Offline TauID and TES Measurements on ATLAS

Shiyi Liang, Yuchen Cai (IHEP)

CLHCP2022

Nov 24, 2022

ATLAS EXPERIMENT

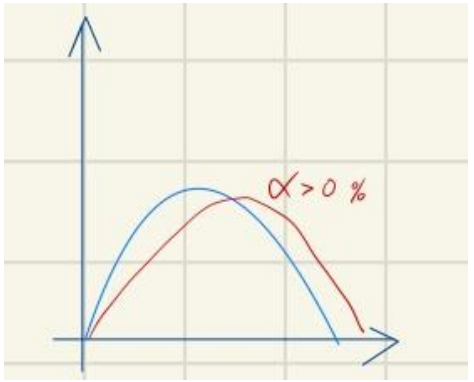
中国科学院高能物理研究所
Institute of High Energy Physics
Chinese Academy of Sciences

■ Motivation

- Correct offline tau p_T (In-situ τ energy scale) & tau ID efficiency (τ ID scale factor) in MC.
- Target @ ATLAS R22 MC20 (for Run-2 reprocessing samples) and MC21 (Run-3 samples)
- **tag & probe** method: $Z \rightarrow \tau_{had} + \tau_{lep} \rightarrow \boldsymbol{\tau_{had} + \mu/e}$
- Dataset: *2018, 58.5* fb^{-1}

■ **In-situ τ energy scale**:

- $p_T' = (1 + \alpha) \cdot p_T$. And α enters the $m_{vis}(\tau, \mu)$ distribution.
- Calculate bin-wise χ^2 sum and minimize it.



■ **τ ID scale factor**:

- $SF = \frac{\epsilon_{Data}}{\epsilon_{MC}}$, where $\epsilon_{ID} = \frac{N_{The\ Truth\ Taus\ pass\ ID}}{N_{Some\ Truth\ Taus}}$
- ϵ_{Data} : Use likelihood fit to estimate true-tau contribution in data.
- ϵ_{MC} : Take MC predicted yields

$$\epsilon(\mathrm{L}) = \frac{LNM+MNT+T}{NL+LNM+MNT+T}$$
$$\epsilon(\mathrm{M}) = \frac{MNT+T}{NL+LNM+MNT+T}$$
$$\epsilon(\mathrm{T}) = \frac{T}{NL+LNM+MNT+T}$$

in-situ TES

■ Strategy:

- In the fit, a scale factor is applied on the $Z \rightarrow \tau_{\mu}\, \tau_{had}$ process.
  - $\chi^2_{(\alpha,f)} = \Sigma_{bins} \frac{(N_{data} - N_{bkg} - f \cdot N_{sig})^2}{\sigma(data)^2 + \sigma(bkg)^2 + f^2 \cdot \sigma(sig)^2}$ (f: float factor)

- **Strategy**:

1-/3-prong Central/Forward Region

SR
μ isolation (Fixed Rad)
OS(τ,μ)
b-jet Veto
$m_T(\mu, xe) < 60$ GeV
$\boldsymbol{\Sigma\, cos(\Delta\phi) > -0.20}$
$\boldsymbol{m_{vis}(\mu, \tau)}$

W scale factor

1-/3-prong Same-/Opposite-Sign

WCR
Muon isolation
b-jet Veto
$m_T(\mu, xe) > 60\ GeV$
$E_T^{miss} > 30\ GeV$
$\Sigma \cos(\Delta\phi) < 0$
$m_{vis}(\mu, \tau)\ in\ (40,90)\ GeV$

W scale factor

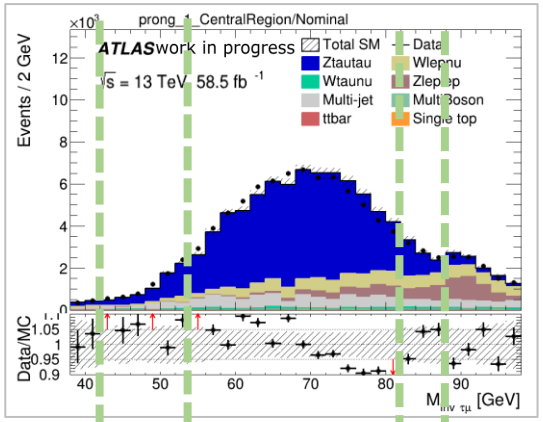
1-/3-prong

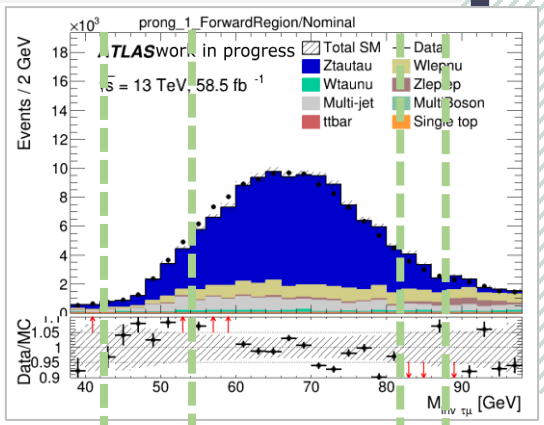
QCD-CR
Not μ isolation
b-jet Veto
$m_{vis}(\mu, \tau)\ in\ (40,90)\ GeV$

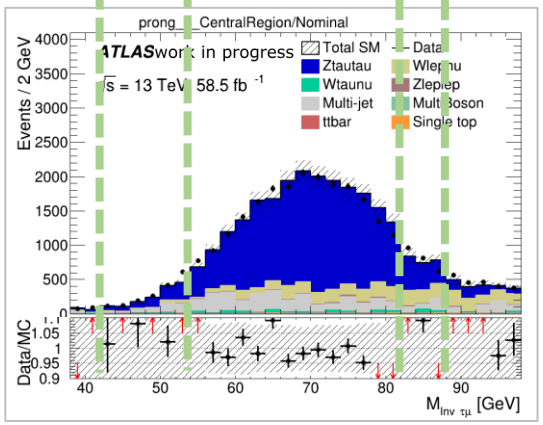
SS-OS method – Transfer factor

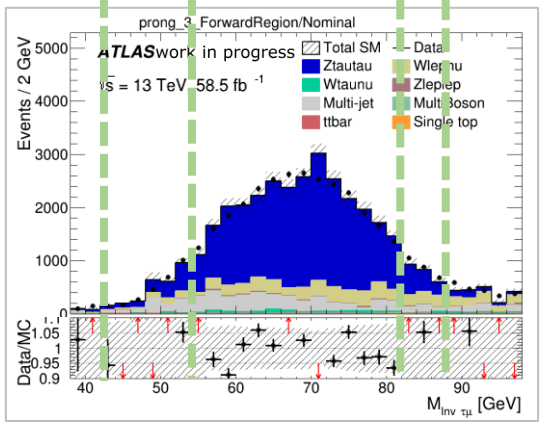
Fit range & Bin width optimization

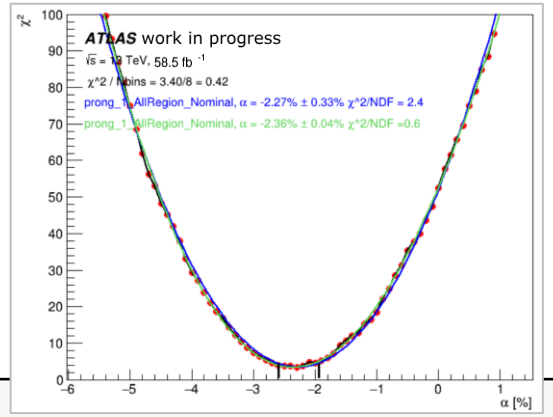
- Add an optimization step and drop the uncertainty.
- Try all the combinations of fit range and bin width:
  - lower range: [42, 54] GeV.
  - Upper range: [82, 88] GeV.
  - Bin width: 4 GeV, 6 GeV, 8 GeV, 10 GeV.
- Select the combination which satisfies:
  - $\frac{\Delta\chi^2_{\text{hist}}}{N_{\text{bins}}} < 0.8$: post fit distributions are within the stat-uncertainty.
  - $\frac{\Delta\chi^2_{\text{parabola}}}{\text{NDF}} < 6$: the $\chi^2 - \alpha$ curve is smooth.
  - Select the intersection of Central and Forward region.

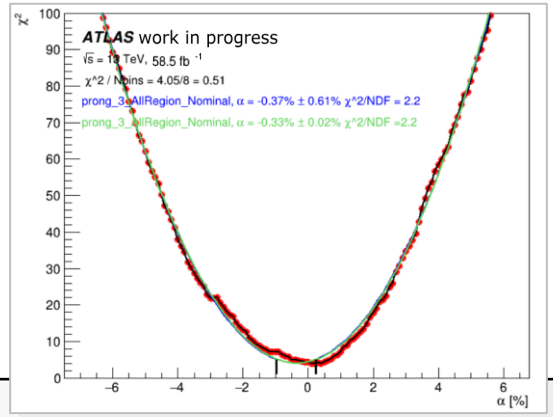




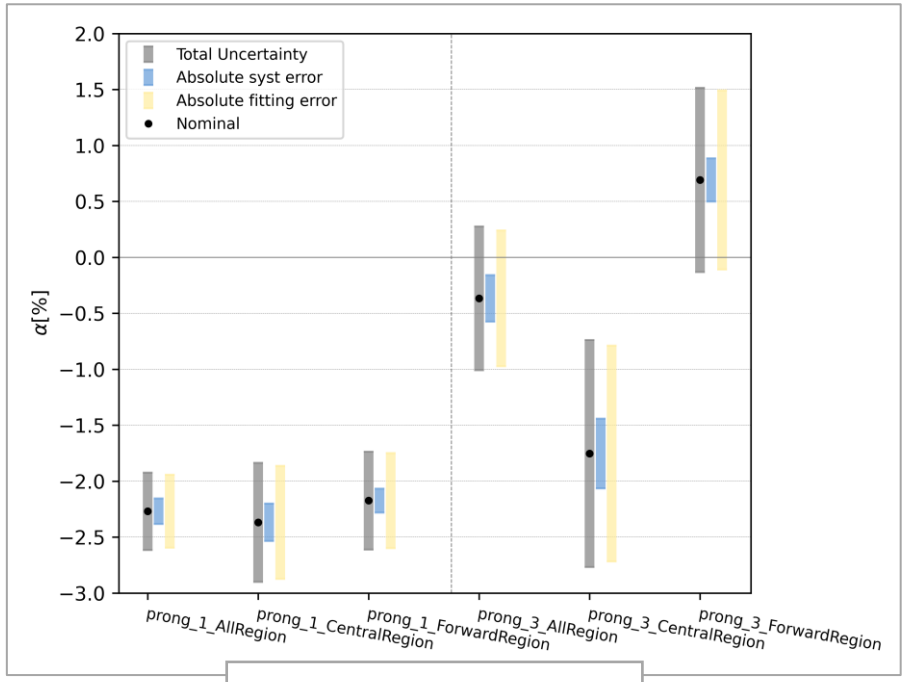




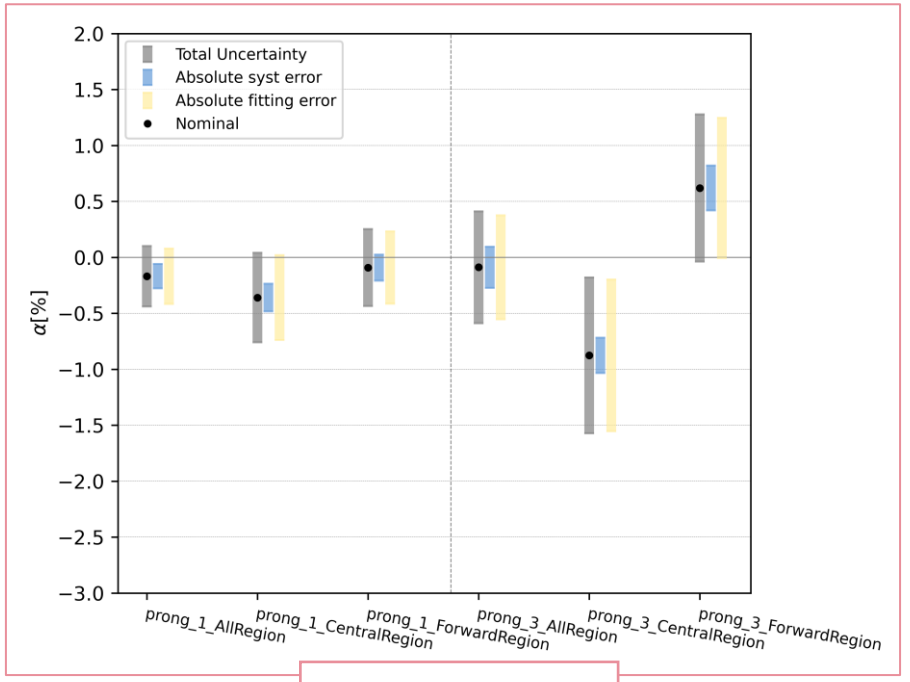




TES results



Sherpa2211

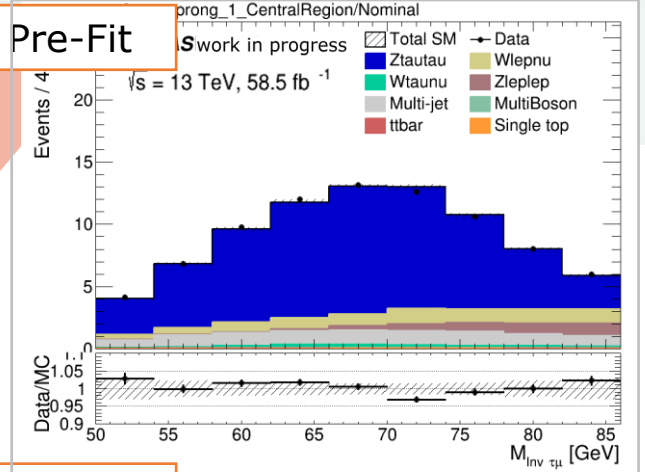


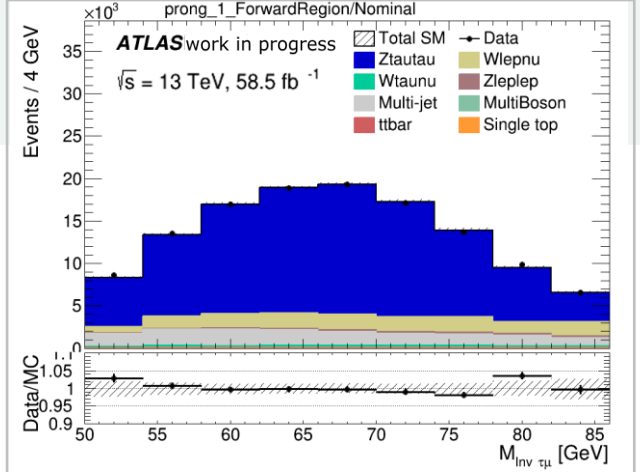
PoPy8

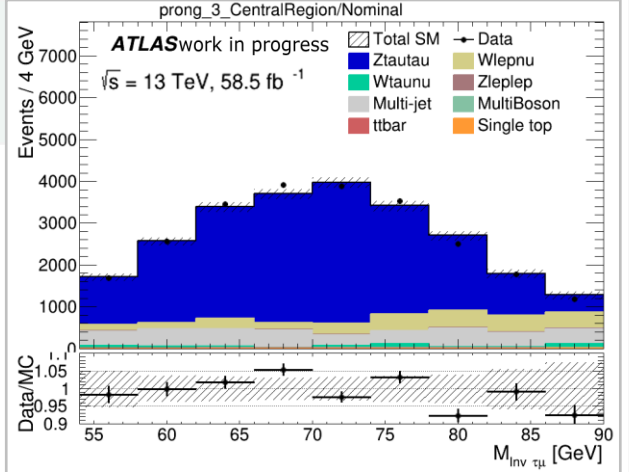
- ~~**Sherpa2211** (Grey)~~ and **PoPy8** (Red)
  - We drop Sh2211 study temporarily due to a bug in Sh2211 Ztautau sample. But we will back to it in the future.
  - Dominant uncertainty is from the chi-2 method.

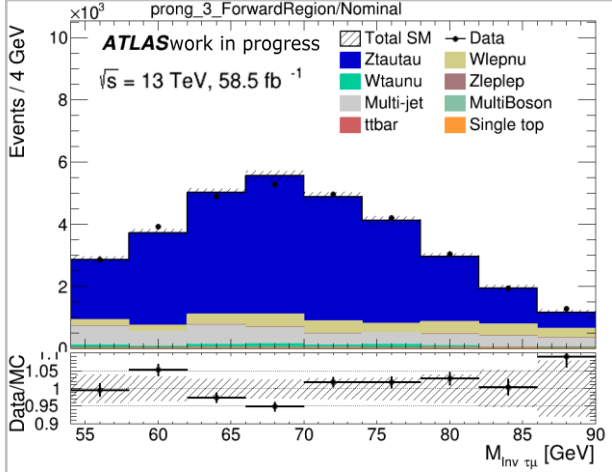
Pre-Fit

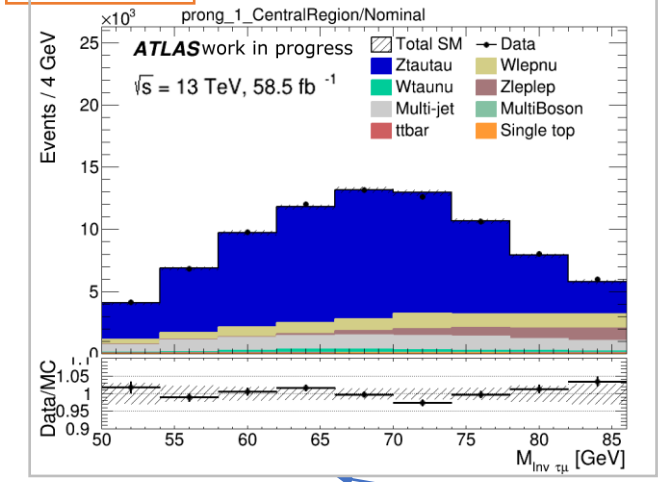
Post-Fit

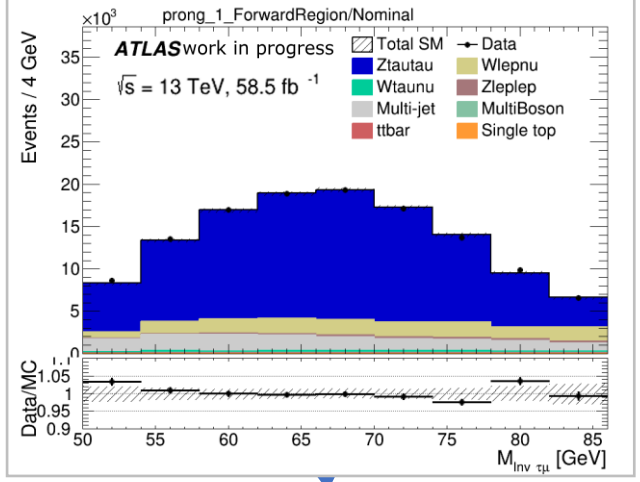


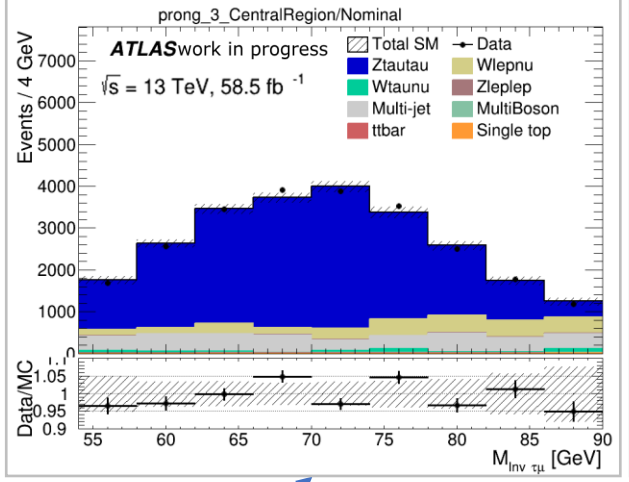


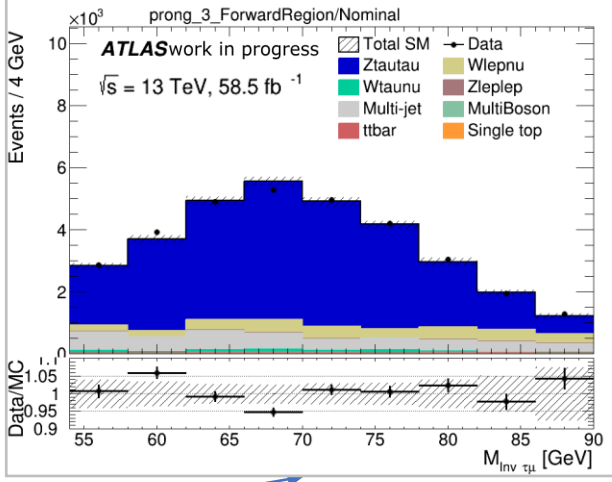


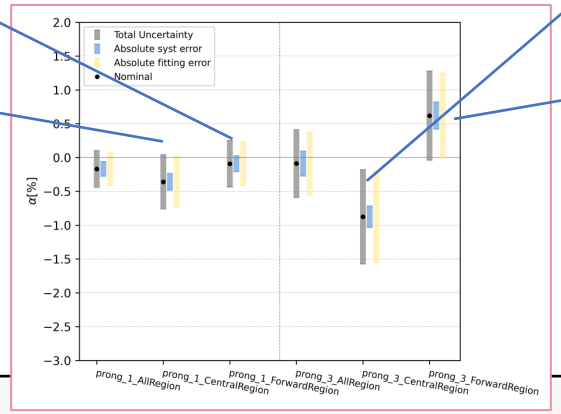








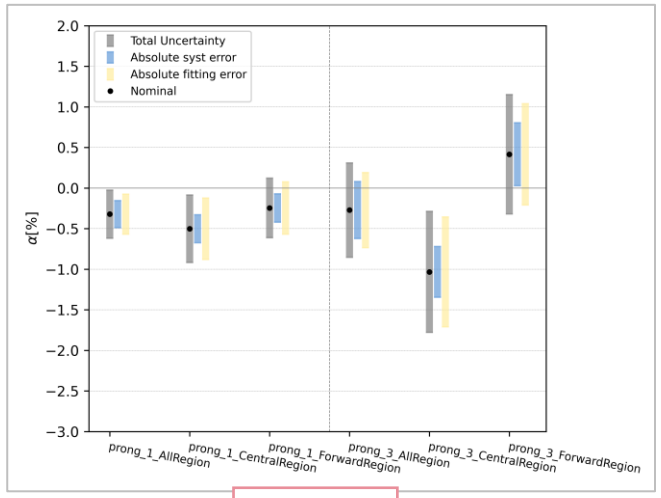




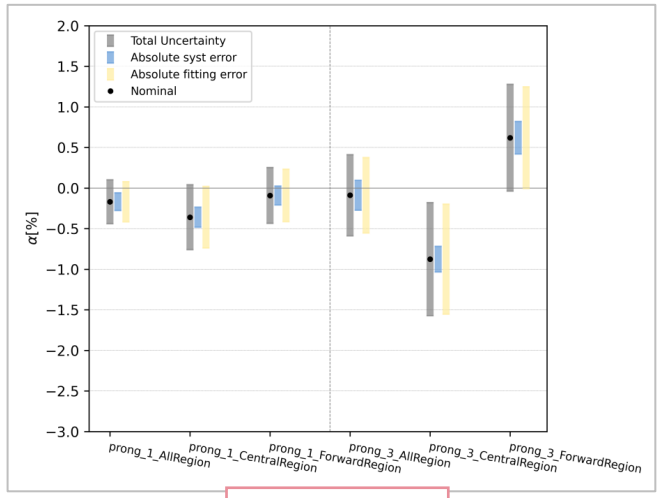
- **PoPy8** pre-/post- fit distributions

tau ID check:

