Joker Talk

Opening
new
horizons

When charm and beauty adjoin the top

First measurement of the cross section of top quark pair production with additional charm jets using the dilepton final state in pp collisions at $\sqrt{s}$=13 TeV

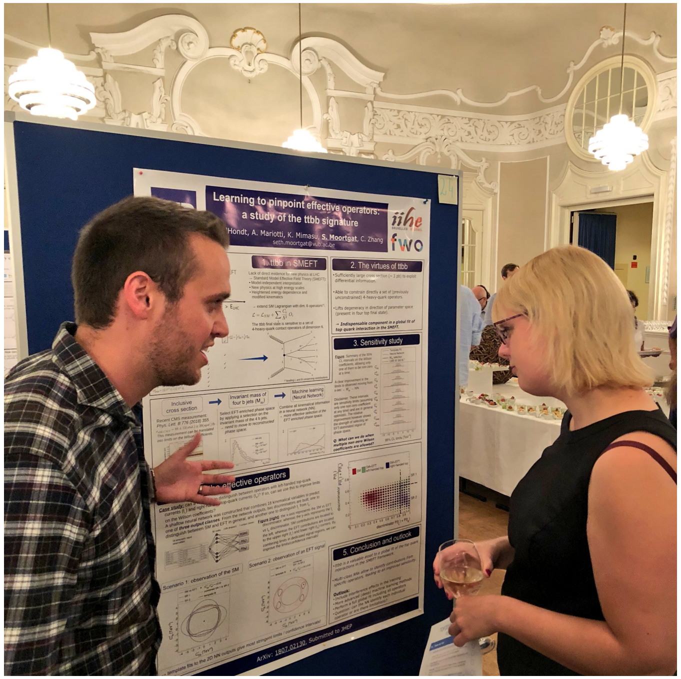
Seth Moortgat

Interuniversity Institute for High Energies (IIHE)
Vrije Universiteit Brussel
seth.moortgat@cern.ch

On behalf of the CMS Collaboration

PAS-TOP-20-003

TOP2020 (virtual), 16-09-2020



Throwback: TOP2018 - Bad Neuenahr

Poster session:

[$t\bar{t}b\bar{b}$ in the SMEFT using ML]

Plenaries:

[Theory progress on $t\bar{t}H(b\bar{b})$ backgrounds]
(S. Pozzorini)

[$t\bar{t}b\bar{b}$ @ CMS and ATLAS] (A. Khanov)

Take-home messages (with personal bias):

1. Theoretical modelling of $t\bar{t}$+heavy-flavour (HF) jets is hard! (but if affects the $t\bar{t}H(b\bar{b})$ measurement)
2. You can not simply consider $t\bar{t}b\bar{b}$ without considering at the same time $t\bar{t}c\bar{c}$ and $t\bar{t}$ + light-flavour jets ($t\bar{t}$LF).
3. Not only b-tagging, but c-tagging is crucial!

First measurement of the inclusive $t\bar{t}c\bar{c}$ cross section

Simultaneously measure $\sigma(t\bar{t}c\bar{c})$, $\sigma(t\bar{t}b\bar{b})$, $\sigma(t\bar{t}LF)$

and $R_{c/b} = \sigma(t\bar{t}+c\bar{c}/b\bar{b}) / \sigma(t\bar{t}+jj)$

Measurement performed in the dilepton channel

Data collected by CMS in 2017, corresponding to 41.5 fb^{-1} of integrated luminosity

Key ingredients:

- Use neural network for matching jets to partons.
- Rely on charm-jet identification to separate the different signals!
- Calibrate the c-tagger discriminants (full shape)

Definition of heavy-flavor jets

Heavy-flavor definition in simulation based on ghost hadron clustering *Phys.Lett.B* 659 (2008) 119-126

Fiducial phase space

- $pp \rightarrow t\bar{t}jj \rightarrow \ell^+\bar{\nu_\ell}b\ell^-\nu_\ell\bar{b}jj$ (dilepton)
- Two generated leptons with p_T > 25 GeV and $|\eta|$<2.4 (electron/muon/tau)
- Two particle-level b jets from top quark decay with p_T > 20 GeV and $|\eta|$<2.4
- At least two additional particle-level jets (not from top quark decay) with p_T > 20 GeV and $|\eta|$<2.4 and ΔR(l,jet)>0.4

Full phase space

- $pp \rightarrow t\bar{t}jj \rightarrow W^+bW^-\bar{b}jj$
- dilepton / single lepton / all-hadronic
- At least two additional particle−level jets (not from top quark decay) with pT > 20 GeV and $|\eta|$<2.4 and ΔR(l,jet)>0.4

Categorization based on flavor of additional jets

- $t\bar{t}b\bar{b}$: $\geq$ 2 add. b jets with at least one b hadron
- $t\bar{t}bL$: 1 add. b jet with at least one b hadron (merged or missing jet)
- $t\bar{t}c\bar{c}$: $\geq$ 2 add. c jets with at least one c hadron (if not $t\bar{t}b\bar{b}$/ $t\bar{t}bL$)
- $t\bar{t}cL$: 1 add. c jet with at least one c hadron (if not $t\bar{t}b\bar{b}$/L, merge/missing jet)
- $t\bar{t}LF$: no add. b or c jets, but 2 add. light jets pass acceptance requirements.
- $t\bar{t}$ other: failing visible/full phase space requirements

Dileptonic top quark pair events + 2 additional jets

Global
- ==2 isolated leptons (e/μ)
- >= 4 jets
- >= 2 b-tagged jets

Electrons
- p_T > 25 GeV
- $|\eta|$ < 2.4
- $|\eta| \notin$ [1.4442 - 1.566] ("transition region")
- Rel. Iso < 0.15

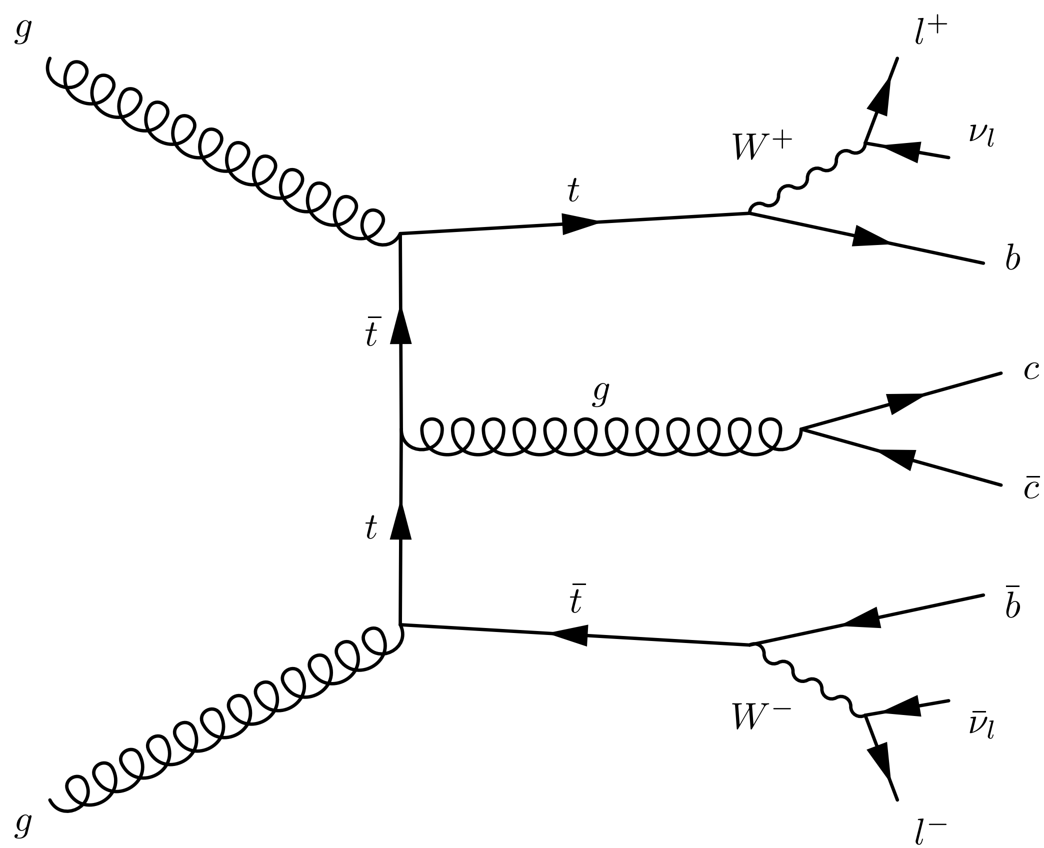
Muons
- p_T > 25 GeV
- $|\eta|$ < 2.4
- Rel. Iso < 0.15

Dilepton invariant mass
- m_{ll} > 12 GeV
- $\mu\mu/ee$: $m_{ll} \notin$ [m_Z -15 GeV , m_Z +15 GeV]

MET
- $\mu\mu/ee$: MET>30 GeV

Jets
- p_T > 30 GeV
- $|\eta|$ < 2.5
- ΔR(lepton,jet)>0.5
- DeepCSV value > 0

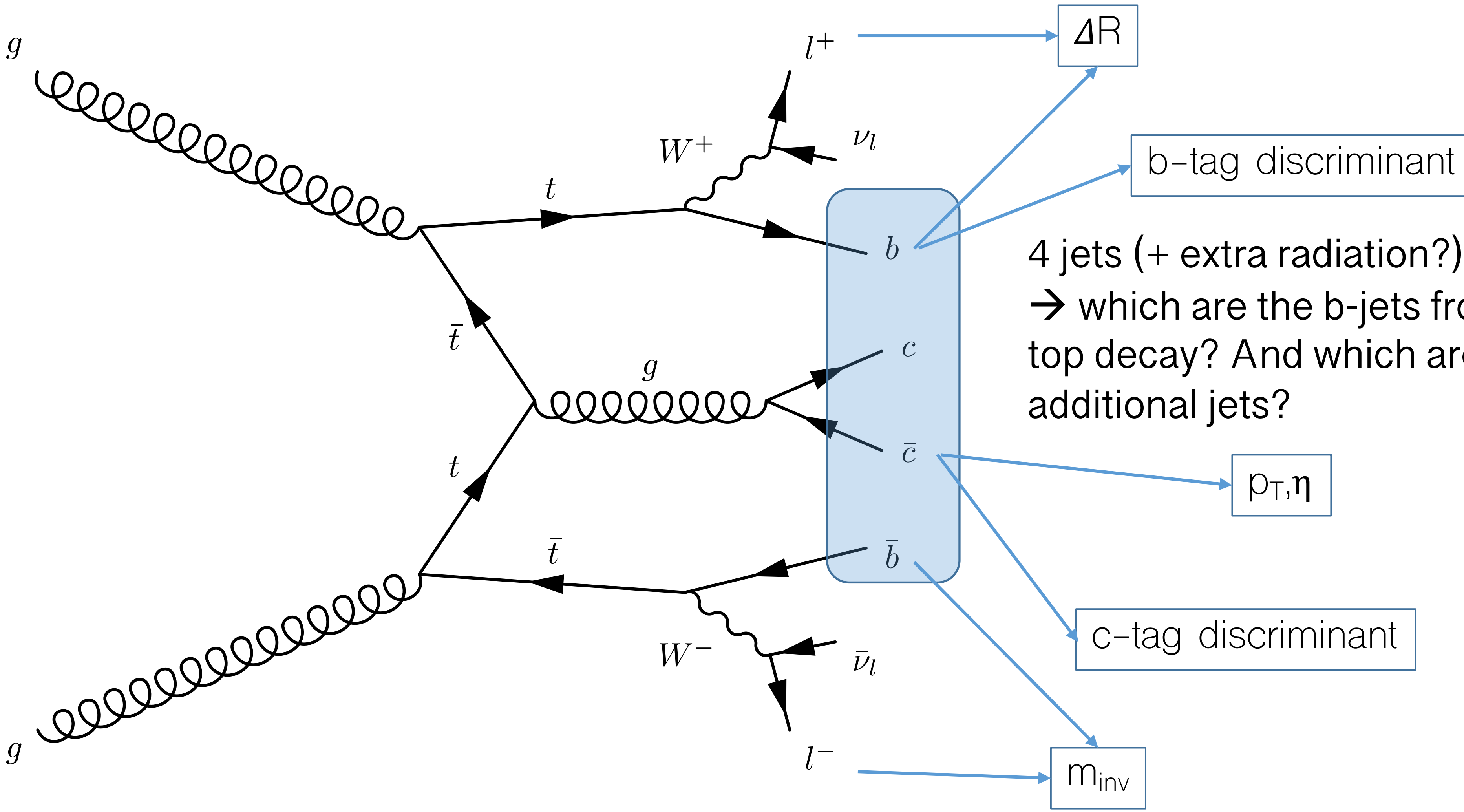
b-jets/c-jets
- 2 top-matched jets: Medium DeepCSV b-tagged



Results in >95% $t\bar{t}$ events

Jet-parton matching

Event kinematics + jet flavour as input to a neural network (NN)



4 jets (+ extra radiation?)
→ which are the b-jets from the top decay? And which are the additional jets?

→ Combine in a NN and pick the best jet-parton assignment

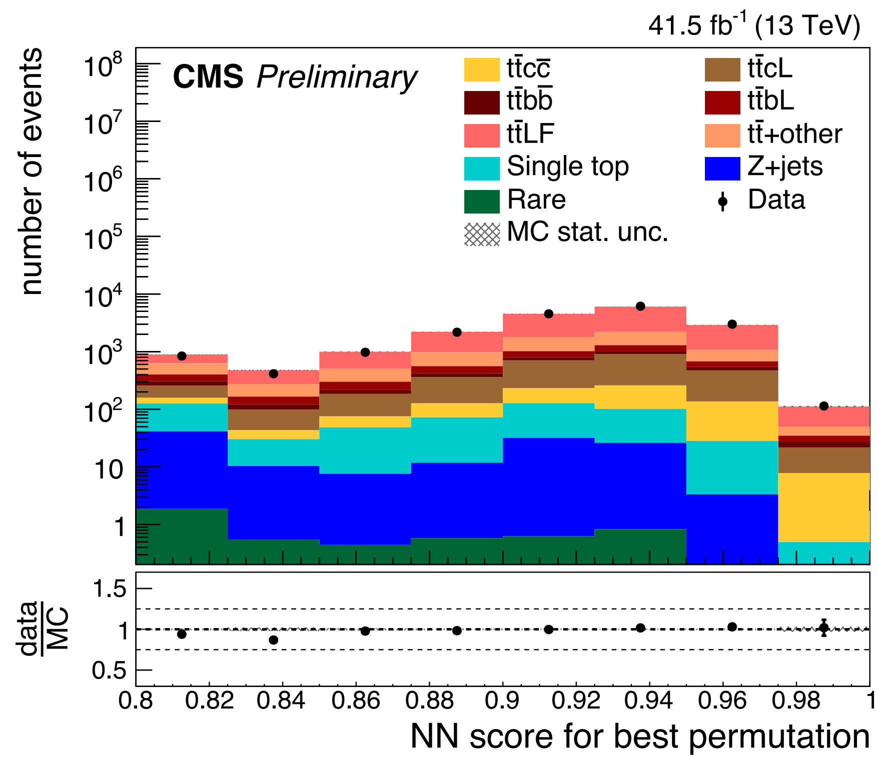
Jet-parton matching

Performance and neural network output

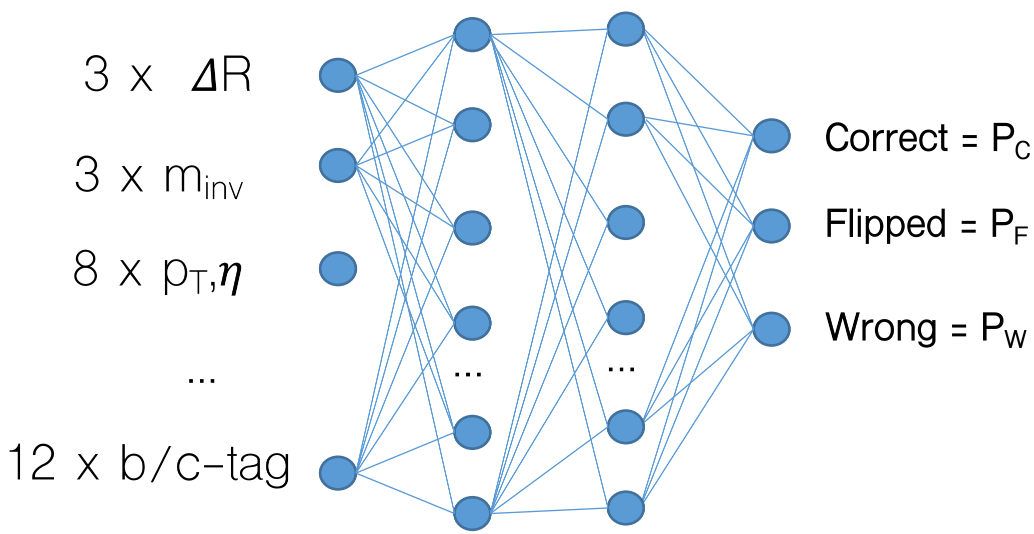
Only ~ 76% of events have two b jets matched to two gen-level b quarks from top quark within $\Delta R<0.3$. Only these are used in the training of the NN.

The network correctly identifies the two additional c (b) jets in **50% (30%)** of the cases for $t\bar{t}c\bar{c}$ ($t\bar{t}b\bar{b}$) events.

Good agreement between the data (black markers) and the simulation (filled histograms).



Two hidden layers that comprise 50 neurons each, with ReLu activation functions and a 10% dropout



NN score for best permutation

$$= \max\left(\frac{P_C}{P_C+P_W}\ ,\ \frac{P_F}{P_F+P_W}\right)$$

Charm jet identification using the DeepCSV algorithm

The DeepCSV heavy-flavour tagging algorithm is a multi-class algorithm that predicts probabilities (P) for jets to originate from a b, c or light-flavour (udsg) quark (or gluon).

This discrimination is based on properties such as track displacement, secondary vertex mass/flight distance, …

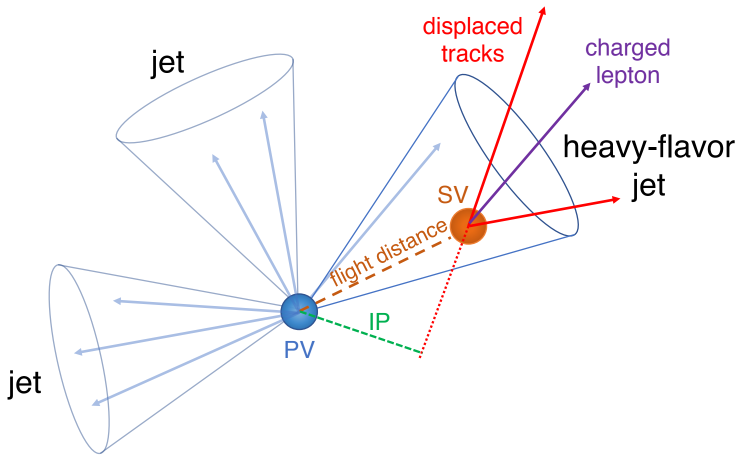
Properties from c jets are distributed midway between those of b or light-flavour jets → two c-tagging discriminants!

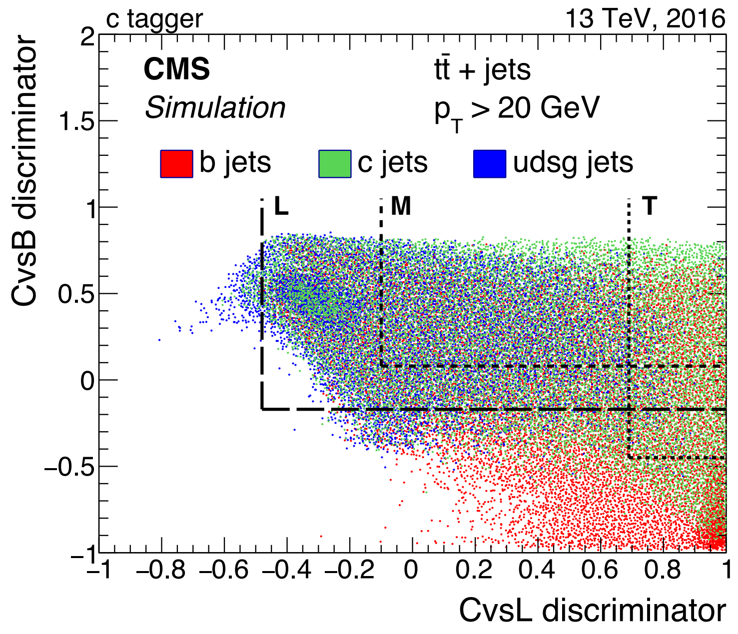
$$\mathrm{P(CvsL)} = \frac{\mathrm{P(c)}}{\mathrm{P(c)} + \mathrm{P(udsg)}}, \qquad \mathrm{P(CvsB)} = \frac{\mathrm{P(c)}}{\mathrm{P(c)} + \mathrm{P(b)} + \mathrm{P(bb)}}.$$

To use these discriminants in a neural network, the 2-dim shape in simulations needs to be calibrated to the data!

Novel shape calibration of the two-dimensional CvsL and CvsB DeepCSV c-tagger discriminators

JINST 13 (2018) P05011

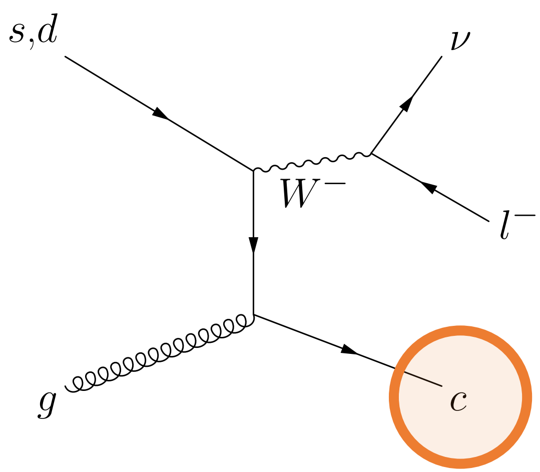




c-tagger calibration

Three control regions for flavor enrichment

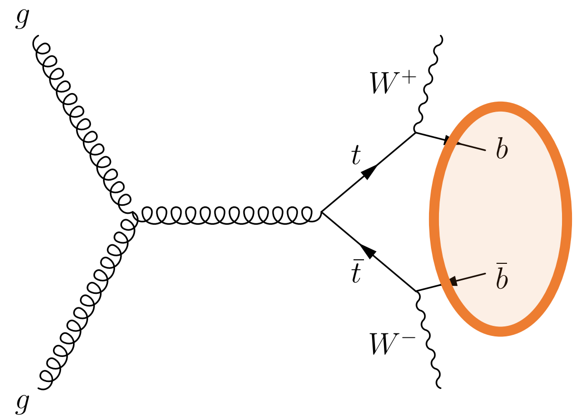
W+charm



c-enriched (93% pure)

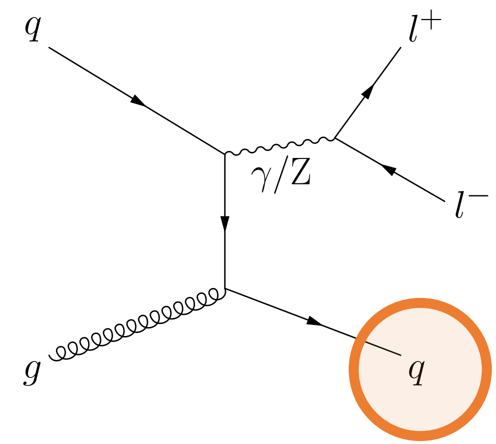
(after OS-SS subtraction)

semi-leptonic $t\bar{t}$



b-enriched (81% pure)

DY + jets



light-enriched (86% pure)

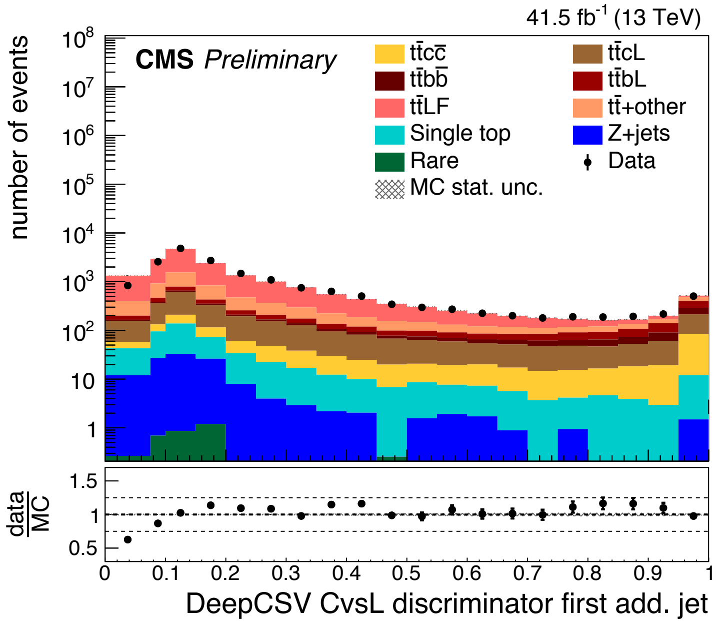
Very good purity in different control regions!

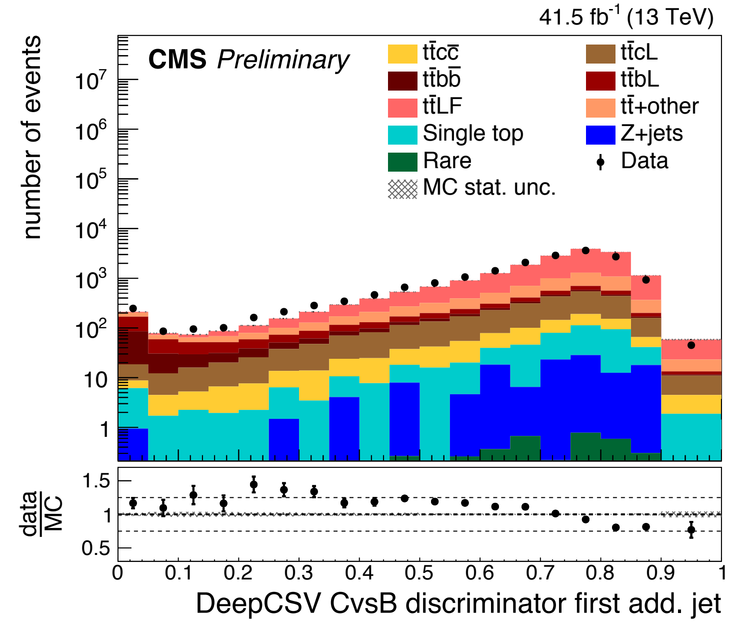
Iterative fitting procedure per (2-dim.) bin, by iterating multiple times over the three control regions → 2-dim SF maps
i.e. SF(CvsL, CvsB, flavour)

c-tagger calibration

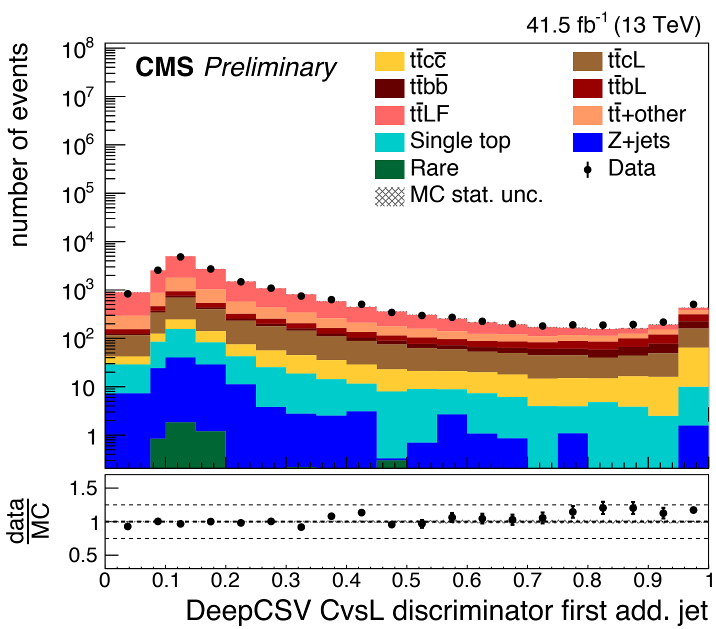
Effect of the calibration on the additional jet CvsL/CvsB

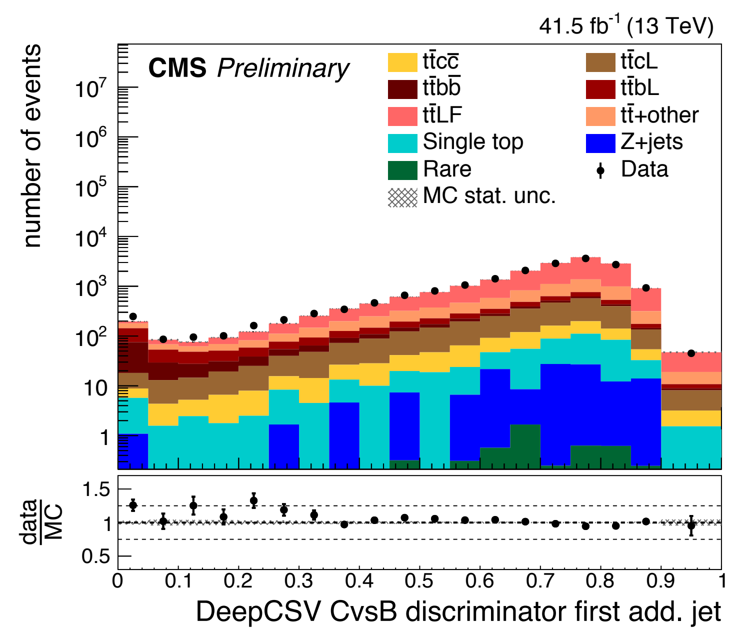
Before





After

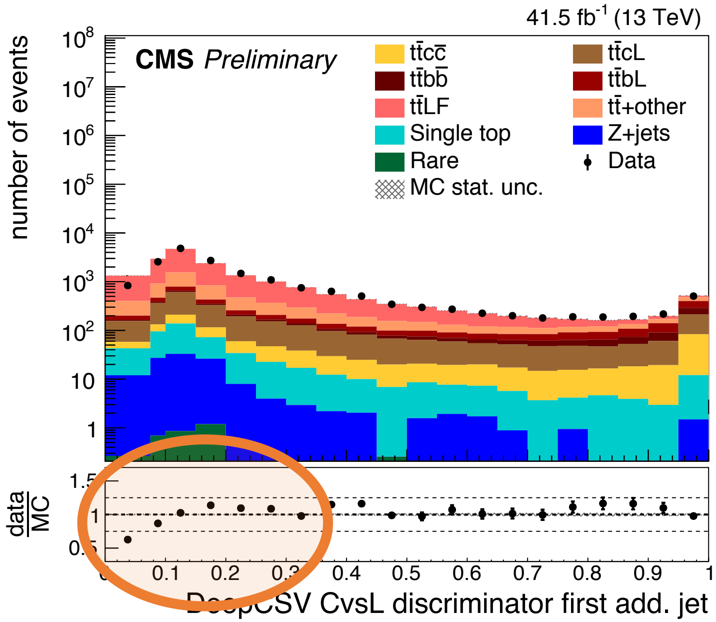


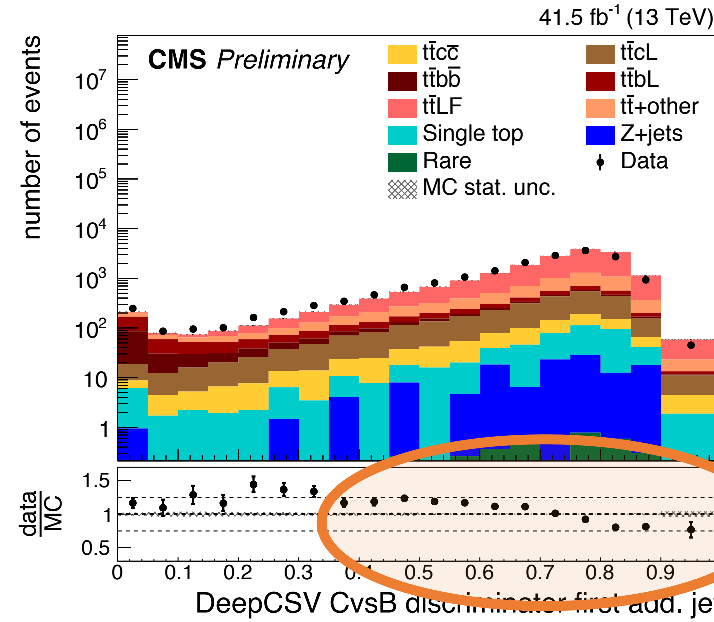


c-tagger calibration

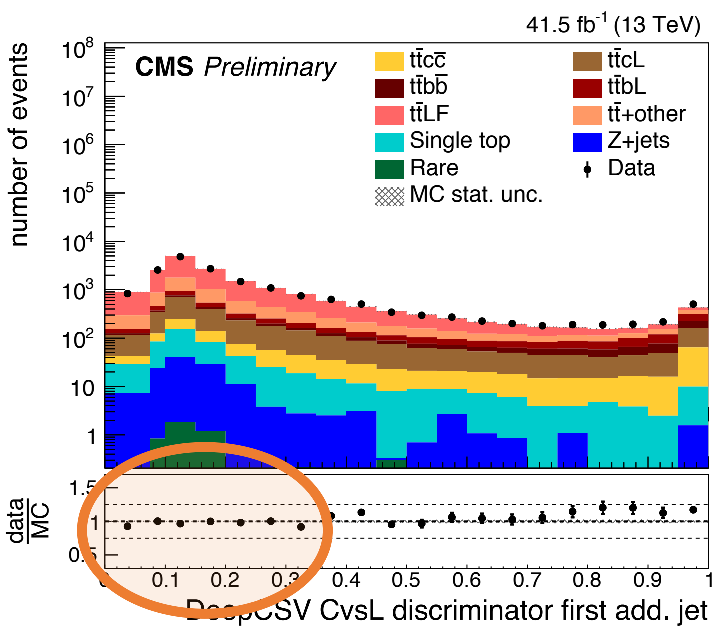
Effect of the calibration on the additional jet CvsL/CvsB

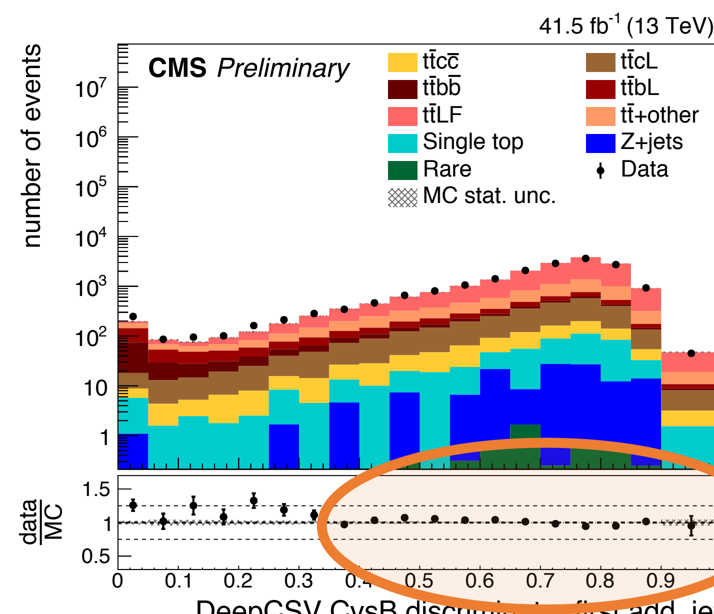
Before





After

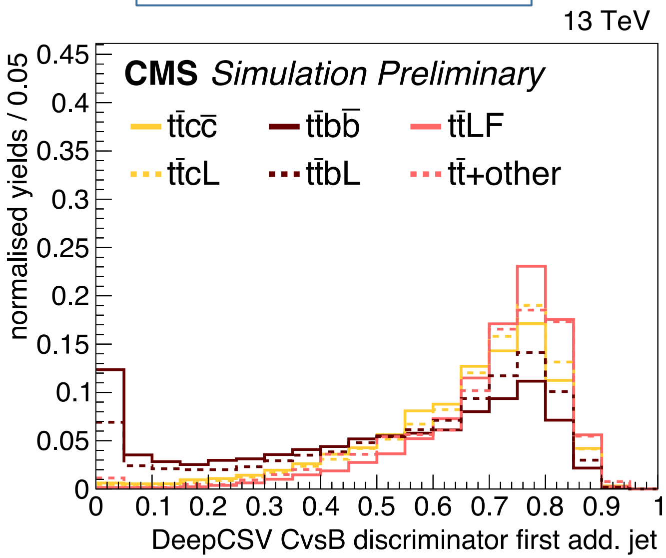




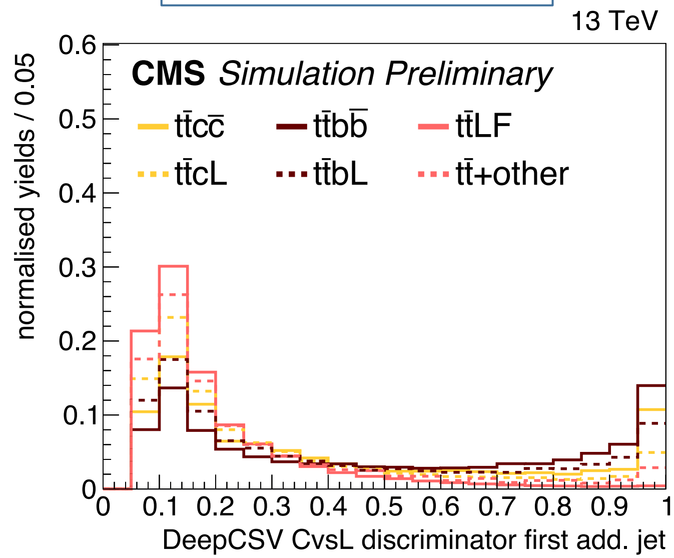
Template fit using NN discriminator

Sensitive observables to distinguish between $t\bar{t}c\bar{c}$, $t\bar{t}b\bar{b}$, $t\bar{t}$LF

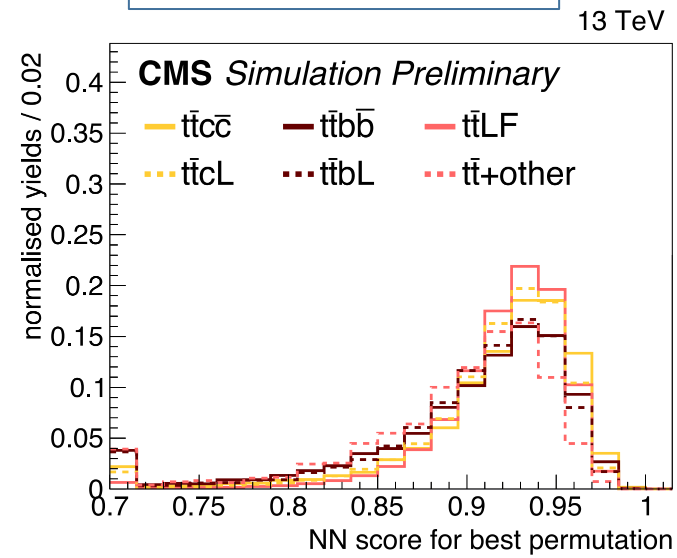
CvsB add. Jet 1



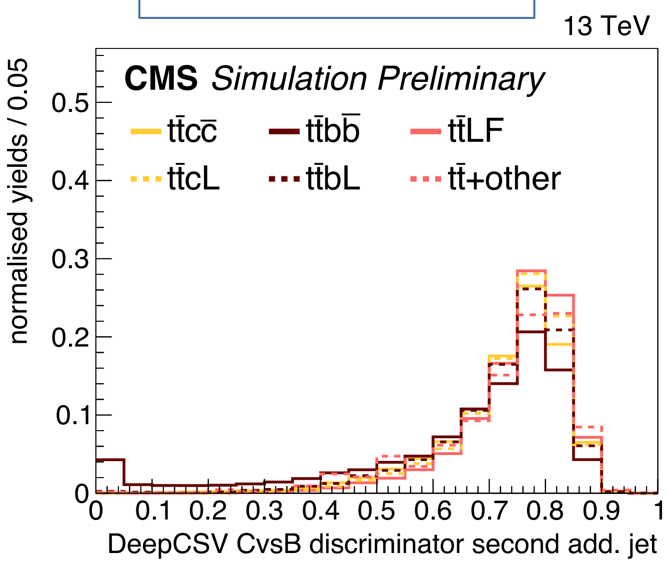
CvsL add. Jet 1



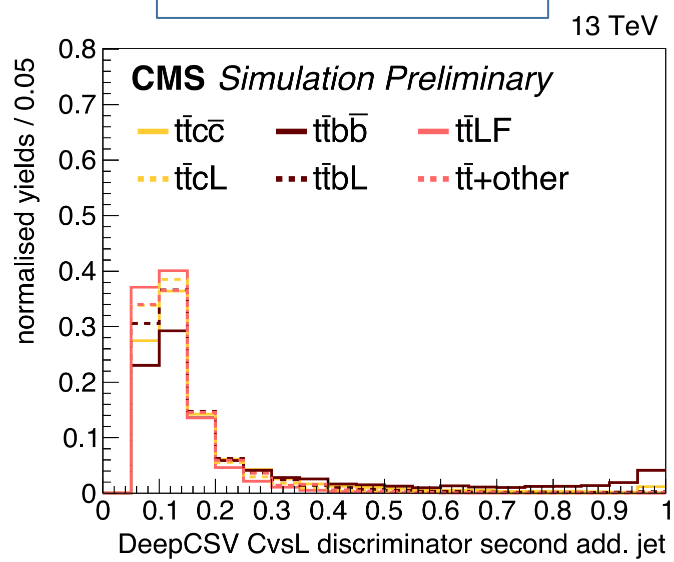
parton match NN



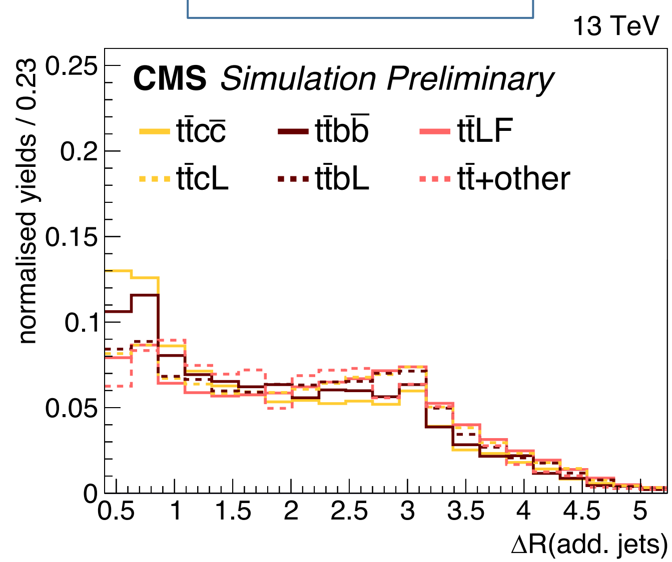
CvsB add. Jet 2



CvsL add. Jet 2

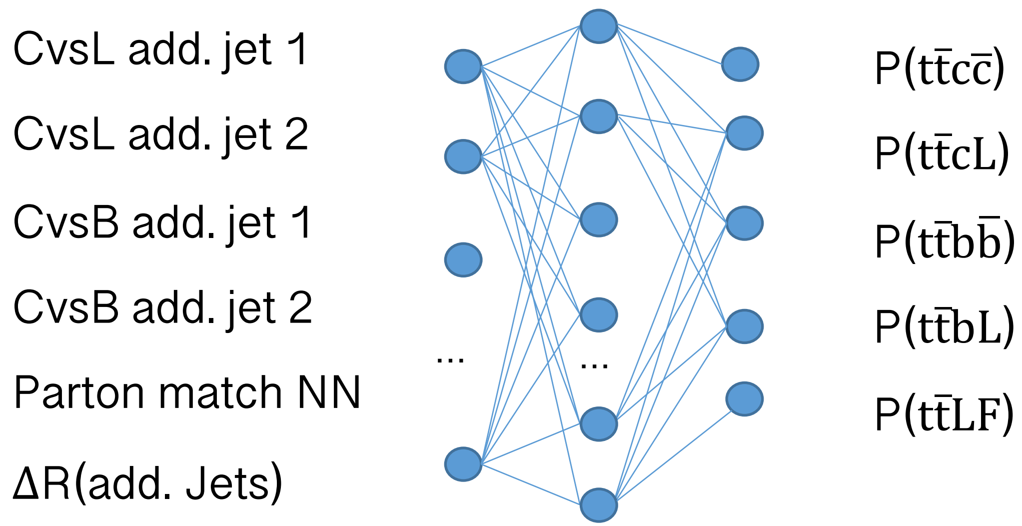


ΔR(add. Jets)



Defining the neural network

one hidden layer that comprises 30 neurons with
ReLu activation functions and a 10% dropout



$$\Delta_b^c = \frac{\mathrm{P}(t\bar{t}c\bar{c})}{\mathrm{P}(t\bar{t}c\bar{c}) + \mathrm{P}(t\bar{t}b\bar{b})}$$

$$\Delta_L^c = \frac{\mathrm{P}(t\bar{t}c\bar{c})}{\mathrm{P}(t\bar{t}c\bar{c}) + \mathrm{P}(t\bar{t}\mathrm{LF})}$$

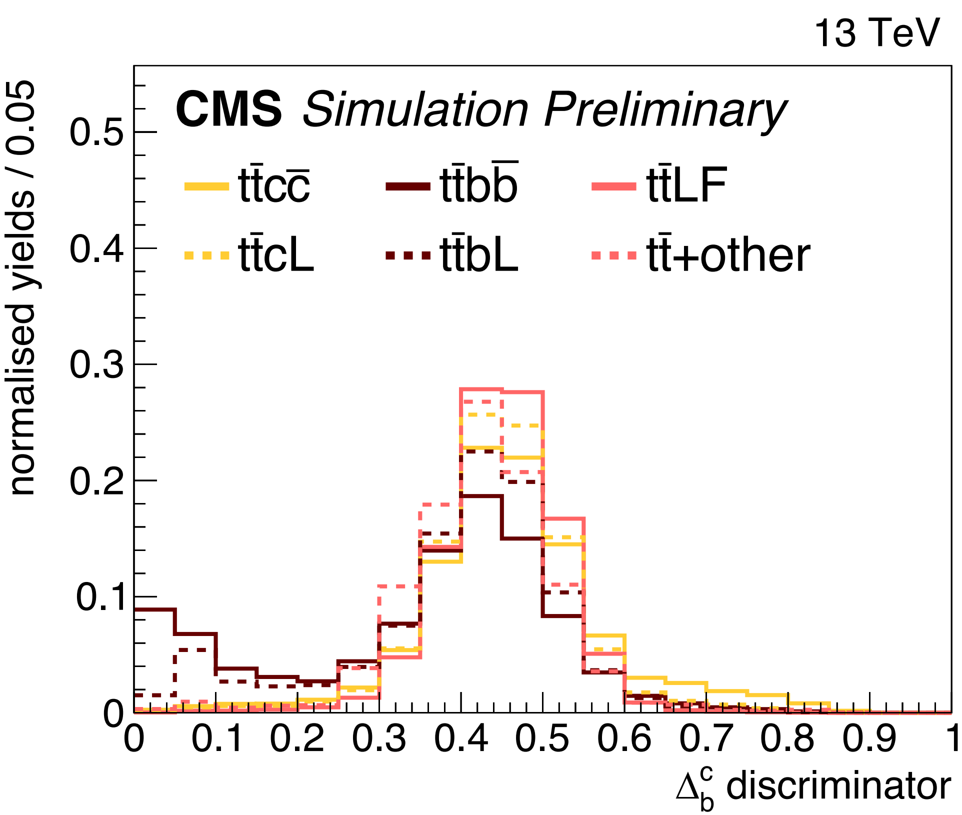
Δ_b^c and Δ_L^c can be interpreted as topology-specific c-tagger discriminants

Information on the flavour of the two additional jets

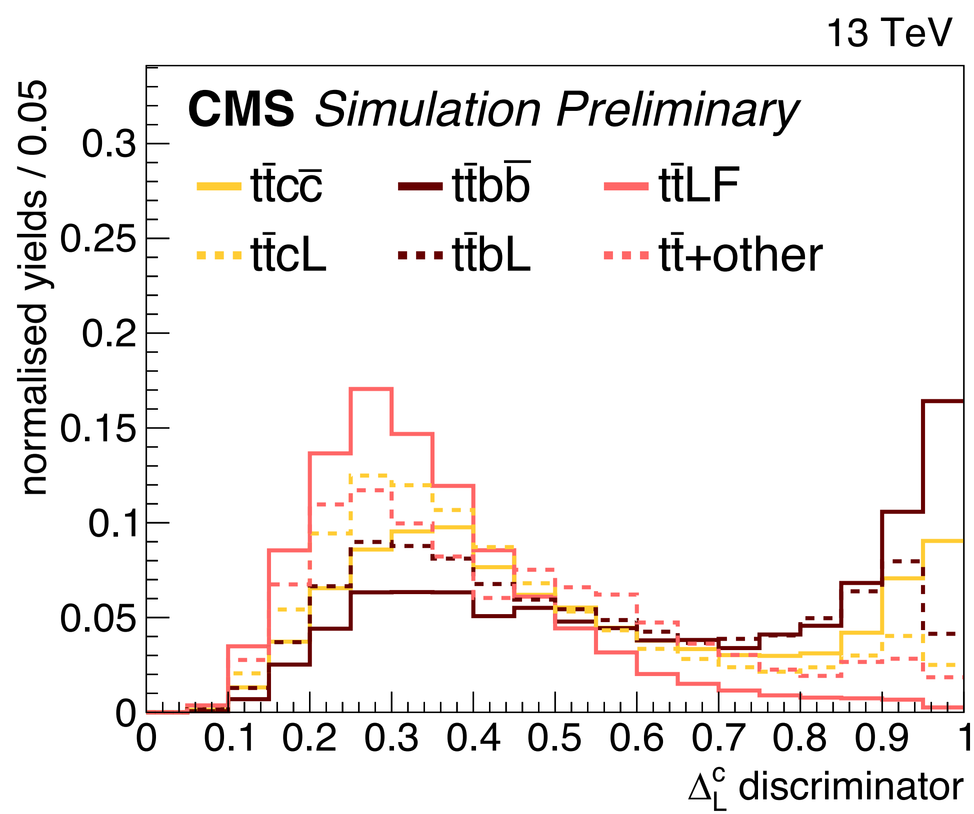
Additional information on the event kinematics to most optimally distinguish different signal categories

Template fit using NN discriminator

Templates from simulated top quark pair events



Constructed to separate $t\bar{t}c\bar{c}$ from $t\bar{t}b\bar{b}$ events



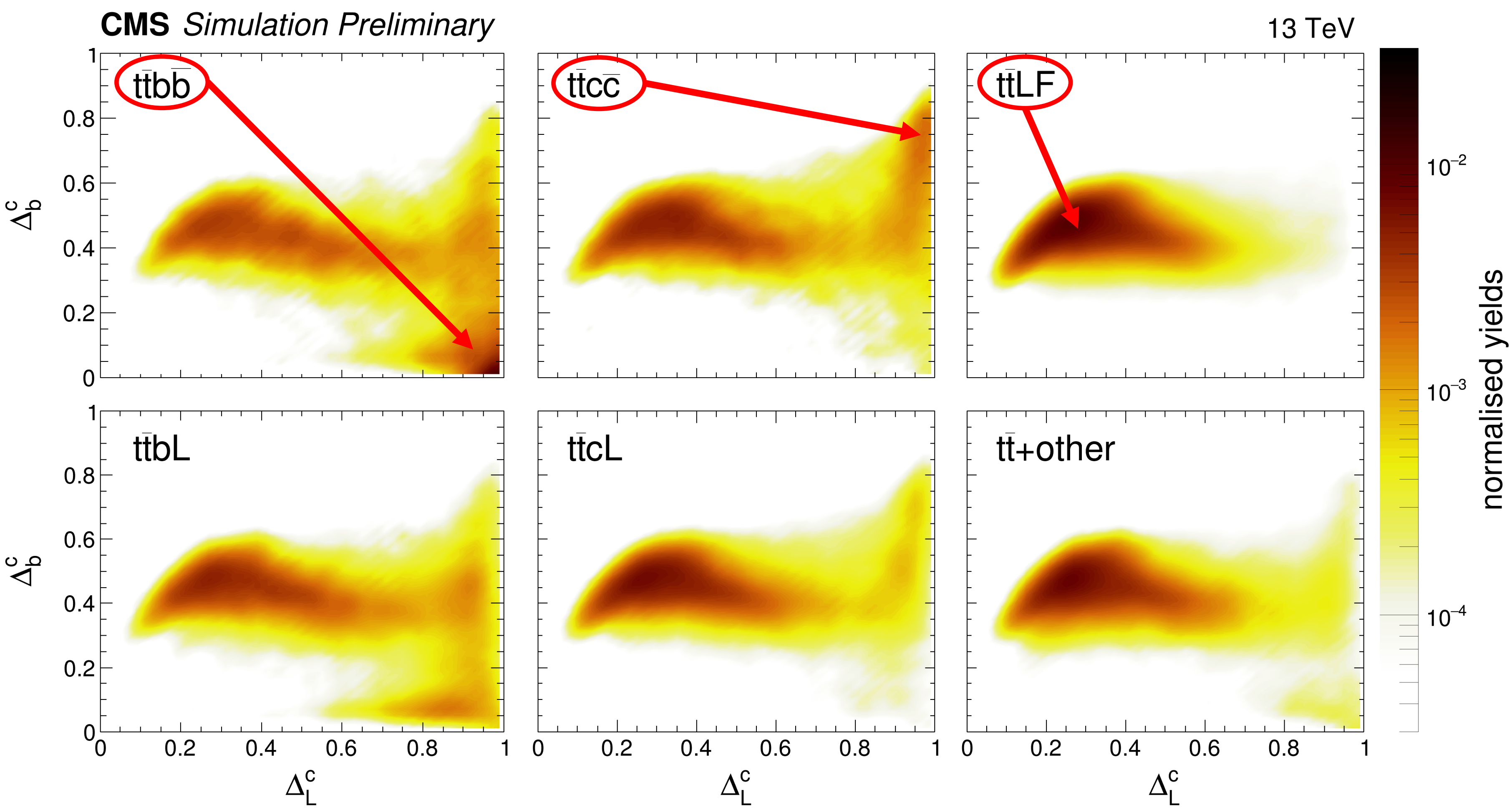
Constructed to separate $t\bar{t}c\bar{c}$ from $t\bar{t}LF$ events

Fitting these templates to the data allows to extract the cross sections for each of the signal processes

Template fit using NN discriminator

Two-dimensional simulated templates used in the fit

The fit is performed on two-dimensional distributions



Clear separation between the $t\bar{t}b\bar{b}$, $t\bar{t}c\bar{c}$ and $t\bar{t}LF$ contributions

Template fit using NN discriminator

Fits to extract inclusive cross sections and their ratios

Full phase space

Fiducial phase space

Reconstructed phase space

Higgs Combine framework

Absolute cross sections

$\sigma_{t\bar{t}c\bar{c}}, \sigma_{t\bar{t}b\bar{b}}, \sigma_{t\bar{t}LF}$

Ratios

$R_b = \sigma_{t\bar{t}b\bar{b}}/\sigma_{t\bar{t}jj}$ and $R_c = \sigma_{t\bar{t}c\bar{c}}/\sigma_{t\bar{t}jj}$

Fit provides results in the fiducial phase space by taking into account an efficiency ϵ.

$t\bar{t}cL$ ($t\bar{t}bL$/ $t\bar{t}$ other) scaled with the same factor as $t\bar{t}c\bar{c}$ ($t\bar{t}b\bar{b}$ / $t\bar{t}LF$), i.e. ratio fixed to MC prediction (with uncertainties)

Event category	$t\bar{t}b\bar{b}$	$t\bar{t}bL$	$t\bar{t}c\bar{c}$	$t\bar{t}cL$	$t\bar{t}LF$
Efficiency ϵ (%)	12.5	8.9	7.1	5.9	4.7
Acceptance $\mathcal{A}$ (%)	2.9	2.5	2.0	2.0	2.3

Results in the fiducial phase space are extrapolated to the full phase space by means of an acceptance A.

Template fit using NN discriminator

Impact of the systematic uncertainties on parameters of interest

numbers in %	fiducial phase space				
	$\Delta\sigma_{t\bar{t}c\bar{c}}$	$\Delta\sigma_{t\bar{t}b\bar{b}}$	$\Delta\sigma_{t\bar{t}LF}$	ΔR_c	ΔR_b
Jet energy scale	7.3	3.3	5.7	3.2	3.4
Jet energy resolution	1.4	0.3	1.2	2.1	1.2
c-tagging calibration	6.7	6.9	2.2	6.9	7.4
Lepton id and isolation	1.3	1.2	1.2	0.2	0.1
Trigger	2.0	2.0	2.0	< 0.1	< 0.1
Pileup	1.2	0.8	0.7	1.6	0.4
Total integrated luminosity	2.4	2.3	2.3	< 0.1	< 0.1
μ_R and μ_F scales in ME	4.3	2.4	0.8	4.1	2.7
Parton shower scale	0.4	1.0	0.1	0.4	0.9
PDF α_s	0.5	< 0.1	0.1	0.4	0.1
Matching ME-PS (hdamp)	6.5	4.9	3.1	2.9	1.4
Underlying event	1.2	1.3	0.7	0.3	0.4
$t\bar{t}bL(cL)/t\bar{t}b\bar{b}(c\bar{c})$ and $t\bar{t}$+other/$t\bar{t}LF$	2.4	1.7	1.2	2.0	1.5
Efficiency (theoretical)	2.0	2.0	2.0	< 0.1	< 0.1
Simulated sample size	4.3	2.7	1.1	4.2	2.7
Background normalisation	0.7	0.1	0.5	0.2	0.5

Dominant experimental uncertainties from c-tagging calibration and JES

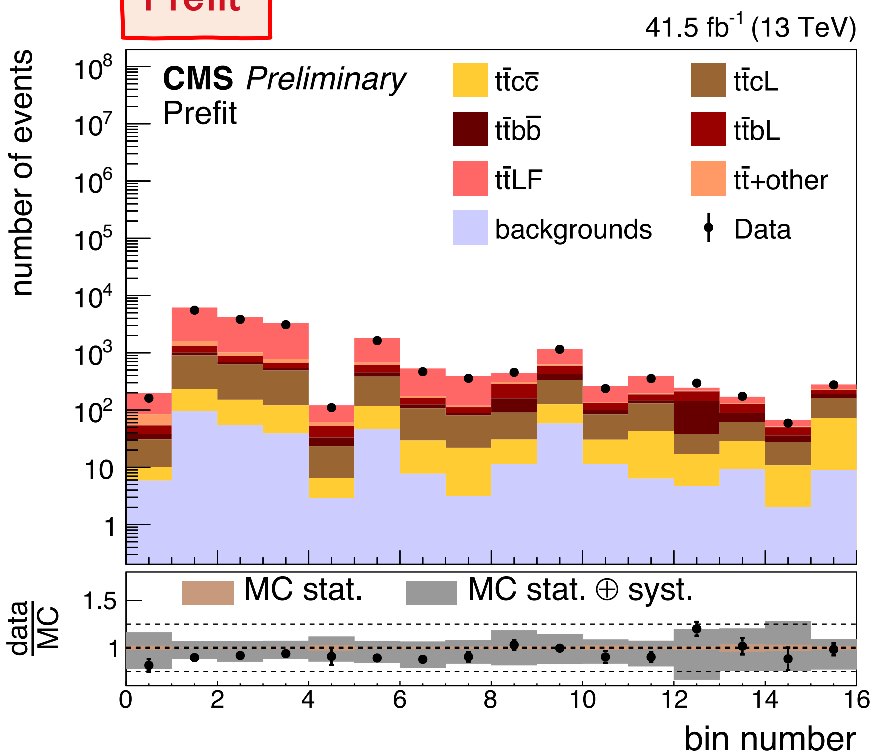
Dominant theoretical uncertainties from QCD scales in the ME and ME-PS matching

Comparison between the prefit and the postfit distributions

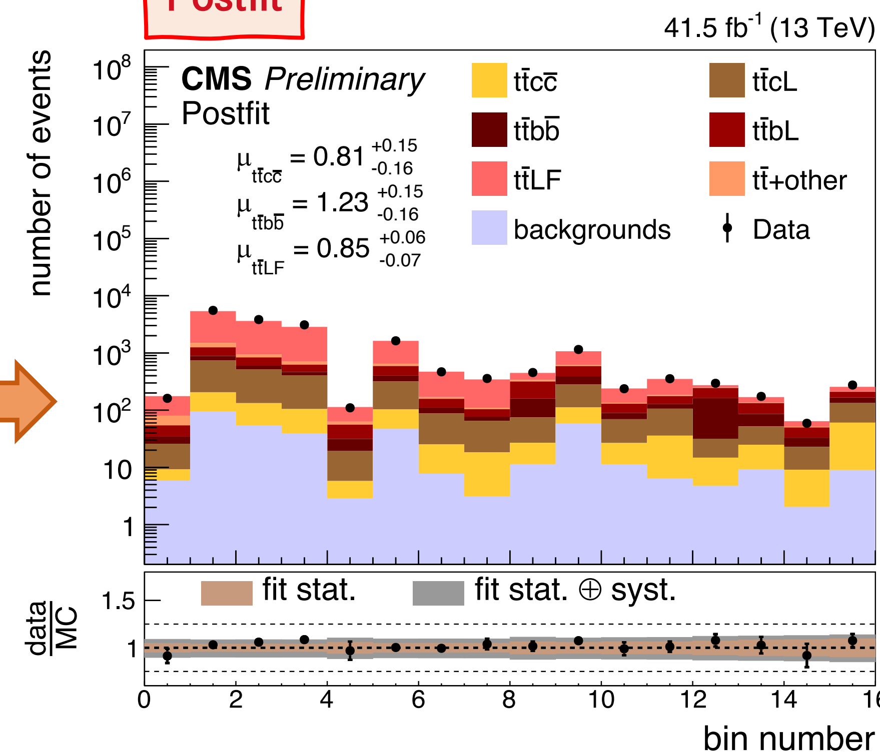
Two-dimensional distributions are unrolled onto a one-dimensional histogram
4x4 binning results in 16 bins with varying flavor composition:

$$\Delta_L^c \otimes \Delta_b^c : [0, 0.45, 0.6, 0.9, 1.0] \otimes [0, 0.3, 0.45, 0.5, 1.0]$$

Prefit



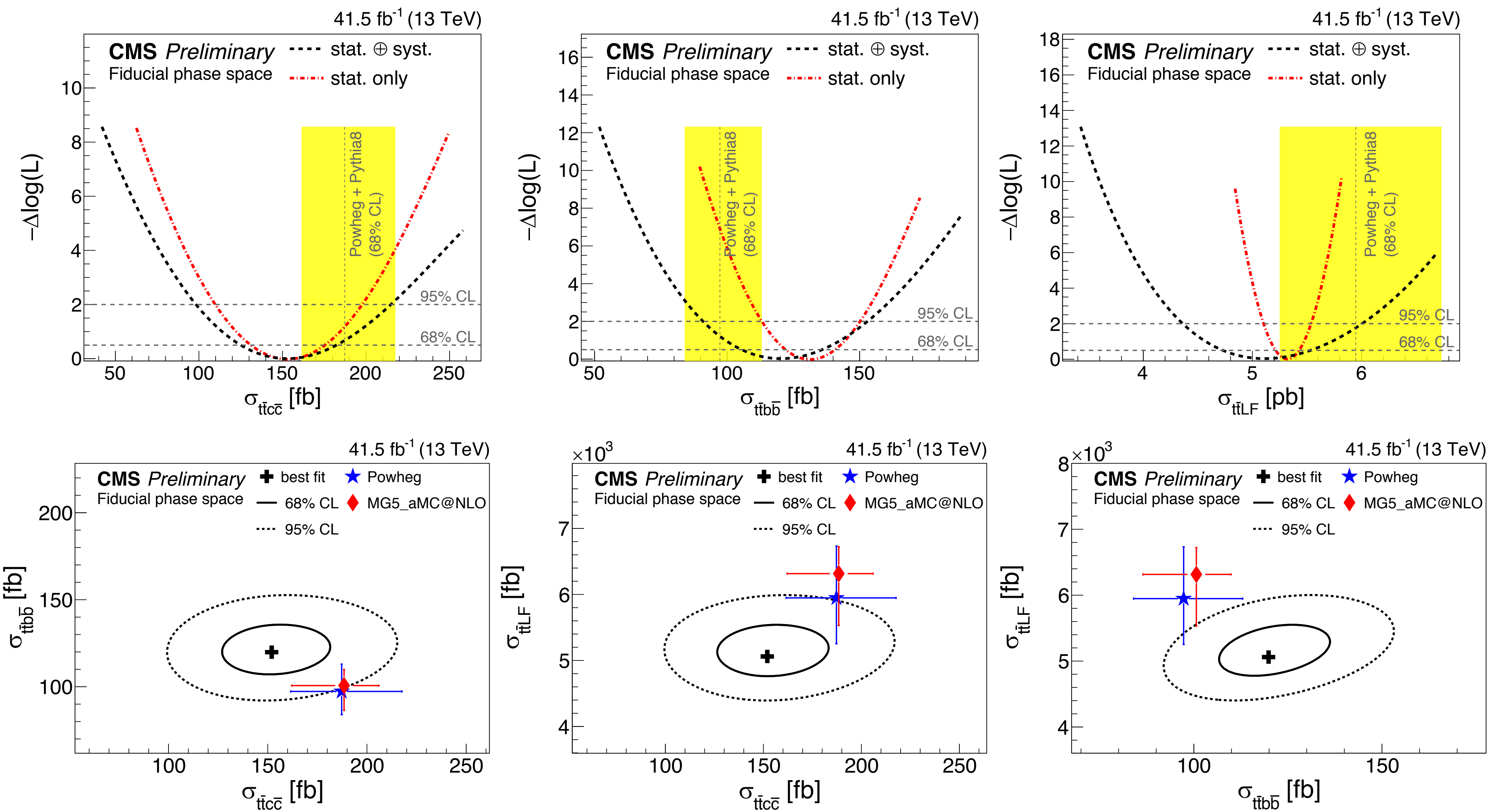
Postfit



μ represent the signal strength, related to the cross section: $\sigma = \dfrac{\mu \times N^{MC}}{\mathcal{L}^{int} \times \epsilon}$

Results

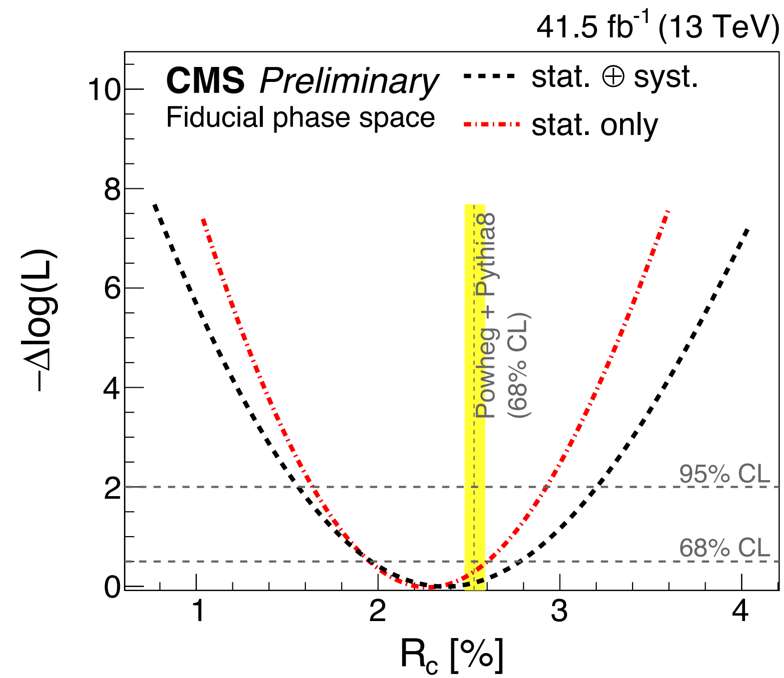
Inclusive cross sections in the fiducial phase space

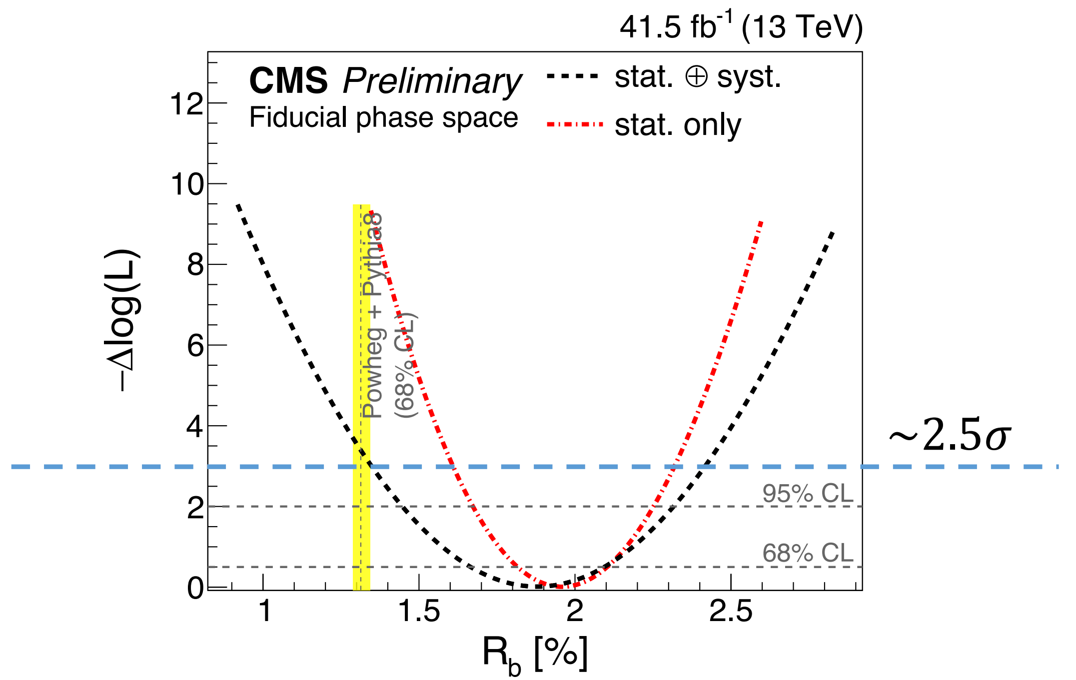


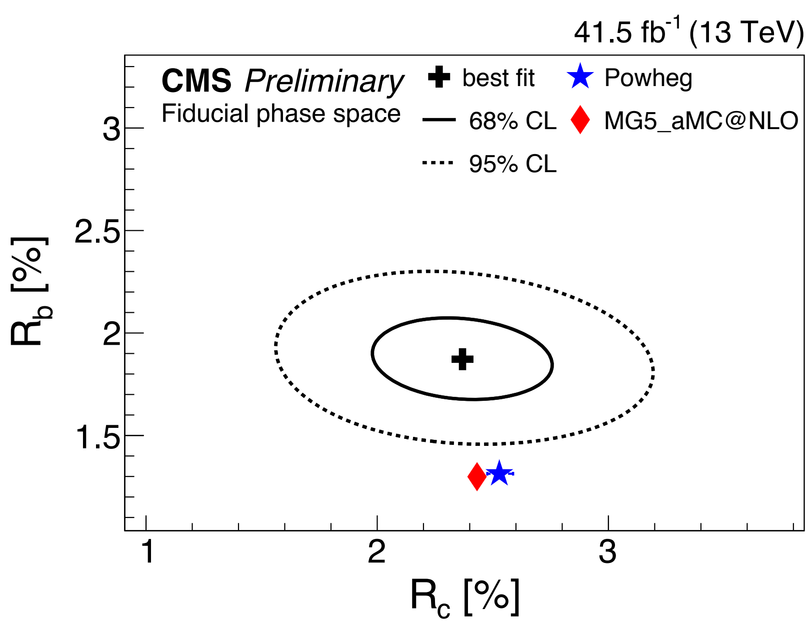
Some tension observed, but overall agreement within 1-2 standard deviations

→measured $t\bar{t}b\bar{b}$ ($t\bar{t}c\bar{c}$ and $t\bar{t}LF$) cross section higher (lower) than predicted.

Ratios R_c and R_b in the fiducial phase space







R_c is in very good agreement with theory prediction.

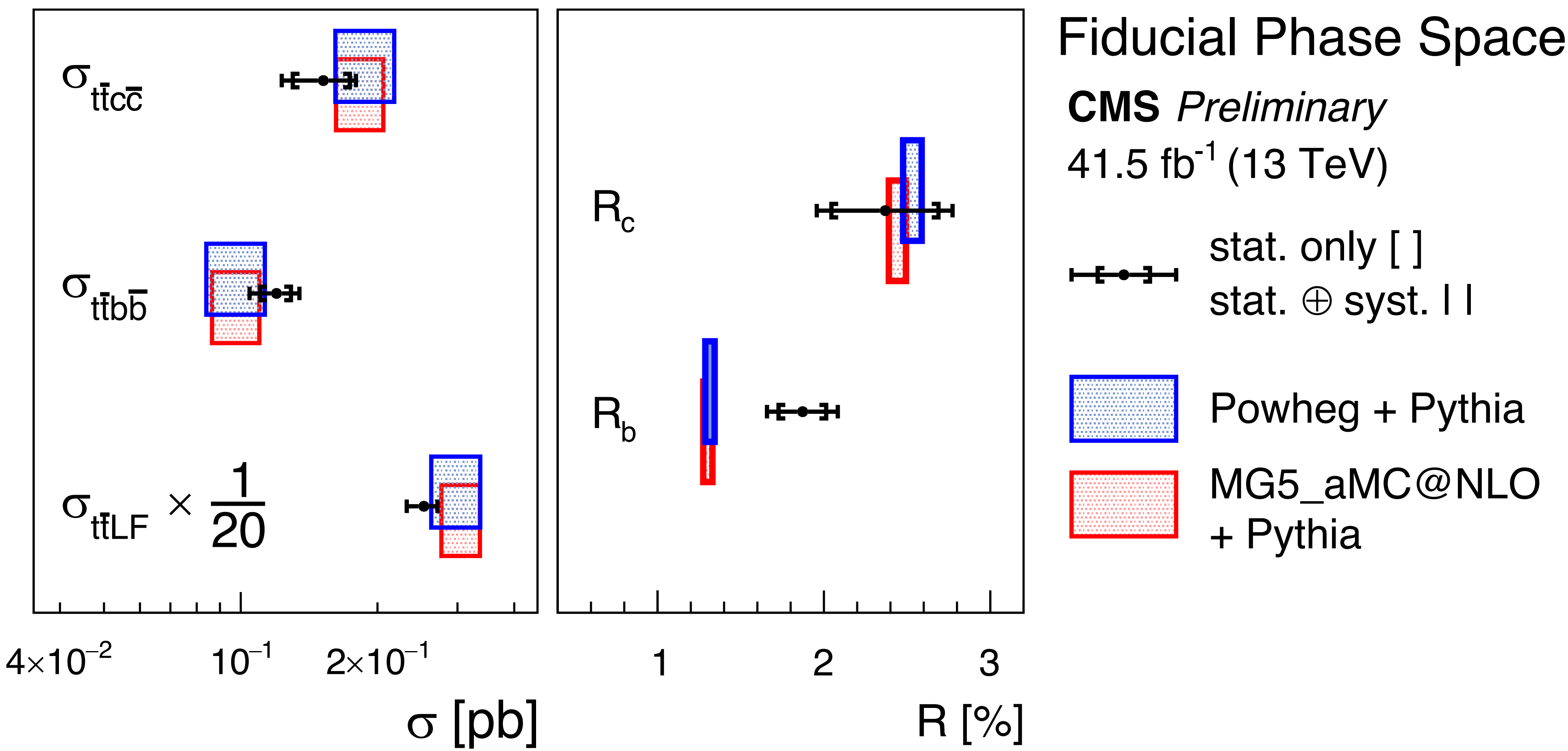
Largest tension observed for R_b

$-\Delta \log L \sim 3 \rightarrow \sim 2.5\sigma$

Numerical values + extrapolation to the full phase space

	Result	Uncertainty	POWHEG	MG5_AMC@NLO	
Fiducial phase space					
$\sigma_{t\bar{t}c\bar{c}}$ [pb]	0.152	$\pm$ 0.022 (stat.) $\pm$ 0.019 (syst.)	0.187 $\pm$ 0.030	0.188 $\pm$ 0.026	~19 %
$\sigma_{t\bar{t}b\bar{b}}$ [pb]	0.120	$\pm$ 0.009 (stat.) $\pm$ 0.012 (syst.)	0.097 $\pm$ 0.016	0.101 $\pm$ 0.014	
$\sigma_{t\bar{t}LF}$ [pb]	5.06	$\pm$ 0.11 (stat.) $\pm$ 0.41 (syst.)	5.95 $\pm$ 0.79	6.32 $\pm$ 0.79	
R_c [%]	2.37	$\pm$ 0.32 (stat.) $\pm$ 0.25 (syst.)	2.53 $\pm$ 0.06	2.43 $\pm$ 0.06	~17 %
R_b [%]	1.87	$\pm$ 0.14 (stat.) $\pm$ 0.16 (syst.)	1.31 $\pm$ 0.03	1.30 $\pm$ 0.03	
Full phase space					
$\sigma_{t\bar{t}c\bar{c}}$ [pb]	7.43	$\pm$ 1.07 (stat.) $\pm$ 0.95 (syst.)	9.15 $\pm$ 1.44	8.92 $\pm$ 1.26	
$\sigma_{t\bar{t}b\bar{b}}$ [pb]	4.12	$\pm$ 0.32 (stat.) $\pm$ 0.42 (syst.)	3.35 $\pm$ 0.54	3.39 $\pm$ 0.49	
$\sigma_{t\bar{t}LF}$ [pb]	217.0	$\pm$ 4.6 (stat.) $\pm$ 18.1 (syst.)	255.1 $\pm$ 32.0	260.6 $\pm$ 32.8	
R_c [%]	2.64	$\pm$ 0.36 (stat.) $\pm$ 0.28 (syst.)	2.82 $\pm$ 0.07	2.72 $\pm$ 0.05	
R_b [%]	1.47	$\pm$ 0.11 (stat.) $\pm$ 0.13 (syst.)	1.03 $\pm$ 0.03	1.03 $\pm$ 0.02	

Summary in the fiducial phase space (visual)



First measurement of the $t\bar{t} + c\bar{c}$ cross section!

Fiducial PS: $\sigma(t\bar{t} + c\bar{c})$ = 152 ± 22 (stat.) ± 19 (syst.) fb (~ 19% uncertainty)
Rc = 2.37 ± 0.32 (stat.) ± 0.25 (syst.) % (~ 17% uncertainty)

Full PS: $\sigma(t\bar{t} + c\bar{c})$ = 7.43 ± 1.07 (stat.) ± 0.95 (syst.) pb
Rc = 2.64 ± 0.36 (stat.) ± 0.29 (syst.) %

Simultaneous extraction $\sigma_{t\bar{t}c\bar{c}}, \sigma_{t\bar{t}b\bar{b}}, \sigma_{t\bar{t}LF}, R_b = {}^{\sigma_{t\bar{t}b\bar{b}}}/_{\sigma_{t\bar{t}jj}}$ and $R_c = {}^{\sigma_{t\bar{t}c\bar{c}}}/_{\sigma_{t\bar{t}jj}}$
→ Fully coherent treatment of different jet flavours in $t\bar{t}$ + 2 jets!

Some tension observed, but cross sections $\sigma(t\bar{t}b\bar{b})$, $\sigma(t\bar{t}c\bar{c})$ and $\sigma(t\bar{t}LF)$ are consistent with Powheg predictions within ~ 1 - 2 σ.

Ratio R_c is consistent with predictions, whereas R_b is found to be higher than predicted, corresponding to ~ 2.5σ.

Higher observed $\sigma_{t\bar{t}b\bar{b}}$ (or R_b) is consistent with previous $t\bar{t}b\bar{b}$ analyses.

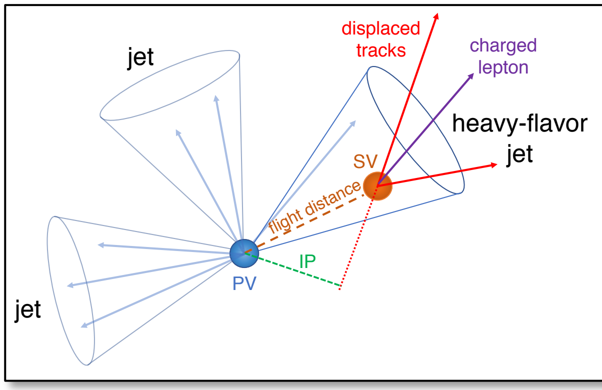
For the first time, we also see that $t\bar{t}c\bar{c}$ and $t\bar{t}LF$ are slightly overestimated in simulations (but within uncertainties)

Backup

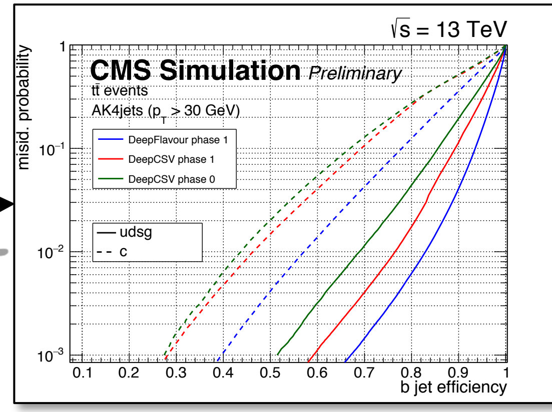
Measurement of $t\bar{t}+c\bar{c}$ production

A roadmap towards a successful measurement

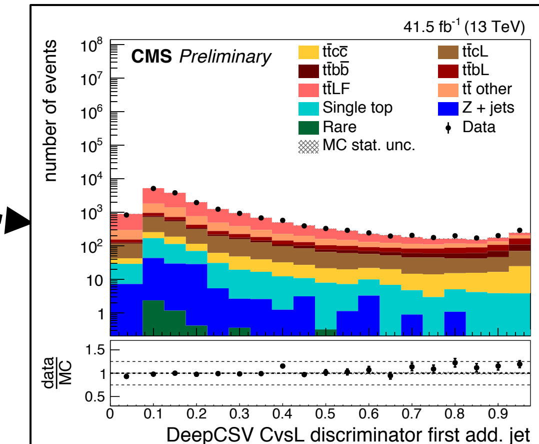
① c, b and t quarks require jets and heavy flavor tagging (new **c-tagger**)



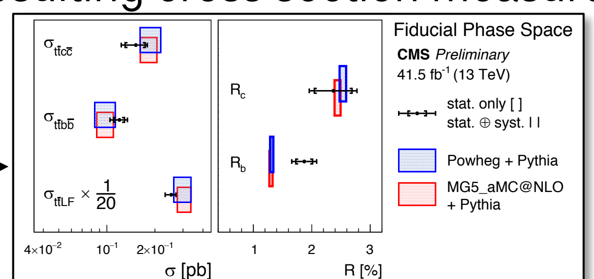
② Improved ML techniques for HF tagging (**DeepCSV**/DeepJet)

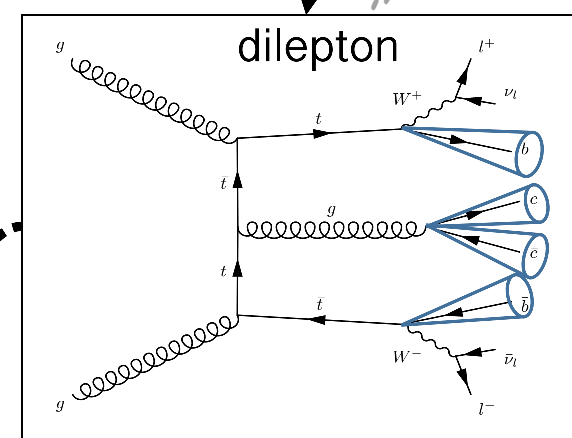


③ Calibration performance (charm tagger shape)

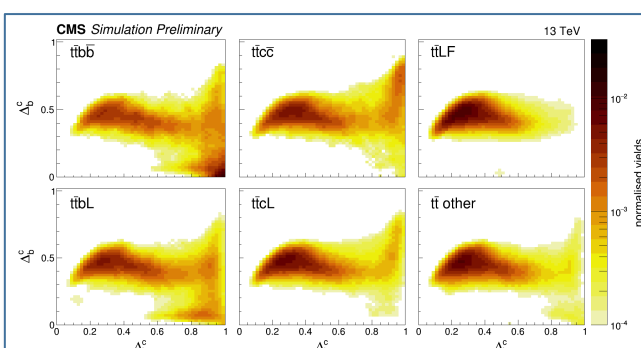


⑥ Resulting cross section measurement





④ Selection and reconstruction of the $t\bar{t}$+HF topology (jet-parton match)



⑤ Differentiating $t\bar{t}$+HF categories (ML classifier)

Motivation for measuring $t\bar{t}$+HF

Theoretical modelling

Theory predictions / Simulation of the $t\bar{t}$+HF final state is highly non-trivial. It deals with very different scales from the top quark mass down to momenta of the relatively soft additional jets

1. Matrix Element vs Parton Shower

2. LO vs NLO (large k-factor, depending on scale choice)

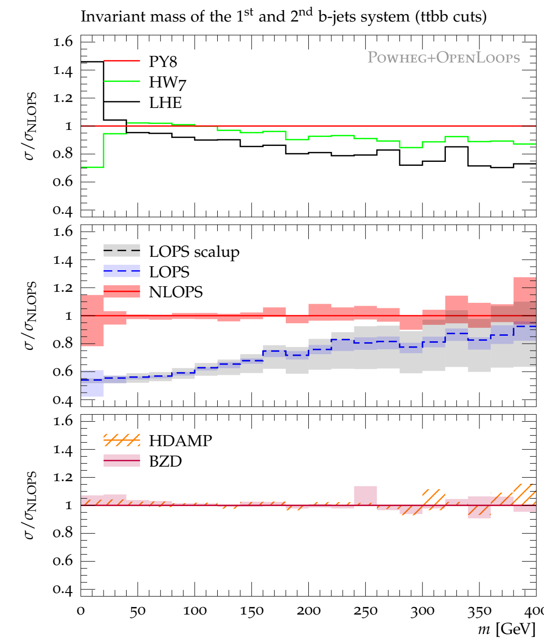
3. Factorization/Renormalization/Shower scales

4. Inclusive $t\bar{t}$+jets versus separate $t\bar{t}b\bar{b}$ and $t\bar{t}c\bar{c}$ sim.

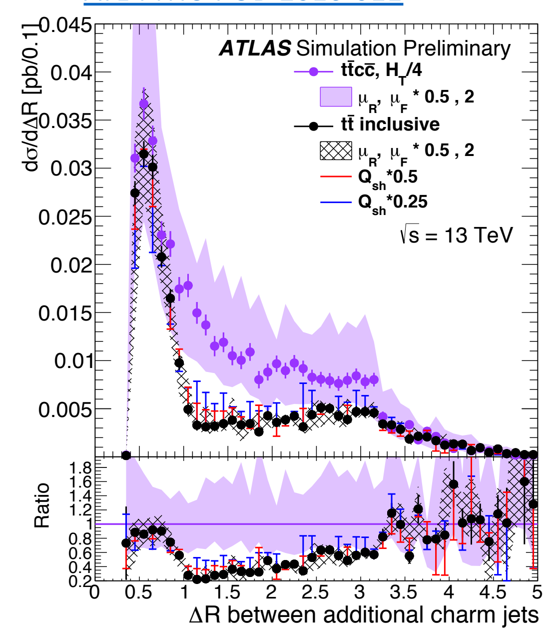
Motivated CMS and ATLAS to measure $t\bar{t}+b\bar{b}$ [arXiv:1411.5621, 1705.10141, 2003.06467, 1909.05306, 1304.6386, 1508.06868, 1811.12113]

$t\bar{t}c\bar{c}$ has not been measured experimentally!

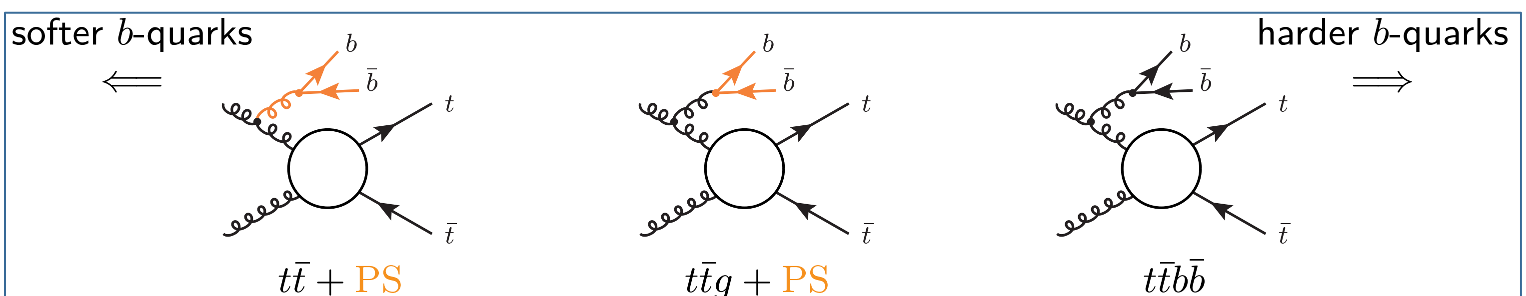
T. Ježo et al., *Eur. Phys. J.* **C78** (2018), no. 6 502, [arXiv:1802.00426]



ATL-PHYS-PUB-2016-011



S. Pozzorini, Theory progress on $t\bar{t}H(b\bar{b})$ background, TOP2018 @ Bad Neuenahr, Germany

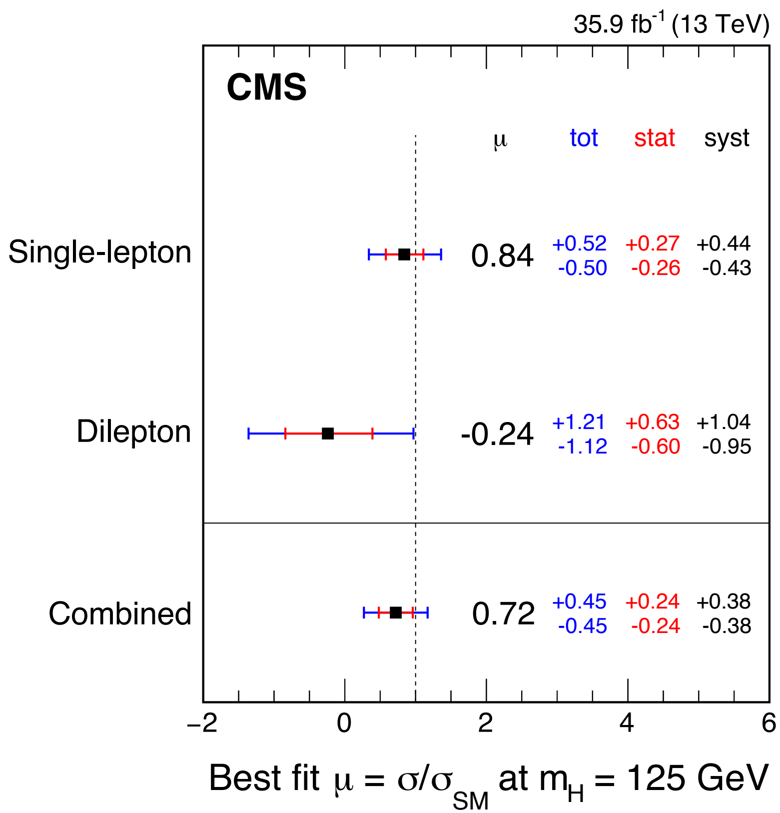


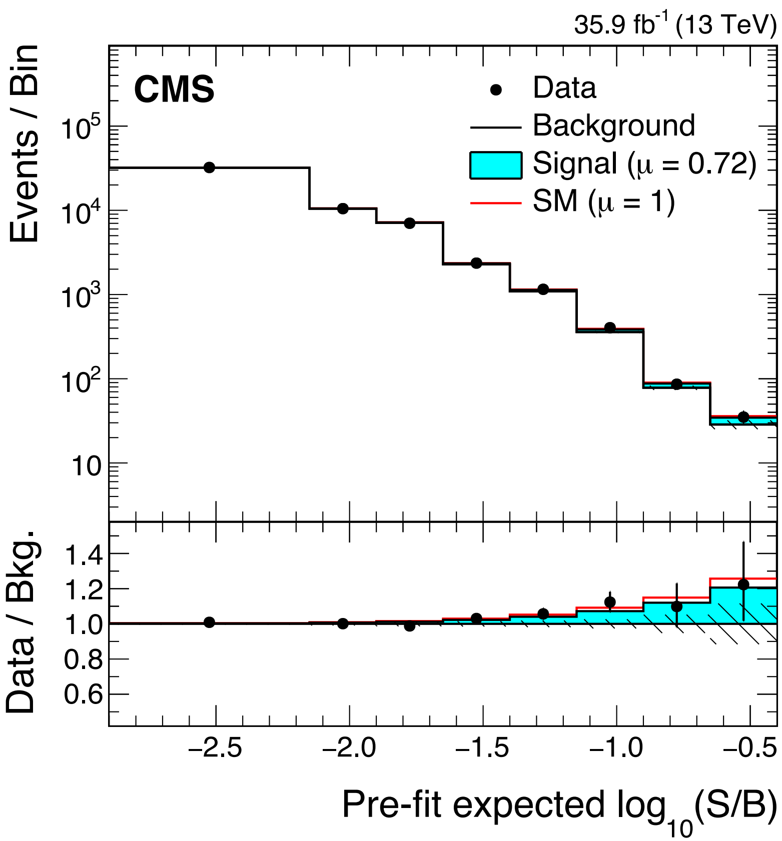
Motivation for measuring $t\bar{t}$+HF

Interplay between Higgs boson and top/bottom quarks

measurement of $t\bar{t}H(H \to b\bar{b})$ [arXiv:1804.03682](https://arxiv.org/abs/1804.03682)

observed (expected) significance of 1.6 (2.2)





$t\bar{t}H$ $(H \to b\bar{b})$ suffers from an irreducible background of (gluon-induced) $t\bar{t}b\bar{b}$ and $t\bar{t}c\bar{c}$ (through mistags) events!

Results
Comparison to other ttbb analyses

PAS-TOP-20-003

	Result	Uncertainty	POWHEG	MG5_AMC@NLO
Fiducial phase space				
$\sigma_{t\bar{t}c\bar{c}}$ [pb]	0.152	$\pm$ 0.022 (stat.) $\pm$ 0.019 (syst.)	0.187 $\pm$ 0.030	0.188 $\pm$ 0.026
$\sigma_{t\bar{t}b\bar{b}}$ [pb]	0.120	$\pm$ 0.009 (stat.) $\pm$ 0.012 (syst.)	0.097 $\pm$ 0.016	0.101 $\pm$ 0.014
$\sigma_{t\bar{t}LF}$ [pb]	5.06	$\pm$ 0.11 (stat.) $\pm$ 0.41 (syst.)	5.95 $\pm$ 0.79	6.32 $\pm$ 0.79
R_c [%]	2.37	$\pm$ 0.32 (stat.) $\pm$ 0.25 (syst.)	2.53 $\pm$ 0.06	2.43 $\pm$ 0.06
R_b [%]	1.87	$\pm$ 0.14 (stat.) $\pm$ 0.16 (syst.)	1.31 $\pm$ 0.03	1.30 $\pm$ 0.03
Full phase space				
$\sigma_{t\bar{t}c\bar{c}}$ [pb]	7.43	$\pm$ 1.07 (stat.) $\pm$ 0.95 (syst.)	9.15 $\pm$ 1.44	8.92 $\pm$ 1.26
$\sigma_{t\bar{t}b\bar{b}}$ [pb]	4.12	$\pm$ 0.32 (stat.) $\pm$ 0.42 (syst.)	3.35 $\pm$ 0.54	3.39 $\pm$ 0.49
$\sigma_{t\bar{t}LF}$ [pb]	217.0	$\pm$ 4.6 (stat.) $\pm$ 18.1 (syst.)	255.1 $\pm$ 32.0	260.6 $\pm$ 32.8
R_c [%]	2.64	$\pm$ 0.36 (stat.) $\pm$ 0.28 (syst.)	2.82 $\pm$ 0.07	2.72 $\pm$ 0.05
R_b [%]	1.47	$\pm$ 0.11 (stat.) $\pm$ 0.13 (syst.)	1.03 $\pm$ 0.03	1.03 $\pm$ 0.02

+2.5σ

TOP-18-002

	$R_{t\bar{t}b\bar{b}/t\bar{t}jj}$	$\sigma_{t\bar{t}jj}$ [pb]	$\sigma_{t\bar{t}b\bar{b}}$ [pb]
Dilepton channel (VPS)			
POWHEG + PYTHIA8	0.013 ± 0.002	2.41 ± 0.21	0.032 ± 0.004
Measurement	$0.017 \pm 0.001 \pm 0.001$	$2.36 \pm 0.02 \pm 0.20$	$0.040 \pm 0.002 \pm 0.005$
Dilepton channel (FPS)			
POWHEG + PYTHIA8	0.014 ± 0.003	163 ± 21	2.3 ± 0.4
MG_aMC@NLO + PYTHIA8 5FS [FxFx]	0.015 ± 0.003	159 ± 25	2.4 ± 0.4
POWHEG + HERWIG++	0.011 ± 0.002	170 ± 25	1.9 ± 0.3
Measurement	$0.018 \pm 0.001 \pm 0.002$	$159 \pm 1 \pm 15$	$2.9 \pm 0.1 \pm 0.5$
Lepton+jets channel (VPS)			
POWHEG + PYTHIA8	0.017 ± 0.002	30.5 ± 3.0	0.52 ± 0.06
Measurement	$0.020 \pm 0.001 \pm 0.001$	$31.0 \pm 0.2 \pm 2.9$	$0.62 \pm 0.03 \pm 0.07$
Lepton+jets channel (FPS)			
POWHEG + PYTHIA8	0.013 ± 0.002	290 ± 29	3.9 ± 0.4
MG_aMC@NLO + PYTHIA8 5FS [FxFx]	0.014 ± 0.003	280 ± 40	4.1 ± 0.4
POWHEG + HERWIG++	0.011 ± 0.002	321 ± 36	3.4 ± 0.5
Measurement	$0.016 \pm 0.001 \pm 0.001$	$292 \pm 1 \pm 29$	$4.7 \pm 0.2 \pm 0.6$

+1.8σ

30 GeV

+2.1σ

TOP-18-011

	Fiducial, parton-independent (pb)	Fiducial, parton-based (pb)	Total (pb)
Measurement	$1.6 \pm 0.1^{+0.5}_{-0.4}$	$1.6 \pm 0.1^{+0.5}_{-0.4}$	$5.5 \pm 0.3^{+1.6}_{-1.3}$
POWHEG ($t\bar{t}$)	1.1 ± 0.2	1.0 ± 0.2	3.5 ± 0.6
POWHEG ($t\bar{t}$) + HERWIG++	0.8 ± 0.2	0.8 ± 0.2	3.0 ± 0.5
MADGRAPH5_aMC@NLO (4FS $t\bar{t}b\bar{b}$)	0.8 ± 0.2	0.8 ± 0.2	2.3 ± 0.7
MADGRAPH5_aMC@NLO (5FS $t\bar{t}$+jets, FxFx)	1.0 ± 0.1	1.0 ± 0.1	3.6 ± 0.3

TOP-16-010

Phase space		$\sigma_{t\bar{t}b\bar{b}}$ [pb]	$\sigma_{t\bar{t}jj}$ [pb]	$\sigma_{t\bar{t}b\bar{b}}/\sigma_{t\bar{t}jj}$
Visible	Measurement	0.088 $\pm$ 0.012 $\pm$ 0.029	3.7 $\pm$ 0.1 $\pm$ 0.7	0.024 $\pm$ 0.003 $\pm$ 0.007
	SM (POWHEG)	0.070 $\pm$ 0.009	5.1 $\pm$ 0.5	0.014 $\pm$ 0.001
Full	Measurement	4.0 $\pm$ 0.6 $\pm$ 1.3	184 $\pm$ 6 $\pm$ 33	0.022 $\pm$ 0.003 $\pm$ 0.006
	SM (POWHEG)	3.2 $\pm$ 0.4	257 $\pm$ 26	0.012 $\pm$ 0.001

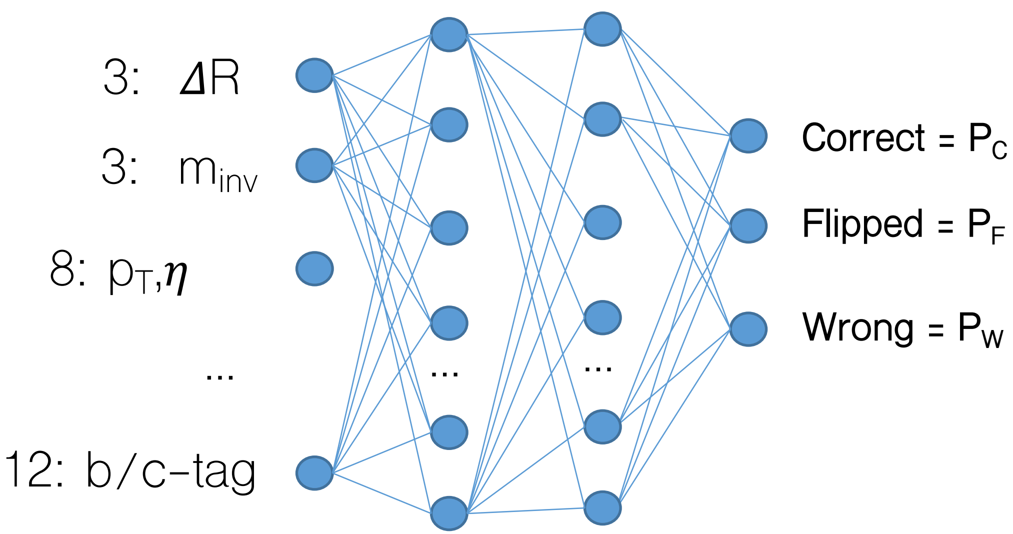
+1.5σ

Jet-parton matching

Performance and neural network output

- Neural network trained with Keras (TensorFlow backend)
- 26 inputs (Standard Normalization, μ=0, σ= 1) → see backup
- 2 hidden layers with 50 neurons each and ReLu activation
- 10% Dropout in each hidden layer (regularization)
- 3 outputs with SoftMax activation
  - Correctly matched
  - Flipped matching
  - Wrong matching
- Loss function = categorical cross-entropy
- Optimizer = Stochastic Gradient Decent
  *learning rate (init) = 0.001, decay = init / (5*n_epoch), nesterov momentum = 0.8*
- n_epoch = 100, batch_size = 128
- Weights added after 30 epochs (ttbb/ttcc = 20, ttbL = 10, ttcL = 5, ttLL = 1)

jet p_T	jet η	b–tag	CvsL c–tag	CvsB c–tag	m_{inv}	ΔR
$p_T(b_t)$	$\eta(b_t)$	BvsAll(b_t)	CvsL(b_t)	CvsB(b_t)	$m_{inv}(b_t,\ell^+)$	ΔR(b_t,ℓ^+)
$p_T(b_{\bar{t}})$	$\eta(b_{\bar{t}})$	BvsAll$(b_{\bar{t}})$	CvsL$(b_{\bar{t}})$	CvsB$(b_{\bar{t}})$	$m_{inv}(b_{\bar{t}},\ell^-)$	ΔR$(b_{\bar{t}},\ell^-)$
$p_T(j_1)$	$\eta(j_1)$	BvsAll(j_1)	CvsL(j_1)	CvsB(j_1)	$m_{inv}(j_1,j_2)$	ΔR(j_1,j_2)
$p_T(j_2)$	$\eta(j_2)$	BvsAll(j_2)	CvsL(j_2)	CvsB(j_2)		



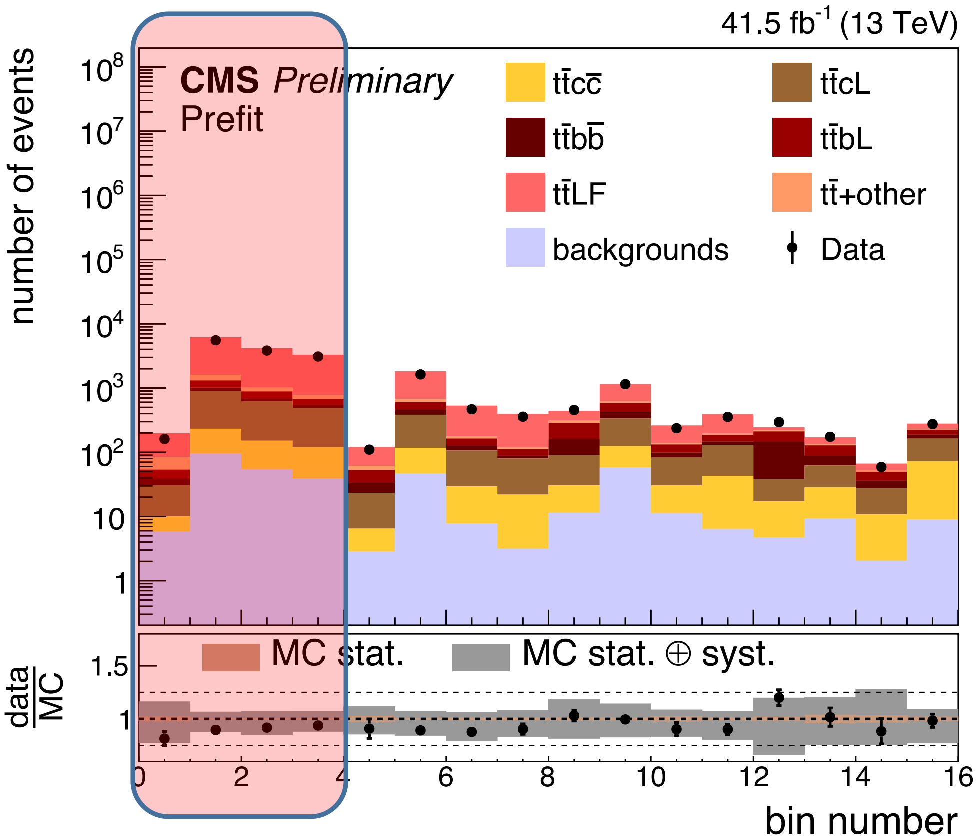
NN matching output =

$$= \max\left(\frac{P_C}{P_C+P_W}\ ,\ \frac{P_F}{P_F+P_W}\right)$$

Unrolling 2D histogram into 1D

4x4 binning results in 16 bins with varying flavor composition

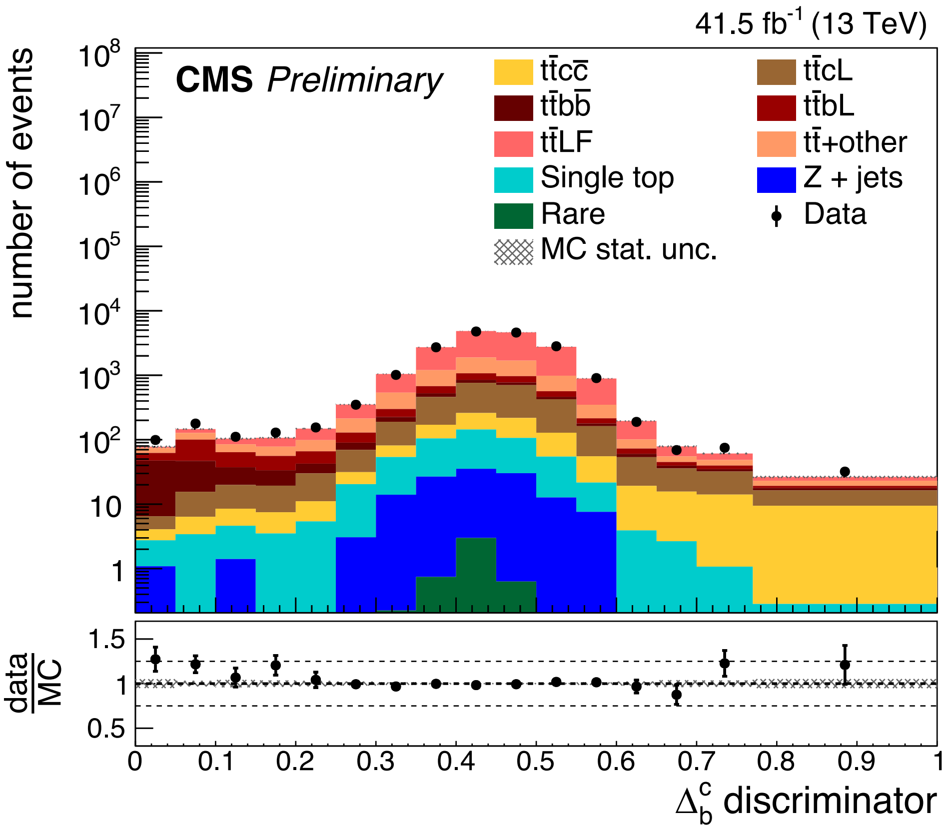
$$\Delta_L^c \otimes \Delta_b^c : [0, 0.45, 0.6, 0.9, 1.0] \otimes [0, 0.3, 0.45, 0.5, 1.0]$$

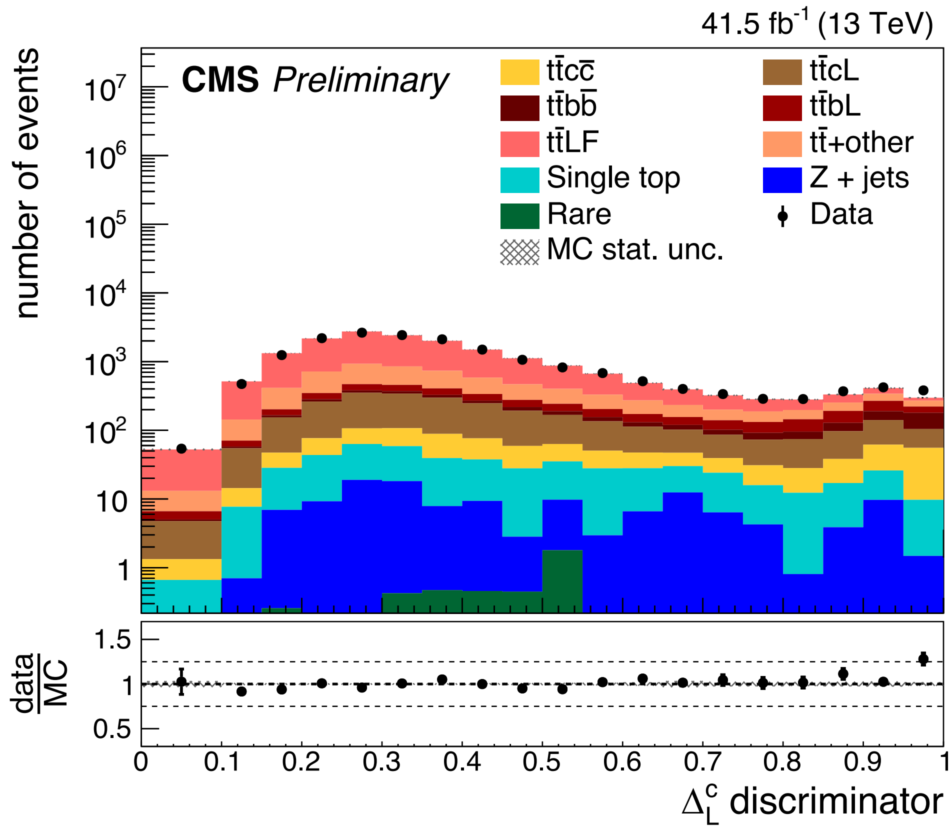


Bins 1 – 4 : $\Delta_L^c \in [0,0.45]$, and increasing bins in Δ_b^c

Template fit using NN discriminator

Comparison between data and simulation (prefit)





Template fit using NN discriminator
Systematic uncertainties

Normalization only:

Luminosity (2.3%)
background normalization (25%)
Efficiency (theoretical) (2%)
Fixed ratios from MC
(ttbL/ttbb, ttcL/ttcc, tt other/ttLF)
Bin-by-bin statistical uncertainty (MC)

These experimental uncertainties **affect the overall efficiency** from the fiducial to the reconstructed phase space, but do not change the shapes of the simulated templates.

Shape + normalization:

JES, JER
lepton ID/iso/reco/trigger
Pileup
c-tagging calibration

On top of affecting the selection efficiency, these experimental uncertainties **also change the shapes of the templates.**

Shape only:

μ_R, μ_F in ME generator
ISR+FSR: α_s in PS
Parton distribution function
Matching between ME/PS
Underlying event Tune
B-Fragmentation (not considered for now)

For these theoretical uncertainties, **only the change in shape of the templates** and their impact on the **acceptance** is considered in the extraction of the results. Their impact on the yield is quoted as an uncertainty on the theory prediction to which the measurement is compared.

Template fit using NN discriminator

Fits to extract inclusive cross sections and their ratios

- Two fits, one to extract the inclusive cross sections, one to extract their ratios
- Systematic uncertainties as nuisance parameters in the fit
- Fit is performed simultaneously in the ee/ μμ and eμ channels
- $\mathrm{t\bar{t}cL}$ ($\mathrm{t\bar{t}bL}$/ $\mathrm{t\bar{t}}$ other) scaled with the same factor as $\mathrm{t\bar{t}c\bar{c}}$ ($\mathrm{t\bar{t}b\bar{b}}$ / $\mathrm{t\bar{t}LF}$), i.e. ratio fixed to MC prediction (with uncertainties)
- Background contribution (<5%) is fixed at MC prediction (with 25% uncertainty)

Absolute cross sections

- Measure $\sigma_{\mathrm{t\bar{t}c\bar{c}}}, \sigma_{\mathrm{t\bar{t}b\bar{b}}}, \sigma_{\mathrm{t\bar{t}LF}}$

$$
\begin{aligned}
f\left(\sigma^{\mathrm{vis}}_{t\bar{t}c\bar{c}}, \sigma^{\mathrm{vis}}_{t\bar{t}b\bar{b}}, \sigma^{\mathrm{vis}}_{t\bar{t}\mathrm{LF}}\right) = \mathcal{L}^{int} \cdot \Bigg\{ & \sigma^{\mathrm{vis}}_{t\bar{t}c\bar{c}} \cdot \epsilon_{t\bar{t}c\bar{c}} \cdot \left(\mathrm{H}^{\mathrm{norm}}_{t\bar{t}c\bar{c}} + \frac{N^{\mathrm{MC}}_{t\bar{t}c\mathrm{L}}}{N^{\mathrm{MC}}_{t\bar{t}c\bar{c}}} \cdot \mathrm{H}^{\mathrm{norm}}_{t\bar{t}cL} \right) \\
& + \sigma^{\mathrm{vis}}_{t\bar{t}b\bar{b}} \cdot \epsilon_{t\bar{t}b\bar{b}} \cdot \left(\mathrm{H}^{\mathrm{norm}}_{t\bar{t}b\bar{b}} + \frac{N^{\mathrm{MC}}_{t\bar{t}b\mathrm{L}}}{N^{\mathrm{MC}}_{t\bar{t}b\bar{b}}} \cdot \mathrm{H}^{\mathrm{norm}}_{t\bar{t}bL} \right) \\
& + \sigma^{\mathrm{vis}}_{t\bar{t}\mathrm{LF}} \cdot \epsilon_{t\bar{t}\mathrm{LF}} \cdot \left(\mathrm{H}^{\mathrm{norm}}_{t\bar{t}\mathrm{LF}} + \frac{N^{\mathrm{MC}}_{t\bar{t}+\mathrm{Other}}}{N^{\mathrm{MC}}_{t\bar{t}LF}} \cdot \mathrm{H}^{\mathrm{norm}}_{t\bar{t}+\mathrm{Other}} \right) \\
& + \sigma_{\mathrm{DY}} \cdot \mathrm{H}^{\mathrm{norm}}_{\mathrm{DY}} + \sigma_{\mathrm{ST}} \cdot \mathrm{H}^{\mathrm{norm}}_{\mathrm{ST}} + \sigma_{\mathrm{Rare}} \cdot \mathrm{H}^{\mathrm{norm}}_{\mathrm{Rare}} \Bigg\}
\end{aligned}
$$

Ratios

- Measure $\mathrm{R_b} = {}^{\sigma_{\mathrm{t\bar{t}b\bar{b}}}}/_{\sigma_{\mathrm{t\bar{t}jj}}}$ and $\mathrm{R_c} = {}^{\sigma_{\mathrm{t\bar{t}c\bar{c}}}}/_{\sigma_{\mathrm{t\bar{t}jj}}}$

$$
\begin{aligned}
f\left(\sigma^{\mathrm{vis}}_{\mathrm{t\bar{t}}jj}, \mathrm{R_c}, \mathrm{R_b}\right) = \mathcal{L}^{int} \cdot \sigma^{\mathrm{vis}}_{\mathrm{t\bar{t}}jj} \cdot \Bigg\{ & \mathrm{R_c} \cdot \epsilon_{\mathrm{t\bar{t}c\bar{c}}} \cdot \left(\mathrm{H}^{\mathrm{norm}}_{\mathrm{t\bar{t}c\bar{c}}} + \frac{N^{\mathrm{MC}}_{\mathrm{t\bar{t}cL}}}{N^{\mathrm{MC}}_{t\bar{t}c\bar{c}}} \cdot \mathrm{H}^{\mathrm{norm}}_{t\bar{t}cL} \right) \\
& + \mathrm{R_b} \cdot \epsilon_{\mathrm{t\bar{t}b\bar{b}}} \cdot \left(\mathrm{H}^{\mathrm{norm}}_{\mathrm{t\bar{t}b\bar{b}}} + \frac{N^{\mathrm{MC}}_{\mathrm{t\bar{t}bL}}}{N^{\mathrm{MC}}_{t\bar{t}b\bar{b}}} \cdot \mathrm{H}^{\mathrm{norm}}_{t\bar{t}bL} \right) \\
& + (1 - \mathrm{R_c} - \mathrm{R_{cL}} - \mathrm{R_b} - \mathrm{R_{bL}}) \cdot \epsilon_{\mathrm{t\bar{t}LF}} \cdot \left(\mathrm{H}^{\mathrm{norm}}_{\mathrm{t\bar{t}LF}} + \frac{N^{\mathrm{MC}}_{t\bar{t}+\mathrm{Other}}}{N^{\mathrm{MC}}_{t\bar{t}LF}} \cdot \mathrm{H}^{\mathrm{norm}}_{t\bar{t}+\mathrm{Other}} \right) \Bigg\} \\
& + \mathcal{L}^{int} \cdot \{ \sigma_{\mathrm{DY}} \cdot \mathrm{H}^{\mathrm{norm}}_{\mathrm{DY}} + \sigma_{\mathrm{ST}} \cdot \mathrm{H}^{\mathrm{norm}}_{\mathrm{ST}} + \sigma_{\mathrm{Rare}} \cdot \mathrm{H}^{\mathrm{norm}}_{\mathrm{Rare}} \}
\end{aligned}
$$

$$
\text{with} \quad \begin{cases} \mathrm{R_{cL}} = \mathrm{R_c} \cdot \left(\dfrac{N^{\mathrm{MC}}_{\mathrm{t\bar{t}cL}} \, \epsilon_{\mathrm{t\bar{t}cL}}}{N^{\mathrm{MC}}_{t\bar{t}c\bar{c}} \, \epsilon_{\mathrm{t\bar{t}c\bar{c}}}} \right) \\ \mathrm{R_{bL}} = \mathrm{R_b} \cdot \left(\dfrac{N^{\mathrm{MC}}_{\mathrm{t\bar{t}bL}} \, \epsilon_{\mathrm{t\bar{t}bL}}}{N^{\mathrm{MC}}_{t\bar{t}b\bar{b}} \, \epsilon_{\mathrm{t\bar{t}b\bar{b}}}} \right) \end{cases}
$$

Results in the fiducial phase space are extrapolated to the full phase space by means of acceptance A.

Event category	$\mathrm{t\bar{t}b\bar{b}}$	$\mathrm{t\bar{t}bL}$	$\mathrm{t\bar{t}c\bar{c}}$	$\mathrm{t\bar{t}cL}$	$\mathrm{t\bar{t}LF}$
Efficiency ϵ (%)	12.5	8.9	7.1	5.9	4.7
Acceptance $\mathcal{A}$ (%)	2.9	2.5	2.0	2.0	2.3

Template fit using NN discriminator

Defining the neural network

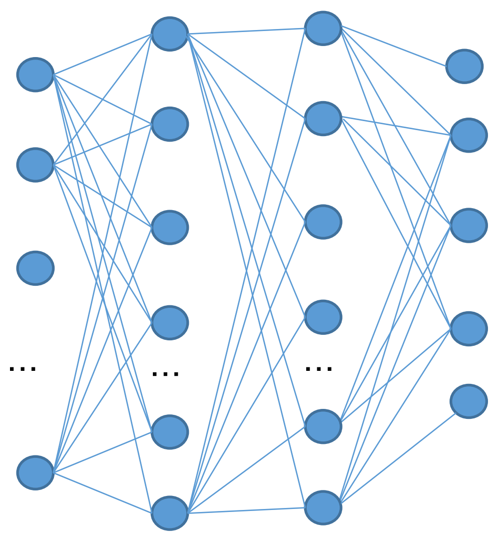
- Neural network trained with Keras (TensorFlow backend)
- 6 inputs (Standard Normalization, μ=0, σ= 1)
- 1 hidden layer with 30 neurons and ReLu activation
- 10% Dropout in hidden layer (regularization)
- 5 outputs with SoftMax activation
- Loss function = categorical cross-entropy
- Optimizer = Stochastic Gradient Decent
  *learning rate (init) = 0.001, decay = init / (5*n_epoch), nesterov momentum = 0.8*
- n_epoch = 100, batch_size = 128

CvsL add. Jets

CvsB add. Jets

Parton match NN

ΔR(add. Jets)



P(ttcc)

P(ttcL)

P(ttbb)

P(ttbL)

P(ttLF)

$$\Delta_b^c = \frac{\mathrm{P}(t\bar{t}c\bar{c})}{\mathrm{P}(t\bar{t}c\bar{c}) + \mathrm{P}(t\bar{t}b\bar{b})}$$

$$\Delta_L^c = \frac{\mathrm{P}(t\bar{t}c\bar{c})}{\mathrm{P}(t\bar{t}c\bar{c}) + \mathrm{P}(t\bar{t}\mathrm{LF})}$$