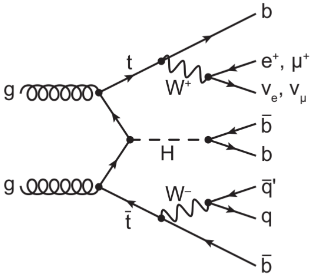
ETH zürich

Measurement of the ttH(bb) process at $\sqrt{s} = 13$ TeV with the CMS experiment

Maren Meinhard on behalf of the CMS collaboration

01 October 2019 - Higgs Couplings 2019, Oxford, UK

Introduction I

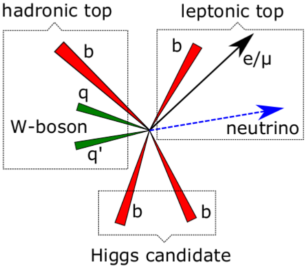


- Top Yukawa coupling can only be directly tested in the $t\bar{t}H$ process (no contributions from loops)
- Small production cross-section: $\sigma(t\bar{t}H) \sim 0.5\,\mathrm{pb}$ ($\sigma(\text{Gluon fusion}) \sim 44\,\mathrm{pb}$)
- Higgs decay to $b\bar{b}$ has the largest branching ratio (58%)

- $t\bar{t}H$, $H \rightarrow b\bar{b}$ process played a major role in the discovery of the $t\bar{t}H$ process [1] and of the Higgs decay to b quarks [2]

[1] Phys. Rev. Lett. 120 (2018) 231801 (CMS), Phys. Lett. B784 (2018) 173 (ATLAS)
[2] Phys. Rev. Lett. 121 (2018) 121801 (CMS), Phys. Lett. B786 (2018) 59 (ATLAS)

Introduction II



- Final state is classified by the decay mode of the $t\bar{t}$ pair:
  - Dileptonic (DL)
  - Single-lepton (SL)
  - Fully hadronic (FH)

  according to the number of top quarks decaying leptonically

- Previous CMS results available for leptonic decay channels [3] and the fully hadronic decay mode of the $t\bar{t}$ pair [4]
- Latest CMS result combines all channels and data from 2016/2017 [5]

$\Rightarrow$ Focus of this talk

[3] JHEP 1903 (2019) 026

[4] JHEP 06 (2018) 101

[5] CMS-PAS-HIG-18-030

Improvements of 2017 analysis

Challenges:

- Complicated final state involving leptons, many jets and MET
- Main background ($t\bar{t}$ + jets) cross-section is three orders of magnitude larger than the signal cross-section ($\sim 832\,\mathrm{pb}$ vs. $0.5\,\mathrm{pb}$)
- Irreducible $t\bar{t}$+bb background
- Large theory uncertainties in the QCD background modelling from MC

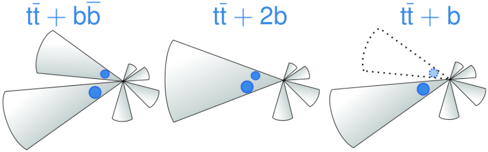
$\Rightarrow$ dedicated analysis strategies in all three channels

Improvements of 2017 analysis:

- Phase 1 upgrade pixel detector with an additional layer close to the beam pipe and new b tagging algorithm $\Rightarrow$ Improved b tagging
- Combination of all three decay channels of the $t\bar{t}$ system
- Improvement of the parton shower modelling uncertainties (now implemented as shape variations via event weights)

$t\bar{t}$ + b jets background

- Considerable background in all analysis channels
- Modelled using Powheg + Pythia8 at NLO
- Additional b jets (not arising from top quark decay) modelled via parton shower
- Involves uncertainties on ISR/FSR, ME-PS matching, UE tune
- Background split according to B hadron content of additional b jets (helps to constrain background in final fit):



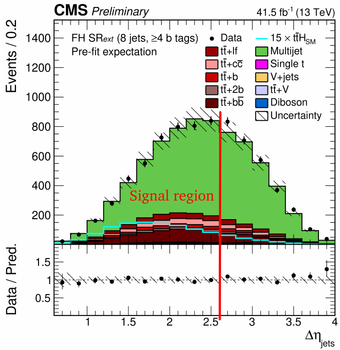
- Uncertainty of 50% on the normalisation separately for each process

Fully hadronic channel (FH) - BR (45%)

- Background dominated by QCD multijet events
- Events categorized by jet and b tagged jet multiplicities → 6 analysis categories, QCD contribution dominant in all categories
- Constrain QCD background by:
  - Testing compatibility of invariant dijet mass m_{qq} with W mass
  - Cutting on angular distance between any two jets:

$$\frac{1}{N_{\text{jets}}} \sum_{\text{jets}} \Delta\eta(\text{jet, jet furthest apart})$$

  - Cutting on quark-gluon likelihood ratio (QGLR) as QCD events contain more gluon-initiated jets as the signal



⇒ 40% QCD rejection at 80% $t\bar{t}H$ efficiency

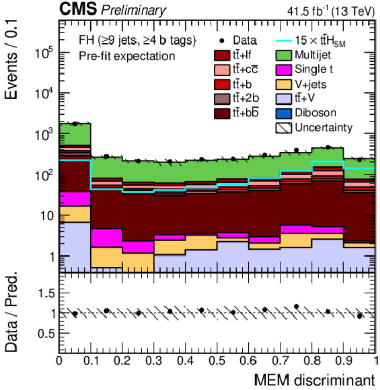
Fully hadronic channel (FH) II

- Matrix element method (MEM) used as final discriminator:
  - Based on the LO scattering amplitude of the process
  - Final discriminant: $P(t\bar{t}H)/(P(t\bar{t}H)+\kappa \cdot P(t\bar{t}{+}bb))$
  - Also provides good discrimination against QCD background

- MEM shape for QCD in signal region is evaluated from control regions in data:

$$\mathrm{MEM_{QCD}} = \mathrm{MEM_{Data}} - \mathrm{MEM_{non\text{-}QCD\ (MC)}}$$

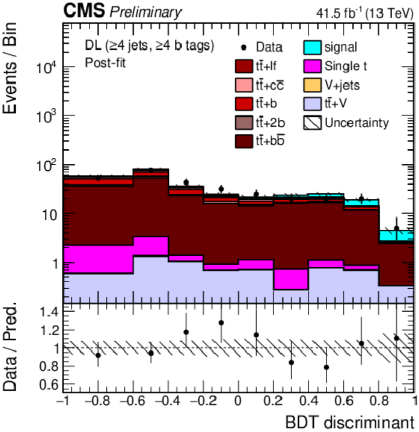
  - Corrections for different b tagging kinematics applied
  - Closure test in a low QGLR validation region
  - Normalization freely floating in final fit



Dilepton channel (DL) - BR (5%)

- Events categorized by jet and b tagged jet multiplicities → 5 analysis categories
- Constrain backgrounds in lower b tagged jet categories

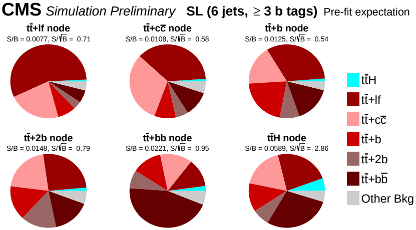
- Boosted decision tree as discriminator in all categories with following input variables:
  - Kinematics of leptons, jets, MET
  - Invariant masses
  - Angular differences
  - b tagging information
  - Matrix Element Method discriminant



Single-lepton channel (SL) - BR (30%)

- Events categorized by jet multiplicities (4,5,$\geq$6 jets)
- Artificial Neural Network (ANN) discriminant used with similar input variables than in DL channel

- Further event classification according to the most probable output node of the ANN:
  - $t\bar{t}H$ signal
  - $t\bar{t}+b\bar{b}$
  - $t\bar{t}+2b$
  - $t\bar{t}+b$
  - $t\bar{t}+c\bar{c}$
  - $t\bar{t}+lf$



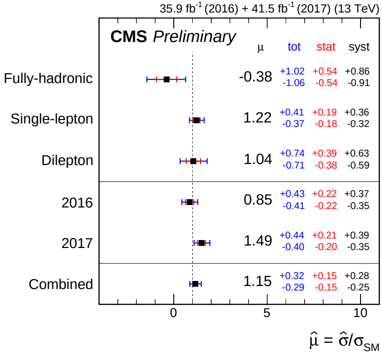
$\rightarrow$ 18 analysis categories

Uncertainties - combined 2016 + 2017 analysis

Uncertainty source	$\Delta\hat{\mu}$
Total experimental	$+0.15/-0.13$
b tagging	$+0.08/-0.07$
jet energy scale and resolution	$+0.05/-0.04$
Total theory	$+0.23/-0.19$
signal	$+0.15/-0.06$
$t\bar{t}$+hf modelling	$+0.14/-0.15$
QCD background prediction	$+0.10/-0.08$
Size of simulated samples	$+0.10/-0.10$
Total systematic	$+0.28/-0.25$
Statistical	$+0.15/-0.15$
Total	$+0.32/-0.29$

- Analysis limited by systematic uncertainties
- Largest contribution from theoretical uncertainties, such as modelling of $t\bar{t}$+hf
- Other major contributions:
  - QCD background prediction
  - Limited size of simulated samples
  - b tagging

Results



- Combined profile likelihood fit
- Compatible with expectations from the SM
- Dominant uncertainties are of systematic kind
- Analysis gives evidence for $t\bar{t}H$, $H \to b\bar{b}$ process

Year	Channel	Best-fit $\hat{\mu}$	Significance obs (exp)
2016	SL + DL	$0.72^{+0.45}_{-0.45}$	1.6 (2.2) σ
2016	FH	$0.9^{+1.5}_{-1.5}$	/
2017	SL + DL + FH	$1.49^{+0.44}_{-0.40}$	3.7 (2.6) σ
2016 + 2017	SL + DL + FH	$1.15^{+0.32}_{-0.29}$	3.9 (3.5) σ

Conclusion

- First inclusive search for $\mathrm{t\bar{t}H}$, $\mathrm{H}\rightarrow \mathrm{b\bar{b}}$ process combining all three decay modes of the $\mathrm{t\bar{t}}$ pair using the 2017 dataset with $41.5\,\mathrm{fb}^{-1}$ of data
- Analysis sensitivity improved because of better b tagging and modelling of the parton shower uncertainties
- Result combined with analysis of 2016 data ($35.9\,\mathrm{fb}^{-1}$)
- Evidence found for $\mathrm{t\bar{t}H}$, $\mathrm{H}\rightarrow \mathrm{b\bar{b}}$ production (3.9σ)

Backup

The Matrix Element Method (MEM)

Probability $P(\mathbf{x}|\boldsymbol{\alpha})$ for an event $\mathbf{x}$ to be the final state of process $\boldsymbol{\alpha}$:

$$P(\mathbf{y}|\boldsymbol{\alpha}) \propto \frac{1}{\sigma_\alpha} \int d\Phi(\mathbf{x})\, |M_\alpha|^2(\mathbf{x})\, W(\mathbf{x}, \mathbf{y})$$

where

- σ_α is the total cross section of process $\boldsymbol{\alpha}$
- $d\Phi(\mathbf{x})$ is the phase-space measure
- $|M_\alpha|^2(\mathbf{x})$ is the LO scattering amplitude squared
- $W(\mathbf{x}, \mathbf{y})$ is the transfer function (probability to obtain a detector response $\mathbf{y}$ for a particle level event $\mathbf{x}$)

Signal ($\mathbf{s}$) to background ($\mathbf{b}$) discriminant:

$$P_{s/b} = \frac{P(\mathbf{y}|\mathbf{s})}{P(\mathbf{y}|\mathbf{s}) + \kappa \cdot P(\mathbf{y}|\mathbf{b})}$$

with κ: scale factor optimized to 0.1