ATLAS Carleton

Suppression Without Shaping: DisCo Neural Network Optimization for $H \to \mu\mu$ Decay Analysis

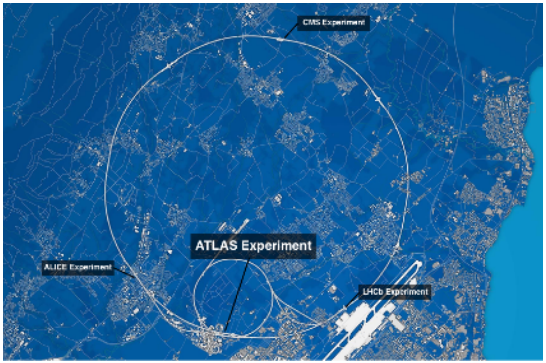
Sam Moir

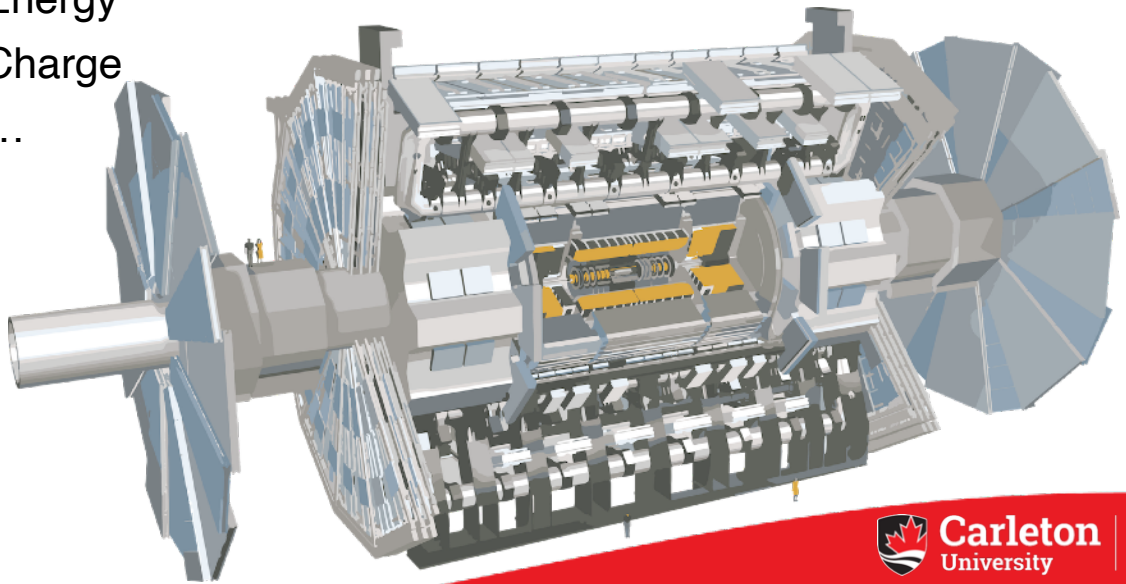
WNPPC - 2025

Carleton University
Department of Physics

LHC and ATLAS

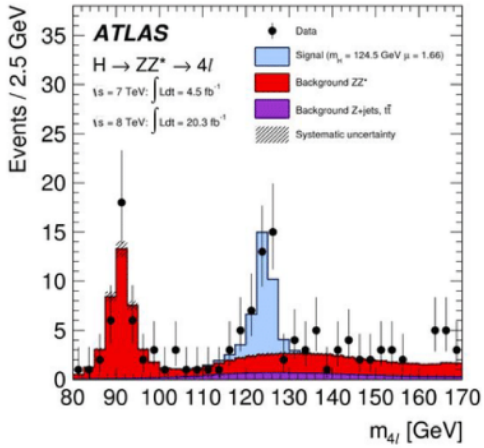
- Collide protons
- Produce Higgs boson
- Measure output:
  - Momentum
  - Energy
  - Charge
  - …



