On $t\bar{t}H$ differential measurements at the LHC

Andrea Giammanco, Hesham El Faham, Pietro Vischia

Solstice meeting, ULB, 19^{th}December 2019

Centre for Cosmology, Particle Physics and Phenomenology (CP3),
Université Catholique de Louvain, Belgium

Outlook

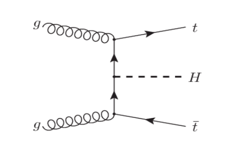
- Why $t\bar{t}H$ production is important?
- How $t\bar{t}H$ differential measurements are important?
- Generic overview of the experimental analysis in CMS.
- Explain our current standing and show some preliminary results.
- Summarize and touch on near-future plans.

Why $t\bar{t}H$ is important?

In the simplest terms

Recall:

- Higgs boson couples to fermions in a Yukawa-type interaction: $y_f \propto m_f$
- The top quark is the heaviest fermion we know, i.e. not kinematically possible to look for $H \to t\bar{t}$ so we look for $t\bar{t}H$ production.



But why knowing $y_{f=t}$ is important?

- y_t should be the strongest compared to other couplings.
- y_t is sensitive to new undiscovered particles, i.e new physics, and therefore,

  if y_t agrees to what we predict then its good $\to$ we understand.

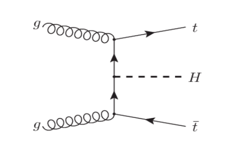
  if y_t doesn't agree to what we predict then its good $\to$ we don't understand.

  if y_t agrees with what we predict within uncertainty $\to$ need more data, precision measurements, precise calculations, etc.

In the simplest terms

Recall:

- Higgs boson couples to fermions in a Yukawa-type interaction: $y_f \propto m_f$
- The top quark is the heaviest fermion we know, i.e. not kinematically possible to look for $H \rightarrow t\bar{t}$ so we look for $t\bar{t}H$ production.



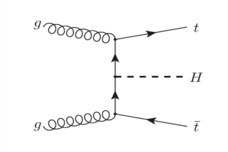
But why knowing $y_{f=t}$ is important?

- y_t should be the strongest compared to other couplings.
- y_t is sensitive to new undiscovered particles, i.e new physics, and therefore,
    - if y_t agrees to what we predict then its good $\rightarrow$ we understand.
    - if y_t doesn't agree to what we predict then its good $\rightarrow$ we don't understand.
    - if y_t agrees with what we predict within uncertainty $\rightarrow$ need more data, precision measurements, precise calculations, etc.

In the simplest terms

Recall:

- Higgs boson couples to fermions in a Yukawa-type interaction: $y_f \propto m_f$
- The top quark is the heaviest fermion we know, i.e. not kinematically possible to look for $H \to t\bar{t}$ so we look for $t\bar{t}H$ production.



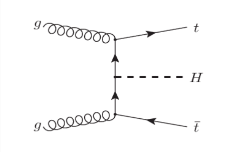
But why knowing $y_{f=t}$ is important?

- y_t should be the strongest compared to other couplings.
- y_t is sensitive to new undiscovered particles, i.e new physics, and therefore,
  - if y_t agrees to what we predict then its good $\to$ we understand.
  - if y_t doesn't agree to what we predict then its good $\to$ we don't understand.
  - if y_t agrees with what we predict within uncertainty $\to$ need more data, precision measurements, precise calculations, etc.

In the simplest terms

Recall:

- Higgs boson couples to fermions in a Yukawa-type interaction: $y_f \propto m_f$
- The top quark is the heaviest fermion we know, i.e. not kinematically possible to look for $H \to t\bar{t}$ so we look for $t\bar{t}H$ production.



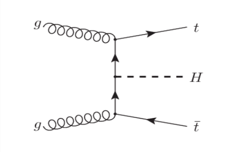
But why knowing $y_{f=t}$ is important?

- y_t should be the strongest compared to other couplings.
- y_t is sensitive to new undiscovered particles, i.e new physics, and therefore,
    - if y_t agrees to what we predict then its good $\to$ we understand.
    - if y_t doesn't agree to what we predict then its good $\to$ we don't understand.
    - if y_t agrees with what we predict **within uncertainty** $\to$ need more data, precision measurements, precise calculations, etc.

