Search for production of a Higgs boson and a single top quark in $\mu\mu$ final states in proton collisions at $\sqrt{s} = 13$ TeV

Departamento de Investigacion en Física
Maestría en Ciencias (Física)
Hiram Ernesto Damián

Universidad de Sonora

March 25, 2019

Contenido

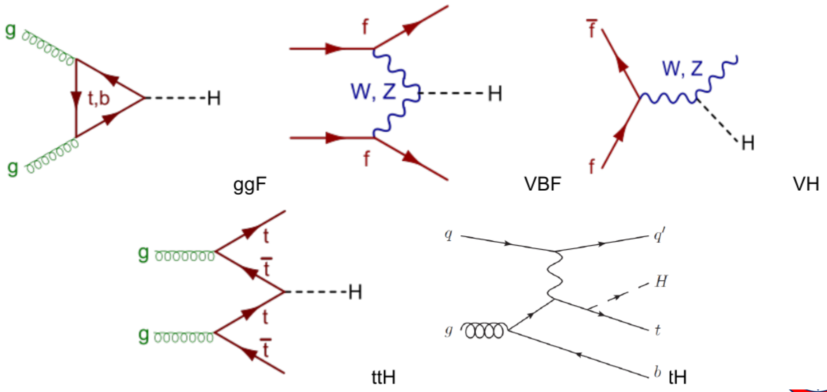
Overview

- Through this project we will investigate the production of Higgs boson in association with a single top quark (tH) in proton-proton collisions with the CMS experiment of the LHC. This mechanism of production of the Higgs boson has not been observed before by any experiment.
- Understanding the production of the Higgs boson, as well as its decays are an important part of the physical program of the CERN international laboratory experiments that try to complete the tests to verify the Standard Model, the theory of the fundamental particles

Motivation for single top Higgs (tH)

- Coupling measurement is essential to establish the nature of the Higgs
- The exploration of Higgs production on the tH channel is subject relatively new. Measurements of CMS and ATLAS are compatible with SM.
- The tH study explores the relative sign of top-Higgs and W-Higgs.
  Small deviations from SM predictions could be associated with physics beyond the standard model (BSM)

Higgs production mechanisms



tH production mechanisms

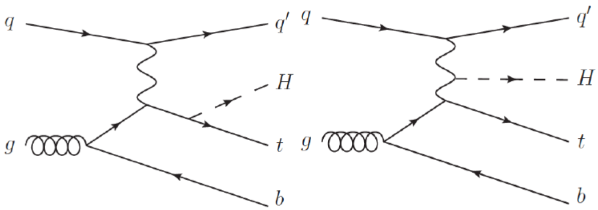


Figure: tH mechanism. Left. Higgs radiated from a top quark. Right.Higgs radiated from a W boson

Cross section

- Classical definition: When two particles interact, cross section is the area transverse to their relative motion within which they must meet in order to scatter from each other.
- Quantum definition: Cross section describes the likelihood of two particles interacting under certain conditions[1]
- Experimentally

$$d\sigma = \frac{\text{number of particles scattered into solid angle}\Delta\Omega}{(\text{number of particles incident})(\text{scattering centers/area})} \tag{1}$$

- Cross sections are expressed in barns , where 1 barn$=10^{-34}$ cm^{-2}

Cross section

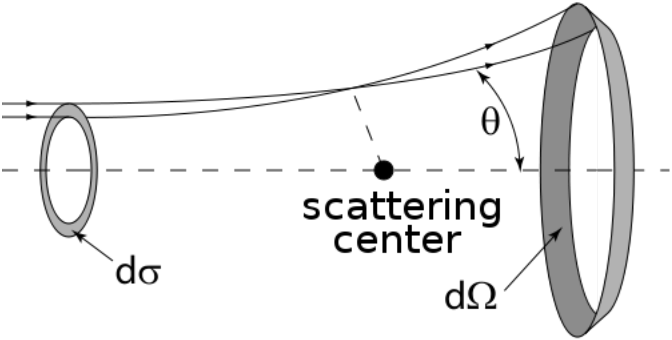


Figure: Drawing of an idealized scattering process showing the differential solid angle $\Delta\Omega$ and the scattering angle θ

Cross section

The reaction rate is connected with the cross section through the number of scatters in the target and the incident flux. The incident flux is called Luminosity and it is measured in $\mathrm{cm}^{-1}\mathrm{s}^{-1}$.

$$N_R = \sigma \mathrm{L} \tag{2}$$

The reaction rate is related to σ and luminosity L.

Higgs production Cross section

Higgs boson production cross sections in pp collisions for $\sqrt{s} = 13$TeV (in pico barn).Integrated luminosity of 35.9 fb^{-1} for Run 2[1]

Production mechanism	σ (picobarns pb)	Number of events
ggF	48.93	1756587
VBF	3.78	135702
WH	1.35	48465
ZH	0.88	31592
ttH	0.50	18255
tH (only)	0.015	560.39

[1] Data taken from The cern collaborarion "Higgs Physics the HL-LHC and HE-LHC" 2019, CERN-LPCC-2018-04

Higgs production cross section

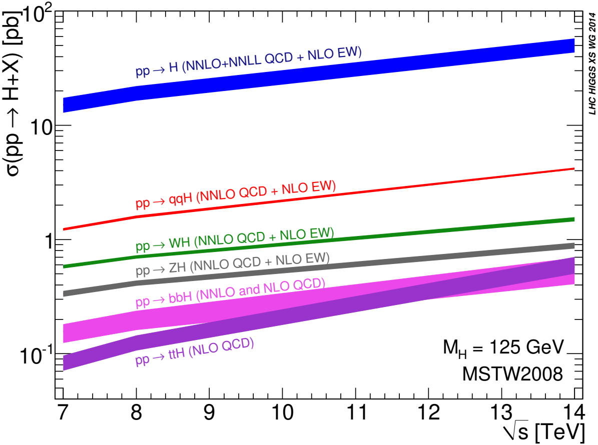


Figure: Higgs boson production cross sections as a function of the

Branching ratio

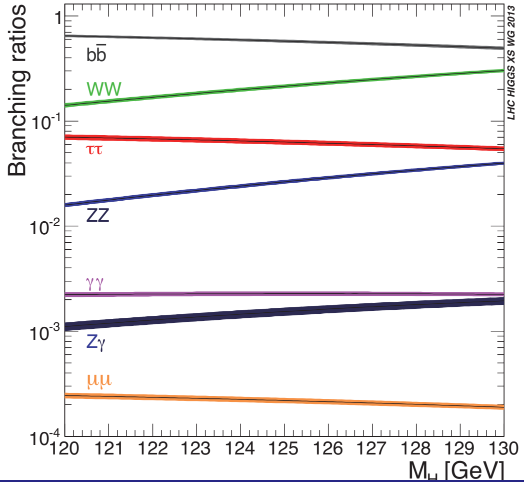
In particle physics, the branching ratio for a decay process is the ratio of the number of particles which decay via a specific decay mode with respect to the total number of particles which decay via all decay modes. It is also equal to the ratio of the partial decay constant to the overall decay[2].

$$\mathrm{BR} = \frac{\Gamma_i}{\sum_i \Gamma_i} \tag{3}$$

Where $\Gamma = \sum_i \Gamma_i$ is the total decay width (sum of all partial widths) of the particle and is related to lifetime of the particle: $\Gamma = 1/\tau$ [0]

[0] Cleaves H.J. (2011) Branching Ratio. In: Gargaud M. et al. (eds) Encyclopedia of Astrobiology. Springer, Berlin, Heidelberg

Higgs Branching ratio



Higgs Branching ratios per channel

SM Higgs boson branching ratios for M_H =125 GeV

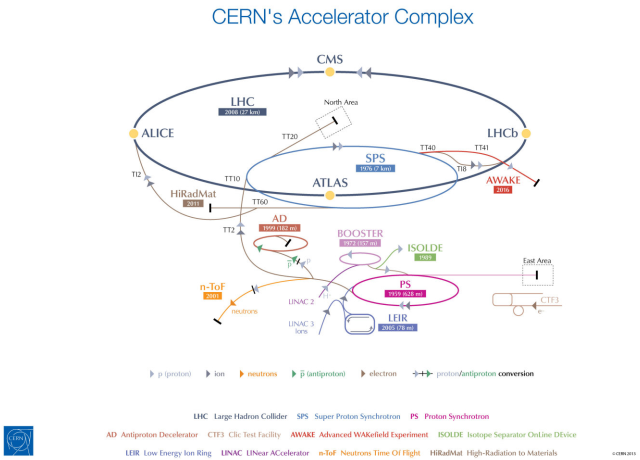
Higgs decay	Branching ratio (BR)
H $\rightarrow$ b$\bar{b}$	5.82×10^{-1}
H $\rightarrow$ W^+W^-	2.15×10^{-1}
H $\rightarrow$ $\tau\bar{\tau}$	6.27×10^{-2}
H $\rightarrow$ ZZ	2.61×10^{-2}
H $\rightarrow$ $\gamma\gamma$	2.27×10^{-3}
H $\rightarrow$ Zγ	1.53×10^{-3}
H $\rightarrow$ $\mu^+\mu^-$	2.17×10^{-4}

$\mu\mu$ same sign decay rate

Table of decay chains for tH. Expected number of final events assuming 560 produced tH events.l represents $\mu^{\pm}$, e^{-}, $\tau^{\pm}$. h represent hadrons ($\pi^{\pm}$, $K^{\pm}$)

