Prospects for the measurement of $t\bar{t}H$ production in the opposite-sign dilepton channel at $\sqrt{s} = 14$ TeV at the High-Luminosity LHC

Charis Kleio Koraka - University of Wisconsin-Madison
On behalf of the CMS Collaboration

The CMS experiment © the HL-LHC

The Large Hadron Collider (LHC) is the most powerful collider built to date. It consists of two proton rings that run to collide in four different interaction points, including one where the CMS experiment is located.

The LHC has successfully delivered $\sim 140$ fb$^{-1}$ of proton-proton collision data during Run-2 (2016-2018). The High-Luminosity LHC (HL-LHC) upgrade is planned to increase the delivered instantaneous luminosity to a total of 3000 fb$^{-1}$ of data:
- It is important to study and improve the physics reach of the HL-LHC by exploring new analysis techniques.

The $t\bar{t}H$ production as a direct probe of the $y_t$

Top quark Yukawa coupling ($y_t$) is a fundamental parameter of the SM that plays a central role in Higgs phenomenology and is sensitive to physics beyond the Standard Model.

Why study the $t\bar{t}H$ associated production?
- Direct probe for measuring $y_t$
- $t\bar{t}H$ production cross-section controlled by $y_t$

Why $H \to b\bar{b}$ and leptonic decays (DL) of top quarks?
- $H \to b\bar{b}$ has the largest branching fraction.
- DL is the cleanest final state from background processes.

Higgs invariant mass reconstruction

- The final state particles measured by the detector consist of b-jets, leptons and missing energy.
- The mass reconstruction is performed by solving analytically the systems kinematic equations.
- For each object assignment permutation the proton momentum fractions are calculated:

\[ x_1 = \frac{1}{2}(E_{AH} - P_{AB}), \quad x_2 = \frac{1}{2}(E_{AH} + P_{AB}) \]

- The solution that best reconstructs the Higgs mass is selected based on the maximum PDF($x_1$, $x_2$) weight.

Data-driven background estimation

Modelling of the $b\bar{b}$ background is one of the dominant systematic uncertainties of the $t\bar{t}H$ measurement. Introducing a data-based background estimation:
- Allows to omit use of simulation.
- Has the potential to significantly reduce background-related uncertainties.

Tag-Rate-Function method (TRF)
- b-tagging not performed by a direct cut on the b-jet identification variable.
- Probability of each jet being b-tagged is estimated using parameterized efficiencies $\epsilon$ ($p_T$,$\eta$).

The shape and normalization of our observable can be extrapolated in regions with high b-jet multiplicity using events of regions with lower b-jet multiplicities by applying probability weights.
- The probability of an event containing N jets to contain M b-jets is estimated by the relation:

\[ P_M = \sum_{k=0}^{M} \prod_{n=1}^{k} (1 - \epsilon(p_T, \eta)) \prod_{n=k+1}^{M} \epsilon(p_T, \eta) \]

Two sources of systematic uncertainty are introduced:
- Rate uncertainty based on agreement of prediction in data and simulation.
- Shape of extrapolation factor.

Uncertainties extrapolation strategy

Different scenarios are employed to scale the uncertainties with the integrated luminosity ($L$):
- Statistical uncertainties only
- Run-2 uncertainties scenario
- Conservative scenario
- Realistic scenario

The analysis specific uncertainties depend on the number of data and simulated events and depending on the projection scenario are scaled as $1/\sqrt{L}$.

Analysis strategy

- Signal region is populated by events with at least 4 b-tagged jets.
- Control/validation regions:
  - Events with 2 b-tagged jets are used to extrapolate the background in the signal region.
  - Events with 3 b-tagged are used to validate the background prediction method.

The signal strength ($\mu$) is extracted in a template fit of the expected background and signal distributions to the Asimov data.

Results

- A measurement utilizing the full HL-LHC integrated luminosity can result to an $13\%$ total uncertainty on the signal cross section from the DL channel alone.
- A similar level of precision is achieved for measuring $y_t$ and will allow us to probe deviations from the SM expectation.

Conclusions

References: