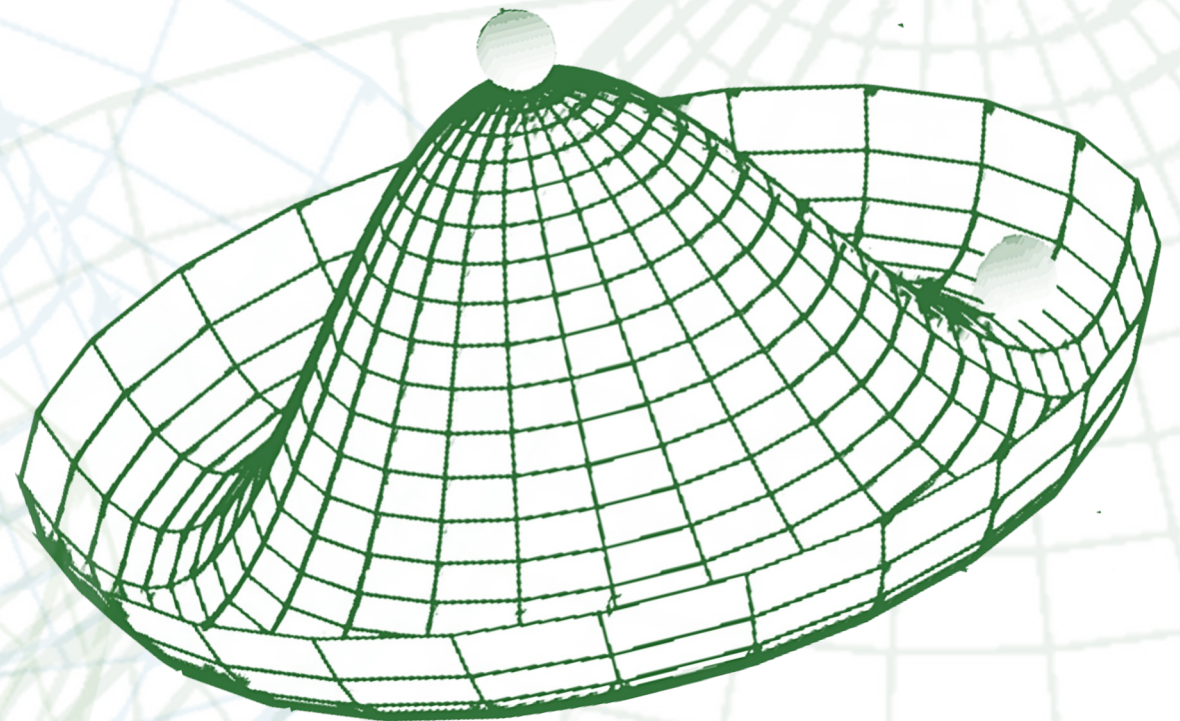




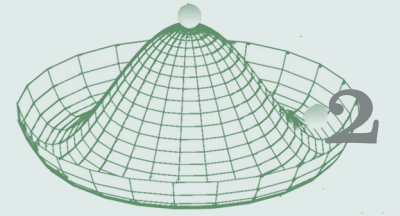
# Searching for Beyond the Standard Model Higgs Bosons



**UNIVERSITY  
OF WARWICK**

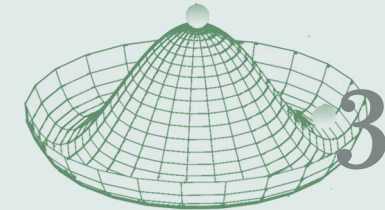


# Outline



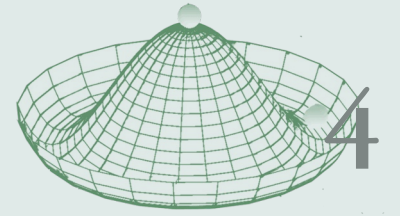
1. Background & Motivation
2. Analysis Outline
3. Signal and Background Modelling
4. Statistical Analysis
5. Summary & What's Next

# Outline



1. Background & Motivation
2. Analysis Outline
3. Signal and Background Modelling
4. Statistical Analysis
5. Summary & What's Next

# Two Higgs Doublet Models



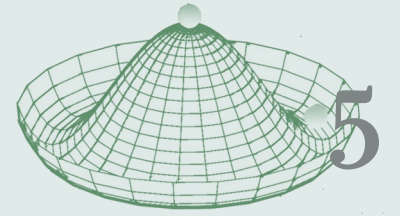
- One of the simplest extensions of the SM, introducing another  $SU(2)_L$  Higgs doublet.
- This leads to 8 fields, which give the masses of the  $W^\pm$  and  $Z^0$ , and the prediction of **5 total Higgs bosons** - the  $h$  (SM Higgs),  $A$  (pseudoscalar),  $H$  (scalar) and  $H^\pm$ .



$$\Phi_a = \begin{pmatrix} \Phi_a^\pm \\ \frac{1}{\sqrt{2}}(v_a + \rho_a + i\eta_a) \end{pmatrix}, \quad a = 1, 2.$$



# Two Higgs Doublet Models



$$\tan(\beta) = \frac{v_1}{v_2} \quad : \text{ratio of the v.e.v.s of the 2 doublets}$$

$\beta$  : rotation angle that diagonalises the mass-squared matrices of the charged scalars and pseudoscalars.

**Note:**

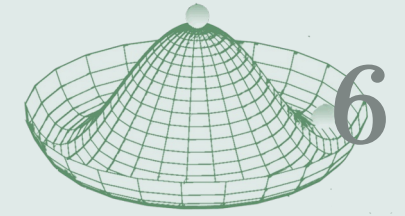
Mixing in a general 2HDM can give rise to **flavour changing neutral currents (FCNC)**.

This can be forbidden by applying a **discrete  $Z_2$  symmetry** (e.g  $\Phi_1 \rightarrow -\Phi_1$ ) on the Lagrangian.

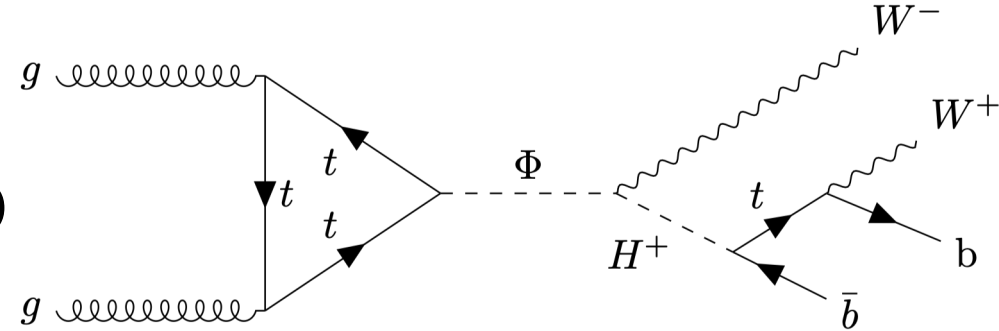
Model	$u_R^i$	$d_R^i$	$e_R^i$	$\rho^U$	$\rho^D$	$\rho^L$
Type-I	$\Phi_2$	$\Phi_2$	$\Phi_2$	$\kappa_U \cot(\beta)$	$\kappa_D \cot(\beta)$	$\kappa_L \cot(\beta)$
Type-II	$\Phi_2$	$\Phi_1$	$\Phi_1$	$\kappa_U \cot(\beta)$	$-\kappa_D \tan(\beta)$	$-\kappa_L \tan(\beta)$
Lepton Specific	$\Phi_2$	$\Phi_2$	$\Phi_1$	$\kappa_U \cot(\beta)$	$-\kappa_D \tan(\beta)$	$\kappa_L \cot(\beta)$
Flipped	$\Phi_2$	$\Phi_1$	$\Phi_2$	$\kappa_U \cot(\beta)$	$\kappa_D \cot(\beta)$	$-\kappa_L \tan(\beta)$

Table 1. Fermion coupling to doublets in 2HDMs with a discrete  $Z_2$  symmetry, and the corresponding Yukawa couplings.

# Search Motivation



- Search for  $\Phi \rightarrow H^\pm W^\mp \rightarrow \dots \rightarrow bbWW$  ( $\Phi = A/H$ )
- $\text{BR}(H^\pm \rightarrow tb) \sim 100\%$
- Allowed decay in Type-I 2HDMs (less explored than type-II)
- Decay mode has never been searched for before



## Gildener-Weinberg 2HDM

- Naturally aligned model. *~SM Higgs (h) couplings, and small mixing between h and H*
- Mass linkage at 1-loop level, modified at 2-loop level

## Electroweak Baryogenesis in aligned 2HDM ★

$A \rightarrow t\bar{t}$  results in [CMS/ATLAS](#)

- Kinematically favoured in certain phase spaces
- Expected cross section  $O(\text{a couple})$  pb at  $\tan(\beta) = 2$

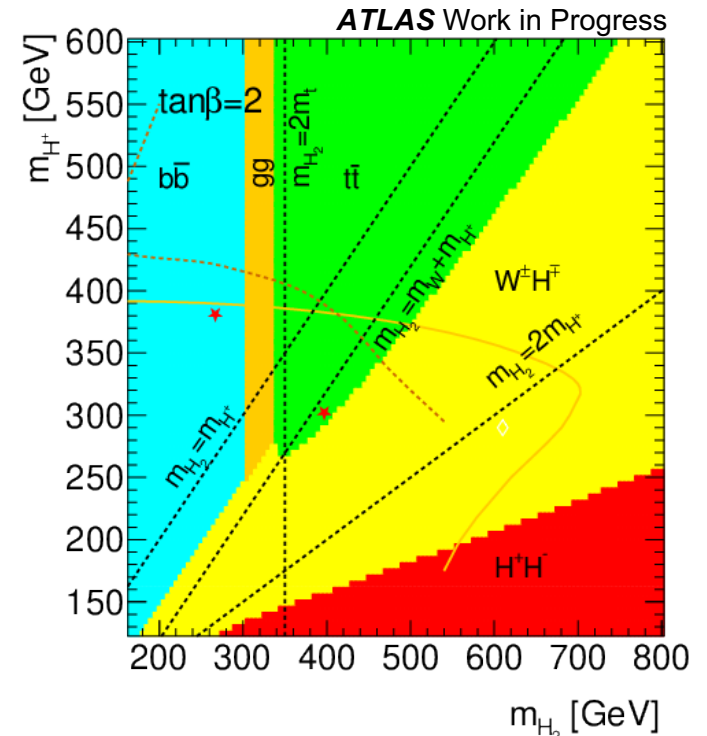
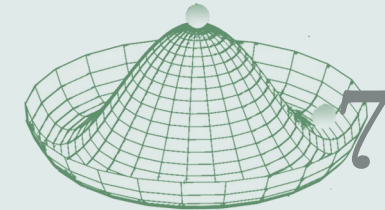


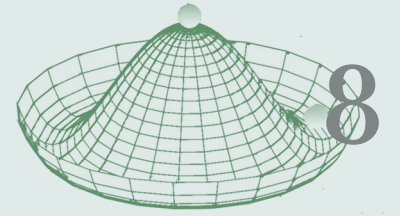
Fig.1. Dominant decay mode of the scalar heavy Higgs for varying  $m_H, m_{H^\pm}$  Plot by B. Murray

# Outline



1. Background & Motivation
2. Analysis Outline
3. Signal and Background Modelling
4. Statistical Analysis
5. Summary & What's Next

# Analysis Outline



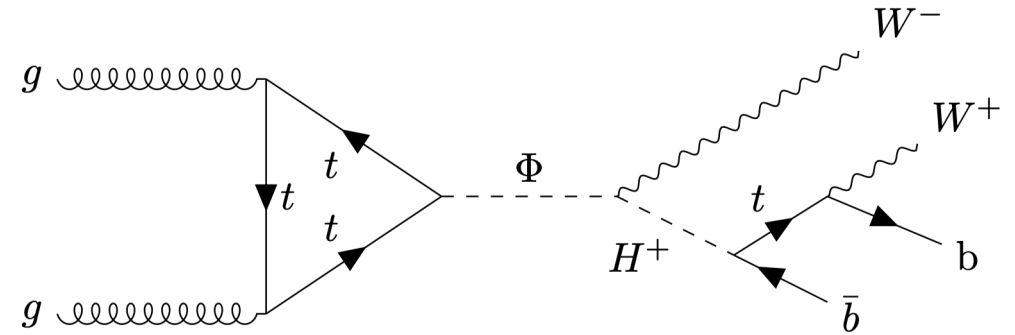
Looking at ggF production mode, and semi-leptonic WW decay

Run 2 ATLAS data, 2015-2018

Signal mass range:

$$m_{H^\pm} \in [200, 800] \text{ GeV}$$

$$m_\Phi \in [300, 1000] \text{ GeV}$$

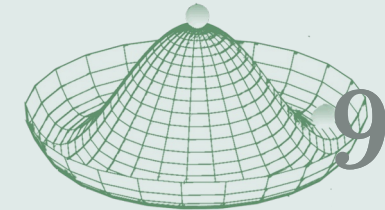


**Selection:**  $\geq 2$  light jets,  $\geq 2$  b jets, 1 lepton ( $e/\mu$ )

**A/H<sup>±</sup> Reconstruction:** Chi-squared fit to reconstruct H<sup>±</sup> and  $\Phi$  bosons (and other objects)

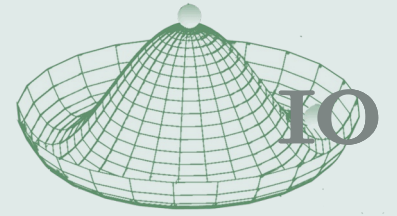
**Statistical fit:** Fit 2D distribution of  $m_{H^\pm}$  vs  $(m_\Phi - m_{H^\pm})$

# Outline



1. Background & Motivation
2. Analysis Outline
3. Signal and Background Modelling
4. Statistical Analysis
5. Summary & What's Next

# Signal Modelling



Monte Carlo simulations are used to model the signal across the phase space where the decay to  $H^\pm W^\mp$  is dominant.

