Search for supersymmetry with τ leptons in the CMS experiment using MVA approach

Ilya Bobovnikov, Alexis Kalogeropoulos,

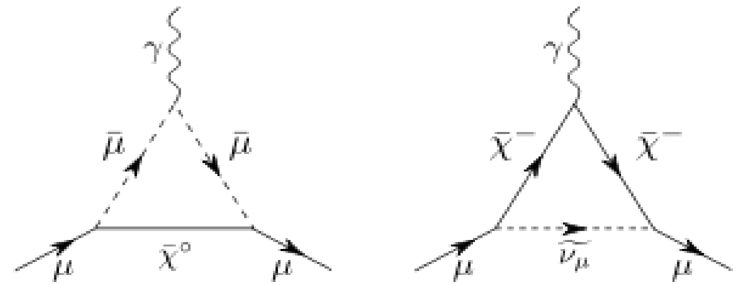
Isabell Melzer-Pellmann, Alexei Raspereza, Ceren Güzelgün, on behalf of the CMS collaboration

Outline

- Motivation for the light stau
- Background estimation strategy
- Cut based approach and results for stau search
- MVA (BDT) strategy
- Comparison of Cut based and BDT performance
- Conclusions

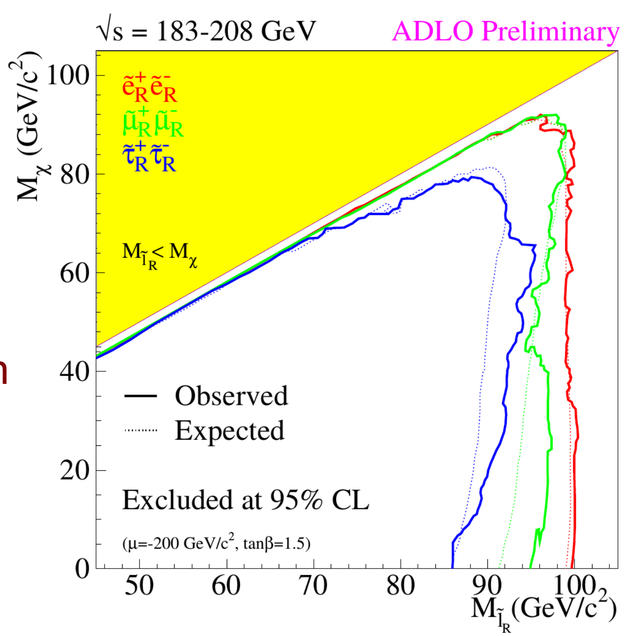
Introduction

- LSP co-annihilation with light stau could bring the neutralino relic density to the observed value
- SUSY can explain ~3σ deviation of muon g – 2 from SM prediction → light electroweak sector
- Likelihood analysis of experimental constraints predicts light staus arXiv:1710.11091v2



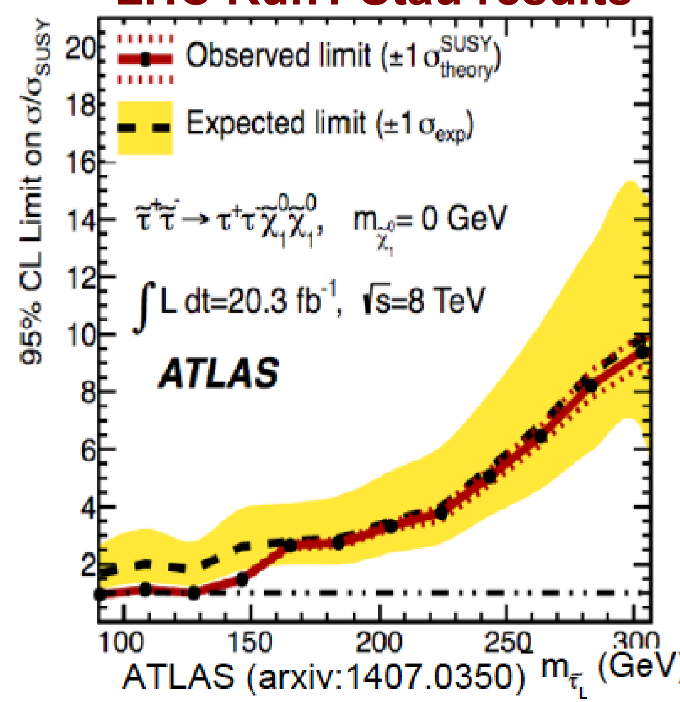
LEP results

LEP put strong limits on slepton masses



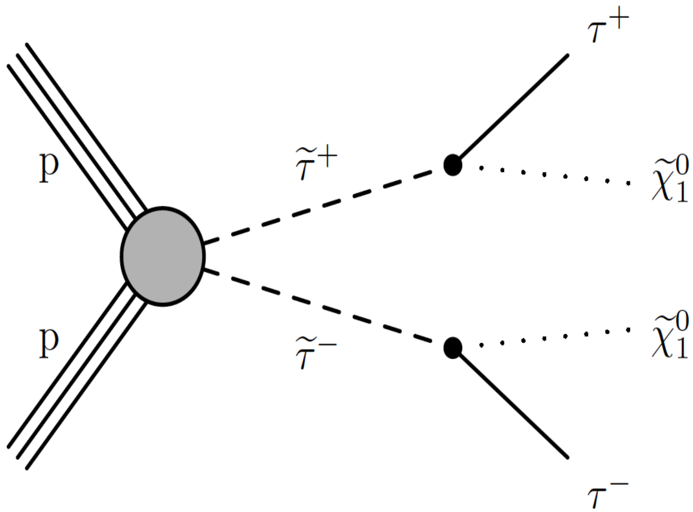
Only one additional point excluded at 8TeV @ 20 fb^{-1}. Analysis is challenging!

LHC Run1 Stau results

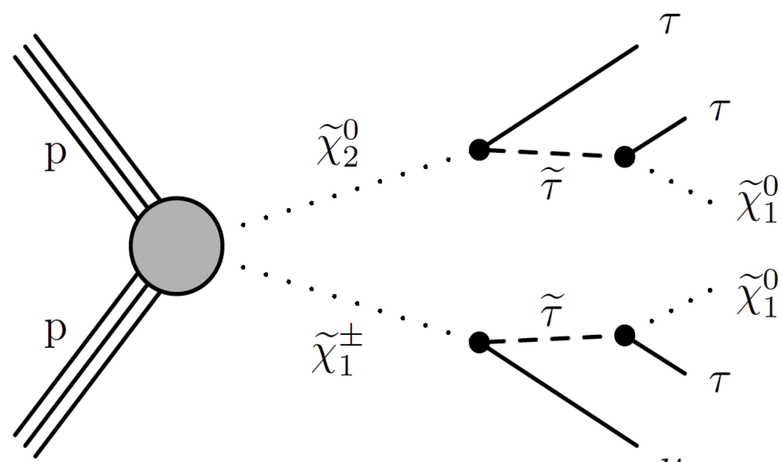


Models and topologies

direct stau production



indirect stau production



Experimental signature

- Missing transverse energy depends on the model parameters
- Small number of jets and no b-tagged jets

Channel	Signature	$\mathcal{BR}$
$0-\ell$	$2\tau_h + \not{E}_T$	$0.65^2 = 0.42$
$1-\ell$	$\tau_\ell \tau_h + \not{E}_T$	$2 \times (0.35 \cdot 0.65) = 0.46$
$2-\ell$	$2\tau_\ell + \not{E}_T$	$0.35^2 = 0.12$

Covered by this analysis
$\mu\tau_h$, $e\tau_h$, $e\mu$

SUS-17-002

Search strategy $\mu\tau_h$, $e\tau_h$, $e\mu$

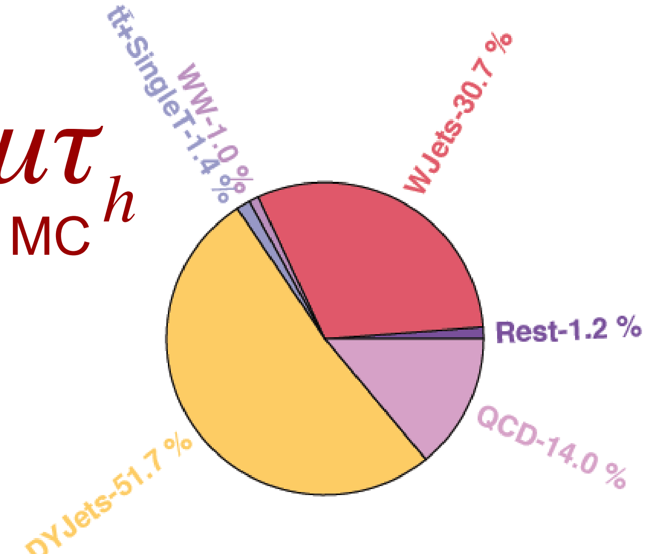
SUS-17-002

Selection (36 /fb of Run 2 2016 data)

- Opposite charge pair of identified isolated leptons
- No additional leptons
- Number of jets ≤ 1, number of b-tagged jets = 0

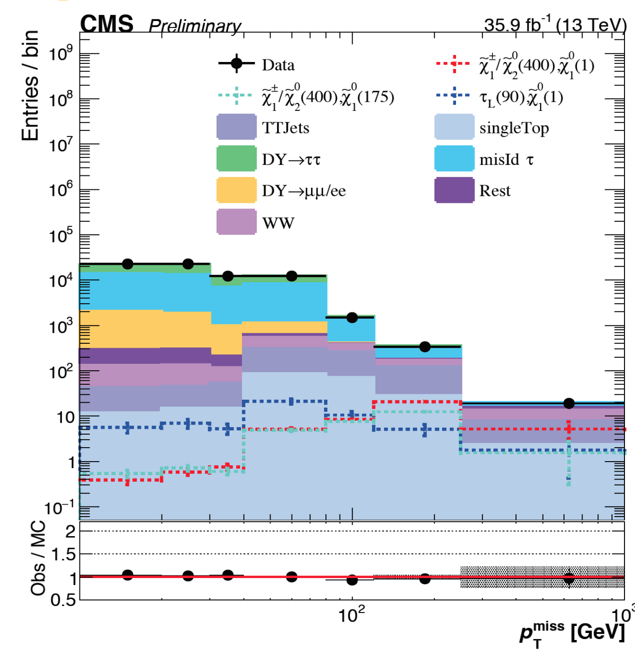
$\mu\tau_h$

MC



Background estimation

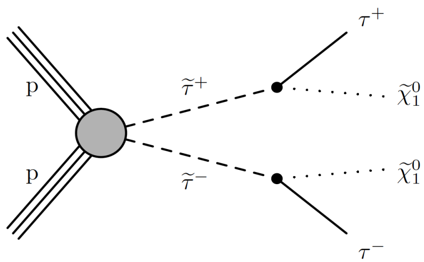
- Z+jets and top pair production: shape from MC is corrected by data in dimuon (mu-el) CR and scaled to normalization from data CR
- Jets misidentified as τ_h (only for $\mu\tau_h$ and $e\tau_h$) and QCD multijet (only for $e\mu$): shape is estimated from data CR and transfer factor is calculated as a ratio of yields in orthogonal CRs
- Other rare backgrounds taken from simulation

