Search for light charged Higgs bosons in $t \to H^{\pm}(cb)b$ decays with the ATLAS detector

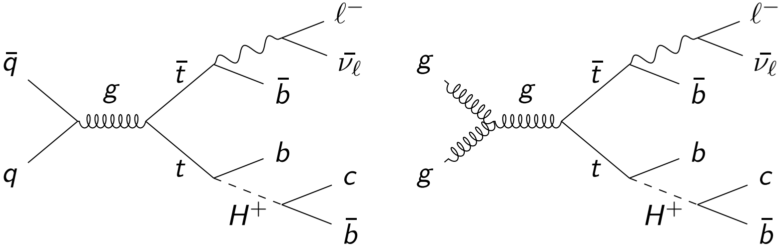
Chen Peng

on behalf of the ATLAS Collaboration

Charged-Higgs@LHC
Aug 31, 2021

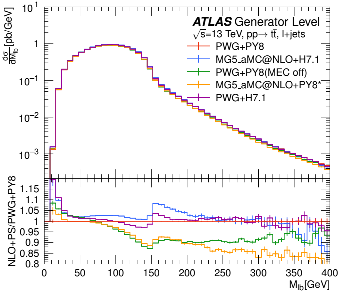
Introduction

- Some beyond Standard Model (BSM) theories with extend Higgs sector (Two-Higgs-Doublet Models (2HDMs), 3HDMs, MSSM, etc.) predict the existence of charged Higgs ($H^{\pm}$).
- 3HDMs feature two $H^{\pm}$ that can be lighter than the top quark $m_{H^{\pm}} < m_t$.
- Such low mass $H^{\pm}$ can decay predominantly to a charm an a bottom quark ($H^{\pm} \to cb$).
- This research is focused on low mass $H^{\pm}$ ($60 < m_{H^{\pm}} < 160$ GeV) produced from top decays with $t\bar{t} \to WbH^{\pm}b$.
  - first time for a search in this channel within ATLAS.
- Model independent upper limits on $\mathscr{B} = \mathscr{B}(t \to H^{\pm}b) \times \mathscr{B}(H^{\pm} \to cb)$ are reported as a function of $m_{H^{\pm}}$.



Introduction

- Main background comes from SM $t\bar{t} \to WbWb$, splitting into flavour components: $t\bar{t} + \geq 1b$, $t\bar{t} + \geq 1c$, $t\bar{t}$ + light.
    - Nominal $t\bar{t}$ are simulated with POWHEG+PYTHIA8.
    - NLO generator and PS uncertainties are estimated by comparing nominal $t\bar{t}$ with MG5_aMC@NLO+PYTHIA8 and POWHEG+HERWIG7 separately.
    - Uncertainties on initial and final state radiation are included.



[ATL-PHYS-PUB-2020-023]

- For signal, the final state has large b-jet multiplicity.
    - Using the DNN based b-tagging algorithm (DL1r) with tightest operating point (60%) to suppress $t\bar{t}$ + light [1907.05120].
    - b/c-tagging efficiencies and light-jet mis-tagging rate are corrected to data.
    - Statistical and systematic uncertainties (from detector calibration and modelling) on efficiencies or mis-tagging rate measurements are included.

Selection

This research exploits following selection to suppress background:

Leptons

- Electrons:
  - $p_{\mathrm{T}} > 28$ GeV
  - $|\eta| < 2.47$, excluding $1.37 < |\eta| < 1.52$
- Muons:
  - $p_{\mathrm{T}} > 28$ GeV
  - $|\eta| < 2.5$

Jets

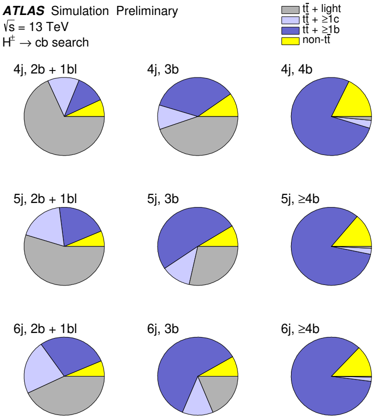
- Jets:
  - $p_{\mathrm{T}} > 25$ GeV
  - $|\eta| < 2.5$
- B-taggged jets:
  - DL1r at 60% OP.

- Tight b-tagging WP @60% combining with a loose b-tagging WP @70% is used for analysis region definition.

Pre-selection	
Trigger	single-lepton trigger
Leptons	$= 1$ isolated e or μ
Jets	≥ 4 jets
B-tagged jets	≥ 2 b-tagged jets
MET	> 20 GeV
MET $+ m_{\mathrm{T}}^{W}$	> 60 GeV

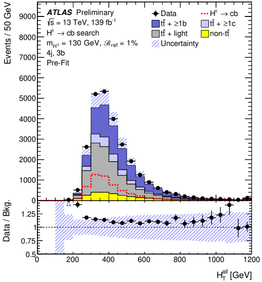
Analysis strategy: event categorisation

- Categorisation: nine regions based on N jets (4j, 5j and 6j) and N b-tagged jets (2b + 1bl, 3b and ≥4b).
  - 2b + 1bl regions: derive data-based corrections to improve modelling of the $t\bar{t}$.
  - (4j, 3b), (5j, 3b) and (6j, 3b) regions: main signal regions.
  - (4j, 4b) regions: recover acceptance for signal events with a mis-tagged c-quark.
  - (5j, ≥4b) and (6j, ≥4b): $t\bar{t}+\geq 1b$ background control regions.