Over 30  $(m_A, m_{H^\pm})$  values are generated, with  $\tan(\beta) = 2$

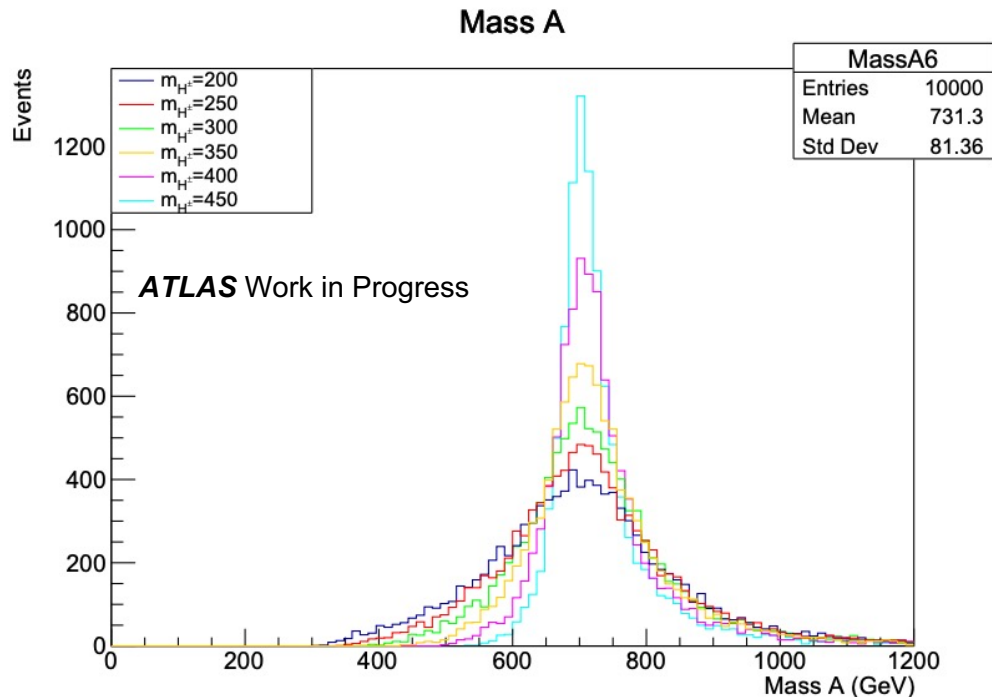


Fig.3. Plot of  $m_A = 700$  GeV, showing the width variation for varying  $m_{H^\pm}$

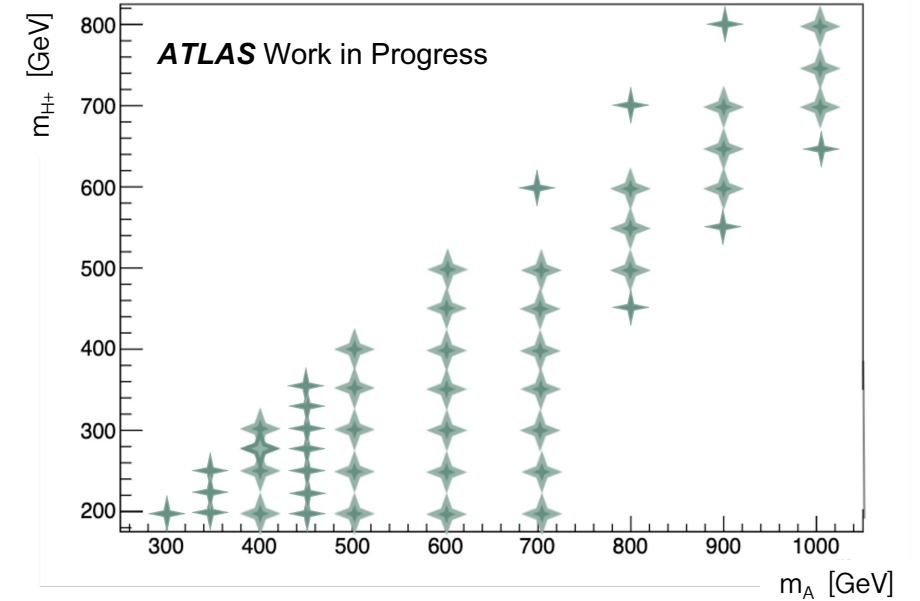
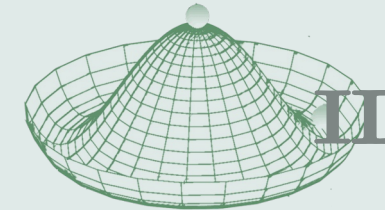


Fig.2. Generated Monte Carlo signal points plotted in  $m_A$  vs  $m_{H^\pm}$


- Widths and cross sections are dependent on  $\tan(\beta)$ , masses, and whether an A/H
- Re-weighting mass distributions to study both A/H decays, and a range of  $\tan(\beta)$  values


# Background Modelling



  $t\bar{t}$


 Single top

 W+jets

 Z+jets

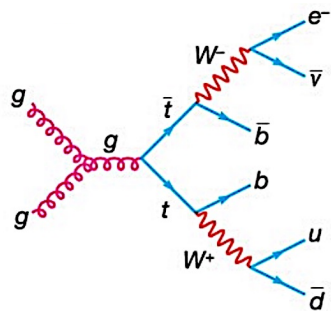
 Di-boson

 Toponium

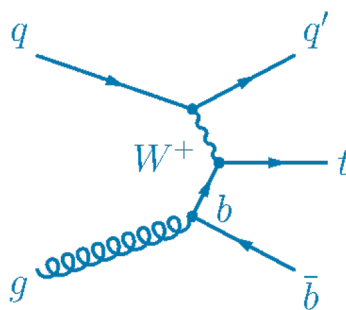
 QCD-multijet

Monte Carlo

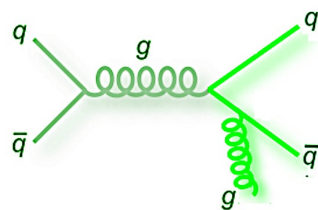
Data-driven



$t\bar{t}$  process



single top process



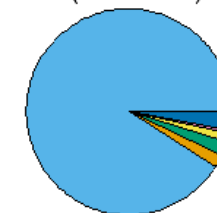
QCD multijet process

Signal regions mainly composed of  $t\bar{t}$  ...

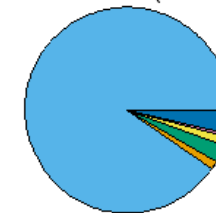
ATLAS Work in Progress



SR 0 (low mass)



Loose SR 0 (low mass)



W+Jets CR (low mass)



ttbar CR (low mass)

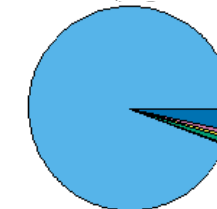
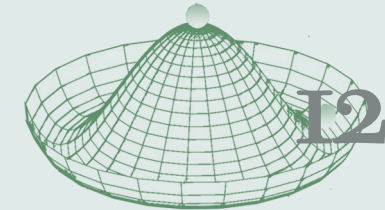


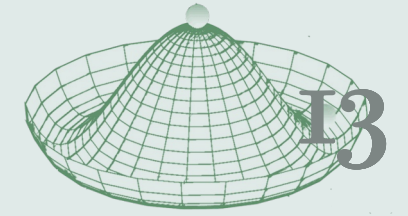
Fig.4. Background composition of the W+jets and  $t\bar{t}$  CRs, and 2 of the signal regions used in the statistical analysis

# Outline



1. Background & Motivation
2. Analysis Outline
3. Signal and Background Modelling
4. Statistical Analysis
5. Summary & What's Next

# Statistical Analysis



- 🍷 Train an XGBoost multi-classifier BDT with kinematic variables, to distinguish between signal, top and W+jets
- 🍷 Split into 2 signal mass regimes for optimum training.
- 🍷 Outputs probability of an event being signal, top or W+jets
  - 🍷 Use the log ratio of these to define fit signal, control (and validation) regions for each mass range

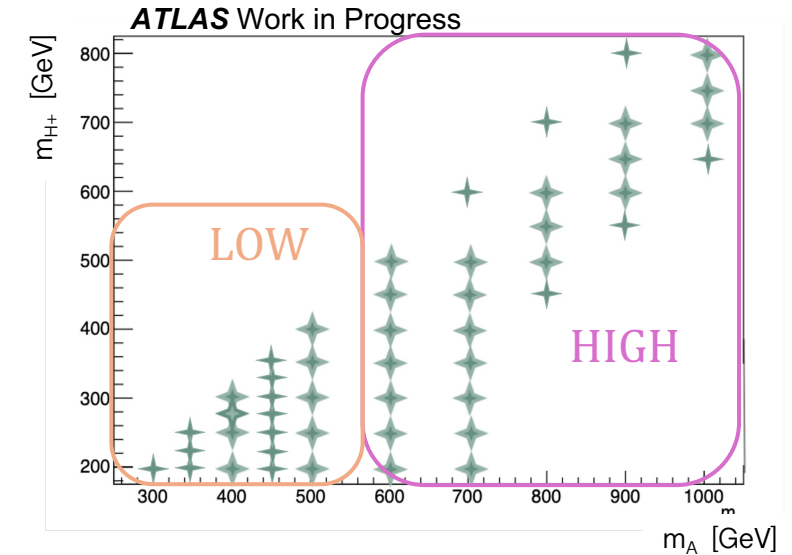


Fig.5. Generated Monte Carlo signal points with mass regimes overlaid

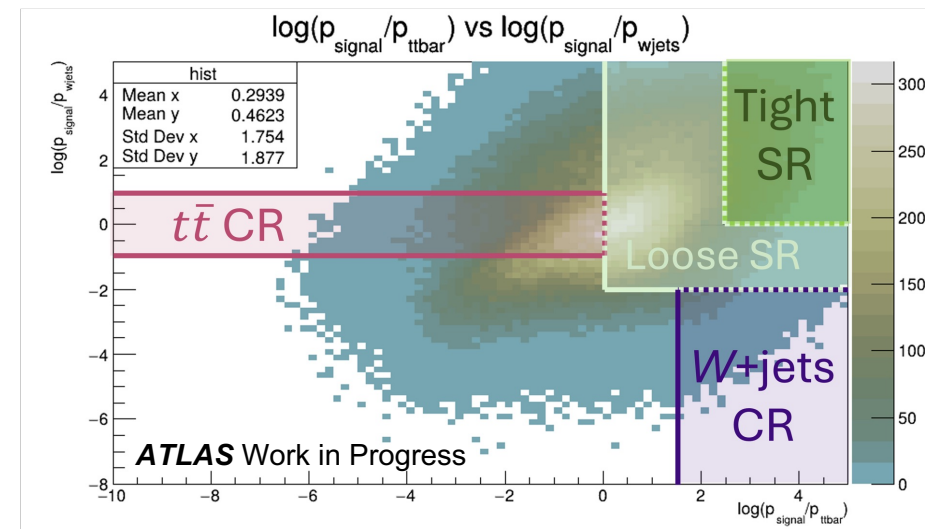
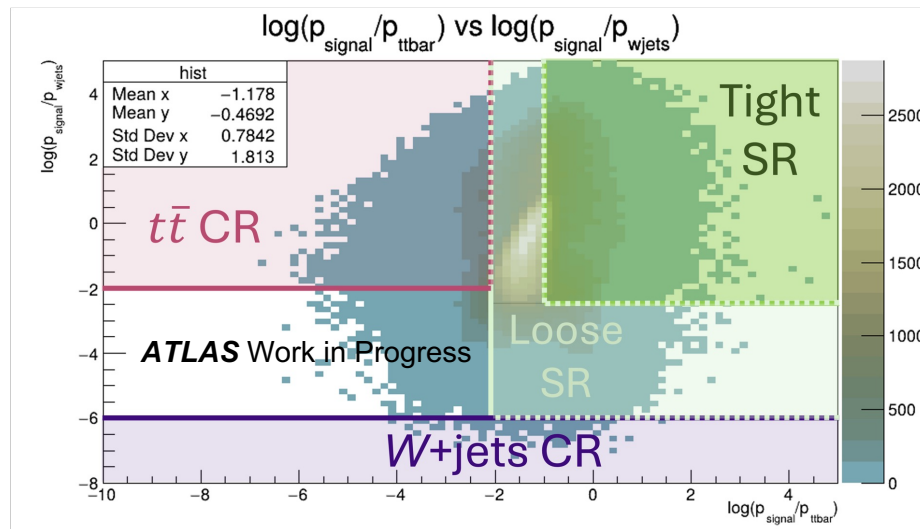
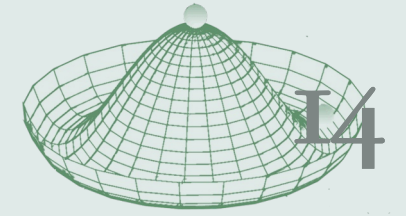



Fig.6 2D distributions of the log ratios of the MVA scores for low and high signal, with the SR/CR definitions overlaid

# Statistical Analysis



 The 2D distribution of  $m_{H^\pm}$  vs  $(m_A - m_{H^\pm})$  in the loose and tight SRs is binned so that the  $t\bar{t}$  background is  $\sim$  flat

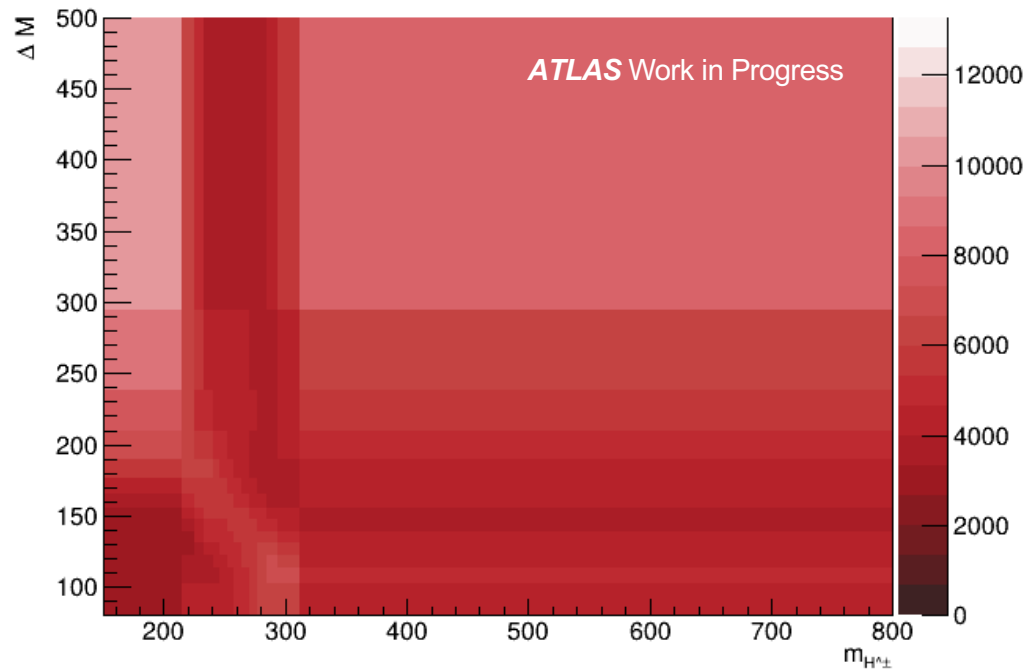


Fig.7a 2D distribution of  $m_{H^\pm}$  vs  $(m_A - m_{H^\pm})$  for  $t\bar{t}$  in the low mass regime tight SR

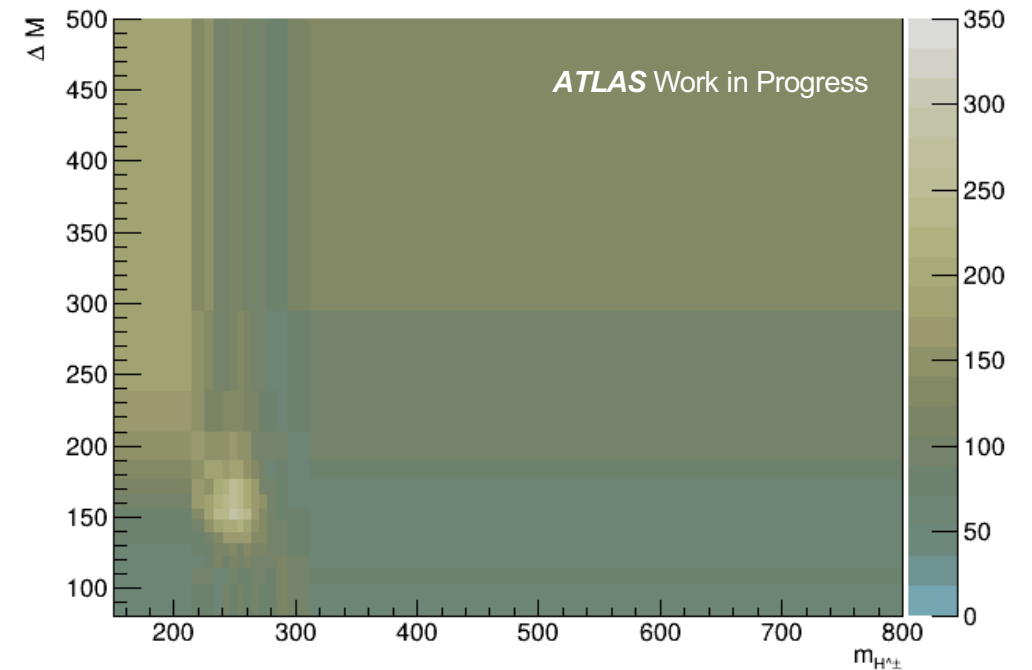
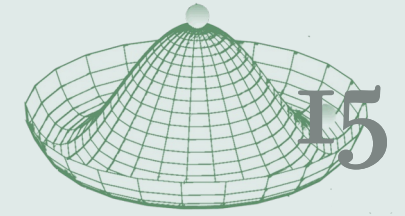


Fig.7b. 2D distribution of  $m_{H^\pm}$  vs  $(m_A - m_{H^\pm})$  for  $(m_A, m_{H^\pm}) = (400, 250)$  in the low mass regime tight SR

# Statistical Analysis



- The 2D distributions are sliced into 1D histograms for each  $\Delta M = (m_A - m_{H^\pm})$  bin
- These 1D histograms are used as signal regions in a binned likelihood fit

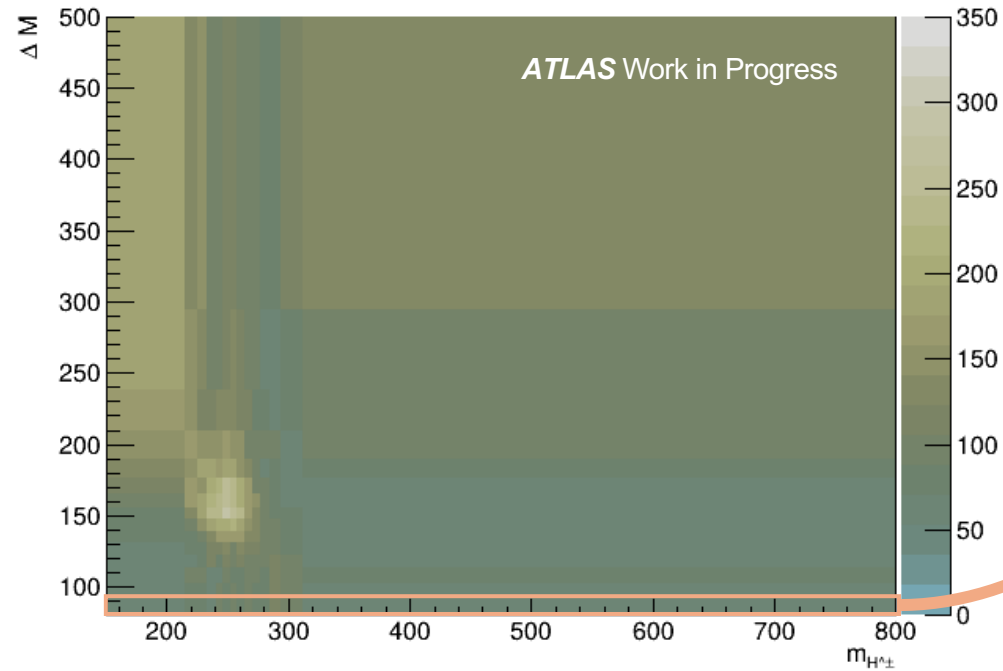


Fig.8a. 2D distribution of  $m_{H^\pm}$  vs  $(m_A - m_{H^\pm})$  for  $(m_A, m_{H^\pm}) = (400, 250)$  in the low mass regime tight SR

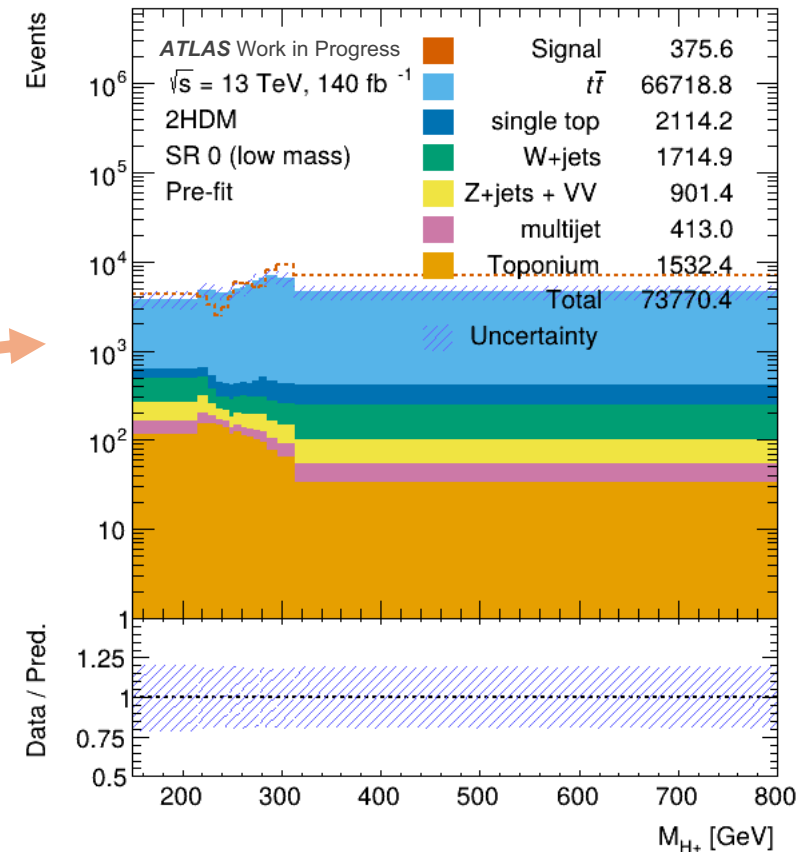
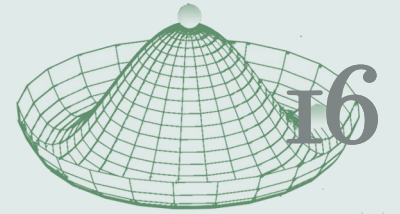


Fig.8b. Signal region for  $(m_A, m_{H^\pm}) = (400, 250)$ , corresponding to the first bin in  $(m_A - m_{H^\pm})$

# Statistical Analysis



  $t\bar{t}$  CR: 1D  $m_A$  distribution to constrain the  $t\bar{t}$  normalization

 W+jets CR: single bin to constrain the W+jets normalization

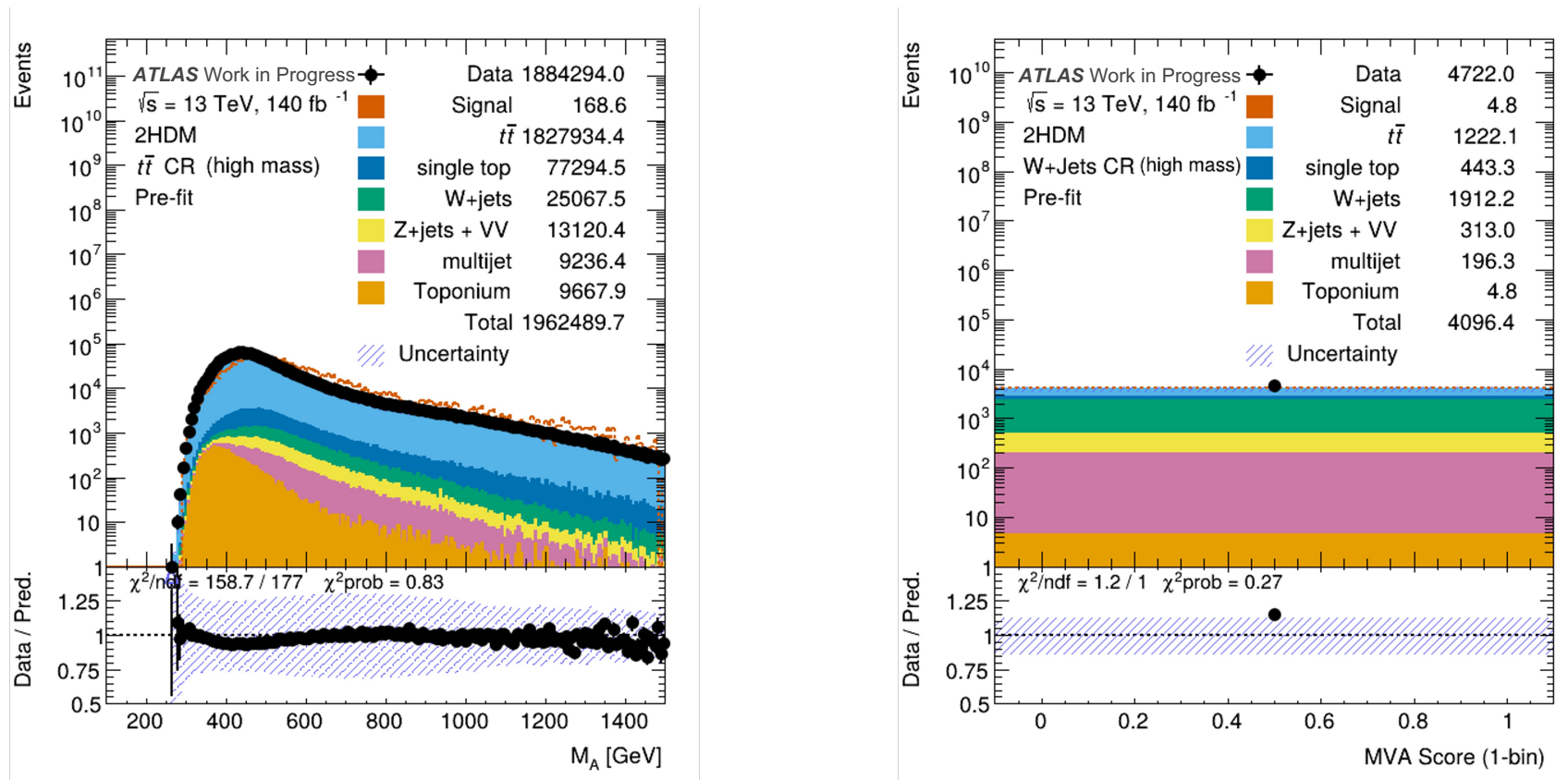
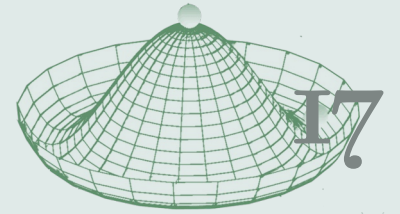


Fig.9  $t\bar{t}$  (left) and W+jets (right) distributions for  $(m_A, m_{H^\pm}) = (700, 500)$  (high mass regime)

# Fit Results



- 95% CL limit results, with blinded SRs (unblinded CRs)
- Majority of our systematics now included
  - We are missing some  $t\bar{t}$  modelling systematics, and electron isolation systematics, but expect the effect to be small
- Sensitivity up to  $m_A \sim 700$  GeV at  $\tan(\beta) = 2$

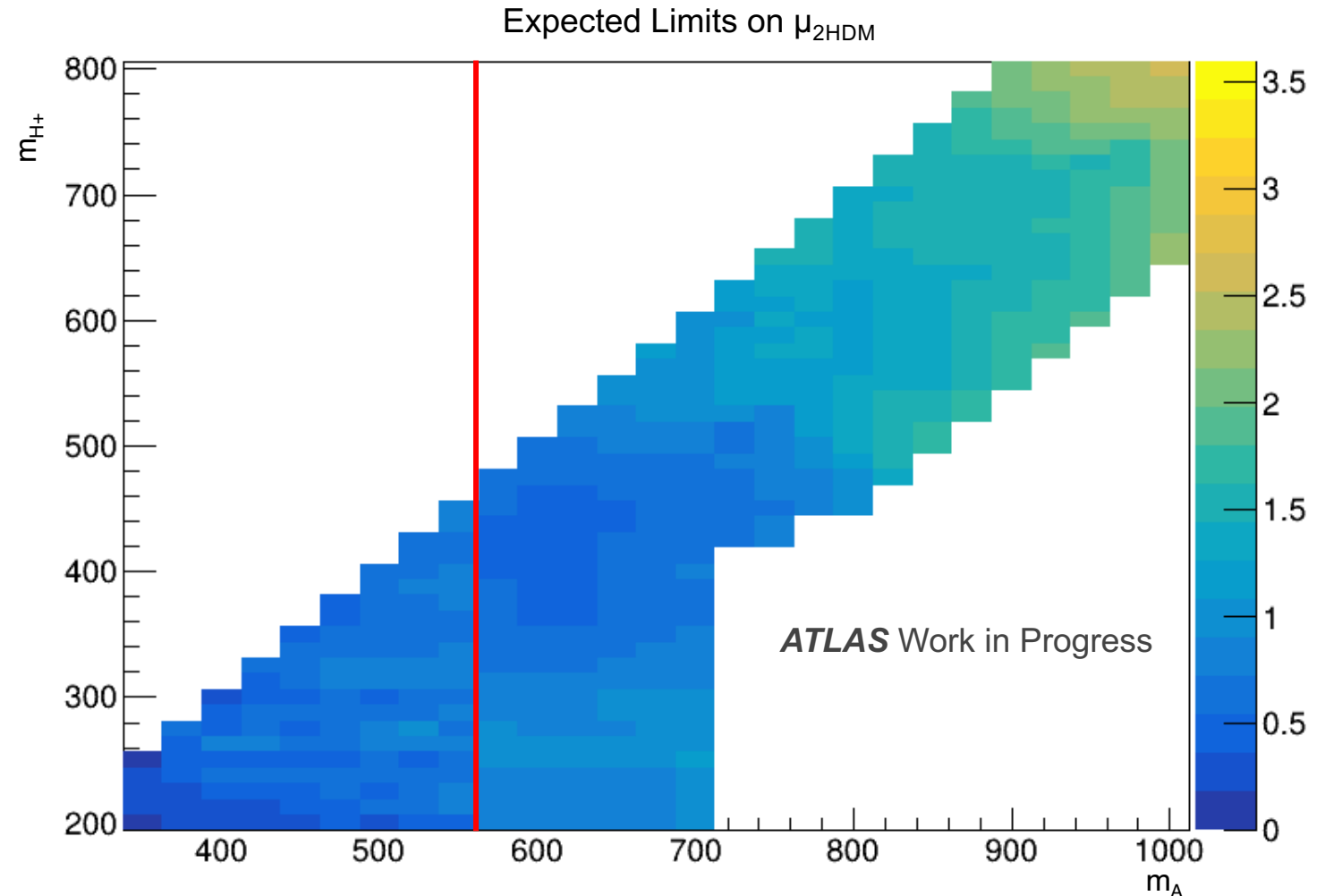


Fig.10. Expected limits on signal strength at  $\tan(\beta) = 2$ , plotted in the  $[m_A, m_{H^+}]$  plane.

# Fit Results

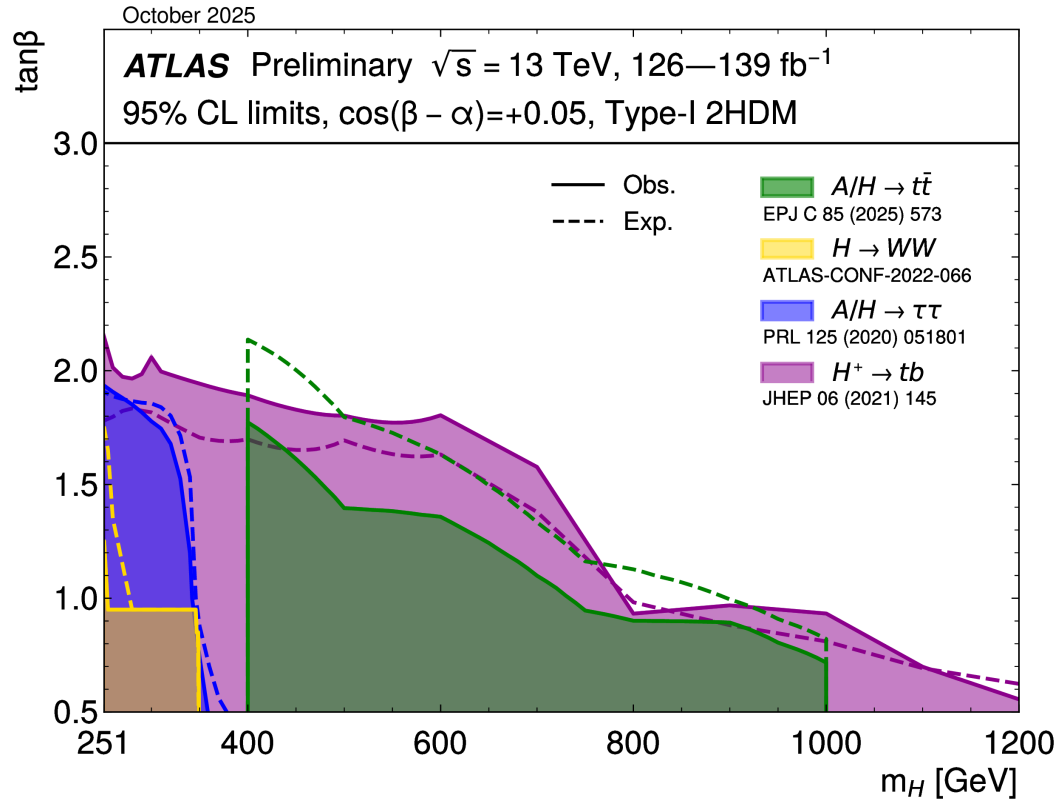
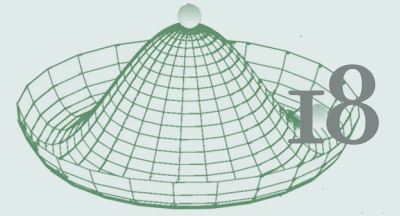


Fig.11. Regions of the  $[m_H, \tan\beta]$  plane excluded in the type-I 2HDM via direct searches for heavy Higgs bosons with  $\cos(\beta - \alpha) = +0.05$  [ATL-PHYS-PUB-2025-042](#)

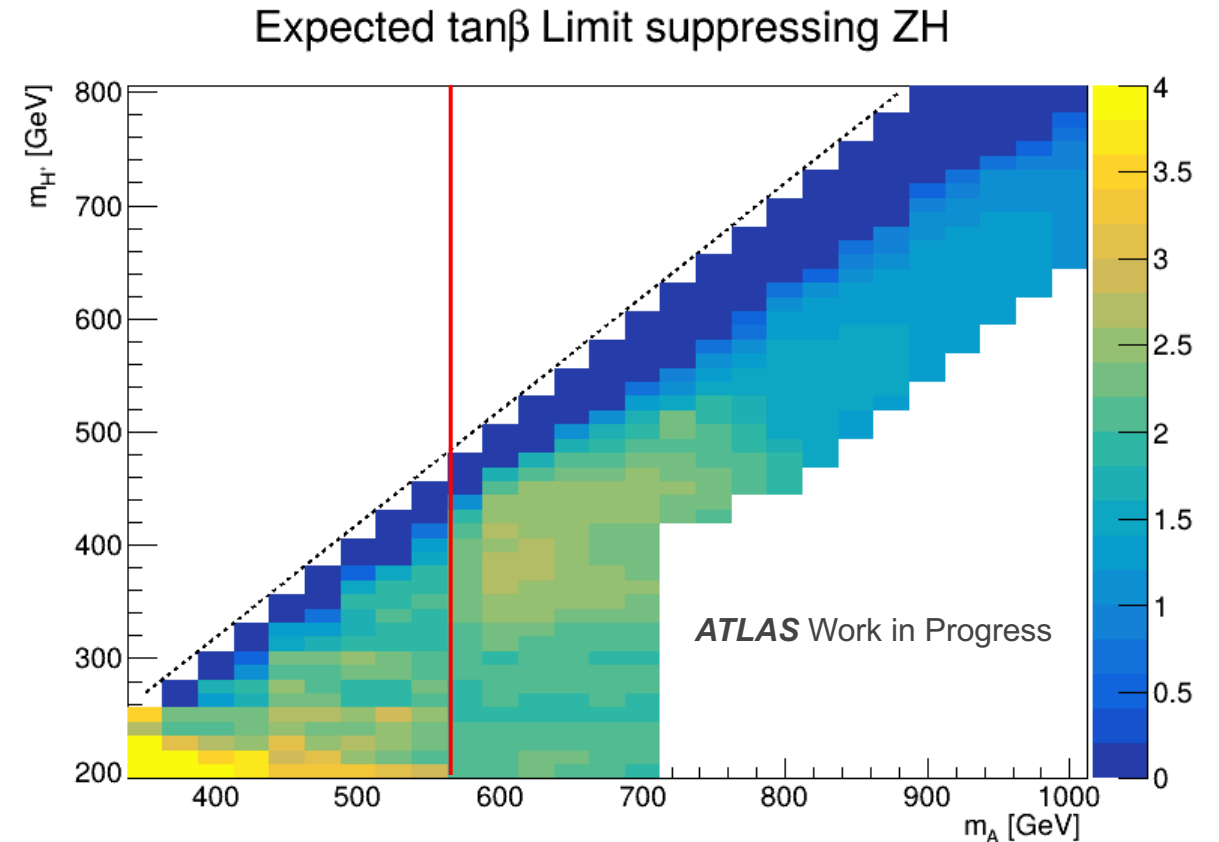
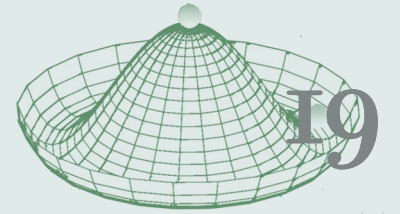


Fig.12. Expected  $\tan\beta$  limits in the type-I 2HDM from this search, plotted in the  $[m_A, m_{H^+}]$  plane.


# Summary



 New search, will either find something (!) or provide new limits for 2HDMs.

 Analysis framework is all in place.

