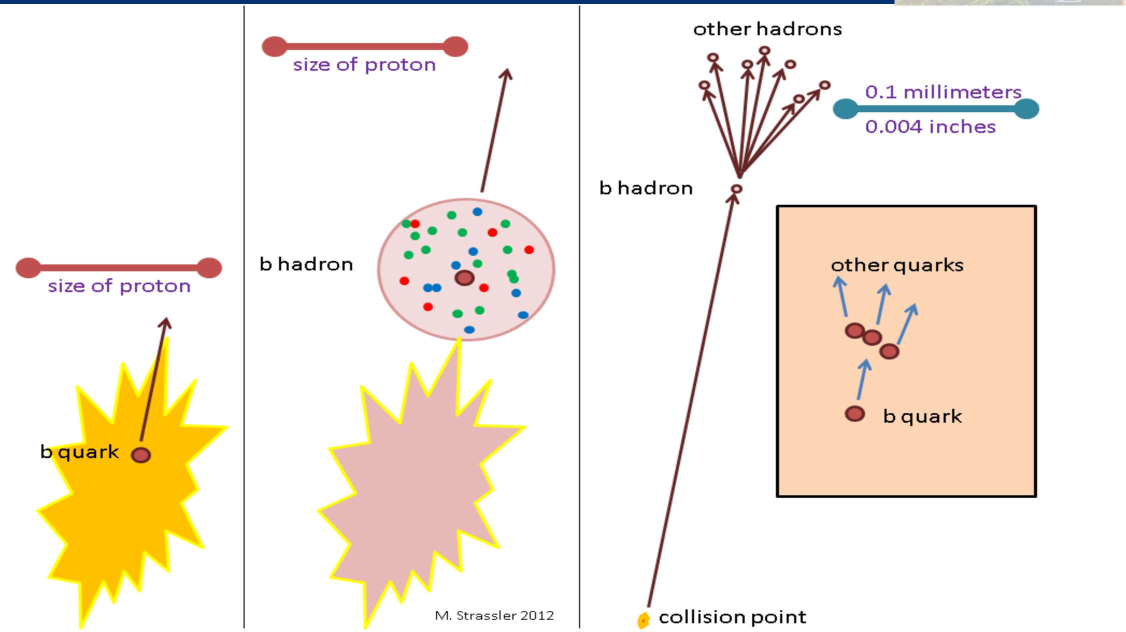
Amplifier Housing

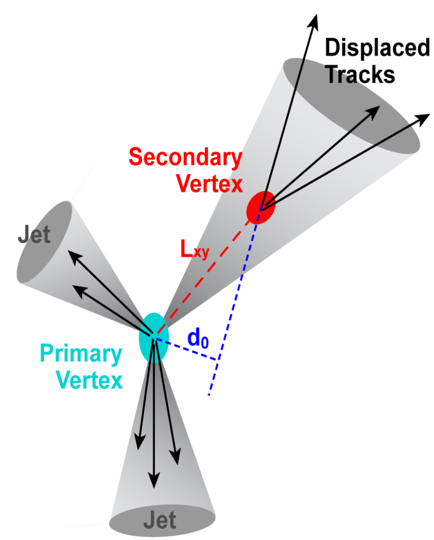


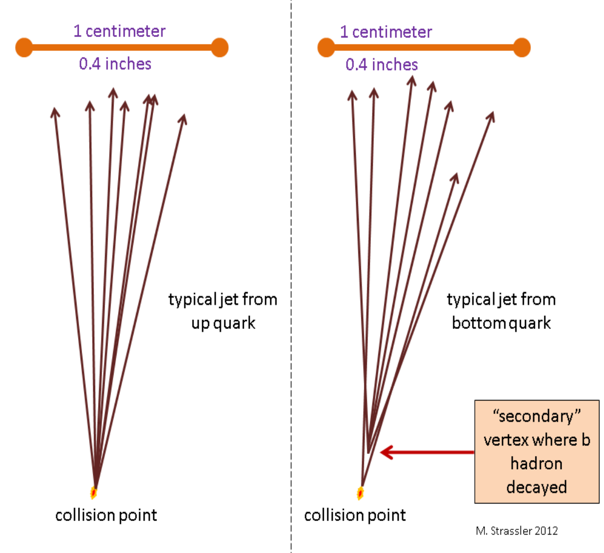
NI PXI crate, DAQ for automated measurements

b-Tagging

- CKM matrix has top->bottom quark at 99.9%
- b-quarks have low probability of decaying to other quarks, on average, longer decays
- Macroscopically observable time of flight - observe a displaced secondary vertex
- Several algorithms available to attempt to “tag”, a jet as being from a b-quark
- Very important for ID of top-quarks, due to high CKM mixing probability

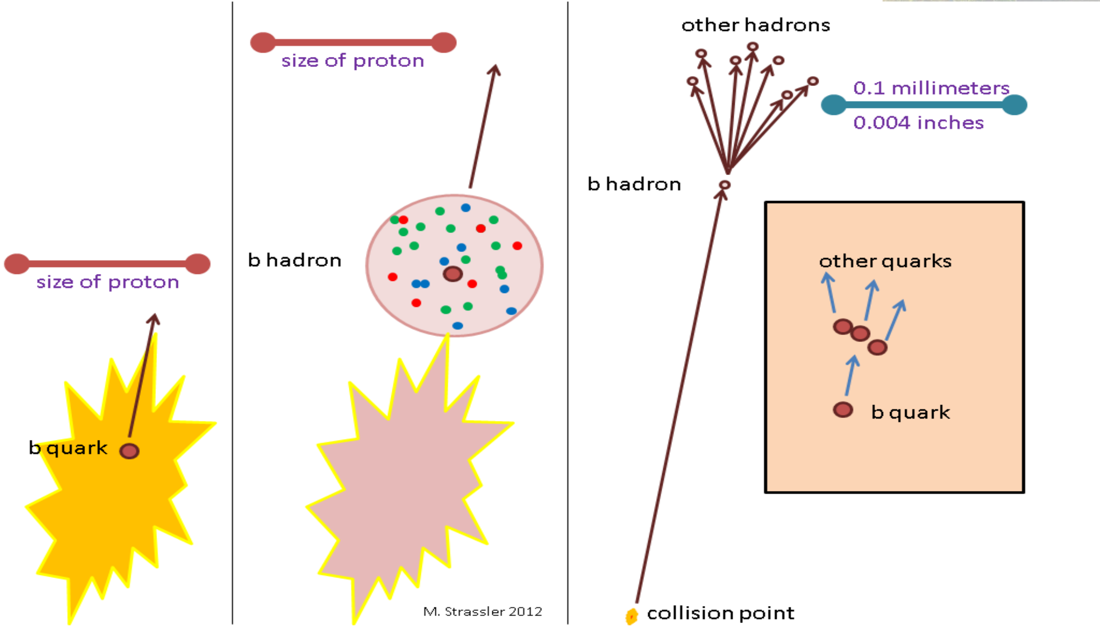


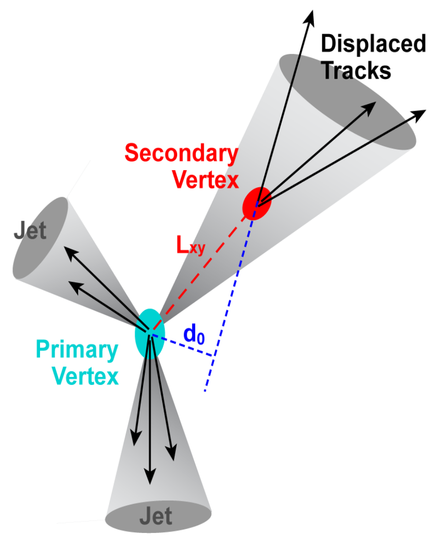


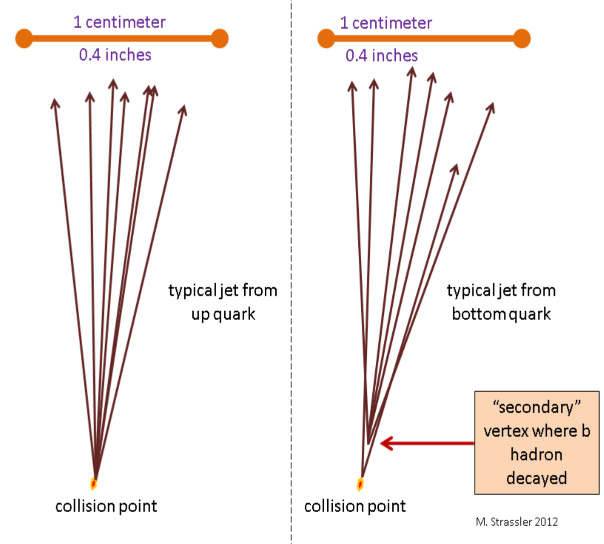


b-Tagging

- **Combined Secondary Vertex (CSV)** is our chosen algorithm. MVA with input variables:
  - invariant mass of 2^{nd} vertex
  - jet multiplicity
  - significance of 2^{nd} vertex distance to primary
  - ratio of energies of $1^{st}/2^{nd}$ vertex
  - rapidities of particles associated with 2^{nd} vertex
  - track impact significance of first track exceeding charm threshold (1.5GeV)

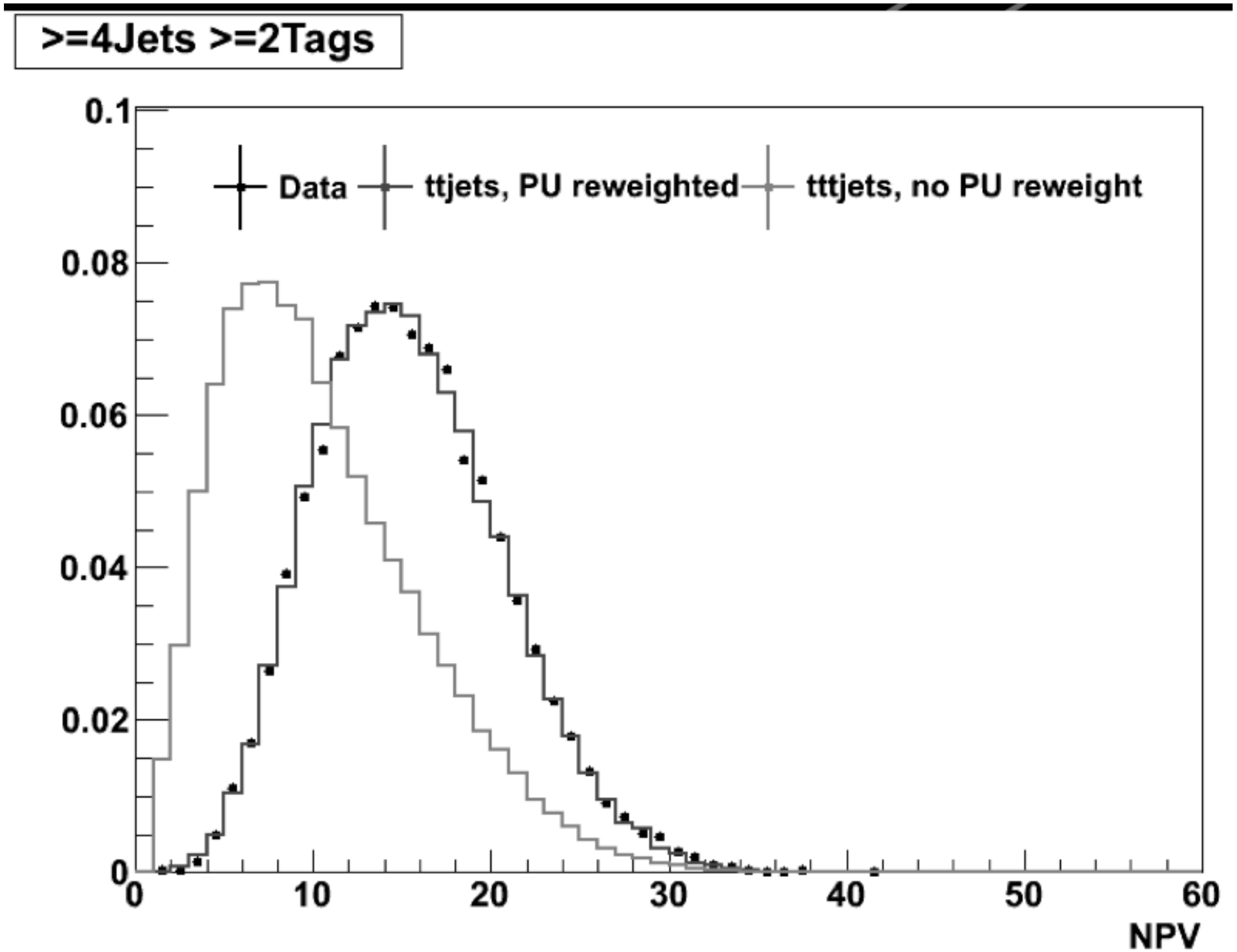






Pileup Reweighting – 5 fb^{-1} 8 TeV

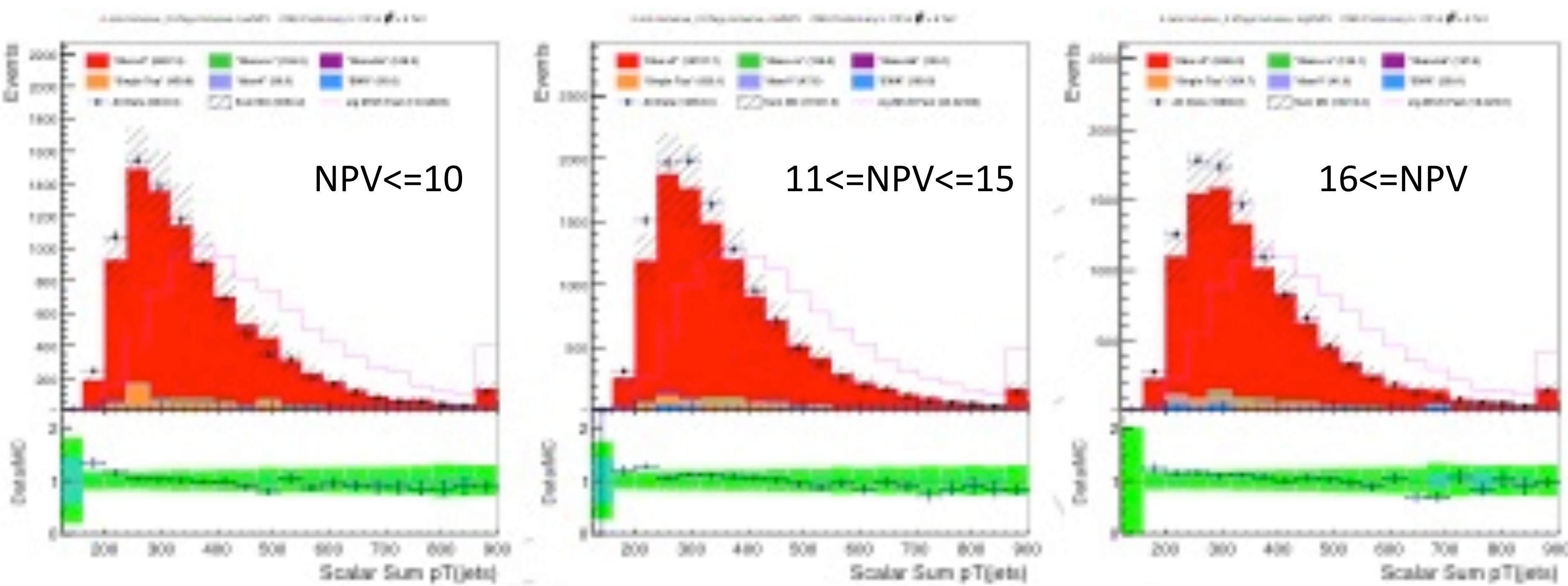
- Left – Distribution of Number of Primary Vertices for each event for tt+jets enriched sample
- Red – Monte Carlo is generated with a different pileup profile than data, so events are re-weighted using a sample of min-bias events from data
- Blue – Reweighted monte carlo sample of tt+jets enriched events
- Black – Data from a sample of tt+jets enriched events



Additional Pileup Reweighting – 5 fb^{-1} 8 TeV

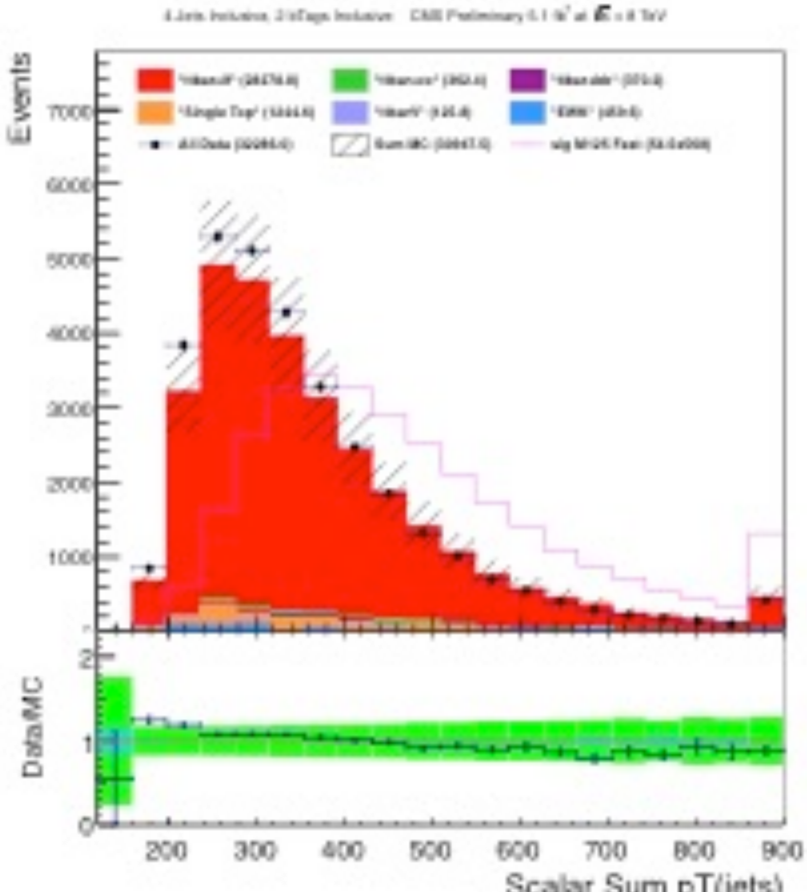
- Scalar Sum of all Jets in the event (Ht) – for different pileup binnings
  - After PF2PAT overlap removal, charged hadron subtraction, and jet energy corrections are applied, distinct trend in discrepancy between data and monte carlo remains

