Karlsruher Institut für Technologie

Searches for $t\bar{t}H$ and tH with $H \rightarrow b\bar{b}$

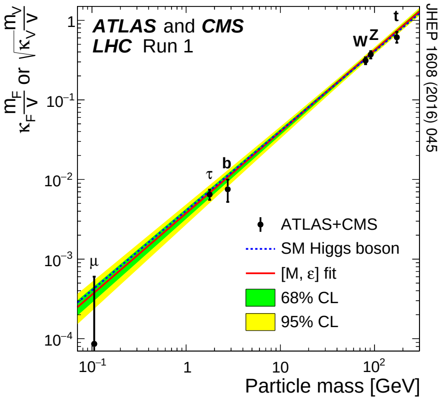
The 9th International Workshop on Top Quark Physics TOP 2016, Olomouc, Czech Republic

Matthias Schröder
on behalf of the ATLAS and CMS Collaborations | September 21, 2016

INSTITUT FÜR EXPERIMENTELLE KERNPHYSIK (IEKP)

A Higgs Boson at the LHC

- What are the properties of the Higgs boson?
- Coupling to top quark especially interesting: large
  - Important in loop processes
  - Contraints on BSM physics
- Current value dominated by indirect, model-dependent measurements of gluon-fusion production and $\gamma\gamma$-decay channel



$t\bar{t}$- and t-associated H production: **direct access to top-Higgs coupling**

The urban dictionary's definition of 'TTH':

TTH

"Trying TOO hard".

The urban dictionary's definition of ‘TTH’:

TTH

"Trying ~~TOO~~ hard".

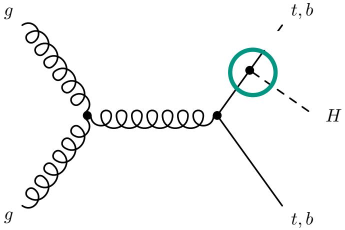
Outline

Latest results on direct searches in H $\rightarrow$ b$\bar{\text{b}}$ channel
by ATLAS and CMS with 13 TeV data:

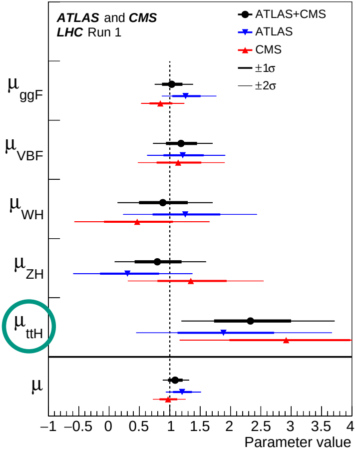
1. Search for t$\bar{\text{t}}$H(b$\bar{\text{b}}$) production
   - Analysis strategy
   - Search with 2.7 fb^{-1} at CMS
   - Search with 13.2 fb^{-1} at ATLAS (NEW)
2. Search for single-tH(b$\bar{\text{b}}$) production
   - Search with 2.7 fb^{-1} at CMS (NEW)

- Also at this conference:
  - *Searches for t$\bar{\text{t}}$H and tH with H $\rightarrow$ $\gamma\gamma$*, Diane Cinca
  - *Searches for t$\bar{\text{t}}$H and tH with H $\rightarrow$ leptons*, Charles Nicholas Mueller

$t\bar{t}H$ Production at the LHC



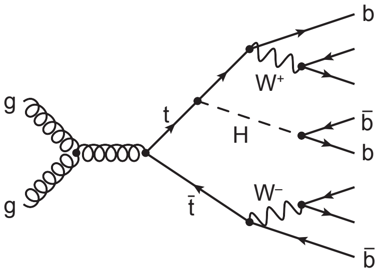
- Run-I: mild excess $\mu_{t\bar{t}H} = 2.3^{+0.7}_{-0.6}$
- Run-II: factor 4 increase of $\sigma_{t\bar{t}H}$
  - 0.13 pb (8 TeV) $\rightarrow$ 0.5 pb (13 TeV)
  - ggF production: factor 2

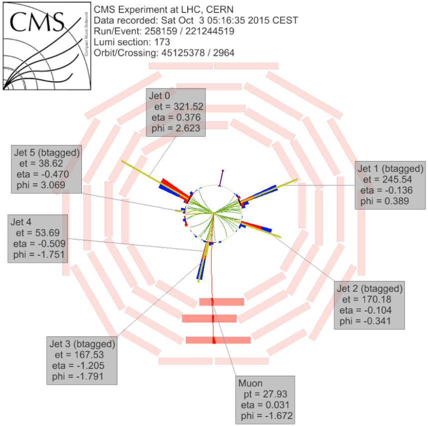


JHEP 1608 (2016) 045

Relatively weak constraints on $t\bar{t}H$ production so far
Expect leap in sensitivity in Run-II

$H \rightarrow b\bar{b}$ Final State





single-lepton-channel candidate event

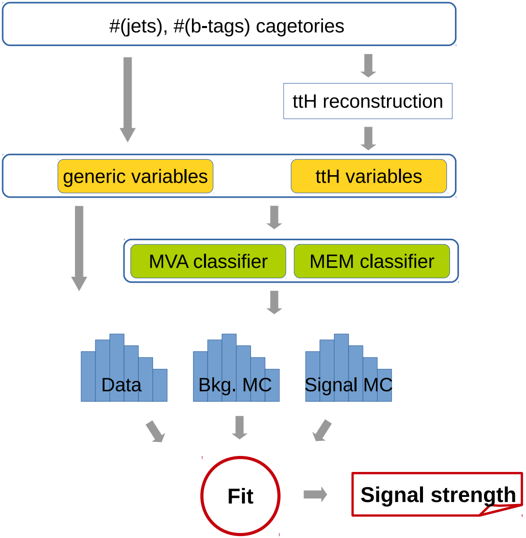
CMS-PAS-HIG-16-004 (suppl. material)

- Largest branching ratio but challenging final state
  - Huge combinatorics in event reconstruction
  - Difficult $t\bar{t} + b\bar{b}$ background
- At 13 TeV, both di- and single-lepton $t\bar{t}$ final states considered[1]
  - Cleaner signature, suppression of QCD-multijet background

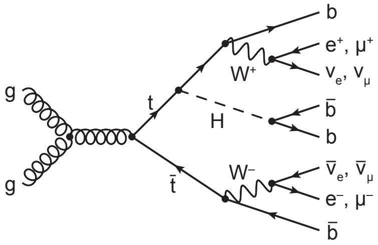
[1] All-hadronic final state analyzed at 8 TeV, e. g. ATLAS J. High Energ. Phys. (2016) 160

Analysis Strategy

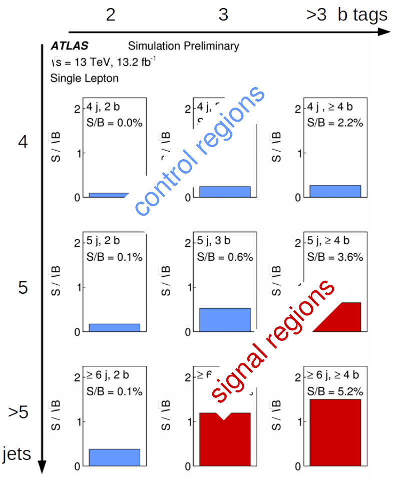
Similar approaches by ATLAS and CMS



Event Selection

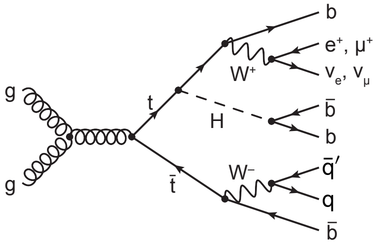


- **Dilepton** channel
  - $= 2$ opposite-charge leptons
  - ≥ 3 jets, ≥ 2 b-tagged jets
- Single-lepton
  - $= 1$ lepton
  - ≥ 4 jets, ≥ 2 b-tagged jets
- Categories in #(jets) and #(b tags)

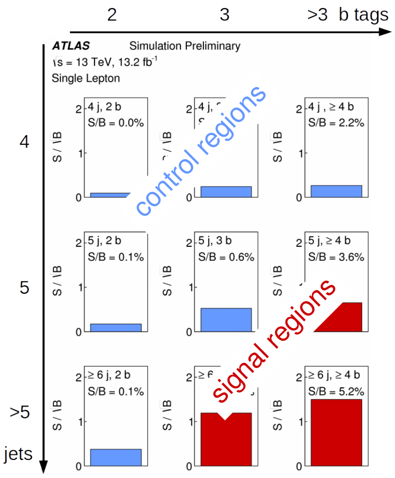


single-lepton example

Event Selection

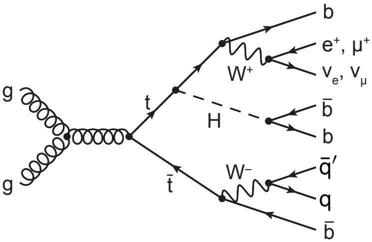


- Dilepton channel
  - $=2$ opposite-charge leptons
  - ≥ 3 jets, ≥ 2 b-tagged jets
- **Single-lepton**
  - $=1$ lepton
  - ≥ 4 jets, ≥ 2 b-tagged jets
- Categories in #(jets) and #(b tags)

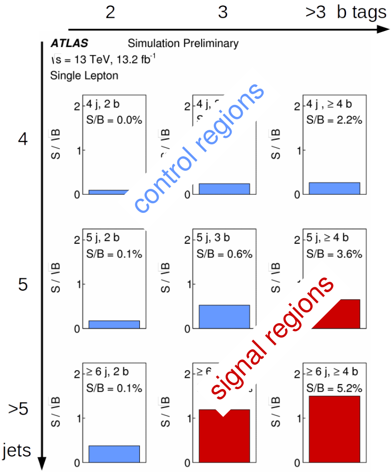


single-lepton example

Event Selection



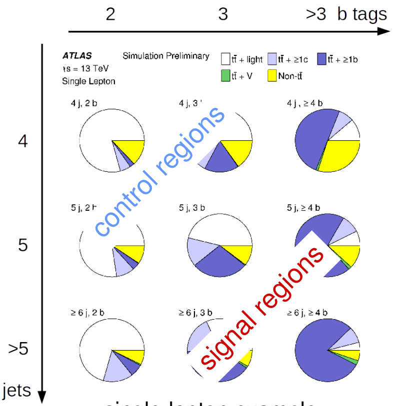
- Dilepton channel
  - $=2$ opposite-charge leptons
  - ≥ 3 jets, ≥ 2 b-tagged jets
- Single-lepton
  - $=1$ lepton
  - ≥ 4 jets, ≥ 2 b-tagged jets
- **Categories in #(jets) and #(b tags)**



single-lepton example

Background

- Dominating background processes
  - $t\bar{t} + \geq 1$ b jets
  - $t\bar{t} + \geq 1$ c jets
  - $t\bar{t}$ + light jets
- Modelled using simulation
  - $t\bar{t}$ + HF background challenging to separate + difficult to model
  - Only relatively weak constraints from data
  - Dominant uncertainty
- Different composition in categories
  - Simultaneous fit helps to constrain processes and reduce uncertainties



single-lepton example

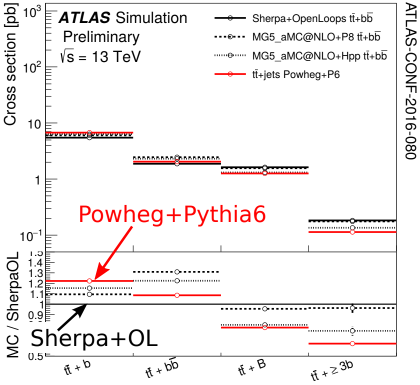
Background Modelling

Approach by CMS

- **Powheg+Pythia8**, normalized to **NNLO** prediction
- Separate templates for $t\bar{t}+b$, $t\bar{t}+b\bar{b}$, $t\bar{t}+2b$, $t\bar{t}+c\bar{c}$, $t\bar{t}+LF$
  - **50% rate uncertainty per $t\bar{t}+$ jets process**, uncorrelated in final fit
  - Add. sources include parton shower, hadronisation, PDF, ISR/FSR

