Measurement of the top-Higgs coupling at CMS

Evan Wolfe
University of Virginia
April 13 2016

Overview

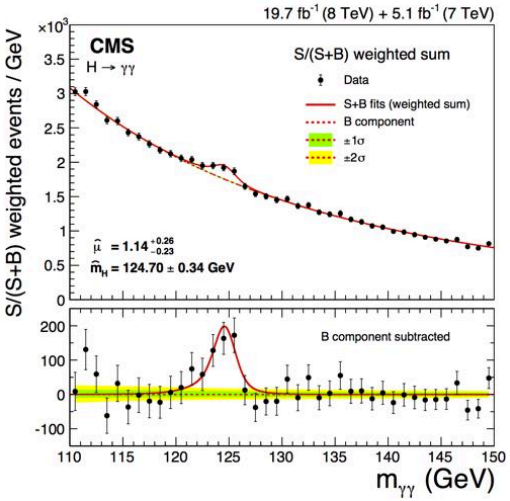
- Motivation
- Experimental (LHC and CMS Detector Overview)
- Searching for ttH
- Event Categories
- Reconstructing Events at CMS
- Analysis Techniques
- Results

UNIVERSITY of VIRGINIA

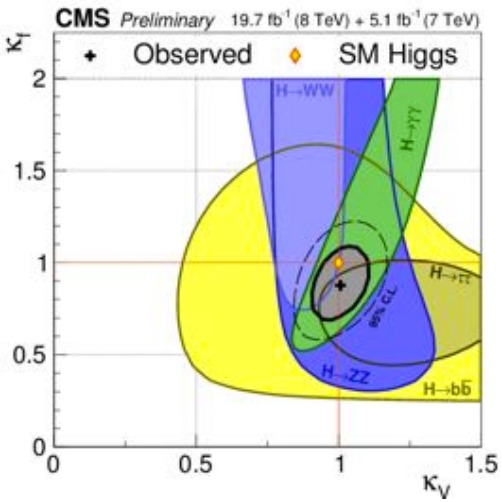
Announcement on July 4, 2012

Motivation

- Discovery of 125 GeV Higgs
  - We now move from the era of searching to characterization
  - Thus far appears to be SM like…

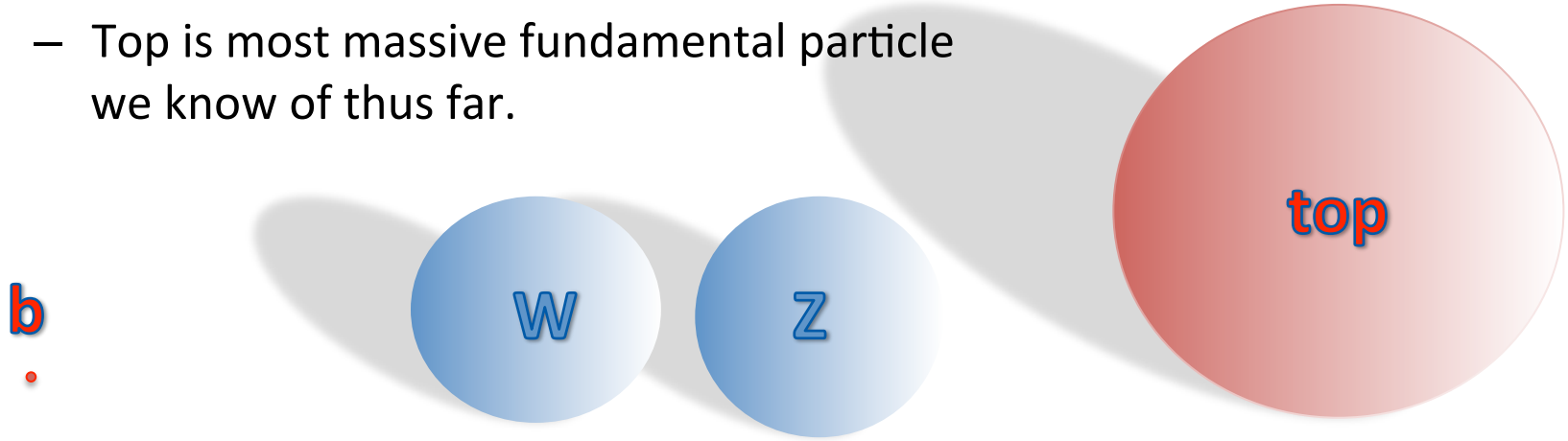


- One of the most important characteristics that remains to be measured:
  - The coupling between the top quark and the Higgs

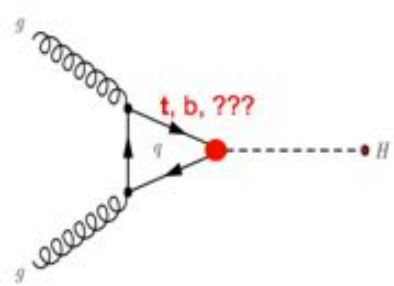


Importance of top-Higgs coupling

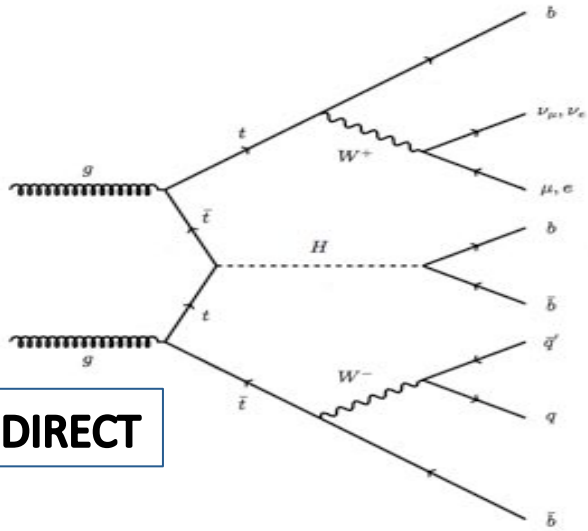
- Why is this coupling important?
  - Top is most massive fundamental particle we know of thus far.



- Thus far only indirect information
  - ttH provides a direct probe for measurement.



INDIRECT

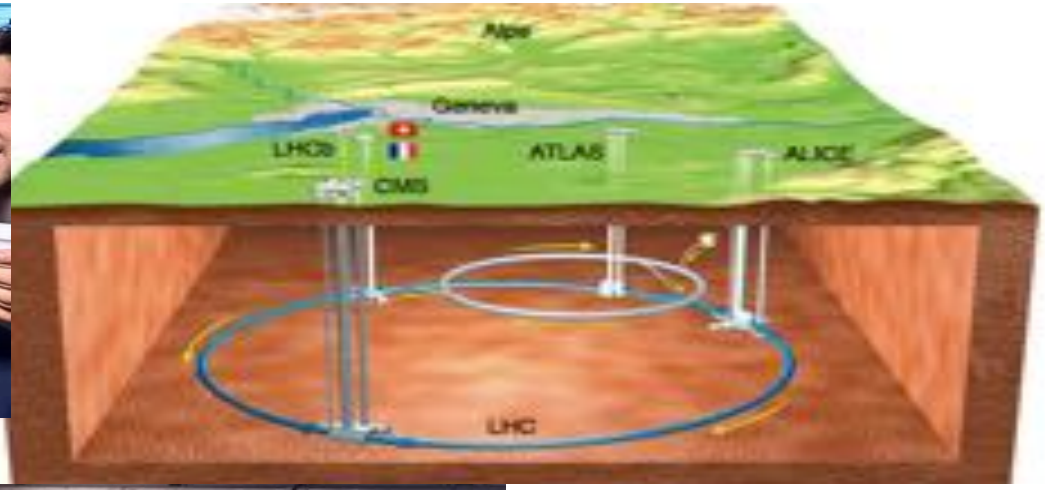


DIRECT

UNIVERSITY of VIRGINIA

The Large Hadron Collider

- 27 km (17 miles) in circumference beamlines spanning into France and Switzerland located 100 m underground
- 1232 dipole magnets, 392 quadrapole magnets
- 13 TeV collisions (2015)



UNIVERSITY of VIRGINIA

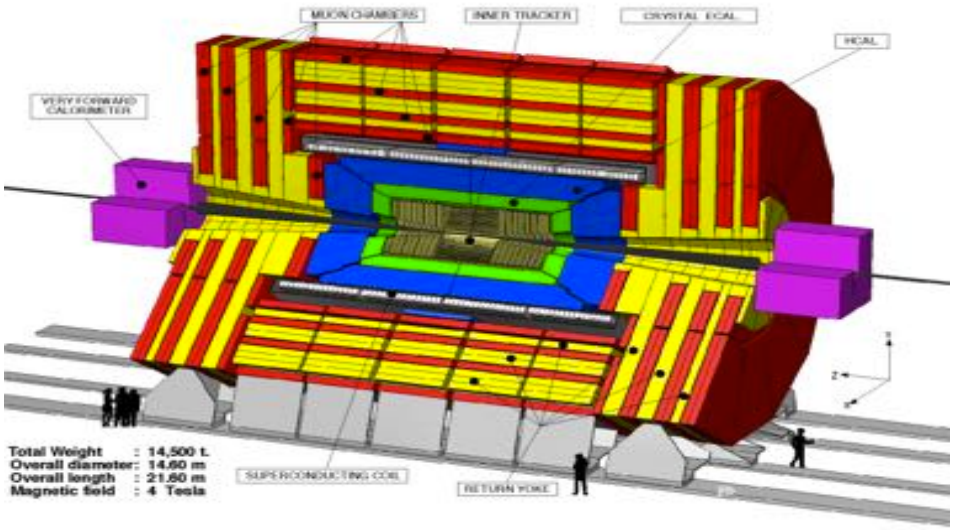
The CMS Detector

Multiple layered detector

- **PIXELS:** 66 million channels
- **TRACKER:** 10 million channels
- **ECAL**: PbWO_4 crystals 75k channels
- **HCAL**: Brass/Scintillator 7k channels

-3.8 T Solenoid (2.3 GJ)

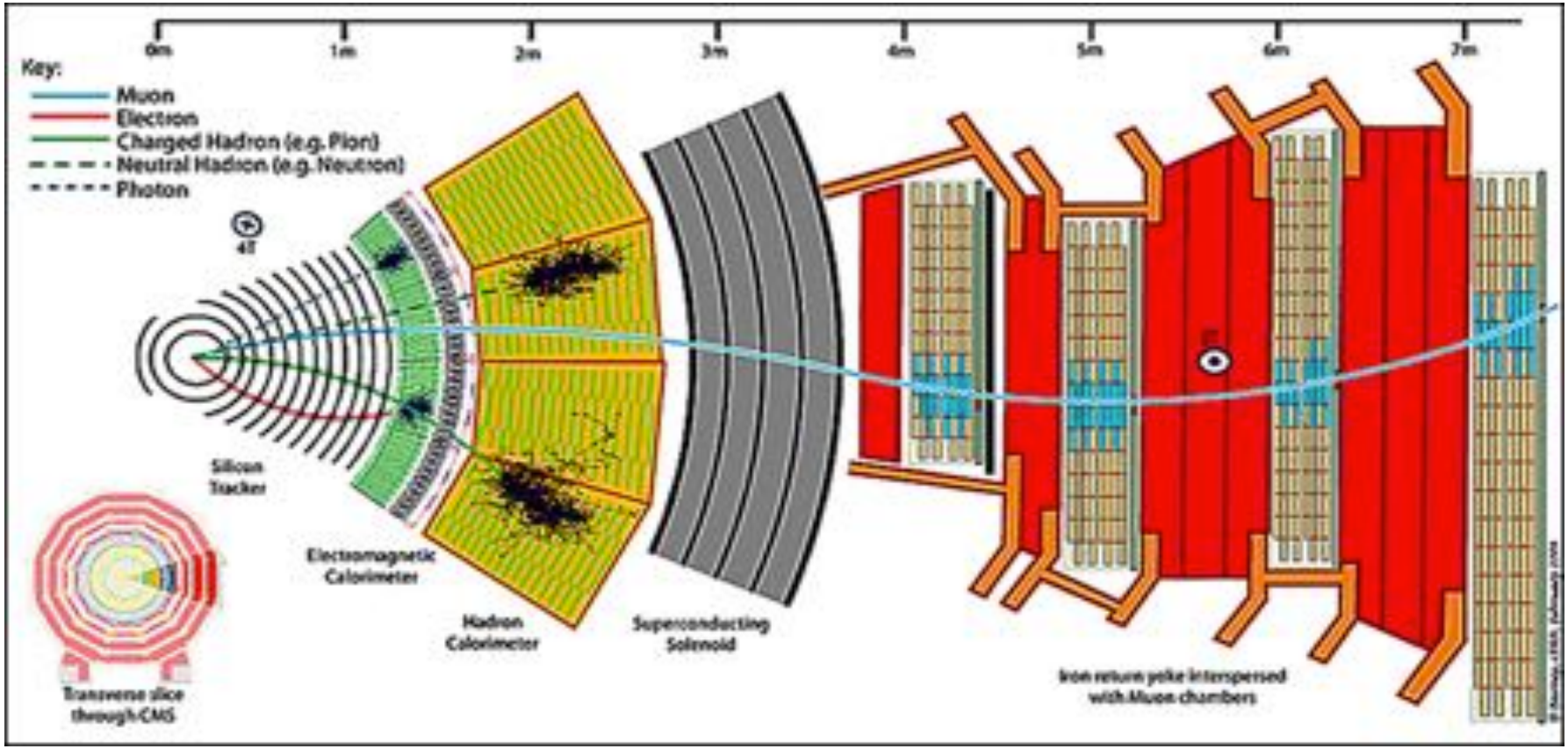
- **MUON** (250 DT, 480 RPC)



UNIVERSITY of VIRGINIA

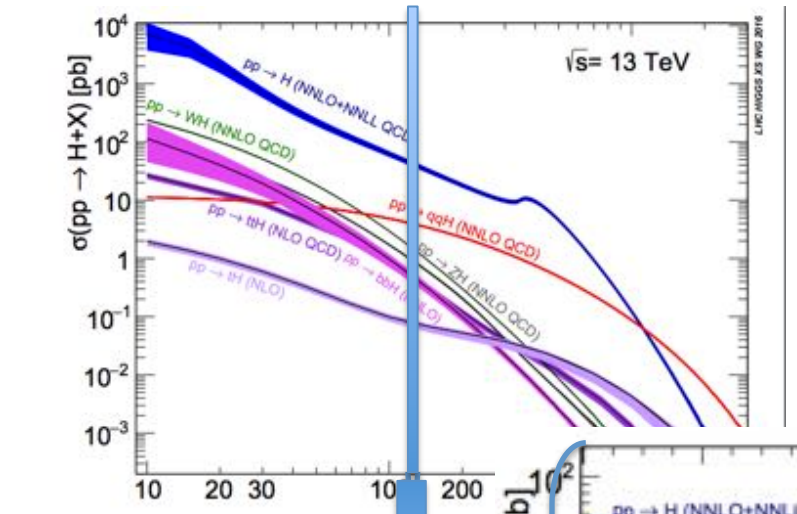
Reconstructing Objects in CMS

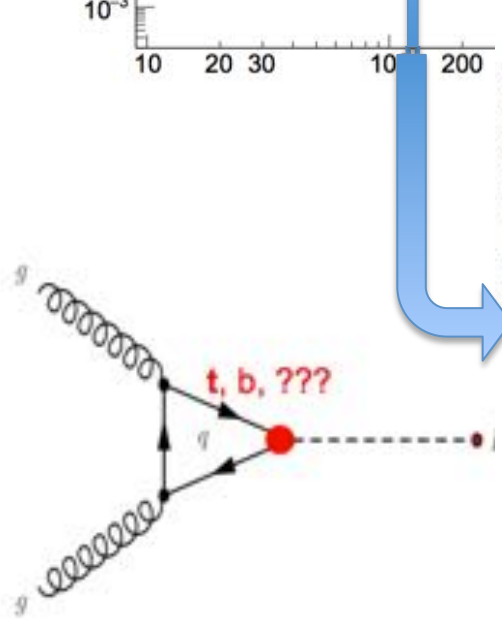
- Particle Flow algorithm uses calorimeter deposits and tracks to give events particle id and description

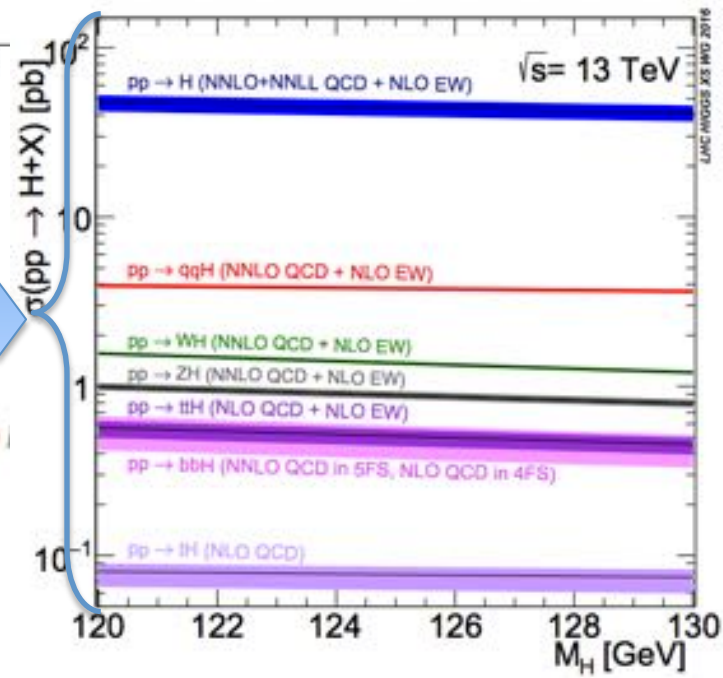


Searching for top-Higgs Coupling at CMS: Higgs Production

- Main production mechanism at LHC though gluon-gluon fusion
- Mediated through loop diagrams assuming SM
- Doesn't allow direct measurement of top quark and Higgs coupling

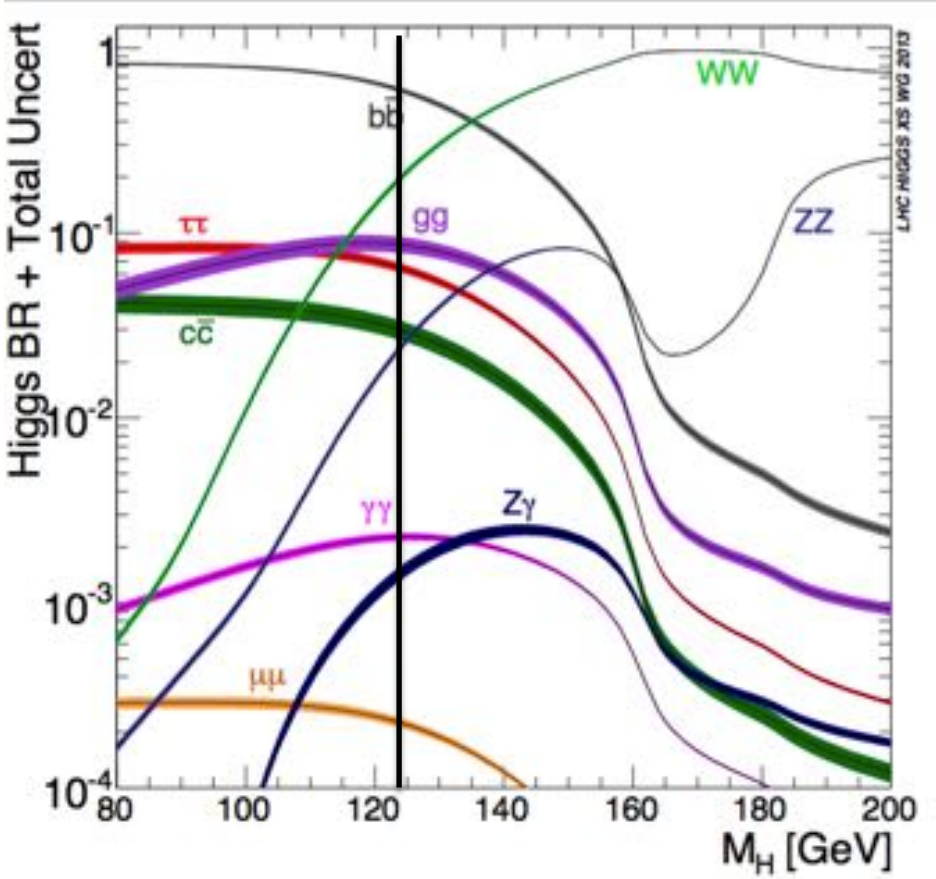






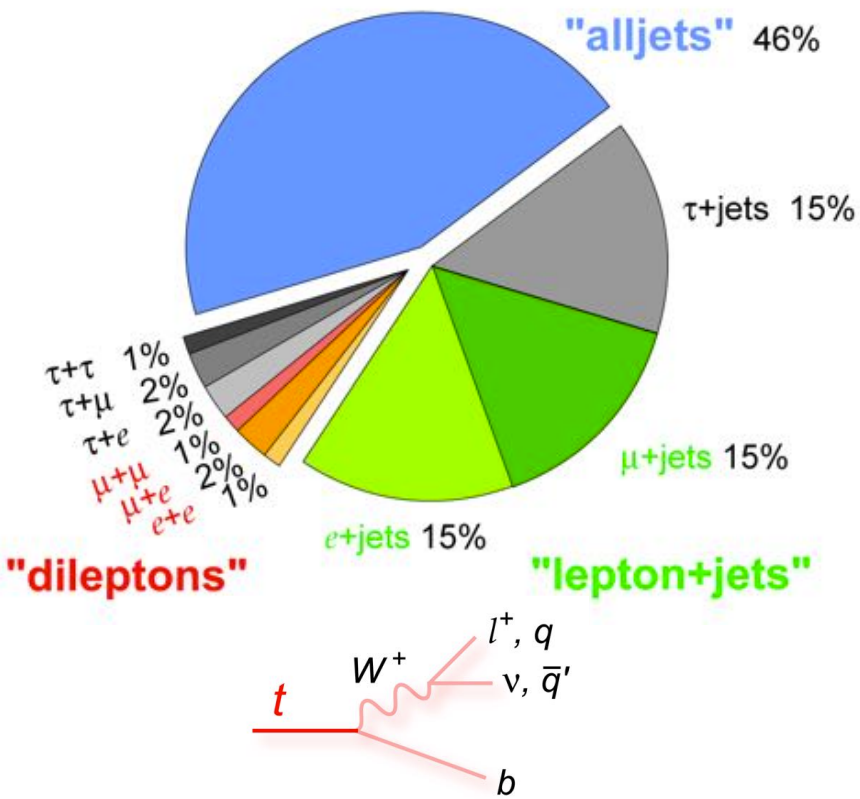
Searching for top-Higgs Coupling at CMS: Higgs Decays

- 125 GeV Higgs has highest branching ratio to bb pair
- Top-Higgs coupling not accessible through Higgs decay
- Best chance is Higgs production in association with top-quark pair and decay to bb…

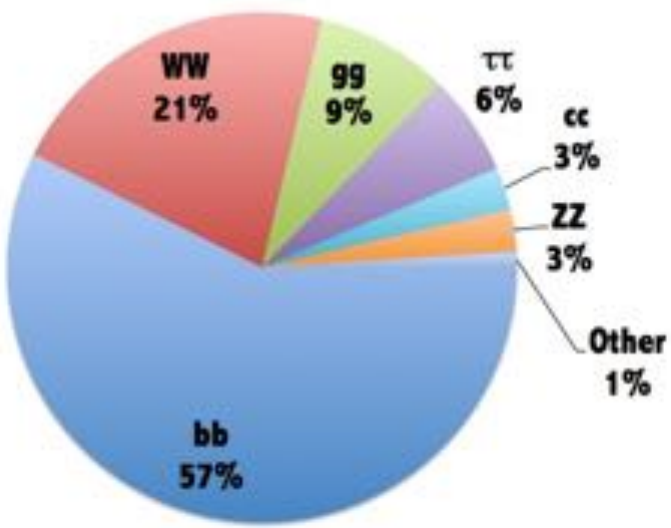


Searching for top-Higgs Coupling at CMS: top + Higgs

- Combinations from both top and Higgs cause many complex final states and combinatorial headaches.



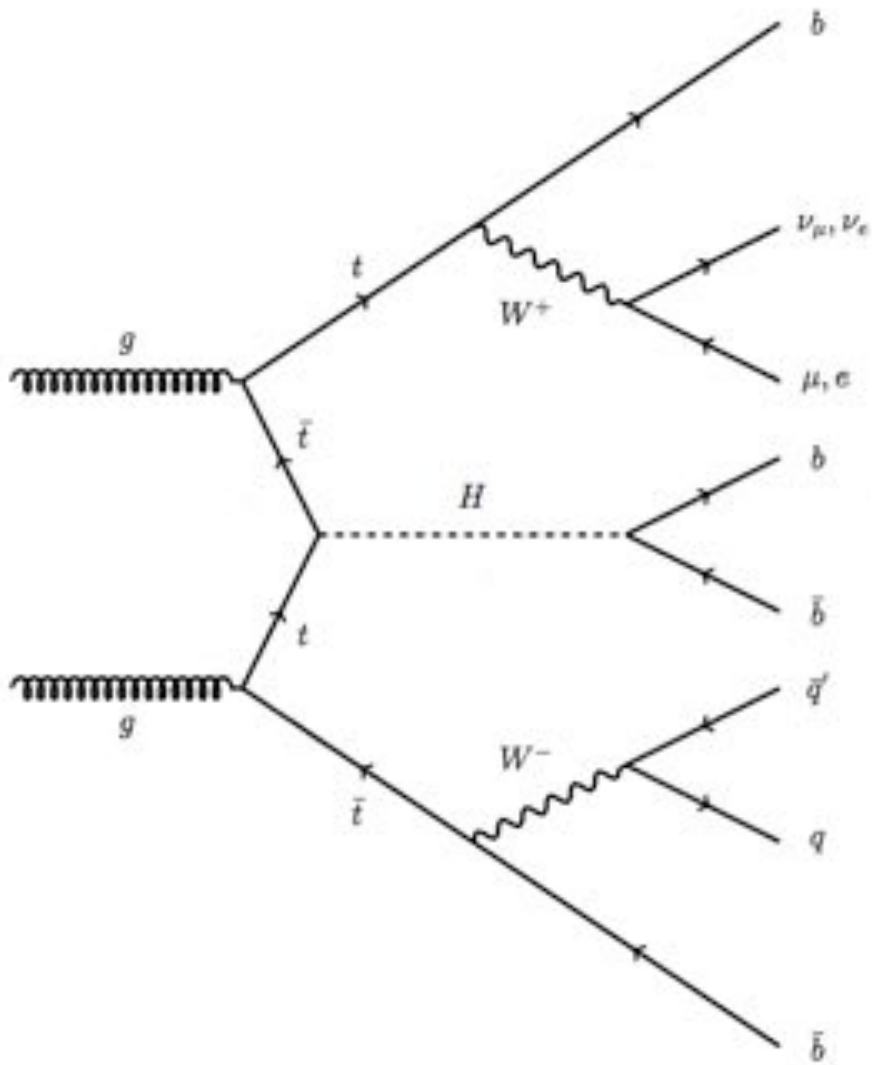
X



Searching for ttH

- ttH, H→bb provides direct measurement of top-Higgs coupling
  - Gluon-Gluon fusion
  - No loop mediation
  - Higgs decay to bb

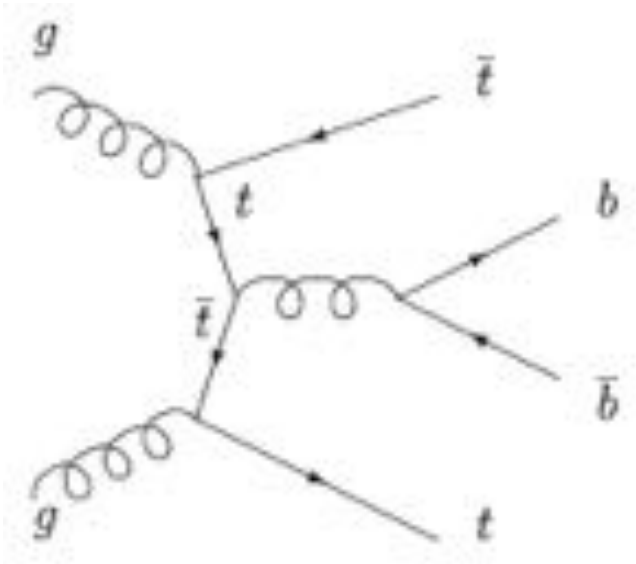
- Complex signature with many final states.