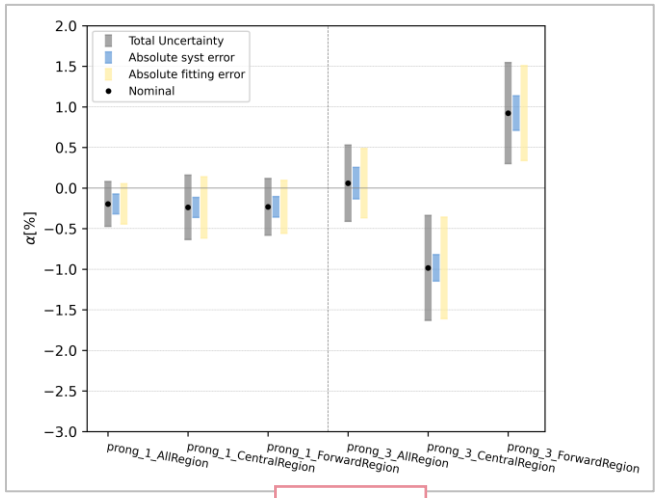
PoPy8



Loose



Medium

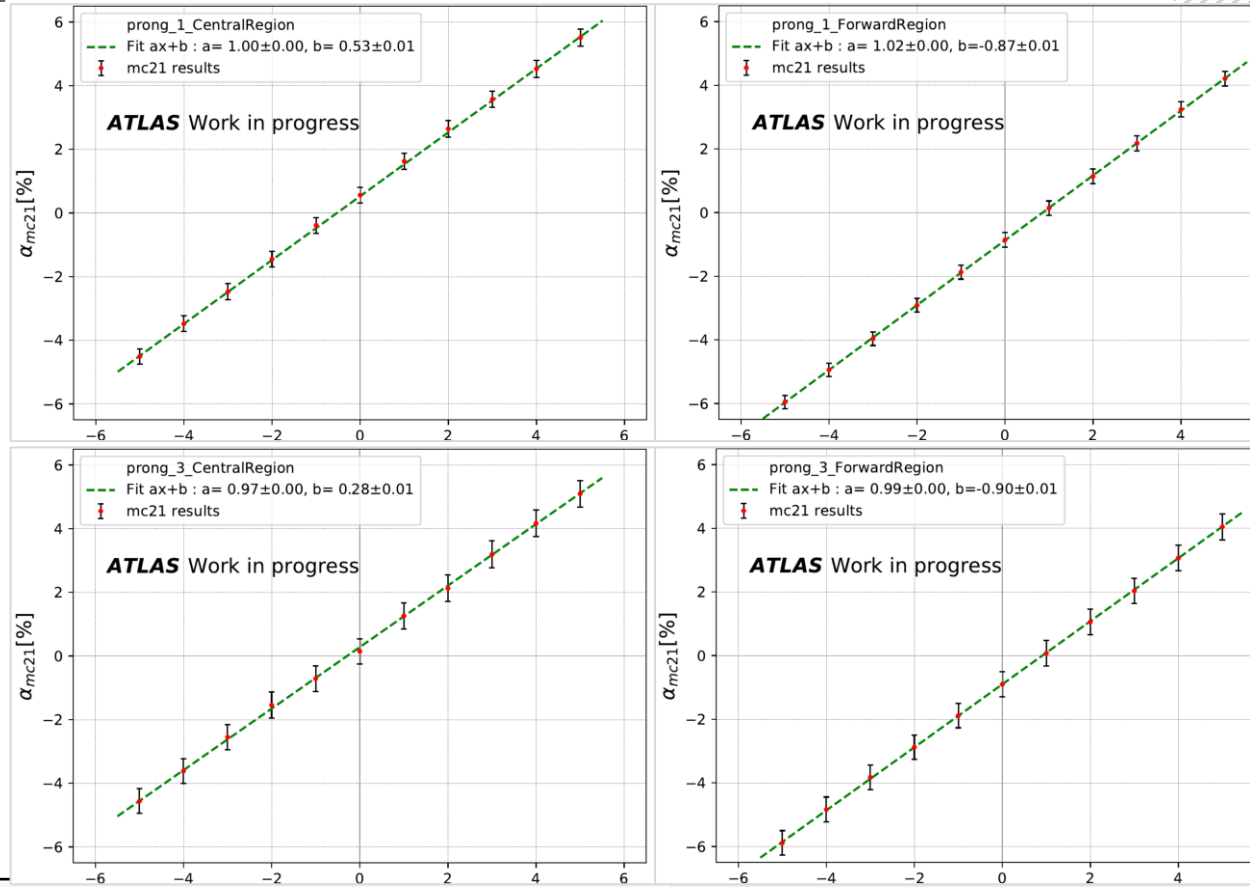


Tight

- Nearly the same.
- Did not consider it as uncertainty.

Run-3 pre-recommendation

- Fit MC to MC to speed up the measurement with limited available data.
  - Full measurement: $\chi^2_{(\alpha,f)} = \Sigma_{bins} \frac{\left(N_{data} - N_{bkg} - f \cdot N_{sig}\right)^2}{\sigma(data)^2 + \sigma(bkg)^2 + f^2 \cdot \sigma(sig)^2}$ (f: float factor)
  - For Run3 pre-recommendation:
    - $\chi^2_{(\alpha,f)} = \Sigma_{bins} \frac{(N_{MC20\ Ztt} - f \cdot N_{MC21\ Ztt})^2}{\sigma(MC20\ Ztt)^2 + f^2 \cdot \sigma(MC21\ Ztt)^2}$ (f: float factor)
    - Shift Run-2 signals with any α_{mc20}.
    - Obtain the Run-3 signal α_{mc21}.
    - Fit linear function to $(\alpha_{mc20},\ \alpha_{mc21})$.
- Intercepts are not close to zero.



Tau ID Scale Factor

■ Strategy:

- ***e* channel and *μ* channel**
- Input histograms
  - No electron rejection, no muon OLR
  - use RNN score as discriminating variables
  - 4 ID requirements: NL, LNM, MNT, T
  - decay mode: 1 prong, 3prong

$$\epsilon(L)=\frac{LNM+MNT+T}{NL+LNM+MNT+T}$$
$$\epsilon(M)=\frac{MNT+T}{NL+LNM+MNT+T}$$
$$\epsilon(T)=\frac{T}{NL+LNM+MNT+T}$$

➢ Input NFs: 8 NFs

RNN score	NL	LNM	MNT	T
1-prong	0-0.15	0.15-0.25	0.25-0.4	0.4-1
3-prong	0-0.25	0.25-0.4	0.4-0.55	0.55-1

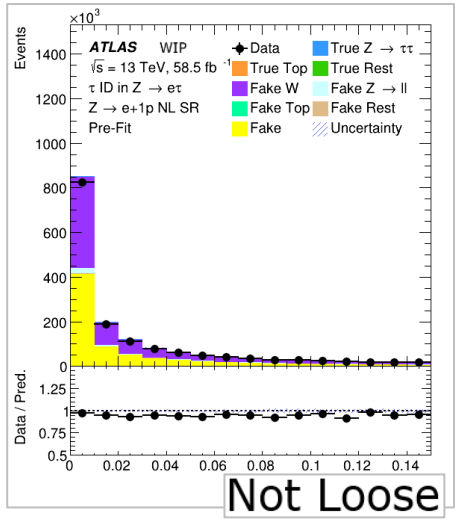
- τ pt slice Region: 20 GeV – 30 GeV, 30 GeV – 40 GeV, > 40 GeV; no τ pt slice (inclusive).

■ Background Estimation: Similar to in-situ TES.

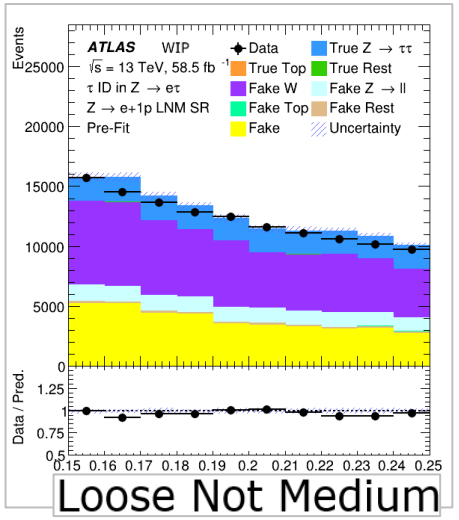
- W+jets: WCRs for SRs and QCD CRs
- QCD: SS-OS method

Pre-fit distributions

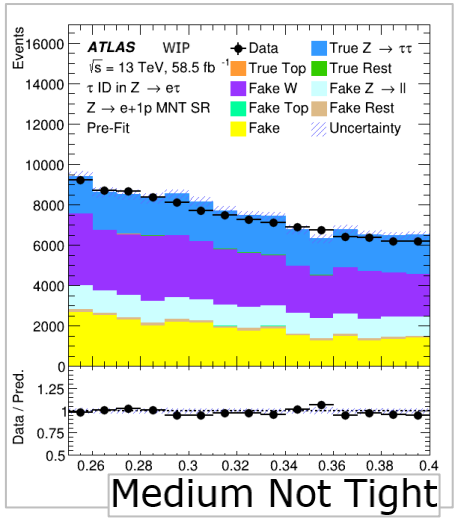
- Pre-fit plots in 1-prong e channel(inclusive)



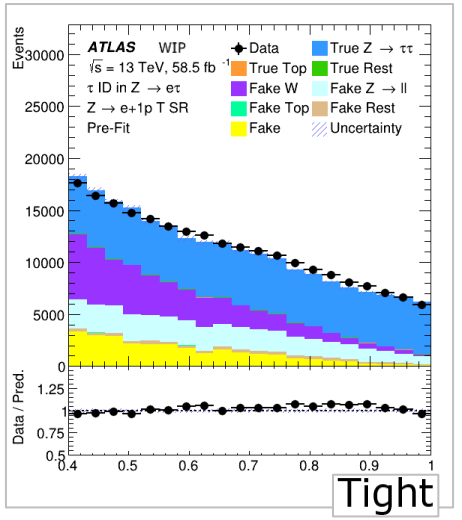
Not Loose



Loose Not Medium

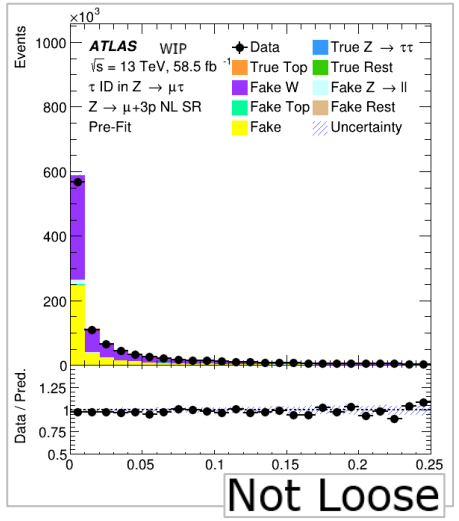


Medium Not Tight

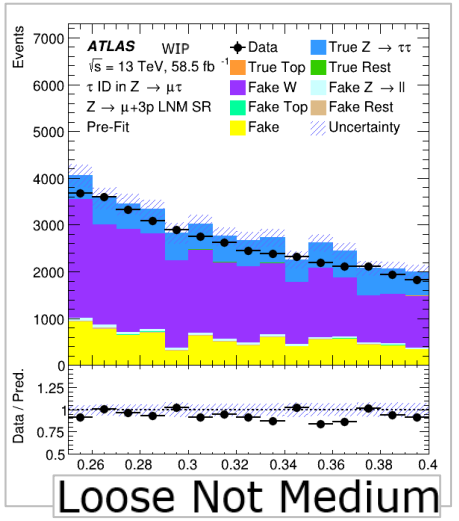


Tight

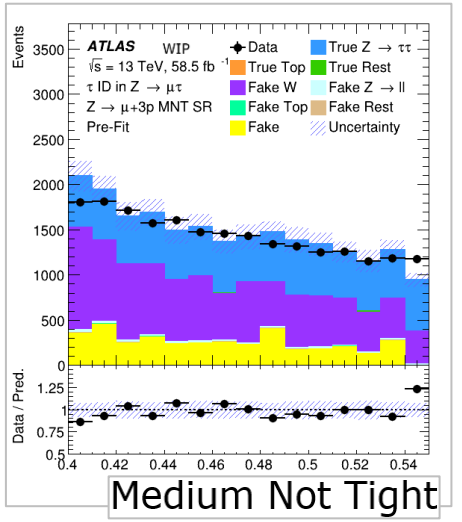
- Pre-fit plots in 3-prong μ channel(inclusive)



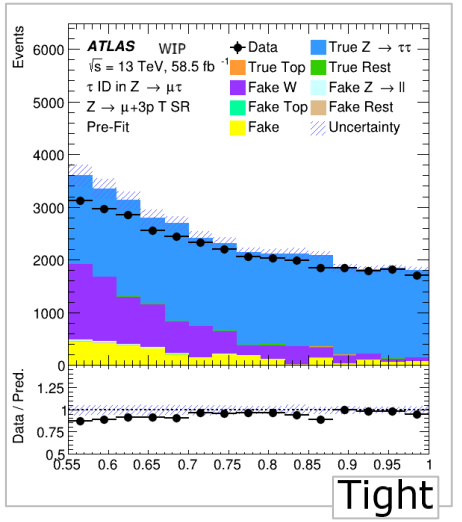
Not Loose



Loose Not Medium

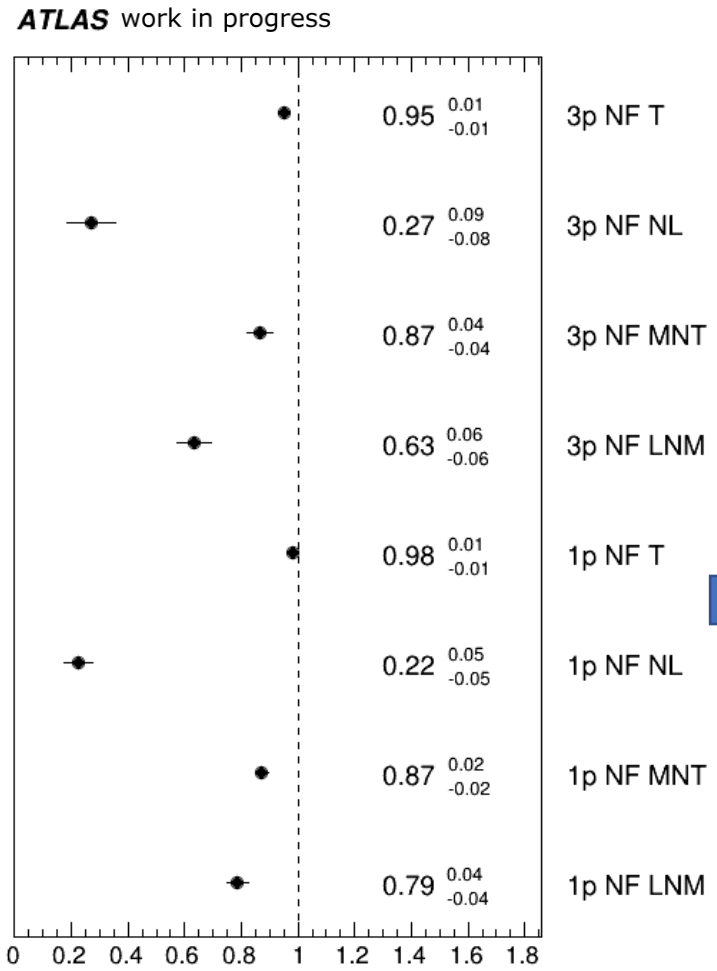


Medium Not Tight



Tight

Fit result (pT inclusive)



- τ ID scale factors are influenced by modelling in the NL region. But the NL NFs are unstable due to very small signal contribution.
- **set NL NF = 1** to calculate τ ID scale factor(assign **50%** uncertainty to fixed-not-loose yields).

NF = 1

	τ ID scale factor(set NL=1)	
	1p	3p
Loose	0.99+-0.09	0.96+-0.13
Medium	1.01+-0.09	1.03+-0.14
Tight	1.04+-0.10	1.05+-0.15

- More reasonable results.

Fit result (binned pT)

(20,30) GeV	**τ ID scale factor**(set NL=1)	
	1p	3p
Loose	0.98+-0.10	0.93+-0.14
Medium	1.02+-0.10	1.04+-0.16
Tight	1.06+-0.10	1.09+-0.17

(30,40) GeV	**τ ID scale factor**(set NL=1)	
	1p	3p
Loose	1.00+-0.08	1.00+-0.12
Medium	0.99+-0.08	1.01+-0.13
Tight	1.00+-0.08	1.00+-0.12

(40,-) GeV	**τ ID scale factor**(set NL=1)	
	1p	3p
Loose	1.00+-0.08	1.02+-0.12
Medium	0.98+-0.08	1.00+-0.12
Tight	0.97+-0.07	0.97+-0.12

- τ ID scale factors are close to one for all τ ID both in 1p and 3p with different pT slices.
- Statistical uncertainty only.
- 3-prong shows more dependence on pT.
  - The central value decreases in the Tight ID category when p_T goes high.
  - All the results are within the uncertainty.

Summary

- We have shown the results of the in-situ TES & tau ID scale factor.
  - Scale factors are close to 1, but the uncertainty needs improvement.
  - Post-fit distributions are fine.
  - Results are stable.
- The ATLAS pre-recommendations for taus have been approved and submitted.

Backup

$e\ \mu$ combine

- Combine e, mu channel to get τ ID scale factors is better?

Likelihood fits results(set NL NF=1)

tau ID SF	e1p	e3p	mu1p	mu3p
Loose	1.00+-0.09	0.96+-0.14	0.98+-0.09	0.97+-0.13
Medium	1.03+-0.09	1.03+-0.15	1.00+-0.10	1.02+-0.14
Tight	1.05+-0.09	1.08+-0.16	1.02+-0.10	1.01+-0.14

	τ ID scale factor(set NL=1)	
	1p	3p
Loose	0.99+-0.09	0.96+-0.13
Medium	1.01+-0.09	1.03+-0.14
Tight	1.04+-0.10	1.05+-0.15

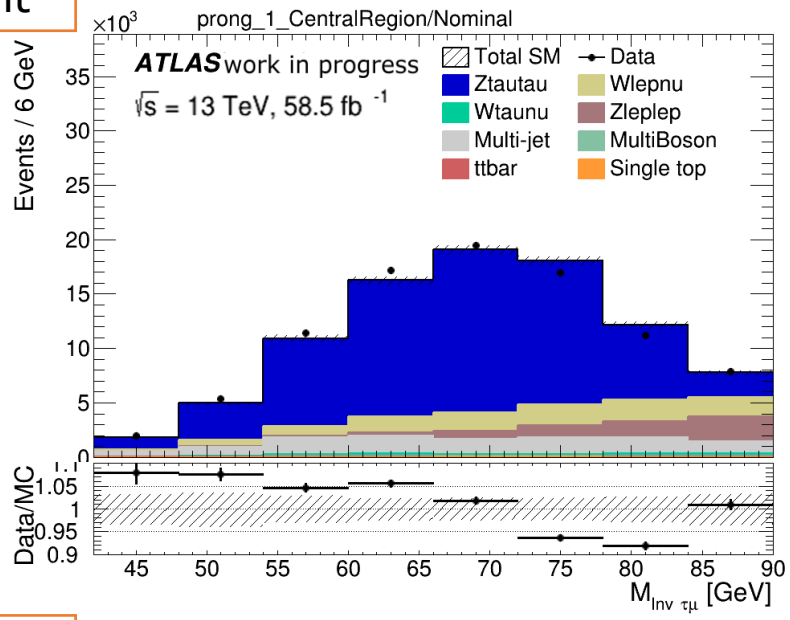
- Separating e and μ channels does not improve the precision of τ ID scale factors.

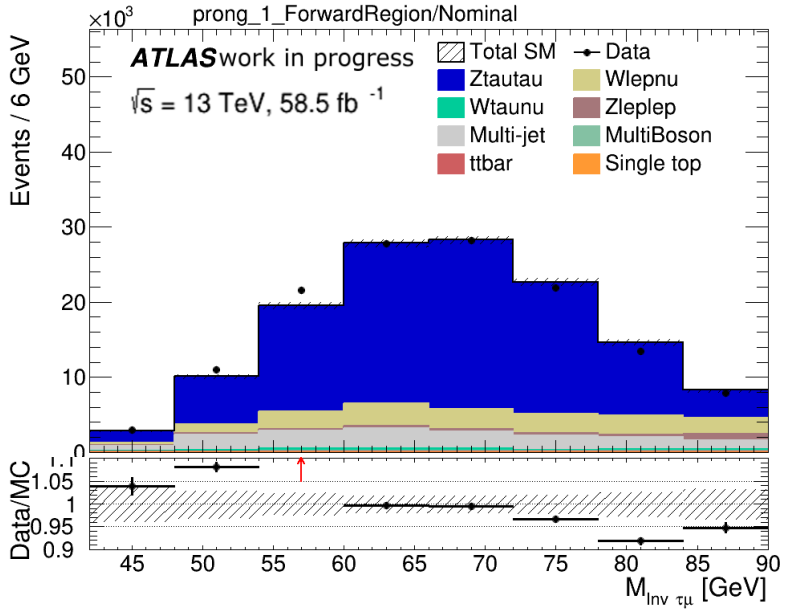
Bin Setting

- New settings:
  - **Sherpa2211**:
    - 1P & 3P: [42,90] GeV, 6 GeV width.
  - **PoPy8**:
    - 1P: [50,86] GeV, 4 GeV width
    - 3P: [54,90] GeV, 4 GeV width
  - **Sherpa221**:
    - 1P & 3P: [52,92] GeV, 8 GeV width.

