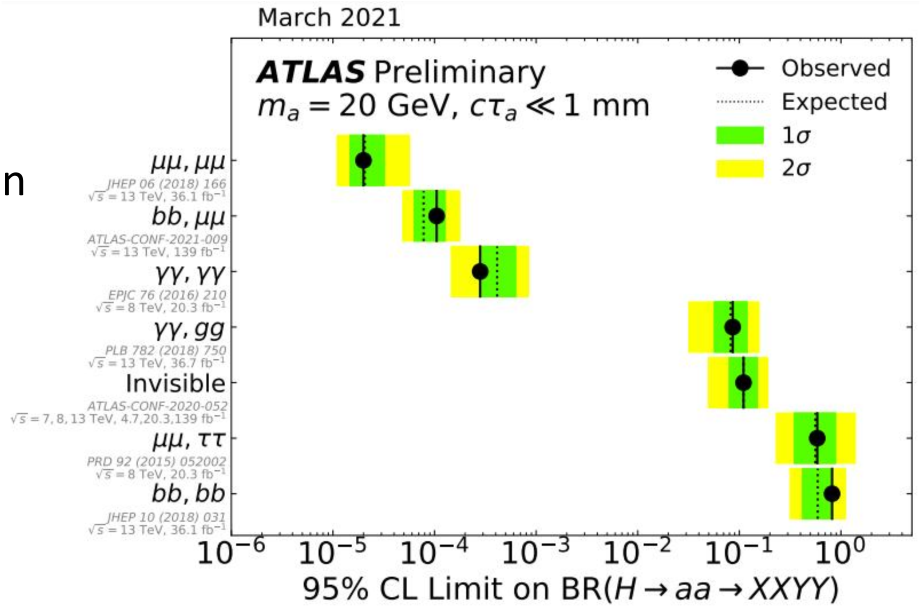
Search for Higgs Boson Decays to a Z Boson and a Light Hadronically Decaying Resonance

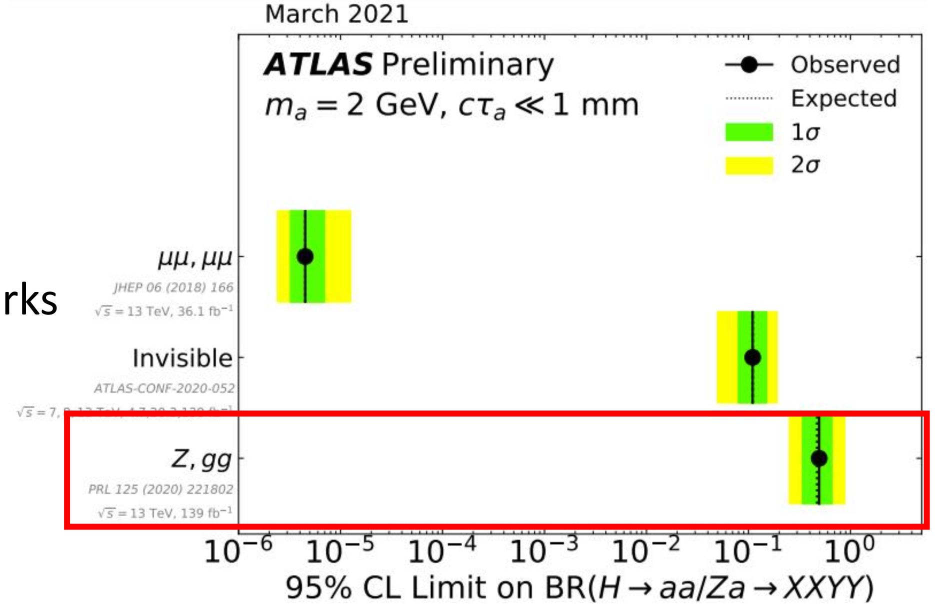
Chonghao Wu

Joint APP, HEPP and NP IOP conference 2024

Motivation

- **New pseudoscalar (a) predicted by various BSM**
  - a could have a large coupling to the observed Higgs boson
  - a can decay to SM particles(qq, gg, γγ, ll)
- **Axions**
  - Solving the strong CP problem
  - Axion-like particle (ALP): more general particles sharing some properties with axions
- **Extensions to the SM Higgs sector**
  - Motivated by SUSY, CP problem in QCD...
  - 2HDM(+S): The two Higgs-doublet model (with an additional scalar singlet)
- **Current searches**
  - Mostly focusing on decays of a to leptons or heavier quarks
  - Few searches for H⟶Za
- **Previous round of analysis published in 2020**
  - [Phys. Rev. Lett. 125 (2020) 221802]
  - Analysis sensitivity limited by background MC statistics

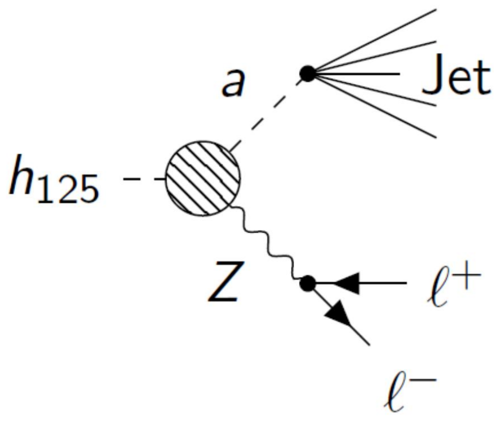




ATL-PHYS-PUB-2021-008

Aim

- **Search for a light resonance (m < 4 GeV) produced through H→Za**
  - a ⟶ hadronic decay, reconstructed as a single jet
  - Z ⟶ leptonic decay
  - Background: mainly from Z + jets



- Full Run 2 data from ATLAS
- Higher statistics POWHEG Z+jets MC samples instead of SHERPA
- Dedicated neural network for background reweighting
- New strategy for systematic uncertainty analysis

Strategy

previous: 3-D histograms

11-D Neural Network

Pre-selection

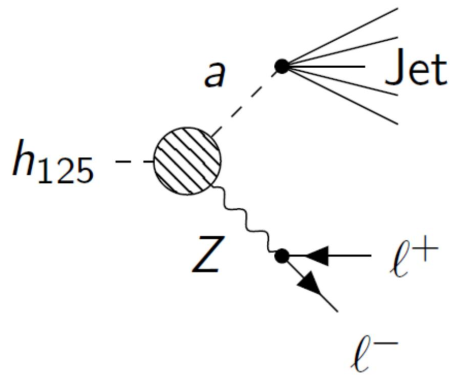
Background reweighting

Regression NN

Classification NN

Statistical analysis

Exclusion limits



>2 tracks

Same-Flavor Opposite-Sign lepton pair

a mass prediction

Discriminate signal from bkg

Background Reweighting

- Use high statistics POWHEG Z+jets MC samples instead of Sherpa
  - Powheg is less accurate for Z+jets than Sherpa, but it has many more events

- Reweight the bkg to match the data, improve the modelling of event kinematics and jet variables

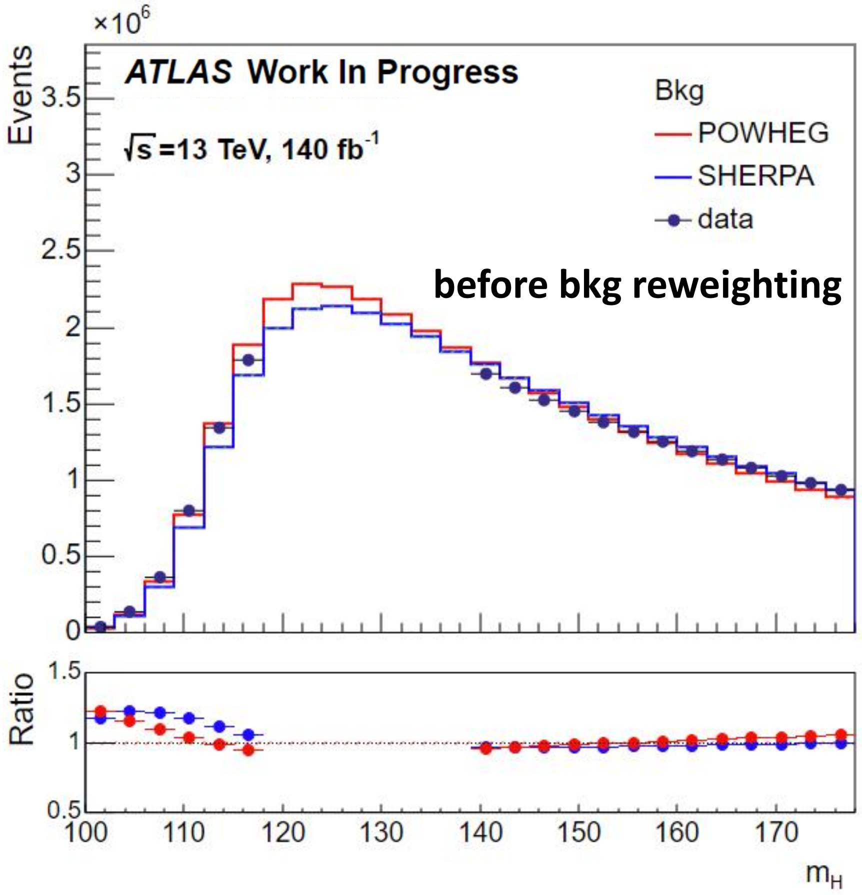
Neural Network

- The aim of NN is to estimate the ratio of data to bkg probability density functions: $r(X)=f_{data}(X)/f_{bkg}(X)$