In the simplest terms

Recall:

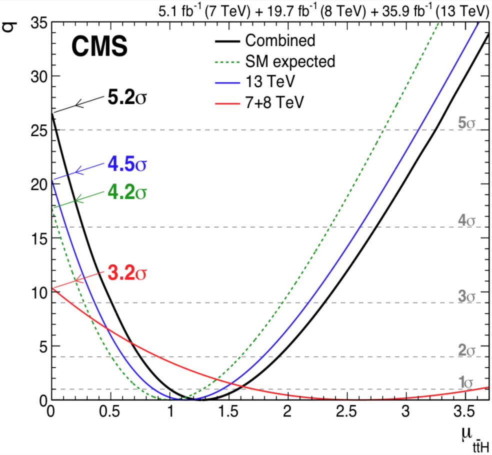
- Higgs boson couples to fermions in a Yukawa-type interaction: $y_f \propto m_f$
- The top quark is the heaviest fermion we know, i.e. not kinematically possible to look for $H \to t\bar{t}$ so we look for $t\bar{t}H$ production.



But why knowing $y_{f=t}$ is important?

- y_t should be the strongest compared to other couplings.
- y_t is sensitive to new undiscovered particles, i.e new physics, and therefore,
    - if y_t agrees to what we predict then its good $\to$ we understand.
    - if y_t doesn't agree to what we predict then its good $\to$ we don't understand.
    - if y_t agrees with what we predict **within uncertainty** $\to$ keep hiring PhD students...

The observation of $t\bar{t}H$ at CMS CMS-HIG-17-035



- Excess of events is observed with significance of 5.2σ over the expectation from the background-only hypothesis.
- Reported signal strength $\mu_{\bar{t}tH} = 1.26^{+0.31}_{-0.26}$, with SM expectation being $\mu_{\bar{t}tH} = 1$.

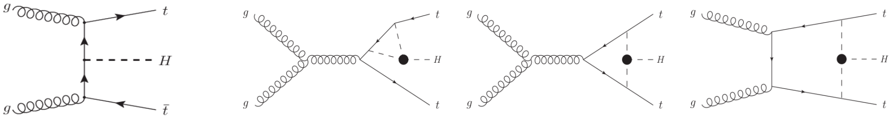
How $t\bar{t}H$ differential measurements are important?

$t\bar{t}H$ differential measurements constrain λ_3

arxiv 1709.08649v2 and 1607.04251v3

$$V(H) = \frac{1}{2}m_H^2 H^2 + \lambda_3 v H^3 + \frac{1}{4}\lambda_4 H^4 + O(H^5) \quad (1)$$

At low energies, new physics might alter the trilinear coupling, λ_3. Single Higgs processes are sensitive to λ_3 via one-loop corrections.



A generic observable, Σ

$$\Sigma_{\lambda_3}^{BSM} \propto Z_H^{BSM}\Sigma_{LO}(1 + k_3 C_1), \qquad \lambda_3 = k_3\lambda_3^{SM}, \qquad Z_H^{BSM} = \frac{1}{1-(k_3^2-1)\delta Z_H.}^{1} \quad (2)$$

$$\Sigma_{\lambda_3}^{SM} \propto \Sigma_{LO}(1 + C_1), \qquad k_3 \to 1 \quad (3)$$

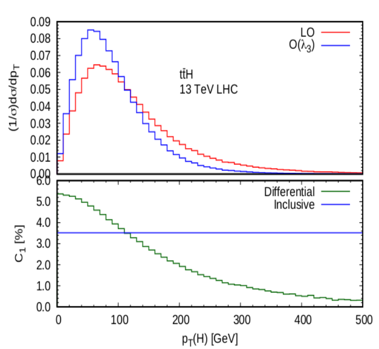
and C_1 is process- and kinematic-dependent component.

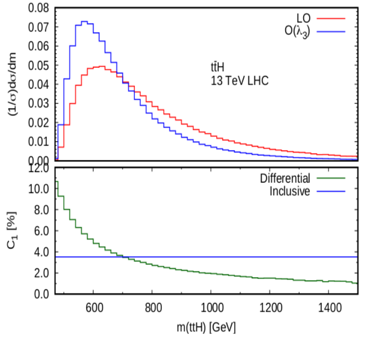
[1] Corresponds to wave-function renormalization where new physics is resummed at one loop.

$t\bar{t}H$ differential measurements constrain λ_3

arxiv 1709.08649v2

Differential distributions have non-flat dependence on λ_3, i.e. relative corrections due to λ_3 is $\propto C_1$,

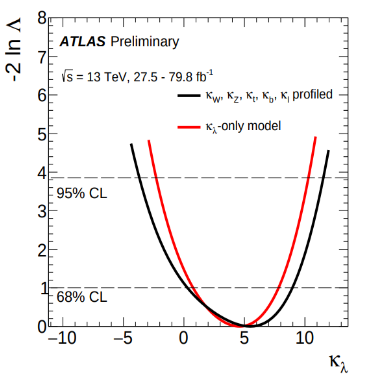




- C_1 at inclusive level for $t\bar{t}H$ is 3.52%.
- $C_1 \sim 5\%$ for P_T and $\sim 10\%$ for invariant mass distributions on differential level.

Indirect effects from modified Higgs self-coupling are significant in the $t\bar{t}H$ channel.

Experimental limits on λ_3 ATLAS-CONF-2019-049



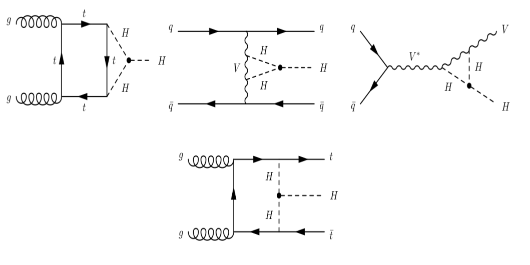


Fig: examples of one loop λ_3-dependent diagrams in single-Higgs production. Double-Higgs diagrams are not shown.

- LHC global fit of single and double Higgs analyses combined.
- k_λ-only model[2] uses the assumption that new physics is expected to appear only as a modification of Higgs self-coupling.
- More generic models (profiled parameters) allows to test BSM models that modify other Higgs couplings at the same time.
- For $t\bar{t}H$ production no differential information was used.

[2] k_λ here is equivalent to the coupling modifier, k_3, shown earlier.

To SMEFT

arxiv 1612.00283v1

Standard Model Effective Field theory (SMEFT): the SM augmented by higher dimension operators encapsulating new physics effects at scale Λ well above the electroweak scale,

$$\mathcal{L}_{EFT} = \mathcal{L}_{SM} + \Sigma_i \frac{C_i^{(6)}}{\Lambda^2} \mathcal{O}_i^{(6)} + \mathcal{O}^{(8)}(\Lambda^{-4}) + h.c. \tag{4}$$

At 1-loop level for single Higgs processes, anomalous coupling and SMEFT frameworks are equivalent[3] *in determining* λ_3 *via* $\mathcal{O}_6 = -\lambda(\phi^\dagger\phi)^3$.

Kinematical distributions provide an important handle on the determination of Higgs properties

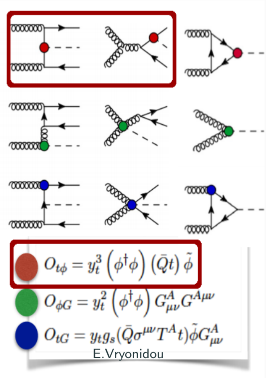
[3] Single Higgs processes are not sensitive to λ_4 at one loop.

$t\bar{t}H$ differential measurements constrain SMEFT

See arxiv 1802.07237v1 for EFT analysis recipe and 1607.05330v2 for the discussion on $t\bar{t}H$

Constrain the $t\bar{t}H$ relevant operators by probing Top-Higgs interaction

- Interested in the re-scaling of the Yukawa coupling in the SM



Define observables at the particle-level in fiducial phase-space.

For each observable, compute the linear and the quadratic contribution of the dim-6 operator.

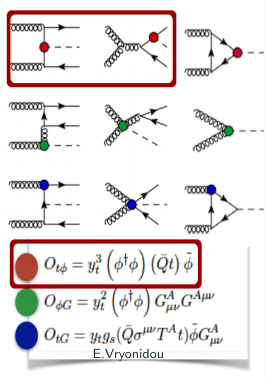
Compare predictions with the measured observable and the estimated SM contribution to constrain the Wilson coefficients.