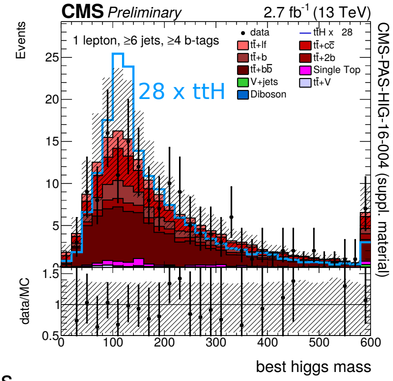
Approach by ATLAS

- **Powheg+Pythia6**, $p_T(t)$ & $p_T(t\bar{t})$ corrected to **NNLO**
- **$t\bar{t} \geq 1b$** processes corrected to **Sherpa+OpenLoops NLO 4-flavour-scheme** calculation
- Normalization of **$t\bar{t}+\geq 1b/c$ freely floating** in final fit
- Uncertainties include choice of generator, parton shower and hadronisation, PDF, ISR/FSR



Signal and Background Separation

- Single variable not very sensitive
  - Jet energy resolution
  - Combinatorics in jet assignment
- Different choices and combinations of more advanced classifiers



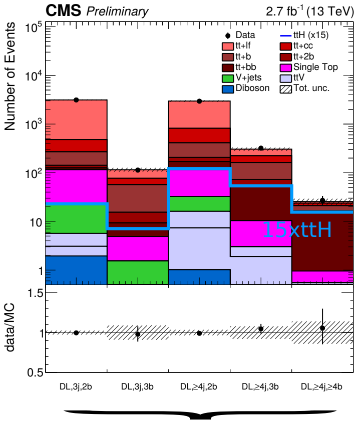
1. **Boosted-Decision-Trees** (BDTs) and Neural Networks (NNs)
   - Combination of various input variables, trained per category
   - Separation against all $t\bar{t}$ +X processes
   - Used in signal vs. background separation and in event reconstruction
2. **Matrix-Element-Method** (MEM) classifiers
   - Likelihood of event kinematics under signal or background hypothesis
   - Particularly powerful against difficult $t\bar{t} + b\bar{b}$ background
   - Used in signal vs. background separation

Search with 2.7 fb^{-1} of 2015 Data at CMS

First $t\bar{t}H$ result with 13 TeV data CMS-PAS-HIG-16-004

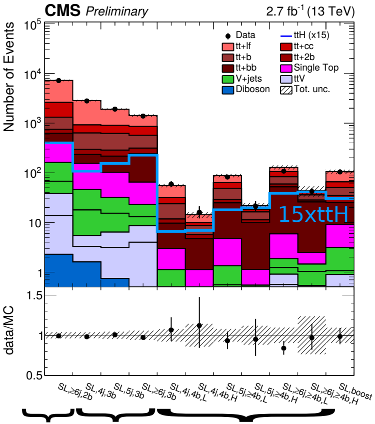
see also poster by Predrag Cirkovic

dilepton channel



BDT

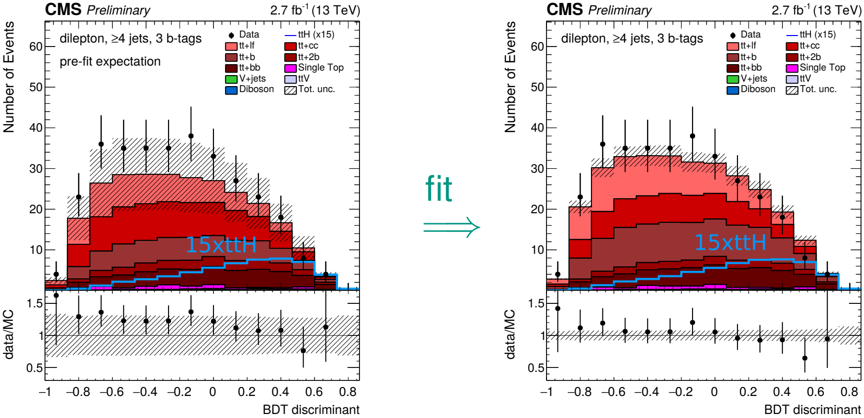
single-lepton channel



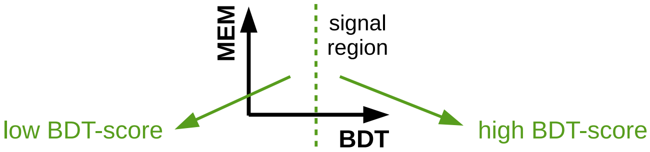
BDT | BDT (incl. MEM) | 2D (BDT,MEM)