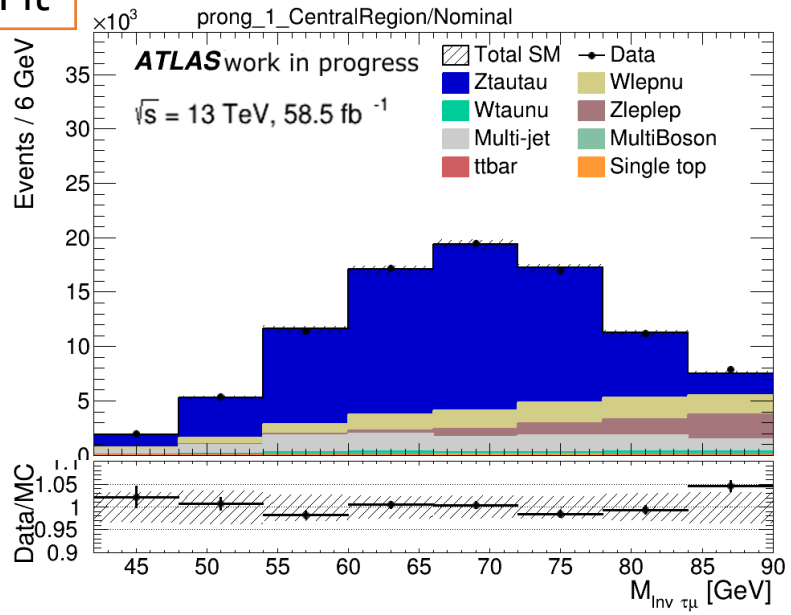
Sherpa2211 1-prong

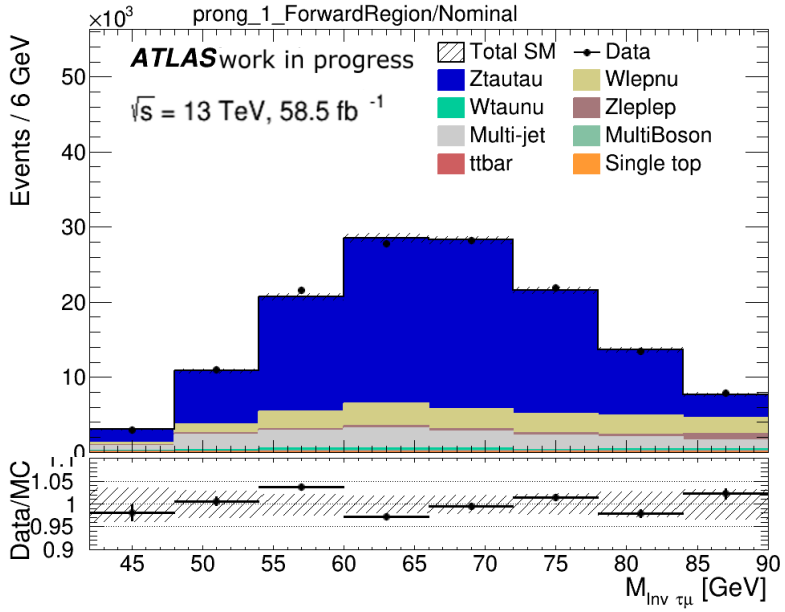
Pre-Fit

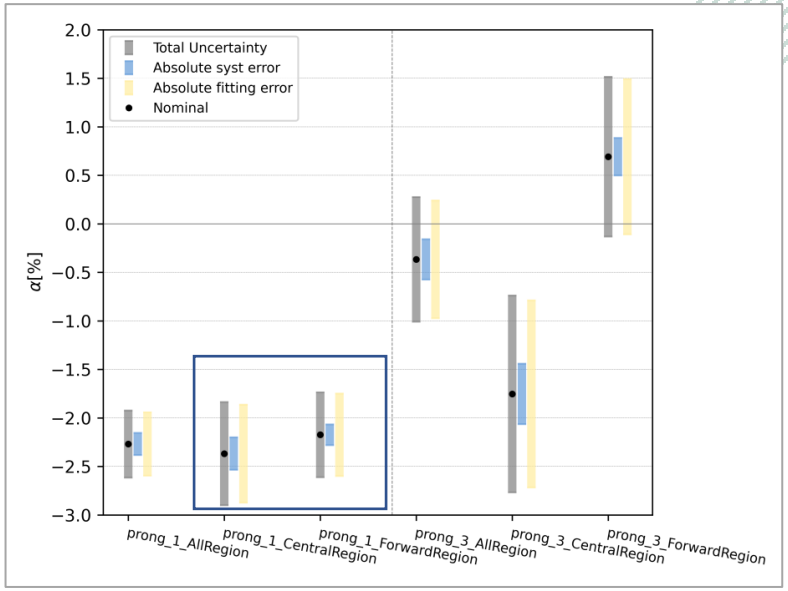




Post-Fit

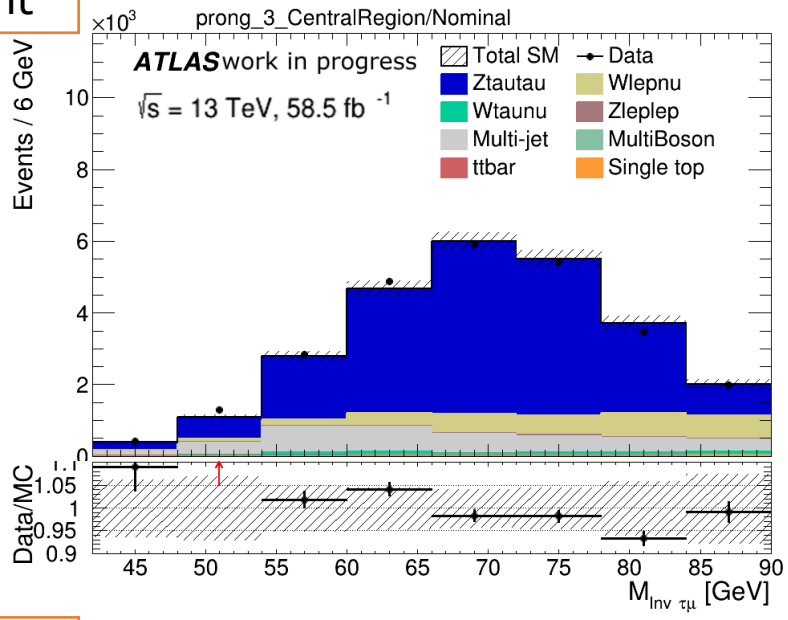


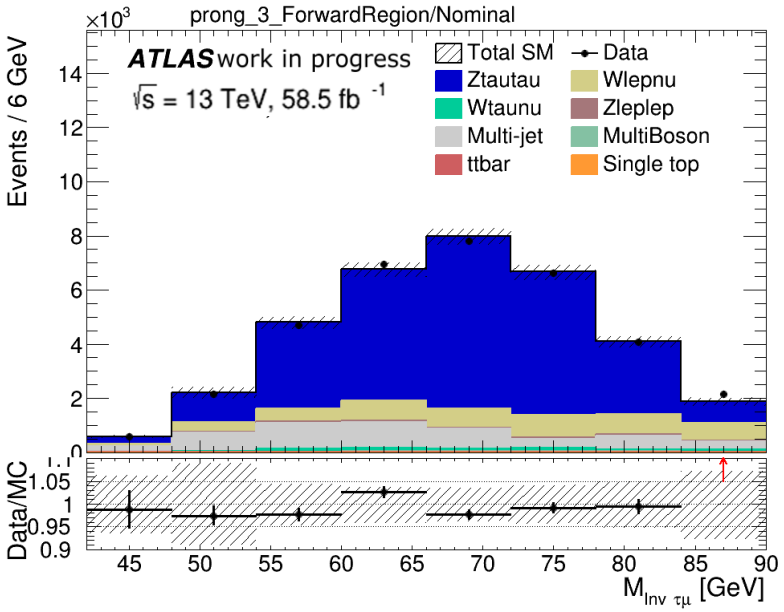




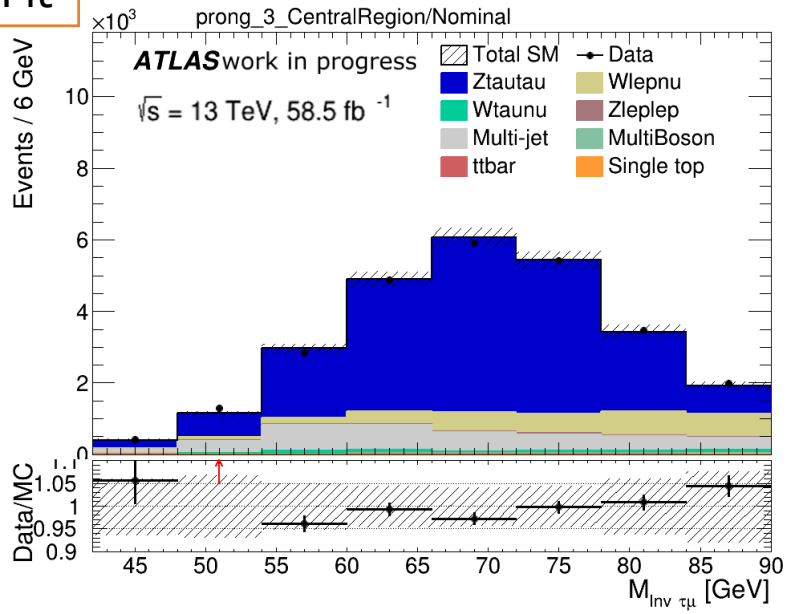
Sherpa2211 3-prong

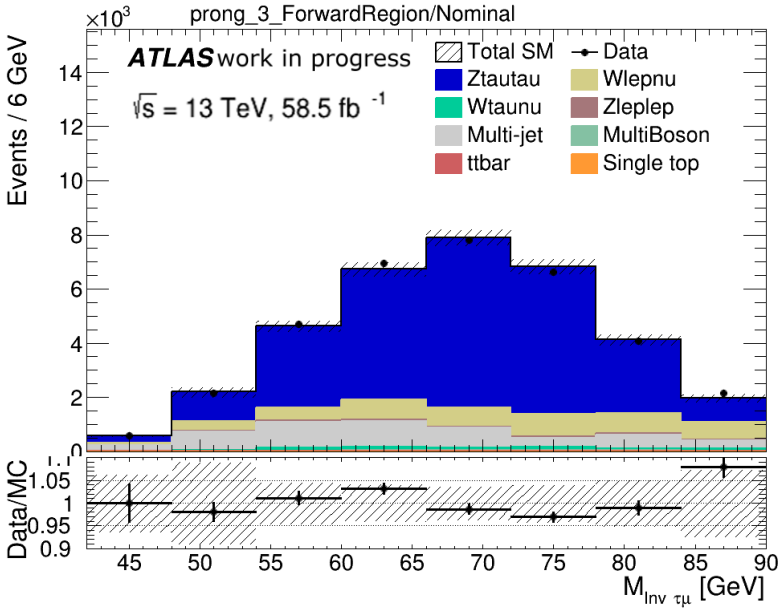
Pre-Fit

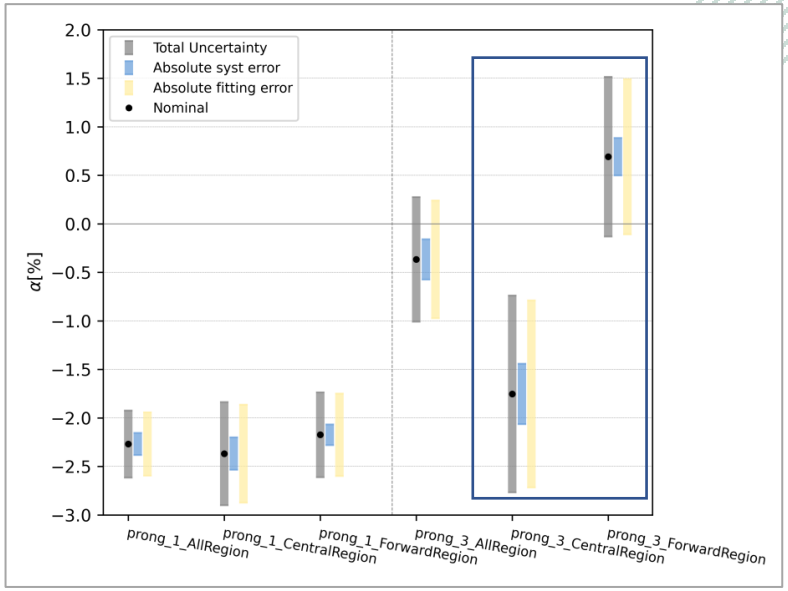




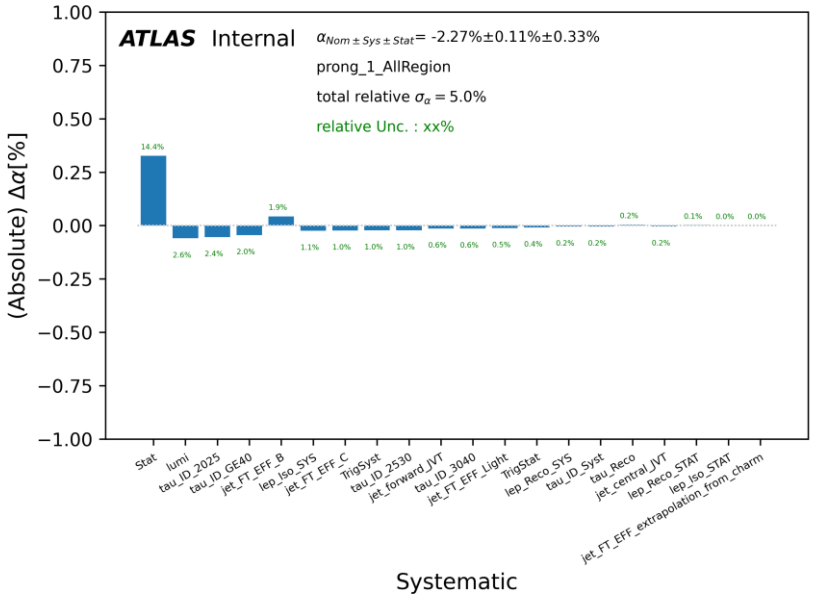
Post-Fit

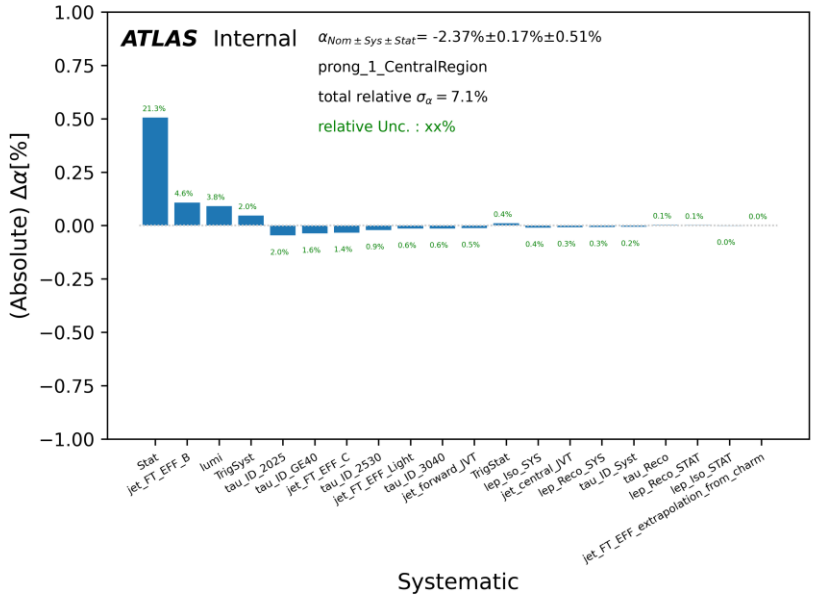


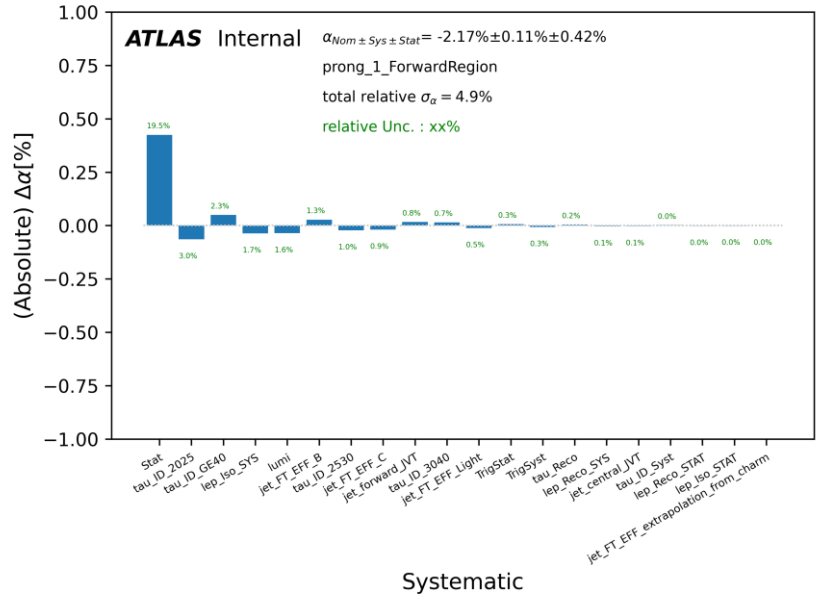


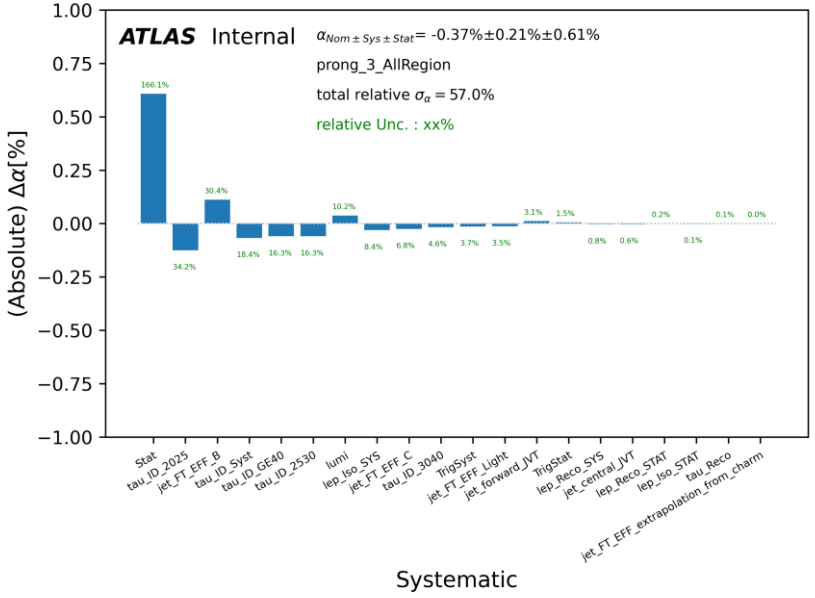


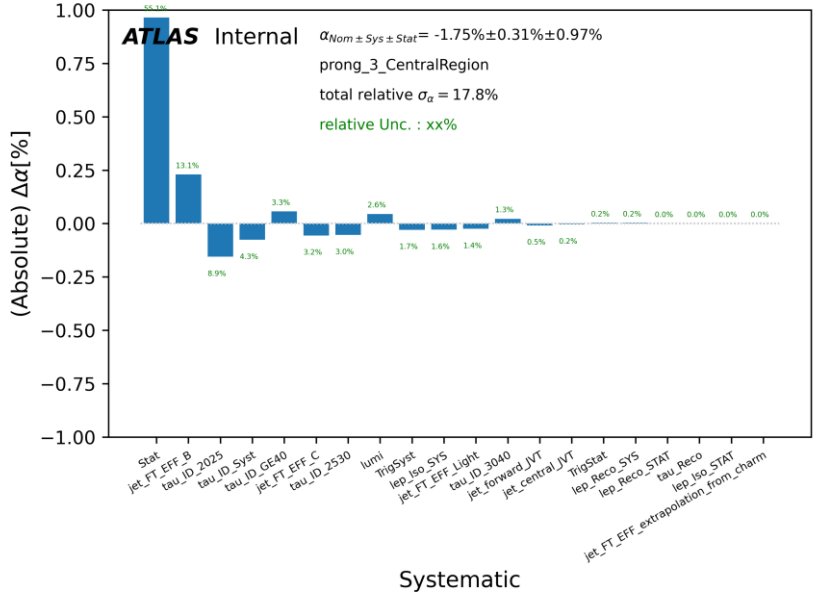
Sherpa 2211

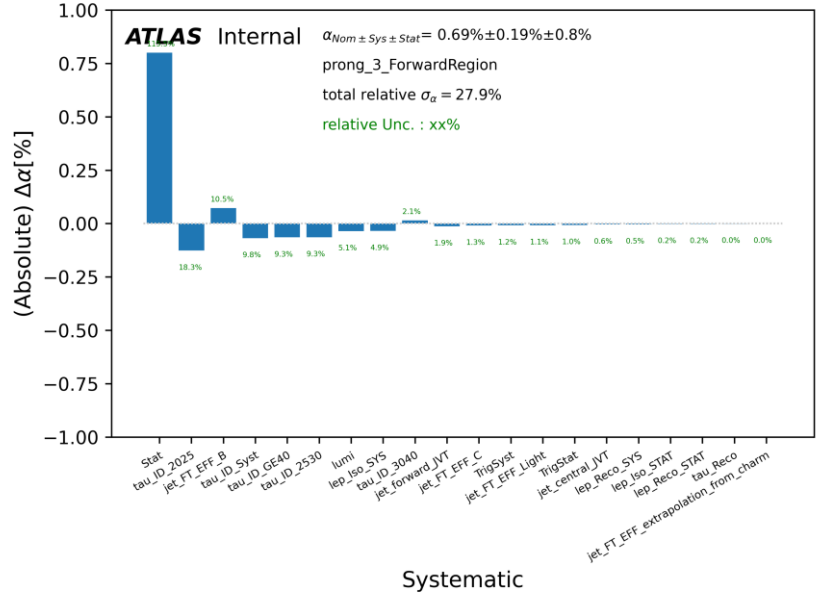

ATLAS Internal
$\alpha_{Nom \pm Sys \pm Stat}$= -2.27%±0.11%±0.33%
prong_1_AllRegion
total relative σ_α = 5.0%
relative Unc. : xx%
(Absolute) Δα[%]
Systematic


ATLAS Internal
$\alpha_{Nom \pm Sys \pm Stat}$= -2.37%±0.17%±0.51%
prong_1_CentralRegion
total relative σ_α = 7.1%
relative Unc. : xx%
(Absolute) Δα[%]
Systematic

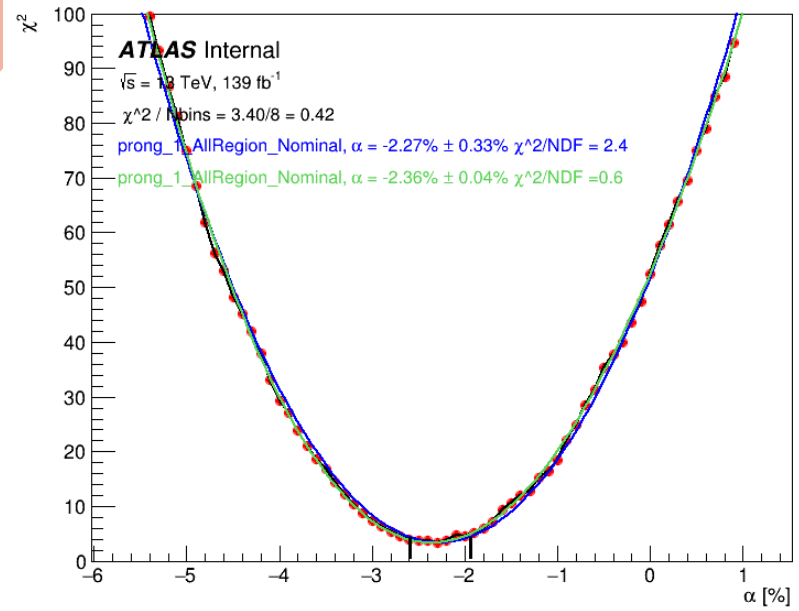

ATLAS Internal
$\alpha_{Nom \pm Sys \pm Stat}$= -2.17%±0.11%±0.42%
prong_1_ForwardRegion
total relative σ_α = 4.9%
relative Unc. : xx%
(Absolute) Δα[%]
Systematic

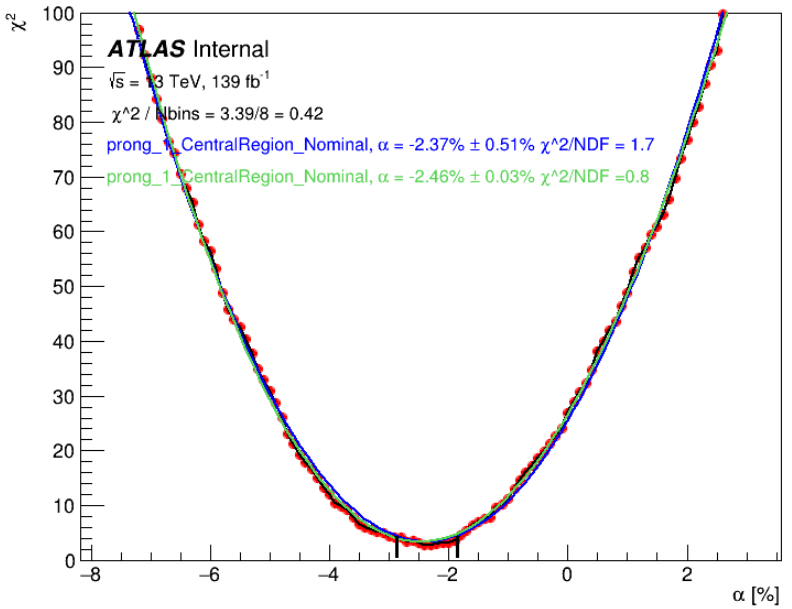

ATLAS Internal
$\alpha_{Nom \pm Sys \pm Stat}$= -0.37%±0.21%±0.61%
prong_3_AllRegion
total relative σ_α = 57.0%
relative Unc. : xx%
(Absolute) Δα[%]
Systematic

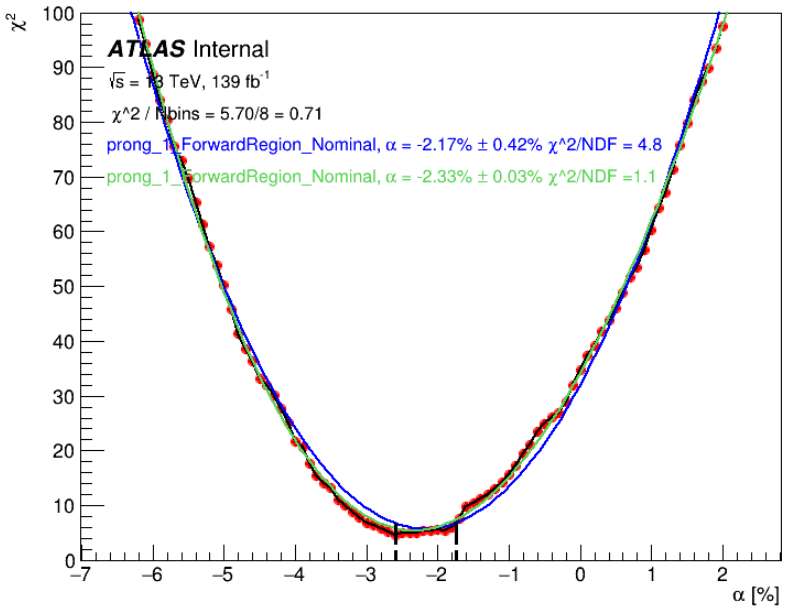

ATLAS Internal
$\alpha_{Nom \pm Sys \pm Stat}$= -1.75%±0.31%±0.97%
prong_3_CentralRegion
total relative σ_α = 17.8%
relative Unc. : xx%
(Absolute) Δα[%]
Systematic

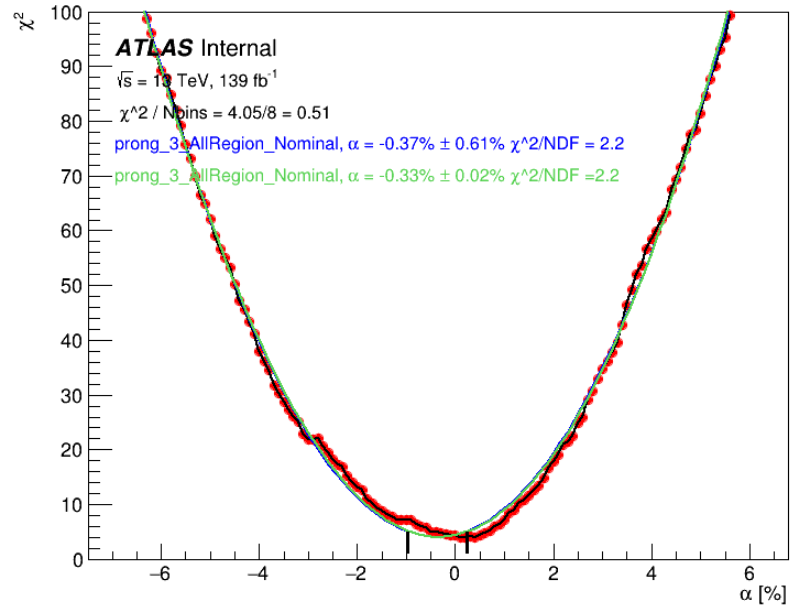

ATLAS Internal
$\alpha_{Nom \pm Sys \pm Stat}$= 0.69%±0.19%±0.8%
prong_3_ForwardRegion
total relative σ_α = 27.9%
relative Unc. : xx%
(Absolute) Δα[%]
Systematic

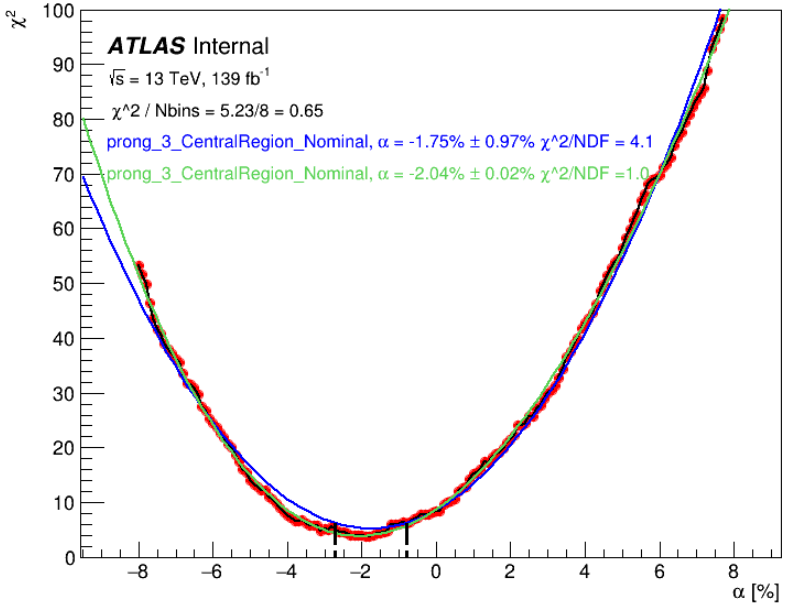
Sherpa 2211

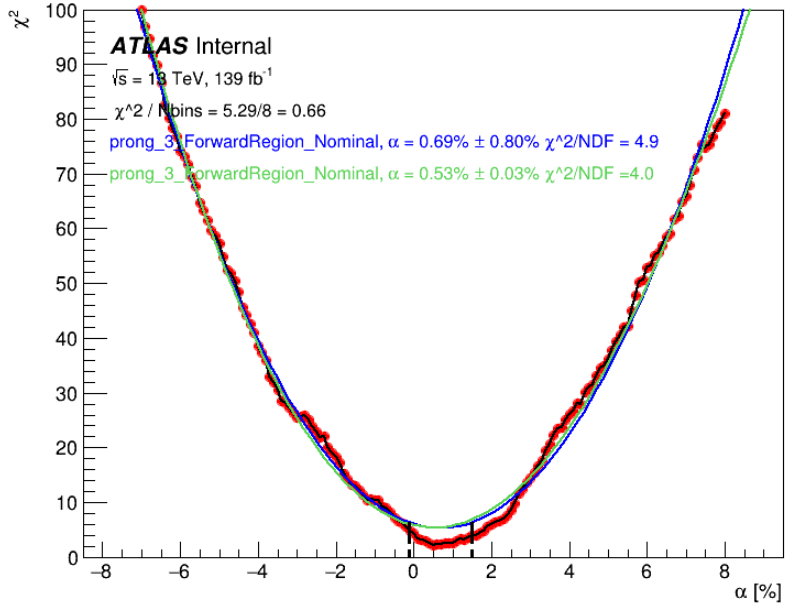

ATLAS Internal
√s = 13 TeV, 139 fb-1
χ^2 / Nbins = 3.40/8 = 0.42
prong_1_AllRegion_Nominal, α = -2.27% ± 0.33% χ^2/NDF = 2.4
prong_1_AllRegion_Nominal, α = -2.36% ± 0.04% χ^2/NDF =0.6
χ2
α [%]


ATLAS Internal
√s = 13 TeV, 139 fb-1
χ^2 / Nbins = 3.39/8 = 0.42
prong_1_CentralRegion_Nominal, α = -2.37% ± 0.51% χ^2/NDF = 1.7
prong_1_CentralRegion_Nominal, α = -2.46% ± 0.03% χ^2/NDF =0.8
χ2
α [%]

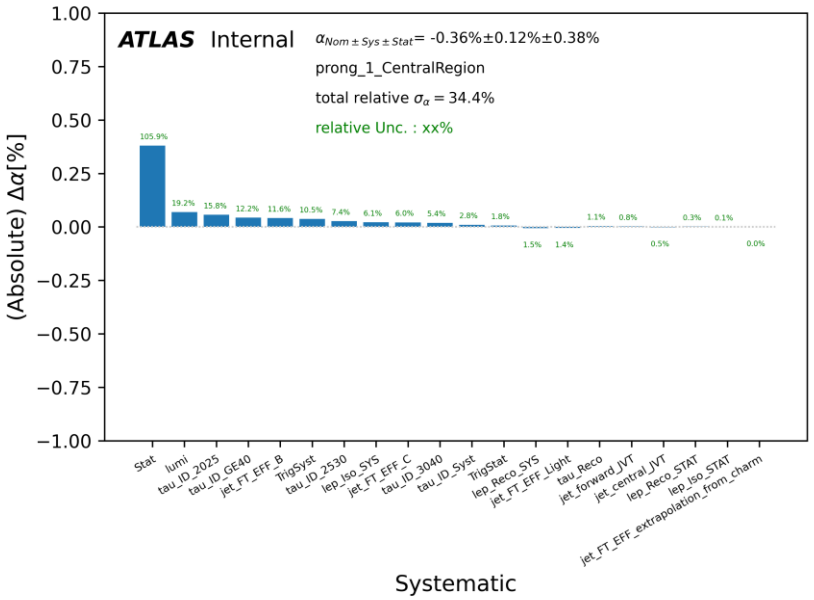

ATLAS Internal
√s = 13 TeV, 139 fb-1
χ^2 / Nbins = 5.70/8 = 0.71
prong_1_ForwardRegion_Nominal, α = -2.17% ± 0.42% χ^2/NDF = 4.8
prong_1_ForwardRegion_Nominal, α = -2.33% ± 0.03% χ^2/NDF =1.1
χ2
α [%]

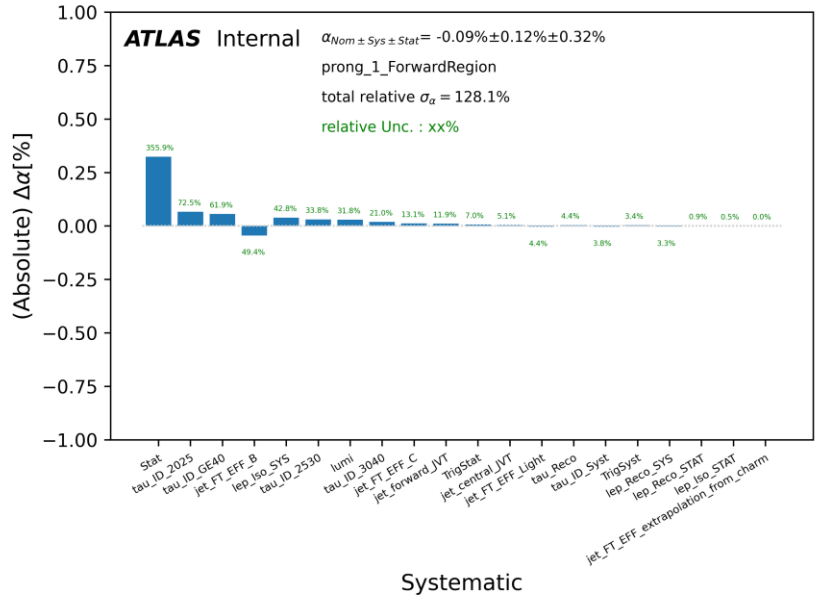

ATLAS Internal
√s = 13 TeV, 139 fb-1
χ^2 / Nbins = 4.05/8 = 0.51
prong_3_AllRegion_Nominal, α = -0.37% ± 0.61% χ^2/NDF = 2.2
prong_3_AllRegion_Nominal, α = -0.33% ± 0.02% χ^2/NDF =2.2
χ2
α [%]

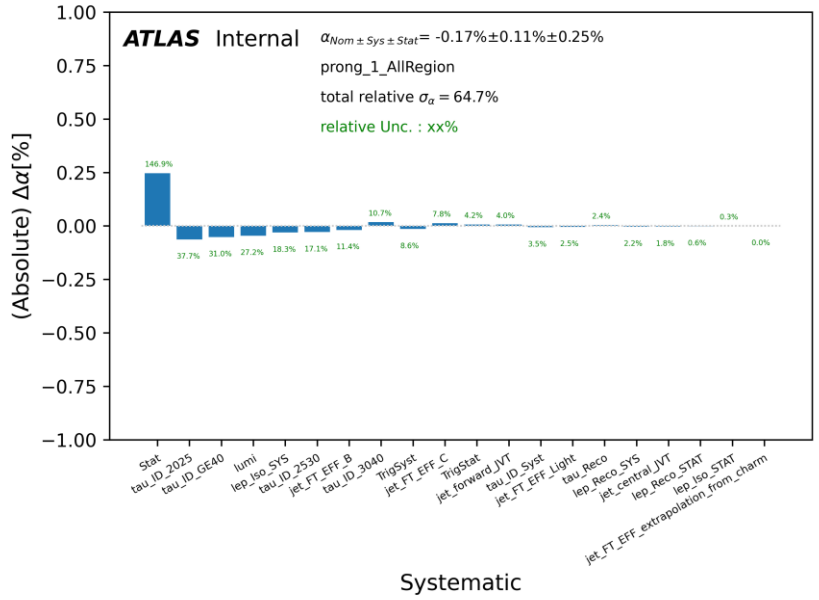

ATLAS Internal
√s = 13 TeV, 139 fb-1
χ^2 / Nbins = 5.23/8 = 0.65
prong_3_CentralRegion_Nominal, α = -1.75% ± 0.97% χ^2/NDF = 4.1
prong_3_CentralRegion_Nominal, α = -2.04% ± 0.02% χ^2/NDF =1.0
χ2
α [%]