Decay chain	BR	Events
$tH \rightarrow W^+bW^+W^- \rightarrow \mu^+\ \nu_\mu b\mu^+\nu_\mu\ h^-$	2.096×10^{-3}	1.173
$tH \rightarrow W^+bW^+W^- \rightarrow \mu^+\nu_\mu b\mu^+\nu_\mu\ l\nu_l$	3.37×10^{-4}	0.899
$tH \rightarrow W^+b\ W^+W^- \rightarrow \mu^-\bar{\nu_\mu}\nu_\tau\nu_\tau\ b\mu^-\bar{\nu_\mu}h^+$	1.890×10^{-4}	0.105
$tH \rightarrow W^+b\ \tau\bar{\tau} \rightarrow \mu^+\ \nu_\mu b\ h^-\ \nu_\tau\ \mu^+\ \nu_\mu\bar{\nu_\tau}$	1.681×10^{-4}	0.094
$tH \rightarrow W^+b\ \tau\bar{\tau} \rightarrow \mu^+\nu_\mu\ b\ \mu^-\bar{\nu_\mu}\nu_\tau\mu^+\nu_\mu\bar{\nu_\tau}$	1.519×10^{-4}	0.085
$tH \rightarrow \tau\nu_\mu b\ W^+W^- \rightarrow \mu^-\bar{\nu_\mu}\nu_\tau\nu_\tau\ b\mu^-\bar{\nu_\mu}l\nu_l$	3.045×10^{-5}	0.017
$tH \rightarrow q\bar{q}bZZ \rightarrow q\bar{q}\ b\ \mu^+\mu^-\mu^+\mu^-$	1.966×10^{-5}	0.011
$tH \rightarrow \tau\nu_\tau b\ \tau\bar{\tau} \rightarrow \mu^-\bar{\nu_\mu}\nu_\tau b\ \mu^-\bar{\nu_\mu}\nu_\tau\ h^+$	1.549×10^{-5}	0.008
$tH \rightarrow W^+bZ\gamma \rightarrow \mu^+\nu_\mu\ b\ \mu^+\mu^-\gamma$	6.888×10^{-6}	0.003
$tH \rightarrow W^+bZZ \rightarrow \mu^+\nu_\mu\ b\ \mu^+\mu^-l^+l^-$	3.962×10^{-6}	0.002

LHC


CERN's Accelerator Complex
CMS
LHC
2008 (27 km)
North Area
ALICE
LHCb
TT20
TT40
TT41
SPS
1976 (7 km)
TI8
TI2
TT10
ATLAS
AWAKE
2016
HiRadMat
2011
TT60
AD
1999 (182 m)
TT2
BOOSTER
1972 (157 m)
ISOLDE
1989
East Area
n-ToF
2001
PS
1959 (628 m)
CTF3
neutrons
LINAC 2
LINAC 3
Ions
LEIR
2005 (78 m)
p (proton)
ion
neutrons
p̄ (antiproton)
electron
proton/antiproton conversion
LHC Large Hadron Collider SPS Super Proton Synchrotron PS Proton Synchrotron
AD Antiproton Decelerator CTF3 Clic Test Facility AWAKE Advanced WAKEfield Experiment ISOLDE Isotope Separator OnLine DEvice
LEIR Low Energy Ion Ring LINAC LINear ACcelerator n-ToF Neutrons Time Of Flight HiRadMat High-Radiation to Materials
© CERN 2013

LHC parameters

Quantity	Number
Circumference	26 659 m
Dipole operating temperature	1.9 K (-271.3°C)
Number of magnets	9593
Number of main dipoles	1232
Number of main quadrupoles	392
Number of RF cavities	8 per beam
Nominal energy, protons	6.5 TeV
Nominal energy, ions	2.56 TeV/u (energy per nucleon)
Nominal energy, protons collisions	13 TeV
No. of bunches per proton beam	2808
No. of protons per bunch (at start)	1.2 x 10^{11}
Number of turns per second	11245
Number of collisions per second	1 billion

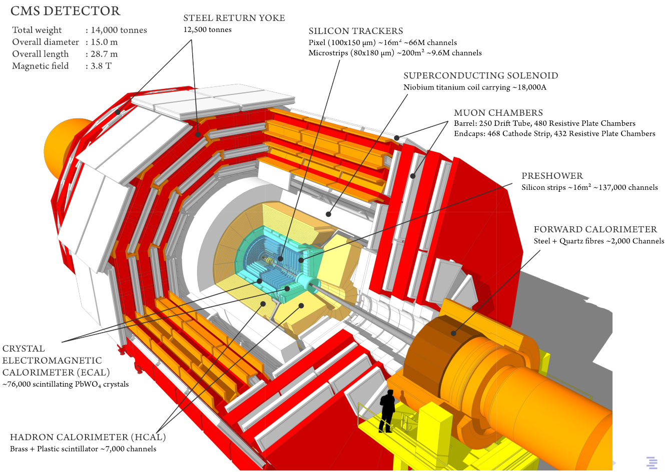
Table. LHC characteristics for run 2.

LHC

Accelerator operation energies

Accelerator	Energy
Linac 2	50 MeV
PS Booster	1.4 GeV
Proton Scyncroton (PS)	25 GeV
SPS	450 GeV
LHC	6.5 TeV

CMS detector



CMS integrated luminosity

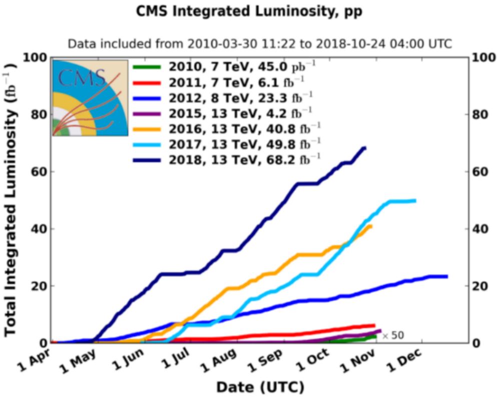
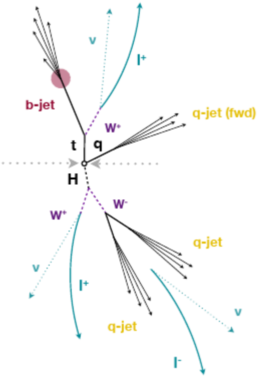


Figure: Integrated luminosity for CMS experiment

Topology of events tH

The characteristics of the signal tHq:

- $t \to W^+ b \to \mu^+ \nu_\mu$
- $H \to W^+W^-$
- $W \to l^\pm \nu$, where l can be $\mu^\pm$,e^- or $\tau^\pm$ or $W \to q\bar{q}$
- W bosons decay leptonically with equal electrical charge, resulting in a signature of two same-sign leptons with two light-quark jets.[2]



Event selection

The main analysis strategy is to obtain a selection of events compatible with certain characteristics of the signal at pre-selection level.

It is required:

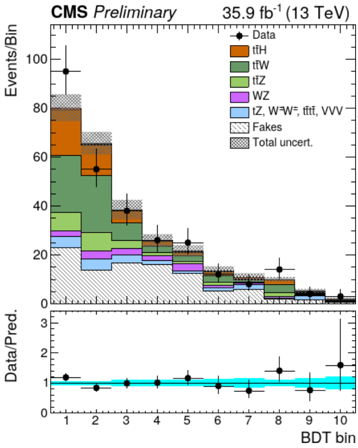
- The events are selected those that contain two leptons $(\mu^{+}\mu^{-})$ with the same sign.
- Transverse moment $p_t > 25$ and 15 GeV, for the muons.
- A front jet with $p_t > 40$ GeV, $\mid \eta \mid > 2.4$
- One or more b-jets with $(\mid \eta \mid <2{,}4)$[1]

Previous results

- Direct searches for tHq production using all relevant Higgs decay modes have previously been carried out by CMS in the 8 TeV dataset and in the 2015 13 TeV dataset using the $H \to b\bar{b}$ channel .
- $\mu\mu$ final states with two same-sign leptons target the case where the Higgs boson decays to a pair of W bosons, τ leptons, or Z bosons, and where the top quark decays leptonically[1]

Previous results

Process	$\mu\mu$
$t\bar{t}W^{\pm}$	68.03 ± 0.61
$t\bar{t}Z/t\bar{t}\gamma$	25.89 ± 1.12
WZ	15.07 ± 1.19
ZZ	1.16 ± 0.29
$W^{\pm}W^{\pm}qq$	3.96 ± 0.52
$W^{\pm}W^{\pm}$(DPS)	2.48 ± 0.42
VVV	2.99 ± 0.34
tttt	2.32 ± 0.45
tZq	5.77 ± 2.24
tZW	2.13 ± 0.13
γ conversions	–
Non-prompt	80.94 ± 2.02
Charge flips	–
Total Background	210.74 ± 3.61
$t\bar{t}H$	24.18 ± 0.48
tHq (SM)	1.43 ± 0.04
tHW (SM)	0.71 ± 0.03
Total SM	237.06 ± 3.64



Post-fit categorized BDT classifier outputs as used in

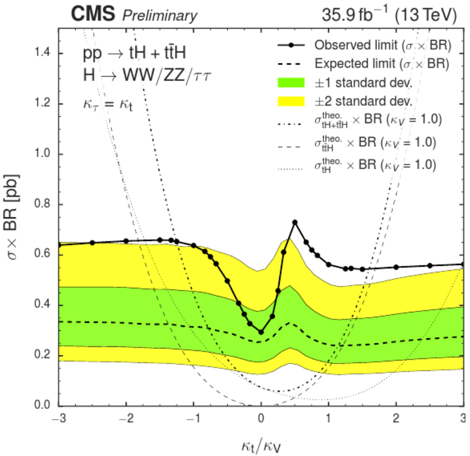


Figure: Observed and expected 95% C.L. upper limit on the tH + ttH cross section times H → WW + $\tau\tau$ + ZZ branching fraction for different 20 values of the coupling ratio κ_t/κ_v . The expected limit is derived from

Main backgrounds

In the leptonic channels, the main backgrounds are expected to arise from the production of top quarks

- In the dominant $t\bar{t}$ mode, where same-sign dilepton signatures can occur when a non-prompt lepton from heavy-flavor decay passes the signal selection, or in associated production with a W/Z or Higgs boson.
- Processes with single top quarks also contribute, mostly in the associated production with a Z boson (tZ) or when produced with both a W and a Z boson (WZ)

Main backgrounds

- Backgrounds from associated production of $t\bar{t}$ pairs and electroweak bosons ($t\bar{t}W^{\pm}$ and $t\bar{t}Z$) are estimated directly from simulated events, which are corrected for data/MC differences and inefficiencies in the same way as signal events.
- Diboson production with leptonic Z decays and additional jet radiation in the final state can lead to signatures very similar to that of the signal. Due to the larger cross section, the main contribution arises from WZ production.

Main backgrounds

All background contributions are obtained from ttH same sign $\mu\mu$ analysis.

- WZ
- ttZ
- ttH
- tV,VVV,WW,tttt
- Non prompt leptons are included

Boosted decision tree (BDT)

A decision tree takes a set of input features and splits input data recursively based on those features. Boosting is a method of combining many weak learners (trees) into a strong classifier.The features can be a mix of categorical and continuous data.

1. BDT training using MC samples for signal and backgrounds for $t\bar{t}V$ ($t\bar{t}W^{\pm}$ and $t\bar{t}Z$):
2. Against combined $t\bar{t}Z$ and $t\bar{t}W$: prompt lepton type background.
3. Extract the signal contribution in a second analysis step, using multivariate discriminators against the main backgrounds of $t\bar{t}W^{\pm}$ /$t\bar{t}Z$ and non prompt leptons from $t\bar{t}$[2].

BDT Variables

- Trailing lepton p_t
- Total charge of tight leptons
- min ΔR (lepton pairs)
- $\Delta\phi$ between highest p_t lepton pair
- Number of jets with $|\eta| < 2.4$
- Number of non b-tagged jets with $|\eta| > 1.0$
- Maximum $|\eta|$ for jets
- $\Delta\eta$ (most forward light jet, closest lepton)
- $\Delta\eta$ (most forward light jet, hardest loosely b-tagged jet)
- $\Delta\eta$ (most forward light jet, 2nd hardest loosely b-tagged jet)

Sources of uncertainty on event yields

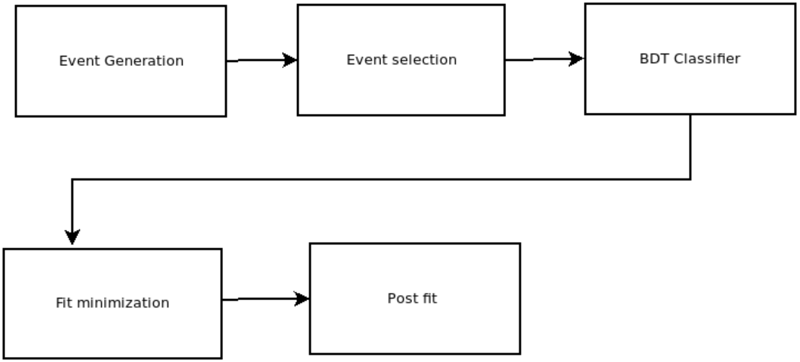
- Luminosity measurement: 2.6%.
- Data/MC scale factors for lepton selection (ID, iso) and trigger efficiencies 5% per lepton.
- Choice of PDF set:
  - 3.7% for tHq
  - 4% for tHW $t\bar{t}$W, $t\bar{t}$Z, ttH

Sources of uncertainty on event yields

Systematic uncertainties (background normalization)

- tth 5%
- t$\bar{t}$Z 10.7 %
- t$\bar{t}$W 12.6 %
- tZ 50%
- WZ 50%
- Non prompt leptons /fakes 40%

Fitting


Event Generation
Event selection
BDT Classifier
Fit minimization
Post fit

Results

Likelihood scan

The likelihood function is the product of Poisson probabilities for all bins

$$L(\mu, \theta) = \prod_{j=1}^{N} \frac{(\mu s_j + b_j)^{n_j}}{n_j!} e^{-(\mu s_j + b_j)} \tag{4}$$

- N=number of bins
- μ=parameter of signal
- s=signal
- b=background
- n=number of events

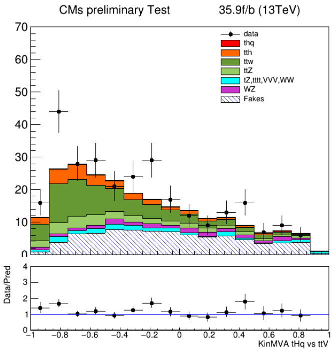
Results

Table: Prefit and postfit table for each yield.

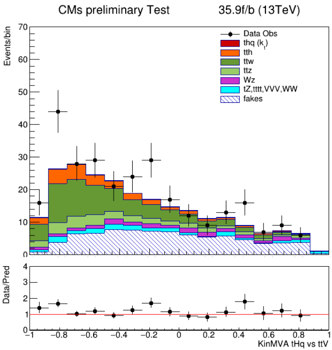
Process	Number of events prefit	Number of events Postfit
tH	2.13 ± 0.05	10.28 ± 26.17
ttH	24.18 ± 0.10	24.31 ± 0.13
$t\bar{t}W$	68.03 ± 8.60	75.57 ± 7.85
$t\bar{t}Z$	25.89 ± 2.78	26.43 ± 2.77
tZ	15.04 ± 7.52	16.25 ± 7.53
WZ	15.07 ± 7.53	15.95 ± 7.46
fakes	80.94 ± 32.37	96.80 ± 25.58

Prefit uncertainty is statistical only. Postfit uncertainty is statistical + systematic.

Results



Pre-fit signal and background yields for tH process. In the box below each distribution, the ratio of the observed and predicted event yields is shown



Post-fit signal and background yields for tH process. In the box below each distribution, the ratio of the observed and predicted event yields is shown

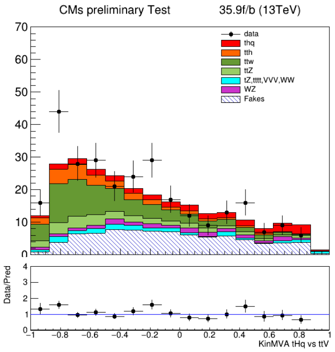
Results

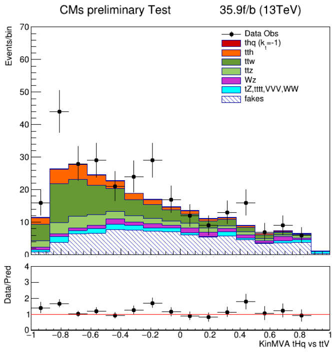
Table: Prefit and postfit table for each yield. k_t =-1.

Process	Number of events prefit	Number of events Postfit
tH	26.2 ± 0.27	6.49 ± 25.38
ttH	24.18 ± 1.31	24.32 ± 0.14
$t\bar{t}W$	68.03 ± 8.60	75.69 ± 8.178
$t\bar{t}Z$	25.89 ± 2.78	26.44 ± 2.78
tZ	15.04 ± 7.52	16.40 ± 7.44
WZ	15.07 ± 7.53	16.10 ± 7.53
fakes	80.94 ± 32.37	99.45 ± 25.80

Prefit uncertainty is statistical only. Postfit uncertainty is statistical + systematic.

Results





Pre-fit signal and background yields for tH process for k_t=-1.
In the box below each distribution, the ratio of the observed and predicted event yields is shown

Post-fit signal and background yields for tH process for k_t=-1.
In the box below each distribution, the ratio of the observed and predicted event yields is shown

Results

Likelihood scan

- Likelihood function (often simply the likelihood) is a function of the parameters of a statistical model, given specific observed data.
- Likelihood functions play a key role in frequentist inference, especially methods of estimating a parameter from a set of statistics.
- In informal contexts, "likelihood" is often used as a synonym for probability.

Results

Likelihood scan

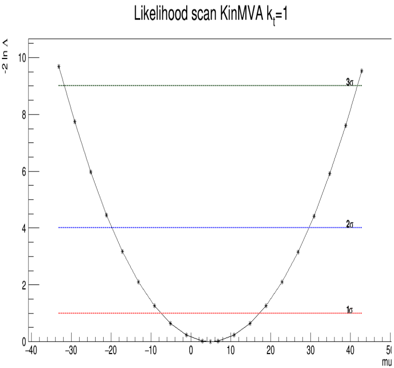
To test a hypothesized value of μ we consider the profile likelihood ratio

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\theta}})}{L(\hat{\mu}, \hat{\theta})} \tag{5}$$

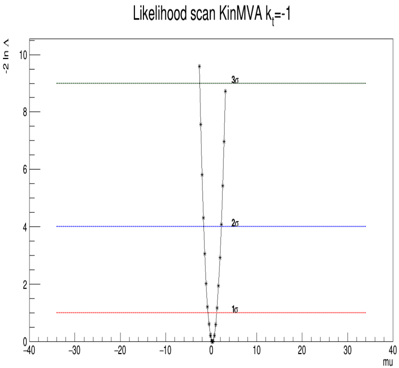
- Here $\hat{\hat{\theta}}$ in the numerator denotes the value of θ that maximizes L for the specified μ, it is the conditional maximum-likelihood (ML) estimator of $\hat{\theta}$ (and thus is a function of μ). The denominator is the maximized (unconditional) likelihood function, i.e., $\hat{\mu}$ and $\hat{\theta}$ are their ML estimators[3].
- The presence of the nuisance parameters broadens the profile likelihood as a function of μ relative to what one would have if their values were fixed. This reflects the loss of information about μ due

