LASTEST TTH(BB) RESULTS @ CMS

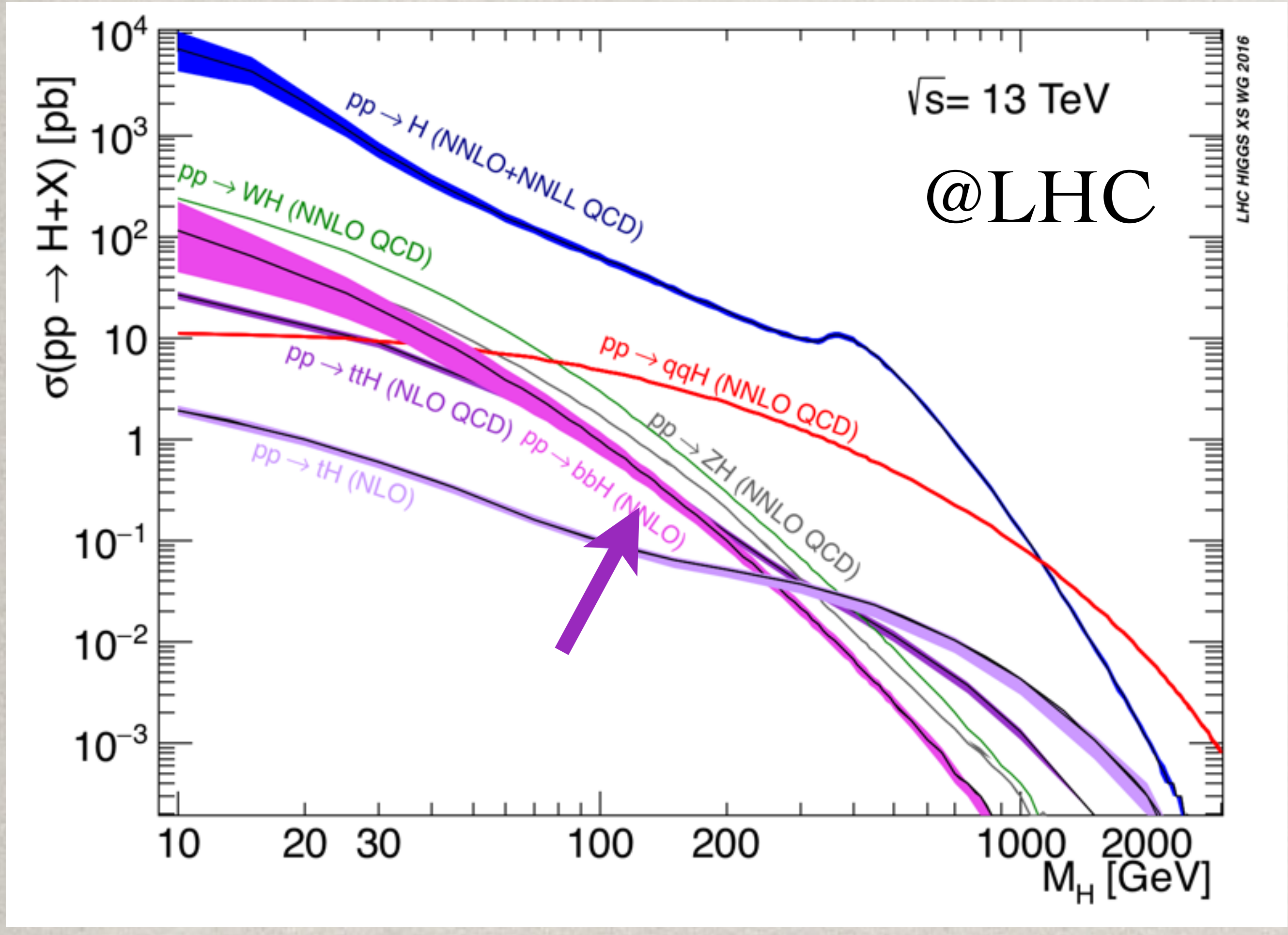
Wuming Luo

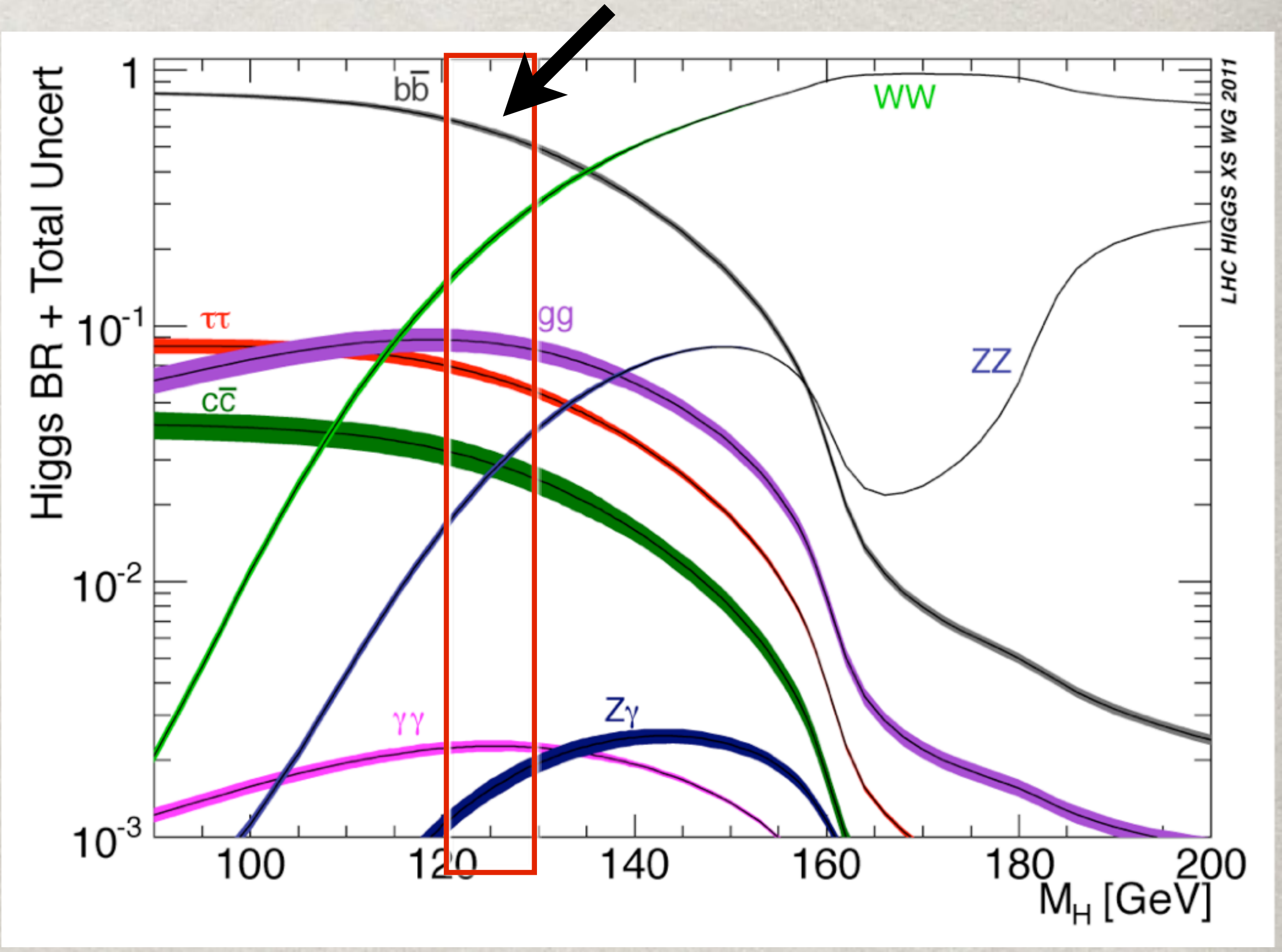
THE OHIO STATE UNIVERSITY

CMS Compact Muon Solenoid

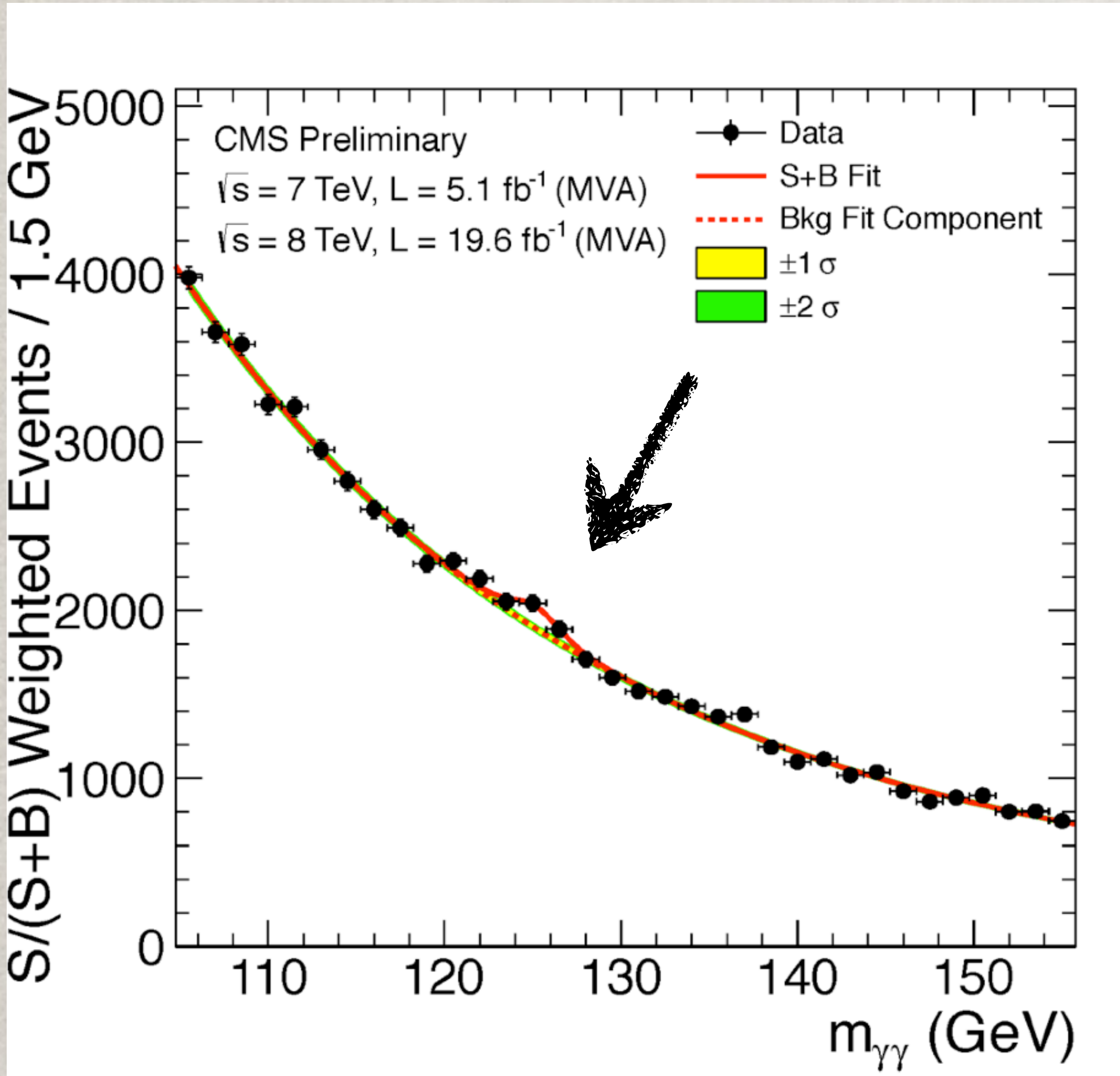
THE 3 "BIG" QUESTIONS

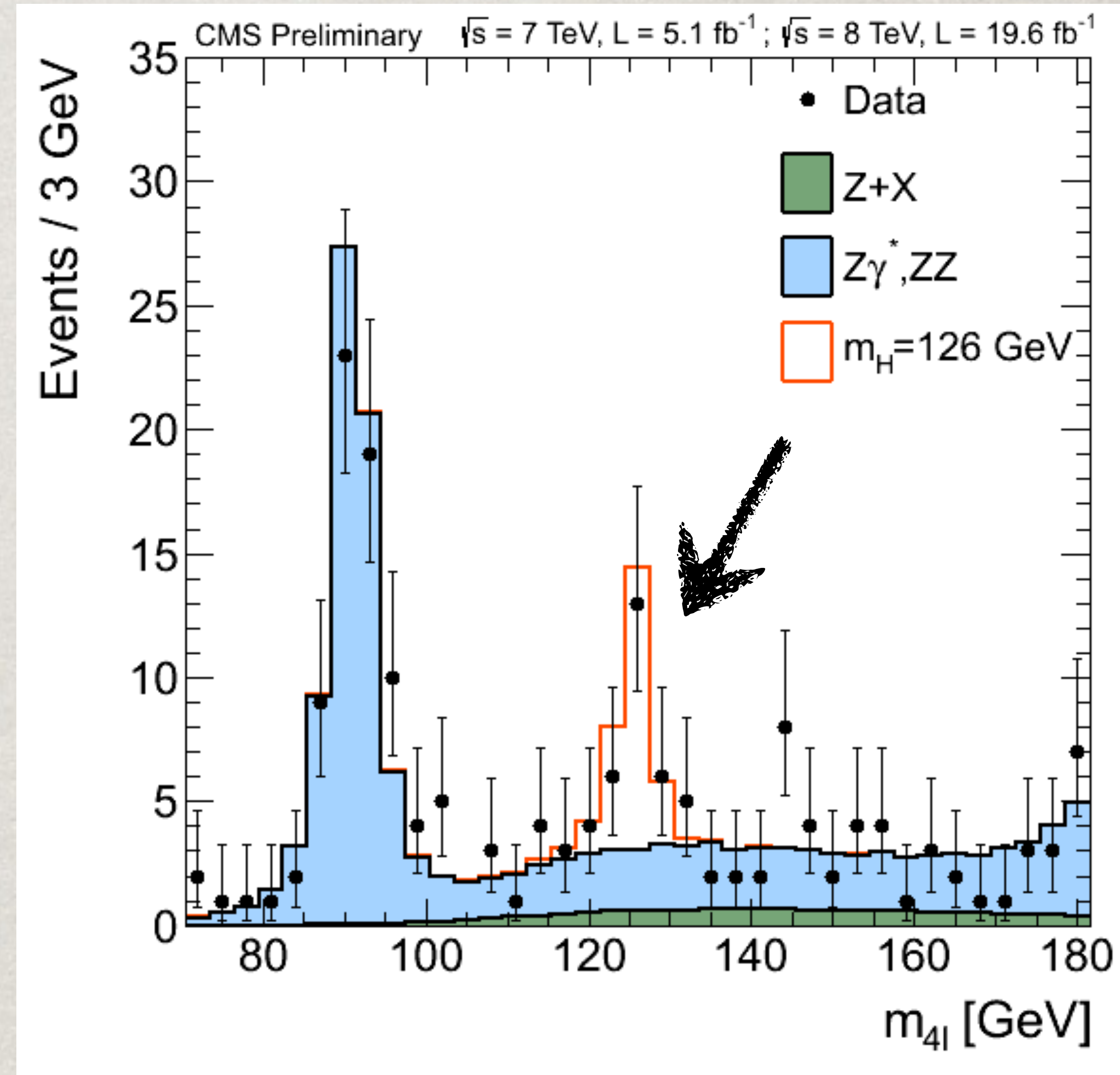
- What is it? — Higgs Boson
- Where does it come from? — Production
- Where does it go? — Decay



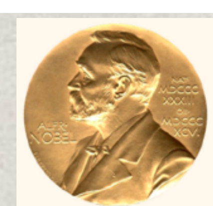


HIGGS BOSON DISCOVERY





- We found "A" Higgs boson!



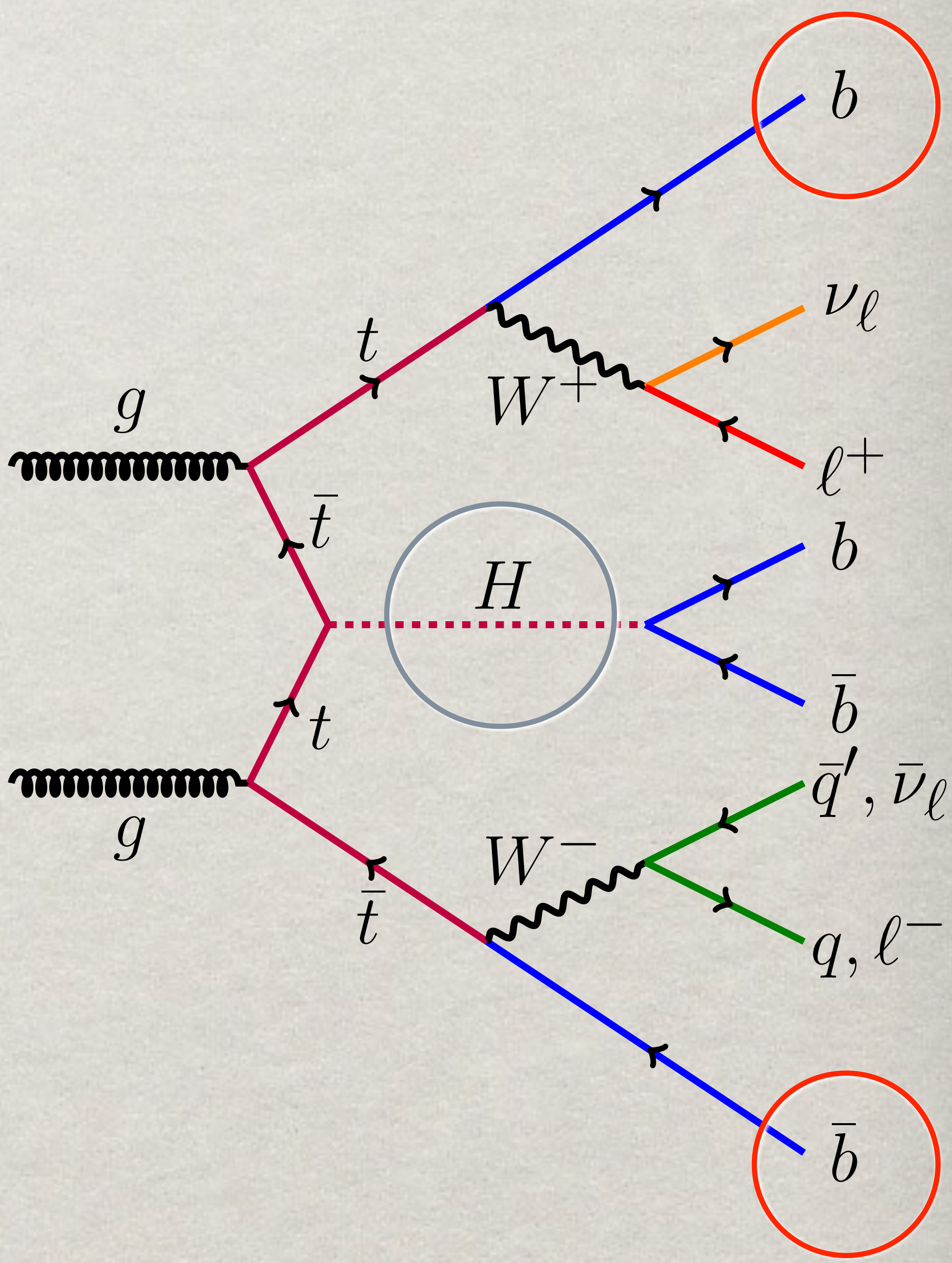
- Many of its measured characteristics seem to indicate it's "THE" Higgs boson, but there is more to measure

