$t\bar{t}H(H \to c\bar{c})$
in the 0-lepton channel

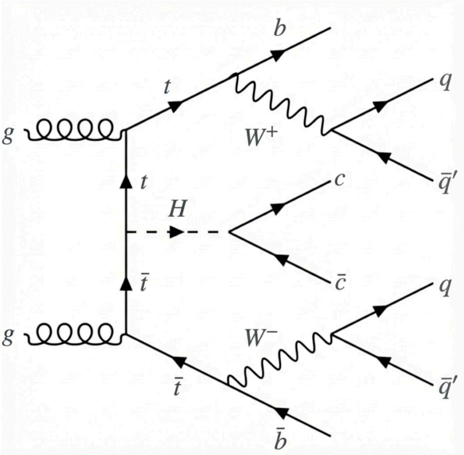
Trigger modelling, a 10-class classifier, and the first fit

Krvavec

September 28, 2026

Andraž Tomšič
Jožef Stefan Institute
Advisor: Miha Muškinja

The 0-lepton channel



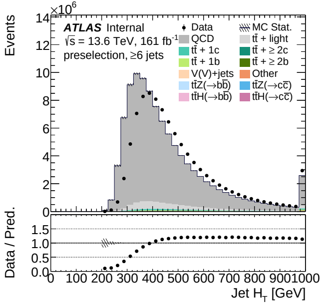
$t\bar{t}H(c\bar{c})$ 0L: both $W \to q\bar{q}'$, eight quarks in the final state.

Largest branching fraction, not used before.

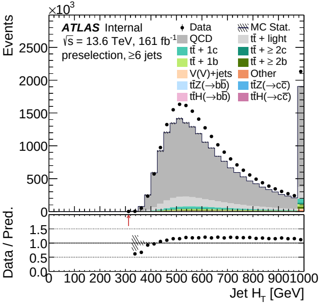
Needs multi-b-jet triggers, and carries a **heavy, poorly modelled QCD multijet background** not present in the other channels.

- Data: Run 2, 140 fb^{-1} at 13 TeV, plus partial Run 3 (2022–24), 161 fb^{-1} at 13.6 TeV.
- Preselection: 0 leptons, ≥ 6 jets, b-jet trigger with kinematic matching, ≥ 2 b-tagged, ≥ 3 b+c-tagged $\to$ 1.36×10^7 data events.

The 0L problem: trigger modelling



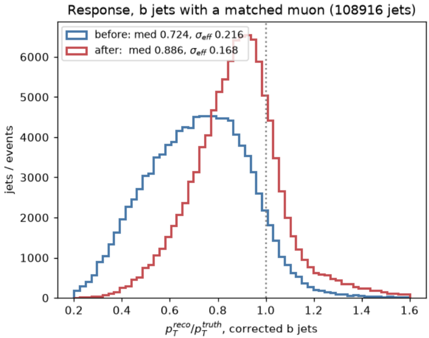
no trigger matching

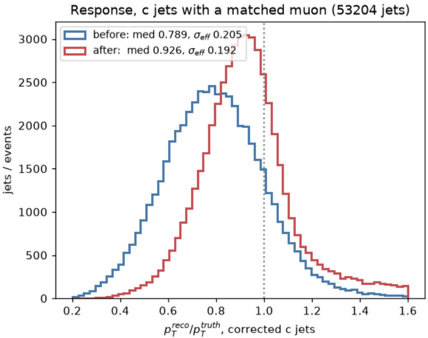


after $1.5\times$ kinematic matching

- Jet H_T (scalar $\sum p_T$ of jets): mismodelled at low H_T.
- Investigated with an ML QCD cut: partly **QCD mismodelling**, not only turn-on.
- Fix: kinematic matching, two trigger menus matched separately (p_T and b-tag WP), $1.5\times\ p_T$ threshold. Less signal loss than an H_T cut; threshold may be loosened.
- Signals: $t\bar{t}H(c\bar{c})$, $t\bar{t}H(b\bar{b})$, $t\bar{t}Z(c\bar{c})$, $t\bar{t}Z(b\bar{b})$. Backgrounds: $t\bar{t}$, single/rare top, $t\bar{t}W$, $V(V)$+jets, all fine; QCD multijet is the “rotten egg”.