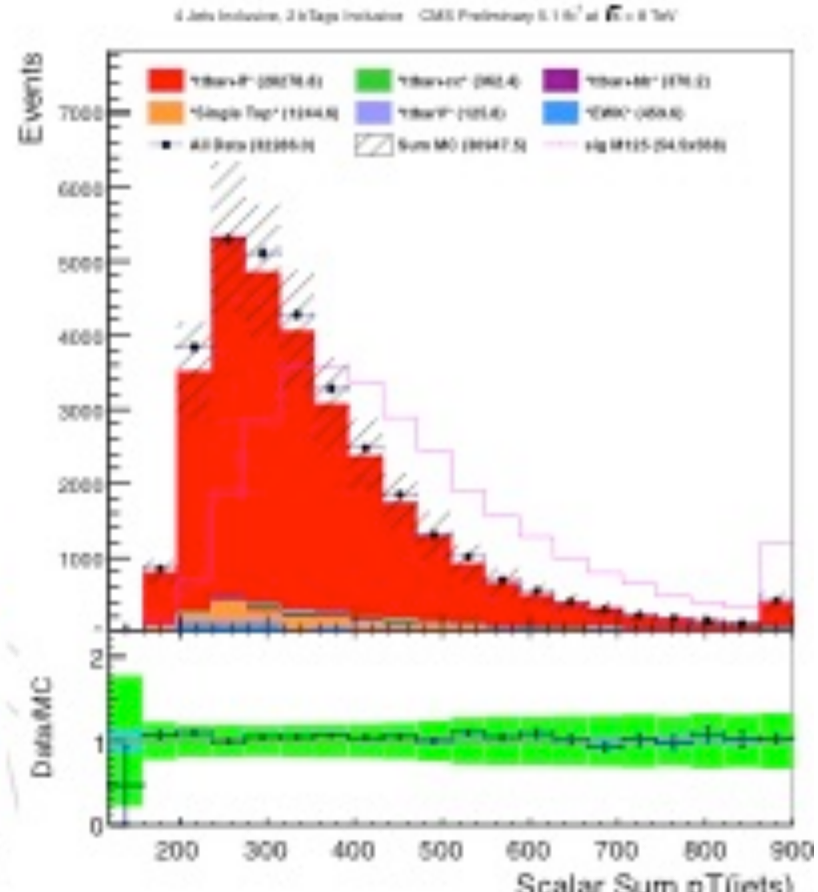
- Correction based on >=4Jets, >=2Tags, Lepton+Jets data and monte carlo comparison
  - Bin-by-bin correction factor in each of the three pileup bins



Additional Pileup Reweighting – 5 fb^{-1} 8 TeV

- Below: Left – before Ht correction; Right – after Ht correction
  - Correction factors derived with variable binning for sufficient stats in tail region, so plot on bottom right is not exactly flat
  - S/B is 0.002, so no worry of signal bias
  - Additional systematic uncertainty for given by no corrections and x2 correction factor

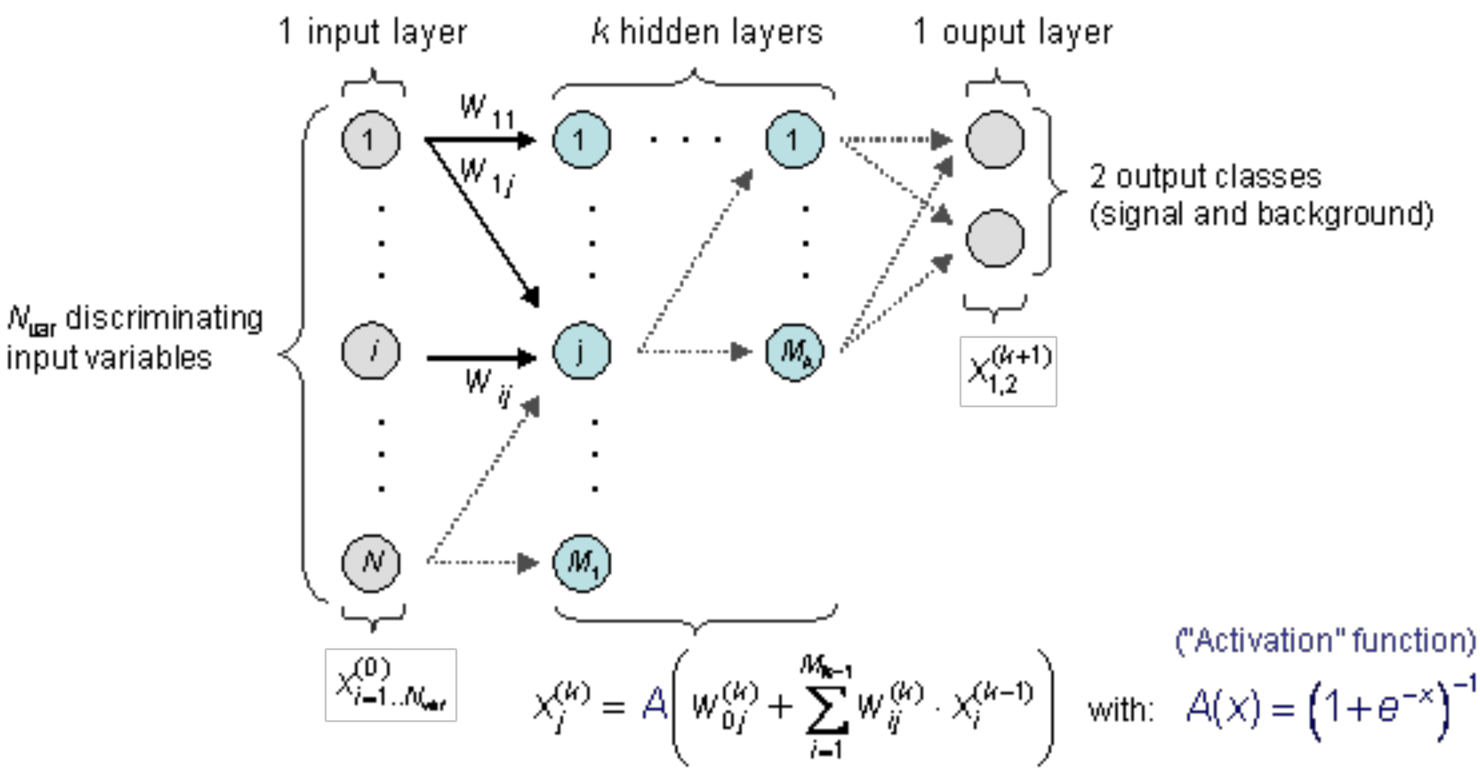




Multi-Variate Analysis Techniques

- **Artificial Neural Network (ANN)** – specific implementation of an MVA. Ours uses 10 input variables, and two hidden layers with 10, 9 nodes respectively

- **Training events** are given to the ANN from a known signal and background source, ANN output is compared to desired result (0=background, 1=signal) and adjusts weights to try and maximize output towards signal/background-like response



Systematic Uncertainties - Lumi

- Luminosity is calculated by using the inner pixel detector
  - 7e7 pixels used
  - Low occupancy per bunch crossing, one pixel for one track
  - # of pixels is linearly related to # of interactions per bunch crosses

- Sources of uncertainty
  - Physical distribution of pixel clusters
  - Resolution on interaction region (length scale calibration)
  - Uncertainties on Beam characteristics
    - Width Evolution
    - BetaStar
    - Scan-to-Scan uncertainties
  - Afterglow

- *Below:* Table of complete list of uncertainties, % change to normalization, and whether or not uncertainty is used in shape comparison

Source	Rate Uncertainty	Shape	Remarks
Luminosity (7 TeV)	2.2%	No	All signal and backgrounds
Luminosity (8 TeV)	4.4%	No	All signal and backgrounds
Lepton ID/Trig	4%	No	All signal and backgrounds
Pileup	1%	No	All signal and backgrounds
Additional Pileup Corr.	–	Yes	All signal and backgrounds
Jet Energy Resolution	1.5%	No	All signal and backgrounds
Jet Energy Scale	0-60%	Yes	All signal and backgrounds
b-Tag SF (b/c)	0-33.6%	Yes	All signal and backgrounds
b-Tag SF (mistag)	0-23.5%	Yes	All signal and backgrounds
MC Statistics	–	Yes	All backgrounds
PDF (gg)	9%	No	For gg initiated processes ($t\bar{t}$, $t\bar{t}Z$, $t\bar{t}H$)
PDF ($q\bar{q}$)	4.2-7%	No	For $q\bar{q}$ initiated processes ($t\bar{t}W$, W, Z).
PDF (qg)	4.6%	No	For qg initiated processes (single top)
QCD Scale ($t\bar{t}H$)	15%	No	For NLO $t\bar{t}H$ prediction
QCD Scale ($t\bar{t}$)	2-12%	No	For NLO $t\bar{t}$ and single top predictions
QCD Scale (V)	1.2-1.3%	No	For NNLO W and Z prediction
QCD Scale (VV)	3.5%	No	For NLO diboson prediction
Madgraph Scale ($t\bar{t}$)	0-20%	Yes	$t\bar{t}$+jets/$b\bar{b}$/$c\bar{c}$ uncorrelated. Varies by jet bin.
Madgraph Scale (V)	20-60%	No	Varies by jet bin.
$t\bar{t}+b\bar{b}$	50%	No	Only $t\bar{t}+b\bar{b}$.

Systematic Uncertainties – Lepton ID

- Lepton ID/Trigger Efficiency uncertainty
  - Tag & Probe Method
  - Select two electrons/muons in an invariant mass window around Z peak
  - Choose events where one electron/muon passes trigger and ID cuts, this is the "tag"
  - Assume partner is identical (since both should be coming form Z decay)
  - Check is partner passes trigger/ID, this is "probe"
  - Efficiency = #probe/#tag which pass trigger ID selection
- Sources of uncertainty
  - Purity of sample
  - Statistical uncertainty

- *Below:* Table of complete list of uncertainties, % change to normalization, and whether or not uncertainty is used in shape comparison

Source	Rate Uncertainty	Shape	Remarks
Luminosity (7 TeV)	2.2%	No	All signal and backgrounds
Luminosity (8 TeV)	4.4%	No	All signal and backgrounds
Lepton ID/Trig	4%	No	All signal and backgrounds
Pileup	1%	No	All signal and backgrounds
Additional Pileup Corr.	–	Yes	All signal and backgrounds
Jet Energy Resolution	1.5%	No	All signal and backgrounds
Jet Energy Scale	0-60%	Yes	All signal and backgrounds
b-Tag SF (b/c)	0-33.6%	Yes	All signal and backgrounds
b-Tag SF (mistag)	0-23.5%	Yes	All signal and backgrounds
MC Statistics	–	Yes	All backgrounds
PDF (gg)	9%	No	For gg initiated processes ($t\bar{t}$, $t\bar{t}Z$, $t\bar{t}H$)
PDF ($q\bar{q}$)	4.2-7%	No	For $q\bar{q}$ initiated processes ($t\bar{t}W$, W, Z).
PDF (qg)	4.6%	No	For qg initiated processes (single top)
QCD Scale ($t\bar{t}H$)	15%	No	For NLO $t\bar{t}H$ prediction
QCD Scale ($t\bar{t}$)	2-12%	No	For NLO $t\bar{t}$ and single top predictions
QCD Scale (V)	1.2-1.3%	No	For NNLO W and Z prediction
QCD Scale (VV)	3.5%	No	For NLO diboson prediction
Madgraph Scale ($t\bar{t}$)	0-20%	Yes	$t\bar{t}$+jets/$b\bar{b}$/$c\bar{c}$ uncorrelated. Varies by jet bin.
Madgraph Scale (V)	20-60%	No	Varies by jet bin.
$t\bar{t}+b\bar{b}$	50%	No	Only $t\bar{t}+b\bar{b}$.

Systematic Uncertainties – Pileup

- Pileup is the effect of having multiple interaction points per bunch crossing
  - MC is generated with an arbitrary number of pileup events
  - Has to be adjusted to match data distribution
  - Use control region in data from your HLT to get pileup distribution from data
  - Scale MC to match
- Sources of uncertainty
  - Min bias Xsection
    - Number of actual pileup of events in each measurement
- *Below:* Table of complete list of uncertainties, % change to normalization, and whether or not uncertainty is used in shape comparison

Source	Rate Uncertainty	Shape	Remarks
Luminosity (7 TeV)	2.2%	No	All signal and backgrounds
Luminosity (8 TeV)	4.4%	No	All signal and backgrounds
Lepton ID/Trig	4%	No	All signal and backgrounds
Pileup	1%	No	All signal and backgrounds
Additional Pileup Corr.	–	Yes	All signal and backgrounds
Jet Energy Resolution	1.5%	No	All signal and backgrounds
Jet Energy Scale	0-60%	Yes	All signal and backgrounds
b-Tag SF (b/c)	0-33.6%	Yes	All signal and backgrounds
b-Tag SF (mistag)	0-23.5%	Yes	All signal and backgrounds
MC Statistics	–	Yes	All backgrounds
PDF (gg)	9%	No	For gg initiated processes ($t\bar{t}$, $t\bar{t}$Z, $t\bar{t}$H)
PDF ($q\bar{q}$)	4.2-7%	No	For $q\bar{q}$ initiated processes ($t\bar{t}$W, W, Z).
PDF (qg)	4.6%	No	For qg initiated processes (single top)
QCD Scale ($t\bar{t}$H)	15%	No	For NLO $t\bar{t}H$ prediction
QCD Scale ($t\bar{t}$)	2-12%	No	For NLO $t\bar{t}$ and single top predictions
QCD Scale (V)	1.2-1.3%	No	For NNLO W and Z prediction
QCD Scale (VV)	3.5%	No	For NLO diboson prediction
Madgraph Scale ($t\bar{t}$)	0-20%	Yes	$t\bar{t}$+jets/$b\bar{b}$/$c\bar{c}$ uncorrelated. Varies by jet bin.
Madgraph Scale (V)	20-60%	No	Varies by jet bin.
$t\bar{t}+b\bar{b}$	50%	No	Only $t\bar{t}+b\bar{b}$.

Systematic Uncertainties – Add. Pileup

- Jet pT distributions showed a sloped bias in the data to MC distributions which varied over Number of Primary Vertex (NPV) regions
- Use Scalar Sum of all jet pTs in event, Ht
- In inclusive selection, reweight this distribution to match data
- Weights distributed to jet/tag categories
- Yields normalized to # of events before this reweighting
- Uncertainty comes from applying no correction to applying x2 the correction

- *Below:* Table of complete list of uncertainties, % change to normalization, and whether or not uncertainty is used in shape comparison

Source	Rate Uncertainty	Shape	Remarks
Luminosity (7 TeV)	2.2%	No	All signal and backgrounds
Luminosity (8 TeV)	4.4%	No	All signal and backgrounds
Lepton ID/Trig	4%	No	All signal and backgrounds
Pileup	1%	No	All signal and backgrounds
Additional Pileup Corr.	–	Yes	All signal and backgrounds
Jet Energy Resolution	1.5%	No	All signal and backgrounds
Jet Energy Scale	0-60%	Yes	All signal and backgrounds
b-Tag SF (b/c)	0-33.6%	Yes	All signal and backgrounds
b-Tag SF (mistag)	0-23.5%	Yes	All signal and backgrounds
MC Statistics	–	Yes	All backgrounds
PDF (gg)	9%	No	For gg initiated processes ($t\bar{t}$, $t\bar{t}Z$, $t\bar{t}H$)
PDF ($q\bar{q}$)	4.2-7%	No	For $q\bar{q}$ initiated processes ($t\bar{t}W$, W, Z).
PDF (qg)	4.6%	No	For qg initiated processes (single top)
QCD Scale ($t\bar{t}H$)	15%	No	For NLO $t\bar{t}H$ prediction
QCD Scale ($t\bar{t}$)	2-12%	No	For NLO $t\bar{t}$ and single top predictions
QCD Scale (V)	1.2-1.3%	No	For NNLO W and Z prediction
QCD Scale (VV)	3.5%	No	For NLO diboson prediction
Madgraph Scale ($t\bar{t}$)	0-20%	Yes	$t\bar{t}$+jets/$b\bar{b}$/$c\bar{c}$ uncorrelated. Varies by jet bin.
Madgraph Scale (V)	20-60%	No	Varies by jet bin.
$t\bar{t}+b\bar{b}$	50%	No	Only $t\bar{t}+b\bar{b}$.

Systematic Uncertainties – JER

- Jet Energy Resolution (JER) is different between MC and data
- Need a scale factor for jet 4-vector, to "smear" resolution of jet energy and momentum in MC to match data
- Photon+Jet, Z+Jet, and diJet events are used to study energy resolution in pT, eta, phi bins in the detector
- Uncertainty comes from
  - Purity of sample
  - Uncertainty on invariant mass fits in resolution analysis

- *Below:* Table of complete list of uncertainties, % change to normalization, and whether or not uncertainty is used in shape comparison

Source	Rate Uncertainty	Shape	Remarks
Luminosity (7 TeV)	2.2%	No	All signal and backgrounds
Luminosity (8 TeV)	4.4%	No	All signal and backgrounds
Lepton ID/Trig	4%	No	All signal and backgrounds
Pileup	1%	No	All signal and backgrounds
Additional Pileup Corr.	–	Yes	All signal and backgrounds
Jet Energy Resolution	1.5%	No	All signal and backgrounds
Jet Energy Scale	0-60%	Yes	All signal and backgrounds
b-Tag SF (b/c)	0-33.6%	Yes	All signal and backgrounds
b-Tag SF (mistag)	0-23.5%	Yes	All signal and backgrounds
MC Statistics	–	Yes	All backgrounds
PDF (gg)	9%	No	For gg initiated processes ($t\bar{t}$, $t\bar{t}Z$, $t\bar{t}H$)
PDF ($q\bar{q}$)	4.2-7%	No	For $q\bar{q}$ initiated processes ($t\bar{t}W$, W, Z).
PDF (qg)	4.6%	No	For qg initiated processes (single top)
QCD Scale ($t\bar{t}H$)	15%	No	For NLO $t\bar{t}H$ prediction
QCD Scale ($t\bar{t}$)	2-12%	No	For NLO $t\bar{t}$ and single top predictions
QCD Scale (V)	1.2-1.3%	No	For NNLO W and Z prediction
QCD Scale (VV)	3.5%	No	For NLO diboson prediction
Madgraph Scale ($t\bar{t}$)	0-20%	Yes	$t\bar{t}$+jets/$b\bar{b}$/$c\bar{c}$ uncorrelated. Varies by jet bin.
Madgraph Scale (V)	20-60%	No	Varies by jet bin.
$t\bar{t}+b\bar{b}$	50%	No	Only $t\bar{t}+b\bar{b}$.

Systematic Uncertainties – JES

- Jet Energy Scale corrections are based on corrections made to jet energies during reconstruction
- Energy offset (L1 correction) is performed to remove energy from pileup deposits
- Eta correction (L2 correction) performed to ensure uniform detector response in eta
- Pt correction (L3 correction) performed to ensure uniform detector response from kinematically similar events
- Corrections derived by studying Z-jetjet using tag & probe, and other diJet processes
- Uncertainty determined by varying Jet Energy Scale correction factors +/- 1 sigma and reproducing ANN output shapes

- *Below:* Table of complete list of uncertainties, % change to normalization, and whether or not uncertainty is used in shape comparison

Source	Rate Uncertainty	Shape	Remarks
Luminosity (7 TeV)	2.2%	No	All signal and backgrounds
Luminosity (8 TeV)	4.4%	No	All signal and backgrounds
Lepton ID/Trig	4%	No	All signal and backgrounds
Pileup	1%	No	All signal and backgrounds
Additional Pileup Corr.	–	Yes	All signal and backgrounds
Jet Energy Resolution	1.5%	No	All signal and backgrounds
Jet Energy Scale	0-60%	Yes	All signal and backgrounds
b-Tag SF (b/c)	0-33.6%	Yes	All signal and backgrounds
b-Tag SF (mistag)	0-23.5%	Yes	All signal and backgrounds
MC Statistics	–	Yes	All backgrounds
PDF (gg)	9%	No	For gg initiated processes ($t\bar{t}$, $t\bar{t}Z$, $t\bar{t}H$)
PDF ($q\bar{q}$)	4.2-7%	No	For $q\bar{q}$ initiated processes ($t\bar{t}W$, W, Z).
PDF (qg)	4.6%	No	For qg initiated processes (single top)
QCD Scale ($t\bar{t}H$)	15%	No	For NLO $t\bar{t}H$ prediction
QCD Scale ($t\bar{t}$)	2-12%	No	For NLO $t\bar{t}$ and single top predictions
QCD Scale (V)	1.2-1.3%	No	For NNLO W and Z prediction
QCD Scale (VV)	3.5%	No	For NLO diboson prediction
Madgraph Scale ($t\bar{t}$)	0-20%	Yes	$t\bar{t}$+jets/$b\bar{b}$/$c\bar{c}$ uncorrelated. Varies by jet bin.
Madgraph Scale (V)	20-60%	No	Varies by jet bin.
$t\bar{t}+b\bar{b}$	50%	No	Only $t\bar{t}+b\bar{b}$.