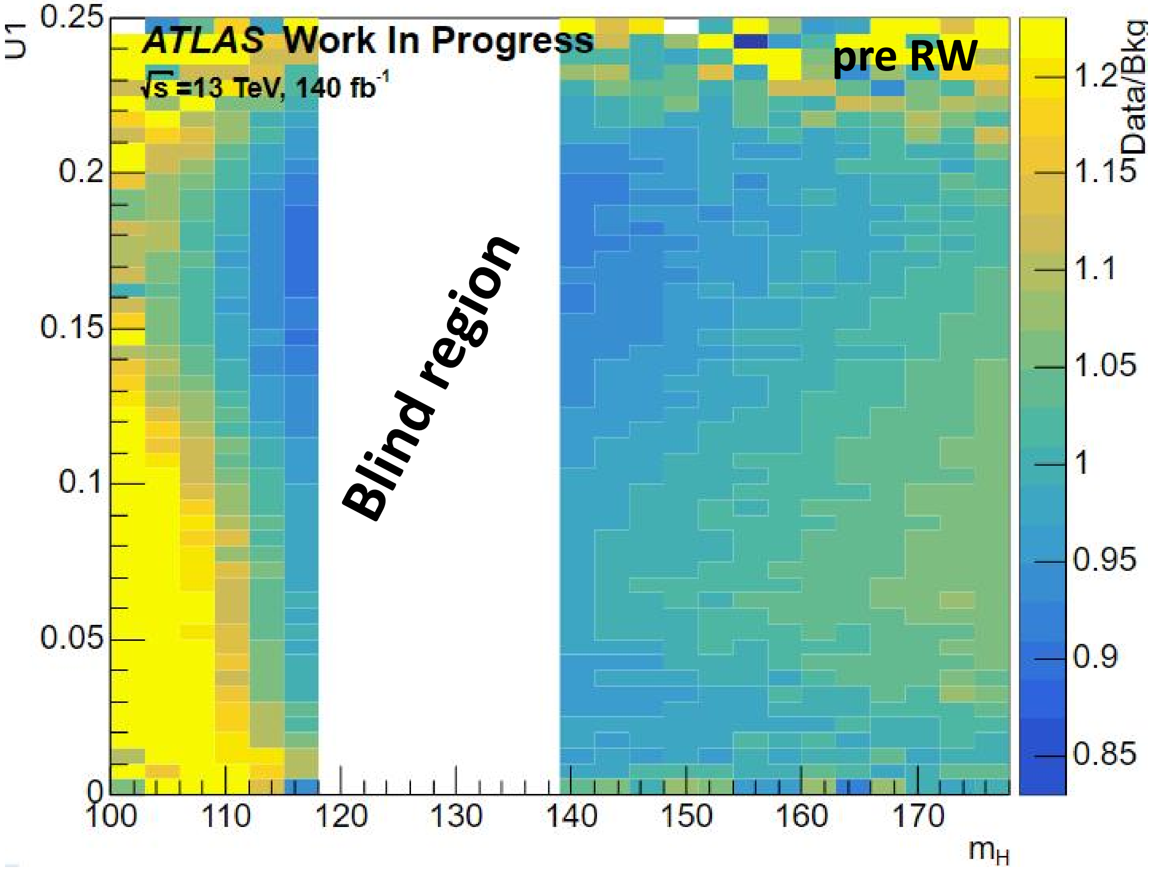
- Dedicated loss function for Log-Likelihood Ratio Estimation, requires no knowledge of pdfs of bkg and data(**arXiv:1911.00405 (2019)**)

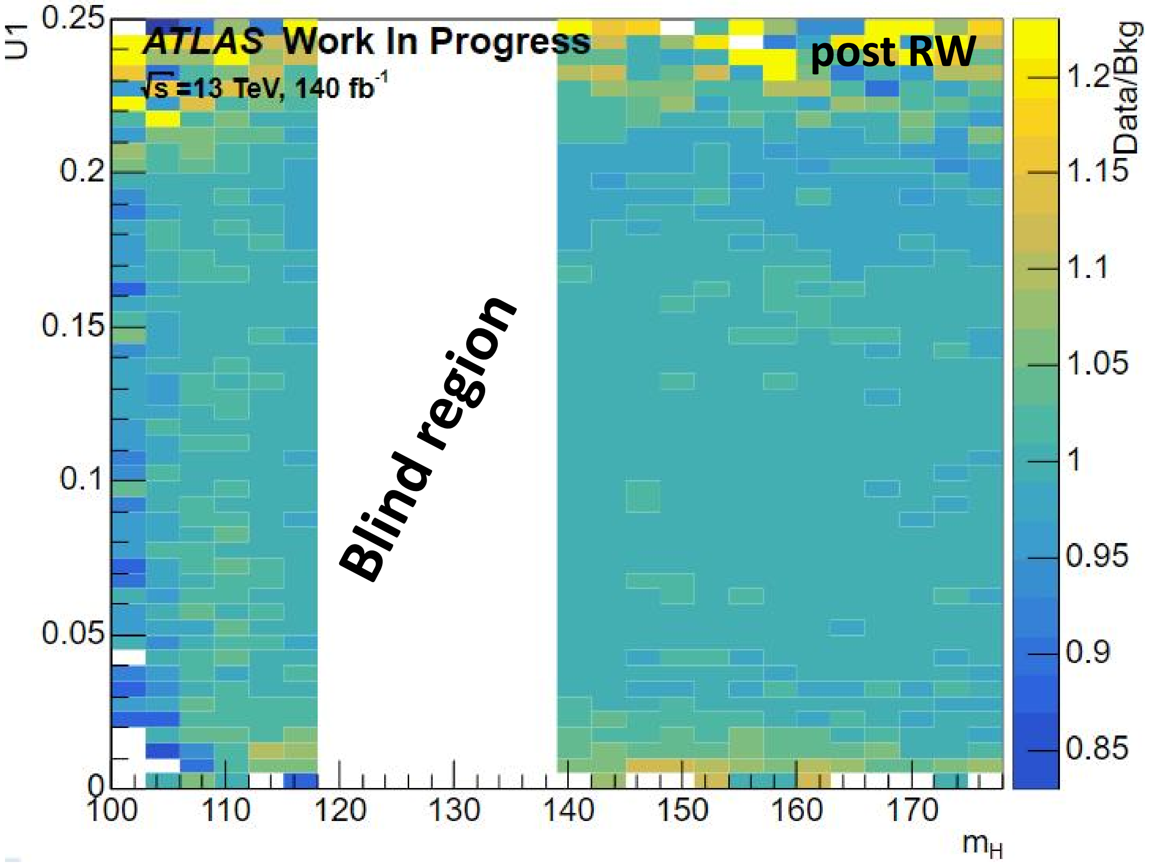
- Blind region: 120 GeV < m_H < 140 GeV, to avoid bias from possible signals

- Training variables: 11 in total.
  - Final state invariant mass, kinematic variables, 6 jet substructure variables



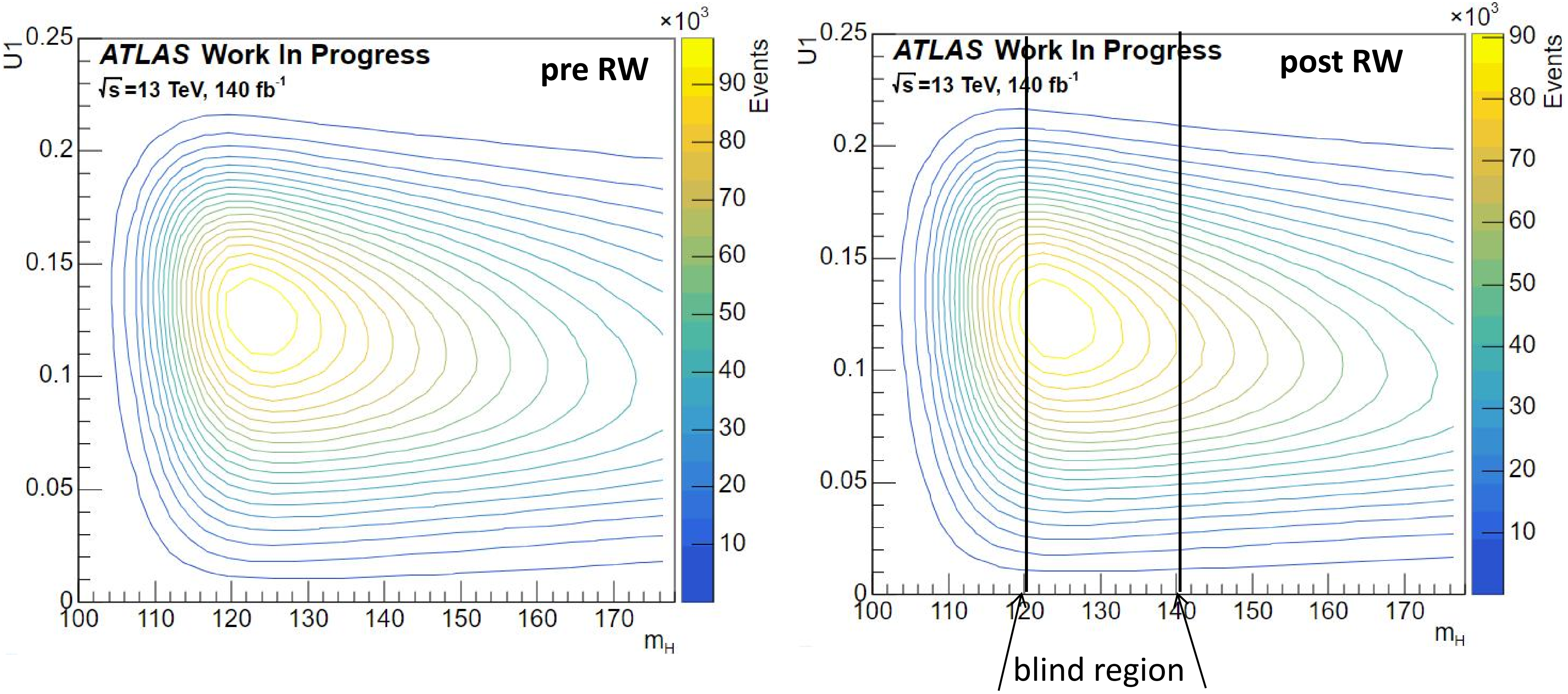
Background Reweighting





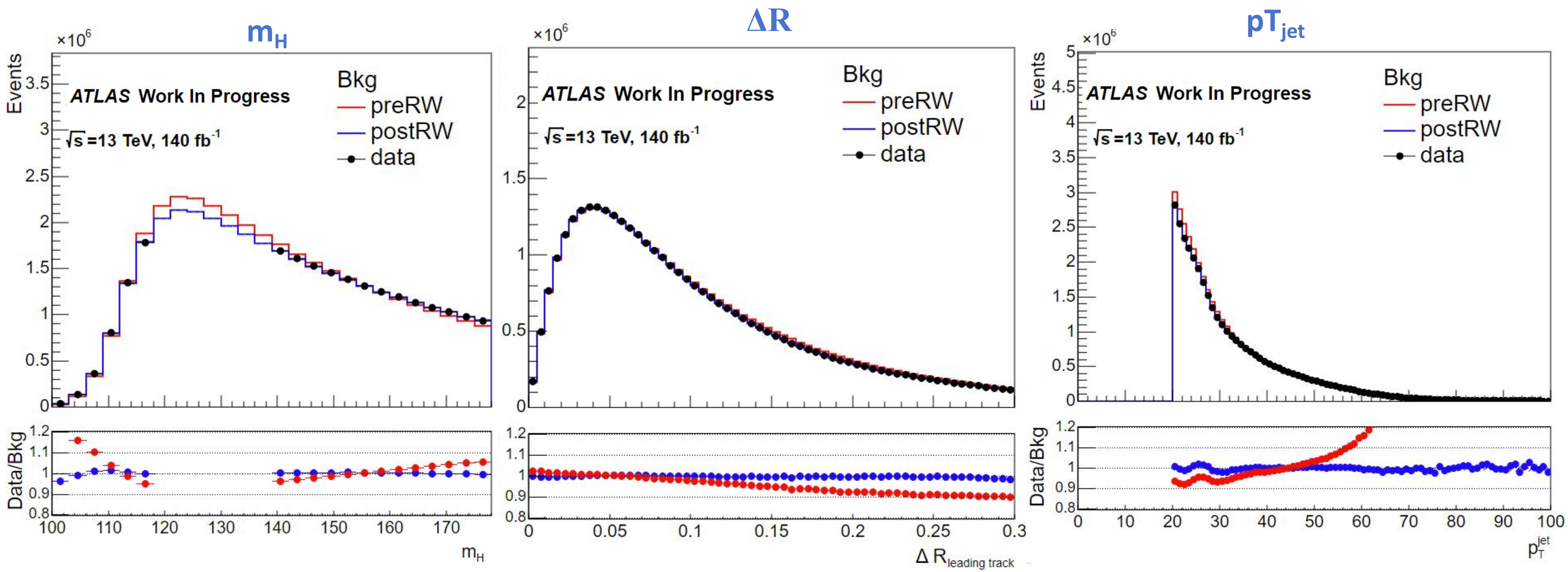
- Data/bkg ratio before and after reweighting
  - The RW improves the data-MC agreement a lot

Background Reweighting



- Z-axis: number of MC events
- Even the events in the blind region are excluded for the NN training, the NN can still understand the structure and give reasonable results in this region

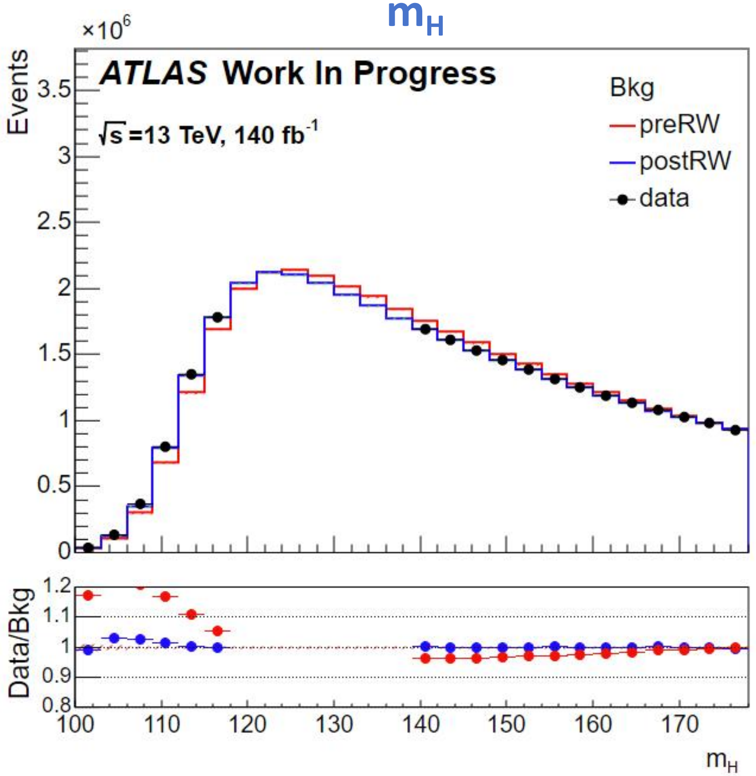
Reweighting result (Powheg)