WHY TTH?

- Hasn't been studied thoroughly before
- Complementary to other H searches
- Directly probes the Yukawa Coupling between Higgs boson and top quarks.
  - Key component to evaluate the consistency of the new boson with SM expectations.
- It could be sensitive to Beyond SM physics.

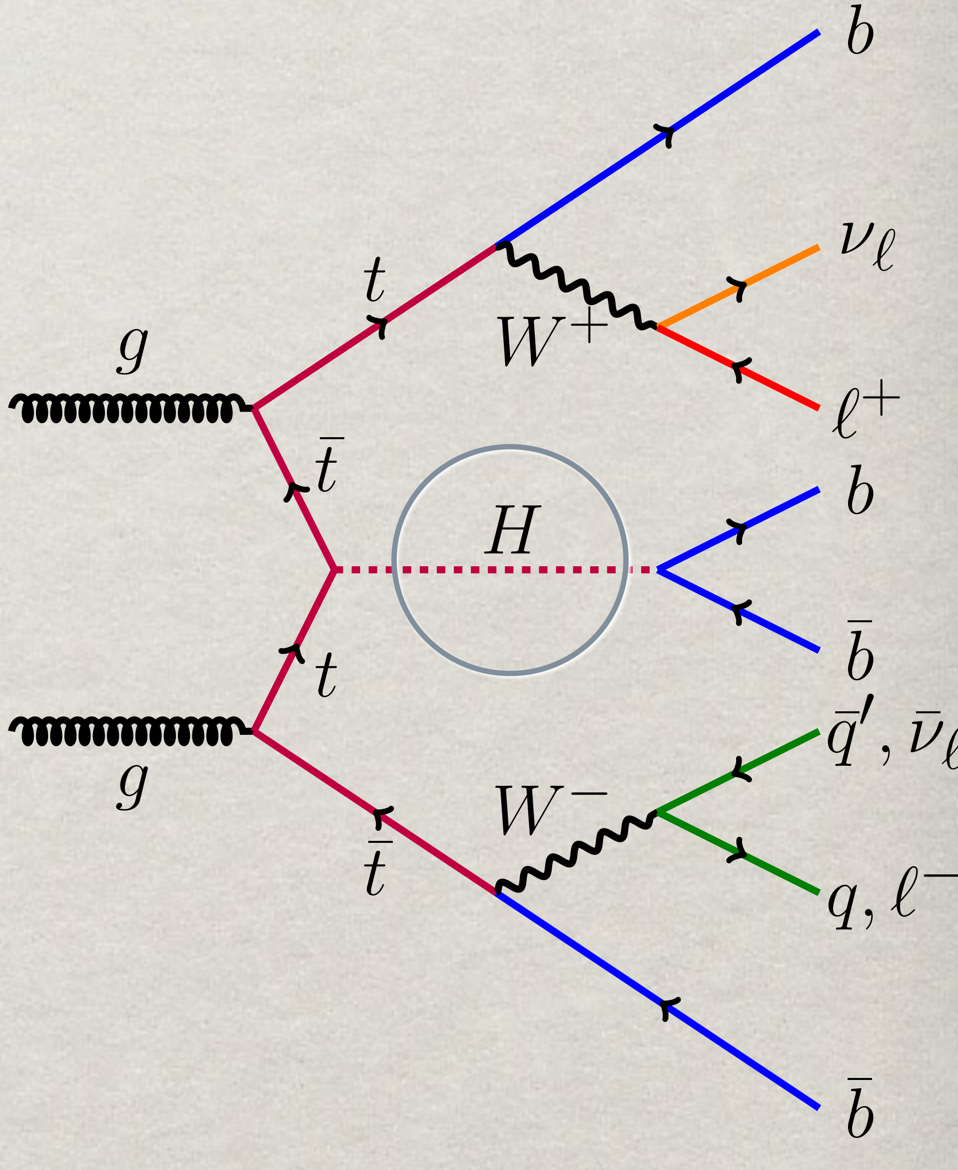
ttH(bb)

- Advantage:
  - Highest branching fraction
- Challenges:
  - 1. Tiny production cross section
  - 2. Higgs invariant mass hard to reconstruct
  - 3. Difficult irreducible background ttbb



