Search for lepton flavor violation in pp collisions at $\sqrt{s}$ = 13 TeV with ATLAS detector

Xiangke Zhang

The 6th China LHC Physics Workshop

Shanghai JiaoTong University

Nov 8, 2020

Outline

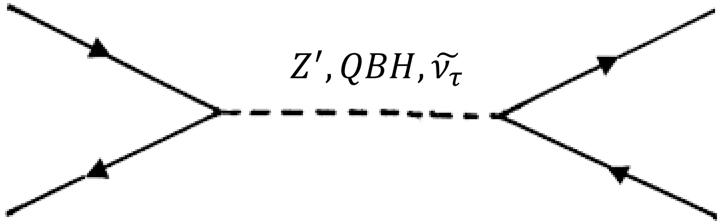
- Introduction
  - motivation and strategy
- Objects and events selection
- Background estimation
  - $W + jets$ background estimation
  - multi-jets estimation
  - $t\bar{t}$ background estimation
- Statistical analysis
- Summary

Introduction: motivation & strategy

- Direct charged-lepton flavor violation (LFV) is forbidden in the Standard Model

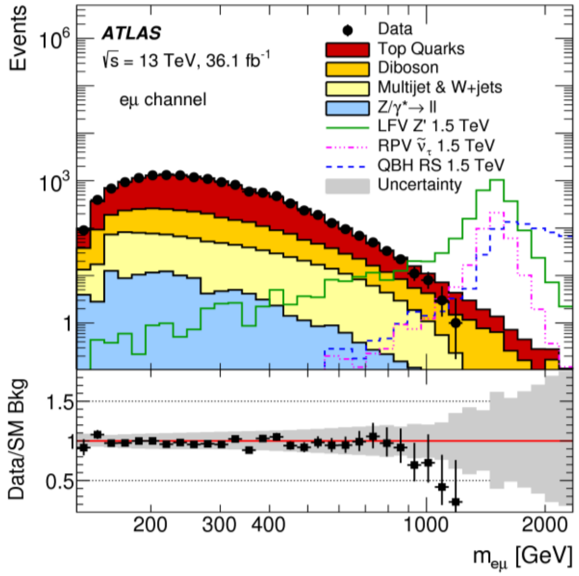
But it's allowed in hypothetic new physics models

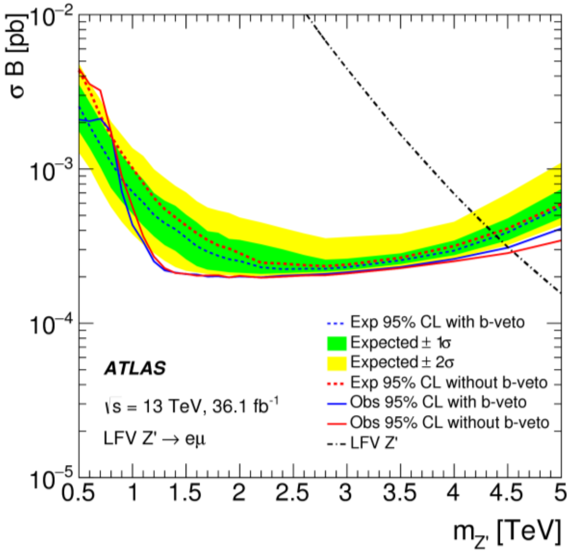
models / Production & reason	SM	models with additional gauge symmetries	R-parity violating(RPV) SUSY	quantum black hole QBH
lepton pairs with different flavor (LFV)	Not allowed	allowed	allowed	allowed
Reason	LFC	Z'	sneutrino τ resonance	$pp \to QBH \to l^{\mp} l'^{\pm}$
Final state		$e\mu$, $e\tau$, $\mu\tau$		



Introduction: motivation & strategy

- Aim of this analysis
  - search for a new resonance with two leptons of different flavor ($e\mu$, $e\tau$, $\mu\tau$) in high mass region
  - otherwise, set limits on the parameters of new physics models
- Clear experimental signature
  - low background from SM processes
  - the invariant mass of the heavy neutral particle can be reconstructed
- Previous publications
  - [Phys. Rev. D 98 (2018) 092008]
  - int. luminosity $36\ fb^{-1}$





Objects & event selection

muon

- $p_T > 65GeV \ and \ |\eta| < 2.5$
- **ID**: high-pT
- **isolation**: FixedCutLoose
- **track**: $|d_0/\sigma_{d_0}| < 3 \ and \ |\Delta z_0 sin\theta| < 0.5mm$

electron

- $p_T > 65GeV \ and \ |\eta| < 2.5$
- **ID**: LHTight
- **isolation**: FixedCutTight
- **track**: $|d_0/\sigma_{d_0}| < 5 \ and \ |\Delta z_0 sin\theta| < 0.5mm$

tau

- $p_T > 65GeV \ and \ |\eta| < 2.5$
- **ID**: RNN Medium
- **track**: 1 or 3 prongs
- **charge**: ± 1

- exclude crack region $1.36 < |\eta| < 1.52$ for electron and tau

event selection

- **pre-selection**: remove events with error status flag on detectors
- **trigger**: pass single-electron or single-muon triggers & at least 1 trigger matched lepton
- **opposite charge**
- **3rd lepton veto**: events with an additional good lepton are rejected
- **b-veto**: DL1r:FixedCutBEff_85
- $\Delta\Phi_{ll} > 2.7$
- $m_{ll} > 130GeV$
- select $e\mu$, $e\tau$, $\mu\tau$ pairs

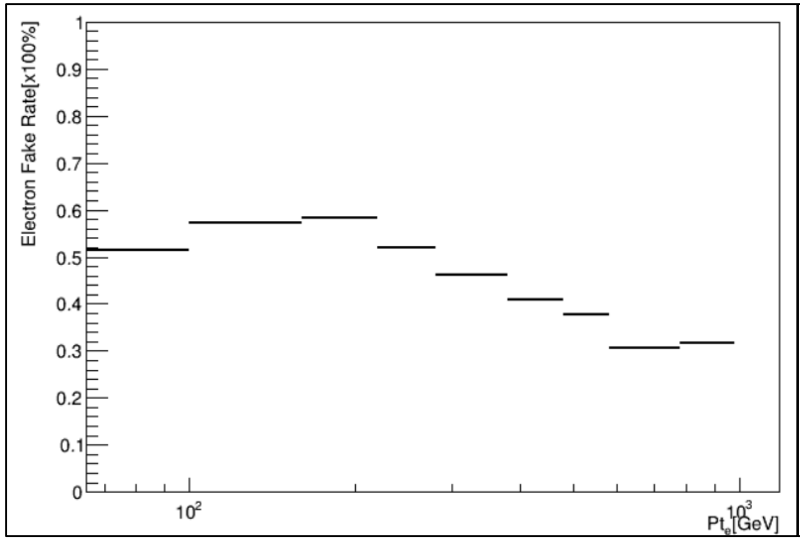
Background estimation

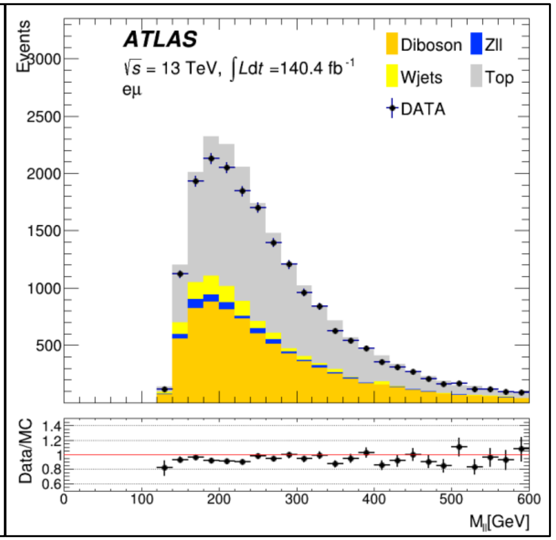
- Background contributions
  - irreducible background
    - includes 2 prompt leptons in the final states
    - $t\bar{t}$: MC simulation + extrapolation (for high-mass)
    - di-boson: MC simulation
    - drell-yan: MC simulation
  - reducible background
    - includes fake leptons in the final states
    - W+jets: MC simulation + data-driven
    - multi-jets: data-driven
    - fake estimation (W+jets and multi-jets)
      - $e\mu$ channel: estimated simultaneously with Matrix Method
      - $e\tau/\mu\tau$ channel: W+jets and multi-jets estimated separately

Irreducible backgrounds	**Reducible backgrounds**
SM Drell-Yan($q\bar{q} \rightarrow Z/\gamma^* \rightarrow ll$) $t\bar{t}$ & Single top Diboson	QCD W+jets
With two prompt leptons in the final state	Jets or non-prompt leptons are reconstructed as prompt leptons
MC simulation	Data-driven

Background estimation: fake in $e\mu$ channel

- Estimated using the Matrix Method
  - all events considered here pass at least loose selection
  - muon fake rate found to be negligible
    - assume all selected muons are real muons
    - $\begin{bmatrix} N_{RR} & N_{RF} \\ N_{FR} & N_{FF} \end{bmatrix} \rightarrow \begin{bmatrix} N_{RR} \\ N_{FR} \end{bmatrix}$
      - 4 → 2 different kinds of events to be investigated
      - N_{FR} (one fake electron + one real muon) will be the number of reducible background events
  - $\begin{bmatrix} N_{TT} \\ N_{LT} \end{bmatrix} = \begin{bmatrix} r_e & f_e \\ 1 & 1 \end{bmatrix} \begin{bmatrix} N_{RR} \\ N_{FR} \end{bmatrix}$
    - N_{xy}: number of events with x(electron) or y(muon) that
      - pass T(tight) or L(loose) ID
      - is R(real) or F(fake)
    - r_e: electron real efficiency
      - the probability of a "loose" electron(looser ID and without isolation requirement) matched to a generated electron to pass the full object selection
      - estimated by Zee MC samples
    - f_e: electron fake rate
      - the probability that a jet is misidentified as an electron
      - estimated in a multi-jet CR