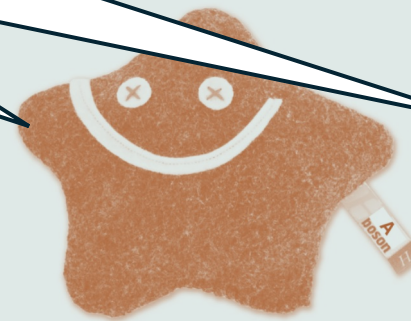
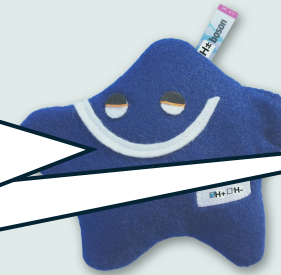
 **To do:**

 Implement final few systematics (don't expect these to be large)

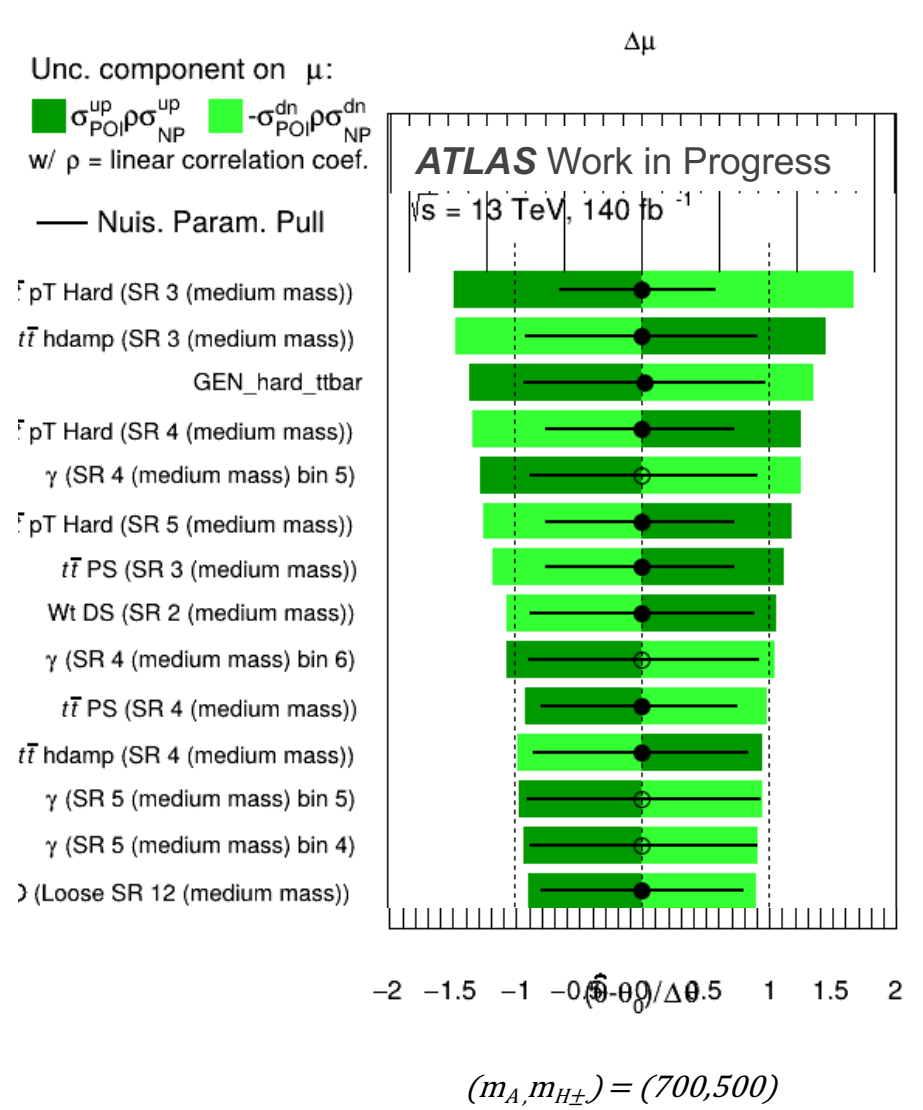
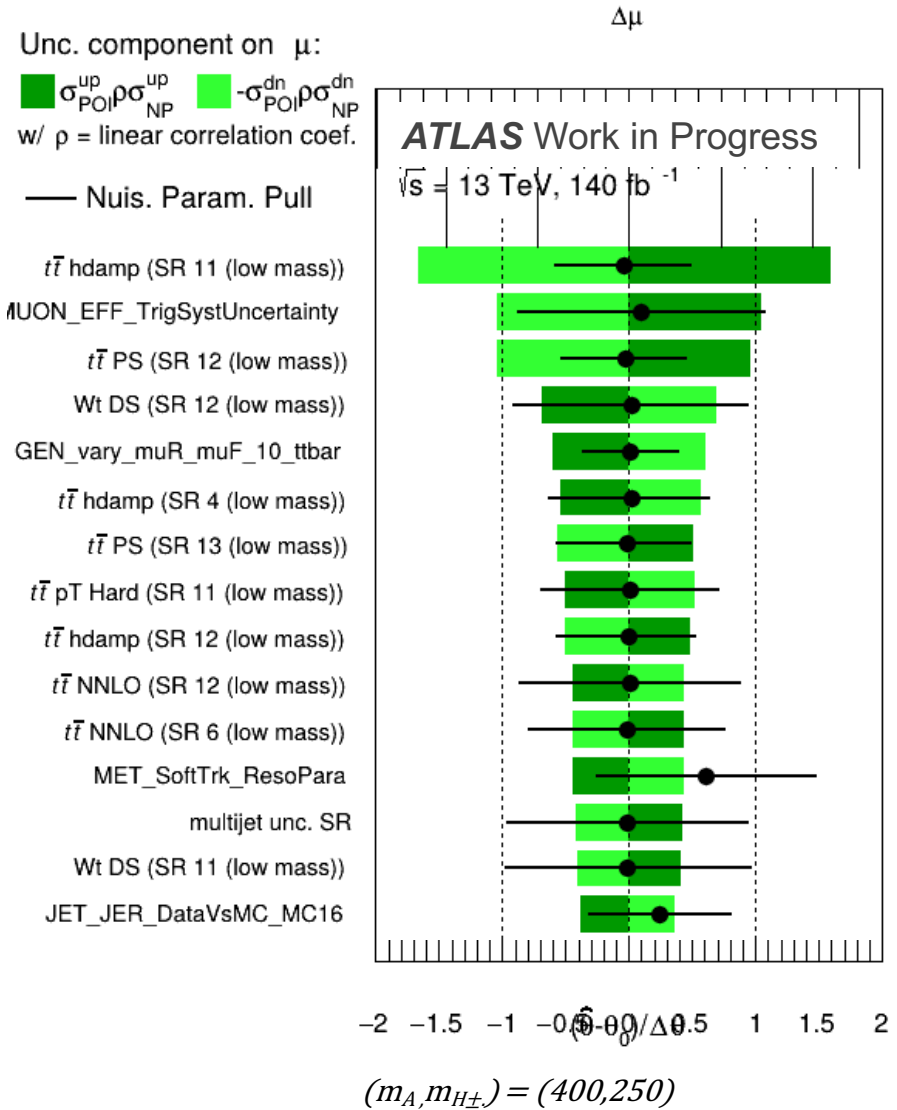
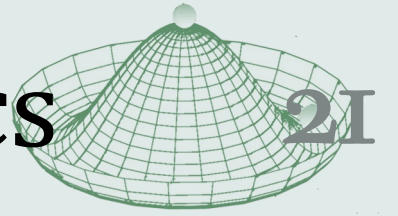
 Unblind SRs



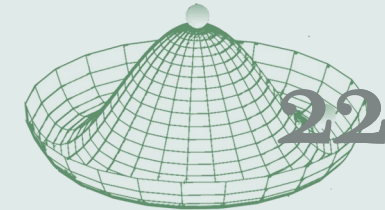
Thank you  
for  
listening!



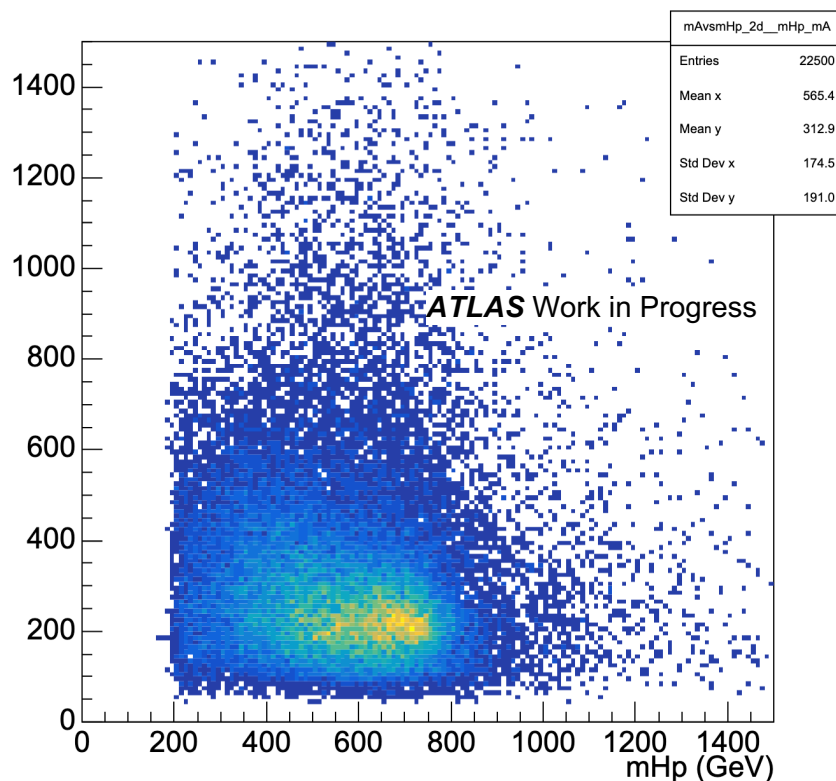
# Backup - Top 15 Ranked Systematics



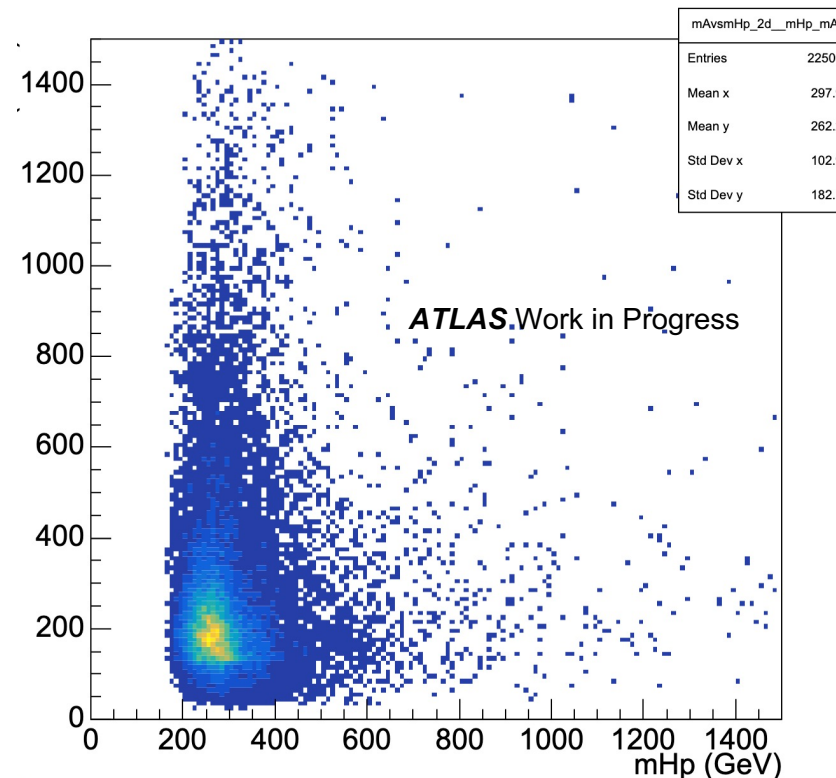
# Backup



- Interpolate between the simulated signal distributions with a morphing framework
- Uses a grid/triangle of simulated 2D distributions and outputs an interpolated distribution between these

