Search for Higgs boson production in association with a top-antitop quark pair in CMS
**INTRODUCTION**

**HIGGS PRODUCTION @ LHC**

- Highest Higgs production rate due to gluon fusion
- Coupling to top quarks is involved in this process through top loops
  ⇒ Indistinguishable from new physics contributions

- Process predicted by SM with \( \sigma_{ttH} \propto g_{tH}^2 \)
- Experimentally very challenging
  - small cross section (0.5 pb)
  - complex estimation of backgrounds
INTRODUCTION

- Very rich experimental signatures
  - Top quarks can either decay hadronically or leptonically
- $b\bar{b}$ decay mode dominant for Higgs boson
- Leptonic decays through $W$, $Z$ and $\tau$ have a non-negligible contribution
- Golden channel $H \rightarrow \gamma\gamma$ provides very low rate
THE MULTILEPTONIC CHANNEL

- $H \rightarrow bb \Rightarrow$ highest branching ratio but overwhelming background
- $H \rightarrow \gamma\gamma \Rightarrow$ clear resonant signature, but low branching ratio
- Multileptonic channel is the middle point in the background size / branching ratio trade-off

- Irreducible backgrounds: $t\bar{t}W$ and $t\bar{t}Z$ $\sigma \sim \mathcal{O} (1 \text{ pb})$
- Reducible backgrounds: $t\bar{t}$ ($\sigma \sim 800 \text{ pb}$) (fake or non-prompt leptons)
**Event Selection**

- Two same-sign leptons (e or $\mu$)
- At least 4 jets
- Two “loose” or one “tight” $b$-tagged jets
- MET LD $> 0.2$ (ee)

- At least three leptons (e or $\mu$)
- At least two jets
- Two “loose” or one “tight” $b$-tagged jets
- MET LD $> 0.2$ if less than 5 jets
BACKGROUND REDUCTION

NON-PROMPT LEPTONS

- Non-prompt leptons those not coming from W or Z decay
  - misidentified jets or photons
  - $\mu$ produced in $b$ decays
- Have built lepton identification criteria optimized to reject non-prompt leptons
- BDT is used to discriminate between:
  - prompt leptons, coming from leptonic W, Z and $\tau$ decays
  - leptons coming from other sources: misidentified jets, $b$ quark decays

- Trained in $t\bar{t}H$ and $t\bar{t}$ samples
- Using observables related to the lepton reconstruction and variables related to particles reconstructed around the lepton

![Lepton BDT Score Distribution](image)
BACKGROUND SEPARATION

REDUCIBLE BACKGROUNDS

- BDTs in order to discriminate between $t\bar{t}H$, $t\bar{t}$ and $t\bar{t}V$
- Trained using kinematic variables of the objects, matrix element weights, BDTs targeting top and Higgs decays
**BACKGROUND SEPARATION**

**HADRONIC TOP AND HIGGS TAGGERS - 2LSS**

- Discriminant targeting hadronic top decays
- Constructed over all jet permutations in the event, trained against $t\bar{t}$

- Discriminant targeting $H \rightarrow WW \rightarrow l\nu + \text{jets}$ decay
- Constructed over all jet permutations in the event, trained against $ttV$
The matrix element weight is used as an input in the 3l channel.

Computationally expensive method.

Very well performant against $t\bar{t}W$.

$$w_{i,a}(\Phi') = \frac{1}{\sigma_a} \int d\Phi \cdot \sigma^4 \left( p_1^\mu + p_2^\mu - \sum_{k \geq 2} p_k^\mu \right) \cdot \frac{f(x_1, \mu_F) f(x_2, \mu_F)}{x_1 x_2^s} \cdot \left| M_a(p_k^\mu) \right|^2 \cdot W(\Phi' | \Phi_a)$$
BACKGROUND ESTIMATION

- Background due to nonprompt and charge misidentified leptons estimated using data-driven methods
- Nonprompt is estimated from control regions with loose lepton identification criteria
- Charge misidentification is estimated from control regions with two opposite-sign dilepton
- Contribution from $t \bar{t}W$ and $t \bar{t}Z$ estimated using MonteCarlo simulation
- Estimation validated in dedicated control regions
CATEGORIZATION AND SIGNAL EXTRACTION

CHANNELS

- Events are further categorized
  - dilepton, three lepton and four lepton channel
  - charge, lepton flavor (2l) and presence of tighter b-jets
Further categorization is performed in bins of a likelihood score.
The likelihood discriminant is a mapping of the 2D plane spanned by the two BDTs.
Results

- Measurement of the $t\bar{t}H$ production cross-section of a 50% uncertainty
- Observed (expected) significance w.r.t. $\mu(t\bar{t}H) = 0$ hypothesis: 3.3 (2.4) $\sigma$
HALLENGES FOR THE FUTURE

- Analysis very optimized for the current detector conditions
  - Using elaborate data-analysis techniques (BDT + likelihood discriminator)
  - ...trained with non-trivial variables (matrix elements, BDTs)
  - Using data-driven techniques for many of the main backgrounds
- ...leading us to a result of 3-$\sigma$-sensitivity analysis

- However this might not be enough to drive us to a 5-$\sigma$ observation with higher luminosity
- Simpler cut-and-count analysis might be more suitable to use for recast or reinterpretations
Thanks for your attention!
Questions?