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 15, 2019



Contenido

- Overview and motivation
- 2 tH mechanisms
- Production cross section
- 4 Higgs Branching ratios
- 6 Previous results
- 6 Fitting
- Results
- References
- Supportive slides



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

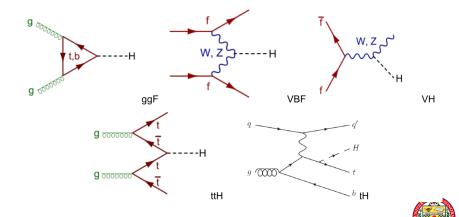


Motivation for single top Higgs (tH)

- 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

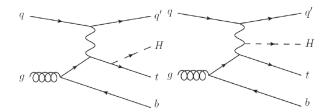


Figure: tH mechanism. Left. Higgs radiated from a top quark. Right.Higgs radiated from a W boson



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

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

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



Cross section

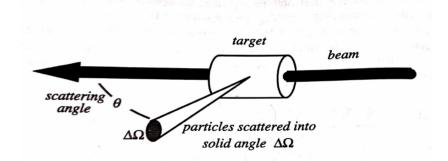


Figure: Drawing of an idealized scattering process showing the differential solid angle $\Delta\Omega$ and the scattering angle $\theta[1]$

Cross section

The transition rate is connected with the cross section through the number of scatterers in the target and the incident flux. The incident flux is called Luminosity and it is measured in cm⁻¹s⁻¹. The reaction rate (event rate) is related to σ and luminosity $\mathcal L$

$$N_R = \sigma \mathcal{L}$$
 (2)



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 2¹

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





Higgs production cross section

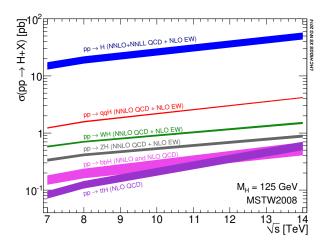


Figure: Higgs boson production cross sections as a function of the centre-of-mass-energies



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 decay².

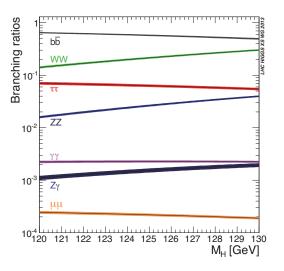
Branching ratio =
$$\frac{\Gamma_i}{\sum_i \Gamma_i}$$
 (3)

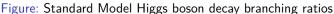
Where $\Gamma = \sum_{i} \Gamma_{i}$ is the total decay width (sum of all partial widths) of the particle and is related to lifetime of the particle: $\Gamma = 1/\tau^0$

⁰Cleaves 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 and number of events per decay for tH process $M_{H} = 125~{\rm GeV}$

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



$\mu\mu$ same sign decay rate

Table of decay chains for tH. Number of total events for tH=560.39. $I=\mu^{\pm},e^{-}\tau$