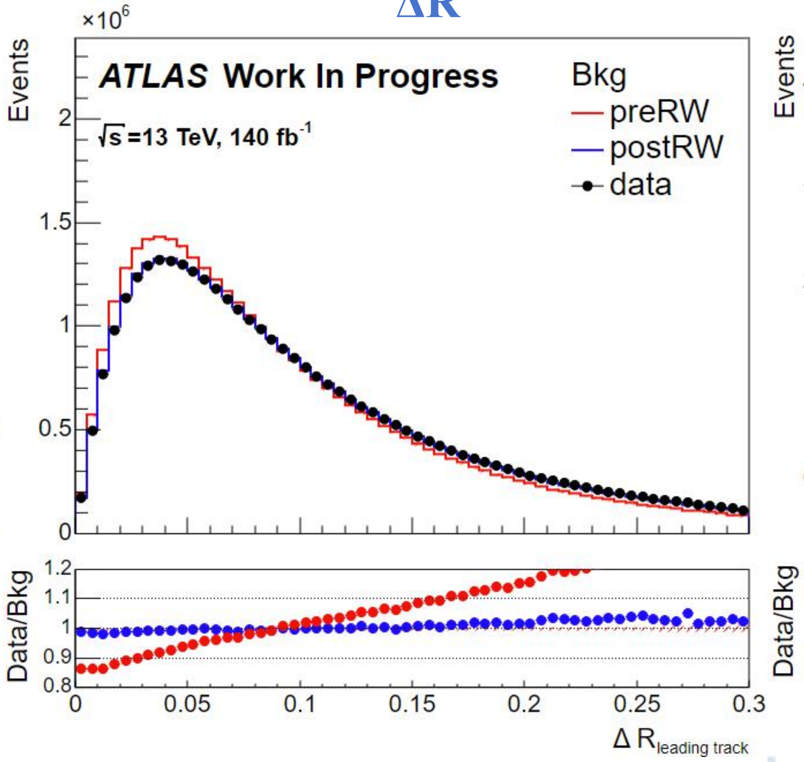
- The reweighting NN works well for all 11 training variables

Reweighting result (Sherpa)

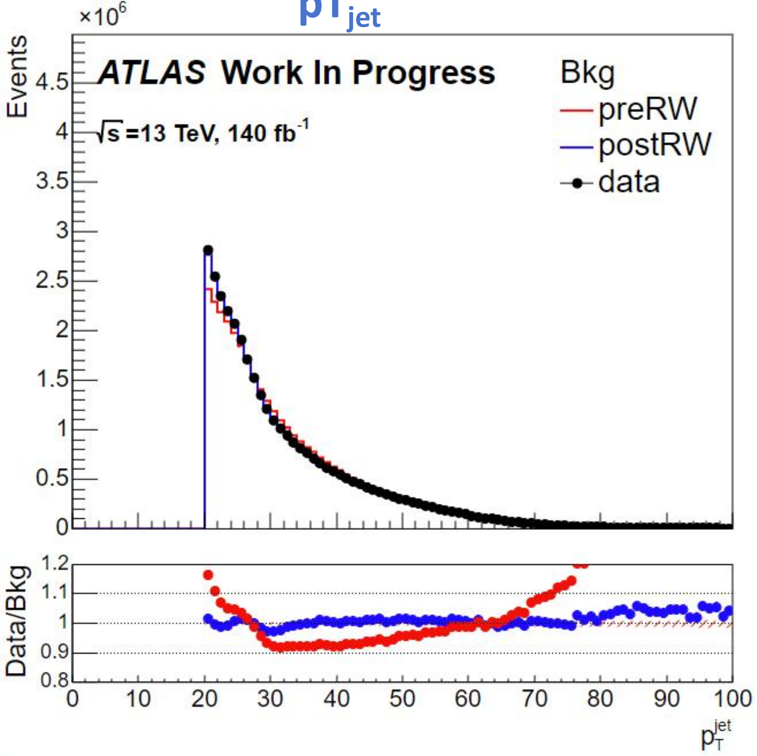
m_H



ΔR



pT_{jet}



- Trained another reweighting NN for Sherpa background samples
- The NN can achieve similar level of data-MC agreement as nominal
- The difference between Sherpa and Powheg is added as one of systematic uncertainties

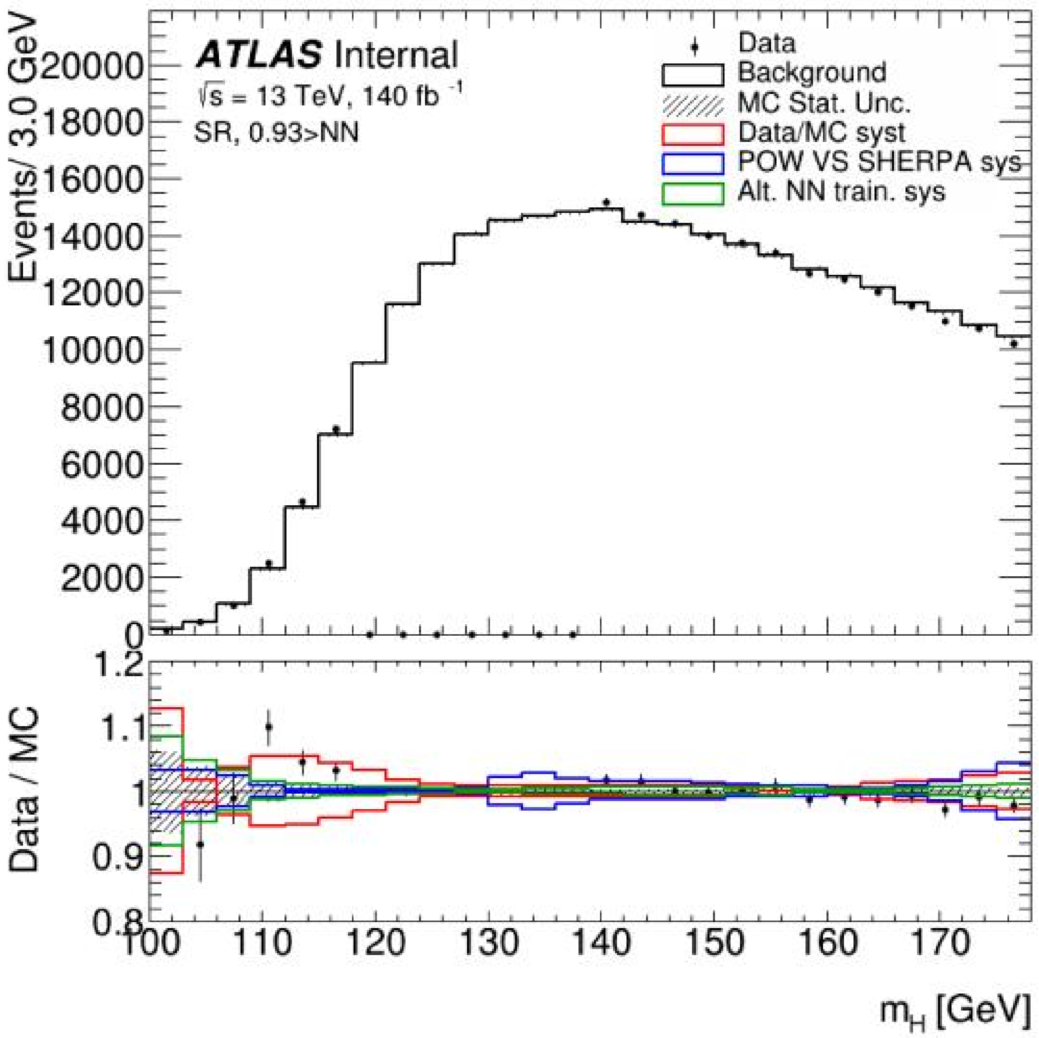
Systematic Uncertainties

- **Background systematic uncertainties**:
  - The experimental and theoretical uncertainties are replaced by 3 **background modelling uncertainties**:
  - Data-driven, estimated from the data-MC difference in the Control Region.
  - From different choice of generator
  - From different choice of the reweighting NN

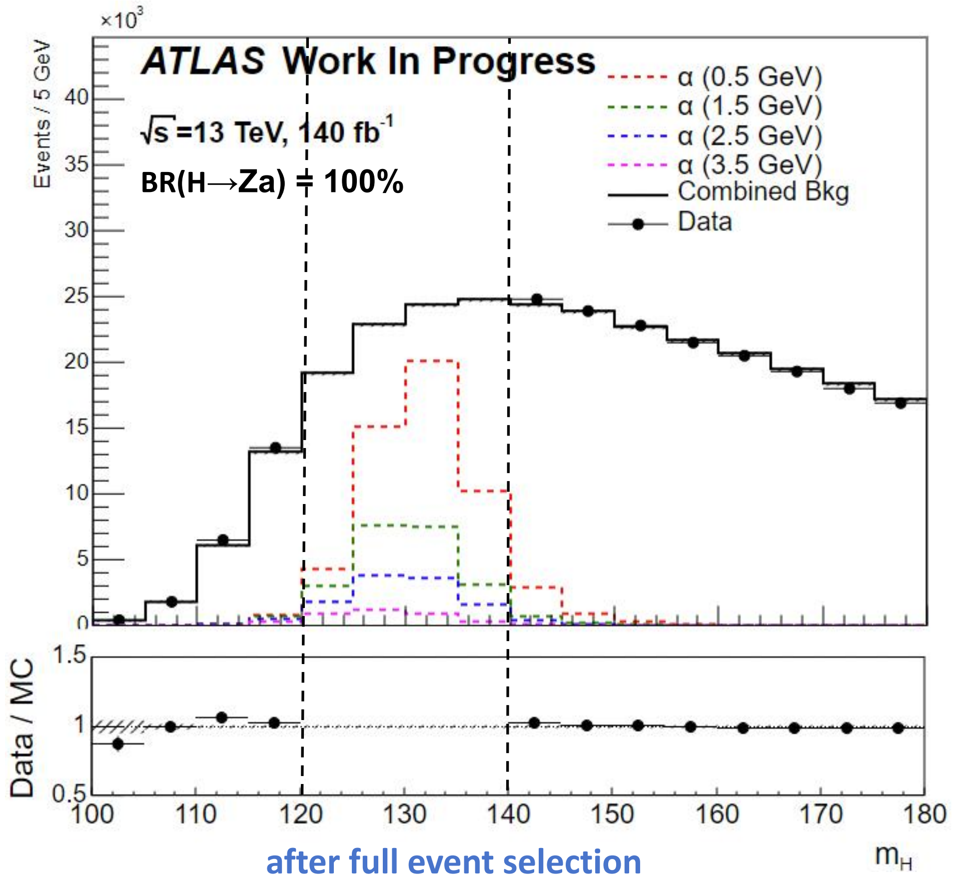
  - The impact of **statistical uncertainty** on bkg estimation reduced from 3.5% to 0.22% (negligible)

- **Signal systematic uncertainties:**
  - **Experimental**: Jet, tracking, pile-up, leptons, Trigger and vertex scale factors uncertainties. Following the latest recommendations

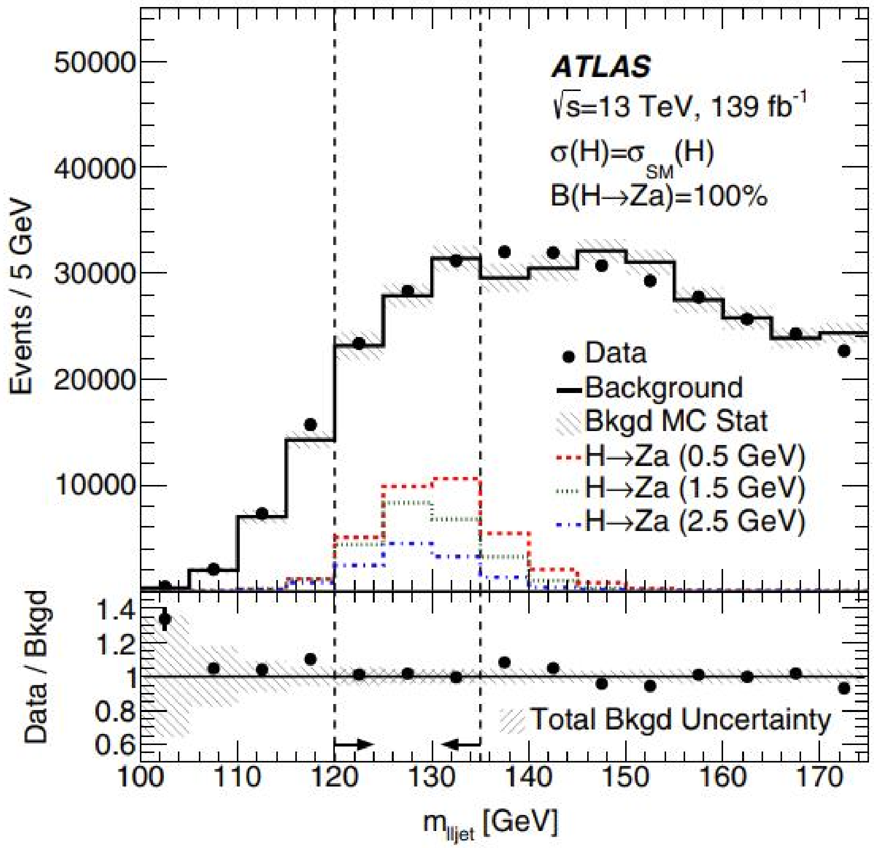
  - **Theoretical:** Parton Shower and Hadronization



Final State Invariant Mass



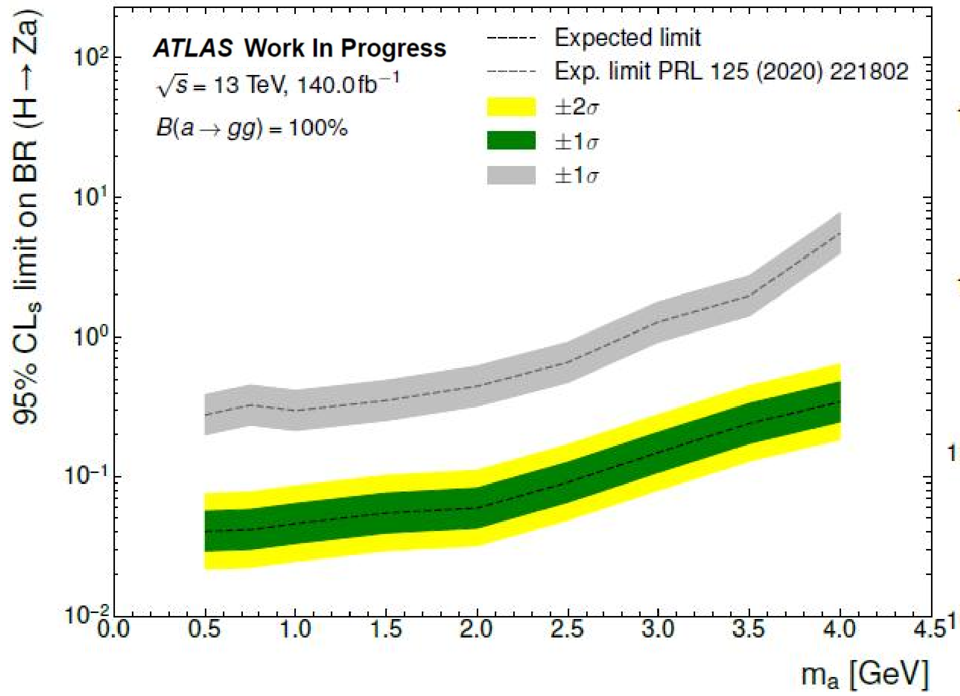
after full event selection
(pre-selection & Classification)

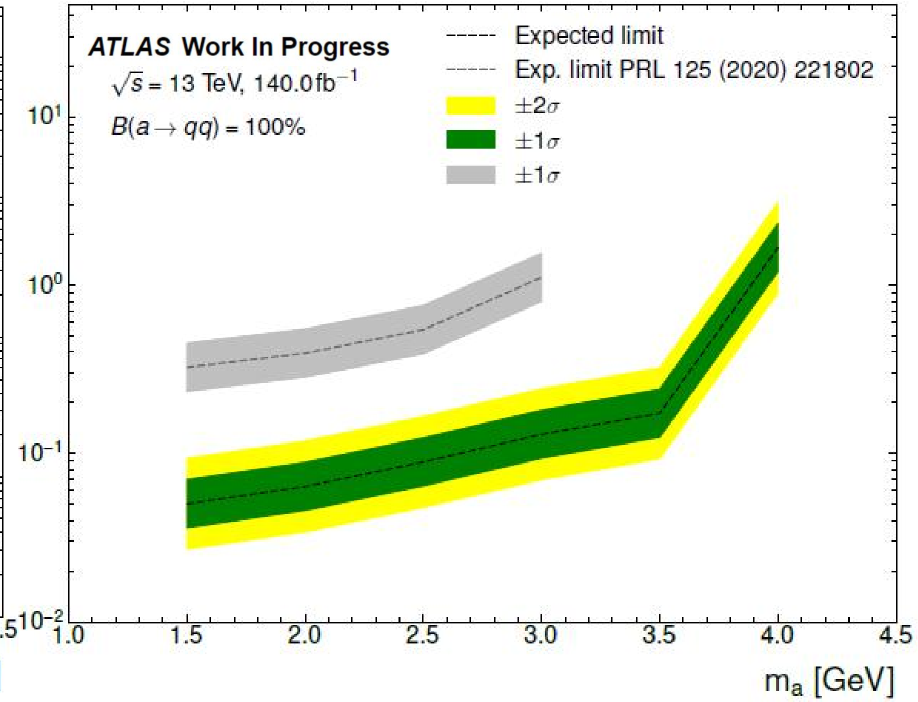


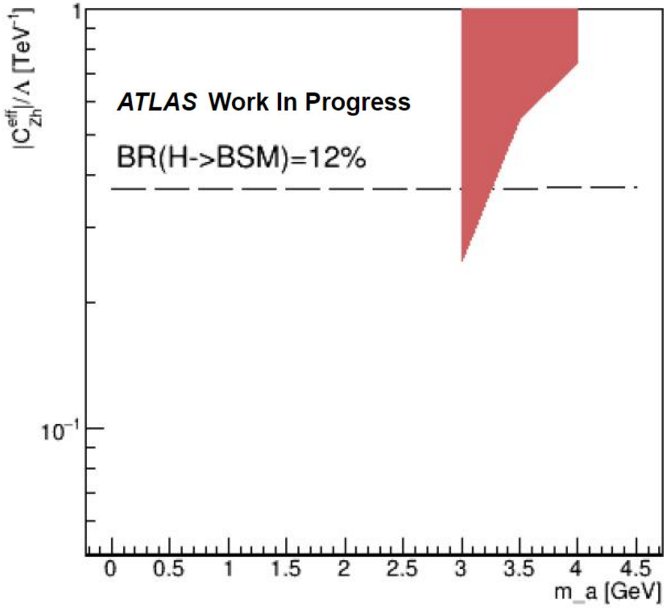
PhysRevLett.125.221802

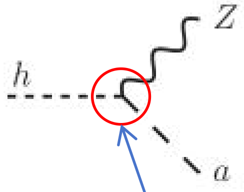
- Classification: set cut at NN output, bkg rejection 99% → 99.3%
- The significance (S/√B) in the SR increased
  - Bkg: 83k → 92k
  - 0.5 GeV signal: 27k → 50k

Exclusion Limits and ALP Interpretation









- **~5 times improvement in the expected limits on the BR(H→Za)**
  - Higher statistics MC sample
  - Novel tools for NN
  - Instead of using cut-and-count method, we used shape-fits

- Set expected limits on the effective coupling C_{Zh}^{eff}/Λ for the Axion-like particle

Summary

- Search performed for H→Za→ll+jet (m_a < 4GeV and hadronically decay)
- Dedicated NN used for background reweighting
- Significantly reduced the background statistical uncertainty, which is the main factor limiting the previous analysis sensitivity
- Upper limits set for exclusive gluon or quark decays, ~5 times improvement
- Set expected limits on the effective coupling for the Axion-like particle

UNIVERSITY OF
BIRMINGHAM

Back up

Changes

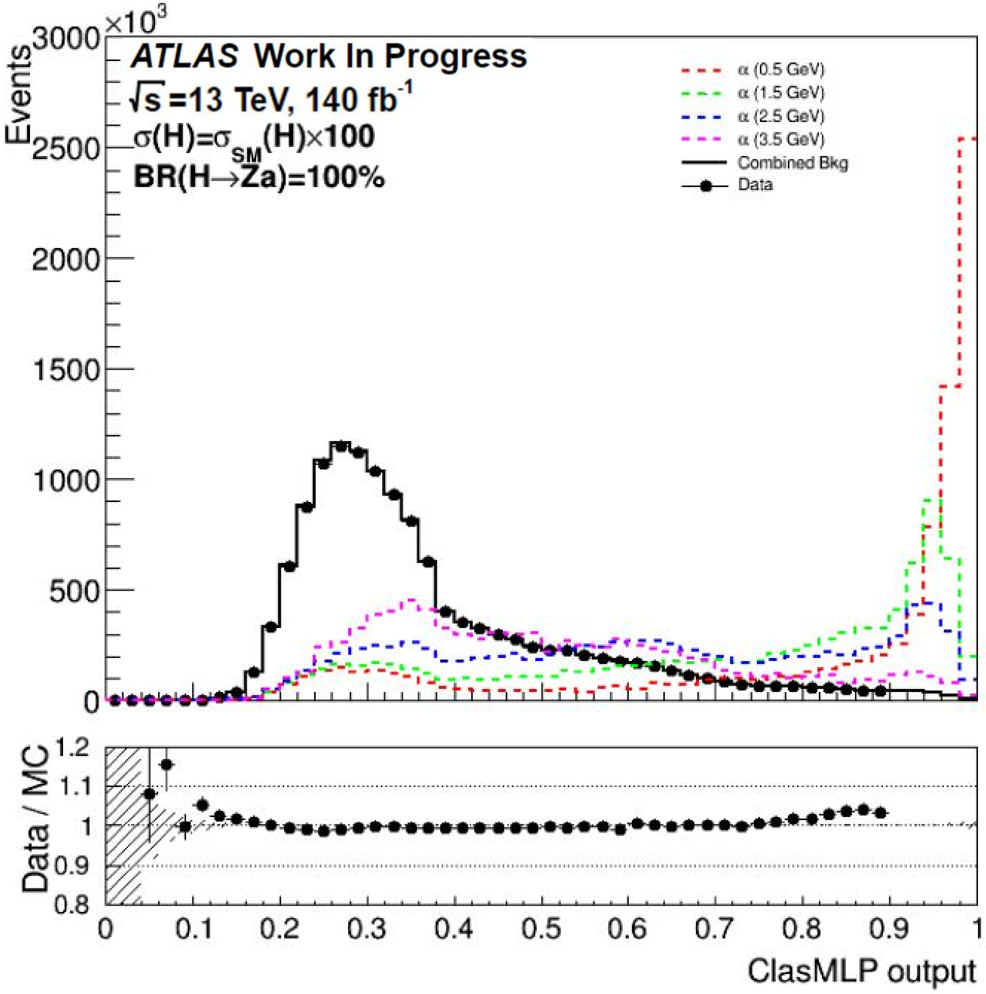
- **Bkg modelling**
  - SHERPA Z+jets ⟶ POWHEG Z+jets
- **Signal modelling**
  - NLO ⟶ NNLO
- **Derivation**
  - FTAG2 ⟶ HDBS3
- **Event selection**
  - EMTopo ⟶ EMPflow
  - recent recommended tools
- **Reweighting**
  - 3D Histograms ⟶ 11D Neural Network
- **Regression & Classification**
  - TMVA ⟶ Keras
- **Bkg estimation**
  - ABCD Method ->Control region
- **Fit**
  - cut&count -> shape fit
- **Additional interpretation**
  - Axion

Reweight variables

Variable	Description
m_{llj}	Invariant mass
n_{tracks}	Number of tracks
p_{T_H}	Transverse momentum of reconstructed Higgs boson
p_{T_Z}	Transverse momentum of reconstructed Z boson
p_T^{jet}	Transverse momentum of the calorimeter jet
$p_{\mathrm{T}}^{\text{lead track}}/p_{\mathrm{T}}^{\text{tracks}}$	Ratio of transverse momentum of the leading track to total
$\Delta R^{\text{lead track, calo jet}}$	ΔR between the leading track and the calorimeter jet axis
τ_2	NSubJettiness 2
$U_1(0.7)$	Modified energy correlation function
$M_2(0.3)$	Ratio of modified energy correlation functions
$angularity(2)$	Angularity