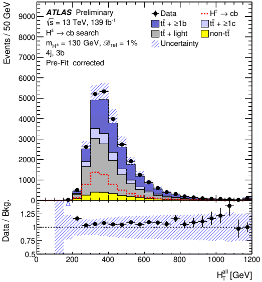


Analysis strategy: $t\bar{t}$ modelling

- The $t\bar{t}$ simulation does not perfectly model jet multiplicity and p_{T} in data.
- A data-based approach is used to correct $t\bar{t}$ simulation.
  - Corrections are derived in three 2b + 1bl regions independently as a function of $H_{\mathrm{T}}^{\mathrm{all}}$ and jet multiplicity defined as: $C(H_{\mathrm{T}}^{\mathrm{all},i}, j^i) = \frac{N^{\mathrm{data}}(H_{\mathrm{T}}^{\mathrm{all},i}, j^i) - N^{\mathrm{non}-t\bar{t}}(H_{\mathrm{T}}^{\mathrm{all},i}, j^i)}{N^{t\bar{t}}(H_{\mathrm{T}}^{\mathrm{all},i}, j^i)}$.
  - Small signal contamination (less than 1.5%) in 2b + 1bl region.
  - Derived corrections are close to one for $H_{\mathrm{T}}^{\mathrm{all}} > 800$ GeV and increase monotonically toward lower $H_{\mathrm{T}}^{\mathrm{all}}$, reaching 1.2 for $H_{\mathrm{T}}^{\mathrm{all}} = 200$ GeV.
  - Correction factors are parameterised with a rational function.



before corrections

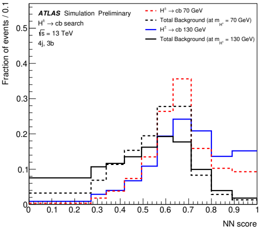


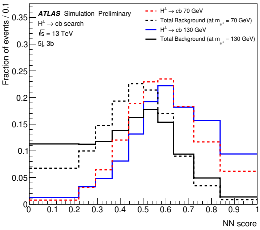
after corrections

- $H_{\mathrm{T}}^{\mathrm{all}} = \sum_{l,\mathrm{jets},\mathrm{MET}} p_{\mathrm{T}}$

Analysis strategy: neural network discriminant

- Building a neural network to separate signal from background.
- Selecting events with $\geq$4j and $\geq$3b to train the neural network.
- Parameterising the neural network as a function of $m_{H^\pm}$ by incorporating the signal mass $m_{H^\pm}$ as a label in the training.
- Input variables (jets are ordered by pseudo-continuous b-tagging scores, then using p_{T} ordering for jets with degenerate b-tagging scores):
    - p_{T}, η and ϕ of first six leading jets.
    - b-tagging score for the fourth, fifth and sixth jets.
    - lepton p_{T}, η, ϕ, MET and ϕ of MET.
    - Di-jet mass variables of fourth leading jet (expected to originate from c) and a selected b-jet ($M_{jb}^{\mathrm{Leading bjet}}$, $M_{jb}^{\mathrm{Max}\Delta R}$, $M_{jb}^{\mathrm{Min}\Delta R}$).





Fit to data

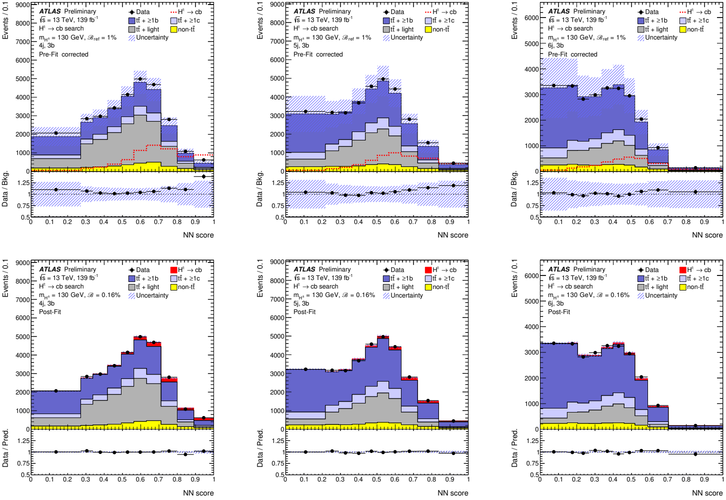
- Fit is performed across regions with ≥ 3 b-jets simultaneously under signal-plus-background hypothesis, maximising a binned likelihood function.
  - Using 10 bins in $3b$ regions and 1 bin in $\geq 4b$ regions.
  - Likelihood function is constructed as a product of Poisson probability terms over all bins.
  - A complete set of systematic uncertainties are considered, including uncertainties on object reconstruction and background modelling.
  - Statistical uncertainties in each bin of the predicted NN score are taken into account.
  - The signal strength μ defined as $\mathscr{B} = 1\% \times \mu$ is set as a free parameter.