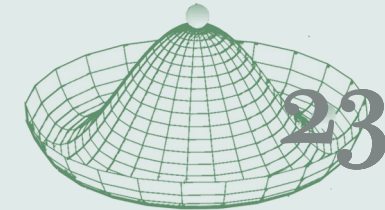


Morphed distribution  $(m_A, m_{H^+}) = (950, 725)$



Morphed distribution  $(m_A, m_{H^+}) = (400, 275)$

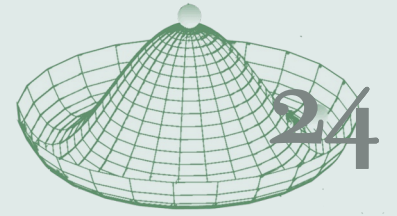
# Backup



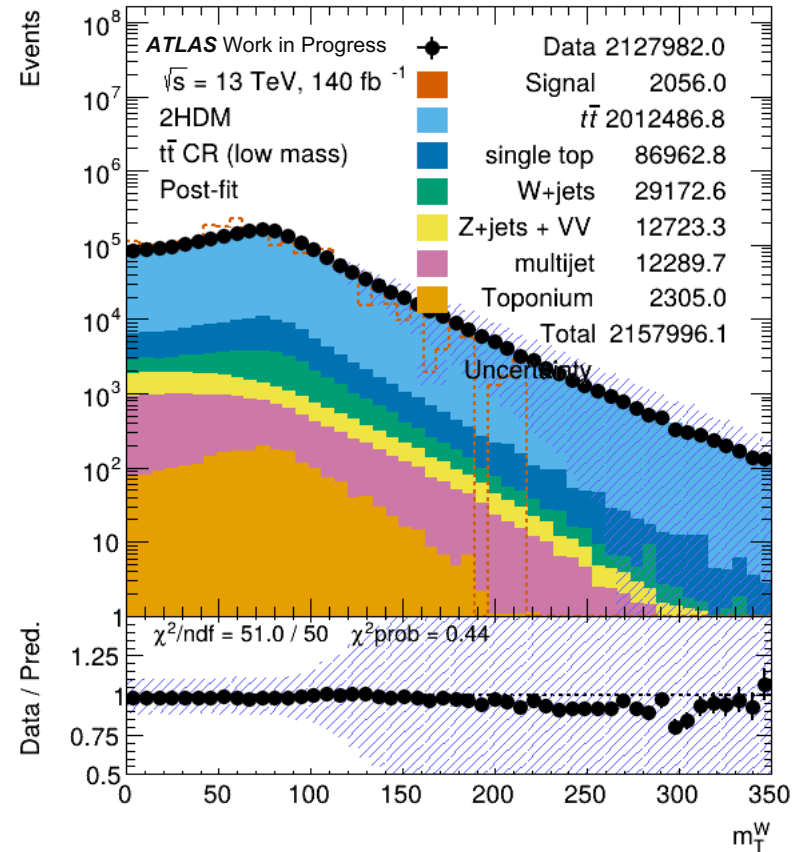
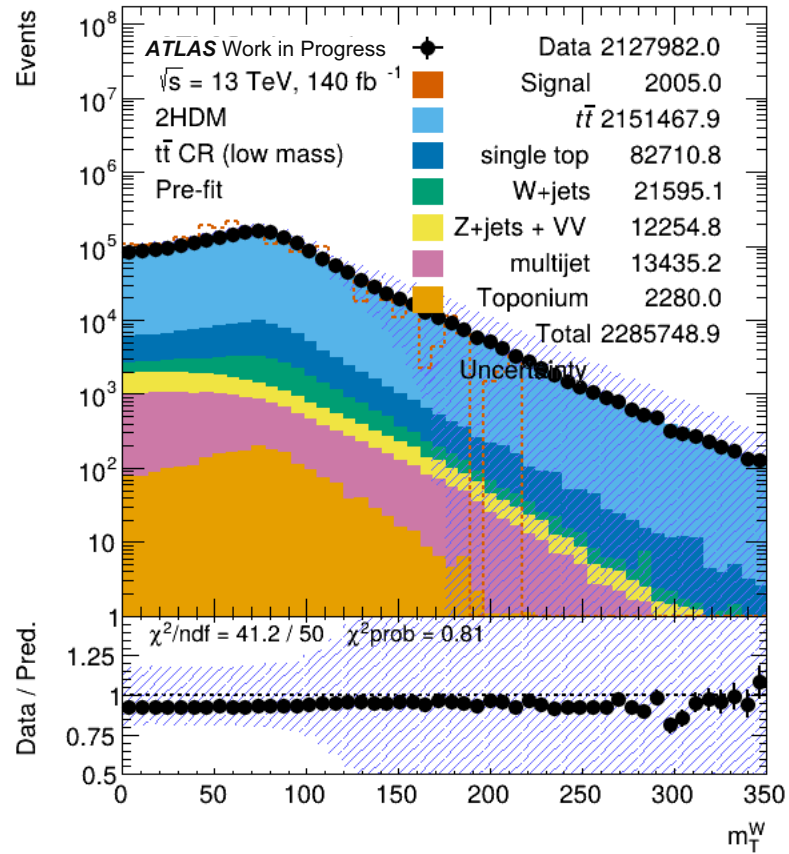
List of variables used in MVA training:

```
features_vec = ['jet_ht_NOSYS', 'A_mass_fitted_tight_NOSYS',  
               'Wb_nonTop_mass_fitted_tight_NOSYS', 'Hp_pt_fitted_tight_NOSYS',  
               'Wb_nonTop_pt_fitted_tight_NOSYS', 'ljet1_pt_NOSYS', 'bjet2_pt_NOSYS',  
               'A_pt_fitted_tight_NOSYS', 'WfromH_pt_fitted_tight_NOSYS',  
               'top_pt_fitted_tight_NOSYS', 'bjet1_pt_NOSYS',  
               'WfromTop_pt_fitted_tight_NOSYS', 'ljet2_pt_NOSYS',  
               'chi2_fitted_tight_NOSYS', 'FoxWolfram_Moment1_NOSYS', 'centrality_NOSYS',  
               'bjet1_phi_NOSYS', 'bjet2_phi_NOSYS', 'dR_bjet1_bjet2_NOSYS', 'dR_bjet2_ljet1',  
               'mWt_NOSYS', 'lepton_pt_NOSYS', "cos_theta_star_A_top_tight_NOSYS",  
               "cos_theta_star lep_Wlep_tight_NOSYS", "cos_theta_star_Whad_ljet1_tight_NOSYS",  
               "dPhi_lep_MET_NOSYS",  
               "Hp_mass_fitted_tight_NOSYS", "MET_NOSYS", "min_dR_bjet1_ljet_NOSYS",  
               "min_dR_bjet2_ljet_NOSYS",  
               "ratio_ptLep_WfromHMass_NOSYS"]
```

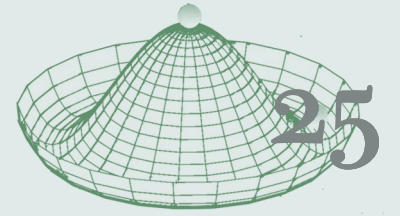
# Backup



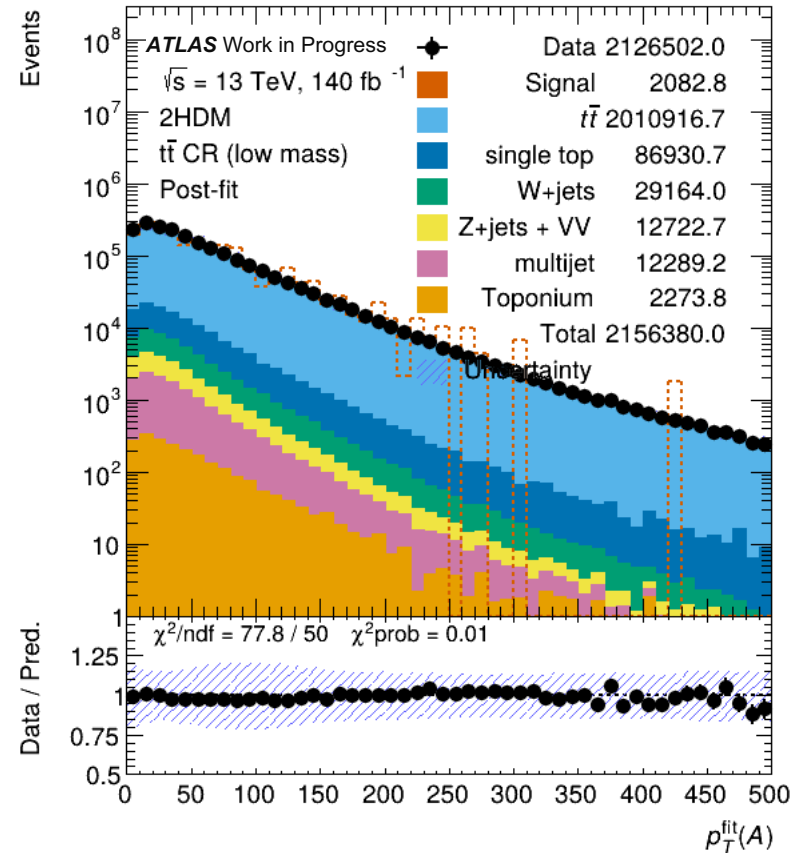
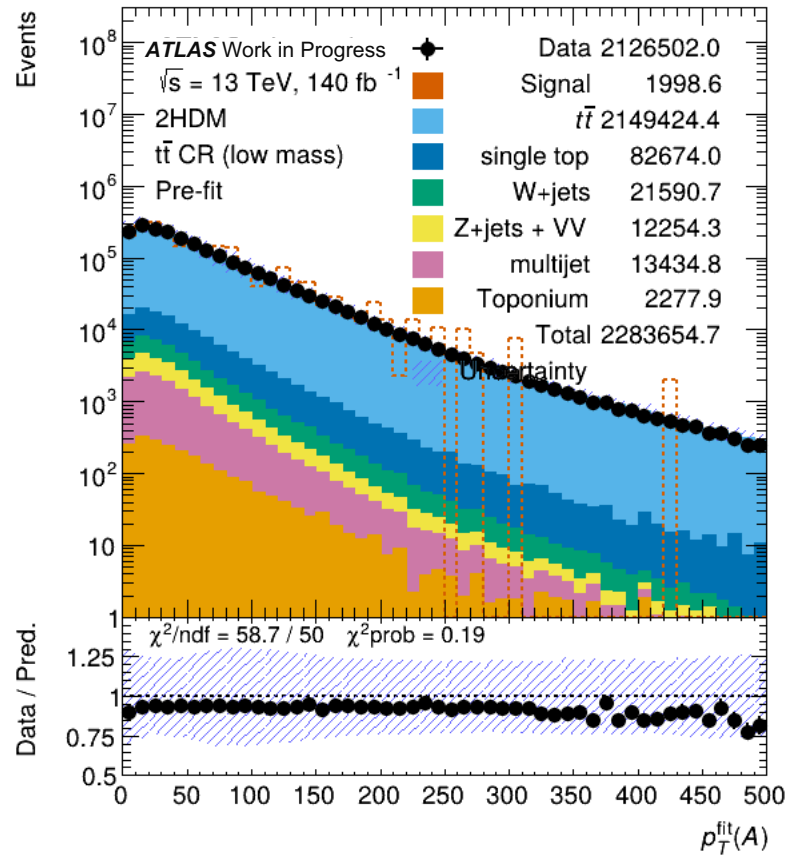
Some kinematic variables used in MVA training, in the low mass  $t\bar{t}$  control region, pre-fit and post-fit:



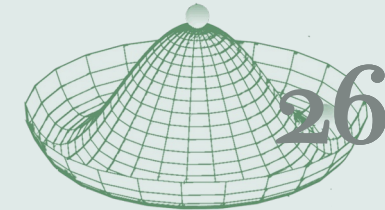
# Backup



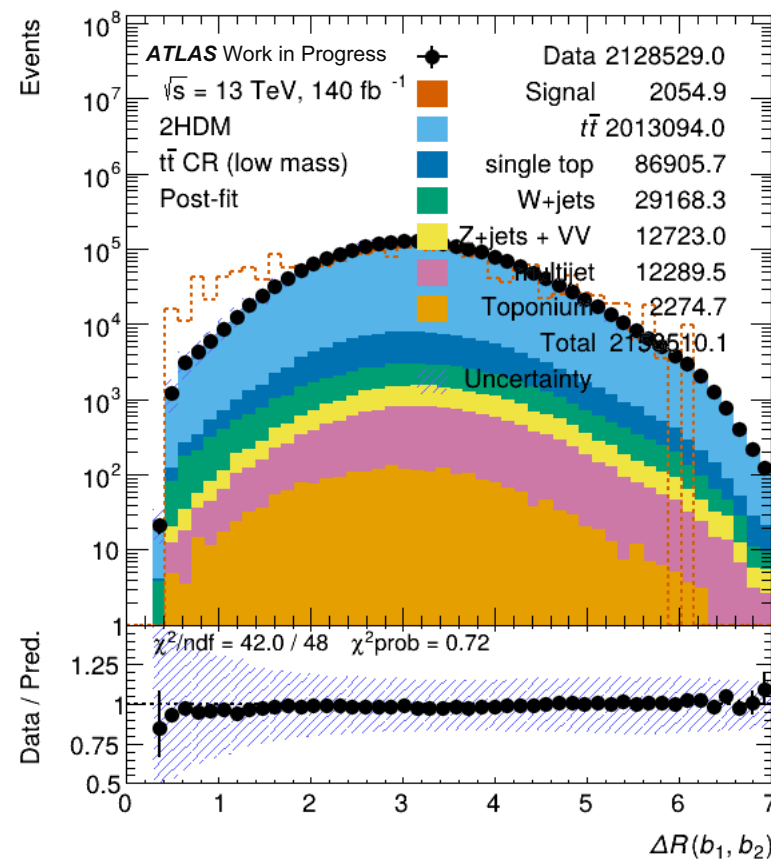
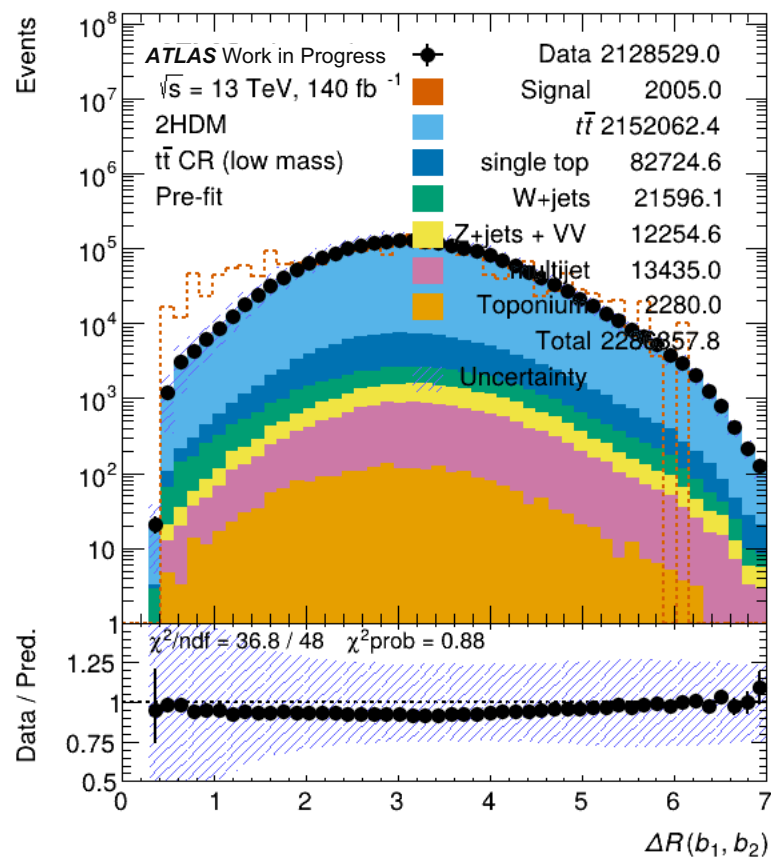
Some kinematic variables used in MVA training, in the low mass  $t\bar{t}$  control region, pre-fit and post-fit:



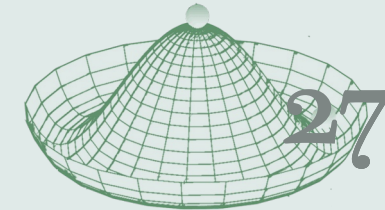
# Backup



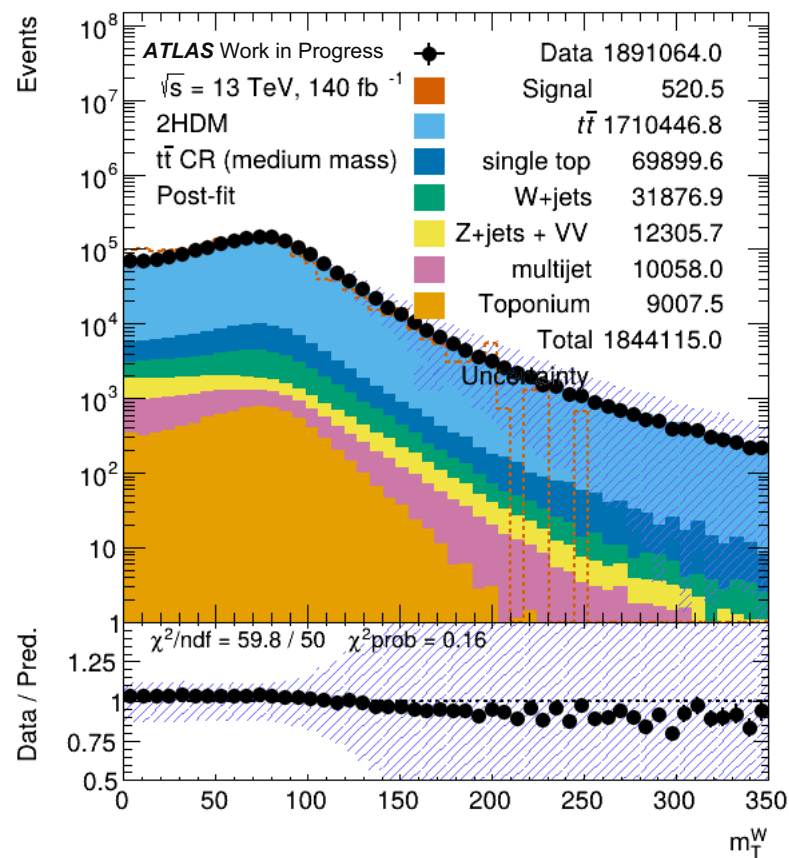
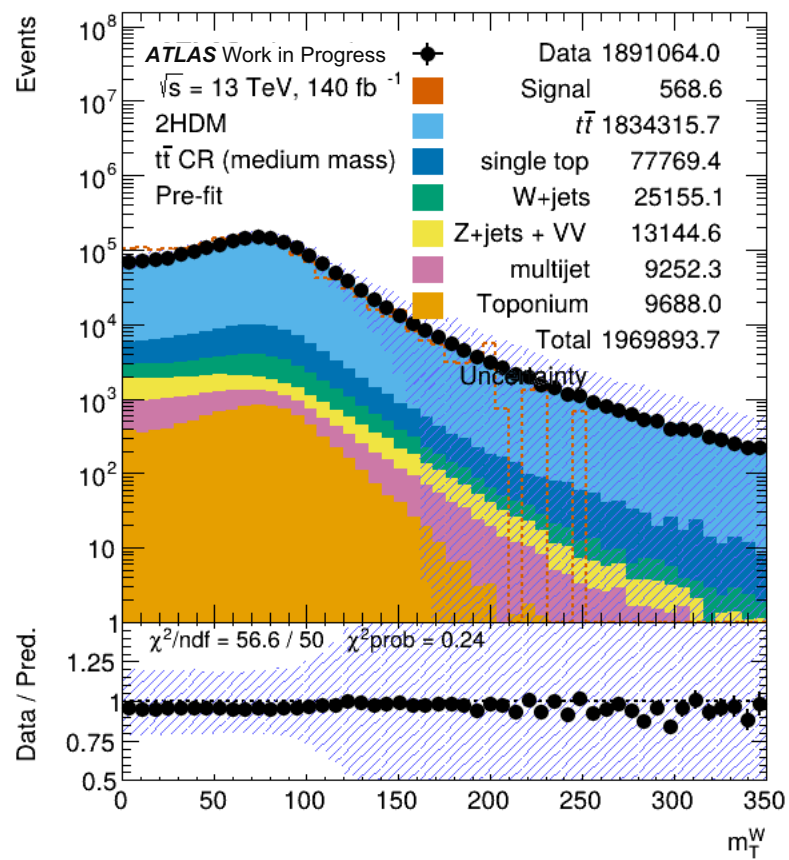
Some kinematic variables used in MVA training, in the low mass  $t\bar{t}$  control region, pre-fit and post-fit:



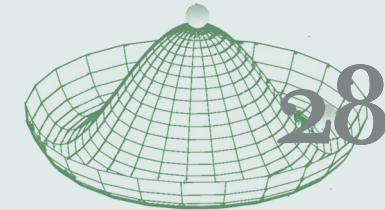
# Backup



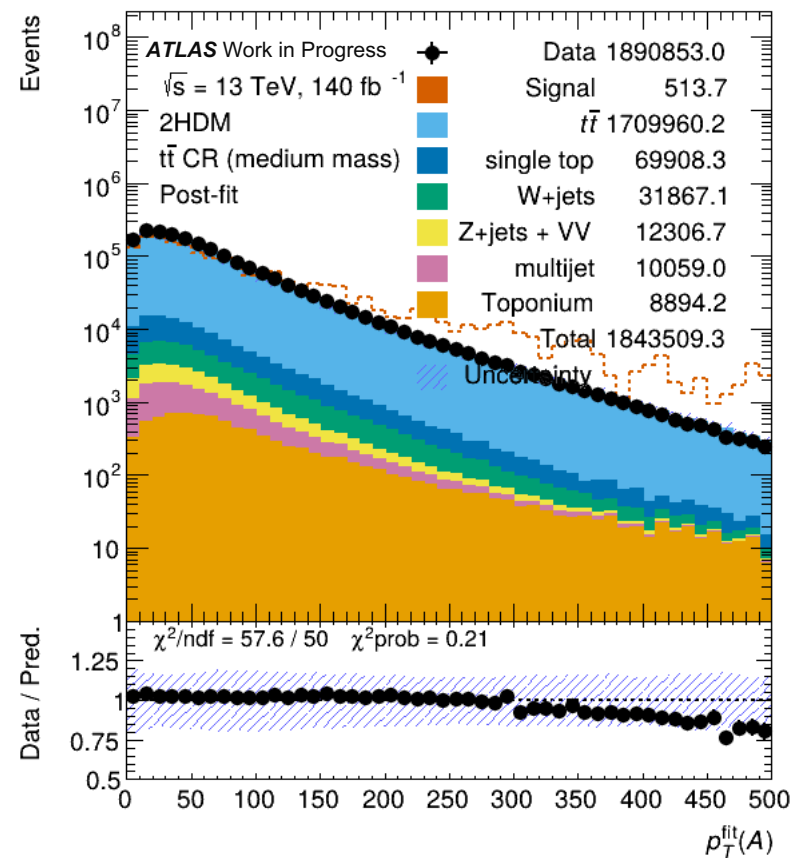
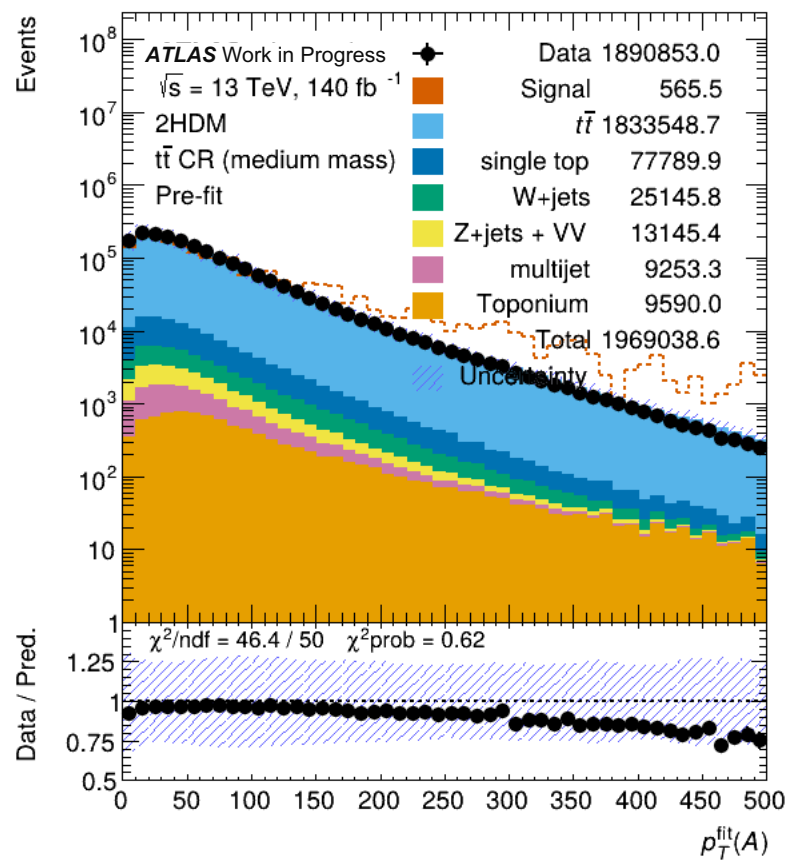
Some kinematic variables used in MVA training, in the high mass  $t\bar{t}$  control region, pre-fit and post-fit:



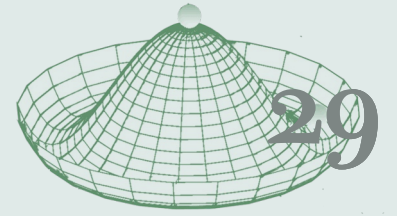
# Backup



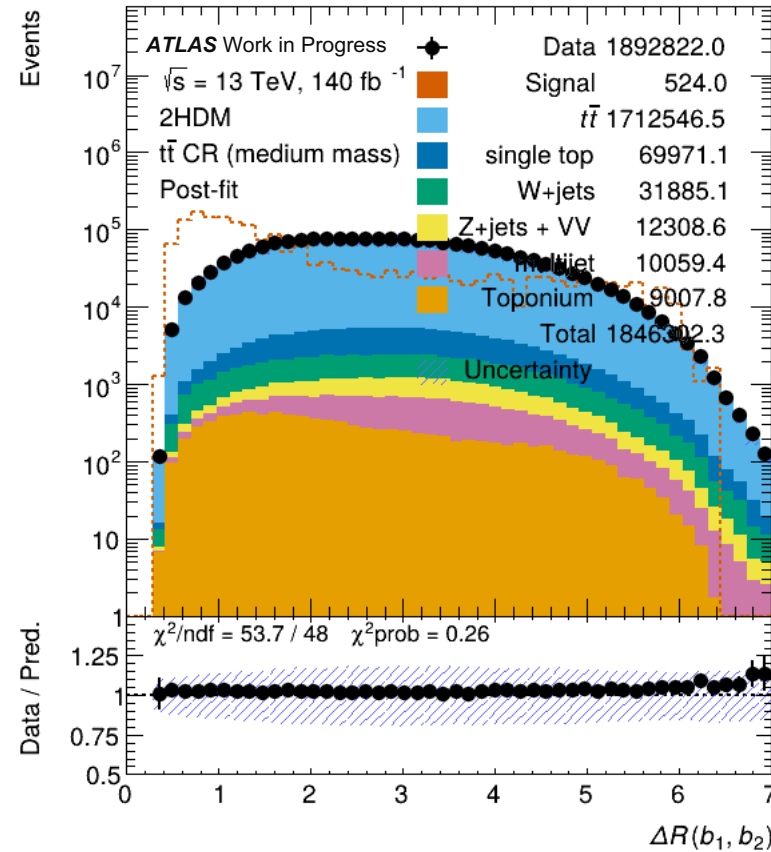
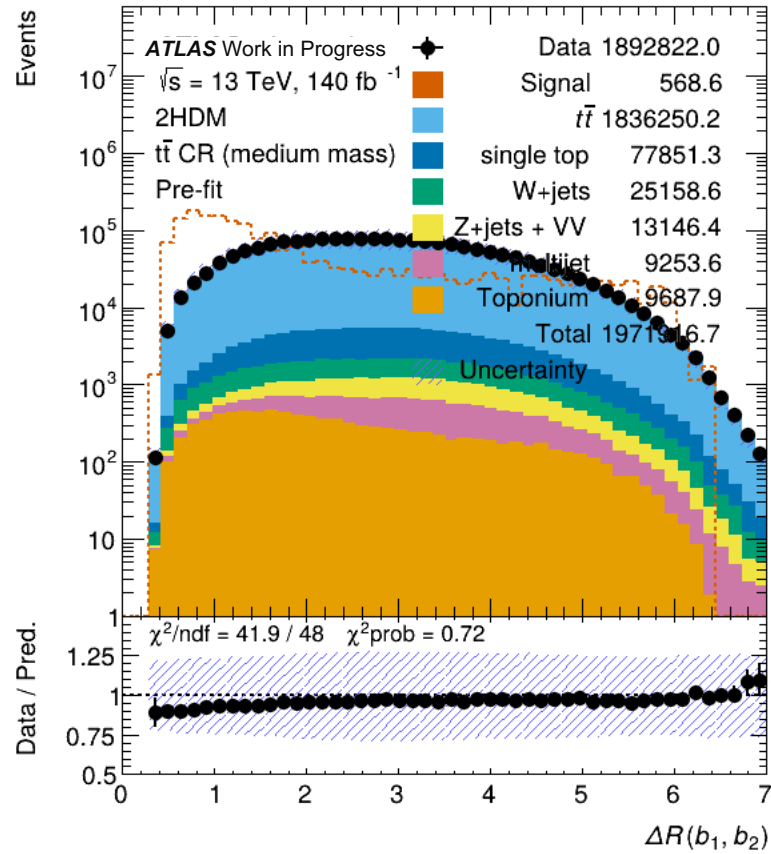
Some kinematic variables used in MVA training, in the high mass  $t\bar{t}$  control region, pre-fit and post-fit:



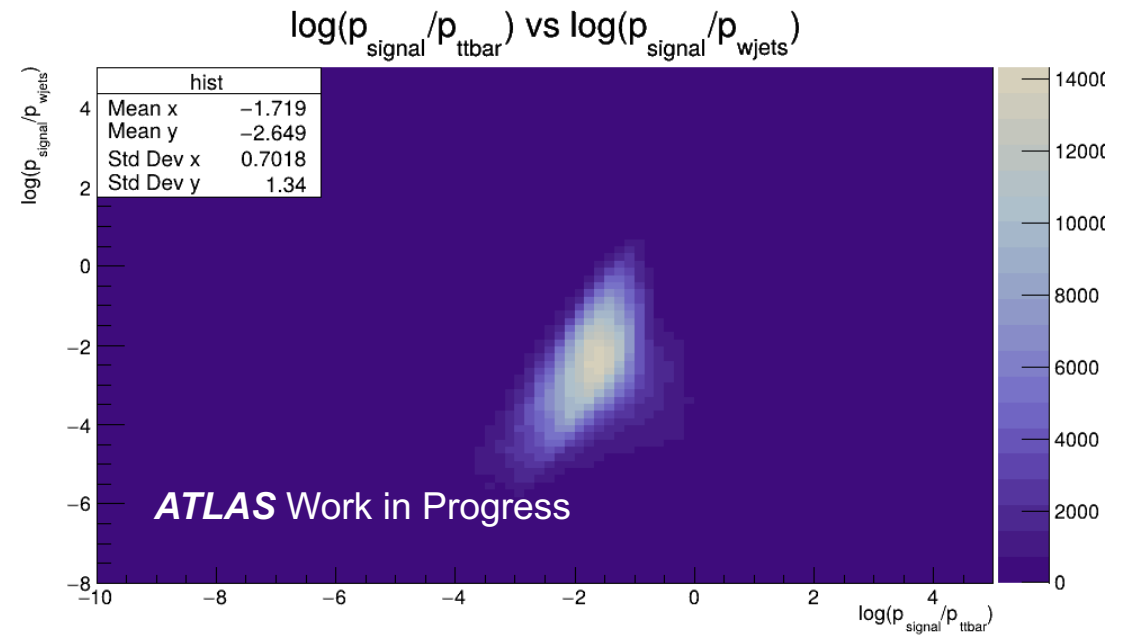
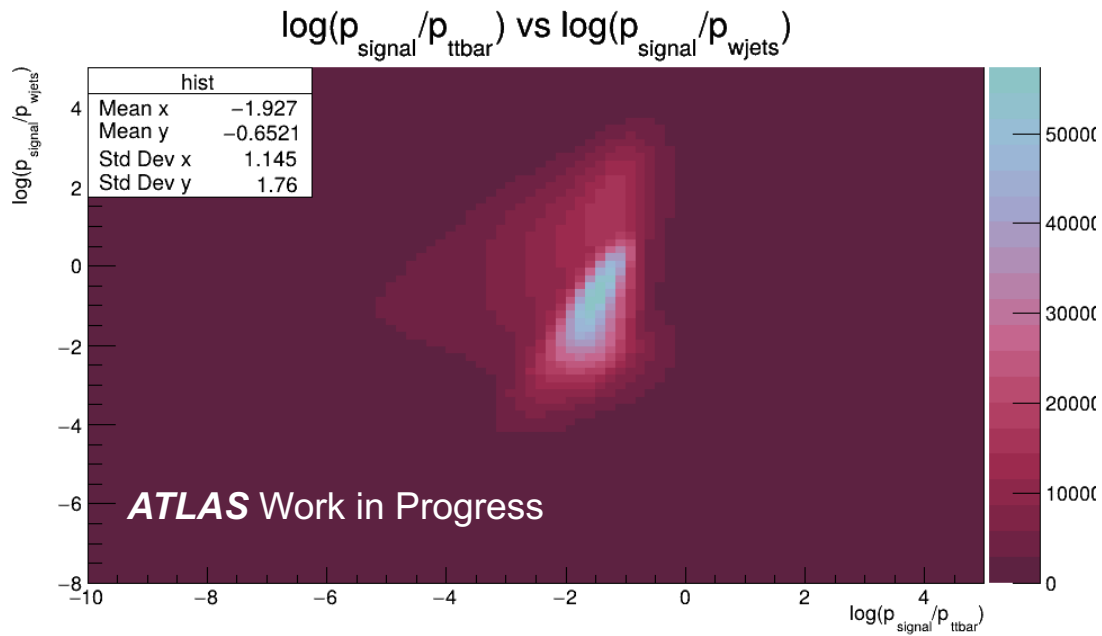
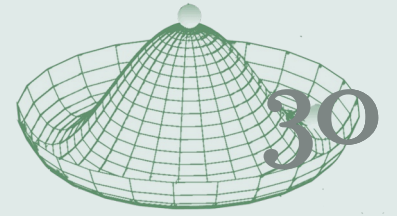
# Backup



Some kinematic variables used in MVA training, in the high mass  $t\bar{t}$  control region, pre-fit and post-fit:

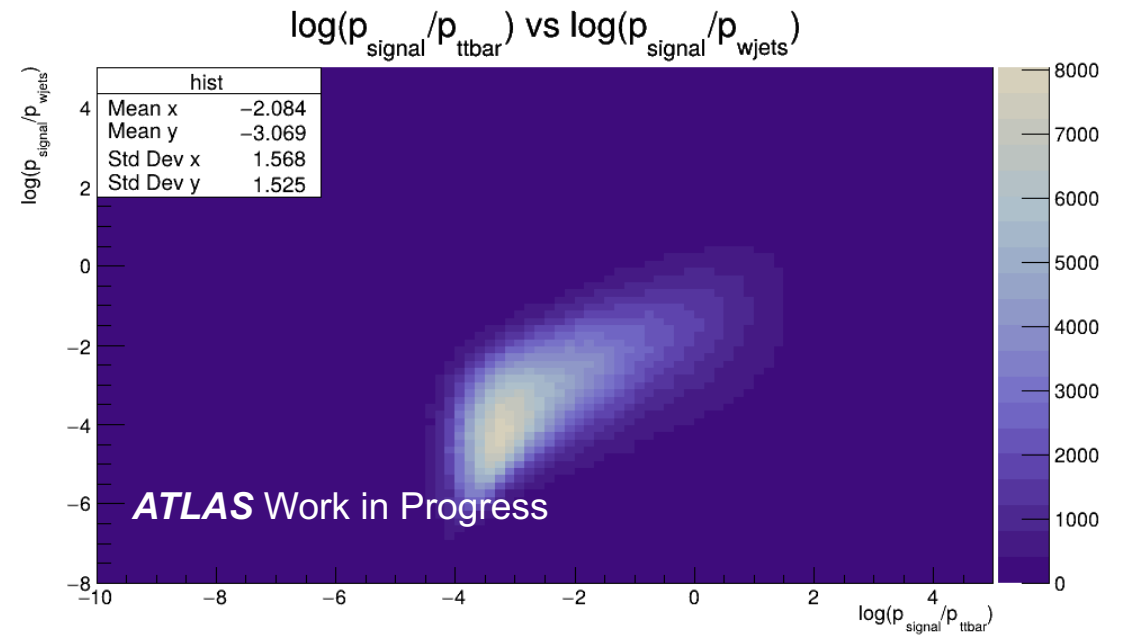
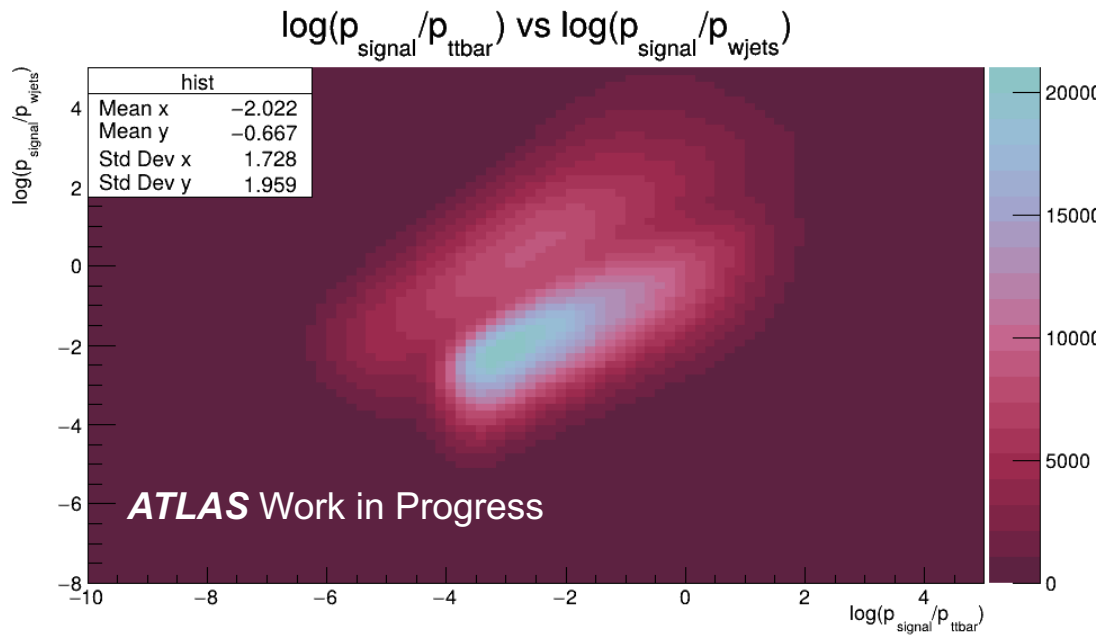
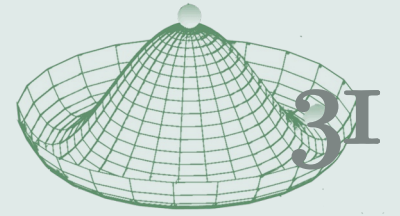


# Backup



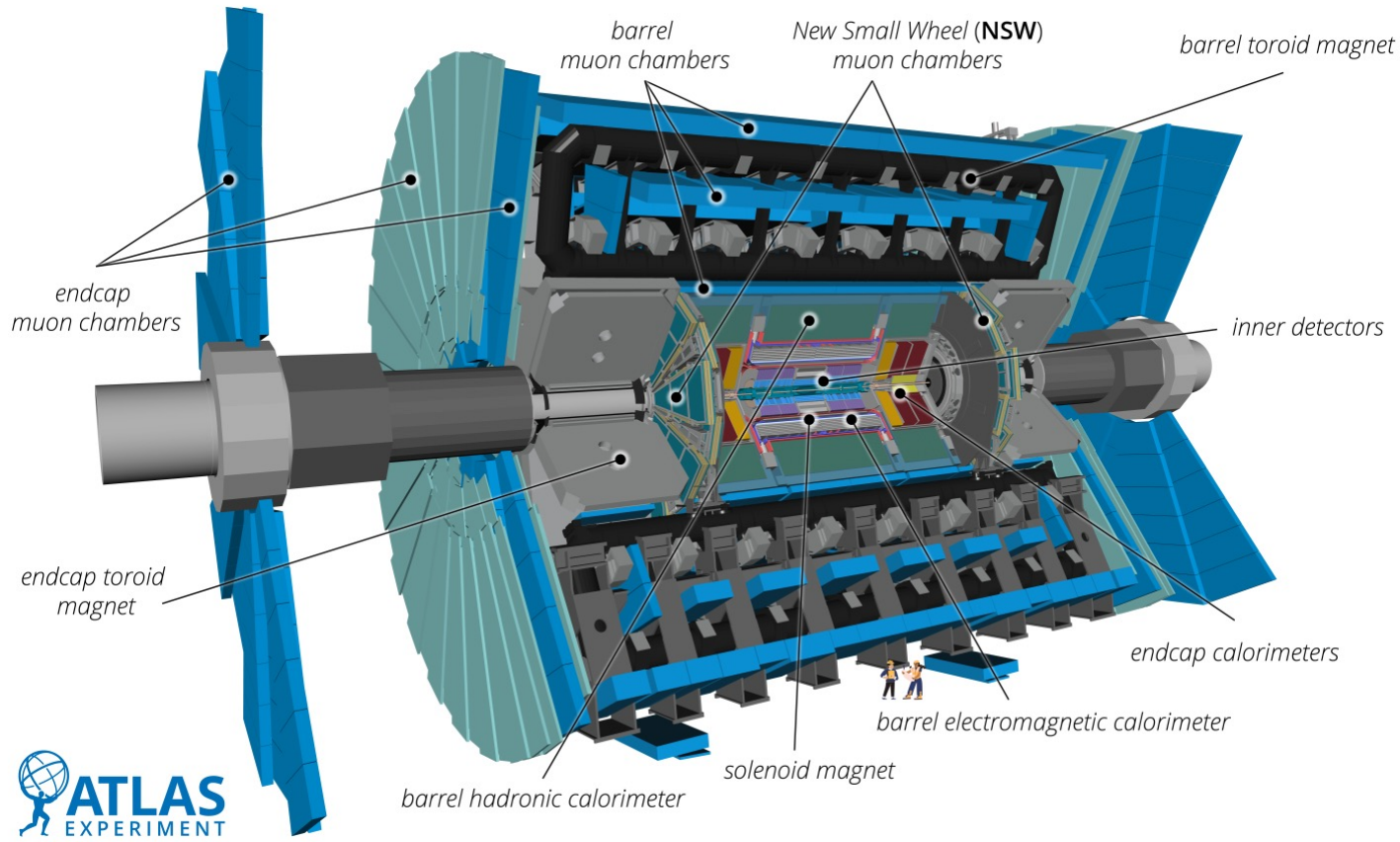
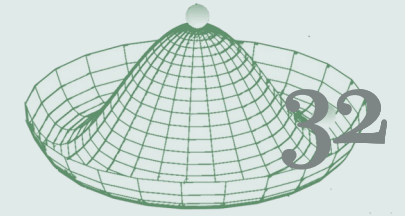
*2D distributions of the log ratios of the MVA scores for a) ttbar and b) W+jets, in the low mass range.*

# Backup



*2D distributions of the log ratios of the MVA scores for a) ttbar and b) W+jets, in the high mass range.*

# Backup



- 🌀 LHC bunch crossing rate  $\sim 40$  MHz
- 🌀 Up to 50 proton-proton collisions per bunch crossing in Run 2
- 🌀 Collision events recorded by ATLAS
  - 🌀 L1 hardware trigger
  - 🌀 HLT software trigger