$t\bar{t}H$ differential measurements constrain SMEFT

See arxiv 1802.07237v1 for EFT analysis recipe and 1607.05330v2 for the discussion on $t\bar{t}H$

Constrain the $t\bar{t}H$ relevant operators by probing Top-Higgs interaction

- Interested in the re-scaling of the Yukawa coupling in the SM



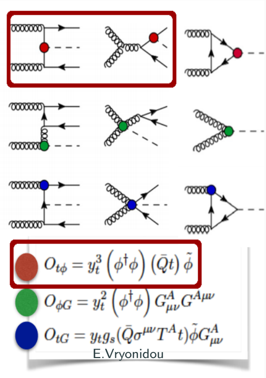
- Define observables at the particle-level in fiducial phase-space.
- For each observable, compute the linear and the quadratic contribution of the dim-6 operator.
- Compare predictions with the measured observable and the estimated SM contribution to constrain the Wilson coefficients.

$t\bar{t}H$ differential measurements constrain SMEFT

See arxiv 1802.07237v1 for EFT analysis recipe and 1607.05330v2 for the discussion on $t\bar{t}H$

Constrain the $t\bar{t}H$ relevant operators by probing Top-Higgs interaction

- Interested in the re-scaling of the Yukawa coupling in the SM



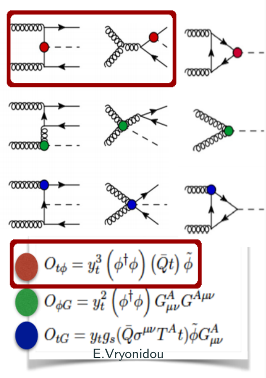
- Define observables at the particle-level in fiducial phase-space.
- For each observable, compute the linear and the quadratic contribution of the dim-6 operator.
- Compare predictions with the measured observable and the estimated SM contribution to constrain the Wilson coefficients.

$t\bar{t}H$ differential measurements constrain SMEFT

See arxiv 1802.07237v1 for EFT analysis recipe and 1607.05330v2 for the discussion on $t\bar{t}H$

Constrain the $t\bar{t}H$ relevant operators by probing Top-Higgs interaction

- Interested in the re-scaling of the Yukawa coupling in the SM

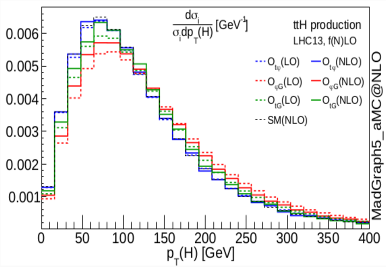


- Define observables at the particle-level in fiducial phase-space.
- For each observable, compute the linear and the quadratic contribution of the dim-6 operator.
- Compare predictions with measured observable and the estimated SM contribution to constrain the Wilson coefficients.

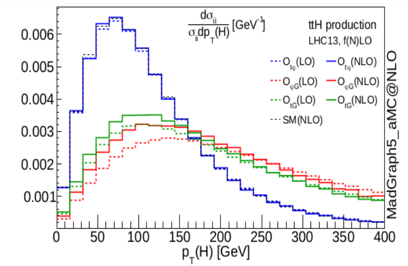
$t\bar{t}H$ differential measurements constrain SMEFT

arxiv 1607.05330v2

Interference contribution



Squared contribution



$$\sigma = \sigma_{SM} + \Sigma_i \frac{C_i^{(6)}}{\Lambda^2}\sigma_i(1\mathrm{TeV}^2) + \Sigma_{ij}\frac{C_i^{(6)}C_j^{(6)}}{\Lambda^4}\sigma_{ij}(1\mathrm{TeV}^4) \quad (6)$$

- $p_T(H)$ is a discriminating observable in a differential measurement.
- Squared contributions can be used to distinguish between different operators contributions. Squared contributions are subdominant to interference ones.

$t\bar{t}H$ in multilepton final states at CMS (HIG-18-019)

Introduction and selection cuts

This analysis focuses on $t\bar{t}H$ final states with electrons, muons, and hadronically decaying tau leptons $\rightarrow$ we will focus here on two leptons same-sign final state (2*lss*).