OVERVIEW

- Split channels by top pair decay
  - Lepton + jets (e, μ): LJ channel
  - Dilepton: DIL channel
- For each channel, separate events into categories by number of b-tags and number of jets
- Use BDTs/MEM to separate S/B and fit simultaneously all categories to data to extract signal



DATA AND MC SAMPLES

- Data: total integrated luminosity of L = 12.9/fb
  - Recorded by CMS in early 2016 at $\sqrt{s}$ = 13 TeV
  - Collected using single-lepton or double-lepton triggers
- Signal and background MC: leptons + jets + neutrinos
  - Signal ttH: Powheg+Pythia8, all Higgs/top decays allowed
  - Main background ttJets: Powheg+Pythias8, separated by extra jet content: tt+lf/bb/b/B/cc
  - Other relevant bkg MC: tt+Z/W, WJets, ZJets, WW, WZ, ZZ, single top

CHALLENGE 1

TINY PRODUCTION CROSS SECTION

EVENT SELECTION

* Details in backup

- Leptons*:
  - Tight and Loose: different cuts on P_T, $|\eta|$ or Isolation
  - LJ: exactly 1 tight lepton, no additional loose leptons
  - DIL: exactly 2 leptons(≥1 tight), opposite sign
- Jets*:
  - Require jets not overlap with leptons spatially
  - Use CSV(Combined Secondary Vertex) algorithm to identify jets coming from b-quarks(b-tags)
  - ≥4jets + ≥2b-tags for LJ and ≥2jets + 1b-tag for DIL
- Corrections to MC: Pile up, lepton scale factors, jet energy scale/resolution, b-tagging reweighting

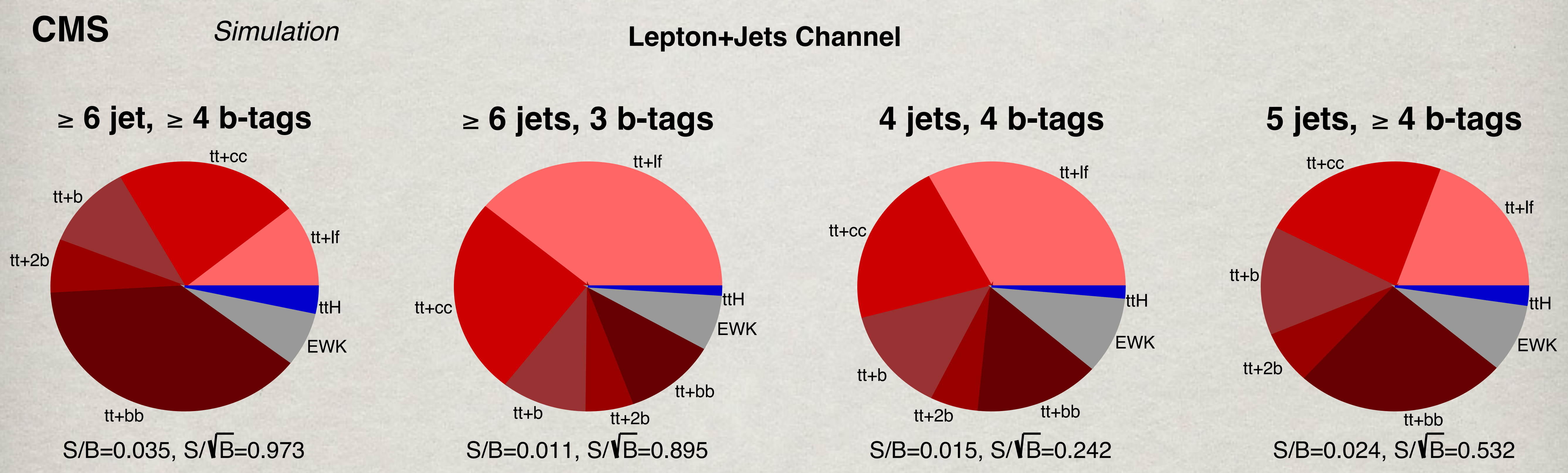
EVENT CATEGORIZATION

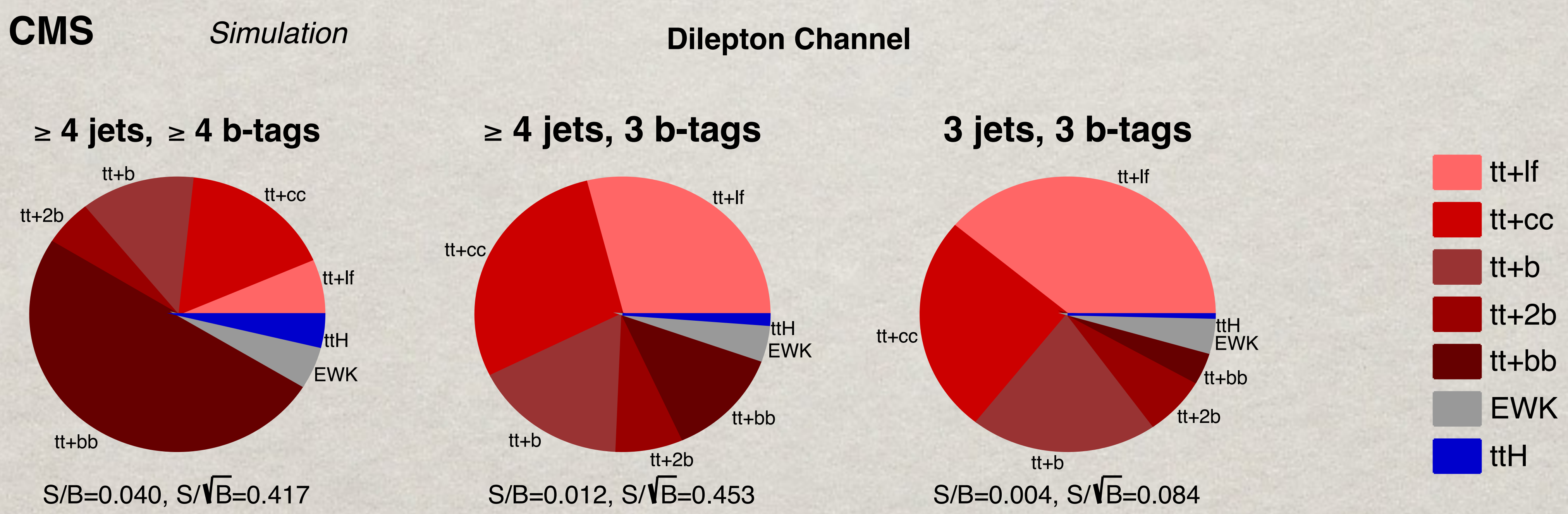
- Different channels based on top pair decay
  - LJ channel and DIL channel
- For each channel, categorize events based on number of jets and number of b-tags