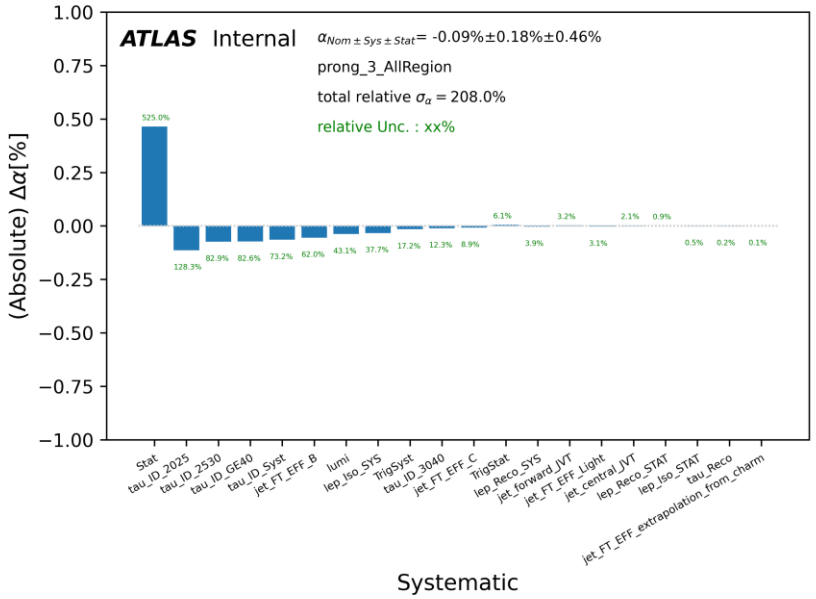

ATLAS Internal
√s = 13 TeV, 139 fb-1
χ^2 / Nbins = 5.29/8 = 0.66
prong_3_ForwardRegion_Nominal, α = 0.69% ± 0.80% χ^2/NDF = 4.9
prong_3_ForwardRegion_Nominal, α = 0.53% ± 0.03% χ^2/NDF =4.0
χ2
α [%]

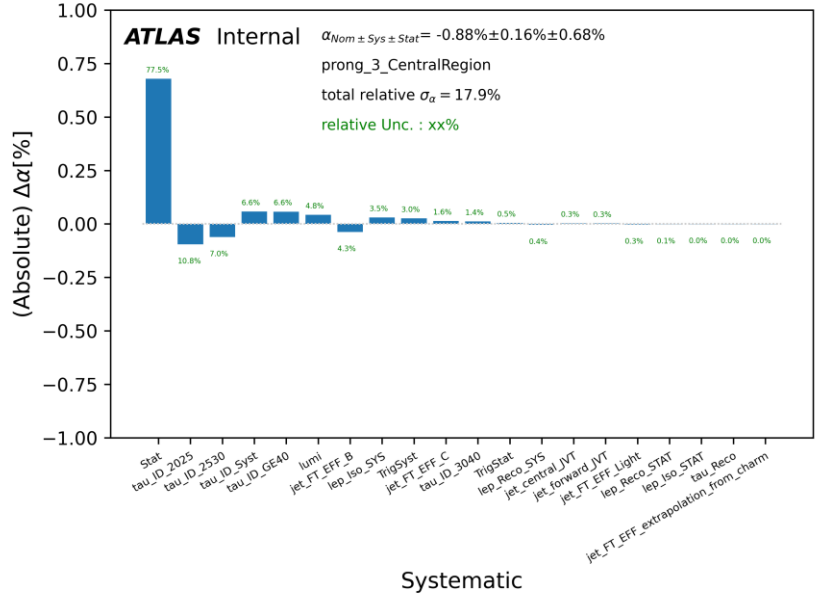
PoPy8

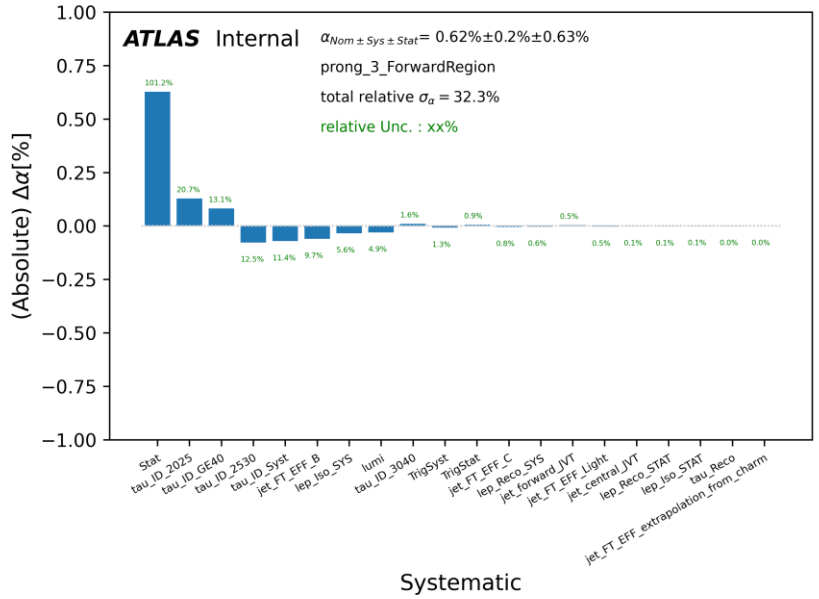

ATLAS Internal
$\alpha_{Nom \pm Sys \pm Stat}$= -0.36%±0.12%±0.38%
prong_1_CentralRegion
total relative σ_α = 34.4%
relative Unc. : xx%
(Absolute) Δα[%]
Systematic


ATLAS Internal
$\alpha_{Nom \pm Sys \pm Stat}$= -0.09%±0.12%±0.32%
prong_1_ForwardRegion
total relative σ_α = 128.1%
relative Unc. : xx%
(Absolute) Δα[%]
Systematic

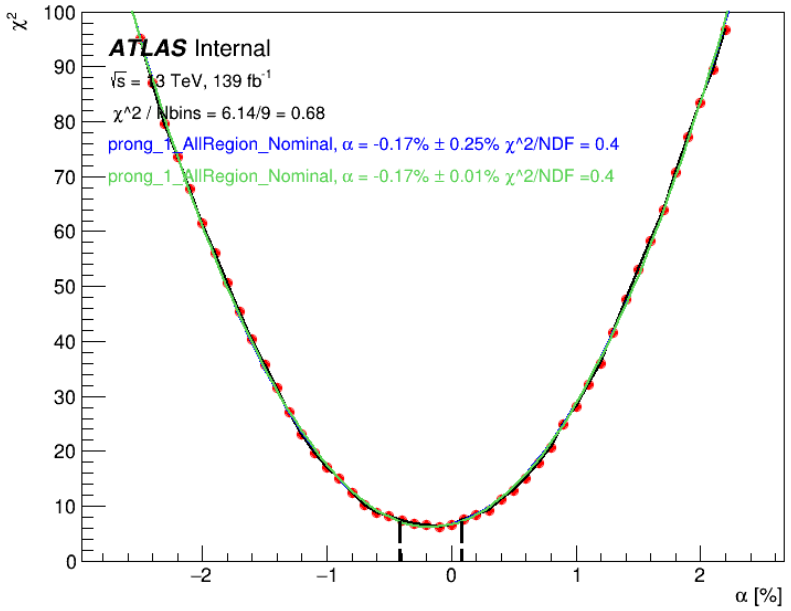

ATLAS Internal
$\alpha_{Nom \pm Sys \pm Stat}$= -0.17%±0.11%±0.25%
prong_1_AllRegion
total relative σ_α = 64.7%
relative Unc. : xx%
(Absolute) Δα[%]
Systematic

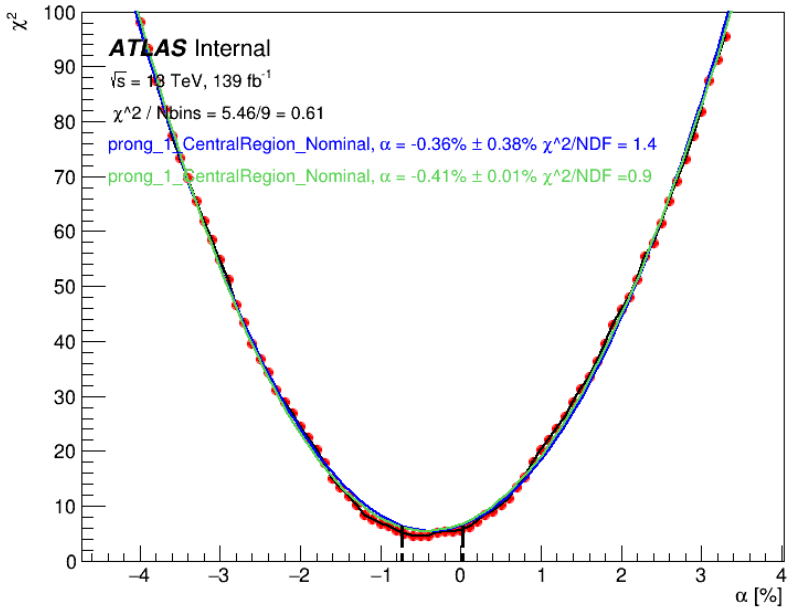

ATLAS Internal
$\alpha_{Nom \pm Sys \pm Stat}$= -0.09%±0.18%±0.46%
prong_3_AllRegion
total relative σ_α = 208.0%
relative Unc. : xx%
(Absolute) Δα[%]
Systematic

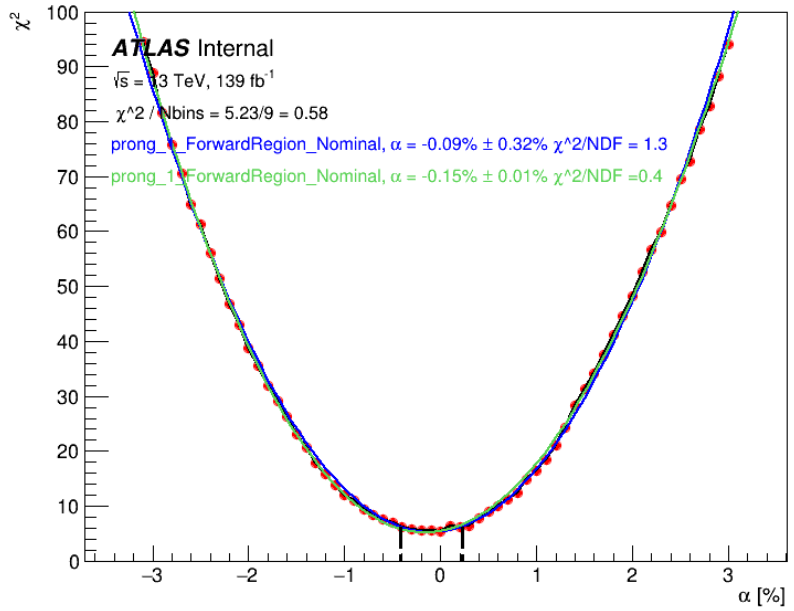

ATLAS Internal
$\alpha_{Nom \pm Sys \pm Stat}$= -0.88%±0.16%±0.68%
prong_3_CentralRegion
total relative σ_α = 17.9%
relative Unc. : xx%
(Absolute) Δα[%]
Systematic

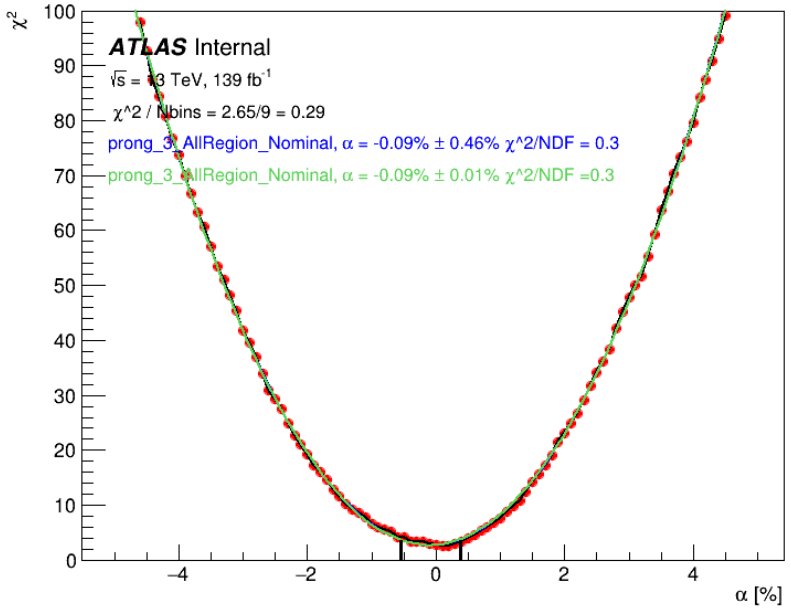

ATLAS Internal
$\alpha_{Nom \pm Sys \pm Stat}$= 0.62%±0.2%±0.63%
prong_3_ForwardRegion
total relative σ_α = 32.3%
relative Unc. : xx%
(Absolute) Δα[%]
Systematic

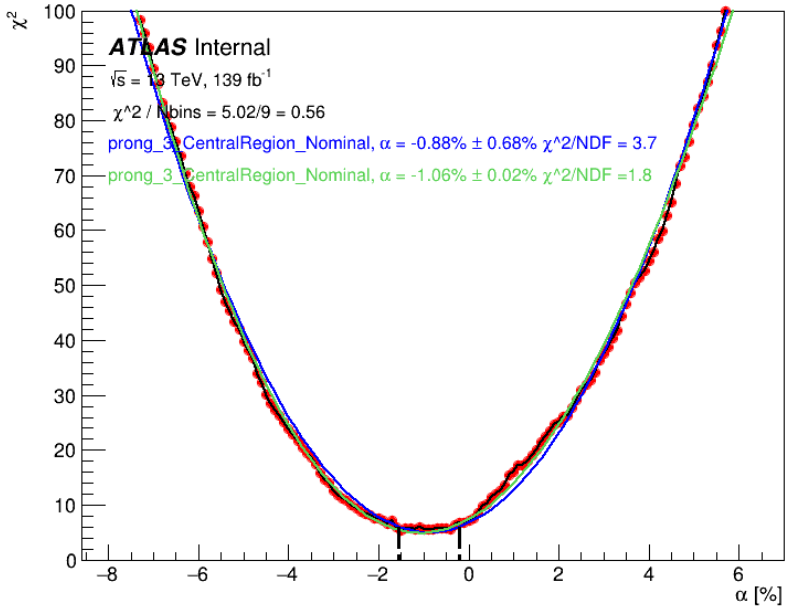
PoPy8

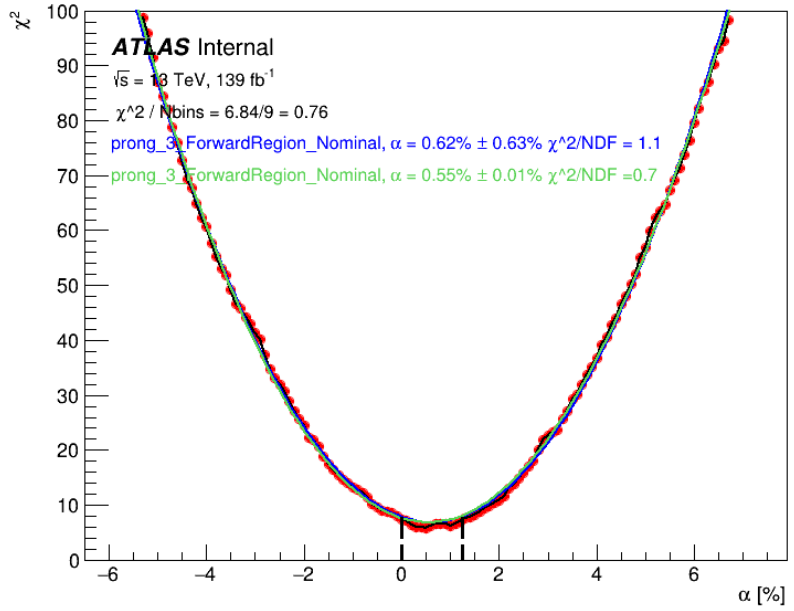

ATLAS Internal
√s = 13 TeV, 139 fb-1
χ^2 / Nbins = 6.14/9 = 0.68
prong_1_AllRegion_Nominal, α = -0.17% ± 0.25% χ^2/NDF = 0.4
prong_1_AllRegion_Nominal, α = -0.17% ± 0.01% χ^2/NDF =0.4
χ2
α [%]


ATLAS Internal
√s = 13 TeV, 139 fb-1
χ^2 / Nbins = 5.46/9 = 0.61
prong_1_CentralRegion_Nominal, α = -0.36% ± 0.38% χ^2/NDF = 1.4
prong_1_CentralRegion_Nominal, α = -0.41% ± 0.01% χ^2/NDF =0.9
χ2
α [%]

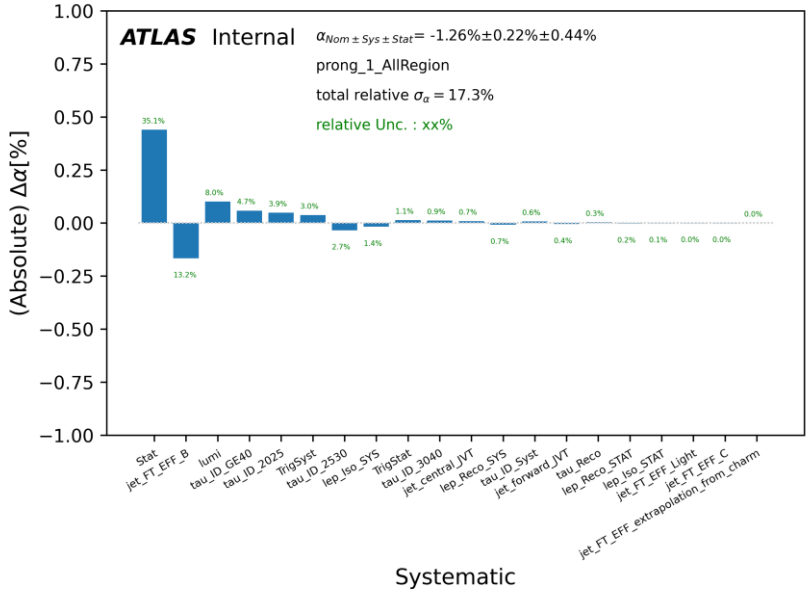

ATLAS Internal
√s = 13 TeV, 139 fb-1
χ^2 / Nbins = 5.23/9 = 0.58
prong_1_ForwardRegion_Nominal, α = -0.09% ± 0.32% χ^2/NDF = 1.3
prong_1_ForwardRegion_Nominal, α = -0.15% ± 0.01% χ^2/NDF =0.4
χ2
α [%]

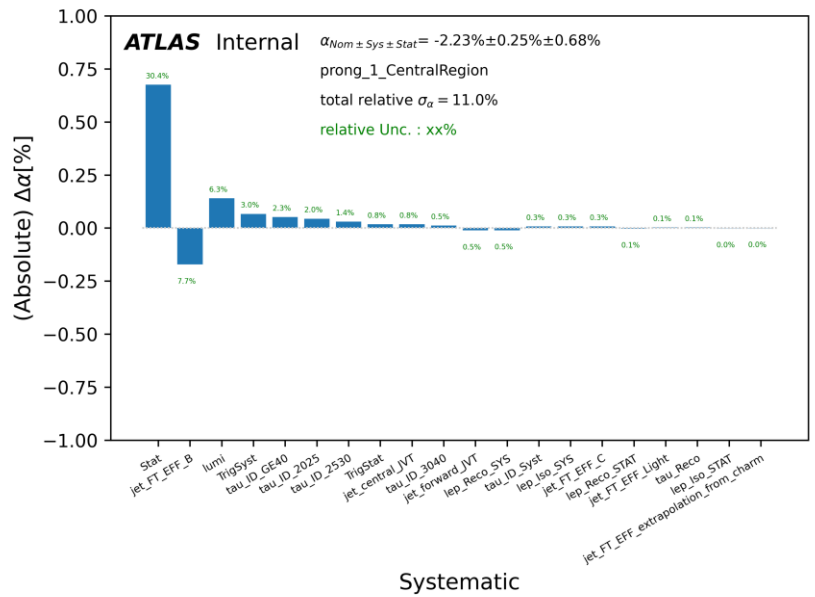

ATLAS Internal
√s = 13 TeV, 139 fb-1
χ^2 / Nbins = 2.65/9 = 0.29
prong_3_AllRegion_Nominal, α = -0.09% ± 0.46% χ^2/NDF = 0.3
prong_3_AllRegion_Nominal, α = -0.09% ± 0.01% χ^2/NDF =0.3
χ2
α [%]

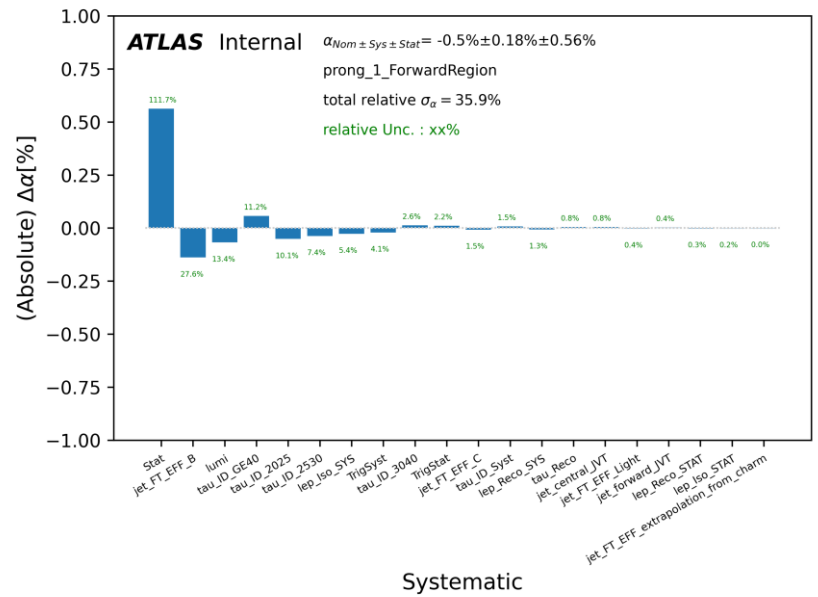

ATLAS Internal
√s = 13 TeV, 139 fb-1
χ^2 / Nbins = 5.02/9 = 0.56
prong_3_CentralRegion_Nominal, α = -0.88% ± 0.68% χ^2/NDF = 3.7
prong_3_CentralRegion_Nominal, α = -1.06% ± 0.02% χ^2/NDF =1.8
χ2
α [%]

