$t\bar{t}H$ Production with $H \rightarrow b\bar{b}$ at the CMS Experiment

ABHISEK DATTA

UNIVERSITY OF CALIFORNIA, LOS ANGELES (UCLA)

ON BEHALF OF THE CMS COLLABORATION

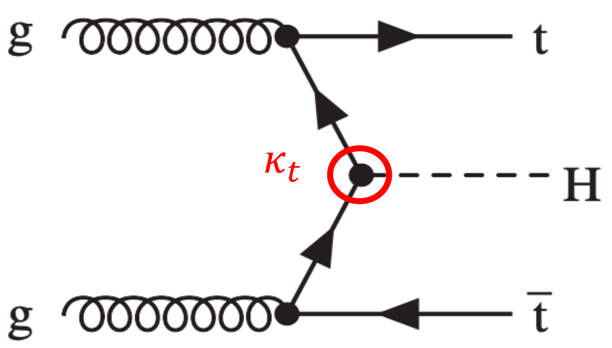
DECEMBER 14, 2023

Why $t\bar{t}H$?

Precision measurements of coupling of Higgs boson to other particles crucial for validation of the Standard Model (SM) and to search for new physics:

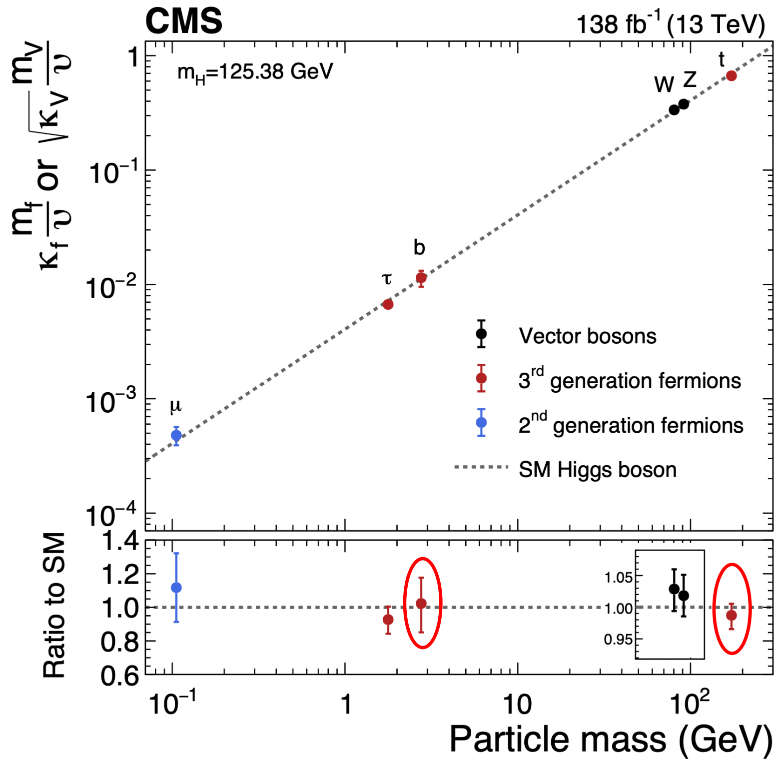
- κ_b and κ_t: many BSM models can lead to modifications of Higgs boson couplings to fermions → higher precision needed

$t\bar{t}H$ is the **best direct probe** of the **Top-Higgs coupling** (κ_t) at **tree level**



$\sigma_{t\bar{t}H} \approx 0.503$ pb

Observing a $t\bar{t}H$ production rate different from the Standard Model prediction can indicate the presence of Beyond the Standard Model (BSM) physics



Nature 607, 60–68 (2022)

Latest measurements of coupling strengths between the Higgs boson and vector bosons & fermions

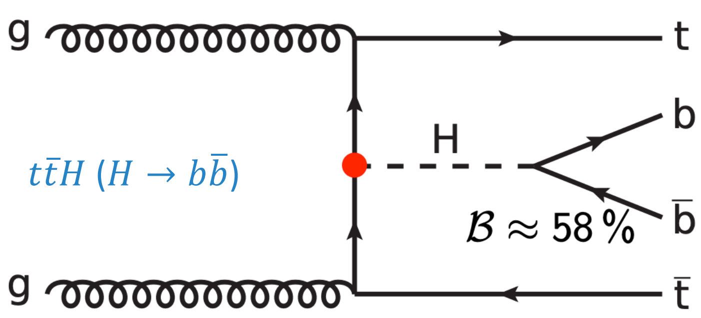
$t\bar{t}H$ with $H \to b\bar{b}$

Many possible decay channels for the Higgs boson:

- $H \to b\bar{b}$
- $H \to WW^*, \tau\tau, ZZ^*$
- $H \to \gamma\gamma$

$H \to b\bar{b}$ final state chosen for this analysis:

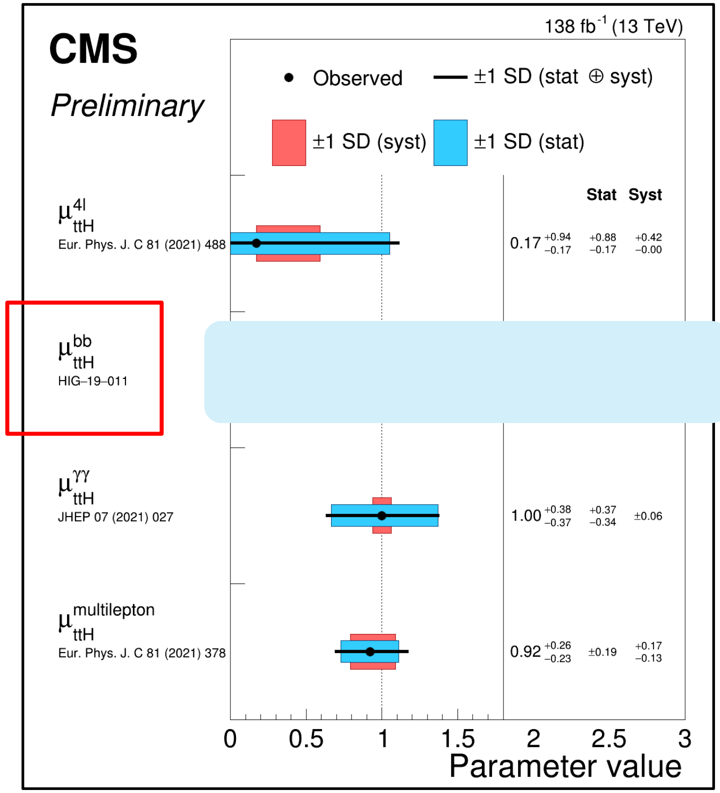
- Largest branching fraction of 58%
- Fully reconstructable Higgs boson final state
- All Higgs-fermion (even 3rd generation – t and b) vertices



$t\bar{t}H$ observation by the combination of all decay channels at both CMS and ATLAS in 2018

Latest status of CMS results:

Focus of today's talk



$$\mu_{t\bar{t}H} = \frac{\sigma(t\bar{t}H)}{\sigma_{SM}(t\bar{t}H)} \cdot \frac{\mathcal{B}(H \to b\bar{b})}{\mathcal{B}_{SM}(H \to b\bar{b})}$$

$t\bar{t}H(H \rightarrow b\bar{b})$ Using Full Run-2 Data

Following the $t\bar{t}H$ discovery, the focus is now on more precise measurements of $t\bar{t}H$ production in each decay channel

For the $t\bar{t}H(H \rightarrow b\bar{b})$ channel, using full Run-2 (2016 – 2018 : 138 fb^{-1}) data

Released in August (CMS-PAS-HIG-19-011) → **Focus of today's talk**

Major improvements in the analysis:

- Better modeling of the major irreducible $t\bar{t} + b\bar{b}$ background for $t\bar{t}H(H \rightarrow b\bar{b})$ with improved simulation
- Refined neural network classifiers (for signal to background discrimination)
- New triggers to increase signal efficiency in the different final states
- Better identification of jets arising from b-quarks from both improved algorithms and also upgrades in the pixel tracking detector of CMS
- Additional interpretations including differential measurements

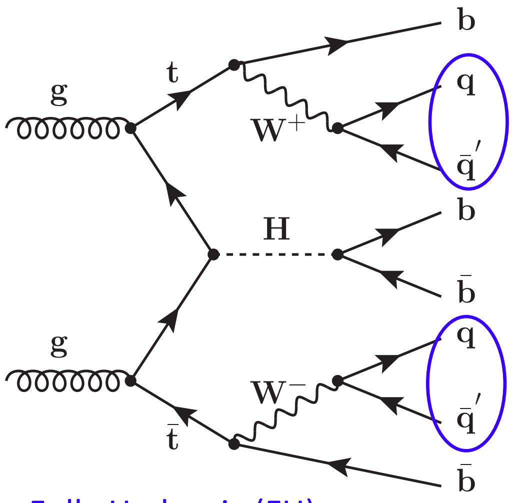
Final States Signatures

Different final states depending on $t\bar{t}$ decay :

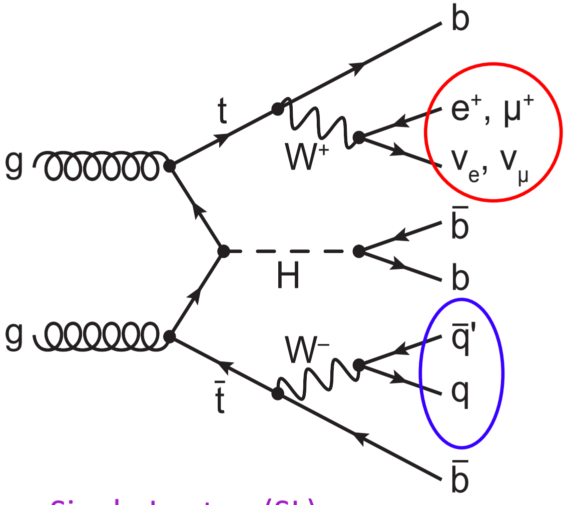
- Fully Hadronic (FH)
- Single Lepton (SL) : e , μ
- Dilepton (DL) : ee , $e\mu$, $\mu\mu$

Event Selection requires :

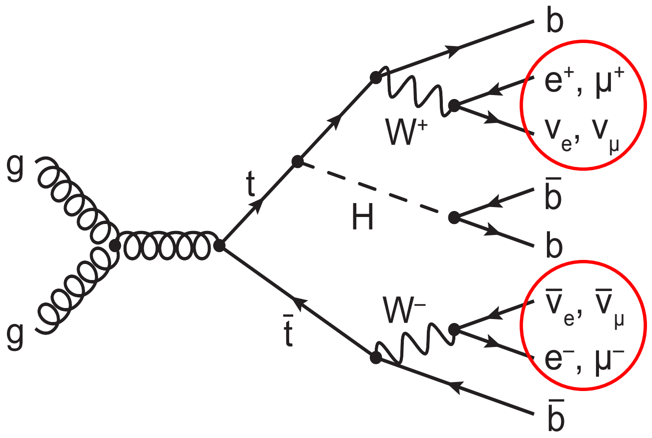
- 0/1/2 leptons depending on the channel
- Multiple jets (including b-tagged jets)



Fully Hadronic (FH)



Single Lepton (SL)



Dilepton (DL)

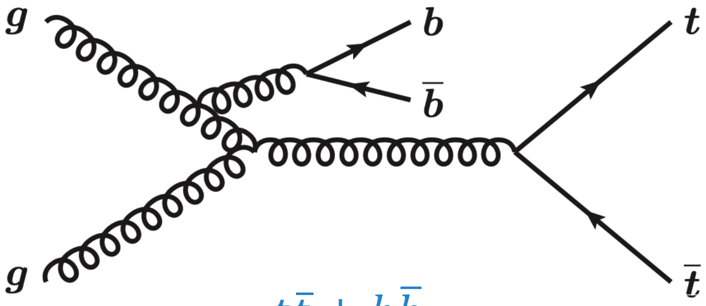
Challenging final state:

- Huge combinatorics in event reconstruction
- Small signal ($t\bar{t}H$) cross-section compared to large irreducible backgrounds ($t\bar{t} + b\bar{b}$)

Major Background Processes

$t\bar{t}$ +jets (all channels):

- Modeled from simulation
- Divided into three sources depending on flavor of additional particle-level jets:
  - $t\bar{t} + B$: >= 1 additional b-jet - irreducible background (mostly $t\bar{t} + b\bar{b}$):
    - Modeling challenging due to complex multi-parton states and multiple, very different scales ($t\bar{t}$, $b\bar{b}$)
    - Large modeling uncertainties
    - Current measurements $\sim 20 - 40\%$ larger than prediction
  - $t\bar{t} + C$: >= 1 additional c-jet but no b-jet
  - $t\bar{t} + LF$: all other events (LF: light flavor)



$t\bar{t} + b\bar{b}$

QCD Multijet (Fully Hadronic channel):

- Dedicated background rejection
- Data-driven background estimation using Control Regions

Minor backgrounds (all channels):

- Single-top, diboson, $t\bar{t} + V$, V +jets
- Modeled from simulation

$t\bar{t} + B$ Background Model

Improved modeling of the $t\bar{t} + B$ irreducible background:

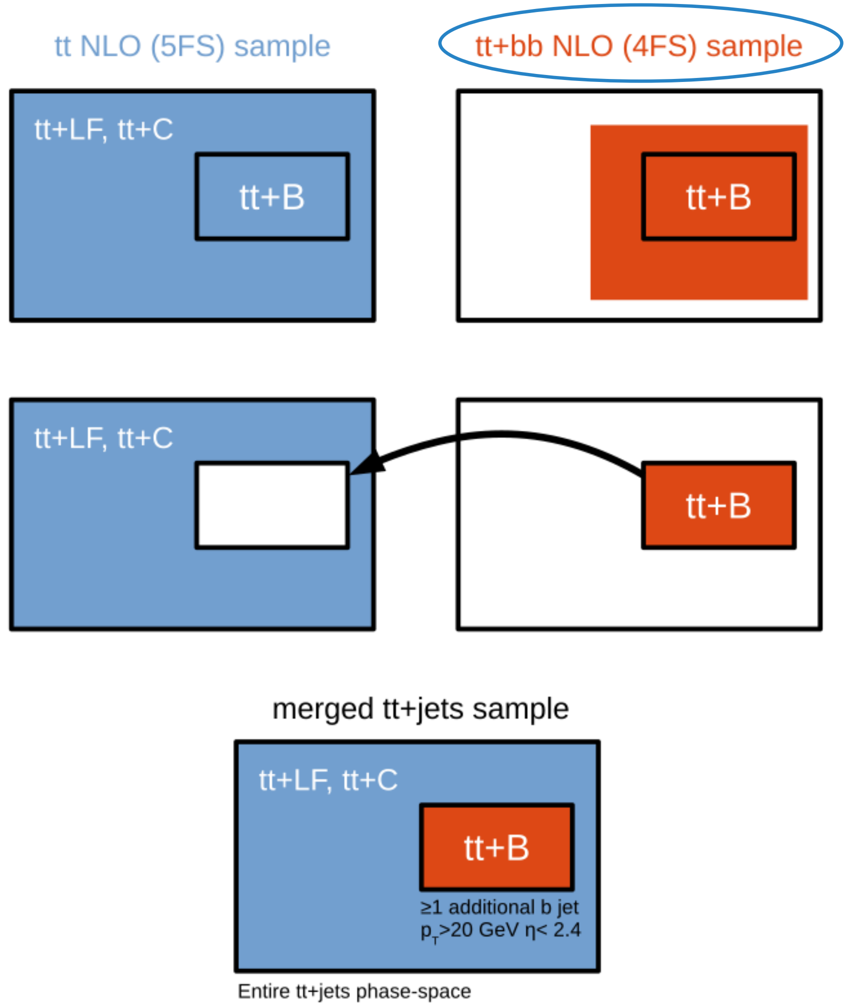
- New dedicated $t\bar{t}b\bar{b}$ simulation* with $t\bar{t} + B$ processes described directly by a $t\bar{t}b\bar{b}$ Matrix Element at Next-to-Leading Order (NLO)
- **Theoretically preferred option**: better description of event kinematics

* Eur. Phys. J. C 78 (2018) 502 Jezo et al Buccioni et al

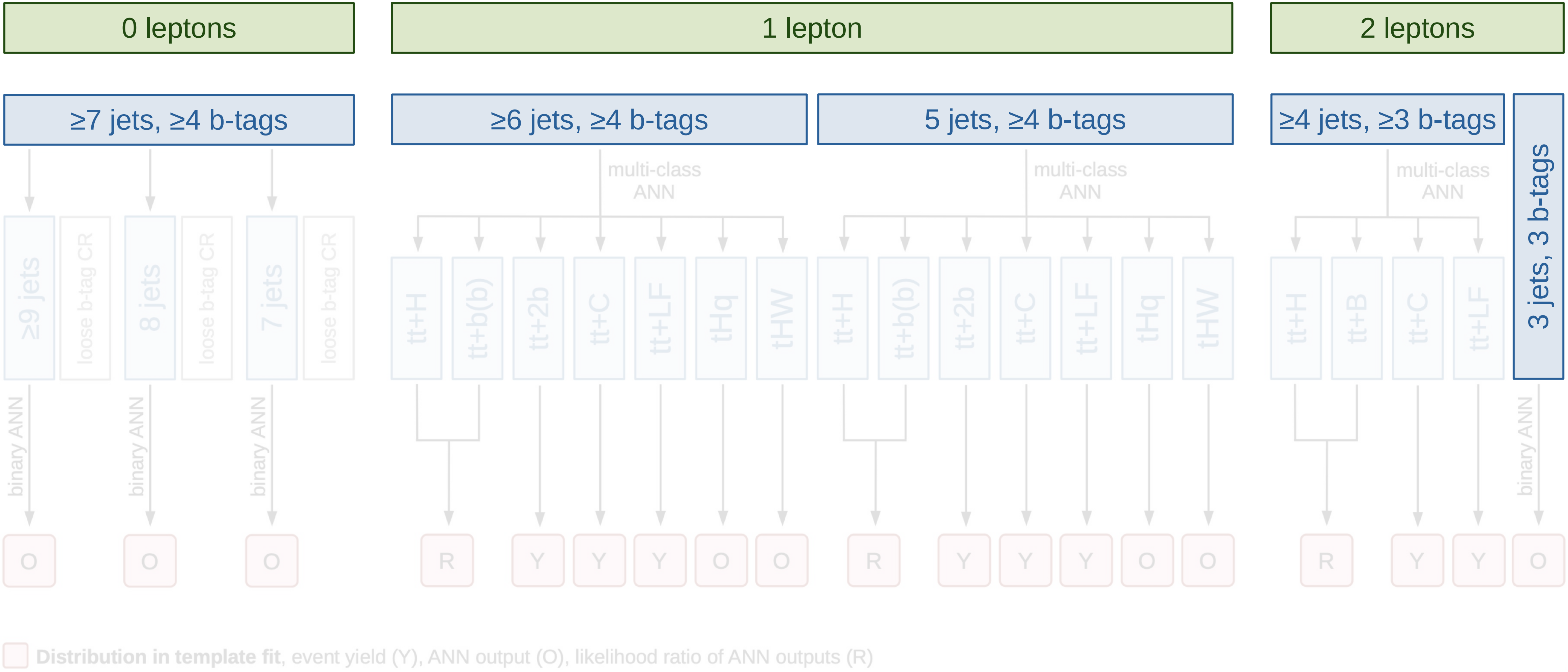
New Merged $t\bar{t}$ + Jets Background Model:

- $t\bar{t} + B$ taken from the new $t\bar{t}b\bar{b}$ simulation
- $t\bar{t} + C, t\bar{t} + LF$ taken from the inclusive $t\bar{t}$+jets simulation
- $t\bar{t} + B$ and $t\bar{t} + C$ normalizations freely floating in the final fit

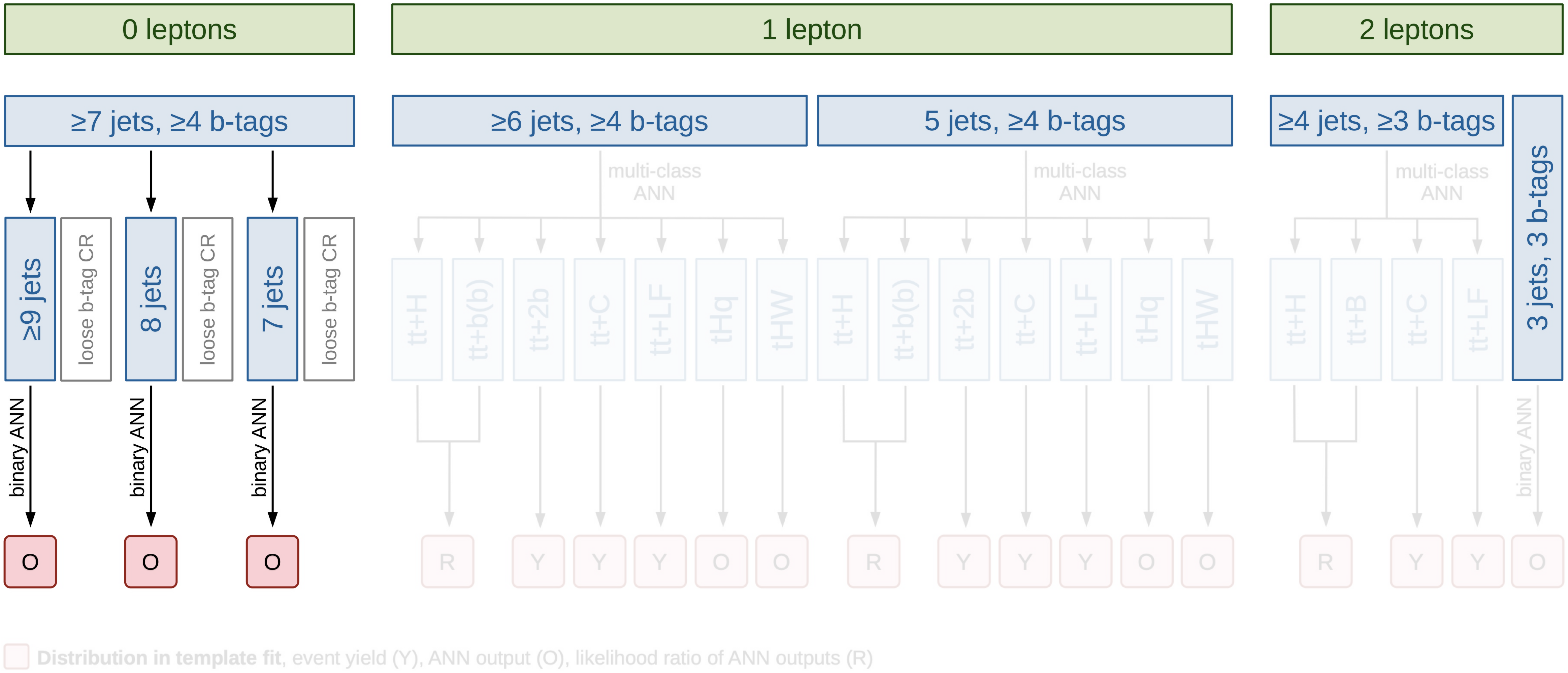
Robustness of background model validated with bias tests using toy data



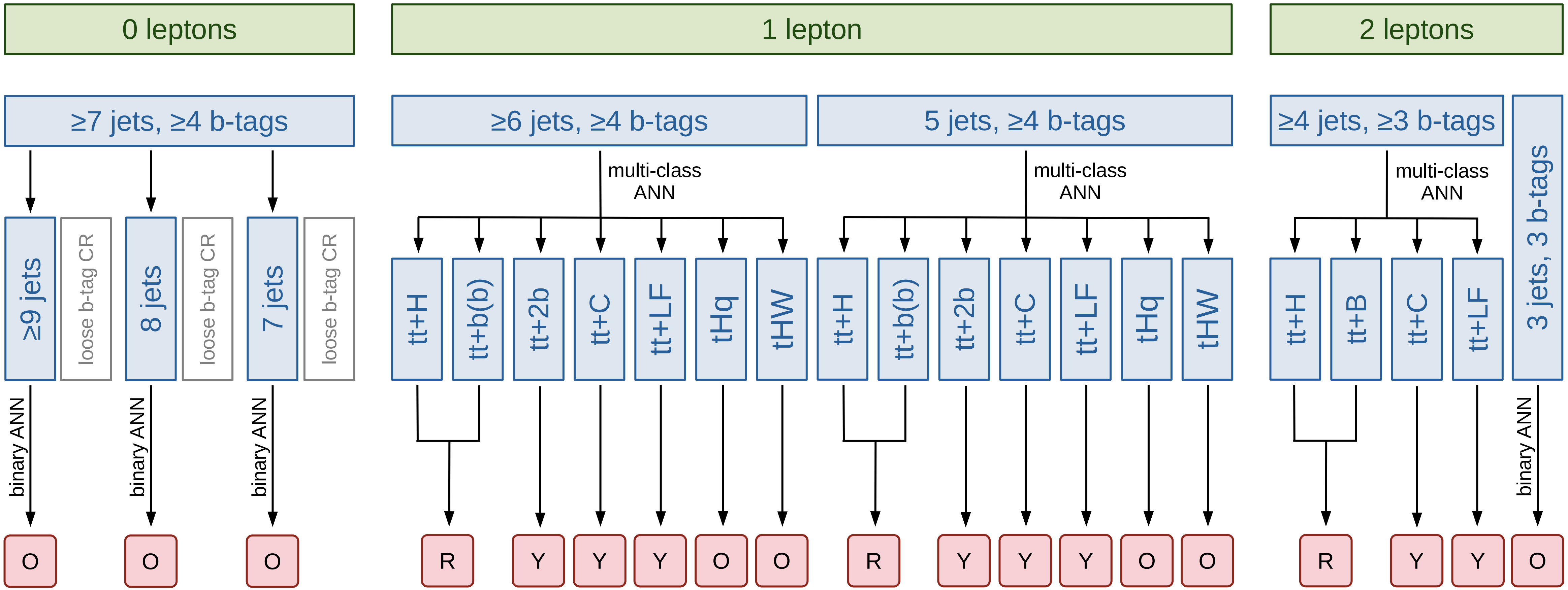
Inclusive $t\bar{t}H$ Measurement Strategy



Inclusive $t\bar{t}H$ Measurement Strategy



Inclusive $t\bar{t}H$ Measurement Strategy



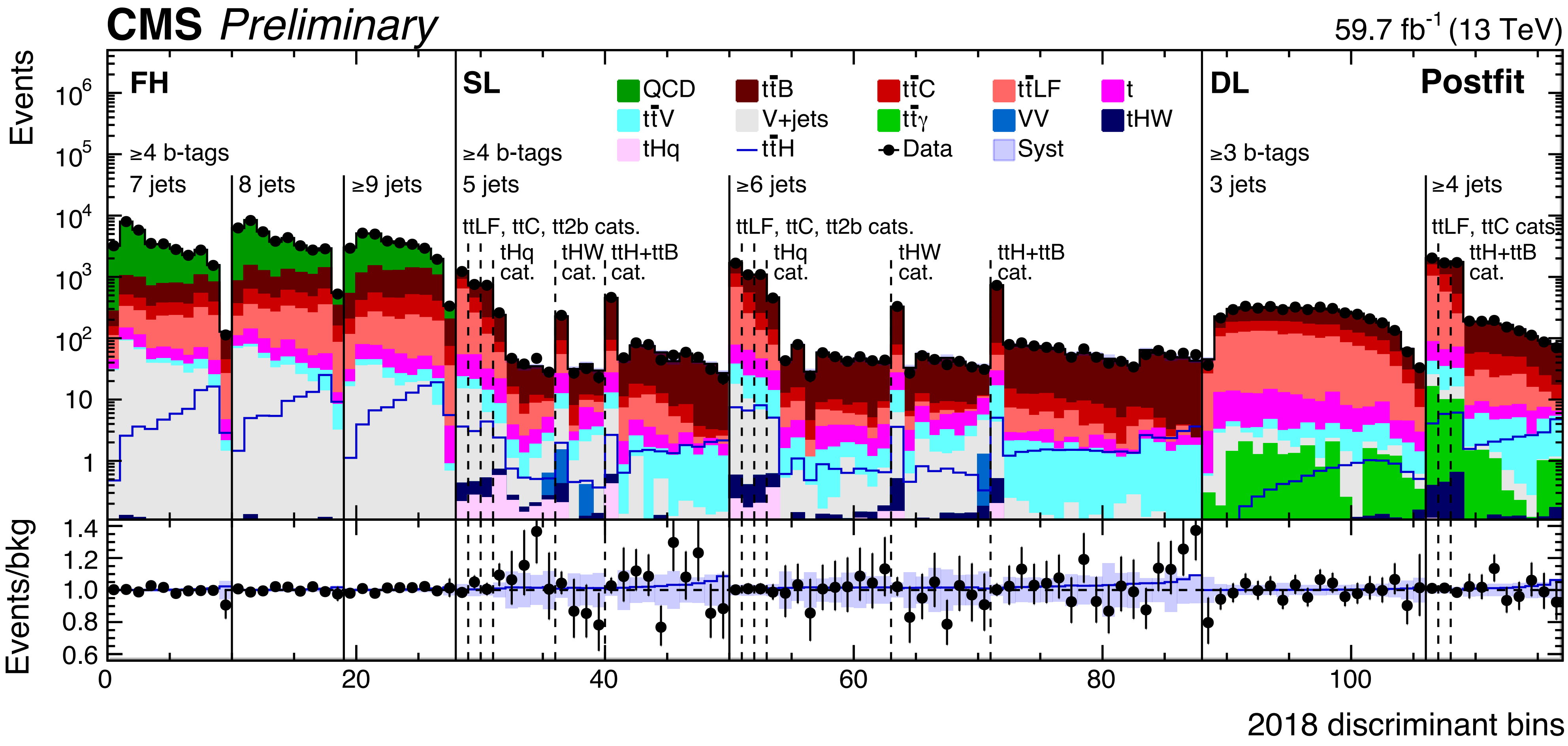
Distribution in template fit, event yield (Y), ANN output (O), likelihood ratio of ANN outputs (R)

$$R_{SL} = \frac{O(t\bar{t}H)}{O(t\bar{t}H) + O(t\bar{t} + b(\bar{b})) + O(t\bar{t} + 2b)}$$

$$R_{DL} = \frac{O(t\bar{t}H)}{O(t\bar{t}H) + O(t\bar{t}B)}$$

Inclusive $t\bar{t}H$ Results: Postfit Distributions

Postfit distributions from 2018 (2016 and 2017 in backup):

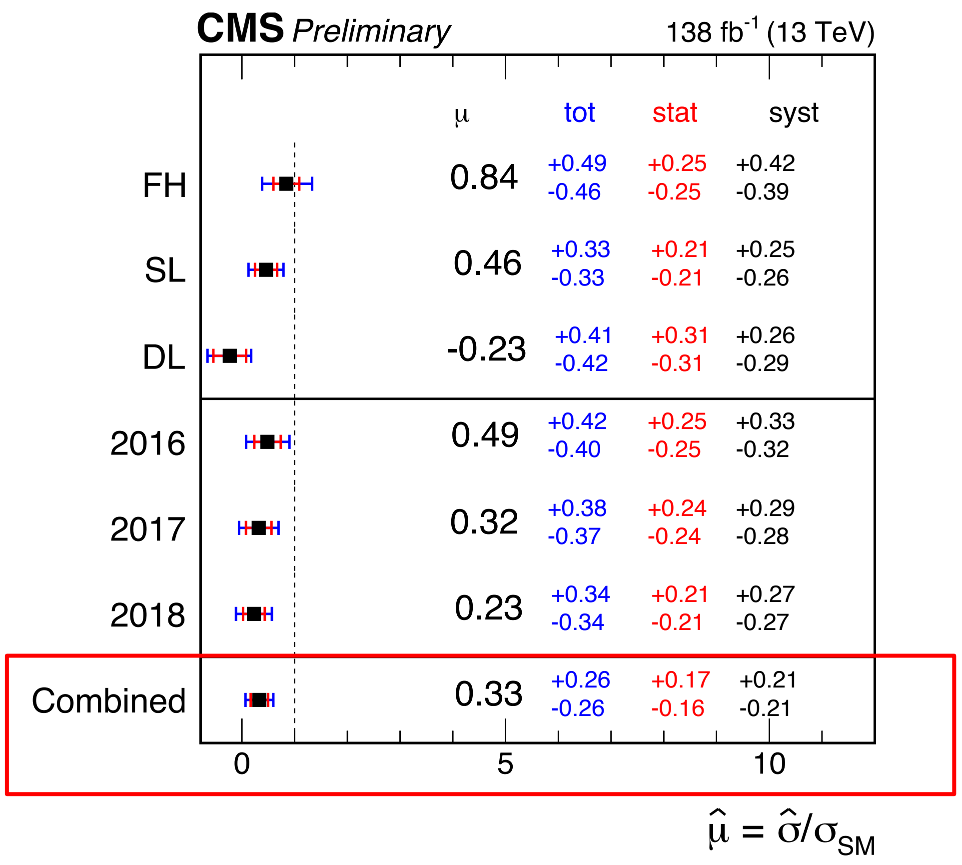


Expect a **total** of $\sim 1100\ t\bar{t}H$ events In 2018

Fitted observables in this plot are the event yields, ANN outputs and ratio observables

Inclusive $t\bar{t}H$ Results: Signal Strength

Full Run-2 Results



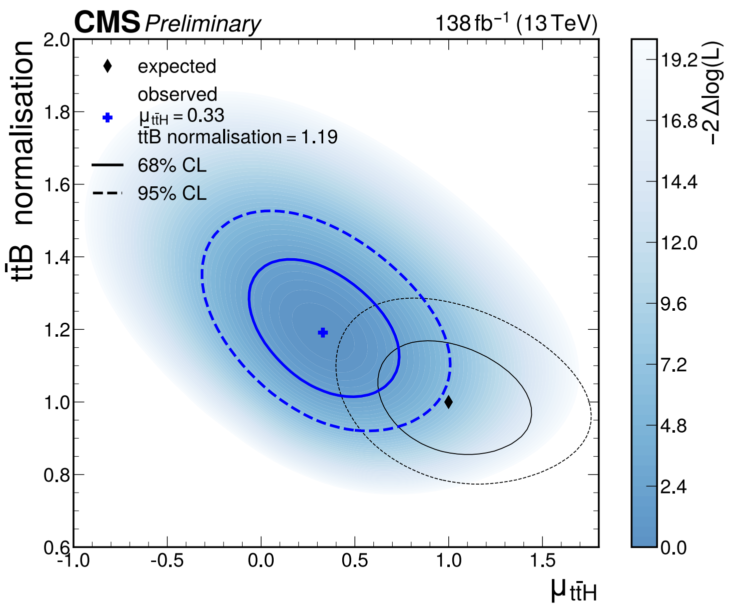
Uncertainties are correlated among channels and years

- $t\bar{t}H$ signal strength:
  - $\mu_{t\bar{t}H} = 0.33 \pm 0.26$, 1.3 σ obs. (4.1 σ exp.) significance
  - SM compatibility p-value: 2% (2.4 σ)
  - Compatibility to 2016 CMS publication (SL+DL): 41% (0.8 σ)

Agreement with ATLAS Full Run-2 result:

- $\mu_{t\bar{t}H} = 0.35^{+0.35}_{-0.34}$

J. High Energ. Phys. 2022, 97 (2022)



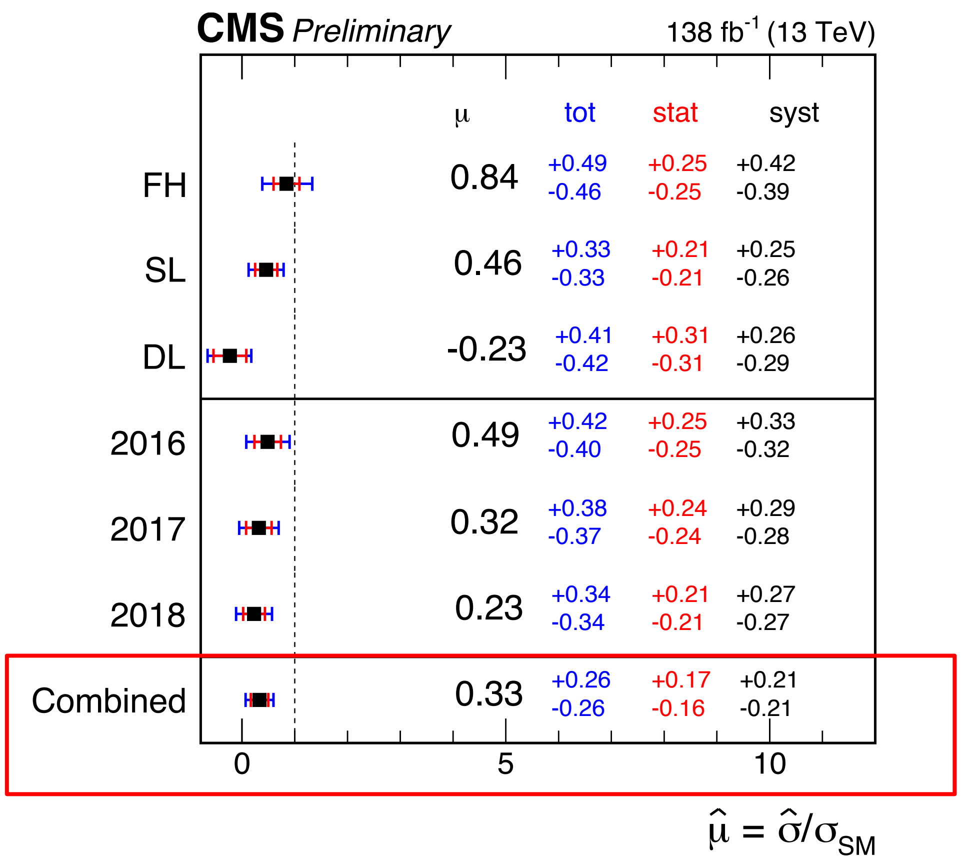
Background normalizations from the $t\bar{t}H$ measurement:

- $t\bar{t} + B$: $1.19^{+0.13}_{-0.12}$
- $t\bar{t} + C$: $1.07^{+0.20}_{-0.19}$

In agreement with dedicated measurements (CMS-PAS-TOP-22-009)

Inclusive $t\bar{t}H$ Results: Systematic Uncertainties

Full Run-2 Results



Uncertainties are correlated among channels and years

Major sources of systematic uncertainties:

Uncertainty source	$\Delta\mu_{t\bar{t}H}$ (observed)	$\Delta\mu_{t\bar{t}H}$ (expected)
Total experimental	$+0.10/-0.10$	$+0.11/-0.10$
jet energy scale and resolution	$+0.08/-0.07$	$+0.09/-0.09$
b tagging	$+0.07/-0.06$	$+0.06/-0.02$
luminosity	$+0.02/-0.02$	$+0.01/-0.01$
Total theory	$+0.16/-0.16$	$+0.18/-0.14$
$t\bar{t}$ + jets background	$+0.15/-0.16$	$+0.12/-0.11$
signal modelling	$+0.06/-0.01$	$+0.13/-0.06$
Size of the simulated event samples	$+0.13/-0.12$	$+0.10/-0.10$
Total systematic	$+0.20/-0.21$	$+0.23/-0.19$
Statistical	$+0.17/-0.16$	$+0.17/-0.17$
background normalisation	$+0.13/-0.13$	$+0.13/-0.13$
$t\bar{t}B$ and $t\bar{t}C$ normalisation	$+0.12/-0.12$	$+0.12/-0.12$
QCD normalisation	$+0.01/-0.01$	$+0.01/-0.01$
Total	$+0.26/-0.26$	$+0.28/-0.25$

$t\bar{t} + jets$ uncertainties most important

$t\bar{t}H$ Measurement in Higgs Boson p_{T} Bins

$t\bar{t}H$ cross-section measured in 5 Higgs boson p_{T} bins (generator level)

Simplified Template Cross-Section (STXS) approach:

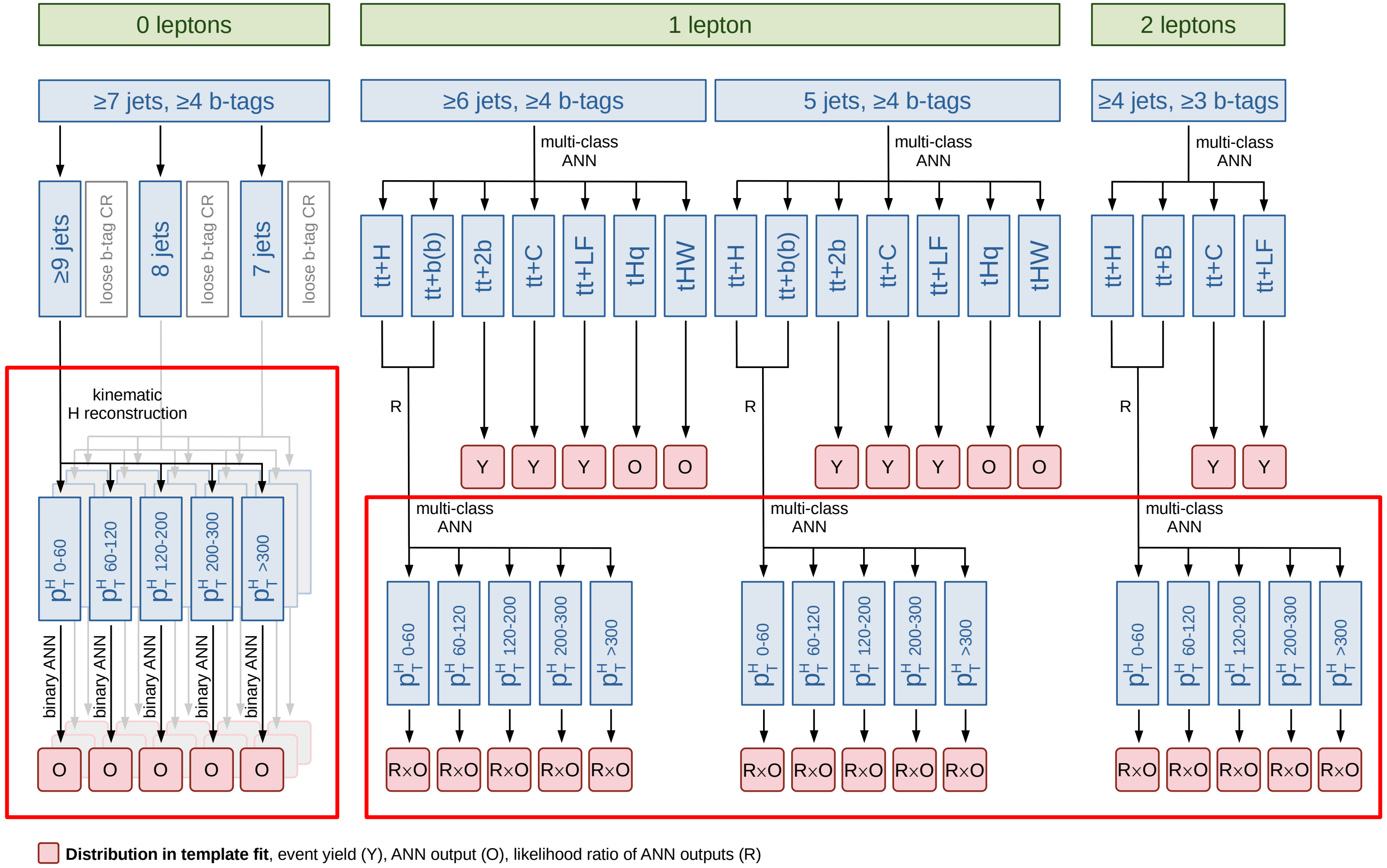
- [0,60[, [60,120[, [120,200[, [200,300[and [300,∞]

Perform reconstruction of bins:

- FH: χ^2 reconstruction of the Higgs from b-jet pairs
- SL and DL: multi-class ANN trained on $t\bar{t}H(b\bar{b})$

5 independent signal templates (for each generator-level $p_{\mathrm{T}}^{\mathrm{H}}$) fit simultaneously

Fitted observable is the output of the Higgs p_{T} ANN times the ratio observable from the inclusive ANN



$t\bar{t}H$ Measurement in Higgs Boson p_{T} Bins

$t\bar{t}H$ cross-section measured in 5 Higgs boson p_{T} bins (generator level)

Simplified Template Cross-Section (STXS) approach:

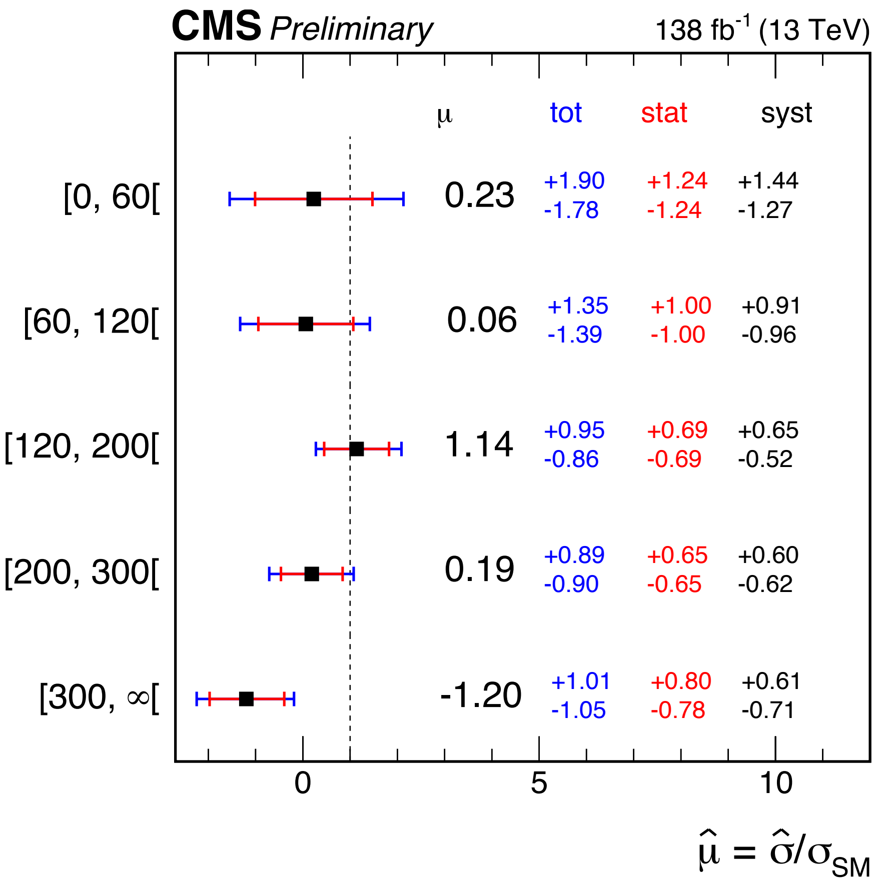
- [0,60[, [60,120[, [120,200[, [200,300[and [300,∞]

Perform reconstruction of bins:

- FH: χ^2 reconstruction of the Higgs from b-jet pairs
- SL and DL: multi-class ANN trained on $t\bar{t}H(b\bar{b})$

5 independent signal templates (for each generator-level $p_{\mathrm{T}}^{\mathrm{H}}$) fit simultaneously

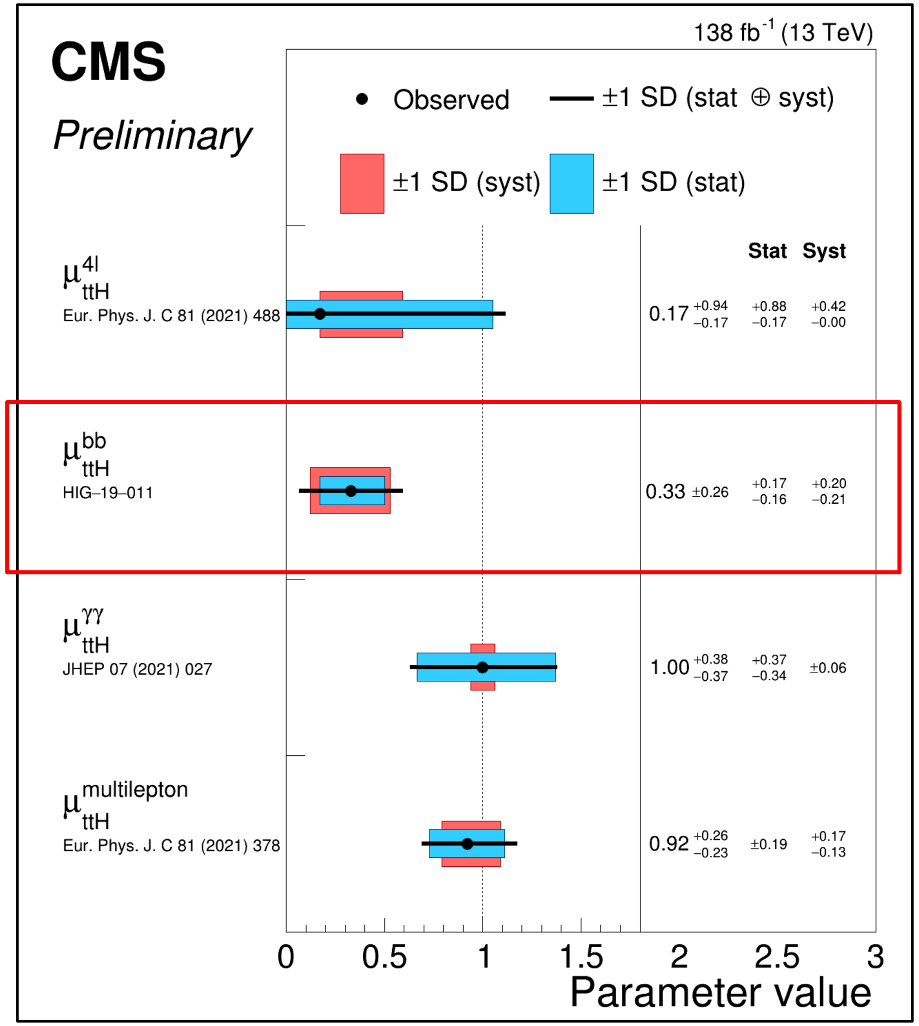
Fitted observable is the output of the Higgs p_{T} ANN times the ratio observable from the inclusive ANN



Results completely compatible with inclusive results: p-value of 0.67 ($0.4\ \sigma$)

Summary

- $t\bar{t}H$ provides a direct probe for the Top-Higgs Yukawa coupling
- Measurement of $t\bar{t}H$ production rate (both inclusive and in Higgs boson p_{T}) presented using full Run-2 data ($138\ \mathrm{fb}^{-1}$)
- $t\bar{t}H$ production rate observed to be smaller than SM expectations
- Necessitates updated measurements of $t\bar{t}H$ and tH production rates with more data and further scrutiny of the $t\bar{t} + B$ background



This work was supported in part by U.S. Dept. of Energy (Award DE-SC0009937)

Backup

Previous Results on $t\bar{t}H$

$t\bar{t}H$ observation by the combination of all decay channels at both CMS and ATLAS in 2018

Using partial Run-2 ($\sqrt{s} = 13$ TeV) + Run-1 ($\sqrt{s} = 7$ and 8 TeV) data :

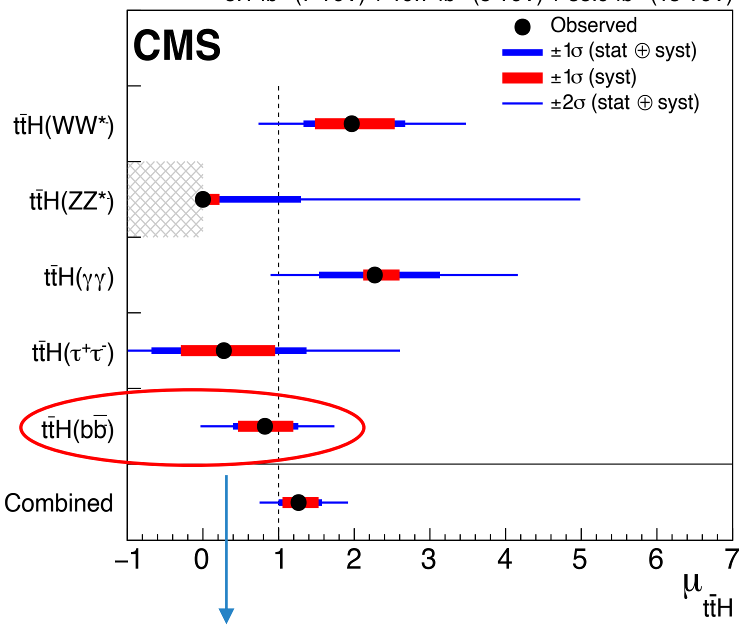
Results	Observed Significance
CMS [1]	$5.2\ \sigma$
ATLAS [2]	$6.3\ \sigma$

[1] Phys. Rev. Lett. 120 (2018) 231801 (CMS)
[2] Physics Letters B 784 (2018) 173–191 (ATLAS)

Best-fit signal strength from CMS: $\mu_{t\bar{t}H} = 1.26^{+0.31}_{-0.26}$

$t\bar{t}H$ signal strength consistent with SM expectations

Observation from CMS [1]



$t\bar{t}H(H \to b\bar{b})$ results using 2016 data

J. High Energ. Phys. **2019**, 26 (2019)
J. High Energ. Phys. 2018, 101 (2018)

Preliminary result for $t\bar{t}H\ (H \to b\bar{b})$ using 2016 + 2017 data only: **published in 2019** (CMS-PAS-HIG-18-030)

Final State Selection

Baseline Event Selection in the three channels:

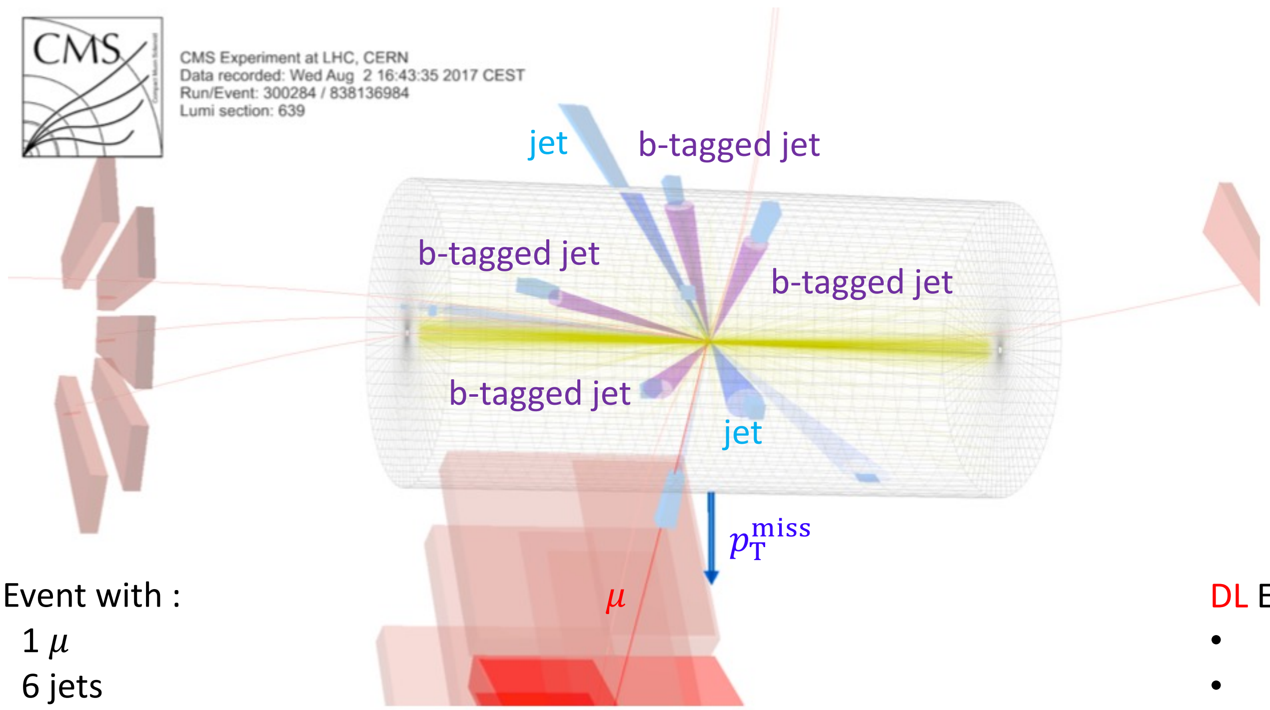
- Use standard PF objects with recommended corrections
- B-tagging using DeepJet medium WP
  - For the data-driven QCD estimation in the FH channel, control regions with loose b-tagged jets
- In each channel, events are categorized according to jet and b tag multiplicity

	FH channel	SL channel	DL channel
Number of leptons	0	1	2
Sign and flavour of leptons	—	$e^{\pm}, \mu^{\pm}$	$e^{+}e^{-}, \mu^{\pm}e^{\mp}, \mu^{+}\mu^{-}$
Min. p_T of leading electron (GeV)	—	29/30/30	25
Min. p_T of leading muon (GeV)	—	26/29/26	25
Min. p_T of additional leptons (GeV)	—	—	15
Max. p_T of additional leptons (GeV)	15	15	—
Max. $\|\eta\|$ of leptons	2.4	2.4	2.4
Min. $m_{\ell\ell}$ (GeV)	—	—	20
$m_{ee/\mu\mu}$ (GeV)	—	—	< 76 or > 106
Min. number of jets	7	5	3
Min. p_T of jets (GeV)	30	30	30
Min. p_T of 6^{th} jet (GeV)	40	—	—
Max. $\|\eta\|$ of jets	2.4	2.4	2.4
Min. number of b-tagged jets	2	4	3
m_{qq} (GeV)	> 30 and < 250	—	—
Min. H_T (GeV)	500	—	—
Min. p_T^{miss} (GeV)	—	20	40

Candidate $t\bar{t}H(H \to b\bar{b})$ Events in CMS

Candidate $t\bar{t}H(H \to b\bar{b})$ events after reconstruction and selections

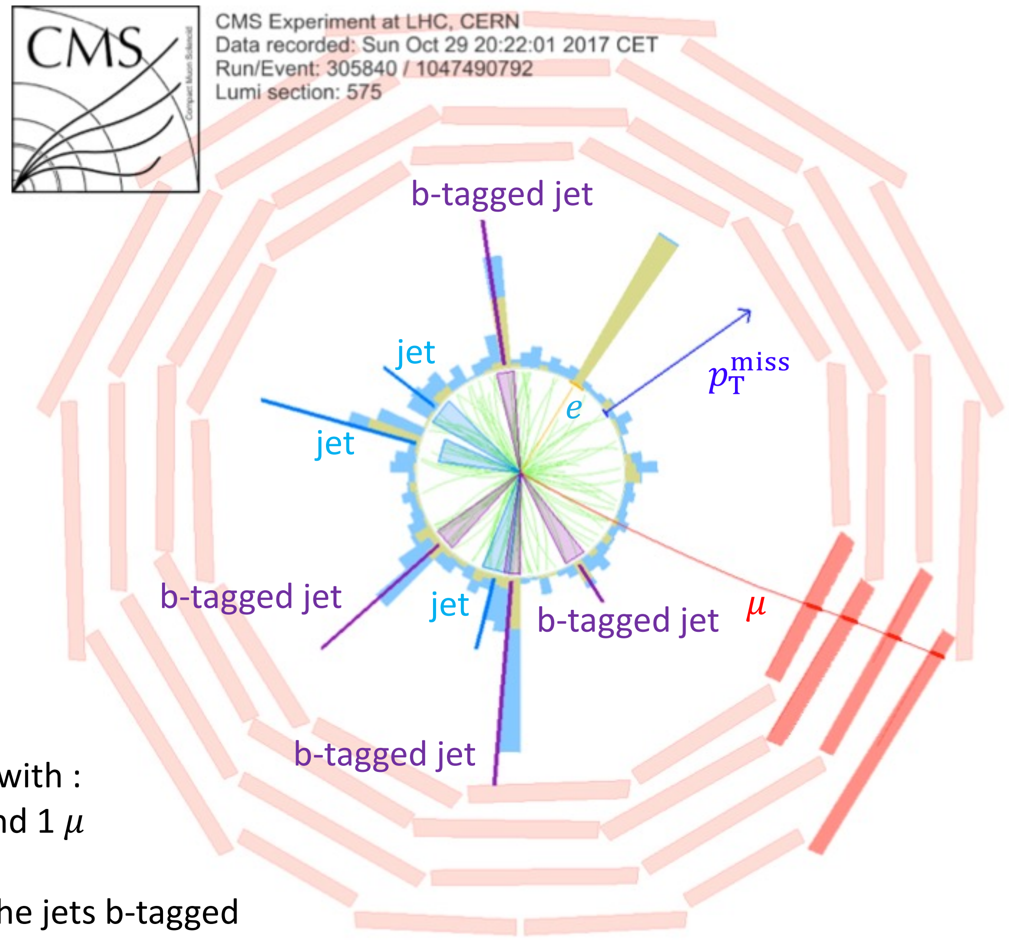
Single Lepton (SL) Channel



SL Event with :

- 1 μ
- 6 jets
- 4 of the jets b-tagged

Dilepton (DL) Channel



DL Event with :

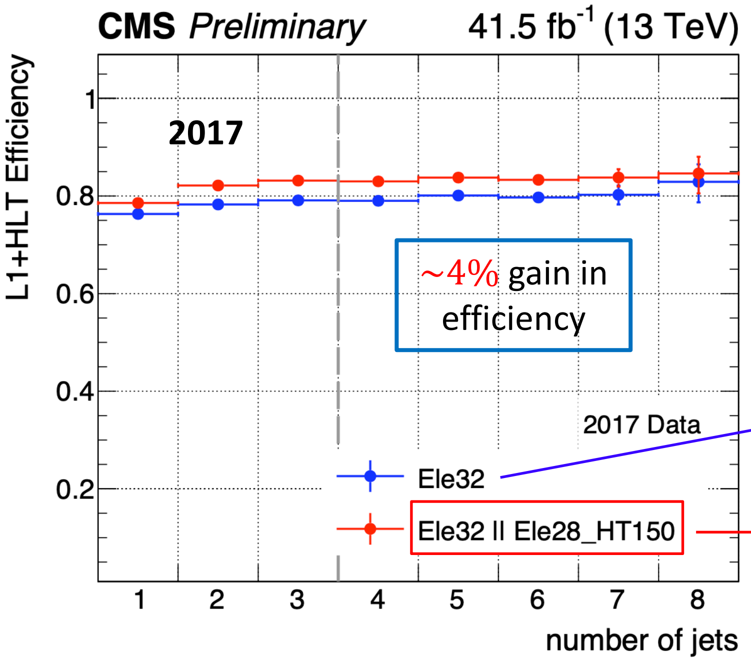
- 1 e and 1 μ
- 7 jets
- 4 of the jets b-tagged

Development of Single Electron Triggers

Developed new Single Electron triggers for 2017 and 2018 data taking for the $t\bar{t}H(H \to b\bar{b})$ analysis :

- Used **both** electron and jet information in the final states to design new triggers
- Allows keeping the p_T threshold low for the electron
- Retains signal efficiency at higher luminosity with negligible increase in data rates

This trigger also **used** in other Higgs and Top analyses



Efficiency vs Jet Multiplicity

H_T : sum of selected jet p_T

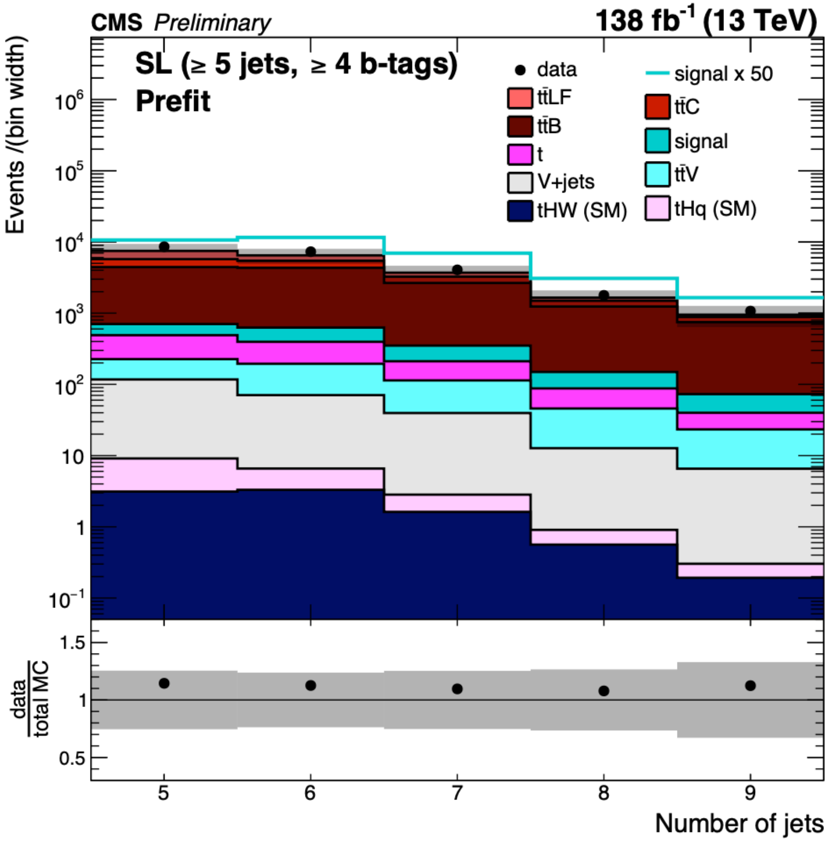
without jet information

electron + jet information

Results public : CMS DP -2019/026

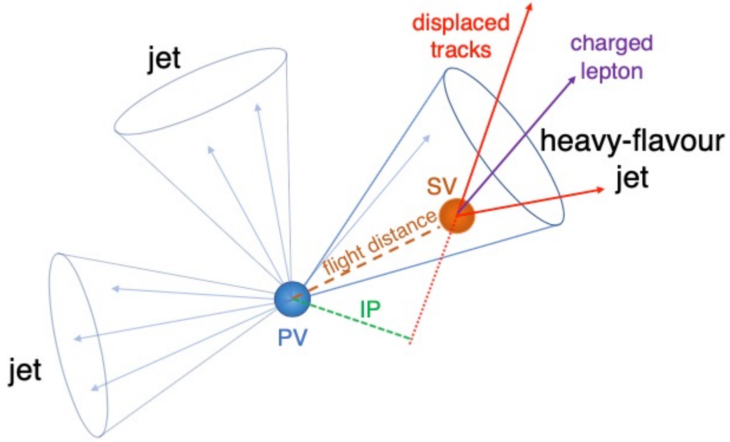
Challenging Final State

- Very busy final state with lots of jets and b-jets
- Small signal ($t\bar{t}H$ and tH) cross-section compared to large irreducible backgrounds ($t\bar{t} + b\bar{b}$)



Identifying jets originating from b-quarks essential

b-tagging algorithms based on:
- Long lifetime of B-hadrons
- Secondary vertex displaced (~ 0.5 mm) from the interaction point



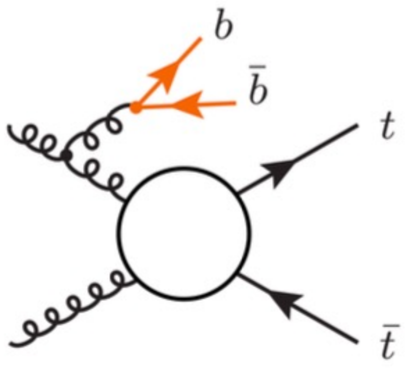
Full Run-2 analysis uses improved DeepJet b-tagging algorithm:

- **Improves** b-tagging efficiency by 5-10% at same mis-tag probability
- Operate at 75-80% signal efficiency, 1.5-2% mis-tag probability for light-flavored jets

More details here and here

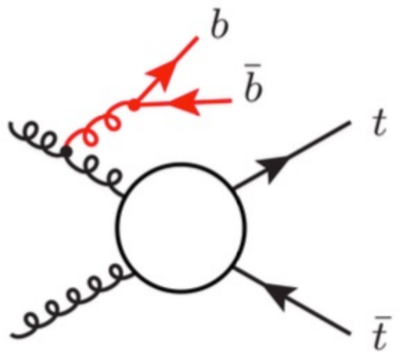
$t\bar{t} + B$ Background Model

One of the major improvements of the Full Run-2 analysis is the improved modeling of the $t\bar{t} + B$ irreducible background



In earlier versions of the analysis (including the 2016+2017 published analysis), $t\bar{t} + B$ processes described by:

- $t\bar{t}$+jet Matrix Element (ME) at NLO (5FS): additional b-jets from parton shower (PS)
- Subject to PS and large/not well-defined uncertainties of PS tuning parameters



In the full Run-2 analysis, $t\bar{t} + B$ processes described by:

- $t\bar{t}b\bar{b}$ ME at NLO (4FS)*: additional b-jets from ME
- NLO+PS accuracy for $t\bar{t}$ + 1 and 2 b-jet observables

Theoretically preferred option for $t\bar{t} + B$ modeling: improvement (w.r.t. to $t\bar{t}$ 5FS) in event kinematics

4FS: 4 flavor scheme
5FS: 5 flavor scheme

NLO: Next to Leading Order

* Eur. Phys. J. C 78 (2018) 502

$t\bar{t} + B$ Background Model

Details on the new $t\bar{t} + B$ MC sample:

	$\mathrm{t\bar{t}}$ sample	$\mathrm{t\bar{t}b\bar{b}}$ sample
POWHEG version	Powheg v2	Powheg-Box-Res
PYTHIA version	8.230	8.230
Flavour scheme	5	4
PDF set	NNPDF3.1	NNPDF3.1
m_{t}	172.5 GeV	172.5 GeV
m_{b}	0	4.75 GeV
μ_{R}	$\sqrt{\frac{1}{2}\left(m^2_{\mathrm{T,t}} + m^2_{\mathrm{T,\bar{t}}}\right)}$	$\frac{1}{2}\sqrt[4]{m_{\mathrm{T,t}} \cdot m_{\mathrm{T,\bar{t}}} \cdot m_{\mathrm{T,b}} \cdot m_{\mathrm{T,\bar{b}}}}$
μ_{F}	μ_{R}	$\frac{1}{4}\left[m_{\mathrm{T,t}} + m_{\mathrm{T,\bar{t}}} + m_{\mathrm{T,b}} + m_{\mathrm{T,\bar{b}}} + m_{\mathrm{T,g}}\right]$
h_{damp}	$1.379 \cdot m_{\mathrm{t}}$	$1.379 \cdot m_{\mathrm{t}}$
Tune	CP5	CP5

Based on theory recommendations, new $t\bar{t}b\bar{b}$ simulation (4FS): NLO accuracy simulation using Powheg-Box-Res (Jezo et al) with OpenLoops (Buccioni et al) in the 4FS

The scale choice in our $t\bar{t} + b\bar{b}$ MC sample is different by a factor of 2 from Eur. Phys. J. C 78 (2018) 502, motivated by later studies (J. High Energ. Phys. 2019, 15 (2019))

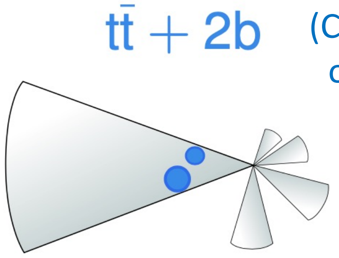
Comparison to ATLAS:

- Same nominal background model, with the exception that the scale is different by a factor 2 (but that should be covered by the scale uncertainties)
- Very different uncertainty models

$t\bar{t} + B$ Background Model

$t\bar{t}$ + Jets Modeling Uncertainties:

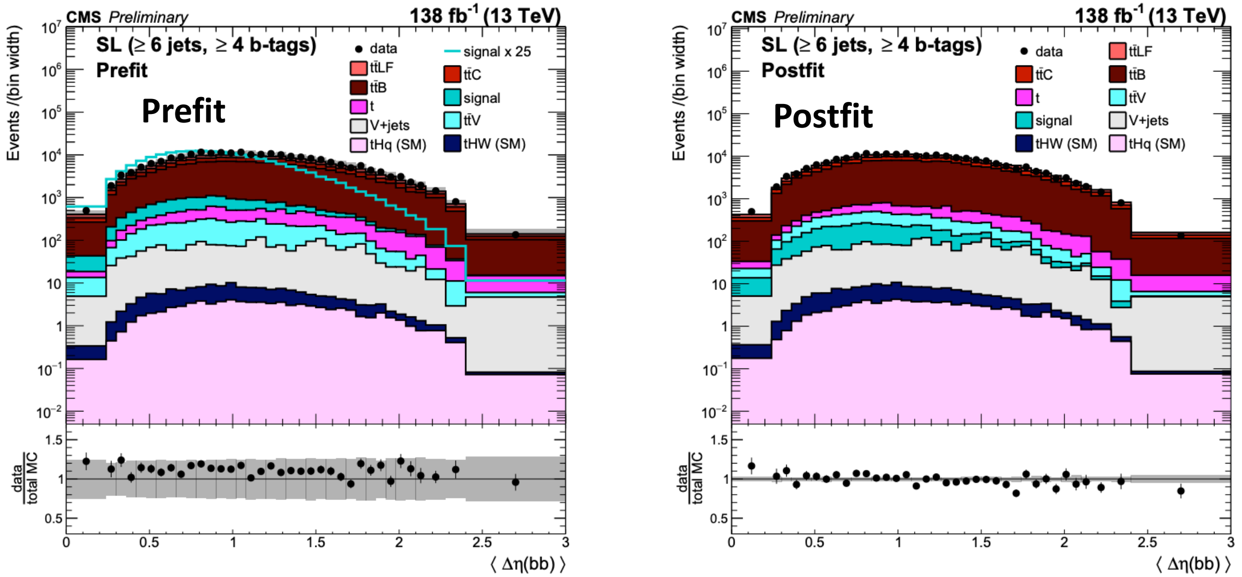
- ISR, FSR, μ_{R}, μ_{F} scale and PDF uncertainties decorrelated between $t\bar{t} + B$ and other $t\bar{t}$ events.
- ME-PS uncertainty decorrelated among $t\bar{t} + B$, $t\bar{t} + C$, $t\bar{t} + LF$
- Conservative uncertainty (100%) on the $t\bar{t} + 2b$ component (collinear $g \rightarrow b\bar{b}$ splitting)



(Component of $t\bar{t} + B$)

Good description of event kinematics on using $t\bar{t}b\bar{b}$ NLO (4FS)

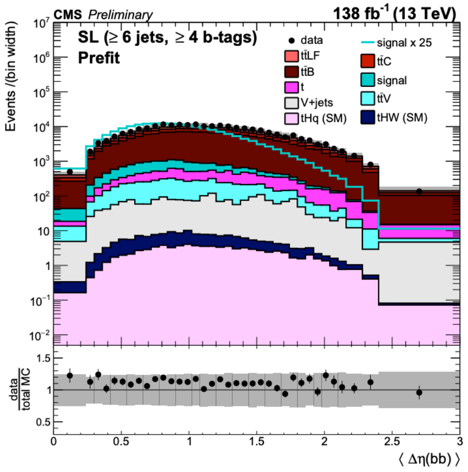
$\Delta\eta(b\bar{b})$ in SL Channel

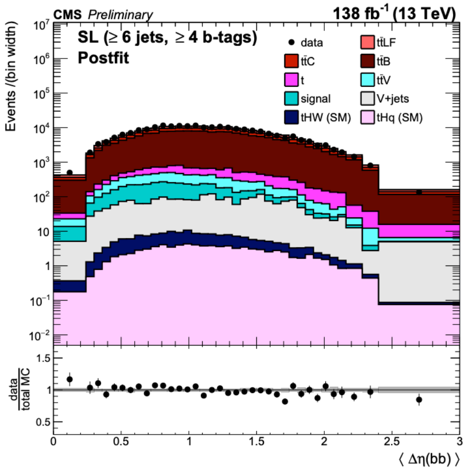


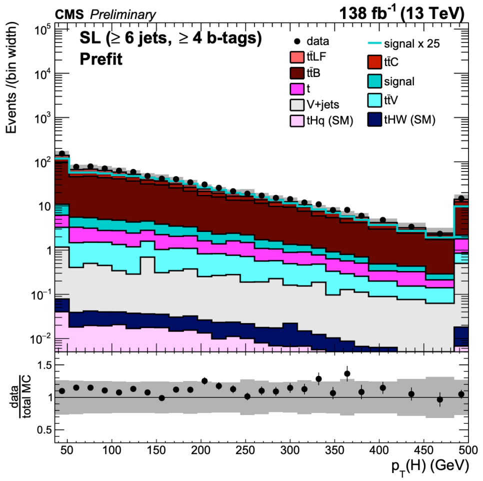
- Careful validation of the modeling through:
  - Goodness-of-fit tests
  - Bias tests on the signal strength
    - Test for potentially mismodelled $t\bar{t} + B$ background using toy data
    - Fit model was found to be robust against potential deviations of the $t\bar{t} + B$ in data from the nominal prediction
  - Pulls and impacts of nuisance parameters

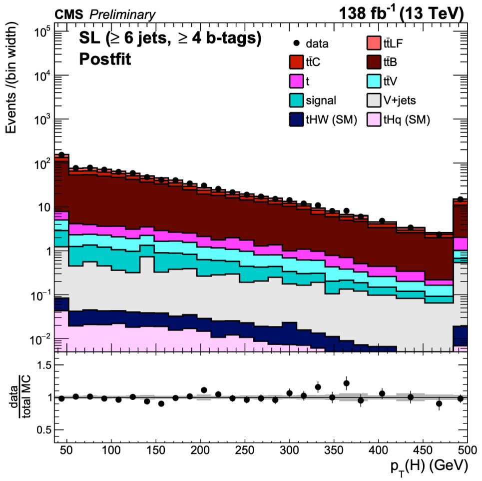
$t\bar{t} + B$ Background Model

SL Channel

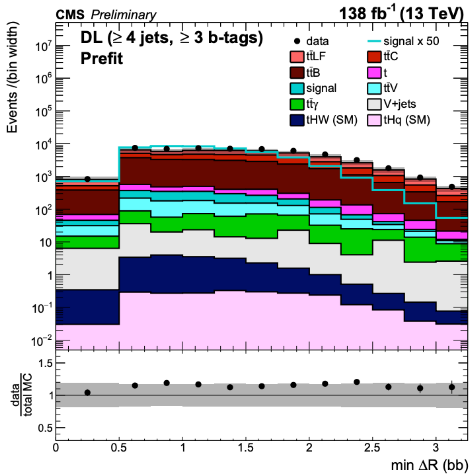


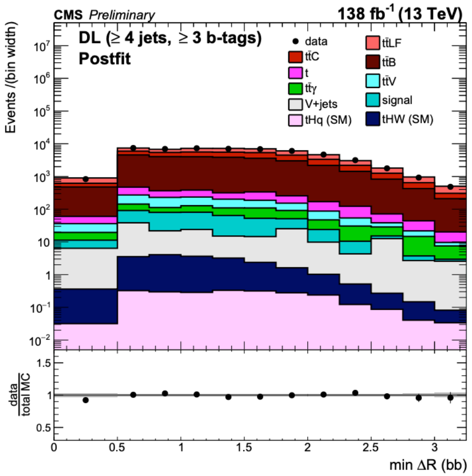


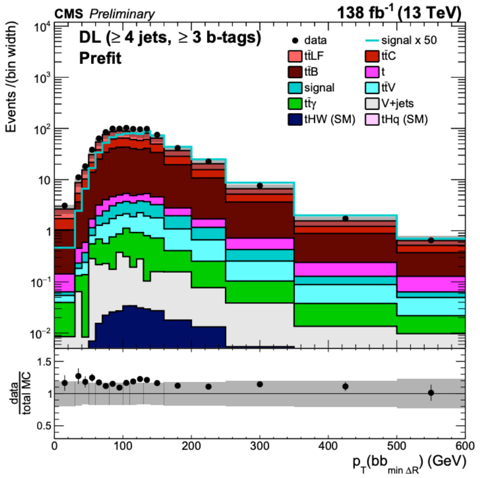


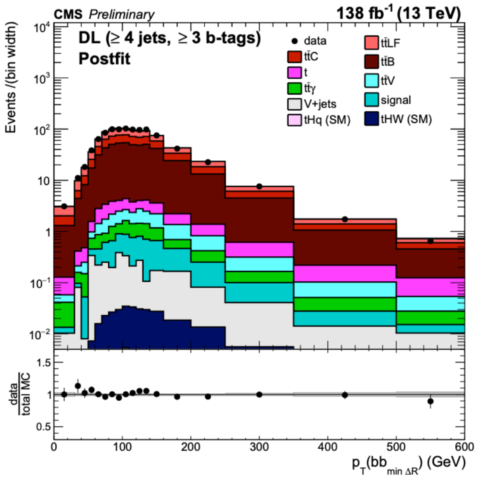


DL Channel









Validation of the $t\bar{t} + B$ Background Model

Checked the stability and robustness of the statistical model against statistical fluctuations in the data and against potentially mismodelled $t\bar{t} + B$ background using **toy data** (performed for **SL + DL**)

- Toys generated with injected $\mu_{t\bar{t}H} = 1$
- Figure of merit are the mean post-fit signal strength and mean bias

	mean value $\pm$ RMS		
$\mathrm{t\bar{t}B}$ component in pseudo data	$\mu_{\mathrm{t\bar{t}H}}$	$\mathrm{t\bar{t}B}$ norm	$\mathrm{t\bar{t}C}$ norm
$\mathrm{t\bar{t}b\bar{b}}$ sample (nominal)	1.03 ± 0.30	1.01 ± 0.09	1.01 ± 0.18
$\mathrm{t\bar{t}b\bar{b}}$ sample, $\mathrm{t\bar{t}B} \times 1.2$	1.03 ± 0.32	1.21 ± 0.15	1.01 ± 0.18
$\mathrm{t\bar{t}}$ sample	1.06 ± 0.30	1.03 ± 0.11	0.77 ± 0.18
$\mathrm{t\bar{t}}$ sample, $\mathrm{t\bar{t}B} \times 1.2$	1.06 ± 0.32	1.18 ± 0.12	0.85 ± 0.20

- Different sets of toy data were generated
- The nominal statistical model is fit to these toy data sets

- In all cases, the mean $\mu_{t\bar{t}H}$ is well compatible with the injected signal strength of 1, with biases well below 1σ
- The $t\bar{t} + B$ and $t\bar{t} + C$ background normalization parameters behave as expected

The fit model is therefore robust against potential deviations of the $t\bar{t} + B$ data from the nominal prediction, and it is able to compensate for the expected underprediction of the $t\bar{t} + B$ cross section

Analysis Strategy

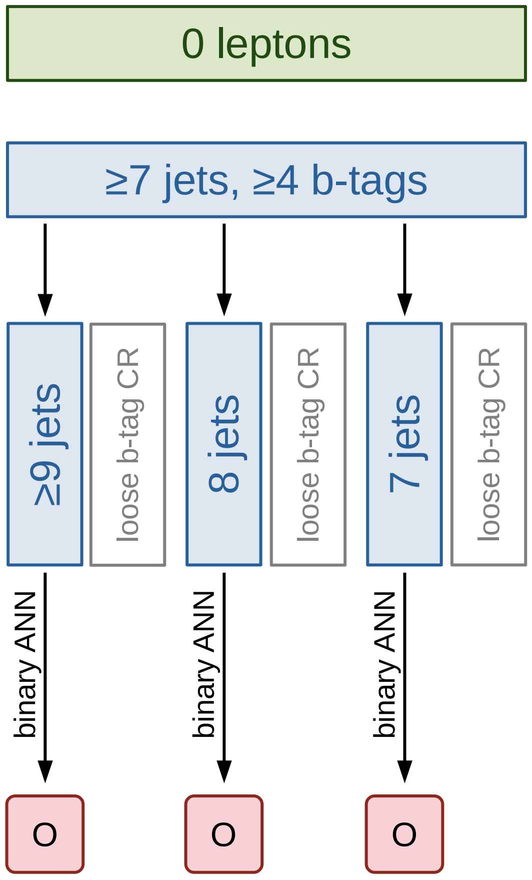
The analysis strategy relies on Event Categorization and Artificial Neural Networks (ANN)

- Event categorization to form signal and control regions (to constrain background):
  - Based on jet and b-tag multiplicity
  - Based on multi-class ANNs
- Artificial Neural Networks (ANN):
  - Trained to separate signal from dominant background
  - Binary or multi-class depending on channel/category
  - Used for event categorization and as final discriminants

ANN Training:

- Trained on several kinematic variables
  - Modelling of input variables validated with goodness-of-fit tests
- Usually trained on Monte-Carlo, except QCD (trained in a QCD enriched control region)
- One ANN training valid for all years in each channel and category

Inclusive $t\bar{t}H$ Measurement Strategy



Full Hadronic Channel:

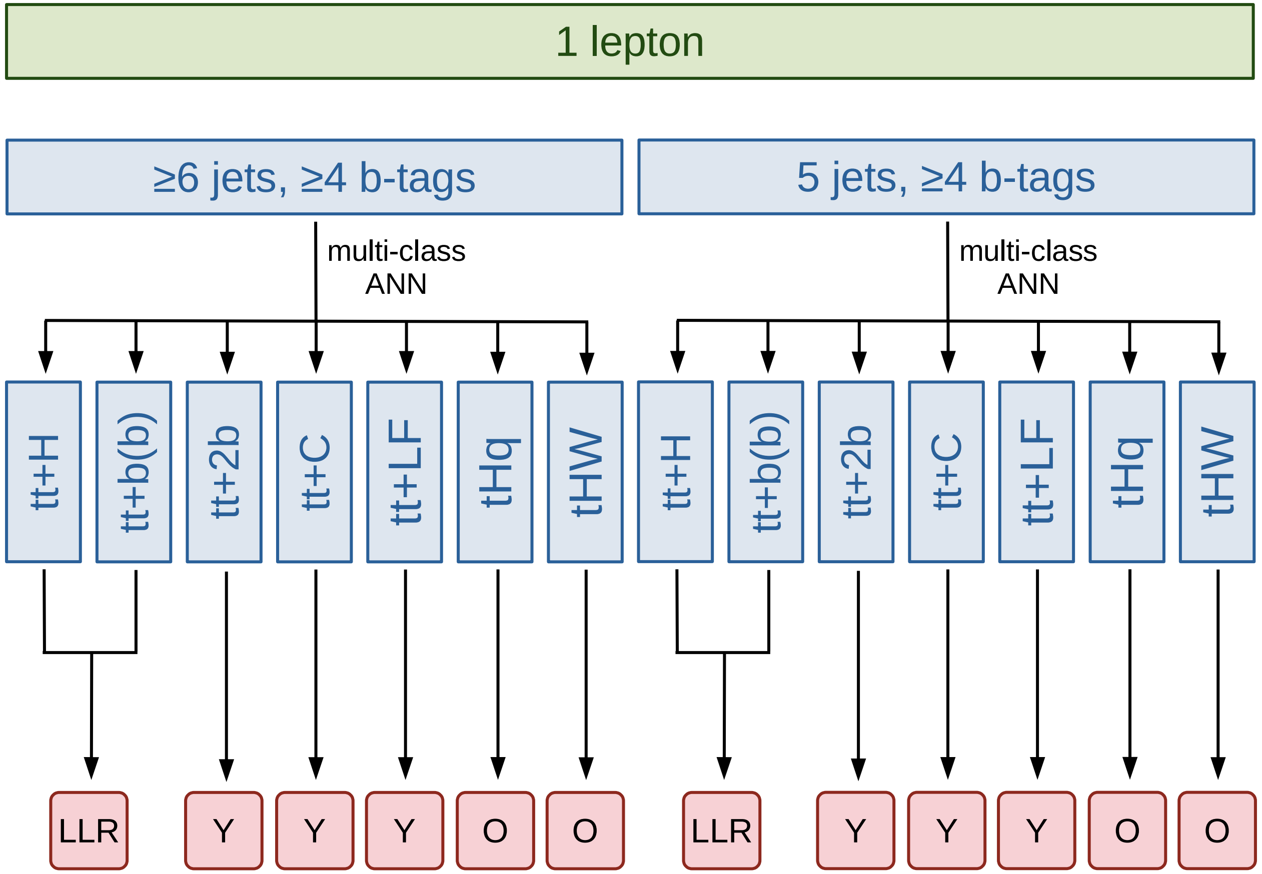
- Event categorization based on number of jets and b-tags
- Control regions based on loosely-tagged b-jets for ANN training and data-driven QCD estimation
- Binary ANN to separate signal from background (mainly QCD):
  - ANN output used as final fit observable

Distribution in template fit, event yield (Y), ANN output (O), likelihood ratio of ANN outputs (LLR)

Inclusive $t\bar{t}H$ Measurement Strategy

Single Lepton Channel:

- Event categorization based on number of jets and b-tags
- Further multiclassification based on ANN



Final fit observables:

- Event yields in background enriched nodes
- ANN output for tHq and tHW signal regions
- Ratio observable in combined $t\bar{t}H$ and $t\bar{t}B$ category

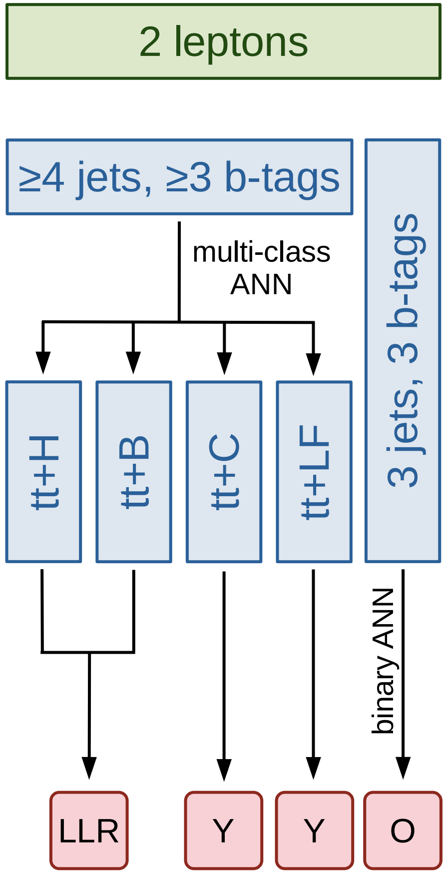
$$R_{SL} = \frac{O(t\bar{t}H)}{O(t\bar{t}H) + O(t\bar{t} + b(\bar{b})) + O(t\bar{t} + 2b)}$$

Distribution in template fit, event yield (Y), ANN output (O), likelihood ratio of ANN outputs (LLR)

Inclusive $t\bar{t}H$ Measurement Strategy

Dilepton Channel:

- Event categorization based on number of jets and b-tags
- Further multiclassification based on ANN in the more signal sensitive category
- **Final fit observables:**
  - Event yields in background enriched nodes
  - Binary ANN output in less signal sensitive region
  - Ratio observable in combined $t\bar{t}H$ and $t\bar{t}B$ category (more signal sensitive)

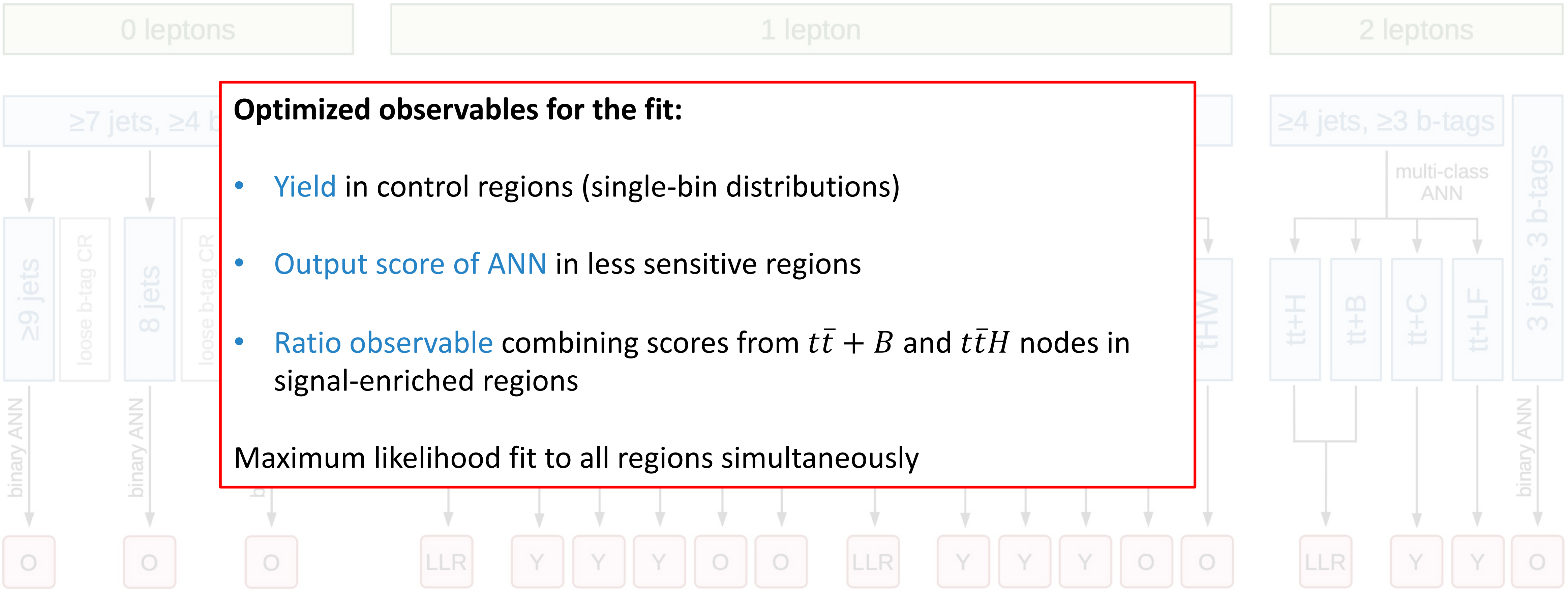


Distribution in template fit, event yield (Y), ANN output (O), likelihood ratio of ANN outputs (LLR)

$$R_{DL} = \frac{O(t\bar{t}H)}{O(t\bar{t}H) + O(t\bar{t}B)}$$

$$R_{SL} = \frac{O(t\bar{t}H)}{O(t\bar{t}H) + O(t\bar{t} + b(\bar{b})) + O(t\bar{t} + 2b)}$$

Inclusive $t\bar{t}H$ Measurement Strategy



Optimized observables for the fit:

- Yield in control regions (single-bin distributions)
- Output score of ANN in less sensitive regions
- Ratio observable combining scores from $t\bar{t} + B$ and $t\bar{t}H$ nodes in signal-enriched regions

Maximum likelihood fit to all regions simultaneously

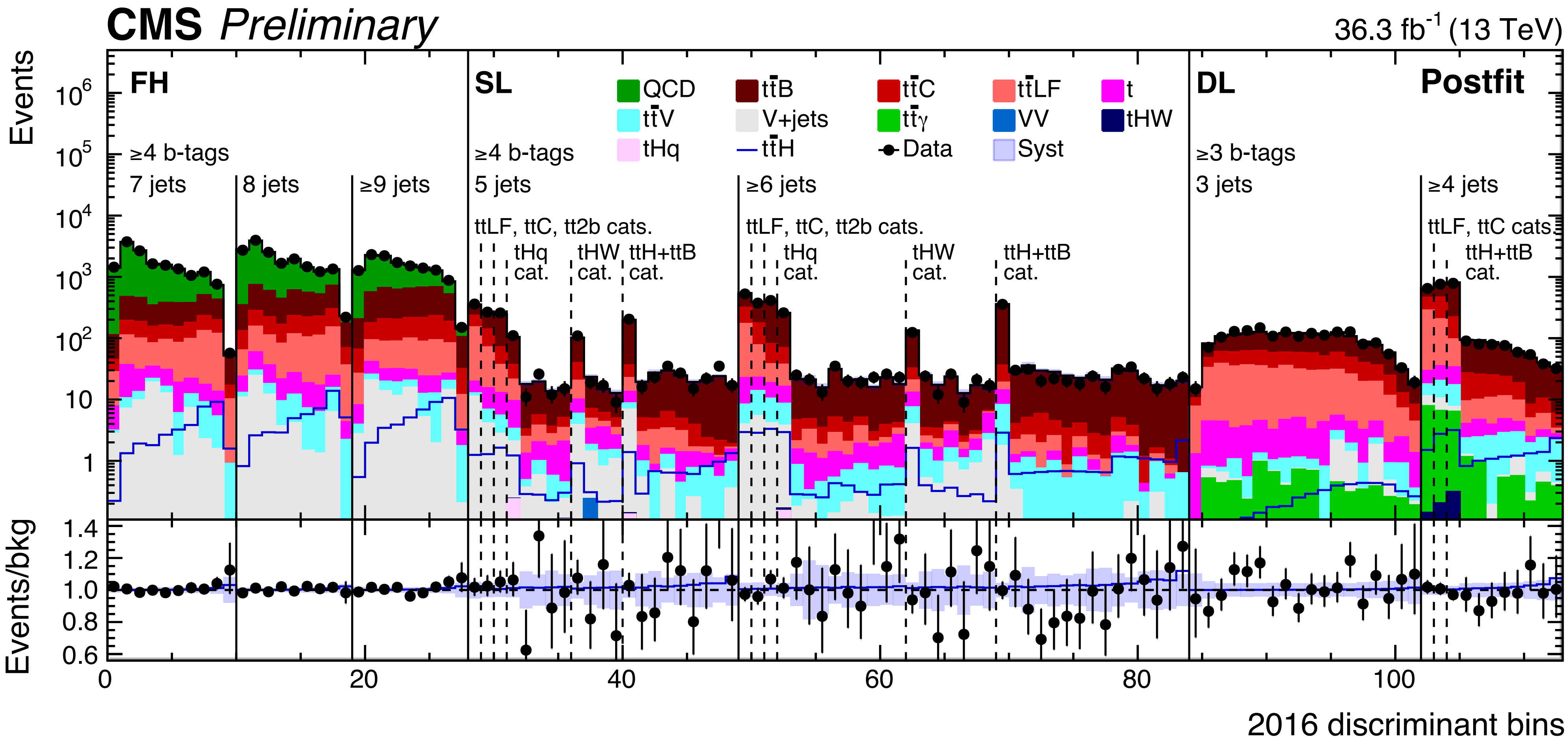
Distribution in template fit, event yield (Y), ANN output (O), likelihood ratio of ANN outputs (LLR)

$$R_{DL} = \frac{O(t\bar{t}H)}{O(t\bar{t}H) + O(t\bar{t}B)}$$

$$R_{SL} = \frac{O(t\bar{t}H)}{O(t\bar{t}H) + O(t\bar{t} + b(\bar{b})) + O(t\bar{t} + 2b)}$$

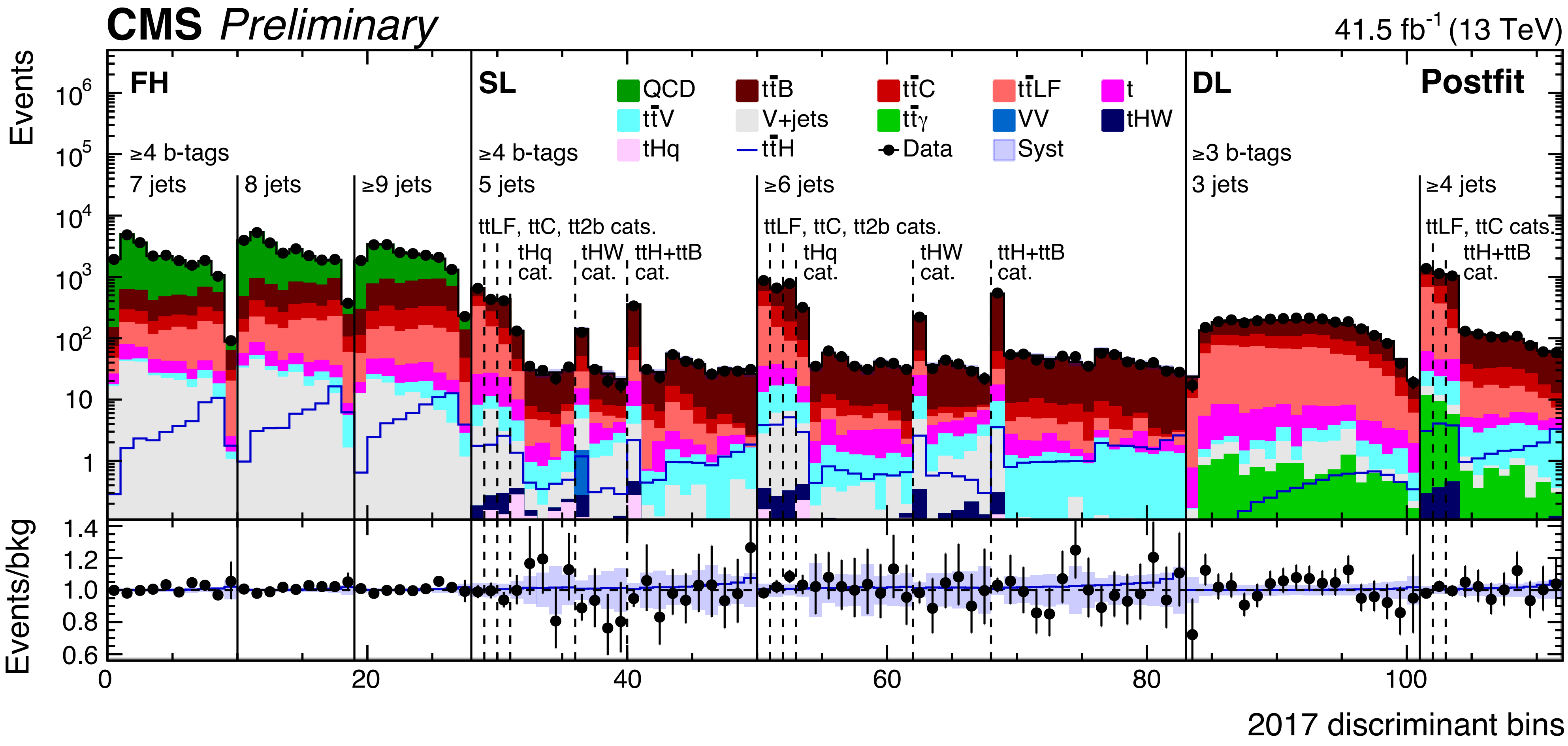
Inclusive $t\bar{t}H$ Results: Postfit Distributions

Postfit distributions from 2016:



Inclusive $t\bar{t}H$ Results: Postfit Distributions

Postfit distributions from 2017:

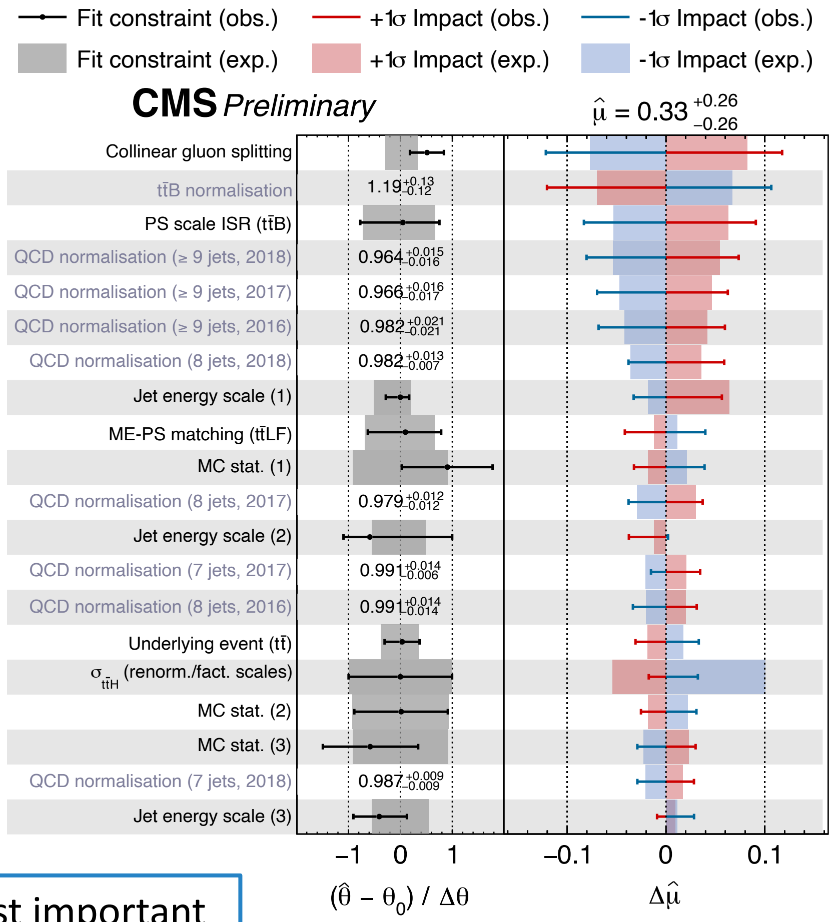


Inclusive $t\bar{t}H$ Results: Systematic Uncertainties

Major sources of systematic uncertainties:

Uncertainty source	$\Delta\mu_{t\bar{t}H}$ (observed)	$\Delta\mu_{t\bar{t}H}$ (expected)
Total experimental	$+0.10/-0.10$	$+0.11/-0.10$
jet energy scale and resolution	$+0.08/-0.07$	$+0.09/-0.09$
b tagging	$+0.07/-0.06$	$+0.06/-0.02$
luminosity	$+0.02/-0.02$	$+0.01/-0.01$
Total theory	$+0.16/-0.16$	$+0.18/-0.14$
$t\bar{t}$ + jets background	$+0.15/-0.16$	$+0.12/-0.11$
signal modelling	$+0.06/-0.01$	$+0.13/-0.06$
Size of the simulated event samples	$+0.13/-0.12$	$+0.10/-0.10$
Total systematic	$+0.20/-0.21$	$+0.23/-0.19$
Statistical	$+0.17/-0.16$	$+0.17/-0.17$
background normalisation	$+0.13/-0.13$	$+0.13/-0.13$
$t\bar{t}B$ and $t\bar{t}C$ normalisation	$+0.12/-0.12$	$+0.12/-0.12$
QCD normalisation	$+0.01/-0.01$	$+0.01/-0.01$
Total	$+0.26/-0.26$	$+0.28/-0.25$

Impacts and pulls of systematic uncertainties:

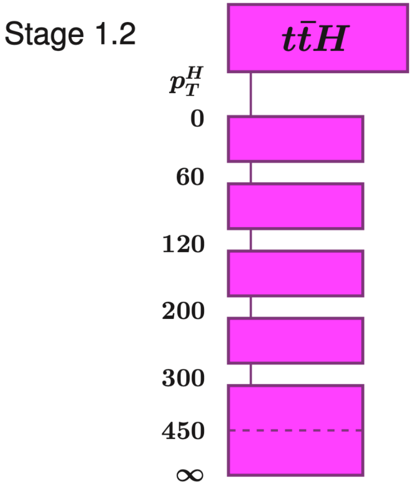


$t\bar{t}$ + *jets* uncertainties most important

$t\bar{t}H$ Measurement in Higgs Boson p_{T} Bins

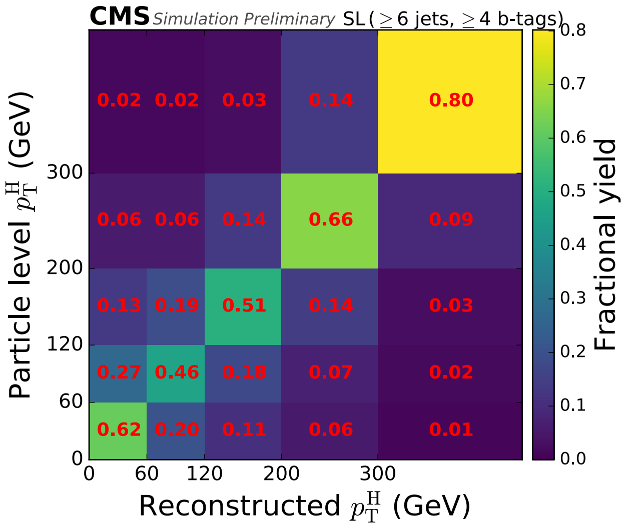
$t\bar{t}H$ cross-section measured in 5 Higgs boson p_{T} ($p_{\mathrm{T}}^{\mathrm{H}}$) bins:
- $t\bar{t}H$ signal split using generator level $p_{\mathrm{T}}^{\mathrm{H}}$

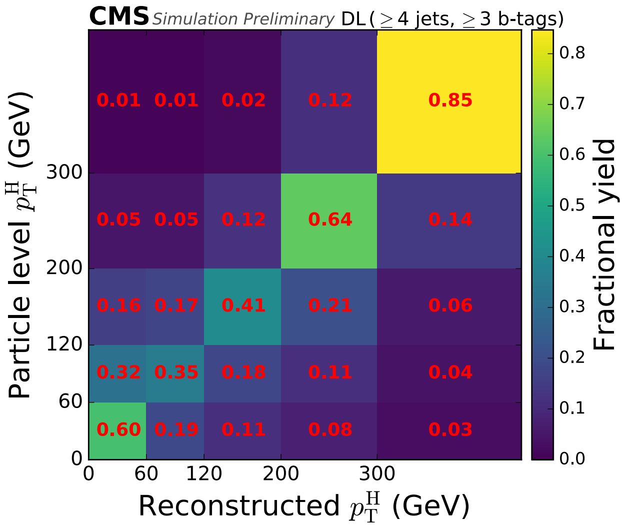
Simplified Template Cross-Section (STXS) approach



Perform reconstruction of the Higgs boson p_{T} bins:
- In FH channel: χ^2 reconstruction of the Higgs from b-jet pairs
- In SL and DL channels: multi-class ANN trained on $t\bar{t}H(b\bar{b})$

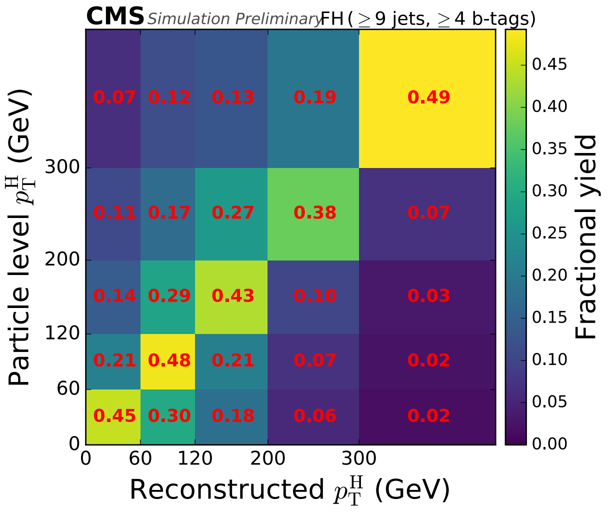
Assignment efficiency between 35-85%, depending on p_T bin and category

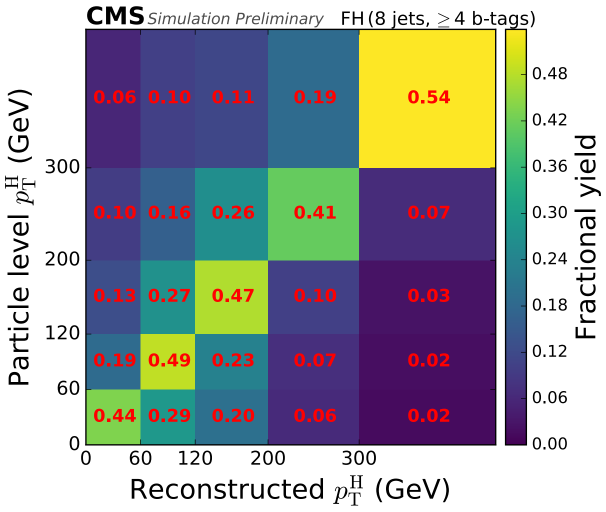


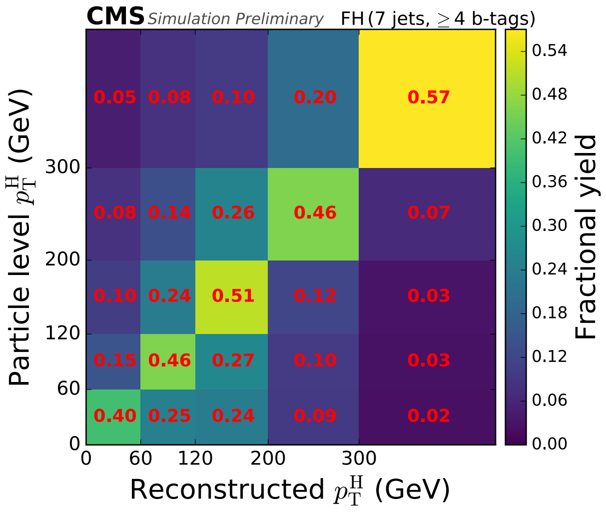


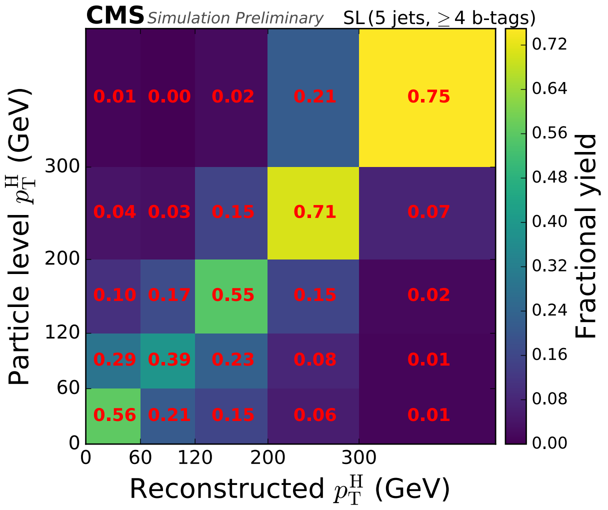
Additional systematic uncertainties:
- $\mu_{\mathrm{R}}/\mu_{\mathrm{F}}$ scale uncertainties merged into μ
- Migration uncertainties introduced:
  - Replace renorm./ fact. scale for signal
- ISR/FSR/μ split between low and high p_{T}
- Partial decorrelation of $t\bar{t} + jets$ parameters:
  - ISR/FSR, ME-PS matching
  - Collinear gluon splitting
  - Freely-floating $t\bar{t} + B$ normalization

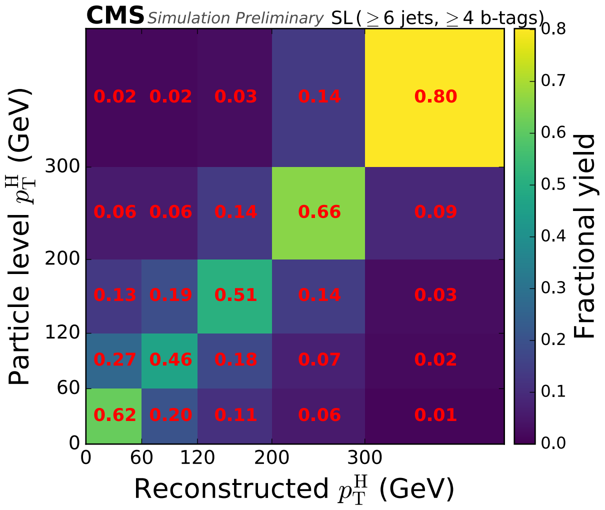
$t\bar{t}H$ Measurement in Higgs Boson p_{T} Bins