Systematic Uncertainties – bTag hfSF

- Begin with high purity sample of bQuarks
- Muon-Jet method, multi-jet production
  - Strong mixing to top-quark gives larger semi-leptonic BR vs other quarks
  - Look for jets with a muon within deltaR<0.4
  - Alternative use high-purity ttbar sample, but we need those events in our analysis
- Several methods to determine efficiency
  - "System8", tag & probe style method, ID a muon-jet as probe, use jet passing some number of alt. tagging algos
- Measure efficiency in data to be ~70%, need MC to match this
- Uncertainties due to purity of sample, pileup, gluon splitting, muon ID, bias of other tagger

- *Below:* Table of complete list of uncertainties, % change to normalization, and whether or not uncertainty is used in shape comparison

Source	Rate Uncertainty	Shape	Remarks
Luminosity (7 TeV)	2.2%	No	All signal and backgrounds
Luminosity (8 TeV)	4.4%	No	All signal and backgrounds
Lepton ID/Trig	4%	No	All signal and backgrounds
Pileup	1%	No	All signal and backgrounds
Additional Pileup Corr.	–	Yes	All signal and backgrounds
Jet Energy Resolution	1.5%	No	All signal and backgrounds
Jet Energy Scale	0-60%	Yes	All signal and backgrounds
b-Tag SF (b/c)	0-33.6%	Yes	All signal and backgrounds
b-Tag SF (mistag)	0-23.5%	Yes	All signal and backgrounds
MC Statistics	–	Yes	All backgrounds
PDF (gg)	9%	No	For gg initiated processes ($t\bar{t}$, $t\bar{t}Z$, $t\bar{t}H$)
PDF ($q\bar{q}$)	4.2-7%	No	For $q\bar{q}$ initiated processes ($t\bar{t}W$, W, Z).
PDF (qg)	4.6%	No	For qg initiated processes (single top)
QCD Scale ($t\bar{t}H$)	15%	No	For NLO $t\bar{t}H$ prediction
QCD Scale ($t\bar{t}$)	2-12%	No	For NLO $t\bar{t}$ and single top predictions
QCD Scale (V)	1.2-1.3%	No	For NNLO W and Z prediction
QCD Scale (VV)	3.5%	No	For NLO diboson prediction
Madgraph Scale ($t\bar{t}$)	0-20%	Yes	$t\bar{t}$+jets/$b\bar{b}$/$c\bar{c}$ uncorrelated. Varies by jet bin.
Madgraph Scale (V)	20-60%	No	Varies by jet bin.
$t\bar{t}+b\bar{b}$	50%	No	Only $t\bar{t}+b\bar{b}$.

Systematic Uncertainties – bTag lfSF

- Similar procedure to calculate mis-ID SF uncertainty, ie uncertainty on SF for lf jets tagged as b-Jets
- Invert away-jet tagger cut, still probe with muon jet
- Systematic uncertainties
  - b/c quark fractions
  - Gluon fractions
  - Long-lived meson decays
  - Mis-ID on tracks
  - Pileup
  - Purity of sample
- Vary results of SF due to these different uncertainties, get +/- 1 sigma on SF, recreate ANN distributions
- Probability to mis-ID lf jet as b for CSVM ~1.5%

- *Below:* Table of complete list of uncertainties, % change to normalization, and whether or not uncertainty is used in shape comparison

Source	Rate Uncertainty	Shape	Remarks
Luminosity (7 TeV)	2.2%	No	All signal and backgrounds
Luminosity (8 TeV)	4.4%	No	All signal and backgrounds
Lepton ID/Trig	4%	No	All signal and backgrounds
Pileup	1%	No	All signal and backgrounds
Additional Pileup Corr.	–	Yes	All signal and backgrounds
Jet Energy Resolution	1.5%	No	All signal and backgrounds
Jet Energy Scale	0-60%	Yes	All signal and backgrounds
b-Tag SF (b/c)	0-33.6%	Yes	All signal and backgrounds
b-Tag SF (mistag)	0-23.5%	Yes	All signal and backgrounds
MC Statistics	–	Yes	All backgrounds
PDF (gg)	9%	No	For gg initiated processes ($t\bar{t}$, $t\bar{t}Z$, $t\bar{t}H$)
PDF ($q\bar{q}$)	4.2-7%	No	For $q\bar{q}$ initiated processes ($t\bar{t}W$, W, Z).
PDF (qg)	4.6%	No	For qg initiated processes (single top)
QCD Scale ($t\bar{t}H$)	15%	No	For NLO $t\bar{t}H$ prediction
QCD Scale ($t\bar{t}$)	2-12%	No	For NLO $t\bar{t}$ and single top predictions
QCD Scale (V)	1.2-1.3%	No	For NNLO W and Z prediction
QCD Scale (VV)	3.5%	No	For NLO diboson prediction
Madgraph Scale ($t\bar{t}$)	0-20%	Yes	$t\bar{t}$+jets/$b\bar{b}$/$c\bar{c}$ uncorrelated. Varies by jet bin.
Madgraph Scale (V)	20-60%	No	Varies by jet bin.
$t\bar{t}+b\bar{b}$	50%	No	Only $t\bar{t}+b\bar{b}$.

Systematic Uncertainties – PDF

- Uncertainty from knowledge on initial Parton Distribution Function of protons being collided
- Studied from earlier Tevatron, and earlier runs at CMS' knowledge of the proton
- Initial state energies of particles are essentially varied, giving different rates for background processes
- Adjust rates +/- 1 sigma from quoted uncertainty used across CMS

- *Below:* Table of complete list of uncertainties, % change to normalization, and whether or not uncertainty is used in shape comparison

Source	Rate Uncertainty	Shape	Remarks
Luminosity (7 TeV)	2.2%	No	All signal and backgrounds
Luminosity (8 TeV)	4.4%	No	All signal and backgrounds
Lepton ID/Trig	4%	No	All signal and backgrounds
Pileup	1%	No	All signal and backgrounds
Additional Pileup Corr.	–	Yes	All signal and backgrounds
Jet Energy Resolution	1.5%	No	All signal and backgrounds
Jet Energy Scale	0-60%	Yes	All signal and backgrounds
b-Tag SF (b/c)	0-33.6%	Yes	All signal and backgrounds
b-Tag SF (mistag)	0-23.5%	Yes	All signal and backgrounds
MC Statistics	–	Yes	All backgrounds
PDF (gg)	9%	No	For gg initiated processes ($t\bar{t}$, $t\bar{t}Z$, $t\bar{t}H$)
PDF ($q\bar{q}$)	4.2-7%	No	For $q\bar{q}$ initiated processes ($t\bar{t}W$, W, Z).
PDF (qg)	4.6%	No	For qg initiated processes (single top)
QCD Scale ($t\bar{t}H$)	15%	No	For NLO $t\bar{t}H$ prediction
QCD Scale ($t\bar{t}$)	2-12%	No	For NLO $t\bar{t}$ and single top predictions
QCD Scale (V)	1.2-1.3%	No	For NNLO W and Z prediction
QCD Scale (VV)	3.5%	No	For NLO diboson prediction
Madgraph Scale ($t\bar{t}$)	0-20%	Yes	$t\bar{t}$+jets/$b\bar{b}$/$c\bar{c}$ uncorrelated. Varies by jet bin.
Madgraph Scale (V)	20-60%	No	Varies by jet bin.
$t\bar{t}+b\bar{b}$	50%	No	Only $t\bar{t}+b\bar{b}$.

Systematic Uncertainties – QCD Scale

- Theoretical uncertainties on production cross-sections from theory
- Estimated by considering effects of next order in perturbation theory on results from current level of calculation
- Processes Effected:
  - ttH (signal)
  - tt+lf,c,b
  - V+X events
  - VV+X events

- *Below:* Table of complete list of uncertainties, % change to normalization, and whether or not uncertainty is used in shape comparison

Source	Rate Uncertainty	Shape	Remarks
Luminosity (7 TeV)	2.2%	No	All signal and backgrounds
Luminosity (8 TeV)	4.4%	No	All signal and backgrounds
Lepton ID/Trig	4%	No	All signal and backgrounds
Pileup	1%	No	All signal and backgrounds
Additional Pileup Corr.	–	Yes	All signal and backgrounds
Jet Energy Resolution	1.5%	No	All signal and backgrounds
Jet Energy Scale	0-60%	Yes	All signal and backgrounds
b-Tag SF (b/c)	0-33.6%	Yes	All signal and backgrounds
b-Tag SF (mistag)	0-23.5%	Yes	All signal and backgrounds
MC Statistics	–	Yes	All backgrounds
PDF (gg)	9%	No	For gg initiated processes ($t\bar{t}$, $t\bar{t}Z$, $t\bar{t}H$)
PDF ($q\bar{q}$)	4.2-7%	No	For $q\bar{q}$ initiated processes ($t\bar{t}W$, W, Z).
PDF (qg)	4.6%	No	For qg initiated processes (single top)
QCD Scale ($t\bar{t}H$)	15%	No	For NLO $t\bar{t}H$ prediction
QCD Scale ($t\bar{t}$)	2-12%	No	For NLO $t\bar{t}$ and single top predictions
QCD Scale (V)	1.2-1.3%	No	For NNLO W and Z prediction
QCD Scale (VV)	3.5%	No	For NLO diboson prediction
Madgraph Scale ($t\bar{t}$)	0-20%	Yes	$t\bar{t}$+jets/$b\bar{b}$/$c\bar{c}$ uncorrelated. Varies by jet bin.
Madgraph Scale (V)	20-60%	No	Varies by jet bin.
$t\bar{t}+b\bar{b}$	50%	No	Only $t\bar{t}+b\bar{b}$.

Systematic Uncertainties – Q^2

- Q^2 is the factorization and renormalization uncertainties from using MC
- Uncertainty from not matching data exactly, thus energy scale, ie strength of coupling constants, kinematic distributions from decays are all affected
- Vary Q^2 in the initial MC samples, reproduce ANN output distribution to have rate and shape information about how ANN should change due to this effect

- *Below:* Table of complete list of uncertainties, % change to normalization, and whether or not uncertainty is used in shape comparison

Source	Rate Uncertainty	Shape	Remarks
Luminosity (7 TeV)	2.2%	No	All signal and backgrounds
Luminosity (8 TeV)	4.4%	No	All signal and backgrounds
Lepton ID/Trig	4%	No	All signal and backgrounds
Pileup	1%	No	All signal and backgrounds
Additional Pileup Corr.	–	Yes	All signal and backgrounds
Jet Energy Resolution	1.5%	No	All signal and backgrounds
Jet Energy Scale	0-60%	Yes	All signal and backgrounds
b-Tag SF (b/c)	0-33.6%	Yes	All signal and backgrounds
b-Tag SF (mistag)	0-23.5%	Yes	All signal and backgrounds
MC Statistics	–	Yes	All backgrounds
PDF (gg)	9%	No	For gg initiated processes ($t\bar{t}$, $t\bar{t}Z$, $t\bar{t}H$)
PDF ($q\bar{q}$)	4.2-7%	No	For $q\bar{q}$ initiated processes ($t\bar{t}W$, W, Z).
PDF (qg)	4.6%	No	For qg initiated processes (single top)
QCD Scale ($t\bar{t}H$)	15%	No	For NLO $t\bar{t}H$ prediction
QCD Scale ($t\bar{t}$)	2-12%	No	For NLO $t\bar{t}$ and single top predictions
QCD Scale (V)	1.2-1.3%	No	For NNLO W and Z prediction
QCD Scale (VV)	3.5%	No	For NLO diboson prediction
Madgraph Scale ($t\bar{t}$)	0-20%	Yes	$t\bar{t}$+jets/bb/$c\bar{c}$ uncorrelated. Varies by jet bin.
Madgraph Scale (V)	20-60%	No	Varies by jet bin.
$t\bar{t}$ + bb	50%	No	Only $t\bar{t}$ + bb.

Systematic Uncertainties – tt+bb

- Since the tt+bb background is kinematically very similar to our signal, it is hard to discriminate against
- In order to be conservative in our estimates of ttH production, we added an additional 50% uncertainty on top of the existing cross-section uncertainty

- *Below:* Table of complete list of uncertainties, % change to normalization, and whether or not uncertainty is used in shape comparison

Source	Rate Uncertainty	Shape	Remarks
Luminosity (7 TeV)	2.2%	No	All signal and backgrounds
Luminosity (8 TeV)	4.4%	No	All signal and backgrounds
Lepton ID/Trig	4%	No	All signal and backgrounds
Pileup	1%	No	All signal and backgrounds
Additional Pileup Corr.	–	Yes	All signal and backgrounds
Jet Energy Resolution	1.5%	No	All signal and backgrounds
Jet Energy Scale	0-60%	Yes	All signal and backgrounds
b-Tag SF (b/c)	0-33.6%	Yes	All signal and backgrounds
b-Tag SF (mistag)	0-23.5%	Yes	All signal and backgrounds
MC Statistics	–	Yes	All backgrounds
PDF (gg)	9%	No	For gg initiated processes ($t\bar{t}$, $t\bar{t}Z$, $t\bar{t}H$)
PDF ($q\bar{q}$)	4.2-7%	No	For $q\bar{q}$ initiated processes ($t\bar{t}W$, W, Z).
PDF (qg)	4.6%	No	For qg initiated processes (single top)
QCD Scale ($t\bar{t}H$)	15%	No	For NLO $t\bar{t}H$ prediction
QCD Scale ($t\bar{t}$)	2-12%	No	For NLO $t\bar{t}$ and single top predictions
QCD Scale (V)	1.2-1.3%	No	For NNLO W and Z prediction
QCD Scale (VV)	3.5%	No	For NLO diboson prediction
Madgraph Scale ($t\bar{t}$)	0-20%	Yes	$t\bar{t}$+jets/$b\bar{b}$/$c\bar{c}$ uncorrelated. Varies by jet bin.
Madgraph Scale (V)	20-60%	No	Varies by jet bin.
$t\bar{t}+b\bar{b}$	50%	No	Only $t\bar{t}+b\bar{b}$.

Limit Setting Technique

- In the case of no clear signal, ask what is the upper limit on ttH production possible that can still be explained by this dataset

- CLs is modified frequentist technique (~Bayesian, relies on background prediction)

- Templates for signal and background processes, nominal and +/- 1σ for shape uncertainties

- Perform pseudoexperiments - vary template shapes within constraints of systematics

- Test signal+background and background only hypothesis by performing a fit across each jet/tag category simultaneously
  - Signal-poor categories constrain background fit in signal-rich jet/tag categories

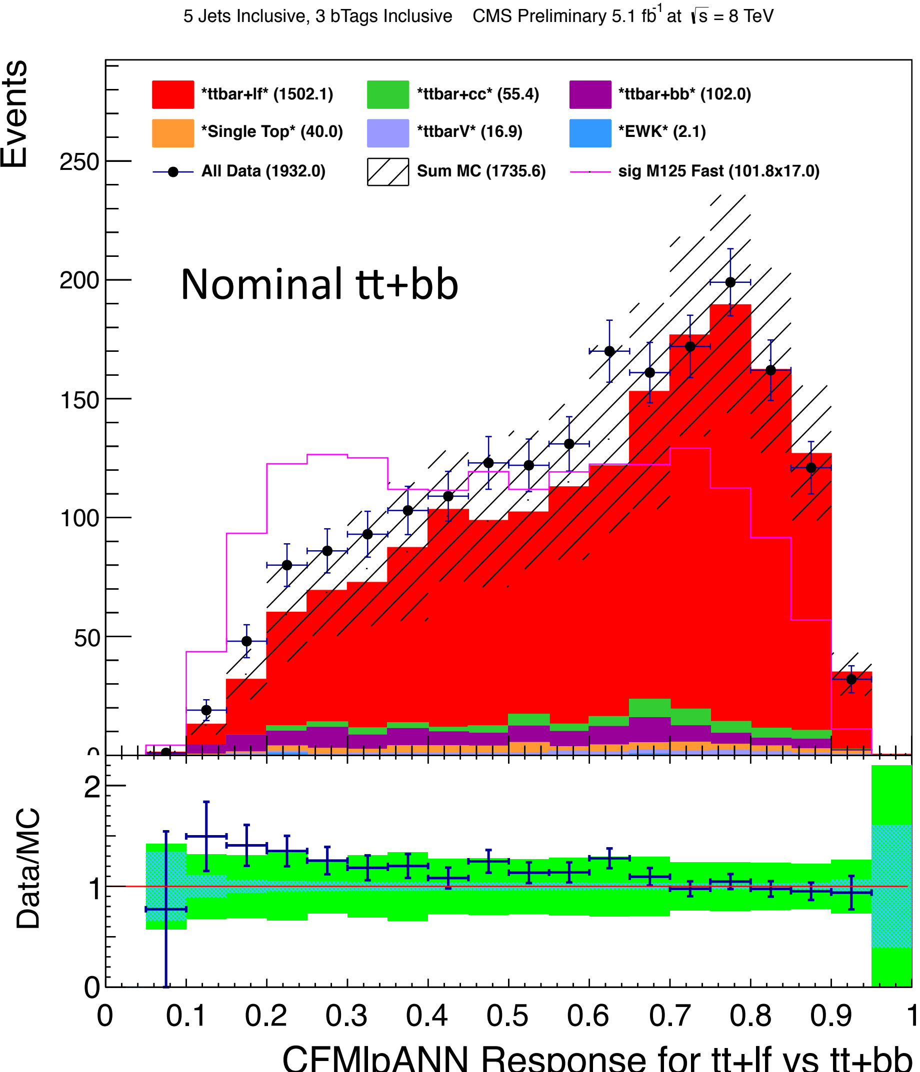
- Repeat many times to get distribution of +/- 1,2σ uncertainty bands