Lepton + Jets(LJ)			
	4jets	5jets	≥6jets
2tags	x	x	x
3tags	x	x	✓
≥4tags	✓	✓	✓

Dilepton(DIL)		
	3jets	≥4jets
2tags	x	x
3tags	x	✓
≥4tags	✓	✓

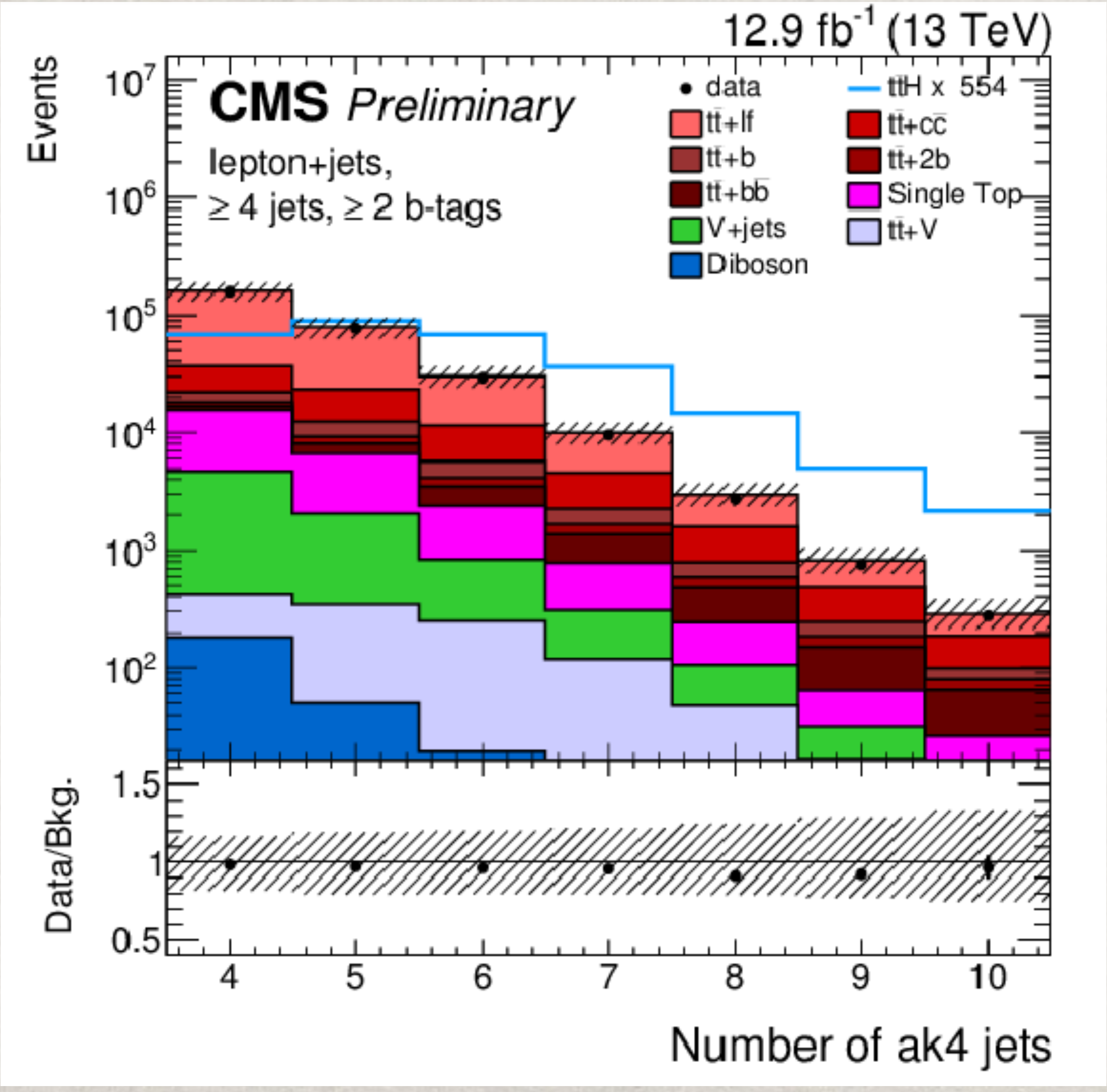
CATEGORIZATION

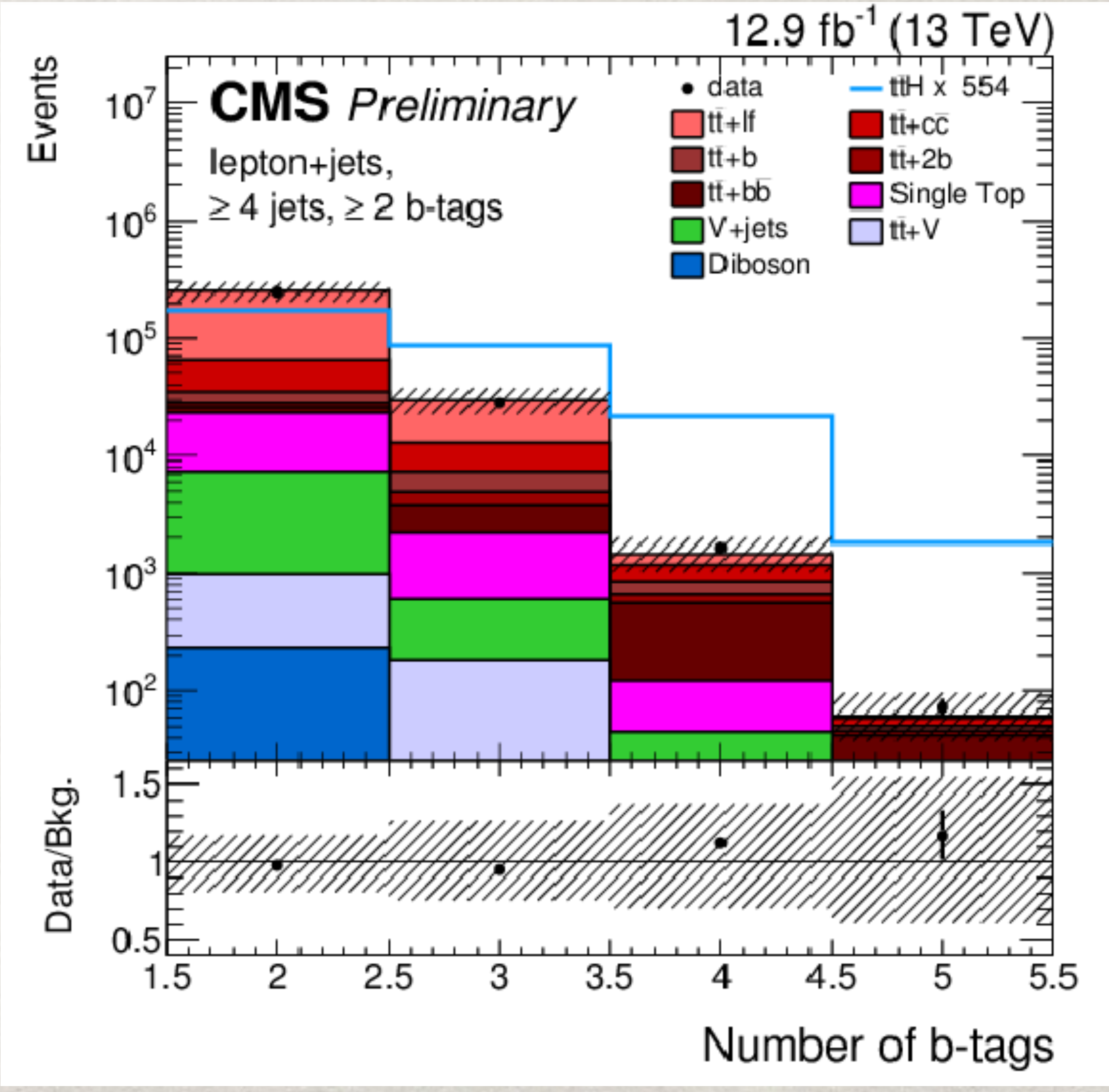




DATA/MC AGREEMENT

- All corrections to MC applied
- Good agreement between Data and MC



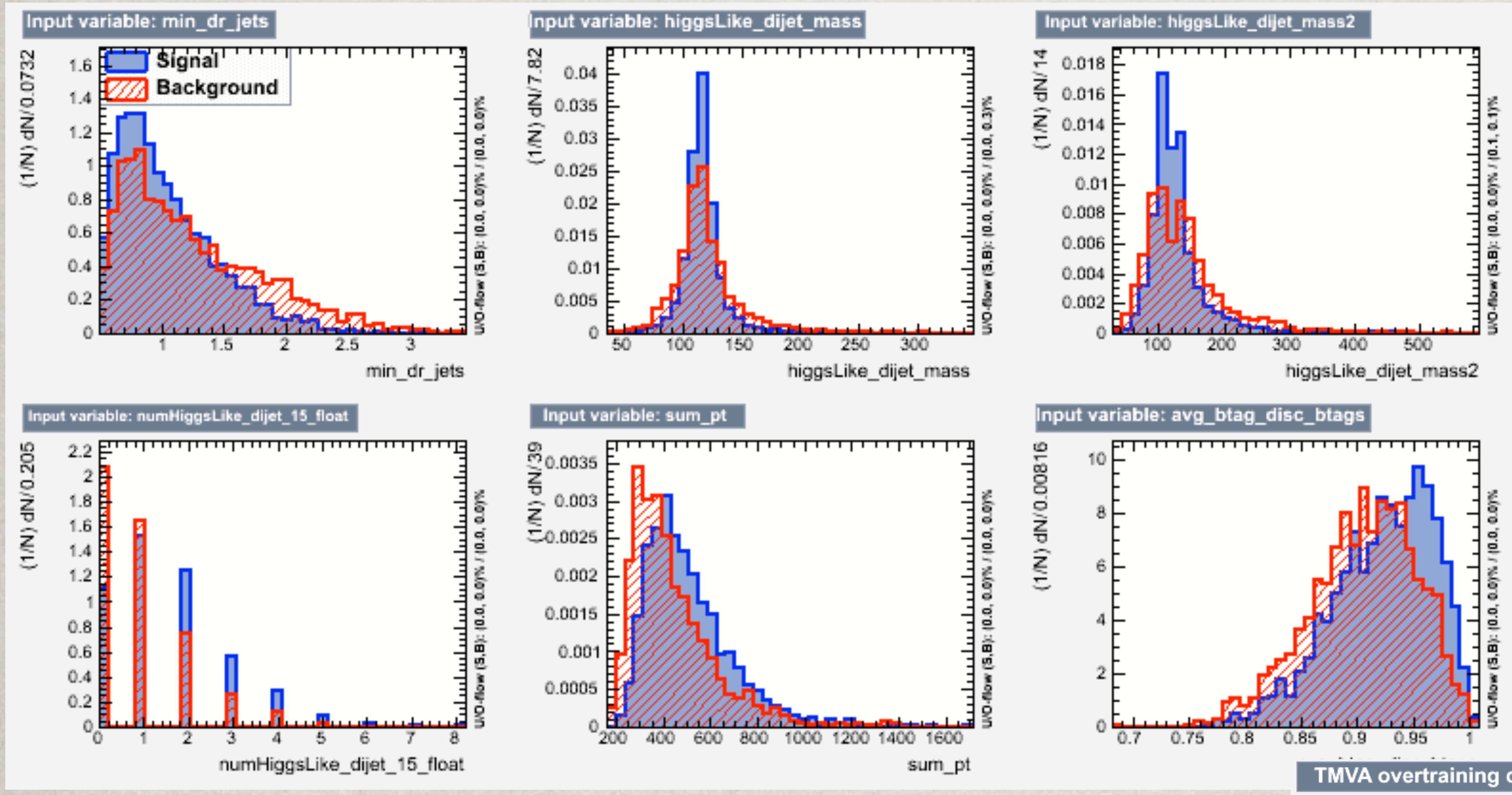


challenge 2
Higgs invariant mass hard to reconstruct

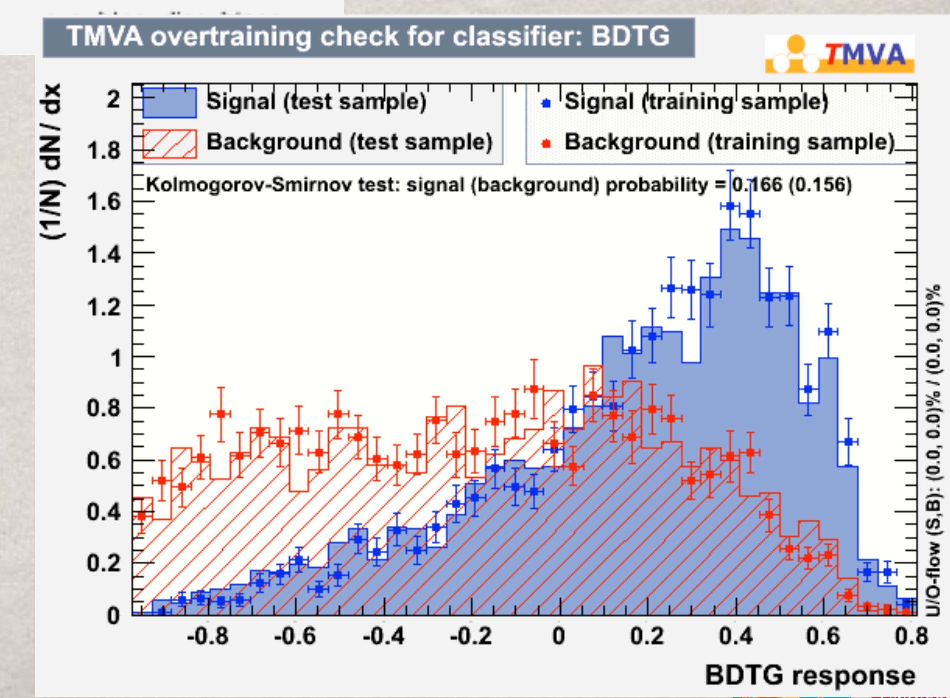
MULTIVARIATE ANALYSIS

- Can't use Higgs invariant mass as discriminant like other analyses ($H \rightarrow \gamma\gamma$ or $H \rightarrow ZZ \rightarrow 4l$).
- Background very similar to Signal
- Multi-Variate Analysis:
  - Combine several variables' discriminating power
  - Use Boosted Decision Tree (BDT)
- Train separate BDT for each jet/tag category
- Fit BDT output discriminators from all categories simultaneously to extract signal

BDT EXAMPLE



input variables



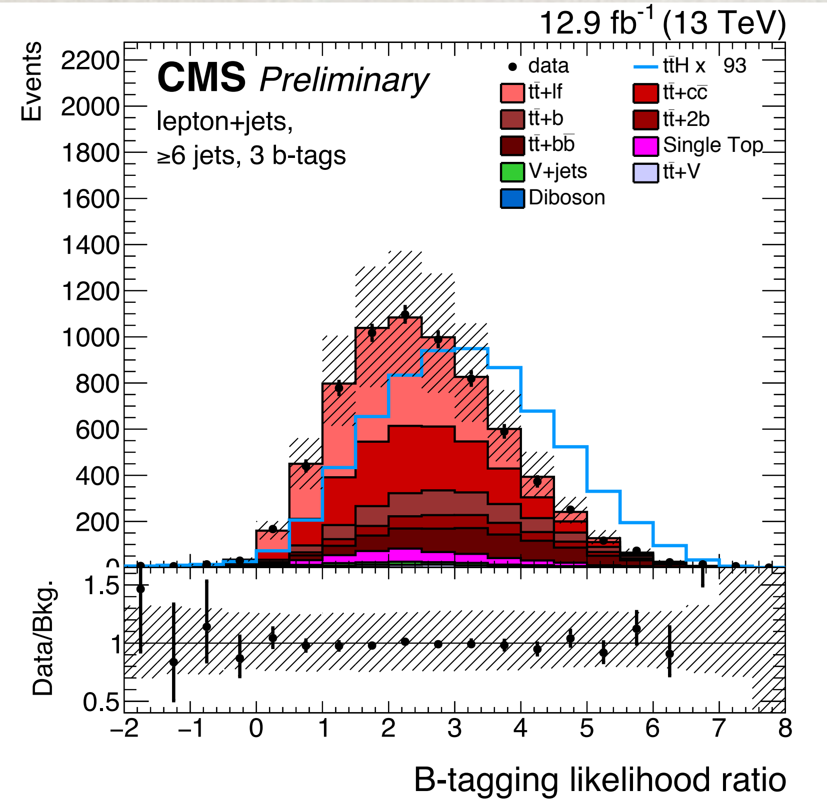
output discriminator

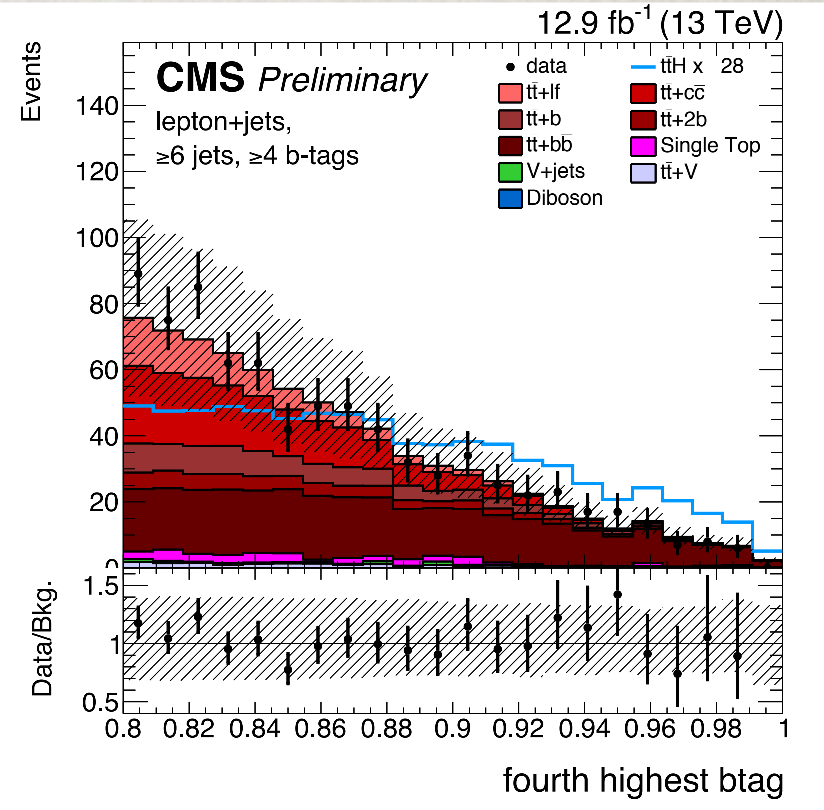
INPUT VARIABLE CHOICE

- Begin from a selection of well-modeled variables of several types:
  - CSV tagging, invariant mass, angular correlations, event shapes, and jet Pt variables
- Rank these variables based on 1D separation between S and B and pick the top ones
- Train MVA with appropriate number of variables
  - removing some would worsen limit, adding won't help much and causes overtraining

