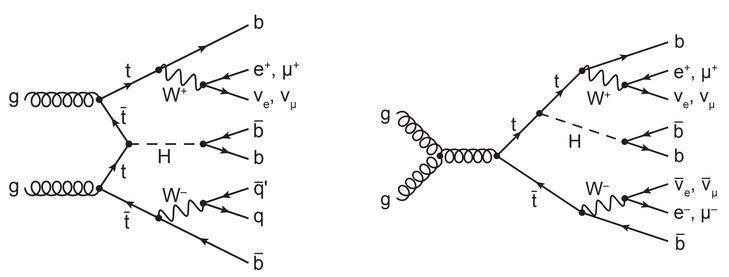
Search for $t\bar{t}H$ production in the $H \to b\bar{b}$ channel with leptonic $t\bar{t}$ decays in proton-proton collisions at $\sqrt{s}$=13 TeV

Andrej Saibel (DESY, Universität Hamburg)
on behalf of the CMS collaboration

Introduction

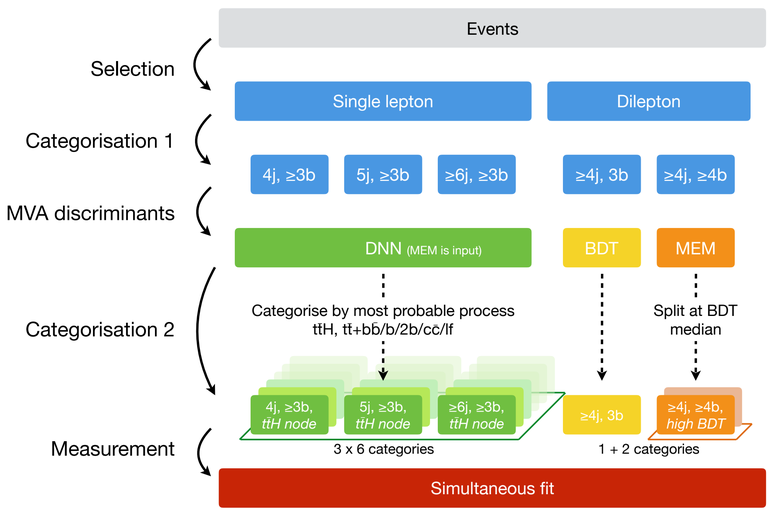


- $t\bar{t}H$: direct access to top-Higgs Yukawa coupling
- H->$b\bar{b}$: largest branching ratio
- Lepton requirement: QCD suppression

Reference

CMS Collaboration. Search for $t\bar{t}H$ production in the $H \to b\bar{b}$ decay channel with leptonic $t\bar{t}$ decays in proton-proton collisions at $\sqrt{s} = 13$ TeV. 2018. CMS-HIG-17-026, arXiv:1804.03682, Subm. to JHEP.

Analysis Strategy



Event Selection

Object	Semileptonic	Dileptonic
Leptons	$\lvert\eta\rvert < 2.1$	$\lvert\eta\rvert < 2.4$
$e^{\pm}$	$p_T > 30$ GeV	$p_T > 25(15)$ GeV
$\mu^{\pm}$	$p_T > 26$ GeV	"
Jets	$\geq$4, $\geq$2 b-tags	$\geq$2, $\geq$1 b-tags
	$p_T > 30$ GeV	$p_T > 30(20)$ GeV
	$\lvert\eta\rvert < 2.4$	$\lvert\eta\rvert < 2.4$
MET	> 20 GeV	> 40 GeV (ee,$\mu\mu$)

- b-tagging: Combined Secondary Vertex
- "$> 30(20)$ GeV": leading(subleading)

Challenges and Solutions

- $\sigma(t\bar{t}\text{+jets}) \approx 1600 \cdot \sigma(t\bar{t}H)$.
- Particular challenge: **$t\bar{t}$+b-jets** irreducible background with large modelling uncertainties.
- **Jet-b-tag categorization**: constrain backgrounds from background rich categories and extract signal strength from signal enriched categories.
- **MVA methods** for best possible signal-background discrimination

Multivariate Methods

- **Semilepton jet-process categorization**: In each of the jet-tag categories, each event gets assigned a probability for being $t\bar{t}H$,$t\bar{t}$+b-jets,$t\bar{t}$+cc, $t\bar{t}$+lf and classified according to highest probability by the **DNN**.
- **Dilepton** channel split into two jet-tag categories. In each category, a dedicated **BDT** used to separate $t\bar{t}H(b\bar{b})$ from background processes.
- **Dilepton** 4jet, 4b-tag category is split further into background- and signal-like subcategories according to BDT output. **Matrix Element Method**, constructed to separate $t\bar{t}H$ and $t\bar{t}$+$b\bar{b}$, is used as final discriminant in each subcategory.

