Simultaneous probe of the charm and bottom quark Yukawa couplings using ttH events

Marina Kholodenko

15th Course on Physics at the LHC 2026 - July 15th, 2026

https://arxiv.org/pdf/2509.22535

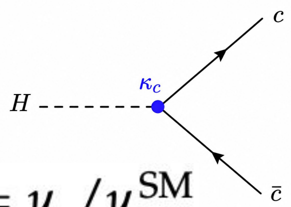
Motivation

Higgs couplings

- **Precise measurements of couplings**
  - The key to verify electroweak symmetry breaking
  - Or to uncover physics beyond the SM
- Higgs couplings to bosons, fermions and quarks, loop-induced effective couplings, higgs self-couplings

The charm quark Yukawa coupling y_c

- Measurements in $H \rightarrow c\bar{c}$ decay



$$\kappa_c = y_c / y_c^{SM}$$

<u>The last measurements:</u>

- VH associated production
  - ATLAS (Run 2) : 95% CL $|k_c|$<4.2 (4.1)
  - CMS (Run 2) : 1.1<$|k_c|$<5.5 ($|k_c|$<3.4)
- gg fusion production mode - large QCD multijet background

New approach:

- **$t\bar{t}H$ production**
- CMS Run 2 data
- Challenging in identifying jets origin and reconstruction top quark and Higgs

Main backgrounds

- **$t\bar{t}$ + jets**
  - Jets from from hadronization b and c quarks
    - b(c)-jets: at least one b(c-hadron)
    - Other jets - light jets
  - Event classification
    - **$t\bar{t}$+b, $t\bar{t}$+≥2b** ($t\bar{t}b\bar{b}$ MC sample, NLO, POWHEG-BOX-RES, OPENLOOPS)
    - **$t\bar{t}$+c, $t\bar{t}$+≥2c, $t\bar{t}$+light** ($t\bar{t}$ inclusive MC sample, NLO, POWHEG v2)
- **Single top quark in t (s) channels and Wt production**
  - NLO, POWHEG v2 (MADGRAPH5 aMC@NLO v2.6.5)
- **$t\bar{t}$W, $t\bar{t}$Z (tWZ)**
  - MADGRAPH5 aMC@NLO, NLO (LO)

Selections

Triggers:
High jet and b-jet multiplicities
Presence of one (two) isolated leptons

- p_T jets > 25 GeV
- $|\eta| < 2.4$

0L channel (fully hadronic)

- ≥ 7 jets (p_T > 40 GeV for six)
- p_T scalar sum > 500 GeV

1L channel (single lepton)

- One isolated lepton
  - p_T > 26-30 GeV
- ≥ 5 jets

2L channel (dilepton)

- Two oppositely charged leptons
  - One with p_T > 25 GeV
- ≥ 4 jets
- **To suppress Drell-Yan:**
  - $m_{ee/\mu\mu}$ > 20 GeV
  - Veto: $76 < m_{ee/\mu\mu} < 106$ GeV

Across all channels:

- ≥ 3 b(or c)-tagged jets
- ≥ 1 b-tagged jet

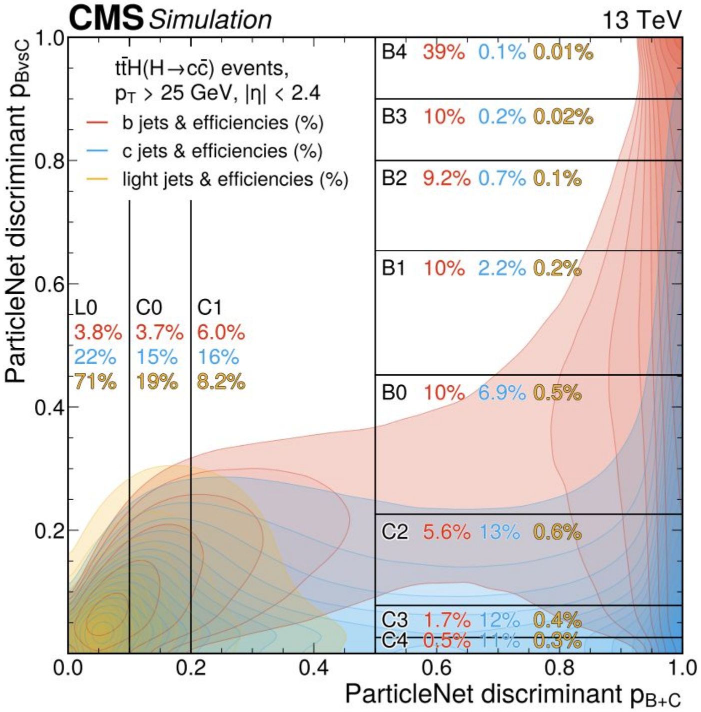
Selections

Selection criteria	0L channel	1L channel	2L channel
Number of leptons	0	1	2
Sign and flavor of leptons	—	$e^{\pm}$, $\mu^{\pm}$	$e^{+}e^{-}$, $\mu^{\pm}e^{\mp}$, $\mu^{+}\mu^{-}$
Min. p_T of p_T-leading electron [GeV]	—	29/30/30	25
Electron identification	—	Tight (80% efficiency)	Loose (90% efficiency)
Electron isolation	—	Tight (80% efficiency)	Loose (90% efficiency)
Min. p_T of p_T-leading muon [GeV]	—	26/29/26	25
Max. muon relative isolation	—	0.15	0.25
Min. $m_{ee/\mu\mu}$ [GeV]	—	—	20
Range of $m_{ee/\mu\mu}$ [GeV]	—	—	<76, >106
Min. p_T of jets [GeV]	25	25	25
Min. p_T of 6^{th} jet [GeV]	40	—	—
Max. $\|\eta\|$ of jets	2.4	2.4	2.4
Min. number of jets	7	5	4
Min. number of b-tagged jets	1	1	1
Min. number of HF (b- or c-tagged) jets	3	3	3
Min. H_T [GeV]	500	—	—
Min. p_T^{miss} [GeV]	—	20	20