UNIVERSITY of VIRGINIA

Searching for ttH: Background

- Major background from top pair production with jets
  - Specifically b-jets
- tt + bb events have similar signatures to ttH



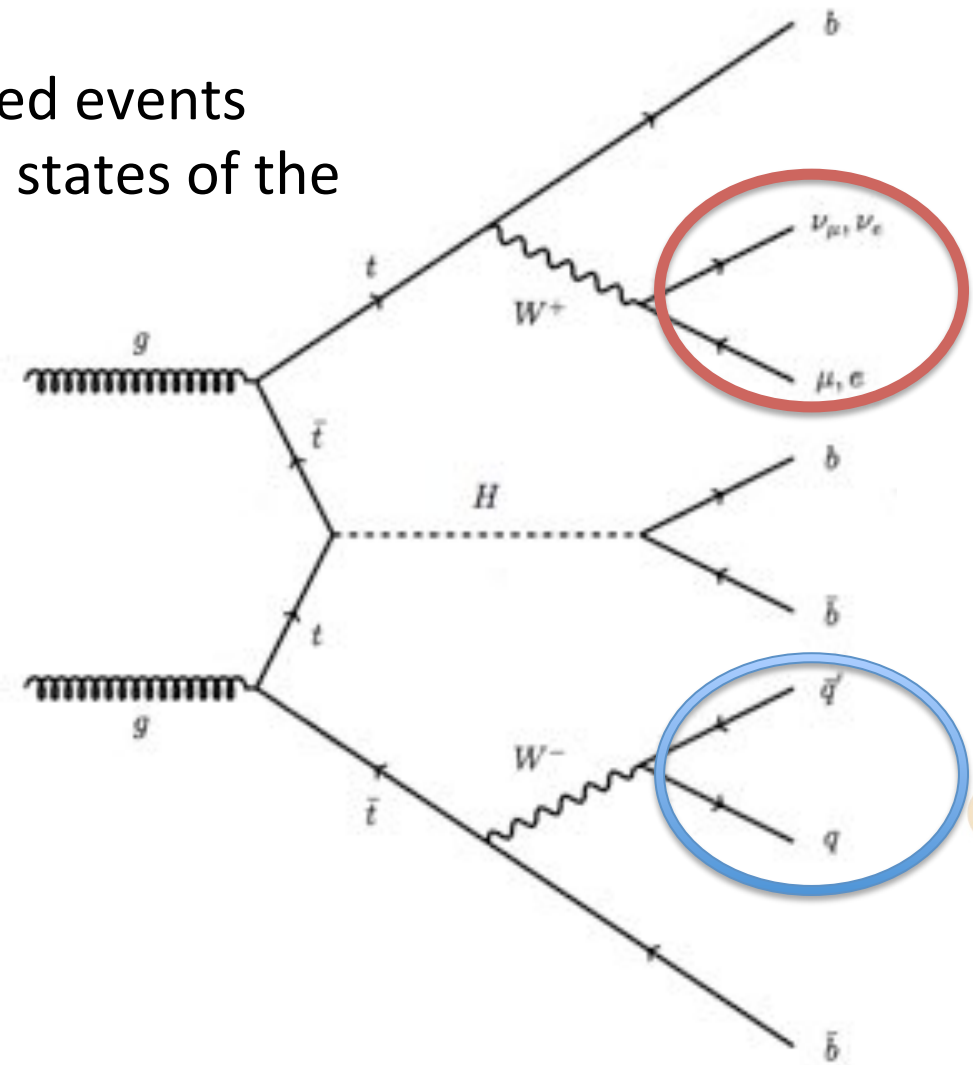
- Cross section of tt+jets is **HUGE** compared to ttH

	7 TeV	14 TeV
σ[ttH(125)]	86 fb	611 fb
σ [tt+jets]	177000 fb	950000 fb

UNIVERSITY of VIRGINIA

Searching for ttH: Decay Channels

- Focused analysis on categorized events according to two specific final states of the W decays.
- For Higgs → bb
  - Semileptonic(lepton+jets)
    - W → l ν
    - W → q q'
  - Dilepton
    - Both W → l ν



UNIVERSITY of VIRGINIA

Semileptonic Channel

- Semileptonic Events:
  - Require each data event passes either:

Dataset	Trigger Name
SingleMu	HLT_IsoMu20_v*
SingleEle	HLT_Ele27_eta2p1_WPLoose_Gsf_v*

  - exactly 1 lepton(e or μ)
    - Veto events with additional identifed leptons to ensure no overlap with dileptonic events.
    - Pt > 25GeV
    - ΔR = 0.3 isolation
  - At least 4 jets among which atleast 2 are b-tagged.
  - OR fulfill boosted event selection

UNIVERSITY of VIRGINIA

Dileptonic Channel

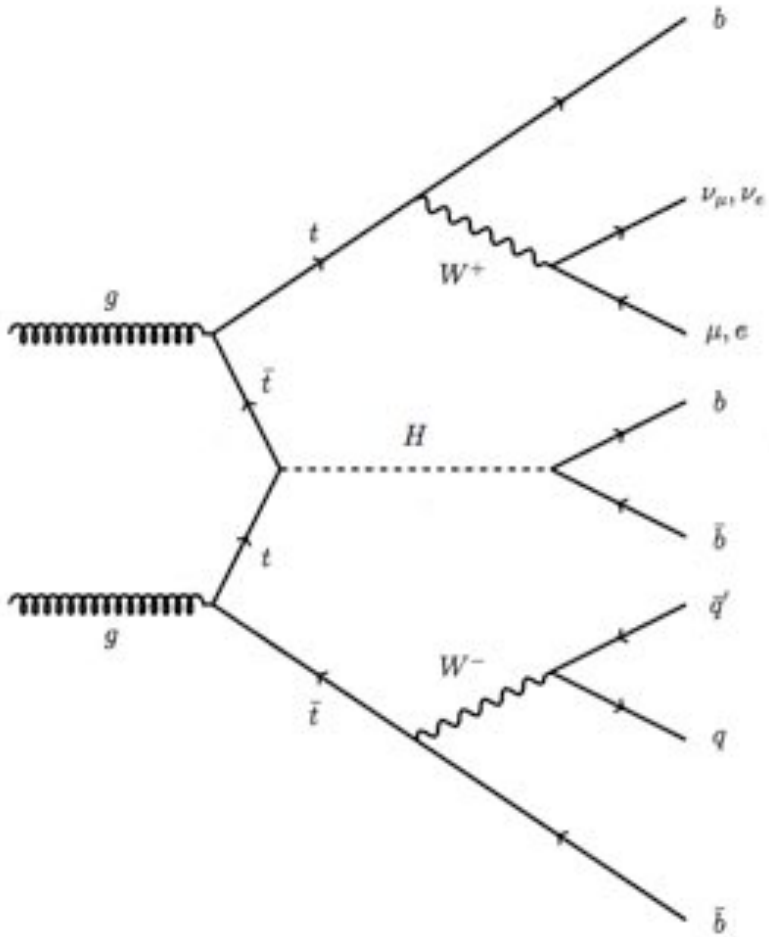
- Dileptonic Events:
  - Require each data event passes triggers:

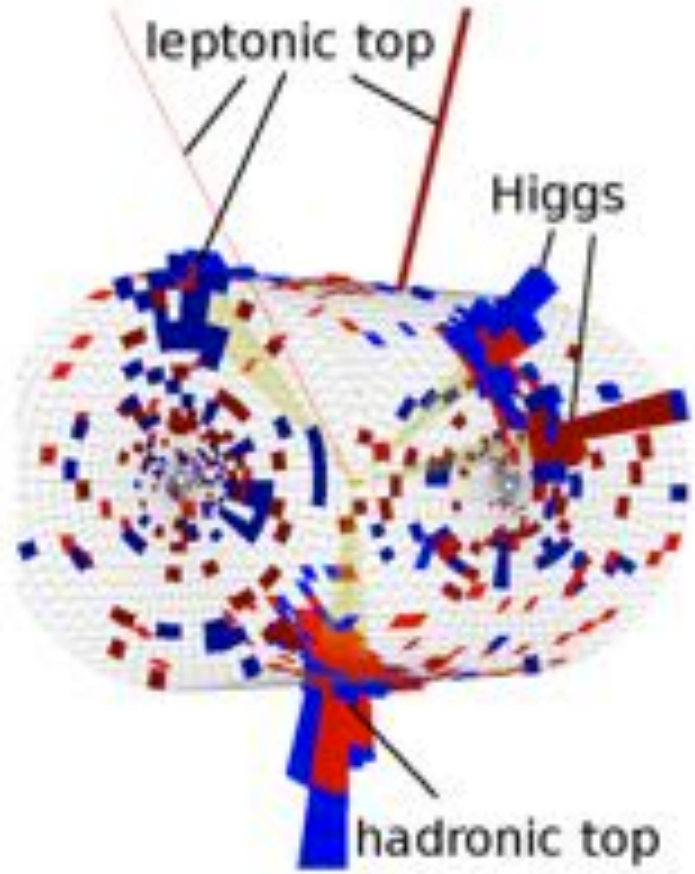
Channel	Trigger Name
$\mu^+\mu^-$	HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ_v
$\mu^+\mu^-$	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_v*
e^+e^-	HLT_Ele17_Ele12_CaloIdL_TrackIdL_IsoVL_v*
$\mu^\pm e^\mp$	HLT_Mu17_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_v*
$\mu^\pm e^\mp$	HLT_Mu8_TrkIsoVVL_Ele17_CaloIdL_TrackIdL_IsoVL_v*

  - Exactly 2 oppositely charge leptons(e or μ)
    - Leading lepton pt > 20GeV
    - Subleading lepton pt > 15GeV
    - ΔR = 0.3 isolation
  - Invariant mass of lepton pair > 20 GeV
  - Events in dimuon and dielectron channels:
    - m_{ll} > 106GeV or m_{ll} < 76 GeV
    - missing Et > 40 GeV
  - At least 2 jets with p_T > 30GeV and |η| < 2.4
    - Other jets > 20GeV

UNIVERSITY of VIRGINIA

Reconstructing Objects in CMS

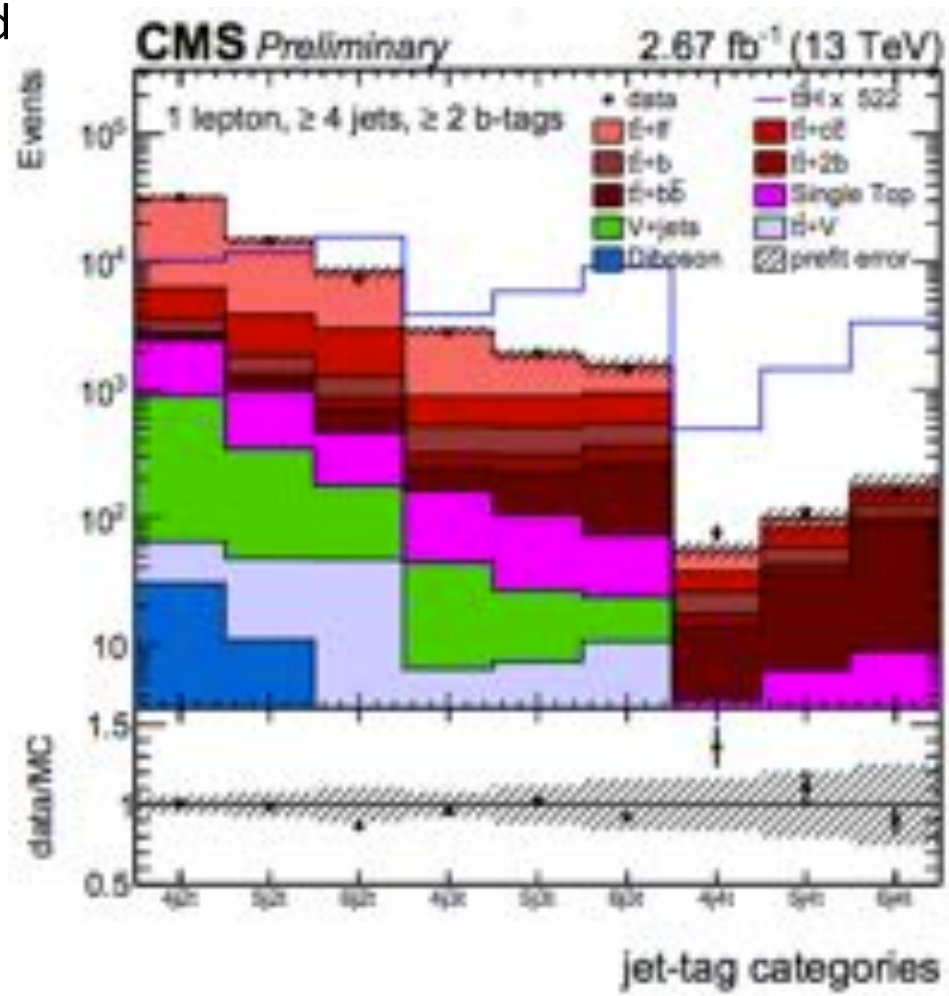




Event Selection and Reconstruction

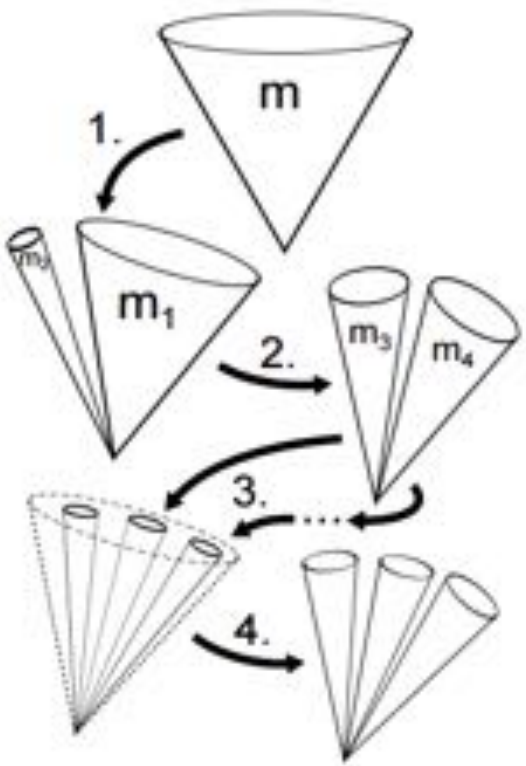
- Signal and Background both modeled with Monte Carlo
  - ttH Signal modeled with PYTHIA8
  - tt+jets, ttW,ttZ,W+jets modeled with MADGRAPH