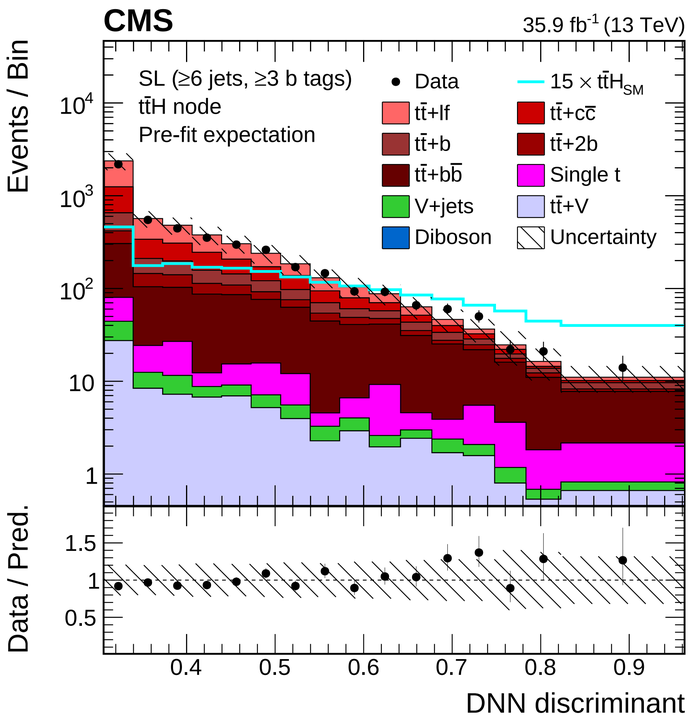


Figure 1: DNN discriminant in jet-process category $\geq$6 Jets-$t\bar{t}H$-Node (pre-fit).

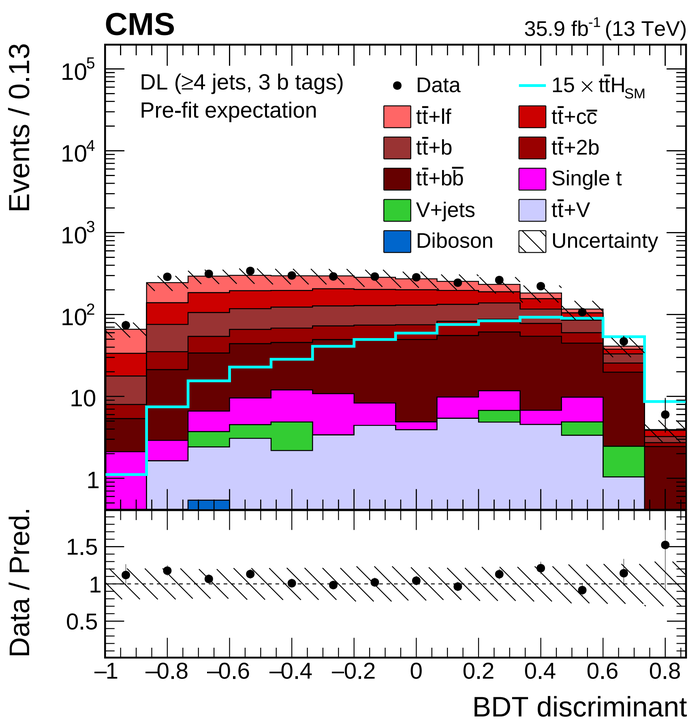


Figure 2: BDT discriminant in dilepton channel.

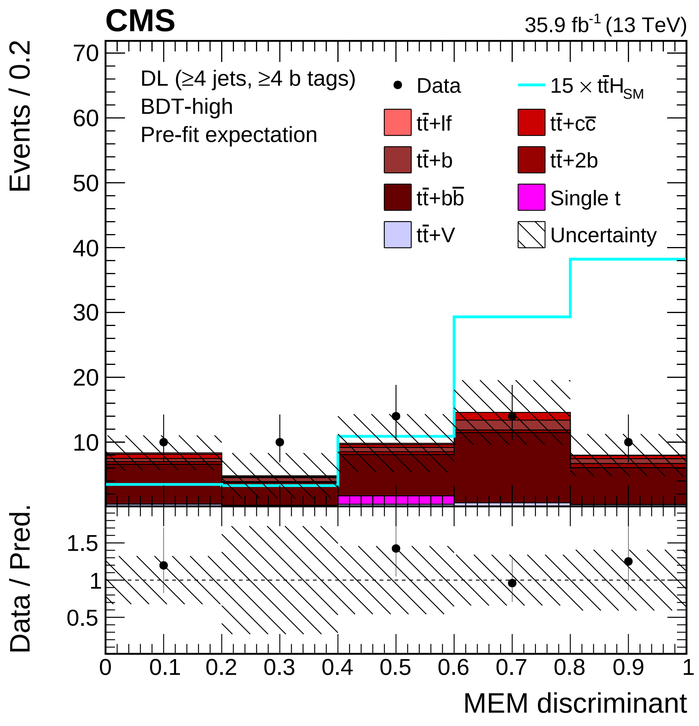


Figure 3: Matrix Element Method discriminant in high BDT output category (pre-fit).

Results

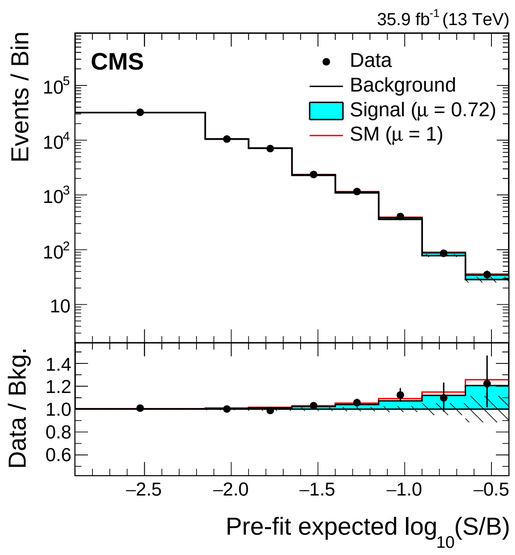


Figure 4: Bins of the final discriminants reordered by the pre-fit expected signal-to-background ratio.

- Best fit $\mu = 0.72 \pm 0.24(\text{stat}) \pm 0.38(\text{syst})$
- **1.6(2.2)σ** observed (expected) significance above background-only hypothesis.
- Measurement dominated by systematic uncertainties. Theoretical uncertainties on $t\bar{t}$+b-jets process by far the largest.
- Experimental uncertainties dominated by flavour tagging and Monte Carlo statistics.

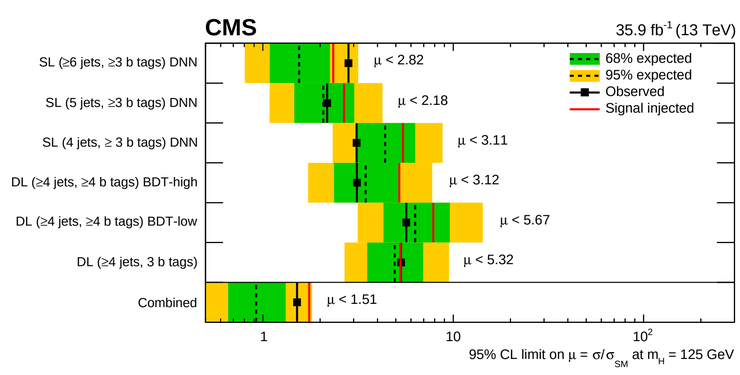


Figure 5: Median expected and observed upper limits on μ. Green and yellow lines indicate regions for 68% and 95% of expected limits under background-only hypothesis.