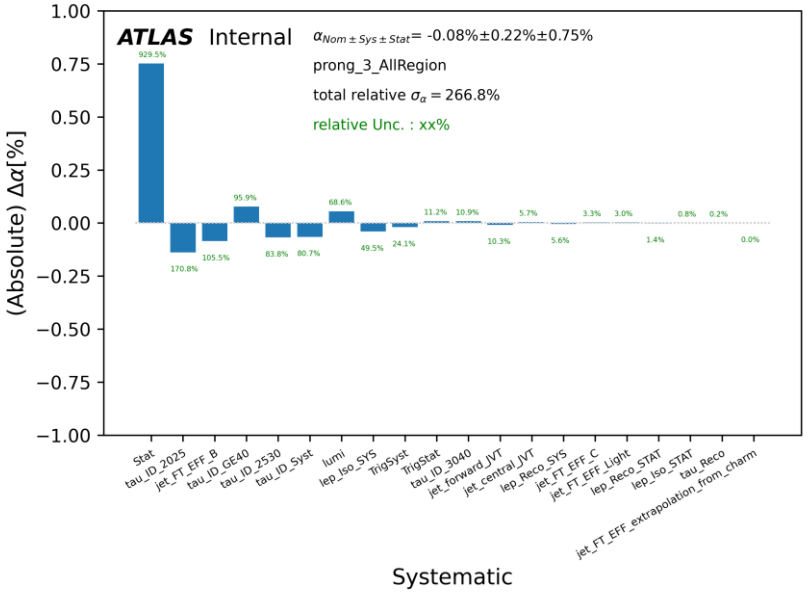

ATLAS Internal
√s = 13 TeV, 139 fb-1
χ^2 / Nbins = 6.84/9 = 0.76
prong_3_ForwardRegion_Nominal, α = 0.62% ± 0.63% χ^2/NDF = 1.1
prong_3_ForwardRegion_Nominal, α = 0.55% ± 0.01% χ^2/NDF =0.7
χ2
α [%]

