Higgs results with direct top and b-Yukawas with CMS

LHCP 2017
SHANGHAI

Christian Contreras (DESY)
on behalf of the CMS collaboration

May 12, 2017

Overview

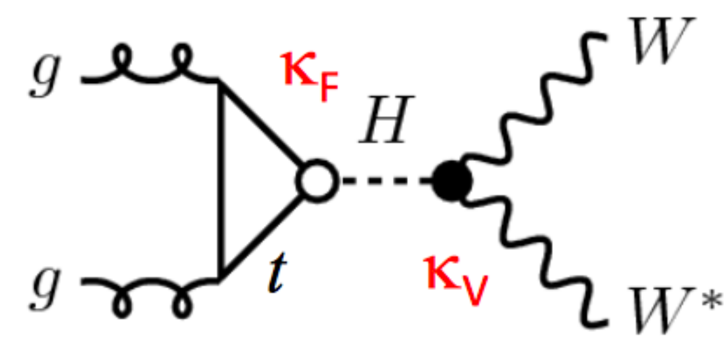
- A search for the Higgs boson decaying to 2 b quarks is performed at CMS in 3 production channels:
  - Associate production with top quarks ($t\bar{t}$H, tH)
  - Vector boson fusion (VBF)
  - Associate production with a vector boson (VH)

- The search for $t\bar{t}$H production is performed in 3 broad decay channels:
  - H→b$\bar{\text{b}}$: Analysis targeting production in the leptonic, dilepton, & hadronic final states (PAS HIG-16-038, presented @ Higgs Coupling 16')
  - H→ multileptons: Analysis targeting in leptonic (e,μ) final
  - states from H→WW, ττ, ZZ (PAS HIG-17-004, presented @ Moriond 17')
  - H→ γγ: Analysis targeting in leptonic & hadronic final states (PAS HIG-16-020, presented @ LHC Day 16')

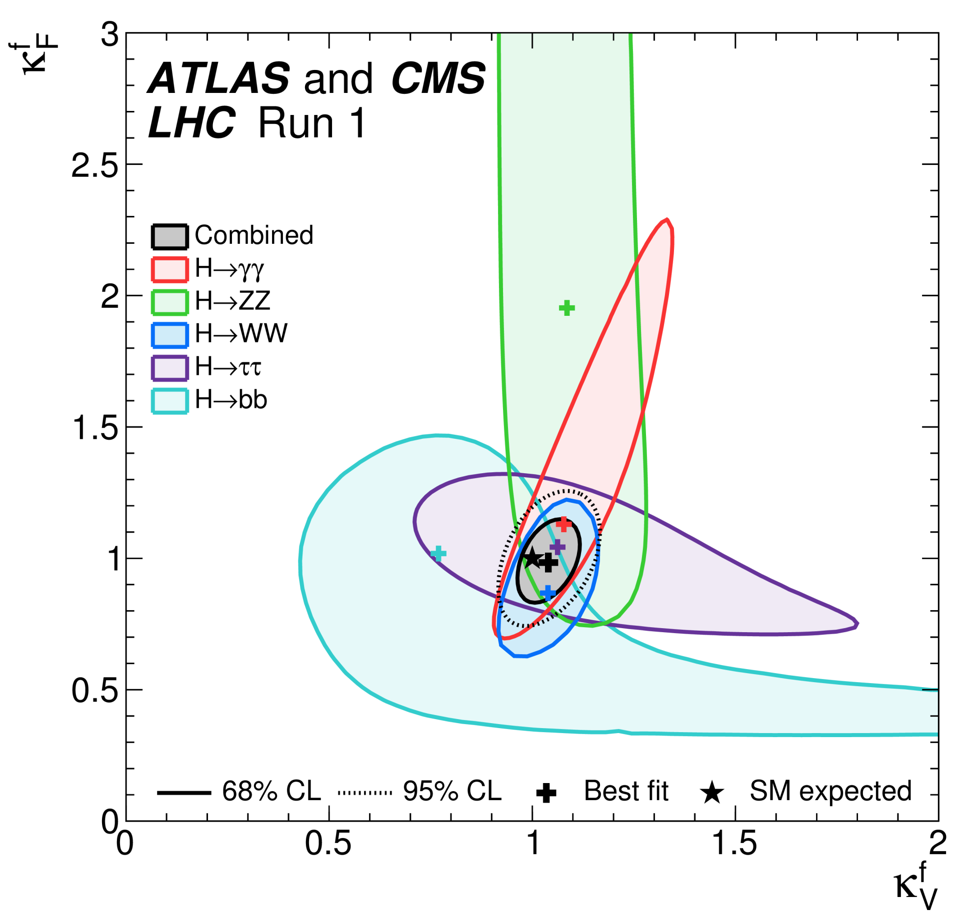
- $t\bar{t}$H and tH employ MVA techniques while VBF and VH a regression method is used to reconstruct the invariant mass of the two b jets

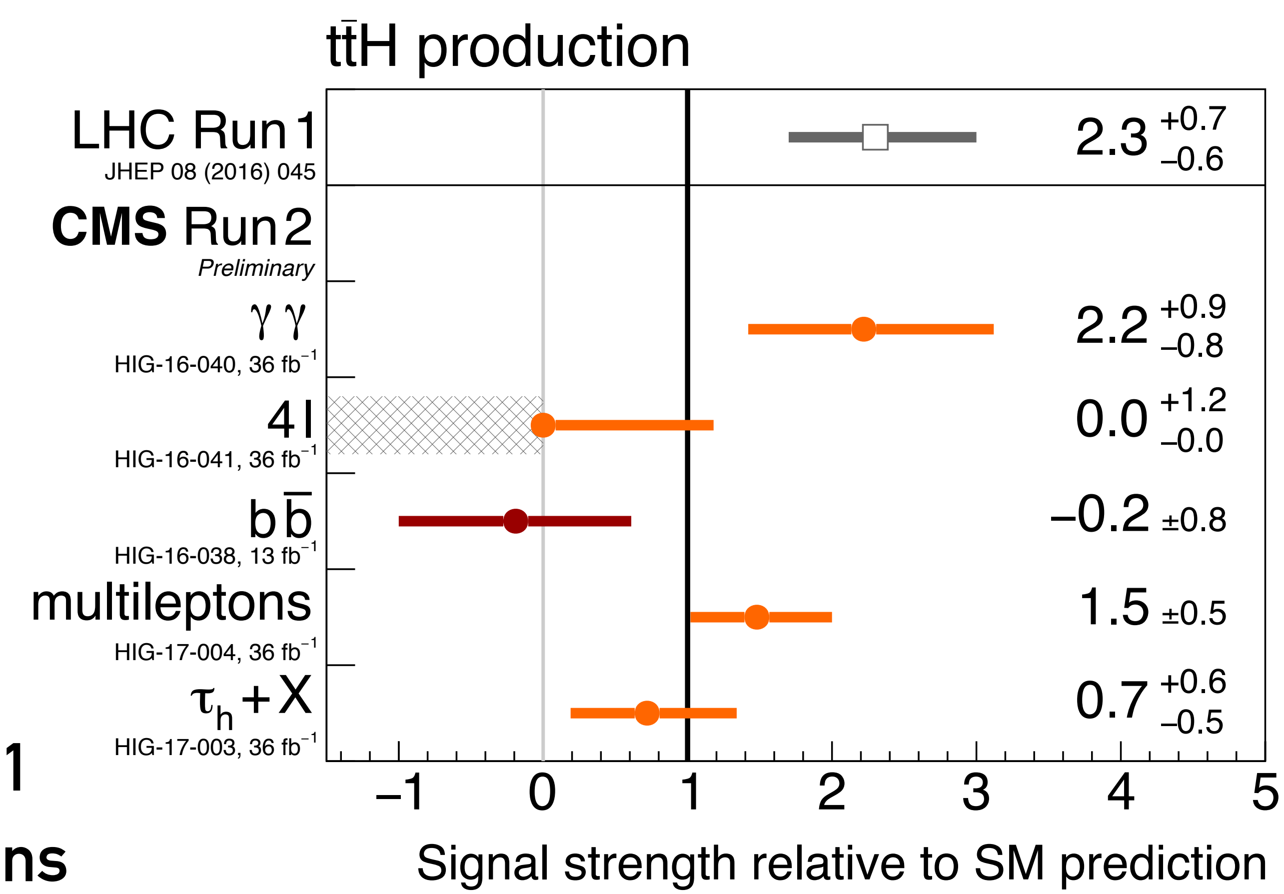
Run 1 coupling results

- Couplings to fermions and boson strongly constrained by Run-I measurements
- Scale factors k_j introduced to quantify deviation from SM
- One benchmark uses 2 scale factors k_V for vector boson and k_F for fermions



- In combination, search for $H \rightarrow \tau^+\tau^-$ exceeds 5σ
- But, despite being the dominant decay mode, coupling to $b\bar{b}$ not yet observed





- All channels compatible with k_V=1 and k_F=1
- Result is consistent with the SM expectations

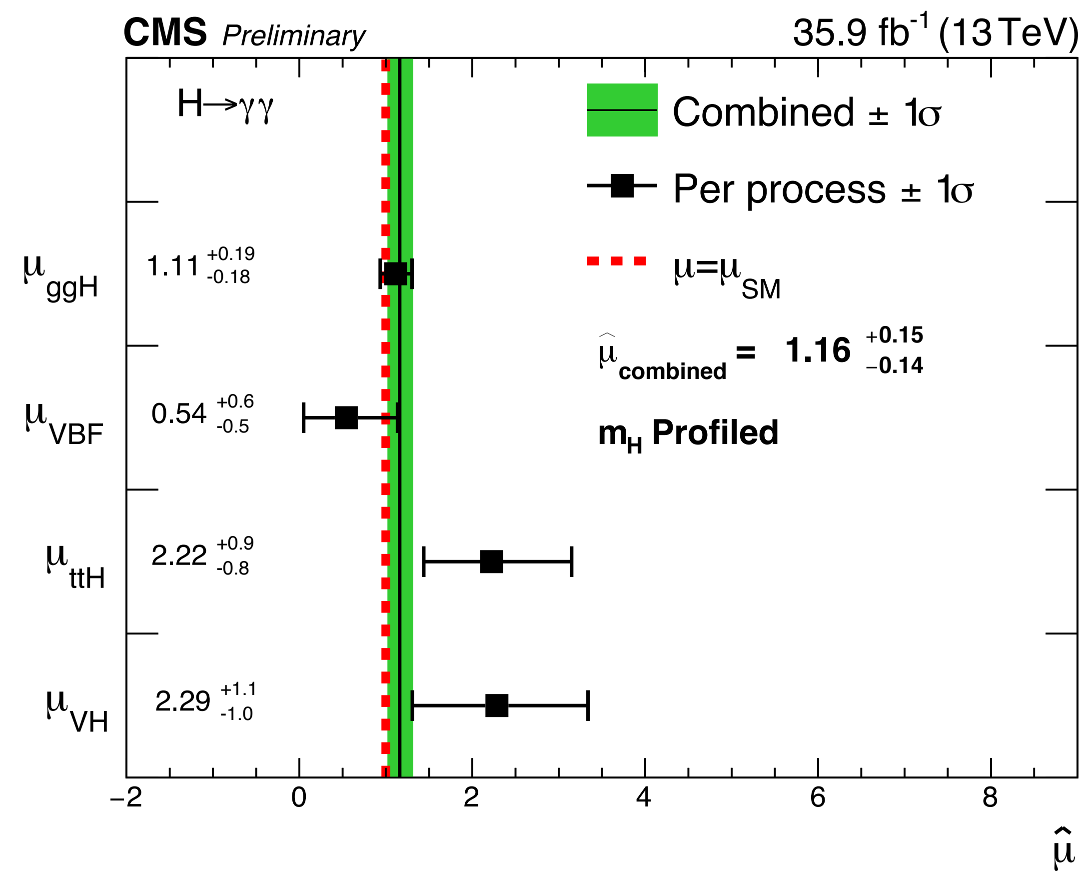
The Hunt for $t\bar{t}$ H @ 13 TeV

ttH why so relevant, why now?

- The combined H($b\bar{b}$) and H($\tau\tau$) result establishes strong evidence for coupling of the Higgs boson to down-type 3rd generation fermions

- Indirect and direct results on $t\bar{t}$ H coupling is also evident for a coupling to up-type fermions [arXiv:1401.6527]

- The $t\bar{t}$ H cross section increases by a factor of ~4 @ 13 TeV
  - Direct measurement is a key process to determine top Yukawa coupling



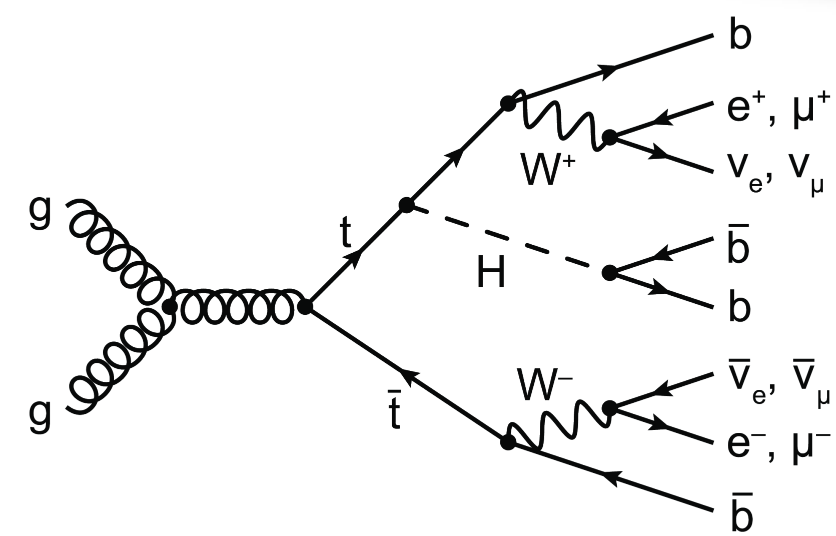
- Small cross section that grows substantially from 7 to 8 and to 13 TeV for Higgs @ 125 GeV (NLO QCD+EW):

→ $\sqrt{s}$=7 TeV: σ(ttH)≃89 fb^{-1}
→ $\sqrt{s}$=8 TeV: σ(ttH)≃133 fb^{-1}
→ $\sqrt{s}$=13 TeV: σ(ttH)≃507 fb^{-1}

$t\bar{t}H(b\bar{b})$

CMS-PAS-HIG-16-038

- Large H → $b\bar{b}$ Branching fraction
- Dominant background: tt+jets
  - Irreducible contribution: tt+bb (theoretically challenging)
- Many jets with similar kinematics and limited mass resolution for H → $b\bar{b}$



Analysis strategy

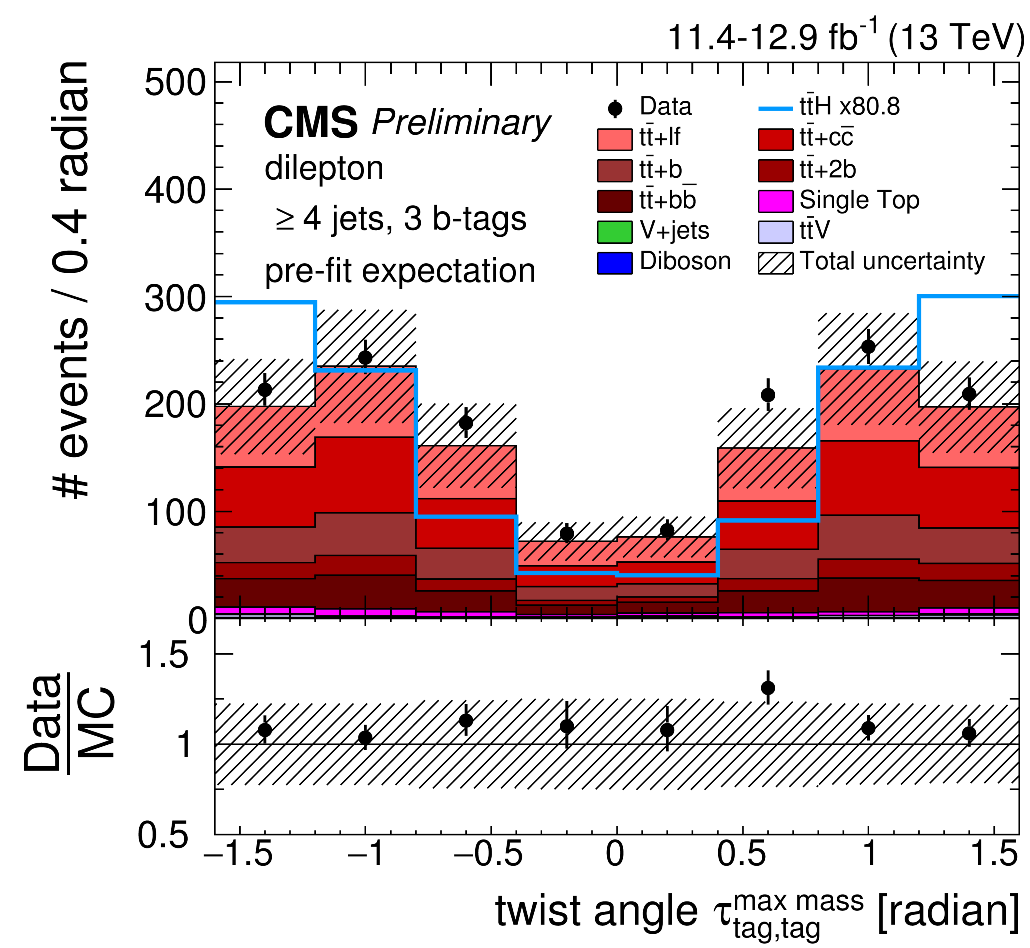
- Obtain good signal separation & constrain background
- Event categories: 11 (5) lepton+jets (dilepton)
- Lepton triggers and offline event selection

Lepton+jets

- exactly 1 lepton
- At least 4 jets
- At least 2 b-tagged jets

Dileptons

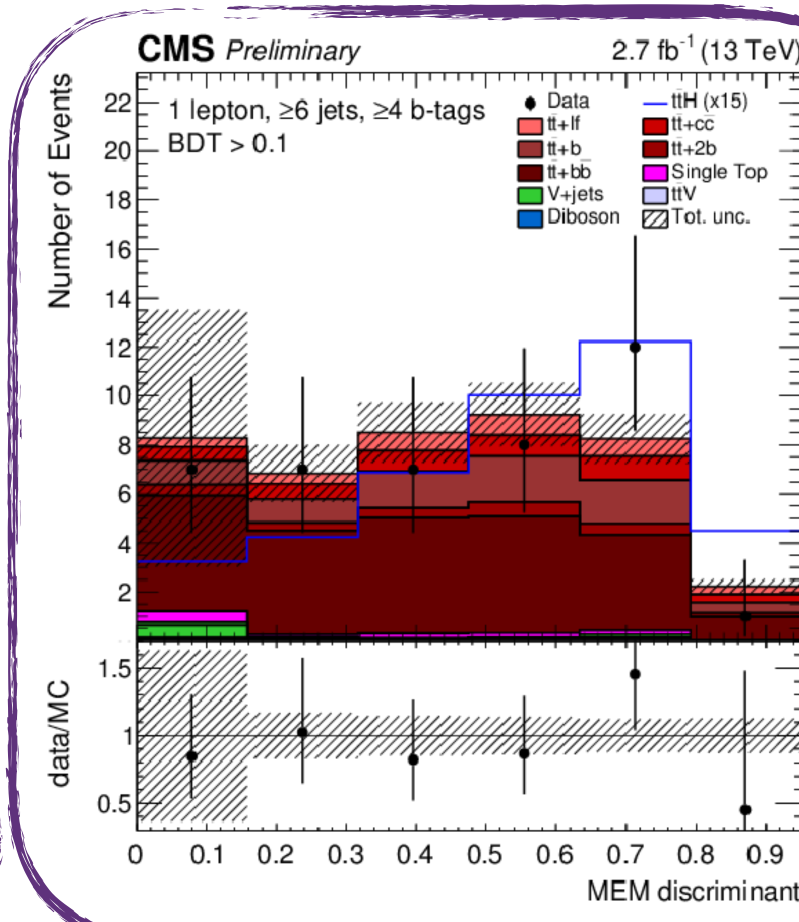
- 2 opposite sign leptons
- At least 3 jets
- At least 2 b-tagged jets