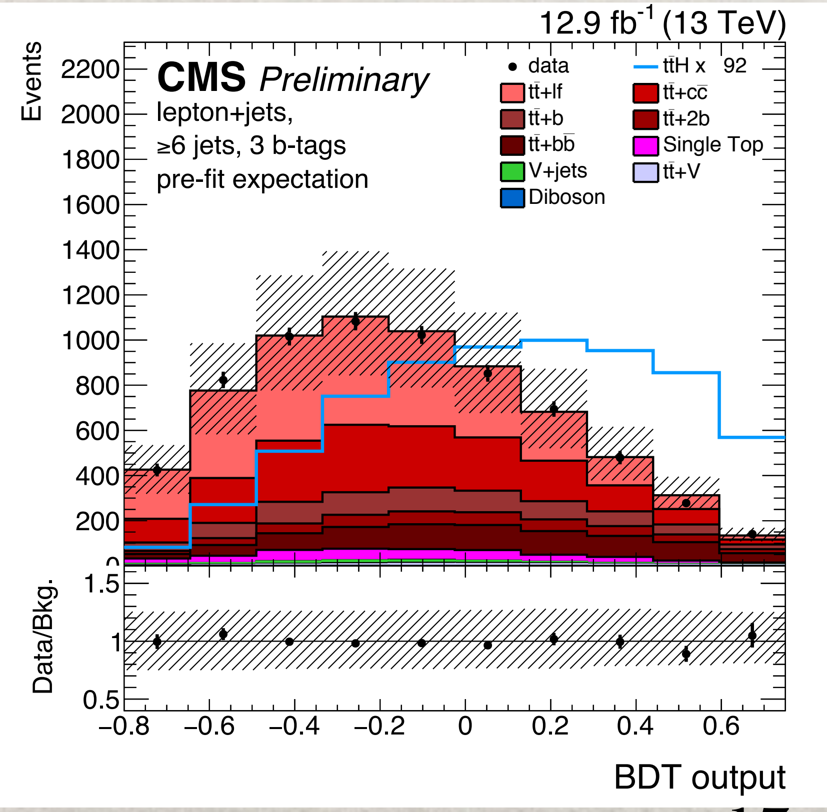
BDT INPUT/OUTPUT

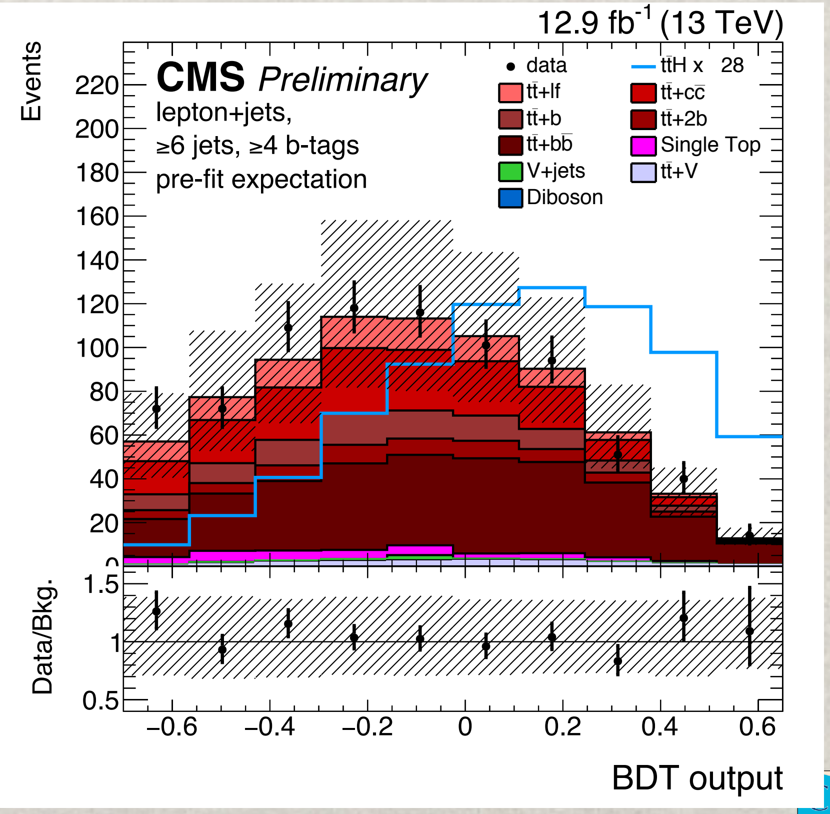
Input variables





Output variable





Challenge 3
Difficult irreducible background ttbb

MATRIX ELEMENT METHOD

- Irreducible background ttbb
- Use Matrix Element Method(MEM) to further distinguish ttbb from ttH

Numerical integration

Momentum conservation

Resolution function (allow ISR)

$$w(\vec{y}|\mathcal{H}) = \sum_{i=1}^{N_a} \int \frac{\mathrm{d}x_a \mathrm{d}x_b}{2x_a x_b s} \int \prod_{k=1}^{8} \left(\frac{\mathrm{d}^3 \vec{p}_k}{(2\pi)^3 2E_k} \right) (2\pi)^4 \, \delta^{(E,z)} \left(p_a + p_b - \sum_{k=1}^{8} p_k \right) \mathcal{R}^{(x,y)} \left(\vec{\rho}_{\mathrm{T}}, \sum_{k=1}^{8} p_k \right) \times g(x_a, \mu_{\mathrm{F}}) g(x_b, \mu_{\mathrm{F}}) |\mathcal{M}_{\mathcal{H}} (p_a, p_b, p_1, \ldots, p_8)|^2 W(\vec{y}, \vec{p}) \, ,$$

Parton density functions

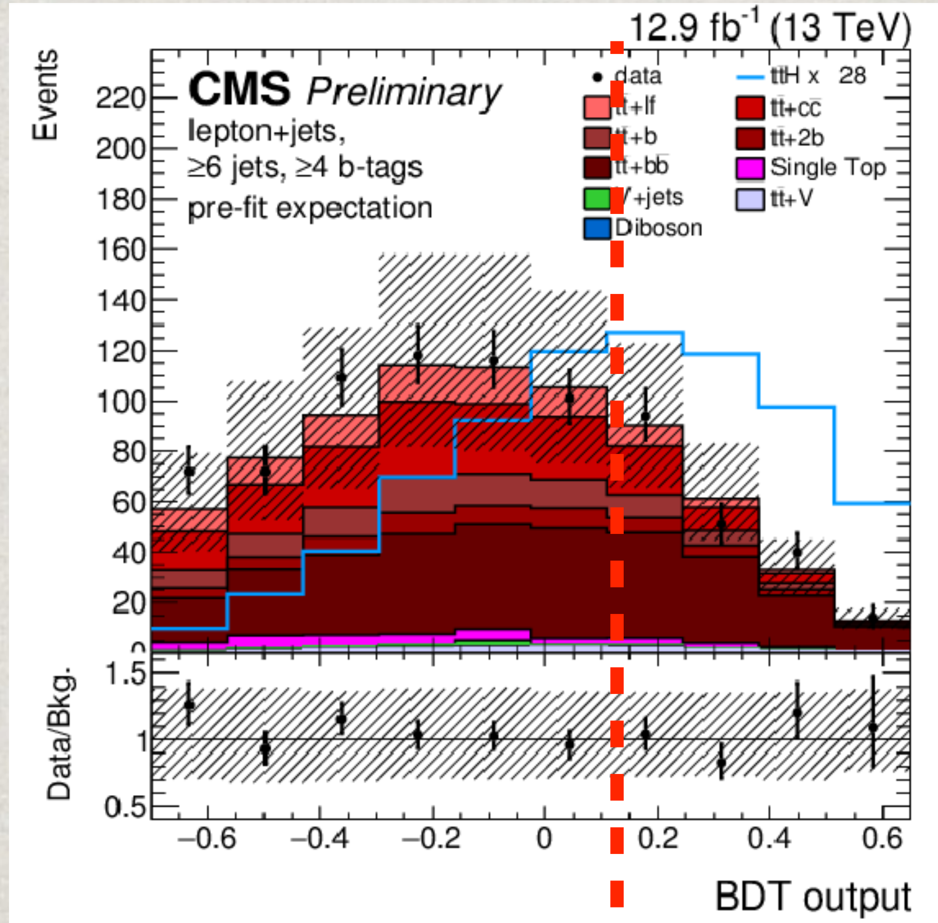
LO Scattering amplitude (Open Loops)

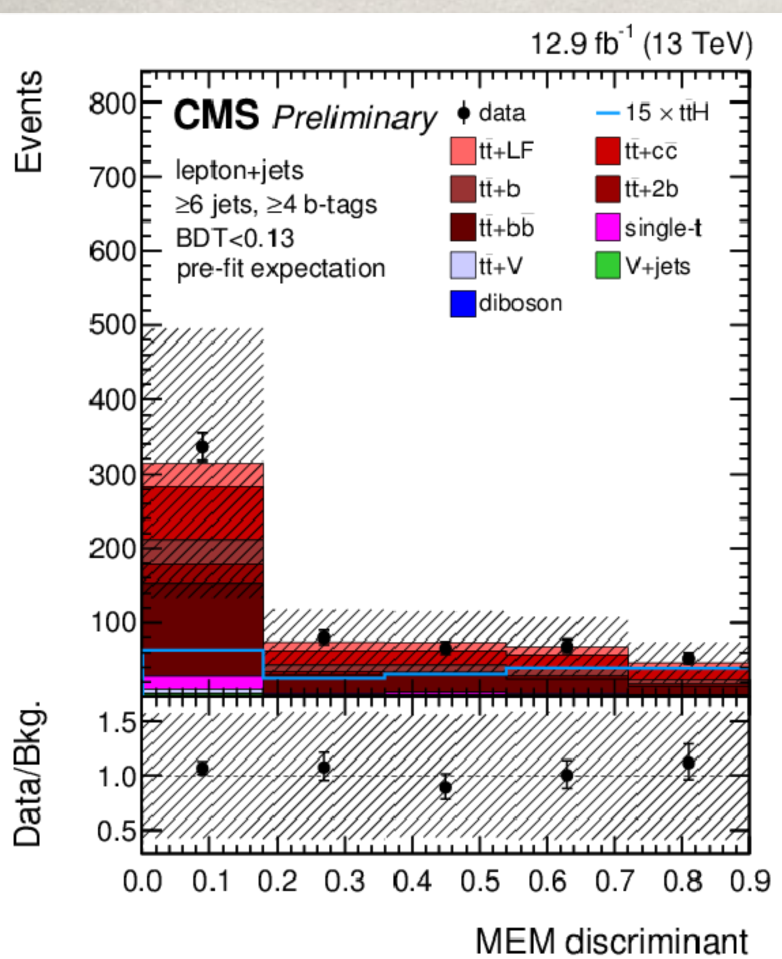
Detector transfer function

$$P_{\mathrm{s/b}} = \frac{w(\vec{y}|\mathrm{t\bar{t}H})}{w(\vec{y}|\mathrm{t\bar{t}H}) + k_{\mathrm{s/b}} w(\vec{y}|\mathrm{t\bar{t}{+}b\bar{b}})}$$