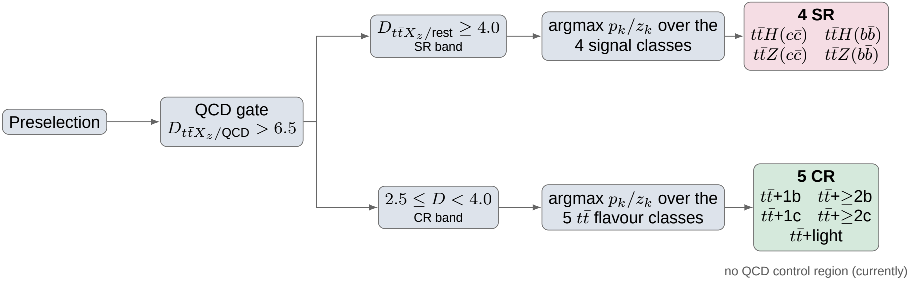
Muon-in-jet correction



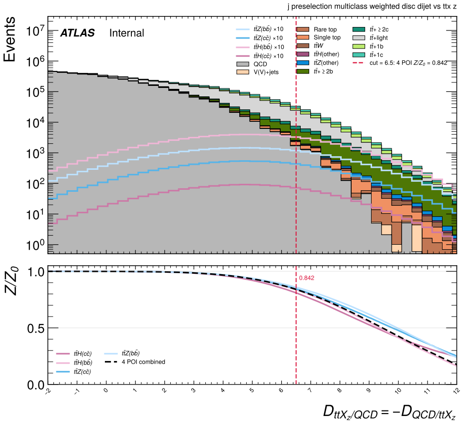


- Semileptonic b, c decays put a muon in the jet, whose energy the jet misses. Fix: **add the soft muon back**.
- Example: $B \to D\,\mu\,\nu$ ($b \to c$), then $D \to K\,\mu\,\nu$ ($c \to s$).
- Difference between reco and truth narrows, as the plots show.
- Corrected: 11.7% of b-jets, 5.5% of c, only 0.6% of light.

Ten classes, nine used in the fit



- ParticleTransformer, ≈ 2M parameters, jets plus a E_T^{miss} token. Discriminants from the class posteriors: $D_{A/B} = \log\left(\sum_{k\in A} p_k \big/ \sum_{k\in B} p_k\right)$, $t\bar{t}X$ = the four signals.
- The tenth class, multijet, builds the gate but no region: QCD is still poorly modelled, so it is cut away rather than modelled.



	$t\bar{t}H(c\bar{c})$	$t\bar{t}H(b\bar{b})$	$t\bar{t}Z(c\bar{c})$	$t\bar{t}Z(b\bar{b})$
ε(sig.)	23%	24%	24%	24%
Z/Z_0	0.808	0.835	0.843	0.858

Gate on $t\bar{t}X$ vs QCD, signals weighted by cross section (the z in $t\bar{t}X_z$): one cut for four POIs.

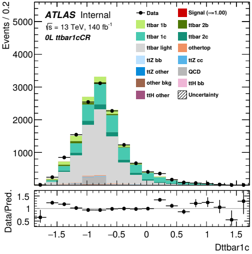
Lose 3/4 of the signal, keep $\geq 80\%$ of Z.

0.036% of the QCD remains, 0.85% of the data.

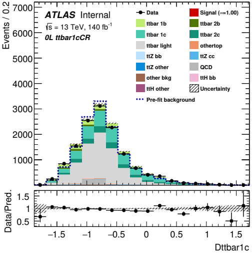
- $Z^2 = \sum_{\text{bins}} 2[(s+b)\ln(1+s/b) - s]$, on $D_{t\bar{t}H(c\bar{c})/\text{rest}}$.
- Second cut on $D_{t\bar{t}X_z/\text{rest}}$ (backup), splits the SR from a CR band below it: Z/Z_0 almost 1. Both thresholds pending optimisation.

