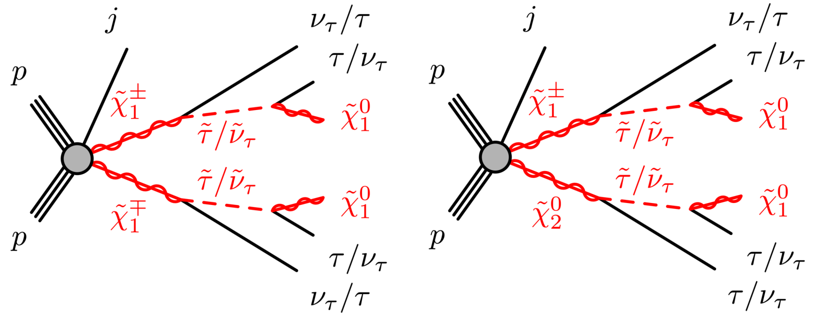
Search for staus and electroweakinos production in compressed region using the final states with two hadronic taus or one hadronic tau one lepton

Chengxin Liao
On behalf of the EWK-2tau team
liaocx@ihep.ac.cn

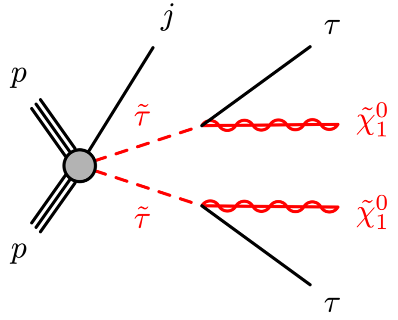
May, Fri 16, 2025

Target Signal

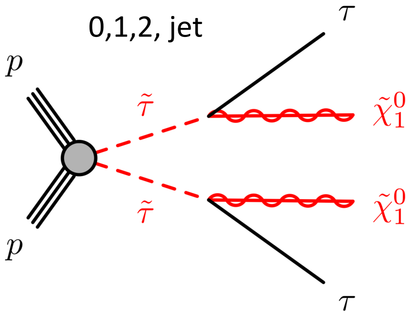
C1C1, C1N2 via stau with $\geq 2\tau + E_T^{miss} + ISR$

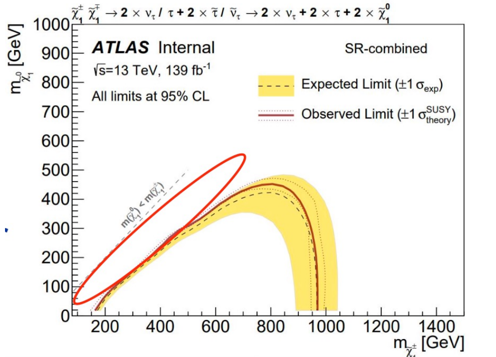


1 or 2 jets, leading jet at least 50 GeV

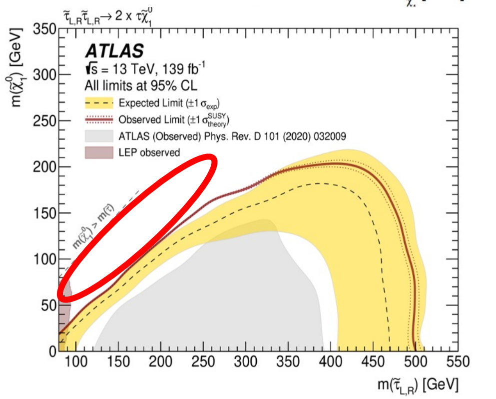


0,1,2, jet





Direct stau with $2\tau + E_T^{miss}$
ISR channel and inclusive channel



Previous paper: JHEP 05 (2024) 150

Motivation:

an explanation for observed excess in the Muon g-2 experiment and the observed dark matter relic density
ISR can boost the SUSY system and improve the sensitivity in compressed region

Glance Analysis entry: ANA-HMBS-2024-07
Internal Note: ANA-HMBS-2024-07-INT1

Ntuple setup

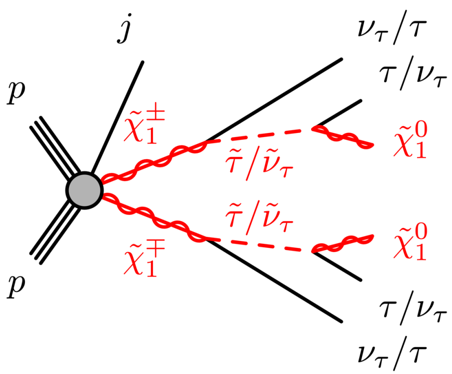
- MelAnalysis framework
- DAOD PHYS
- ptag: signal p6244, bkg p6490/p6491, data p6479/p6269 (run2 run3)
- background list: mc20, mc23

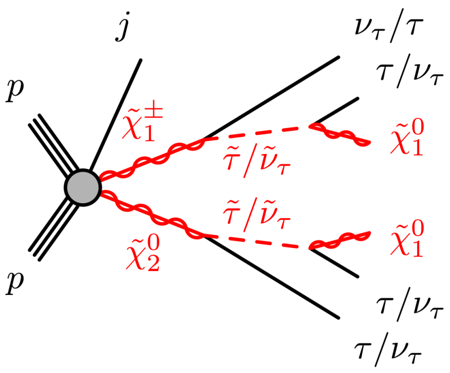
Type	dsid	sample
$t\bar{t}$	410470	PhPy8EG_A14_ttbar_hdamp258p75_nonallhad
	410471	PhPy8EG_A14_ttbar_hdamp258p75_allhad
single top	410644,410645	PowhegPythia8EvtGen_A14_singletop_schan_lept
	410658,410659	PhPy8EG_A14_tchan_BW50_lept
	601352,601355	PhPy8EG_tW_dyn_DR_incl
Rare top	304014	MadGraphPythia8EvtGen_A14NNPDF23_3top_SM
	410276	aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_ttee_mll_1_5
	410277	aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_ttmumu_mll_1_5
	410278	aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_tttautau_mll_1_5
$t\bar{t}X$	410081	MadGraphPythia8EvtGen_A14NNPDF23_ttbarWW
	504330	aMCPy8EG_NNPDF30NLO_A14N23LO_ttee
	504334	aMCPy8EG_NNPDF30NLO_A14N23LO_ttmumu
	504338	aMCPy8EG_NNPDF30NLO_A14N23LO_ttZqq
	504342	aMCPy8EG_NNPDF30NLO_A14N23LO_tttautau
	504346	aMCPy8EG_NNPDF30NLO_A14N23LO_ttZnunu

Type	dsid	sample
Higgs	346343	PhPy8EG_A14NNPDF23_NNPDF30ME_ttH125_allhad
	346344	PhPy8EG_A14NNPDF23_NNPDF30ME_ttH125_semilep
	346345	PhPy8EG_A14NNPDF23_NNPDF30ME_ttH125_dilep
	345097	PowhegPythia8EvtGen_NNLOPS_nnlo_30_ggH125_mumu
	345120	PowhegPy8EG_NNLOPS_nnlo_30_ggH125_tautaul13l7
	345121	PowhegPy8EG_NNLOPS_nnlo_30_ggH125_tautaulm15hp20
	345122	PowhegPy8EG_NNLOPS_nnlo_30_ggH125_tautaulp15hm20
	345123	PowhegPy8EG_NNLOPS_nnlo_30_ggH125_tautauh30h20
	345949	PowhegPythia8EvtGen_NNPDF30_AZNLOCTEQ6L1_VBFH125_bb
	346190	PowhegPy8EG_NNPDF30_AZNLOCTEQ6L1_VBFH125_tautaul13l7
	346191	PowhegPy8EG_NNPDF30_AZNLOCTEQ6L1_VBFH125_tautaulm15hp20
	346192	PowhegPy8EG_NNPDF30_AZNLOCTEQ6L1_VBFH125_tautaulp15hm20
	346193	PowhegPy8EG_NNPDF30_AZNLOCTEQ6L1_VBFH125_tautauh30h20
	345053	PowhegPythia8EvtGen_NNPDF3_AZNLO_WmH125J_MINLO_lvbb_VpT
	345054	PowhegPythia8EvtGen_NNPDF3_AZNLO_WpH125J_MINLO_lvbb_VpT
	345055	PowhegPythia8EvtGen_NNPDF3_AZNLO_ZH125J_MINLO_llbb_VpT
	345056	PowhegPythia8EvtGen_NNPDF3_AZNLO_ZH125J_MINLO_vvbb_VpT
	345057	PowhegPythia8EvtGen_NNPDF3_AZNLO_ggZH125_llbb
	345098	PowhegPythia8EvtGen_NNPDF3_AZNLO_ggZH125_Hmumu_Zinc
	345103	PowhegPythia8EvtGen_NNPDF30_AZNLO_ZH125J_Hmumu_Zincl_MINLO
	345104	PowhegPythia8EvtGen_NNPDF30_AZNLO_WpH125J_Hmumu_Wincl_MINLO
	345105	PowhegPythia8EvtGen_NNPDF30_AZNLO_WmH125J_Hmumu_Wincl_MINLO
	345109	PowhegPythia8EvtGen_NNPDF3_AZNLO_WmH125J_MINLO_lvcc_VpT
	345110	PowhegPythia8EvtGen_NNPDF3_AZNLO_WpH125J_MINLO_lvcc_VpT
	345111	PowhegPythia8EvtGen_NNPDF3_AZNLO_ZH125J_MINLO_llcc_VpT
	345112	PowhegPythia8EvtGen_NNPDF3_AZNLO_ZH125J_MINLO_vvcc_VpT
	345113	PowhegPythia8EvtGen_NNPDF3_AZNLO_ggZH125_llcc
	345211	PowhegPy8EG_NNPDF30_AZNLO_WmH125J_Winc_MINLO_tautau
	345212	PowhegPy8EG_NNPDF30_AZNLO_WpH125J_Winc_MINLO_tautau
	345217	PowhegPy8EG_NNPDF30_AZNLO_ZH125J_Zinc_MINLO_tautau
	346329	PowhegPy8EG_NNPDF30_AZNLO_ggZH125_lltautau_filt

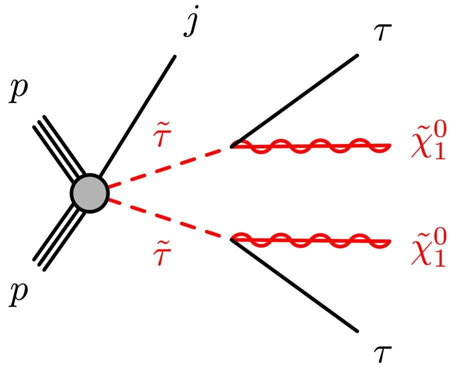
Type	dsid	sample
Z+jets	700320-700322, 700467-700469	Sh_2211_Zee_maxHTpTV2
	700323-700325, 700470-700472	Sh_2211_Zmumu_maxHTpTV2
	700792-700794, 700901-700903	Sh_2214_Ztautau_maxHTpTV2
	700358	Sh_2211_Zee2jets_Min_N_TChannel
	700359	Sh_2211_Zmm2jets_Min_N_TChannel
	700360	Sh_2211_Ztt2jets_Min_N_TChannel
W+jets	700338-700340	Sh_2211_Wenu_maxHTpTV2
	700341-700343	Sh_2211_Wmunu_maxHTpTV2
	700344-700349	Sh_2211_Wtaunu_*_maxHTpTV2
	700362	Sh_2211_Wenu2jets_Min_N_TChannel
	700363	Sh_2211_Wmunu2jets_Min_N_TChannel
	700364	Sh_2211_Wtaunu2jets_Min_N_TChannel
Diboson	345705,345706	Sherpa_222_NNPDF30NNLO_ggllll
	345718	Sherpa_222_NNPDF30NNLO_gglllvvWW
	345723	Sherpa_222_NNPDF30NNLO_ggllvvZZ
	364288	Sherpa_222_NNPDF30NNLO_llll_lowMllPtComplement
	364289	Sherpa_222_NNPDF30NNLO_lllv_lowMllPtComplement
	364290	Sherpa_222_NNPDF30NNLO_llvv_lowMllPtComplement
	364302,364303	Sherpa_222_NNPDF30NNLO_ggZ*Zqq
	364304,364305	Sherpa_222_NNPDF30NNLO_ggW*lvW*qq
	700600	Sh_2212_llll
	700601	Sh_2212_lllv
	700602,700603	Sh_2212_llvv_*s
	700604	Sh_2212_lvvv
	700605	Sh_2212_vvvv
	701000	Sh_2214_lllljj
	701005	Sh_2214_lllvjj
	701010,701015	Sh_2214_llvvjj_*s
	701020	Sh_2214_lllljj_Int
	701025	Sh_2214_lllvjj_Int
	701030,701035	Sh_2214_llvvjj_*s_Int
	701085	Sh_2214_ZqqZll
	701090	Sh_2214_ZbbZll
	701095	Sh_2214_ZqqZvv
	701100	Sh_2214_ZbbZvv
	701105	Sh_2214_WqqZll
	701110	Sh_2214_WqqZvv
	701115	Sh_2214_WlvZqq
	701120	Sh_2214_WlvZbb
	701125	Sh_2214_WlvWqq
Triboson	364242	Sherpa_222_NNPDF30NNLO_WWW_3l3v_EW6
	364243-364244	Sherpa_222_NNPDF30NNLO_WWZ
	364245-364246	Sherpa_222_NNPDF30NNLO_WZZ
	364247-364249	Sherpa_222_NNPDF30NNLO_ZZZ

Signal samples

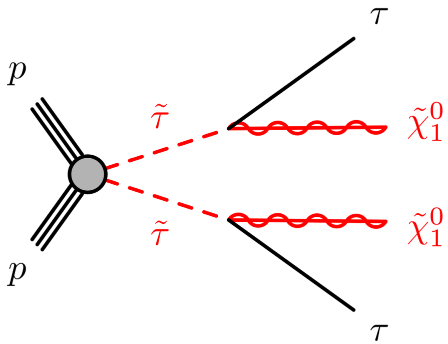




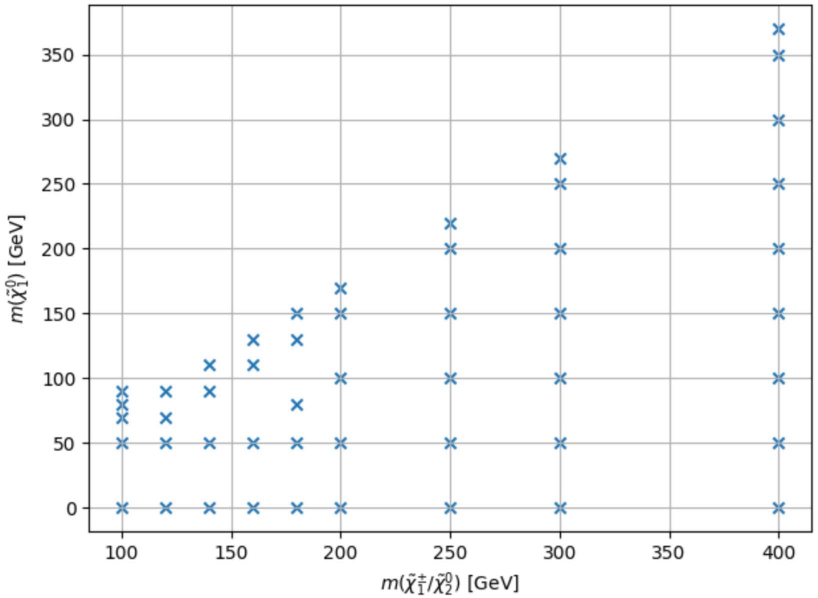
$\tilde{\chi}_1^{\pm}$ and $\tilde{\chi}_2^0$ mass degenerate and pure wino
$\tilde{\tau}/\tilde{\nu}$ mass degenerate and lie midway between $\tilde{\chi}_1^{\pm}$ and $\tilde{\chi}_2^0$

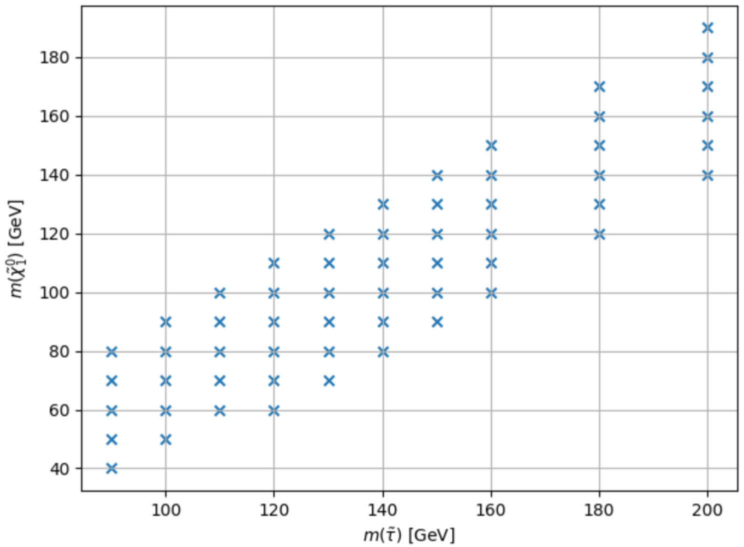


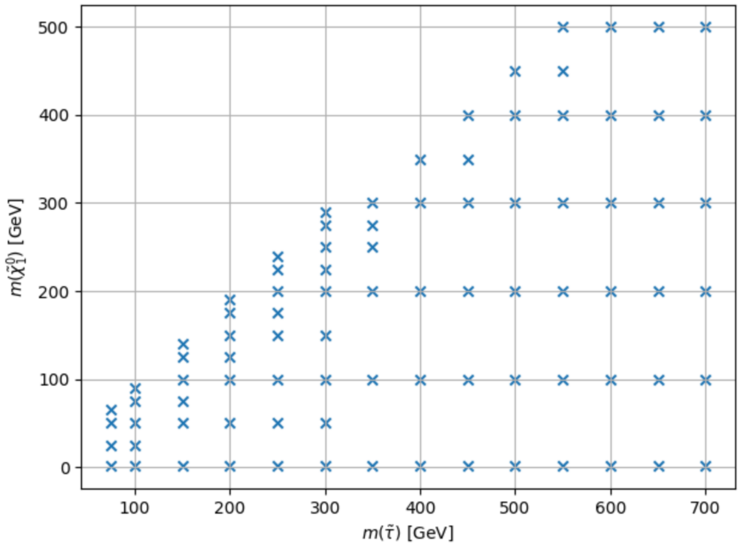
$\tilde{\tau}_L/\tilde{\tau}_R$ mass degenerate
Bino-like $\tilde{\chi}_1^0$



$\tilde{\tau}_L/\tilde{\tau}_R$ mass degenerate
Bino-like $\tilde{\chi}_1^0$







Object Definition

- Full ST config: Run2, Run3
- Follow the CP recommendation

Selection	Electron	Muon	Tau	Jet
Baseline	$p_T > 4.5$ GeV $\lvert\eta\rvert < 2.47$ $\lvert z_0 \cdot \sin\theta\rvert < 0.5$ mm ID: LooseAndBLayerLLH impact parameter	$p_T > 5$ GeV $\lvert\eta\rvert < 2.5$ $\lvert z_0 \cdot \sin\theta\rvert < 0.5$ mm ID: Medium	$p_T > 15$ GeV $\lvert\eta\rvert < 1.37$ or $1.52 < \lvert\eta\rvert < 2.5$ $\lvert Q\rvert = 1$ 1 or 3 tracks Ele ID: RNN Loose Tau ID: VeryLoose	AntiKt4EMPFlow $p_T > 20$ GeV $\lvert\eta\rvert < 4.5$ JVT: FixedEffPt if $p_T < 60$ GeV bTag: GN2v01, 77% WP
Signal	ID: TightLLH Run2 ISO: Loose_VarRad, HighPtCaloOnly Run3 ISO: Tight_VarRad $\lvert d_0/\sigma_{d_0}\rvert < 5$	ISO: PflowLoose_VarRad $\lvert d_0/\sigma_{d_0}\rvert < 3$	Tau ID: Medium	

Overlap Removal

- ST default overlap removal
- Do Tau OLR

Reject	Against	Criteria
electron	electron	shared track, $p_{T1} < p_{T2}$
tau	electron	$\Delta R < 0.2$
tau	muon	$\Delta R < 0.2$
muon	electron	is calo-muon & shared ID track
electron	muon	shared ID track
jet	electron	$\Delta R < 0.2$
electron	jet	$\Delta R < \min(0.4,\ 0.04 + 10\ \text{GeV}/p_T^{\text{ele}})$
jet	muon	NumTrack < 3 & (ghost-associated or $\Delta R < 0.2$)
muon	jet	$\Delta R < \min(0.4,\ 0.04 + 10\ \text{GeV}/p_T^{\mu})$
jet	tau	$\Delta R < 0.2$

Trigger Strategy

Stau:

MET trigger for MET >= 200,

Single lepton trigger for MET <= 200

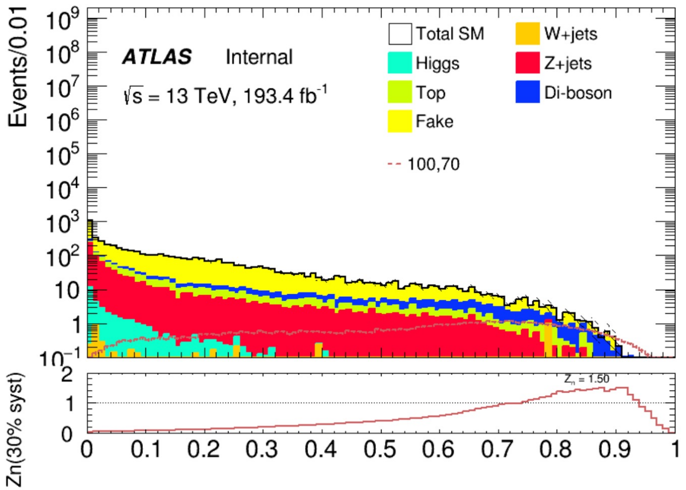
C1C1/C1N2: MET trigger

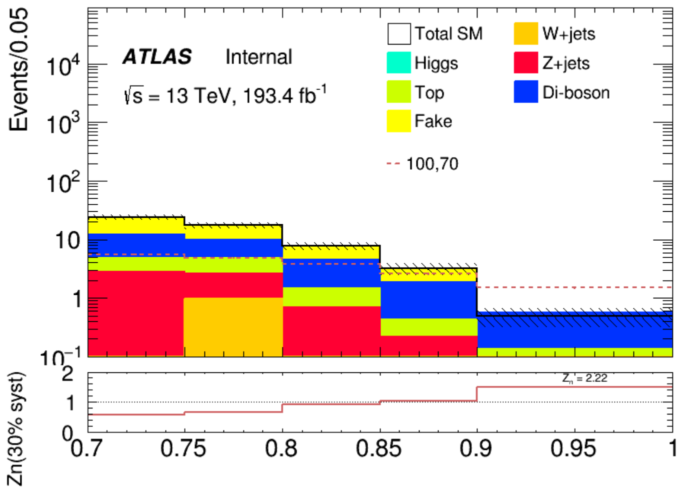
Trigger	Trigger name	Year	HLT	Offline
MET trigger	HLT_xe70_mht	2015	70	200
	HLT_xe90_mht_L1XE50	2016	90	
	HLT_xe100_mht_L1XE50	2016	100	
	HLT_xe110_pufit_L1XE55	2017	110	
	HLT_xe110_pufit_L1XE50	2017	110	
	HLT_xe110_pufit_xe70_L1XE50	2018	70	
	HLT_xe110_pufit_xe65_L1XE50	2018	65	
	HLT_xe65_cell_xe90_pfopufit_L1XE50	2022,2023	90	
single electron trigger	HLT_e24_lhmedium_L1EM20VH	2015	24	25
	HLT_e60_lhmedium	2015	60	61
	HLT_e120_lhloose	2015	120	121
	HLT_e26_lhtight_nod0_ivarloose	2016-2018	26	27
	HLT_e60_lhmedium_nod0	2016-2018	60	61
	HLT_e140_lhloose_nod0	2016-2018	140	141
	HLT_e26_lhtight_ivarloose_L1EM22VHI	2022-2023	26	27
	HLT_e60_lhmedium_L1EM22VHI	2022-2023	60	61
	HLT_e140_lhloose_L1EM22VHI	2022-2023	140	141
single muon trigger	HLT_mu20_iloose_L1MU15	2015	20	21
	HLT_mu26_ivarmedium	2016-2018	26	27.3
	HLT_mu50	2015-2018	50	52.5
	HLT_mu24_ivarmedium_L1MU14FCH	2022-2023	24	25.2
	HLT_mu50_L1MU14FCH	2022-2023	50	52.5

Direct Stau ISR signal region definition(HH)

- Preselection

HH Pre-selection
>= 2 medium taus (OS)
0 base lepton
bveto
MET trigger
MET ≥ 200 GeV
$\mathbf{M_{\tau\tau}^{reco} < 40\ GeV\ or\ M_{\tau\tau}^{reco} > 130\ GeV}$





- SR selection using neuron network score
  - Fake tau estimation using fake factor method
  - DNN signal score > 0.7

SR
Pre-selection
DNN score > 0.7

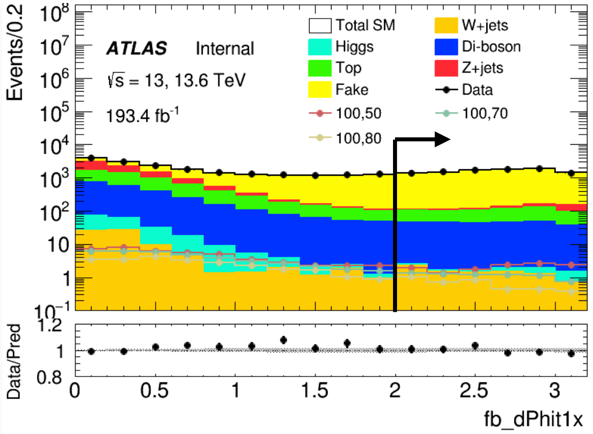
- Background composition

Process	[0.70,0.75]	[0.75,0.80]	[0.80,0.85]	[0.85,0.90]	[0.90,1.00]	Combined
TotalBkg	24.17±2.35	17.35±2.07	7.69±1.13	3.20±0.77	0.49±0.17	52.90±4.84
Leading Fake	11.95±2.23	7.35±1.66	3.24±1.04	1.33±0.72	-0.07±0.09	23.80±3.05
subleading VV	7.32±0.52	5.18±0.52	2.97±0.27	1.45±0.22	0.42±0.08	17.33±0.82
Top	2.12±0.49	2.23±0.52	0.80±0.31	0.21±0.14	0.11±0.10	5.47±0.80
Zjets	2.69±0.23	1.61±0.21	0.66±0.17	0.20±0.08	0.02±0.07	5.18±0.38
Wjets	0.00±0.00	0.98±0.98	0.00±0.00	0.00±0.00	0.00±0.00	0.98±0.98
Higgs	0.09±0.03	0.00±0.00	0.02±0.01	0.01±0.01	0.01±0.01	0.14±0.03
StauStauISR-100-70	5.61±0.10	4.90±0.09	3.84±0.08	2.59±0.07	1.53±0.05	18.48±0.18
ZnSignificance	0.58	0.65	0.91	1.03	1.50	2.22

Direct Stau ISR fake estimation

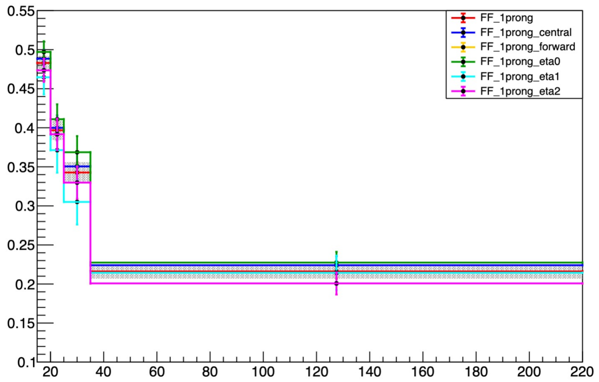
- CRs (fake factor computation)
  - METtrig
  - MET>=200
  - bveto
  - 1 signal lepton
  - $\Delta\phi(\tau, MET) > 2$ improve fake purity
  - **ID: >= 1 medium tau**
  - **antiID: >= 1 VeryLoose tau, 0 medium tau**
- SRs
  - preselection
  - 2ID: >= 2 medium tau
  - 1ID1antiID: >= **2 VeryLoose tau** , 1 medium tau
  - 2antiID: >= **2 VeryLoose tau** , 0 medium tau

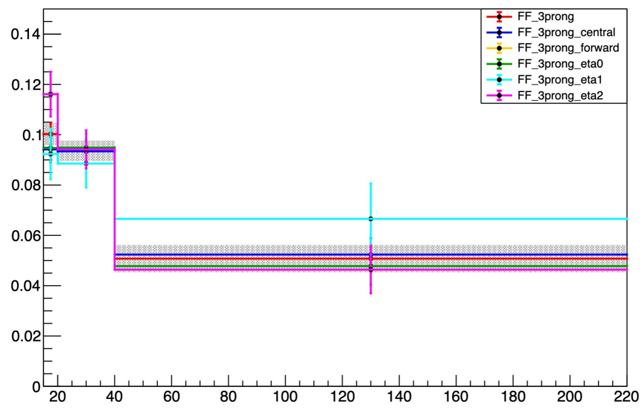
$\Delta\phi(\tau, MET)$ N-1 plots



<table>
<tr><td>FF determination
Data</td><td>FF application
Data</td></tr>
<tr><td rowspan="2">CR
Pass ID</td><td>SR
Pass ID</td></tr>
<tr><td>SR
but
1 tau Fail ID</td></tr>
<tr><td>CR
Fail ID</td><td>SR
but
2 tau Fail ID</td></tr>
</table>

- **Binned in prongness, tau eta, tau pT**
  - Eta bins
    - 2 bins: central [0,1.37], forward [1.52,2.5]
    - 3 bins: eta0,1,2 for [0,1), [1, 1.37], [1.52,2.5]
- **Auto binning:**
  - > 10% of events in nominator and denominator
  - Add bins to bin i until it is not consistent anymore with bin i - 1
    - Relative stat uncertainty on ratio smaller than 50%
    - >10% events in nominator and denominator

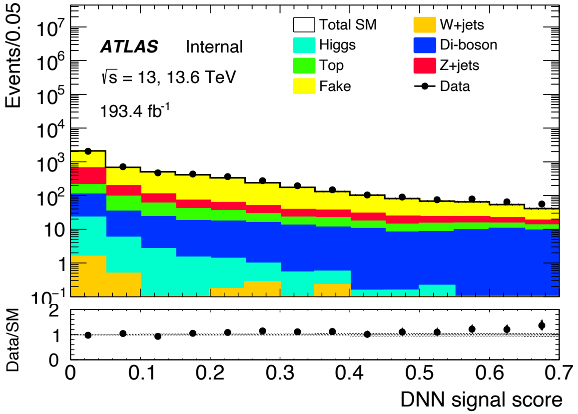




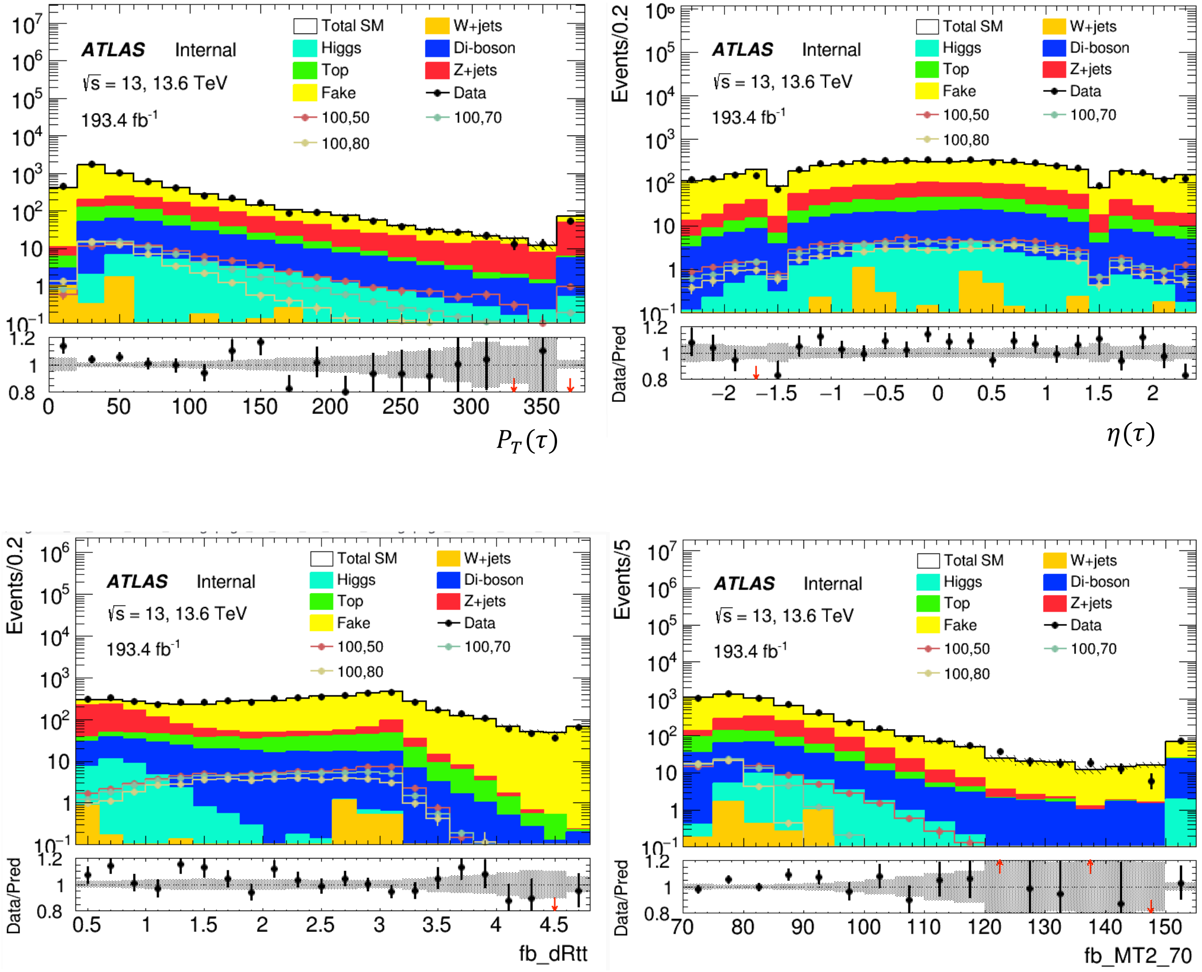
Direct Stau ISR fake estimation validation(HH)

- Data-driven fakes in preselection region

Score distribution

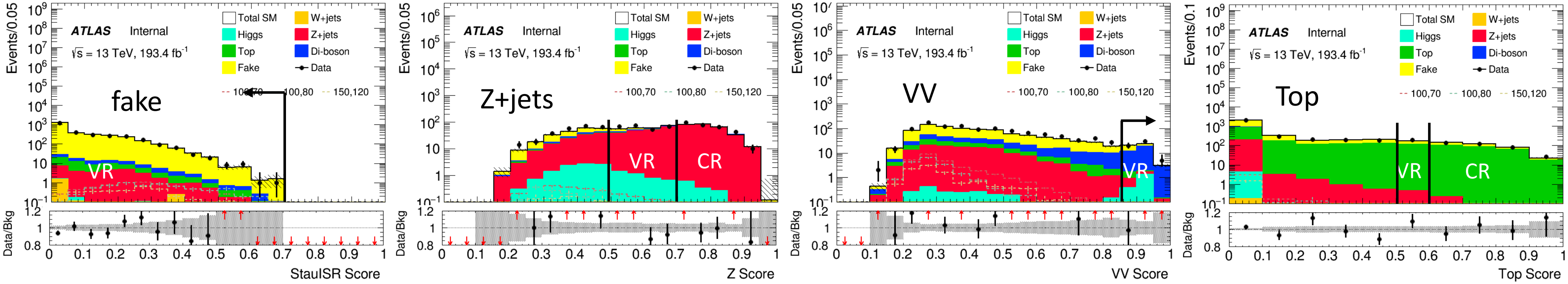


Kinematic distribution



Direct Stau ISR background estimation (HH)

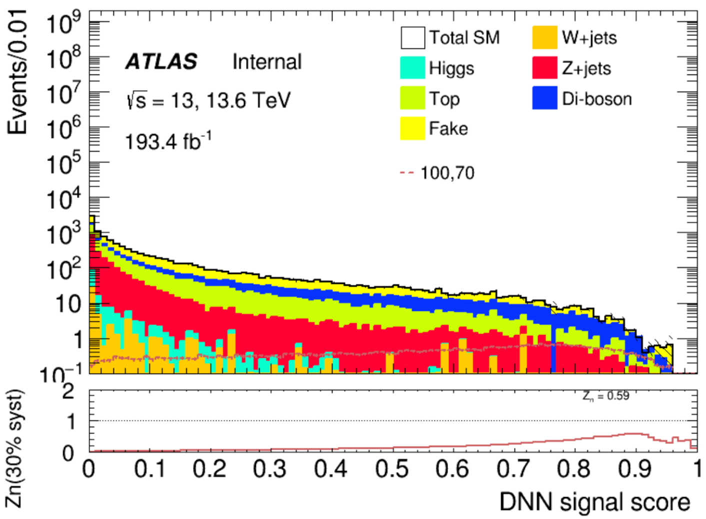
Region	Selections		Total Bkg	Dominant Bkg	Purity	Data	Data/Bkg
Fake VR	**signal score < 0.7, max bkg score: Fake**		2870 +- 40	2770 +- 40	**0.96**	2794	0.972
Z CR	**signal score < 0.7, max bkg score: Z**	Z Score > 0.7	276 ± 7	257 ± 6	**0.93**	295	1.07
Z VR		Z Score ∈ (0.5, 0.7)	251 ± 7	198 ± 5	**0.79**	264	1.05
VV VR	**signal score < 0.7, max bkg score: VV**	VV Score > 0.85	48 ± 2.8	30 ± 1.1	**0.62**	55	1.15
Top CR	HH **bTag**	Top Score > 0.6	374 +- 9	271 +- 5	**0.72**	375	1
Top VR		Top Score ∈ (0.5, 0.6)	178 +- 6	178 +- 6	**0.64**	195	1.09

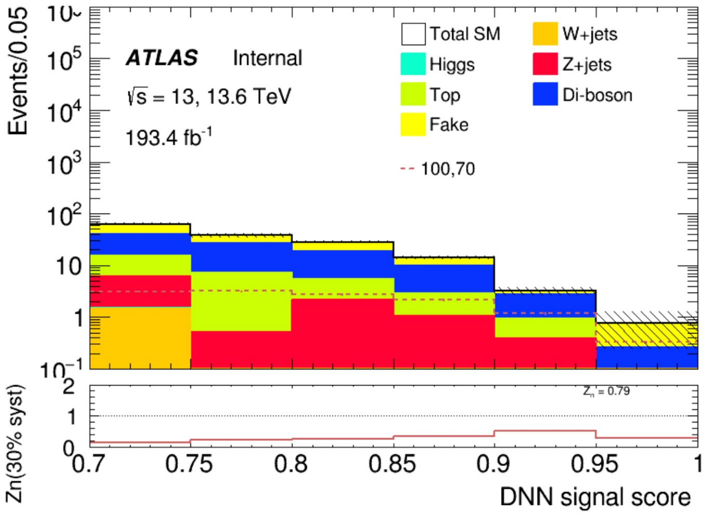


Direct Stau ISR signal region definition(LH)

- Preselection

LH Pre-selection
>= 1 medium taus
1base lepton, 1 signal lepton
Opposite-sign
bveto
MET trigger
MET ≥ 200
$\Delta\phi(\tau, MET) < 2$
$M_{T2,70} < 100$
$\Delta R_{\tau\ell} \in (0.6, 3.6)$





- SR selection using neuron network score
  - Fake tau estimation using fake factor method
  - DNN signal score > 0.7

SR
Pre-selection
DNN score > 0.7

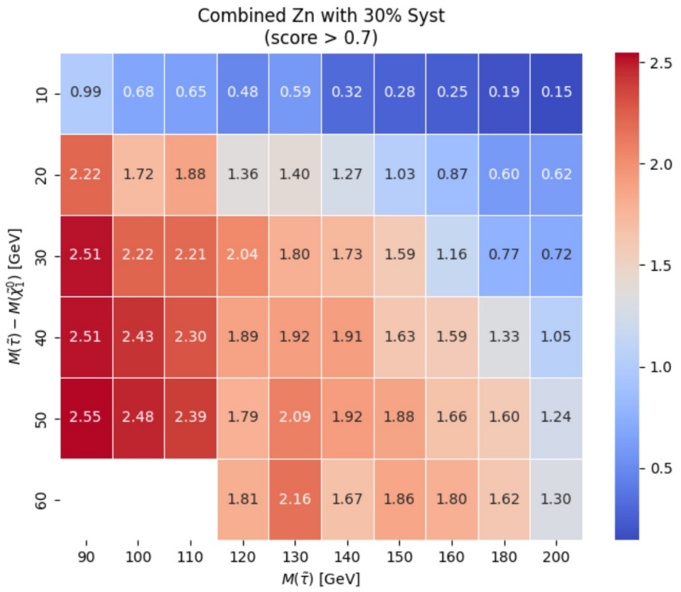
- Background composition

Process	[0.70,0.75]	[0.75,0.80]	[0.80,0.85]	[0.85,0.90]	[0.90,0.95]	[0.95,1.00]	Combined
TotalBkg	63.19±3.25	38.56±3.70	28.54±2.06	14.29±1.34	3.17±0.54	0.75±0.51	148.50±5.56
VV	25.27±0.86	19.64±0.81	12.90±0.63	6.97±0.45	1.77±0.24	0.21±0.04	66.76±1.43
Fake	22.41±2.65	11.87±1.87	10.13±1.85	4.40±1.16	0.45±0.40	0.49±0.51	49.75±3.96
Top	9.28±1.05	6.53±0.86	3.32±0.60	1.86±0.48	0.56±0.26	0.00±0.00	21.56±1.58
Zjets	4.74±0.29	3.33±0.25	2.16±0.24	1.03±0.16	0.37±0.10	0.05±0.02	11.68±0.49
Higgs	0.06±0.03	0.11±0.04	0.03±0.01	0.03±0.01	0.00±0.00	0.00±0.00	0.22±0.06
Wjets	1.44±1.28	-2.92±2.95	0.00±0.00	0.00±0.00	0.01±0.01	0.00±0.00	-1.47±3.22
StauStauISR-100-70	3.13±0.07	3.17±0.07	2.75±0.07	2.17±0.06	1.18±0.04	0.32±0.02	12.71±0.15
ZnSignificance	0.15	0.23	0.26	0.36	0.53	0.29	0.79

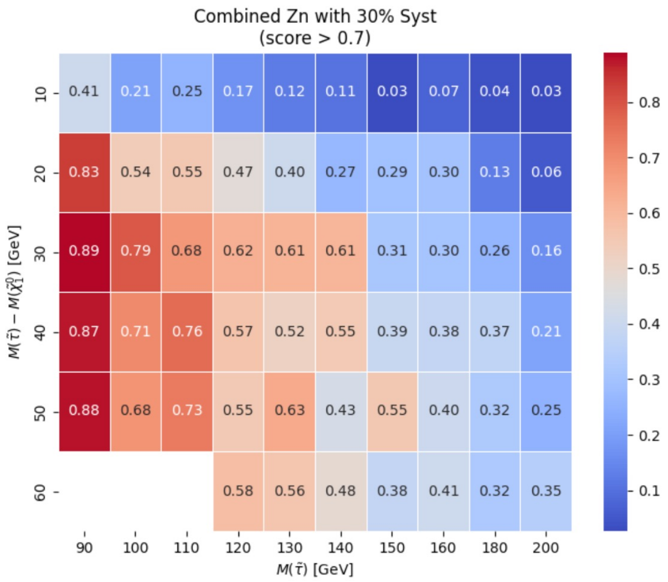
Direct Stau ISR signal region definition

- Expected sensitvity
  - 30% flat systematic uncertainty

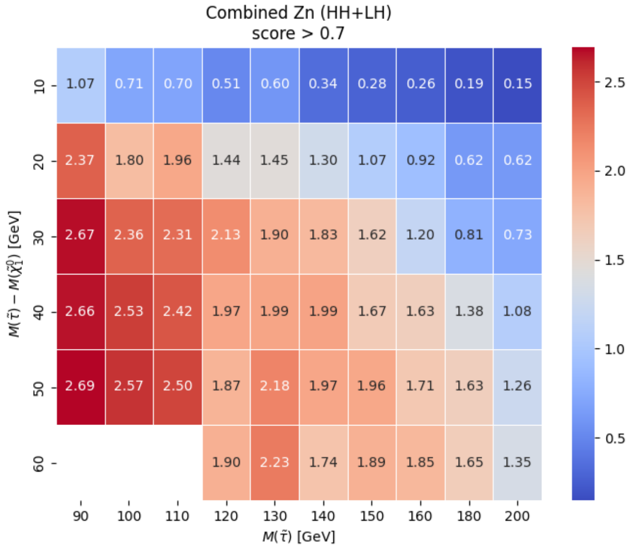
HH channel



LH channel



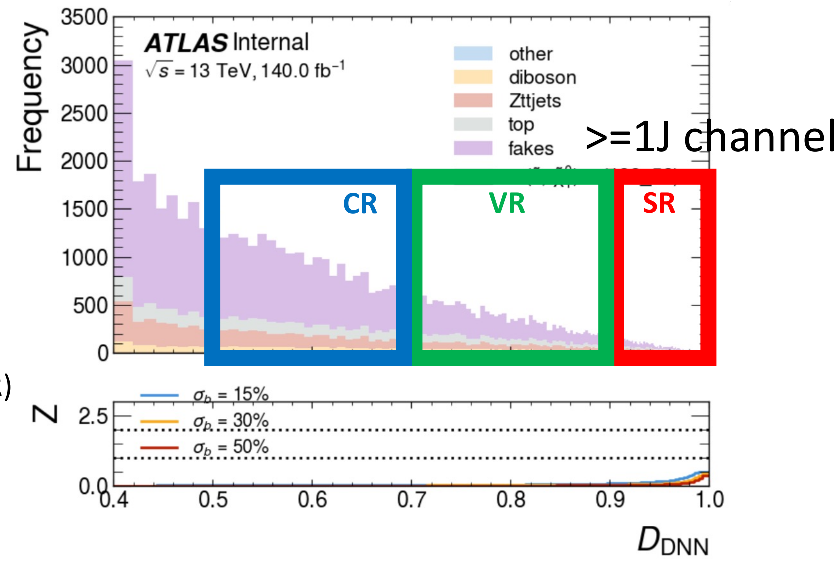
Combined channel

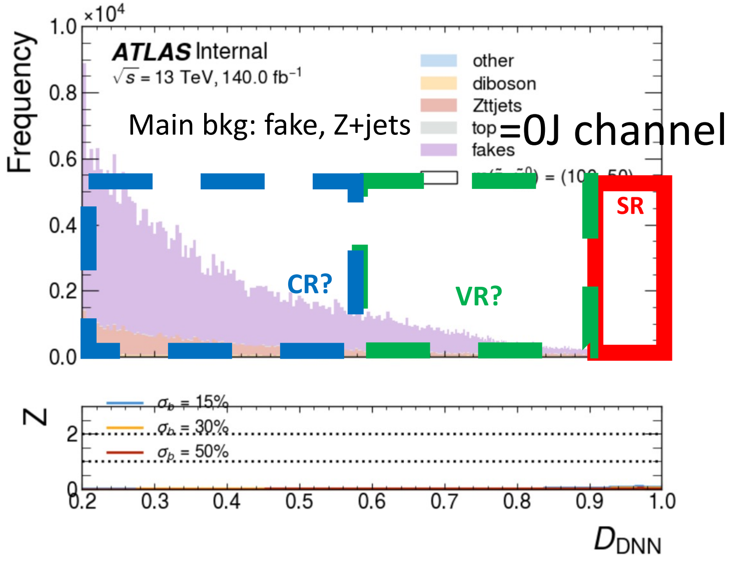


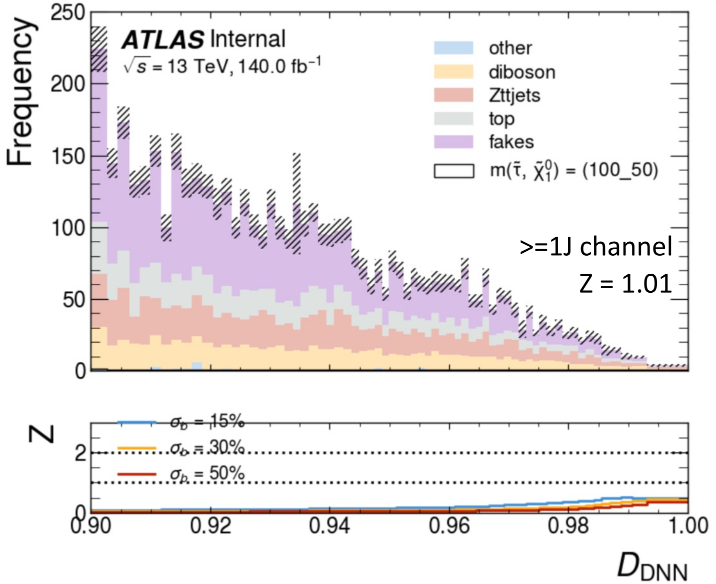
Direct Stau signle-lep trigger SR definition(LH)

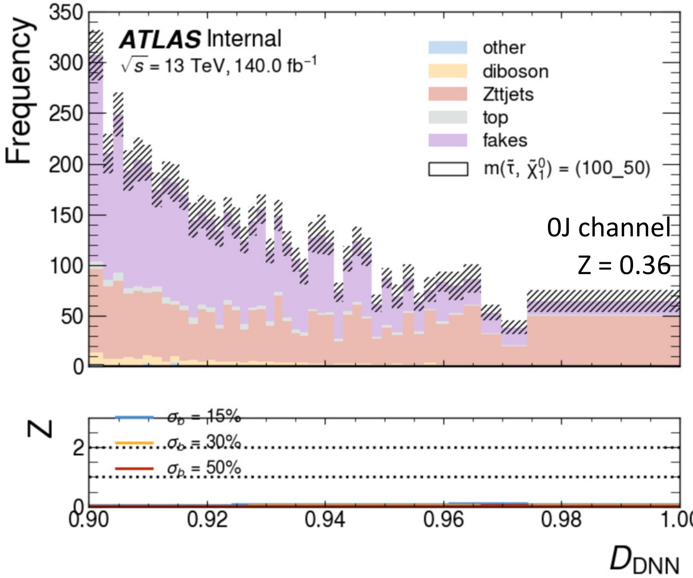
LH Pre-selection
== 1 Tight taus (OS)
1 signal lepton
bveto
Single-lepton trigger
15 GeV < MET < 200 GeV (orthogonal with ISR SR)
METSig > 2
MT2 < 90
$dR(lep, tau) < 3.6$
custom MT cut*
$\phi_{\{1,2\}}$ cut*

SR	
Pre-selection	
>=1 Jet	== 0 Jet
DNN score > 0.9	









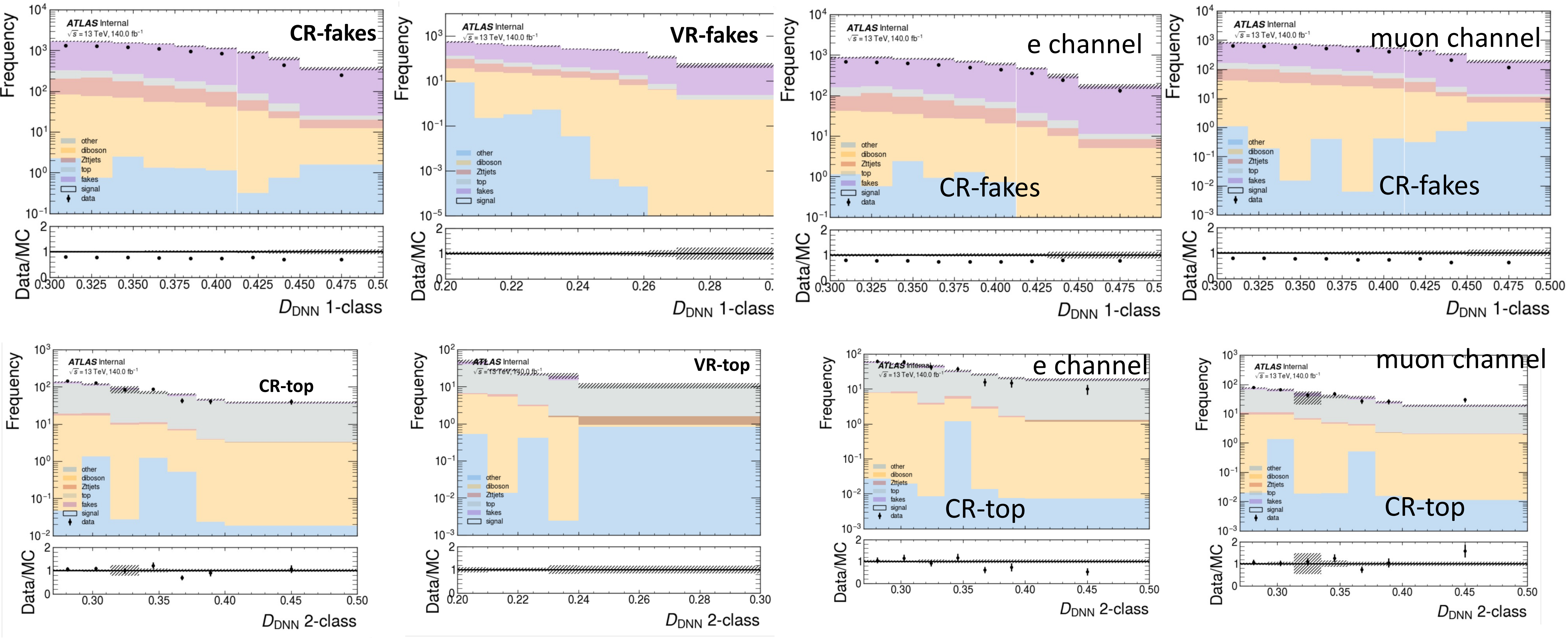
Direct Stau single-lep trigger background estimation(LH)

>= 1J channel

Region	Selections		channel	Total Bkg		Dominant Bkg		Purity	
Fake CR	**0.5 < signal score < 0.7, max bkg score: Fake**	**fake score > 0.3**	e-channel	5557	10628	4746	9029	0.85	0.85
			muon-channel	4283		4283		0.84	
Top CR	**0.5 < signal score < 0.7, max bkg score: Top**	**Top score > 0.27**	e-channel	256	555	203	449	0.79	0.82
			muon-channel	299		245		0.82	
Z CR	**0.5 < signal score < 0.7, max bkg score: Z**	**Z score > 0.28**	e-channel	457	775	374	631	0.82	0.82
			muon-channel	317		256		0.81	
VV CR	**0.0 < signal score < 0.7, max bkg score: VV**	VV Score > 0	e-channel	145	264	72	134	0.50	0.50
			muon-channel	118		61		0.52	
Fake VR	**0.7 < signal score < 0.9, max bkg score: Fake**	**fake score > 0.2**	--	2601		2166		0.83	
Top VR	**0.7 < signal score < 0.9, max bkg score: Top**	**Top score > 0.2**	--	128		95		0.74	
Z VR	**0.7 < signal score < 0.9, max bkg score: Z**	**Z score > 0.165**	--	247		189		0.76	
VV VR	**0.7 < signal score < 0.9, max bkg score: VV**	**VV score > 0.06**	--	166		84		0.51	

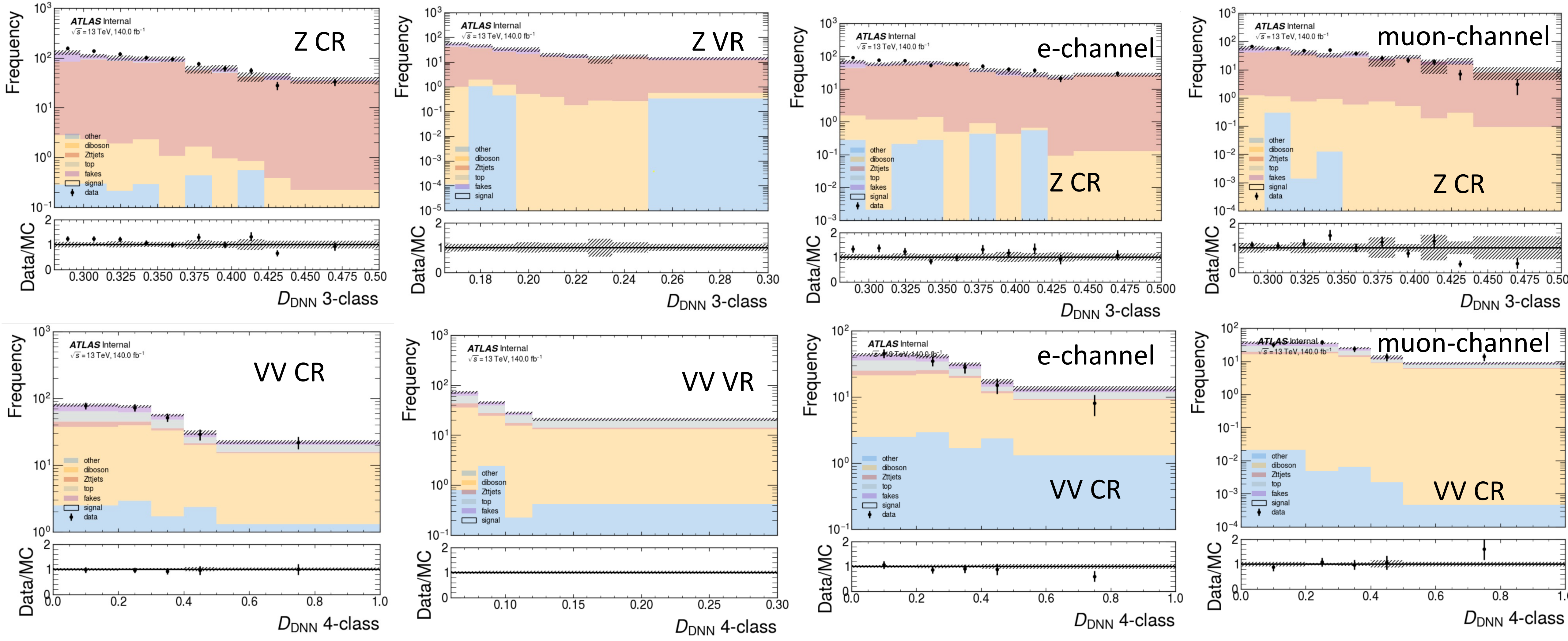
***(Note: MC fakes are only place holders, will be replaced with data driven background in final fit and CR will be added to VR)**

Direct Stau single-lep trigger background estimation (LH)



*(Note: MC fakes are only place holders, will be replaced with data driven background in final fit and CR will be added to VR)

Direct Stau single-lep trigger background estimation(LH)

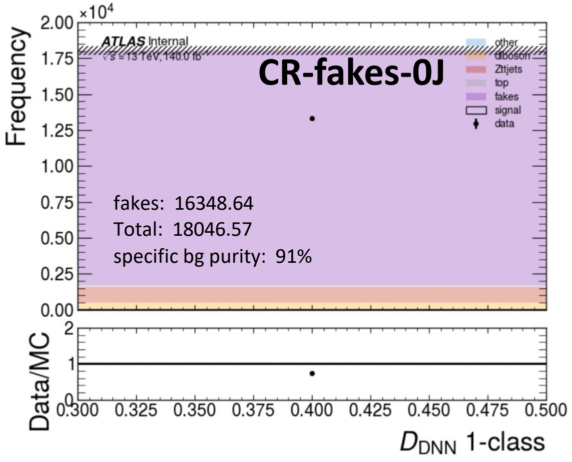


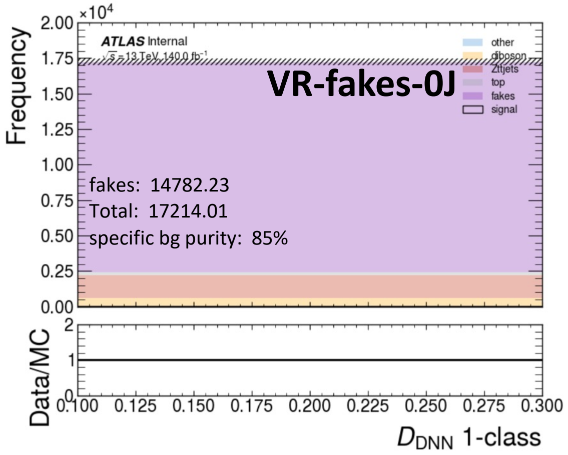
***(Note: MC fakes are only place holders, will be replaced with data driven background in final fit and CR will be added to VR)**

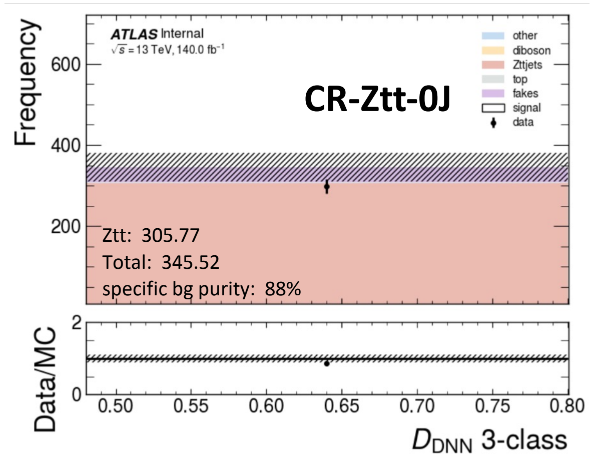
Direct Stau single-lep trigger background estimation(LH)

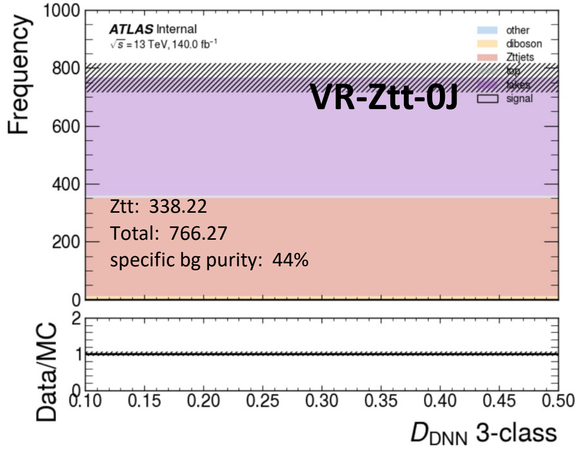
= 0J channel

Region	Selections		Total Bkg	Dominant Bkg	Purity
Fake CR	0.5 < signal score < 0.7, max bkg score: Fake	fake score > 0.3	18046	16348	0.91
Z CR	0.2 < signal score < 0.5, max bkg score: Z	Z score > 0.48	345	305	0.88
Fake VR	0.7 < signal score < 0.9, max bkg score: Fake	fake score > 0.1	17214	14782	0.85
Z VR	0.5 < signal score < 0.9, max bkg score: Z	Z score > 0.1	766	338	0.44







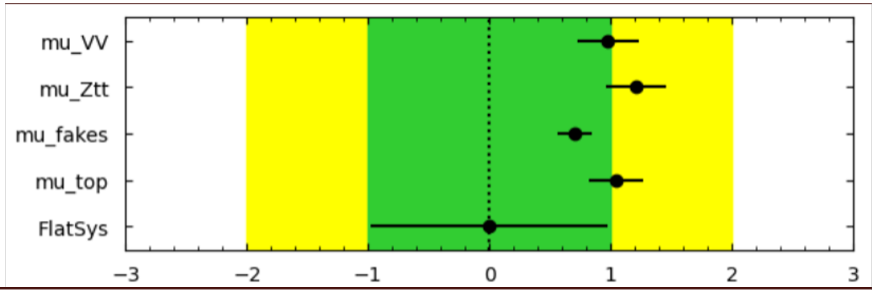


***(Note: MC fakes are only place holders, will be replaced with data driven background in final fit and CR will be added to VR)**

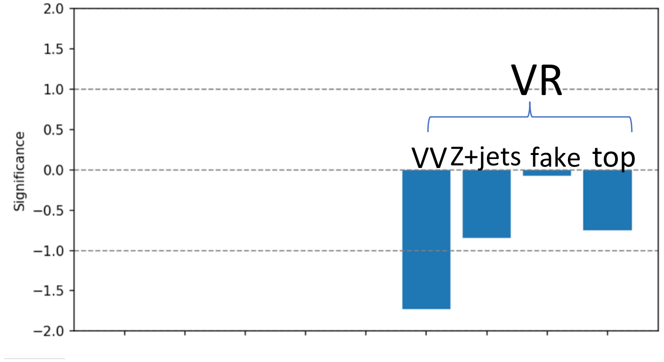
Direct Stau single-lep trigger background only fit (LH)

Norm Factor	results	Norm Factor	results
mu_VV	0.9763+-0.2523	mu_VV_e	0.8181+-0.1909
		mu_VV_m	1.1690+-0.2165
mu_Ztt	1.2063+-0.2497	mu_Ztt_e	1.2519+-0.0929
		mu_Ztt_m	1.1379+-0.1224
mu_fake	0.7051+-0.1398	mu_fakes_e	0.7060+-0.0259
		mu_fakes_m	0.7044+-0.0268
mu_top	1.0451+-0.2208	mu_top_e	1.1251+-0.1277
		mu_top_m	0.9486+-0.0945
FaltSys	-0.0009+-0.9745	FaltSys	0.0008+-0.9934

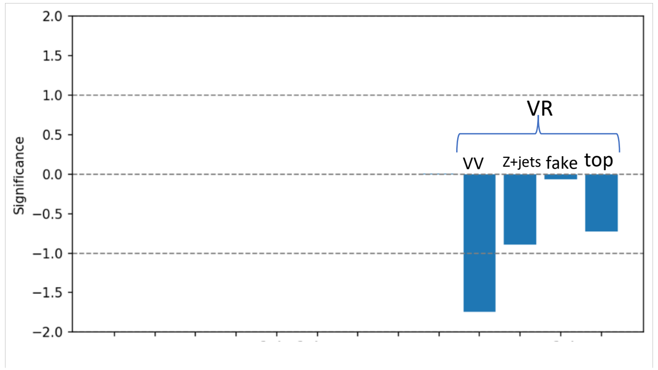
- Norm factors indicate overprediction for fake
  - Hence data-driven method
- VRs have slight over prediction (but all under 2-sigma)

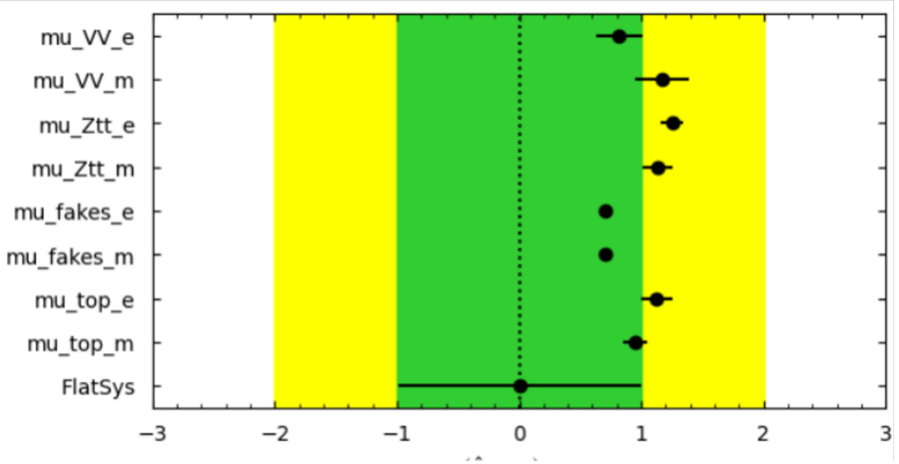


>= 1 J channel(combined)



>= 1J channel e/mu separate

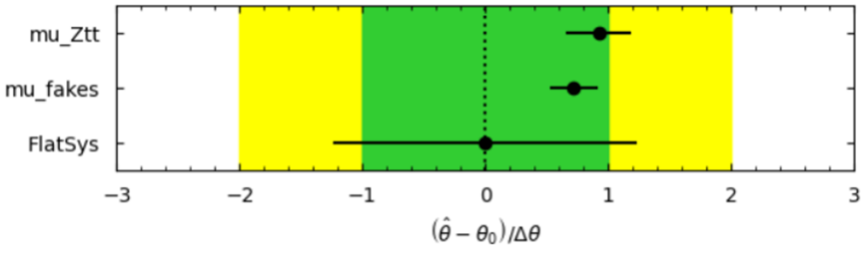


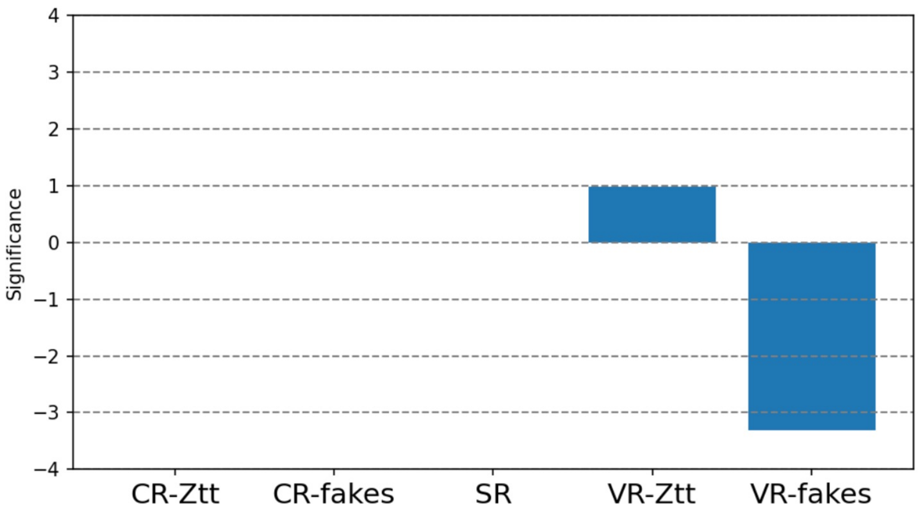


Direct Stau single-lep trigger background only fit (LH)

= 0 J channel

Norm Factor	results
mu_Ztt	0.9223+-0.2675
mu_fake	0.7189+-0.1905
FaltSys	-0.0024





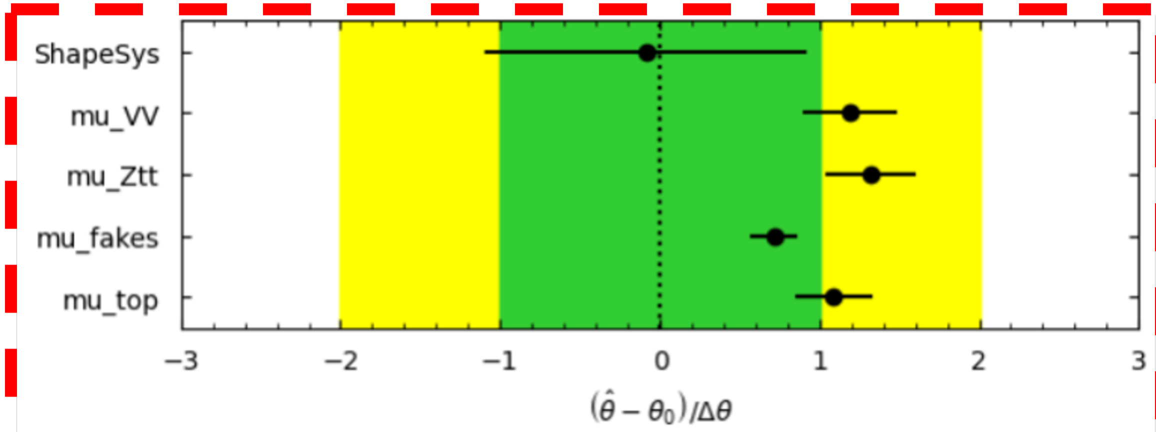
- Observe >3 sigma overestimation for fake
  - Futher solidifies need for data-driven fakes
- Z+jets agreement in VR very good
- Top and boson minor backgrounds in this channel

Direct Stau single-lep trigger Exclusion Fit(LH)

>= 1J channel

- expected CLs very robust to size of systematics
- Sensitivity worse than estimated from simple Z score formula
- (Mu factors preliminary since dominant fake background MC will be replaced)*

Type of systematic	% Syst.	[	Expected CLs	]
Normalisation	1	0.238	0.437	0.699
Normalisation	10	0.238	0.437	0.699
Normalisation	20	0.238	0.437	0.699
Normalisation	50	0.240	0.439	0.701
Norm+Shape	1	0.238	0.437	0.699
Norm+Shape	20	0.238	0.437	0.699
Norm+Shape	50	0.240	0.439	0.701

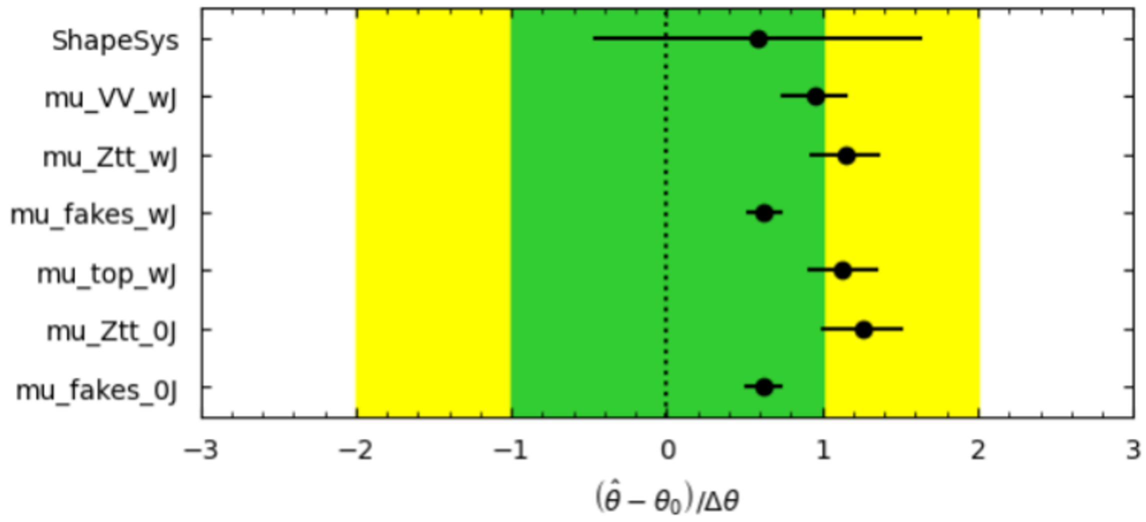


Direct Stau single-lep trigger Exclusion Fit(LH)

0J and >= 1J channel combined

- a slight decrease in the expected CLs, but not much
- DNN studies showed >=1J channel drives sensitivity

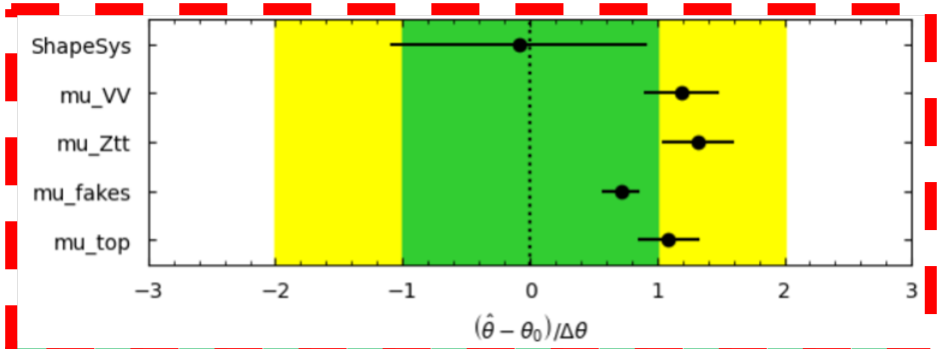
Type of systematic	% Syst.	[	Expected CLs	]
Normalisation	20	0.231	0.429	0.693
Norm+Shape	20	0.231	0.429	0.693

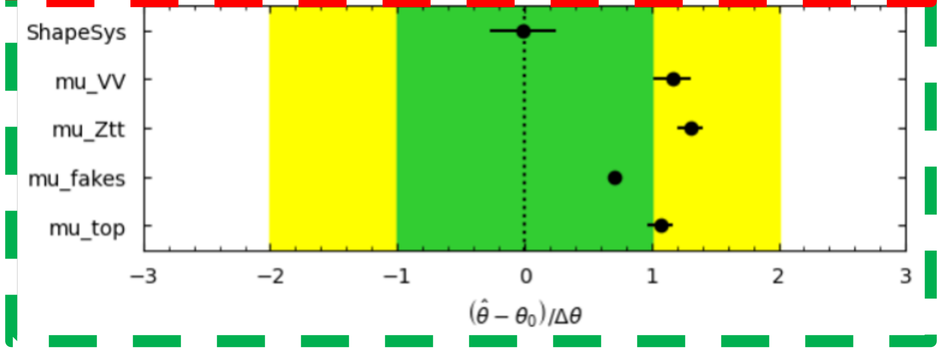


Direct Stau single-lep trigger OS-SS Fit(LH)

- Idea to use same-sign (SS) region in the fit to constrain systematics
- small impact on resulting CLs value
- Small signal statistics remains main issue

Type of fit	% Syst.	[	Expected CLs	]
OS only	20	0.242	0.442	0.703
OS-SS	20	0.242	0.441	0.702

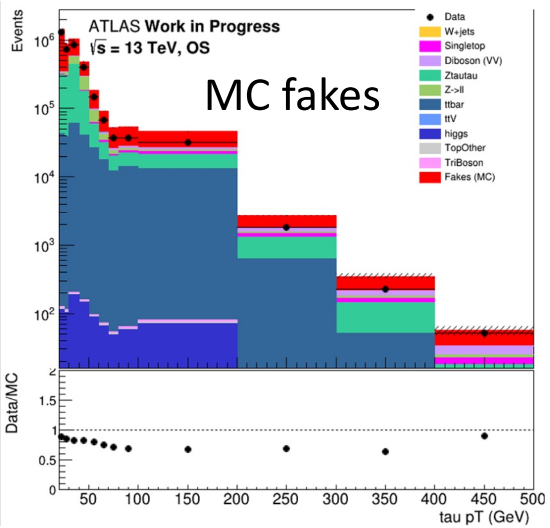




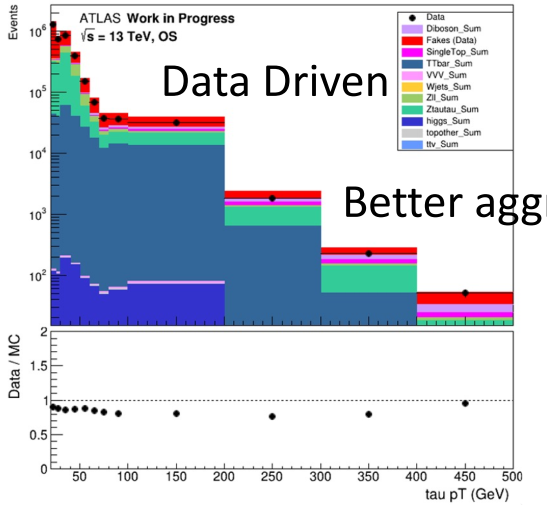
Direct Stau fake estimation(LH)

- CRs (fake factor computation)
  - Single lepton trigger
  - OS/SS
  - 1 signal lepton
  - **ID: >= 1 tight tau**
  - **antiID: > 0.05 RNN score, no tight tau**
- SRs
  - Single lepton trigger
  - MET > 15GeV
  - B-veto
  - OS/SS
  - **ID: >= 1 tight tau**
  - **antiID: > 0.05 RNN score, no tight tau**
- Focus on W+jets and ttbar
- **Binned in prongness, tau pT**

FF determination Data	FF application Data
CR Pass ID	SR Pass ID
CR Fail ID	SR Fail ID

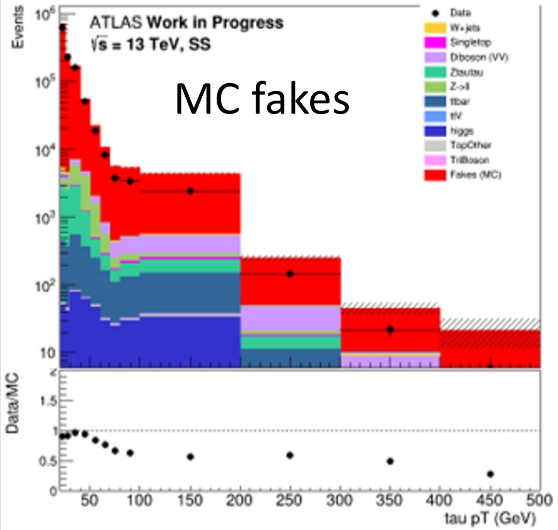


MC fakes

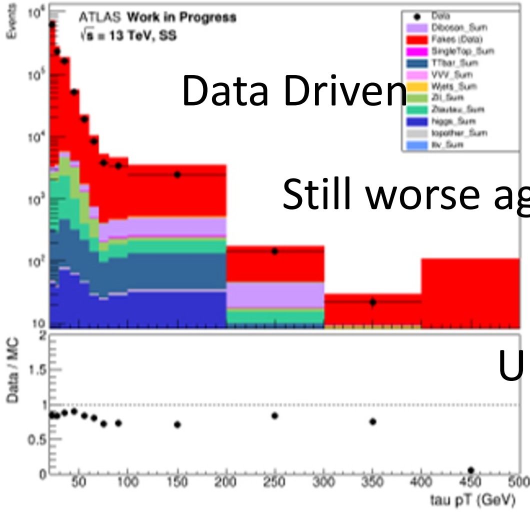


Data Driven

Better aggrement for high pt



MC fakes



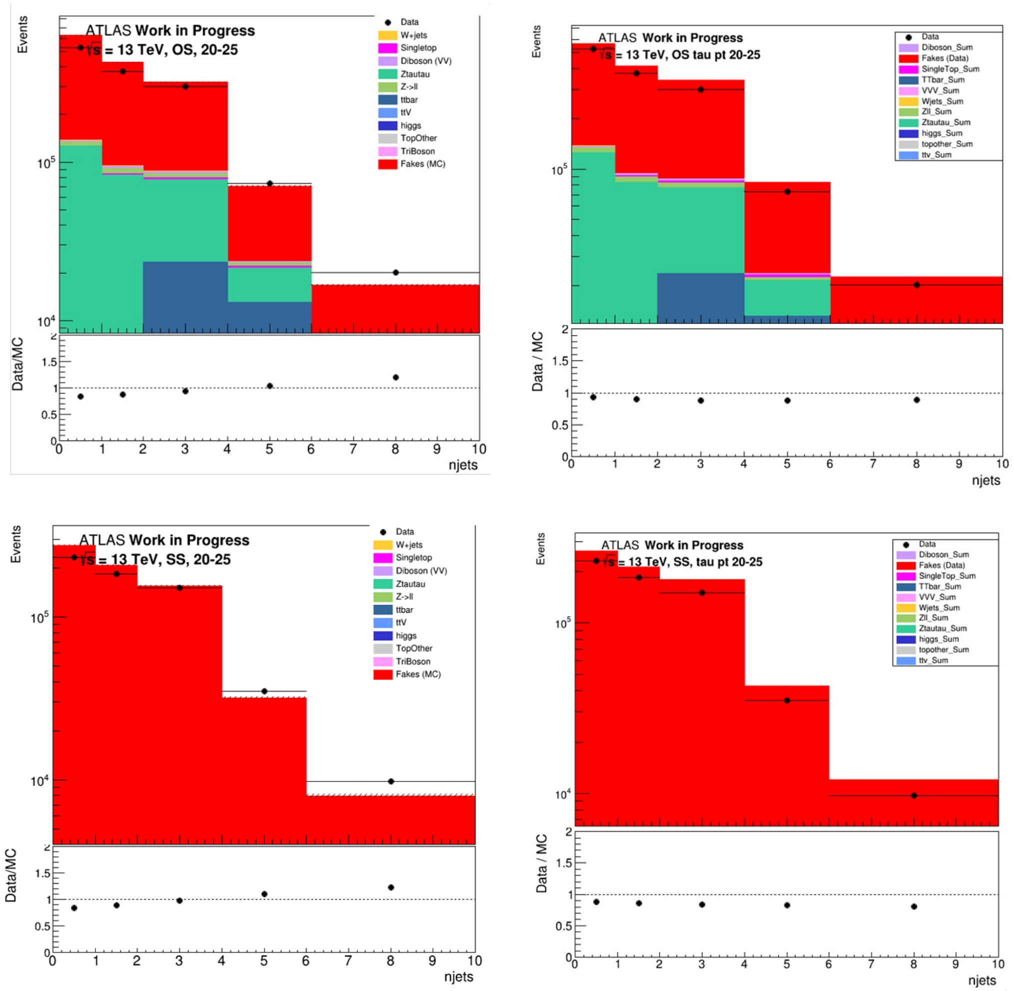
Data Driven

Still worse aggrement for high pt

Under investigated

Direct Stau fake estimation(LH)

- Focus on events with taus with pT 20-25 GeV
- Top: OS
- Bottom: SS
- Better agreement with data-driven fake taus compared to MC for different numbers of jets in the event

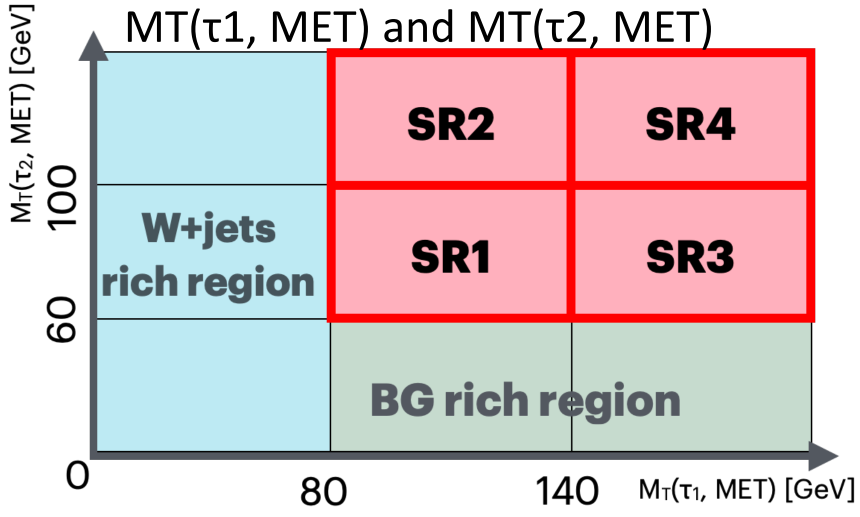


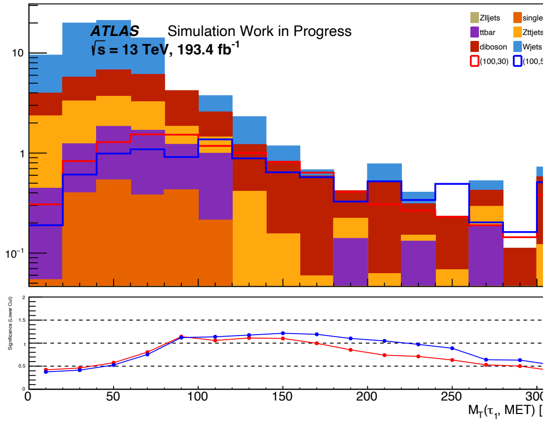
Direct Stau ISR Cut and Count(HH)

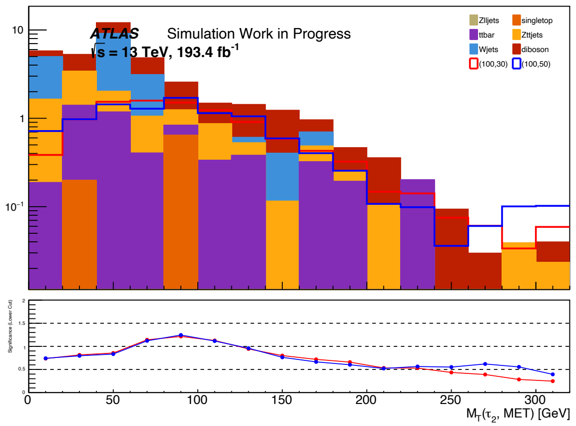
HH Pre-selection
>= 2 medium taus (OS)
0 base lepton
bveto
MET trigger
MET ≥ 200 GeV
Primary Jet pt > 200GeV
$\tau_{pt} > 20GeV$
Collinear mass($\tau\tau$) > 250GeV

SR before binned
Pre-selection
$\boldsymbol{M_{T2}(0) < 5GeV}$
$\boldsymbol{M_{T2}(70) < 85GeV}$
$\boldsymbol{d\phi(MET, Jet) > 0.4}$
$\boldsymbol{d\eta(\tau_1, \tau_2) < 2.0}$
$\boldsymbol{dR(\tau_1, \tau_2) < 3.2}$
$\boldsymbol{dR(\tau, MET) > 1}$
$\boldsymbol{M_T(\tau_1, MET) > 80GeV}$
$\boldsymbol{M_T(\tau_2, MET) > 60GeV}$

Signal Region definition:
Binned for the combination of
MT(τ1, MET) and MT(τ2, MET)



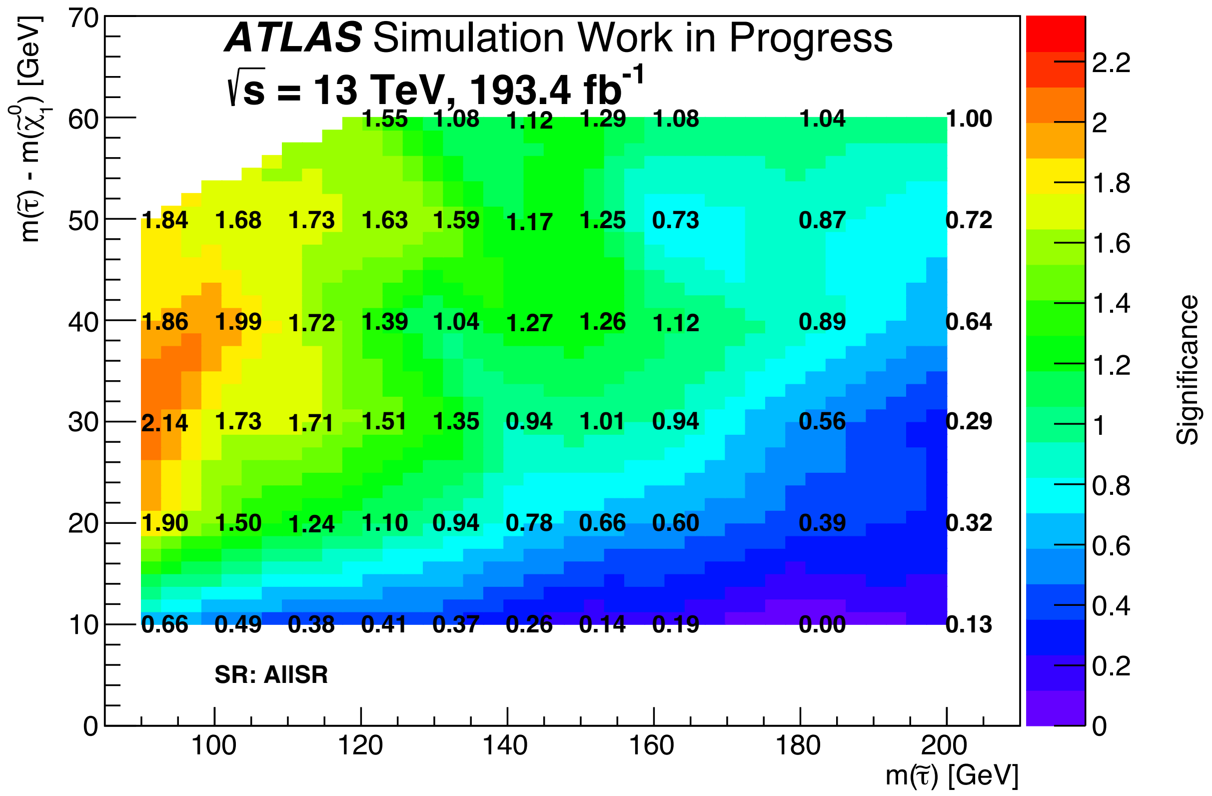




Direct Stau ISR Cut and Count(HH)

	SR1	SR2	SR3	SR4	All
VV	2.88	2.49	2.58	1.05	9
Top	2.05	1.16	0.20	0.32	3.73
Wjets	0.32	0.40	0.48	0.0	1.2
Zjets	1.06	0.79	0.57	0.32	2.74
BG Sum	6.31	4.84	3.83	1.69	**16.67**
Signal (100, 70)	2.35	2.15	2.00	1.73	**8.23**
Significance	0.69	0.75	0.80	1.14	**1.73**

Sensitivity for any pair production of $\tilde{\tau}_l + \tilde{\tau}_R$ pair-production with 30% flat systematic uncertainty

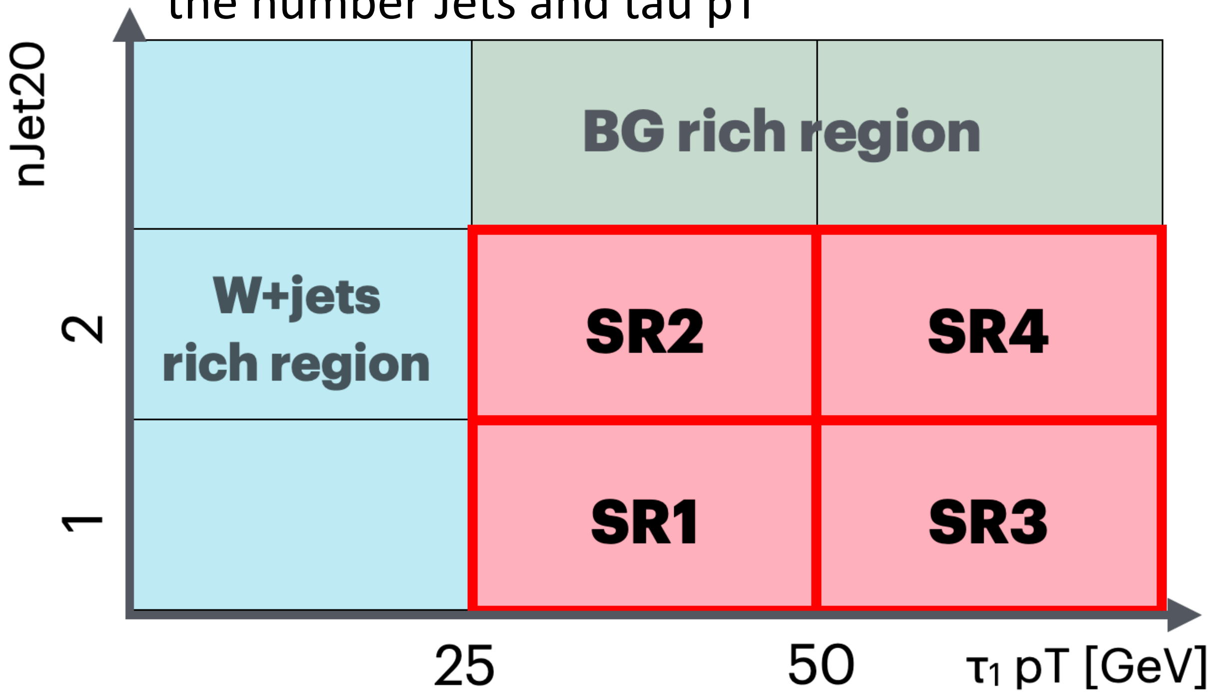


Direct Stau ISR Cut and Count(LH)

LH Pre-selection
>= 1 medium taus (OS)
1 base lepton, 1 signal lepton
bveto
MET trigger
MET ≥ 130 GeV
Primary Jet pt > 50GeV
$\tau_{pt} > 25GeV$
Collinear mass(τl) > 130GeV

SR
Pre-selection
$\boldsymbol{M_{T2}(40) < 55GeV}$
$\boldsymbol{\frac{pt_{lep}}{M_T(\tau_1, MET)} < 0.1}$
$\boldsymbol{M_T(\tau_1, MET) > 90GeV}$

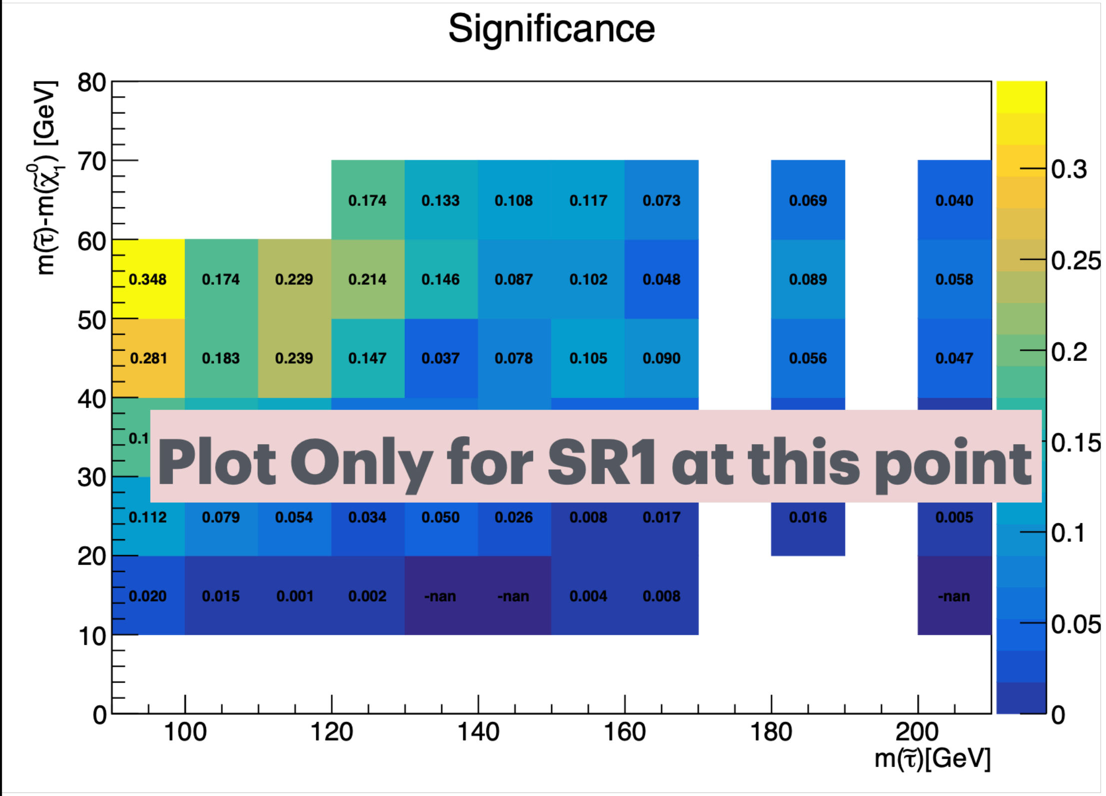
Signal Region definition:
Binned for the combination of
the number Jets and tau pT



Direct Stau ISR Cut and Count(LH)

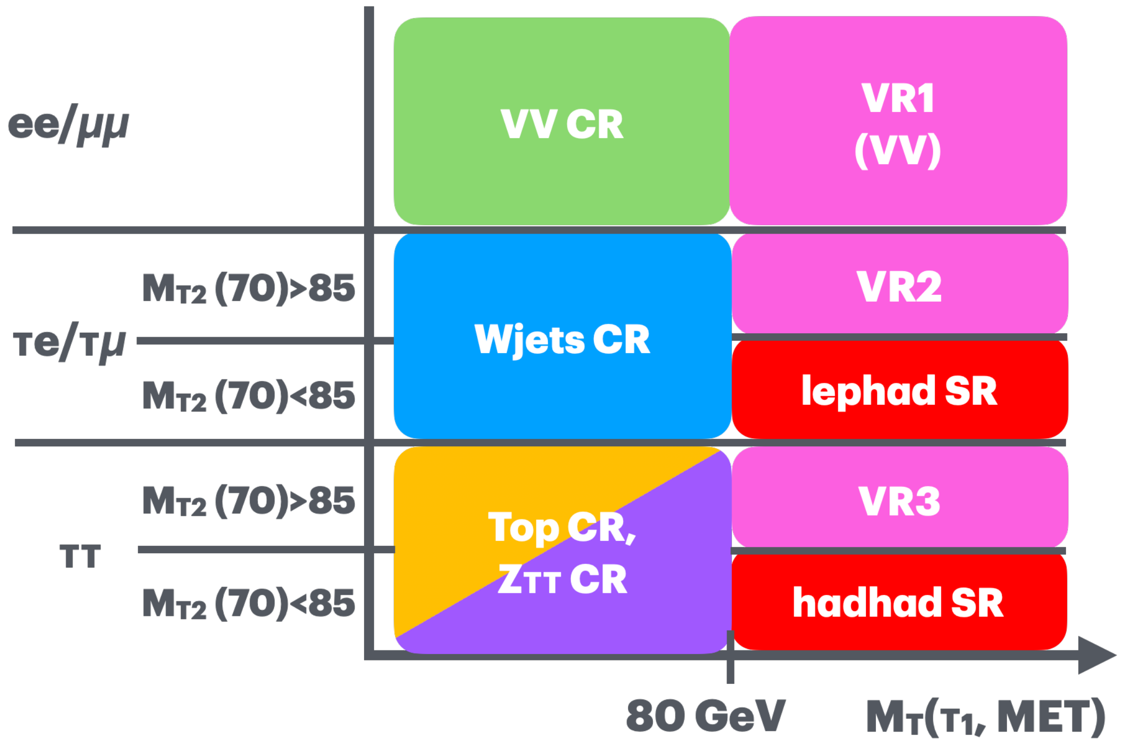
Sensitivity map with 30% systematic uncertainty

	SR1	SR2	SR3	SR4	All
VV	6.24	11.01	4.27	5.83	27.35
Top	5.93	11.02	4.48	7.80	29.23
Wjets	13.48	28.78	13.60	18.49	74.35
Zjets	5.71	7.67	1.71	4.17	19.26
BG Sum	31.36	58.49	24.06	36.29	**150.2**
Signal (100, 70)	1.70	1.22	1.08	0.84	**4.84**
Significance	0.15	0.06	0.12	0.07	**0.4**



Direct Stau ISR Background Estimation

- Common estimation for both had-had and lep-had
- No large difference in distributions of $M_T(\tau_1, MET)$
- Prepare CR for each of the BG
  - Use leplep / lephad mode
    - W decay and lepton ID (including τ) are modeled well in MC
  - $M_T(\tau_1, MET)$ distribution matches as well
    → Extrapolation along M_T and lepton flavor
  - Reverse b-veto or colinear mass to obtain top/Z rich CRs
- VR: Reverse MT2 cut
- Fake estimation:
  - Estimation using MC, setting Wjets CR
  - Validation with Fake Factor method planned



Direct Stau ISR Background Estimation

- Common Preselection: Pass MET Trigger
- Each SR is binned and has specific cuts (see SR definitions)

<table>
<tr><th></th><th>SR (HH)</th><th>SR (LH)</th><th>CR (VV)</th><th>CR (Wjets)</th><th>CR (Z$\tau\tau$)</th><th>CR (Top)</th><th>VR (VV)</th><th>VR (HH)</th><th>VR (LH)</th></tr>
<tr><td>MET</td><td>> 200 GeV</td><td>> 130 GeV</td><td colspan="4">> 200 GeV</td><td colspan="2">> 200 GeV</td><td>> 130 GeV</td></tr>
<tr><td>nJet20</td><td><= 4</td><td><= 2</td><td colspan="4"><= 4</td><td colspan="2"><= 4</td><td><= 2</td></tr>
<tr><td>nBJet20 (WP 90%)</td><td colspan="2">= 0</td><td colspan="3">= 0</td><td>>= 1</td><td colspan="3">= 0</td></tr>
<tr><td>Primary Jet pT</td><td>> 200 GeV</td><td>> 50 GeV</td><td colspan="4">> 200 GeV</td><td colspan="2">> 200 GeV</td><td>> 50 GeV</td></tr>
<tr><td>(nLightLeptons, nHadronic Taus)</td><td>= (0, 2)</td><td>= (1, 1)</td><td>= (2, 0)</td><td>= (1, 1)</td><td colspan="2">= (0, 2)</td><td>= (2, 0)</td><td>= (0, 2)</td><td>= (1, 1)</td></tr>
<tr><td>tau pT</td><td>> 20 GeV</td><td>> 25 GeV</td><td>-</td><td colspan="3">> 20 GeV</td><td>-</td><td>> 20 GeV</td><td>> 25 GeV</td></tr>
<tr><td>M_T (τ_1, MET)</td><td>> 80 GeV</td><td>> 90 GeV</td><td colspan="4">< 80 GeV</td><td colspan="2">> 80 GeV</td><td>> 90 GeV</td></tr>
<tr><td>Collinear mass</td><td>> 250 GeV</td><td>> 130 GeV</td><td colspan="2">> 250 GeV</td><td>∈ [70, 110] GeV</td><td>> 250 GeV</td><td colspan="2">> 250 GeV</td><td>> 130 GeV</td></tr>
<tr><td>M_{T2}(70)</td><td>< 85 GeV</td><td>-</td><td colspan="4">-</td><td colspan="2">> 85 GeV</td><td>-</td></tr>
<tr><td>M_{T2}(40)</td><td>-</td><td>< 55 GeV</td><td colspan="4">-</td><td colspan="2">-</td><td>> 55 GeV</td></tr>
</table>

C1C1 ISR signal region optimization

- BDT method for signal region optimization
  - 30% flat systemic uncertainty
  - Figure of merit: Zn
  - 5fold to Train(Cross-Validation)
- Preselection based on final state($\tau\ \tau$ or $\tau\ l$)

HH Pre-selection	LH Pre-selection
>= 2 medium taus	>= 1 medium taus
0 base lepton	1 base lepton, 1 signal lepton
MET ≥ 150; pass MET trigger	MET ≥ 200; pass MET trigger
1≤nJet	1≤nJet
Opposite-sign hadronic-hadronic tau pair	Opposite-sign lepton-hadronic tau pair
bveto	bveto
jet pt>100 GeV	jet pt>100 GeV

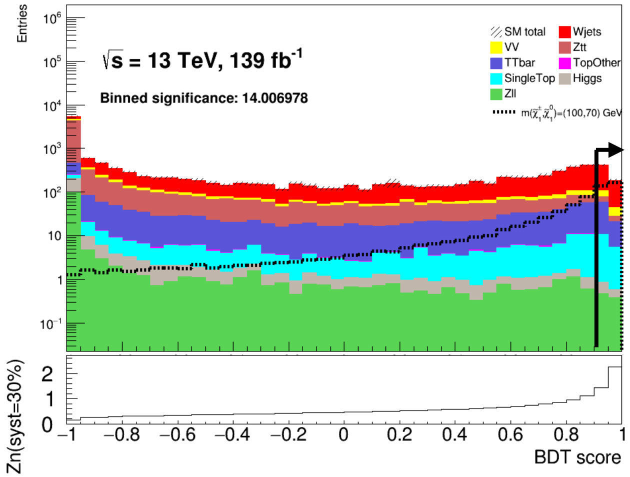
Table 3: Best performance for HH and LH channel after grid search

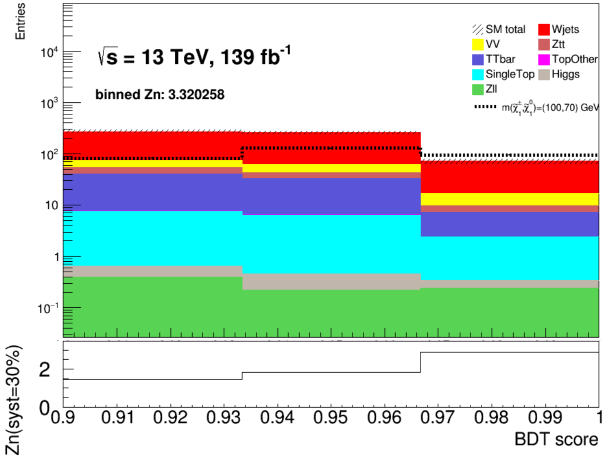
Channel	Ntrees	MaxDepth	MinNode	Learning rate
had-had channel	400	8	1	0.1
lep-had channel	300	6	1	0.05

C1C1 ISR signal region definition(HH)

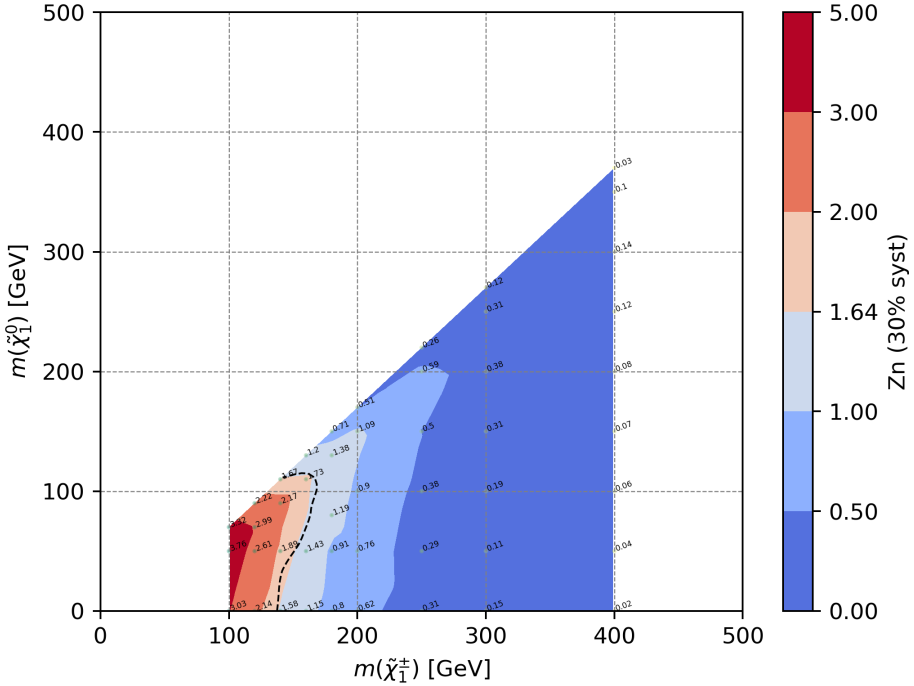
SR
Pre-selection
BDT score > 0.9

Binned significance: 3.32





Sensitive map

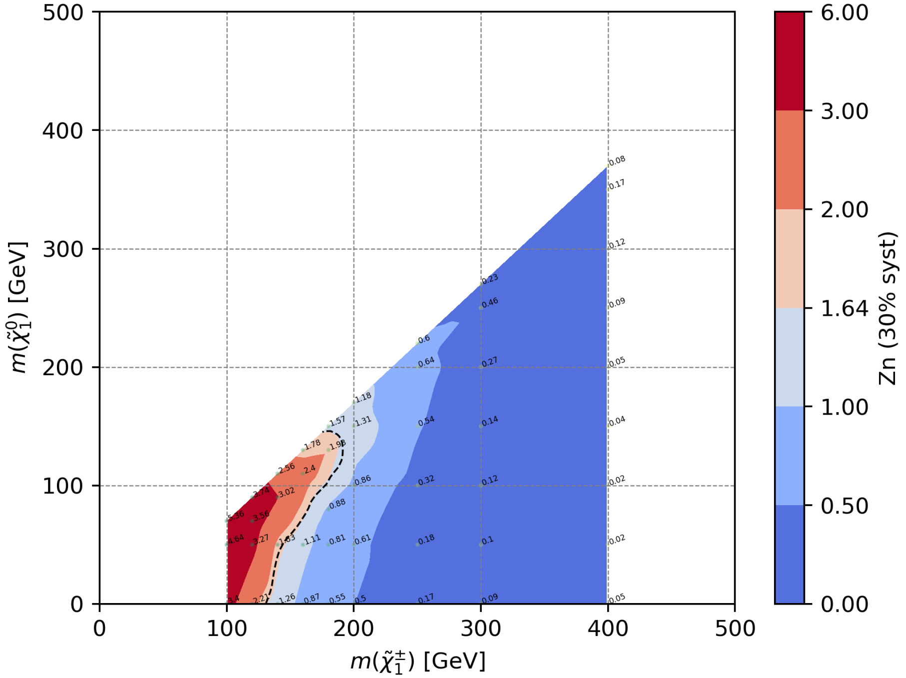


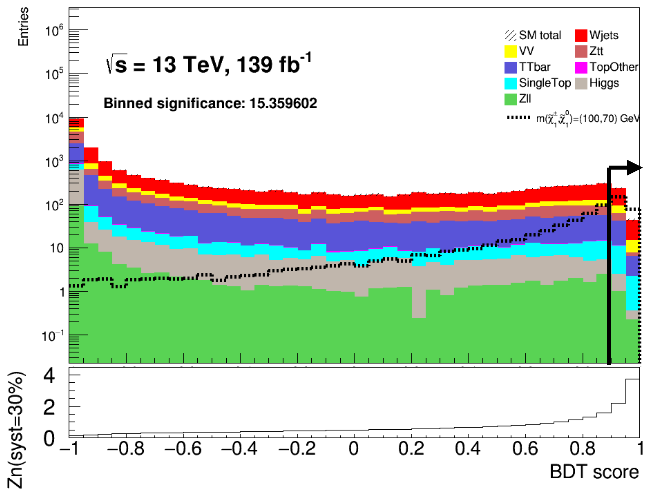
SM process	SR2	SR2-Bin1	SR2-Bin2	SR2-Bin3
Wjets	448 ± 28	197 ± 18	55 ± 9	197 ± 19
VV	49.1 ± 1.2	20.7 ± 0.9	7.4 ± 0.4	21.0 ± 0.9
Ztt	25.3 ± 2.9	9.8 ± 1.6	2.3 ± 1.1	13.1 ± 2.1
TTbar	64.5 ± 3.1	26.7 ± 2.0	4.9 ± 0.9	32.9 ± 2.2
TopOther	0.38 ± 0.07	0.15 ± 0.04	0.031 ± 0.028	0.21 ± 0.05
SingleTop	14.7 ± 1.3	5.8 ± 0.8	2.1 ± 0.5	6.8 ± 0.9
Higgs	0.60 ± 0.06	0.24 ± 0.04	0.100 ± 0.026	0.26 ± 0.04
Zll	0.86 ± 0.29	0.22 ± 0.07	0.24 ± 0.18	0.40 ± 0.20
SM total	603 ± 28	271 ± 19	260 ± 19	72 ± 9
$m(\tilde{\chi}_1^\pm, \tilde{\chi}_1^0)$ =(100,70)	307.3 ± 2.4	82.9 ± 1.2	129.1 ± 1.5	95.3 ± 1.3

C1C1 ISR signal region definition(LH)

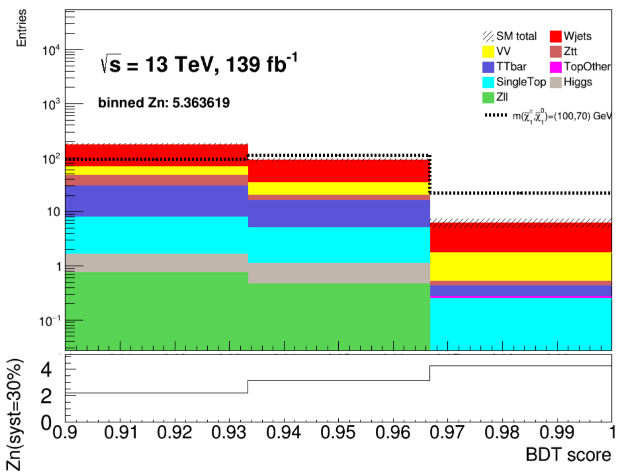
SR
Pre-selection
BDT score > 0.9

Sensitive map





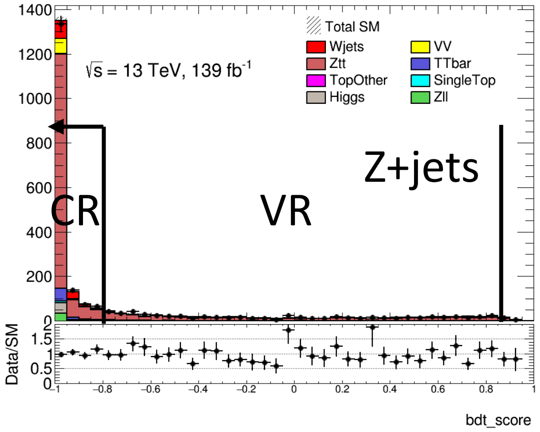
Binned significance: 5.36

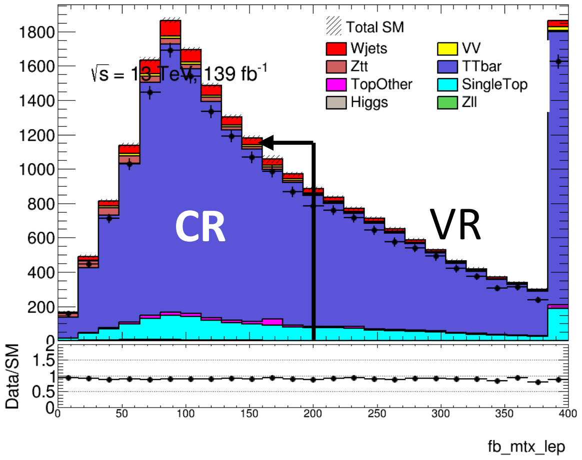


SM process	SR1	SR1-Bin1	SR1-Bin2	SR1-Bin3
Wjets	167 ± 8	56 ± 4	4.4 ± 1.1	106 ± 7
VV	35.90 ± 0.95	13.81 ± 0.54	1.25 ± 0.17	20.8 ± 0.8
Ztt	22.4 ± 2.6	4.5 ± 0.7	0.10 ± 0.04	17.8 ± 2.6
TTbar	33.2 ± 2.3	11.0 ± 1.3	0.15 ± 0.15	22.0 ± 1.8
TopOther	0.27 ± 0.05	0.061 ± 0.025	0.022 ± 0.019	0.18 ± 0.04
SingleTop	10.5 ± 1.0	4.0 ± 0.7	0.24 ± 0.11	6.3 ± 0.8
Higgs	1.59 ± 0.29	0.66 ± 0.19	0.008 ± 0.005	0.92 ± 0.22
Zll	1.22 ± 0.20	0.46 ± 0.13	0.007 ± 0.005	0.75 ± 0.16
SM total	272 ± 9	175 ± 7	91 ± 5	6.2 ± 1.2
$m(\tilde{\chi}_1^\pm, \tilde{\chi}_1^0)$ =(100,70)	226.0 ± 2.0	92.9 ± 1.3	110.6 ± 1.4	22.5 ± 0.6

C1C1 ISR Z+jets estimation(LH)

Region	Selections		Total Bkg	Dominant Bkg	Purity	Data	Data/Bkg
Top CR	**pre-selection(b-veto->bTag)**	$M_T(lep, MET) > 200$	8367+-34	8043+-34	0.96	7404	0.89
Z CR	**Pre-selection** $40 < M_{\tau\tau}^{reco} < 100$	**BDT score < -0.8**	1615+-16	1272+-15	0.79	1611	0.99
Top VR	**pre-selection(b-veto->bTag)**	$120 < M_T(lep, MET) < 200$	5674+-35	5345+-34	0.94	5141	0.90
Z VR	**Pre-selection** $40 < M_{\tau\tau}^{reco} < 100$	**-0.8 < BDT score < 0.9**	623+-12	472+-10	0.76	619	0.99





C1N2 ISR signal region optimization

- BDT method for signal region optimization
  - 30% flat systemic uncertainty
  - Figure of merit: Zn
  - 5fold to Train(Cross-Validation)
- Preselection based on final state($\tau\ \tau$ or $\tau\ l$)

HH Pre-selection	LH Pre-selection
>= 2 medium taus	>= 1 medium taus
0 base lepton	>=1 base lepton, >=1 signal lepton
MET ≥ 200; pass MET trigger	MET ≥ 200; pass MET trigger
1≤nJet	1≤nJet
Opposite-sign hadronic-hadronic tau pair	Opposite-sign lepton-hadronic tau pair
bveto	bveto
jet pt>100 GeV	jet pt>100 GeV
$\mathrm{M}_{\tau\tau}^{\mathrm{reco}} < 40$ GeV or $\mathrm{M}_{\tau\tau}^{\mathrm{reco}} > 130$ GeV	$\mathrm{M}_{\tau\tau}^{\mathrm{reco}} < 40$ GeV or $\mathrm{M}_{\tau\tau}^{\mathrm{reco}} > 130$ GeV

Grid Search for the best model

Hyperparameter	Scan Range
NTrees	200, 300, 400
MaxDepth	4, 6, 8
MinNode	1, 3, 5
Learning rate	0.01, 0.03, 0.05, 0.08, 0.1

Penalty function to balance the AUC and overfit

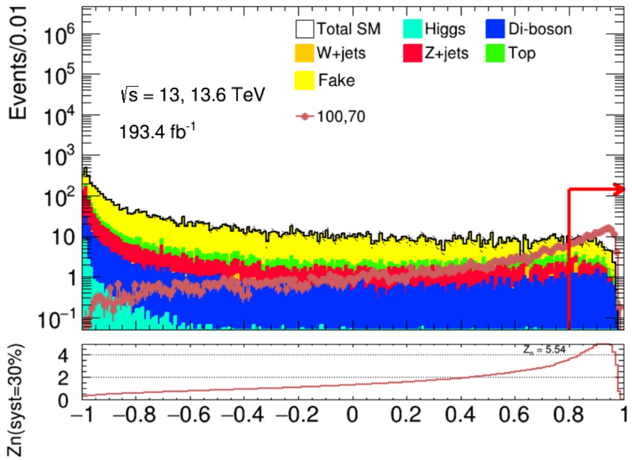
$$\mathcal{F} = AUC_{validation} - 0.3 \times AUC_{gap}$$

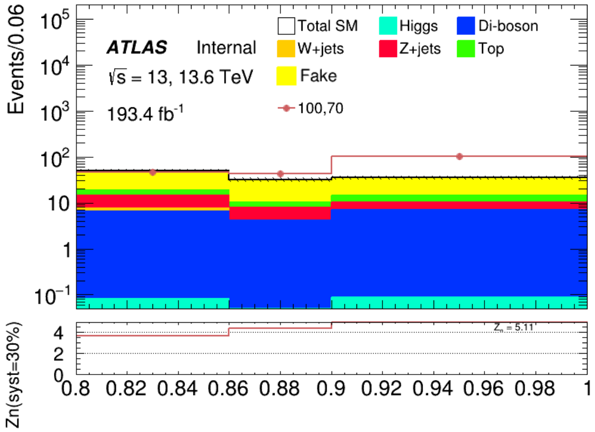
$$AUC_{gap} = |AUC_{train} - AUC_{validation}|$$

C1N2 ISR signal region definition(HH)

SR definition using BDT score(HH)

- Fake tau estimation using fake factor method
- Three bins
  [0.80, 0.86], [0.86, 0.90], [0.90, 1.00]
- Fake estimation is same with direct stau
  only change signal lep to at least signal lep





HH Pre-selection
>= 2 medium taus
0 base lepton
MET ≥ 150; pass MET trigger
1≤nJet
Opposite-sign hadronic-hadronic tau pair
bveto
jet pt>100 GeV
$\mathbf{M_{\tau\tau}^{reco} < 40}$ **GeV or** $\mathbf{M_{\tau\tau}^{reco} > 130}$ **GeV**

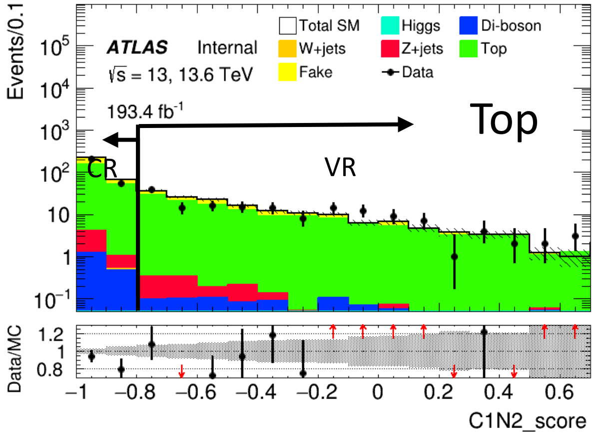
SR
Pre-selection
BDT score > 0.8

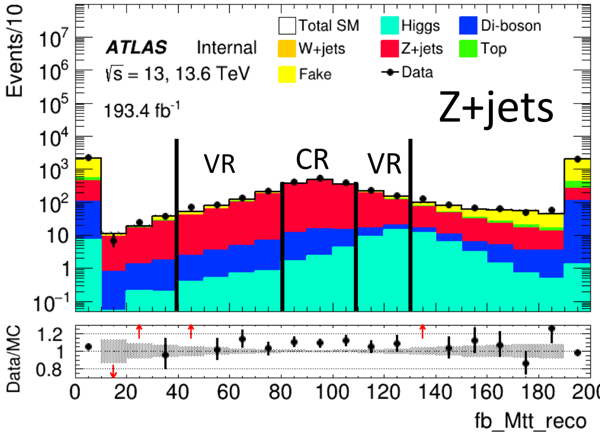
SM Process	[0.80, 0.86]	[0.86, 0.90]	[0.90, 1.00]
VV	6.508 ± 0.368	4.124 ± 0.341	6.855 ± 0.398
Top	4.719 ± 0.686	2.663 ± 0.546	4.331 ± 0.731
Fake	31.456 ± 3.898	21.525 ± 3.292	20.438 ± 2.976
Higgs	0.080 ± 0.023	0.045 ± 0.019	0.086 ± 0.031
Z+jets	6.659 ± 0.416	3.564 ± 0.286	3.207 ± 0.250
W+jets	0.978 ± 0.978	0.000 ± 0.000	0.000 ± 0.000
Total Bkg	50.402 ± 4.042	31.921 ± 3.354	34.917 ± 3.118
$m(\tilde{\chi}_1^{\pm}, \tilde{\chi}_2^0)$=(100,70)	46.569 ± 1.327	43.128 ± 1.283	101.059 ± 1.961
Z_n	3.65	4.37	5.11

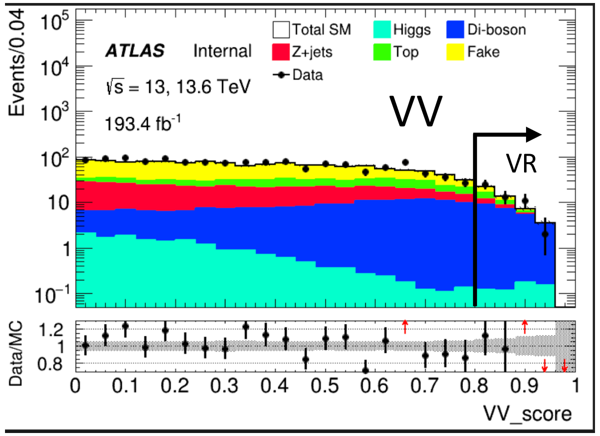
C1N2 ISR background estimation(HH)

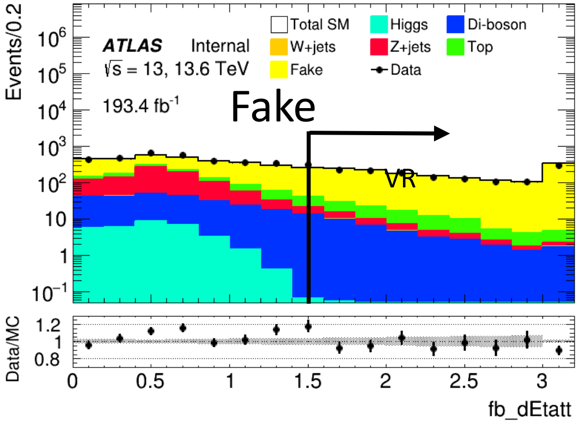
Selection for control region and validation region

Process	Top		Z+jets		Multi-bosons	Fake
	TCR	TVR	ZCR	ZVR	MBVR	FakeVR
Charge combination				OS		
Trigger				MET trigger, $E_T^{miss} \geq 200 GeV$		
N medium τ				≥ 2		
N lep				=0		
$nBaseJet$				≥ 1		
$Jet \quad p_T$[GeV]				≥ 100		
N b-jets	≥ 1			= 0		
$m(\tau_1, \tau_2)$ [GeV]	≤ 40 or ≥ 130		[80,110]	[40,80] or [110,130]	≤ 40 or ≥ 130	
$M_T(\tau_1, E_T^{\text{miss}})$	≥ 200		–	–	–	–
$d\eta(\tau_1, \tau_2)$	–	–	–	–	–	≥ 1.5
C1N2 score	[-1, -0.8]	[-0.8, 0.7]		≤ 0.7		
VV score	–	–	–	–	≥ 0.80	–
Total bkg	290+-7	162+-5	1420+-10	1286+-11	46+-2	2000+-28
Dominant Bkg	206+-4	131+-3	1221+-5	993+-4	24+-1	1851+-28
Purity	0.71	0.80	0.86	0.77	0.52	0.93
Data	264	161	1559	1427	51	1950
Data/MC	0.91	0.99	1.09	1.10	1.10	0.98







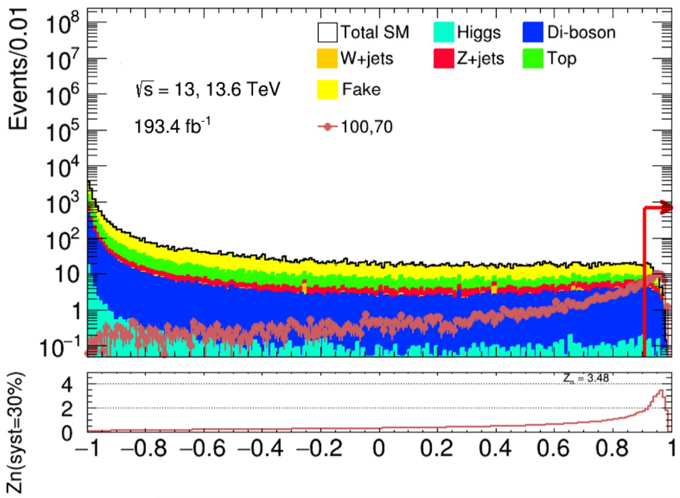


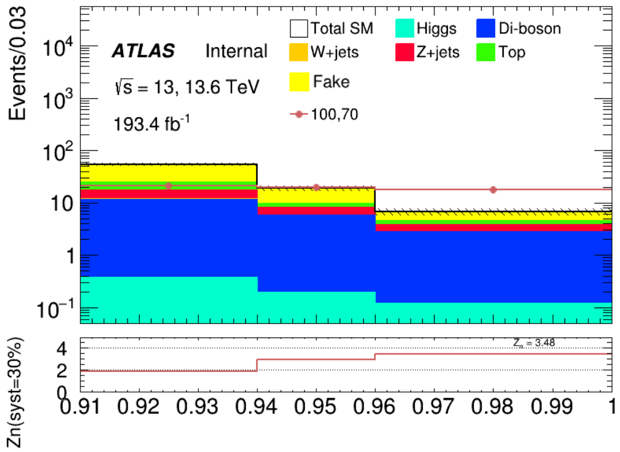
C1N2 ISR signal region definition(LH)

SR definition using BDT score(LH)

- Fake tau estimation using fake factor method
- BDT score > 0.9
- Three bins
  [0.91, 0.94], [0.94, 0.96], [0.96, 1.00]

LH Pre-selection
>= 1 medium taus
>=1 base lepton, >=1 signal lepton
MET ≥ 200; pass MET trigger
1≤nJet
Opposite-sign lepton-hadronic tau pair
bveto
jet pt>100 GeV
$\mathbf{M_{\tau\tau}^{reco} < 40\ GeV\ or\ M_{\tau\tau}^{reco} > 130\ GeV}$
SR
Pre-selection
BDT score > 0.91



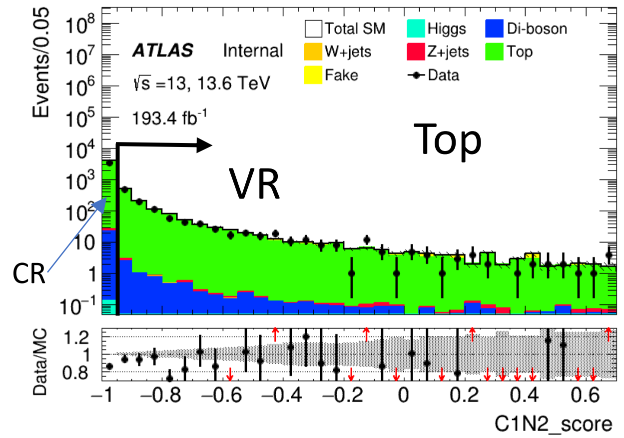


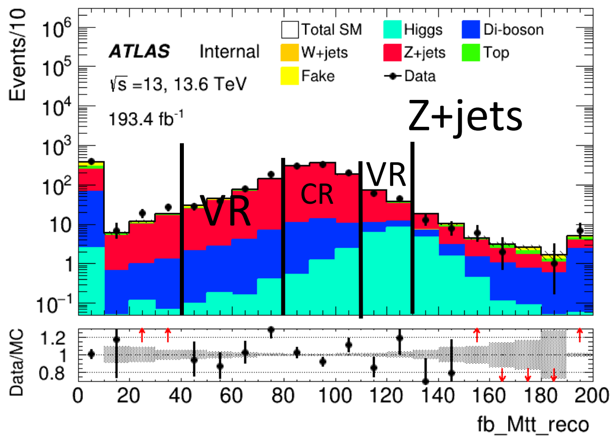
Process	[0.91, 0.94]	[0.94, 0.96]	[0.96, 1.00]
VV	11.105 ± 0.565	5.550 ± 0.367	2.666 ± 0.263
Top	7.432 ± 0.956	1.524 ± 0.441	0.578 ± 0.237
Fake	28.625 ± 3.295	9.404 ± 1.929	2.400 ± 1.009
Higgs	0.376 ± 0.057	0.189 ± 0.052	0.118 ± 0.040
Z+jets	5.515 ± 0.357	2.251 ± 0.222	1.026 ± 0.118
W+jets	0.173 ± 0.142	0.000 ± 0.000	0.000 ± 0.000
Total Bkg	53.226 ± 3.497	18.918 ± 2.018	6.788 ± 1.067
C1N2 (100,70)	21.447 ± 0.901	19.678 ± 0.874	17.543 ± 0.820
Z_n	1.90	2.94	3.47

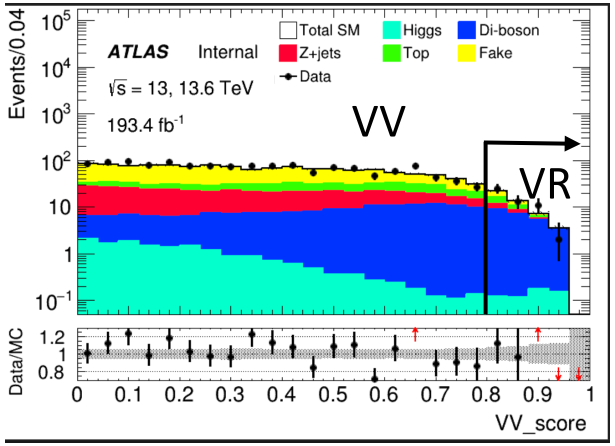
C1N2 ISR background estimation(LH)

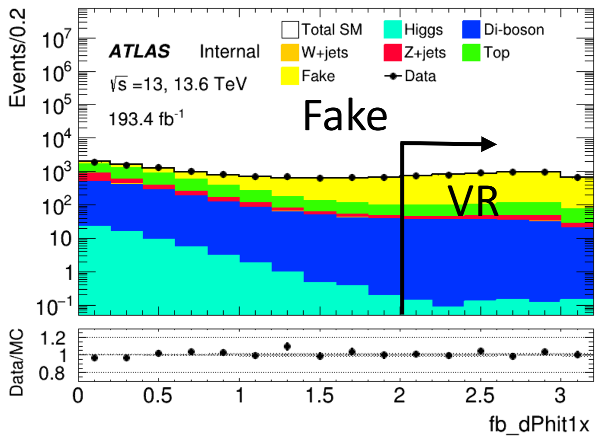
Selection for control region and validation region

Process	Top		Z+jets		Multi-bosons	Fake
	TCR	TVR	ZCR	ZVR	MBVR	FakeVR
Charge combination	OS					
Trigger	MET trigger, $E_T^{miss} \geq 200GeV$					
N medium τ	≥ 1					
N lep	≥ 1					
$nBaseJet$	≥ 1					
$Jet \quad p_T$[GeV]	≥ 100					
N b-jets	≥ 1		$= 0$			
$m(\tau_1, l)$ [GeV]	≤ 40 or ≥ 130		[80,110]	[40,80] or [110,130]	≤ 40 or ≥ 130	
$M_{inv}(l, MET)$	≥ 300		–	–	–	–
$\Delta\phi(MET, \tau_1)$	–	–	–	–	–	≥ 2
$dR(\tau, l)$	–	–	≤ 0.6		–	–
C1N2 score	[-1, -0.95]	[-0.95, 0.7]	≤ 0.7			–
VV score			–	–	≥ 0.80	–
Total bkg	4151+-11	1212+-13	917+-6	823+-4	78+-2	4983+-40
Dominant Bkg	3781+-20	1102+-11	823+-4	606+-3	53+-1	4390+-40
Purity	0.91	0.90	0.90	0.84	0.68	0.88
Data	3626	1119	908	745	72	5043
Data/MC	0.87	0.91	0.99	1.03	0.92	1.01





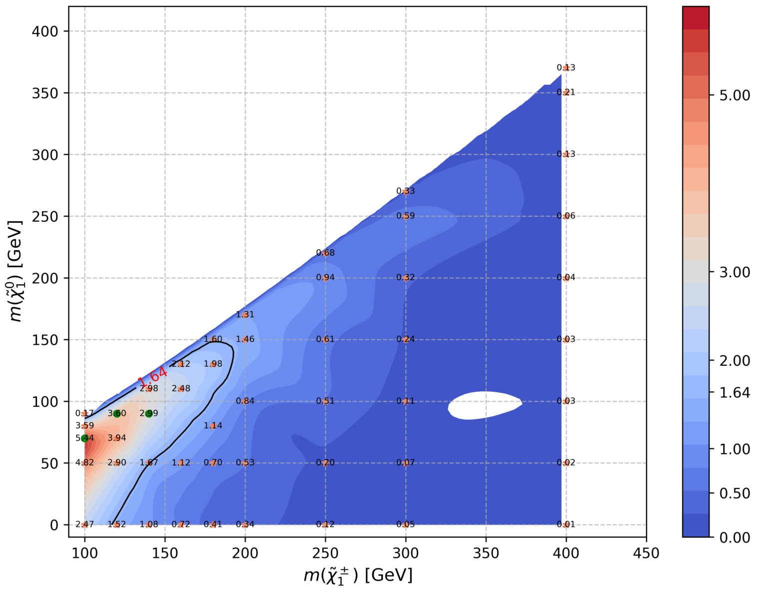




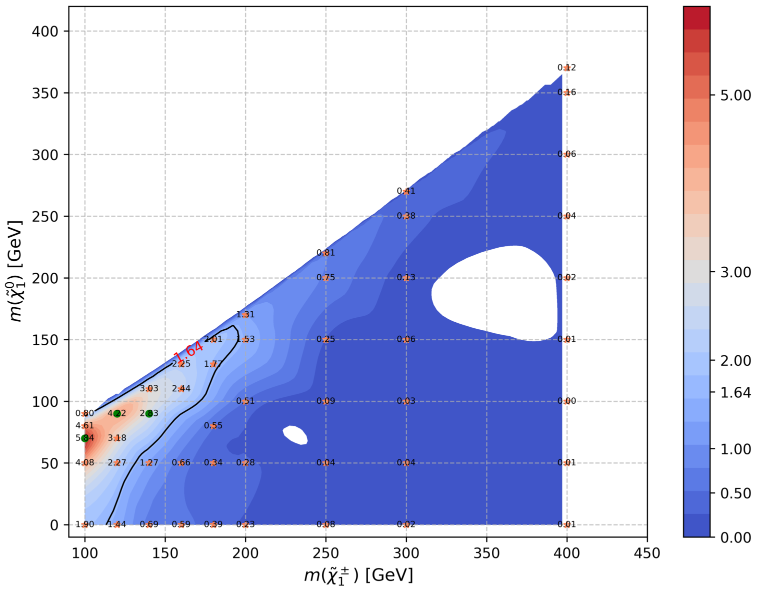
C1N2 ISR signal region definition

- Expected sensitivity
  - 30% flat systematic uncertainty
  - gap caused by interpolation algorithm
  - Result only for full run2 sample, will update to full run3 sample later

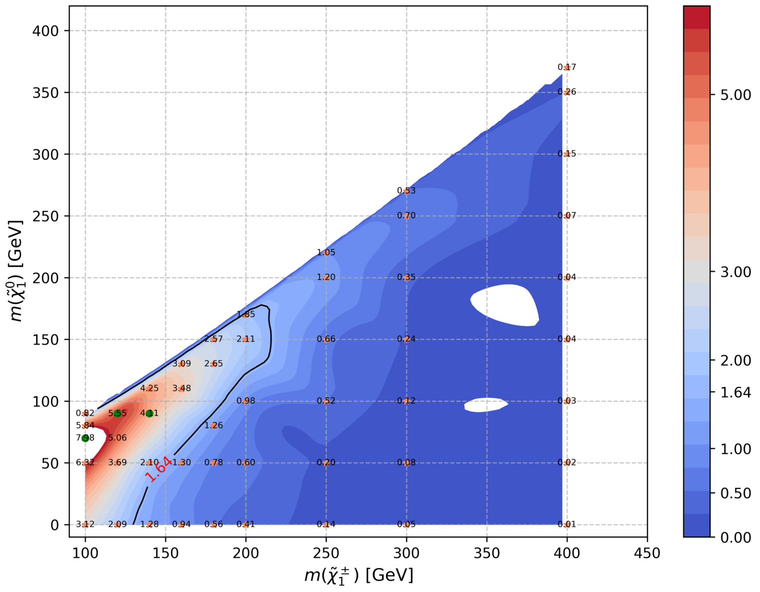
HH channel



LH channel



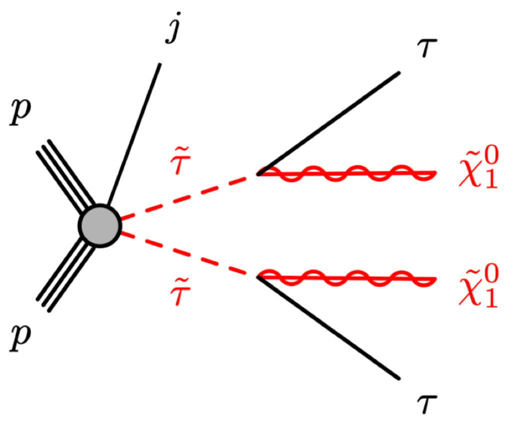
Combined channel



Backup

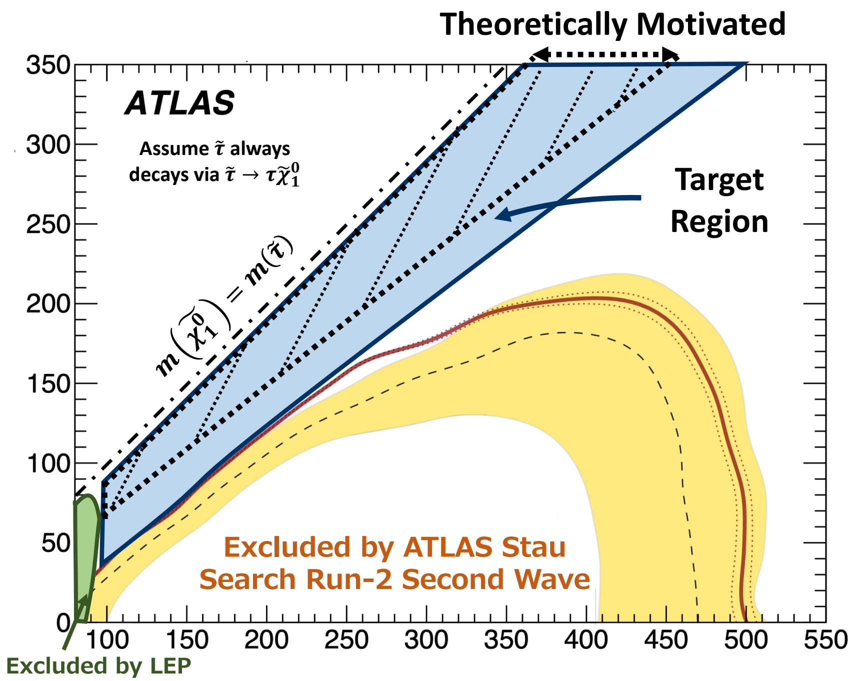
Direct stau ISR motivation

Direct stau with $2\tau + E_T^{miss}$ + ISR
Boosted by ISR and has a large MET due to two missing $\tilde{\chi}_1^0$



Basic Strategy:

- Trigger: MET Trigger
- Search channel:
  - had-had channel
  - lep-had channel
- Follow the NN approach for samples
  - Reference point $(\tilde{\tau}, \tilde{\chi}_0)$ = (100,70)

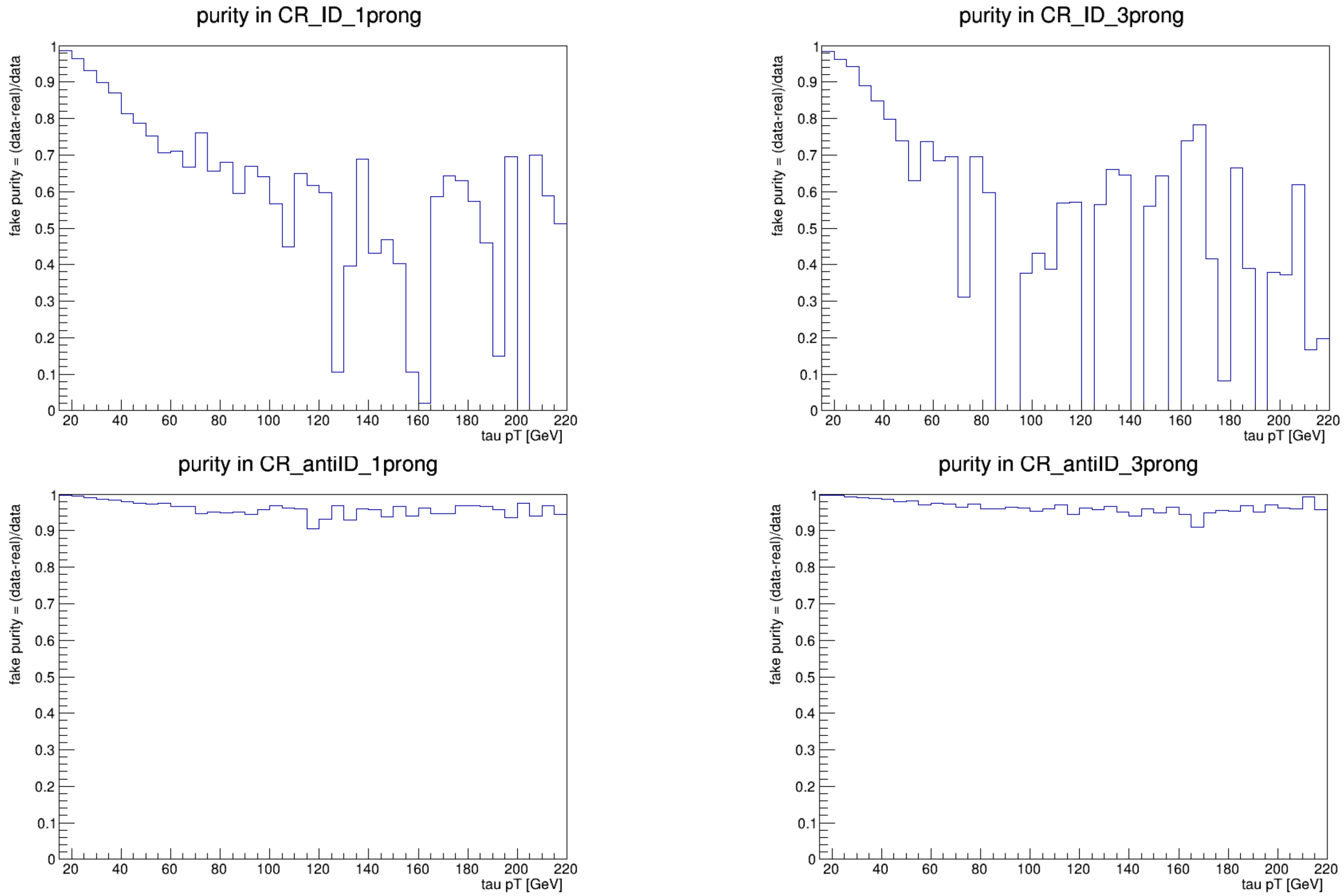


Direct Stau ISR Fake purity in CR

1-prong | **3-prong**

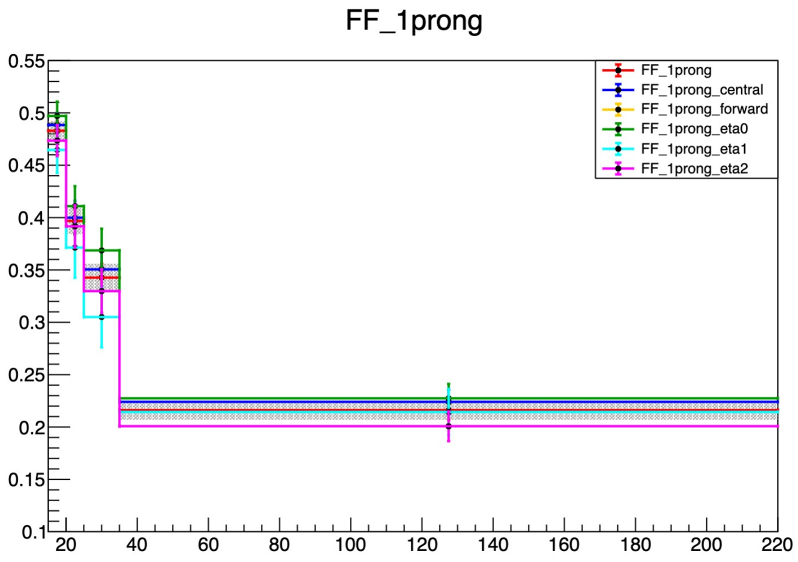
ID

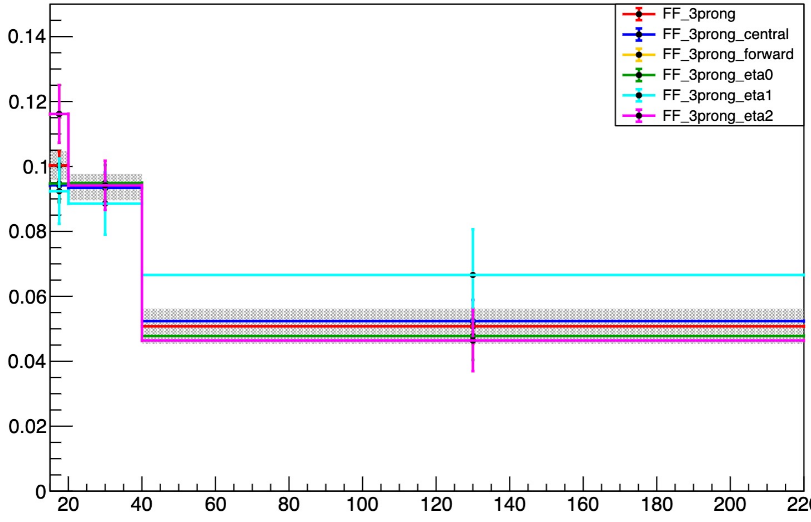
anti-ID



Direct Stau ISR Fake Factor

- **Binned in prongness, tau eta, tau pT**
  - Eta bins
    - 2 bins: central [0,1.37], forward [1.52,2.5]
    - 3 bins: eta0,1,2 for [0,1), [1, 1.37], [1.52,2.5]
- **Auto binning:**
  - > 10% of events in nominator and denominator
  - Add bins to bin i until it is not consistent anymore with bin i - 1
    - Relative stat uncertainty on ratio smaller than 50%
    - >10% events in nominator and denominator
- 1-prong: 3 eta bins
- 3-prong: central, forward regions

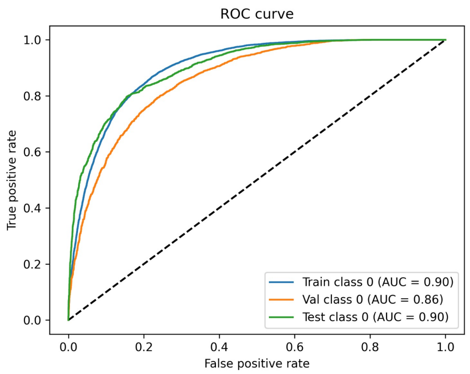


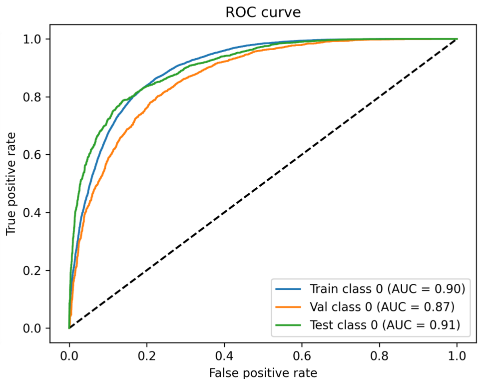


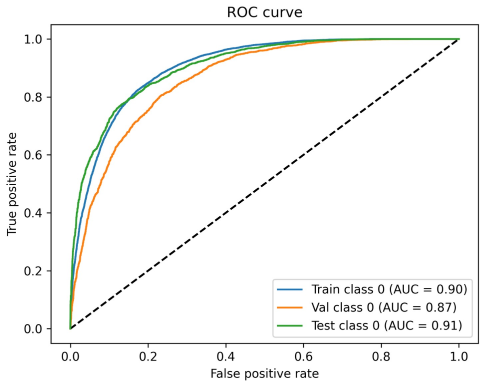
Direct Stau ISR Neuron network(HH)

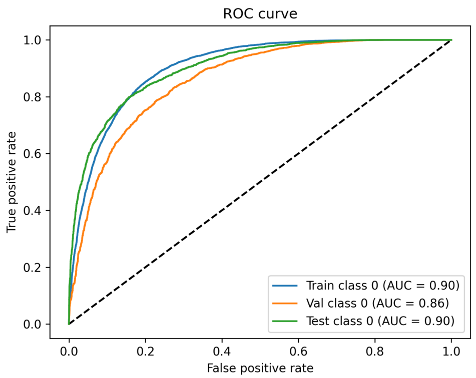
- **Model Architecture**
  - **Neurons:** [256, 128, 128, 64, 32]
  - **Activation Function:** Leaky ReLU
  - **Batch Normalization**
  - **Dropout Rate:** 0.3
  - **L2 Regularization (λ):** 1e-4
- **Training Configuration**
  - **Optimizer:** Adam
    - **Learning Rate:** 1e-3
  - **Batch Size:** 128
  - **Epochs:** 600

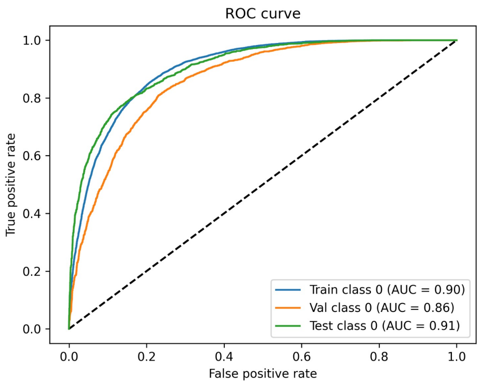
- **Reduce Learning Rate on Plateau:**
  - Monitor: *val_loss*
  - Factor: 0.5
  - Patience: 20
- Class weight to solve imbalance issue
- Normalization before training
  - $\frac{X-\mu}{\sigma}$
- Drop Negative weight







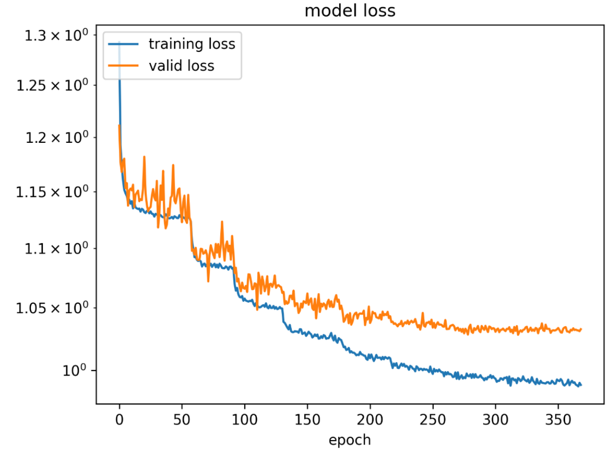


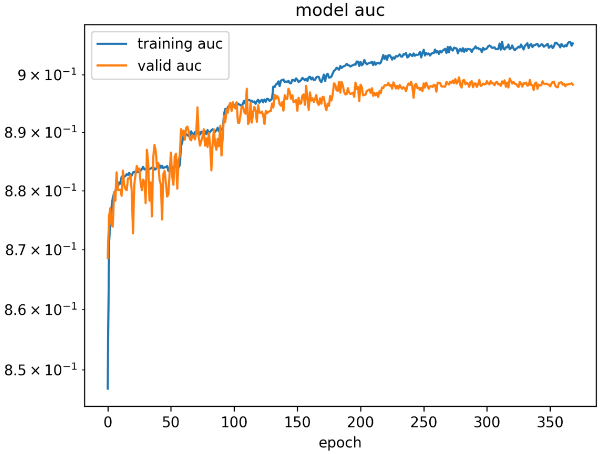


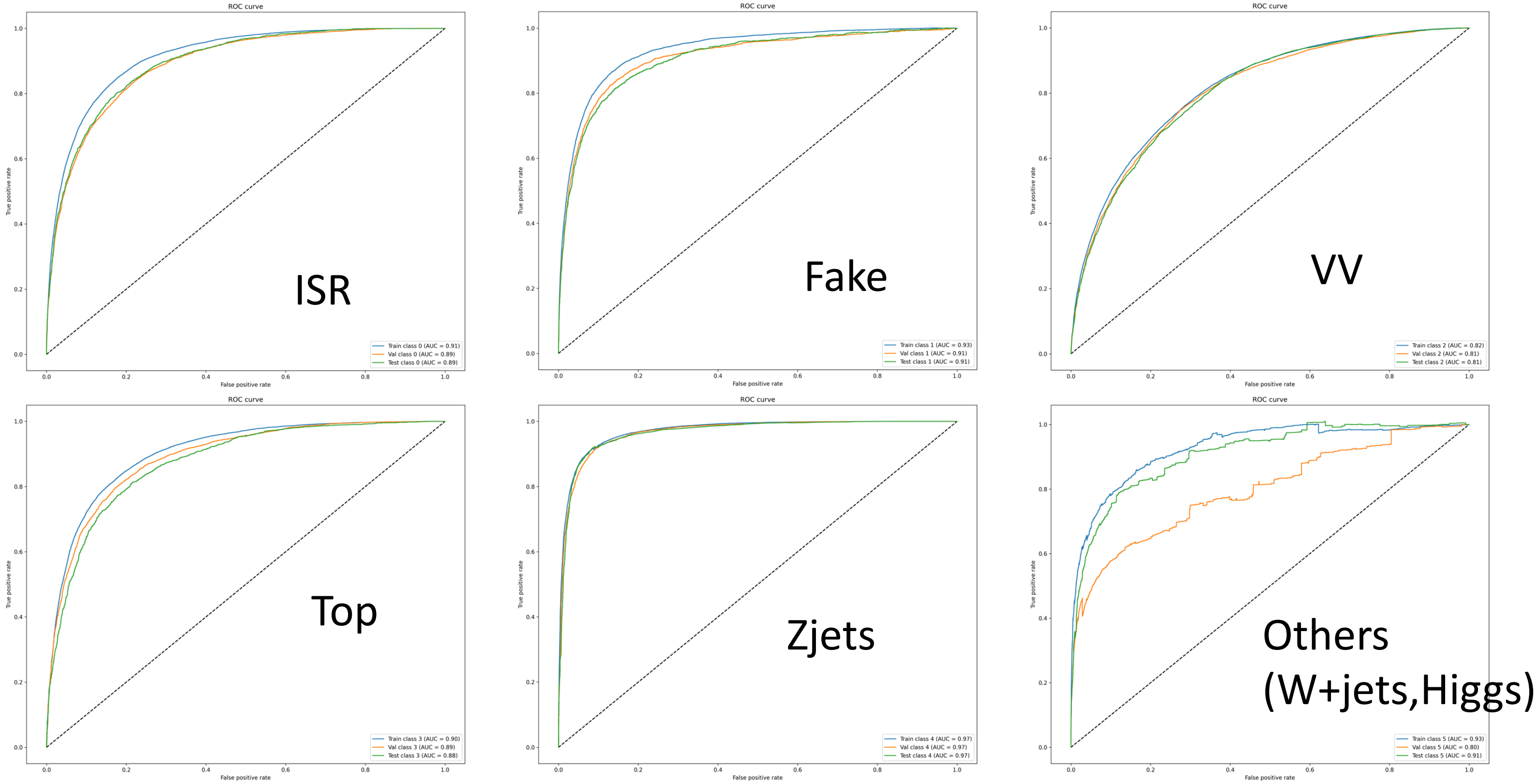
46

Direct Stau ISR Neuron network(LH)

AUC for 6 output scores







Direct Stau ISR Class Weight

- $w_i = I_i \cdot \frac{N_{all}}{n_{class} \times N_i}$
  - N_{all} : The total sum of weight in the entire dataset (across all classes).
  - N_i : The sum of weight belonging to class i.
  - n_{class} : The total number of distinct classes in the dataset.
  - I_i: importance factor

Direct Stau ISR Top, Ztt CR

- Top CR:
  - **b-jet >= 1 (GN2v01, WP 90)**
  - Colinear mass (τ,τ) > 250 GeV

Yield, purity table (Top CR)

MT or MT2 Data/MC Comparison plot (Top CR)

- Ztautau CR
  - b-veto (GN2v01, WP 90)
  - **70 < Collinear mass (τ,τ) < 110 GeV**

Yield, purity table (Ztautau CR)

MT or MT2 Data/MC Comparison plot (Ztautau CR)

Direct Stau ISR Diboson CR

- Wjets CR:
  - 1 Medium taus and 1 Tight light lepton with pT > 20 GeV
  - OS for tau and light lepton
  - M_T(lep1, MET) < 80 GeV

Yield, purity table Wjets CR

MT or MT2 Data/MC Comparison plot Wjets CR

- Diboson CR:
  - Tight SFOS light lepton with pT > 20 GeV
  - M_T(lep1, MET) < 80 GeV

Yield, purity table VV CR

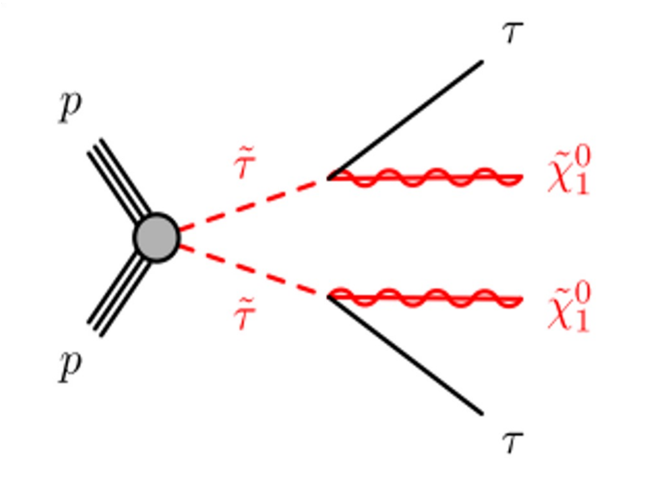
MT or MT2 Data/MC Comparison plot VV CR

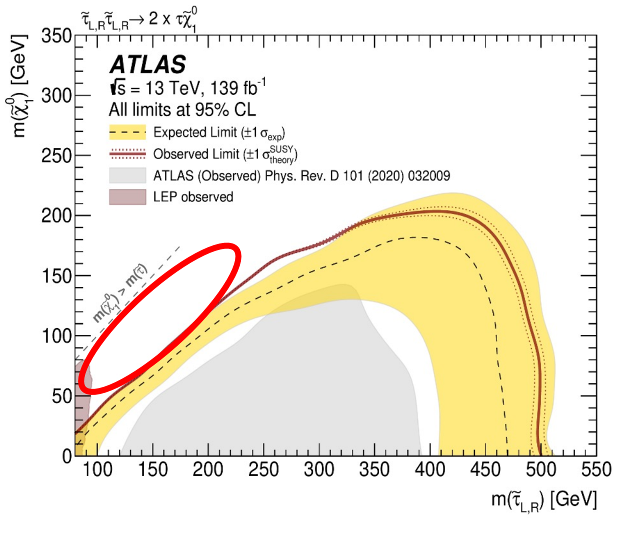
Direct Stau ISR VRs

- VR1 (VV):
  - Two SFOS leptons with pT > 20 GeV
- VR2 (LH):
  - MET > 130 GeV
  - nJet <=2
  - Primary Jet pT > 50 GeV
  - 1 Medium tau with pT > 25 GeV
  - 1 Tight light lepton
  - OS for tau and light lepton
  - $M_{T2}(40) > 55$ GeV
- VR3 (HH):
  - 2 Medium OS taus with pT > 20 GeV
  - $M_{T2}(70) > 85$ GeV

Yield, purity table (VR1)	**MT or MT2 Data/MC Comparison plot (VR1)**
Yield, purity table (VR2)	**MT or MT2 Data/MC Comparison plot (VR2)**
Yield, purity table (VR3)	**MT or MT2 Data/MC Comparison plot (VR3)**

Direct Stau single-lep trigger motivation





Sensitivity to direct stau production is difficult
- Low cross-sections
- Hadronically-decaying taus are difficult objects

Possible final states		… and problems
$\tau_{had}\ \tau_{had}$	BR = 42%	trigger on 2x high pT taus, large Δm
$\tau_{had}\ \tau_{lep}$	BR = 46%	large W+jets background
$\tau_{lep}\ \tau_{lep}$	BR = 12%	leptons are soft (Ben H study)

$\tau_{had}\ \tau_{had}$ has been the main focus for ATLAS so far, with potential improvements at low mass from Dom's ML efforts.

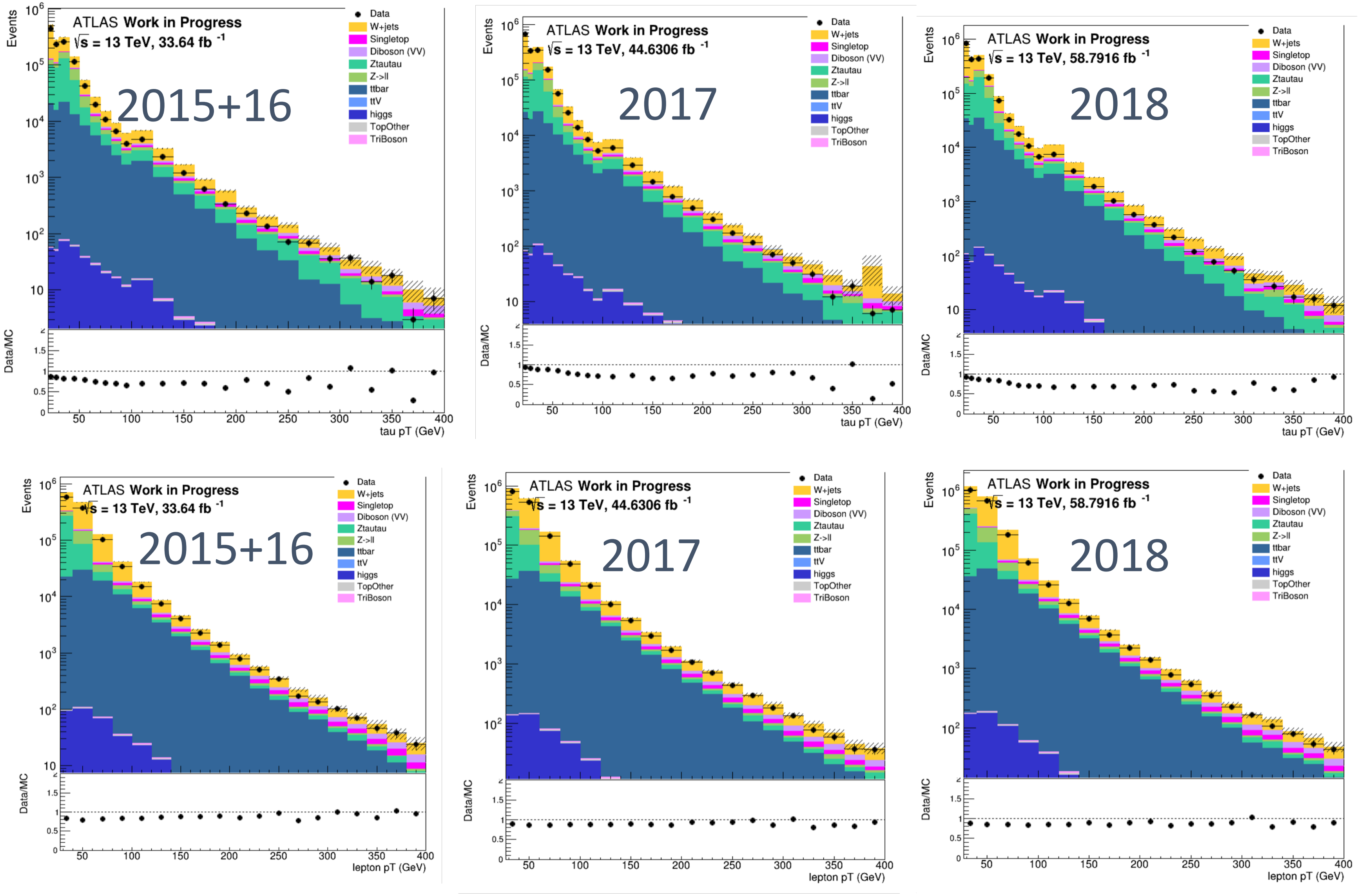
2-tau triggering 95,75 or 75,40+MET 150 kills moderate Δm signals.

Huge uncovered region from moderate Δm → low Δm.

Target with the $\tau_{had}\ \tau_{lep}$ channel and use single lep triggers

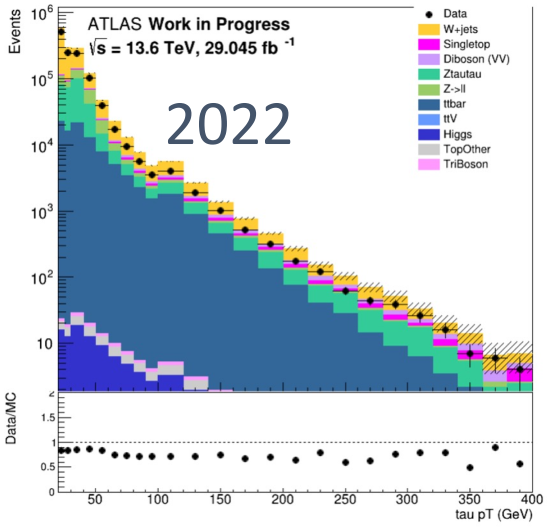
Direct Stau single-lep trigger MC modeling check

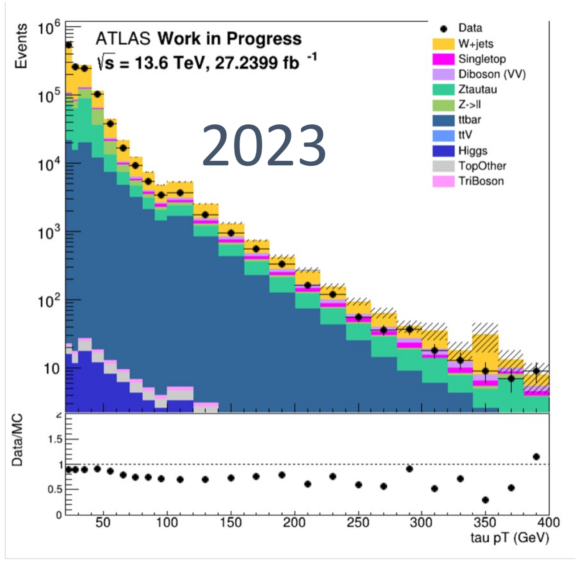
- Requirements: 1e/mu passing trigger, at least 1 tight signal tau
- MC overestimates data

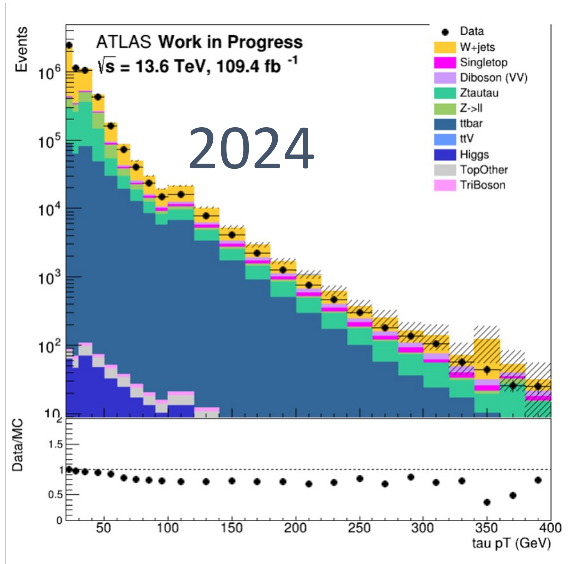


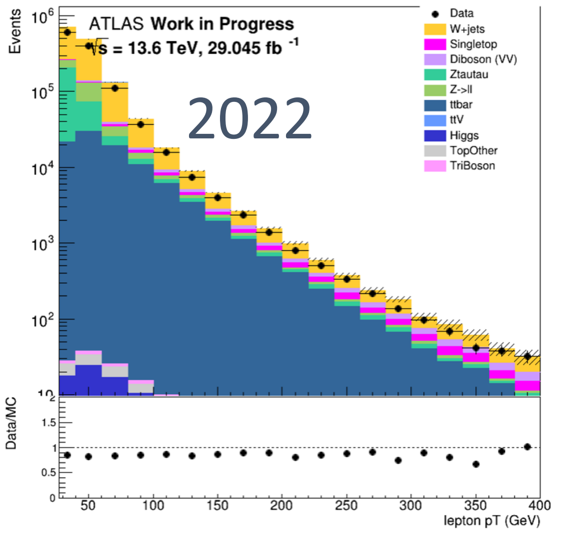
Direct Stau single-lep trigger MC modeling check

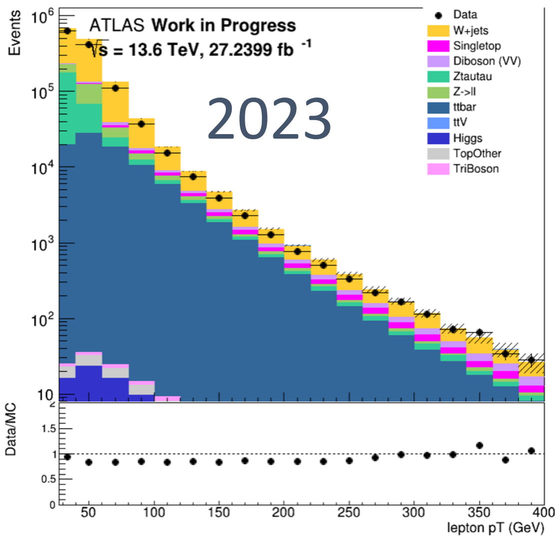
- Requirements: 1e/mu passing trigger, at least 1 tight signal tau
- MC overestimates data

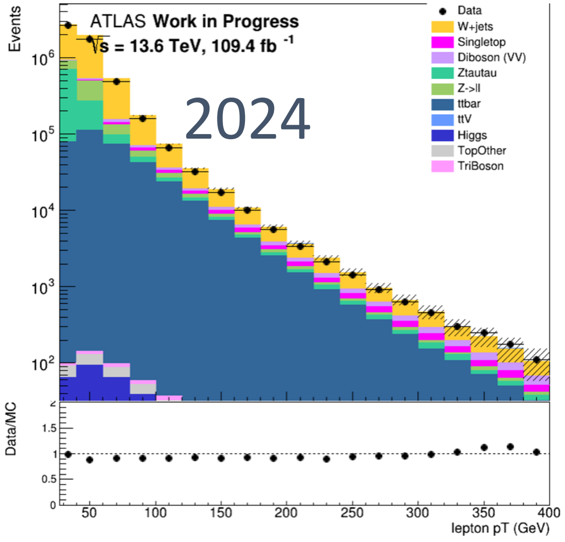












Direct Stau single-lep trigger Pre-selection(LH)

LH Pre-selection
== 1 Tight taus (OS)
1 signal lepton
bveto
Single-lepton trigger
15 GeV < MET < 200 GeV
METSig > 2
MT2 < 90
$dR(lep, tau) < 3.6$
custom MT cut*
$\boldsymbol{\phi_{\{1,2\}}\ cut^*}$

*custom MT cut (removes red shape):

- Targets Ztt

(mT_sum > 70_000) &
((mT_lep_met > 20_000) | (mT_tau_met > 90_000)) &
((mT_lep_met > 90_000) | (mT_tau_met > 20_000))

*phi_{1,2} cut (removes red circle slice):

- Targets fakes

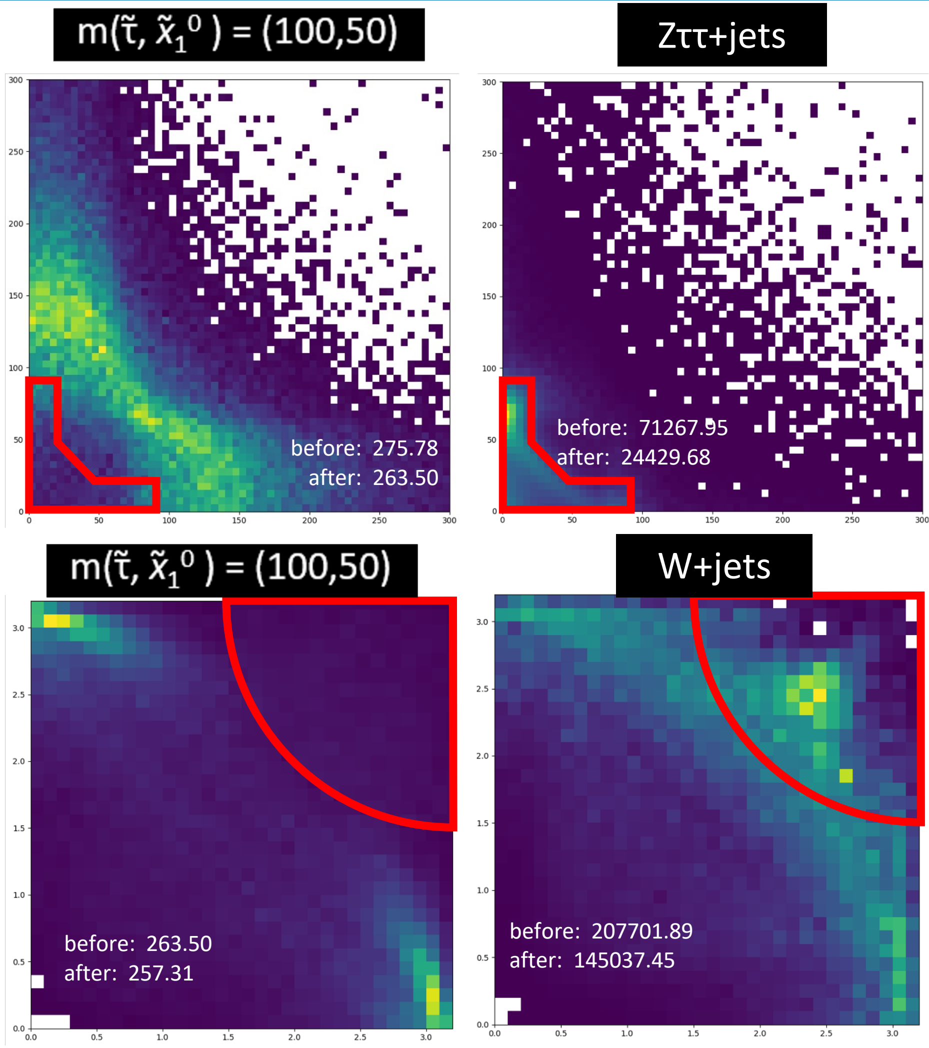
Define:

$\phi_{max} = max(\phi_{lep}, \phi_\tau)$ and $\phi_{min} = min(\phi_{lep}, \phi_\tau)$

$\phi_1 = \phi_{max} - \phi_{MET}$ and $\phi_2 = \phi_{MET} - \phi_{min}$

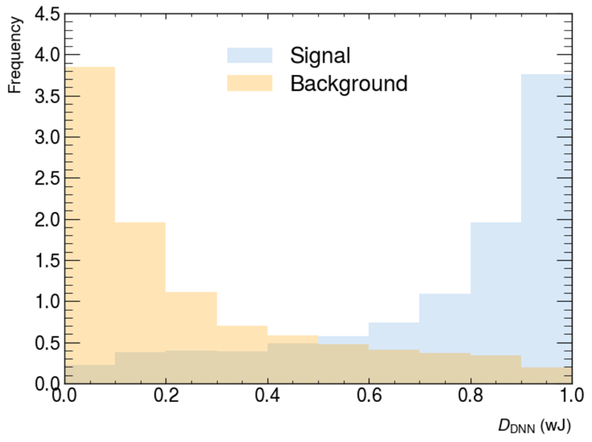
(abs(phi1)-a)**2 + (abs(phi2)-b)**2 >= r**2

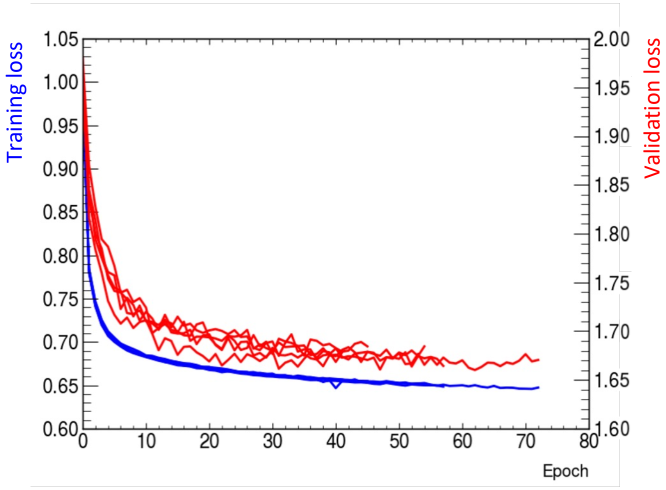
where a=pi, b=pi, and r = pi/2



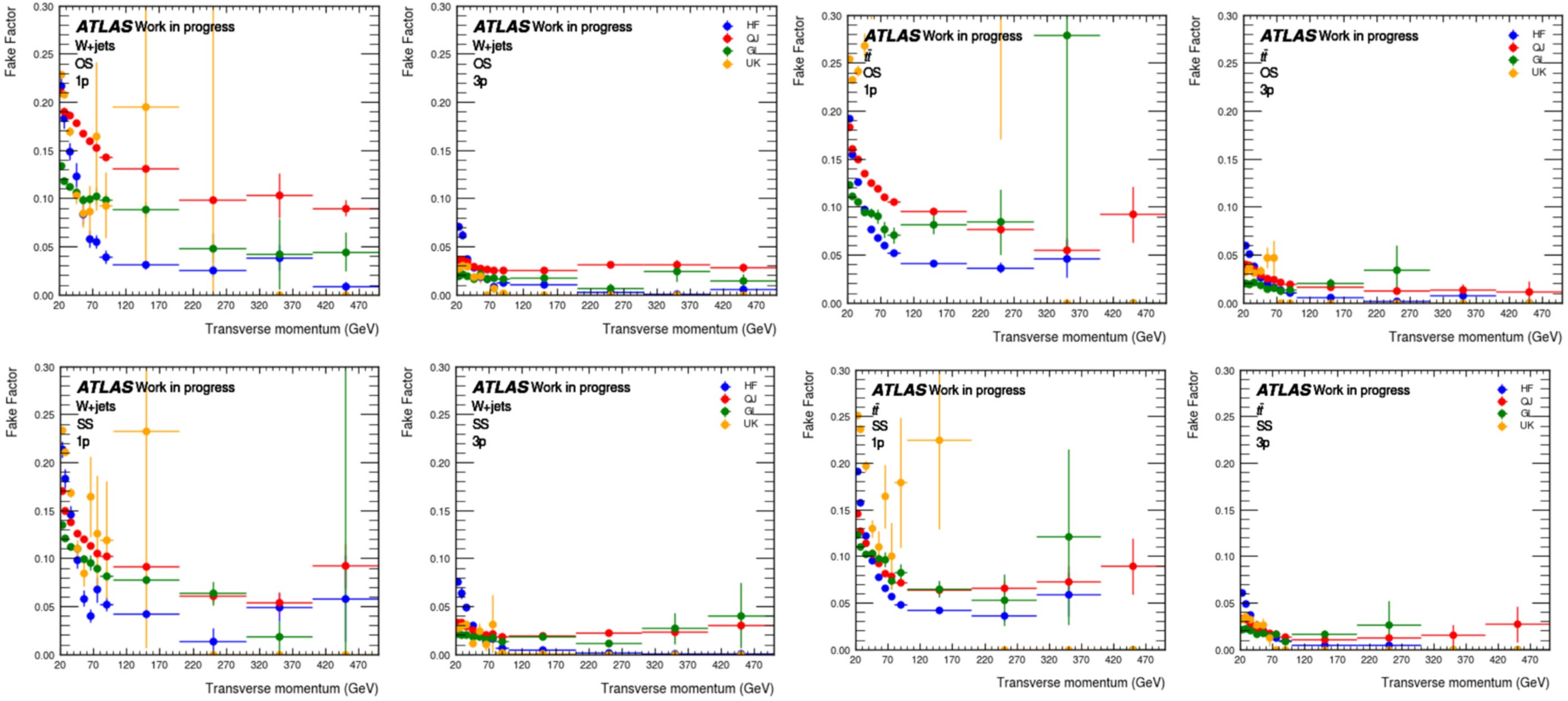
Direct Stau single-lep trigger DNN Training

- Train deep neural network (DNN) to separate signal from background and cut on network score targeting (100, 50) & (150,100) mass points (high stats points with 1M events generated) and trained 0J and >=1J channels separately.
  - Training converges well (across all training folds) and very good signal/background separation is achieved
- Use multiclass DNN with 6 output classes: Signal, fake-tau background, 4 real-backgrounds (top, Ztt, VV, other)
- Input variables:
  - Basic object variables: p_T, η, MET, charge etc.
  - High-level variables: ΔR, $\Delta\eta$, $\Delta\phi$, m_{T2}, M_{eff}, M_{inv}, Σm_T, m_{CT}, balance + more!

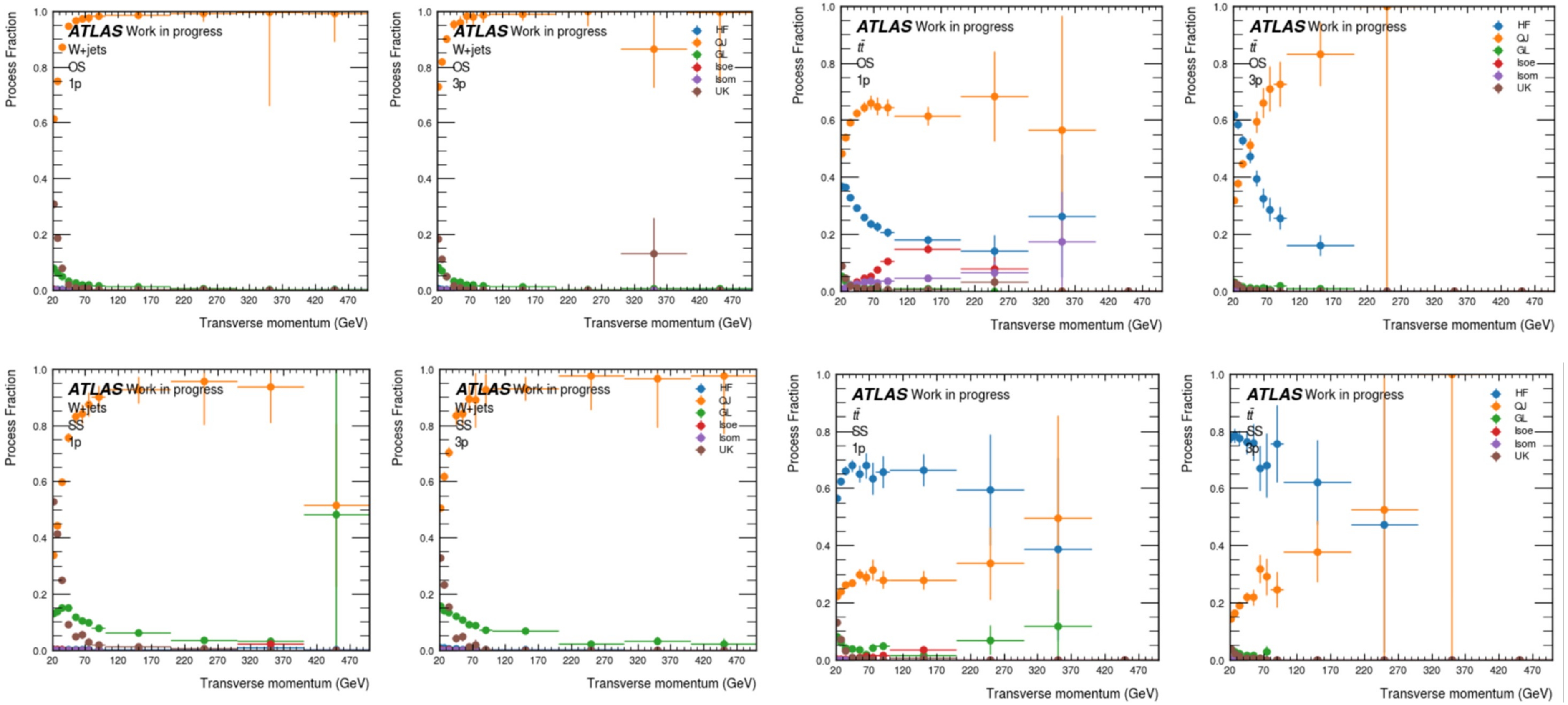




Direct Stau single-lep trigger Fake Factors



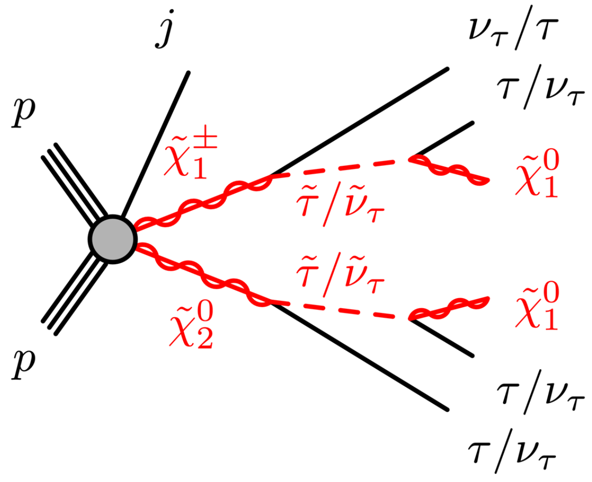
Direct Stau single-lep trigger Fake Factors

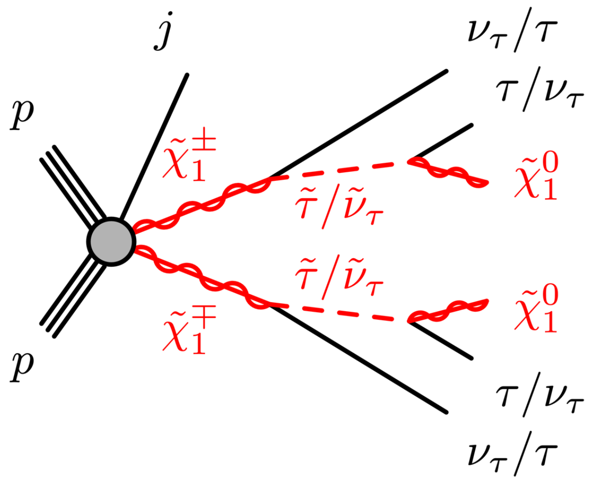


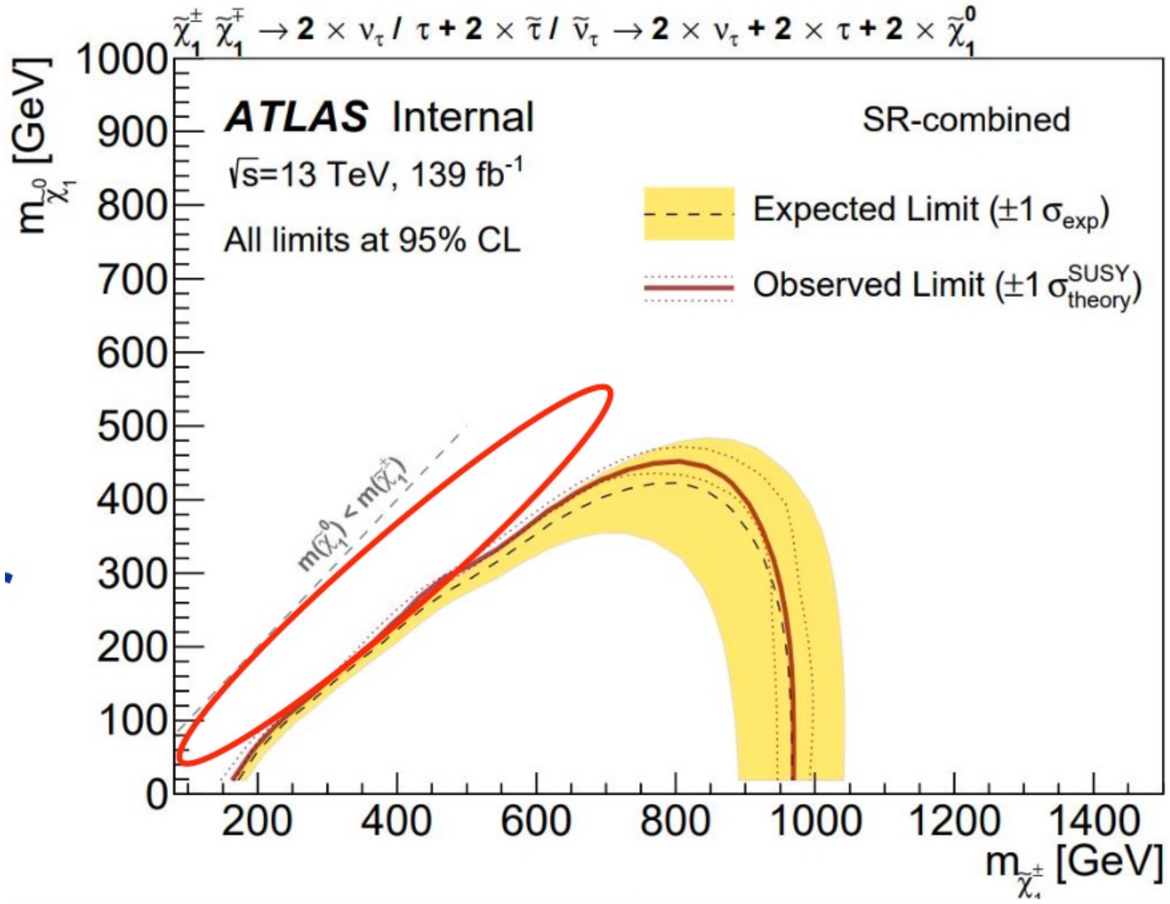
C1C1/C1N2 Motivation

C1C1, C1N2 via stau with $\geq 2\tau + E_T^{miss} + ISR$

Boosted by ISR and has a large MET due to two miss $\tilde{\chi}_1^0$







C1C1 ISR signal region optimization(HH)

- Input:
  - data: MC events passing pre-selection
    - signal: 39382 events C1C1 + ISR, M(C1, N1) = (100, 70) GeV
    - Bkg: 1089081 events.

  feature:Pzetaj1,Pzetae,MCTtt,m_jet,dPhiISRV,dPhiVI,nS_tau,PtV,dPhit2x,maxdPhit1j,maxdPhit2j,METOPttau2,METOHTtt,METOHTjet,Ptjet1OPttau1,MT2tt_110,mtx_jet,mtx_lep,mt_lep,mt_C1C1,MTtau2met,MTtaumin,MTsum,Mtsum,Mtsumj,Mll,MT2,MET,METsig

- Strategy
  - TMVA.Types.kBDT
  - Cross-Validation(5-fold)
  - hyper paramter: NTrees=400,learning rate=0.1,max depth=8, MinNodeSize=1%

  other set-up: nCuts=20:BoostType=Grad:UseBaggedBoost=true:BaggedSampleFraction=0.5:NegWeightTreatment=Pray

C1C1 ISR signal region optimization(LH)

- Input:
  - data: MC events passing pre-selection
    - signal: ISRC1C1(with C1 mass=100GeV, N1 mass=70 GeV)(39336 events)
    - bkg(1559557 events).

  feature: dRt1x,dRtt,dPhitt,e_lep,mt_lep,ht_tau,METsig,nBaseJet,mindPhije,maxdPhit2j,MTtau1met,MTtau2met,MTtaumin, MTtot,MTsum,Minvtt,Minvt2j,MT2tt_110,MCTtt,Pzetaj1,Pzetatt,METOPttau1,METOPttau2,METOHTtt,METOHTjet,Ptjet1OPttau1,Ptjet1OPttau2,Ptjet1OHTtt,PtV
- Strategy
  - TMVA.Types.kBDT
  - 5-fold train
  - hyper paramter: NTrees=300,learning rate=0.05,max depth=6, MinNodeSize=1%

  other set-up: nCuts=20:BoostType=Grad:UseBaggedBoost=true:BaggedSampleFraction=0.5:NegWeightTreatment=Pray

C1N2 ISR Binary class(HH)

Hyperparameters: Ntrees = 200, MaxDepth = 6, MinNodeSize = 2%, Learning rate = 0.03(initial setting)

Feature engineering:

Select a simple model and put all features into model, choose Top 30 vars based on importance list, drop high correlated vars

Final feature list:

Rank	Variable	Variable Importance
1	fb_dEtatt	5.153e-02
2	fb_dRtt	4.318e-02
3	fb_dRMax_xt	4.248e-02
4	fb_dPhitt	4.228e-02
5	fb_MIA	4.205e-02
6	fb_METsig	3.979e-02
7	fb_dPhizxe	3.972e-02
8	fb_dPhiztt	3.942e-02
9	fb_frac_MET_tau1	3.735e-02
10	fb_dPhiMin_xt	3.513e-02
11	fb_dPhiMin_tj1	3.512e-02
12	fb_MT2_150	3.494e-02
13	fb_frac_MET_MeffInc_40	3.474e-02
14	fb_dRMin_tj	3.467e-02
15	fb_eta_tau2	3.454e-02
16	fb_frac_MET_tt	3.452e-02
17	fb_frac_MET_Meff	3.408e-02
18	fb_dPhit2x	3.277e-02
19	fb_dPhiMax_xt	3.207e-02
20	fb_dRt2x	3.131e-02
21	fb_dPhit1x	3.089e-02
22	fb_frac_MET_tau2	3.085e-02
23	fb_Mll	2.960e-02
24	fb_MET_Jet	2.734e-02
25	fb_sum_cos_dphi	2.530e-02
26	fb_pt_Vframe	2.272e-02
27	fb_Pt_tt	1.912e-02
28	fb_MstauA	1.881e-02
29	fb_Proj_t1	1.594e-02
30	fb_Proj_tt	1.427e-02
31	fb_MCT	1.345e-02

Weight choose: no weight, abs(weight)

No weight have better performance
but abs(weight) fit our anaysis requirment

Split strategy: Separate entries by using mod 5, for Fake bkg, if separate follow sequence, all weighted entry will split into first fold

C1N2 ISR Binary class(HH)

Class: C1N2, bkg
Test_auc = \sum{Test_auc_class}
Train_auc = \sum{Train_auc_class}

Hyperparameter tune:
use optuna to auto-optimze

constraint:

average of AUC need to >= 0.6
penalty function: score = test_auc – 0.3*auc_gap (auc_gap = abs(train_auc- test_auc))
maximum(score)

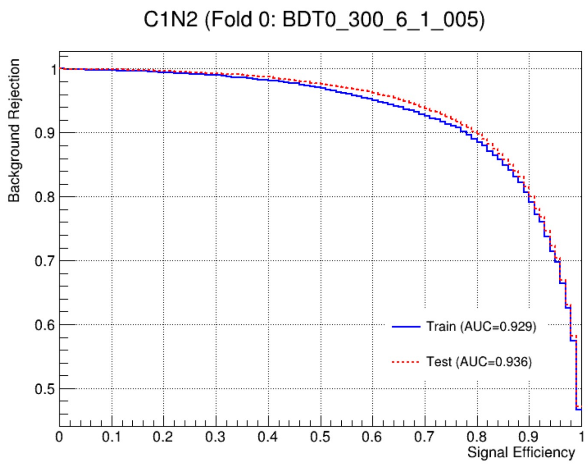
Grid Search
Ntrees: [200, 300, 400]
MaxDepth: [4, 6, 8, 10]
MinNode: [1, 3, 5, 7]
Learning rate: [0.001, 0.005, 0.01, 0.05, 0.1]

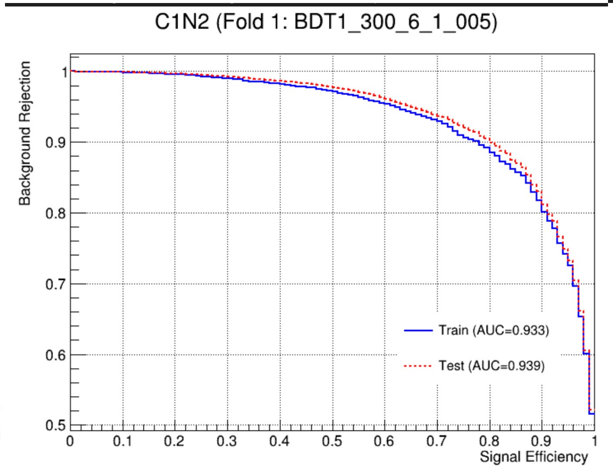
Best one: Ntree=300, MaxDepth=6, MinNode=1%, Learning Rate=0.05

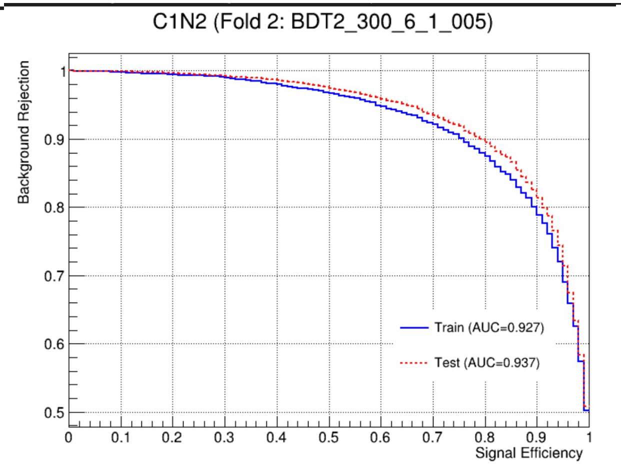
There still have rooms to optimize for lr

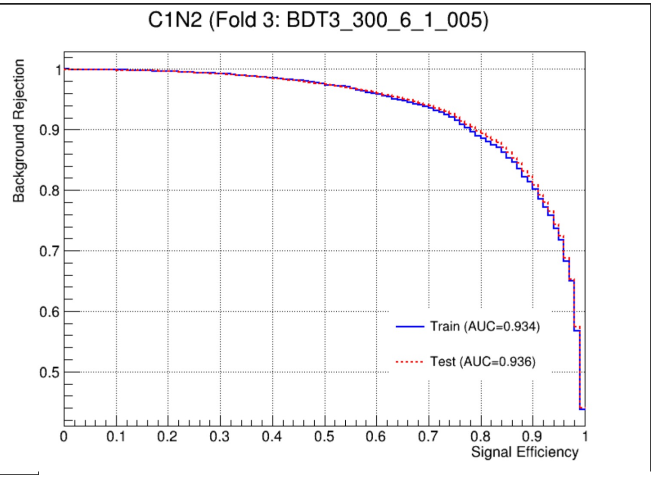
C1N2 ISR Binary class(HH)

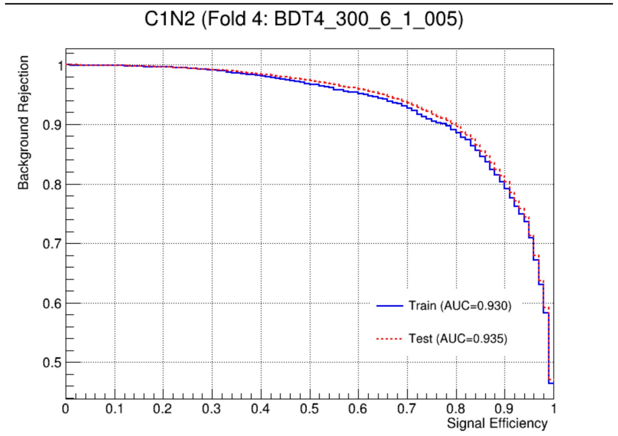
Overfit Check





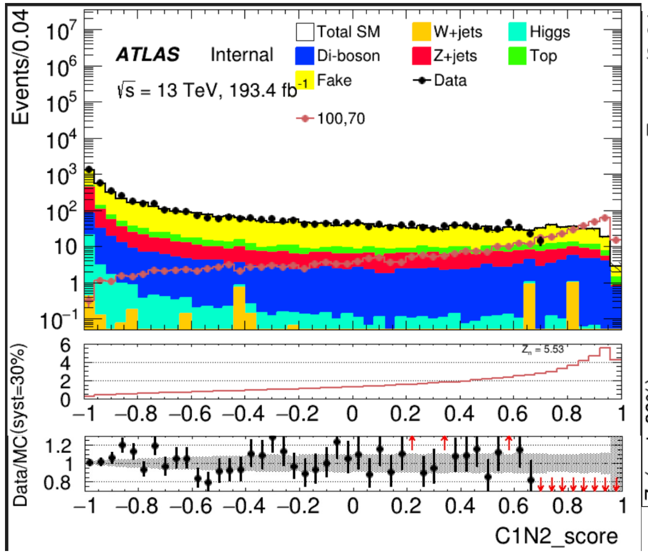




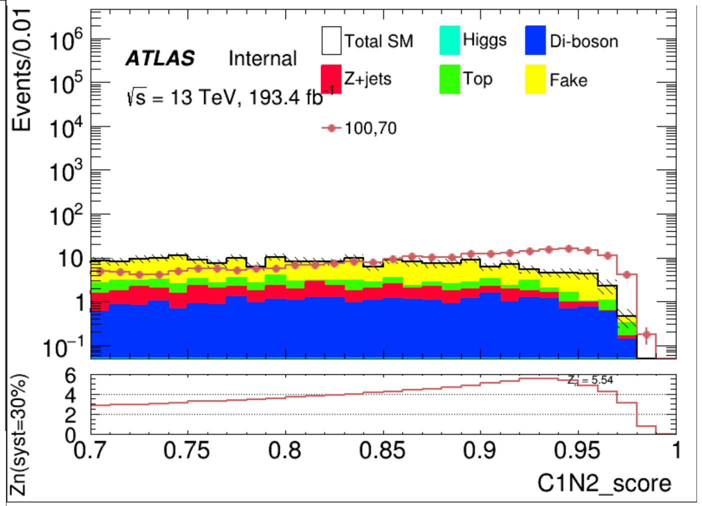


C1N2 ISR Binary class(HH)

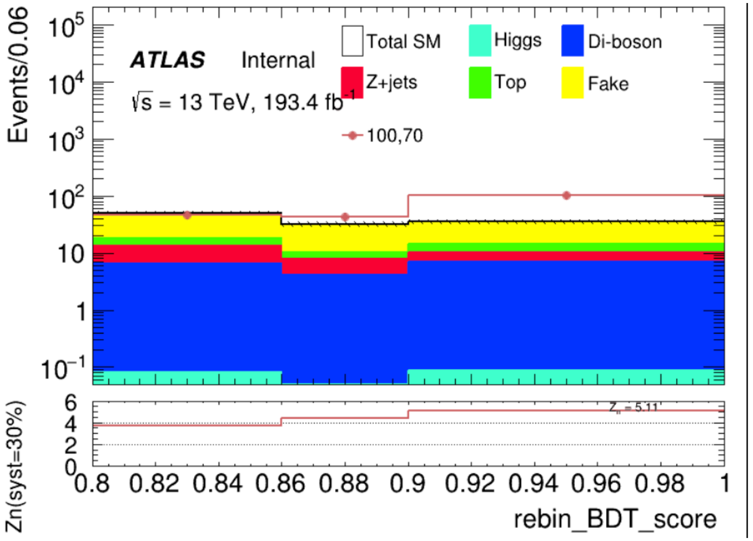
50 bins



Cut at 0.8



rebin



Sum Zn = 7.65

Bin Range	Zn	C1N2 (100_70) Yield ± Error	VV Yield ± Error	Top Yield ± Error	Fake Yield ± Error	Higgs Yield ± Error	Zjets Yield ± Error	Wjets Yield ± Error	Total Bkg Yield ± Error
[0.80,0.85]	3.65	46.569 ± 1.327	6.508 ± 0.368	4.719 ± 0.686	31.458 ± 3.898	0.080 ± 0.023	6.659 ± 0.416	0.978 ± 0.978	**50.402 ± 4.042**
[0.85,0.90]	4.37	43.128 ± 1.283	4.124 ± 0.341	2.663 ± 0.546	21.525 ± 3.292	0.045 ± 0.019	3.564 ± 0.286	0.000 ± 0.000	**31.921 ± 3.354**
[0.90,1.00]	5.11	101.059 ± 1.961	6.855 ± 0.398	4.331 ± 0.731	20.438 ± 2.976	0.086 ± 0.031	3.207 ± 0.250	0.000 ± 0.000	**34.917 ± 3.118**

C1N2 ISR Binary class(LH)

Hyperparameters: Ntrees = 200, MaxDepth = 6, MinNodeSize = 2%, Learning rate = 0.03(initial setting)

Feature engineering:

Select a simple model and put all features into model, choose Top 30 vars based on importance list, drop high correlated vars

Final feature list:

Rank	Variable	Variable Importance
1	fb_frac_MET_tau2	8.270e-02
2	fb_dRtt	6.684e-02
3	fb_dPhitt	6.226e-02
4	fb_frac_MET_tt	5.197e-02
5	fb_frac_jet_tau2	5.179e-02
6	fb_MT2_50	5.077e-02
7	fb_dPhiMax_tj	4.779e-02
8	fb_dPhiMin_xj	4.343e-02
9	fb_mt_taumin	3.547e-02
10	fb_Mll	3.511e-02
11	fb_mtx_tau1	3.408e-02
12	fb_nBaseJet	3.146e-02
13	fb_frac_jet_tt	3.110e-02
14	fb_mtx_tau2	2.941e-02
15	fb_frac_MET_tau1	2.898e-02
16	fb_METsig	2.824e-02
17	fb_pt_Vframe	2.726e-02
18	fb_Mwh	2.684e-02
19	fb_Proj_j	2.678e-02
20	fb_frac_MET_sqrtHT_40	2.560e-02
21	fb_frac_jet_tau1	2.518e-02
22	fb_MCT	2.254e-02
23	fb_Mwl	2.185e-02
24	fb_mt_quad_sum	2.165e-02
25	fb_Proj_tt	2.038e-02
26	fb_ht_tau	1.992e-02
27	fb_e_tau2	1.819e-02
28	fb_mt_sum_ttj	1.624e-02
29	fb_mt_tau2	1.618e-02

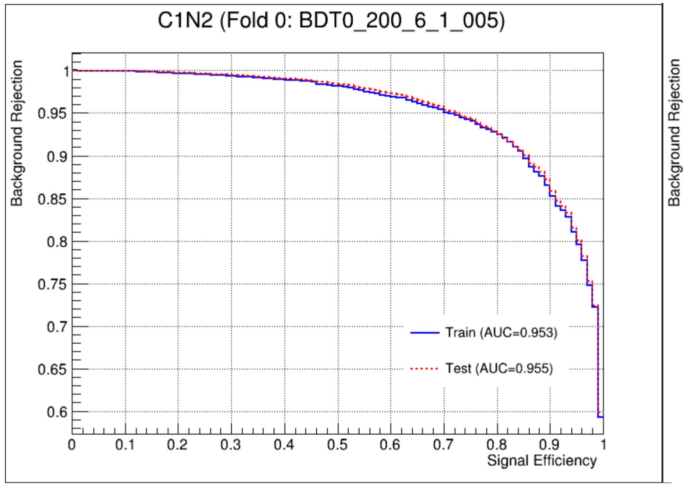
Weight choose: no weight, abs(weight)

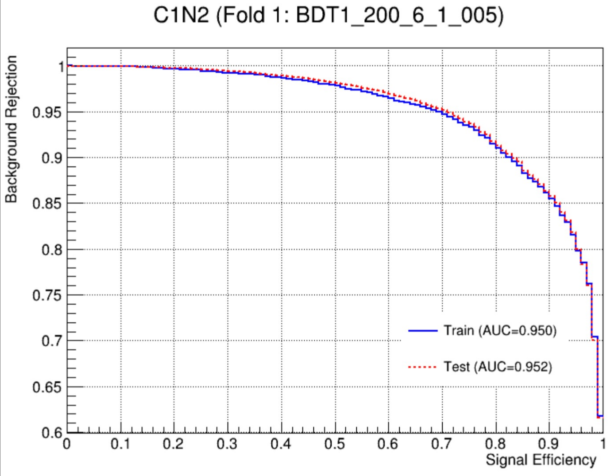
No weight have better performance
but abs(weight) fit our anaysis requirment

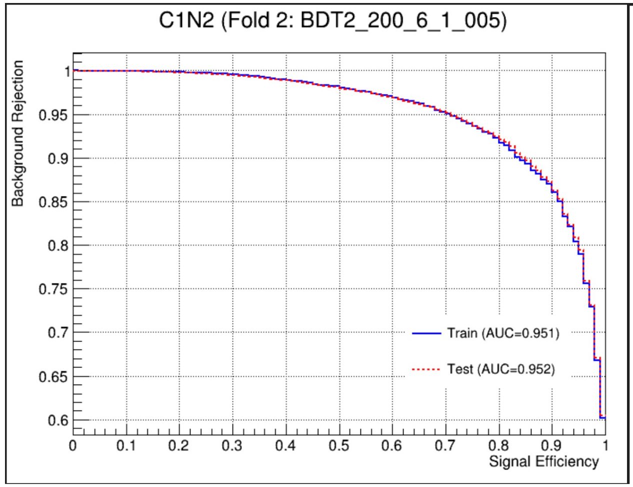
Split strategy: Separate entries by using mod 5, for Fake bkg, if separate follow sequence, all weighted entry will split into first fold

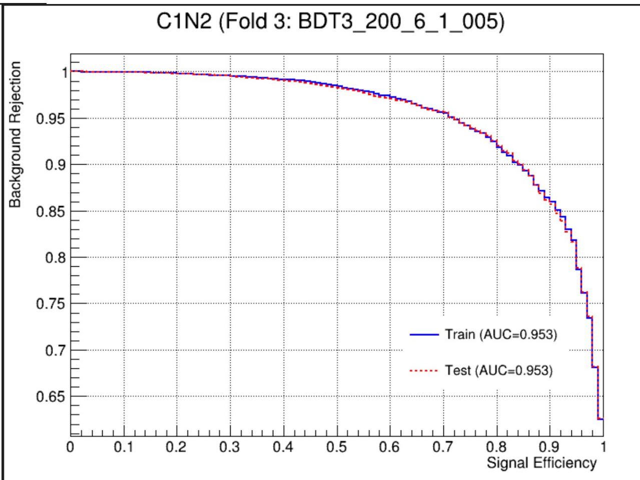
C1N2 ISR Binary class(LH)

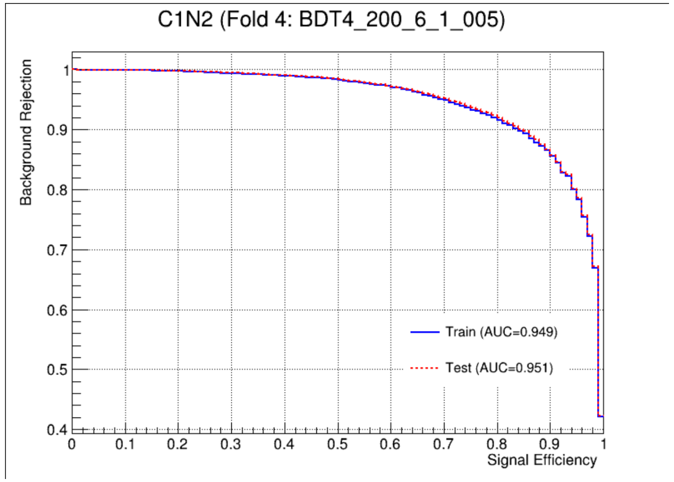
Overfit Check





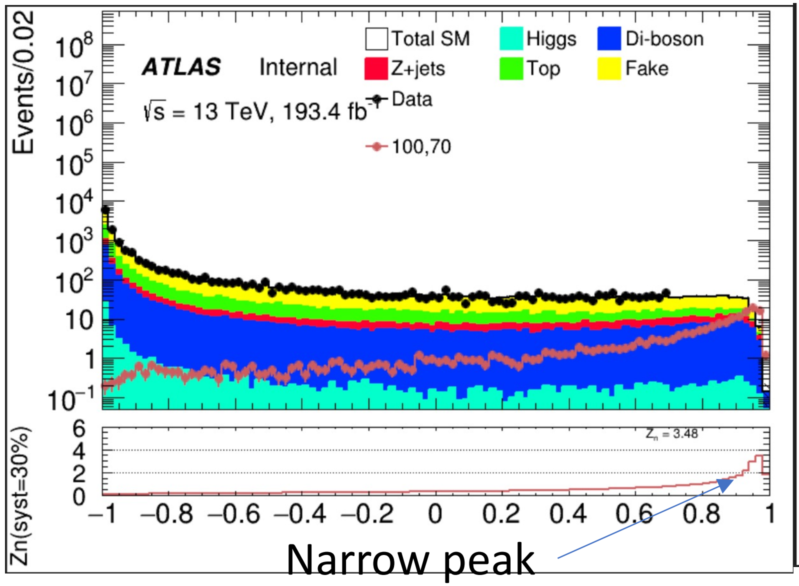






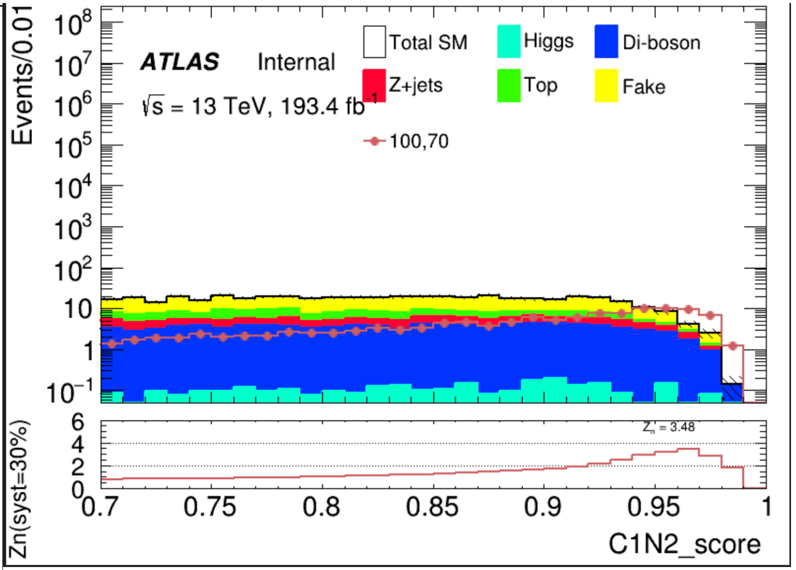
C1N2 ISR Binary class(LH)

100 bins

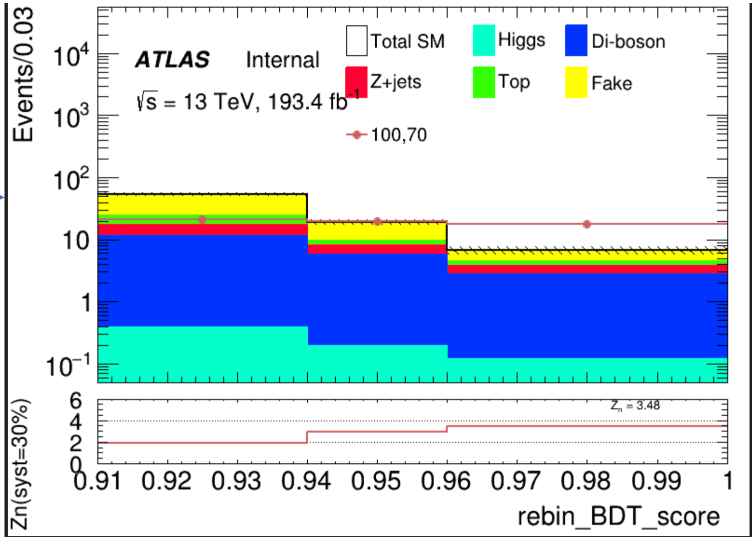


Narrow peak

Cut at 0.91



rebin

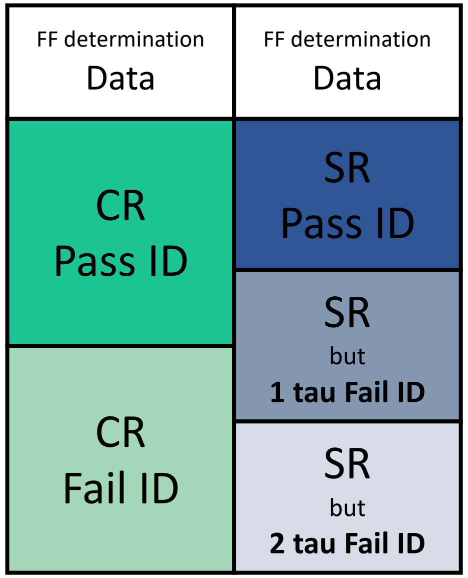


Sum Zn = 4.934

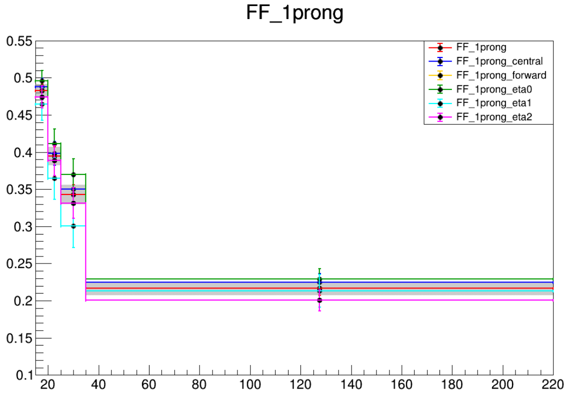
Bin Range	Zn	C1N2 (100_70) Yield ± Error	VV Yield ± Error	Top Yield ± Error	Fake Yield ± Error	Higgs Yield ± Error	Zjets Yield ± Error	Wjets Yield ± Error	Total Bkg Yield ± Error
[0.91,0.94]	1.90	21.447 ± 0.901	11.105 ± 0.565	7.432 ± 0.956	28.625 ± 3.295	0.376 ± 0.057	5.515 ± 0.357	0.173 ± 0.142	**53.226 ± 3.497**
[0.94,0.96]	2.94	19.678 ± 0.874	5.550 ± 0.367	1.524 ± 0.441	9.404 ± 1.929	0.189 ± 0.052	2.251 ± 0.222	0.000 ± 0.000	**18.918 ± 2.018**
[0.96,1.00]	3.47	17.543 ± 0.820	2.666 ± 0.263	0.578 ± 0.237	2.400 ± 1.009	0.118 ± 0.040	1.026 ± 0.118	0.000 ± 0.000	**6.788 ± 1.067**

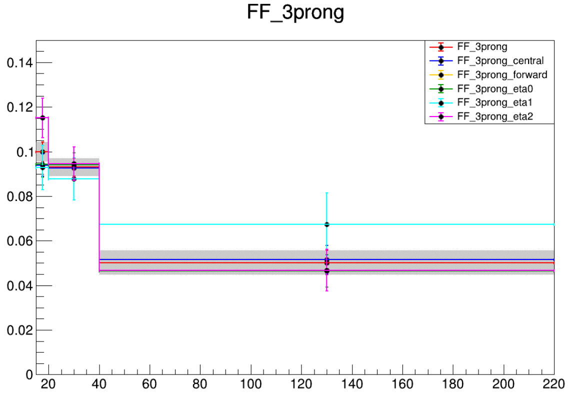
C1N2 ISR fake estimation(HH)

- CRs (fake factor computation)
  - METtrig
  - MET>=200
  - bveto
  - at least 1 signal lepton
  - $\Delta\phi(\tau, MET) > 2$
  - **ID: >= 1 medium tau**
  - **antiID: >= 1 VeryLoose tau, 0 medium tau**
- SRs
  - preselection
  - 2ID: >= 2 medium tau
  - 1ID1antiID: >= **2 VeryLoose tau** , 1 medium tau
  - 2antiID: >= **2 VeryLoose tau** , 0 medium tau



- **Binned in prongness, tau eta, tau pT**
  - Eta bins
    - 2 bins: central [0,1.37], forward [1.52,2.5]
    - 3 bins: eta0,1,2 for [0,1), [1, 1.37], [1.52,2.5]
- **Auto binning:**
  - > 10% of events in nominator and denominator
  - Add bins to bin i until it is not consistent anymore with bin i - 1
    - Relative stat uncertainty on ratio smaller than 50%
    - >10% events in nominator and denominator

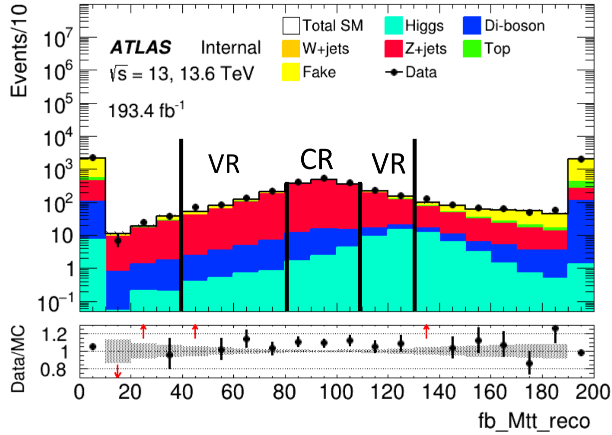


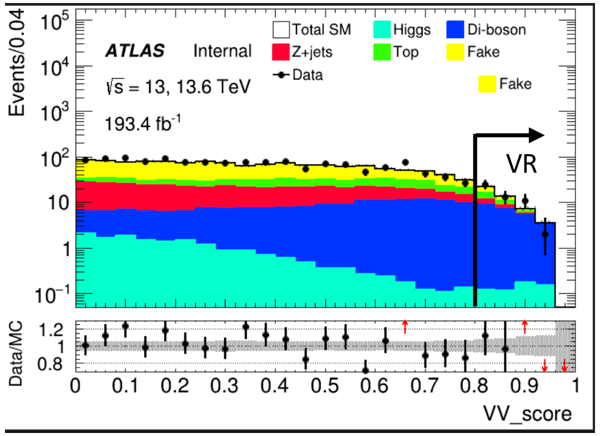


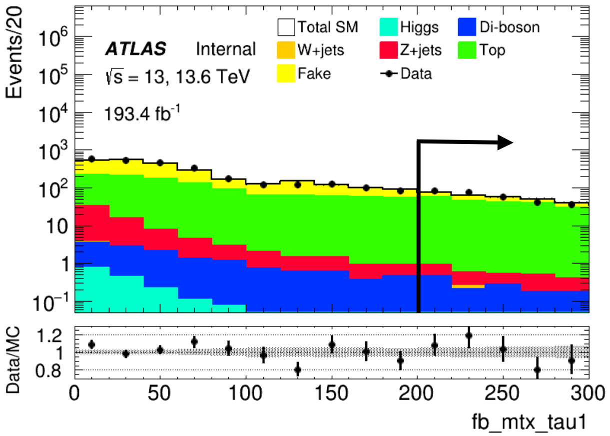
C1N2 ISR background estimation(HH)

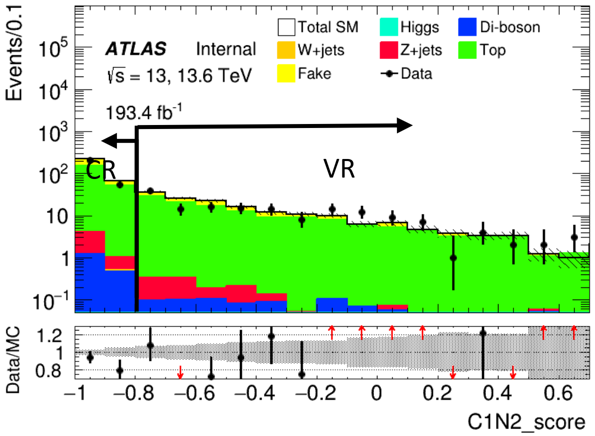
Selection for control region and validation region

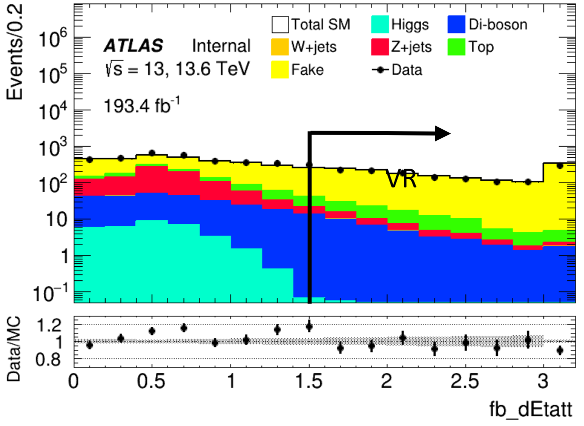
<table>
<tr><th rowspan="2">Process</th><th colspan="2">Top</th><th colspan="2">Z+jets</th><th>Multi-bosons</th><th>Fake</th></tr>
<tr><th>TCR</th><th>TVR</th><th>ZCR</th><th>ZVR</th><th>MBVR</th><th>FakeVR</th></tr>
<tr><td>Charge combination</td><td colspan="6">OS</td></tr>
<tr><td>Trigger</td><td colspan="6">MET trigger, $E_T^{miss} \geq 200GeV$</td></tr>
<tr><td>N medium τ</td><td colspan="6">≥ 2</td></tr>
<tr><td>N lep</td><td colspan="6">=0</td></tr>
<tr><td>$nBaseJet$</td><td colspan="6">≥ 1</td></tr>
<tr><td>$Jet \quad p_T$[GeV]</td><td colspan="6">≥ 100</td></tr>
<tr><td>N b-jets</td><td colspan="2">≥ 1</td><td colspan="4">$= 0$</td></tr>
<tr><td>$m(\tau_1, \tau_2)$ [GeV]</td><td colspan="2">≤ 40 or ≥ 130</td><td>[80,110]</td><td>[40,80] or [110,130]</td><td colspan="2">≤ 40 or ≥ 130</td></tr>
<tr><td>$M_T(\tau_1, E_T^{\text{miss}})$</td><td colspan="2">$\geq 200$</td><td>–</td><td>–</td><td>–</td><td>–</td></tr>
<tr><td>$d\eta(\tau_1, \tau_2)$</td><td>–</td><td>–</td><td>–</td><td>–</td><td>–</td><td>≥ 1.5</td></tr>
<tr><td>C1N2 score</td><td>[-1, -0.8]</td><td>[-0.8, 0.7]</td><td colspan="4">≤ 0.7</td></tr>
<tr><td>VV score</td><td>–</td><td>–</td><td>–</td><td>–</td><td>≥ 0.80</td><td>–</td></tr>
<tr><td>Total bkg</td><td>290+-7</td><td>162+-5</td><td>1420+-10</td><td>1286+-11</td><td>46+-2</td><td>2000+-28</td></tr>
<tr><td>Dominant Bkg</td><td>206+-4</td><td>131+-3</td><td>1221+-5</td><td>993+-4</td><td>24+-1</td><td>1851+-28</td></tr>
<tr><td>Purity</td><td>0.71</td><td>0.80</td><td>0.86</td><td>0.77</td><td>0.52</td><td>0.93</td></tr>
<tr><td>Data</td><td>264</td><td>161</td><td>1559</td><td>1427</td><td>51</td><td>1950</td></tr>
<tr><td>Data/MC</td><td>0.91</td><td>0.99</td><td>1.09</td><td>1.10</td><td>1.10</td><td>0.98</td></tr>
</table>











C1N2 ISR background estimation(LH)

Selection for control region and validation region

Process	Top		Z+jets		Multi-bosons	Fake
	TCR	TVR	ZCR	ZVR	MBVR	FakeVR
Charge combination	OS					
Trigger	MET trigger, $E_T^{miss} \geq 200GeV$					
N medium τ	≥ 1					
N lep	≥ 1					
$nBaseJet$	≥ 1					
$Jet \quad p_T$[GeV]	≥ 100					
N b-jets	≥ 1		$= 0$			
$m(\tau_1, l)$ [GeV]	≤ 40 or ≥ 130		[80,110]	[40,80] or [110,130]	≤ 40 or ≥ 130	
$M_{inv}(l, MET)$	≥ 300		–	–	–	–
$\Delta\phi(MET, \tau_1)$	–	–	–	–	–	≥ 2
$dR(\tau, l)$	–	–	≤ 0.6		–	–
C1N2 score	[-1, -0.95]	[-0.95, 0.7]	≤ 0.7			
VV score			–	–	≥ 0.80	–
Total bkg	4151+-11	1212+-13	917+-6	823+-4	78+-2	4983+-40
Dominant Bkg	3781+-20	1102+-11	823+-4	606+-3	53+-1	4390+-40
Purity	0.91	0.90	0.90	0.84	0.68	0.88
Data	3626	1119	908	745	72	5043
Data/MC	0.87	0.91	0.99	1.03	0.92	1.01