- Leptons + jets: high statistics
- Dilepton: minimal non-$t\bar{t}$ background, and jet combinatorics
- Classify events based on jet, b-tag multiplicities and boosted jets (leptons+jets)

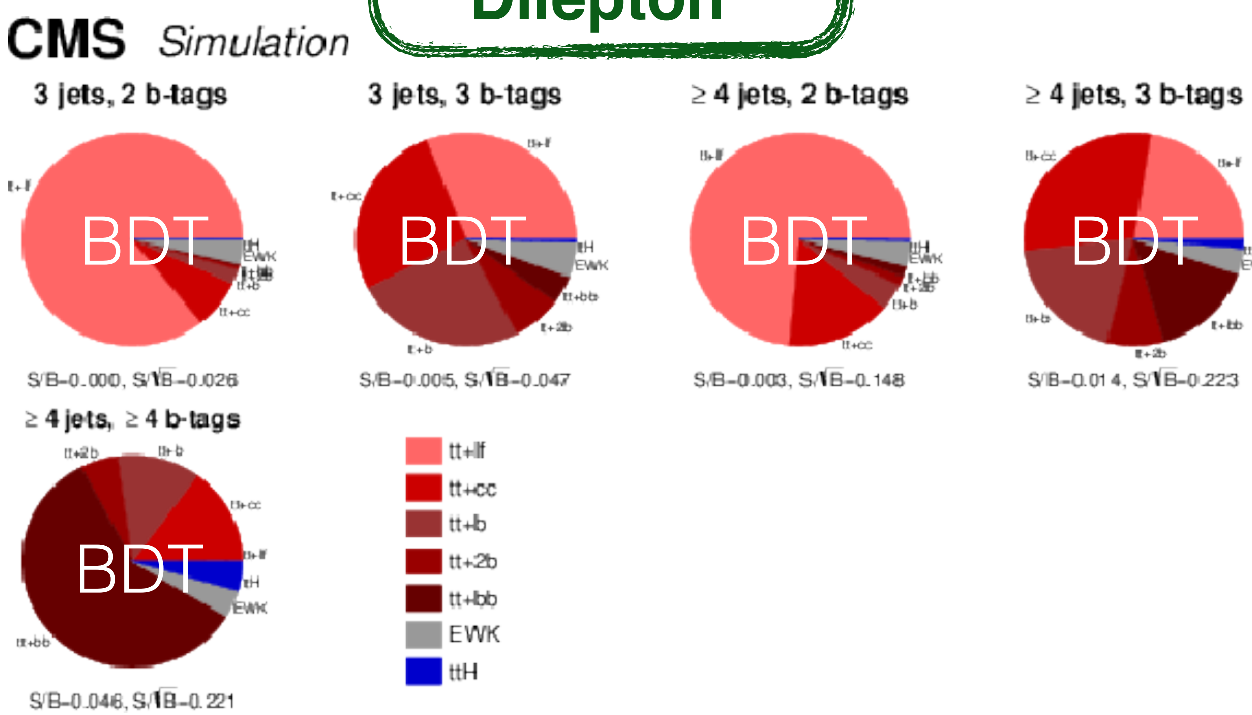
ttH(b$\bar{b}$) signal seperation

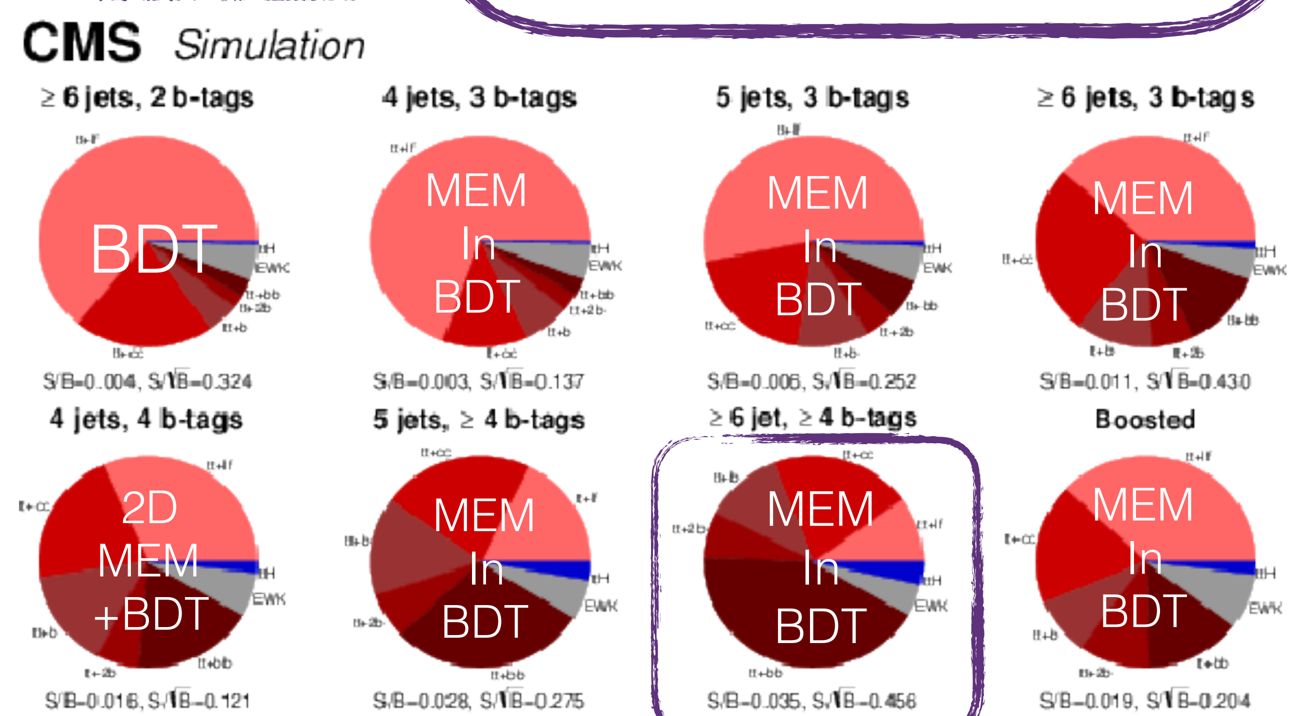
- Each sub-category discriminator is optimized to improve sensitivity
  - Dilepton: use Boosted Decision Tree (BDT)
  - Lepton+jets: use Matrix Element Method (MEM)
    - $t\bar{t}$ H vs $t\bar{t}$ +bb background hypothesis, permuting over all b-quark association
    - MEM as input to BDT modeling
    - 2Dim MEM+BDT analysis