Results

Likelihood scan



Likelihood scan for k$_t$=1 (SM)



Likelihood scan for k$_t$=-1

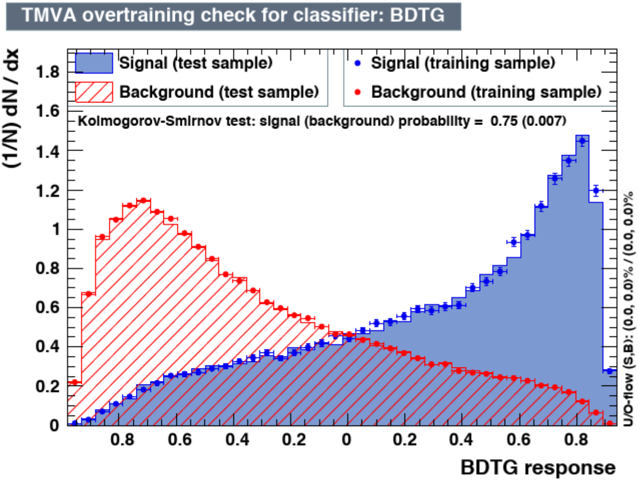
References

- Gross F. *Relativistic quantum mechanics and field theory* 1994 Wiley
- The CMS collaboration, *Search for production of a Higgs boson and a single top quark in multilepton final states in proton collisions at 13 TeV*, CMS-PAS-HIG-17- 005
- Verkerke W *Dealing with systematic uncertainties* 2014. From https://indico.cern.ch/event/287744/contributions/1641261/attachments/535763/738679/Verkerke_Statistics_3.pdf
- Cowan G. , Cranmer K., Gross E. , Vitells O. *Asymptotic formulae for likelihood-based tests of new physics* 2013 , arXiv:1007.1727

Back up

BDT parameters

- Gradient boosted (BDTG)
- No. of trees = 800
- No. of cuts = 50
- Maximum depth = 3
- Found to be most discriminating, minimal overtraining.



ttV BDT output (2lss)

List of histograms used in the analysis

Data taken from the file `plots-thq-2lss-kinMVA.root` 2016 CERN

```
thqMVA_ttv_2lss_40_tZq
thqMVA_ttv_2lss_40_ttZ
thqMVA_ttv_2lss_40_VVV
thqMVA_ttv_2lss_40_ttW
thqMVA_ttv_2lss_40_data_fakes
thqMVA_ttv_2lss_40_ttH
thqMVA_ttv_2lss_40_tHW_hww
thqMVA_ttv_2lss_40_WWss
thqMVA_ttv_2lss_40_tttt
thqMVA_ttv_2lss_40_WZ
thqMVA_ttv_2lss_40_tHq_hww
```

Backgrounds and signal histograms

tHq: Signal (tH)

- `thqMVA_ttv_2lss_40_tHq_hww`
- `thqMVA_ttv_2lss_40_tHW_hww`

Backgrounds

ttW

- `thqMVA_ttv_2lss_40_ttW`

ttZ

- `thqMVA_ttv_2lss_40_ttZ`

WZ

- `thqMVA_ttv_2lss_40_WZ`

Backgrounds and signal histograms

Backgrounds

tZ, VVV,tttt,WW:

- thqMVA_ttv_2lss_40_tZq
- thqMVA_ttv_2lss_40_WWss
- thqMVA_ttv_2lss_40_VVV
- thqMVA_ttv_2lss_40_tttt

ttH

- thqMVA_ttv_2lss_40_ttH

Non prompt leptons (fakes)

- thqMVA_ttv_2lss_40_data_fakes

Sources of systematic uncertainty in HEP (high energy physics)

Detector-simulation related uncertainty

- Calibrations (electron, jet energy scale)
- Efficiencies (particle ID, reconstruction)
- Resolutions (jet energy, muon momentum)

Theoretical uncertainties

- Factorization/Normalization scale of MC generators
- Choice of MC generator (ME and/or PS, e.g. Herwig vs Pythia)

Monte Carlo Statistical uncertainties

- Statistical uncertainty of simulated samples[2]

Observed events

- After applying the event pre-selection on the dataset, 280 events are observed in the same-sign $\mu\mu$ channel
- The events are then sorted into ten categories depending on the output of the two BDT classifiers according to an optimized binning strategy, resulting in a one-dimensional histogram with ten bins.
- In each point, the tH and ttH production cross sections and the Higgs decay branching ratios are modified with the Higgs-top (κ_t) and Higgs-vector boson (κ_V) coupling strength.
- The Higgs-tau coupling strength modifier (κ_τ) is assumed to be equal to κ_t .
- All other parameters are assumed to be at the values predicted by the standard model[1]