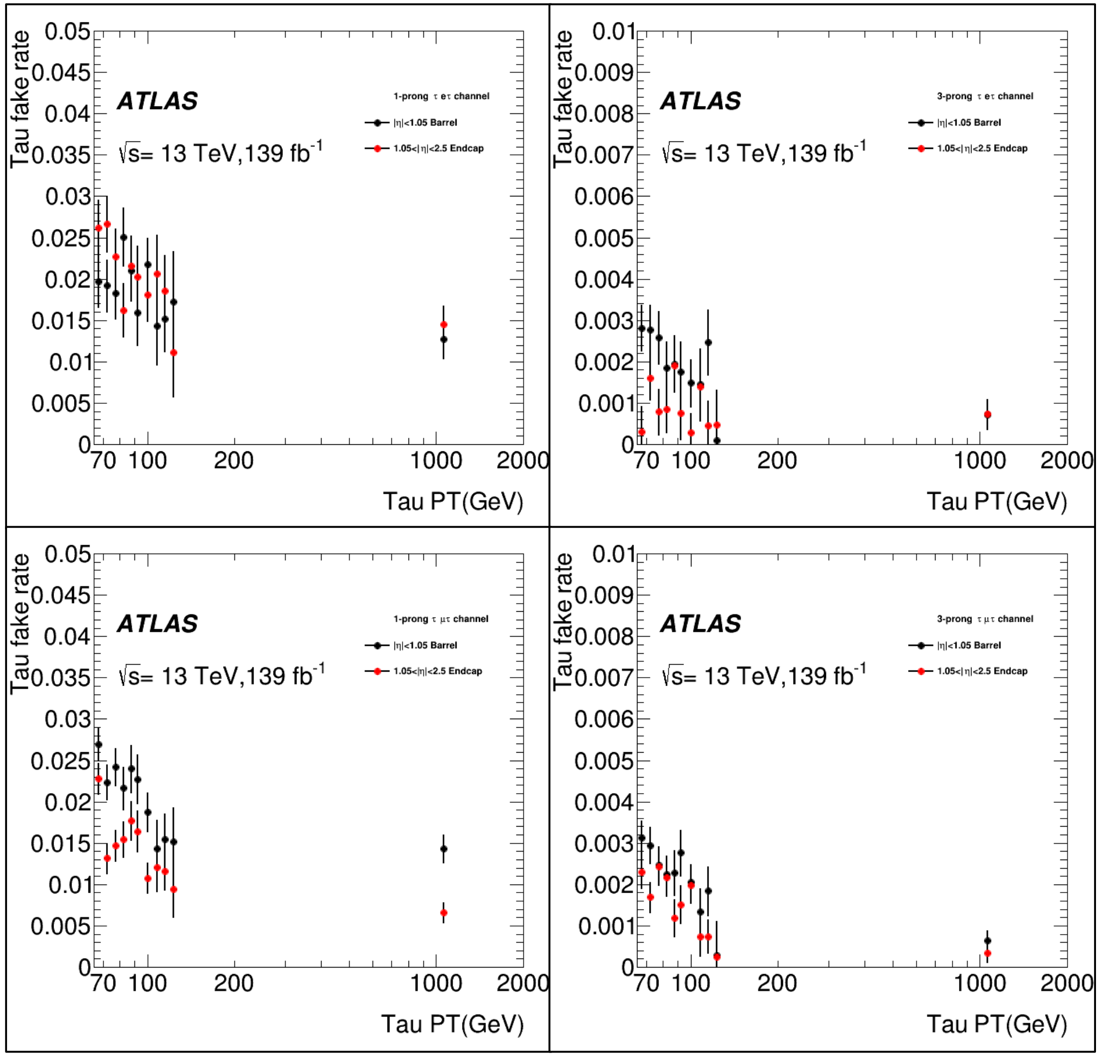


data	di-boson	top	Zll	Wjets+ multi-jets	total MC
18832	7534.16	11062.8	455.32	1013.99 (5.1%)	20056.28

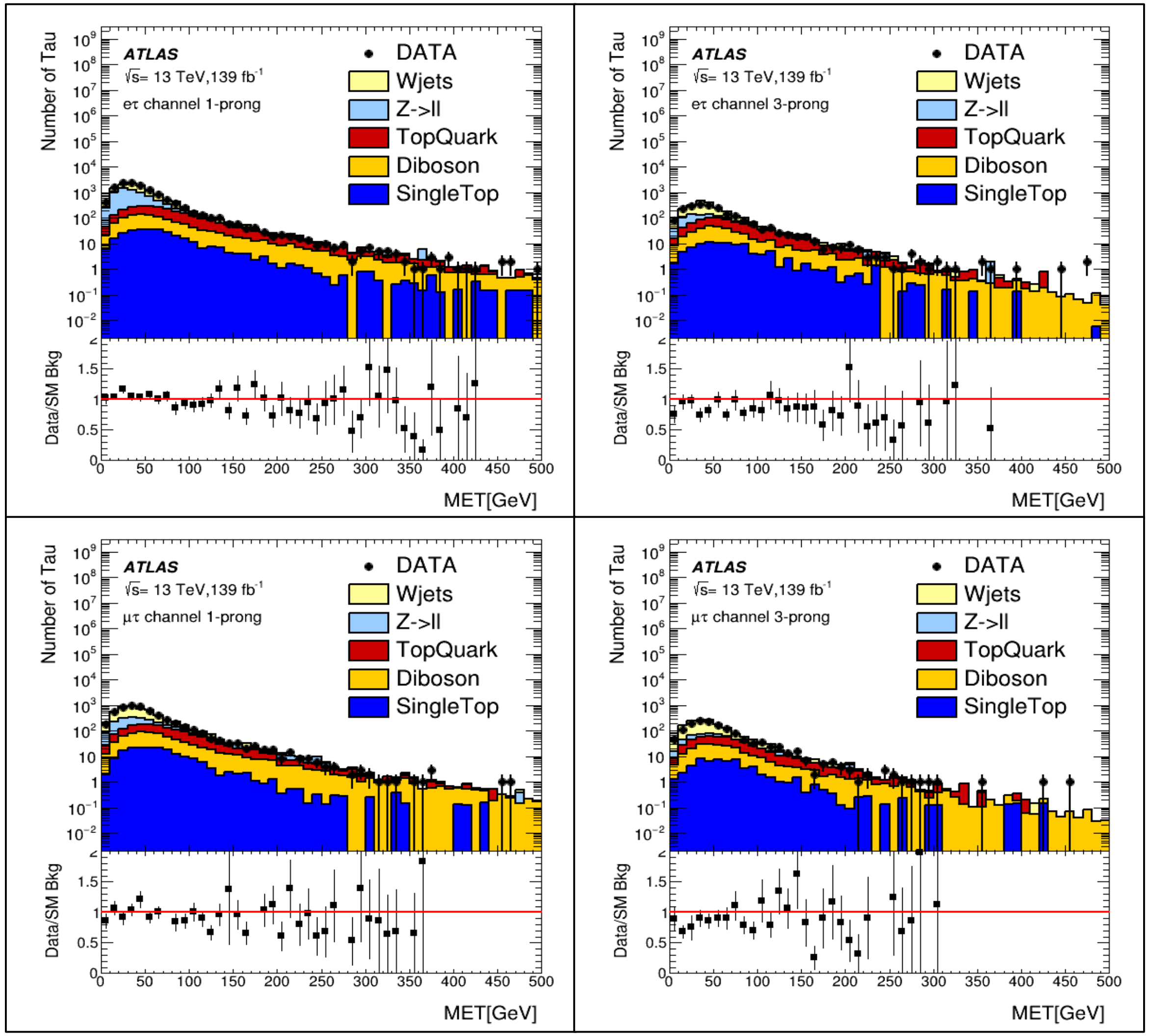
- ➢ the total electron fake rate after b-veto is 52.5%, close to the last run(56.4%)
- ➢ will be cross-checked with the IFF tool
- ➢ systematic uncertainty study and validation is ongoing

Background estimation: Wjets in $\boldsymbol{e\tau/\mu\tau}$ channel

- Wjets is a reducible background
  - mainly jet misidentified as τ
- Wjets control region definition
  - only $e\tau/\mu\tau$ channel
    - $e\tau$: $E_{missing}^{T} > 40GeV$
    - $\mu\tau$: no $E_{missing}^{T}$ cut
  - $130GeV < m_{ll} < 600GeV$
  - $\Delta\Phi_{\mathrm{ll}} < 2.7$
    - reverse the $\Delta\Phi_{ll}$ to ensure orthogonality with SR
  - B-jet veto
  - no more transverse mass cut
    - was $m_T(l_{lead}, E_{missing}^{T}) > 80GeV$
  - compared with/without Tau ID requirement
- Estimation method
  - data-driven: $\tau_{fake-rate} = \frac{N_{\tau}^{passID}}{N_{\tau}^{total}}$
  - tau fake rate as a function of tau p_T