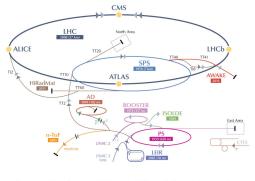
Decay chain	BR	Events
$tH \rightarrow W^+bW^+W^- \rightarrow \mu^+\nu_{\mu}b\mu^+\nu_{\mu} I\nu$	3.37×10^{-4}	0.899
$tH \rightarrow \mathrm{W^+bW^+W^-} \rightarrow \mu^+ \; \nu_\mu \mathrm{b} \mu^+ \nu_\mu \mu^- \nu_\mu$	3.235×10^{-4}	0.1888
$tH \rightarrow W^+ b \tau^+ \tau^- \rightarrow \mu^+ \nu_\mu b \nu_\mu \nu_\tau \mu^- \nu_\mu \nu_\tau$	2.540×10^{-4}	0.1423
$tH \rightarrow \tau \nu_{\mu} b W^{+}W^{-} \rightarrow \mu^{+} \mu^{\nu}_{\mu} \nu_{\tau} b + \mu^{-} \mu^{+} \mu^{-}$	2.981×10^{-5}	0.016
$tH \rightarrow W^+ bZ\gamma \rightarrow \mu^+ \nu_\mu b\mu^- \nu_\mu \mu^+ \nu_\mu \gamma$	6.902×10^{-6}	0.0038
$tH \rightarrow W^+ bZZ \rightarrow I \nu b \mu^+ \mu^- \mu^+ \mu^-$	3.963×10^{-6}	0.0022
$tH \rightarrow W^+ bZZ \rightarrow \mu^+ \nu_\mu b\mu^+ \mu^- \mu^+ \mu^-$	3.962×10^{-6}	0.0022
$tH \rightarrow \tau \nu_{\mu} bZZ \rightarrow \mu^{+} \nu_{\mu} b\mu^{+} \nu_{\mu} \mu^{-} \nu_{\mu}$	3.650×10^{-7}	0.0002





LHC

CERN's Accelerator Complex



 $\blacktriangleright \ p \ (proton) \qquad \blacktriangleright \ ion \qquad \blacktriangleright \ neutrons \qquad \blacktriangleright \ \overline{p} \ (antiproton) \qquad \blacktriangleright \ electron \qquad \lnot \blacktriangleright \bot \blacktriangleright \ proton/antiproton \ conversion$



AD Antiproton Decelerator CTF3 Clic Test facility AWAKE Advanced WAKefield Experiment ISOLDE Isotope Separator OnLine Device

LEIR Low Energy Ion Ring UNAC UNear Accelerator n-ToF Neutrons Time Of Flight HiRadMat High-Radiation to Materials







LHC

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)	1.2×10 ¹¹
Number of turns per second	11245
Number of collisions per second	1 billion

Table. LHC characteristics for run 2.



CMS

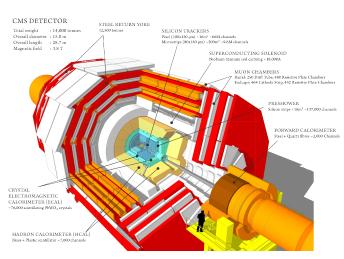


Figure: Compact muon solenoid



CMS

CMS Integrated Luminosity, pp

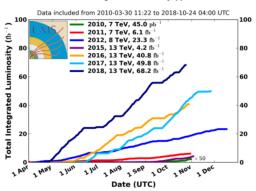


Figure: Integrated luminosity for CMS experiment





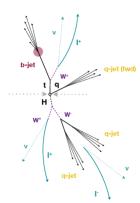
Topology of events tH

The characteristics of the signal tHq:

 A Higgs Bosons decays two W boson H→ WW Same-sign dilepton (2lss): one W from Higgs decays hadronically, others decay Leptonically.

• W Boson decays to a lepton and a neutrino $W \rightarrow I \nu$

 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

- \bullet The events are selected those that contain two leptons ($\mu\mu$) with the same sign.
- 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:

- Transverse moment $p_t > 25$ and 15 GeV, for the muons.
- A front jet with $p_t > 40$ GeV, $|\eta| > 2.4$
- One or more b-jets with ($\mid \eta \mid <2,4)(1)$

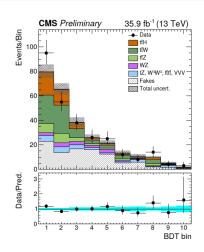


Previous results

- 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 $H \rightarrow b\bar{b}$ channel .
- In the full 2016 13 TeV dataset, a search for ttH production in multilepton final states recently produced first evidence for associated production of top quarks and Higgs bosons
- Multilepton final states with either 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	μμ
t t W [±]	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 t H	24.18 ± 0.48
tHq (SM)	1.43 ± 0.04
tHW (SM)	0.71 ± 0.03
Total SM	237.06 ± 3.64
$tHq (\kappa_V = 1 = -\kappa_t)$	18.48 ± 0.22
tHW ($\kappa_{\rm V} = 1 = -\kappa_{\rm t}$)	7.72 ± 0.17
Data	280



Post-fit categorized BDT classifier outputs as used in the maximum likelihood fit for the $\mu\mu$ channel for 35 fb $^{-1}$.In the box below each distribution, the ratio of the observed and predicted event yields is shown.