α values for post fit

```
Floating Parameter      FinalValue +/- Error
--------------------------------------------
--------------------------------------------
Lumi                    1.0000e+00 +/- 1.00e-04
alpha_sample_tth_sys    1.1516e-01 +/- 9.94e-01
alpha_sample_ttw_sys    8.7701e-01 +/- 9.13e-01
alpha_sample_ttz_sys    1.9573e-01 +/- 9.95e-01
alpha_sample_tz_sys     1.6106e-01 +/- 1.00e+00
alpha_sample_wz_sys     1.1766e-01 +/- 9.90e-01
alpha_sample_fakes_sys  4.8992e-01 +/- 7.90e-01
mu                      4.8280e+00 +/- 1.23e+01
```

α values for post fit thq k_t=-1

Floating Parameter	FinalValue +/- Error
Lumi	1.0000e+00 +/- 1.00e-04
alpha_sample_tth_sys	1.1667e-01 +/- 9.96e-01
alpha_sample_ttw_sys	8.9019e-01 +/- 9.51e-01
alpha_sample_ttz_sys	2.0063e-01 +/- 9.99e-01
alpha_sample_tz_sys	1.8157e-01 +/- 9.90e-01
alpha_sample_wz_sys	1.3754e-01 +/- 1.00e+00
alpha_sample_fakes_sys	5.7180e-01 +/- 7.97e-01
mu	2.4808e-01 +/- 9.69e-01

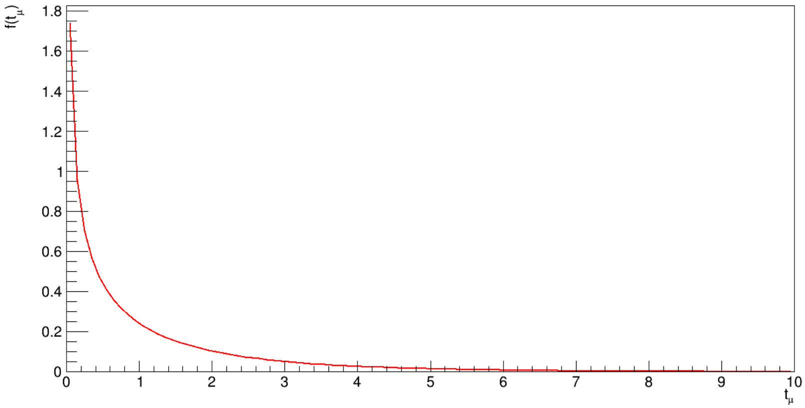
Statistical test

It is convenient to use the statistic

$$t_\mu = -2 \ln \lambda(\mu) \tag{6}$$

as the basis of a statistical test. Higher values of t_μ thus correspond to increasing incompatbility between the data and μ. We may define a test of a hypothesized value of μ by using the statistic t_μ directly as measure of discrepancy between the data and the hypothesis, with higher values of t_μ correspond to increasing disagreement[4]

Statistical test



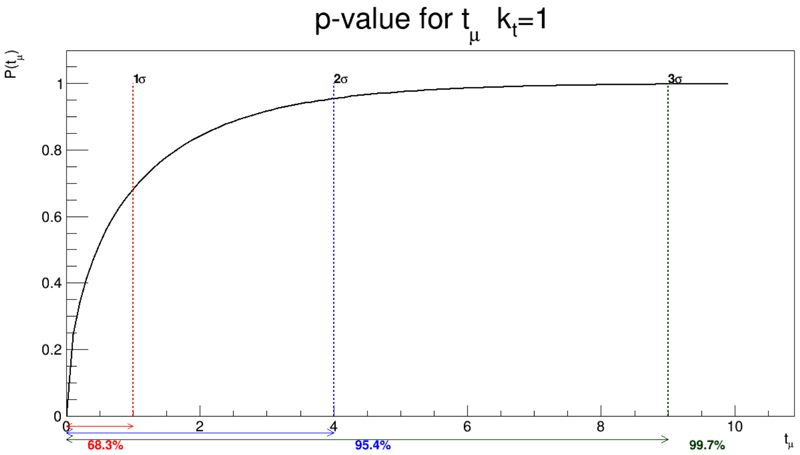
P value

To quantify the level of disagreement we compute the P-value

$$P_{\mu} = \int_{t_{\mu}}^{\infty} f(t\mu|\mu)dt_{\mu} \quad (7)$$

where t μ is the value of the statistic tμ observed from the data and $f(t\mu|\mu)$ denotes the PDF (Probability density function) of tμ under the assumption of the signal strength μ[4]

P value



p-value for t_μ k_t=1
$P(t_\mu)$
t_μ
1σ
2σ
3σ
68.3%
95.4%
99.7%
0
0.2
0.4
0.6
0.8
1
0
2
4
6
8
10