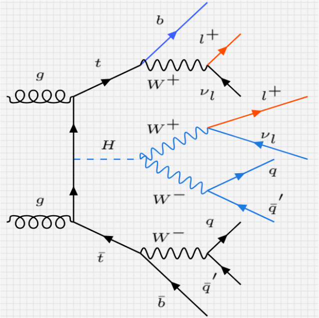


Fig: $t\bar{t}H$ in $2lss$.

Selection:

- At least two b-jets are of $p_T > 25$ GeV and $|\eta| < 2.4$.
- Reject events with dilepton mass $m_{ll} < 12$ GeV.
- Sub(leading) lepton $p_T > 15$ GeV (> 25 GeV).
- No lepton pair with mass close to Z boson.
- $L_D >$ [4] 30 GeV.
- Events with a hadronic tau are vetoed.

[4] A variant of missing transverse momentum that is designed for optimal trade-off between discrimination and sensitivity to pileup.

Background estimation

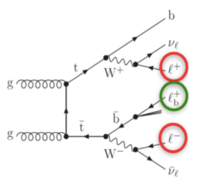
- **Reducible background:** one or more leptons passing the object selection are affected by reconstruction.
    - Fake background: misidentified leptons estimated from data.
    - Charge-flip background: charge of one lepton is mismeasured, estimated from data ($Z \to e^+e^-$).
- **Irreducible background:** backgrounds whose final state has the same particle content as the signal. Dominated by $t\bar{t}V$.

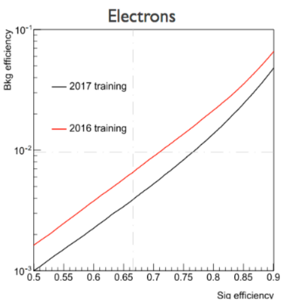
Category	2ℓss									
b-tag	no req.		Loose				Tight			
Leptons	ee		eμ		μμ		eμ		μμ	
Charge	−	+	−	+	−	+	−	+	−	+
tt̄H	3.3 ± 1.5	3.4 ± 1.8	6.5 ± 3.2	6.1 ± 3.0	4.2 ± 2.2	4.4 ± 2.2	4.56 ± 2.2	4.7 ± 2.5	3.1 ± 1.7	2.8 ± 1.4
tt̄W + tt̄WW	13.6 ± 2.4	20.8 ± 3.9	24.1 ± 4.5	40.8 ± 7.9	14.9 ± 2.6	25.9 ± 5.0	16.5 ± 2.8	31.5 ± 5.3	11.7 ± 2.0	18.7 ± 3.4
tH	0.1 ± 0.0	0.2 ± 0.0	0.3 ± 0.0	0.4 ± 0.0	0.3 ± 0.1	0.5 ± 0.0	0.2 ± 0.0	0.2 ± 0.0	0.1 ± 0.0	0.1 ± 0.0
WZ + ZZ	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
tt̄Z/γ*	17.9 ± 2.9	17.6 ± 2.7	20.0 ± 3.1	23.1 ± 3.8	7.6 ± 1.2	8.7 ± 1.4	14.8 ± 2.4	17.0 ± 2.7	5.9 ± 1.0	5.7 ± 0.9
Misidentified	7.9 ± 2.1	10.4 ± 2.5	26.6 ± 4.3	28.3 ± 4.8	20.6 ± 4.4	18.2 ± 3.9	5.6 ± 1.9	6.6 ± 1.9	3.7 ± 1.6	4.2 ± 1.8
Conversions	1.52 ± 0.5	2.2 ± 1.8	2.5 ± 0.9	3.5 ± 2.1	< 0.05	< 0.05	0.9 ± 0.3	1.1 ± 0.3	< 0.05	< 0.05
Signal flip	4.5 ± 1.1	4.6 ± 1.1	3.7 ± 0.9	3.8 ± 1.0	< 0.05	< 0.05	3.1 ± 0.8	3.1 ± 0.8	< 0.05	< 0.05
Other	1.4 ± 0.6	2.5 ± 1.1	4.6 ± 2.0	5.7 ± 2.2	1.7 ± 0.6	3.3 ± 1.3	2.2 ± 0.9	2.4 ± 1.0	1.1 ± 0.5	1.8 ± 0.8
SM expectation	50.2 ± 4.7	61.7 ± 6.1	88.3 ± 8.0	111.6 ± 10.9	49.2 ± 5.7	61.0 ± 7.0	47.8 ± 4.9	66.5 ± 6.9	25.6 ± 3.2	33.4 ± 4.3
Observed data	54	67	86	108	41	79	50	70	32	27