- Objects reconstructed with Particle Flow algorithm for the two channels
  - Lepton + Jets:
    - e/μ + 4 jets
    - ≥2 b-tagged jets
  - Dilepton:
    - Pair of oppositely charged e/μ
    - ≥3 jets, 2 b-tagged jets



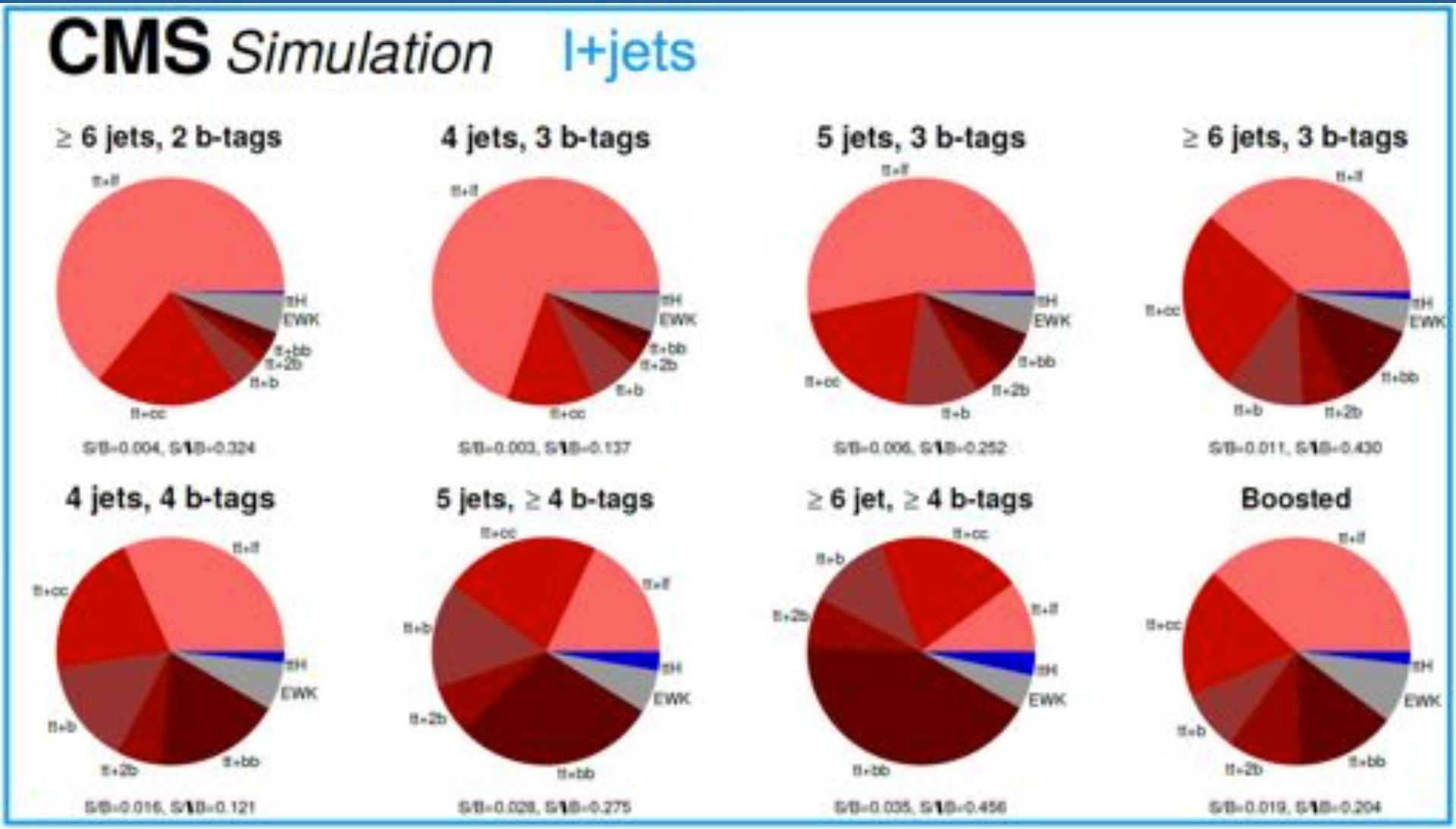
Boosted Analysis

- Large fraction of events at 13TeV include top quarks and Higgs bosons with $p_T \geq 200$ GeV
  - Boosted jet reconstruction techniques favorable
  - Implements HEP Top Tagger + sub-jet filtering
- Candidates require:
  - Exactly one selected lepton (only semi-leptonic)
  - $|\eta| < 2.0$
  - No sharing of subjets between hadronic top and Higgs



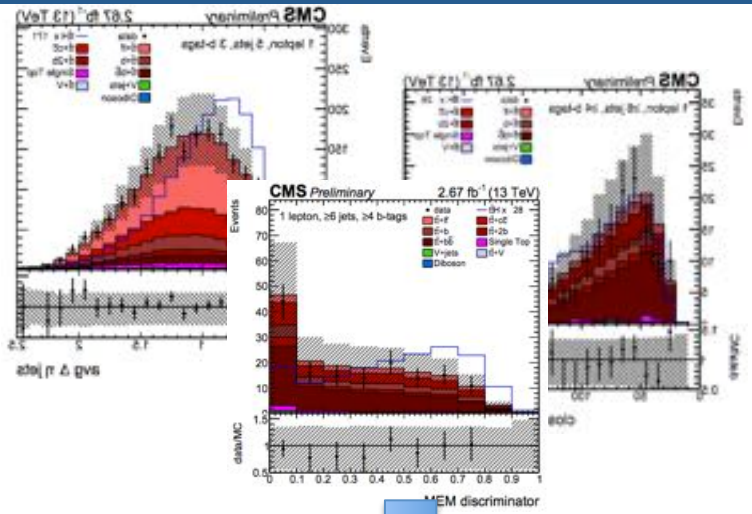
UNIVERSITY of VIRGINIA

Event Classification

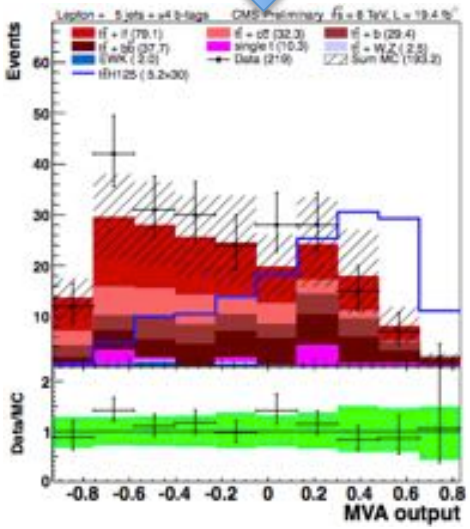


Multivariate Analysis Techniques

- For identifying ttH signal amidst large tt+jets background have handful of not so powerful discriminants.
- Use MVA to exploit higher-dimensional correlations among them.
- New discriminating variable has more power in identifying signal than individual variables alone.



Feeble 1D Discriminants



Powerful 1D Discriminant

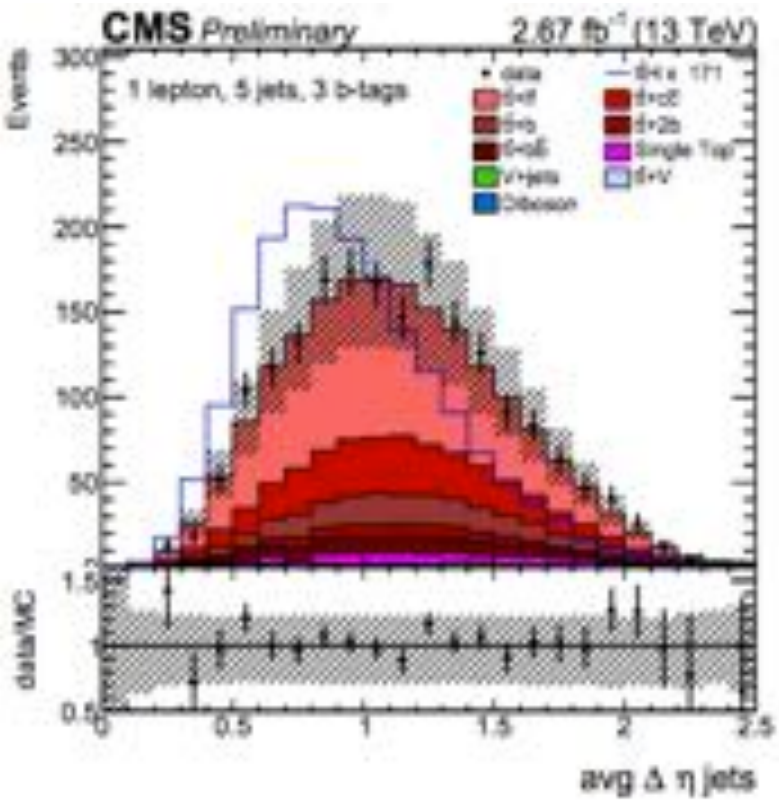
Boosted Decision Trees (BDT)

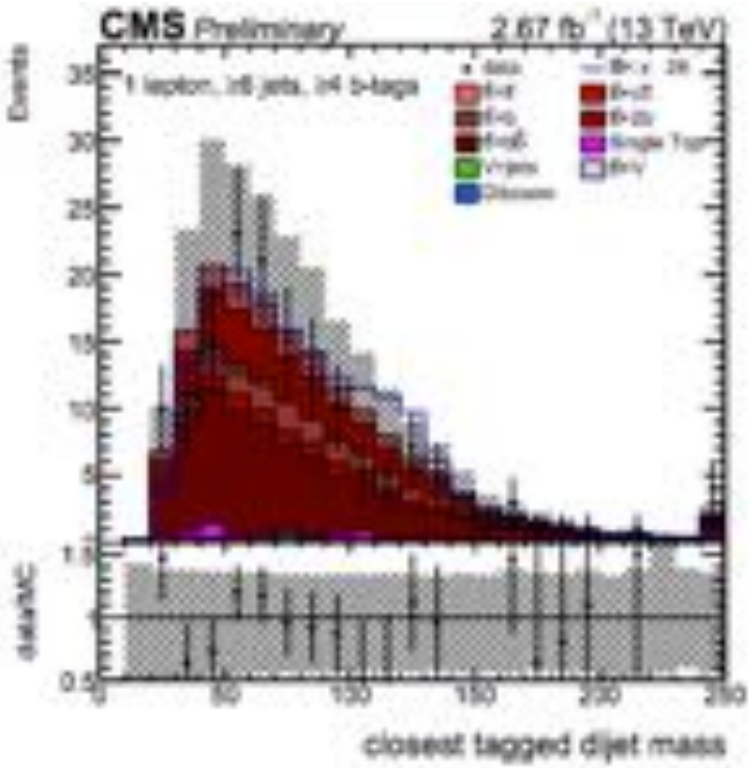
- Each category receives it's own BDT
  - 8 in lepton + jets channels
  - 5 in dilepton channels
- Training is performed using MC samples for ttH and tt +jets events
  - Weighted to yield equal events in signal and background
- To prevent overoptimization signal and background events are split in half
  - ½ to train the BDT's
  - ½ to test performance

UNIVERSITY of VIRGINIA

Boosted Decision Trees (BDT)

- Input Variables:
  - Object, Event Kinematics
  - CSVv2IVF b-tag
  - Event Shape
  - Boosted object and event reconstruction
  - MEM Discriminant