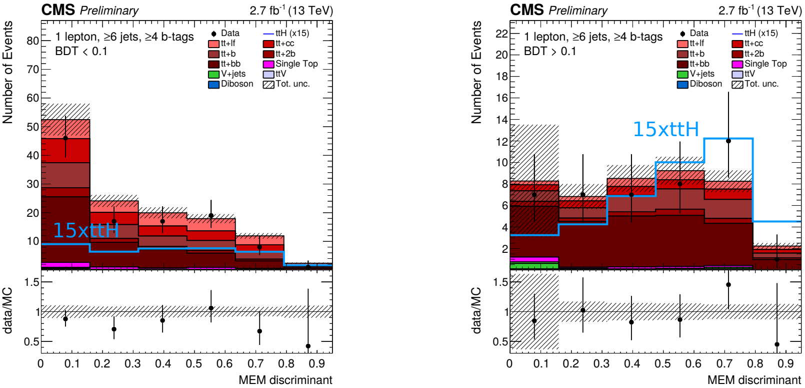
Signal Separation: BDT

- Algorithmic optimization of BDT configuration and choice of input variables
  - b-tagging output, angular variables, and MEM among most sensitive



Signal Separation: 2D (BDT,MEM)

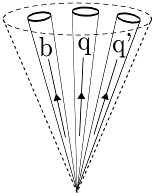




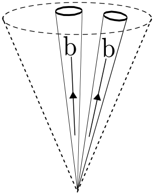
post-fit distributions

‘Boosted’ Category

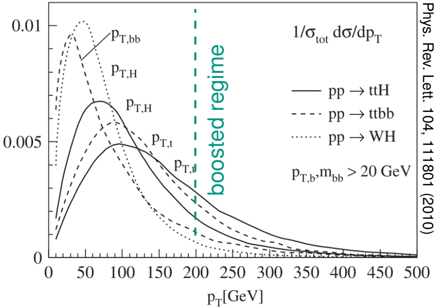
- Events with boosted H and t_{had}: decay products merged in fat jet
- Fat-jet substructure analysis to reconstruct event topology



boosted t_{had}

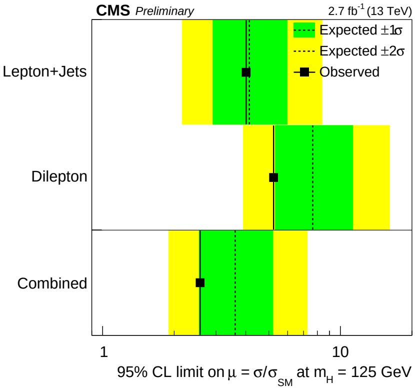


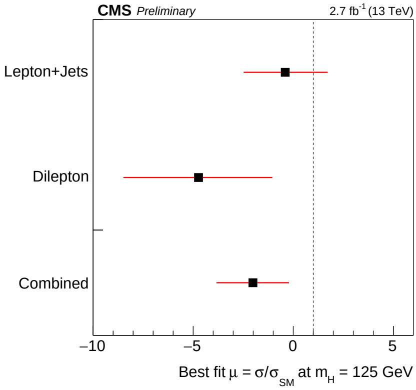
boosted H



Higher p_T in $t\bar{t}H$: separation from background (in particular $t\bar{t} + b\bar{b}$)
Reduced combinatorics: better reconstruction efficiency

Results





Upper limit on $\sigma_{\mathrm{t\bar{t}H}}$ of 2.6 (obs.) and 3.6 (exp.) $\times$ SM expectation

Best-fit signal strength of $\mu = -2.0^{+1.8}_{-1.8}$

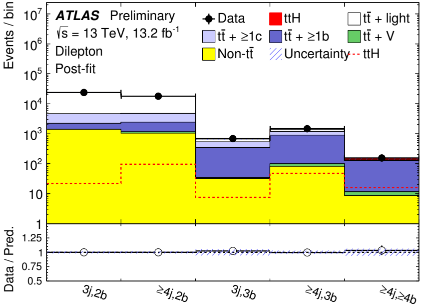
Search with 13.2 fb^{-1} of 2015/16 Data at ATLAS

Karlsruher Institut für Technologie

NEW: first $t\bar{t}H$ result with 2016 data ATLAS-CONF-2016-080

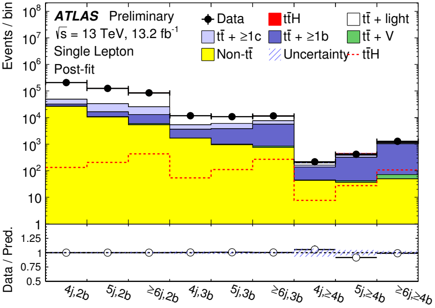
see also poster by Nedaa Alexandra Asbah

