Probing the CP nature of the Higgs' couplings in $t\bar{t}H$ events at the LHC

Duarte Azevedo
Supervised by: António Onofre (antonio.onofre@cern.ch)
Frank Filthaut (F.Filthaut@science.ru.nl)

Departamento de Física
Universidade do Porto

May 26, 2017

FACULDADE DE CIÊNCIAS
UNIVERSIDADE DO PORTO

Radboud Universiteit Nijmegen

LIP

Outline

Higgs Mechanism

Where did Electroweak Theory failed?

- Not possible to add *ad hoc* mass terms to the EW Lagrangian without explicitly breaking $SU(2)_L \times U(1)_Y$ invariance.

$$\begin{aligned}
\psi_R &\xrightarrow{SU(2)_L\times U(1)_Y} \psi'_R = \exp\left(i\beta\hat{Y}\right)\psi_R \\
\psi_L &\xrightarrow{SU(2)_L\times U(1)_Y} \psi'_L = \exp\left(i\alpha^i\hat{\tau}_i + i\beta\hat{Y}\right)\psi_L \\
W^a_\mu &\xrightarrow{SU(2)_L\times U(1)_Y} W^{a\prime}_\mu = W^a_\mu + f_{bc}{}^a\alpha^b W^c_\mu + \partial_\mu\alpha^a \\
B_\mu &\xrightarrow{SU(2)_L\times U(1)_Y} B'_\mu = B_\mu + \partial_\mu\beta
\end{aligned} \tag{1}$$

Mass terms would not be invariant

$$\begin{aligned}
&m_f(\overline{\psi}_L\psi_R + h.c.) \to m_f\left[\overline{\psi}_L \exp(i(Y_R - Y_L) - i\alpha_i\hat{\tau}^i)\psi_R + h.c.\right] \neq m_f(\overline{\psi}_L\psi_R + h.c.) \\
&\frac{1}{2}m_X^2 X^\mu X_\mu \to \frac{1}{2}m_X^2(X_\mu + \partial_\mu\beta)(X^\mu + \partial^\mu\beta) \neq \frac{1}{2}m_X^2 X_\mu X^\mu
\end{aligned} \tag{2}$$

thus conservation of weak isospin and hypercharge would not be possible.

- Unitarity is broken for high energies (∼TeV). Consider for instance $WW \to ZZ$ scattering, the cross section scales like

$$\sigma(WW \to ZZ) \propto E^2 \tag{3}$$

Higgs Mechanism

New particles

New Mechanism presented in 1964 independently by R. Brout, F. Englert, P. Higgs, G. Guralnik, C. R. Hagen, T. Kibble.

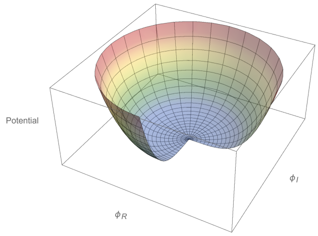
- New field $\Phi(x) = \begin{bmatrix} \phi_a(x) \\ \phi_b(x) \end{bmatrix}$ being a $SU(2)_L$ doublet and $U(1)_Y$ singlet with a non-trivial vacuum expectation value.
- Yukawa couplings between Φ and fermions.

With the following Lagrangian density

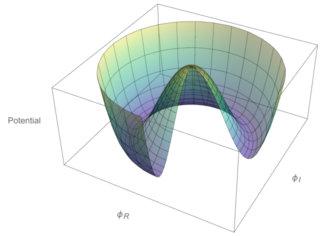
$$\mathcal{L}[\Phi] = (D^\mu \Phi)^\dagger (D_\mu \Phi) - \mu^2 (\Phi^\dagger \Phi) - \lambda (\Phi^\dagger \Phi)^2 \tag{4}$$

where the covariant derivative is given by

$$D^\mu = \partial^\mu + i g_W \hat{\tau}_a W_a^\mu + i g_B \hat{Y} B^\mu \tag{5}$$



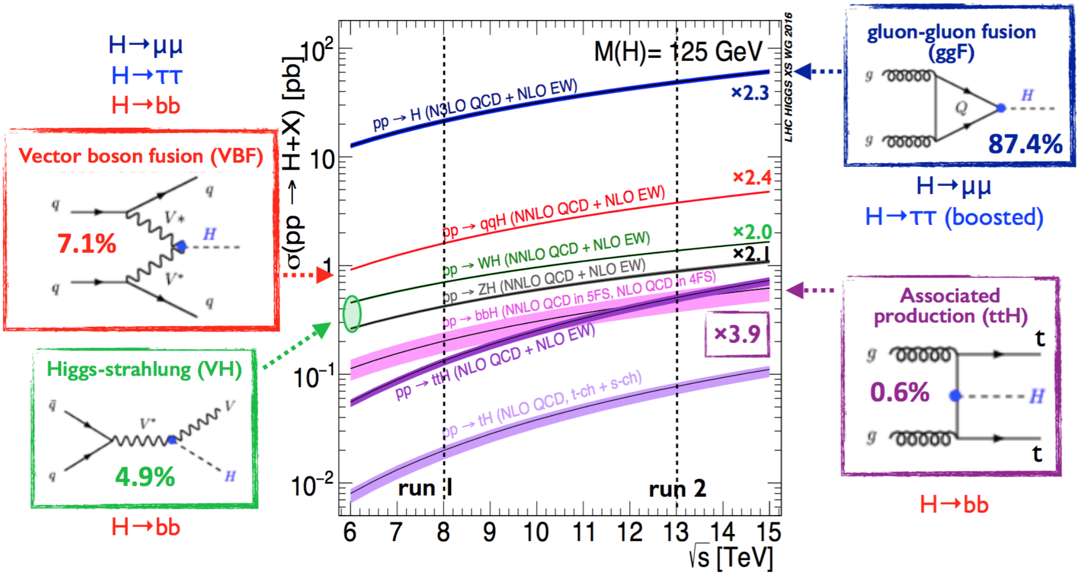
(a) $\mu^2 > 0$



(b) $\mu^2 < 0$

Figure: Higgs potential, simplified with just two degrees of freedom, before and after phase transition. The plot (A) represents an unbroken state whereas (B) represents a broken one.

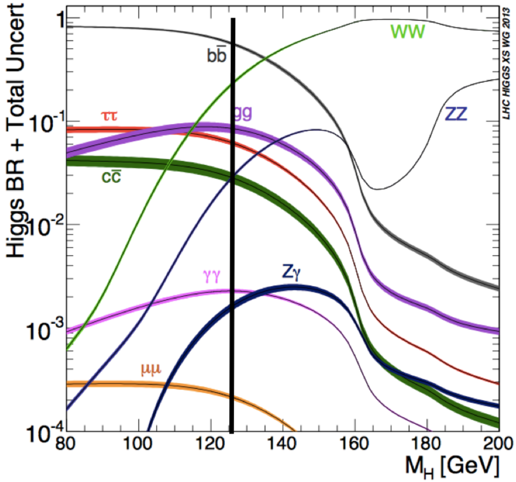
Higgs Production



- VH and ttH production mechanisms haven't been observed yet
- These are the most promising channels to observe Higgs to bottom coupling

Higgs Decays

Branching ratios vs. Higgs mass



Decay channel	Branching ratio [%]
$H \to b\bar{b}$	57.5 ± 1.9
$H \to WW$	21.6 ± 0.9
$H \to gg$	8.56 ± 0.86
$H \to \tau\tau$	6.30 ± 0.36
$H \to c\bar{c}$	2.90 ± 0.35
$H \to ZZ$	2.67 ± 0.11
$H \to \gamma\gamma$	0.228 ± 0.011
$H \to Z\gamma$	0.155 ± 0.014
$H \to \mu\mu$	0.022 ± 0.001

SM BR theory uncertainties
2-5% for most important decays

The width of the Higgs is too small!

$t\bar{t}H$ Channel

Motivation

A SM Lagrangian generalization, without explicit symmetry breaking, is considering a non-pure scalar Higgs:

$$\mathcal{L}_{HQ} = g_i^q \overline{\psi}_i(x)(a_i + ib_i\gamma^5)\psi_i(x)\sigma(x) \qquad (6)$$

understanding the nature of the coupling ($h = H, A$) may solve the matter/antimatter problem, among other.

Why $t\bar{t}H$?

- The top quark has the strongest coupling to the Higgs SM boson ($g_t \sim 1$).
- $\Gamma_t^{SM} = 1.42\text{GeV} \rightarrow$ top quark decays before hadronizing. Decay products are highly correlated:

$$\frac{1}{\sigma}\frac{d^2\sigma}{d(\cos\theta_+)d(\cos\theta_-)} = \frac{1}{4}(1 + B_1\cos\theta_+ + B_2\cos\theta_- - C\cdot\cos\theta_+\cdot\cos\theta_-) \qquad (7)$$