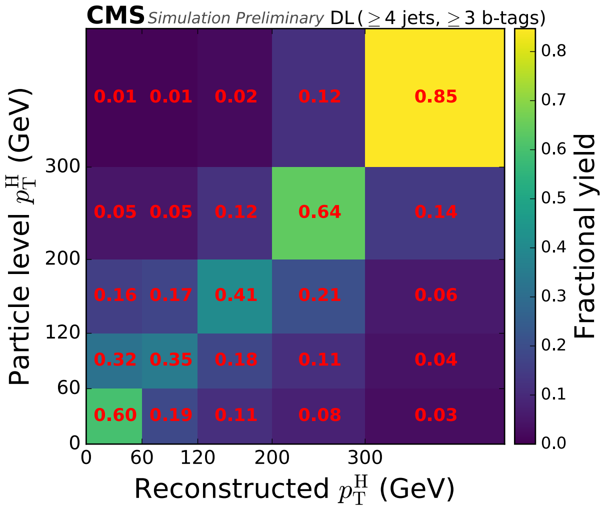






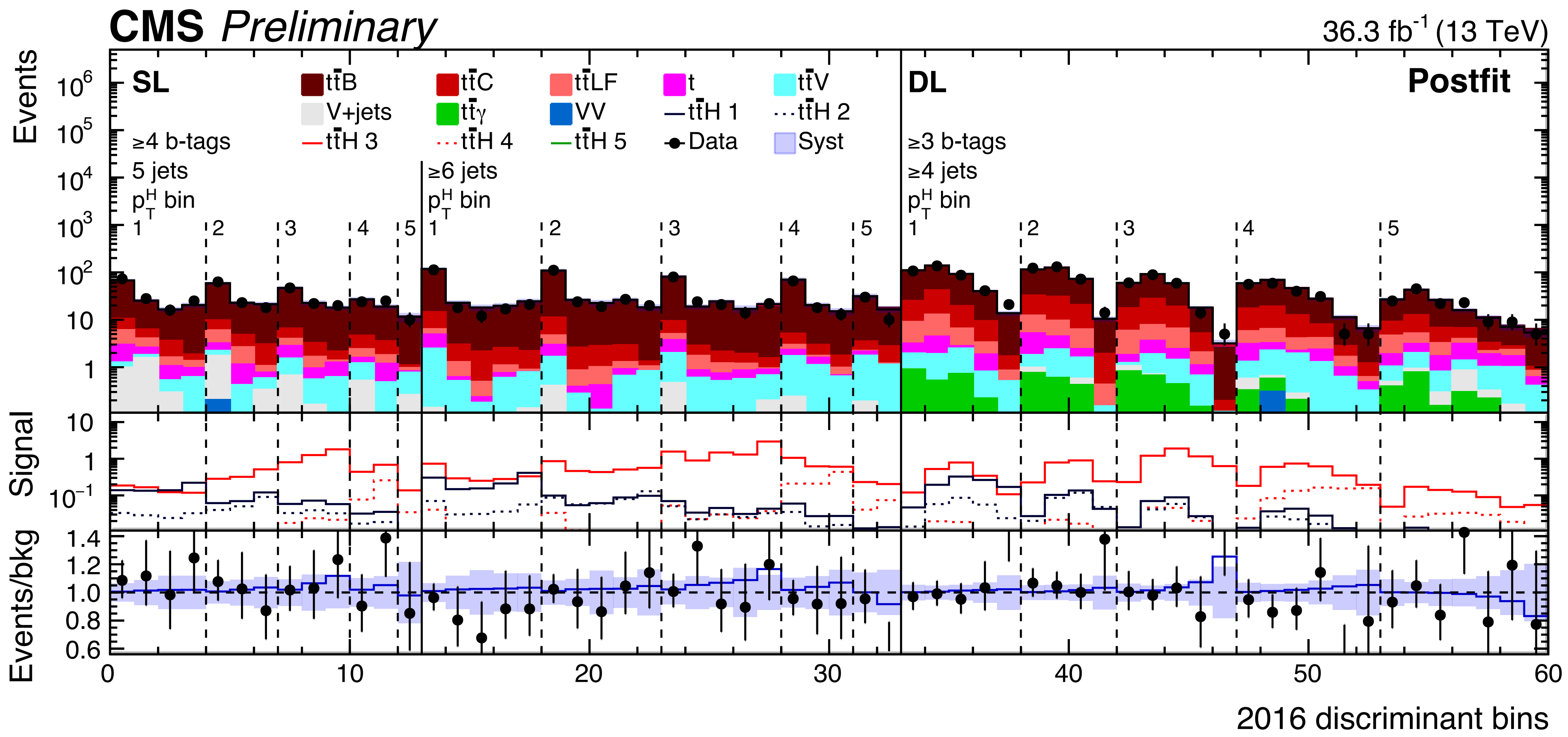






$t\bar{t}H$ Measurement in Higgs Boson p_{T} Bins

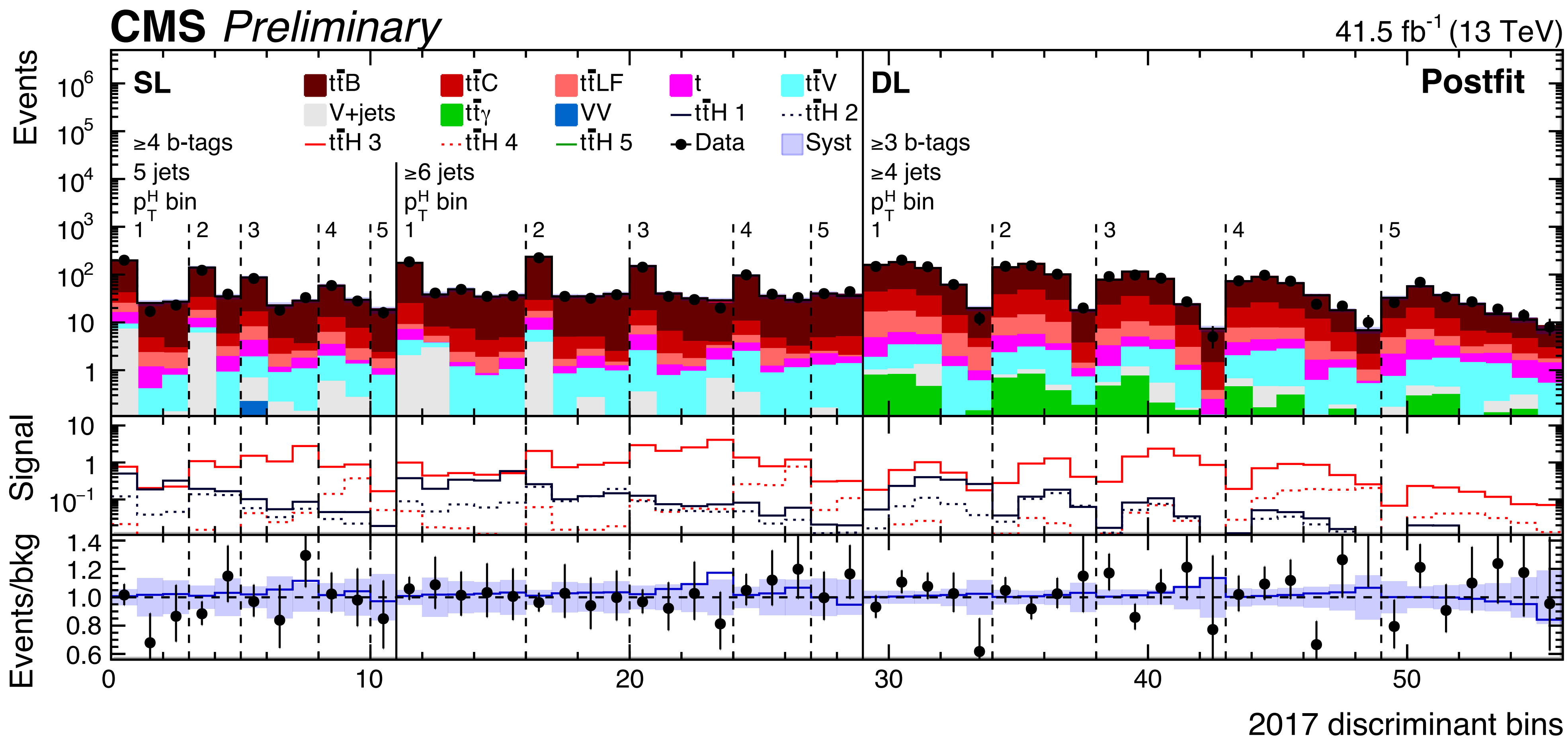
Postfit distributions in SL and DL channels in Higgs p_{T} bins from 2016:



Fitted observables in this plot are the output of the Higgs p_{T} ANN times the ratio observable from the inclusive ANN

$t\bar{t}H$ Measurement in Higgs Boson p_{T} Bins

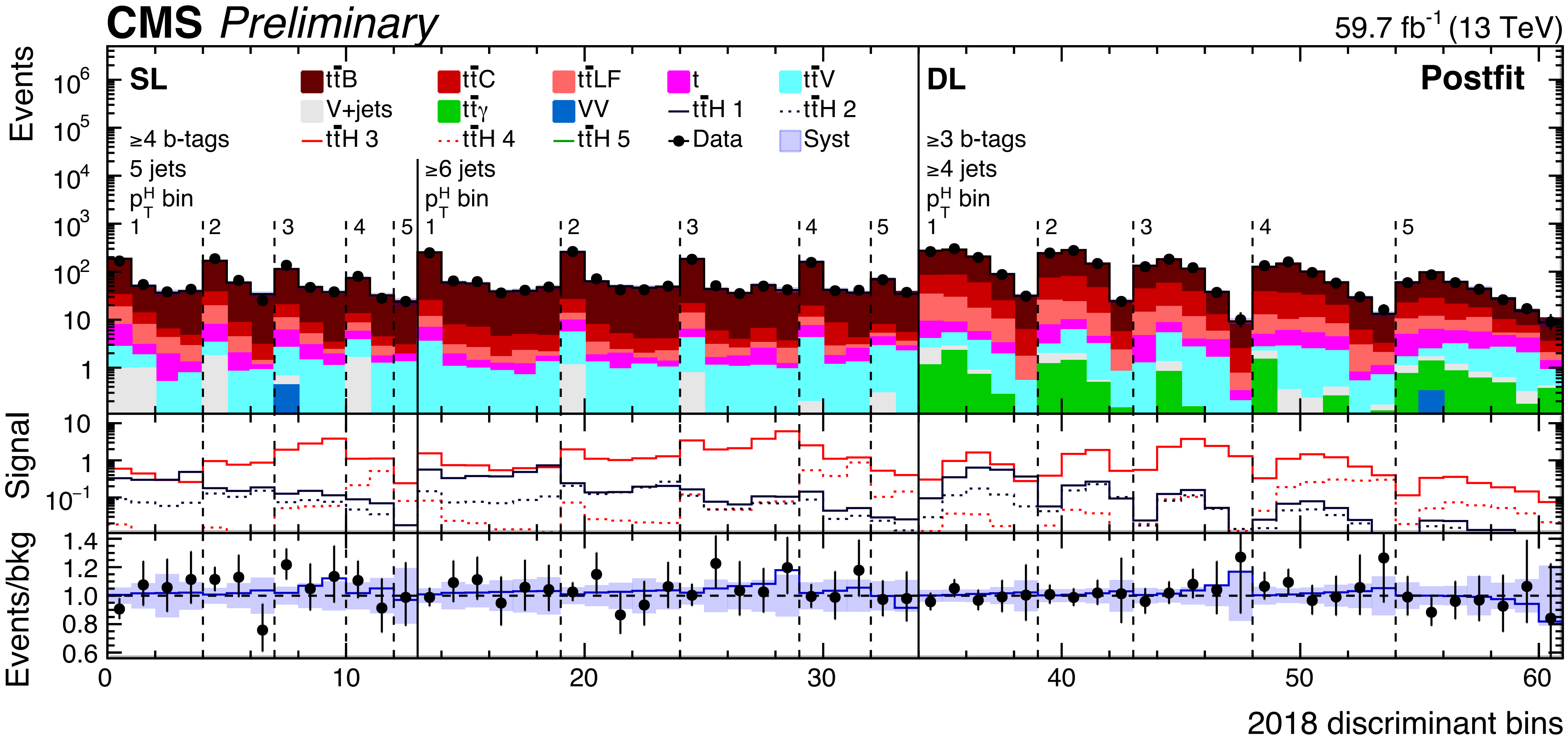
Postfit distributions in SL and DL channels in Higgs p_{T} bins from 2017:



Fitted observables in this plot are the output of the Higgs p_{T} ANN times the ratio observable from the inclusive ANN

$t\bar{t}H$ Measurement in Higgs Boson p_{T} Bins

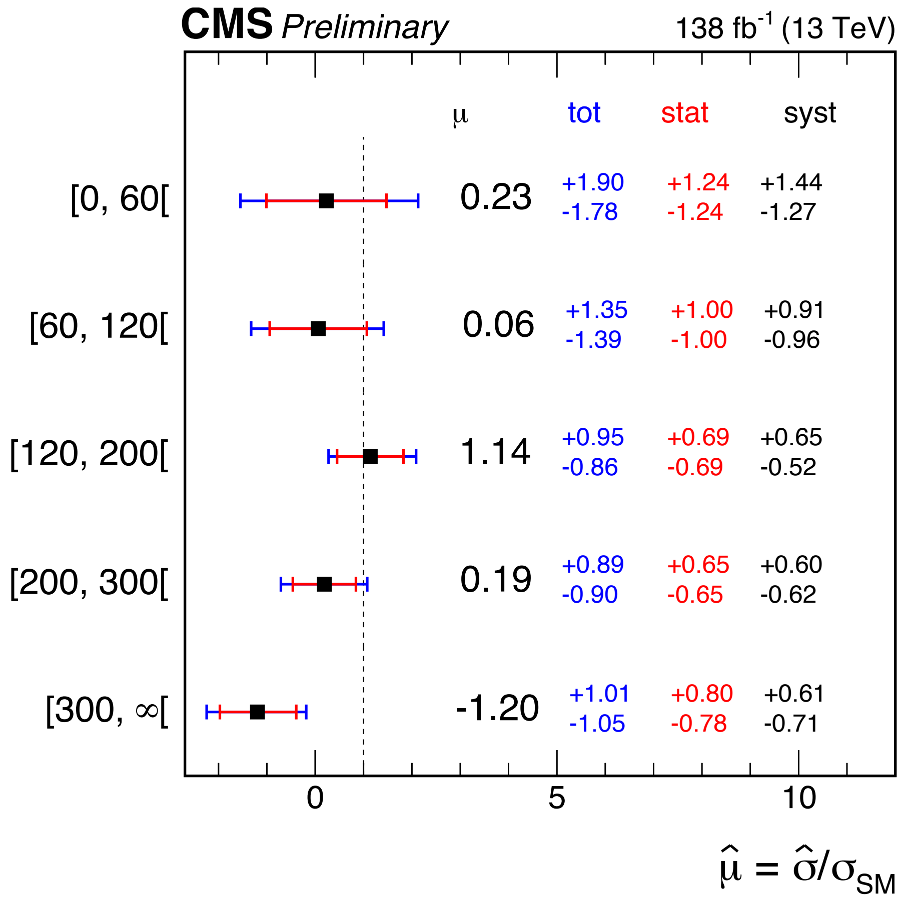
Post-fit distributions in SL and DL channels in Higgs p_{T} bins from 2018:



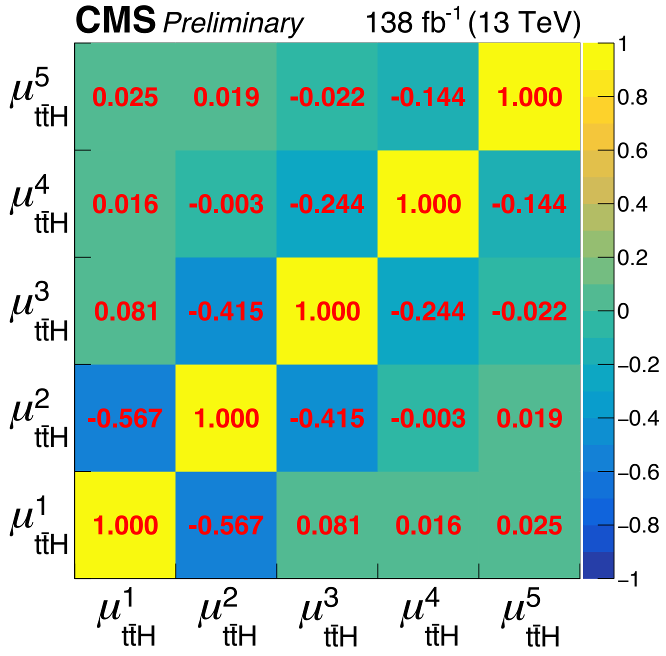
Fitted observables in this plot are the output of the Higgs p_{T} ANN times the ratio observable from the inclusive ANN

$t\bar{t}H$ Measurement in Higgs Boson p_T Bins

Full Run-2 Results



Correlations among the signal strengths



Significant correlations in the first few bins

Compatibility with:

- Inclusive results: p-value of 0.67 ($0.4\ \sigma$)
  - Additional single parameter fit gives signal strength within 3% of the inclusive result: completely compatible
- SM: p-value of 0.21 ($1.3\ \sigma$)

$t\bar{t}H$ Measurement in Higgs Boson p_{T} Bins

Impacts and pulls of systematic uncertainties: