# Search for electroweakinos with small Δm using XGBoost

Knut Vadla (University of Oslo)

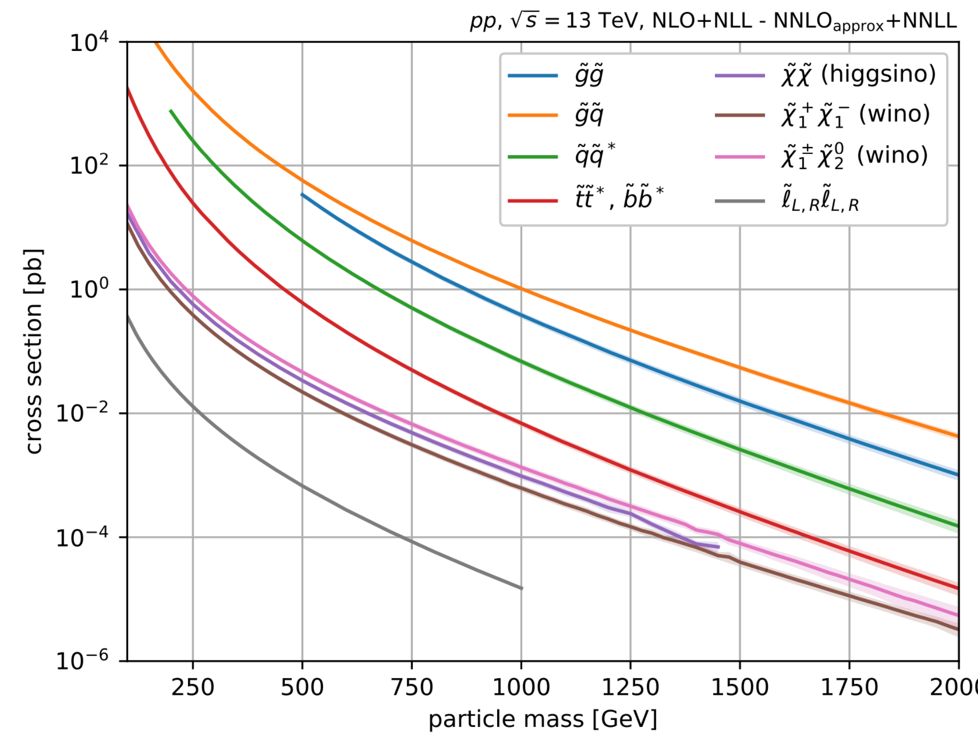
Spåtind 2020 — Nordic Conference on Particle Physics


Photo ref.: ATLAS-PHOTO-2016-009-1

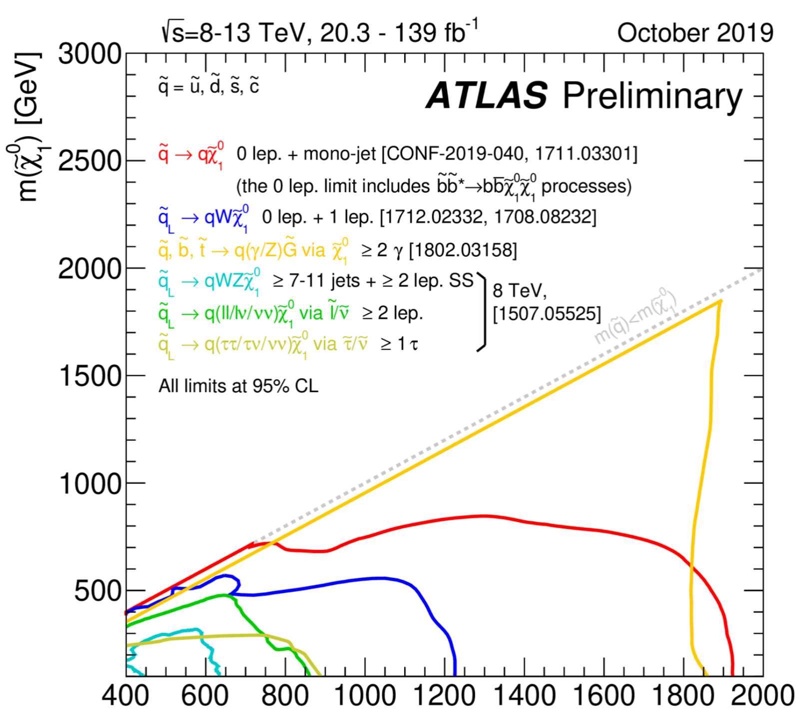
# SUSY production @ LHC

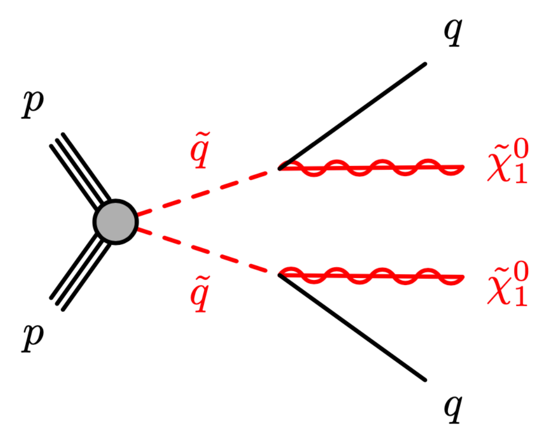
[Plot: https://github.com/fuenfundachtzig/xsec](https://github.com/fuenfundachtzig/xsec)

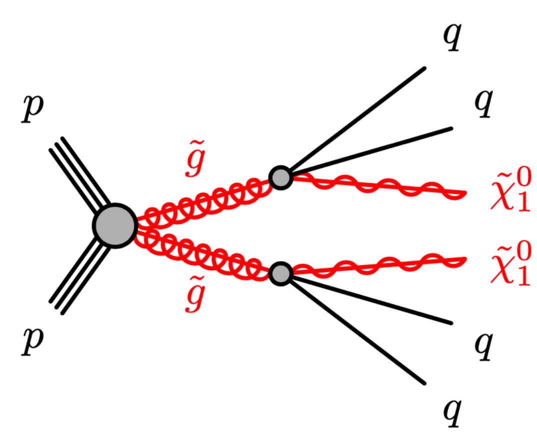
Cross-section refs.: see slide 14

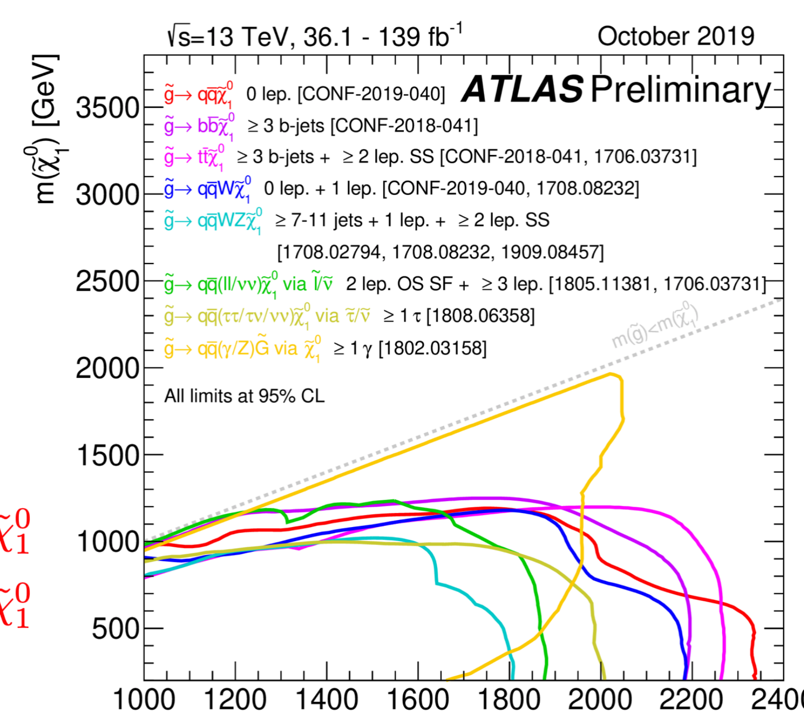



- Cross-sections for strongly interacting sparticles expected to be much higher than weakly interacting sparticles @ LHC
- No signs of strong SUSY so far
- Is SUSY hiding in the EW sector?



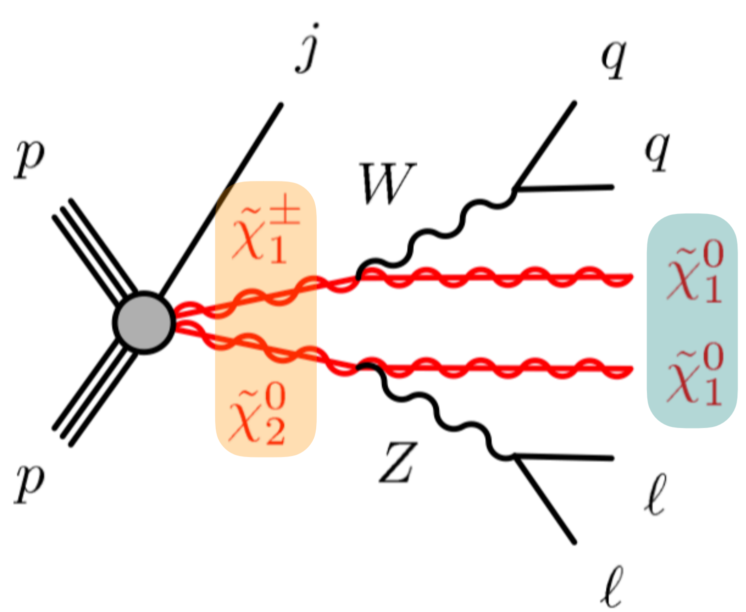

[ATL-PHYS-PUB-2019-044]()

[ATL-PHYS-PUB-2019-044]()

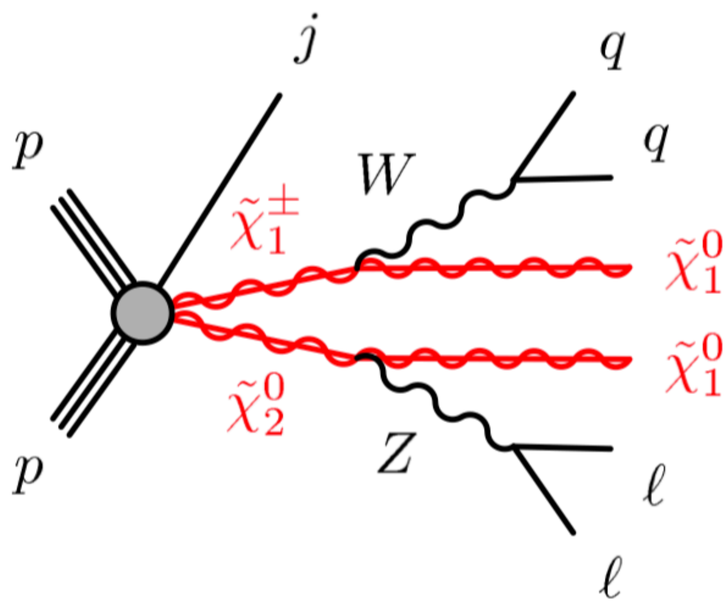
# Search for electroweakinos

- **Electroweakinos**
  - ***Neutralinos***, $\tilde{\chi}_i^0 = N_1\tilde{B}^0 + N_2\tilde{W}^0 + N_3\tilde{H}_u^0 + N_4\tilde{H}_d^0 \qquad i = 1, 2, 3, 4$
    - superpartners of the neutral gauge and CP-even higgs bosons
  - ***Charginos***, $\tilde{\chi}_j^\pm = C_1\tilde{W}^\pm + C_2\tilde{H}_{u/d}^\pm \qquad j = 1, 2$
    - superpartners of the charged gauge and higgs bosons

- **Simplified model**
  - Mass-degenerate $\tilde{\chi}^\pm_1$, $\tilde{\chi}^0_2$ (wino)
  - Decay to $\tilde{\chi}^0_1$ (bino) via W(qq)Z(ll)
  - *Final state:* 2 leptons + 2 (+ISR) jets + $E_T^{miss}$