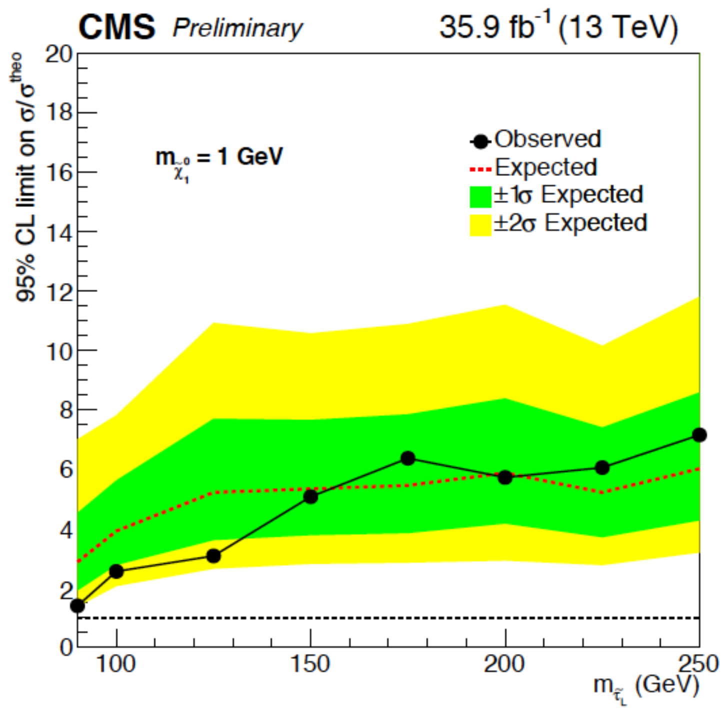


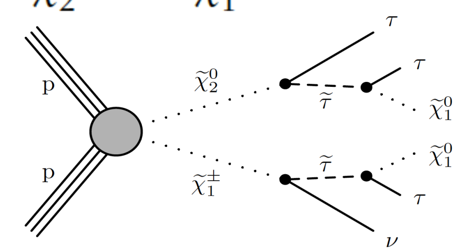
Interpretation $\mu\tau_h, e\tau_h, e\mu$

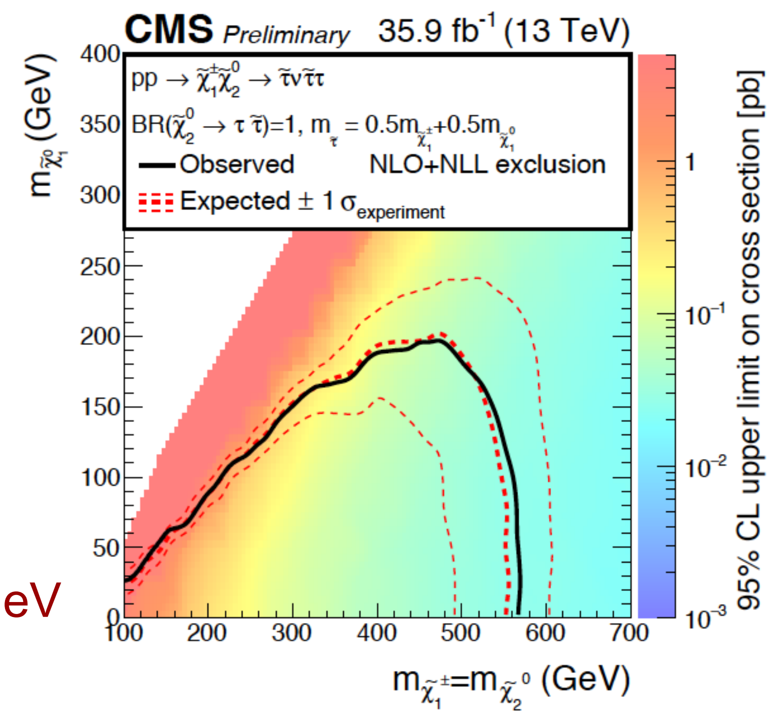
SUS-17-002

For cut based approach 132 search bins are defined with kinematic variables and jet multiplicity.
For chargino-neutralino (C1N2) interpretation $m_{\tilde{\tau}} = m_{\tilde{\chi}_1^0} + x(m_{\tilde{\chi}_2^0} - m_{\tilde{\chi}_1^0})$



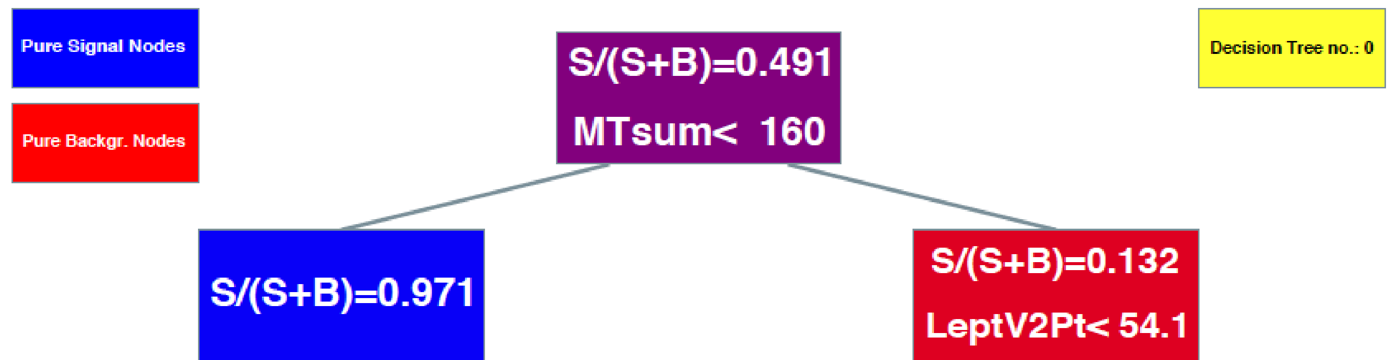


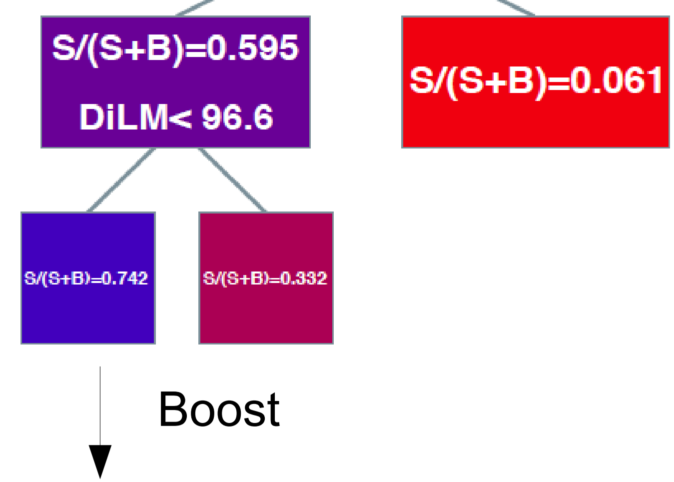




- No significant deviation in any signal region
- Exclude N2 and C1 decaying through staus below 560 GeV for ~ massless LSP
- Direct stau production not yet excluded due to low cross section
- For left-handed stau of around 90 GeV and a massless LSP we exclude 1.5 times the expected SUSY cross-section.

Studies for a BDT approach



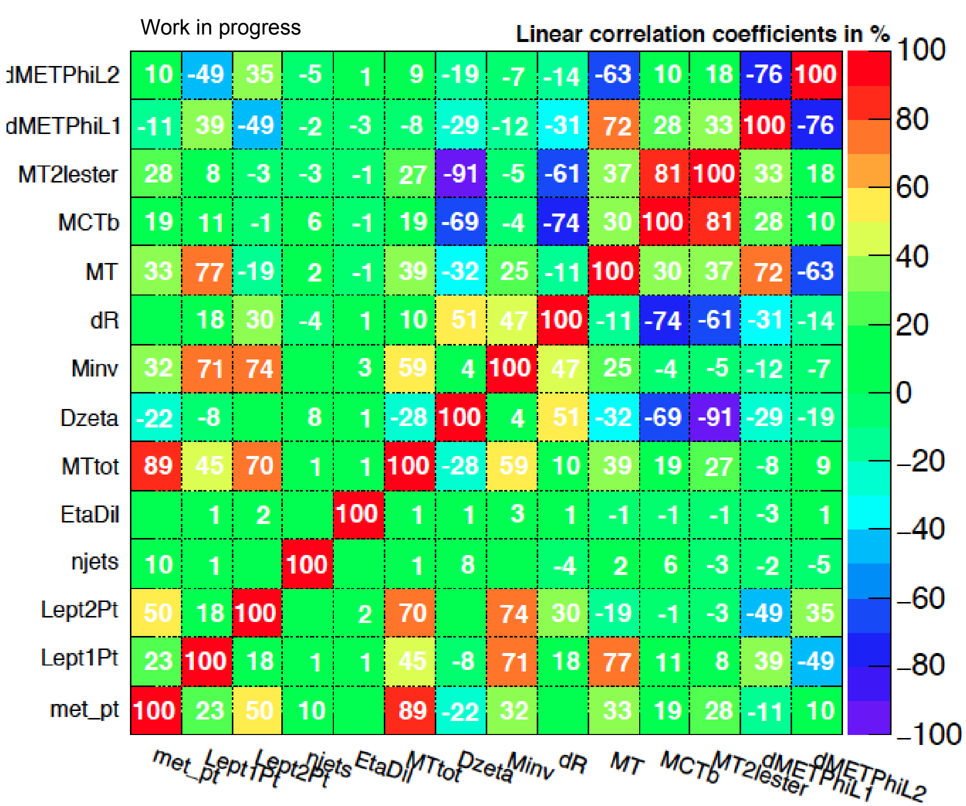


- Boosted Decision Tree (BDT)is used to improve the limit and reach sensitivity to direct stau production.
- Decision tree – sequence of binary splits using the discriminating variable.
- Boosting – reweighting the training data to stabilize response and enhance the performance
- Framework TMVA: http://tmva.sourceforge.net/

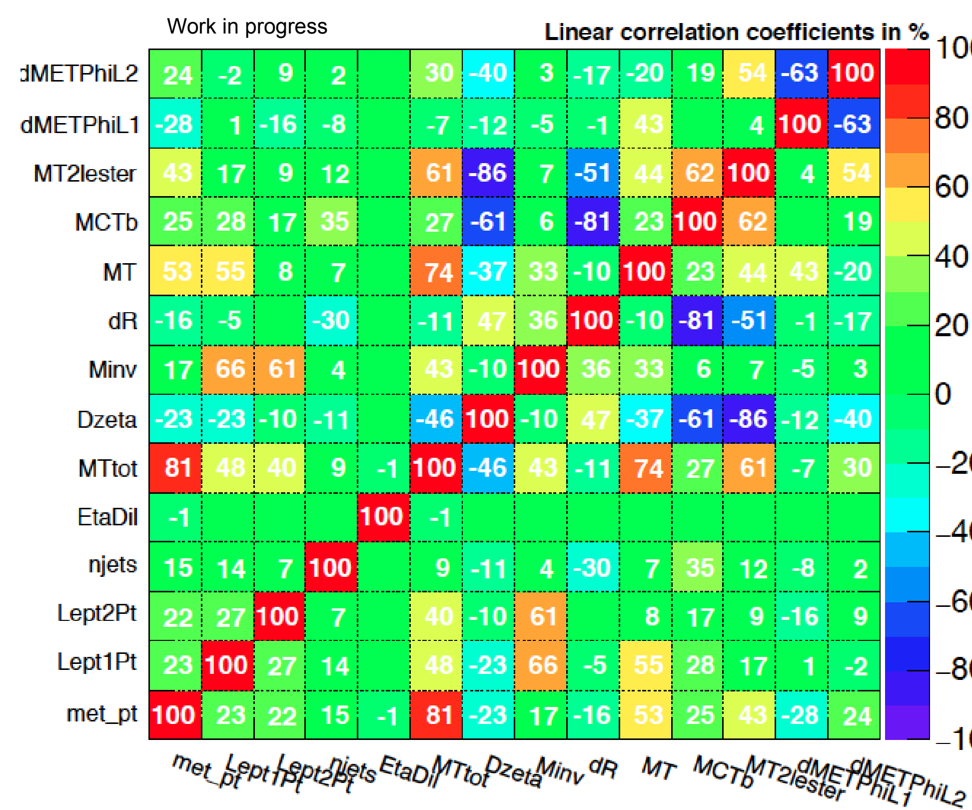
Variables and strategy

- Use 24% for training, 6% for testing and 70% events for limits evaluation
- BDT will be used only for $\mu\tau_h$ and $e\tau_h$
- Train BDT with the several benchmark stau scenarios (Stau 100,150,200,300 GeV and Lsp 1, 50,100 GeV)
- Optimize BDT options, input variables and search region
- Finally the 13 variables were selected

Correlation Matrix (signal)



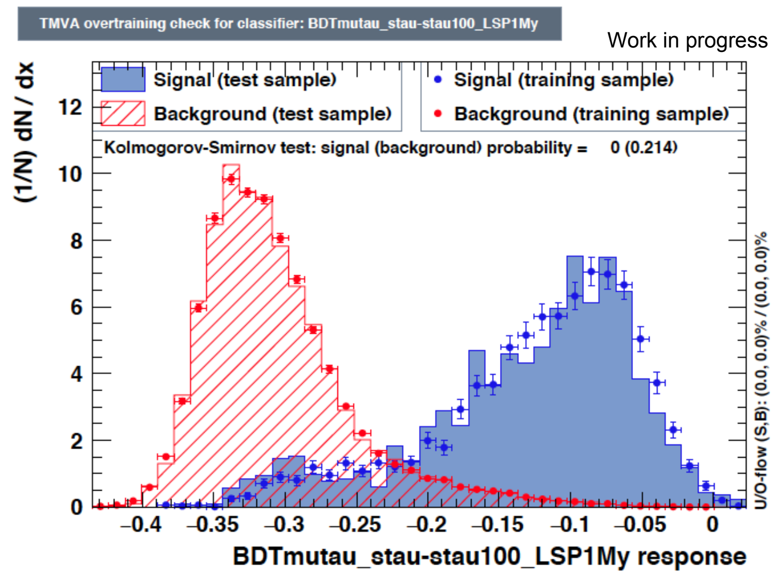
Correlation Matrix (background)

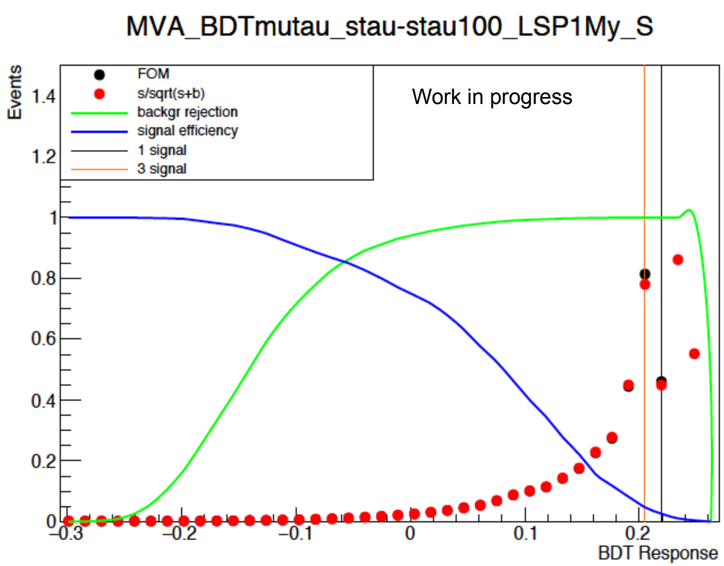


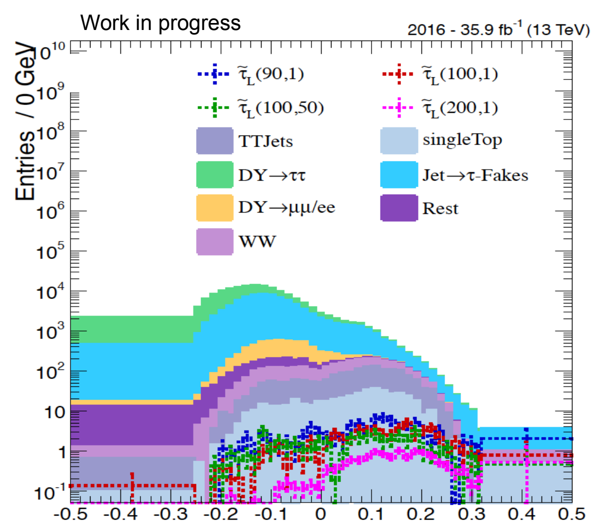
BDT training

- BDT output comparison between training and test samples shows no overtraining
- Binning, the most sensitive last bin, is selected by significance study

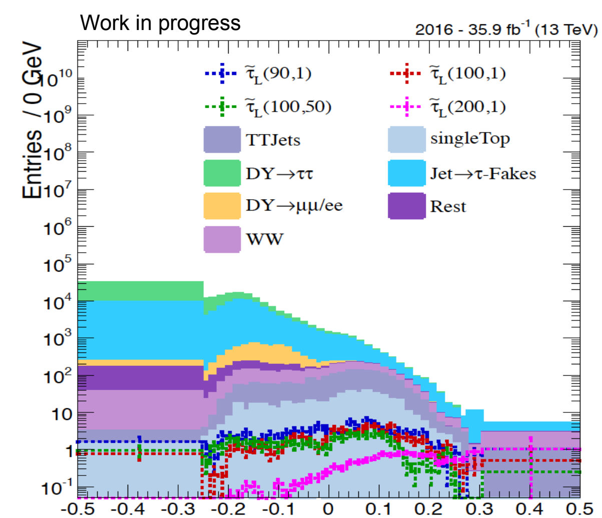
$$Z_A = \left[2 \left((s+b) \ln \left[\frac{(s+b)(b+\sigma_b^2)}{b^2 + (s+b)\sigma_b^2} \right] - \frac{b^2}{\sigma_b^2} \ln \left[1 + \frac{\sigma_b^2 s}{b(b+\sigma_b^2)} \right] \right) \right]^{1/2}$$







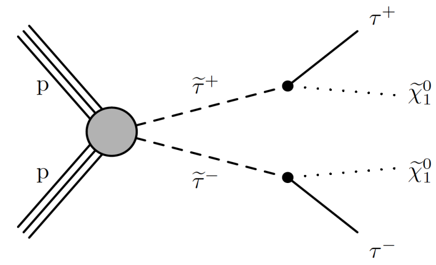
BDT Stau100_LSP1



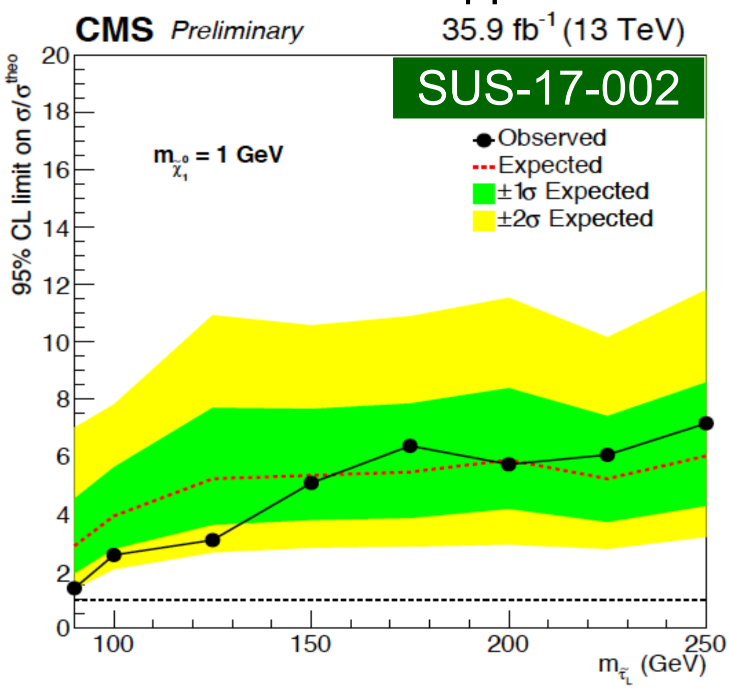
BDT Stau200_LSP1

Results

- A profile likelihood ratio in the asymptotic approximation is used as the test statistic. Limits are then calculated at the 95% confidence level (C.L.) using the asymptotic CLs criterion. The full BDT distribution shape is used in the combine tool.



Cut based approach



Expected limit

	Cut based	BDT (only mutau channel)
Stau mass 90	2.82	2.53
Stau mass 200	5.78	4.5

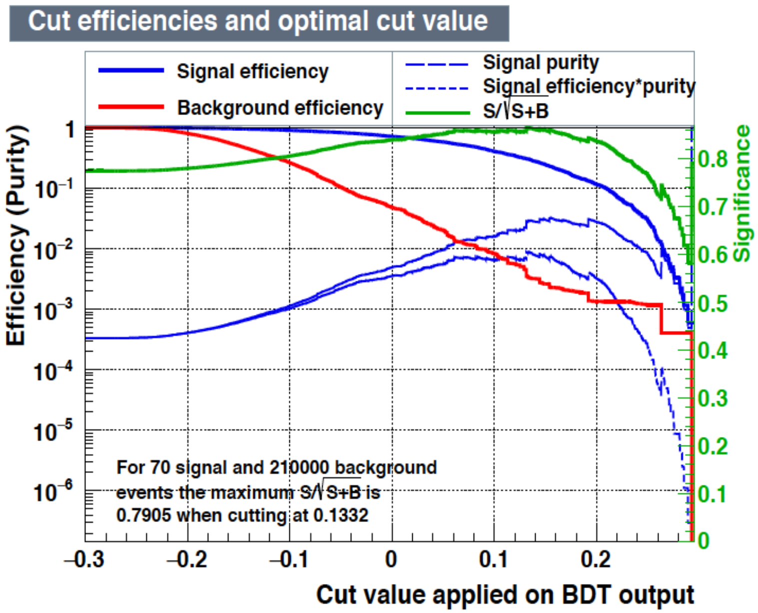
- For now BDT sensitivity is better (comparable) than for cut based approach → we will combine mutau and eltau channels
- Work in progress

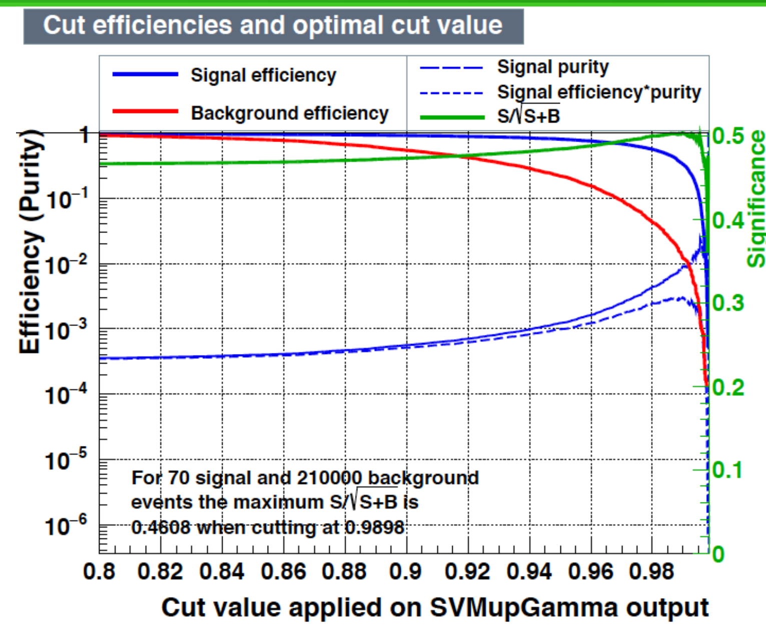
Conclusion

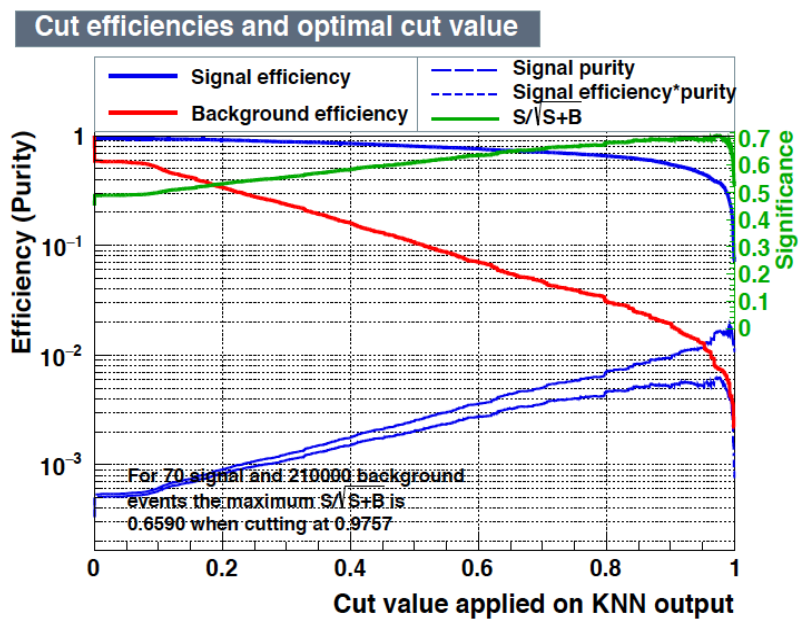
- Cut based search for direct and indirect tau production is being developed and was approved SUS-17-002
- This search is not sensitive to direct stau production
- To enhance sensitivity BDT approach is being developed
- Plan to get better sensitivity for BDT than for cut based approach and include 2017 data

Backup

MVA efficiency

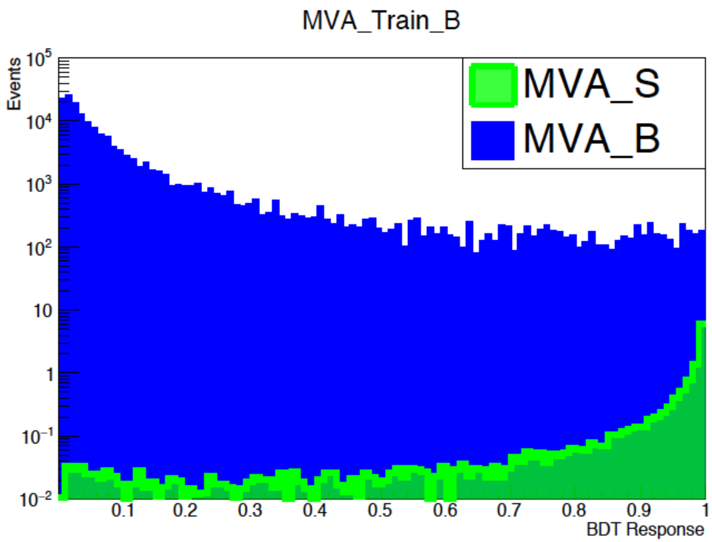




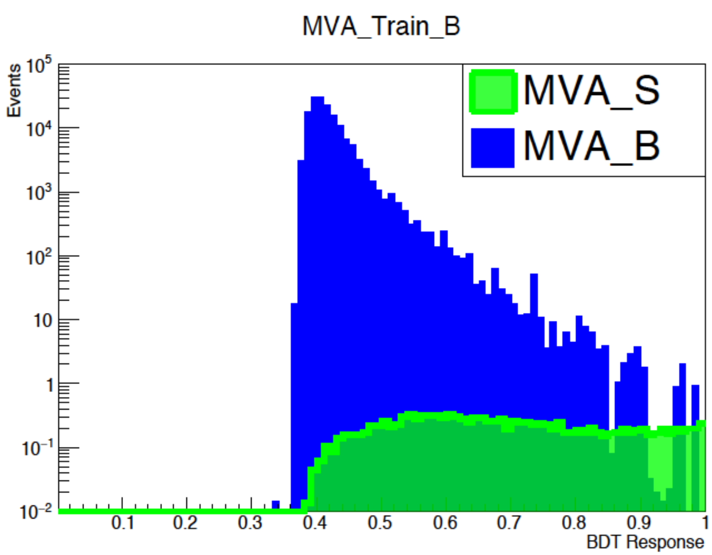


Including ANN and DNN

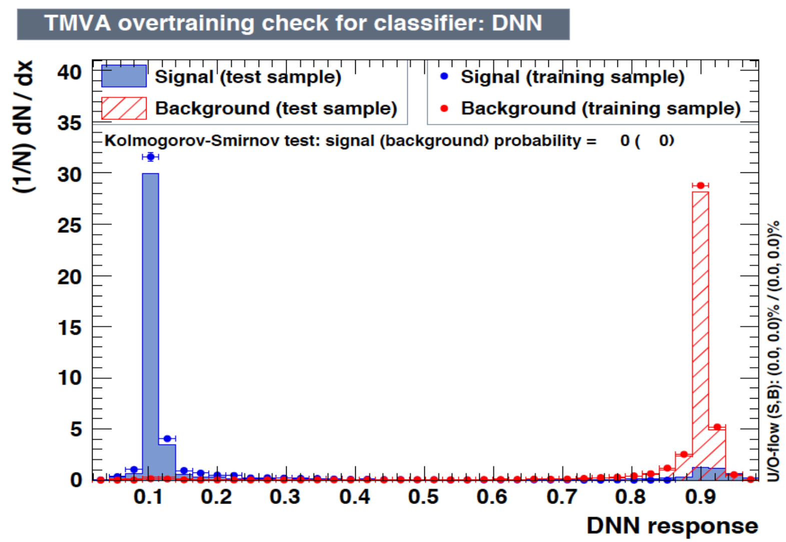
ANN

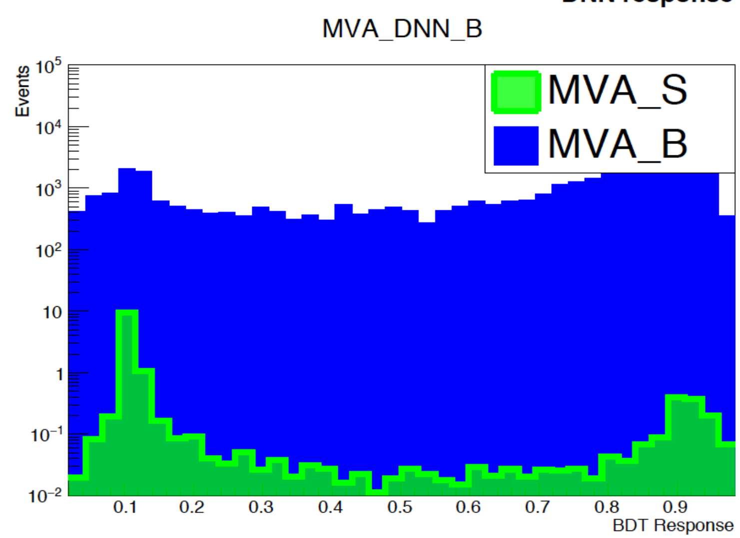


Boosted ANN



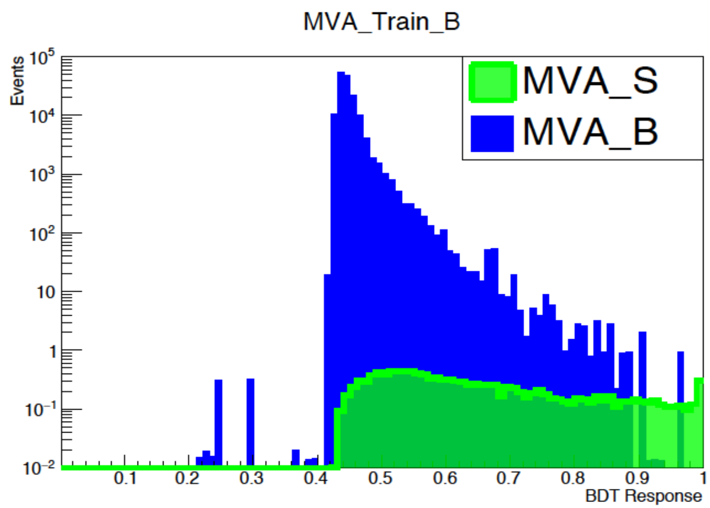
DNN (only after `root-6.08.00`)

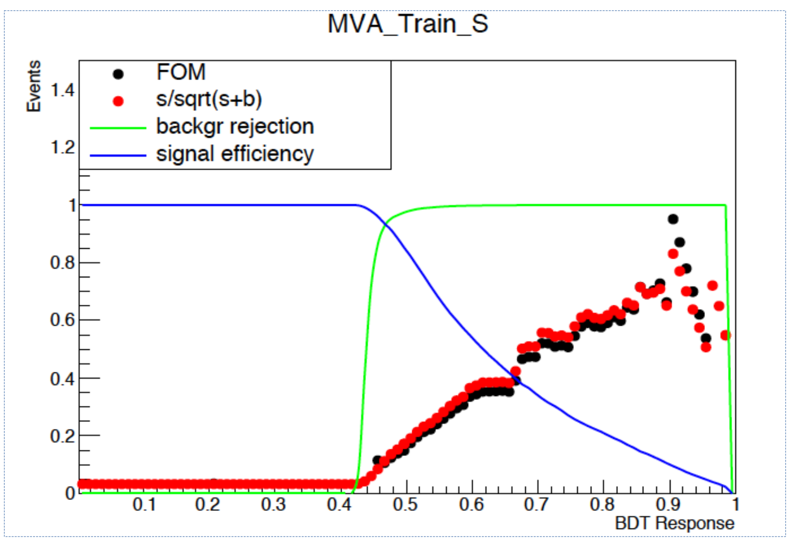




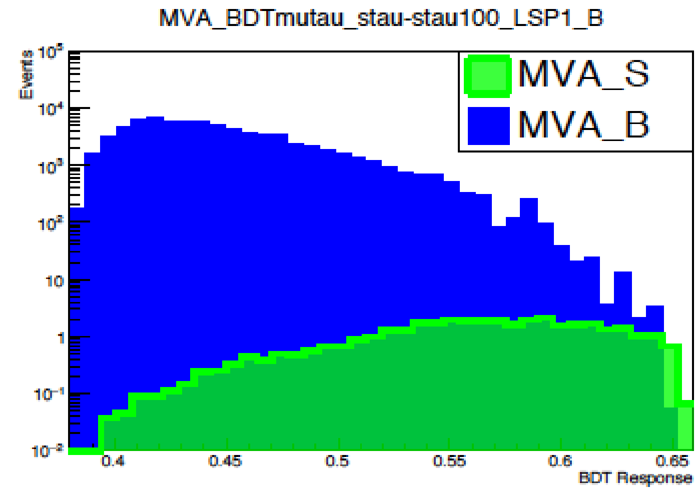
ANN vs BDT

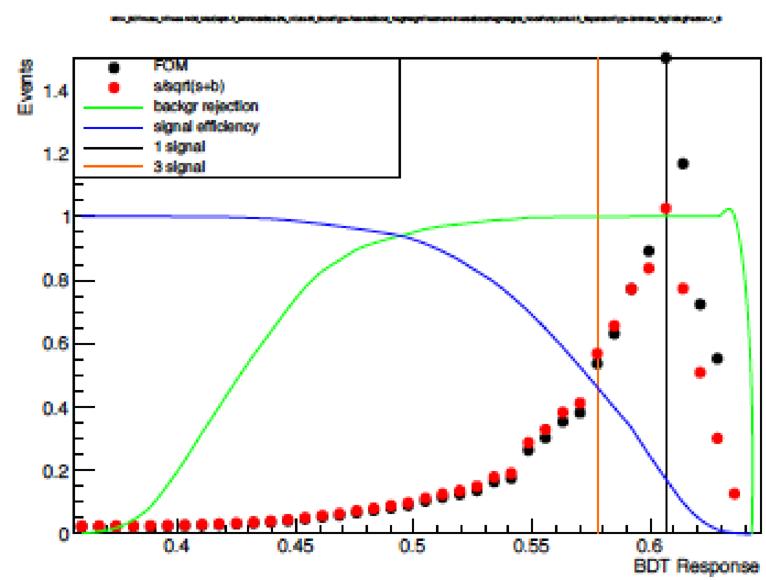
ANN





BDT





BDT is still better

BDT

NTrees

Number of trees in the forest

600, 800, 1000, 1200, 1400, 1600

• MaxDepth

Max depth of the decision tree allowed

3, 5, 7, 9

• MinNodeSize

Minimum percentage of training events required in a leaf node

2%, 3%,5%

• nCuts

Number of grid points in variable range used in finding optimal cut in node splitting

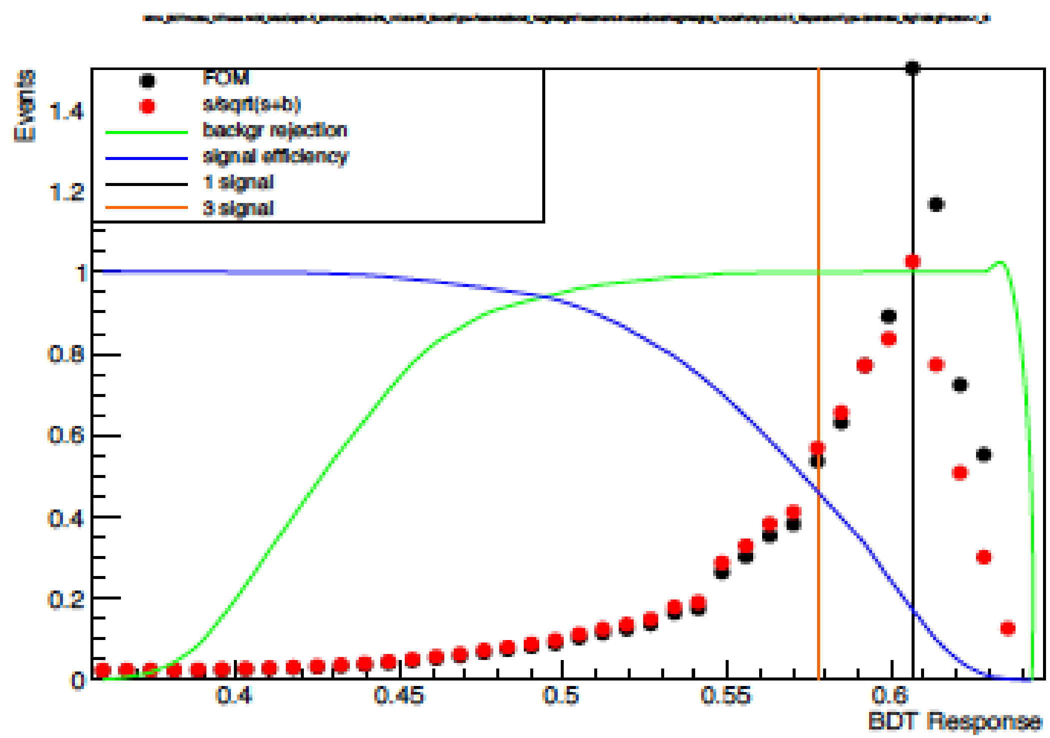
20, 25, 30, 35, 40, 50

• BoostType

Boosting type for the trees in the forest

AdaBoost, RealAdaBoost

$$Z_{\mathrm{A}} = \left[2\left((s+b)\ln\left[\frac{(s+b)(b+\sigma_b^2)}{b^2+(s+b)\sigma_b^2}\right] - \frac{b^2}{\sigma_b^2}\ln\left[1+\frac{\sigma_b^2 s}{b(b+\sigma_b^2)}\right]\right)\right]^{1/2}$$



Details and conclusions

- Train BDT for every signal point ~ 30
- Signal points MC with enough statystics to avoid overtraining were produced
- Increase in sensitivity

	Cuts	BDT
stau100	0.22	**1.05**
stau150	0.4	0.9
stau200	0.41	0.56
stau300	0.21	0.3

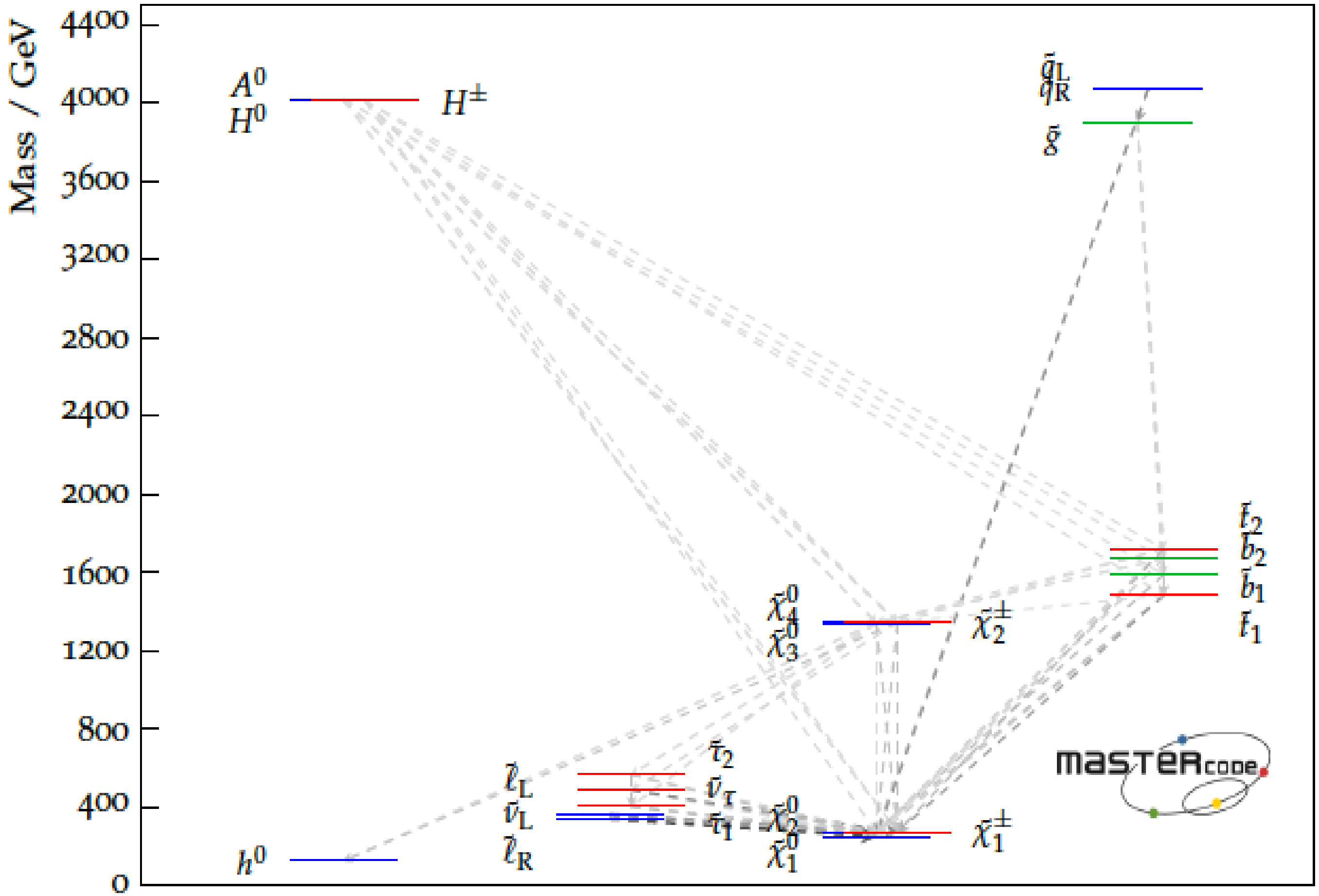
- Could be used in our analysis

Slection $\mu\tau_h$, $e\tau_h$, $e\mu$

Event selection

- Two opposite sign (OS) leptons ($\mu\tau$, eτ, or eμ) with $\Delta R > 0.3$ and no additional leptons
- $n_{\text{jet}} < 2$ and $n_{\text{b-tag}} = 0$
- $M_{\ell_1\ell_2} > 50$ GeV and $M_{\text{Tsum}} > 50$ GeV (low mass-resonances rejection)
- $20 < M_{\text{T}} < 60$ GeV or $M_{\text{T}} > 120$ GeV ($\mu\tau$, eτ channel) (W+jets rejection)
- $90 < M_{\mu e} < 250$ GeV (Z+jets rejection)
- $\text{P}_{\text{T}}(\ell_i) < 200$ GeV (i=e,μ, for the eμ channel only)
- $|d_z(\ell_i)| < 0.2\,\text{cm},|d_{xy}(\ell_i)| < 0.045\,\text{cm}$ (non-prompt ℓ background rejection)
- $\Delta\eta(J_0, \ell_{1,2}) < 3$ (eτ and $\mu\tau$ channels, 1-jet category only)
- $\Delta R(J_0, \tau) < 4$ (eτ and $\mu\tau$ channels, 1-jet category only)

Likelihood analysis



arXiv:1710.11091
v2

Search bins

Bin name	p_T^{miss} [GeV]	M_{T2} [GeV]	$D\zeta$ [GeV]	n_{jet}
$p_T^{miss}{}_A M_{T2A} D\zeta_{B-}$	<40	<40	<-100	0
$p_T^{miss}{}_A M_{T2B+} D\zeta_{A+}$		>40	>-500	
$p_T^{miss}{}_B M_{T2A} D\zeta_{B-}$	[40,80]	<40	<-100	
$p_T^{miss}{}_B M_{T2A} D\zeta_{E}$			>50	
$p_T^{miss}{}_B M_{T2B} D\zeta_{B-}$		[40,80]	<-100	
$p_T^{miss}{}_B M_{T2B} D\zeta_{C+}$			>-100	
$p_T^{miss}{}_B M_{T2C+} D\zeta_{A+}$		>80	>-500	
$p_T^{miss}{}_C M_{T2A} D\zeta_{B-}$	[80,120]	<40	<-100	
$p_T^{miss}{}_C M_{T2A} D\zeta_{C+}$			>-100	
$p_T^{miss}{}_C M_{T2B} D\zeta_{B-}$		[40,80]	<-150	
$p_T^{miss}{}_C M_{T2B} D\zeta_{A+}$			>-150	
$p_T^{miss}{}_C M_{T2B+} D\zeta_{A+}$		>80	>-500	
$p_T^{miss}{}_D M_{T2A} D\zeta_{B-}$	[120,250]	<40	<-100	
$p_T^{miss}{}_D M_{T2A} D\zeta_{C+}$			>-100	
$p_T^{miss}{}_D M_{T2B} D\zeta_{B-}$		[40,80]	<-150	
$p_T^{miss}{}_D M_{T2B} D\zeta_{B}$			[-150,-100]	
$p_T^{miss}{}_D M_{T2B} D\zeta_{C+}$			>-100	
$p_T^{miss}{}_D M_{T2C} D\zeta_{A+}$		[80,100]	>-500	
$p_T^{miss}{}_D M_{T2D} D\zeta_{A+}$		[100,120]	>-500	
$p_T^{miss}{}_D M_{T2E} D\zeta_{A+}$		>120	>-500	
$p_T^{miss}{}_E M_{T2A+} D\zeta_{A+}$	>250	>0	>-500	

Bin name	p_T^{miss} [GeV]	M_{T2} [GeV]	$D\zeta$ [GeV]	n_{jet}
$p_T^{miss}{}_A M_{T2A} D\zeta_{B-}$	<40	<40	<-150	1
$p_T^{miss}{}_A M_{T2A} D\zeta_{B}$			[-150,100]	
$p_T^{miss}{}_A M_{T2A} D\zeta_{D+}$			>0	
$p_T^{miss}{}_A M_{T2A+} D\zeta_{A+}$		>40	>-500	
$p_T^{miss}{}_B M_{T2A} D\zeta_{B-}$	[40,80]		<-100	
$p_T^{miss}{}_B M_{T2A} D\zeta_{E}$		<40	>50	
$p_T^{miss}{}_B M_{T2B} D\zeta_{B-}$		[40,80]	<-100	
$p_T^{miss}{}_B M_{T2B} D\zeta_{C+}$			>-100	
$p_T^{miss}{}_B M_{T2B+} D\zeta_{A+}$		>80	>-500	
$p_T^{miss}{}_C M_{T2A} D\zeta_{B-}$	[80,120]	<40	<-100	
$p_T^{miss}{}_C M_{T2B} D\zeta_{B-}$		[40,80]	<-150	
$p_T^{miss}{}_C M_{T2B} D\zeta_{A+}$			>-150	
$p_T^{miss}{}_C M_{T2CD} D\zeta_{A+}$		[80,120]	>-500	
$p_T^{miss}{}_C M_{T2E} D\zeta_{A+}$		>120	>-500	
$p_T^{miss}{}_D M_{T2A} D\zeta_{B-}$	[120,250]	<40	<-150	
$p_T^{miss}{}_D M_{T2A} D\zeta_{B}$			[-150,-100]	
$p_T^{miss}{}_D M_{T2A} D\zeta_{C+}$			>-100	
$p_T^{miss}{}_D M_{T2B} D\zeta_{B-}$		[40,80]	<-150	
$p_T^{miss}{}_D M_{T2B} D\zeta_{B}$			[-150,-100]	
$p_T^{miss}{}_D M_{T2B} D\zeta_{C+}$			>-100	
$p_T^{miss}{}_D M_{T2C} D\zeta_{A+}$		[80,100]	>-500	
$p_T^{miss}{}_D M_{T2D} D\zeta_{A+}$		[80,120]	>-500	
$p_T^{miss}{}_D M_{T2E} D\zeta_{A+}$		>120	>-500	
$p_T^{miss}{}_E M_{T2C+} D\zeta_{A+}$	>250	>80	>-500	

Table 9: Definition of combined search bins to be used for easier reinterpretation of the results.

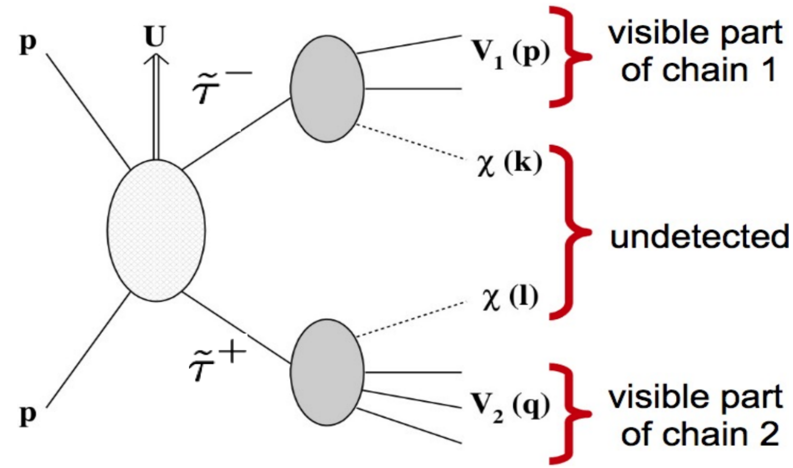
n_{jet}	p_T^{miss}	M_{T2}	Bkg ($e\tau_h$)	Obs. ($e\tau_h$)	Bkg ($\mu\tau_h$)	Obs ($\mu\tau_h$)	Bkg ($e\mu$)	Obs ($e\mu$)
0	> 120 GeV	> 120 GeV	$4.9 \pm 1.5 \pm 1.9$	4	$5.8 \pm 1.8 \pm 2.7$	7.0	$6.8 \pm 2.2 \pm 2.7$	6
1	> 120 GeV	> 100 GeV	$10.8 \pm 2.1 \pm 2.5$	9	$14.4 \pm 2.5 \pm 3.1$	14	$9.7 \pm 2.4 \pm 3.0$	6
1	> 250 GeV	> 80 GeV	$1.6 \pm 0.9 \pm 1.2$	0	$1.5 \pm 0.9 \pm 1.1$	1	$3.3 \pm 2.0 \pm 2.3$	1

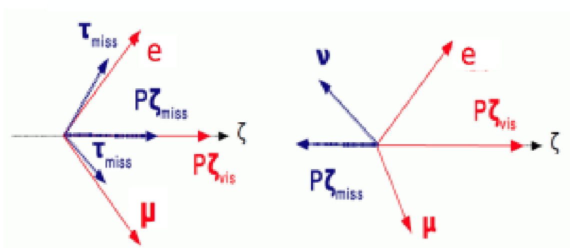
Search variables $\mu\tau_h, e\tau_h, e\mu$

- E_{Tmiss} – missing transverse energy
- M_{T2} – “stransverse” mass

$$M_{T2}^2 = \min_{\vec{k}_T+\vec{l}_T=\text{tot miss }\vec{p}_T} \left\{ \max\left[M_T^2(\text{chain 1}), M_T^2(\text{chain 2}) \right] \right\} \leq m^2_{\tilde{\tau}}$$

- D_ζ – Discriminant used in legacy Higgs searches



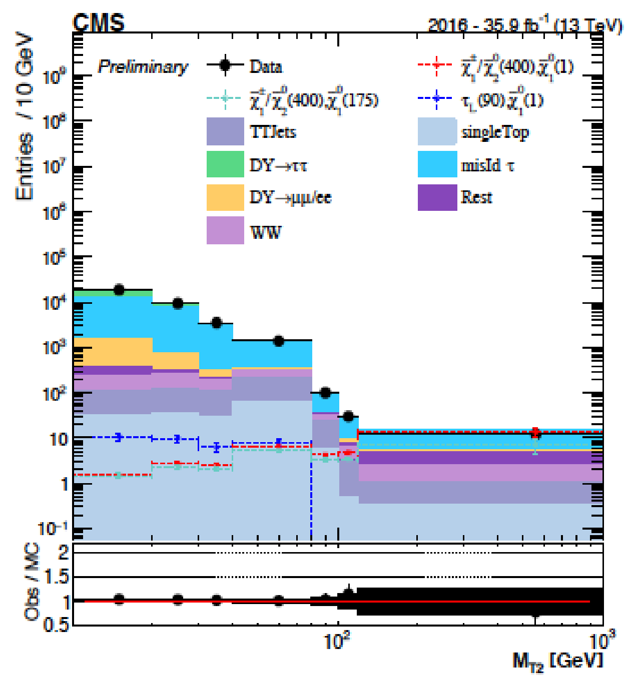


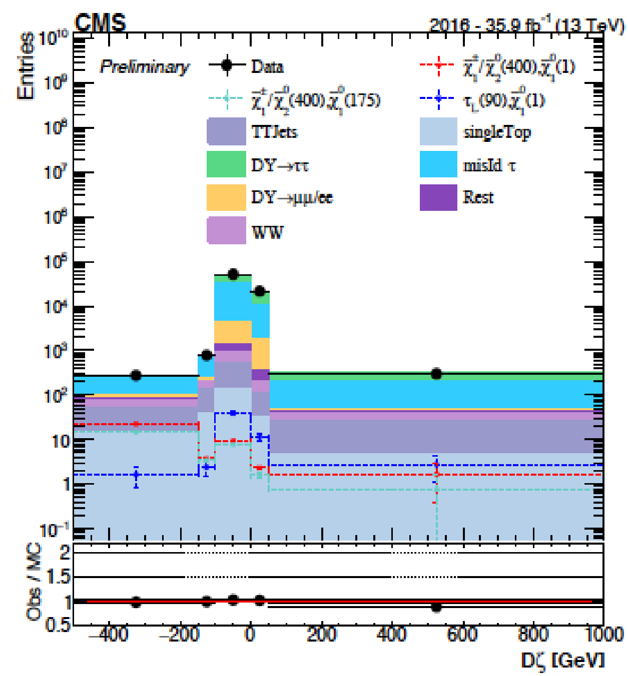
$$D_\zeta = P_{\zeta,\text{mis}} - \alpha \cdot P_{\zeta,\text{vis}}$$

$$P_{\zeta,\text{mis}} = \vec{p}_{T,\text{mis}} \cdot \vec{\zeta}, \quad P_{\zeta,\text{vis}} = (\vec{p}_{T,e} + \vec{p}_{T,\mu}) \cdot \vec{\zeta}$$

ζ – bisector between the direction of the electron and that of the muon

α = 0.85 (optimized value)





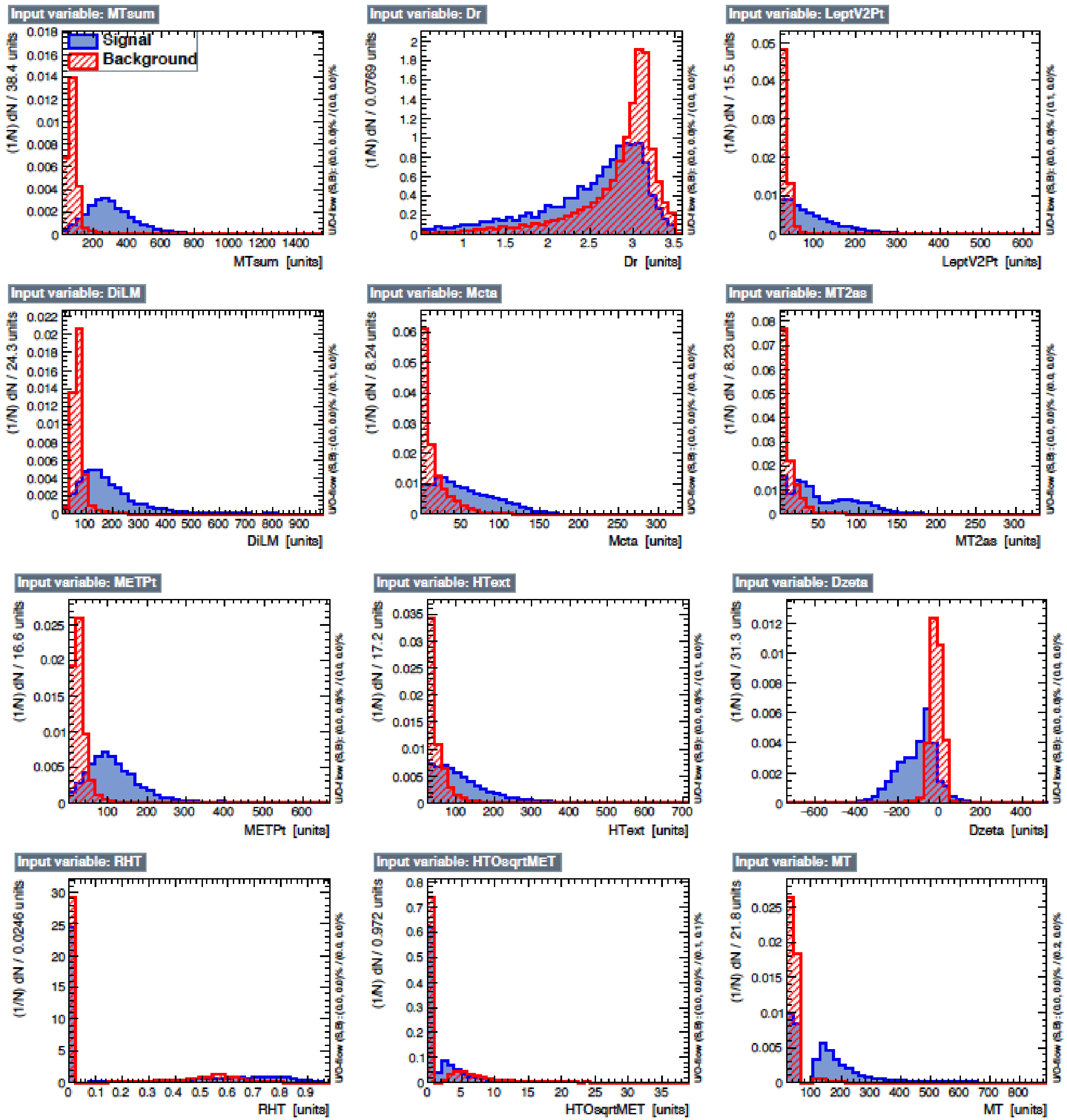
Search bins

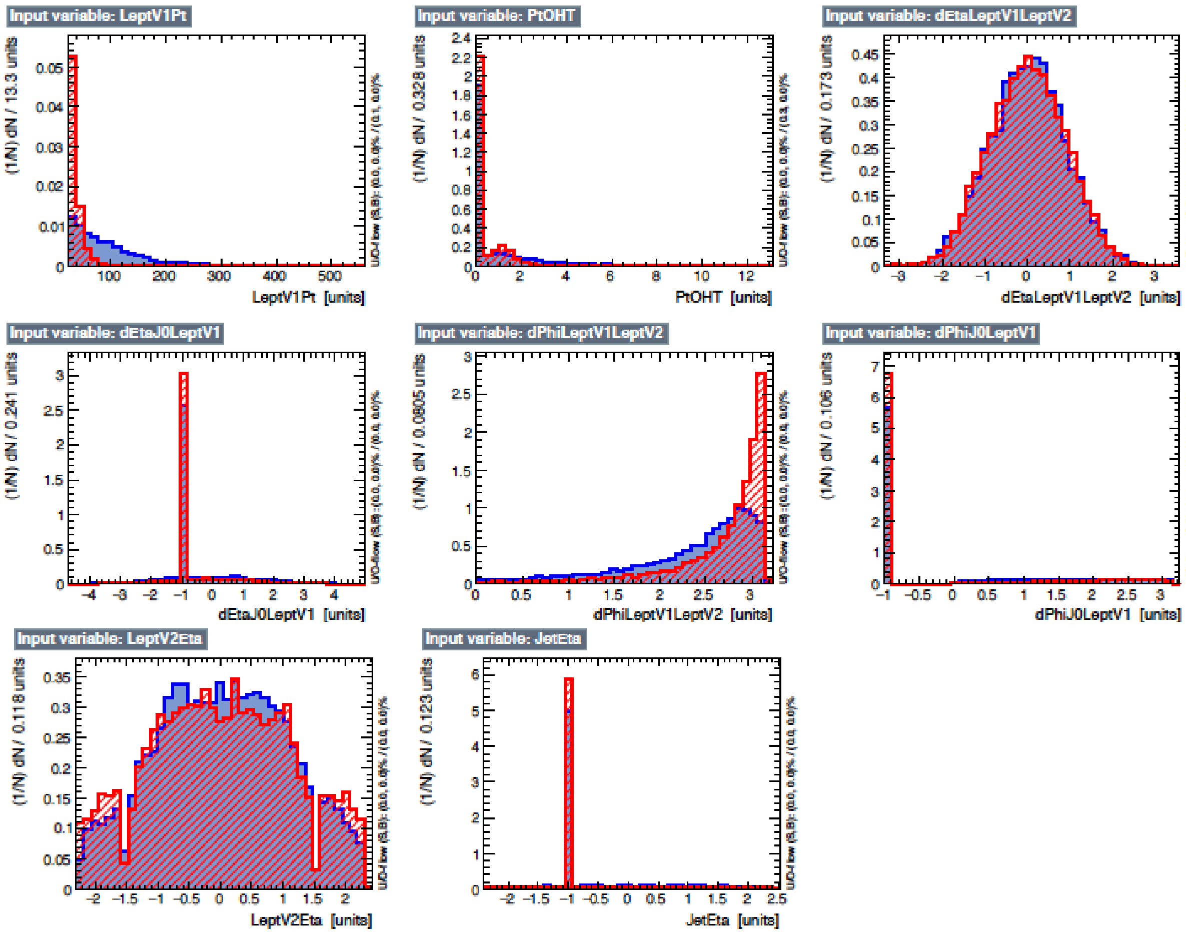
Bin name	p_T^{miss} [GeV]	M_{T2} [GeV]	$D\zeta$ [GeV]	n_{jet}
$p_T^{miss}{}_A M_{T2A} D\zeta_{B-}$	<40	<40	<-100	0
$p_T^{miss}{}_A M_{T2B+} D\zeta_{A+}$		>40	>-500	
$p_T^{miss}{}_B M_{T2A} D\zeta_{B-}$	[40,80]	<40	<-100	
$p_T^{miss}{}_B M_{T2A} D\zeta_{E}$			>50	
$p_T^{miss}{}_B M_{T2B} D\zeta_{B-}$		[40,80]	<-100	
$p_T^{miss}{}_B M_{T2B} D\zeta_{C+}$			>-100	
$p_T^{miss}{}_B M_{T2C+} D\zeta_{A+}$		>80	>-500	
$p_T^{miss}{}_C M_{T2A} D\zeta_{B-}$	[80,120]	<40	<-100	
$p_T^{miss}{}_C M_{T2A} D\zeta_{C+}$			>-100	
$p_T^{miss}{}_C M_{T2B} D\zeta_{B-}$		[40,80]	<-150	
$p_T^{miss}{}_C M_{T2B} D\zeta_{A+}$			>-150	
$p_T^{miss}{}_C M_{T2B+} D\zeta_{A+}$		>80	>-500	
$p_T^{miss}{}_D M_{T2A} D\zeta_{B-}$	[120,250]	<40	<-100	
$p_T^{miss}{}_D M_{T2A} D\zeta_{C+}$			>-100	
$p_T^{miss}{}_D M_{T2B} D\zeta_{B-}$		[40,80]	<-150	
$p_T^{miss}{}_D M_{T2B} D\zeta_{B}$			[-150,-100]	
$p_T^{miss}{}_D M_{T2B} D\zeta_{C+}$			>-100	
$p_T^{miss}{}_D M_{T2C} D\zeta_{A+}$		[80,100]	>-500	
$p_T^{miss}{}_D M_{T2D} D\zeta_{A+}$		[100,120]	>-500	
$p_T^{miss}{}_D M_{T2E} D\zeta_{A+}$		>120	>-500	
$p_T^{miss}{}_E M_{T2A+} D\zeta_{A+}$	>250	>0	>-500	

Bin name	p_T^{miss} [GeV]	M_{T2} [GeV]	$D\zeta$ [GeV]	n_{jet}
$p_T^{miss}{}_A M_{T2A} D\zeta_{B-}$	<40	<40	<-150	1
$p_T^{miss}{}_A M_{T2A} D\zeta_{B}$			[-150,100]	
$p_T^{miss}{}_A M_{T2A} D\zeta_{D+}$			>0	
$p_T^{miss}{}_A M_{T2A+} D\zeta_{A+}$		>40	>-500	
$p_T^{miss}{}_B M_{T2A} D\zeta_{B-}$	[40,80]		<-100	
$p_T^{miss}{}_B M_{T2A} D\zeta_{E}$		<40	>50	
$p_T^{miss}{}_B M_{T2B} D\zeta_{B-}$		[40,80]	<-100	
$p_T^{miss}{}_B M_{T2B} D\zeta_{C+}$			>-100	
$p_T^{miss}{}_B M_{T2B+} D\zeta_{A+}$		>80	>-500	
$p_T^{miss}{}_C M_{T2A} D\zeta_{B-}$	[80,120]	<40	<-100	
$p_T^{miss}{}_C M_{T2B} D\zeta_{B-}$		[40,80]	<-150	
$p_T^{miss}{}_C M_{T2B} D\zeta_{A+}$			>-150	
$p_T^{miss}{}_C M_{T2CD} D\zeta_{A+}$		[80,120]	>-500	
$p_T^{miss}{}_C M_{T2E} D\zeta_{A+}$		>120	>-500	
$p_T^{miss}{}_D M_{T2A} D\zeta_{B-}$	[120,250]	<40	<-150	
$p_T^{miss}{}_D M_{T2A} D\zeta_{B}$			[-150,-100]	
$p_T^{miss}{}_D M_{T2A} D\zeta_{C+}$			>-100	
$p_T^{miss}{}_D M_{T2B} D\zeta_{B-}$		[40,80]	<-150	
$p_T^{miss}{}_D M_{T2B} D\zeta_{B}$			[-150,-100]	
$p_T^{miss}{}_D M_{T2B} D\zeta_{C+}$			>-100	
$p_T^{miss}{}_D M_{T2C} D\zeta_{A+}$		[80,100]	>-500	
$p_T^{miss}{}_D M_{T2D} D\zeta_{A+}$		[80,120]	>-500	
$p_T^{miss}{}_D M_{T2E} D\zeta_{A+}$		>120	>-500	
$p_T^{miss}{}_E M_{T2C+} D\zeta_{A+}$	>250	>80	>-500	

Table 9: Definition of combined search bins to be used for easier reinterpretation of the results.

n_{jet}	p_T^{miss}	M_{T2}	Bkg ($e\tau_h$)	Obs. ($e\tau_h$)	Bkg ($\mu\tau_h$)	Obs ($\mu\tau_h$)	Bkg ($e\mu$)	Obs ($e\mu$)
0	> 120 GeV	> 120 GeV	$4.9 \pm 1.5 \pm 1.9$	4	$5.8 \pm 1.8 \pm 2.7$	7.0	$6.8 \pm 2.2 \pm 2.7$	6
1	> 120 GeV	> 100 GeV	$10.8 \pm 2.1 \pm 2.5$	9	$14.4 \pm 2.5 \pm 3.1$	14	$9.7 \pm 2.4 \pm 3.0$	6
1	> 250 GeV	> 80 GeV	$1.6 \pm 0.9 \pm 1.2$	0	$1.5 \pm 0.9 \pm 1.1$	1	$3.3 \pm 2.0 \pm 2.3$	1


Input variable: MTsum
Signal
Background
(1/N) dN / 38.4 units
MTsum [units]
Input variable: Dr
(1/N) dN / 0.0769 units
Dr [units]
Input variable: LeptV2Pt
(1/N) dN / 15.5 units
LeptV2Pt [units]
Input variable: DiLM
(1/N) dN / 24.3 units
DiLM [units]
Input variable: Mcta
(1/N) dN / 8.24 units
Mcta [units]
Input variable: MT2as
(1/N) dN / 8.23 units
MT2as [units]
Input variable: METPt
(1/N) dN / 16.6 units
METPt [units]
Input variable: HText
(1/N) dN / 17.2 units
HText [units]
Input variable: Dzeta
(1/N) dN / 31.3 units
Dzeta [units]
Input variable: RHT
(1/N) dN / 0.0246 units
RHT [units]
Input variable: HTOsqrtMET
(1/N) dN / 0.972 units
HTOsqrtMET [units]
Input variable: MT
(1/N) dN / 21.8 units
MT [units]


Input variable: LeptV1Pt
(1/N) dN / 13.3 units
LeptV1Pt [units]
Input variable: PtOHT
(1/N) dN / 0.328 units
PtOHT [units]
Input variable: dEtaLeptV1LeptV2
(1/N) dN / 0.173 units
dEtaLeptV1LeptV2 [units]
Input variable: dEtaJ0LeptV1
(1/N) dN / 0.241 units
dEtaJ0LeptV1 [units]
Input variable: dPhiLeptV1LeptV2
(1/N) dN / 0.0805 units
dPhiLeptV1LeptV2 [units]
Input variable: dPhiJ0LeptV1
(1/N) dN / 0.106 units
dPhiJ0LeptV1 [units]
Input variable: LeptV2Eta
(1/N) dN / 0.118 units
LeptV2Eta [units]
Input variable: JetEta
(1/N) dN / 0.123 units
JetEta [units]