Classification NN

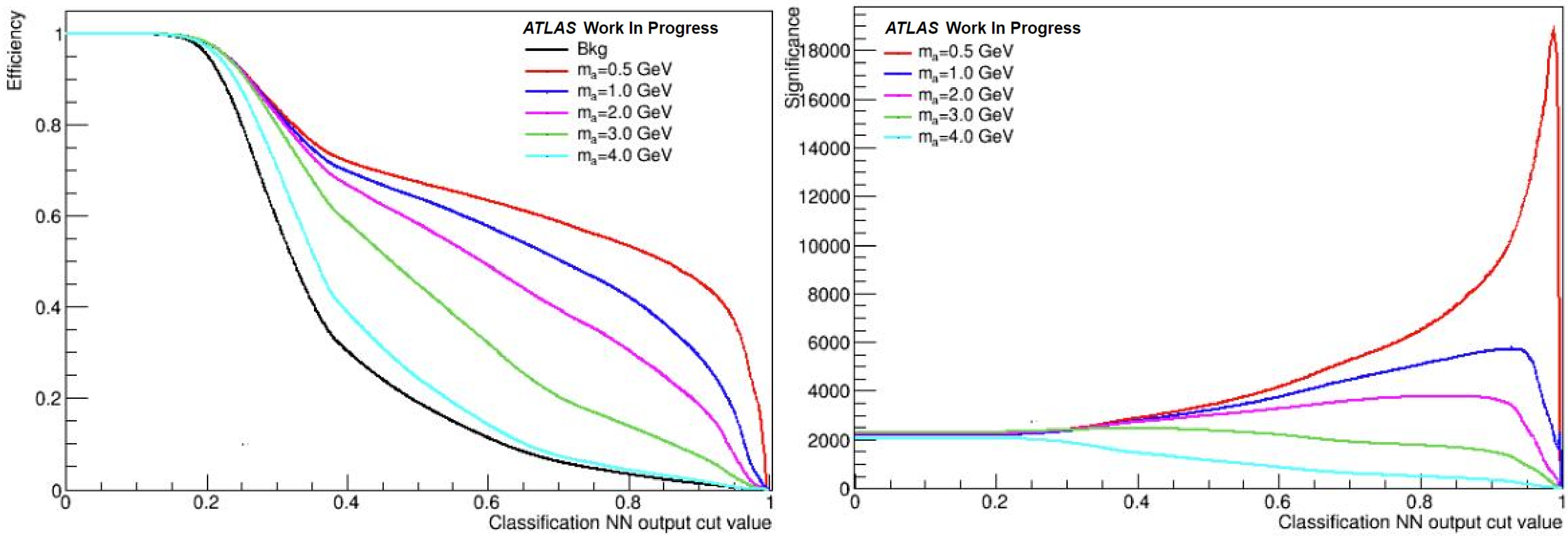
- To discriminate signal from the background
- NN Input: regression output + jet variables



Signal Region: NN output > **0.93**

- ~99% bkg events are rejected
- Relatively high significance for low mass signals

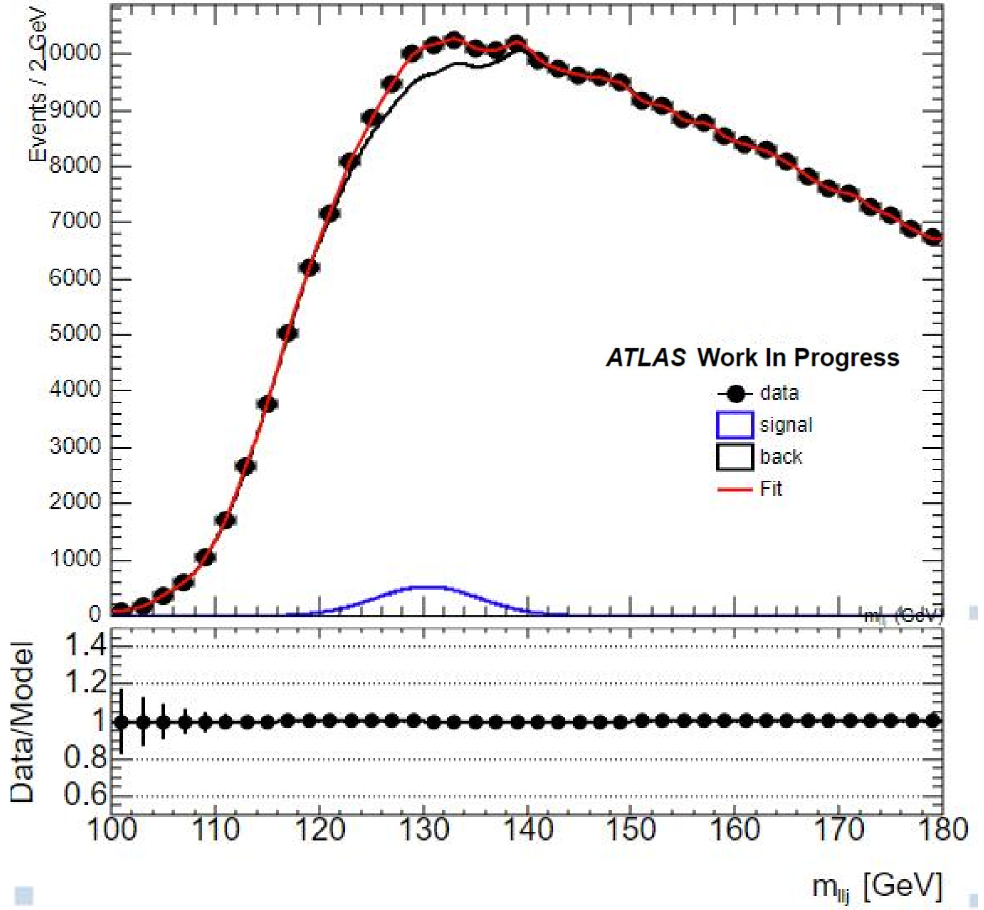
Classification NN



Fit

- **Fit parameters:**
  - μ: signal strengh
  - B: background normalization
  - a_b : background shape uncertainty
  - Δμ, Δσ: uncertainties of mean and sigma of nominal signal histograms
  - a_{Lumi} : Luminosity uncertainty