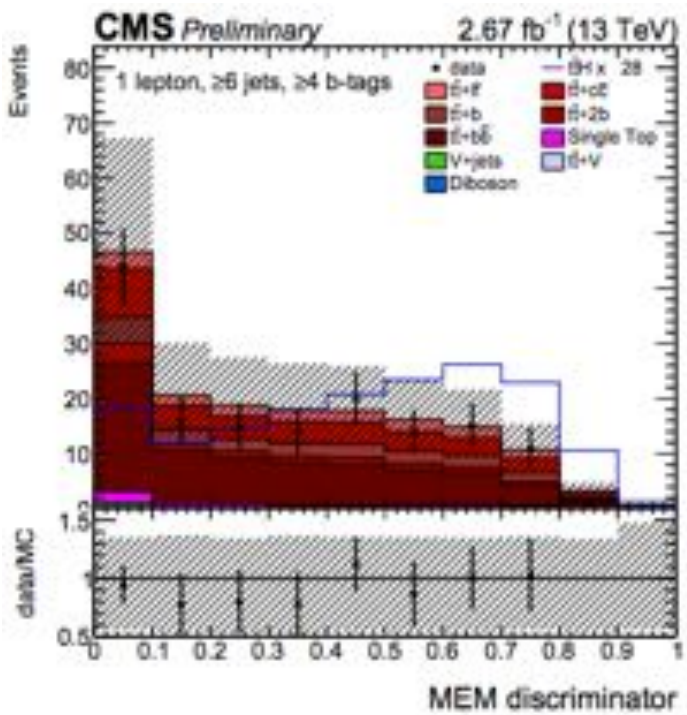




Matrix Element Method (MEM)

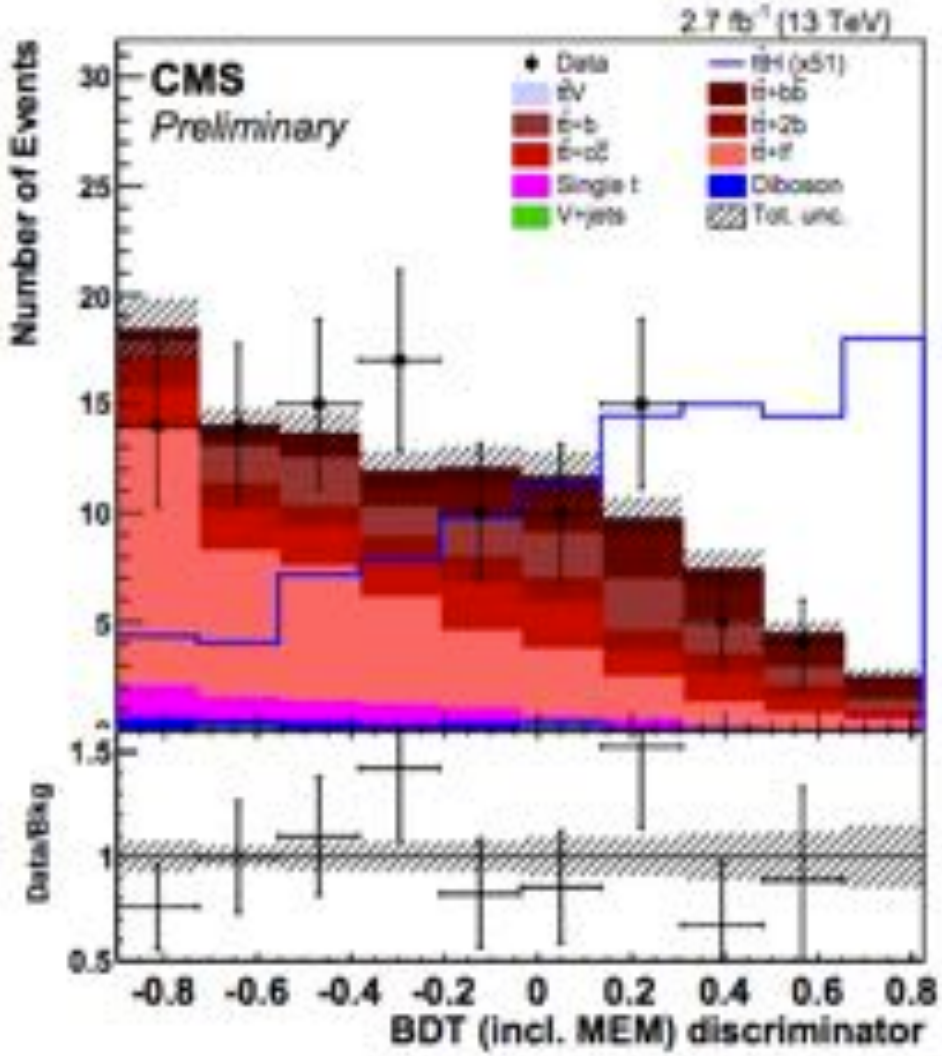
- Calculates the differential probability density of measuring a set of observables (y) under the hypothesis that the event is generated by a specific process (i)
- Provides separation between null hypothesis (signal) and the alternative hypothesis (background)
- Parton showering and detector response simulated with transfer functions from simulation (GEANT4)

$$P(\mathbf{y}|\lambda) = \frac{w_0(\mathbf{y}|\lambda)}{w_0(\mathbf{y}|\lambda) + w_1(\mathbf{y}|\lambda)}$$



BDT and MEM: Final Discriminants

- Final discriminator shape for lepton + jets
  - 6jets, ≥4 b-tags shown right
- Final discriminant outputs provide better separation than any individual input variables alone.



Systematic Uncertainties

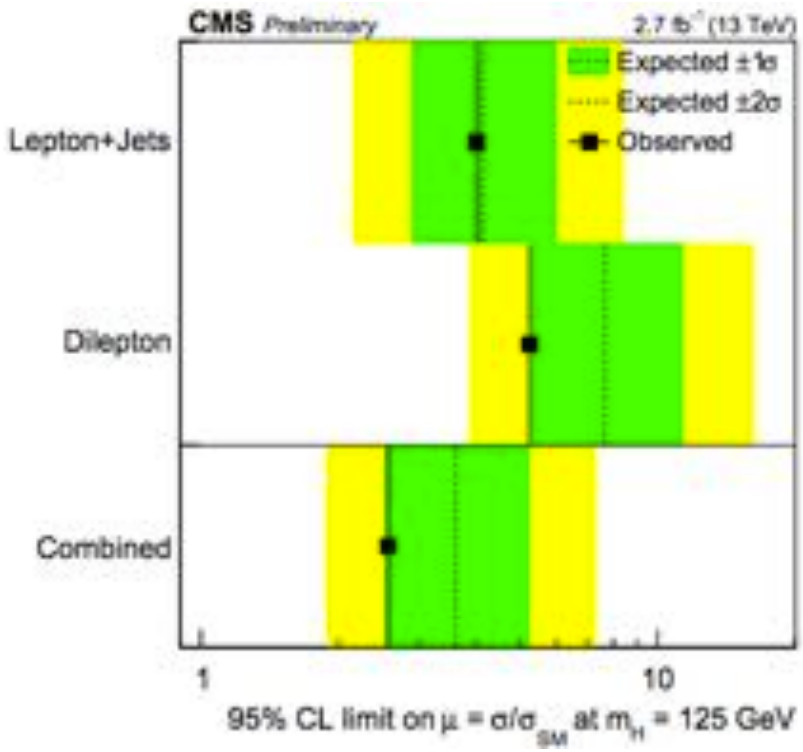
- Sources of systematic uncertainty affecting event yields (signal or background), discriminant shape, or both.

Process	$t\bar{t}$ rate up/down [%]	$t\bar{t}H$ rate up/down [%]
Jet energy Scale	+11.3/ − 10.1	+7.7/ − 7.0
Jet energy Resolution	−0.1/ + 0.1	−0.1/ + 0.1
Pile-Up	−0.1/ + 0.0	+0.1/ − 0.2
Electron Efficiency	+1.6/ − 1.6	+1.6/ − 1.6
Muon Efficiency	+1.2/ − 1.2	+1.2/ − 1.2
b-Tag HF contamination	−3.5/ + 8.4	+0.2/ + 0.6
b-Tag HF stats (linear)	−6.4/ + 6.2	−5.3/ + 4.9
b-Tag HF stats (quadratic)	+4.2/ − 4.4	+3.3/ − 3.6
b-Tag LF contamination	+7.1/ − 5.1	+5.5/ − 4.2
b-Tag LF stats (linear)	−3.2/ + 6.5	−0.6/ + 1.1
b-Tag LF stats (quadratic)	+0.5/ + 1.2	−0.8/ + 1.1
b-Tag charm Uncertainty (linear)	−12.6/ + 16.9	−0.6/ + 0.7
b-Tag charm Uncertainty (quadratic)	+1.4/ − 1.4	+0.0/ − 0.0
Q2 scale ($t\bar{t}$+lf)	−1.9/ + 2.8	–
Q2 scale ($t\bar{t}$+b)	−0.6/ + 0.9	–
Q2 scale ($t\bar{t}$+2b)	−0.5/ + 0.8	–
Q2 scale ($t\bar{t}$+$b\bar{b}$)	−0.9/ + 1.3	–
Q2 scale ($t\bar{t}$+$c\bar{c}$)	−1.6/ + 2.4	–
PS scale ($t\bar{t}$+lf)	4.4/ − 8.7	–
PS scale ($t\bar{t}$+b)	−1.3/ + 0.8	–
PS scale ($t\bar{t}$+2b)	−1.0/ + 0.4	–
PS scale ($t\bar{t}$+$b\bar{b}$)	−2.0/ + 1.3	–
PS scale ($t\bar{t}$+$c\bar{c}$)	−4.3/ + 2.3	–

UNIVERSITY of VIRGINIA

Current Results from 13TeV

- We incorporate BDT discriminate shapes from each category and perform a simultaneous fit across all.
- Seeing no excess, and without sufficient sensitivity to probe the expected-SM-level content, we place upper limits
- 2.7 fb^{-1} of pp collision data
- Observed upper limit on the ttH production cross section
- Signal strength modifier
  - $\mu = \sigma/\sigma_{SM}$



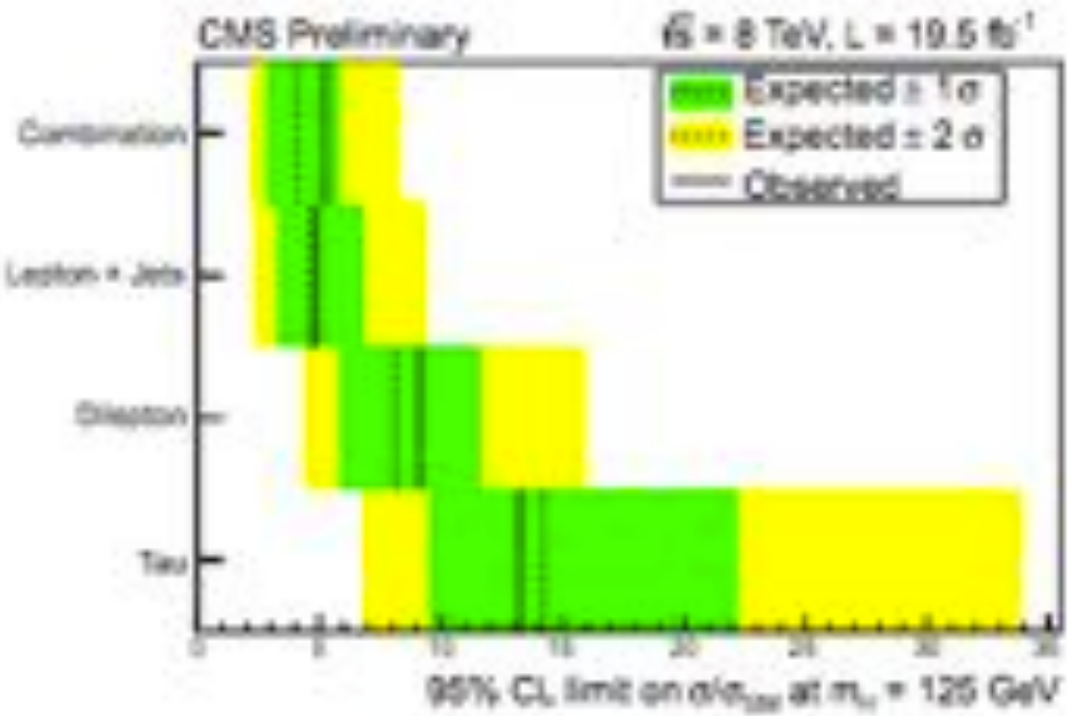
$\mu = 2.6$ at 95% CL

Results from 8TeV

Median limit for m_H = 125 GeV:

μ = 5.2 at 95% CL

- Current upper limit has decreased the upper limit by a factor of 2x



Publications

Analysis Note
and
Physics Analysis Summary

CMS Draft Analysis Note

The content of this note is intended for CMS internal use and distribution only

2016/03/04
Head Id: 330169
Archive Id: 330533
Archive Date: 2016/03/03
Archive Tag: trunk

First results on $t\bar{t}H$ with $H \to b\bar{b}$ at 13 TeV

Maria Aldaya[1], Jim Alexander[2], Lorenzo Bianchini[3], Lea Caminada[4], Christian Contreras Campana[1], Florencia Canelli[4], Abhisek Datta[2], Günther Dissertori[3], Karim El Morabit[5], Sean Flowers[6], Jasone Garay Garcia[1], Marco Harrendorf[5], Satoshi Hasegawa[7], Johannes Hauk[1], Ulrich Husemann[8], Gregor Kasieczka[3], Thomas Klijnsma[3], Kevin Lannon[8], Hongbo Liao[9], Wuming Luo[6], Hannes Mildner[5], Chris Neu[10], Eleni Ntomari[1], Carmen Diez Pardos[1], Joosep Pata[3], Darren Puigh[6], Thorben Quast[11], Felix Riese[5], Aurelijus Rinkevicius[2], Francesco Romeo[9], Andrej Saibel[5], Daniel Salerno[4], Matthias Schröder[5], Korbinian Schweiger[5], Tutanon Sinthuprasith[10], Aniello Spiezia[9], Shao Min Tan[2], Julia Thom[2], Gerrit Van Onsem[1], Rainer Wallny[3], Michael Waßmer[5], Shawn Williamson[5], Brian Winer[6], Evan Wolfe[10], and Huaqiao Zhang[9]