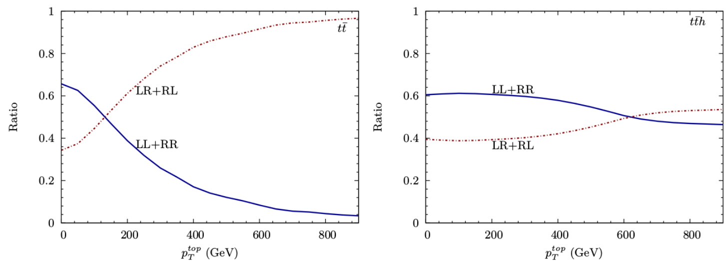


Figure: Ratio of the spin correlation terms for pure top pair- and higgs associated production.

arXiv **1403.1790v1** (2014) S. Biswas, R. Frederix, E. Gabrielli and B. Mele

$t\bar{t}H$ Channel

Semileptonic final states

At the LHC ($\sqrt{s} = 13$ TeV) it is expected that 43.8% [Patrignani, C. and others, 2016 , Particle Data Group, Chin. Phys.] of the top pairs decay to the single lepton channel.

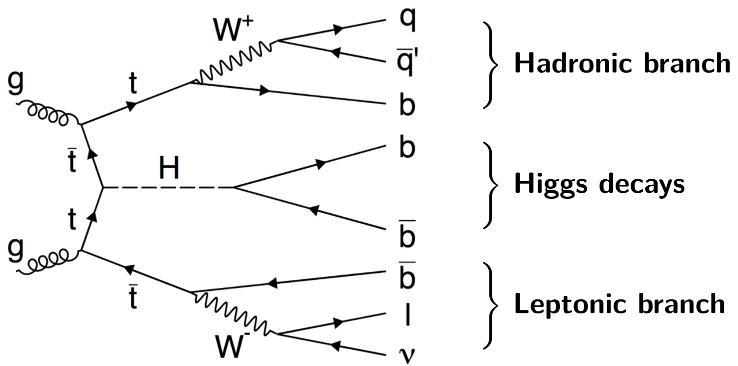
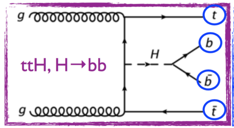


Figure: Leading-order $t\bar{t}H$ single lepton topology diagram.

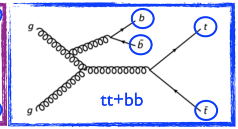
Background Contributions

Measurements by ATLAS @ 13 TeV [ATLAS-CONF-2016-080 (2016) Atlas Collaboration]

Signal



Background



- Single Lepton Channel

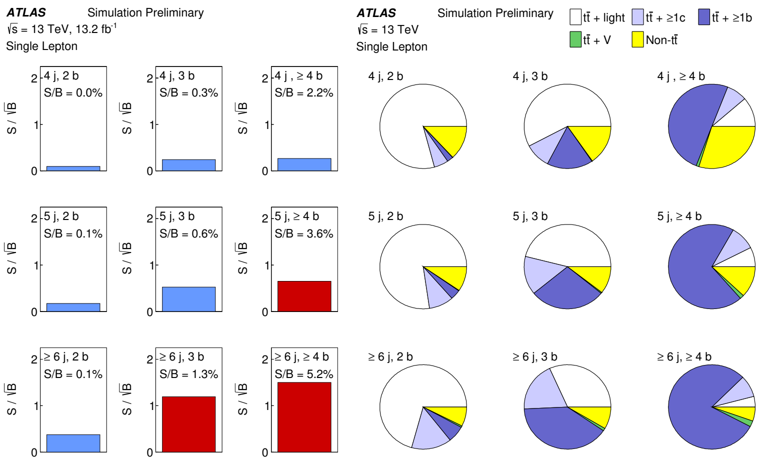


Figure: Analysis regions for the semilepton channel. Each row corresponds to a different jet multiplicity, while each column corresponds to a different b-jet multiplicity.

Generation and Analysis Outline

Single lepton topology: $t\bar{t}H \to (b\bar{b})\left[(bl^+\nu_l)(\bar{b}q\bar{q}) \quad \text{or} \quad (bq\bar{q})(\bar{b}l^-\bar{\nu}_l)\right]$

Event Generation @ 13 TeV:

- **MadGraph5_aMC@NLO** JHEP **1407**, 079 (2014) J. Alwall *et al.*
  †NNPDF2.3 PDF NPB **867**, 079 (2013) R. D. Ball *et al.*
  for $t\bar{t}h$, $h = A, H$ and $t\bar{t}b\bar{b}$ (@NLO)
  other backgrounds @ LO with MLM:
  $t\bar{t}$ + jets, $t\bar{t}V$ + jets, Single top,
  $W(Z)$ + jets, VV + jets
- **MadSpin** JHEP **1303**, 015 (2013) P. Artoisenet *et al.*
  full spin correlations for $t \to bW^+ \to b(l^+\nu||q\bar{q})$
  $\bar{t} \to \bar{b}W^+ \to \bar{b}(l^-\bar{\nu}||q\bar{q})$, $h \to b\bar{b}$
- **Pythia 6** JHEP **0605**, 026 (2006) T. Sjostrand, S. Mrenna, P. Z. Skands
  showering and hadronization

Simulation: DELPHES 3 JHEP **1402**, 057 (2014)

J. de Favereau, C. Delaere, P. Demin, A. Giammanco, V. Lemaître, A. Mertens, M. Selvaggi

Analysis: MadAnalysis5 EPJC **74, no 10, 3103 (2014)**

E. Conte, B. Dumont, B. Fuks, C. Wymant

Kinematic reconstruction: KLFitter NIM **A748** (2014) 18-25

J. Erdmann, S. Guindon, K. Kroeninger, B. Lemmer, O. Nackenhorst, A. Quadt, P. Stolte

$$6 \leq N_{\text{jets}} \leq 8(p_T \geq 20\text{GeV}, |\eta_{\text{light}}| \leq 4.5, |\eta_b| \leq 2.5) \oplus N_{\text{lep}} = 1(p_T \geq 20\text{GeV}, |\eta_{\text{lep}}| \leq 2.5)$$

Analysis

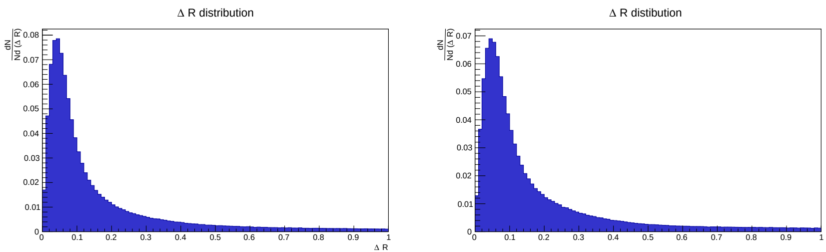
Truth-Match Reconstruction

Associate detected with hard scattering objects with the least angular distance

$$\Delta R_i^\xi = \sqrt{(\eta^\xi_{\text{jet,i}} - \eta_{\text{parton,i}})^2 + (\phi^\xi_{\text{jet,i}} - \phi_{\text{parton,i}})^2}$$
$$\Delta R^\xi_{\text{total}} = \sum_i \Delta R_i^\xi \equiv f(\xi) \quad (8)$$

The solution permutation $\xi^\star$ is the one that minimizes $f(\xi)$. A matched object is considered valid if

$$\Delta R_i^{\xi^\star} \leq 0.4 \quad (9)$$



(a) Hadronic B quark **(b)** Light quark

Figure: Example ΔR distributions of the truth matched reconstructed objects. Most are withing the matching solid angle but some are inevitably mismatched.

Truth-Match Reconstruction

Performance

Leptonic b-quark	χ^2	χ^2_B	χ^2_M	χ^2_{MB}	χ^2_{iM}	χ^2_{iMB}
P_x	1.32	5.18	0.82	5.18	1.53	2.02
P_y	7.94	1.25	1.42	0.91	1.41	2.86
P_z	1.17	1.79	1.74	1.67	1.18	1.23
E	2.67	1.69	3.08	0.77	0.81	2.01
Light quark	χ^2	χ^2_B	χ^2_M	χ^2_{MB}	χ^2_{iM}	χ^2_{iMB}
P_x	1.23	1.26	0.49	0.49	0.69	0.70
P_y	0.98	0.96	0.71	0.69	0.41	0.42
P_z	3.01	3.21	4.11	4.26	4.02	4.10
E	0.55	0.53	0.42	0.43	0.66	1.06

Table: Pearson's reduced χ^2 test values of the closeness of fit of the momentum components of the leptonic b quark and light quark.

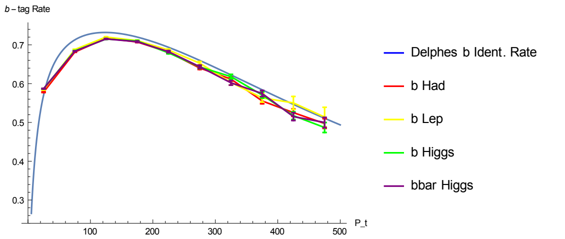
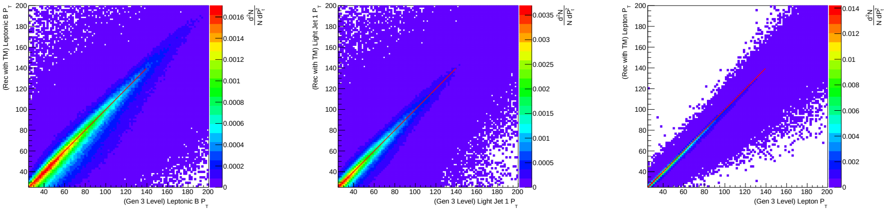


Figure: Delphes b-jet identification rate and ratio of b-tag for matched b-jets imposing the solution to be isolated for comparison.