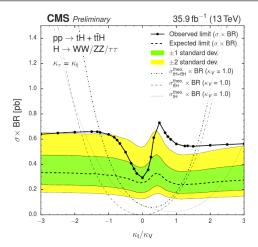


Figure: Observed and expected 95% C.L. upper limit on the tH + ttH cross section times H \rightarrow WW + $\tau\tau$ + ZZ branching fraction for different 20 values of the coupling ratio κ_t/κ_v . The expected limit is derived from a background-only MC dataset.

Signal and background

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 multi-lepton and 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)



Signal and background

Signal: tH for same sign $\mu\mu$ final state Background components

- 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. Pros:

- Fast
- Easy to tune
- Not sensitive to scale (The features can be a mix of categorical and continuous data)
- Good performance

Cons:

• Sensitive to overfitting and noise



Boosted decision tree (BDT) Signal discrimination using BDT

The production cross section for the signal processes tH is only a few fb resulting in a small signal to background ratio even for a tight selection.

Two separate BDT trainings using MC samples for signal and backgrounds:

- **1** Signal is only tHq with $\kappa_t = -1.0$, $\kappa_v = 1.0$.
- ② Against $t\bar{t}$: non-prompt lepton type background.
- Against combined ttZ and ttW: prompt lepton type background.
- $\begin{tabular}{ll} \bullet & Extract the signal contribution in a second analysis step, using multivariate discriminators against the main backgrounds of $$ttW^{\pm}$ /ttZ and non prompt leptons from tt[2]. \end{tabular}$

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 $|\eta| < 2.4$
- ullet 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)
- ullet $\Delta\eta$ (most forward light jet, 2nd hardest loosely b-tagged jet)



Sources of systematic uncertainty

Luminosity measurement: 2.6Data/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, ttW, ttZ, ttH Systematic uncertainties (background normalization)

- tth 5%
- ttZ 10.7 %
- ttW 12.6 %
- tZ 50%
- WZ 50%
- Non prompt leptons /fakes 40%



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 ($\kappa_{ au}$) is assumed to be equal to κ_{t} .
- All other parameters are assumed to be at the values predicted by the standard model[1] .



Fitting

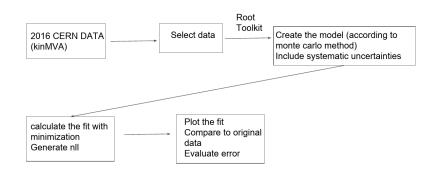
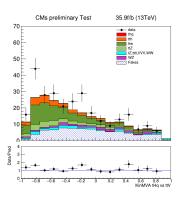


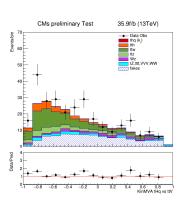


Table: Prefit and postfit table for each yield.

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
ttW	68.03 ± 8.60	75.57 ± 7.85
ttZ	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





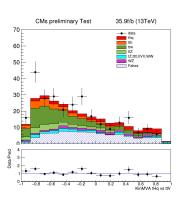


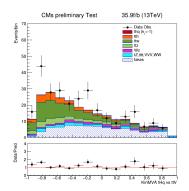
Pre-fit signal and background yields for tH process. In the box below each distribution, the ratio of the observedthe box below each distribution, the ratio of the observed and predicted event yields is shown Post-fit signal and background yields for tH process. In the box below each distribution, the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

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
ttW	68.03 ± 8.60	75.69 ± 8.178
ttZ	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







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



- 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.