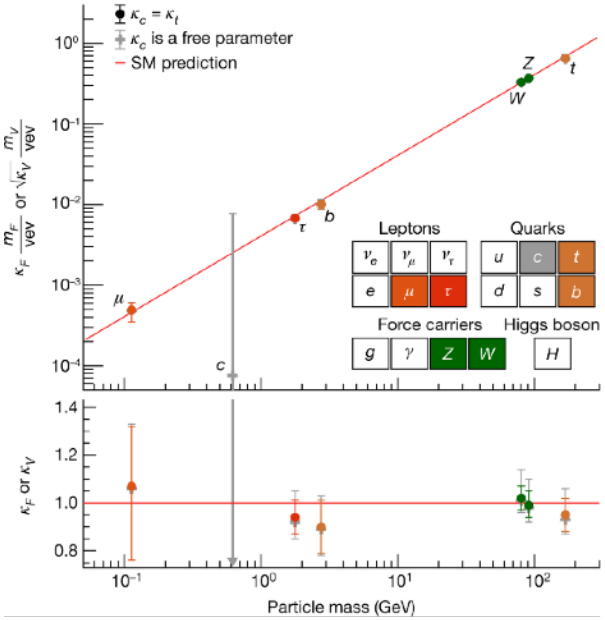


Reconstructing a Higgs Boson Decay

- Reconstruct mass from particles' energy & momentum
- Sharp peak around Higgs mass (125 GeV) indicates Higgs decay



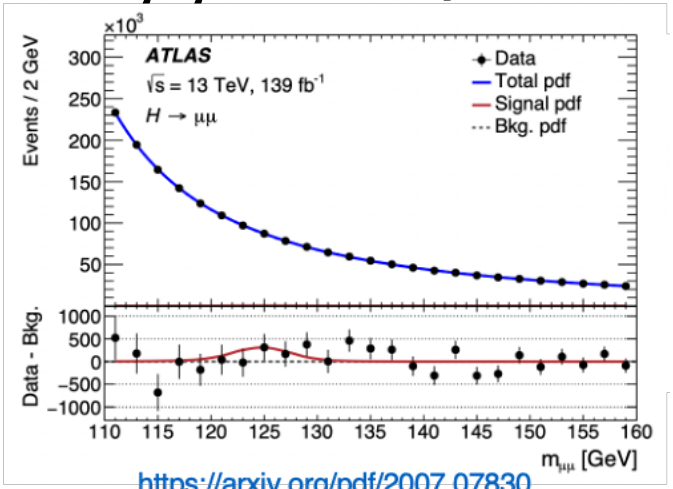
DOI: 10.1088/1742-6596/706/2/022002



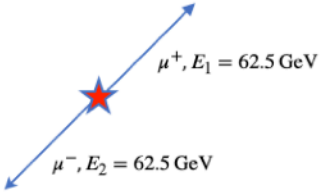
https://arxiv.org/pdf/2207.00092

Reconstructing a $H \to \mu\mu$ decay

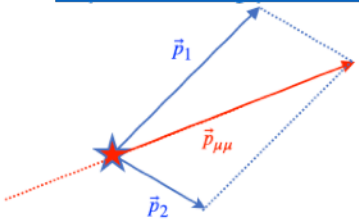
- Reconstruct mass from muons' energy & momentum
- Sharp $m_{\mu\mu}$ peak around Higgs mass (125 GeV) indicates Higgs decay



https://arxiv.org/pdf/2007.07830



Higgs boson rest frame

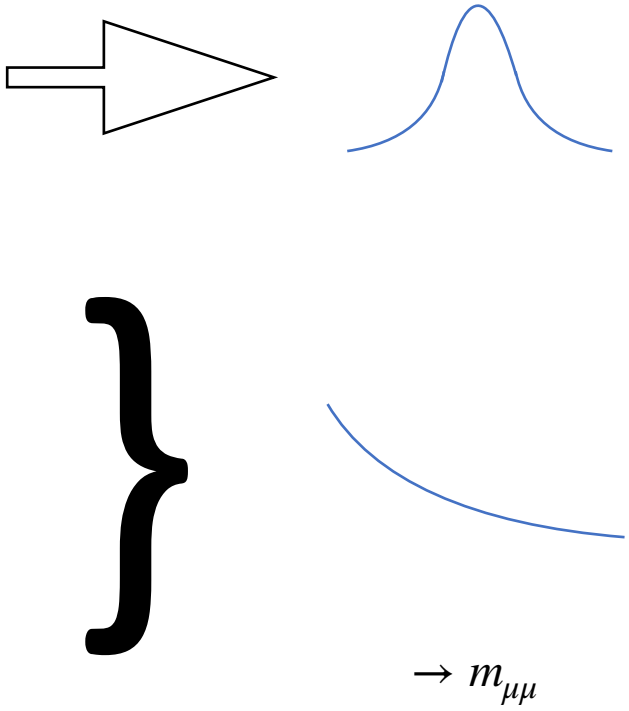


LAB frame

Higgs boson mas momentum

Comparing data and simulations

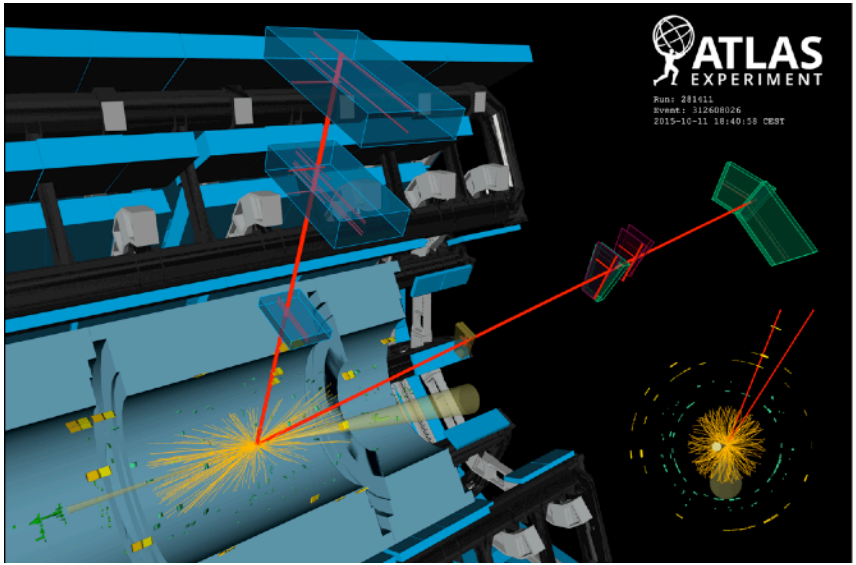
- Signal:
  - $H \to \mu\mu$
- Background processes:
  - $Z \to \mu\mu$
  - $t\bar{t} \to \boldsymbol{\mu\mu}jj\nu\nu$
  - …
  - $ZZ \to \boldsymbol{\mu\mu}\ell\ell$
  - $tt \to \boldsymbol{\mu\mu}jjjj\nu\nu$
  - $WZqq \to \boldsymbol{\mu\mu}\ell\nu jj$



$\to m_{\mu\mu}$

Using ML to extract $H \rightarrow \mu\mu$ signal

- Boosted Decision Trees (BDTs)
  - Achieved 2.0σ measurement
  - Next goal is 3.0σ
- Deep Neural Nets (DNNs)
  - Major issue: background shaping



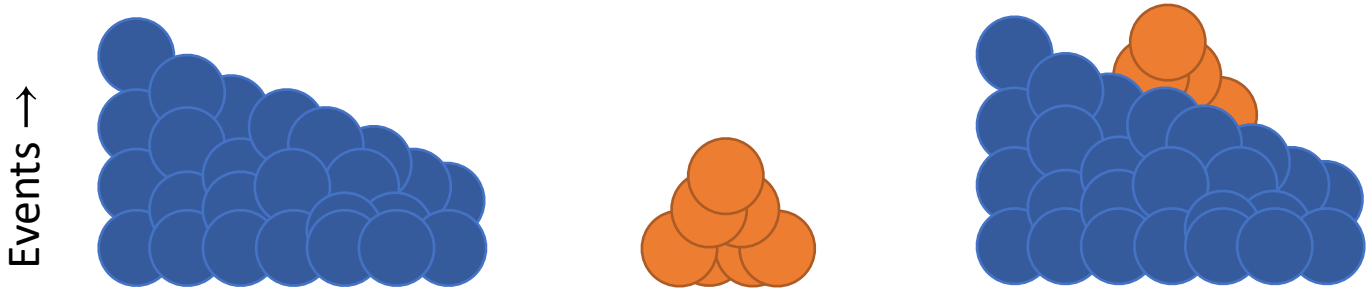
Background shaping

Pre-NN

Background

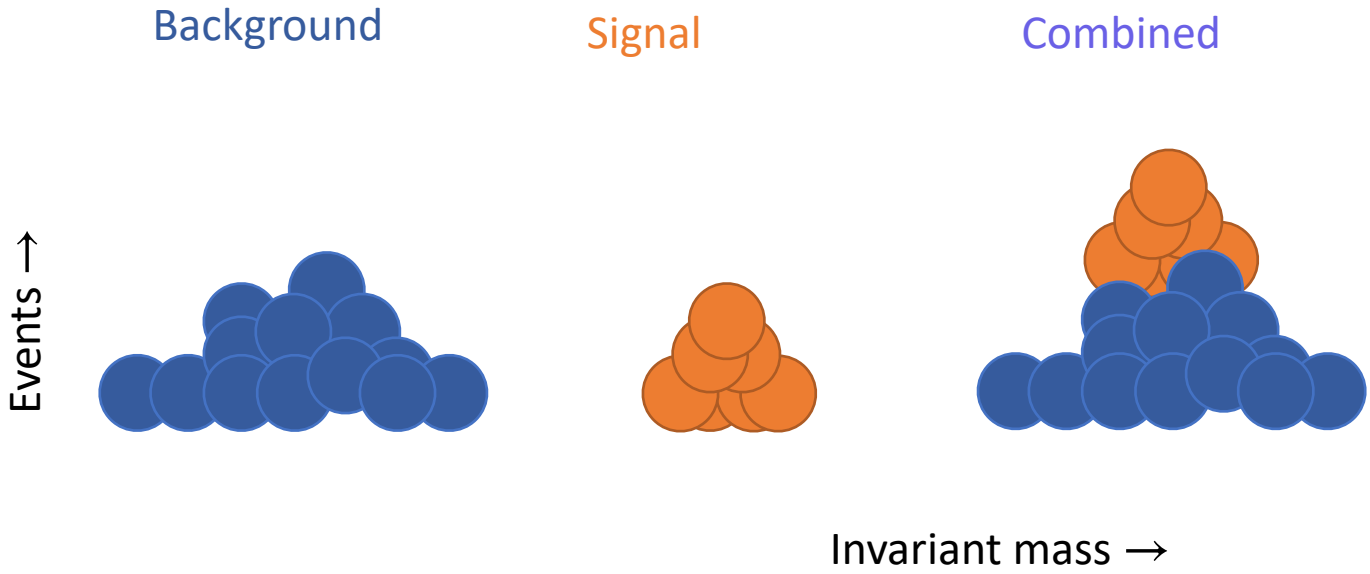
Signal

Combined



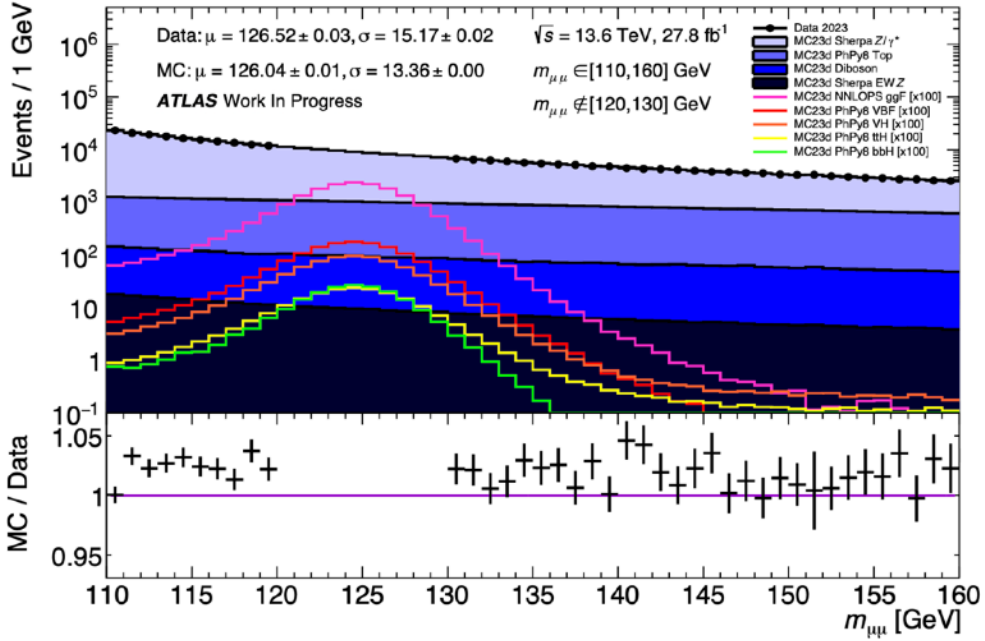
Background shaping

Post-NN: NN selects background in signal region



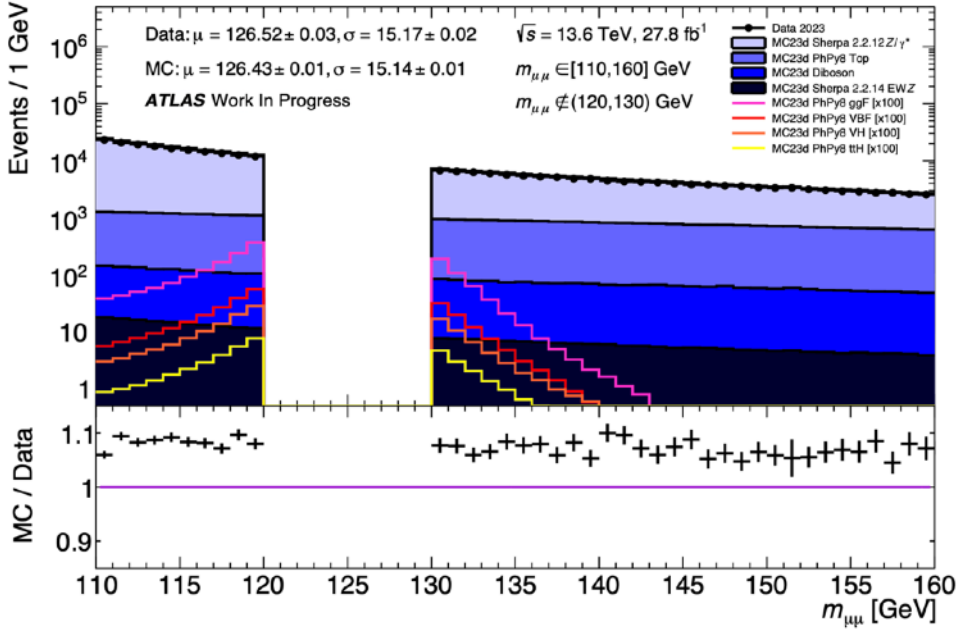
Why is background shaping bad?

- Want to subtract background
  - Need to know background from sidebands



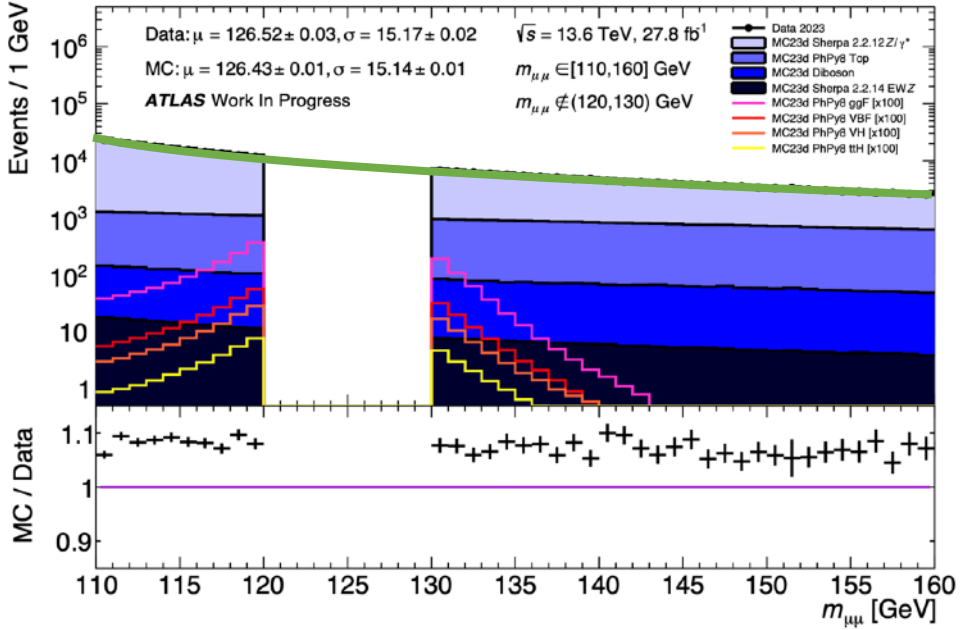
Why is background shaping bad?

- Want to subtract background
  - Need to know background from sidebands



Why is background shaping bad?

- Want to subtract background
  - Need to know background from sidebands



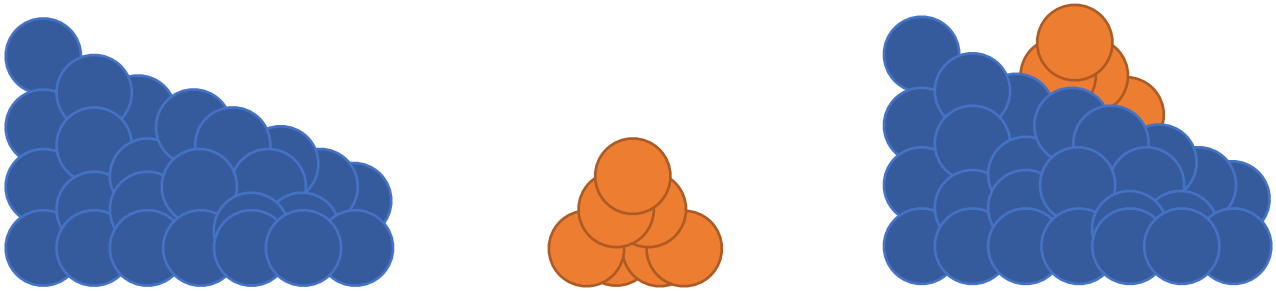
Unshaped background

Pre-NN

Background

Signal

Combined



Events →

Invariant mass →

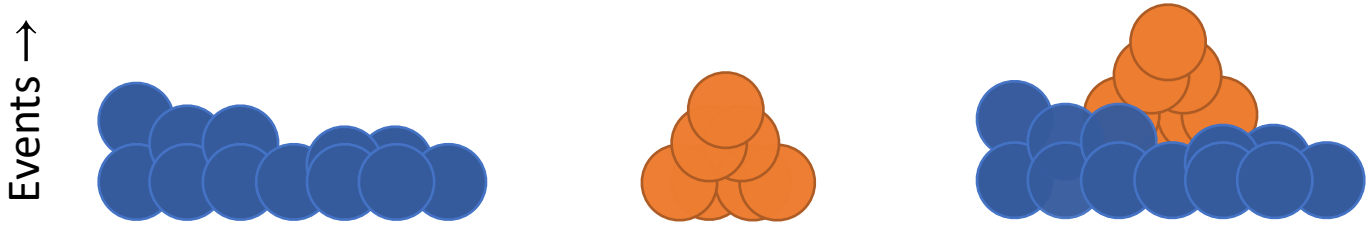
Unshaped background

Post-NN: NN selects background evenly

Background

Signal

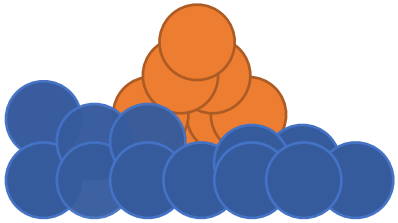
Combined



Unshaped background

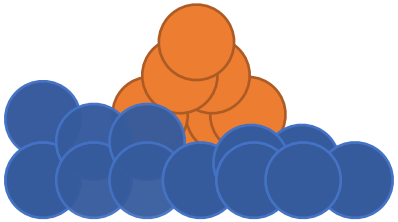
Post-NN: NN selects background evenly

Events →



Invariant mass →

Unshaped background



Events →

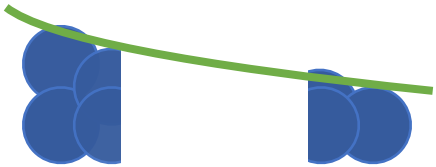
Invariant mass →

Unshaped background

Events →

Invariant mass →

Unshaped background



Events →

Invariant mass →

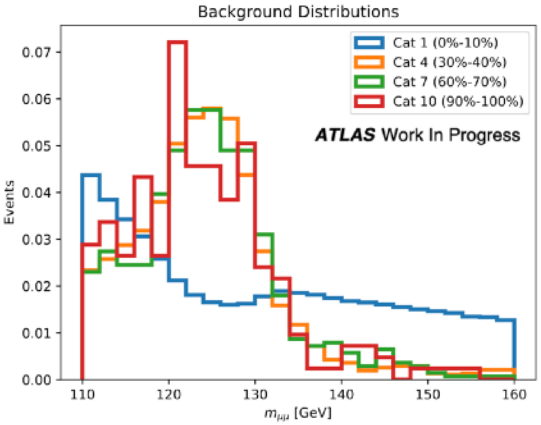
Using ML to extract $H \rightarrow \mu\mu$ signal

- Previous Analyses
  - BDTs
  - DNNs
- New Analysis: DisCo Neural Net (DisCo NN)
  - Adds a distance-correlation term to DNN loss function
  - Punishes shaping in the background

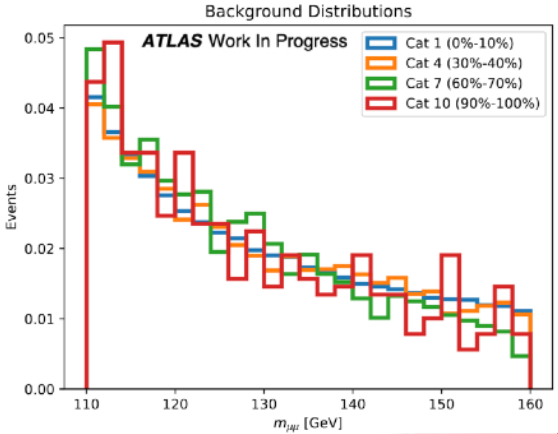
DisCo NN

- Loss function
  - $L(\vec{y}_{\text{pred}}, \vec{y}_{\text{true}}) = L_{\text{BCE}}(\vec{y}_{\text{pred}}, \vec{y}_{\text{true}}) + \alpha \cdot \text{dCorr}^2_{\vec{y}_{\text{true}}=0}(\vec{m}_{\mu\mu}, \vec{y}_{\text{pred}})$
    - L_{BCE}: Standard DNN training term
    - α: DisCo parameter controlling loss due to background shaping
    - dCorr^2: Distance Correlation function

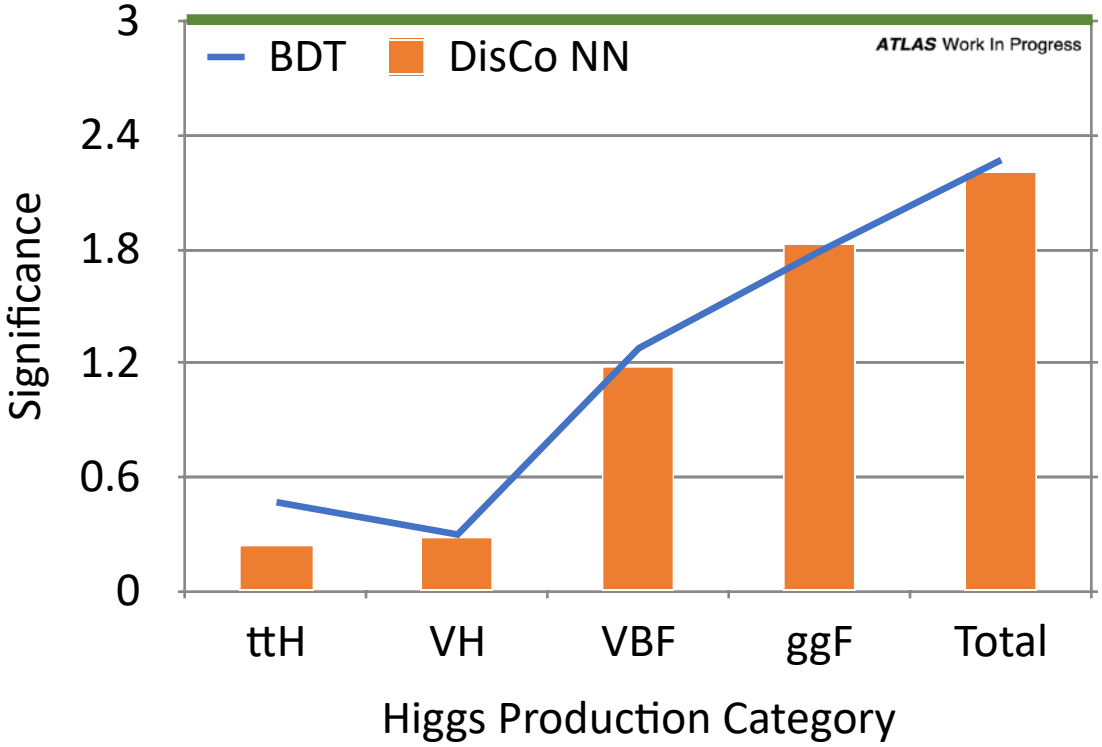
$\alpha = 0.0$



$\alpha = 30.0$



Preliminary Results



Conclusion + Next Steps

- $H \rightarrow \mu\mu$ analysis > next big step in Higgs physics
  - Difficult due to quantity of $\mu\mu$ production in LHC
- ML analyses unsatisfactory (so far)
  - BDTs don't learn as well as DNNs, DNNs produce shaping
- New direction for ML analysis: DisCo
  - Punishes shaping
  - Preliminary results competitive with BDT
- Future steps
  - Continue to optimize meta-architecture (exact value of α, NN hyperparameters, etc.)
  - Train on more parameters (option not available to BDT)

Acknowledgements

Acknowledgements: Bryce Norman, Ian Ramirez-Berend, Dag Gillberg, Manuella Vincter, Thomas Koffas (Supervisor)