(a) Leptonic B quark **(b)** Light quark **(c)** Lepton

Figure: Truth-match reconstructed vs. generator level transverse momenta of different particles.

KLFitter

Maximum Likelihood Estimate

Likelihood

$$
\begin{aligned}
L(\theta|y) = & B(m_{(\mathrm{bHad,LJ1,LJ2})}|m_{\mathrm{top}},\Gamma_{\mathrm{top}}) \cdot B(m_{(\mathrm{LJ1,LJ2})}|m_W,\Gamma_W) \\
& \times B(m_{(\mathrm{bLep},l,\nu)}|m_{\mathrm{top}},\Gamma_{\mathrm{top}}) \cdot B(m_{(l,\nu)}|m_W,\Gamma_W) \\
& \times B(m_{(\mathrm{bH1,bH2})}|m_H,\Gamma_H) \\
& \times \prod_{i=1}^{6} W_{\mathrm{jet}}(E_i^{\mathrm{meas}}|E_i^{\mathrm{true}}) \cdot W_l(E_l^{\mathrm{meas}}|E_l^{\mathrm{true}}) \\
& \times W_{\mathrm{miss}}(E_{\mathrm{miss,x}}^{\mathrm{meas}}|E_{\nu,x}^{\mathrm{true}}) \cdot W_{\mathrm{miss}}(E_{\mathrm{miss,y}}^{\mathrm{meas}}|E_{\nu,y}^{\mathrm{true}})
\end{aligned}
\tag{10}
$$

Two effects are accounted for:

- Radiation.
- Detector resolution.

Transfer Functions

- Obtained with the truth match analysis.
- Parametrization of the energy changes of the electrons, jets and MET (P_t for muons). Double gaussian is the best fit.

KLFitter

Transfer Functions

Leptons

$$G(X^{\text{meas}}, X^{\text{true}}) = \frac{1}{\sqrt{2\pi\sigma^2(X^{\text{true}})}} \exp\left[\frac{(X^{\text{meas}} - X^{\text{true}})^2}{2\sigma^2(X^{\text{true}})}\right] \quad (11)$$

where

- $X = E$ for electrons.
- $X = P_t$ for muons.

Jets and MET NIM **A748** (2014) 18-25

J. Erdmann, S. Guindon, K. Kroeninger, B. Lemmer, O. Nackenhorst, A. Quadt, P. Stolte

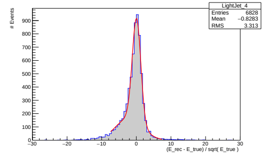
$$D(R) = \frac{1}{\sqrt{2\pi}} \left\{ \frac{C_1}{\sigma_1} e^{-\frac{(R-\mu_1)^2}{2\sigma_1^2}} + \frac{C_2}{\sigma_2} e^{-\frac{(R-\mu_2)^2}{2\sigma_2^2}} \right\} \quad (12)$$

- Jets

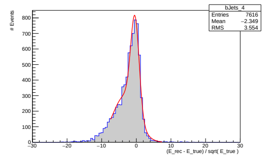
$$R(E_{\text{jet},i}, E_{\text{parton},i}) = \frac{E_{\text{jet},i} - E_{\text{parton},i}}{\sqrt{E_{\text{parton},i}}} \quad (13)$$

- MET

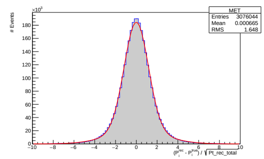
$$R(P_{x,y}^{\nu_{\text{gen}}}, P_{x,y}^{\text{miss}}, \psi \equiv \sum_{\text{rec}} P_t) = \frac{P_{x,y}^{\text{miss}} - P_{x,y}^{\nu_{\text{gen}}}}{\psi^{1/2}} \quad (14)$$



(a) Light quark



(b) Bottom quark



(c) MET

Figure: Transfer function examples.

KLFitter

Performance

Overall results:

- Selected topology impacts the reconstruction efficiency.

	Reconstruction Efficiency [%]				
Topology	Overall	b_{had}	b_{lep}	W_{had}	Higgs
6 Jets (4-btagged)	54.4	84.8	54.4	98.3	63.0
7 Jets (4-btagged)	26.2	60.6	55.7	79.5	44.3
6 Jets (3-btagged)	41.4	59.4	62.7	73.3	52.0
6-8 Jets (3-4 btagged)	**30.1**	**51.2**	**57.1**	**68.0**	**43.1**

Table: Reconstructed efficiencies of the Maximum Likelihood Estimate method on a KLFitter implementation. The values represent the *reconstructed efficiency* which is defined as the fraction of matched events for which the chosen permutation is indeed the correct one.

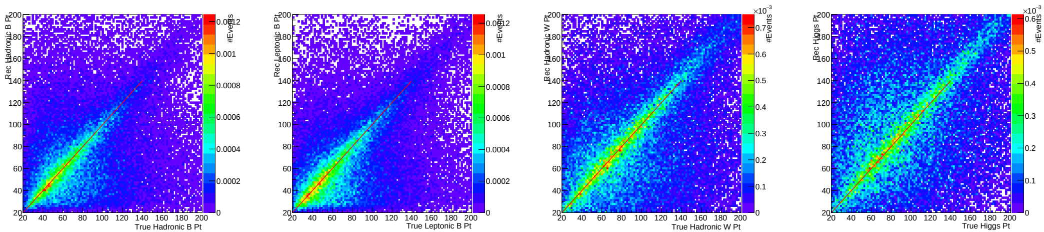


Figure: Non-truth match reconstructed vs. generator level transverse momenta for the selected topology. Left to right: hadronic b-quark, leptonic b-quark, hadronic W boson, Higgs boson.

Event Yields

Cut Flow 1/2

Parton State	Generated Cross section (pb)	$N_j > 6$ & $N_l = 1$ Cross section (pb)	E/P_t & η cuts Cross section (pb)
(NLO) $t\bar{t}b\bar{b}$	4.708×10^{0}	8.984×10^{-1}	7.470×10^{-1}
(LO) $t\bar{t}$ + up to $3j$	2.393×10^{2}	2.488×10^{1}	2.061×10^{1}
(LO) $t\bar{t}V$ + up to j	3.243×10^{0}	7.490×10^{-2}	6.498×10^{-2}
(LO) sT (s-chan)	2.192×10^{0}	7.773×10^{-3}	6.219×10^{-3}
(LO) sT (t-chan)	4.686×10^{1}	4.852×10^{-1}	3.725×10^{-1}
(LO) w + up to $4j$	3.450×10^{4}	3.293×10^{0}	2.779×10^{0}
(LO) wbb + up to $2j$	2.893×10^{2}	7.097×10^{-1}	5.648×10^{-1}
(LO) ww + up to $3j$	8.424×10^{1}	1.927×10^{-1}	1.627×10^{-1}
(LO) wz + up to $3j$	3.793×10^{1}	9.420×10^{-2}	7.926×10^{-2}
(LO) zz + up to $3j$	1.100×10^{1}	8.662×10^{-3}	5.962×10^{-3}
(NLO) $t\bar{t}H$	1.384×10^{-1}	2.661×10^{-2}	2.237×10^{-2}
(NLO) $t\bar{t}A$	5.822×10^{-2}	1.893×10^{-2}	1.524×10^{-2}

Table: Generated Background cross sections and successively applied cuts. The cross sections for each cut is presented. For comparison there is the two Higgs signals at the end of the table.

Event Yields

Cut Flow 2/2

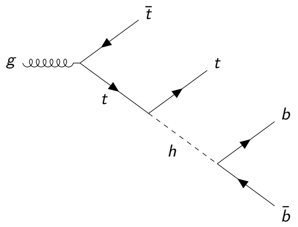
Parton State	Prev. Cuts + $N_j \leq 8$ & $3 \leq N_b \leq 4$ Cross section (pb)	Efficiency
(NLO) $t\bar{t}b\bar{b}$	1.656×10^{-1}	3.518×10^{-2}
(LO) $t\bar{t}$ + up to $3j$	5.655×10^{-1}	2.362×10^{-3}
(LO) $t\bar{t}V$ + up to j	4.133×10^{-3}	1.276×10^{-2}
(LO) sT (s-chan)	1.508×10^{-4}	6.88×10^{-5}
(LO) sT (t-chan)	4.780×10^{-3}	1.023×10^{-4}
(LO) w + up to $4j$	0	0
(LO) wbb + up to $2j$	3.716×10^{-3}	1.286×10^{-5}
(LO) ww + up to $3j$	0	0
(LO) wz + up to $3j$	4.529×10^{-4}	1.195×10^{-5}
(LO) zz + up to $3j$	5.095×10^{-5}	4.632×10^{-6}
(NLO) $t\bar{t}H$	8.846×10^{-3}	6.394×10^{-2}
(NLO) $t\bar{t}A$	6.067×10^{-3}	1.042×10^{-1}

Table: Surviving events' cross sections after all cuts were applied. The efficiency, defined as the ratio of "all cuts" cross sections with its original value, of each background and signal is also shown.