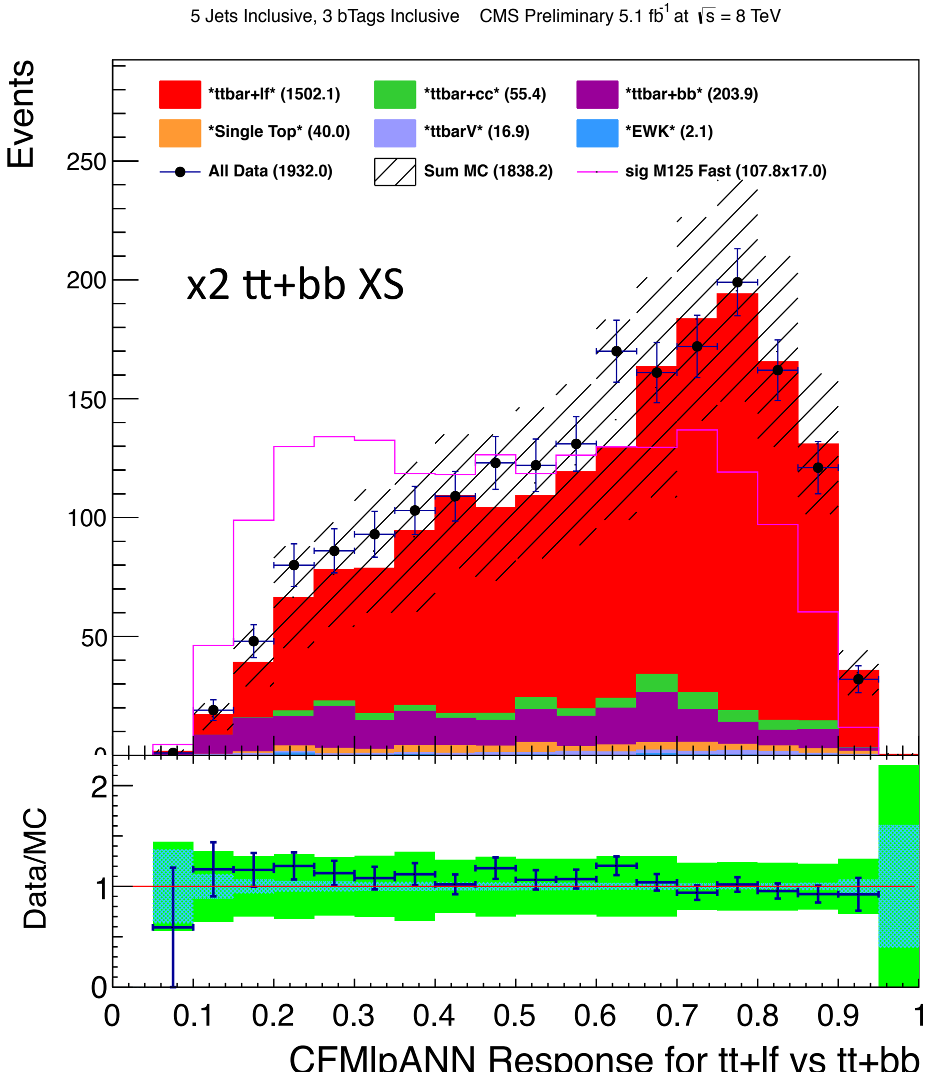
tt+bb vs tt+lf discriminant– 8TeV L+J, mu+ele

- In order to gauge the sensitivity of our analysis to tt+bb, we have trained an ANN to discriminate tt+bb from tt+lf
  - Using CFMlpANN MVA – same type of MVA algorithm used in signal extraction
  - Train tt+bb as signal against tt+lf background
  - Divide samples into 2 sub-categories, so that TMVA has sufficient events in each sample to train the ANN
    - ==5 Jets, >=3 bTags
    - >=6 Jets, >=3 bTags
  - For comparison, have also created plots with the X-Section for tt+bb doubled
  - From these we can look for any discrepancies between data and MC agreement, which could possibly be attributed to erroneous assumptions on the XS or shape of the tt+bb process

tt+bb vs tt+lf discriminant– 8TeV L+J, mu+ele, >=5J >=3T

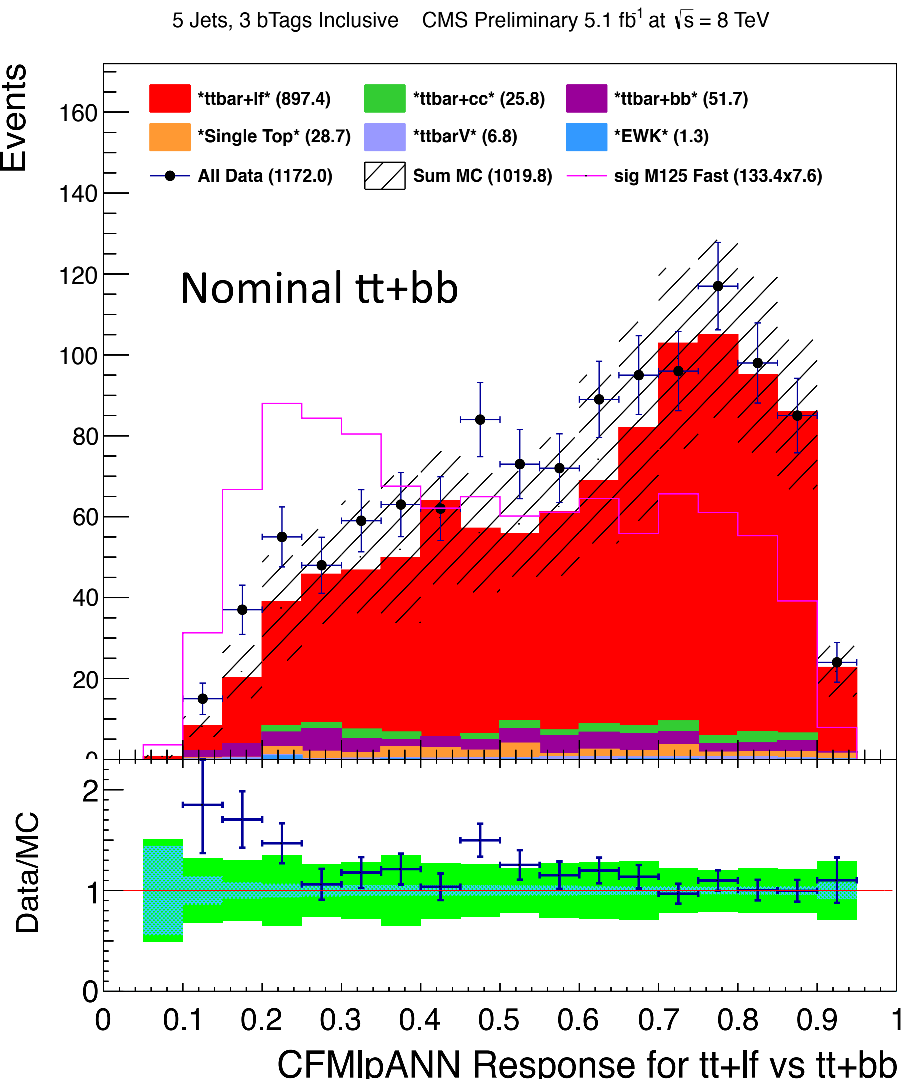
- For the inclusive category, Nominal tt+bb XS accounts for data within error bands
- Doubling tt+bb XS improves agreement only slightly

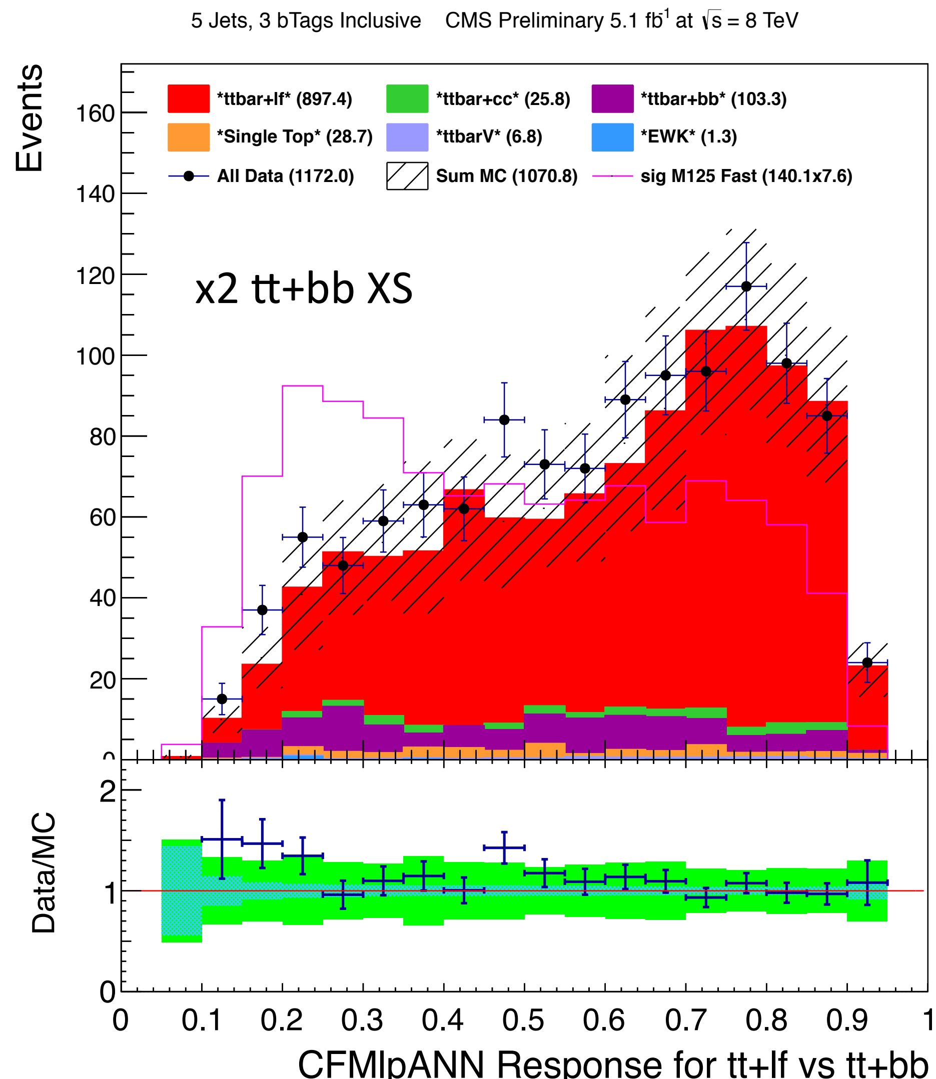




tt+bb vs tt+lf discriminant– 8TeV L+J, mu+ele, ==5J, >=3T

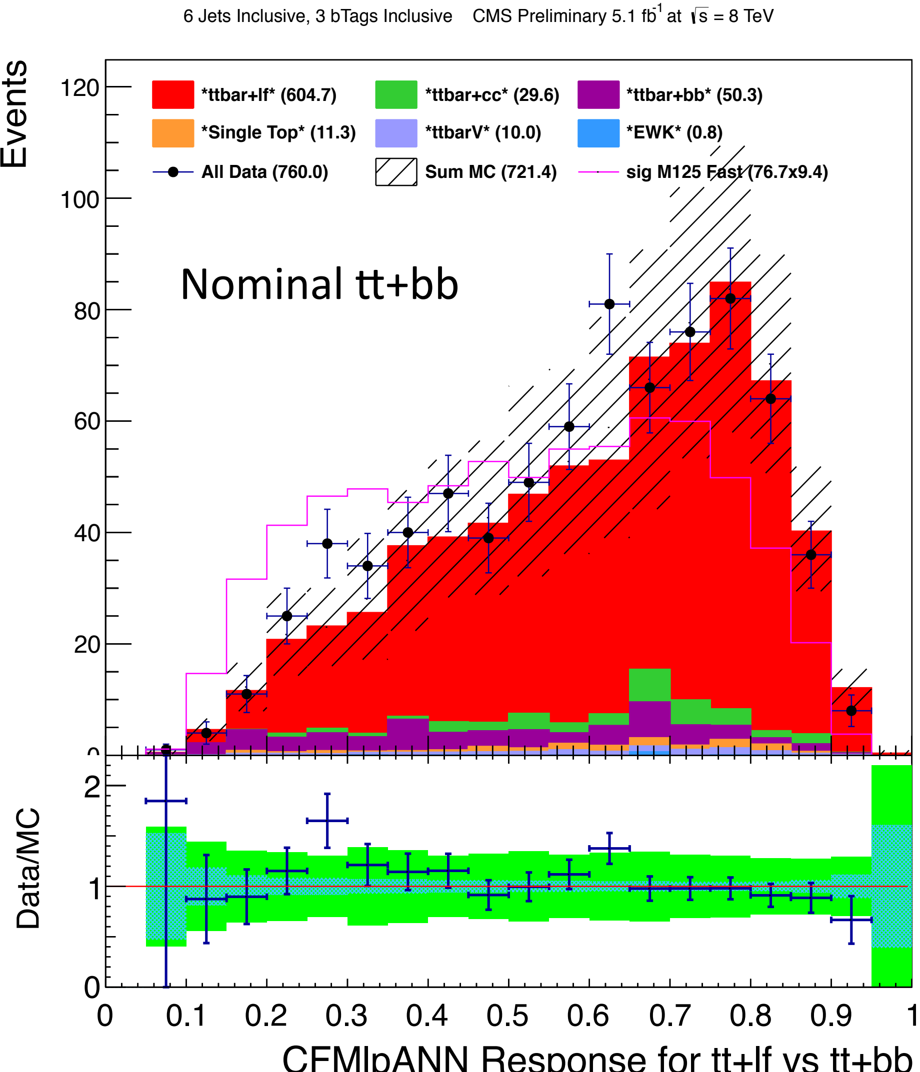
- Nominal tt+bb XS accounts for data within error bands
- In this sub-category, doubling tt+bb XS improves only a few bins

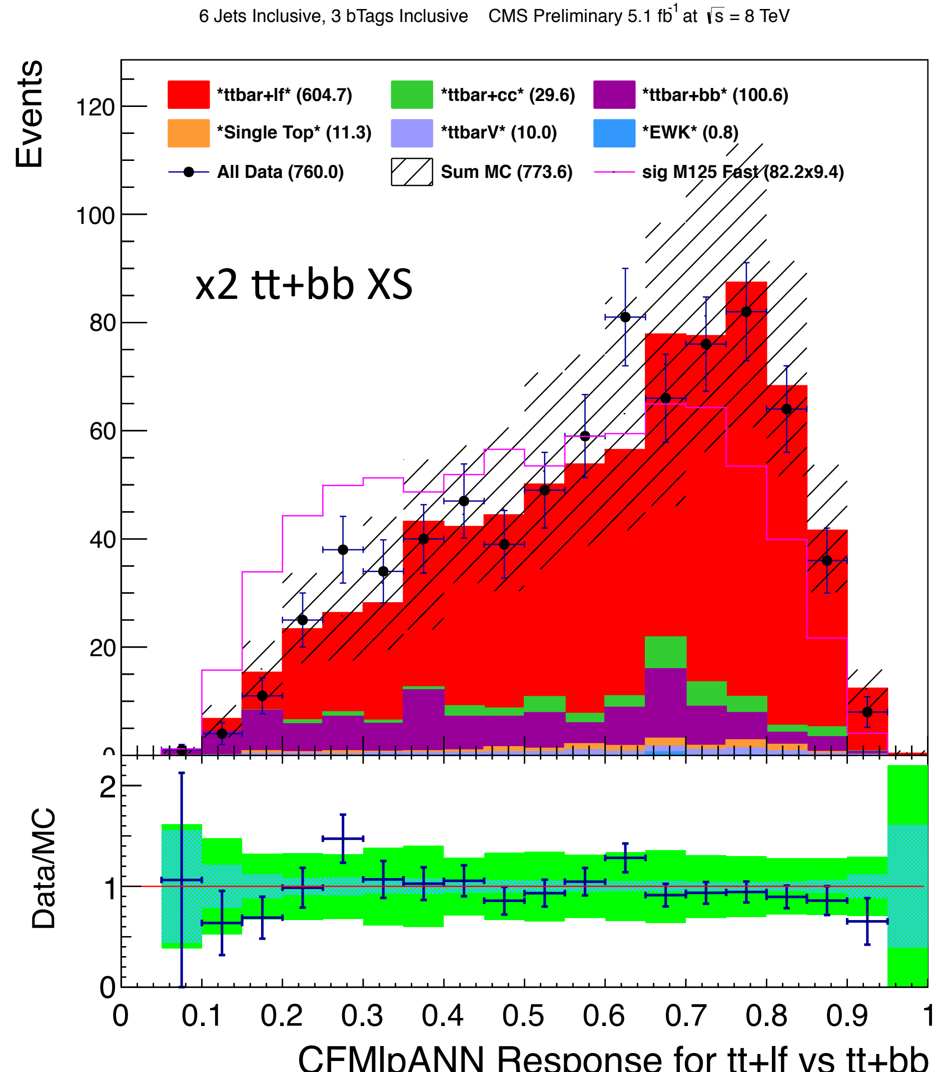




tt+bb vs tt+lf discriminant– 8TeV L+J, mu+ele, >=6J, >=3T

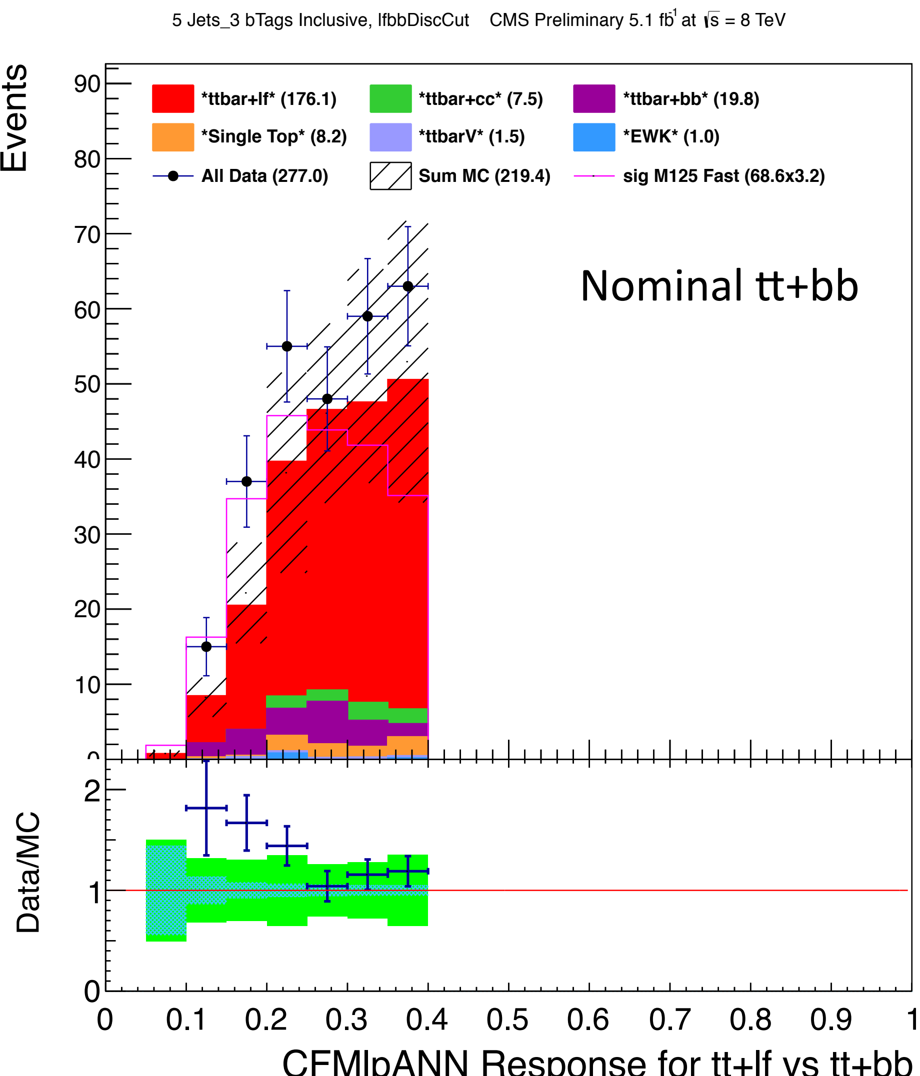
- Here, in the most sensitive category for tt+bb, nominal tt+bb XS shows more consistent levels of agreement than the doubled XS

