Institute of High Energy Physics
Chinese Academy of Sciences

ML in SUSY DiTau and SS/3L analysis

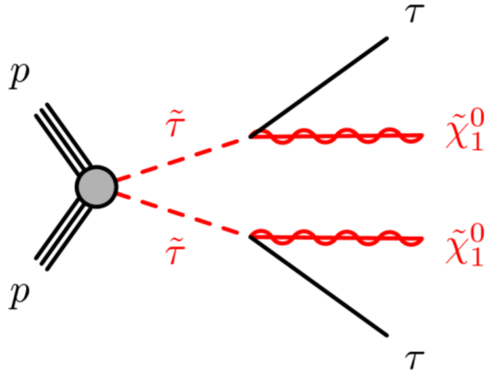
Jiarong Yuan
On behalf of ATLAS SUSY DiTau and SS/3L team

2022/9/13

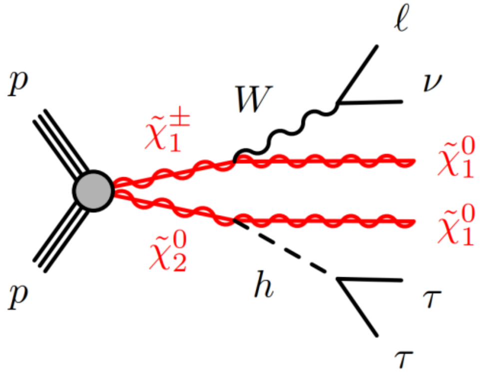
Overview

The application of machine learning in the search of the following SUSY scenarios:

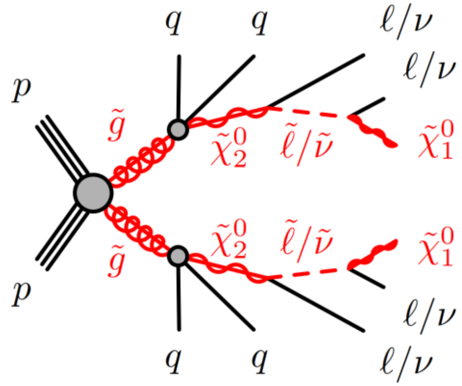
1. Direct stau production [Dominic Jones]



2. C1N2 via Wh to $2\tau1\ell$ [Chenzheng Zhu]

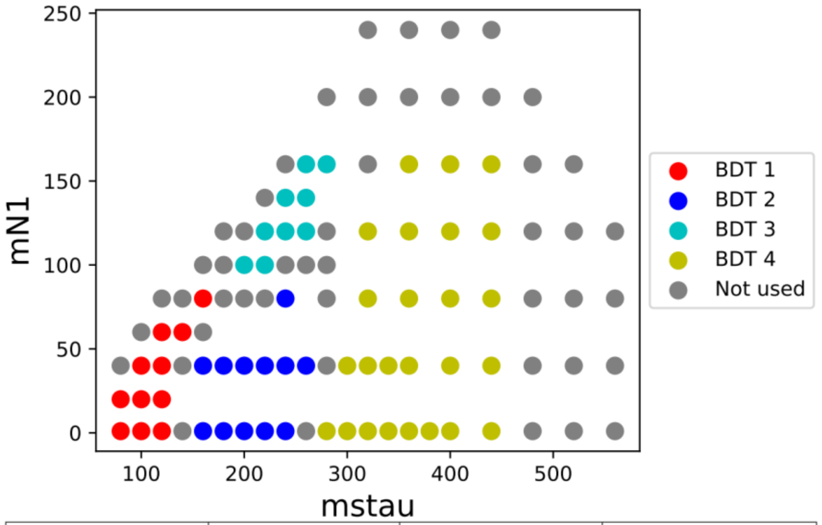


3. GG 2-step via sleptons [Xin Wang]

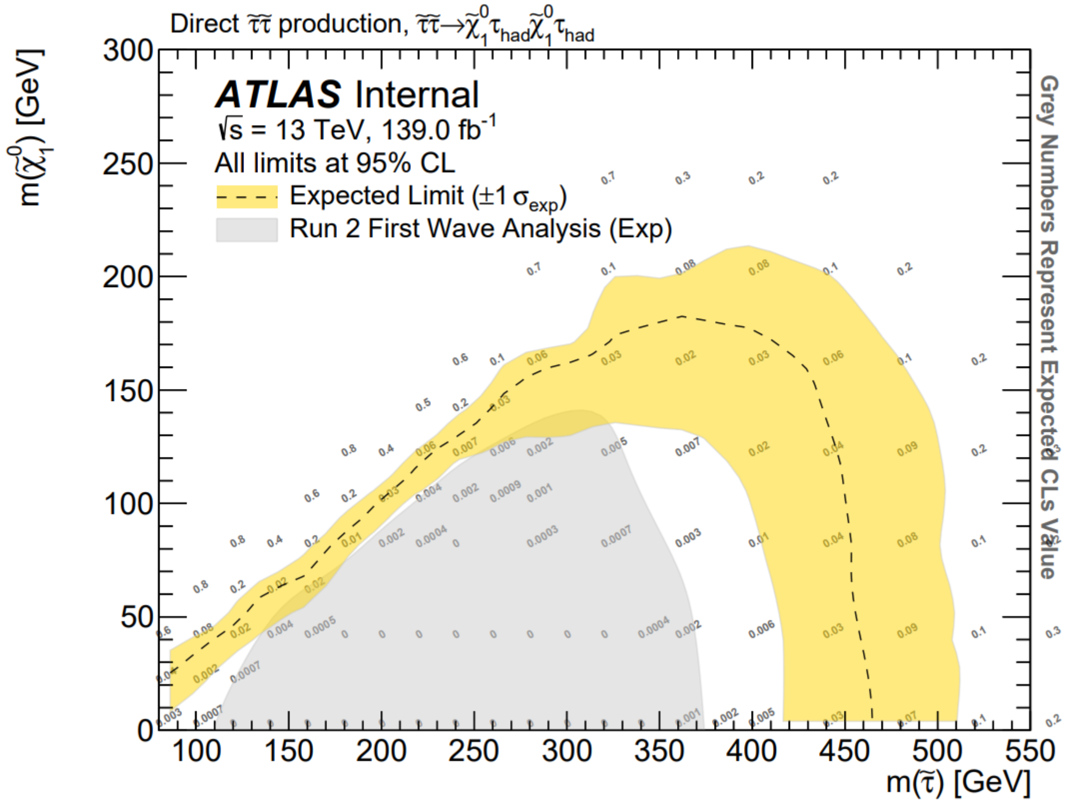


Direct stau production

- The current 2nd wave analysis exclusion limit is improved because of
  - Tau-ID: MVA → RNN
  - Signal region optimization: CutCount → BDT
- Samples are grouped into 4 groups using distances given by UMAP(based on kinematics).
  - adding more points = trade of between kinematics and statistics
- 4 BDTs are trained across the whole grid, for each signal point use BDT which gives best CLs value.



SR BDT 1	SR BDT 2	SR BDT 3	SR BDT 4
2 medium τ, OS			
asymmetric ditau trigger			
3rd baseline τ veto			
light lepton veto			
b-veto			
$m_{T2}(\tau_1, \tau_2) > 30$ GeV			
$p_T(\tau_1) > 95$ GeV			
$p_T(\tau_2) > 65$ GeV			
$m_{inv}(\tau_1, \tau_2) > 100$ GeV			
BDT1score > 0.85	BDT2score > 0.9	BDT3score > 0.96	BDT4score > 0.985



With flat 15% systematic

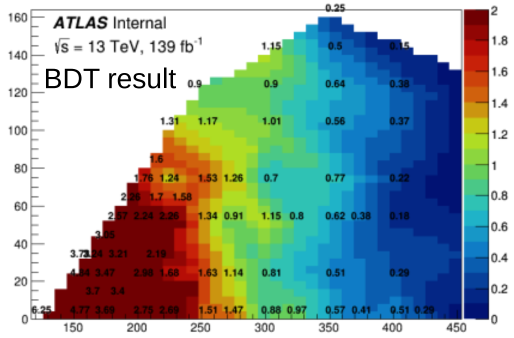
C1N2 via Wh to $2\tau 1\ell$

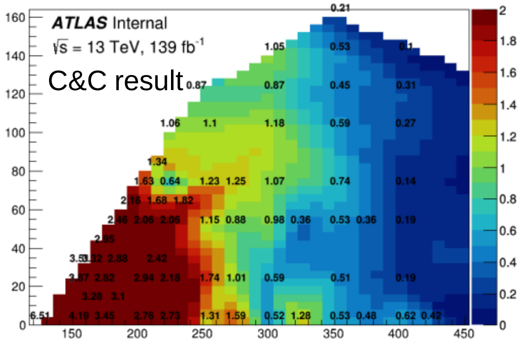
- The signal regions optimized by CutCount:

Preselection	
Pass single lepton trigger	1 base lepton, 1 signal lepton
>= 2 medium tau(OS)	B-jet Veto
$\Delta\phi(\tau_1, \tau_2)$ < 3	m_{T2} > 30 GeV
40 GeV<$m(\tau_1, \tau_2)$<160 GeV	

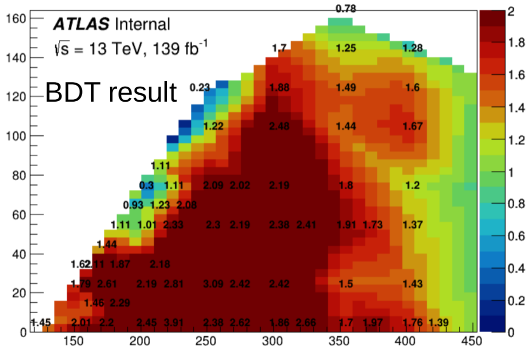
SR-Low	SR-High
>= 2 Tight tau	m_{T2} > 90 GeV
90 GeV < $m(\tau_1, \tau_2)$ < 130 GeV	M_{Tlep} > 80 GeV
$P_{T\tau_1}$ > 60 GeV	$\Delta R(\tau_1, \tau_2)$< 2.2
$P_{T\tau_2}$ > 30 GeV	70 GeV<$m(\tau_1, \tau_2)$<140 GeV
m_{T2} >70 GeV	M_{Tsum}> 450 GeV

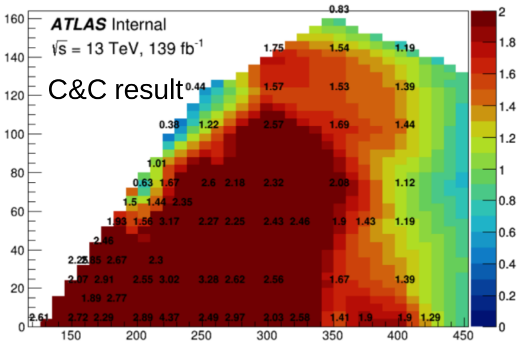
- Two BDTs are trained, one for SR-HighMass and the other for SR-LowMass.
  - SR-LM: BDT results have slightly larger significance than CutCount results.





  - SR-HM: The results are similar as CutCount results due to simple tree.



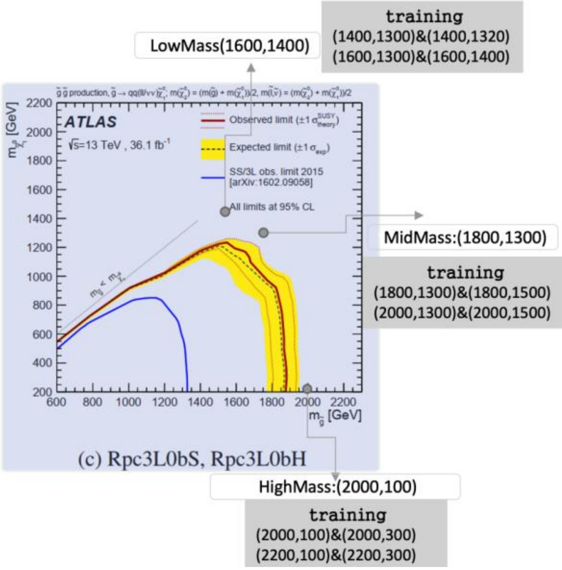


GG 2-step via sleptons

- The discovery signal regions are optimized by **CutCount**:

RpcGGslep1	RpcGGslep2	RpcGGslep3
$N_{\text{lept}}^{\text{signal}} \geq 3$ $N_{\text{jets}}^{40\text{GeV}} \geq 4$ $N_{\text{b-jets}}^{20\text{GeV}} == 0$ veto $81 < m_{\text{SFOS}} < 101$ GeV		
$E_{\text{T}}^{\text{miss}} / \sum p_T^j \geq 0.4$	$E_{\text{T}}^{\text{miss}} / \sum p_T^j \geq 0.3$	-
$E_{\text{T}}^{\text{miss}} / \sum p_T^{\text{lep}} \geq 1.4$	$\Delta\phi(\ell 1 \ell 2, E_{\text{T}}^{\text{miss}}) > 0.7$	-
$p_{\text{T}}^{\ell 2} \geq 30$ GeV	$p_{\text{T}}^{\ell 2} \geq 70$ GeV	$\sum p_T^j \geq 1200$ GeV
-	$E_{\text{T}}^{\text{miss}} \geq 150$ GeV	$E_{\text{T}}^{\text{miss}} \geq 100$ GeV

- Train 3 **BDT**s are trained for different regions:

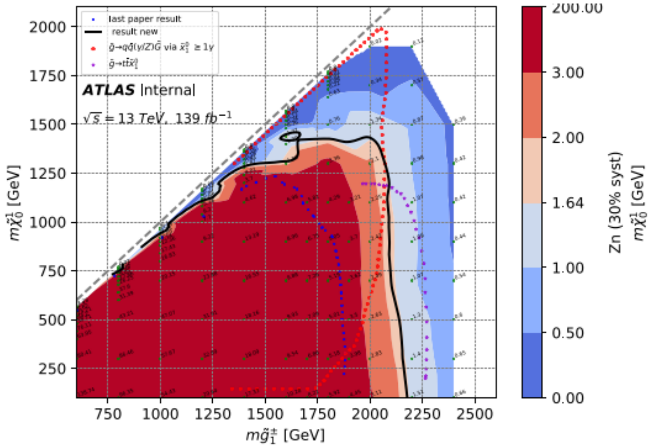


(c) Rpc3L0bS, Rpc3L0bH

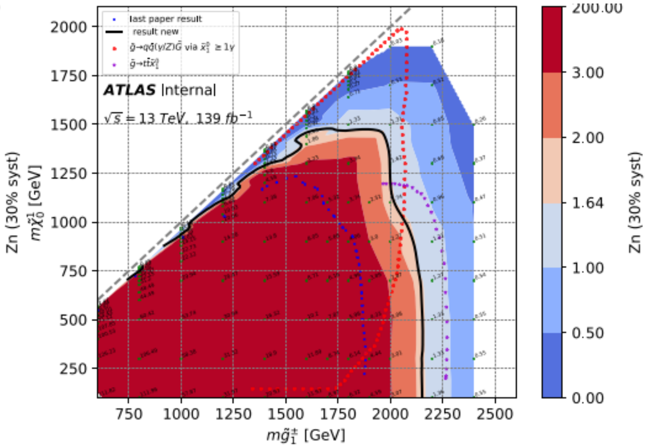
SR-Low	SR-Mid	SR-High
N-SigLep ≥ 3 N-Jet25 ≥ 4 B-veto Z-veto		
BDT-score > 0.85	BDT-score > 0.90	BDT-score > 0.75
$Pt_{subl} > 30$ GeV	-	$E_T^{miss} \geq 100$ GeV

- Small improvement in the compressed region

CutCount



BDT



Summary

1. For the signal with small cross-section(Direct stau production), challenging signal(GG 2-step via slepton compressed region), the results of BDT method are better than the results of CutCount.
2. For general processes(C1N2 via Wh, GG 2-step via slepton), the results of BDT method is similar with the results of CutCount. This may be due to the use of simple BDT.
3. In the new challenging compressed ditau search, we plan to use machine learning to get better results.

The End