CP Sensitive Variables

Angular Distributions

From the helicity formalism...



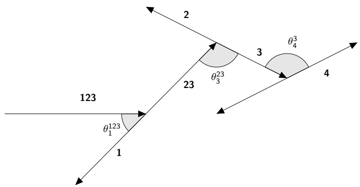


Figure: $t\bar{t}H$ production as a decay chain, in the helicity formalism.

Considered family of functions:

$$f(\theta_1^{123})g(\theta_4^3) \quad \text{and} \quad f(\theta_3^{23})g(\theta_4^3)$$

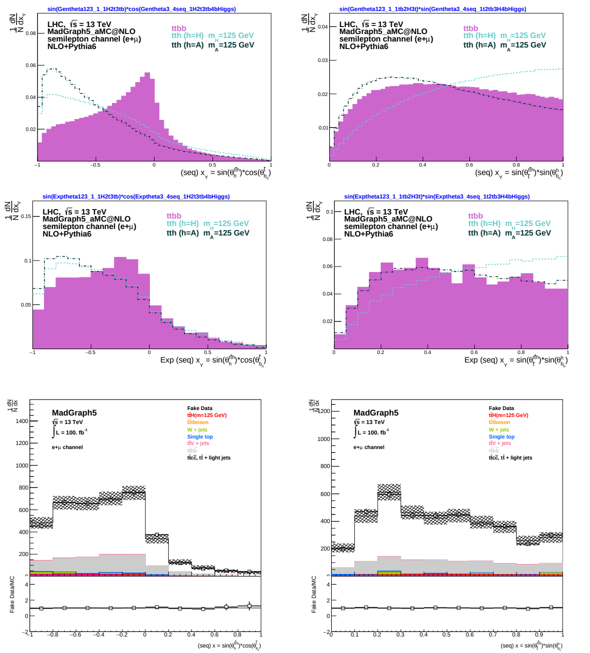
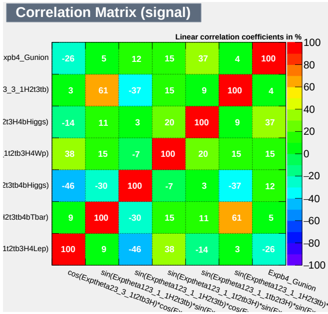


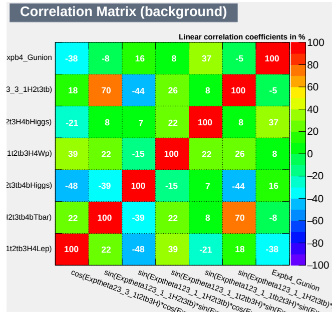
Figure: Angular distributions at generator, model levels and with background.

Multivariate Analysis

Chosen variables should be the least correlated.



(a) Signal



(b) Background

Figure: Correlation matrices of the set of chosen variables for (A) signal and (B) background events.

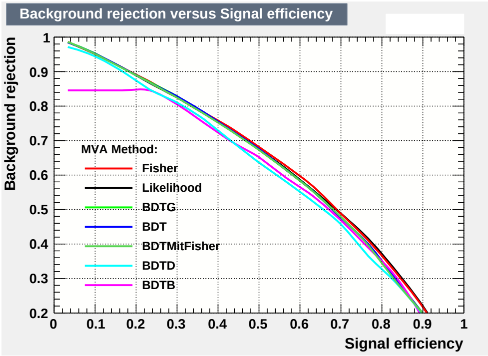
Using the set of best angular variables is possible to create a singular probability distribution function that contrasts the most between signal and background. Methods used:

- Likelihood.
- Fisher's Method.
- Several Boosted Decision Trees (BDT).

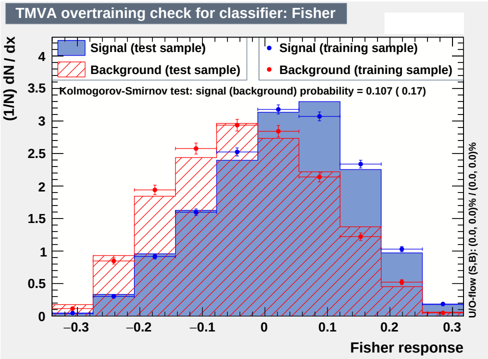
Multivariate Analysis

Best Results

Fisher's Method



(a) ROC curve



(b) Overtraining with Fisher

Figure: Receiver operating characteristic curve (A) and Multivariate training sample and over-training test sample for the Fisher methods (B). The background is represented in red and the signal in blue. Points and uncertainties represent the training samples, filled bins represent the test samples.