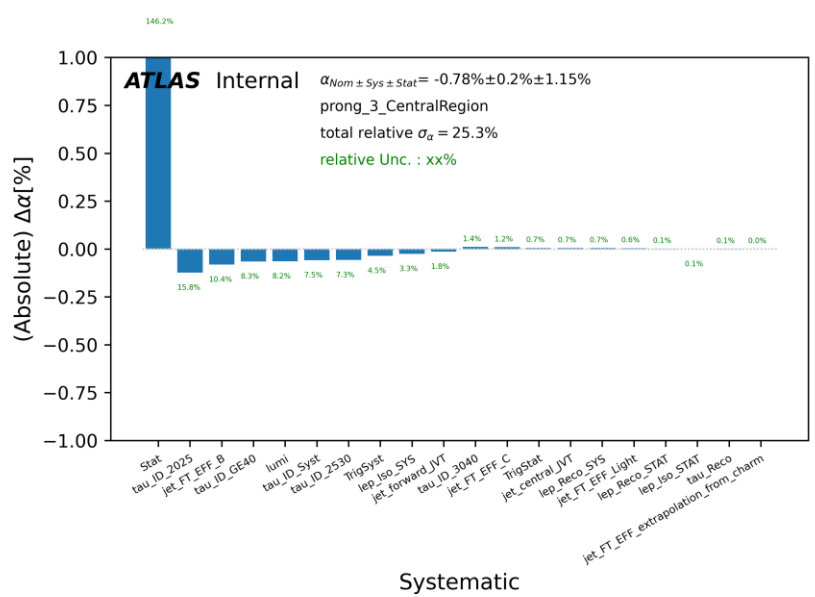
Sherpa221 + Sherpa2211

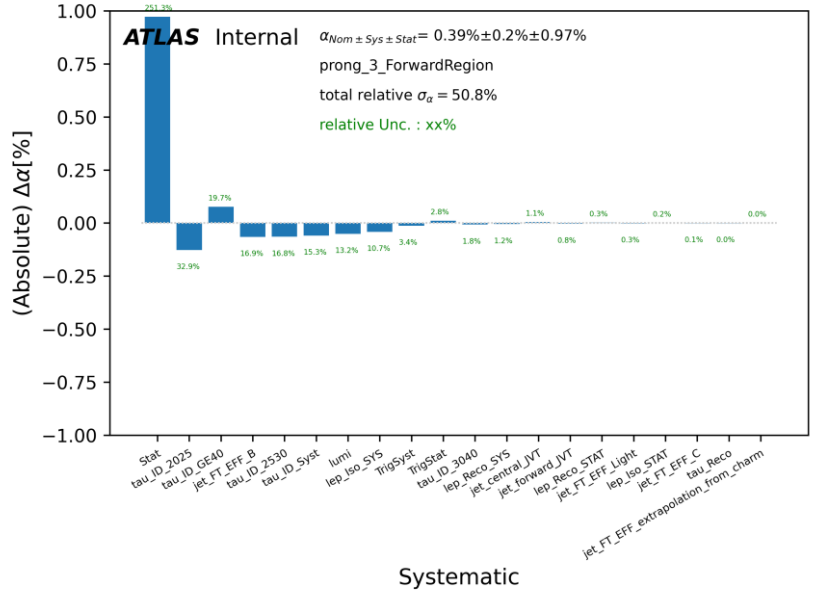




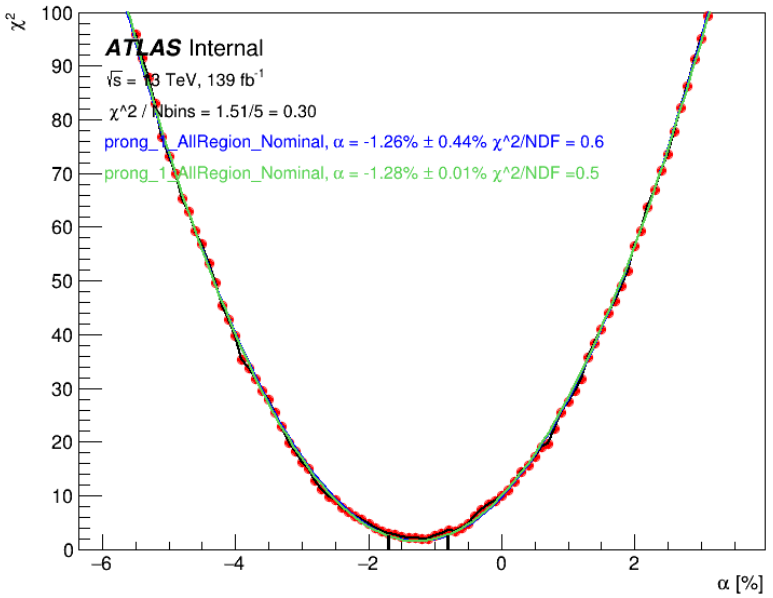


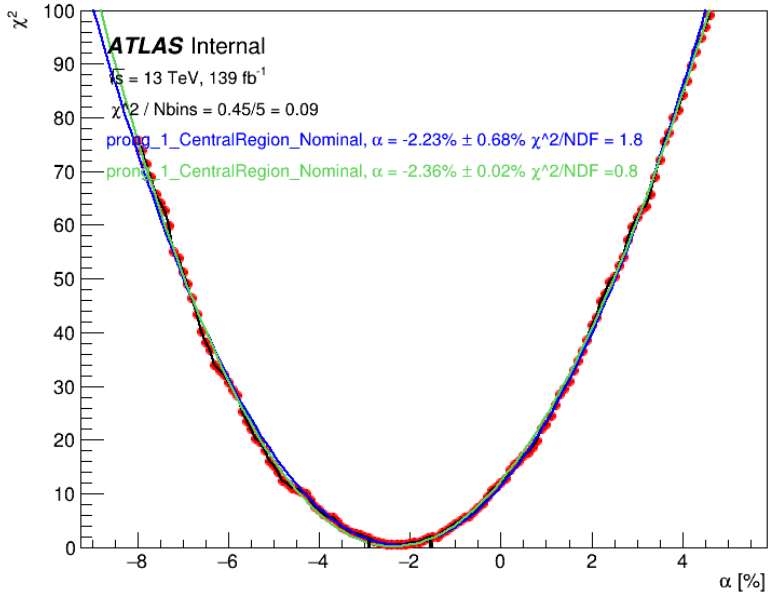


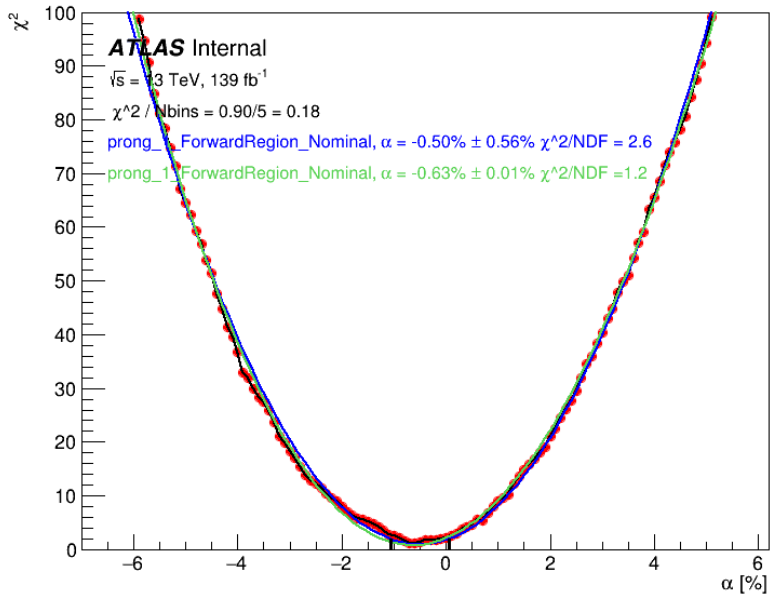


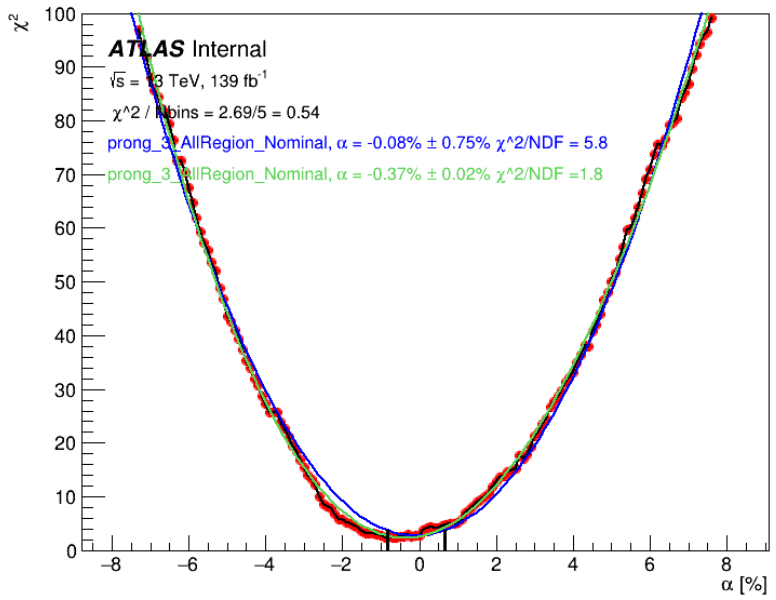


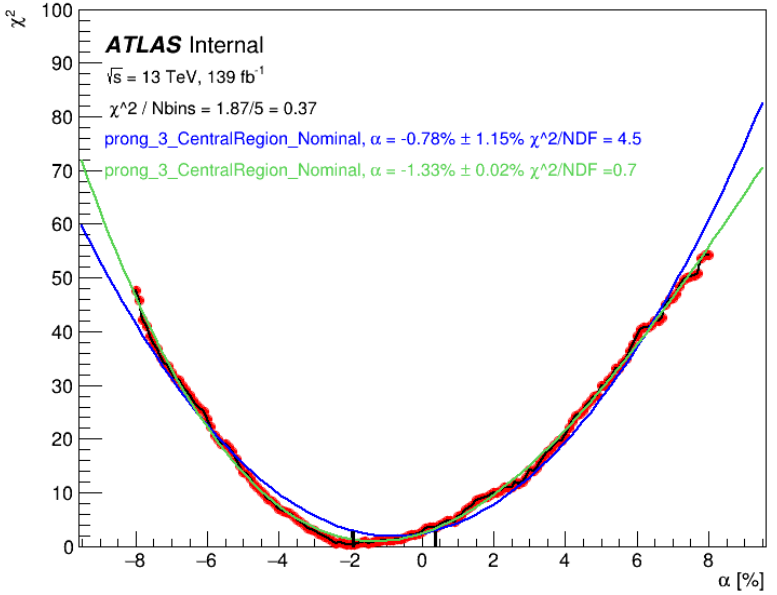
Sherpa221 + Sherpa2211

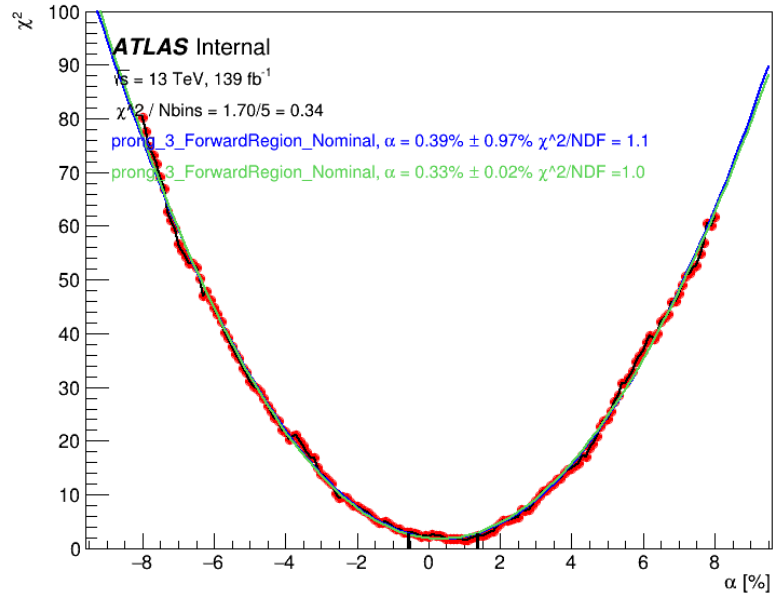




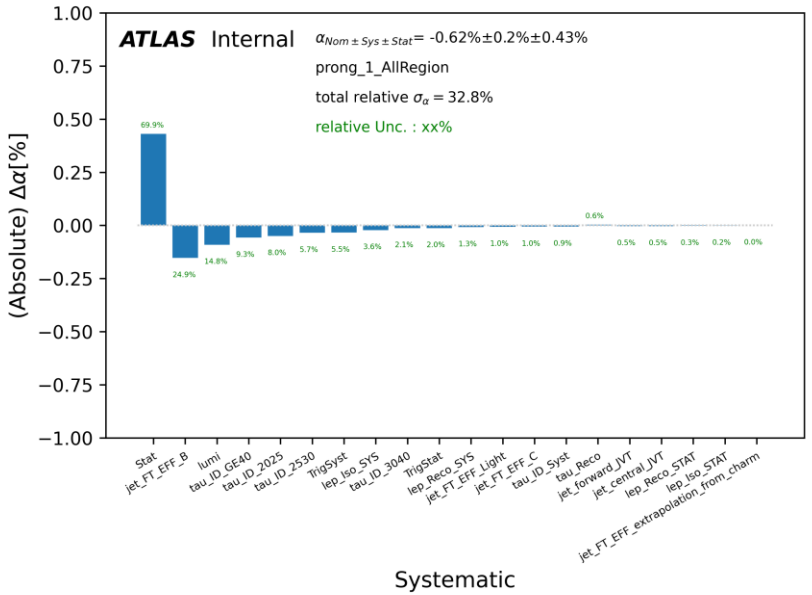


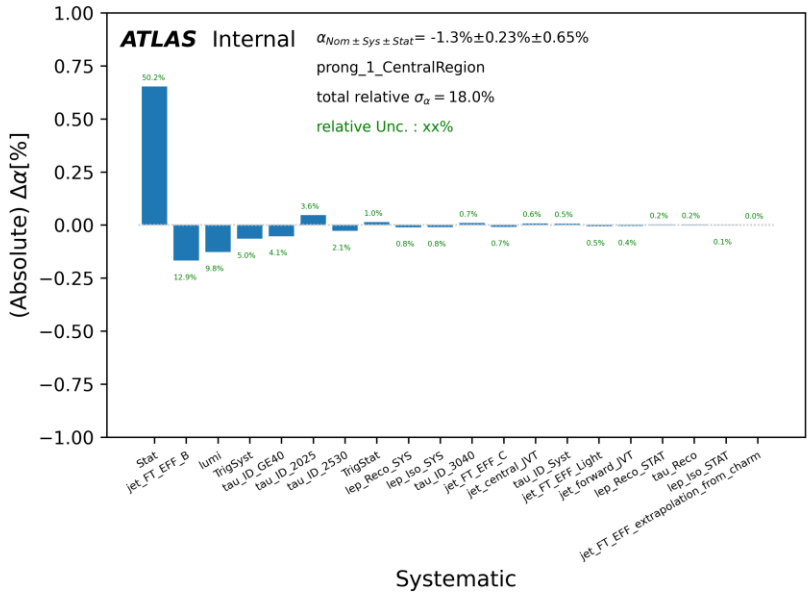


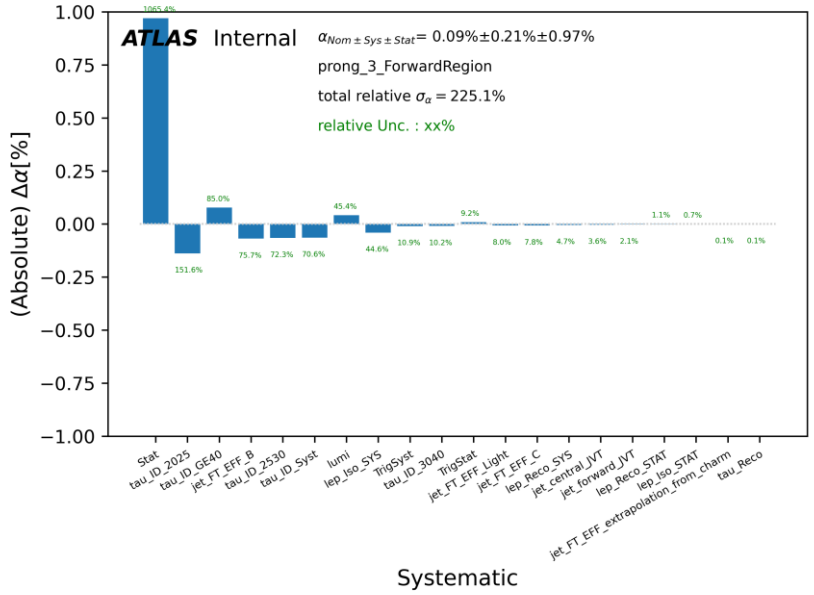


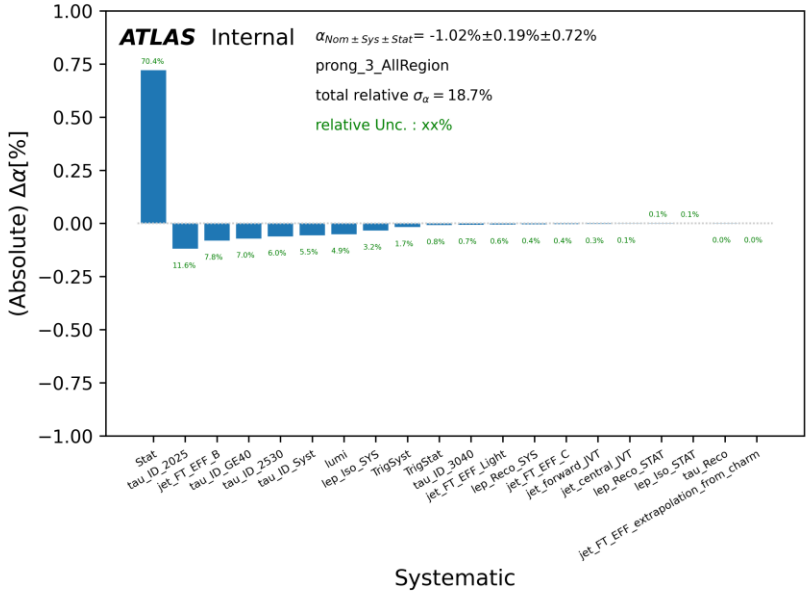


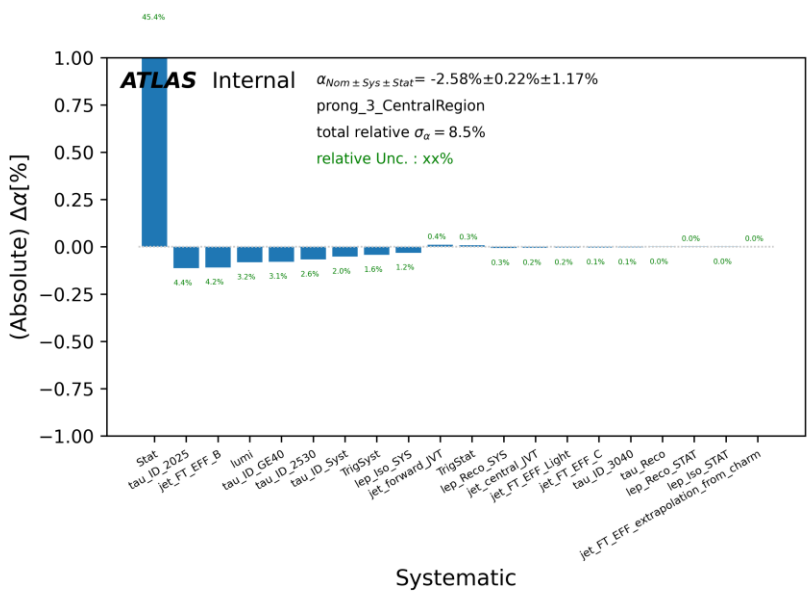
Sherpa221 + PoPy8

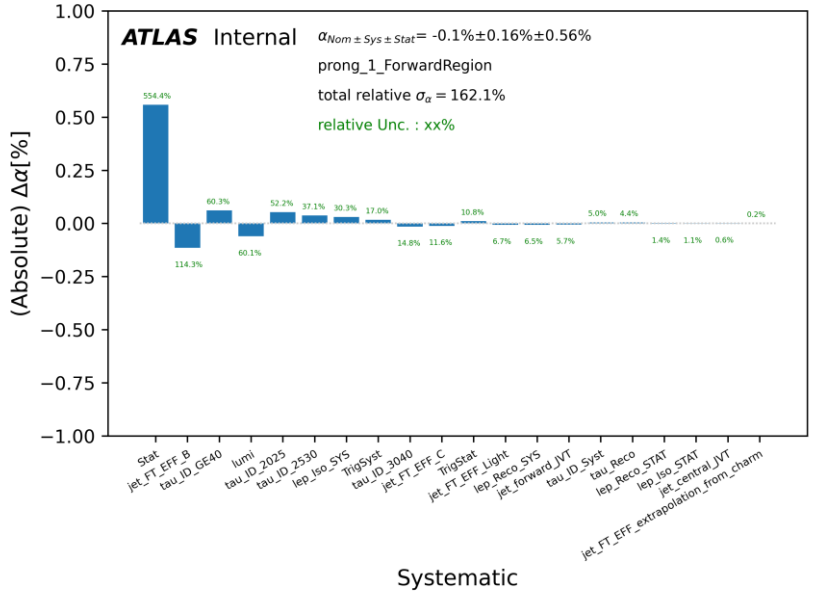




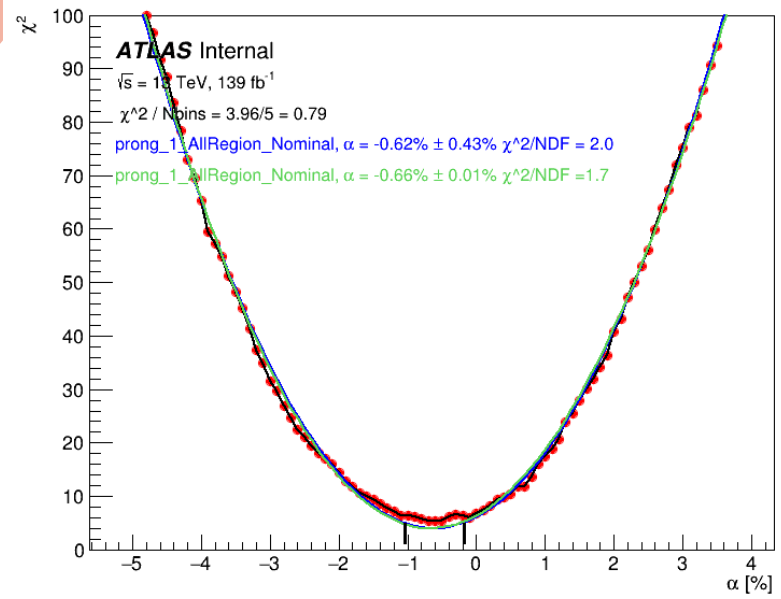


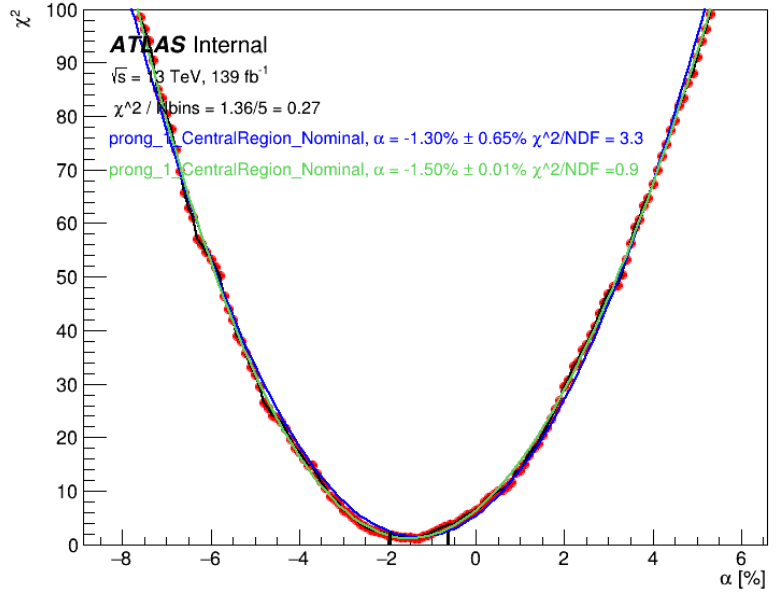


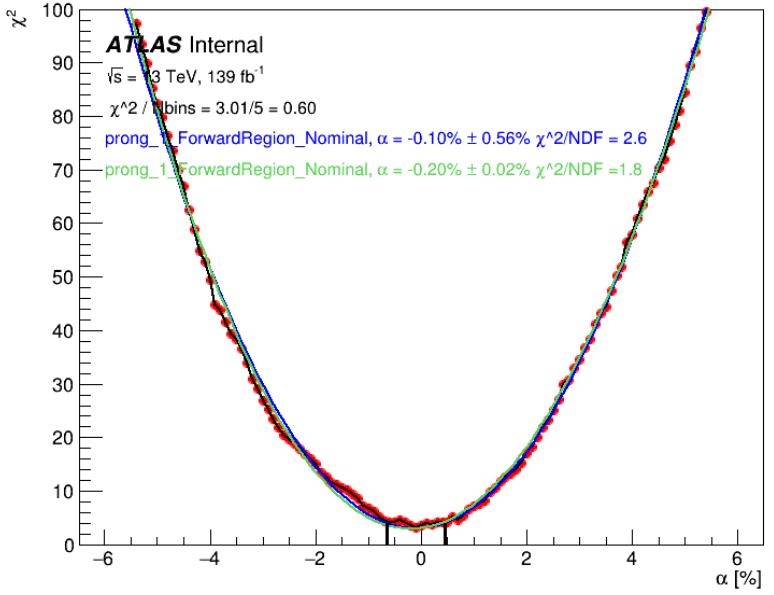


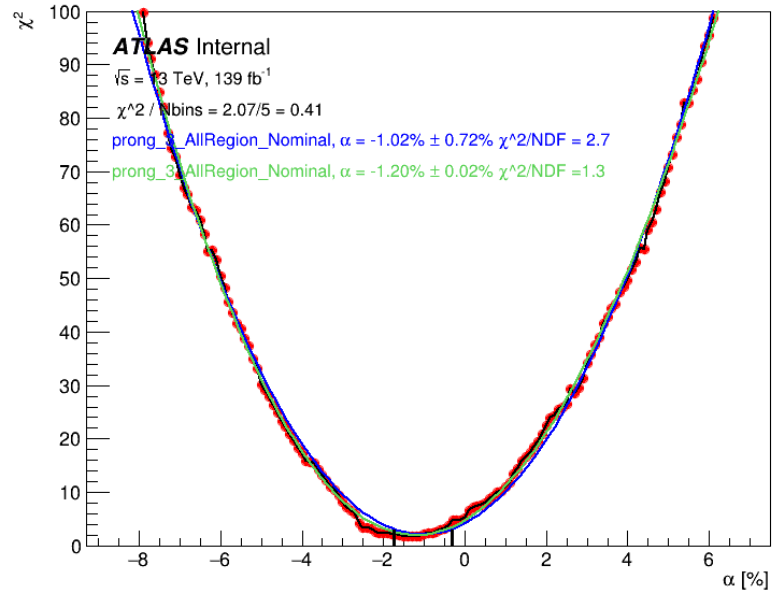


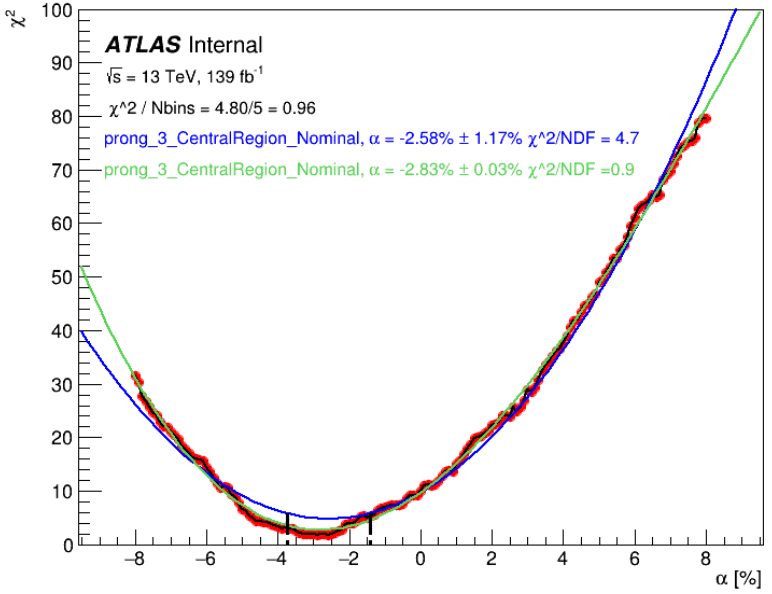
Sherpa221 + PoPy8

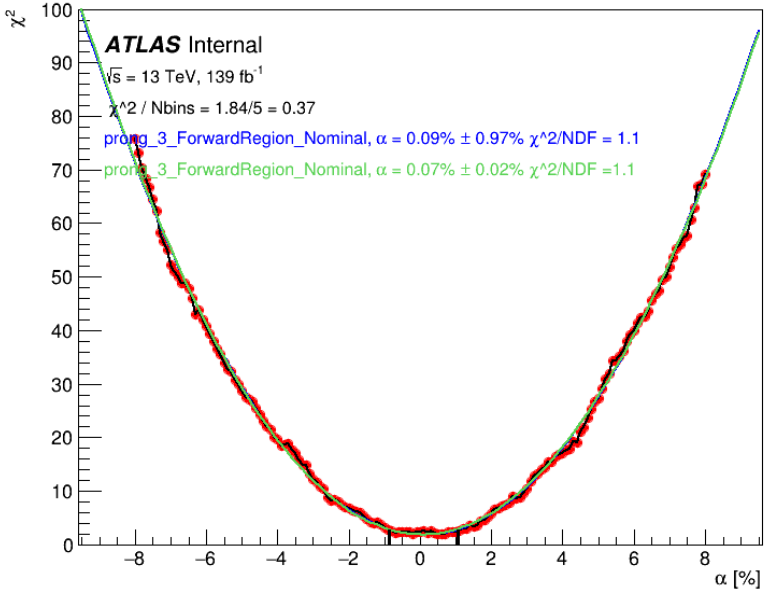




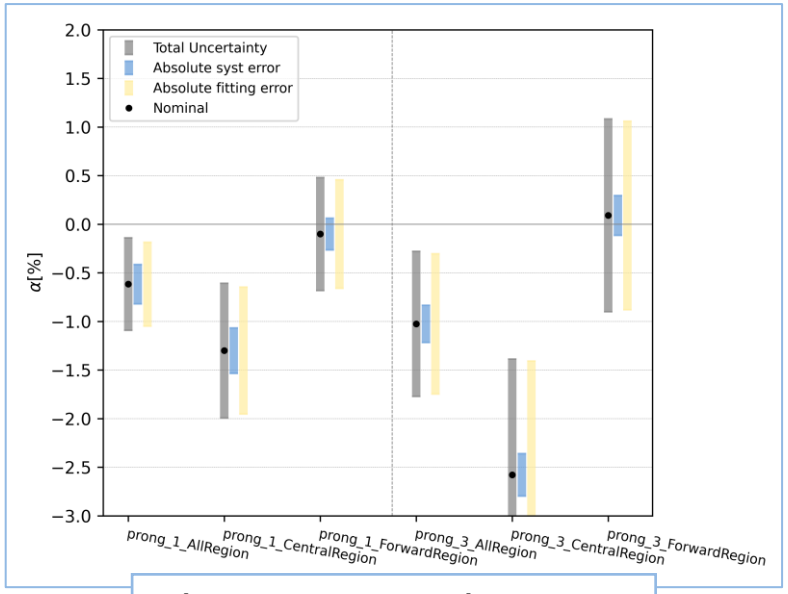




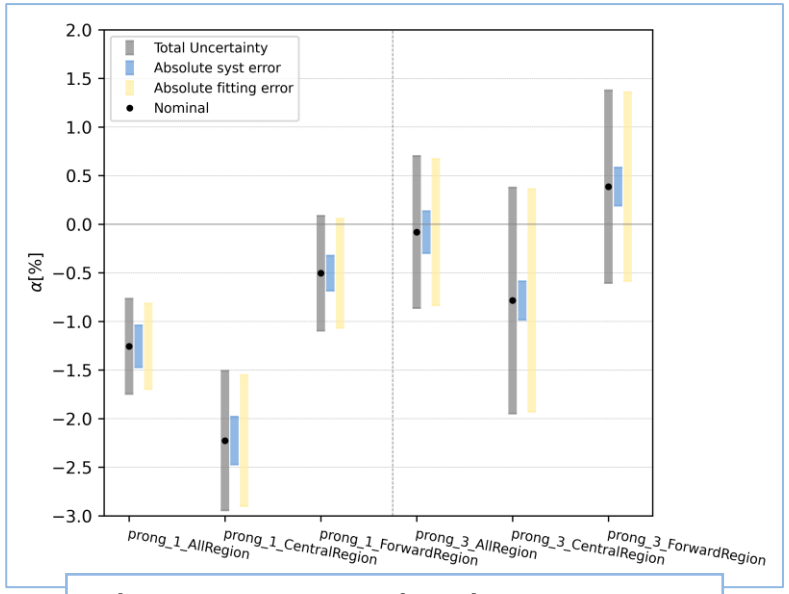




TES results



Sherpa221 with PoPy8

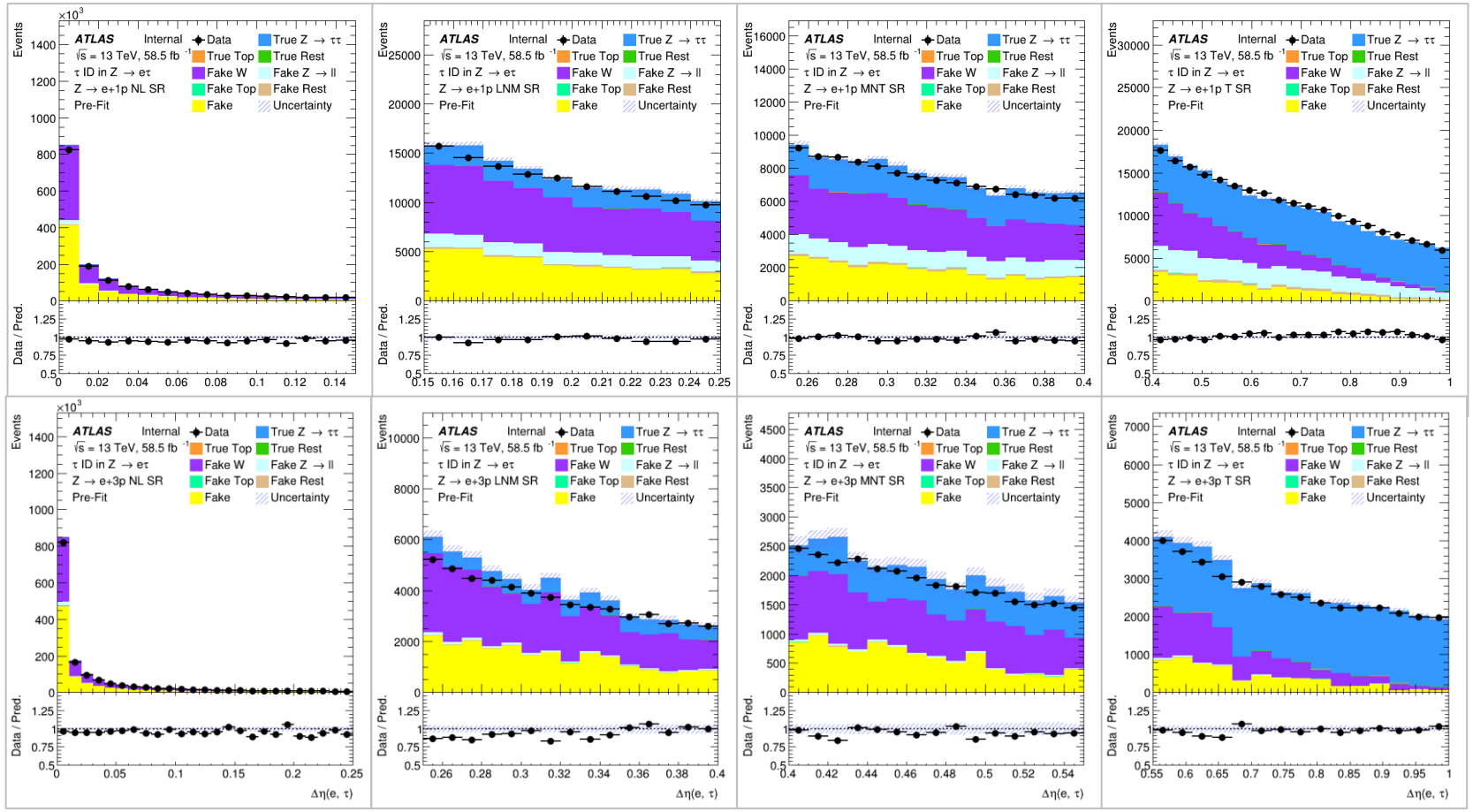


Sherpa221 with Sherpa2211

- **Sherpa221** (Blue): The results are similar in the Forward regions but not close in Central regions when comparing between mc-to-mc results and common fit results. And there are also some differences between Sherpa2211 or PoPy8 in 1P or 3P Central.

Pre-fit plots in e channel(inclusive)

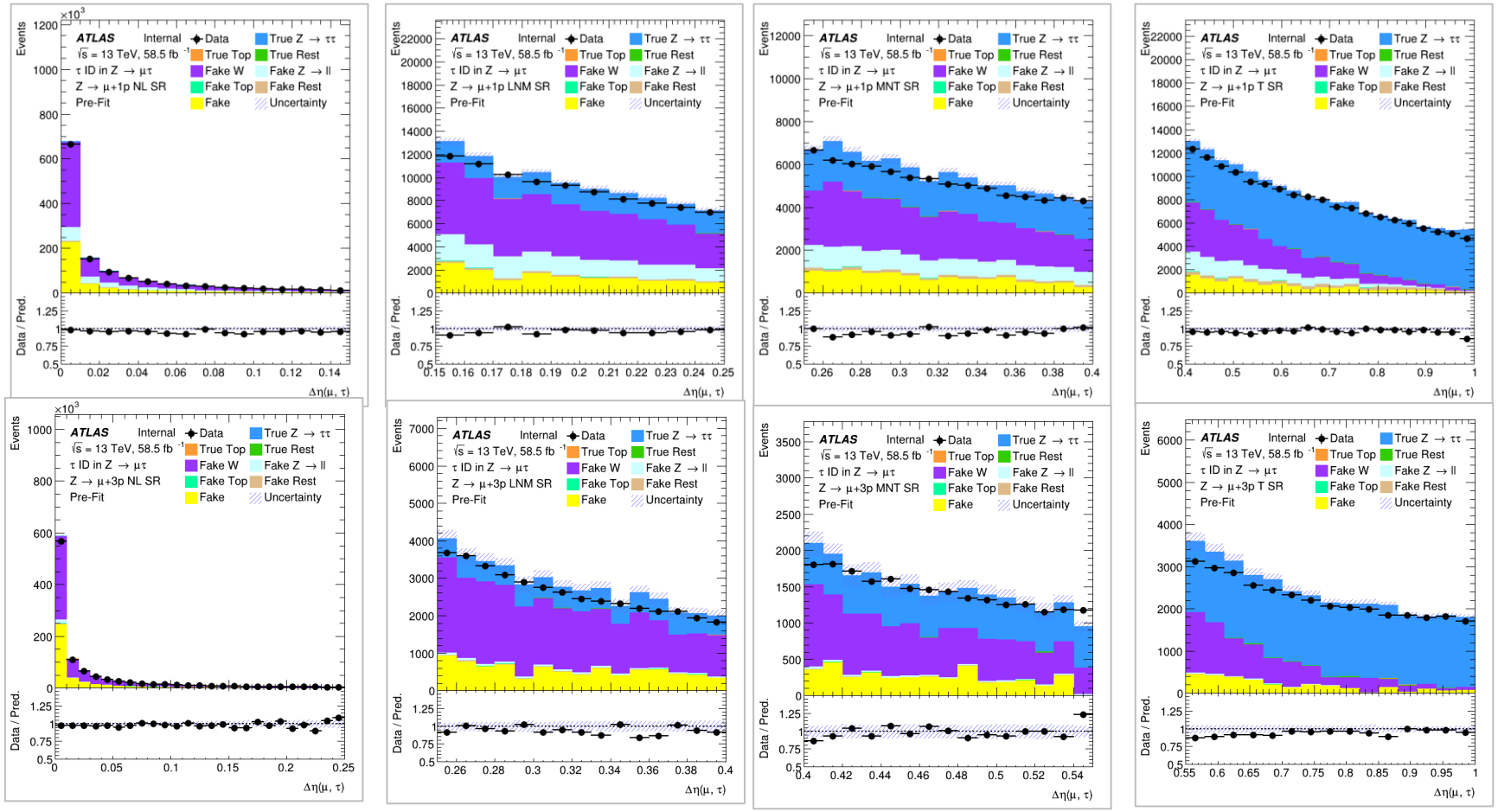
- Pre-fit plots in e channel(inclusive)



- Slight overestimation in all τ ID selections in the 3-prong channel.
- Overestimation in NL, LNM selection in 1prong channel.