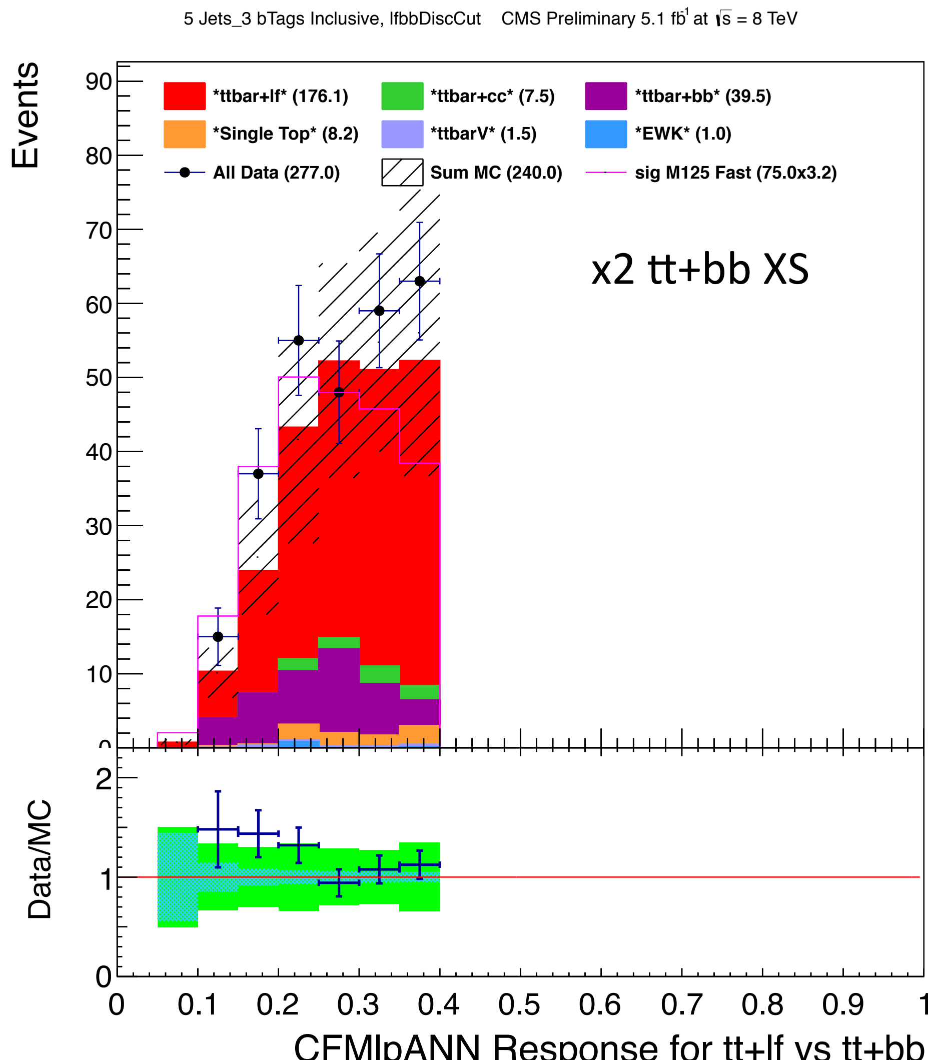




tt+bb vs tt+lf discriminant– 8TeV L+J, mu+ele, ==5J, >=3T

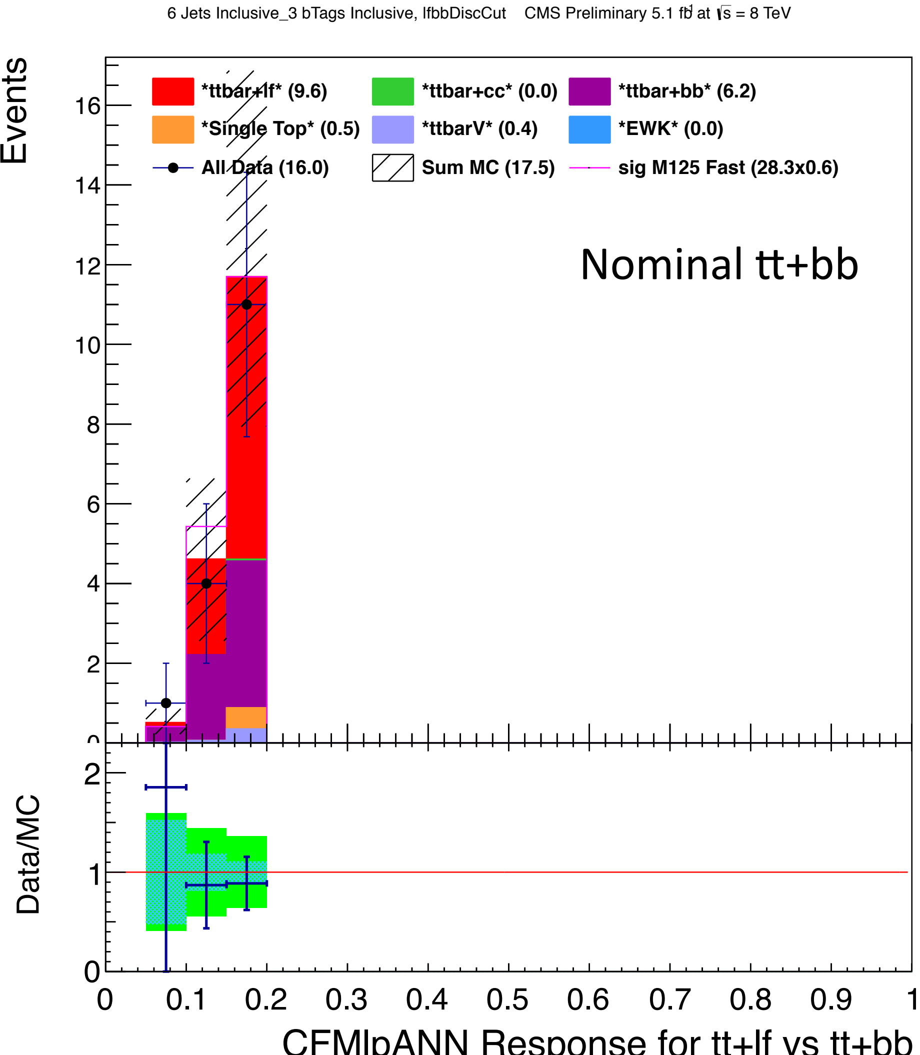
- Looking at the region with best tt+bb/Sum_allBg, doubling XS provides some improvement

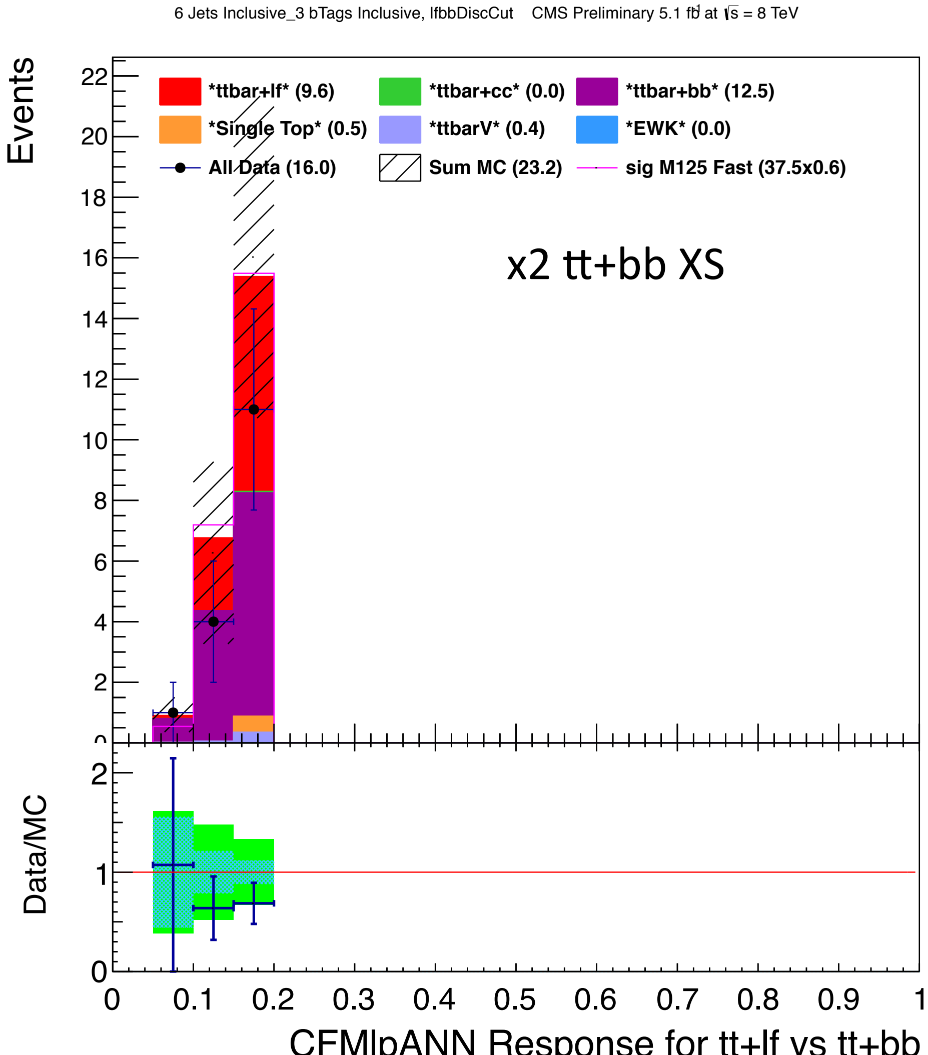




tt+bb vs tt+lf discriminant– 8TeV L+J, mu+ele, >=6J, >=3T

- In this category, which has the best tt+bb/allBg ratio; in the most sensitive region of this discriminant: nominal tt+bb XS provides better agreement





tt+bb vs tt+lf discriminant– 8TeV L+J, mu+ele

- Cutflow tables for nominal, and x2 tt+bb XS cases

Nominal tt+bb

Table 1: Cut Flow Table for ttH→ttbb→lep+Jets

Cut	$t\bar{t}$+lf	$t\bar{t}+c\bar{c}$	$t\bar{t}+b\bar{b}$	Single t	$t\bar{t}+V$	EWK	sig M125 Fast	$\epsilon_{sig}(\%)$	Sum B	Observed
>=5 Jets >=3 bTags	1502.09	55.42	101.95	40.03	16.85	2.10	17.05	2.5992	1718.45±124.38	1932.00
==5 Jets >=3 bTags	897.40	25.80	51.67	28.70	6.82	1.28	7.64	1.1692	1011.66±68.67	1172.00
>=6 Jets >=3 bTags	604.69	29.62	50.28	11.34	10.04	0.82	9.41	1.4300	706.79±63.47	760.00
==5 Jets, >=3 bTags, lfbbDisc<0.4	176.14	7.48	19.77	8.25	1.52	1.00	3.20	0.4927	214.17±30.48	277.00
>=6 Jets >=3 bTags, lfbbDisc<0.2	9.60	0.05	6.23	0.53	0.43	0.00	0.62	0.0975	16.83±4.92	16.00

x2 tt+bb XS

Table 1: Cut Flow Table for ttH→ttbb→lep+Jets

Cut	$t\bar{t}$+lf	$t\bar{t}+c\bar{c}$	$t\bar{t}+b\bar{b}$	Single t	$t\bar{t}+V$	EWK	sig M125 Fast	$\epsilon_{sig}(\%)$	Sum B	Observed
>=5 Jets >=3 bTags	1502.09	55.42	203.90	40.03	16.85	2.10	17.05	2.5992	1820.40±131.25	1932.00
==5 Jets >=3 bTags	897.40	25.80	103.34	28.70	6.82	1.28	7.64	1.1692	1063.33±71.72	1172.00
>=6 Jets >=3 bTags	604.69	29.62	100.57	11.34	10.04	0.82	9.41	1.4300	757.08±68.02	760.00
==5 Jets, >=3 bTags, lfbbDisc<0.4	176.14	7.48	39.55	8.25	1.52	1.00	3.20	0.4927	233.94±32.75	277.00
>=6 Jets >=3 bTags, lfbbDisc<0.2	9.60	0.05	12.45	0.53	0.43	0.00	0.62	0.0975	23.06±6.13	16.00

tt+bb vs tt+lf discriminant– 8TeV 5fb^{-1} L+J, mu+ele

- ==5 Jets, >=3 bTags
  - This category shows some improvement overall when doubling the tt+bb XS

- >=6 Jets, >=3 bTags
  - This category only shows slight improvement overall when doubling tt+bb XS
  - In the most sensitive region of the discriminant for tt+bb, the nominal XS showed shows the better agreement

- By designing an MVA to look specifically for tt+bb, we can gauge how accurately we believe we are modeling this process

- It is not clear that increasing the tt+bb X-section is the right thing to do for all categories. Hence, we propose to increase our systematic uncertainty on the normalization of this process to a larger number, 100%

- Also, even though we look at these plots and can make statements about the yields/ shapes for these different scenarios, there is no statistically significant discrepancy in any of the nominal shapes. The size of the tt+bb "signal" is approximately the same size as the uncertainty of the background (which is driven by tt+lf)

Pre/Post Fit Comparison – 8 TeV 5 fb^{-1}

- Pre-fit versus post-fit normalizations for all backgrounds and signal
- In general, background uncertainties reduced, signal uncertainty increased
- K-Factor for tt+bb consistent with expectations from NLO comparisons

Category	Sample	Pre-Fit	Post-fit (s+b)	$\frac{Post-fit}{Pre-fit}$
Dilepton ≥ 3 jets, ≥ 3 b-tags	$t\bar{t}H(125)$	3.3 ± 1.7	11.2 ± 9.7	3.4
	$t\bar{t}$+lf	135.5 ± 34.4	150.2 ± 10.5	1.1
	$t\bar{t}+b\bar{b}$	25.1 ± 13.9	32.7 ± 15.3	1.3
	$t\bar{t}+c\bar{c}$	14.1 ± 4.1	14.8 ± 2.7	1.0
	$t\bar{t}W$	1.1 ± 0.3	1.2 ± 0.2	1.1
	$t\bar{t}Z$	2.7 ± 0.6	2.9 ± 0.3	1.1
	Single t	6.2 ± 2.4	5.3 ± 2.2	0.8
	W+jets	< 0.1	< 0.1	-
	Z+jets	29.4 ± 13.3	26.0 ± 10.0	0.9
	Diboson	0.7 ± 0.2	0.7 ± 0.2	1.0
	Total Bkg	214.7 ± 48.0	233.8 ± 18.8	1.1
	Data	251		
Lepton+jets 5 jets, ≥ 4 b-tags	$t\bar{t}H(125)$	1.5 ± 0.8	5.6 ± 4.8	3.6
	$t\bar{t}$+lf	27.6 ± 8.6	32.1 ± 5.0	1.2
	$t\bar{t}+b\bar{b}$	8.4 ± 5.4	11.6 ± 5.6	1.4
	$t\bar{t}+c\bar{c}$	0.8 ± 0.9	0.8 ± 0.9	1.0
	$t\bar{t}W$	0.3 ± 0.1	0.3 ± 0.1	1.0
	$t\bar{t}Z$	0.4 ± 0.2	0.5 ± 0.1	1.2
	Single t	3.1 ± 2.2	2.8 ± 2.0	0.9
	W+jets	< 0.1	< 0.1	-
	Z+jets	< 0.1	< 0.1	-
	Diboson	< 0.1	< 0.1	-
	Total Bkg	40.5 ± 11.8	48.1 ± 7.0	1.2
	Data	56		
Lepton+jets ≥ 6 jets, ≥ 4 b-tags	$t\bar{t}H(125)$	2.5 ± 1.3	8.7 ± 7.5	3.4
	$t\bar{t}$+lf	41.3 ± 14.9	45.3 ± 5.8	1.1
	$t\bar{t}+b\bar{b}$	15.4 ± 9.4	18.9 ± 9.2	1.2
	$t\bar{t}+c\bar{c}$	3.7 ± 1.8	3.9 ± 1.5	1.1
	$t\bar{t}W$	0.3 ± 0.1	0.3 ± 0.1	1.1
	$t\bar{t}Z$	1.2 ± 0.3	1.3 ± 0.2	1.1
	Single t	1.0 ± 0.6	1.1 ± 0.5	1.1
	W+jets	< 0.1	< 0.1	-
	Z+jets	< 0.1	< 0.1	-
	Diboson	< 0.1	< 0.1	-
	Total Bkg	62.9 ± 21.0	70.9 ± 10.2	1.1
	Data	74		

	$\geq$6 jets 2 tags	4 jets 3 tags	5 jets 3 tags	$\geq$6 jets 3 tags	4 jets $\geq$4 tags	5 jets $\geq$4 tags	$\geq$6 jets $\geq$4 tags
$t\bar{t}$+lf	0.9	1.1	1.1	1.0	1.3	1.2	1.1
$t\bar{t}+b\bar{b}$	1.1	1.3	1.2	1.3	1.4	1.4	1.2
$t\bar{t}+c\bar{c}$	1.0	1.1	1.1	1.0	1.0	1.0	1.1

Input Variables, all Categories

5 jets, ≥ 4 tags	≥ 6 jets, 3 tags	≥ 6 jets, ≥ 4 tags
ave ΔR(tag,tag)	tagged dijet mass closest to 125	H_3
max $\Delta\eta$ (tag, ave tag η)	$(\Sigma$ jet $p_T)/(\Sigma$ jet E$)$	ave ΔR(tag,tag)
$(\Sigma$ jet $p_T)/(\Sigma$ jet E$)$	$\sqrt{\Delta\eta(t^{lep}, bb) \times \Delta\eta(t^{had}, bb)}$	closest tagged dijet mass
tagged dijet mass closest to 125	H_1	sphericity
H_1	H_3	max $\Delta\eta$ (tag, ave jet η)
H_3	M3	max $\Delta\eta$ (tag, ave tag η)
$\sum p_T$(jets,lepton,MET)	max $\Delta\eta$ (tag, ave tag η)	mass(lepton,jet,MET)
fourth-highest CSV (tags)	max $\Delta\eta$ (tag, ave jet η)	$(\Sigma$ jet $p_T)/(\Sigma$ jet E$)$
aplanarity	max $\Delta\eta$ (jet, ave jet η)	abs $\Delta\eta$ (leptonic top, bb)
MET	abs $\Delta\eta$ (hadronic top, bb)	abs $\Delta\eta$ (hadronic top, bb)
	abs $\Delta\eta$ (leptonic top, bb)	$\sqrt{\Delta\eta(t^{lep}, bb) \times \Delta\eta(t^{had}, bb)}$
	sphericity	ave CSV (tags)
	aplanarity	best ΔR(b,b)
	min ΔR(tag,tag)	best Higgs boson mass
	jet 3 p_T	median inv. mass (tag pairs)

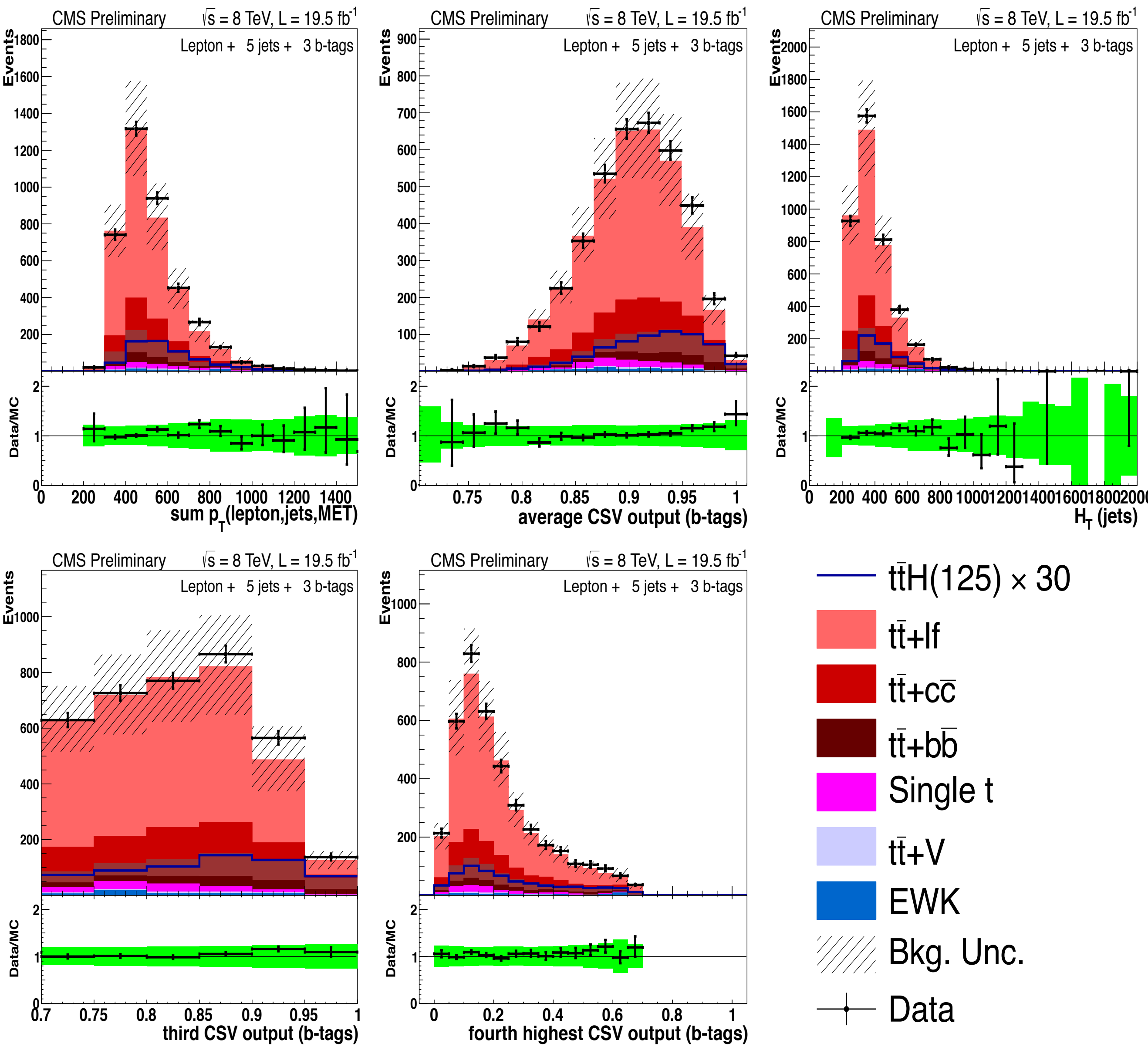
4 jets, 3 b-tags	4 jets, 4 b-tags
jet 1 p_T	jet 1 p_T
jet 2 p_T	jet 2 p_T
jet 3 p_T	jet 4 p_T
jet 4 p_T	HT
M3	$\sum p_T$(jets,lepton,MET)
$\sum p_T$(jets,lepton,MET)	M3
HT	ave CSV (tags)
lowest CSV (tags)	second-highest CSV (tags)
MHT	third-highest CSV (tags)
MET	lowest CSV (tags)

5 jets, 3 b-tags	5 jets, ≥ 4 b-tags
jet 1 p_T	max $\Delta\eta$ (tag, ave jet η)
jet 2 p_T	$\sum p_T$(jets,lepton,MET)
jet 3 p_T	$(\Sigma$ jet $p_T)/(\Sigma$ jet E$)$
jet 4 p_T	ave ΔR(tag,tag)
$\sum p_T$(jets,lepton,MET)	ave CSV (tags)
$(\Sigma$ jet $p_T)/(\Sigma$ jet E$)$	dev from ave CSV (tags)
HT	second-highest CSV (tags)
ave CSV (tags)	third-highest CSV (tags)
third-highest CSV (tags)	lowest CSV (tags)
fourth-highest CSV (jets)	ttbb/ttH BDT

≥ 6 jets, 2 tags	≥ 6 jets, 3 tags	≥ 6 jets, ≥ 4 tags
$\sum p_T$(jets,lepton,MET)	H_0	$(\Sigma$ jet $p_T)/(\Sigma$ jet E$)$
HT	sphericity	ave ΔR(tag,tag)
mass(lepton,closest tag)	$(\Sigma$ jet $p_T)/(\Sigma$ jet E$)$	product($\Delta\eta$(leptonic top, bb), $\Delta\eta$(hadronic top, bb))
max $\Delta\eta$ (jet, ave jet η)	max $\Delta\eta$ (jet, ave jet η)	closest tag mass
min ΔR(lepton,jet)	$\sum p_T$(jets,lepton,MET)	max $\Delta\eta$ (tag, ave tag η)
H_2	ave CSV (tags)	ave CSV (tags)
sphericity	second-highest CSV (tags)	third-highest CSV (tags)
$(\Sigma$ jet $p_T)/(\Sigma$ jet E$)$	third-highest CSV (tags)	fourth-highest CSV (tags)
third-highest CSV (jets)	fourth-highest CSV (jets)	best Higgs boson mass
fourth-highest CSV (jets)	ttbb/ttH BDT	ttbb/ttH BDT

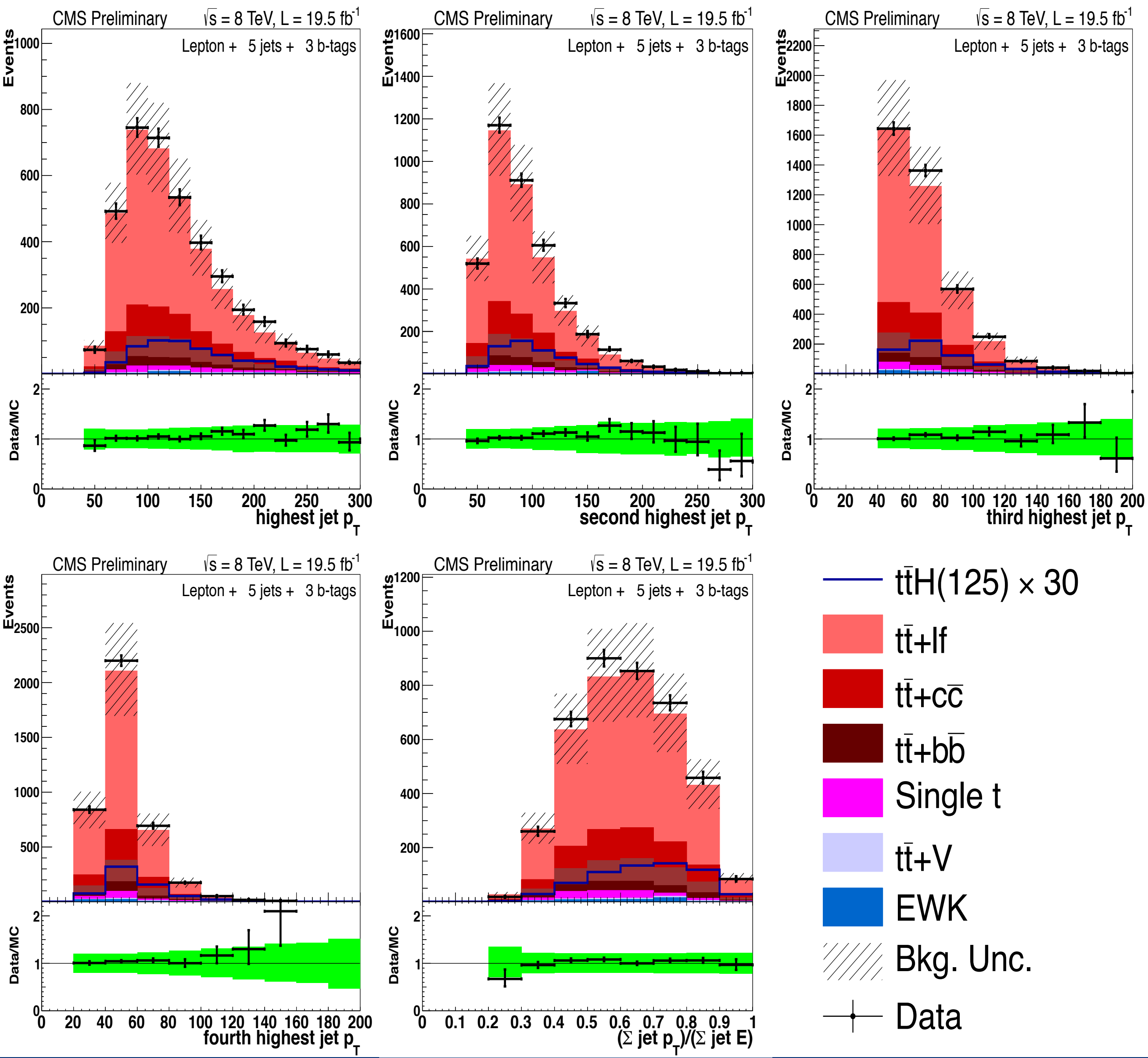
- Above: Variables used for the ttbb/ttH discriminant in the most signal sensitive categories
- Right: Variables used for tt+Jets/ttH discriminant used in signal extraction and limit setting

Boosted Decision Tree - Input



- ❖ **Train BDT,** 10 variables depending on category (eg. ==5j, ==3t, signal scaled x 30):
- ✧ Sum of pT for lepton, jets, MET
- ✧ avg b-tag discriminant value (CSV) for tagged jets
- ✧ Scalar Sum of pT, Jets (Ht)
- ✧ Third highest b-tag CSV
- ✧ Fourth highest CSV
- ✧ Highest pT jet
- ✧ Second highest pT jet
- ✧ Third Highest pT jet
- ✧ Fourth highest pT jet
- ✧ Sum of Jet pT / Sum of jet Energy

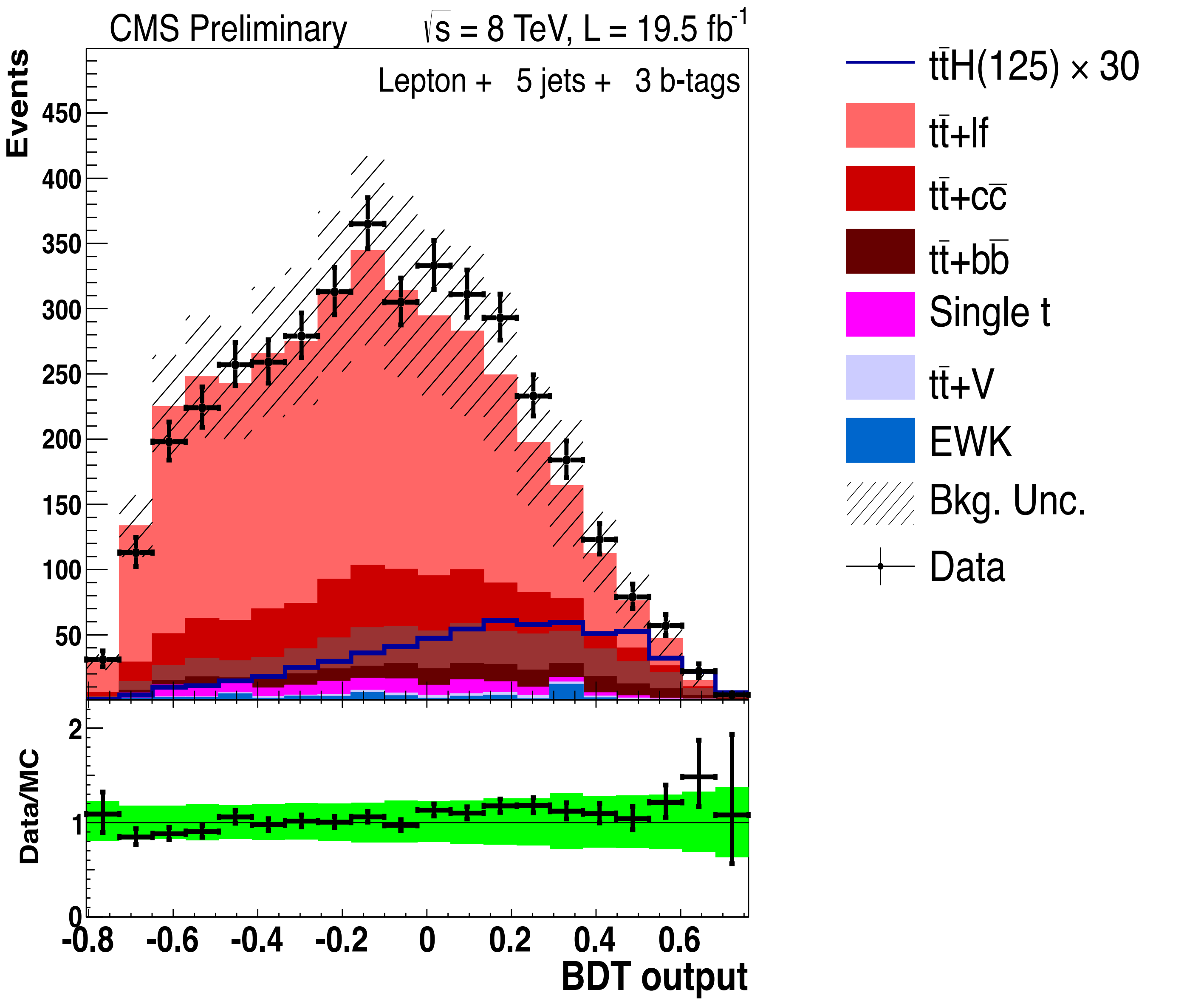
Boosted Decision Tree - Input



- **Train BDT,** 10 variables depending on category (eg. ==5j, ==3t, signal scaled x 30):
  - Sum of pT for lepton, jets, MET
  - avg b-tag discriminant value (CSV) for tagged jets
  - Scalar Sum of pT, Jets (Ht)
  - Third highest b-tag CSV
  - Fourth highest CSV
  - Highest pT jet
  - Second highest pT jet
  - Third Highest pT jet
  - Fourth highest pT jet
  - Sum of Jet pT / Sum of jet Energy

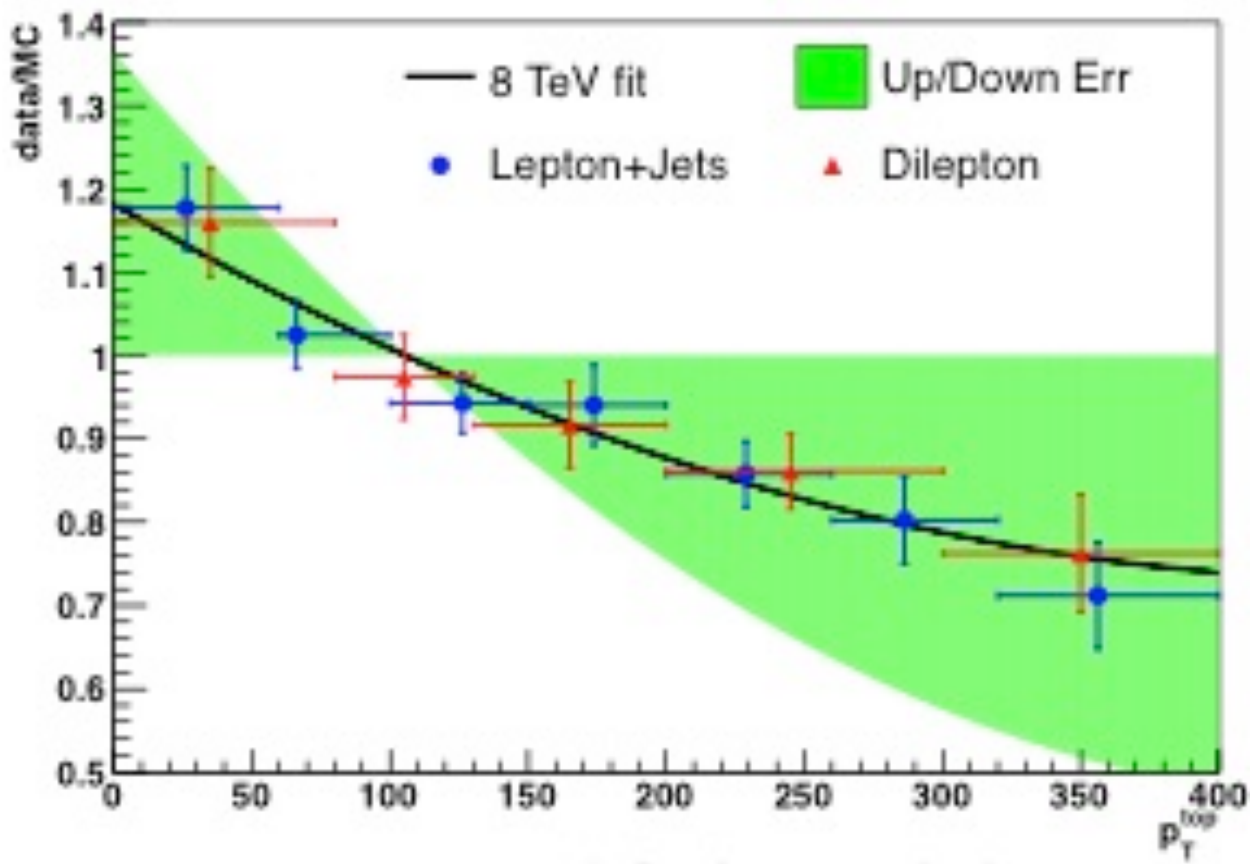
Boosted Decision Tree – Output

- Events which are more similar to largest background, tt+jets, have values closer to -1, events which are most similar to ttH, have values closer to 1
- Example to the right is from the Lepton +Jets analysis: (==5Jets,==3Tags)