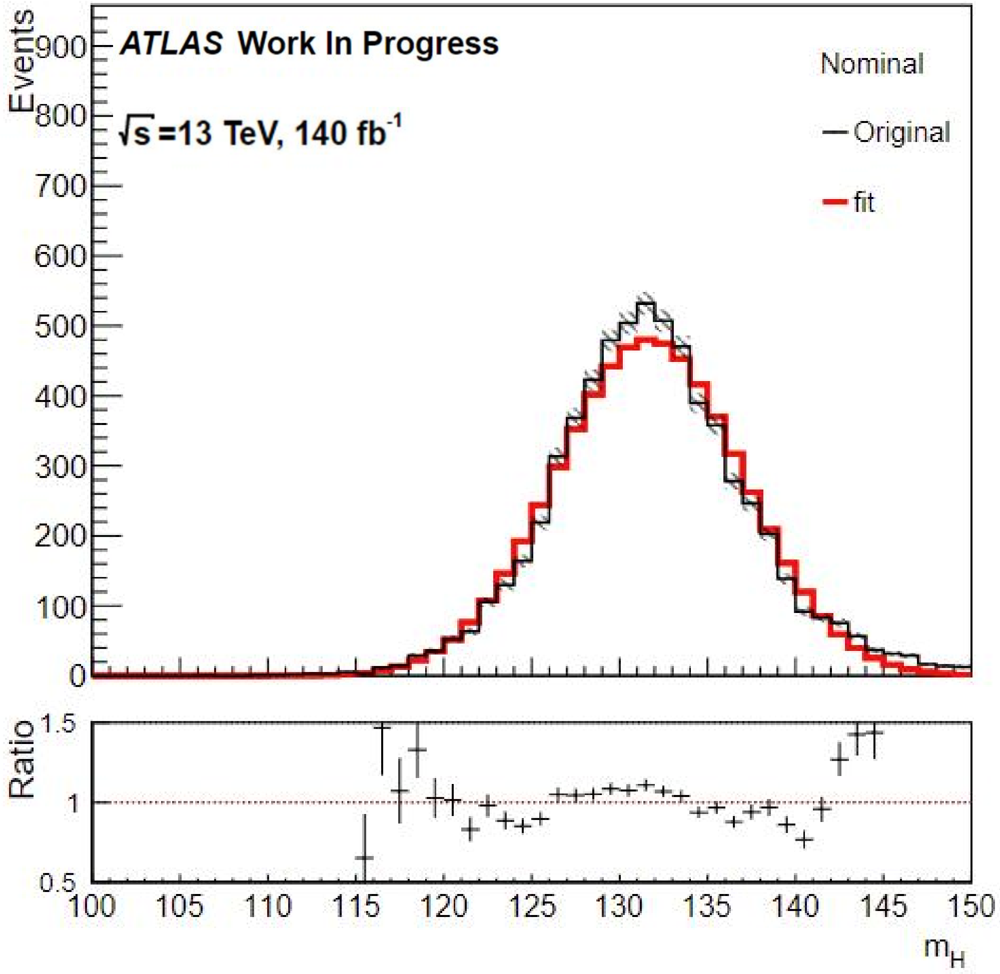
- 0.5 GeV example



Signal Modelling

- Fit the m_{llj} distribution with a Gaussian function for each signal
- Calculate the fit parameters (mean and sigma) of histograms
- The fitted mean and sigma will be added in the fit model

0.5 GeV sample



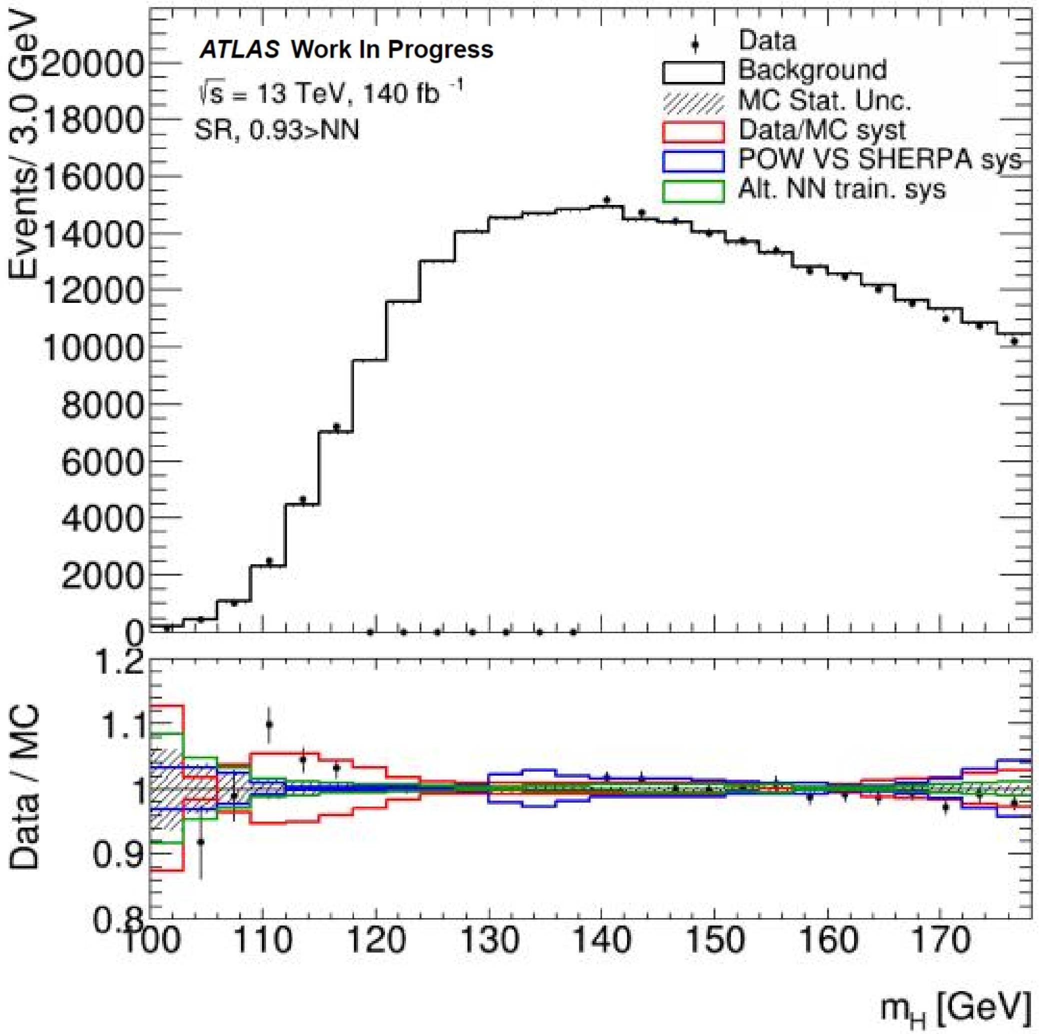
Systematic Uncertainties

3 **background systematic uncertainties**

- Data-driven, estimated from the data-MC difference in the Control Region.
- From different choice of MC generator.
- From different choice of the reweighting NN. Using the 2nd best nominal bkg reweighting NN.

Signal systematic uncertainties:

- Experimental: Luminosity, Pile up, Jet-related uncertainties
- Theoretical: Parton shower and Hadronization



Fit Model

ATLAS Work In Progress

a→gg

Mass (GeV)	Events N ($\cdot 10^3$)	μ	σ	σ_N (%)	σ_μ	σ_σ
0.5	58	131.5	5.0	20	1.5	0.7
1.0	33	130.5	5.1	20	1.5	0.7
1.5	24	130.7	5.1	20	1.5	0.7
2.0	21	130.1	5.3	20	1.5	0.7
2.5	12	129.7	5.5	20	1.5	0.7
3.0	6.5	128.4	5.5	20	1.5	0.7
3.5	3.9	127.4	5.6	20	1.5	0.7
4.0	2.5	126.6	5.2	20	1.5	0.7

a→qq

Mass (GeV)	Events N ($\cdot 10^3$)	μ	σ	σ_N (%)	σ_μ	σ_σ
1.5	28	129.4	5.4	21	1.5	0.7
2.0	19	128.9	5.5	21	1.5	0.7
2.5	13	129.1	5.8	22	1.4	0.7
3.0	7.9	127.8	5.5	24	1.7	0.7
3.5	5.6	127.0	5.0	29	1.4	0.7
4.0	0.61	125.1	5.5	29	1.5	0.7

Reweighting NN

arXiv:1911.00405 (2019)

- Minimize the cost function:

$$\mathcal{J}(\mathrm{u}) = \mathrm{E}_0[\phi(\mathrm{u}(X)) + \mathrm{r}(X)\psi(\mathrm{u}(X))] = \mathrm{E}_0[\phi(\mathrm{u}(X))] + \mathrm{E}_1[\psi(\mathrm{u}(X))]$$

scalar functions

$$\mathrm{r}(X) = \frac{\mathrm{f}_1(X)}{\mathrm{f}_0(X)}$$

E_0, E_1: expectation with respect to f_0, f_1
f_0, f_1: pdfs of bkg and data

ϕ and *ψ* are designed to satisfy: the global minimizer is equal to u(*X*)=*ω*(r(*X*)),
ω(r) is called the transformation function.

- In the case of **Log-Likelihood Ratio Estimation**:

$$\omega(\mathrm{r}) = \log \mathrm{r} \Longrightarrow \phi(z) = e^{0.5z},\ \psi(z) = e^{-0.5z}$$

u(X) estimates log-likelihood ratio, $e^{u(X)}$ estimates likelihood ratio

- In practice,

samples from bkg, data

$$\mathcal{J}(\mathrm{u}) \approx \hat{\mathcal{J}}(\theta) = \frac{1}{n_0}\sum_{i=1}^{n_0}\phi(\mathrm{u}(X_i^0, \theta)) + \frac{1}{n_1}\sum_{i=1}^{n_1}\psi(\mathrm{u}(X_i^1, \theta))$$

θ : NN parameters

u(X,θ): NN output

The cost function only depends on two datasets and *ω*(r), requires no knowledge of pdfs f_0, f_1

Optimization of u(X) → Classical optimation of NN parameters (θ)