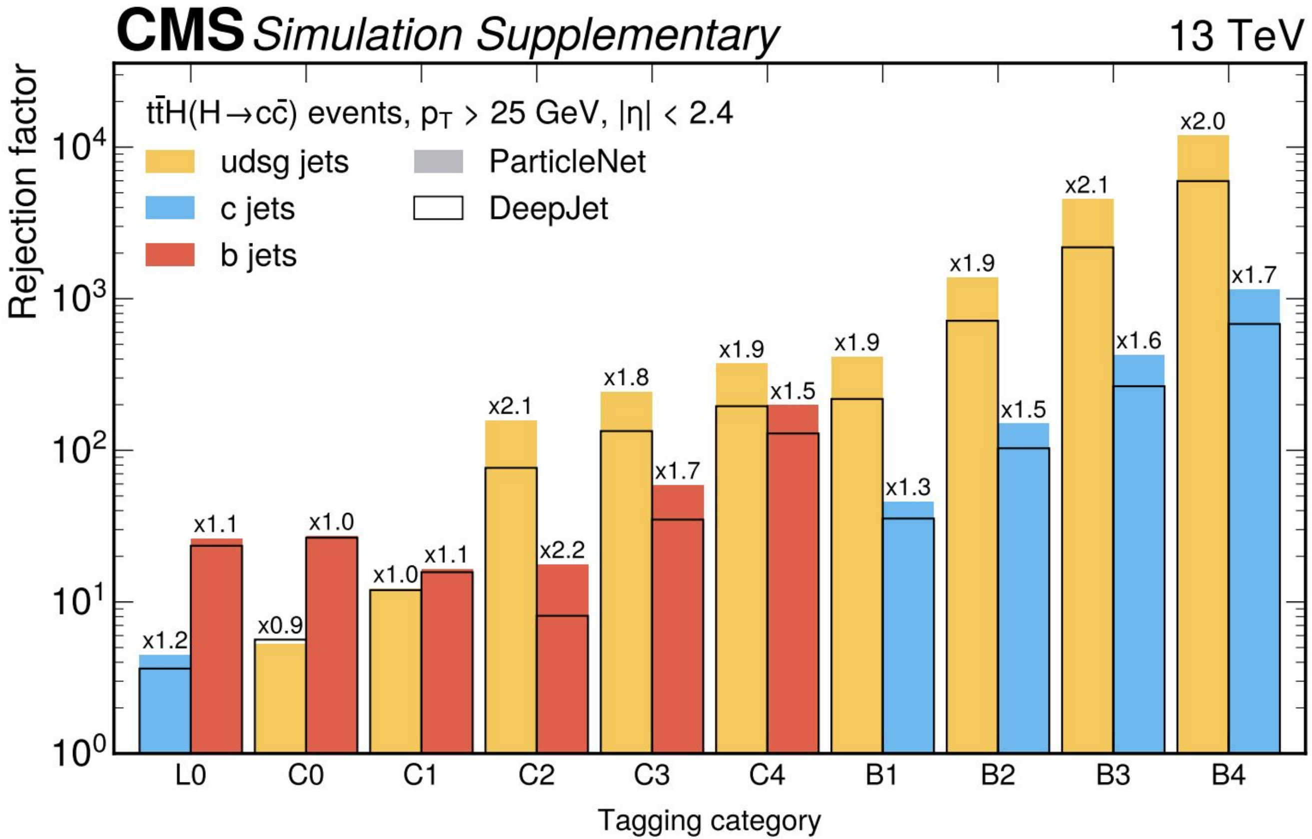
Jet flavour identification

PARTICLENET - graph convolutional neural network

- **Two discriminants:**
  - p_{B+C} - to differentiate heavy-flavour from light jets
  - p_{BvsC} - to distinguish b an c jets
- **11 exclusive tagging categories:**
  - 5 b-tagging (B0-B4)
  - 5 c-tagging (C0-C4)
  - One untagged (L0)
- **To correct mismodelling tagging efficiencies:**
  - Flavour-dependent SFs derived
  - Typically: 0.75 - 1.5
  - Uncertainties up to 15% fo b-jets, up to 50% for c and light jets



Rejection factors



Simulation-to-data SFs

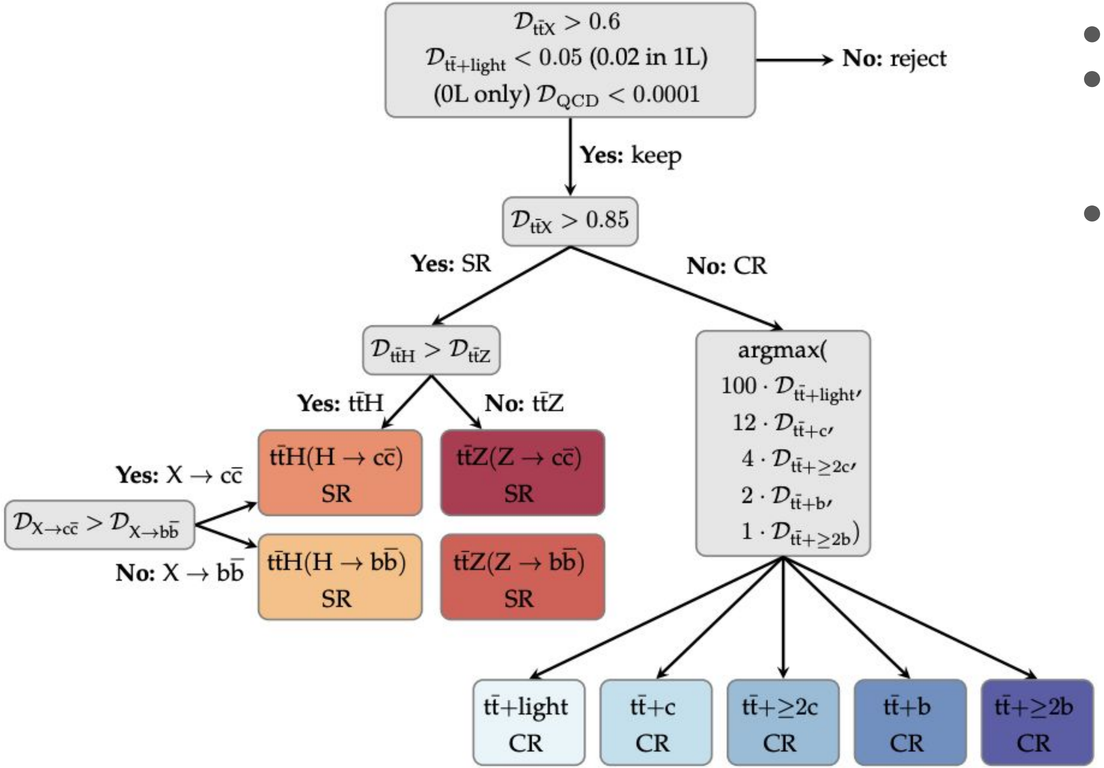
- **Three data samples** enriched in **$t\bar{t}$** (b-jets), **W+jets** (c-jets), **Z+jets** (light-jets)
- **SFs as a function of p_T** [25, 35, 50, 70, 90, 120, ∞] GeV
- **11 tagging bins** (one POI)
  - Except L0: in order to preserve unitarity
    - To keep overall normalization, SFs correct only efficiencies
- When the contribution of certain jet flavour is too small, **SF is not determined well**
  - SF is set to 1
  - Assigned uncertainty covers [0.3,3]
  - Negligible impact on the analysis results
- **Typical values**
  - 0.75 - 1.5
  - Uncertainties up to 15% for b-jets, up to 50% for c and light jets