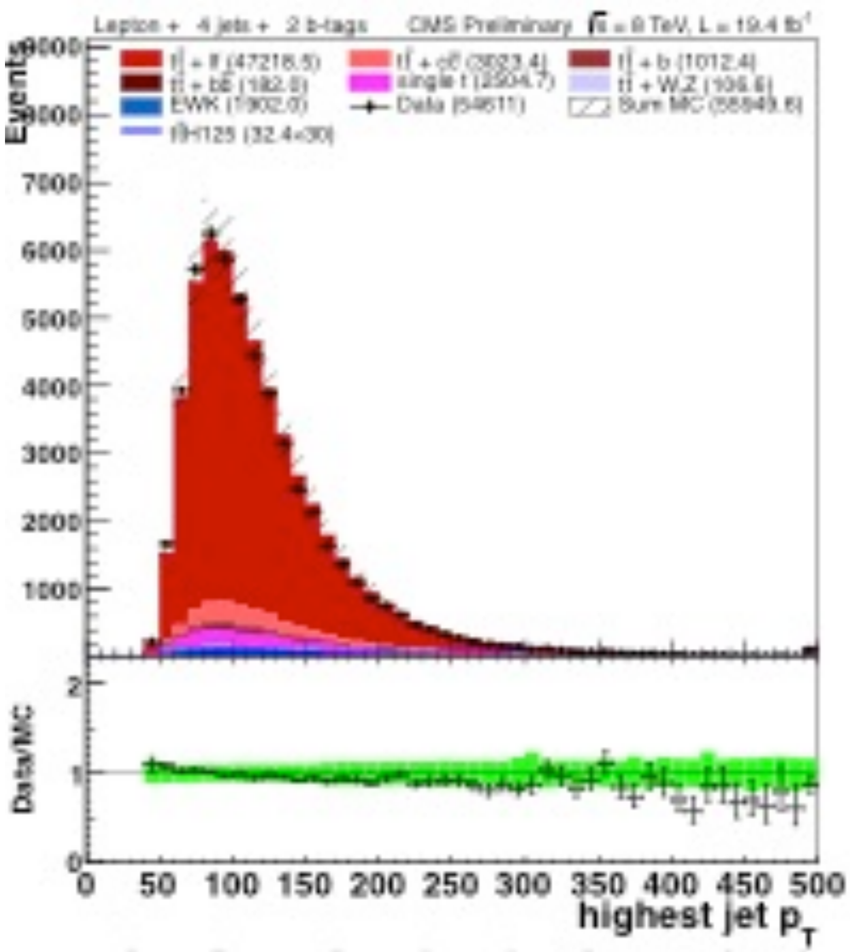
Top pT Reweighting – 19.5 fb^{-1} 8 TeV

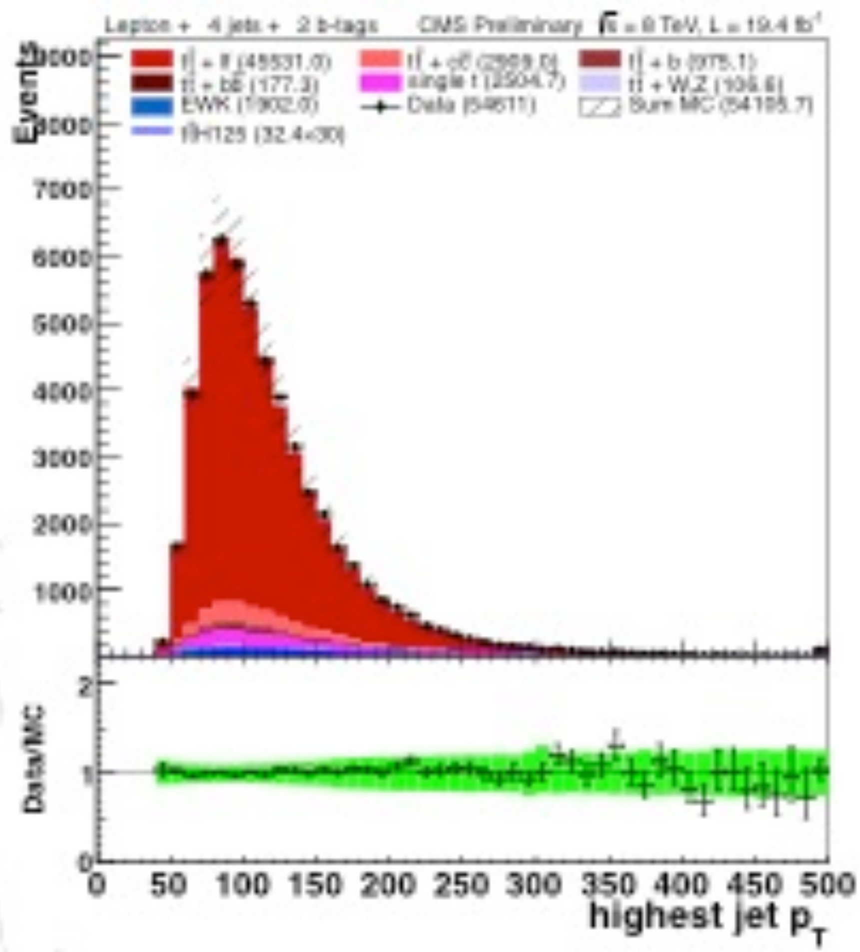
- New Monte Carlo ttJets samples also exhibited the same discrepancy with respect to data that we previously used Ht reweighting to correct
- Top pT cross section measured in data and compared to madgraph monte carlo
- Top pT at genParton level is softer than data at low pT, so correction factor and associated systematic uncertainty added to the analysis



Top pT Reweighting – 19.5 fb^{-1} 8 TeV

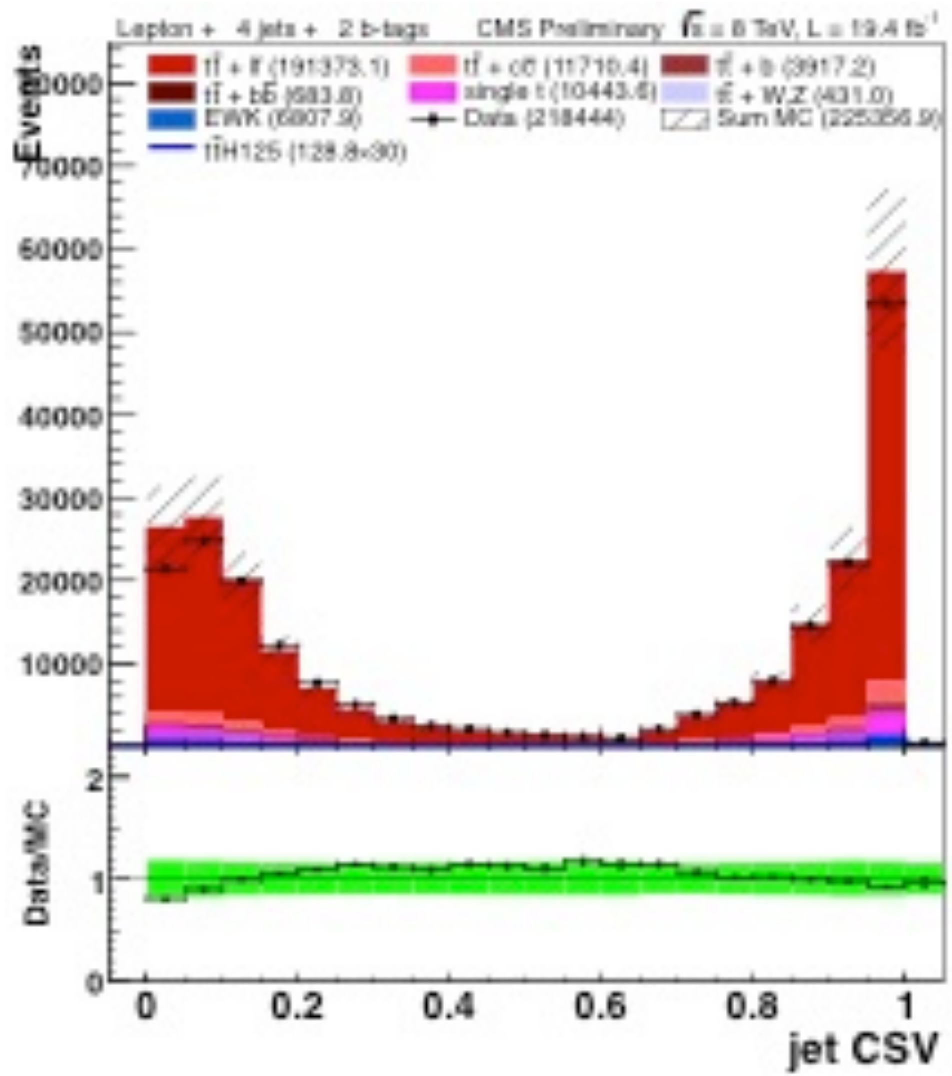
- >=4Jets >=2Tags Data to Monte Comparison before (left) and after (right) top pT reweighting
- Sloped trend clearly reduced, and based off of data-driven measurements
- Replaces Ht re-weighting from previous iteration of analysis





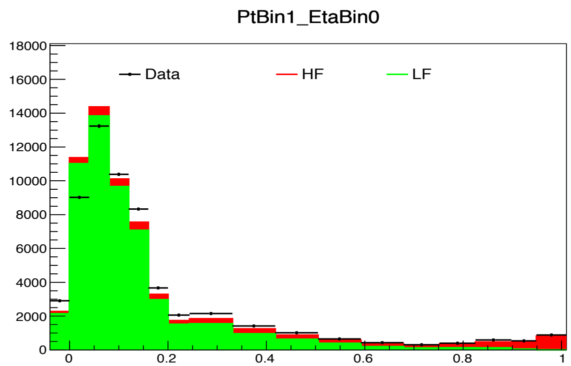
bTag Scale Factors – 19.5 fb^{-1} 8 TeV

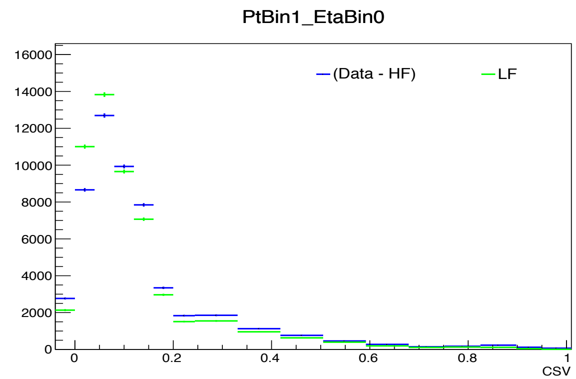
- When updating to the full 8 TeV dataset, set discovered a new problem in the modeling of the bTagging discriminant
- (Right) Data to MC comparisons of jet csv in the Lepton+Jets ==4Jet ==2Tag channel shows a distinctive “bow” in the csv shape
- Since we rely heavily on CSV shape in our signal extraction BDT, it is necessary to model this variable correctly
- Data-Driven correction factors
  - Tag and Probe in tt+jets enriched, 2jet, diLepton sample for efficiency scale factor (HF scale factor)
  - Tag and Probe in Z+jets enriched, 2 jet, diLepton sample for mistag scale factor (LF scale factor)
  - Bin in pT and eta for optimal SF extraction

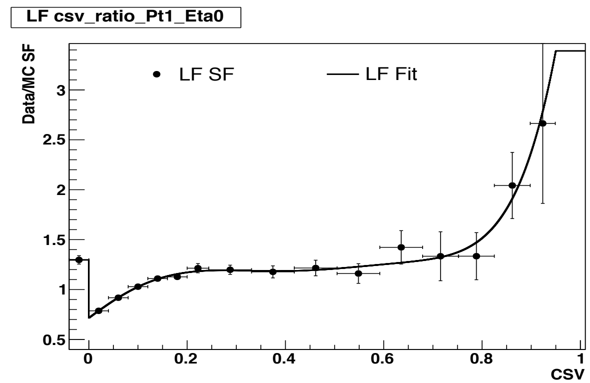


bTag Scale Factors – 19.5 fb^{-1} 8 TeV

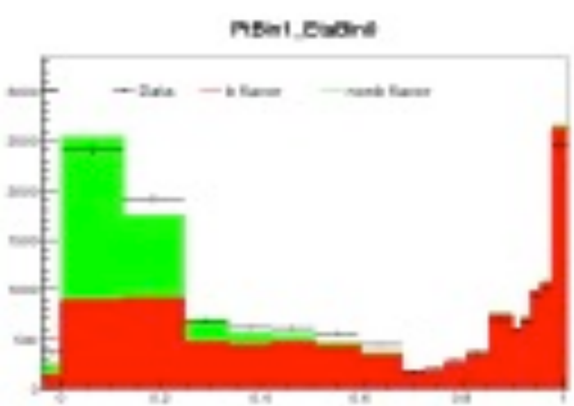
- pT bins of [30,40], [40, 60], [60+]
- |eta| bins of [0, 0.8], [0.8, 1.6], [1.6, 2.4]
- (Above) Light Flavor (Below) Heavy Flavor

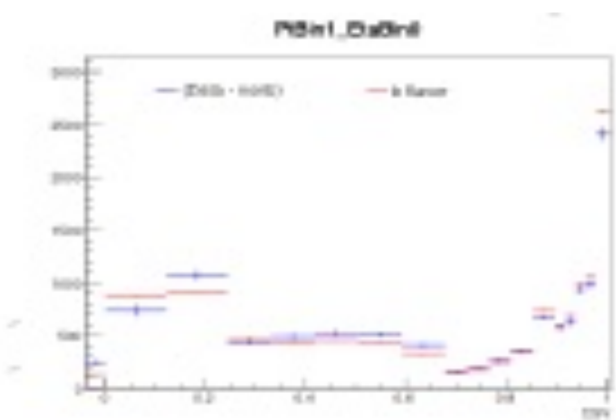


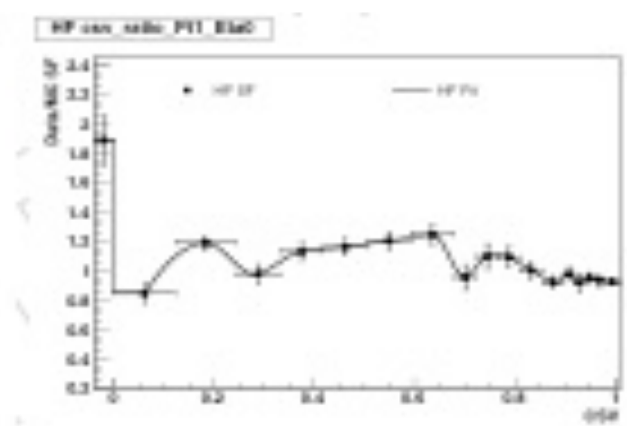




- (Left) Data vs MC comparisons of LF/HF regions
- (Center) Isolating LF/HF in data – subtract HF mc from LF data, and vice versa
- (Right) Scale Factors after applying tag & probe technique to selected sample

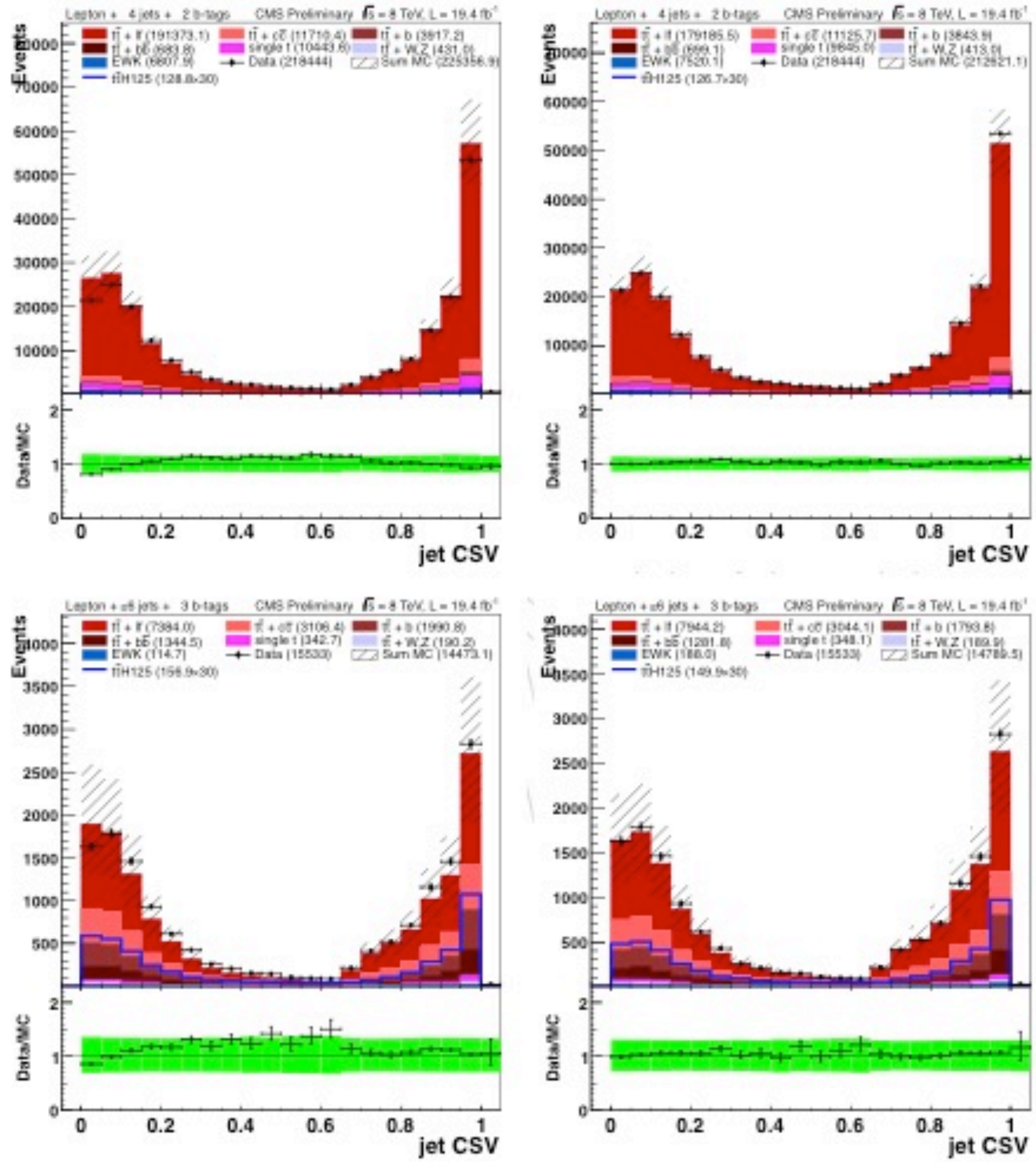






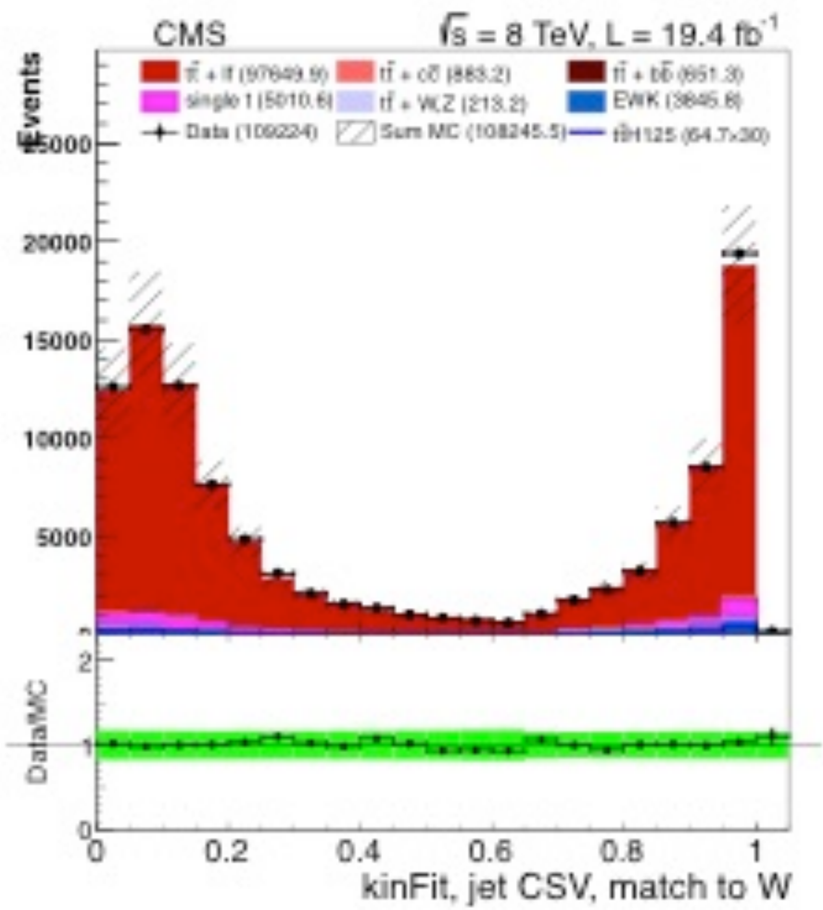
bTag Scale Factors – 19.5 fb^{-1} 8 TeV

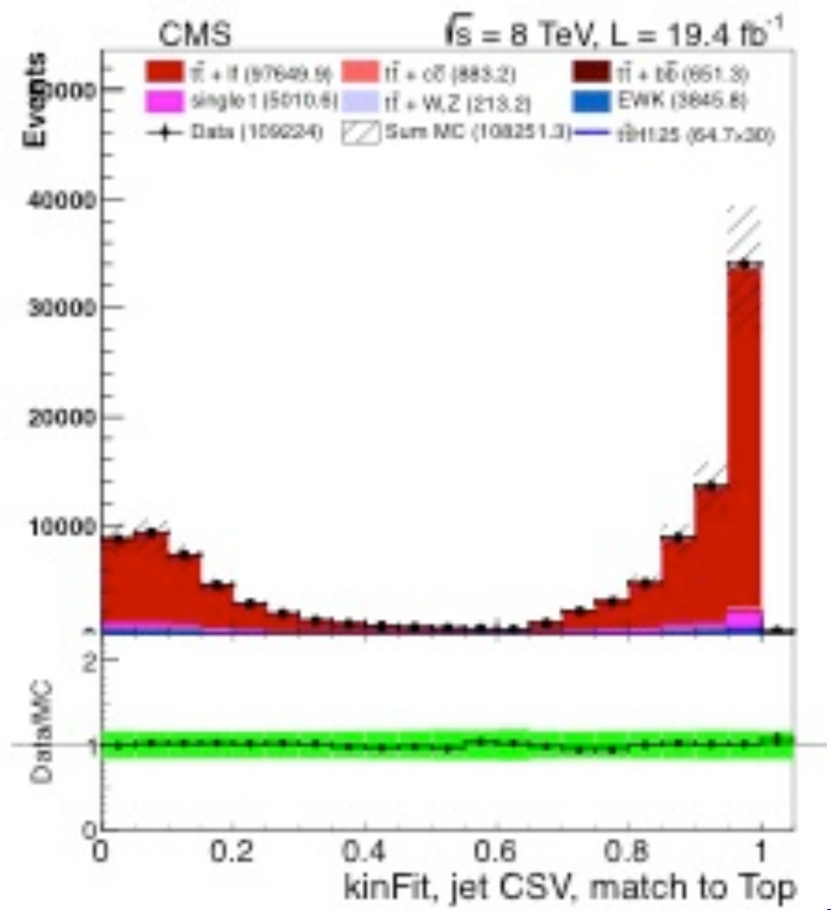
- (Above Left) ==4Jets ==2Tags jet csv distribution without scale factors – note distinct bow in center
- (Above Right) ==4Jets ==2Tags jet csv distribution with scale factors – flat data/mc ratio shows csv factors derived in diLepton work in lepton +jets
- (Below Left) >=6Jets ==2Tags jet csv distribution without scale factors – note distinct bow in center
- (Below Right) >=6Jets ==2Tags jet csv distribution with scale factors – flat data/mc ratio shows csv factors derived in diLepton work in lepton +jets in high jet multiplicity environment



bTag Scale Factors – 19.5 fb^{-1} 8 TeV

- (Below Right) ==4Jets ==2Tags jet csv distribution with scale factors, for selection of jets matched to W boson decays, using a X^2 minimization algorithm- validating LF Scale Factors
- (Below Left) ==4Jets ==2Tags jet csv distribution with scale factors, for selection of jets matched to b-quark decays from top quarks, using a X^2 minimization algorithm – validating HF scale factors



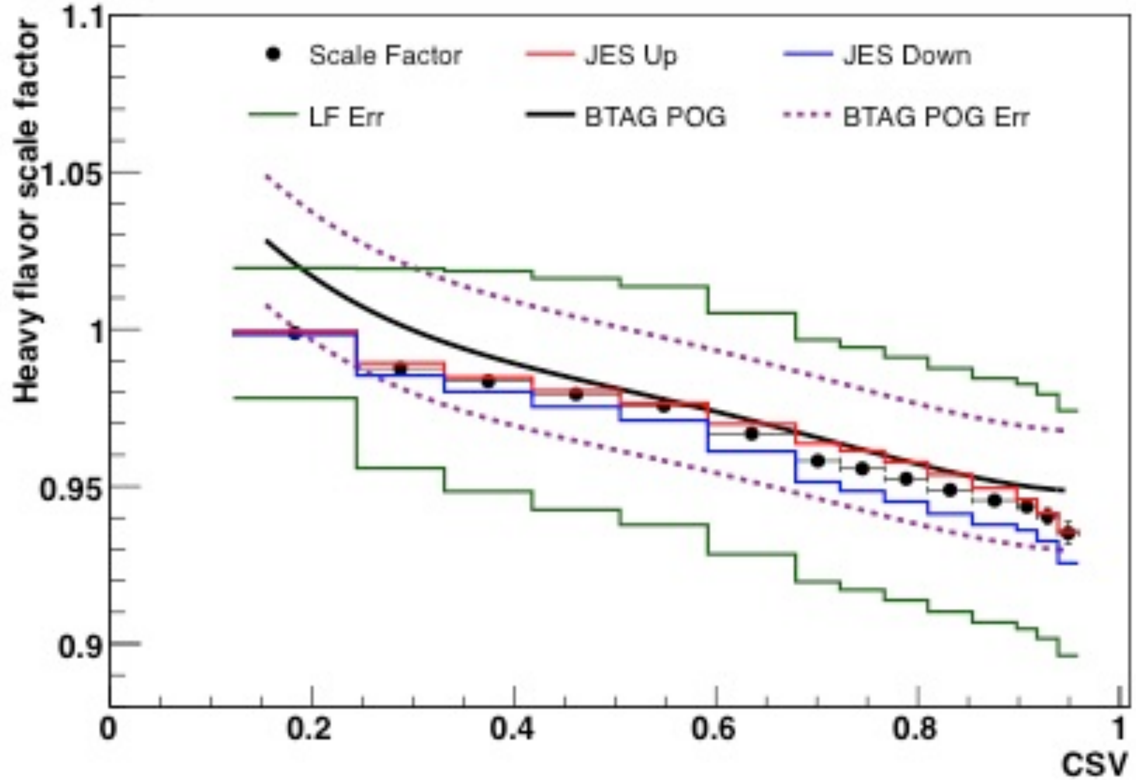


bTag Scale Factors – 19.5 fb^{-1} 8 TeV

- (Right) Comparison of Scale Factors and Uncertainties to bTag POG Scale Factors and uncertainties

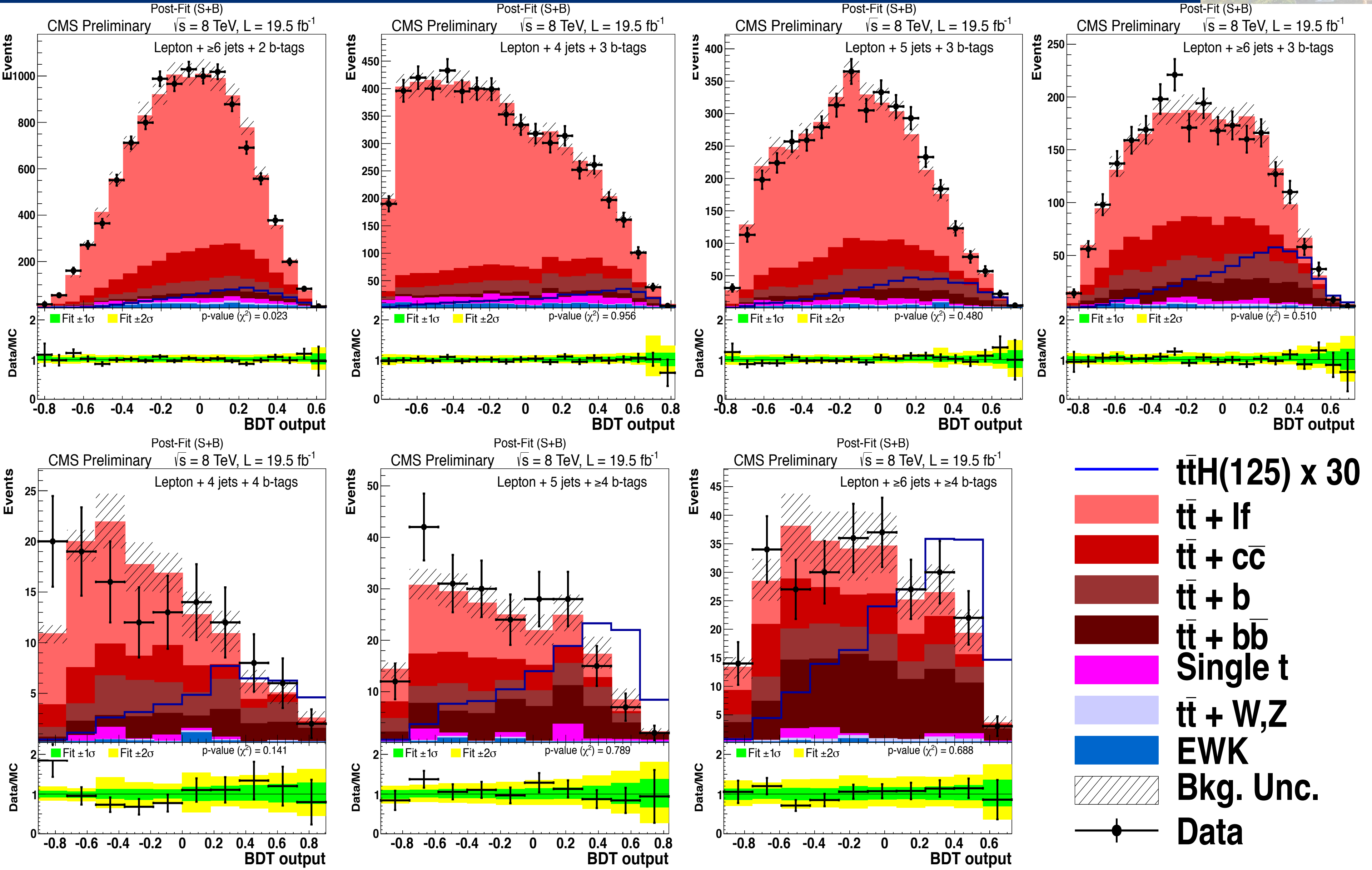
- Measuring Scale factors in differential form comes at a cost of increased uncertainty bands

- Uncertainties based on:
  - JES Scale Up/Down
  - Purity Uncertainty from Subtracting MC from data
  - Statistical uncertainty from using MC



- ***Method described in AN-13-130 and is now an official bTag POG set of scale factors***

Boosted Decision Tree – Output



Systematic Uncertainties

- Uncertainties table for sum of ttbar+X background yields
  - >=6Jet, >=4Tag category
- QCD Scale– ie renormalization and factorization uncertainty – MC only uncertainty, due to matrix element nature of calculations, it is the largest uncertainty in the analysis
- bTagging efficiencies –differences in tagging efficiencies between data and MC
- Other sources – uncertainties on scale of jet energy measurement, momentum distributions of incoming colliding partons, and production rates of SM processes

Uncertainties of the sum of $t\bar{t}$+lf, $t\bar{t}$+b, $t\bar{t}+b\bar{b}$, and $t\bar{t}+c\bar{c}$ events with ≥ 6 jets and ≥ 4 b-tags

Source	Rate	Shape?
QCD Scale (all $t\bar{t}$+hf)	35%	No
QCD Scale ($t\bar{t}+b\bar{b}$)	17%	No
b-Tag bottom-flavor contamination	17%	Yes
QCD Scale ($t\bar{t}+c\bar{c}$)	11%	No
Jet Energy Scale	11%	Yes
b-Tag light-flavor contamination	9.6%	Yes
b-Tag bottom-flavor statistics (linear)	9.1%	Yes
QCD Scale ($t\bar{t}$+b)	7.1%	No
Madgraph Q^2 Scale ($t\bar{t}+b\bar{b}$)	6.8%	Yes
b-Tag Charm uncertainty (quadratic)	6.7%	Yes
Top p_T Correction	6.7%	Yes
b-Tag bottom-flavor statistics (quadratic)	6.4%	Yes
b-Tag light-flavor statistics (linear)	6.4%	Yes
Madgraph Q^2 Scale ($t\bar{t}+2$ partons)	4.8%	Yes
b-Tag light-flavor statistics (quadratic)	4.8%	Yes
Luminosity	4.4%	No
Madgraph Q^2 Scale ($t\bar{t}+c\bar{c}$)	4.3%	Yes
Madgraph Q^2 Scale ($t\bar{t}$+b)	2.6%	Yes
QCD Scale ($t\bar{t}$)	3%	No
pdf (gg)	2.6%	No
Jet Energy Resolution	1.5%	No
Lepton ID/Trigger efficiency	1.4%	No
Pileup	1%	No
b-Tag Charm uncertainty (linear)	0.6%	Yes

Background Composition Comparison, 6j4t

- CMS "Baseline" vs ATLAS results, taken from conference note
  - Error bars on CMS numbers are from ***shape systematics only*** from limit file
- Scaled to 20.3 fb^{-1}, ATLAS lumi
- tt+bb = tt+ 1,2b
- f_{tt+l} = fraction of tt +Inclusive Jets
- Pre-fit yields for both experiments

$\geq$6 jets, $\geq$4 tags	ATLAS	CMS	ATLAS/CMS
$\mathbf{t\bar{t}H, M125}$	16.0 ± 2.8	8.5 ± 2.0	1.9 ± 0.5
$\mathbf{t\bar{t} + lf}$	67.0 ± 20.0	73.7 ± 38.8	0.9 ± 0.6
$\mathbf{t\bar{t} + c\bar{c}}$	80.0 ± 40.0	53.7 ± 12.5	1.5 ± 0.8
$\mathbf{t\bar{t} + b\bar{b}}$	240.0 ± 130.0	115.6 ± 22.1	2.1 ± 1.2
$\mathbf{t\bar{t} + X}$	387.0 ± 137.5	243.0 ± 46.4	1.6 ± 0.6
$\mathbf{f_{t\bar{t}+lf}}$	0.17 ± 0.1	0.30 ± 0.2	0.57 ± 0.4
$\mathbf{f_{t\bar{t}+c\bar{c}}}$	0.21 ± 0.1	0.22 ± 0.1	0.94 ± 0.6
$\mathbf{f_{t\bar{t}+b\bar{b}}}$	0.62 ± 0.4	0.48 ± 0.1	1.30 ± 0.9
$\mathbf{t\bar{t} + V}$	8.4 ± 2.8	6.0 ± 1.0	1.4 ± 0.5
$\mathbf{Non - t\bar{t}}$	22.0 ± 5.0	9.0 ± 1.7	2.4 ± 0.7
Total Bkg	430.0 ± 160.0	258.1 ± 46.4	1.7 ± 0.7
Data	516.0	271.4	1.9

- In this category, fractional composition is rather different for ATLAS and CMS, with ATLAS claiming an increase in ttH signal but also tt+bb background

Background Composition Comparison, jet/lepPt>25, 6j4t

- CMS jet/lep pT>25 vs ATLAS results
  - Error bars on CMS numbers are from ***shape systematics only*** from limit file
- Scaled to 20.3 fb^{-1}, ATLAS lumi
- tt+bb = tt+ 1,2b
- f_{tt+l} = fraction of tt +Inclusive Jets
- Pre-fit yields for both experiments

$\geq$6 jets, $\geq$4 tags

	ATLAS	CMS	ATLAS/CMS
$t\bar{t}H$, M125	16.0 ± 2.8	13.2 ± 3.0	1.2 ± 0.3
$t\bar{t} + lf$	67.0 ± 20.0	138.7 ± 74.9	0.5 ± 0.3
$t\bar{t} + c\bar{c}$	80.0 ± 40.0	91.6 ± 19.9	0.9 ± 0.5
$t\bar{t} + b\bar{b}$	240.0 ± 130.0	202.5 ± 37.2	1.2 ± 0.7
$t\bar{t} + X$	387.0 ± 137.5	432.8 ± 86.0	0.9 ± 0.4
$f_{t\bar{t}+lf}$	0.17 ± 0.1	0.32 ± 0.2	0.54 ± 0.4
$f_{t\bar{t}+c\bar{c}}$	0.21 ± 0.1	0.21 ± 0.1	0.98 ± 0.7
$f_{t\bar{t}+b\bar{b}}$	0.62 ± 0.4	0.47 ± 0.1	1.33 ± 0.9
$t\bar{t} + V$	8.4 ± 2.8	9.3 ± 1.6	0.9 ± 0.3
Non $- t\bar{t}$	22.0 ± 5.0	16.3 ± 2.8	1.3 ± 0.4
Total Bkg	430.0 ± 160.0	458.4 ± 86.0	0.9 ± 0.4
Data	516.0	472.8	1.1

- Still less ttH and tt+bb than ATLAS, but closer than in baseline analysis
- Fractional composition is ~same as CMS baseline analysis

Summary of Analysis Configuration Studies

- ***Right Col., blind, would guide analysis decisions; Center Col., partial unblind, for ATLAS comparison***

Scenario	Exp. Limit ("Unblinded")	Exp Limit ("Blind", "-t –1")
Baseline	4.83	3.80
jetLep pT>25, new BDT weights	4.45	3.83
Baseline, add HT var for 4j2t, 5j2t	4.14	3.52
jetLep pT>25, new BDT weights, add HT var for 4j2t, 5j2t	3.94	3.70

- ***Many other options have been tested as well, focusing on background composition and other effects***

Scenario	Exp. Limit ("Unblinded")	Exp Limit ("Blind", "-t –1")
tt+lf to ATLAS	--	3.58
tt+cc/bb to ATLAS	--	3.98
tt+lf/cc/bb to ATLAS	--	3.73
ttH, tt+lf/cc/bb to ATLAS	--	3.17
ttH, tt+lf/cc/bb to ATLAS, add HT var for 4j2t, 5j2t	--	3.02
Baseline, add BDT var for 4j2t, 5j2t	4.17	3.52
jetLep pT>25, new BDT weights, add BDT var for 4j2t, 5j2t	3.98	3.70

Summary of Analysis Configuration Studies

- ***Right Col., blind, is all we would have access to when designing the analysis***

Scenario	Exp. Limit ("Unblinded")	Exp Limit ("Blind", "-t –1")
Baseline	4.83	3.80
jetLep pT>25, new BDT weights	4.45	3.83
Baseline, add HT var for 4j2t, 5j2t	4.14	3.52
jetLep pT>25, new BDT weights, add HT var for 4j2t, 5j2t	3.94	3.70

- ***Many other options have been tested as well, focusing on background composition and other effects***

Scenario	Exp. Limit ("Unblinded")	Exp Limit ("Blind", "-t –1")
tt+lf to ATLAS	--	3.58
tt+cc/bb to ATLAS	--	3.98
tt+lf/cc/bb to ATLAS	--	3.73
ttH, tt+lf/cc/bb to ATLAS	--	3.17
ttH, tt+lf/cc/bb to ATLAS, add HT var for 4j2t, 5j2t	--	3.02
Baseline, add BDT var for 4j2t, 5j2t	4.17	3.52
jetLep pT>25, new BDT weights, add BDT var for 4j2t, 5j2t	3.98	3.70