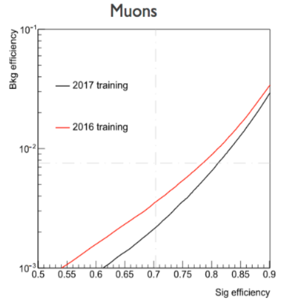
Tab: Selected number of events in the $2lss$ subcategories.

Optimization

Lepton MVA: Dedicated BDT[5] to discriminate prompt leptons coming from W and Z against fakes (non prompt) leptons from light mesons and b decays.







Hadronic top-tagger: BDT discriminator to compute the likelihood of a jet triplet to be compatible with a hadronic top decay.

[5] e.g. for MVA inputs: kinematics (p_T,η)

Systematics

Summary of the main source of systematics and their impact on the signal rate:

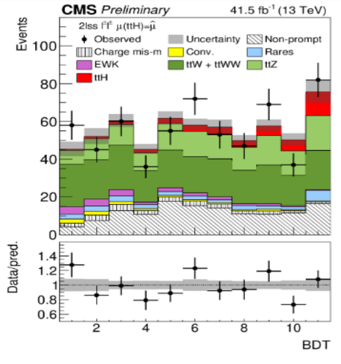
Source	Uncertainty [%]	$\Delta\mu/\mu$ [%] (2017)	$\Delta\mu/\mu$ [%] (Comb.)	Correlations
Theoretical sources	≈ 8	8	9	Correlated
e, μ selection efficiency	3–5	4	3	Correlated
τ_h selection efficiency	5	3	5	Correlated
τ_h energy calibration	1.2	1	2	Correlated
b tagging efficiency	2–15 [48]	10	5	Correlated
Jet energy calibration	2–15 [56]	3	3	Correlated
Fake background yield	$\approx$ 30–50	17	9	Un-correlated

- Rate of $t\bar{t}H$ signal, $\mu = \frac{\sigma_{obs}}{\sigma_{SM}}$, $\rightarrow$ binned Maximum Likelihood (ML) fit to the distribution of a discriminate variable.
- $t\bar{t}Z$ and $t\bar{t}W$ backgrounds are kept floating in the fit[6].

[6] Theory uncertainties on production cross sections don't affect the results.

Signal extraction

Both multivariate discriminant, namely BDTs and Matrix Element Method(MEM) were used for signal extraction:



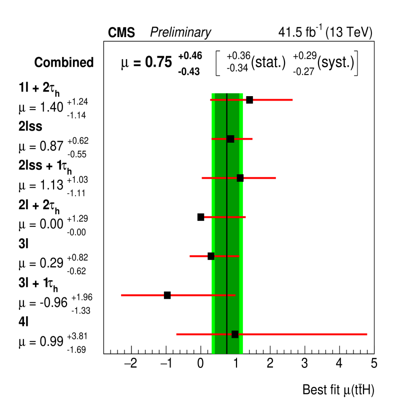


Fig: distributions in the discriminating observable (left) and measured signal rates, normalized to the SM $t\bar{t}H$ production rate (right).

Estimating the $pT(H)$: reconstruction

Reconstruction: the basic principle

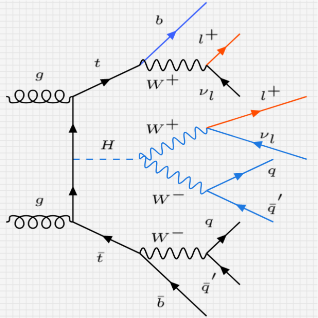


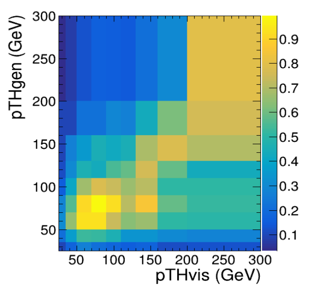
Fig:$t\bar{t}H$ in $2lss$.

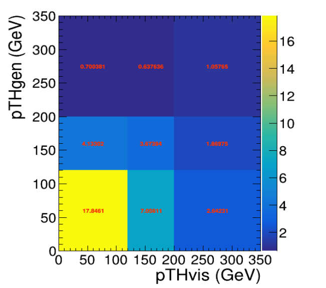
- $H \rightarrow W(l\nu)W(q\bar{q}')$
- Wide window for m_W and $m_{H_{vis}}$
- Reconstruct $pT(H_{vis})$: $(q\bar{q}') \leftarrow W \leftarrow H$ along with the "best" lepton, i.e. the closest to the two hadronic jets in the $\phi - \eta$ plane.
- Tag hadronic t and $\bar{t}$ decays via dedicated BDT hadronic top-tagger.
- Missing Transverse Energy (MET) assignment is expected to play a crucial role in this estimate.

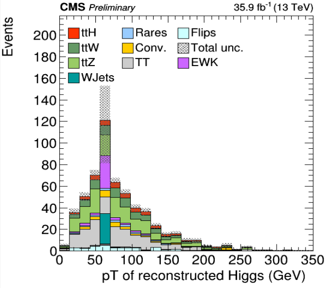
Current results based on 2016 MC

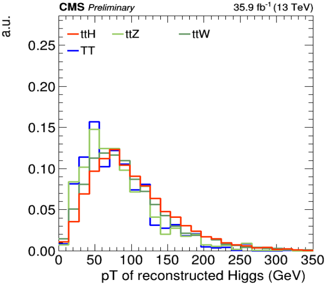
(2017 and 2018 results are in the backup slides)

Current results: 2016 MC samples



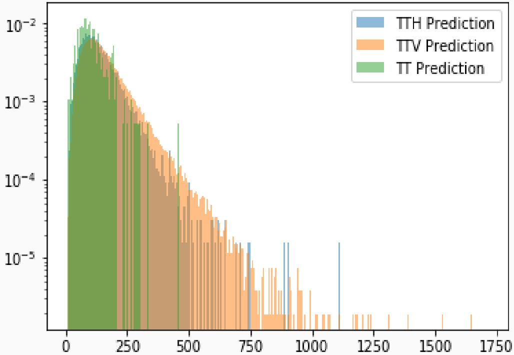






Current results: using neural networks

Regressing the $p_T(H)$ and evaluating the prediction in background events:



- Promising shape-discrimination ($t\bar{t}H$ vs $t\bar{t}V$ vs $t\bar{t}$), which resembles the one at generator level (plot shows discrete densities).
- Improving the network structure is work in progress.

Summary

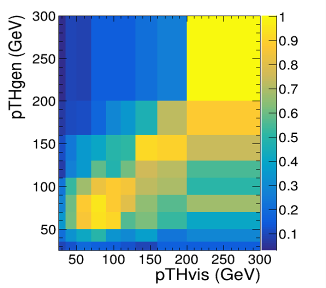
Summary

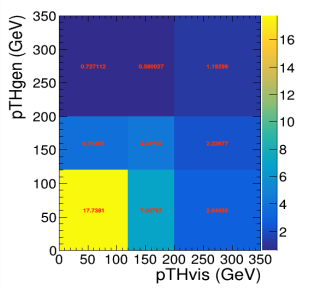
- Using differential information of $t\bar{t}H$ is crucial for λ_3 and SMEFT studies.
- The Higgs p_T plays a key role in determining the Higgs properties.
- Current $pT(H)_{vis}$ to true $pT(H)$ correlation $\sim$ 30%
- MET assignment in this estimate is crucial and is currently work in progress.
- Near future plan is to perform the sought-after full fiducial differential measurement followed by an EFT interpretation.

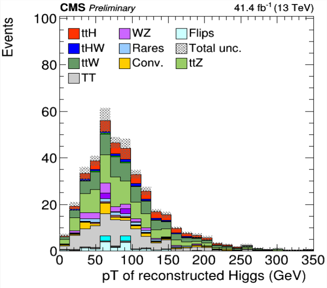
Thanks to Ken Mimasu and Vincent Lemaître for their feedback during the preparation of this talk.