Neural network event classifier

Multiclass event classifier

- **Inputs:**
  - p_T, E, η for each object
  - Jet flavour encoded in 10 boolean values
  - Lepton flag: muon/electron
  - Constructs pairwise features as invariant mass, etc…
- Trained and assigned **likelihood scores**:
  - Two **$t\bar{t}H$** classes ($H \to b\bar{b}$, $H \to c\bar{c}$)
  - Two **$t\bar{t}Z$** classes ($Z \to b\bar{b}$, $Z \to c\bar{c}$)
  - Five **$t\bar{t}$ + jets** classes: $t\bar{t}$+b, $t\bar{t}$+≥2b, $t\bar{t}$+c, $t\bar{t}$+≥2c, $t\bar{t}$+light
  - QCD multijet for 0L channel

Event categorization



- **Four SRs**
- **Five CRs**
  - to estimate the normalizations of various $t\bar{t}$+jets background
- **signal-depleted sideband region**
  - $0.4 < D_{ttX} < 0.6$
  - Validation region

Fit

Bin profiled likelihood fit

- The fitted variable: PART classifier discriminant
- The dominant background $t\bar{t}$+jets estimated from CRs fit
  - $t\bar{t}$+b, $t\bar{t}$+≥2b, $t\bar{t}$+c, $t\bar{t}$+≥2c, $t\bar{t}$+light allowed to float independently
- Minor background (single Top, $t\bar{t}$W, tWZ) estimated from MC

Systematic uncertainties (Via nuisance parameters)

- $t\bar{t}$H(H->c$\bar{c}$)
  - Leading uncertainty - statistical
  - Main systematic uncertainty:
    - Theoretical modelling of tt+c, tt+light (32%)
- $t\bar{t}$H(H->b$\bar{b}$)
  - Leading uncertainty - theoretical
    - Renormalization and factorization scales, flavour scheme, parton shower modelling of $t\bar{t}b\bar{b}$ (60%)

Primary experimental uncertainty:

- jet flavor identification efficiencies

Validation of analysis

$t\bar{t}Z$ signal strength measurements

- Consistent with SM within uncertainties
- Consistent with measurements in leptonic decay channel

$$\mu_{t\bar{t}Z(Z\to c\bar{c})} = 1.02^{+0.79}_{-0.84},$$

$$\mu_{t\bar{t}Z(Z\to b\bar{b})} = 1.47^{+0.45}_{-0.41},$$

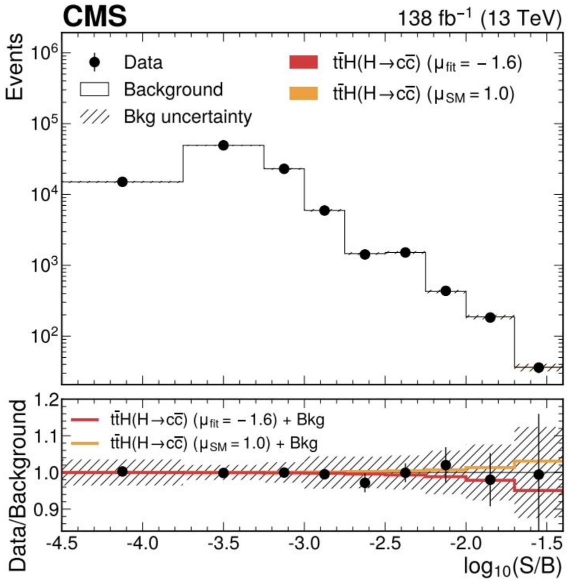
The observed (expected) significances:

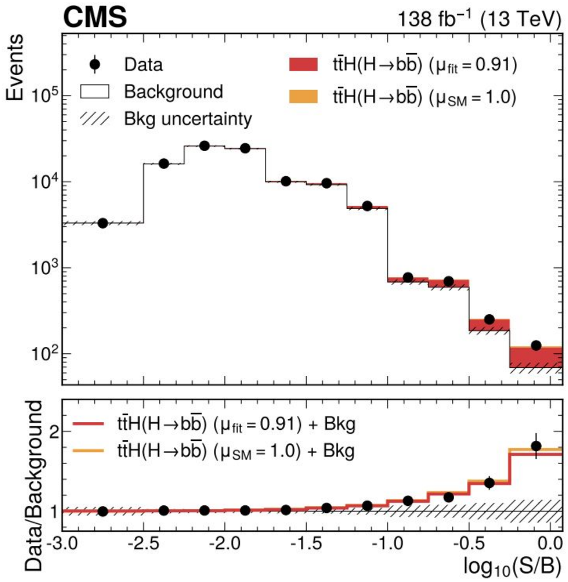
- **$t\bar{t}Z$ ($Z\to b\bar{b}$):** 3.5 (2.4)
- **$t\bar{t}Z$ ($Z\to c\bar{c}$):** 1.2 (1.3)

Results: ttH (H->bb)

[Signal strength](#)

$$\mu_{t\bar{t}H(H\to b\bar{b})} = 0.91^{+0.26}_{-0.22}$$





The observed (expected) significances:

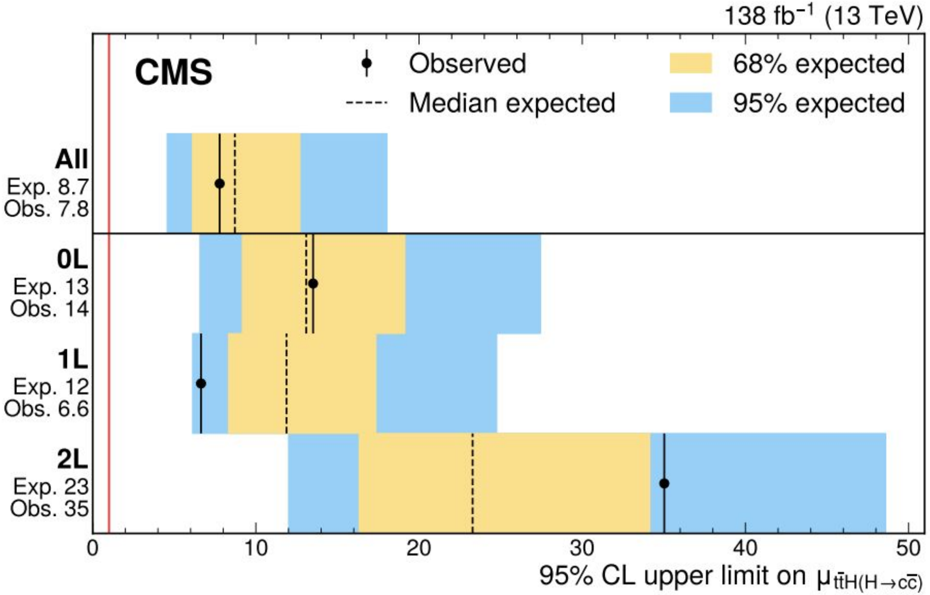
- **tt⁻H (H->bb⁻):** 4.4 (4.5)
- $\mu_{ttH(H\text{->}bb)}$ is closer to SM than [previous measurement](#) (ttH, H->bb)

Results: ttH (H->cc)

Signal strength

$$\mu_{t\bar{t}H(H\to c\bar{c})} = -1.6 \pm 4.5$$

- No excess over background-only hypothesis
- Upper limit is set
- The observed (expected) 95% CL limits:
  - 7.8 (8.7)
- Upper limit on σ(ttH)Br(H->cc)
  - 0.11 (0.13) pb



Kappa framework

Handbook of LHC Higgs cross-sections

- Used to quantify to deviations of Higgs couplings from their SM values
- **$k_i = y_i/y^{SM}_i$ - coupling modifier** (k_i = 1 for SM)
- Production cross sections and partial decay widths are rescaled by the square of the corresponding coupling modifier

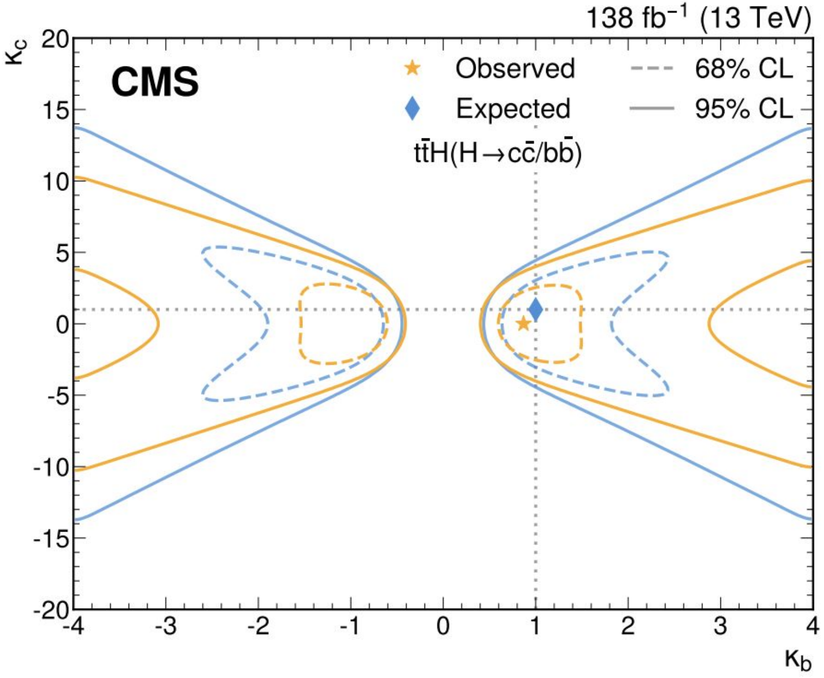
$$\sigma_i = \kappa_i^2\, \sigma_i^{\rm SM}, \qquad \Gamma_f = \kappa_f^2\, \Gamma_f^{\rm SM}. \qquad \Gamma_H = \sum_i \kappa_i^2 \Gamma_i^{SM}$$

$$BR(H \to XX) = \frac{\Gamma(H \to XX)}{\Gamma_H} \qquad BR = \frac{\kappa_X^2 \Gamma_X^{SM}}{\sum_i \kappa_i^2 \Gamma_i^{SM}} \qquad \mu = \frac{\kappa^2_{\rm prod}\kappa^2_{\rm decay}}{\kappa_H^2}$$

k_H is scaled total Higgs width

- **Main assumptions:**
  Higgs is sufficiently narrow that production and decay can be factorized. The production and the decay processes can be treated independently and then connected through the total width.
- Only the strength of the coupling is modified; the Lorentz/CP structure is assumed to remain SM-like.
- Leading-order interpretation

K-framework interpretation



$$\mathcal{B}(H\to c\bar{c}) = \frac{\kappa_c^2\,\mathcal{B}_{SM}^{H\to c\bar{c}}}{1+(\kappa_c^2-1)\,\mathcal{B}_{SM}^{H\to c\bar{c}}+(\kappa_b^2-1)\,\mathcal{B}_{SM}^{H\to b\bar{b}}},$$

$$\mathcal{B}(H\to b\bar{b}) = \frac{\kappa_b^2\,\mathcal{B}_{SM}^{H\to b\bar{b}}}{1+(\kappa_c^2-1)\,\mathcal{B}_{SM}^{H\to c\bar{c}}+(\kappa_b^2-1)\,\mathcal{B}_{SM}^{H\to b\bar{b}}}.$$

The observed (expected) 95% CL $|k_c|$ < 3.0 (3.3)

- k_b is fixed to SM expectation

Combined ttH and VH results

Combined analysis with previous VH measurement performed

- Common experimental uncertainties are correlated
  - Except jet flavour-tagging
- Background modelling uncertainties are uncorrelated
  - Different dominant backgrounds
- Theoretical modelling uncertainties for same processes are correlated

- Expected 95% CL limit on $\mu_{H\to cc}$: 5.6
  - 36% (26%) improvement over ttH (VH) alone
- Observed 95% CL limit on $\mu_{H\to cc}$: 9.3
  - Driven by statistical fluctuation in VH channel
- Observed (expected) 95% CL: $|\kappa_c| < 3.5$ (2.7)

Summary

- Measurement of $H \to c\bar{c}$ via $t\bar{t}H$ production
  - Alongside with $H \to b\bar{b}$
- **New jet-flavour identification tool and event classification techniques**
  - Using advanced machine learning algorithms
- The observed (expected) 95% CL limits on $\mu_{ttH(H\to cc)}$: 7.8 (8.7)
- Combined ttH and VH measurements:
  - Observed (expected) 95% CL: $|\kappa_c| < 3.5$ (2.7)
  - The most stringent constraint to date

Backup

ParticleNET

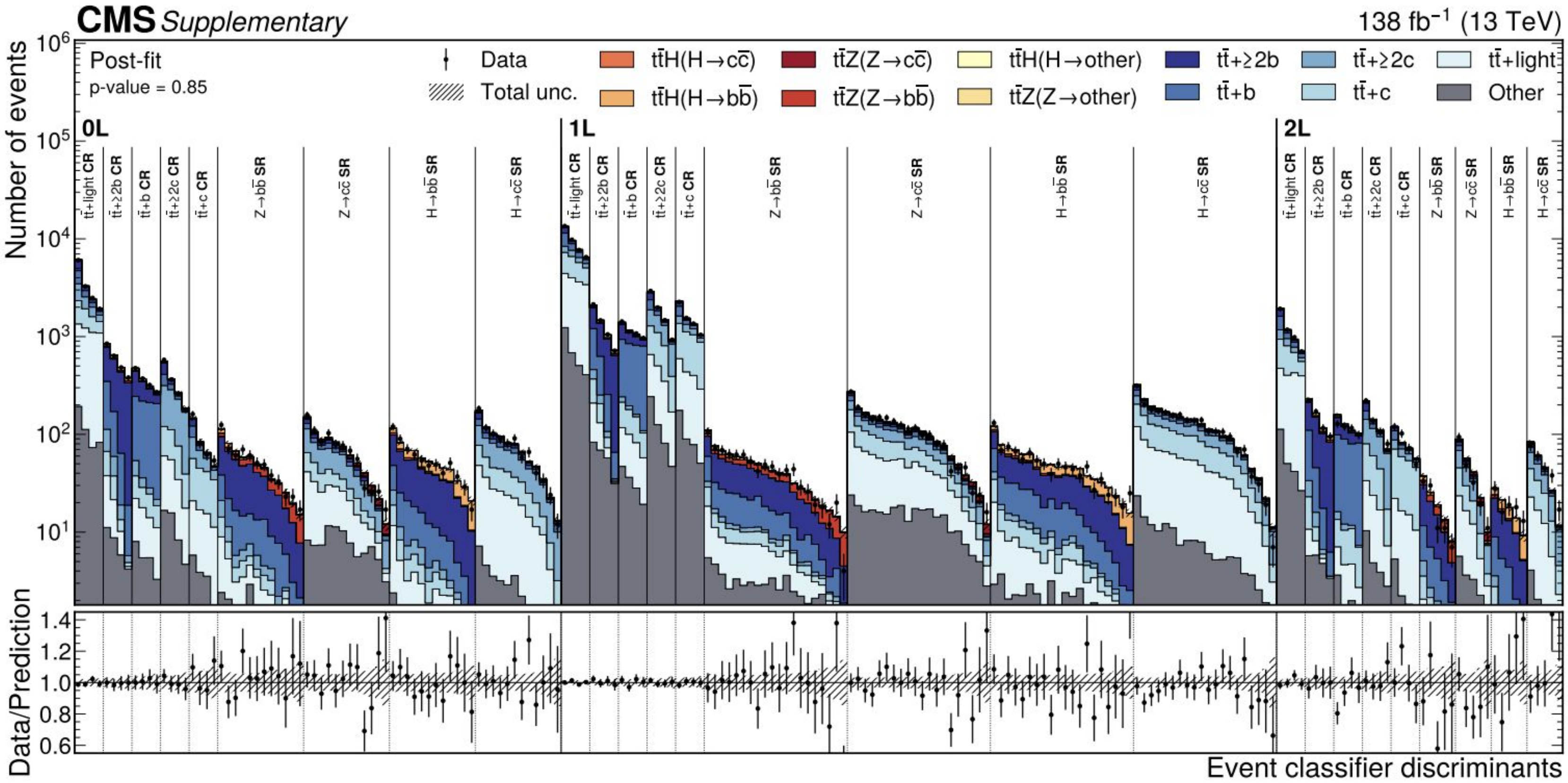
Outputs eight probabilities:

- p_b, p_{bb}, p_c, p_{cc}, p_{uds}, p_g, pileup and other sources

$$p_{\mathrm{B+C}} = \frac{p_{\mathrm{b}} + p_{\mathrm{bb}} + p_{\mathrm{c}} + p_{\mathrm{cc}}}{p_{\mathrm{b}} + p_{\mathrm{bb}} + p_{\mathrm{c}} + p_{\mathrm{cc}} + p_{\mathrm{uds}} + p_{\mathrm{g}}}$$

$$p_{\mathrm{BvsC}} = \frac{p_{\mathrm{b}} + p_{\mathrm{bb}}}{p_{\mathrm{b}} + p_{\mathrm{bb}} + p_{\mathrm{c}} + p_{\mathrm{cc}}}$$

Distribution of PART discriminants



Uncertainties breakdown

Uncertainty source	$\Delta\mu$ $(\Delta\mu/\Delta\mu_{tot})$ $\mu_{t\bar{t}H(H\to c\bar{c})}$		$\mu_{t\bar{t}H(H\to b\bar{b})}$	
Statistical	3.3	(74%)	0.14	(57%)
$t\bar{t}$+jets normalizations	1.4	(32%)	0.06	(26%)
$t\bar{t}Z$ normalizations	0.4	(8.4%)	0.06	(30%)
Theory	2.1	(47%)	0.18	(75%)
Signal	0.7	(15%)	0.11	(47%)
$t\bar{t}+\geq 1b$	0.7	(15%)	0.14	(60%)
$t\bar{t}+\geq 1c$	1.4	(32%)	0.01	(5.8%)
$t\bar{t}$+light	1.3	(29%)	0.01	(5.2%)
Minor backgrounds	0.2	(4.6%)	0.01	(4.6%)
Experimental	2.0	(47%)	0.07	(31%)
Jet flavor tagging	1.7	(39%)	0.07	(28%)
Size of the simulated samples	1.1	(24%)	0.05	(21%)
Jet energy scale and resolution	0.8	(18%)	0.02	(8.6%)
Lepton identification	0.3	(6.0%)	0.02	(6.3%)
Integrated luminosity	0.1	(2.0%)	0.02	(6.2%)
Total	4.5	(100%)	0.24	(100%)