The fit and its expected significance

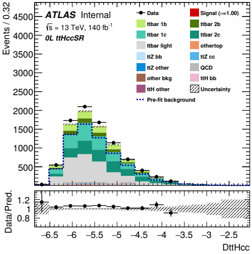
CR $t\bar{t}$+1c, pre-fit



CR $t\bar{t}$+1c, post-fit



SR $t\bar{t}H(c\bar{c})$, post-fit



expected Z		$t\bar{t}H(c\bar{c})$	$t\bar{t}H(b\bar{b})$	$t\bar{t}Z(c\bar{c})$	$t\bar{t}Z(b\bar{b})$
Counting, preselection	per-bin Asimov, no fit	**0.454**	18.18	3.41	8.67
Counting, own SR	same rule, after the cuts	**0.397**	15.14	3.15	7.58
Fit, Asimov, stat only	no norm factors, no MC stat	0.392	14.28	2.99	7.20
Fit, expected pre-fit	full model, NPs at nominal	0.295	9.05	2.33	5.22
Fit, expected post-fit	NPs profiled to data at $\mu = 1$	**0.239**	8.02	2.02	4.60

- Every region has enough statistics, and all nine sit within a factor 3.3 of each other.
- The fit has **no systematics** yet.

Summary, and what needs work

- 0L runs end to end: trigger modelling, objects, a 10-class transformer, nine regions, a fit that converges. **Only systematics are missing**.
- One QCD cut for all four SRs: loses $\sim 3/4$ of the signal and 16% of the significance, while cutting away more than three orders of magnitude of QCD.

- Systematics first. Experimental, then $t\bar{t}$ heavy-flavour modelling.
- Optimisation of the fit: cuts, binning, CR-to-SR closure.
- Modelling improvements for 0L.
- Run 3 fit, and the combined 0L+1L+2L fit.

Backup

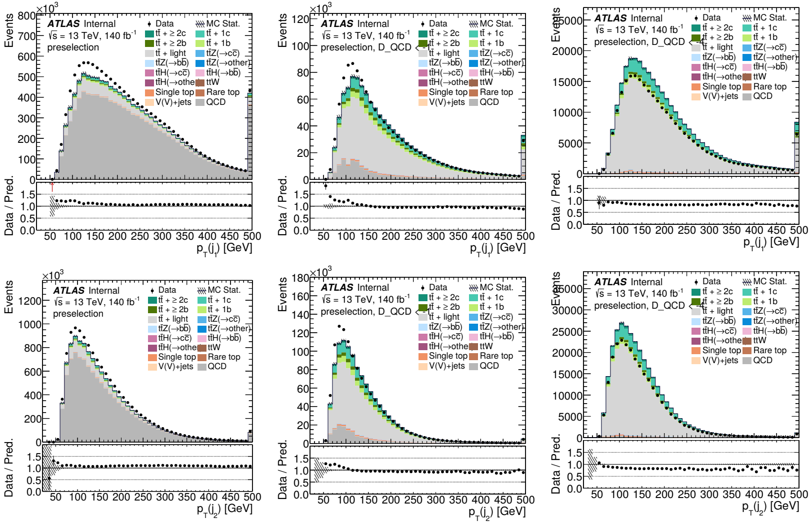
Backup

- Trigger matching and the ML QCD cut, jet p_T.
- The $t\bar{t}X_z$ band, the second cut.
- Post-fit control regions.
- Muon-in-jet: what it changes in 1L.

no QCD cut | $D_{\text{QCD}} < -1$ | $D_{\text{QCD}} < -4$

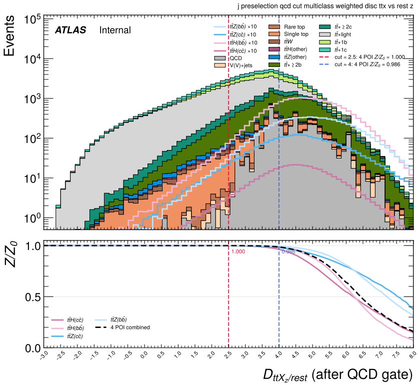
$p_T(j_1)$

$p_T(j_2)$



Mismodelling worse without the QCD cut, at low p_T.