BDT/MEM 2D APPROACH

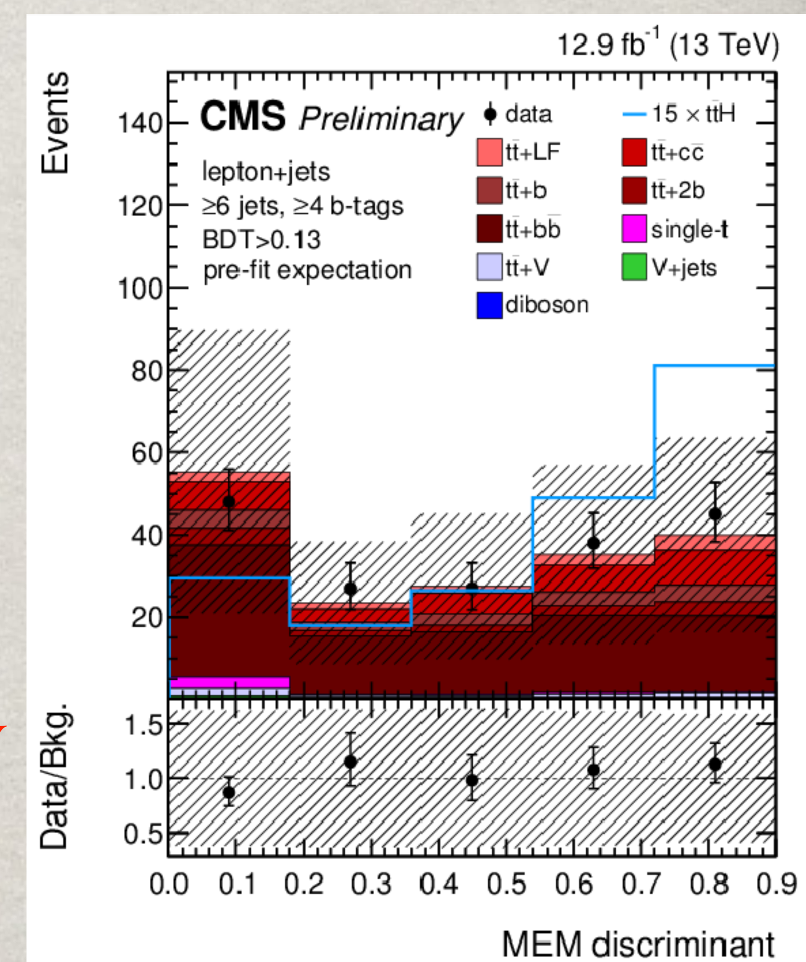
split category at the median of ttH BDT output





low purity

high purity

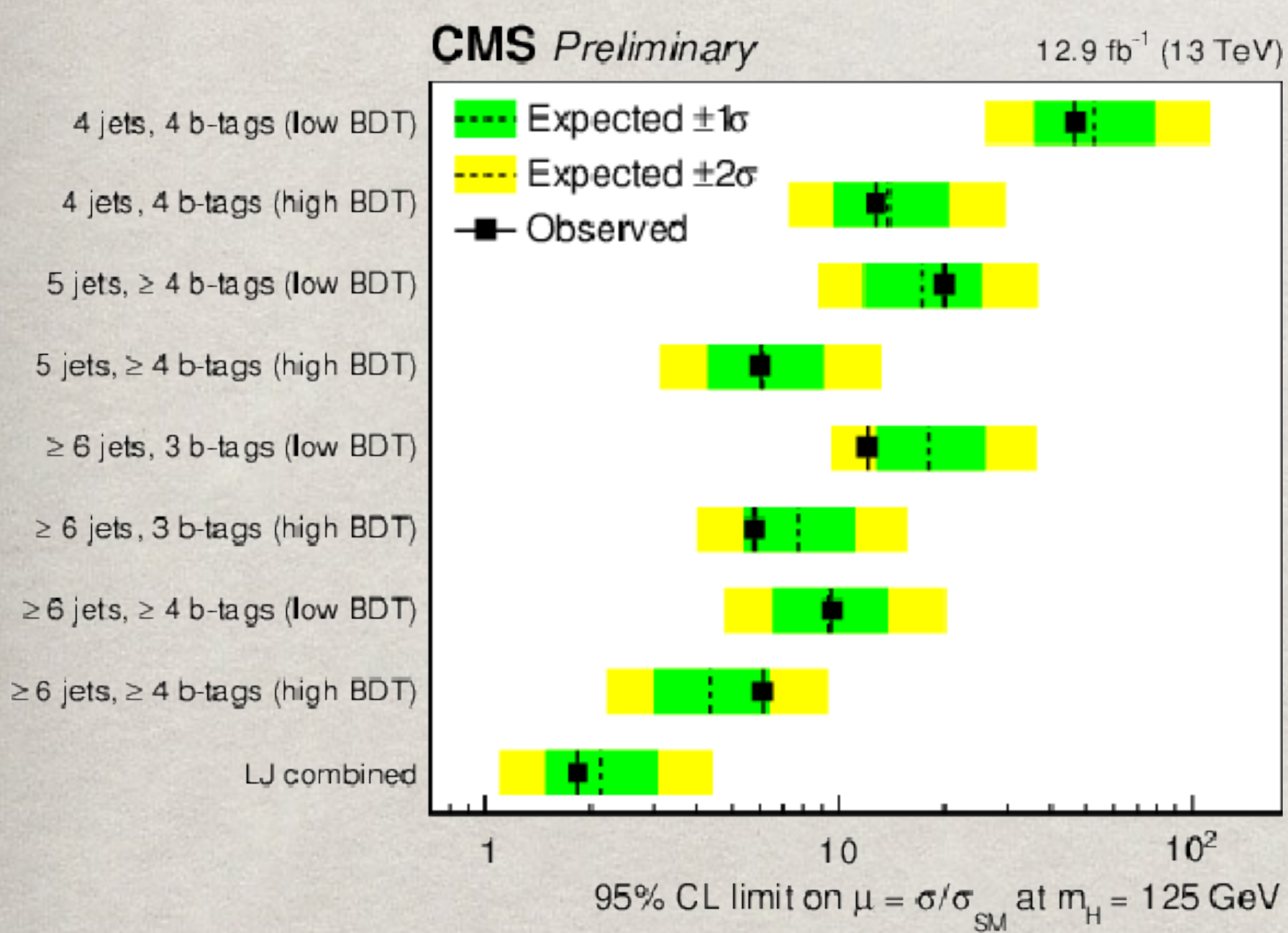


ADDITIONAL TT+HF UNCERTAINTY

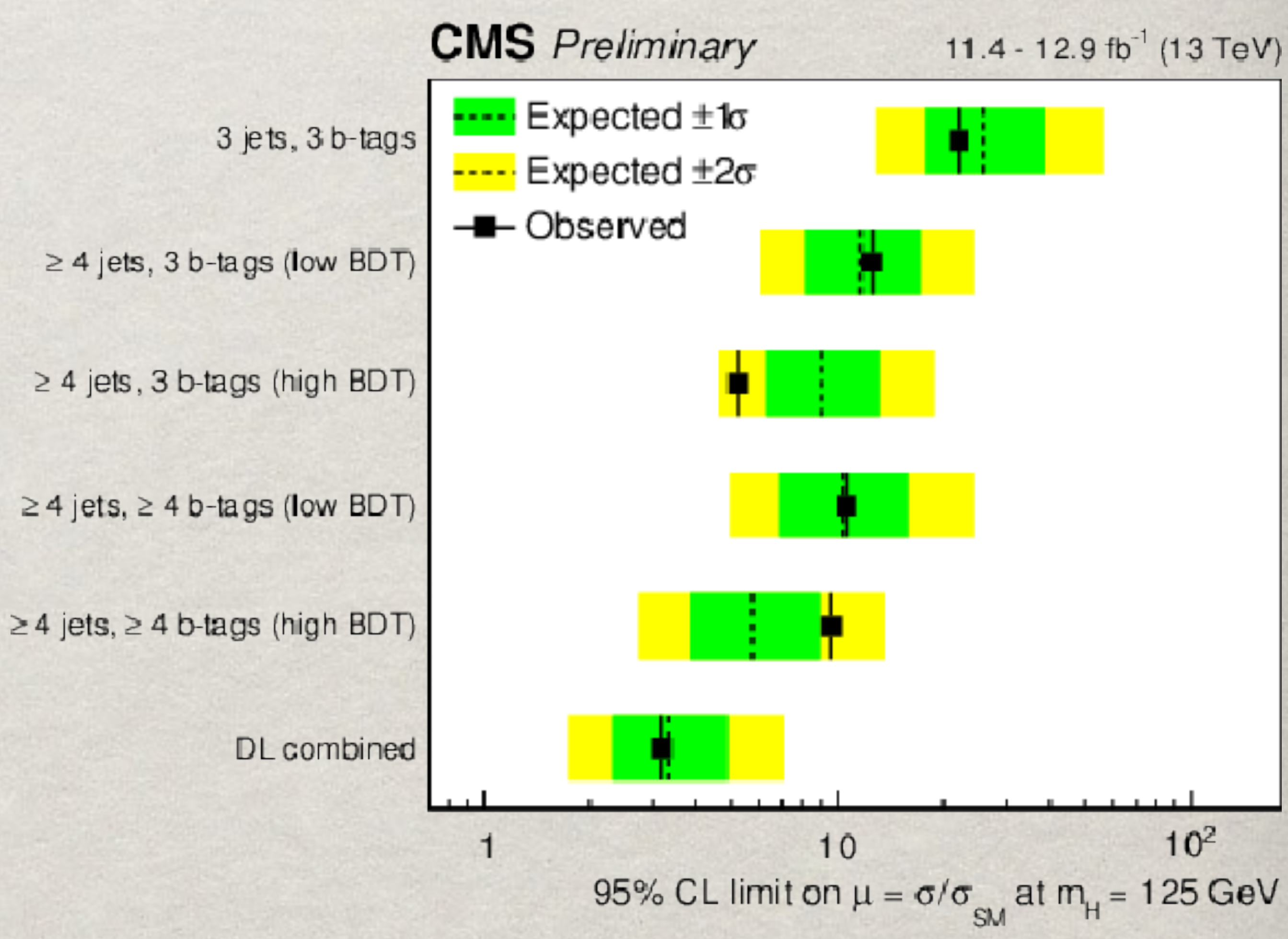
- Contribution from tt+HF very similar to signal
  - uncertainty on rate/shape has a big impact on our search
- Due to lack of more accurate higher order theory predictions, we obtained tt+HF estimate and uncertainty based on the inclusive ttbar sample
- On top of other uncertainty, assign an extra 50% rate uncertainty for tt+bb/b/B/cc independently

LIMITS

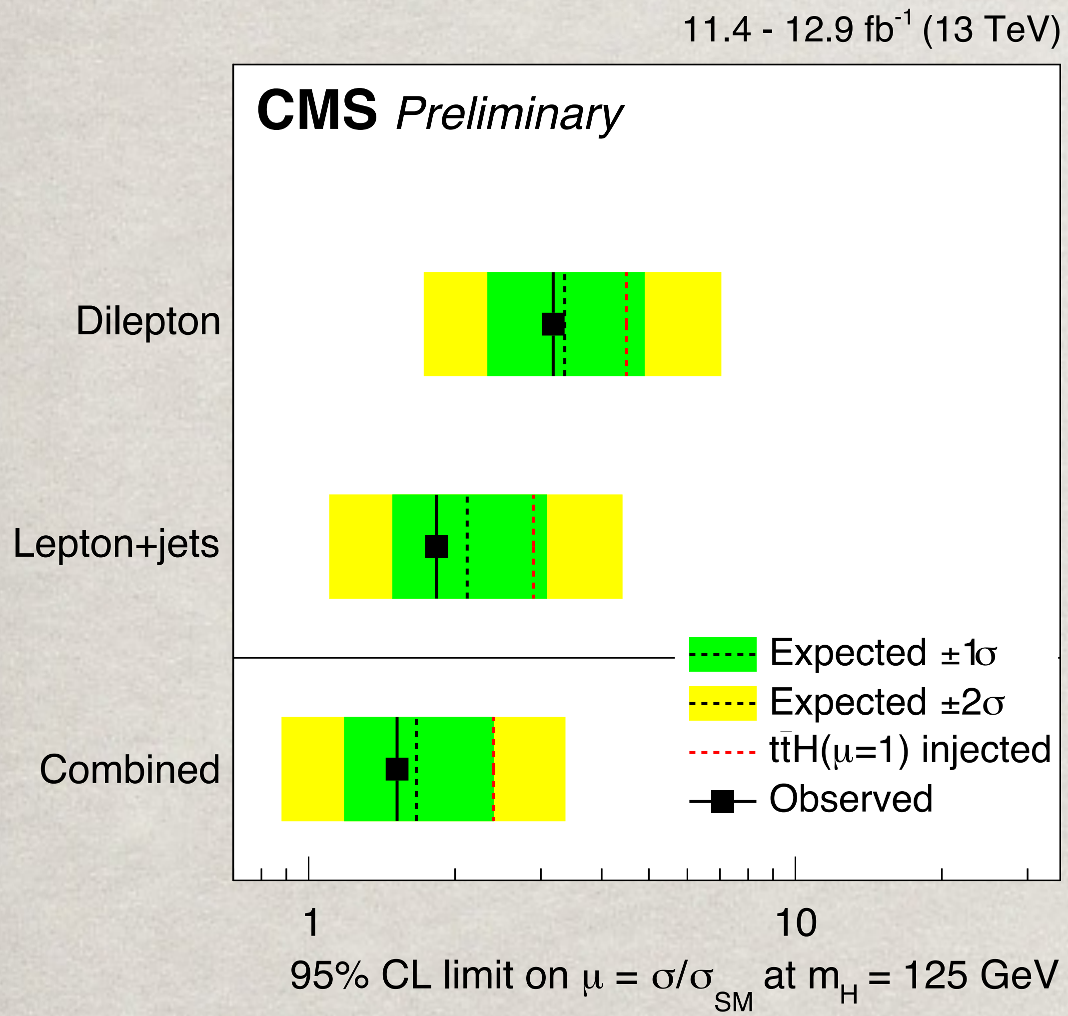
Lepton+Jets

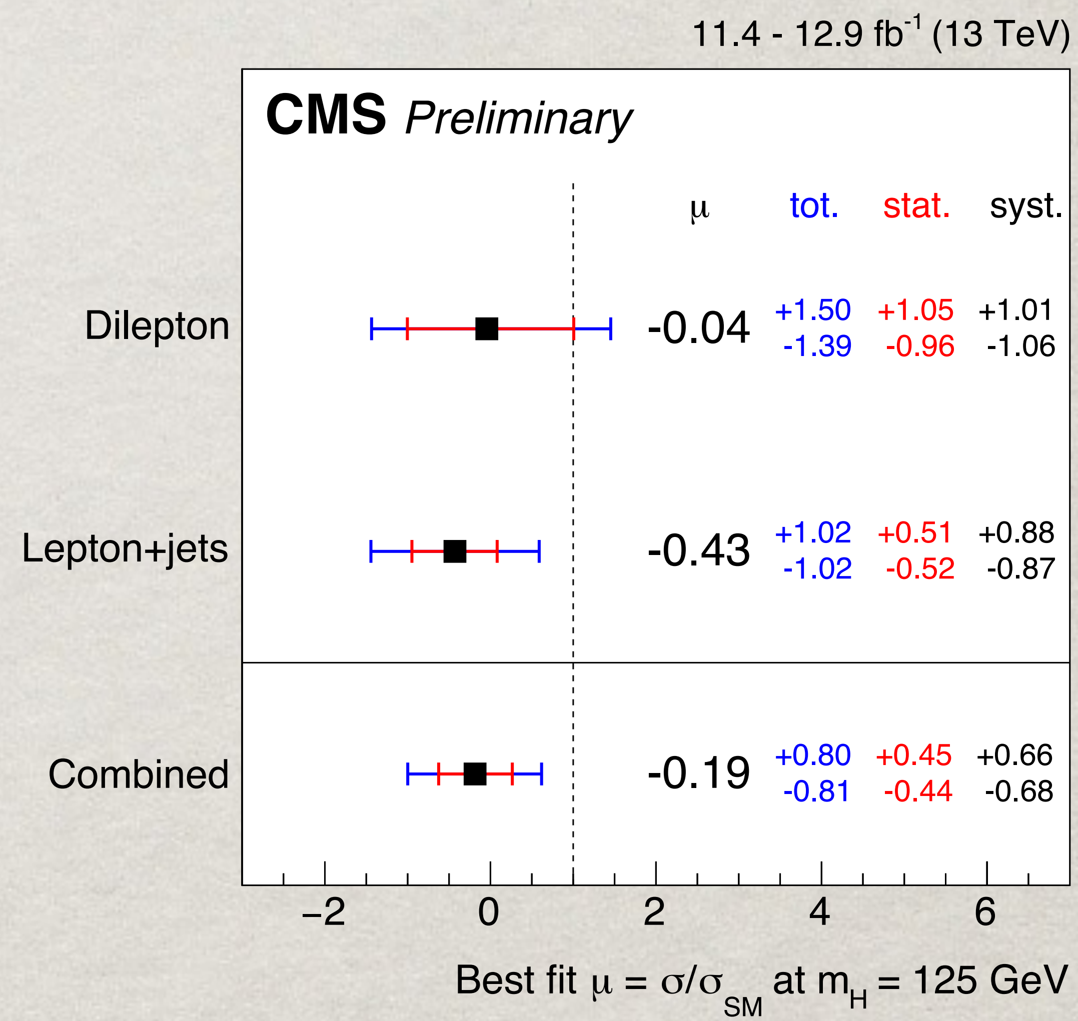


DiLepton



LIMITS AND SIGNAL STRENGTH





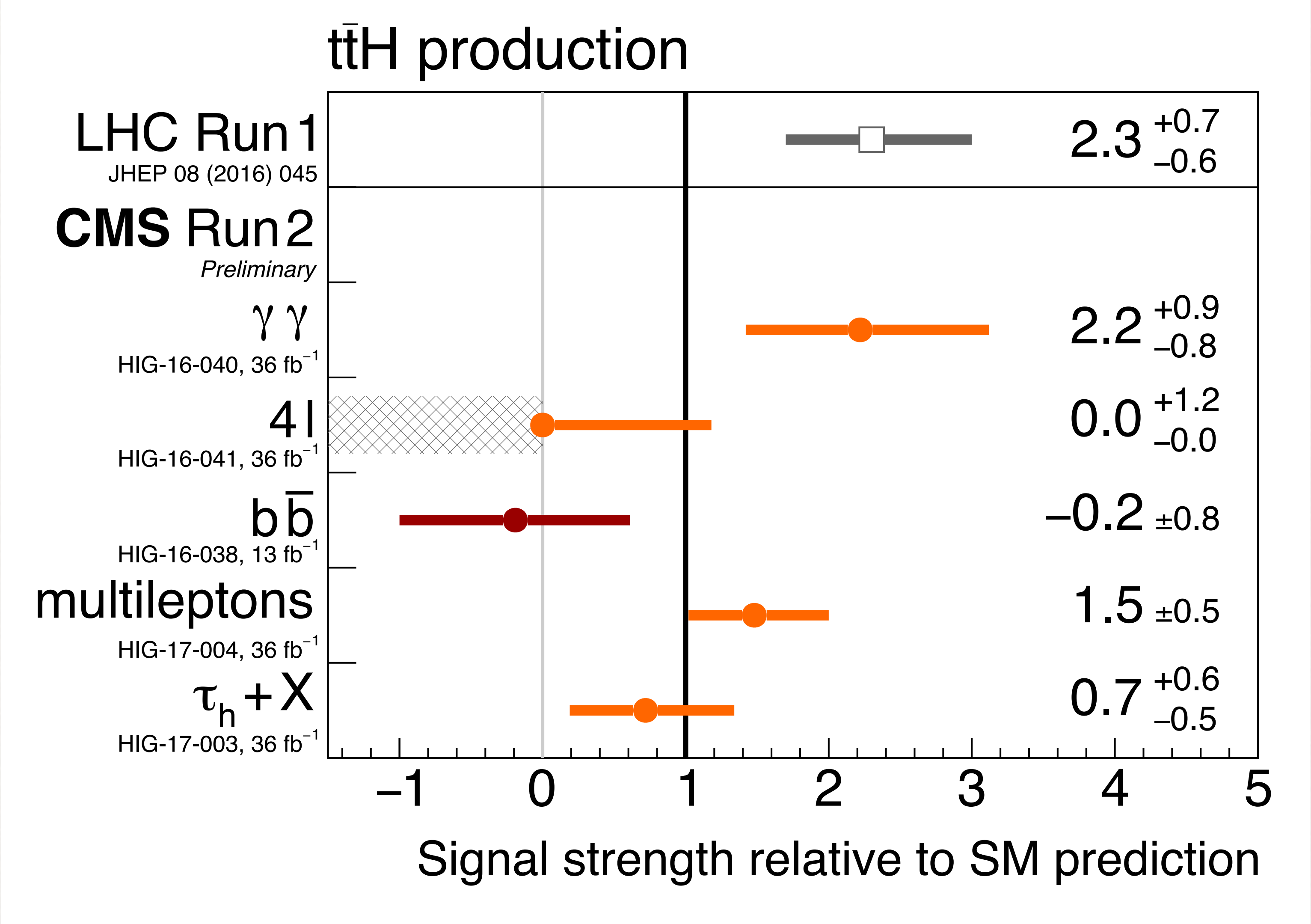
Channel	Observed UL	Expected UL	Best-fit μ
Dilepton	3.2	$3.4^{+1.5}_{-1.0}$	$-0.04^{+1.50}_{-1.39}(\text{tot.})\ ^{+1.05}_{-0.96}(\text{stat.})\ ^{+1.01}_{-1.06}(\text{syst.})$
Lepton+jets	1.8	$2.1^{+1.0}_{-0.6}$	$-0.43^{+1.02}_{-1.02}(\text{tot.})\ ^{+0.51}_{-0.52}(\text{stat.})\ ^{+0.88}_{-0.87}(\text{syst.})$
Combined	1.5	$1.7^{+0.7}_{-0.5}$	$-0.19^{+0.80}_{-0.81}(\text{tot.})\ ^{+0.45}_{-0.44}(\text{stat.})\ ^{+0.66}_{-0.68}(\text{syst.})$

SUMMARY

- ttH(bb) directly probes top-Higgs Yukawa coupling
- It also has a few challenges:
  - Small production XS: split events to channels/categories
  - Higgs invariant mass not applicable: use BDT/MEM to extract signal
  - Difficult tt+HF bkg: MEM, extra uncertainty
- Latest results are approaching SM sensitivity
  - Observed(expected) upper limit: 1.5(1.7)
  - Best fit signal strength: -0.2±0.8

OUTLOOK

- Updated ttH(bb) results with 36/fb data come out soon
- Combined 13 TeV ttH searches(all decay modes) might yield interesting findings