Lepton+jets

Dilepton



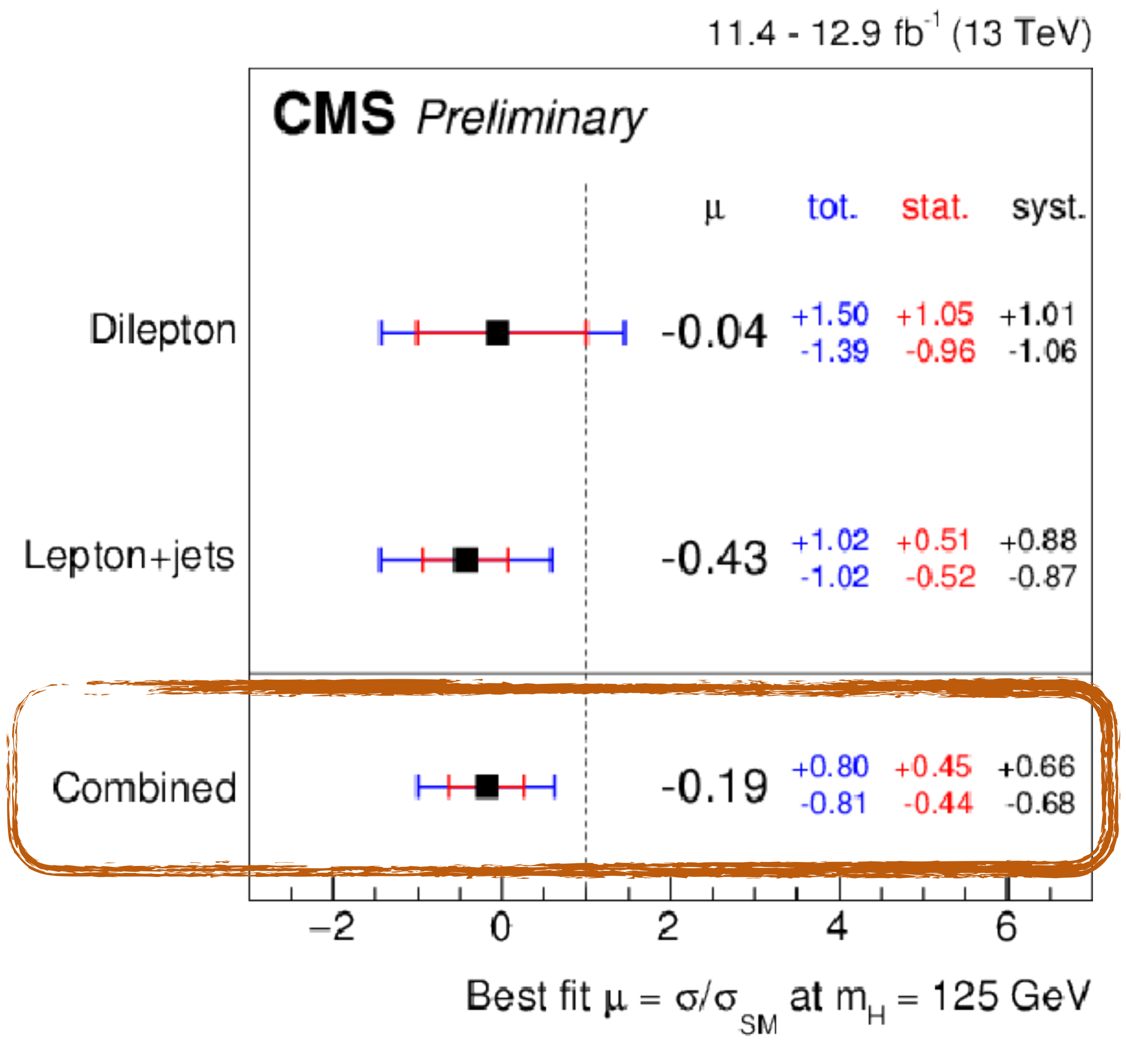


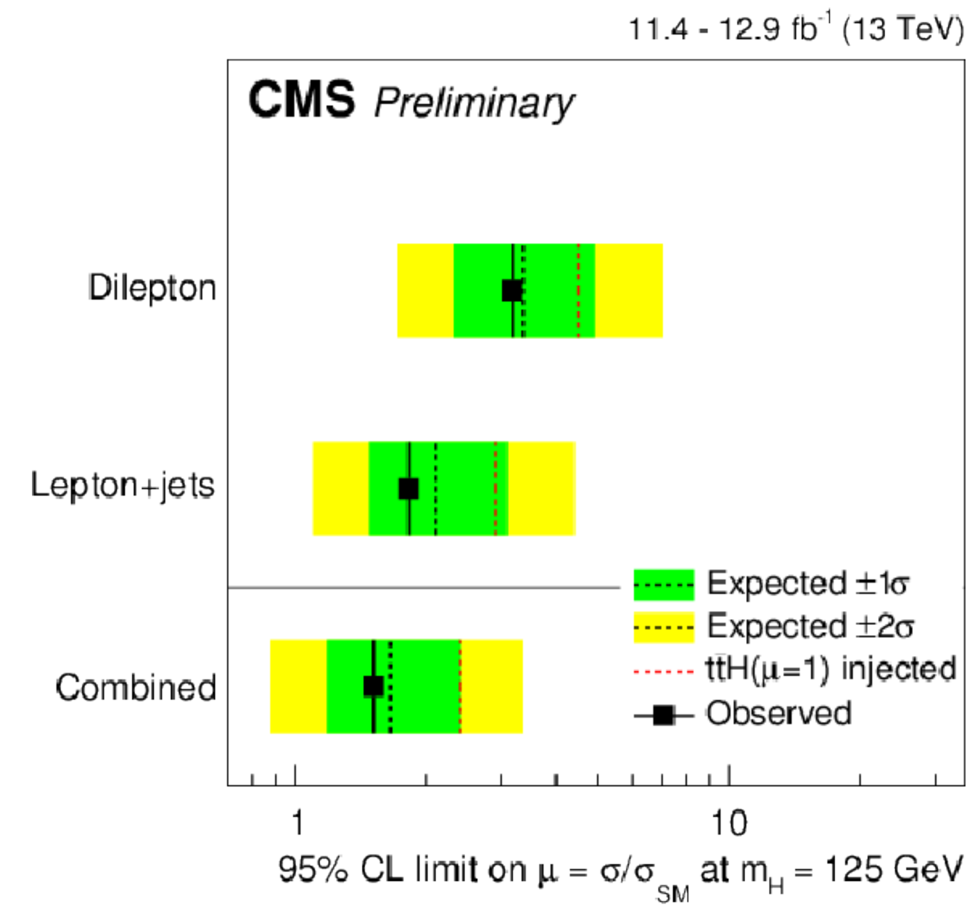
ttH($b\bar{b}$) results

Combined fit of discriminant output across all event categories

- Observe no significant excess
- Limited by systematics
    - dominated by those on $t\bar{t}$ +(b-)jets background

Upper limit at 95% CL

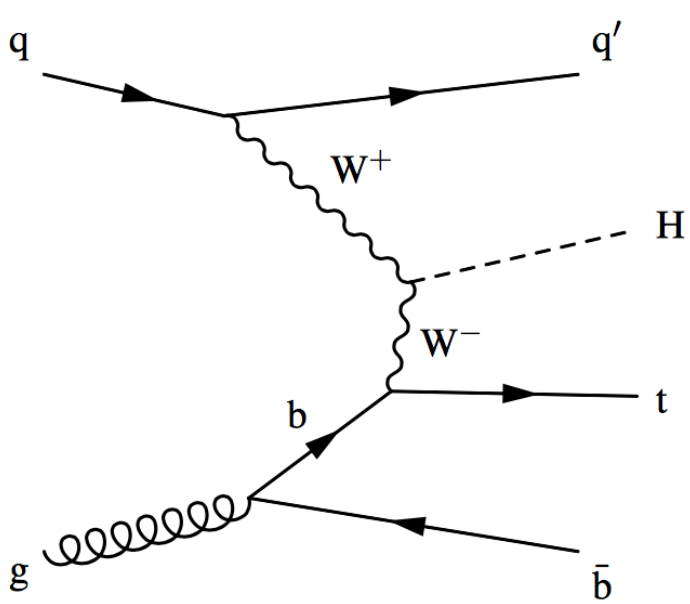




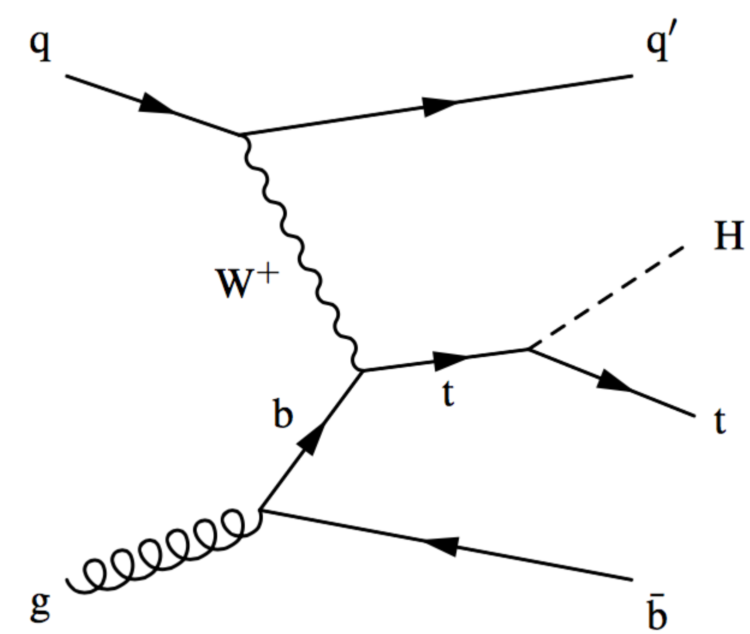
Channel	Observed UL	Expected UL	Best-fit μ
Dilepton	3.2	$3.4^{+1.5}_{-1.0}$	$-0.04^{+1.50}_{-1.39}$(tot.) $^{+1.05}_{-0.96}$(stat.) $^{+1.01}_{-1.06}$(syst.)
Lepton+jets	1.8	$2.1^{+1.0}_{-0.6}$	$-0.43^{+1.02}_{-1.02}$(tot.) $^{+0.51}_{-0.52}$(stat.) $^{+0.88}_{-0.87}$(syst.)
Combined	1.5	$1.7^{+0.7}_{-0.5}$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

Higgs production with single top

- At LO tH can be separated into 3 production modes:
  - t-channel (tHq) (diagrams interfere destructively in SM)
  - Associated tW production (tHW)
  - s-channel (negligible cross section at the LHC)



t-channel



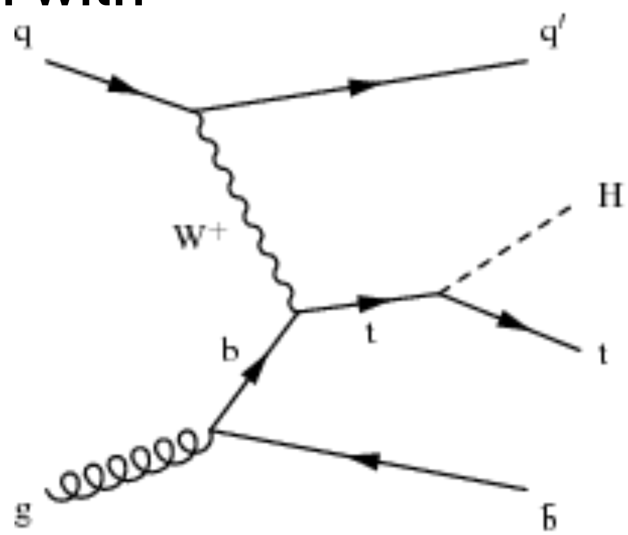
- Sensitive to both the magnitude and sign of top Yukawa coupling
- In BSM scenarios not necessarily (e.g. inverted top coupling scenario)

→ Effective theory with possibly CP violating top Yukawa couplings, and modified couplings to vector bosons (Eur. Phys. J. C 75 (2015), no. 6, 267)

Modified top Yukawa coupling

CMS-PAS-HIG-16-019

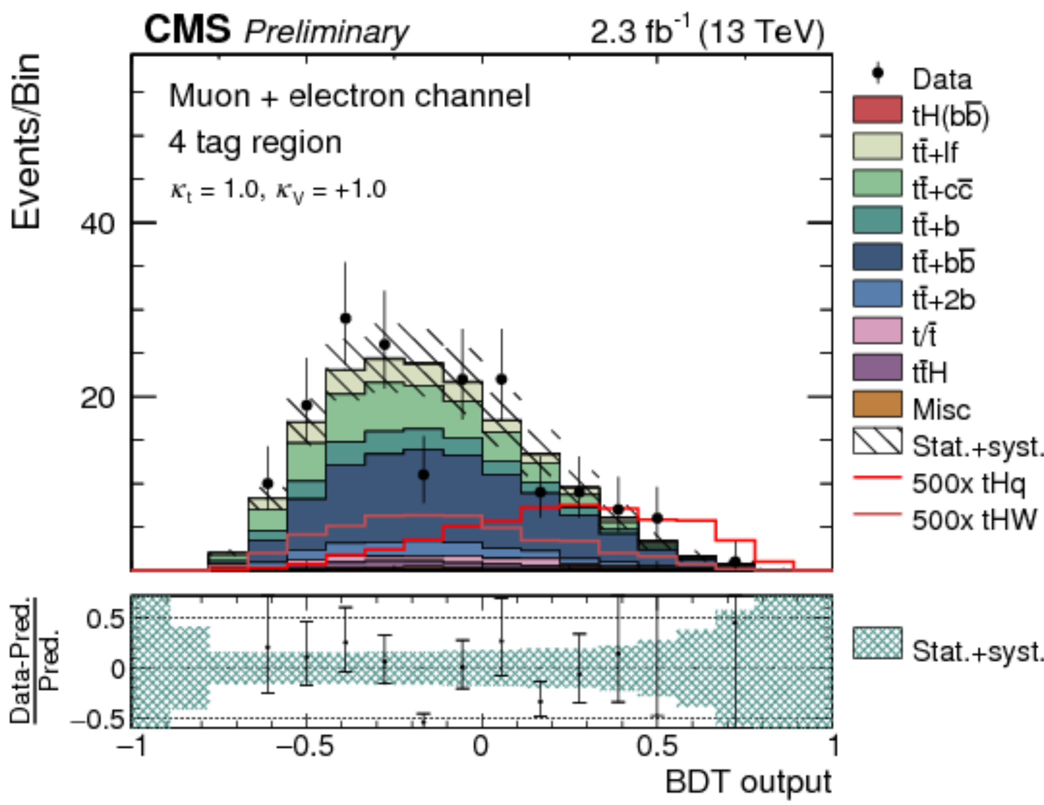
Search for $H \to b\bar{b}$ in association with a single top ($t \to b\, e\nu / b\, \mu\nu$)

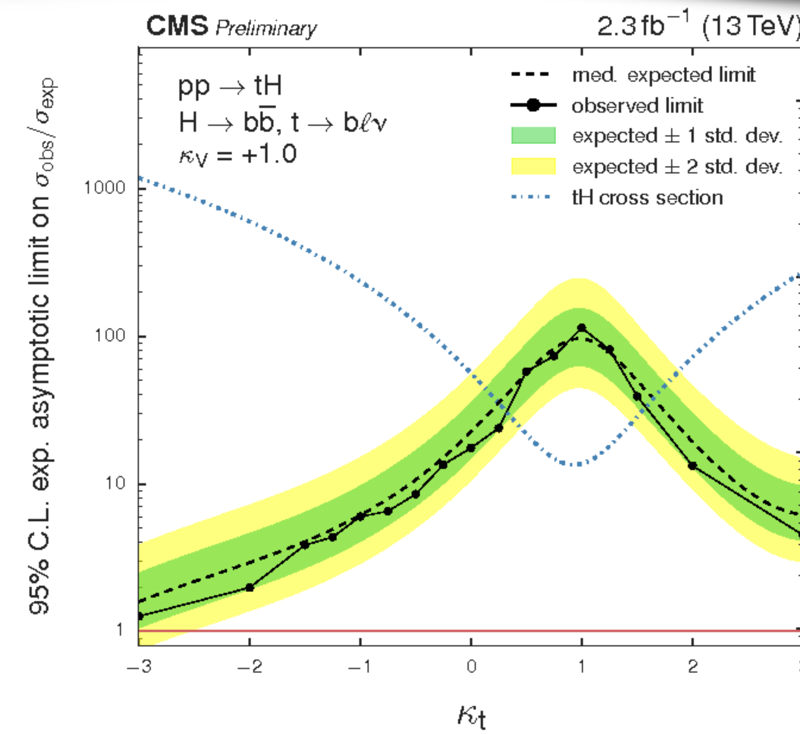


Signal region:

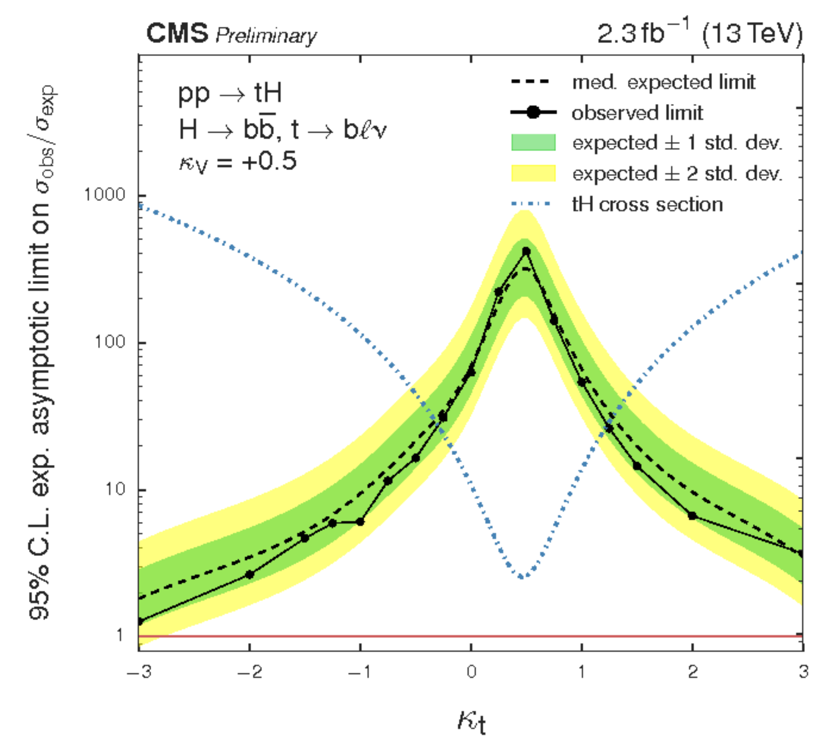
- Final state e/mu+3 or 4 b-tagged jets, 1 non-tagged jet

- BDT to find jet assignment for $t\bar{t}$ and tHq hypothesis
- Final discrimination MVA classifier kinematics + kinematics interpreted in the two hypothesis





SM coupling to W,Z



Reduced coupling to W,Z

Observed (expected) 95% CL upper limit for the SM is $113.7 \times \sigma_{SM}$ ($98.6 \times \sigma_{SM}$) & $6.0 \times \sigma_{ITC}$ ($6.4\times \sigma_{ITC}$)

tHq+tHW, H→multileptons

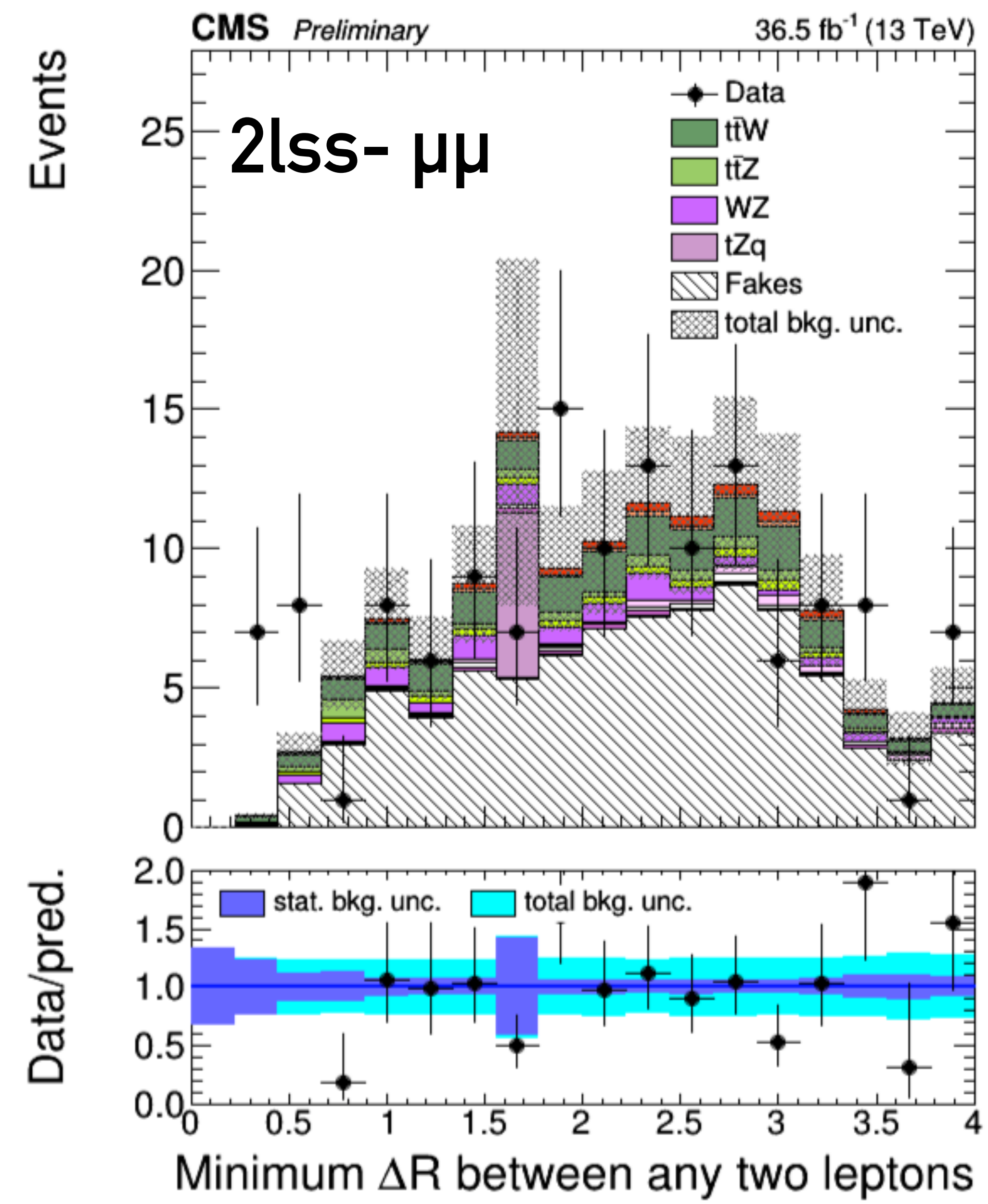
CMS-PAS-HIG-17-005

The process exposes relative sign of top-Higgs and W-Higgs couplings via interference

σ (tHq) = 792.7 fb

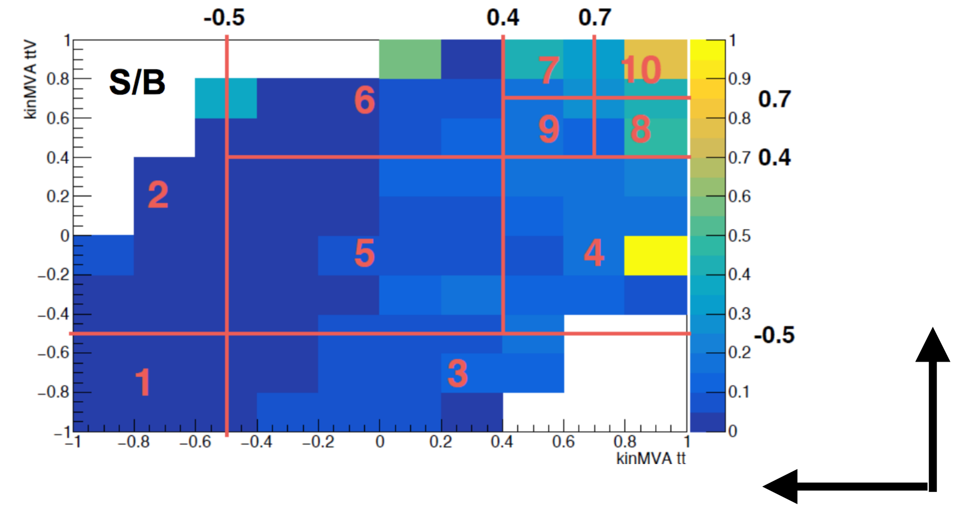
σ_{SM}(tHW) = 15.61 fb

- Irreducible bkg (MC):
  - $t\bar{t}$ +X (X=W/Z/H/γ*)
  - Photon conversions
  - Rare SM tZq, tWZ, tri-bosons, WWqq, tttt
  - Di-bosons WZ, ZZ
- Reducible bkg (data-drive):
  - Fakes due to non-prompt leptons & miss-ID of jets passing lepton selection
  - Charge flips:
    - Charge mis-ID (2lss)
    - Opposite-sign processes (e.g $t\bar{t}$/Z+jets)
- Analysis strategy
  - Same-sign dilepton (2lss): 1 W from Higgs decays hadronically, others decay leptonically
  - Trilepton (3l): All 3 Ws decay leptonically



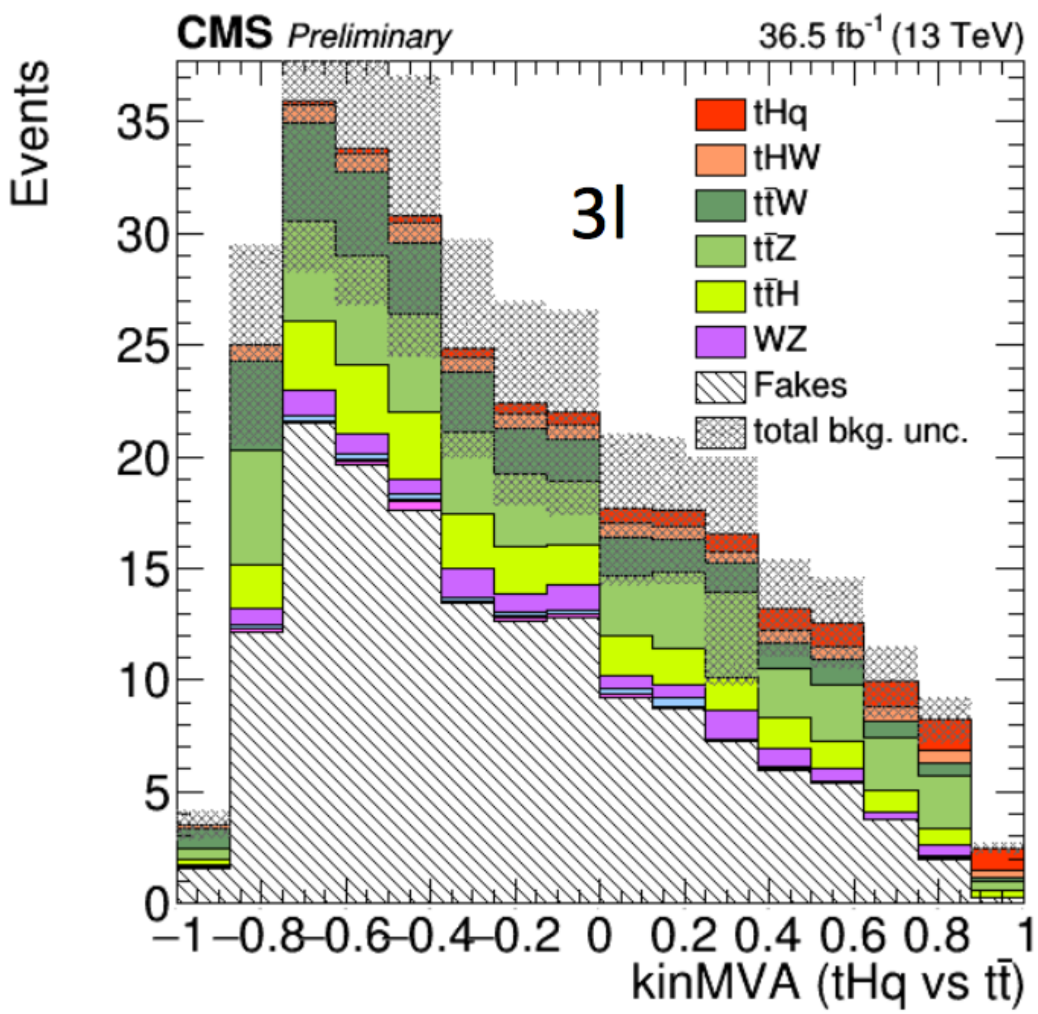
tHq+tHW, H→multileptons signal extraction

- MVA discriminator train $t\bar{t}$V against $t\bar{t}$
  - Divide the plane of BDT $t\bar{t}$ vs BDT $t\bar{t}$V into bins for signal and background
  - Shape fit the MVA binned output to extract the signal yield combine for final result

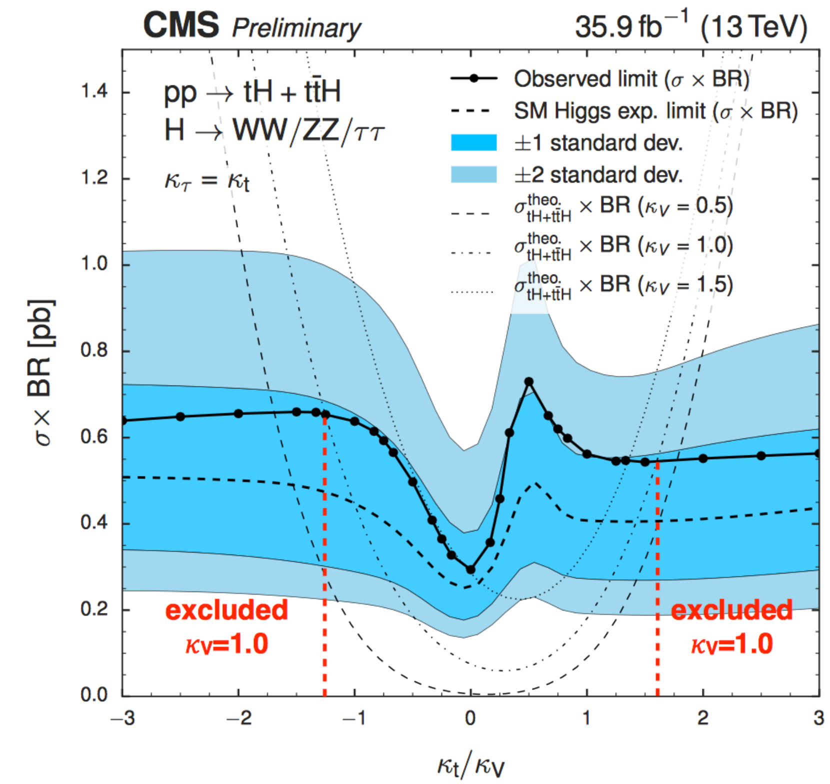


$t\bar{t}$V BDT score

$t\bar{t}$ BDT score



Expected (including SM Higgs) and observed limit on σ*BR



- Excess of about 1 σ of expected SM-like tHq+tHW+ttH signal observed
- Best-fit signal strength for SM: μ = 1.8 ± 0.3 stat. ± 0.6 syst.