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

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

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





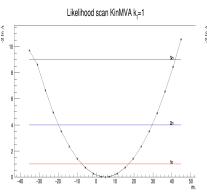
To test a hypothesized value of μ we consider the profile likelihood ratio

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

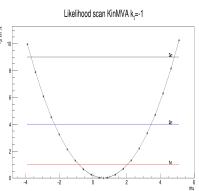
Here $\hat{\theta}$ in the numerator denotes the value of θ that maximizes L for the specified μ , it is the conditional maximum-likelihood (ML) estimator of $\hat{\theta}$ (and thus is a function of μ). The denominator is the maximized (unconditional) likelihood function, i.e., $\hat{\mu}$ and $\hat{\theta}$ 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 to the systematic uncertainties[3][4].





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

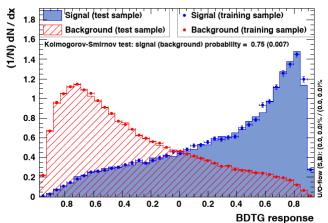


BDT parameters

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



TMVA overtraining check for classifier: BDTG



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 21ss 40 tZq
thqMVA ttv 21ss 40 ttZ
thqMVA ttv 21ss 40 VVV
thqMVA ttv 21ss 40 ttW
thqMVA_ttv_2lss_40_data_fakes
thoMVA ttv 21ss 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 21ss 40 WZ



Backgrounds and signal histograms

Backgrounds

tZ, VVV,tttt,WW:

- thqMVA_ttv_2lss_40_tZq
- thqMVA_ttv_21ss_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]





α values for post fit

Floating Parameter FinalValue +/- Error

Lumi
alpha_sample_B_sys
alpha_sample_F_sys
alpha_sample_H_sys
alpha_sample_T_sys
alpha_sample_W_sys
alpha_sample_Z_sys
mu

1.0000e+00 +/- 1.00e-04 2.0316e-01 +/- 9.90e-01 5.2474e-01 +/- 8.11e-01 1.1984e-01 +/- 9.92e-01 9.1741e-01 +/- 9.44e-01 1.2297e-01 +/- 9.89e-01 7.3604e-02 +/- 9.76e-01 5.2836e+00 +/- 1.17e+01



α values for post fit the k_t =-1

Floating Parameter

FinalValue +/- Error

Lumi
alpha_sample_B_sys
alpha_sample_F_sys
alpha_sample_H_sys
alpha_sample_T_sys
alpha_sample_W_sys
alpha_sample_Z_sys
mu

1.0000e+00 +/- 1.00e-04 2.0279e-01 +/- 9.87e-01 5.2541e-01 +/- 8.02e-01 1.1946e-01 +/- 9.90e-01 9.1778e-01 +/- 9.36e-01 1.2155e-01 +/- 9.73e-01 7.2960e-02 +/- 9.67e-01 6.1286e-01 +/- 1.39e+00



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

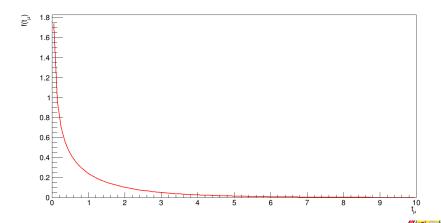


Figure: Statistic test plot $f(t_{\mu})$ vs t_{μ} with $t_{\mu} = -2 \ln \lambda(\mu)$



P value

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

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

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



P value

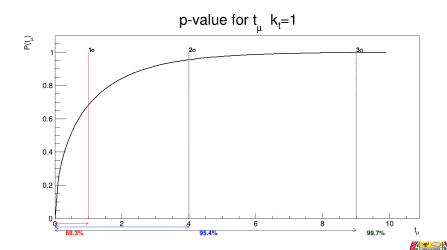


Figure: P_{μ} vs f($t_{\mu}|\mu$)