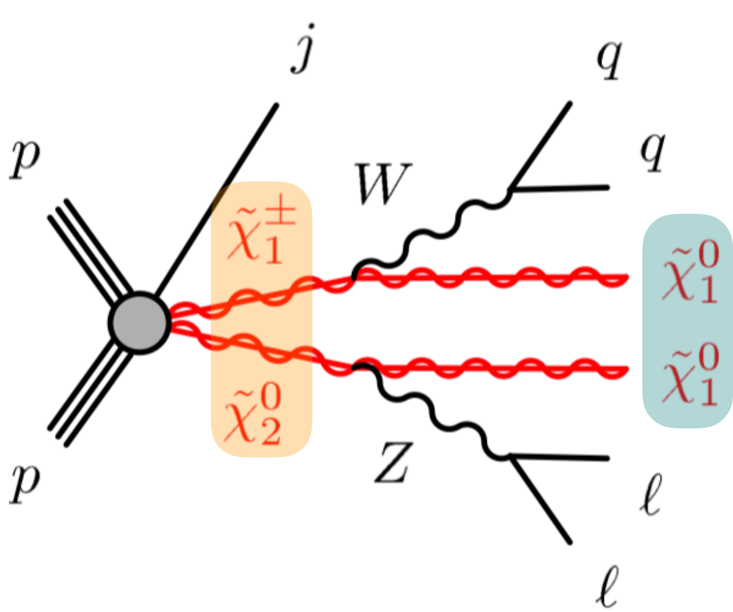
# Preselection of events and objects

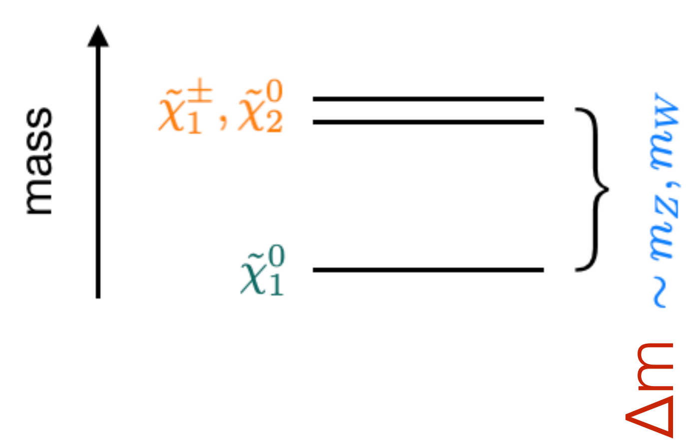


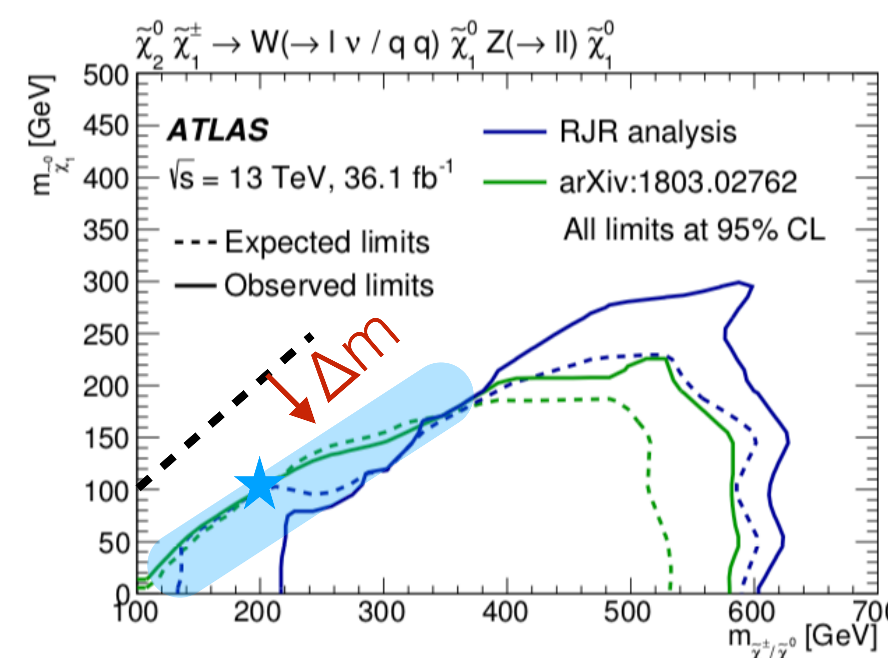

| | **Preselection** |
|---|---|
| **triggers** | ee, μμ, eμ/μe |
| **# SFOS leptons** $p_T > 25$ GeV | 2 |
| **# jets** $p_T > 30$ GeV | ≥ 2 |
| **# b-jets** $p_T > 20$ GeV, 77% eff | 0 |
| **$m_{ll}$ [GeV]** | (81, 101) |
| **$m_{jj}$ [GeV]** | (60, 110) |

# Scenario: $\Delta m(\tilde{\chi}^{\pm}_1/\tilde{\chi}^0_2, \tilde{\chi}^0_1) \sim m_Z, m_W$

- **Signal looks very much like the SM WZ background**
- Conventional cut-and-count analyses have used initial state radiation (ISR) to boost the final-state momenta (recoiling against ISR jet)
- Struggling to be sensitive to $\Delta m \sim 100$ GeV
- Can machine learning (ML) methods help gain sensitivity to these particularly difficult scenarios?




arXiv:1806.02293

# XGBoost BDT

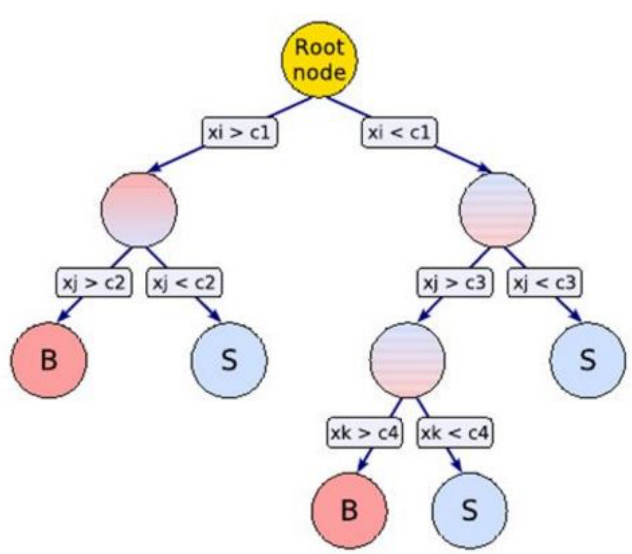
- **Xtreme Gradient Boosting (XGBoost)**
- **Decision tree**
  - Make leaf splits that minimize an **objective function**

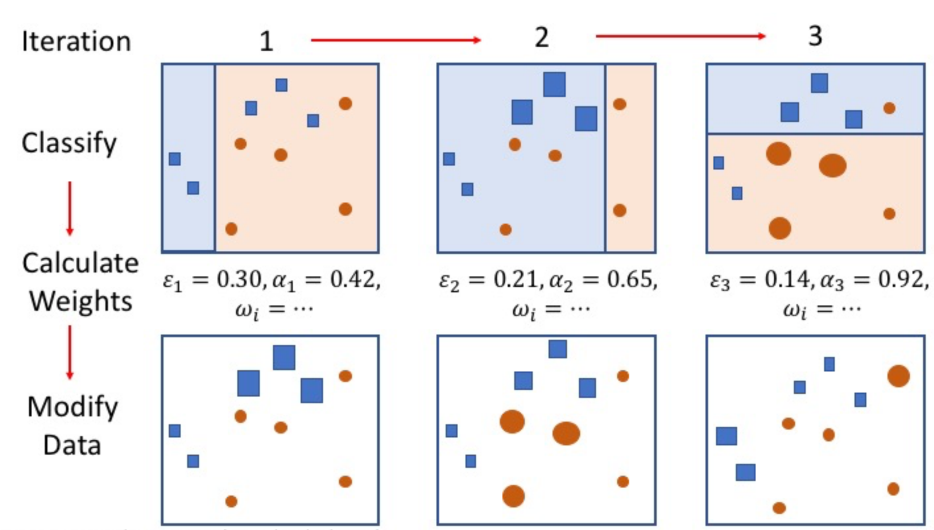
$$\text{obj}^{(t)} = \sum_{i=1}^{n} l(y_i, \hat{y}_i^{(t)}) + \sum_{i=1}^{t} \Omega(f_i), \qquad \Omega(f) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2$$

obj. func. — loss — regularization

$t$ : number of trees
$n$ : number of data points
$y$ : true class label
$\hat{y}$ : predicted class label
$f$ : tree function/complexity
$T$ : number of leaves
$\gamma$ : min split loss
$\lambda$ : $L^2$ norm/weight decay

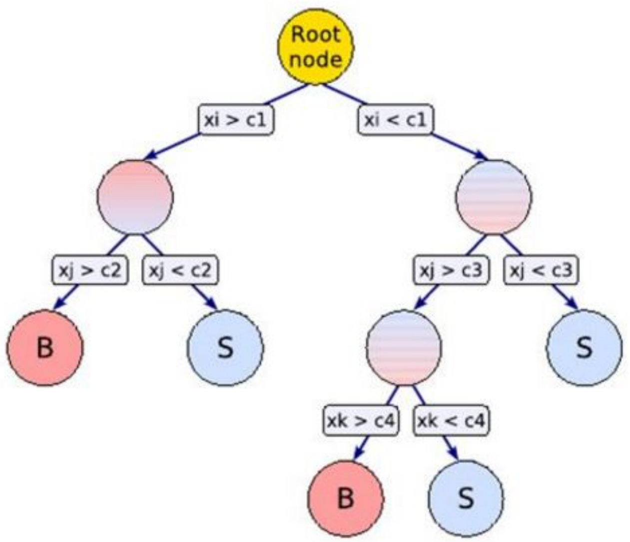
  - Optimize obj. func. using **gradient descent**
  - Combine **multiple trees** that are sequentially trained to **correct mistakes of previous trees (boosting)**




Source: A Tutorial on Boosting (Freund and Schapire)

# XGBoost setup




| | Training parameters |
|---|---|
| **max tree depth**<br>(controls number of variables/cuts allowed) | 10 |
| **learning rate**<br>(size of gradient descent update) | 0.1 |
| **gamma, γ**<br>(required reduction in loss in order to split a leaf) | 20 |
| **subsample**<br>(fraction of events used for training; random per tree) | 0.5 |
| **colsample_bytree**<br>(fraction of variables used for training; random per tree) | 0.5 |

# Training and testing the low-Δm BDT

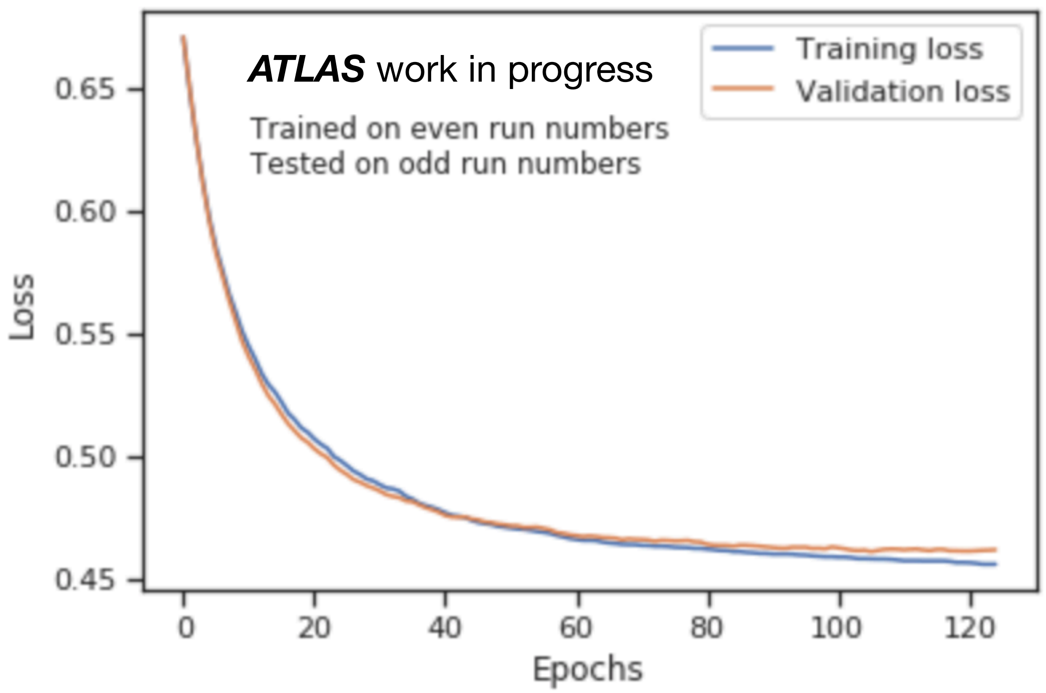
- **Signal (MC)**
  - 21 different samples/mass points
  - $\Delta m(\tilde{\chi}^{\pm}_1/\tilde{\chi}^0_2, \tilde{\chi}^0_1)$ = 50, 100, 150, 200 GeV
  - ~60,000 events in total pass preselection

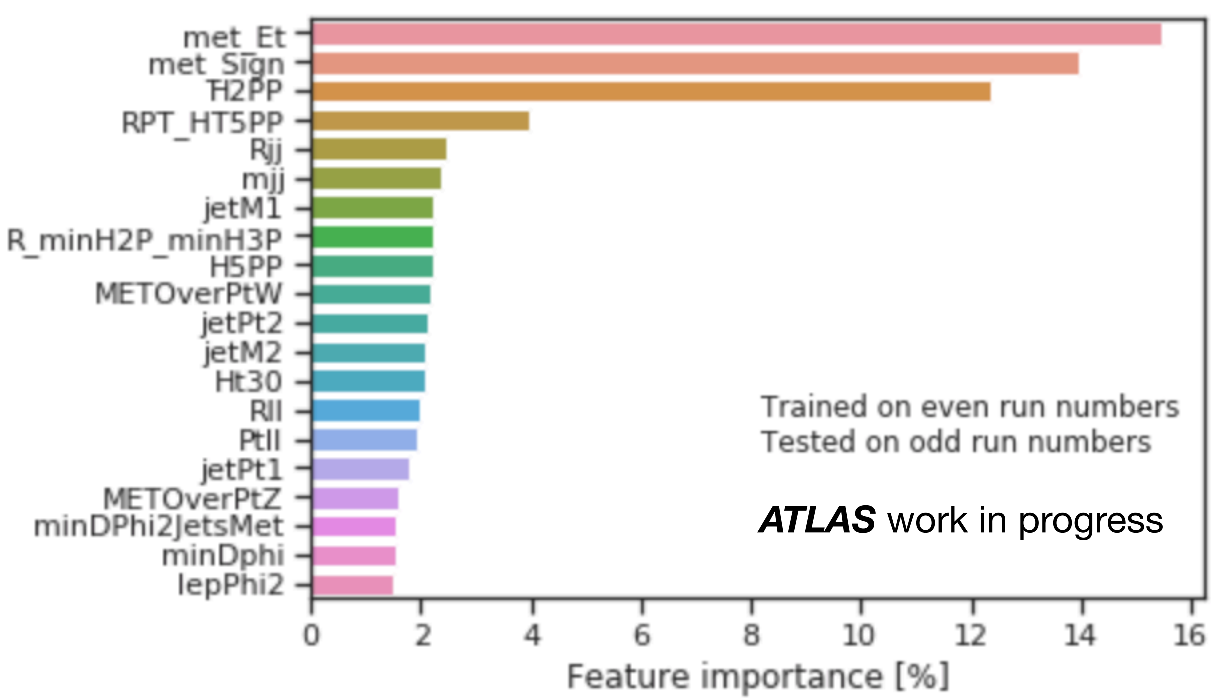
- **Background (MC)**
  - Z+jets, W+jets, diboson, triboson, higgs, ttbar, single top, Wt, 3 top, 4 top
  - Randomly draw ~60,000 events that pass preselection

- **Train 2 BDTs**
  - Odd run numbers — tested on even run numbers
  - Even run numbers — tested on odd run numbers

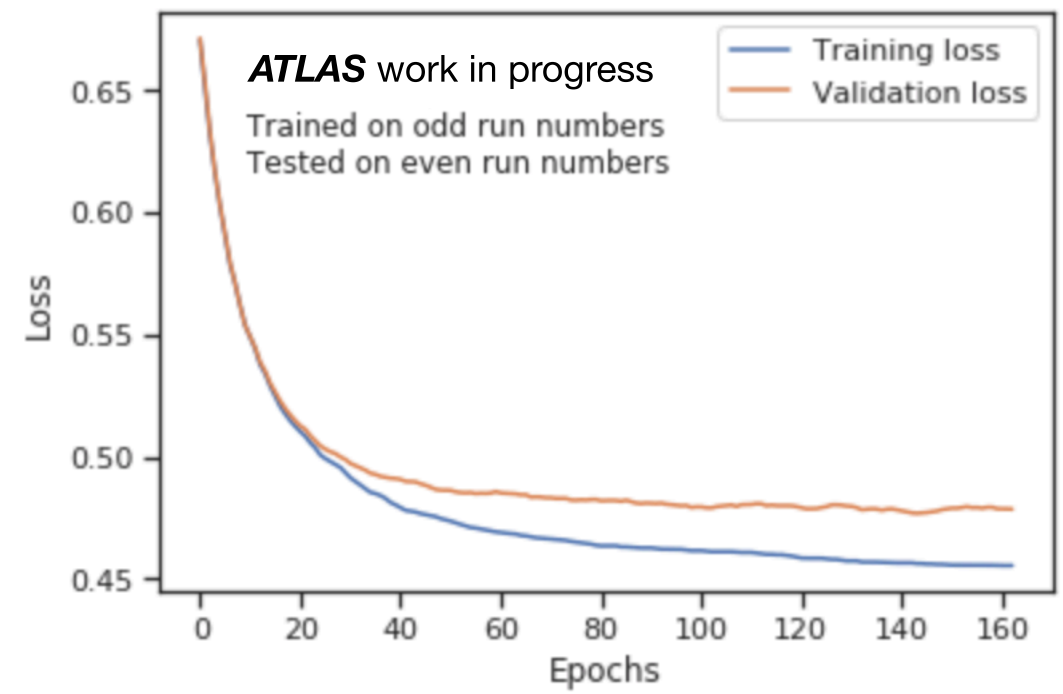
# Training the BDTs

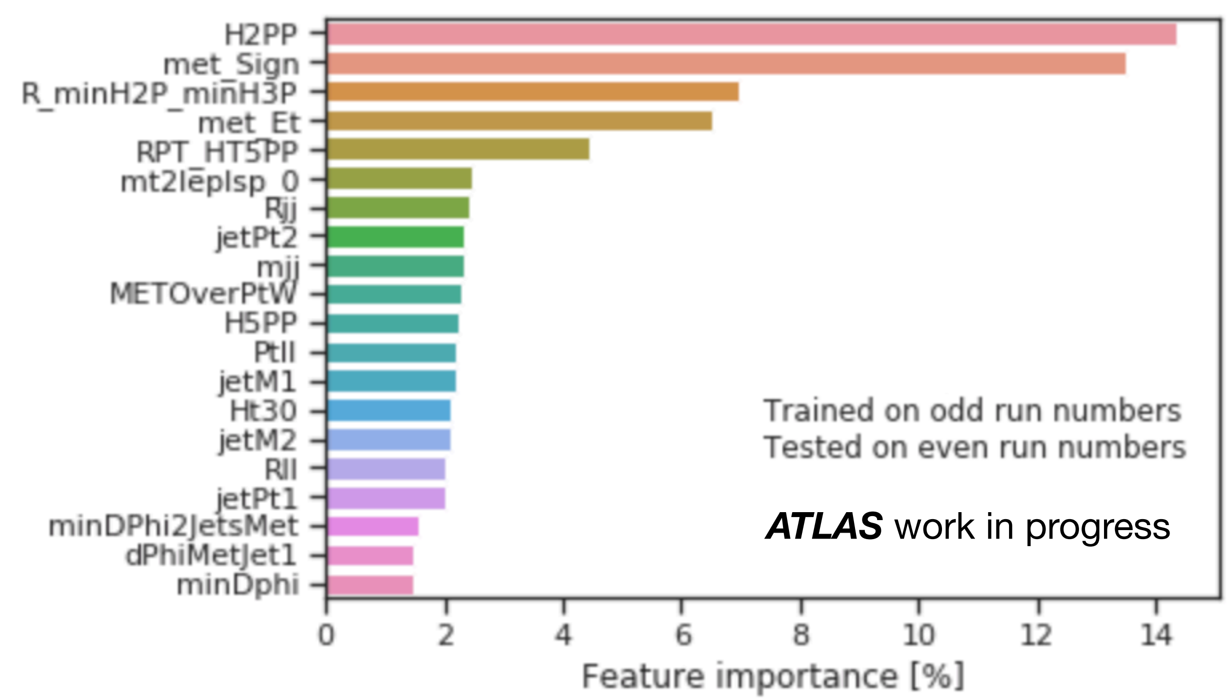
Train even
Test odd



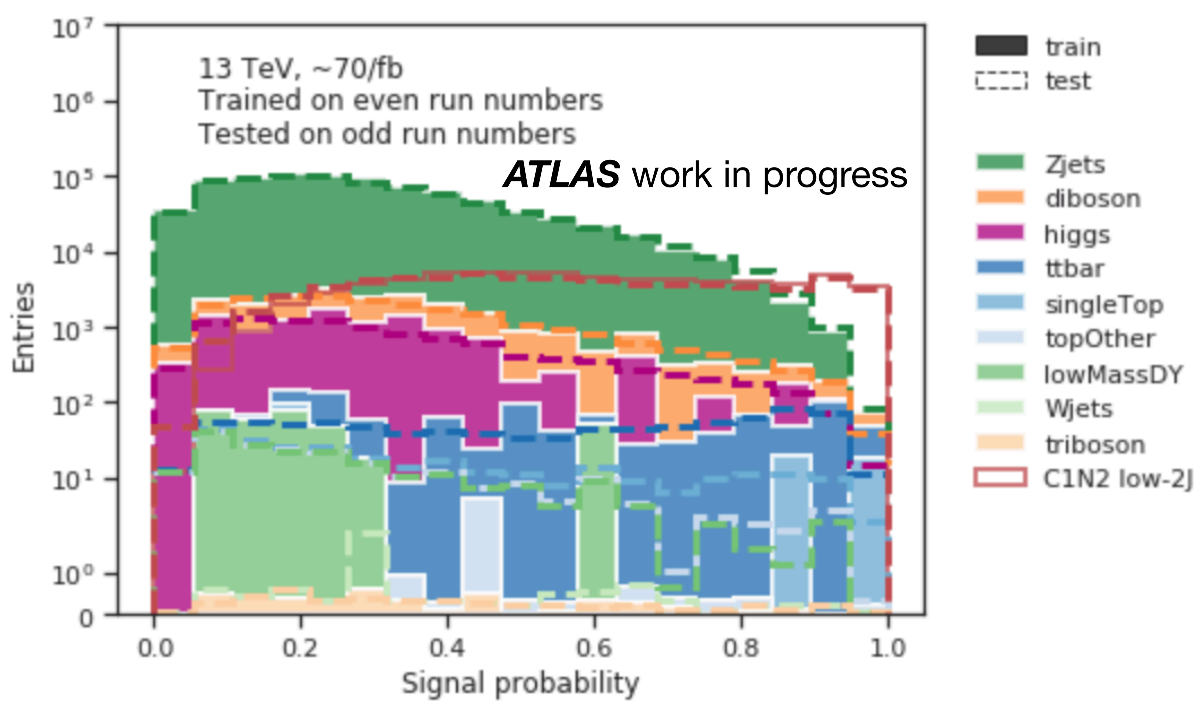

Train odd
Test even



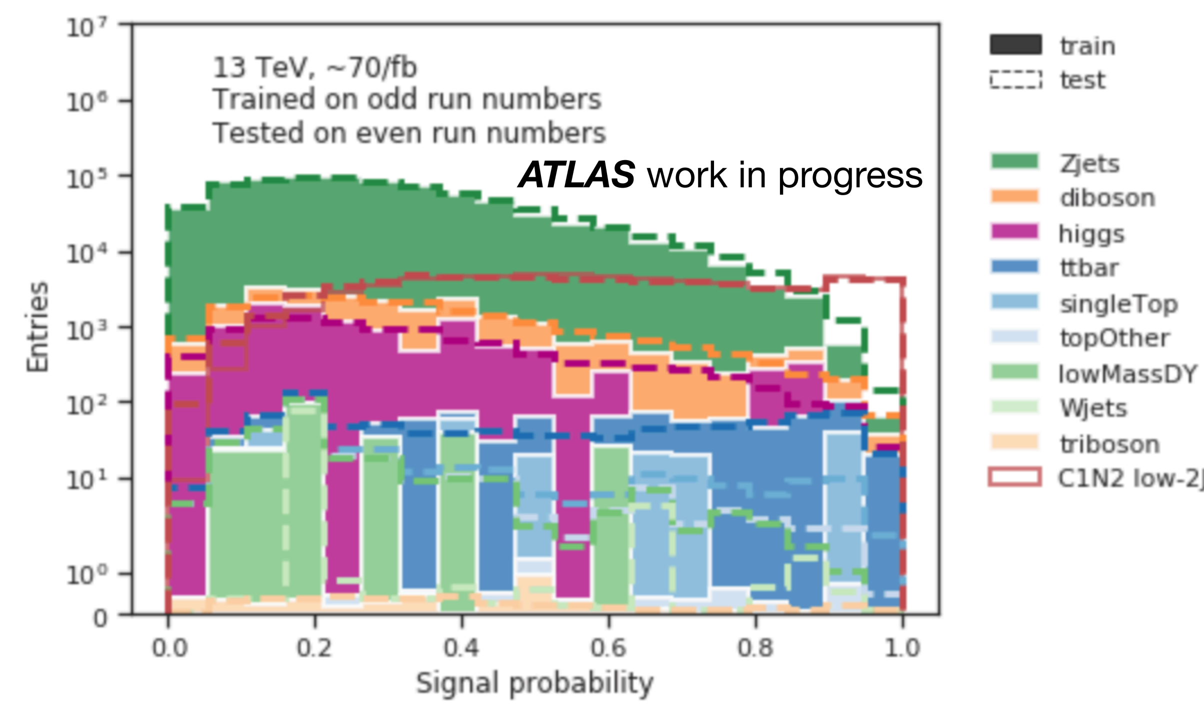



# Training vs. test results

Train even
Test odd



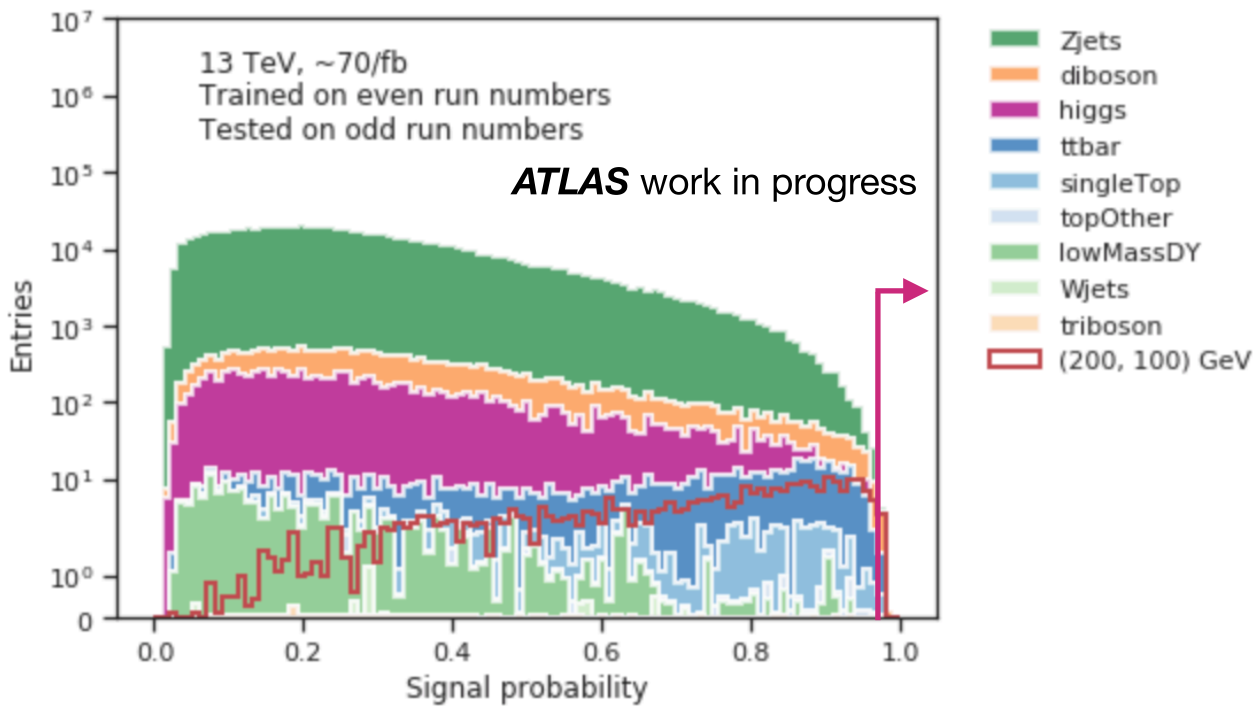

Train odd
Test even



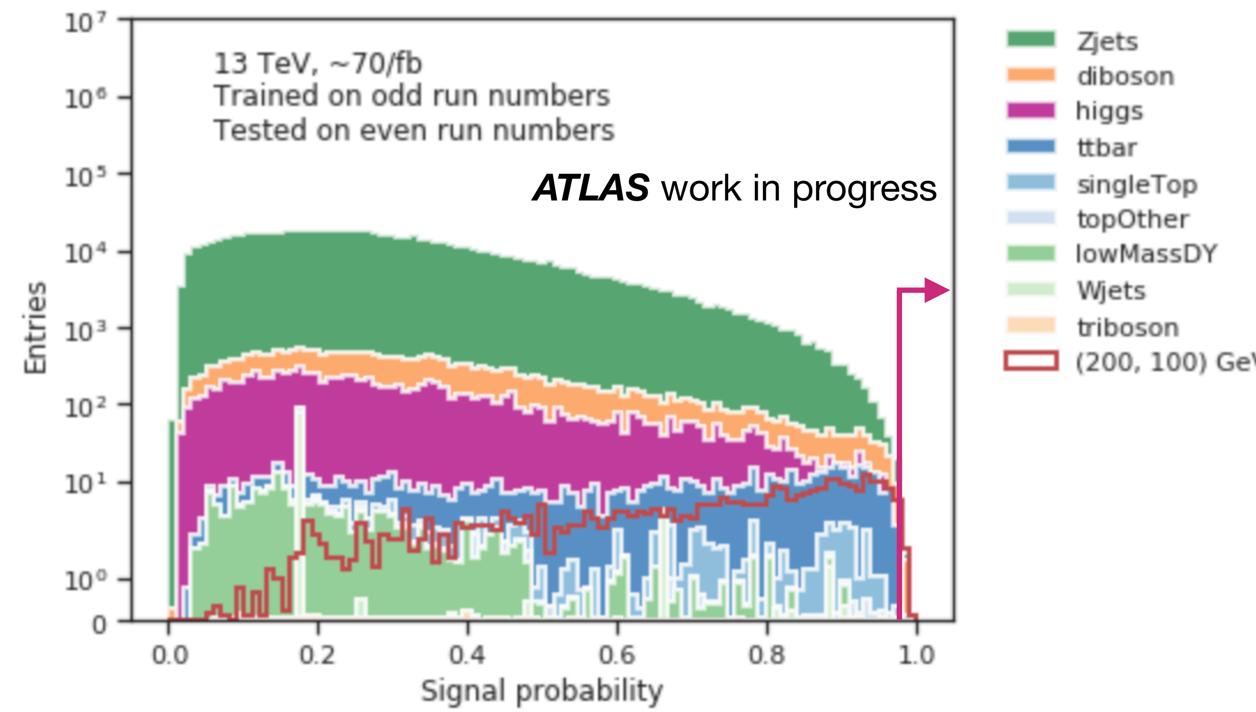

- Test scores are only slightly less accurate than the training scores
  - **No significant overtraining observed**

# Test results

Train even
Test odd




Train odd
Test even

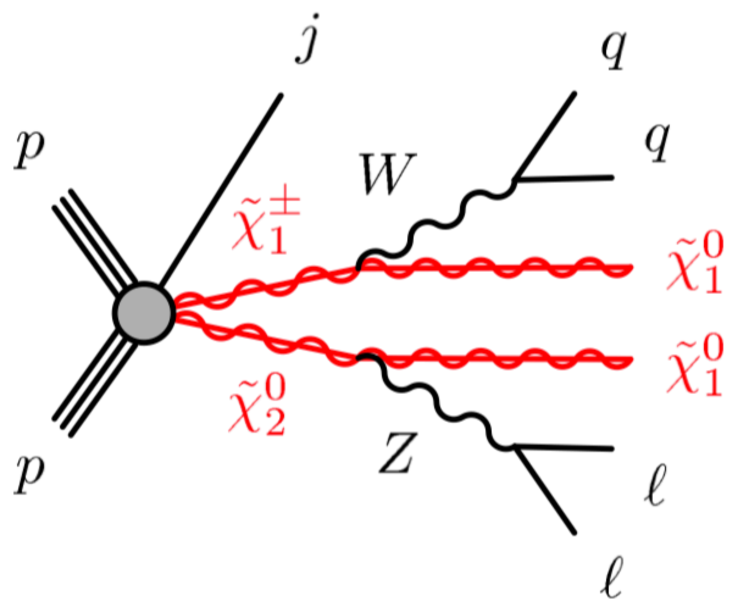



| Odd run no. | Signal prob. > 0.97 |
|---|---|
| **# signal events** | 2.92 |
| **# total bkg. events** | 4.43 |
| **Significance (30% sys.) [σ]** | **0.83** |

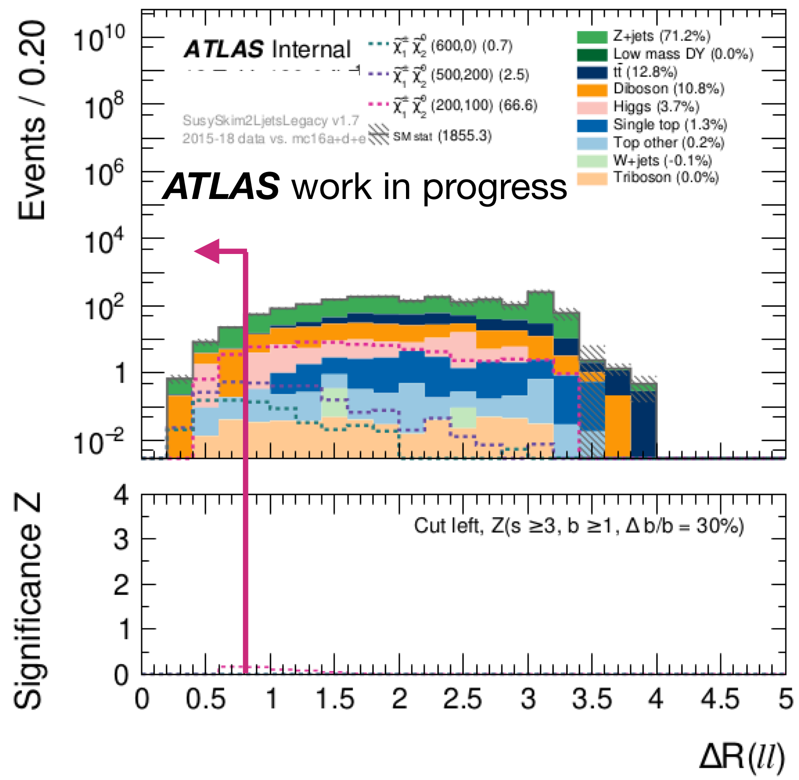
| Even run no. | Signal prob. > 0.97 |
|---|---|
| **# signal events** | 7.39 |
| **# total bkg. events** | 18.70 |
| **Significance (30% sys.) [σ]** | **0.76** |

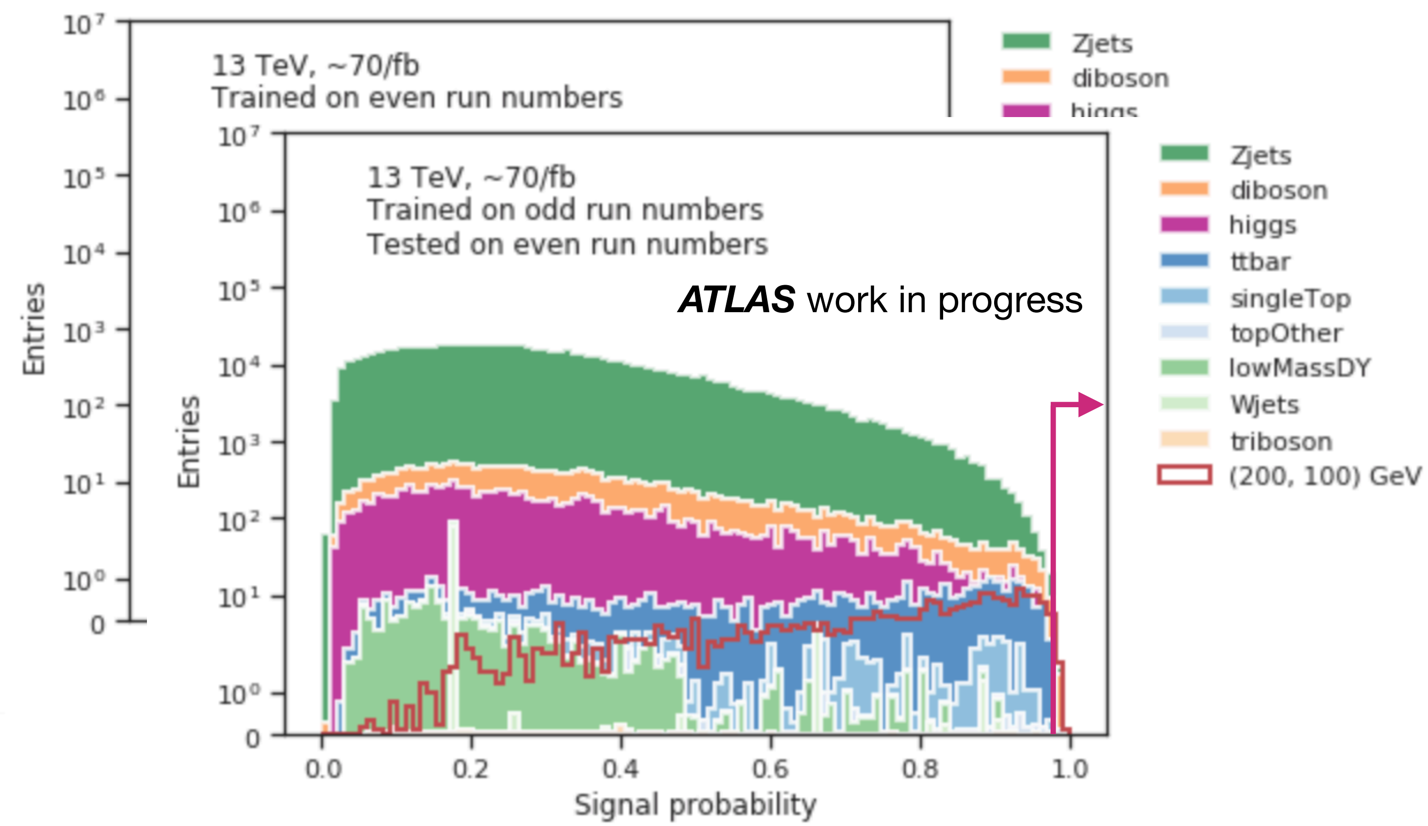
# Conventional cut-and-count region

| Preselection + | SRLow |
|---|---|
| **$E_T^{miss}$ sig.** | (6, 9) |
| **$\Delta\phi(p_{T,ll}, E_T^{miss})$** | $< 0.6$ |
| **$\Delta R(l,l)$** | $< 0.8$ |



# Comparison with conventional analysis




| **SRLow** | **ΔR(l,l) < 0.8** |
| --- | --- |
| **# signal events** | 4.22 |
| **# total bkg. events** | 32.00 |
| **Significance (30% sys.) [σ]** | **0.17** |

| **Even + odd run no.** | **Signal prob. > 0.97** |
| --- | --- |
| **# signal events** | 10.31 |
| **# total bkg. events** | 23.13 |
| **Significance (30% sys.) [σ]** | **0.91** |

**Need 1.64σ to exclude at 95% CL → no exclusion sensitivity expected yet**

# Conclusions

- Small Δm-scenarios are hard to target using conventional cut-and-count
- By estimating a signal probability using an XGBoost BDT, we can increase the sensitivity
- Crude feasibility study not sensitive enough to reach exclusion sensitivity for the benchmark point of m($\tilde{\chi}^{\pm}_1/\tilde{\chi}^0_2, \tilde{\chi}^0_1$) = (200, 100) GeV
- Will most likely benefit from a multi-bin/shape fit to the BDT output distribution

**Outlook/ideas**

- My initial attempts to train a neural network (NN) have not succeeded in matching the performance of the XGBoost BDT
- Would be interesting to train a NN that interpolates between sparticle masses
- Maybe multiclass classification of several background processes could help?

# LHC SUSY cross-section references

**Colored Sector**

- NLO + NLL Tool for 13, 14, 33 and 100 TeV cross sections.
- Squark and gluino production cross sections in pp collisions at \sqrt(s) = 13, 14, 33 and 100 TeV, C. Borschensky, M. Kramer, A. Kulesza, M. Mangano, S. Padhi, T. Plehn, X. Portell, arXiv:1407.5066, published in Eur.Phys.J. C74 (2014) 12.

- NNLO_approx + NNLL Tool for 13 TeV cross sections, uses PDF4LHC15_MC PDF sets
- NNLL-fast: predictions for coloured supersymmetric particle production at the LHC with threshold and Coulomb resummation, Wim Beenakker, Christoph Borschensky, Michael Krämer, Anna Kulesza, Eric Laenen, arXiv:1607.07741, published in JHEP 1612 (2016) 133.
- The full list of references can be found here.

**Electroweak Sector**

NLO+NLL threshold resummed results from Resummino.

For slepton production:

- G. Bozzi, B. Fuks, and M. Klasen, Nucl. Phys. B 777, 157 (2007)
- B. Fuks, M. Klasen, D. R. Lamprea, and M. Rothering, Eur. Phys. J. C 73, 2480 (2013)
- B. Fuks, M. Klasen, D. R. Lamprea and M. Rothering, JHEP 01 (2014) 168
- J. Fiaschi and M. Klasen, JHEP 03 (2018) 094
- W. Beenakker et al., Phys. Rev. Lett. 83 (1999) 3780, Erratum: Phys. Rev. Lett. 100 (2008) 029901

For gaugino production:

- J. Debove, B. Fuks, and M. Klasen, Nucl. Phys. B 842, 51 (2011)
- B. Fuks, M. Klasen, D. R. Lamprea, and M. Rothering, JHEP 10 (2012) 081
- B. Fuks, M. Klasen, D. R. Lamprea, and M. Rothering, Eur. Phys. J. C 73, 2480 (2013)
- J. Fiaschi and M. Klasen, Phys. Rev. D 98 (2018) 055014
- W. Beenakker et al., Phys. Rev. Lett. 83 (1999) 3780, Erratum: Phys. Rev. Lett. 100 (2008) 029901

(The last reference to the paper by W. Beenakker et al. is the reference for the original NLO results.)