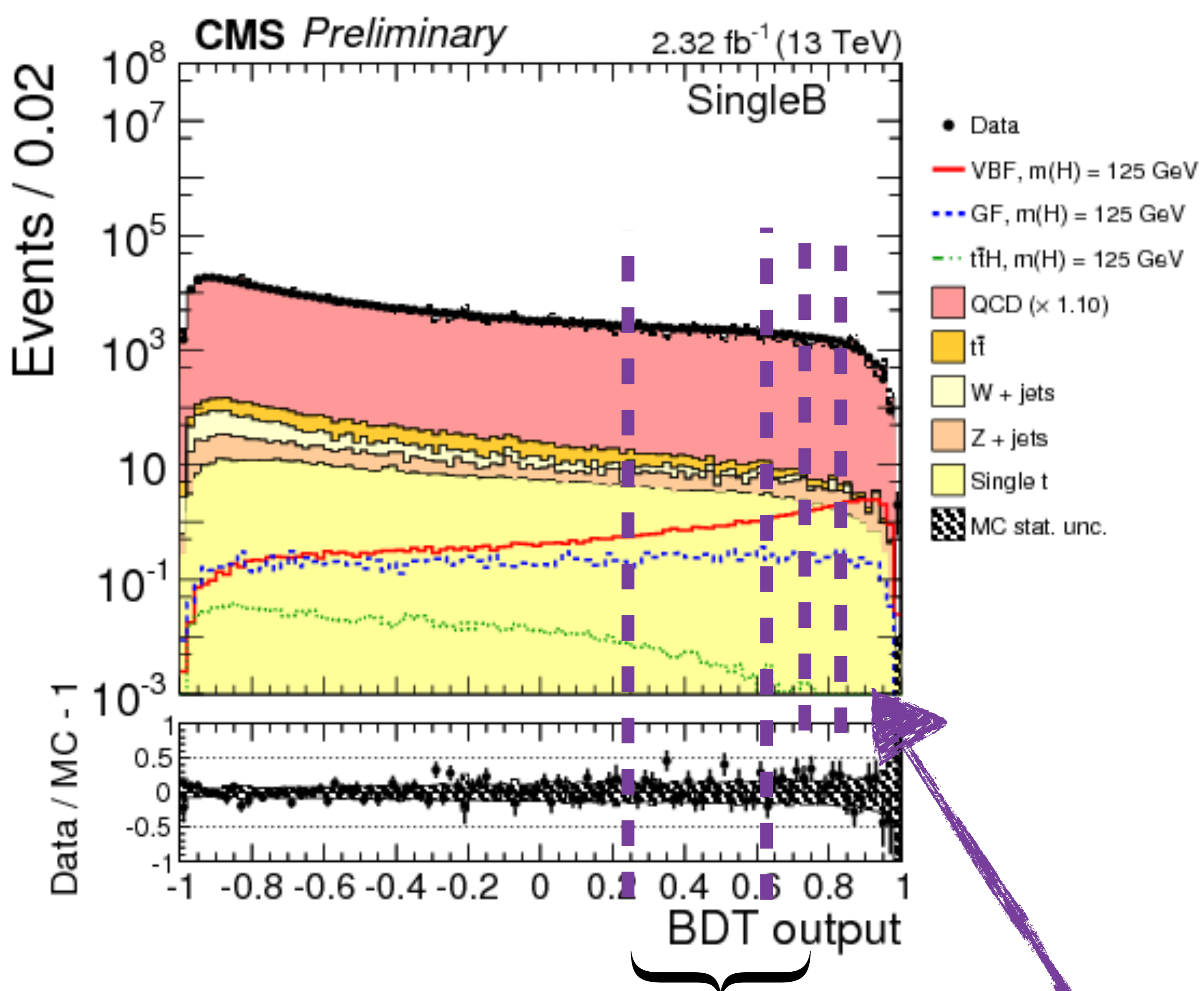
VBF, H → $b\bar{b}$

CMS-PAS-HIG-16-003

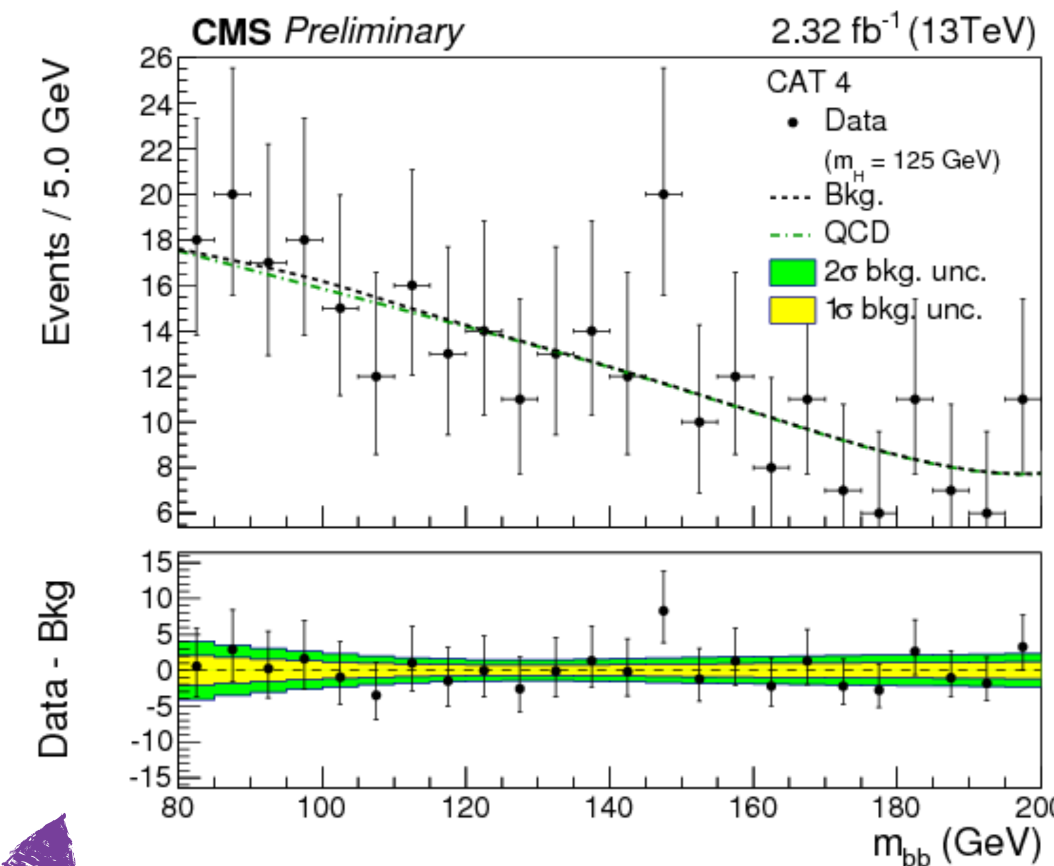
Properties of the VBF H → $b\bar{b}$ channel:

- cross section significantly larger than VH or ttH production
- very large QCD background

- 4-jet signal event topology
  1 or 2 b tag and BDT categorization



CAT 1 . . . CAT 4



- Signal extraction in simultaneous fit to m_{bb} spectrum in all categories
- Result using 2.3 fb–1 @ √s = 13 TeV:
  $\mu = -3.7^{+2.4}_{-2.5}$
- Combination with Run I (18 -19 fb^{-1} @ 8 TeV):
  $\mu = 1.3^{+1.2}_{-1.1}$

m_{bb} resolution

m_{bb} resolution significantly worsened by semi-leptonic b-decays and gluon radiation outside jet "cone"

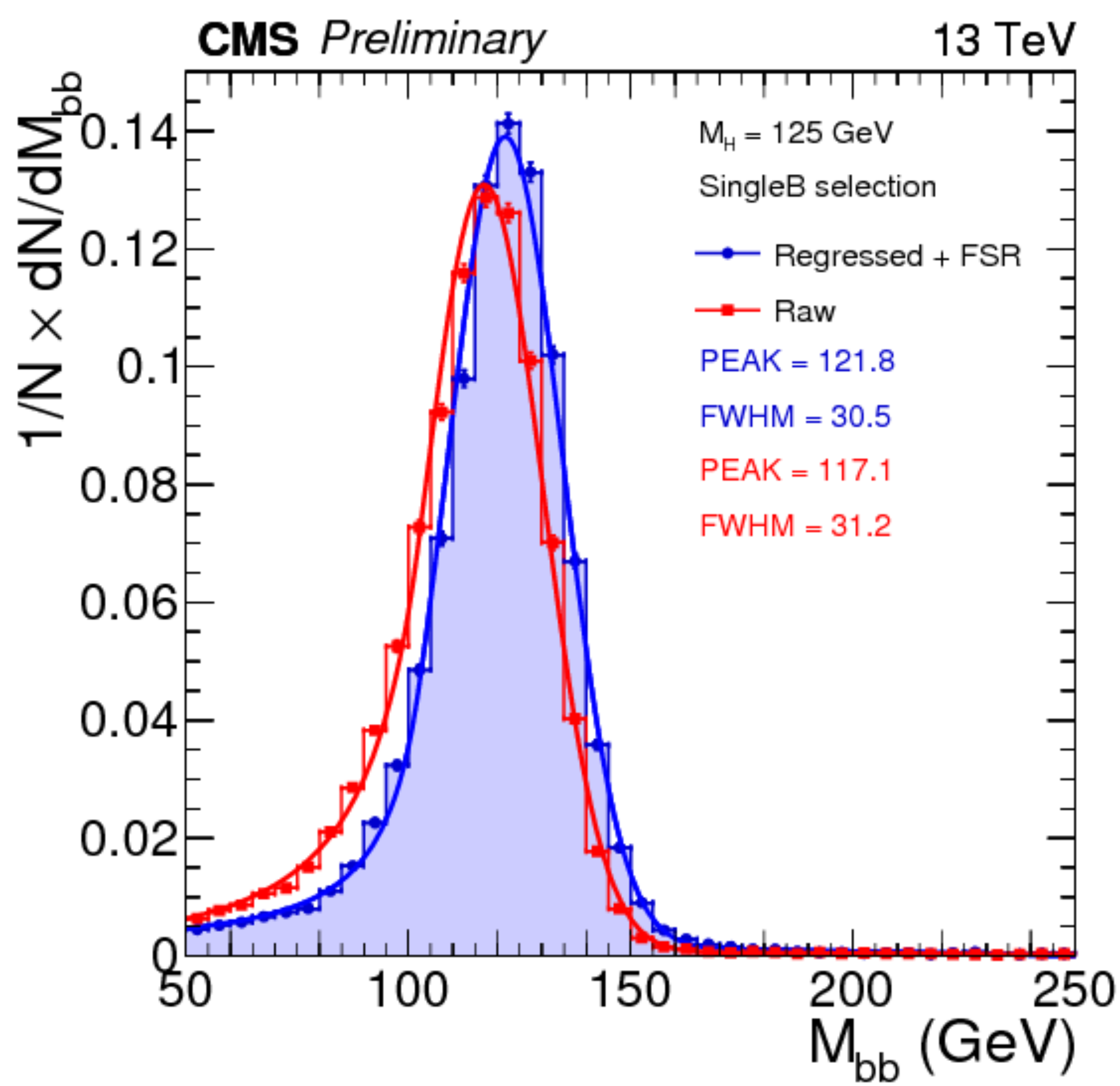
→ improve b-jet energy resolution with regression

Regression inputs

- Jet kinematic
- EM energy fraction
- Information about soft leptons in the jet
- Secondary vertex information
- Pileup

FSR correction:

Add jets within $\Delta R < 0.8$.



Summary and Outlook

- About 39 fb-1 recorded and 2-39 fb^{-1} analyzed at 13 TeV
- Run-2 sensitivity exceeded Run-1 result
- Presented searches for $H \rightarrow b\bar{b}$ using 13 TeV data performed
- Probes Top-Higgs Yukawa coupling directly accessible through associated t(t)H production
  - Important for understanding loop contributions
- Studies involve complex final states with leptons, jets, photons etc.
- Multiple analysis channels contribute sensitivity
- No deviation from the SM prediction observed

Outlook

- Not all analyses updated to all available data
  - updates in the very near future
- Excellent prospects for establishing $t\bar{t}$ H, VBF signal with complete Run-2 data set
- Continuous improvement of the signal extraction methods & modeling of $t\bar{t}$ + (b-)jets (indispensable collaboration with theory & MC experts)

References

Publication

- CMS-PAS-HIG-16-038
  - http://cms-results.web.cern.ch/cms-results/public-results/preliminary-results/HIG-16-038/index.html
- CMS-PAS-HIG-16-003
  - http://cms-results.web.cern.ch/cms-results/public-results/preliminary-results/HIG-16-003/index.html
- CMS-PAS-HIG-16-019
  - http://cms-results.web.cern.ch/cms-results/public-results/preliminary-results/HIG-16-019/index.html
- CMS-PAS-HIG-17-005
  - http://cms.cern.ch/iCMS/analysisadmin/cadilines?id=1868&ancode=HIG-17-005&tp=an&line=HIG-17-005

Conference talks

- https://indico.cern.ch/event/505065/contributions/2166376/attachments/1339192/2019944/LHC-Day-Split_2016_ChristianJCC.pdf
- https://indico.in2p3.fr/event/13763/session/0/contribution/79/material/slides/0.pdf
- https://indico.cern.ch/event/477407/contributions/2200113/subcontributions/198573/attachments/1369935/2077138/ttH_Hbb_CMS_Kasieczka.pdf

Backup

Analysis targeting ttH production

Matrix Element method for ttH(bb) vs ttbb

- Signal extraction via Matrix Element Methods (MEM):
  - Event-by-event discriminator build upon matrix elements, combined with reconstruction-level information

Numerical integration — Momentum conservation — Resolution function (allow ISR)

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

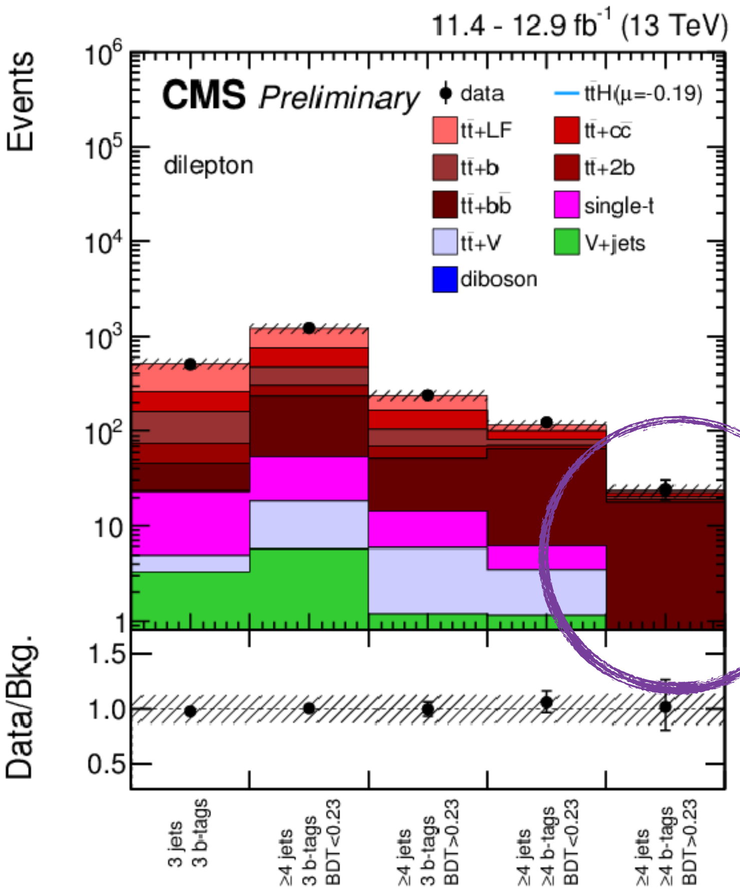
Parton density functions — LO scattering amplitude (Open Loops) — Detector transfer function

- Construct per-event signal/background probabilities using full kinematic information in an analytic approach

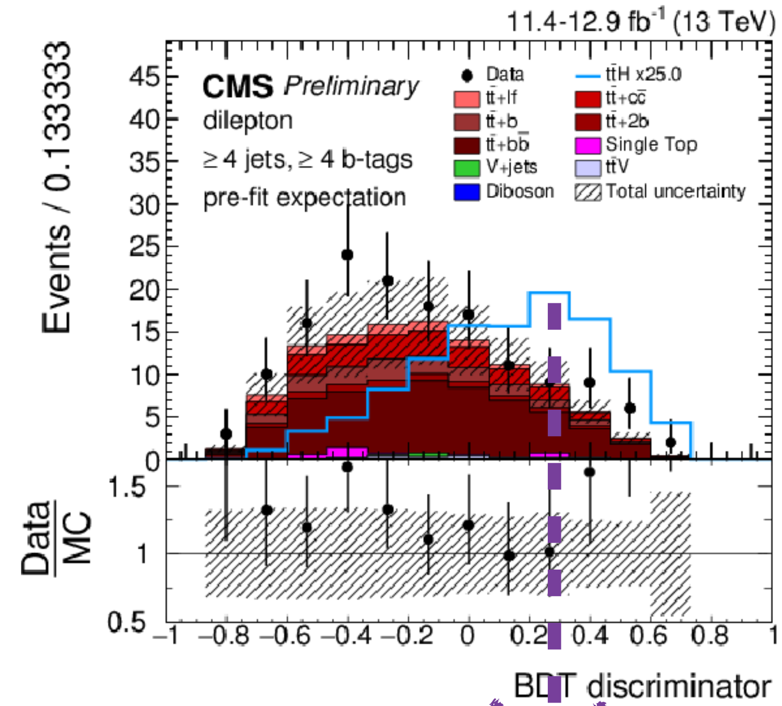
$$P_{s/b} = \frac{w(\vec{y}|t\bar{t}H)}{w(\vec{y}|t\bar{t}H) + k_{s/b} w(\vec{y}|t\bar{t}+b\bar{b})}$$

- tt+bb take as background hypothesis, permuting overall jet assignments
- Works best for final states with multiple reconstructed jets

Combination of BDT & MEM

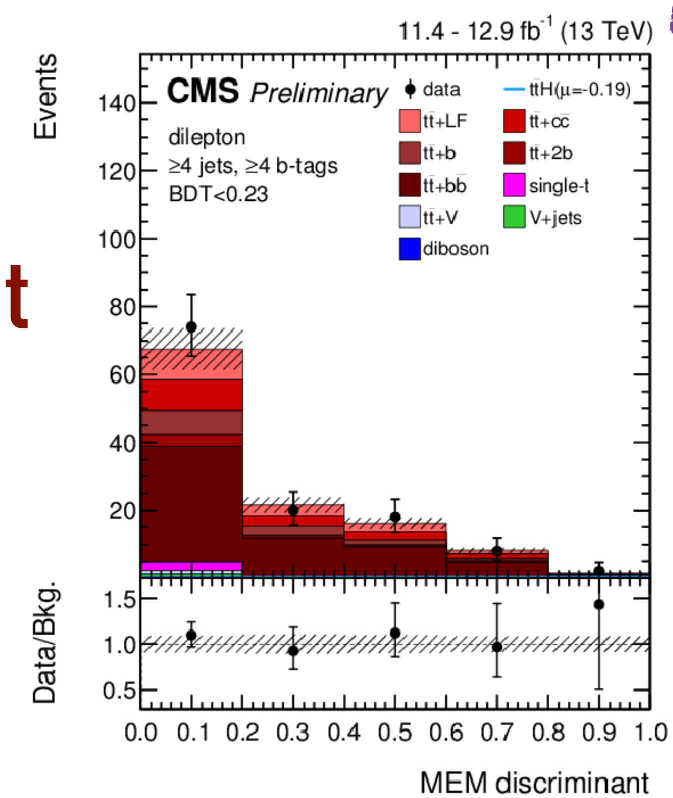


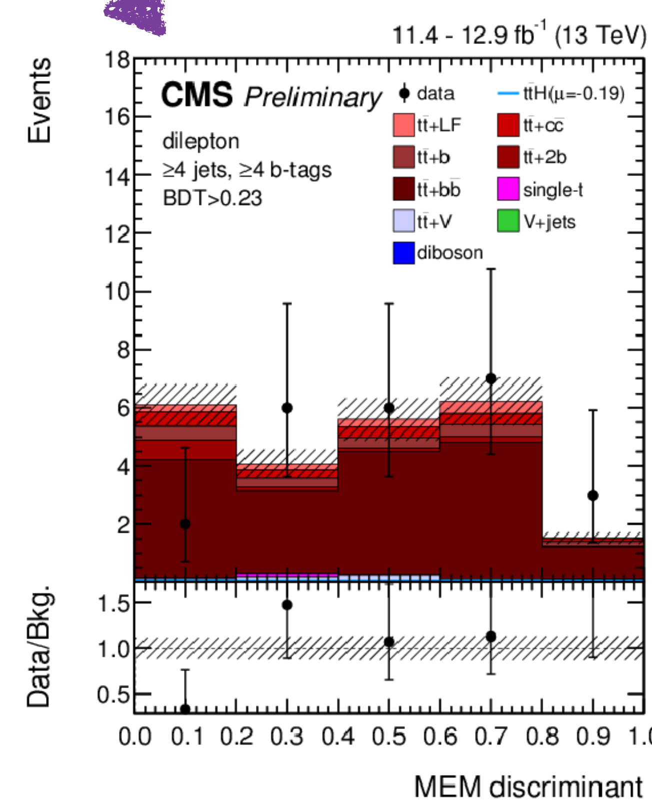
Split category at the median of ttH BDT output



low purity

high purity





MEM as final discriminant in the split categories

Particle Swarm Optimization

- See: Particle swarm optimization (J. Kennedy, R. Eberhart)
  Proceedings of the IEEE International Conference on Neural Networks, 1995.
- Optimization algorithm
- Different BDT setting (i.e. tree structure and variables) form the search-space
- A specific setting corresponds to one point in this search space
- Algorithm:
  - Create swarm of candidate BDTs
  - Each BDT is initialized with a random set of input variables and position in parameter-space
  - Do N iterations
    - Repeatedly train/test at current position.
    - Vary input variables to maximize ROC while KS > threshold
    - Then the BDTs move to new positions, based on their own and