BACK UP

DATA SAMPLES

- Runs B, C, D in 2016, $\sqrt{s}$ = 13TeV
- The total integrated luminosity is: L = 12.9/fb

lepton+jets triggers

Dataset	Trigger Name
SingleMu	HLT_IsoMu22_v*
SingleMu	HLT_IsoTkMu22_v*
SingleEle	HLT_Ele27_eta2p1_WPTight_Gsf_v*

dilepton triggers

Channel	Trigger Name
$\mu^+\mu^-$	HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_v*
$\mu^+\mu^-$	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_v*
e^+e^-	HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v*
$\mu^\pm e^\mp$	HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_v*
$\mu^\pm e^\mp$	HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_v*

SELECTION: LEPTONS

Muons	Single Muon Channel ID	Leading ID Dilpeton	Sub-Leading Dilepton ID Veto ID for Single Muon
p_T [GeV] $>$	25	25	15
$\|\eta\| <$	2.1	2.4	2.4
ID	tight	tight	tight
$\mathrm{Iso}_{\delta\beta}/p_T <$	0.15	0.25	0.25

Electrons	Single Electron Channel ID	Leading ID Dilpeton	Sub-Leading Dilepton ID Veto ID for Single Electron
p_T [GeV] $>$	30	25	15
$\|\eta\| <$	2.1	2.4	2.4
ID	80% eff. non-trig. MVA ID	80% eff. non-trig. MVA ID	80% eff. non-trig. MVA ID
$\mathrm{Iso}_{\rho A}/p_T <$	0.15	0.15	0.15

$\mu^+\mu^-$ and e^+e^- Channel:
$m_{\ell\ell} > 20$ GeV
$m_{\ell\ell} < 76$ GeV or $m_{\ell\ell} > 106$ GeV
MET > 40 GeV

- Lepton Pt > trigger thresholds
- Tight ID and isolation to suppress multi-jet events
- Veto Z+jets events for DL

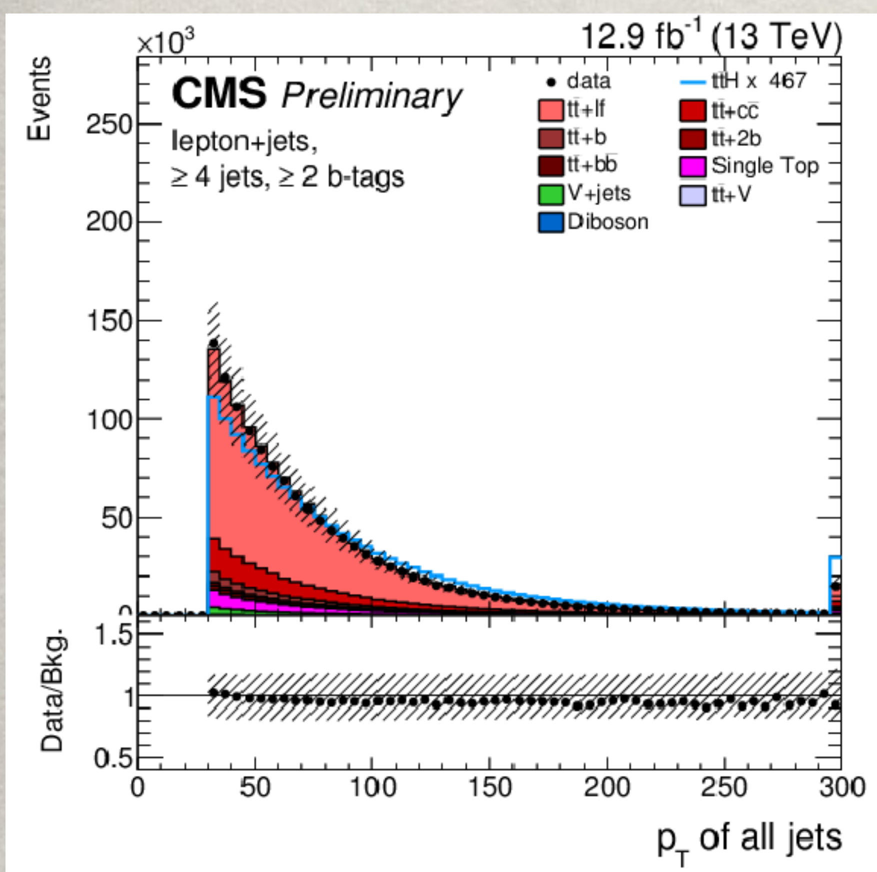
SELECTION: JETS

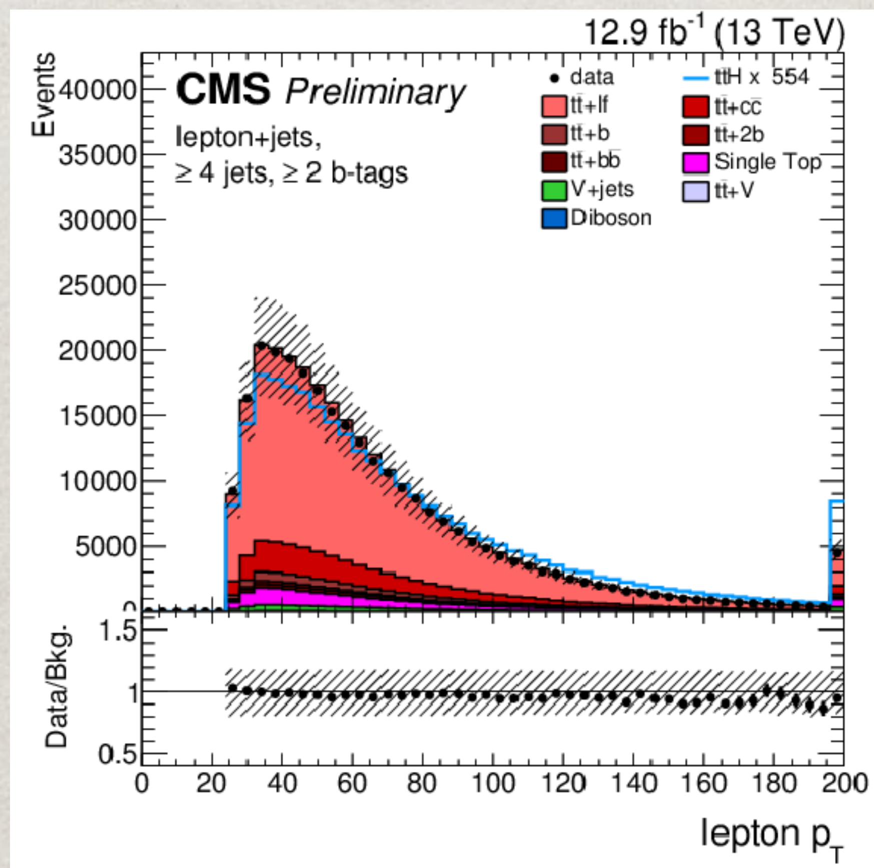
Jets	Single Lepton Channel Leading 2 Jets Dilepton	Dilepton Channel Subleading Jets Dilepton
Type	PFJets, CHS	PFJets, CHS
Algorithm	anti-k_T 0.4	anti-k_T
p_T [GeV] $>$	30	20
$\|\eta\| <$	2.4	2.4
Lepton cleaning	Require $\Delta R(\ell, j) > 0.4$	Require $\Delta R(\ell, j) > 0.4$

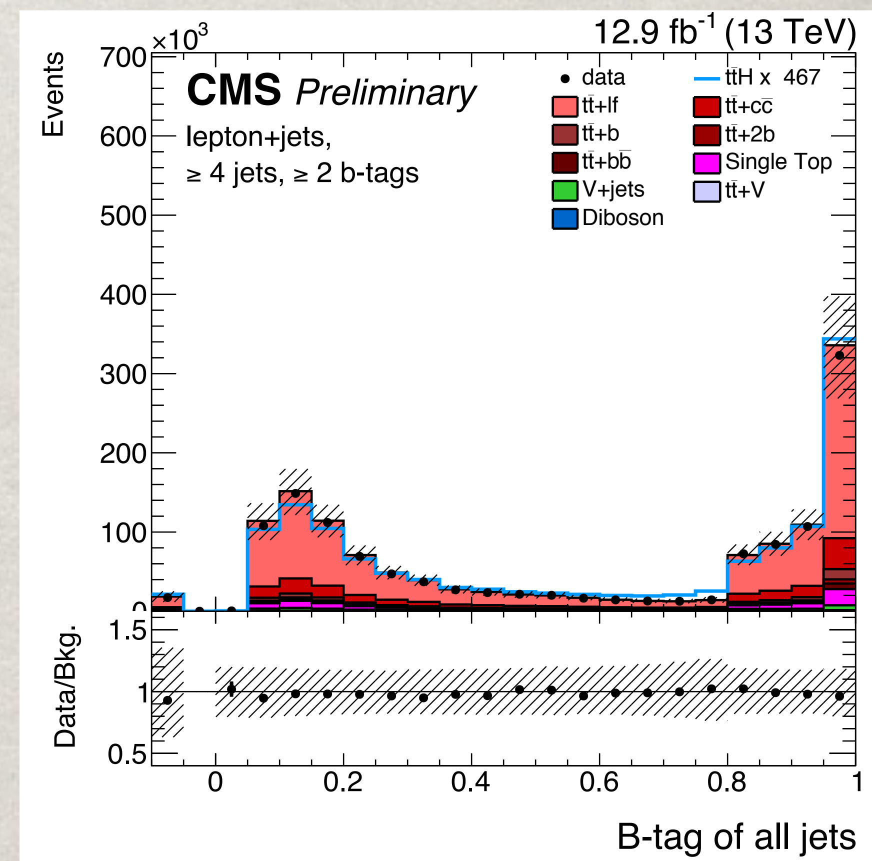
- Jet multiplicity:
  - ≥4 jets in LJ channel
  - ≥2 jets in DL channel
- b-tags: jets originating from b quarks
  - use CSV(Combined Secondary Vertex) algorithm
  - identify as b-jets if passing Medium working point
  - 1(2) b-tags for DL(LJ) inclusive selection

DATA/MC AGREEMENT

- All corrections to MC applied
- Good agreement between Data and MC







III: Matrix Element

Numerical integration

Momentum conservation

Resolution function (allow ISR)

$$w(\vec{y}|\mathcal{H}) = \sum_{i=1}^{N_a} \int \frac{\mathrm{d}x_a \mathrm{d}x_b}{2 x_a x_b s} \int \prod_{k=1}^{8} \left(\frac{\mathrm{d}^3 \vec{p}_k}{(2\pi)^3 2E_k} \right) (2\pi)^4 \, \delta^{(E,z)} \left(p_a + p_b - \sum_{k=1}^{8} p_k \right) \mathcal{R}^{(x,y)} \left(\vec{\rho}_{\mathrm{T}}, \sum_{k=1}^{8} p_k \right) \times$$
$$g(x_a, \mu_{\mathrm{F}}) g(x_b, \mu_{\mathrm{F}}) |\mathcal{M}_{\mathcal{H}} \left(p_a, p_b, p_1, \ldots, p_8 \right)|^2 W \left(\vec{y}, \vec{p} \right),$$

Parton density functions

LO Scattering amplitude (Open Loops)

Detector transfer function

$$P_{\mathrm{s/b}} = \frac{w(\vec{y}|\mathrm{t\bar{t}H})}{w(\vec{y}|\mathrm{t\bar{t}H}) + k_{\mathrm{s/b}} w(\vec{y}|\mathrm{t\bar{t}{+}b\bar{b}})}$$

Construct per-event signal/background probability
using full kinematic information in an analytical approach

Ideal for final states with many reconstructed objects.

Built for ttH(bb) vs ttbb

SYSTEMATIC UNCERTAINTIES

- PileUp re-weighting: use 69.4 mb ± x%
- Lepton SF
  - independently vary id and HLT efficiency
- b-tag SF
- JER and JES
  - JER has a negligible effect on shape or normalization
- Luminosity: x.x%
- Cross section
  - Use CMS standard model cross section uncertainties
- MC statistics
- Q^2 scale/Parton Shower for ttJets
- extra 50% rate uncertainty for tt+HF