Background estimation: Wjets in $e\tau/\mu\tau$ channel



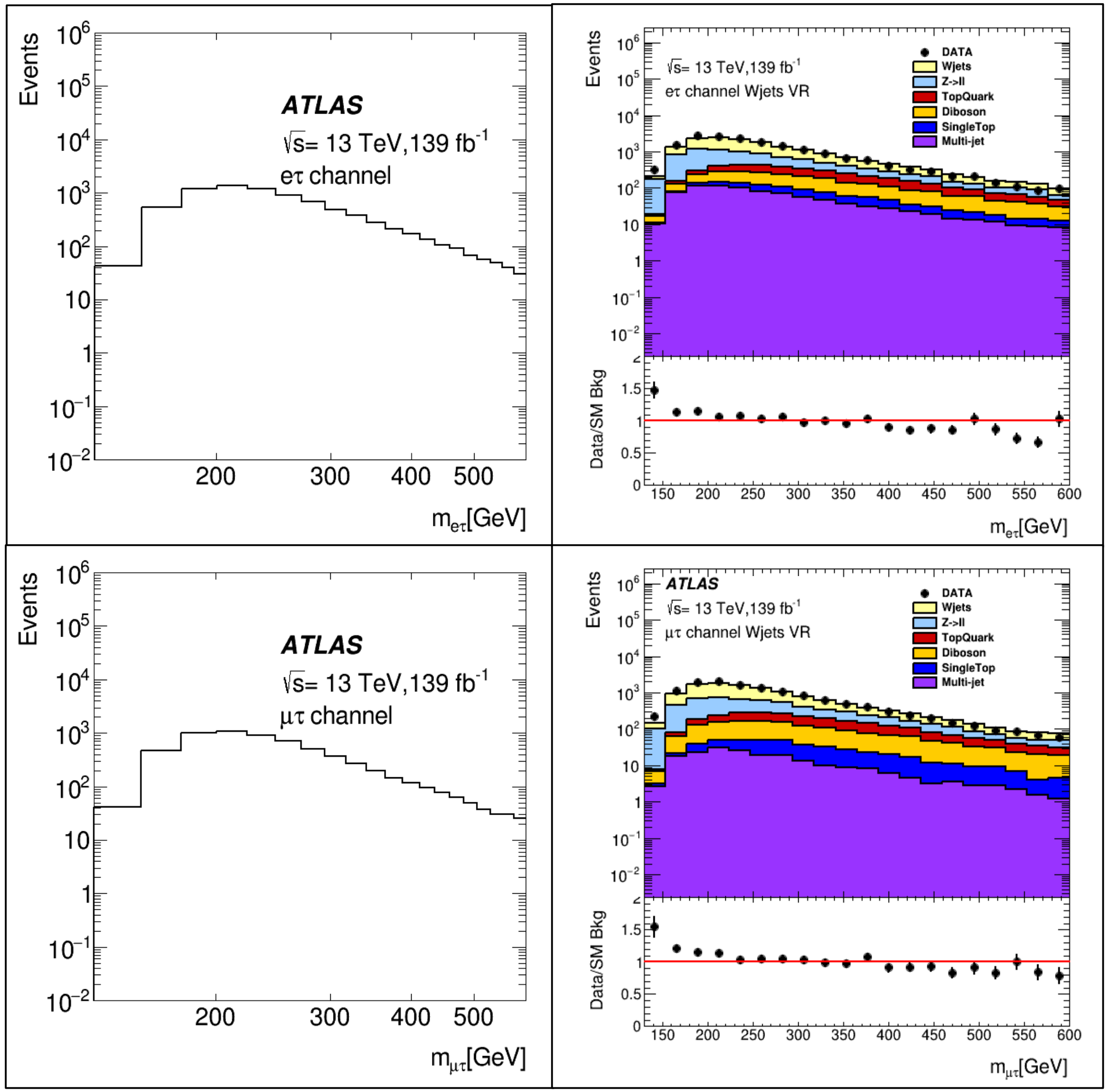
1-prong 3-prong

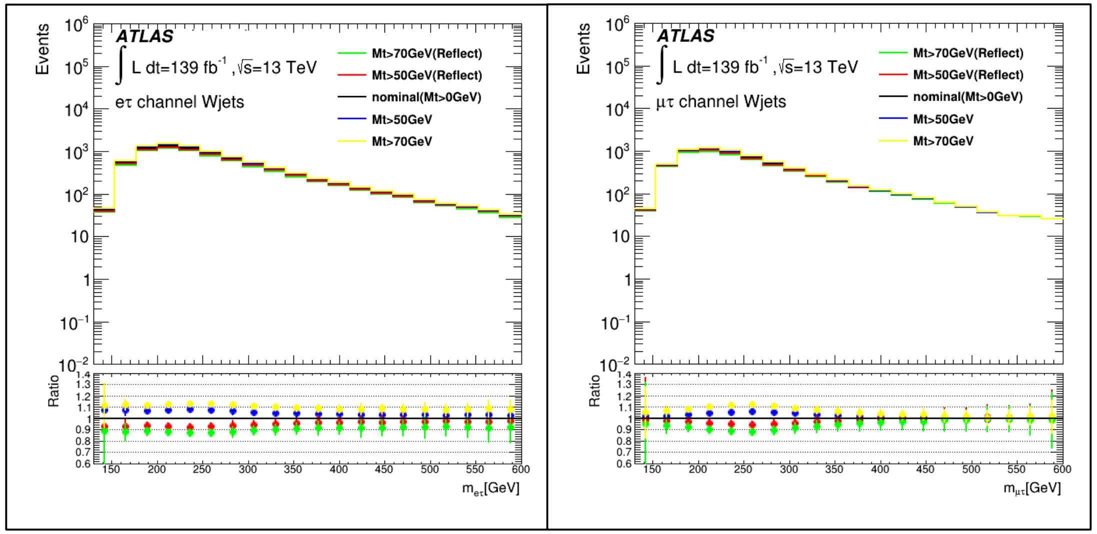
$E^T_{missing}$ cut decision in Wjets CR

- Zll is the second large backgrounds except Wjets
  - need to be minimize for Wjets CR
- 1-prong tau is the major contribution of Zee/Zmumu
- different treatment for two channels
  - $e\tau$ channel
    - Zee is the major contribution
    - cut at 40 GeV to get rid of Zll
  - $\mu\tau$ cahnnel
    - Zmumu is the major contribution
    - not $E^T_{missing}$ cut applied, otherwise will lose many Wjets events

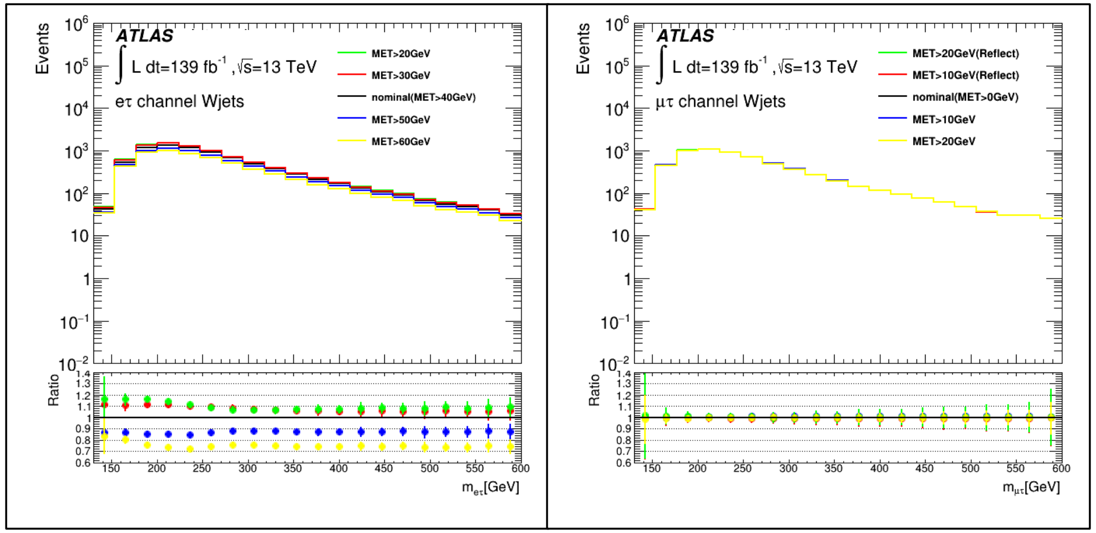
	$e\tau$	$\mu\tau$
Z->ee(361106)	4915.9	0
Z->mumu(361107)	0	833.8
Z->tautau(361108)	454.3	288.9

Background estimation: Wjets in $e\tau/\mu\tau$ channel





↑ fake rate with different m_T cut applied

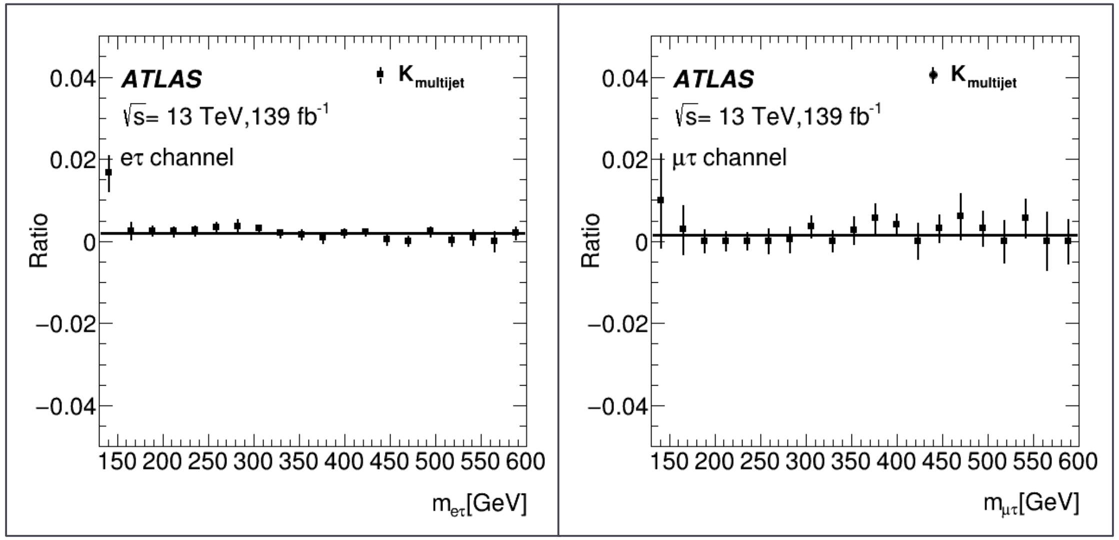


↑ fake rate with different $E^T_{missing}$ cut applied

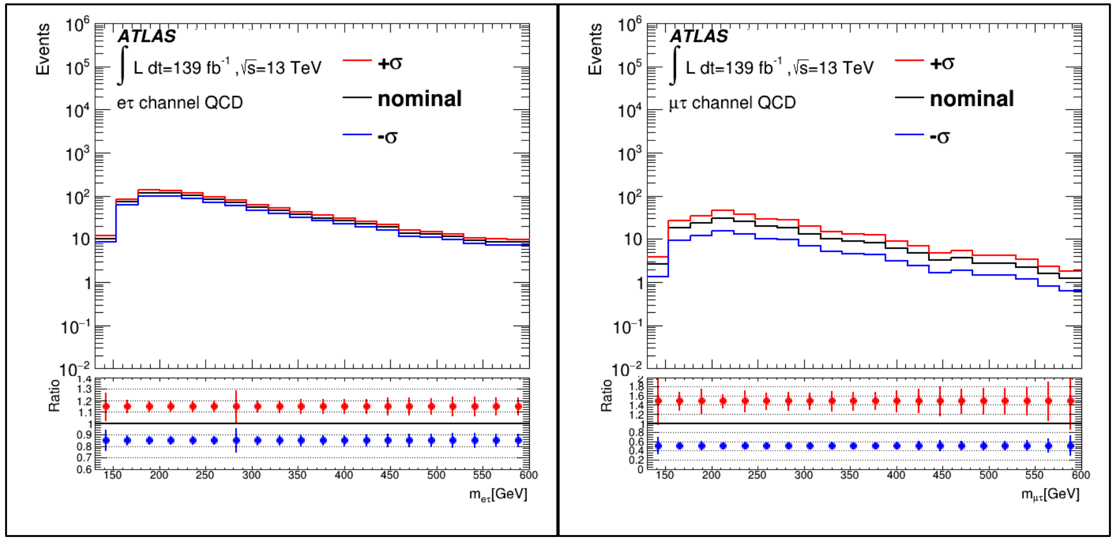
Background estimation: Multi-jet in $\boldsymbol{e\tau/\mu\tau}$ channel

- Define 3 study regions with respect to the signal selection
- Equations to estimate
  - $N_{multi-jet}^{region} = N_{data}^{region} - N_{MC}^{region}$
  - $K_{multi-jet} = N_{multi-jet}^{region2} / N_{multi-jet}^{region1}$
  - $N_{multi-jet}^{SR} = N_{multi-jet}^{region3} \times K_{multi-jet}$

region	same/oppo charge	non/isolated e/μ	require τ ID
1	same	non-isolated	no
2	same	isolated	yes
3	oppo	non-isolated	no



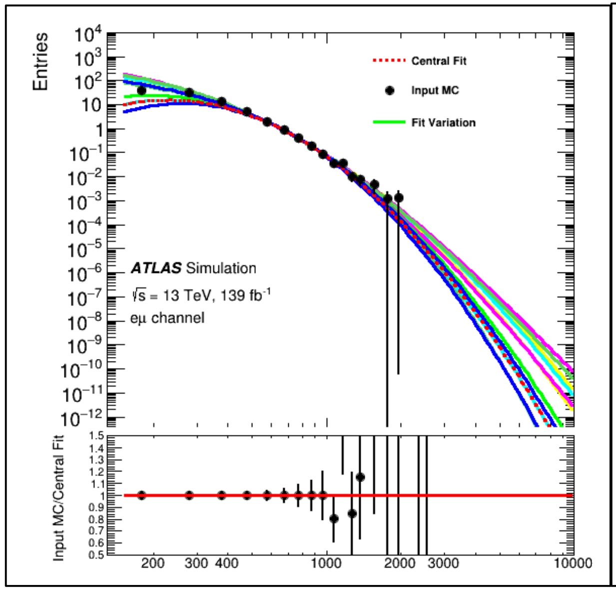
$e\tau$ $\mu\tau$

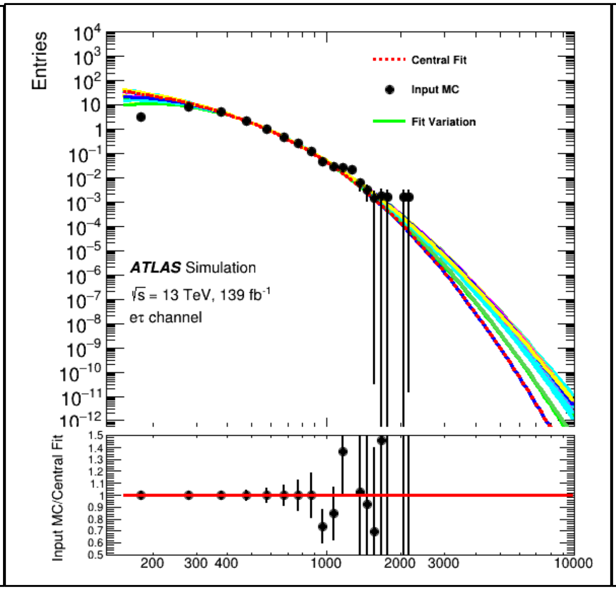


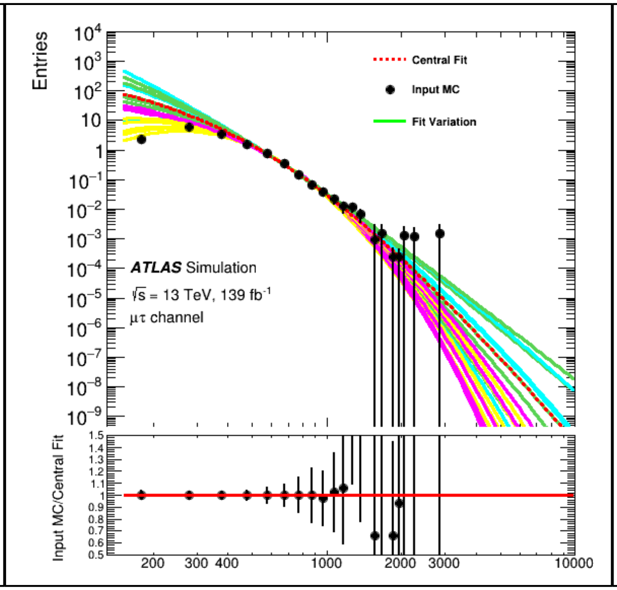
Background estimation: top background estimation

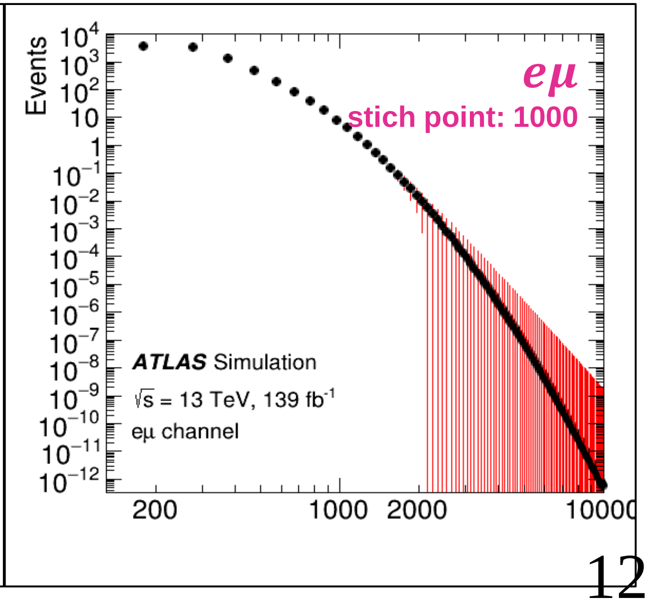
- Top is an irreducible background
  - MC simulation + extrapolation estimation
- Top control region definition
  - same selection as SR except the following
  - final events have
    - one good e/μ + one good τ
- Extrapolation method
  - for high mass region top estimation
    - MC simulation statistically limited in high m_{ll} region
  - two functions were investigated
    - di-jet function: $e^{-a}x^{b}x^{c\ln x}$, monomial: $\frac{a}{(x+b)^{c}}$
- multiple fitting ranges were investigated
  - find best fit among a few hundred fittings
  - criteria: mean, median and best chi2 (give similar results)
  - available MC statistics and extrapolation agree well beyond the stitching point

Unit: GeV	$e\mu$	$e\tau$	$\mu\tau$
lower range	450, 470, … , 530, 550	300, 320, … , 440, 460	450, 470, … , 530, 550
upper range	1000, 1025, …,1175, 1200	1000, 1025, …, 1475, 1500	1000, 1025, …, 1475, 1500
stitch point	1000	900	900







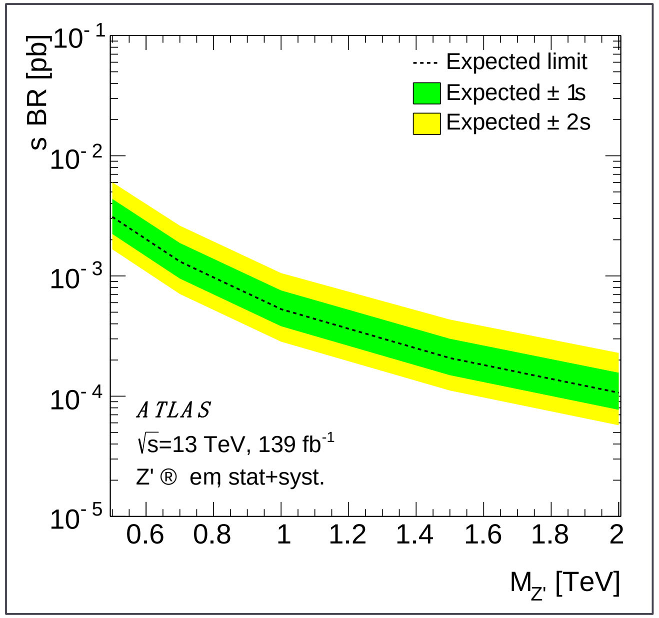


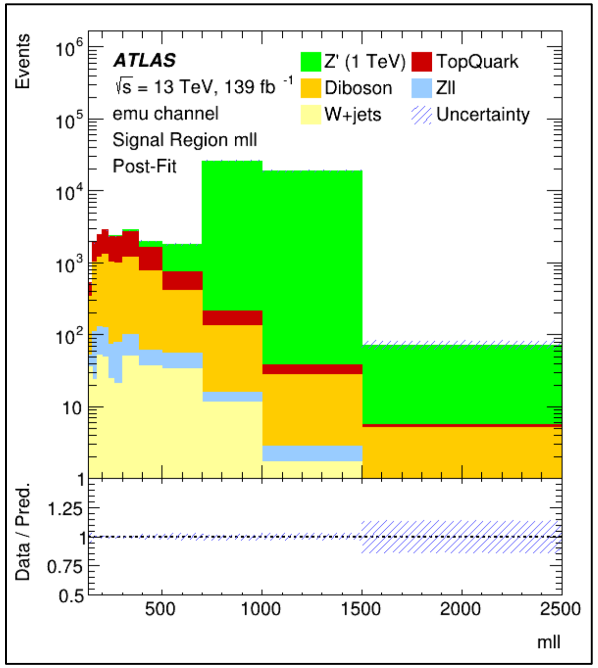
Statistical analysis: setup

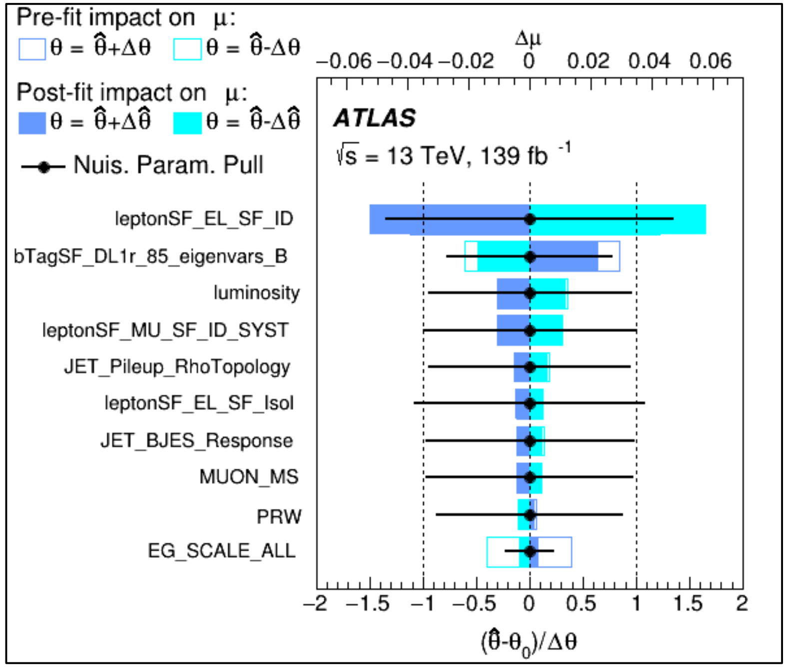
- Signal samples considered
  - Zprime (500 GeV, 700 GeV, 1 TeV, 1.5 TeV, and 2 TeV)
- Framework: TRexFitter
  - binning:
    - 130,150,175,200,230,260,300,380,500,700,1000,1500,2000,2500
  - NPs considered (Instrumental):
    - histograms based experimental syst. on all the bkg/sig
      - theoretical systematics not included yet
    - one luminosity uncertainty ($\pm$0.017) on all the bkg/sig
    - QCD variation from data-driven
    - overall theory uncertainty on the background
      - samples with theory uncertainty is not ready yet

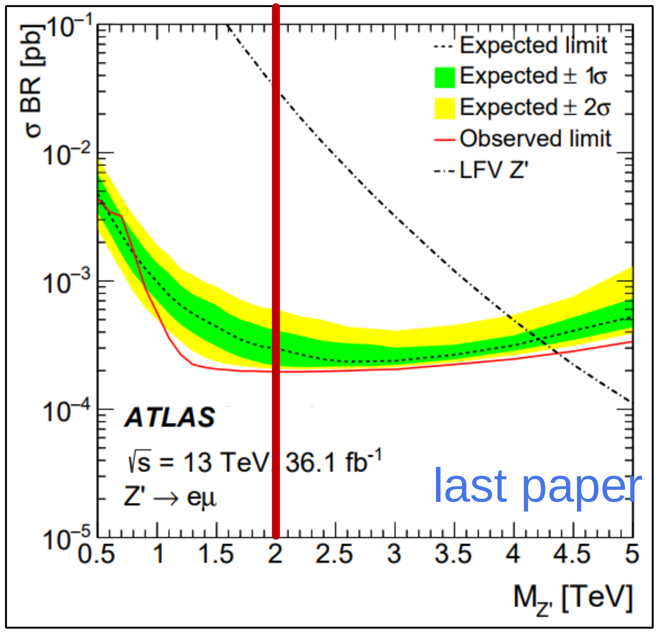
Statistical analysis: limit calculation ($e\mu$)

MP/GeV	XS*BR
500	0.00310
700	0.00132
1000	0.000530
1500	0.000208
2000	0.000107









Summary

- Background estimation
  - top background
    - extrapolation method was investigated and applied above stitch point
    - haven’t be included in limit fitting yet
  - Wjets and multi-jet background
    - Wjets fakes rates were estimated, and data/MC looks consistent in Wjets CR
    - multi-jet estimated with data-driven by fitting a constant $K_{multi-jet}$
- Limit fitting
  - setup using TRexFitter
  - experimental systematics were investigated, need to include theoretical systematics next
- Finalizing the analysis

Back up

Statistical analysis: experimental uncertainty

	up	down
leptonSF_MU_SF_Isol_STAT	**0.36%**	**-9.62%**
bTagSF_DL1r_85_eigenvars_B	**6.76%**	**-4.67%**
GlobalReduction_JET_Pileup_RhoTopology	**-2.46%**	**2.60%**
GlobalReduction_JET_BJES_Response	**-1.87%**	**1.96%**
pileup	**-1.27%**	**-0.92%**
leptonSF_EL_SF_ID	**1.09%**	**-1.09%**
GlobalReduction_JET_EffectiveNP_1	**-0.96%**	**1.04%**
GlobalReduction_JET_EtaIntercalibration_Modelling	**-0.75%**	**0.92%**
GlobalReduction_JET_JER_EffectiveNP_2	**0.68%**	**0.68%**
leptonSF_MU_SF_ID_SYST	**0.78%**	**0.00%**
total	**7.86%**	**-11.47%**

$e\mu$

	up	down
bTagSF_DL1r_85_eigenvars_B	**-4.40%**	**6.08%**
GlobalReduction_JET_Pileup_RhoTopology	**2.92%**	**-2.23%**
GlobalReduction_JET_BJES_Response	**2.14%**	**-1.48%**
tauSF_EFF_RNNID_SYST	**-1.57%**	**1.57%**
TAUS_TRUEHADTAU_SME_TES_PHYSICSLIST	**-1.14%**	**1.39%**
leptonSF_EL_SF_ID	**-1.13%**	**1.13%**
tauSF_EFF_ELEOLR_TOTAL	**-1.08%**	**1.08%**
GlobalReduction_JET_EffectiveNP_1	**1.08%**	**-0.89%**
pileup	**-0.60%**	**-1.11%**
TAUS_TRUEHADTAU_SME_TES_MODEL_CLOSURE	**-0.74%**	**1.00%**
total	**6.66%**	**-7.54%**

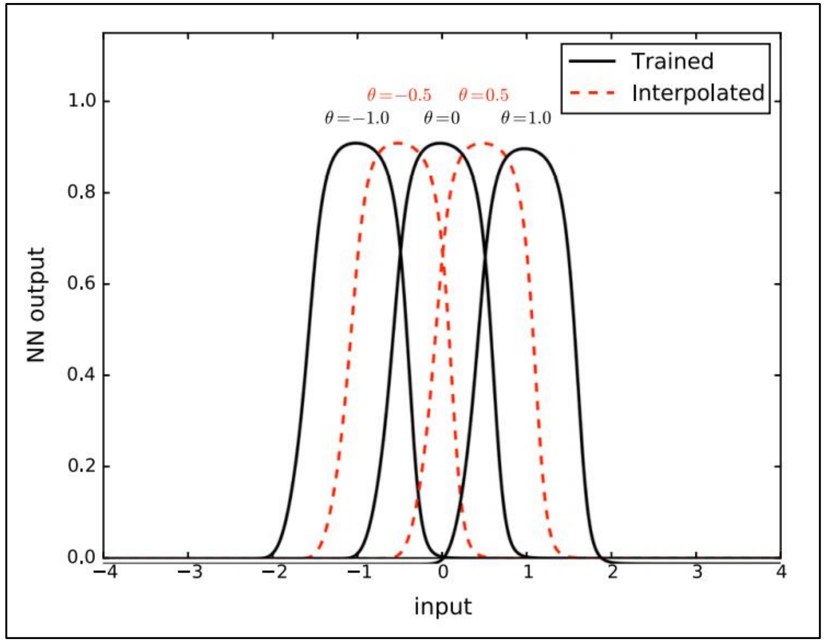
$e\tau$

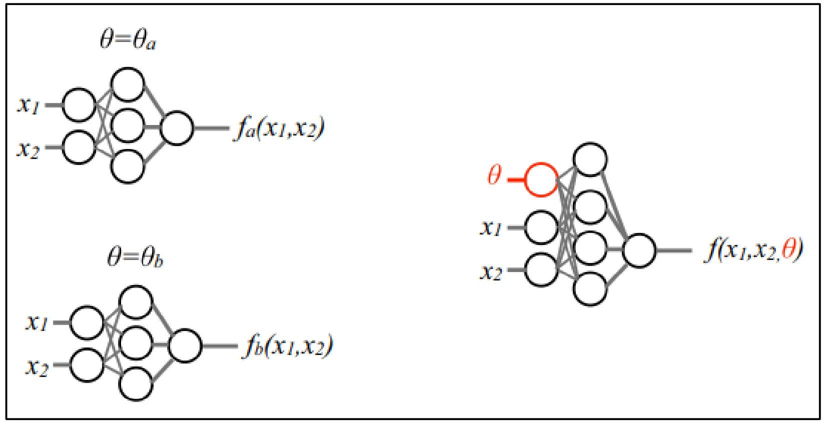
	up	down
leptonSF_MU_SF_Isol_STAT	**0.35%**	**-10.21%**
bTagSF_DL1r_85_eigenvars_B	**6.75%**	**-4.66%**
GlobalReduction_JET_Pileup_RhoTopology	**-2.79%**	**-2.79%**
GlobalReduction_JET_BJES_Response	**-1.73%**	**-1.73%**
tauSF_EFF_RNNID_SYST	**1.57%**	**-1.57%**
TAUS_TRUEHADTAU_SME_TES_PHYSICSLIST	**1.37%**	**1.37%**
pileup	**-1.39%**	**-1.07%**
GlobalReduction_JET_EffectiveNP_1	**-1.12%**	**-1.12%**
tauSF_EFF_ELEOLR_TOTAL	**1.08%**	**-1.08%**
TAUS_TRUEHADTAU_SME_TES_MODEL_CLOSURE	**1.05%**	**1.05%**
total	**8.38%**	**-12.26%**

$\mu\tau$

- leptonSF_MU_SF_Isol_STAT is a tool issue happening on Loose Muon isolation WP
  - has submitted issue, not used in fitting right now
- still need to consider Top quark theoretical uncertainties
- for signals and other backgrounds, see backup slides

MVA optimization: pDNN to Boost the Sensitivity

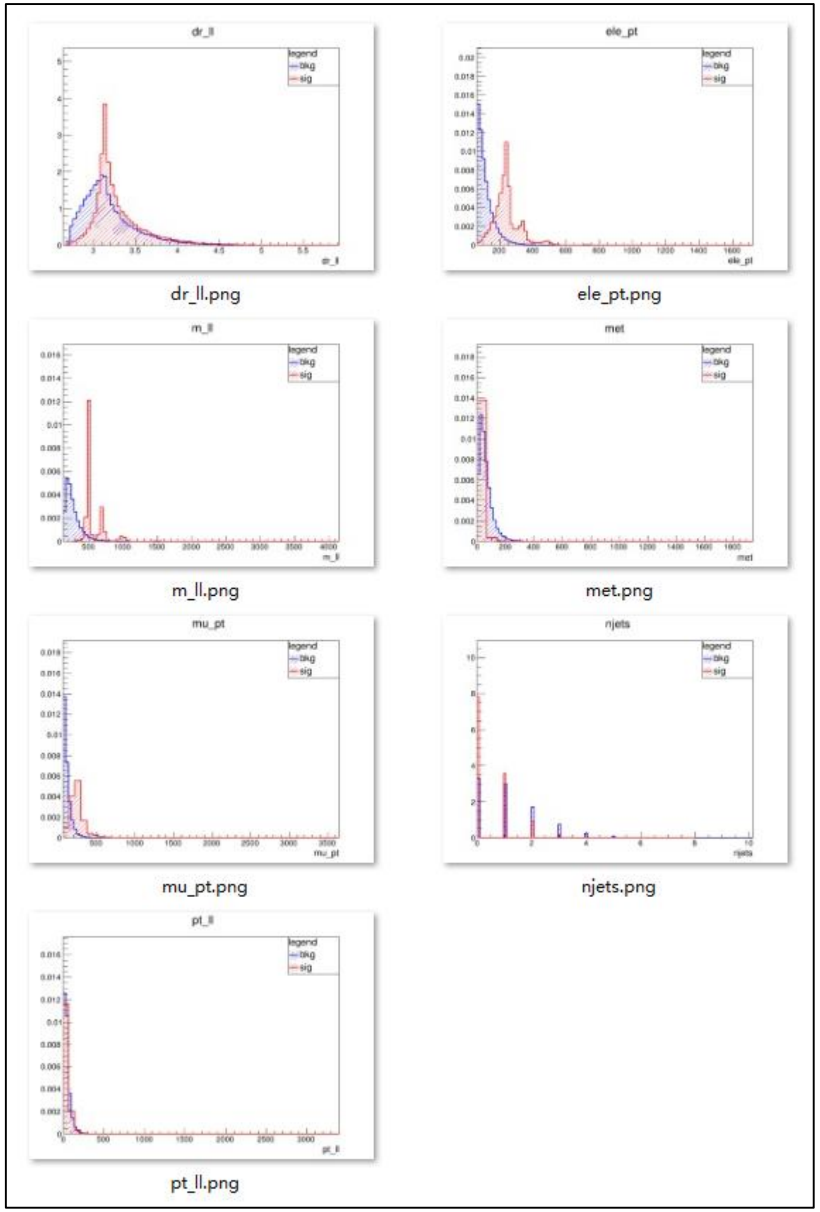




- Use parameterized DNN (pDNN)
  - instead of training different NN classifier at different mass points we can combine all mass points samples together with
    - merge different mass points signal samples together with reasonable weight and normalize
    - certain input variable randomized to have same distribution between signals and backgrounds
  - expect the classifier will learn the trick between mass points
    - distance between mass points shouldn’t be too large, otherwise will expect performance drop

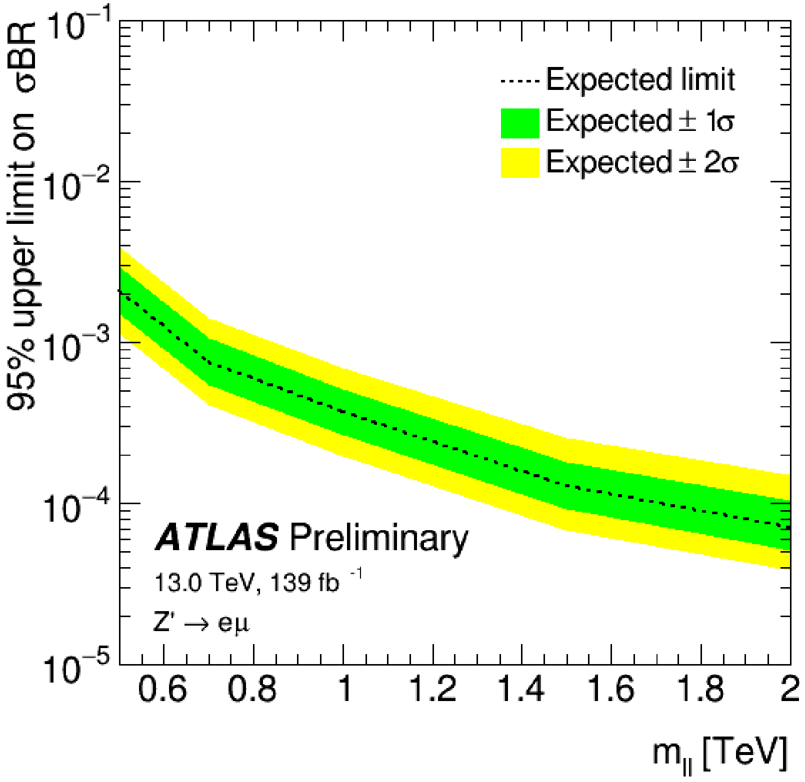
MVA optimization: training set up

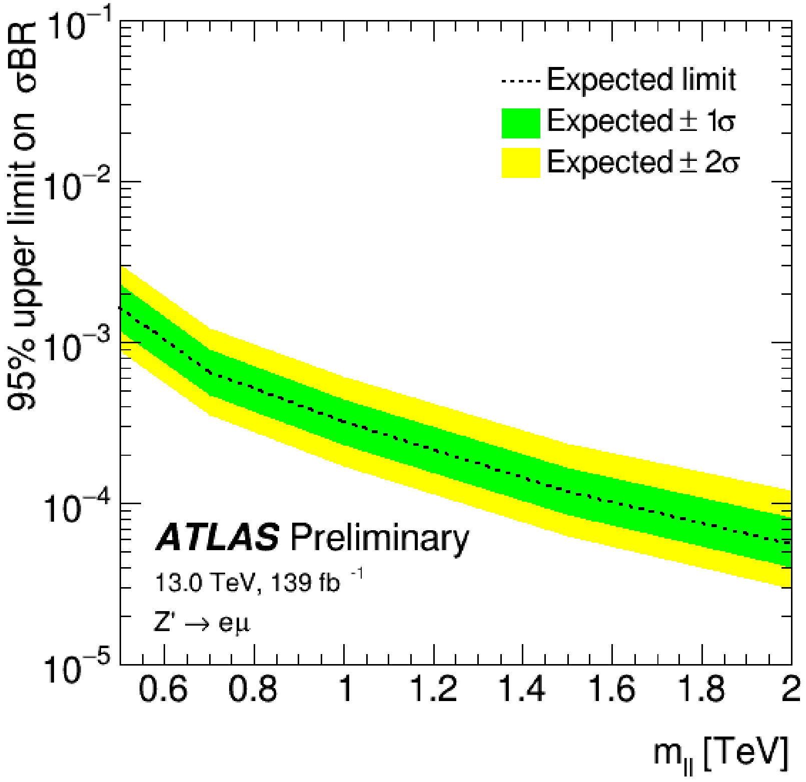
- Signal models
  - currently checking Zprime model, will check RPV & QBH
- Training variables
  - single lepton vars: p_T
  - lepton pair vars: $m, p_T, \Delta R$
  - other vars: $E^T_{missing}, n_{jets}$
- DNN setups
  - use Keras framework with TensorFlow backend
  - each mass point normed to unit and then merged to train against background
  - input variables normalized with mean & standard error
  - training-testing-validation ratio: 0.6-0.2-0.2
  - 5 hidden layers & 128 nodes per layer
  - hyperparameter: manual tune + Bayesian optimization
- Cut value
  - cut at 0.5 for now, need to check performance with different cut value



Statistical analysis: results

- fit m_{ll}
- cut $dnn > 0.5$ and fit m_{ll}





	fit m_{ll}	cut $dnn > 0.5$ fit m_{ll}
	xsec*BR	xsec*BR
500 GeV	2.10e-3	1.64e-3
700 GeV	7.55e-4	6.53e-4
1000 GeV	3.66e-4	3.17e-4
1500 GeV	1.28e-4	1.17e-4
2000 GeV	7.11e-5	5.50e-5

- compare limits with/without DNN cut, DNN helps improve about 10~20% in different mass point

TopQuark sample (extrapolation)

E_pt > 65 GeV ;
| E_eta| < 2.47 and can't at 1.37 < |E_eta| < 1.52 ;
E_isElTight = 1.0 ;
E_isolation_FixedCutTight = true;
|E_delta_z0_sintheta| < 0.5 and |E_d0sig| < 5.0 ;

Mu_pt > 65 GeV;
|Mu_eta| < 2.5;
Mu_isHighPt = true;
Mu_isolation_FixedCutLoose = true;
|Mu_delta_z0_sintheta| < 0.5 and |Mu_d0sig| < 3.0;

Tau_pt > 65GeV;
|Tau_eta| < 2.47 and can't at range 1.37<|Tau_eta|< 1.52;
|Tau_charge|= 1.0;
Tau_nTracks = 1.0 or 3.0;
Tau ID =RNN Medium.

130 GeV < Invariant Mass;
Final events have one good e or mu and one good tau;
Charge are contrary;
DeltaPhi >2.7;

TopQuark signal region

Take etau channel SR for an example:
Propagator.Mag() >130 GeV;
NumberGoodElectrons = 1;
NumberGoodTaus = 1;
NumberGoodMuons = 0;
Charge_ele*Charge_tau == -1.0;
fabs(DeltaPhi) >2.7;

Fit function

Dijet:

TMath::Exp(-1.0*[0])*TMath::Power(x,[1])*TMath::Power(x,[2]*TMath::Log(x))

$$e^{-a}x^{b}x^{c\ln x}$$

Monomial:

[0]/pow(x+[1],[2])

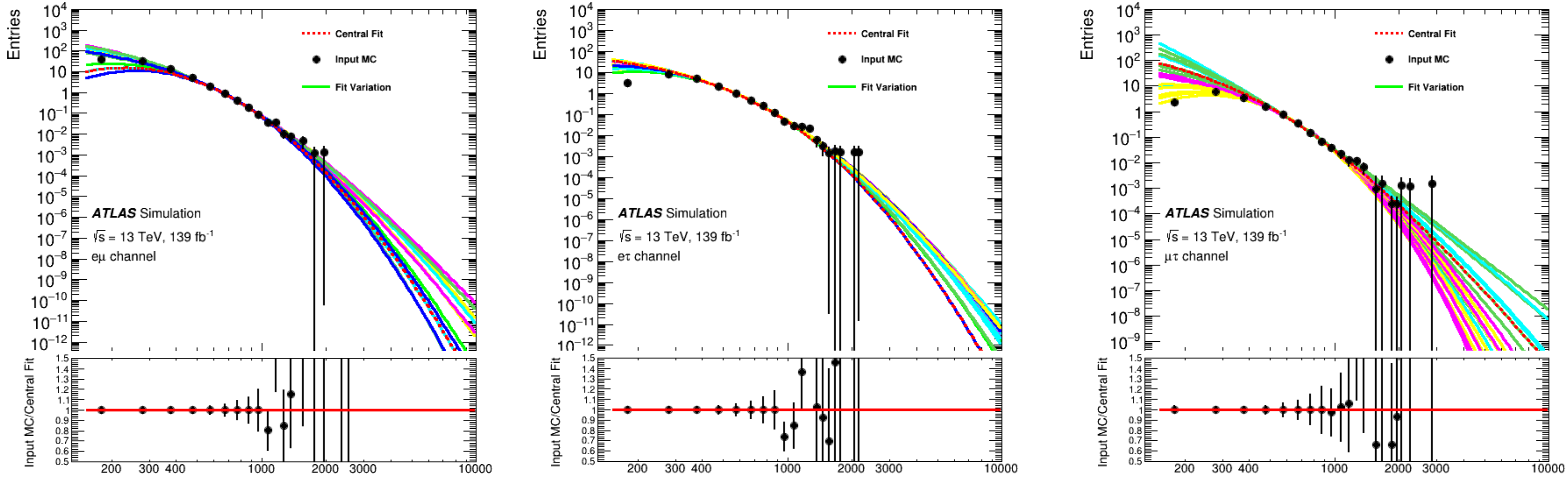
$$\frac{a}{(x+b)^{c}}$$

	emu	etau	mutau
Lower range(step)	450-550(20)[GeV]	300-450(20)[GeV]	450-550(20)[GeV]
Upper range(step)	1000-1200(25)[GeV]	1000-1500(25)[GeV]	1000-1500(25)[GeV]
Stitch point	1000 GeV	900GeV	900GeV

Take emu as an example:450-1000, 450-1025...450-1200, ... 470-1000, ... 450-1200

In total have 6*9=54 and have 2 functions, so will have 54*2= 108 fit.

Fit range variation

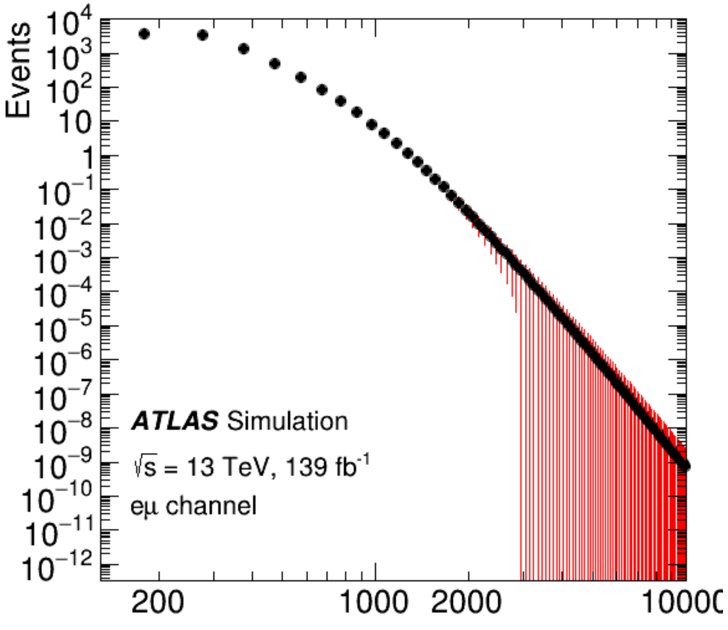


Central fit is the fit result which have the best chi2(minimum value).

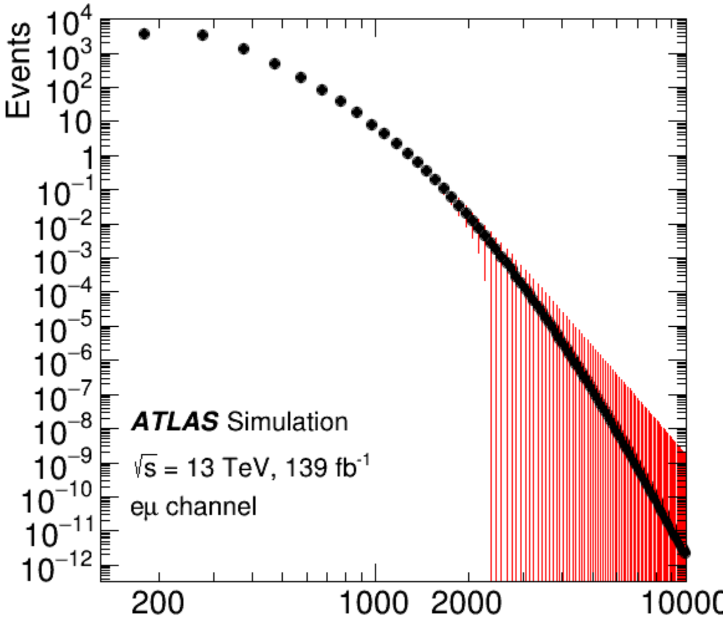
	emu	etau	mutau
Total fit	108	378	252
Converged fit	92	303	252
Percent	85%	80%	100%

Can be used to check fit stability

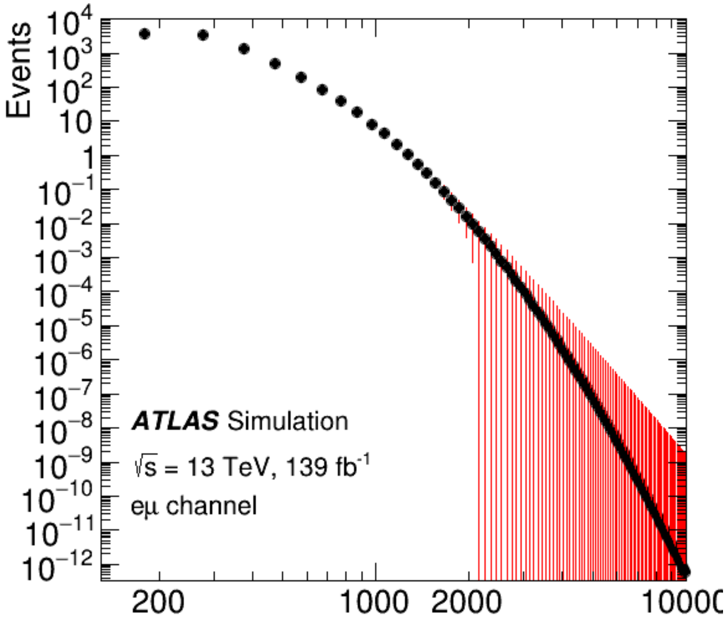
Final fit result in emu channel



mean



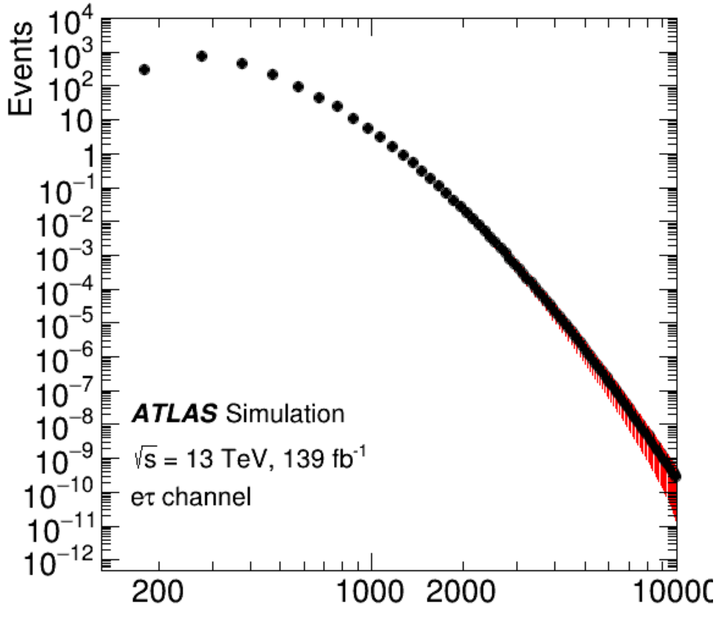
median



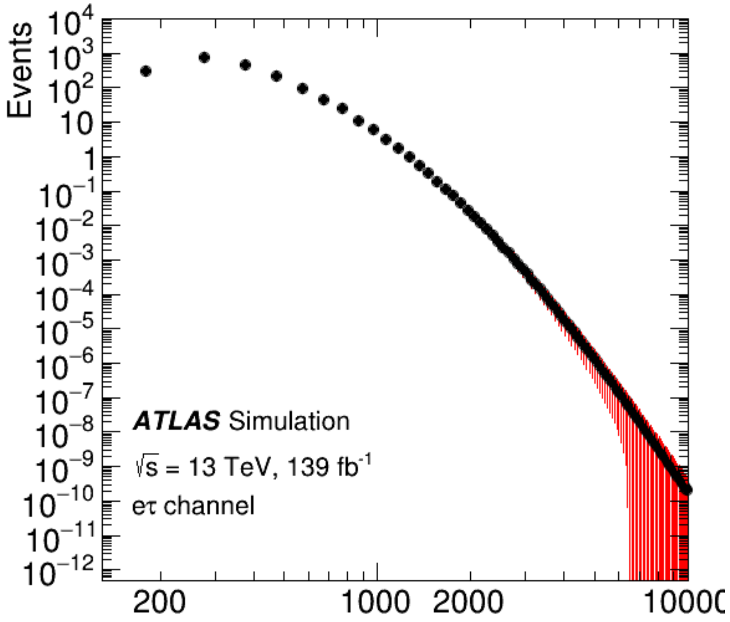
best chi2

The final fit above stitch point which is determined to use according to mean, median or best chi2.

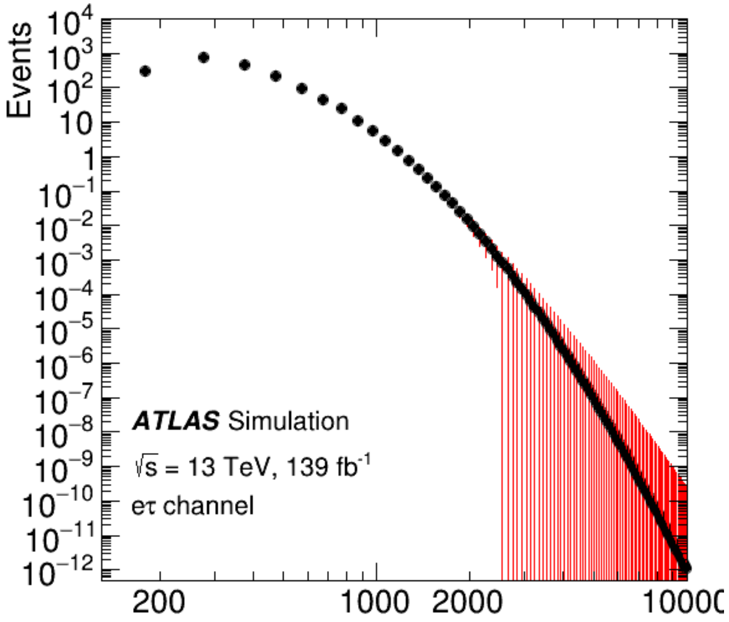
Final fit result in etau channel



mean

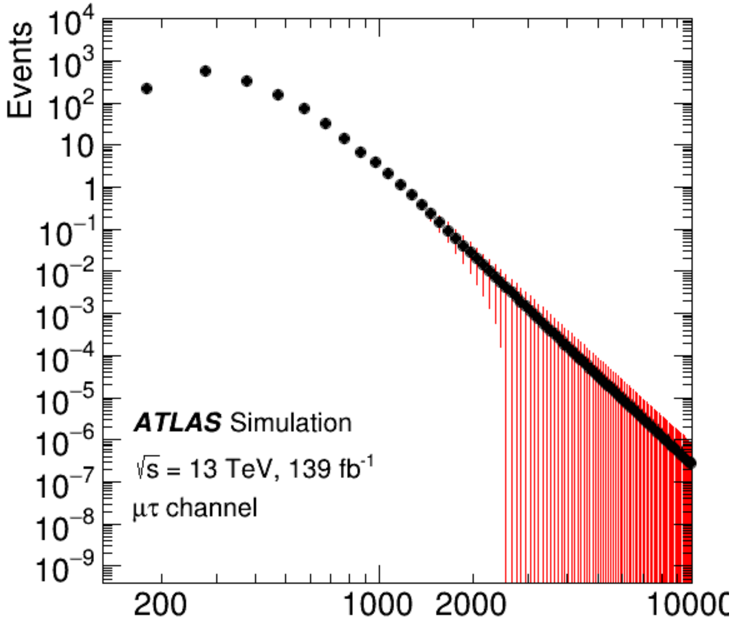


median

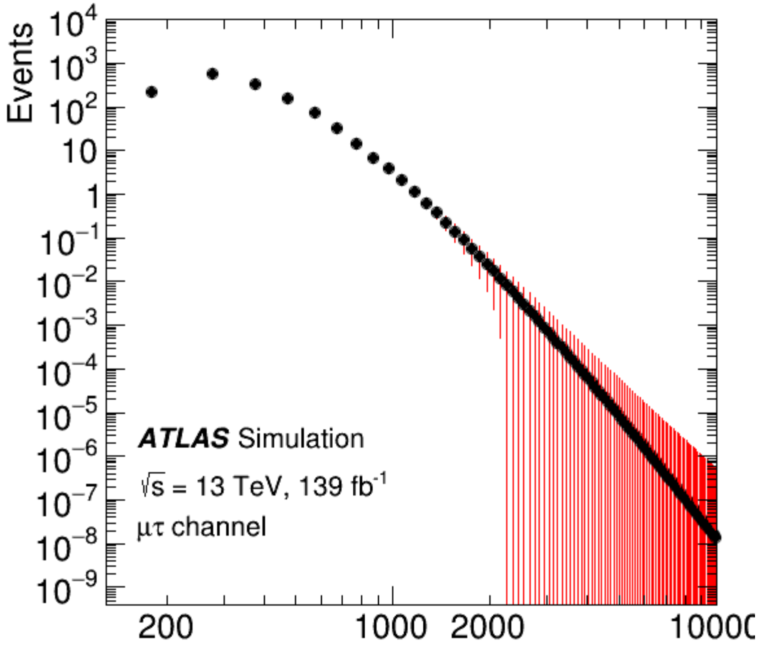


best chi2

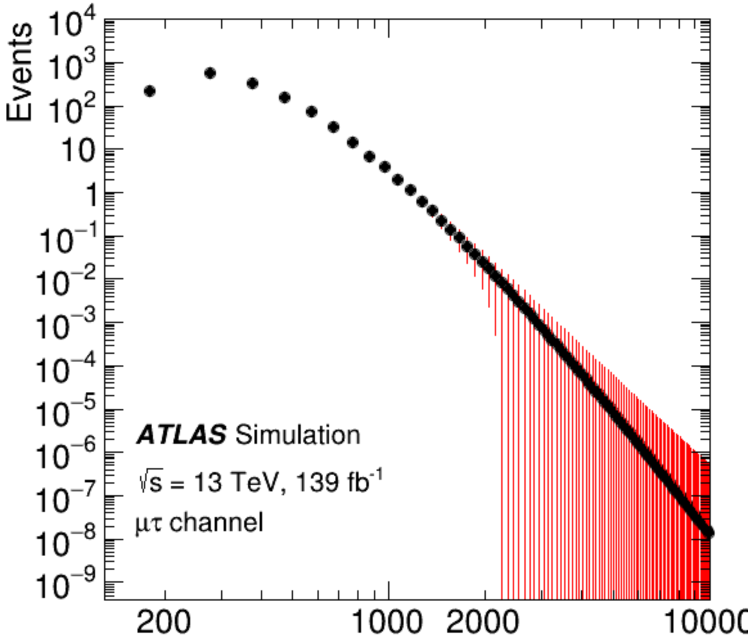
Final fit result in mutau channel



mean



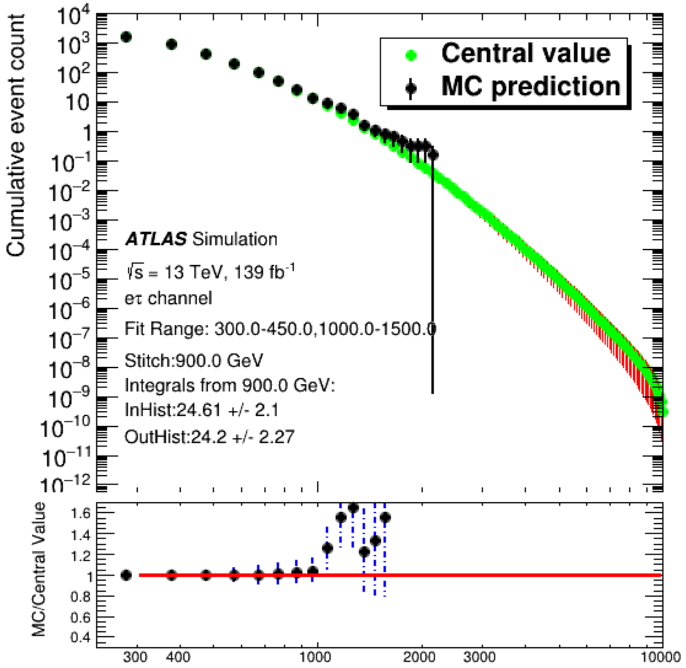
median



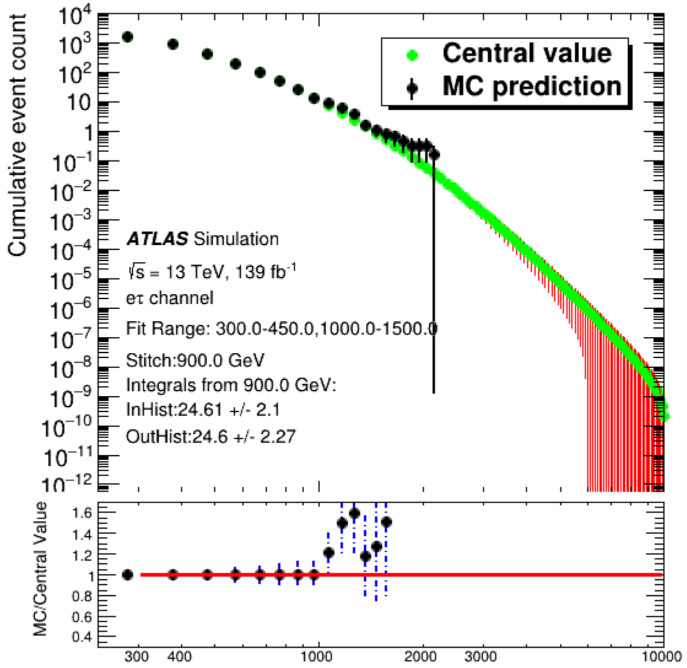
best chi2

In all three channels, the final fit result is similar no matter to use mean value, median value or best chi2 to determine it.

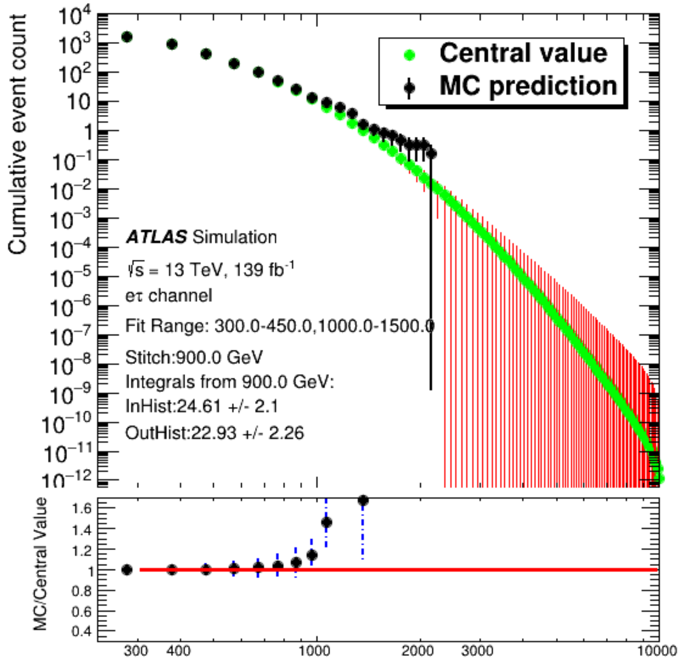
Cumulative integral bin-by-bin in etau channel



mean

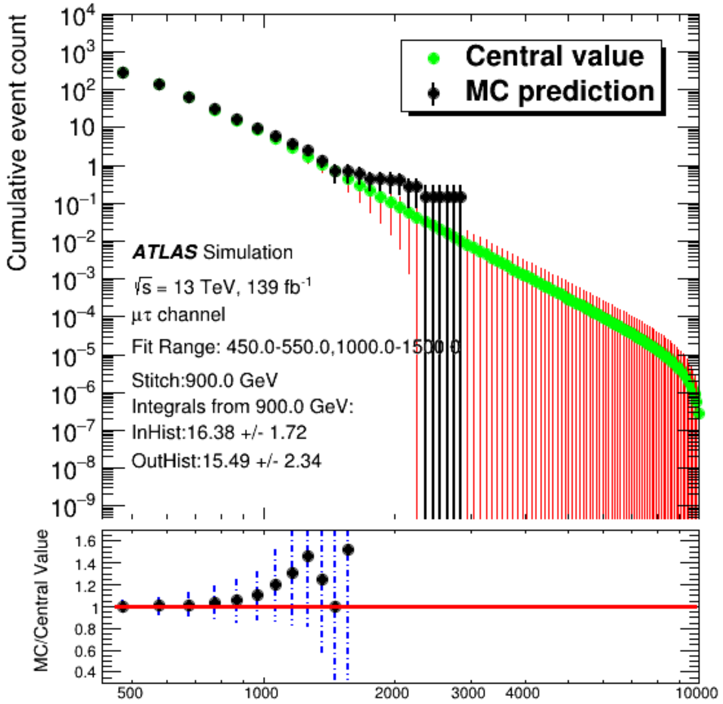


median

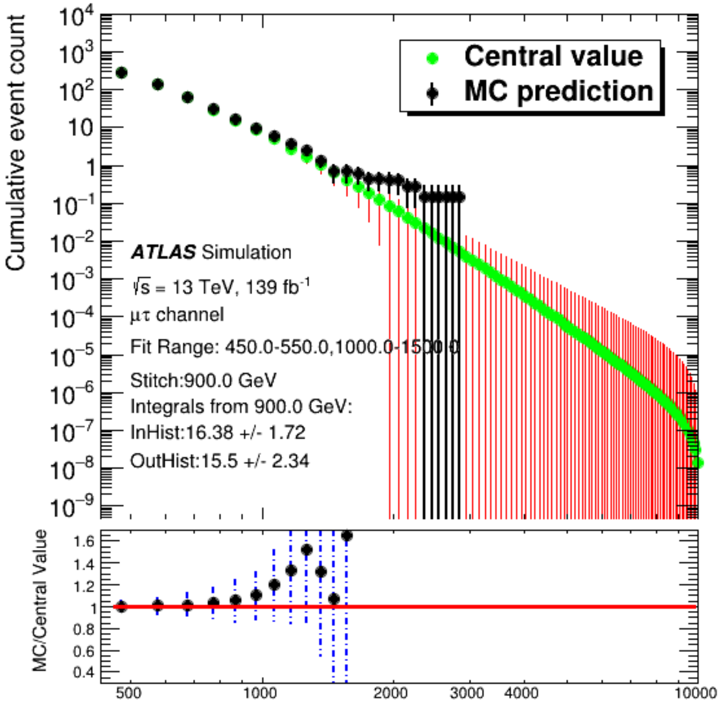


best chi2

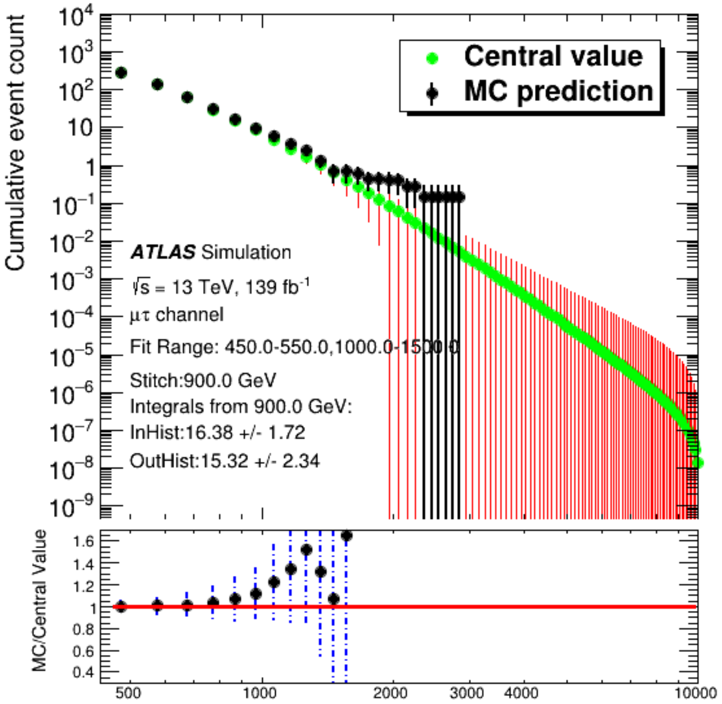
Cumulative integral bin-by-bin in mutau channel



mean



median



best chi2

Background estimation: Wjets in $\boldsymbol{e\tau/\mu\tau}$ channel

m_T cut decision in Wjets CR

- transverse mass cut reduce similar percentage (roughly 50%) of events in different backgrounds
  - not helping enrich Wjets events or reduce other bkg (mainly Zll)
- decision: remove m_T cut

	no $m_T(l_{lead}, E^T_{missing}) > 80GeV$				with $m_T(l_{lead}, E^T_{missing}) > 80GeV$			
	CR (no tau ID)		CR (with tau ID)		CR (no tau ID)		CR (with tau ID)	
	$e\tau$	$\mu\tau$	$e\tau$	$\mu\tau$	$e\tau$	$\mu\tau$	$e\tau$	$\mu\tau$
Data	949724	715972	14638	6979	440913	347861	8553	4063
MC	823622.9	660736.4	14419.2	7103.8	389638	328893.4	8284.1	4006.8
Wjets	658321.3	531238.5	5392.7	3821.1	316040.1	263150.7	2977.6	1952.9
Top	54730.2	42408	2069	1145.9	30182.2	23704.8	1367.4	763.9
Zll	82903.9	65534.1	5370.1	1122.7	29211.1	30740	2856.3	577.3
di-boson	16338.8	12775.8	1170.6	760.3	8336.1	6680.8	801.1	536.8
singleTop	11328.6	8780	416.8	253.9	5868.6	4617.1	281.7	175.8