The $t\bar{t}X_z$ band: the second cut

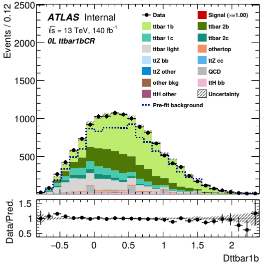


	$t\bar{t}H(c\bar{c})$	$t\bar{t}H(b\bar{b})$	$t\bar{t}Z(c\bar{c})$	$t\bar{t}Z(b\bar{b})$
CR band	26%	20%	24%	18%
SR band	70%	79%	70%	80%
Z/Z_0	0.963	0.984	0.981	0.993

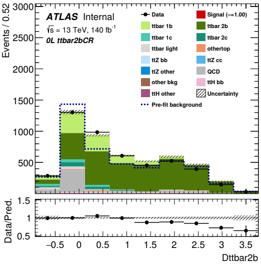
Z/Z_0 at the SR edge. Read after the gate: 99.1% of the background already gone.

CR band $2.5 \leq D < 4.0$, SR band $D \geq 4.0$: adjacent, and they partition what survives.

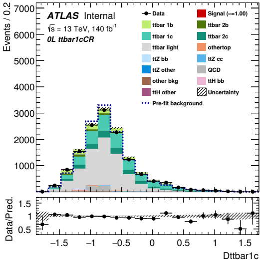
Post-fit control regions



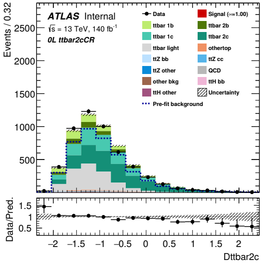
$t\bar{t}$+1b



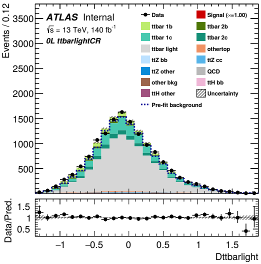
$t\bar{t}$+$\geq$2b



$t\bar{t}$+1c



$t\bar{t}$+$\geq$2c



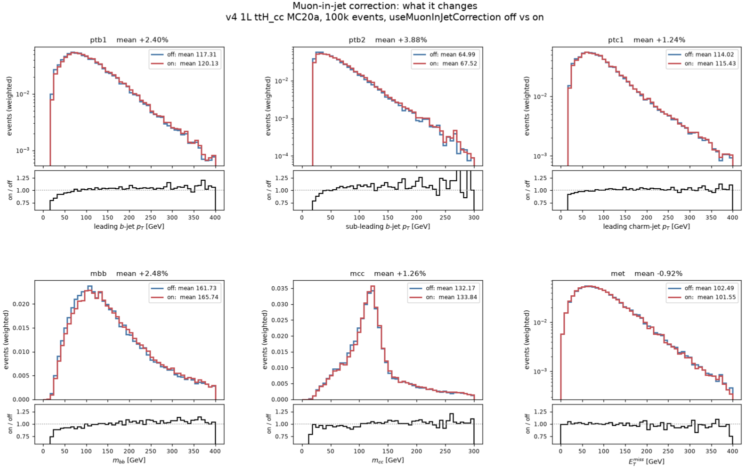
$t\bar{t}$+light

Global fit to data at $\mu = 1$.

Dashed: pre-fit background.

The $\geq$2b region keeps a slope after the fit.

Muon-in-jet: what it changes (1L $t\bar{t}H(c\bar{c})$, v4)



- Cutflows and everything the correction does not touch agree exactly.
- What moves: p_T^{b1} $+2.4\%$, p_T^{b2} $+3.9\%$, p_T^{c1} $+1.2\%$, m_{bb} $+2.5\%$, m_{cc} $+1.3\%$, E_T^{miss} -0.9%.