CMS Physics Analysis Summary

Contact: cms-pag-conveners-higgs@cern.ch 2016/03/16

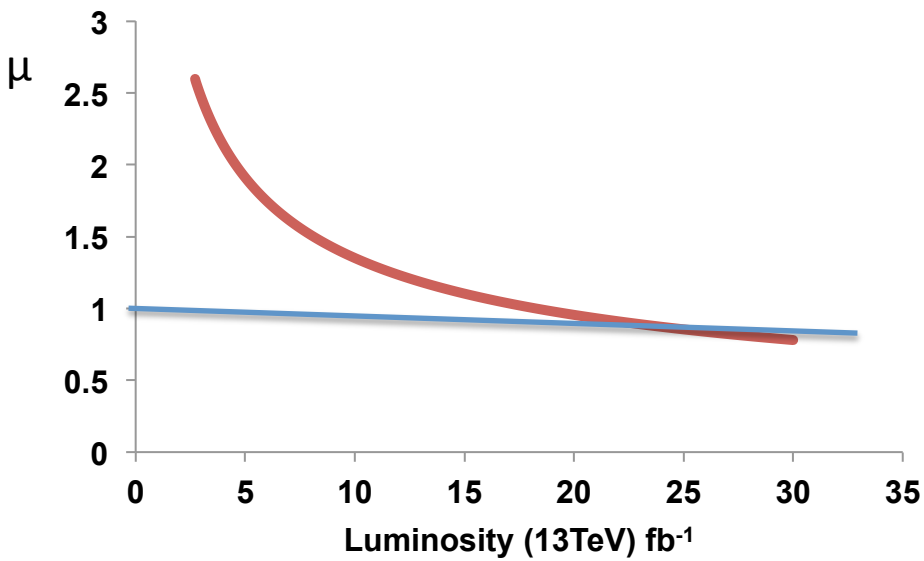
Search for $t\bar{t}H$ production in the $H \to b\bar{b}$ decay channel with $\sqrt{s} = 13$ TeV pp collisions at the CMS experiment

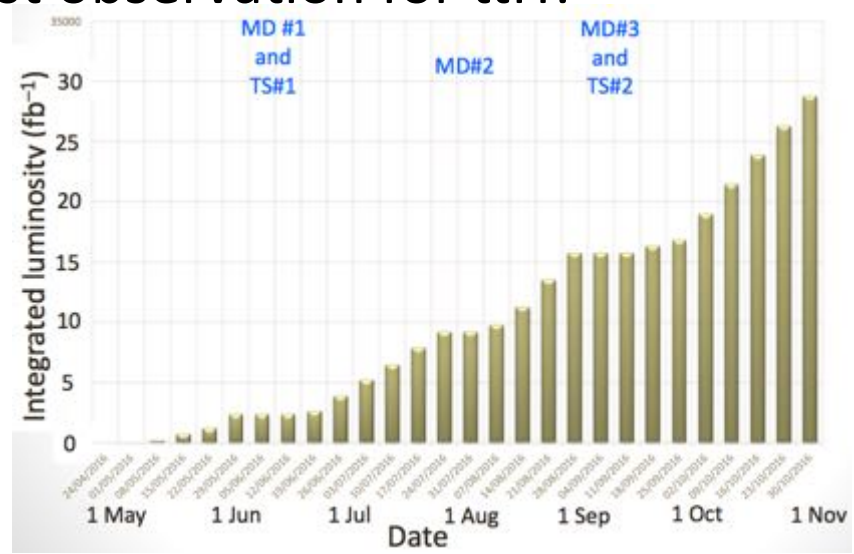
The CMS Collaboration

, Julia Thom[2], Gerrit Van Onsem[1]
Brian Winer[6], Evan Wolfe[10], and

Outlook for Run II

- Integrated luminosity for 2016 is estimated to deliver around 20fb^{-1} ~ 30 fb^{-1} of data.
- Crude estimated sensitivity by end of 2016 yields a decrease in the upper limit by a factor of 3!!!
- 2016 may just be the year of first observation for ttH!





Outlook for Run II

- New conditions and improvements being implemented:
  - Triggers for high luminosity
  - Improved tt+bb background modeling (NLO)
  - Improved ttH signal modeling
    - Adoption of NLO event generator
  - Higher purity b-jet identification
  - Improved signal extraction approaches
  - Better understanding of charm jet identification.
- Evaluation of improvements ongoing…

GOAL: Make first observation of ttH production at 13TeV with Run II dataset at CMS

- Provide first-ever direct measurement of top-Higgs coupling.

UNIVERSITY of VIRGINIA

Summary

- ttH is a crucial element for the understanding and characterization of the Higgs
- Our sensitivity continues to improve and analysis techniques further refined.
- Possibility of the first observation of ttH with the 2016 data sample.

UNIVERSITY of VIRGINIA

BACKUP

- BDT Input Variables
  - Object and Event Kinematics:
    - Mass and pt variables
      - 4 highest transverse jet momenta
      - Scalar Sum of all jet transverse momenta (HT)
    - Angular Variables
      - ΔR between b-tagged jets, ΔR between lepton and closest jet
      - $\Delta\eta$ between b-tagged jets
    - Event Reconstruction related variables
      - Best Higgs Mass
      - Median Invariant Mass of all combinations of b-tagged jet pairs
      - Scalar Sum of transverse momentum of all b-tagged jets and leptons
      - Invariant mass of 3-jet system with largest tansverse momentum where atleast 2 jets are b-tagged
      - Centrality
      - Twist Angle
    - CSVv2IVF b-tag
    - MEM Discriminator
      - Lepton+Jets uses as input

UNIVERSITY of VIRGINIA

Event Yields: Semileptonic

- Predicted and observed event yields after event selection in lepton + jets
- Shows good agreement with all final states across subdivided categories
  - Number of jets and number of b-tags

Process	$\geq$ 6 jets, 2 b-tags	4 jets, 3 b-tags	5 jets, 3 b-tags	$\geq$ 6 jets, 3 b-tags
$t\bar{t}$+lf	5359.3 ± 1226.3	2026.1 ± 651.4	1000.2 ± 352.9	589.5 ± 199.7
$t\bar{t} + c\bar{c}$	1722.2 ± 849.5	363.2 ± 190.9	368.1 ± 191.3	396.6 ± 209.5
$t\bar{t}$+b	393.7 ± 188.2	203.1 ± 92.5	199.6 ± 90.8	170.8 ± 81.4
$t\bar{t}$+2b	165.2 ± 81.2	78.9 ± 38.0	87.2 ± 40.7	97.3 ± 46.8
$t\bar{t} + b\bar{b}$	226.4 ± 113.2	75.8 ± 35.3	114.1 ± 52.3	183.7 ± 86.7
Single Top	283.0 ± 49.0	115.3 ± 30.8	76.2 ± 19.5	47.5 ± 12.7
V+jets	130.5 ± 35.2	38.6 ± 17.8	22.8 ± 10.4	13.6 ± 6.4
$t\bar{t}$+V	43.5 ± 8.2	4.3 ± 1.2	6.4 ± 1.8	10.0 ± 2.7
Diboson	2.8 ± 1.3	2.1 ± 1.3	0.9 ± 0.5	0.2 ± 0.3
Total bkg	8326.7 ± 1788.6	2907.4 ± 836.5	1875.5 ± 534.7	1509.1 ± 423.7
$t\bar{t}H$	29.6 ± 2.1	7.4 ± 1.0	10.9 ± 1.2	16.7 ± 2.1
Data	7185	2793	1914	1386
S/B	0.0036	0.0026	0.0059	0.011
Data/B	0.9 ± 0.2	1.0 ± 0.3	1.0 ± 0.3	0.9 ± 0.3

Process	4 jets, $\geq$ 4 b-tags	5 jets, $\geq$ 4 b-tags	$\geq$ 6 jets, $\geq$ 4 b-tags	boosted
$t\bar{t}$+lf	17.8 ± 10.8	17.7 ± 10.9	17.6 ± 11.3	45.1 ± 9.4
$t\bar{t} + c\bar{c}$	11.6 ± 8.2	22.1 ± 15.4	35.9 ± 24.9	21.8 ± 12.0
$t\bar{t}$+b	8.4 ± 4.4	14.8 ± 7.7	20.0 ± 10.9	10.3 ± 5.5
$t\bar{t}$+2b	3.5 ± 1.9	6.9 ± 3.7	12.3 ± 6.9	12.3 ± 6.6
$t\bar{t} + b\bar{b}$	10.1 ± 4.9	28.8 ± 13.9	73.4 ± 36.6	17.0 ± 8.4
Single Top	2.5 ± 1.1	4.3 ± 1.4	5.5 ± 2.0	7.0 ± 1.7
V+jets	1.0 ± 0.8	0.9 ± 0.8	1.4 ± 0.7	2.5 ± 0.8
$t\bar{t}$+V	0.3 ± 0.1	0.7 ± 0.3	1.6 ± 0.6	0.9 ± 0.3
Diboson	0.0 ± 0.0	0.1 ± 0.1	0.0 ± 0.0	0.1 ± 0.1
Total bkg	55.2 ± 23.0	96.5 ± 37.6	167.6 ± 65.7	117.0 ± 24.9
$t\bar{t}H$	0.9 ± 0.2	2.7 ± 0.6	5.9 ± 1.4	2.2 ± 0.3
Data	75	104	150	104
S/B	0.017	0.028	0.035	0.019
Data/B	1.4 ± 0.5	1.1 ± 0.4	0.9 ± 0.4	0.9 ± 0.2

UNIVERSITY of VIRGINIA

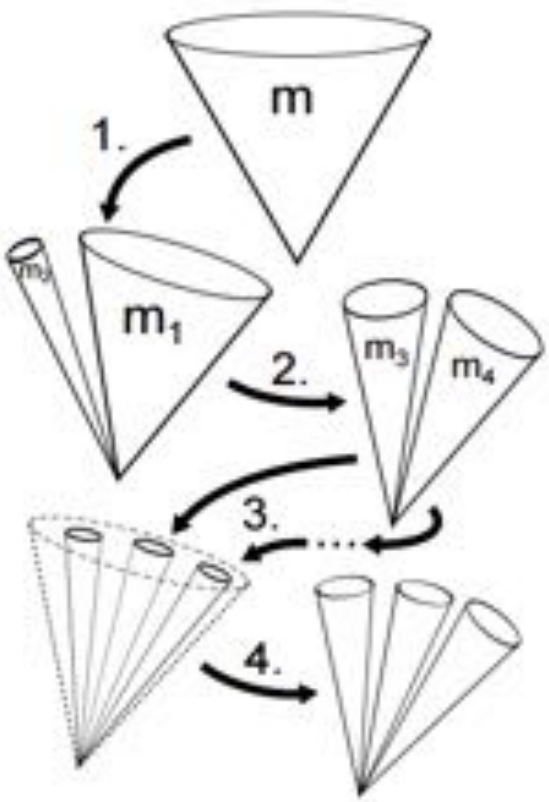
Event Yields: Dilepton

- Predicted and observed event yields in dilepton events
- Again shows good agreement across subdivided categories

	3 jets, 2 b-tags	3 jets, 3 b-tags	$\geq$ 4 jets, 2 b-tags	$\geq$ 4 jets, 3 b-tags	$\geq$ 4 jets, $\geq$ 4 b-tags
$t\bar{t}$+lf	2558.6 ± 542.7	26.6 ± 10.5	2271.6 ± 505.0	60.3 ± 25.6	0.9 ± 0.8
$t\bar{t} + c\bar{c}$	220.9 ± 103.4	22.7 ± 13.6	478.4 ± 234.4	78.4 ± 45.4	3.4 ± 2.9
$t\bar{t}$+b	65.4 ± 28.5	21.4 ± 10.2	126.2 ± 57.7	52.2 ± 25.1	2.7 ± 1.6
$t\bar{t}$+2b	16.9 ± 7.6	6.6 ± 3.1	42.9 ± 20.2	22.3 ± 10.7	1.2 ± 0.7
$t\bar{t} + b\bar{b}$	8.6 ± 4.2	3.6 ± 1.8	48.9 ± 23.7	39.8 ± 18.8	13.4 ± 7.1
Single Top	93.2 ± 16.7	3.0 ± 1.0	87.6 ± 15.8	7.3 ± 2.5	0.4 ± 0.4
V+jets	14.5 ± 11.0	1.3 ± 0.8	16.0 ± 7.4	0.0 ± 0.0	0.0 ± 0.0
$t\bar{t}$+V	3.6 ± 0.9	0.3 ± 0.2	16.4 ± 3.2	3.2 ± 0.9	0.5 ± 0.2
Diboson	1.7 ± 0.9	0.0 ± 0.0	1.2 ± 1.0	0.1 ± 0.0	0.0 ± 0.0
Total bkg	2983.4 ± 590.4	85.6 ± 25.6	3089.2 ± 650.6	263.6 ± 79.9	22.5 ± 9.8
$t\bar{t}H$	1.4 ± 0.2	0.4 ± 0.1	8.1 ± 1.1	3.6 ± 0.6	1.0 ± 0.3
Data	3123	115	2943	319	27
S/B	0.00047	0.0051	0.0026	0.014	0.046
Data/B	1.0 ± 0.2	1.3 ± 0.4	1.0 ± 0.2	1.2 ± 0.3	1.2 ± 0.5

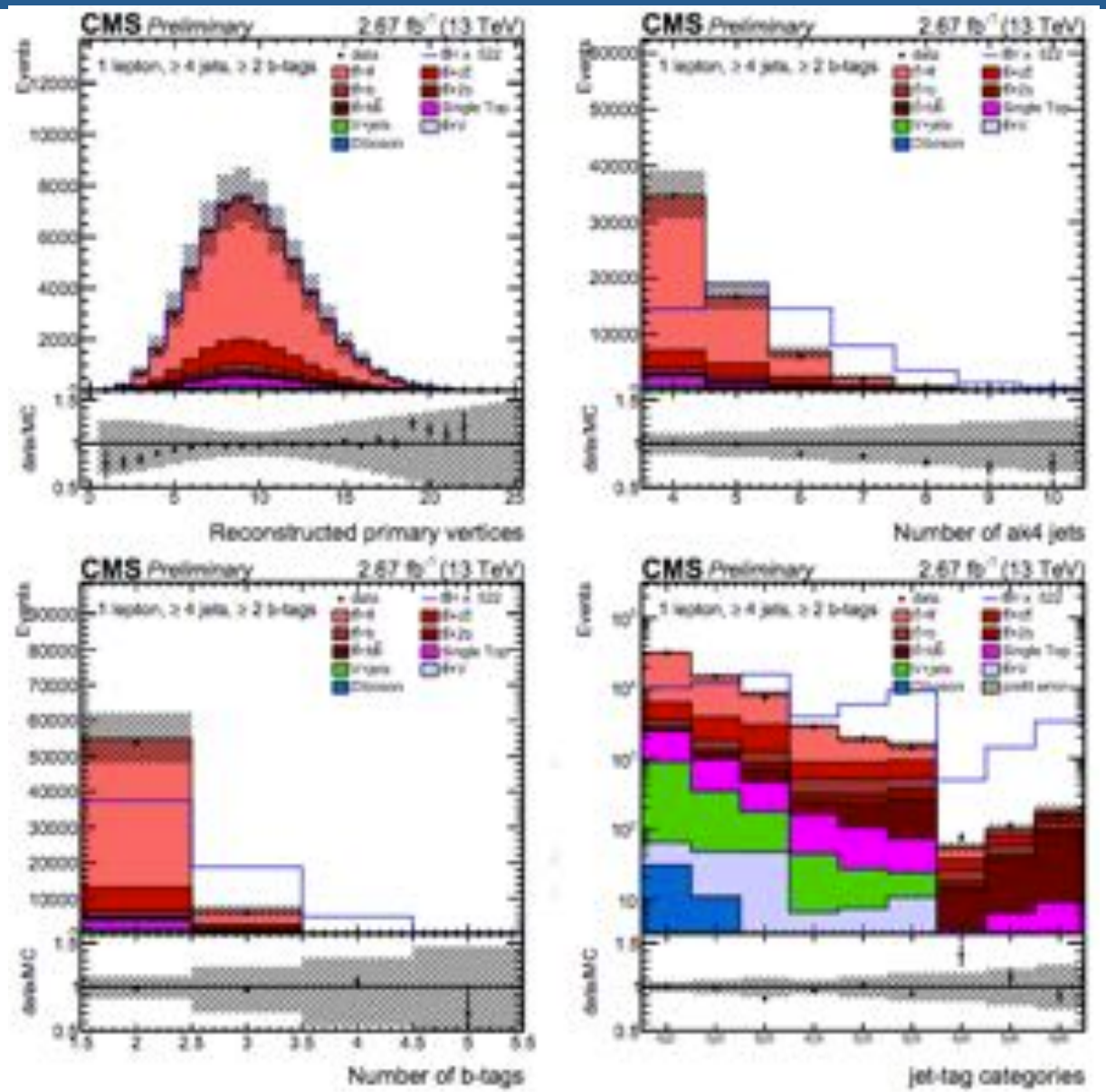
Boosted Analysis: Hadronic Top and Higgs

- Clustering of all decay products to Fat Jet ($\Delta R = 1.5$)
- Looks for large decrease in invariant mass
- Jet Grooming
- Returns 3 subjets assigned to hadronic top decay



UNIVERSITY of VIRGINIA

Lepton + Jets Control Plots



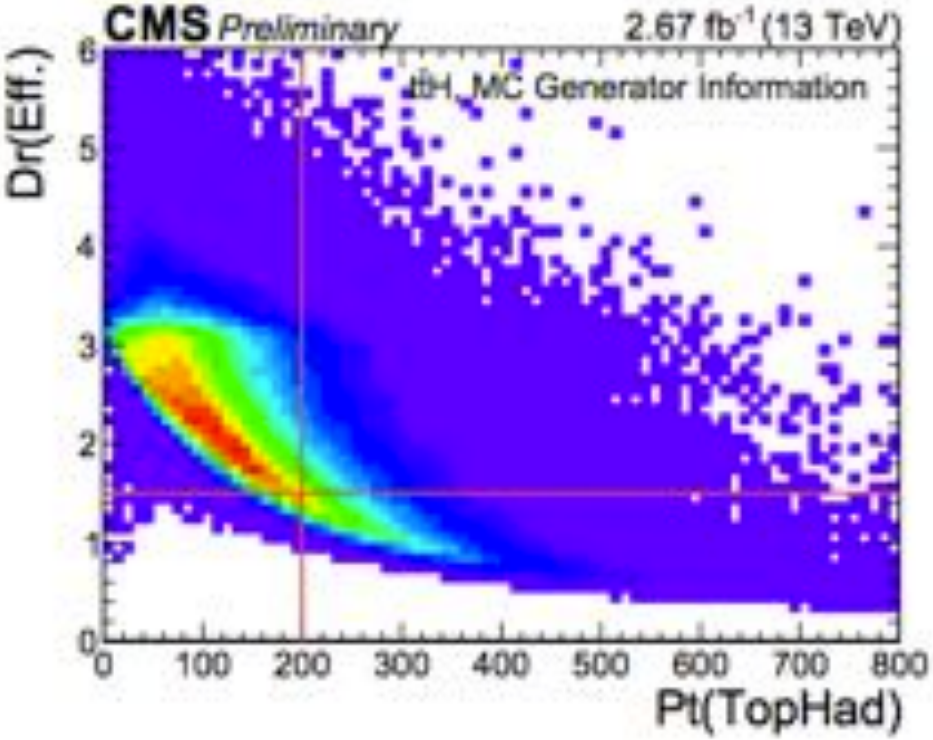
UNIVERSITY of VIRGINIA

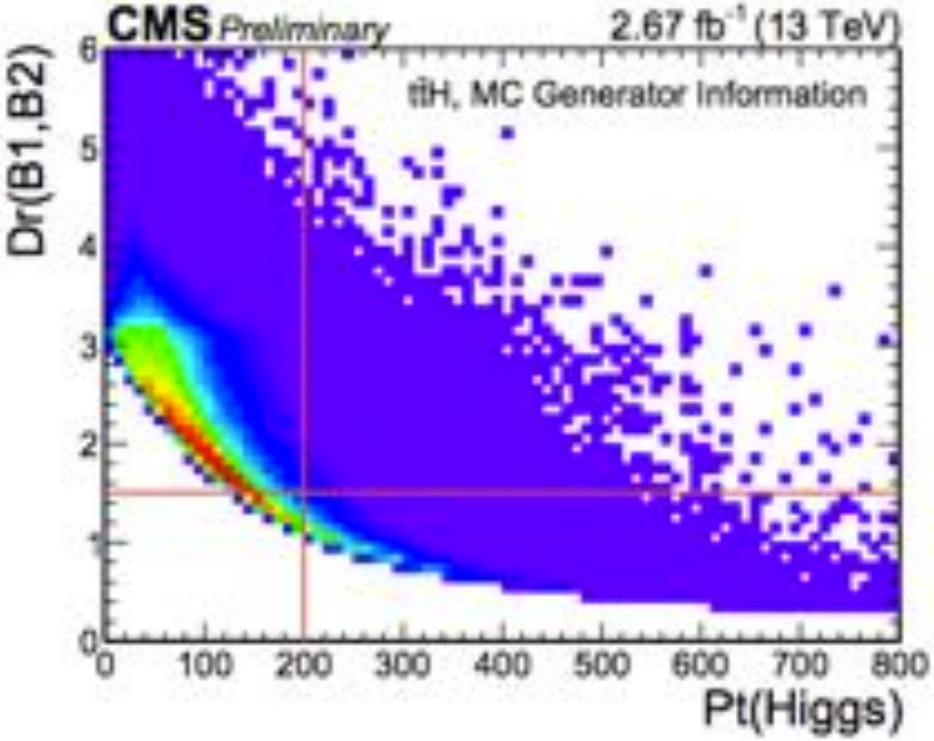
Systematic Uncertainties

Source	Type	Notes
Luminosity	lnN	Signal and all backgrounds
Lepton ID/trigger efficiency	shape	Signal and all backgrounds
Pileup	shape	Signal and all backgrounds
Jet Energy Scale	shape	Signal and all backgrounds
b-Tag HF fraction	shape	Signal and all backgrounds
b-Tag HF stats (linear)	shape	Signal and all backgrounds
b-Tag HF stats (quadratic)	shape	Signal and all backgrounds
b-Tag LF fraction	shape	Signal and all backgrounds
b-Tag LF stats (linear)	shape	Signal and all backgrounds
b-Tag LF stats (quadratic)	shape	Signal and all backgrounds
b-Tag Charm (linear)	shape	Signal and all backgrounds
b-Tag Charm (quadratic)	shape	Signal and all backgrounds
QCD Scale ($t\bar{t}H$)	lnN	Scale uncertainty for NLO $t\bar{t}H$ prediction
QCD Scale ($t\bar{t}$)	lnN	Scale uncertainty for NLO $t\bar{t}$ prediction
QCD Scale ($t\bar{t}$+HF)	lnN	Additional scale uncertainty for NLO $t\bar{t}$+HF predictions
QCD Scale (t)	lnN	Scale uncertainty for NLO single top prediction
QCD Scale (V)	lnN	Scale uncertainty for NNLO W and Z prediction
QCD Scale (VV)	lnN	Scale uncertainty for NLO diboson prediction
pdf (gg)	lnN	Pdf uncertainty for gg initiated processes except $t\bar{t}H$ ($t\bar{t}$, $t\bar{t}Z$)
pdf ($gg t\bar{t}H$)	lnN	Pdf uncertainty for $t\bar{t}H$
pdf ($q\bar{q}$)	lnN	Pdf uncertainty for $q\bar{q}$ initiated processes ($t\bar{t}W$, W, Z).
pdf (qg)	lnN	Pdf uncertainty for qg initiated processes (single top)
Q2 Scale ($t\bar{t}$)	shape	Renormalization and factorization scale uncertainties of the $t\bar{t}$ ME generator, independent for additional jet flavors
PS Scale ($t\bar{t}$)	shape	Renormalization and factorization scale uncertainties of the parton shower (for $t\bar{t}$ events), independent for additional jet flavors