95% CL Limits

Signal Strength

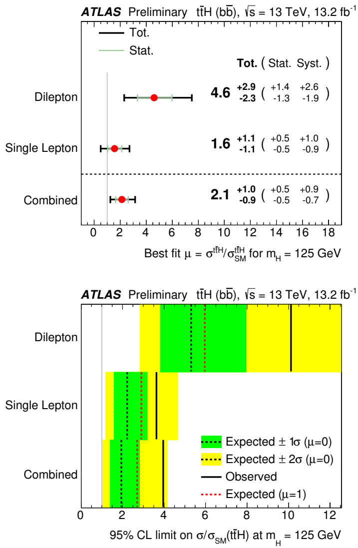


Figure: Signal strength (top) and ratio of observed and expected cross sections (bottom) for $t\bar{t}H(b\bar{b})$ production at $\sqrt{s} = 13$, $\mathcal{L}_I = 13.2$ fb^{-1}.
ATLAS-CONF-2016-080 (2016) ATLAS

The 95% confidence level (CL) limit is given when

$$1 - CL = \frac{\int_0^{X_d} P_{s+b}(X)dX}{\int_0^{X_d} P_b(X)dX} = 0.05 \qquad (15)$$

Signal strength:

$$\mu = \frac{\sigma(t\bar{t}H(A)) \times \text{Br}(H(A) \to b\bar{b})}{\sigma(t\bar{t}H)_{\text{SM}} \times \text{Br}(H \to b\bar{b})_{\text{SM}}} \qquad (16)$$

Signal strength for other luminosities:

$$\mu_{\text{new}} = \mu_{\text{old}} \cdot \sqrt{\frac{\mathcal{L}_{\text{old}}}{\mathcal{L}_{\text{new}}}} \qquad (17)$$

For $\mathcal{L}_I = 100$ fb^{-1} the expected strength is

$$\mu^{\text{fit}}_{\text{extrap}} = 0.5813 \qquad (18)$$

The value computed from my analysis

$$\mu = 0.7092 \qquad (19)$$

95% CL Limits

Fisher

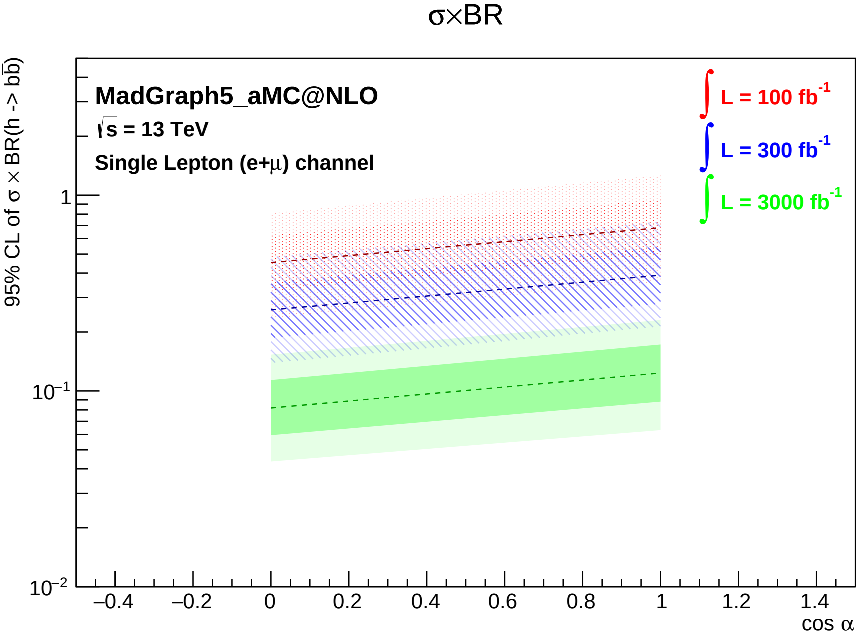


Figure: Obtained 95% CL limits in the background only scenario, $\pm 1\sigma$ and $\pm 2\sigma$ bands for the cross section of $t\bar{t}H(b\bar{b})$ for different luminosities. The value at zero correspond to the pure pseudo-scalar Higgs (A), the value at one to the SM one (H).

What was accomplished

- Phenomenological analysis in the $t\bar{t}H$ single lepton channel.
- Several CP samples generated with MadGraph5, as well as the dominant background processes at the LHC @ 13 TeV.
- Implementation of KLFitter for the $t\bar{t}H$ channel.
- Found best set of angular variables for CP measure.
- Computed the 95% confidence level limits in the absence of signal.