dilepton channel



NN BDTs BDTs

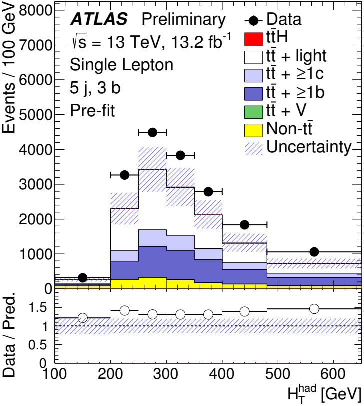
single-lepton channel



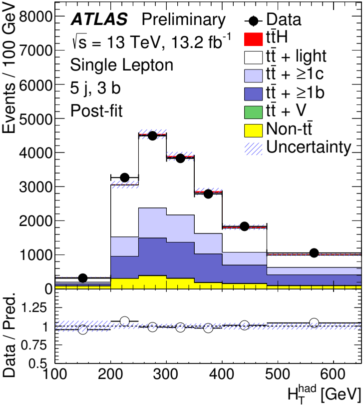
BDTs BDTs BDTs

Final Discriminant in Control Regions

- Scalar sum of all jet (and lepton) p_T in single- (di-) lepton channels



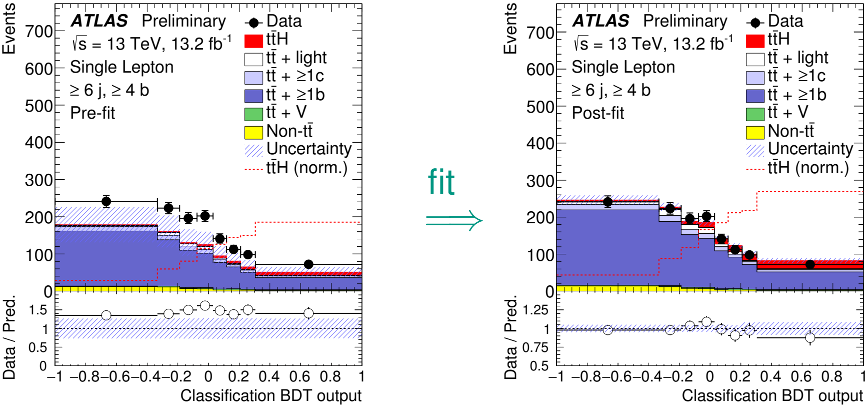
fit $\Longrightarrow$



H_T templates used in final fit to constrain systematic uncertainties

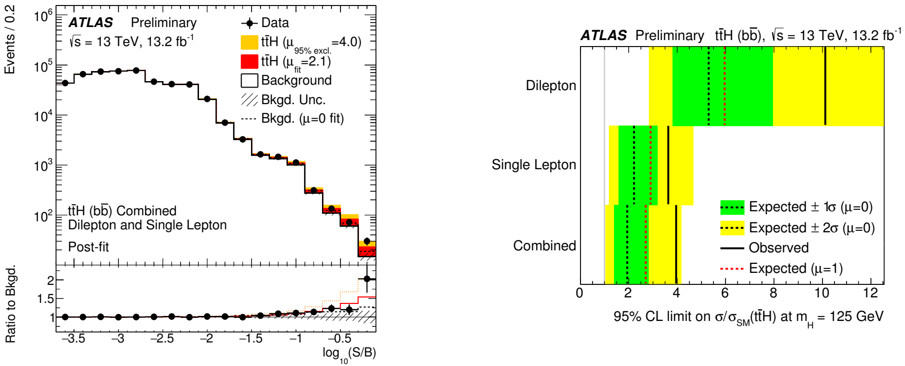
Final Discriminant in Signal Regions

- Two-stage multivariate approach
  1. Reconstruction BDT: match jets to partons from H and t decay
  2. Classification MVA: classify event as more signal- or background-like



Classification output used in final fit to extract signal

Results

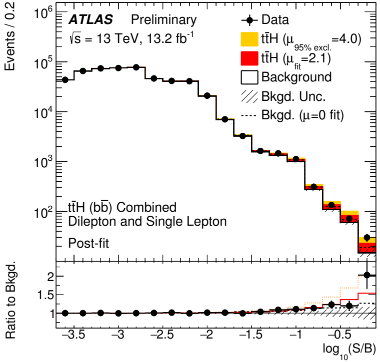


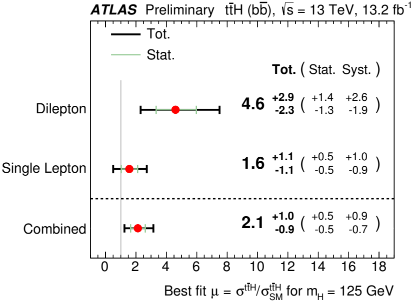
Data consistent with both background-only and SM $t\bar{t}H$ expectation

Upper limit[2] on $\sigma_{t\bar{t}H}$ of 4.0 (obs.) and 1.9 (exp.) $\times$ SM

[2] Under null hypothesis of no ttH production

Results

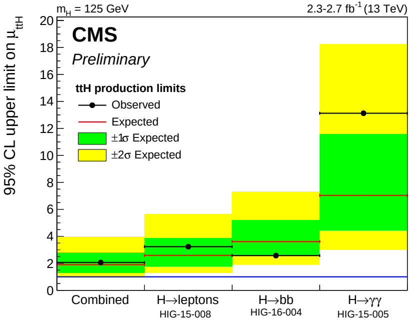


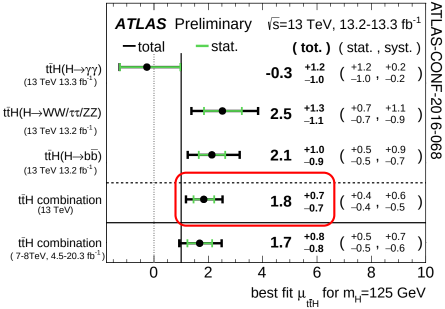


Data consistent with both background-only and SM $t\bar{t}H$ expectation

Best-fit signal strength of $\mu = 2.1^{+1.0}_{-0.9}$

Status on $t\bar{t}H$





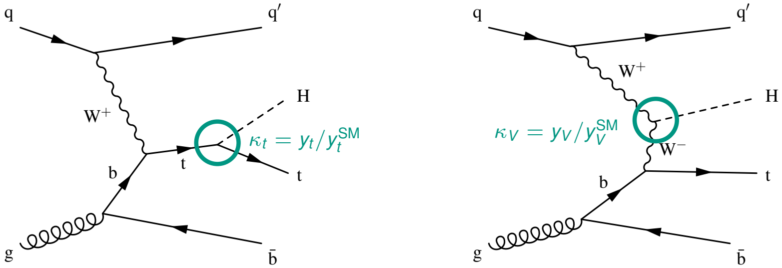
ATLAS and CMS are closing in on $t\bar{t}H$

Run-I sensitivity surpassed in ATLAS combination

$H \to b\bar{b}$ channel among most sensitive

tH Production at the LHC

- Dominating contributions depend on κ_t and κ_V
  - Both t-channel and tW-channel production considered

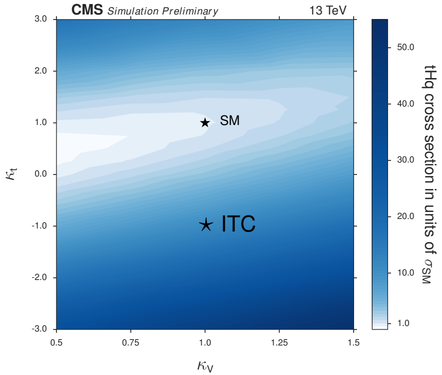


- Interference: $\boxed{\mathcal{A} \propto (\kappa_V - \kappa_t)} \rightarrow \sigma \propto \kappa_V^2 + \kappa_t^2 - 2\kappa_V\kappa_t$
- Destructive interference in SM → small cross section of $\sim 90\,\text{fb}$

tH production sensitive to magnitude and sign of top-Yukawa coupling

NB: indirect sensitivity by $t\bar{t}H(\gamma\gamma)$ process, e. g. ATLAS Phys.Lett. B740 (2015) 222-242

tH Production at the LHC



- Strong dependence of cross section on κ_t
- E.g. 'inverted top coupling' $\kappa_t = -1$ scenario: $\sim 10 \times \sigma_{\mathrm{SM}}(\mathrm{tH})$
  - Not excluded if BSM contributions allowed to loops

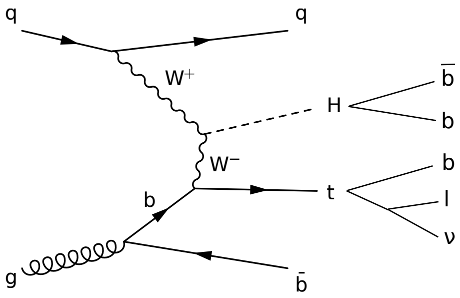
Perform direct search for tH production to constrain sign of κ_t

Search for tH($b\bar{b}$) Production at CMS

NEW: first tH result with 13 TeV data CMS-PAS-HIG-16-019

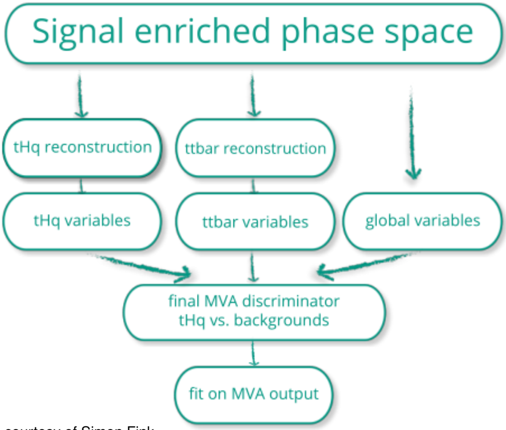
see also posters by Kevin Flöh and Denise Müller

- $H \to b\bar{b}$ final state with leptonically decaying W from t decay
- Dominant background from semi-leptonic $t\bar{t}$ events



- $2.3\,\text{fb}^{-1}$ of data
- **3 tag** signal region
  - $=$ 1 muon or electron
  - $=$ 3 b-tagged jets
  - $\geq$ 4 jets
- **4 tag** signal region
  - $=$ 1 muon or electron
  - $=$ 4 b-tagged jets
  - $\geq$ 5 jets

Analysis Strategy



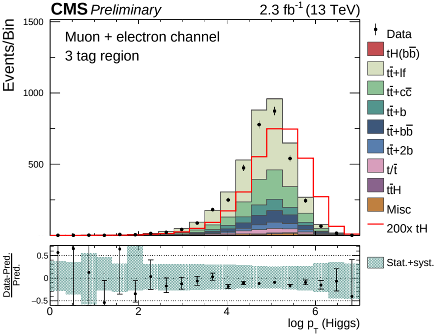
courtesy of Simon Fink

Event Reconstruction

- Event reconstruction under both signal and $t\bar{t}$-background hypothesis
  - BDTs taking into account kinematic and b-tagging information
  - Signal kinematics depend on (κ_t, κ_V): different trainings per point
- Separating variables based on specific event reconstruction

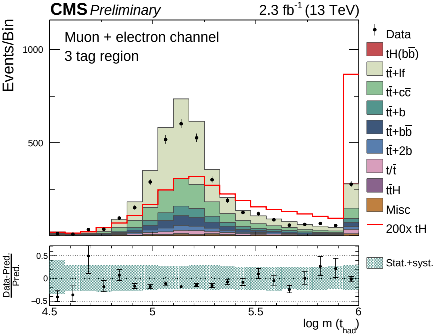
signal reconstruction

e. g. $p_T(H)$



$t\bar{t}$-background reconstruction

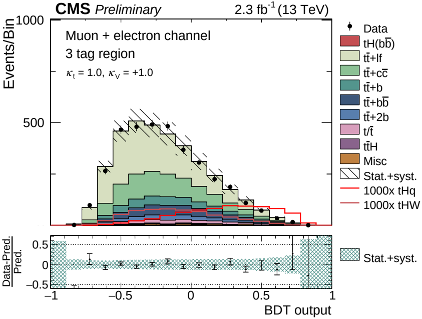
e. g. $m(t_{had})$



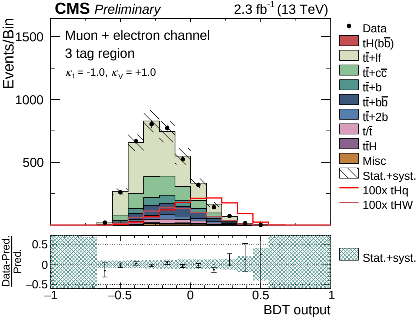
Signal Extraction

- Final signal-vs-background classification by additional BDT, using
  - variables based on signal reconstruction
  - variables based on $t\bar{t}$-background reconstruction
  - reconstruction-independent variables
- Different trainings per (κ_t, κ_V) point