Major systematic uncertainties

- Uncertainties related to c-tagging efficiencies and light-jet mis-tagging rate measure are leading uncertainties for some masses, $\Delta\mathscr{B} \sim 0.03$).
- $t\bar{t}$ modelling uncertainties:
  - 50% $t\bar{t} + \geq 1b$ and $t\bar{t} + \geq 1c$ normalisation uncertainties.
  - NLO generator and PS uncertainties, uncorrelated in jet multiplicity and $t\bar{t}$ categories. NLO generator uncertainty in $4j$ regions is leading uncertainties for several masses ($\Delta\mathscr{B} \sim 0.05\%$).
  - Four-flavour scheme vs five-flavour scheme on $t\bar{t} + \geq 1b$.
  - Initial and final state radiation on $t\bar{t}$.
- A set of uncertainties from data-based $t\bar{t}$ corrections (uncorrelated in N_j):
  - Statistical uncertainty on the parametrisation of the correction factors.
  - Non-$t\bar{t}$ background normalisation uncertainties are propagated through the correction.

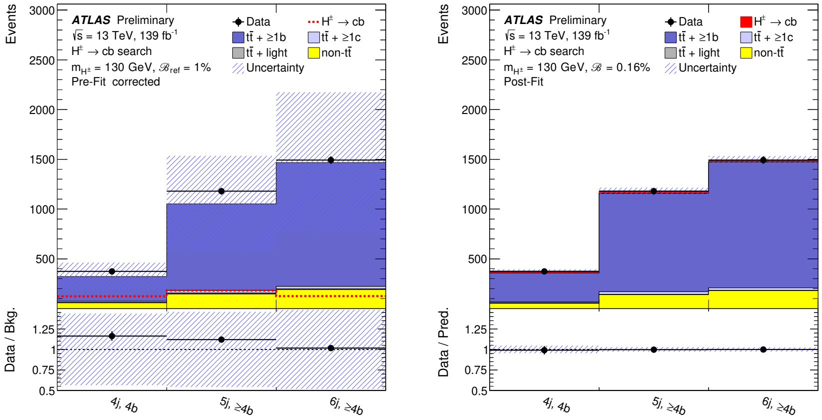
Data/MC agreement before and after fit

- Showing NN score in 3b regions before (signal normalised to $\mathscr{B}_{\text{ref}} = 1\%$) and after ($\mathscr{B} = 0.16\%$) at $m_{H^{\pm}}$=130 GeV.



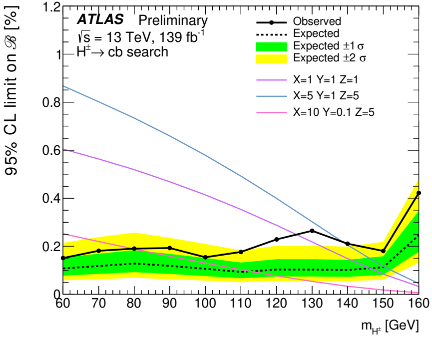
Data/MC agreement before and after fit

- Showing NN score in $\geq 4b$ regions before (signal normalised to $\mathscr{B}_{\mathrm{ref}} = 1\%$) and after ($\mathscr{B} = 0.16\%$) at $m_{H^\pm}$=130 GeV.
- Good data/MC agreement after the likelihood fit to data.



Results

- In the absence of a significant excess of data, 95% CL limits are set on $\mathscr{B}$.
- The observed (expected) limits vary between 0.15% (0.09%) and 0.42% (0.25%).
- The predictions from 3HDM corresponding to three benchmark values of X, Y, Z are shown [1810.05430].
- The largest excess in data is seen at $m_{H^{\pm}} = 130$ GeV, corresponding to $\sim 3\sigma$.
- A global probability for the most significant excess to be observed anywhere is $\sim 2\sigma$.
- The broad excess in data is consistent with the expected $m_H^{\pm}$ resolution.
- The acceptance loss for the b-jet produced from $t \to H^{\pm}b$ increases for $m_H^{\pm}$ close to m_t, resulting in weaker exclusion limits.



Summary

- A search for $H^{\pm}$ in top-quark decays is presented, using full Run 2 data.
- The search exploits the high multiplicity of b-jets and deploys a neural network classifier to separate signal events from background.
- Model independent exclusion limits at 95% confidence level on $\mathscr{B}$ vary between 0.15% (0.09%) and 0.42% (0.25%)
- The largest excess in data of $\sim 3\sigma$ is observed at $m_{H^{\pm}} = 130$ GeV with a global significance of $\sim 2\sigma$.
- The search improves the expected sensitivity by a factor of five compared to CMS results ([1808.06575]) and explores an extended $m_H^{\pm}$ range.