UNIVERSITY of VIRGINIA

Boosted Analysis

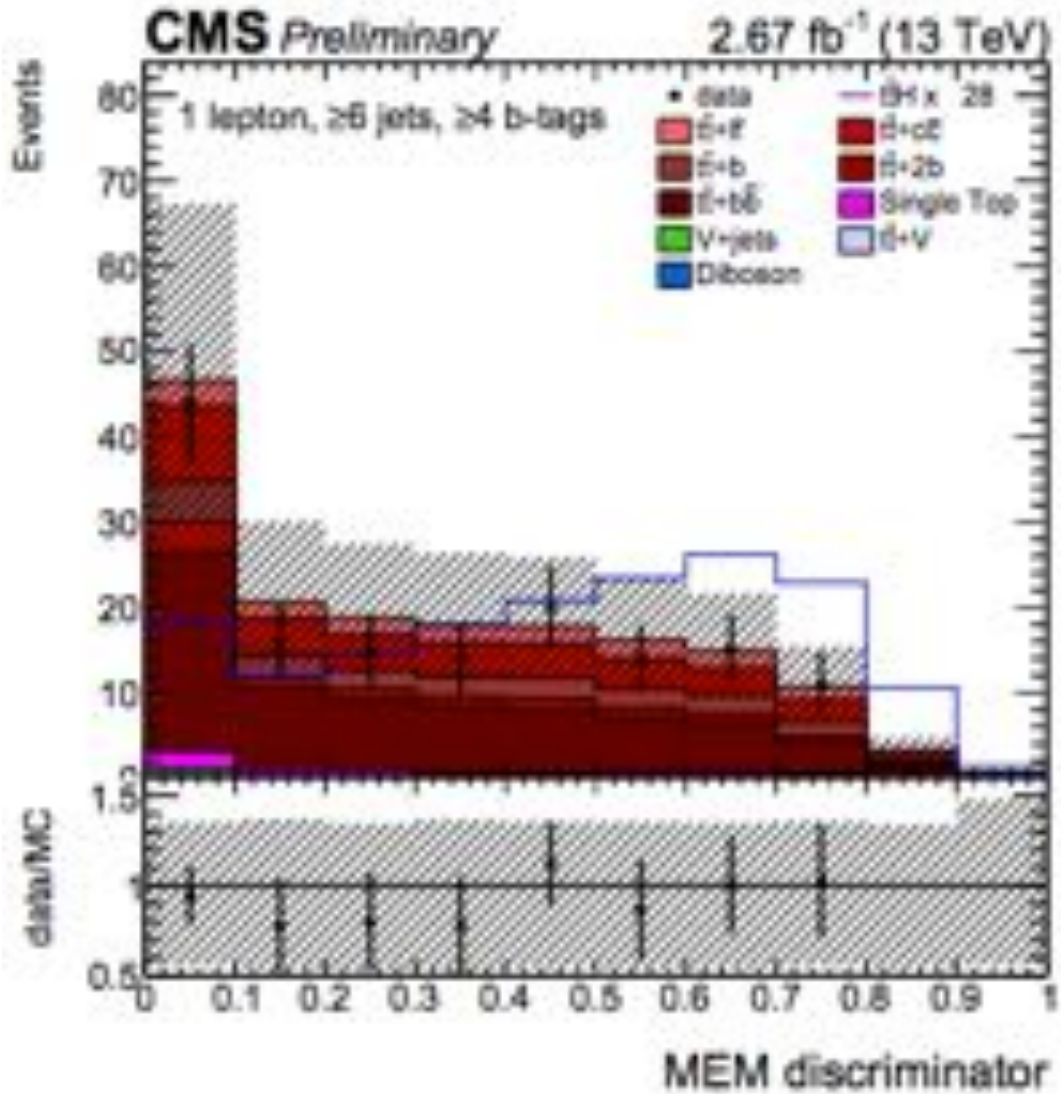




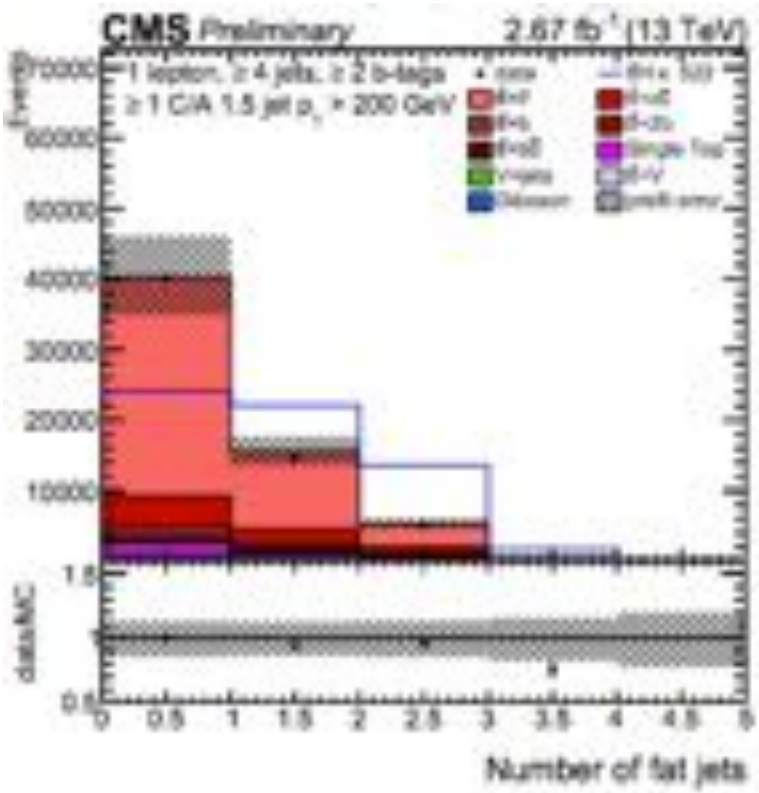
- Distance of decay products from (left)top going hadronic and (right) Higgs decaying to two b-jets

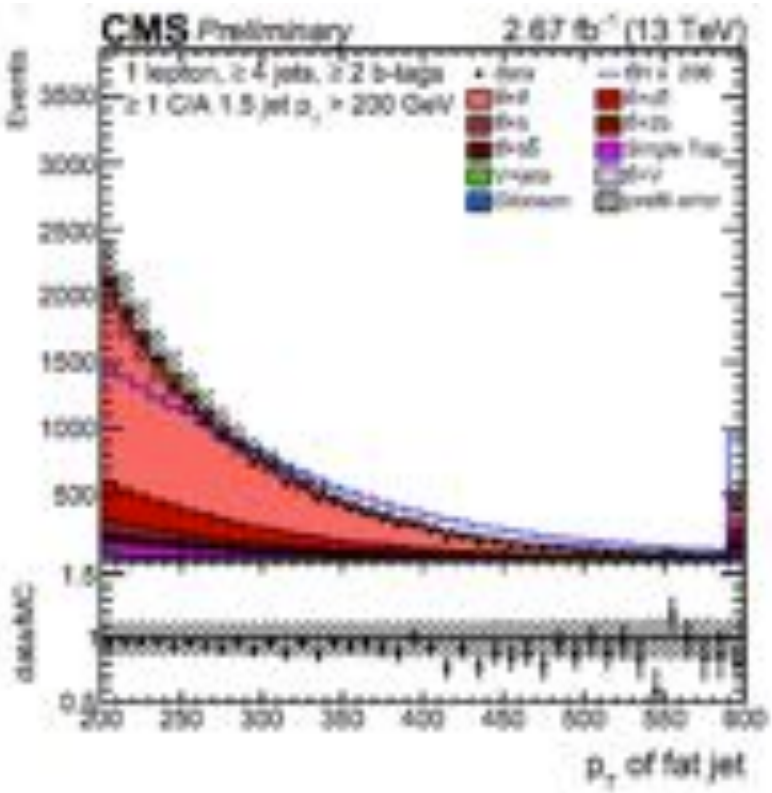
MEM Discriminant

- Lepton + Jets shown right
  - 6 jets, 4 b-tags
- Shown to be good variable choice to separate ttH signal from background processes
- Used as input for BDT in several semileptonic subcategories



Boosted Analysis

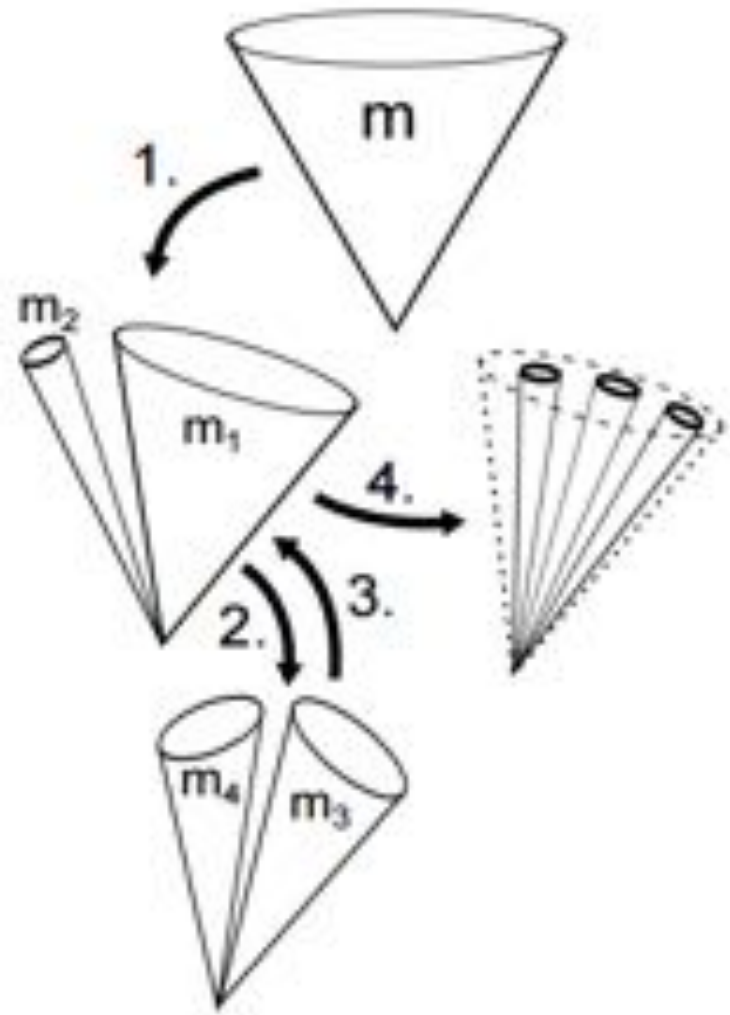




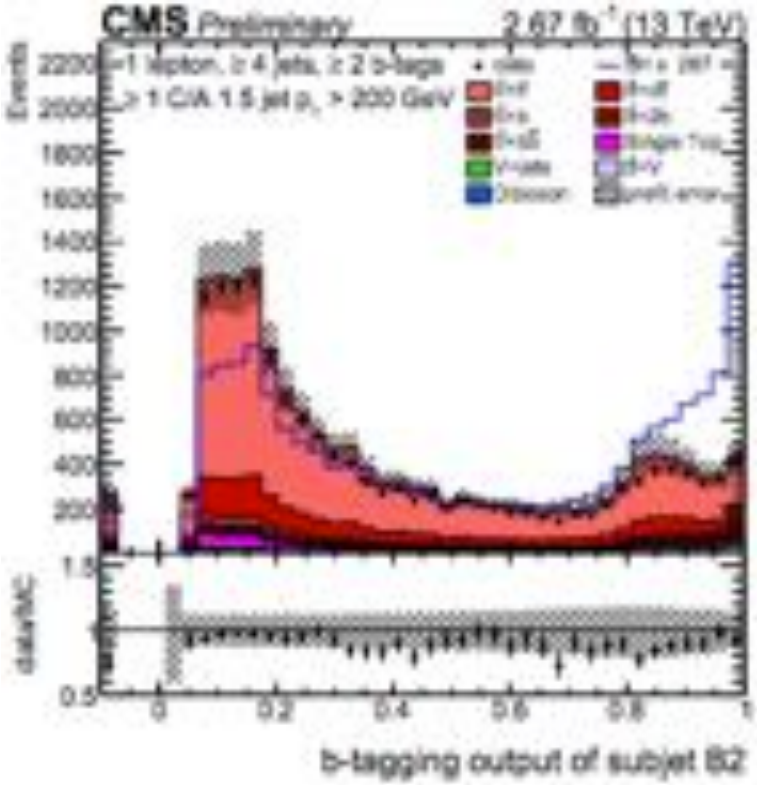
- (Left)Number of selected fat jets and pt of leading fat jet

Boosted Analysis: Higgs

- Declustering fat jets
- Find subjets from invariant mass
- Jet Grooming
- Returns number of jets



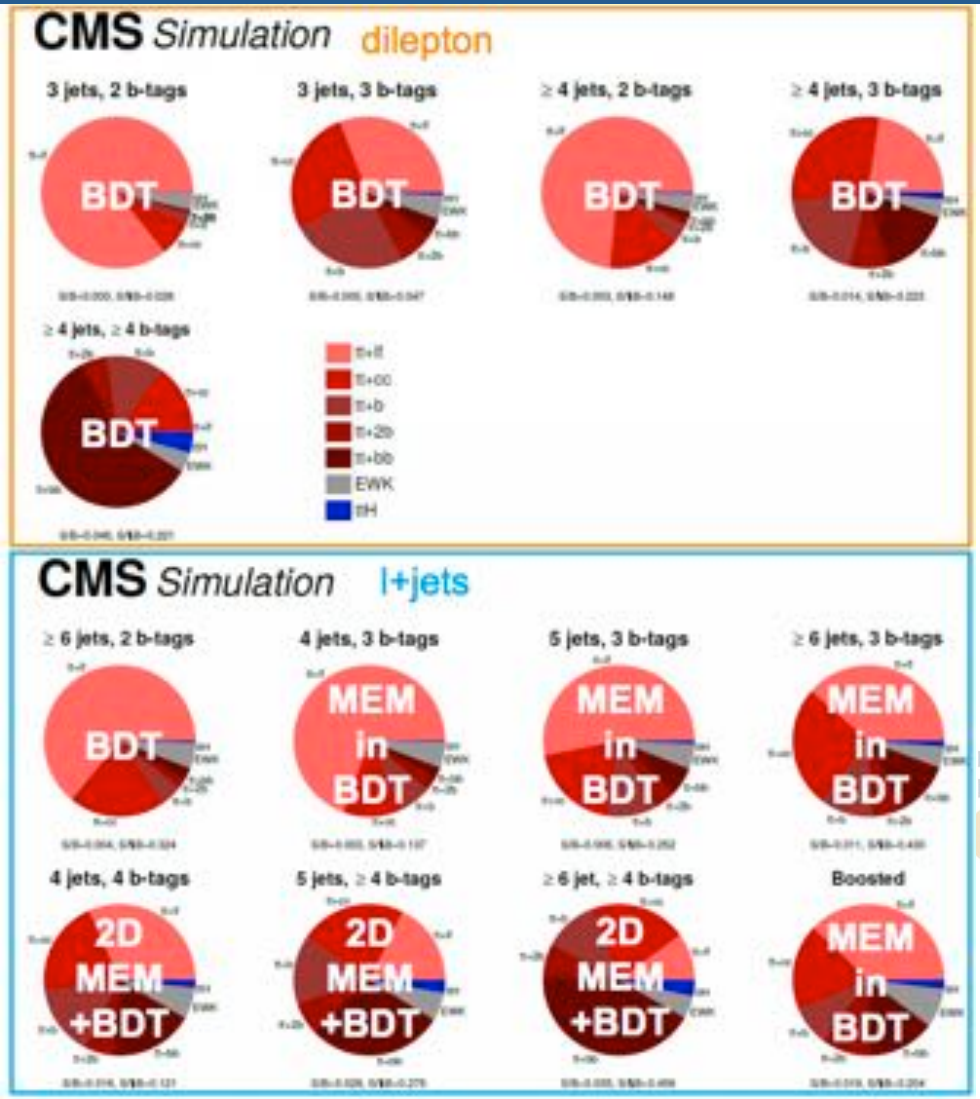
Boosted Analysis: Higgs Discriminator



- CSV(B2) shows highest discriminator value in identifying true boosted Higgs over fake candidates.

BDT and MEM: Final Discriminants

- In lepton + jets (semileptonic) events the BDT and MEM discriminants are combined
- Boosted Category events (both top and Higgs)
  - BDT used as final discriminant
  - Sensitivity increases from reduced combinatorics in reconstructed objects from both top and Higgs



UNIVERSITY of VIRGINIA

References

- "Search for ttH production in the H → bb decay channel with s = 13TeV pp collisions at the CMS experiment" - CMS PAS HIG-16-004
- "First results on ttH with H→bb at 13 TeV" – CMS AN-15-216
- "Search for the associated ttH(→bb) production using the Matrix Element Method" – CMS AN-13-313
- "Search for ttH production in boosted topologies" – GRK Plenary Workshop 2015, Shawn Williamson
- "Results and Prospects for ttH at CMS" – 51st Rencontres de Moriond EW 2016, Johannes Hauk

UNIVERSITY of VIRGINIA