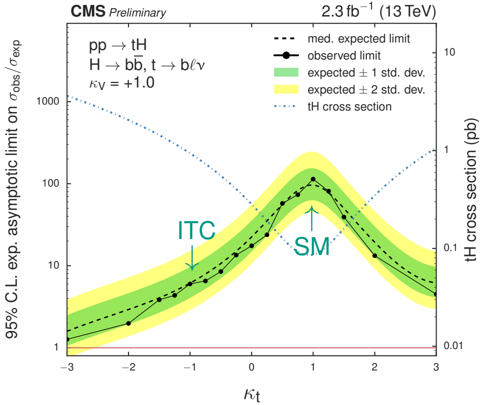
SM hypothesis



ITC hypothesis $\kappa_t = -1$



Results



- Excluding SM tH production above 113.7 (obs.) and 98.6 (exp.) $\times\sigma_{SM}$
- Exclusion at $\kappa_t = -1$: 6.0 (obs.) and 6.4 (exp.) $\times\sigma_{ITC}$

Close to Run-I sensitivity with 1/10 of the statistics
Stay tuned for 2016 data results

Summary & Conclusions

- $t\bar{t}H$ and tH production allow direct measurement of top-Higgs Yukawa coupling
- $H \to b\bar{b}$ final state offers large branching ratio but challenging
  - ATLAS and CMS have developed advanced analysis techniques including multivariate methods and are analyzing the 13 TeV data
  - $H \to b\bar{b}$ final state among most sensitive
- Reaching Run-I sensitivity with fraction of data

Outlook

TTH

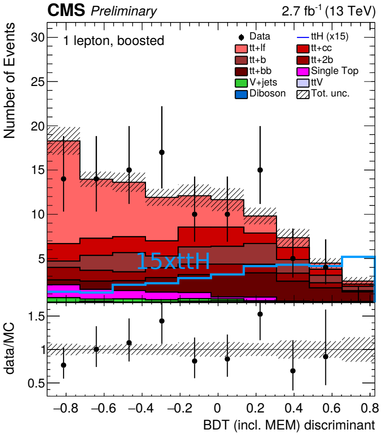
We will be "Trying ~~TOO~~ hard".

More data ahead — stay tuned!

Additional Material

$t\bar{t}H$ (CMS): 'Boosted' Category

CMS-PAS-HIG-16-004



Final discrimination via BDT
Takes into account information of boosted-object reconstruction

$t\bar{t}H$ (CMS): Systematic Uncertainties

KIT Karlsruher Institut für Technologie

CMS-PAS-HIG-16-004

Source	Type	Remarks
Luminosity	rate	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 ($t\bar{t}H$)	rate	Scale uncertainty of NLO $t\bar{t}H$ prediction
QCD scale ($t\bar{t}$)	rate	Scale uncertainty of NLO $t\bar{t}$ prediction
QCD scale ($t\bar{t}$+hf)	rate	Additional scale uncertainty of NLO $t\bar{t}$+hf predictions
QCD scale (t)	rate	Scale uncertainty of NLO single t prediction
QCD scale (V)	rate	Scale uncertainty of NNLO W and Z prediction
QCD scale (VV)	rate	Scale uncertainty of NLO diboson prediction
pdf (gg)	rate	Pdf uncertainty for gg initiated processes except $t\bar{t}H$
pdf (gg $t\bar{t}H$)	rate	Pdf uncertainty for $t\bar{t}H$
pdf ($q\bar{q}$)	rate	PDF uncertainty of $q\bar{q}$ initiated processes ($t\bar{t}$ W, W, Z)
pdf (qg)	rate	PDF uncertainty of qg initiated processes (single t)
Q^2 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

$t\bar{t}H$ (ATLAS): Systematic Uncertainties

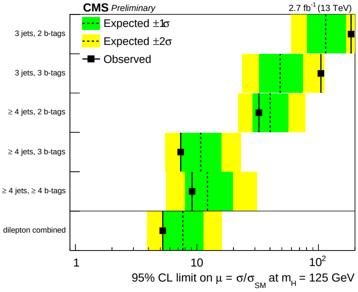
Karlsruher Institut für Technologie

ATLAS-CONF-2016-080

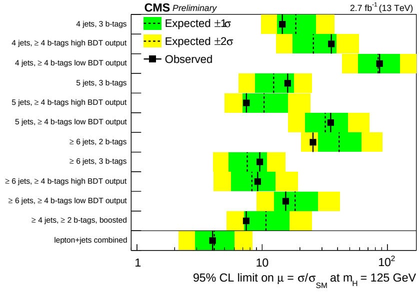
Systematic source	How evaluated	$t\bar{t}$ categories
$t\bar{t}$ cross-section	$\pm 6\%$	All, correlated
NLO generator (*residual*)	Powheg-Box + Herwig++ vs. MG5_aMC + Herwig++	All, uncorrelated
Radiation (*residual*)	Variations of μ_{R}, μ_{F}, and *hdamp*	All, uncorrelated
PS & hadronisation (*residual*)	Powheg-Box + Pythia 6 vs. Powheg-Box + Herwig++	All, uncorrelated
NNLO top & $t\bar{t}$ p_{T}	Maximum variation from any NLO prediction	$t\bar{t}+\geq 1c$, $t\bar{t}$ +light, uncorr.
$t\bar{t}+b\bar{b}$ NLO generator *reweighting*	SherpaOL vs. MG5_aMC + Pythia8	$t\bar{t}+\geq 1b$
$t\bar{t}+b\bar{b}$ PS & hadronis. *reweighting*	MG5_aMC + Pythia8 vs. MG5_aMC + Herwig++	$t\bar{t}+\geq 1b$
$t\bar{t}+b\bar{b}$ renorm. scale *reweighting*	Up or down a by factor of two	$t\bar{t}+\geq 1b$
$t\bar{t}+b\bar{b}$ resumm. scale *reweighting*	Vary μ_{Q} from $H_{\mathrm{T}}/2$ to μ_{CMMPS}	$t\bar{t}+\geq 1b$
$t\bar{t}+b\bar{b}$ global scales *reweighting*	Set μ_{Q}, μ_{R}, and μ_{F} to μ_{CMMPS}	$t\bar{t}+\geq 1b$
$t\bar{t}+b\bar{b}$ shower recoil *reweighting*	Alternative model scheme	$t\bar{t}+\geq 1b$
$t\bar{t}+b\bar{b}$ PDF *reweighting*	CT10 vs. MSTW or NNPDF	$t\bar{t}+\geq 1b$
$t\bar{t}+b\bar{b}$ MPI	Up or down by 50%	$t\bar{t}+\geq 1b$
$t\bar{t}+b\bar{b}$ FSR	Radiation variation samples	$t\bar{t}+\geq 1b$
$t\bar{t}+c\bar{c}$ ME calculation	MG5_aMC + Herwig++ inclusive vs. ME prediction	$t\bar{t}+\geq 1c$