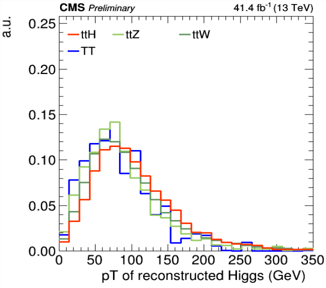
backup

Current results: 2017 MC samples

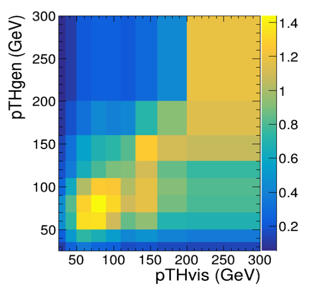


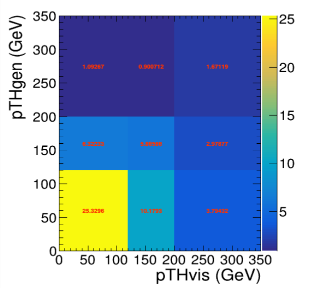


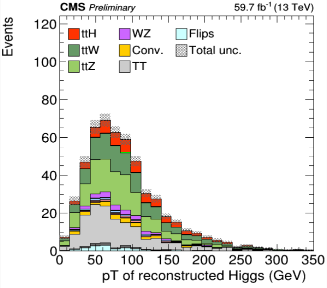


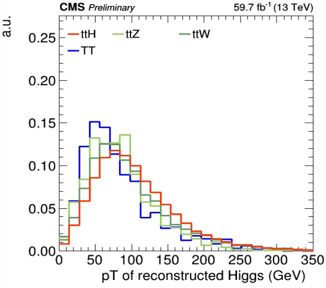


Current results: 2018 MC samples

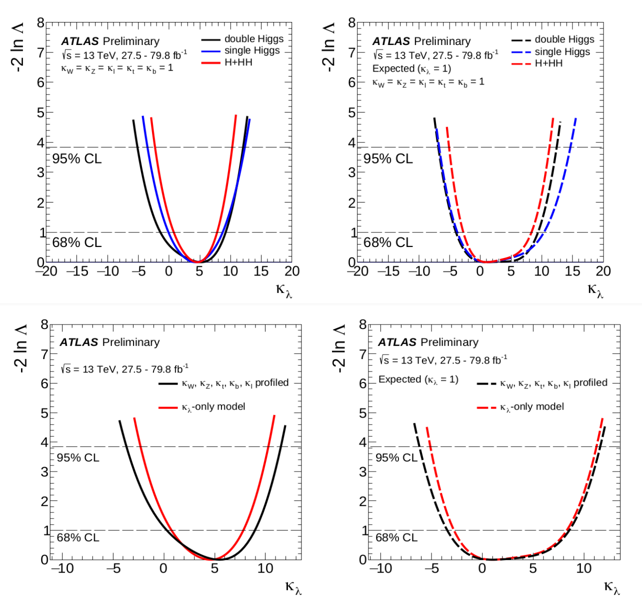








More on the experimental limits by ATLAS on λ_3 ATLAS-CONF-2019-049



The two observables p_T^{miss} and H_T^{miss} are combined into a single linear discriminant,

$$L_D = 0.6 \times p_T^{miss} + 0.4 \times H_T^{miss} \tag{1}$$

where p_T^{miss} is calculated as a the negative of the vector p_T sum of all particles reconstructed by the Particle Flow (PF) algorithm.

H_T^{miss} is the magnitude of the vectorial p_T sum of electrons, muons, τ_h, and jets,

$$H_T^{miss} = \|\Sigma_{leptons} \vec{P}_{Tl} + \Sigma_{\tau_h} \vec{P}_{T\tau} + \Sigma_{jets} \vec{P}_{Tj}\| \tag{2}$$

- Leptons,τ_h and jets predominantly originate from the hard scattering interaction and rarely from pileup interactions $\rightarrow$ H_T^{miss} less sensitive to variations in pileup conditions.
- The variable exploits the fact that p_T^{miss} and H_T^{miss} are less correlated in events in which the reconstructed p_T^{miss} is due to instrumental effects compare to genuine p_T^{miss} from neutrinos.
- The coefficients of the linear combination have been optimized to provide the best rejection against the Z+jets background.