$t\bar{t}H$ (CMS): Results per Channel

dilepton channel



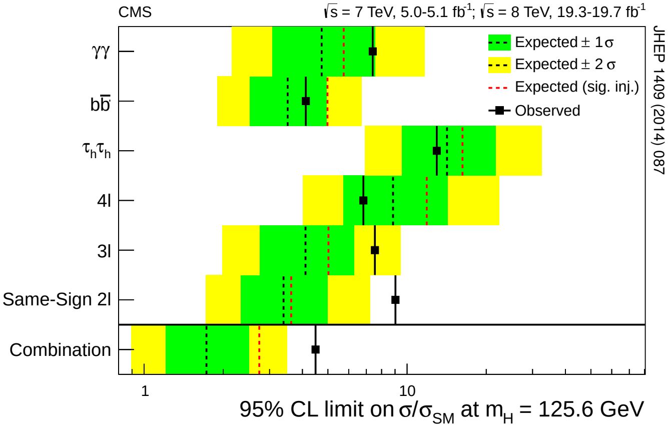
single-lepton channel



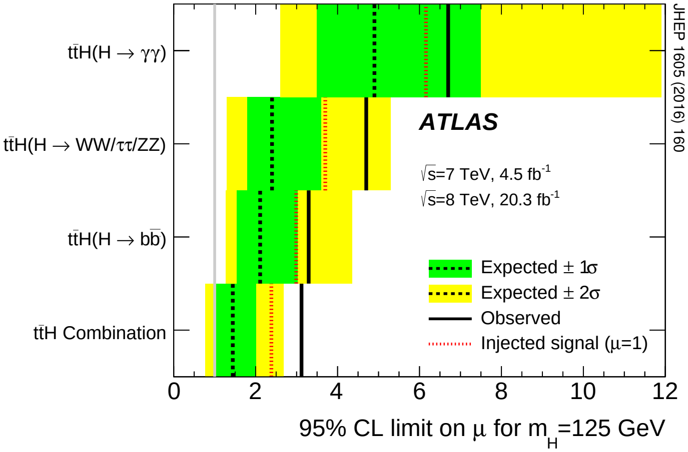
- Highest sensitivity in $\geq$ 3 b tags and boosted categories
  - Other categories constrain uncertainties
- Uncertainty dominated by normalization and modelling of $t\bar{t}$ + HF bkg

Excluding $\sigma_{t\bar{t}H}$ above 5.2 (dilepton) and 4.0 (single-lepton) $\times$ SM expectation at 95% CL

$t\bar{t}H$ (CMS): Run-I Results

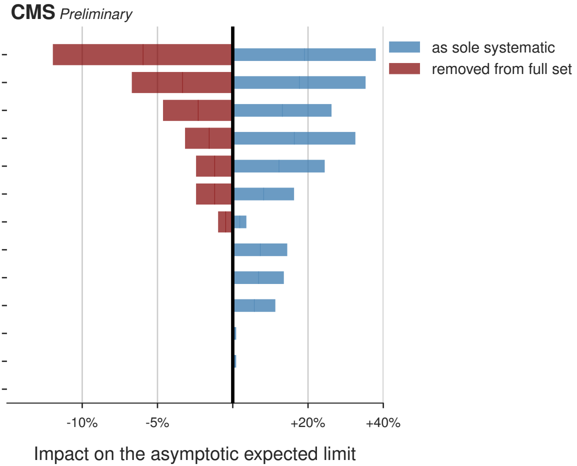


$t\bar{t}H$ (ATLAS): Run-I Results



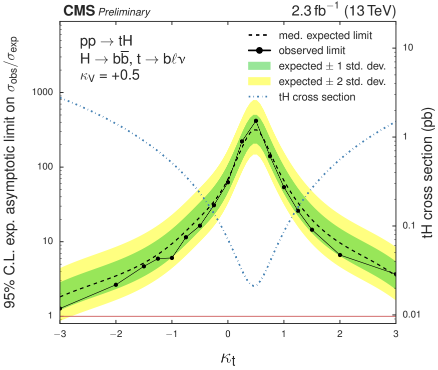
tH (CMS): Systematic Uncertainties

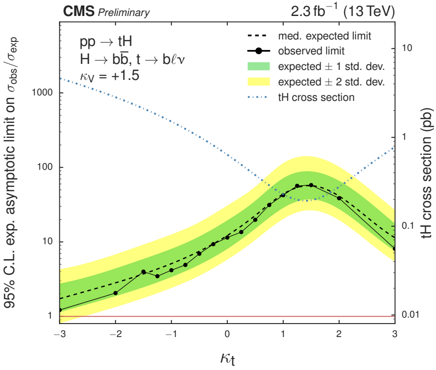
CMS-PAS-HIG-16-019



tH (CMS): Further Results

CMS-PAS-HIG-16-019





tH (CMS): Run-I Results