Pre-fit plots in μ channel(inclusive)

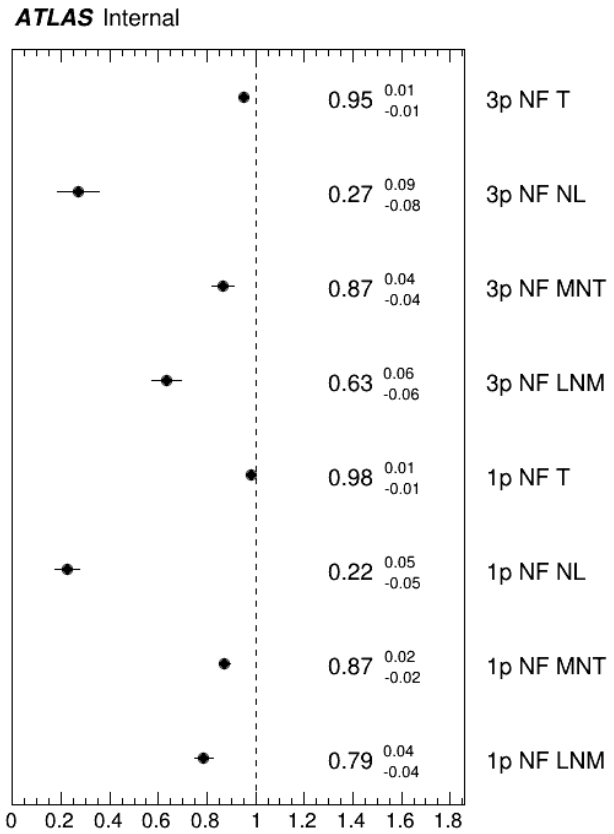
- Pre-fit plots in μ channel(inclusive)



- Slight overestimation in all τ ID selection.

τ ID scale factor: likelihood fits strategy

➢ Fit result(inclusive)



	τ ID scale factor	
	1p	3p
Loose	1.15+-0.01	1.21+-0.02
Medium	1.18+-0.01	1.29+-0.03
Tight	1.21+-0.01	1.31+-0.03

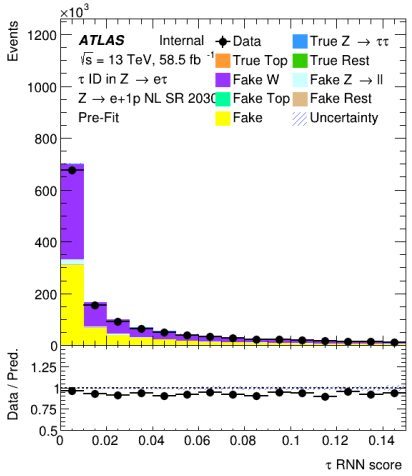
- Small NL norm factor due to slight overshoot.
- Leads to a large post-fit-efficiency, and make SF>1.
- **the NL NFs are unstable due to very small signal contribution**

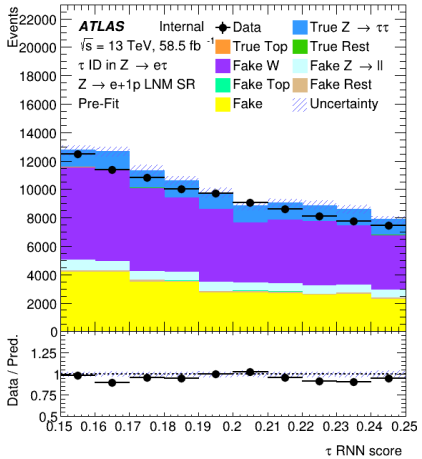
	τ ID scale factor(set NL=1)	
	1p	3p
Loose	0.99+-0.09	0.96+-0.13
Medium	1.01+-0.09	1.03+-0.14
Tight	1.04+-0.10	1.05+-0.15

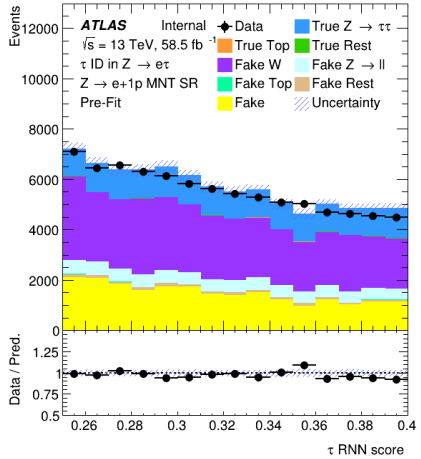
- **set NL NF = 1** to calculate τ ID scale factor(assign **50%** uncertainty to fixed-not-loose yields).

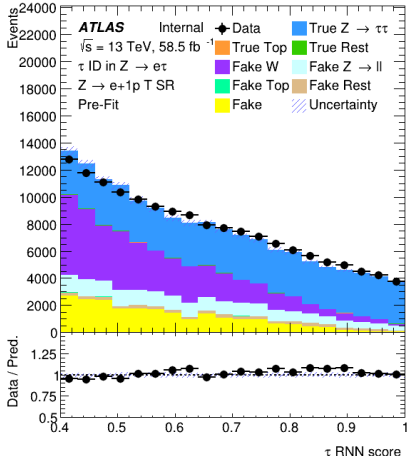
τ ID scale factor: likelihood fits strategy

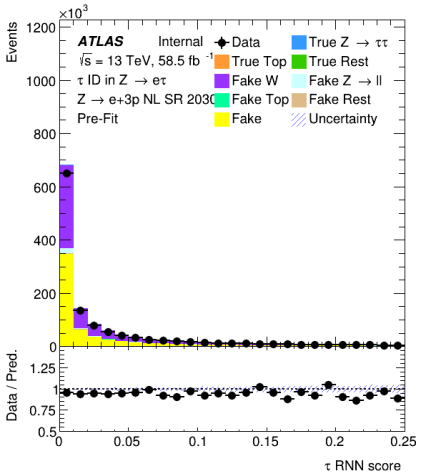
➢ Pre-fit plots in e channel(pt:20GeV-30GeV)

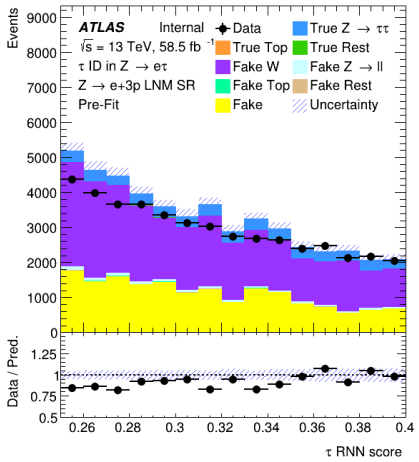


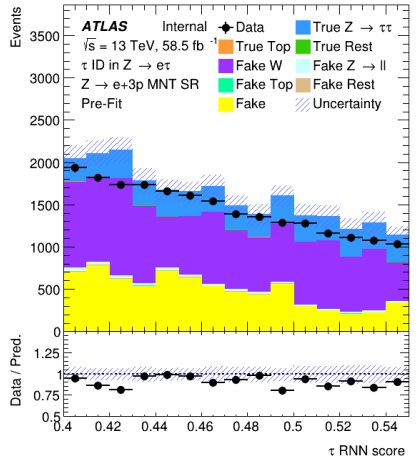


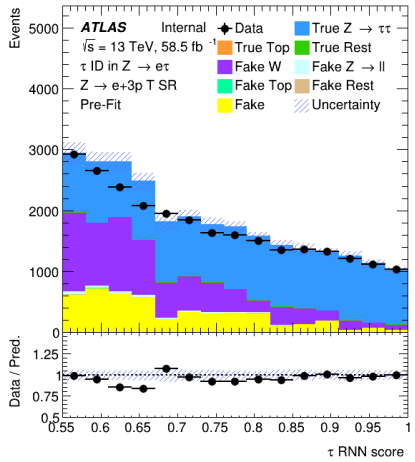








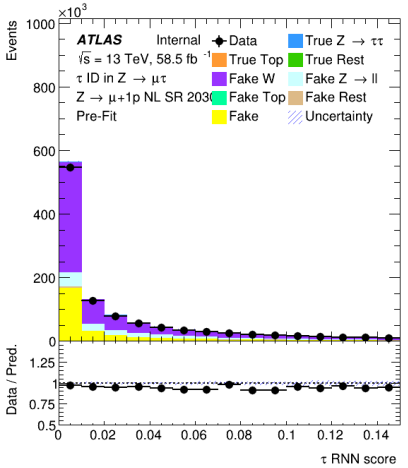


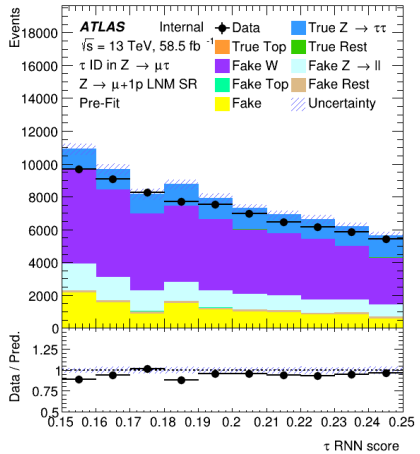


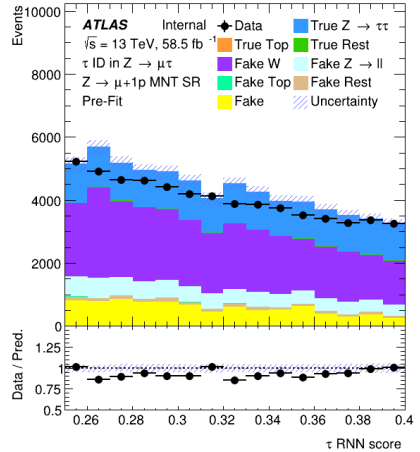
- Slight overestimation in NL, LNM, MNT seclection.

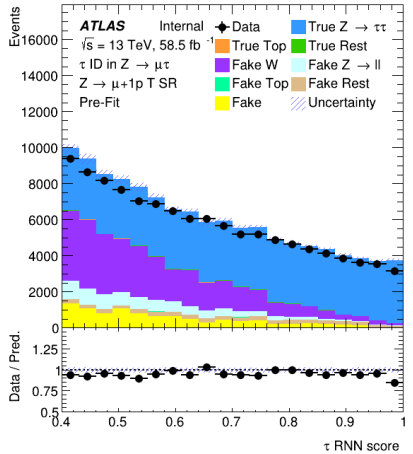
τ ID scale factor: likelihood fits strategy

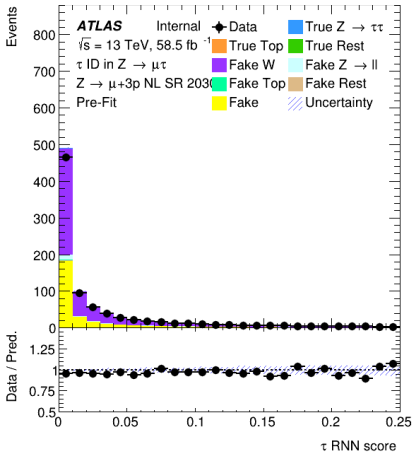
➢ Pre-fit plots in μ channel(pt:20GeV-30GeV)

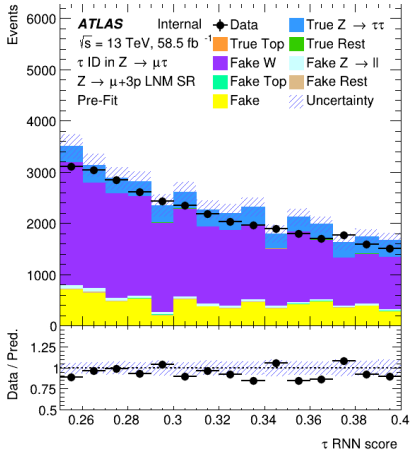


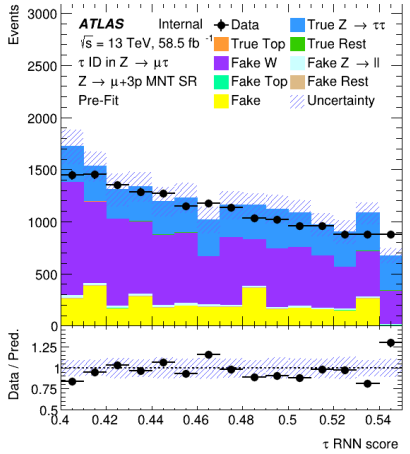


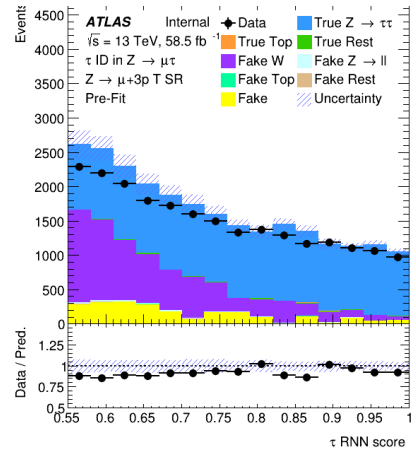








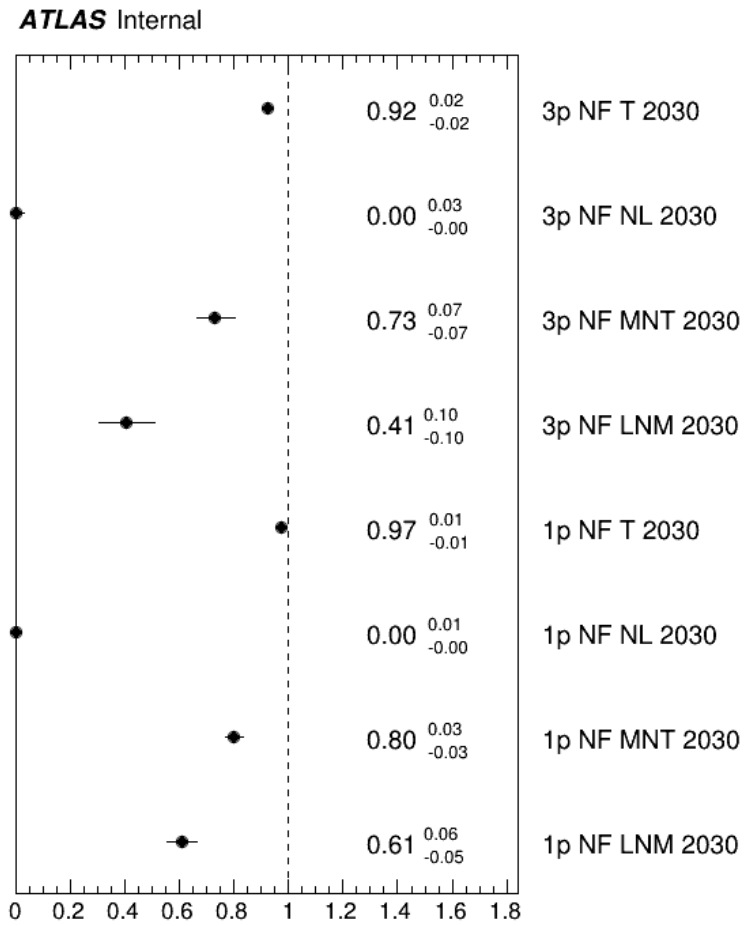




- Slight overestimation in all τ ID seclection

τ ID scale factor: likelihood fits strategy

➢ Fit result(pt:20GeV-30GeV)



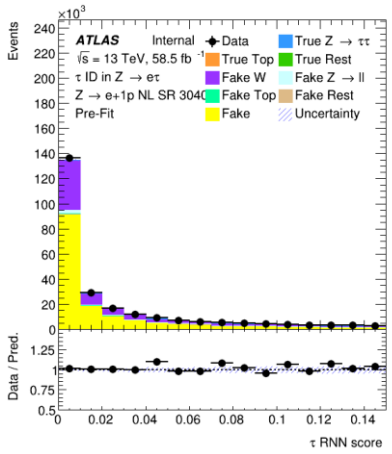
	τ ID scale factor	
	1p	3p
Loose	1.22+-0.00	1.34+-0.01
Medium	1.27+-0.01	1.50+-0.03
Tight	1.32+-0.01	1.57+-0.03

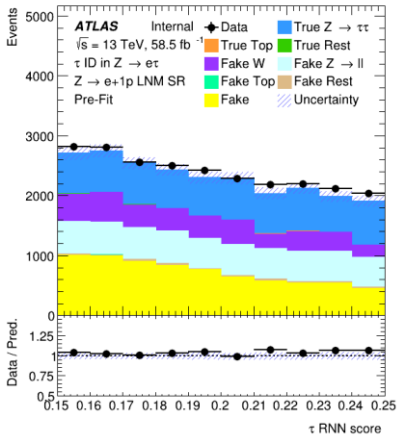
	τ ID scale factor(set NL=1)	
	1p	3p
Loose	0.98+-0.10	0.93+-0.14
Medium	1.02+-0.10	1.04+-0.16
Tight	1.06+-0.10	1.09+-0.17

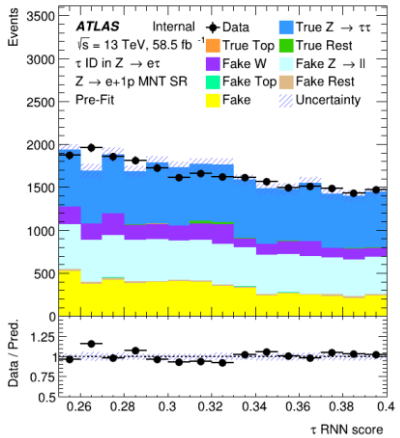
- Small NL norm factor leads to a large post-fit-efficiency, and make SF>1.

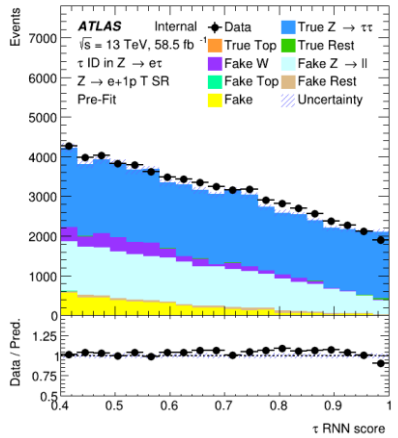
τ ID scale factor: likelihood fits strategy

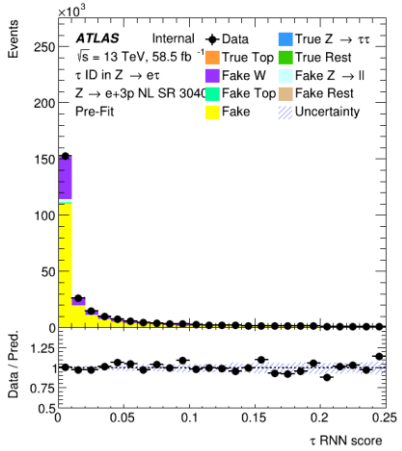
➢ Pre-fit plots in e channel(pt:30GeV-40GeV)

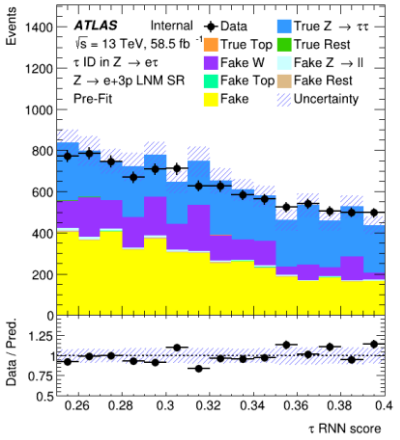


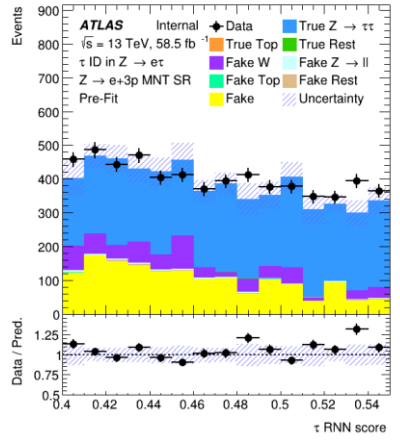


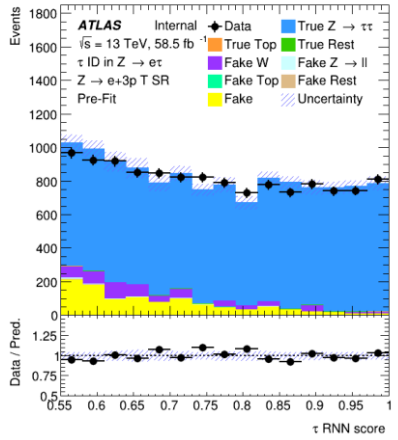








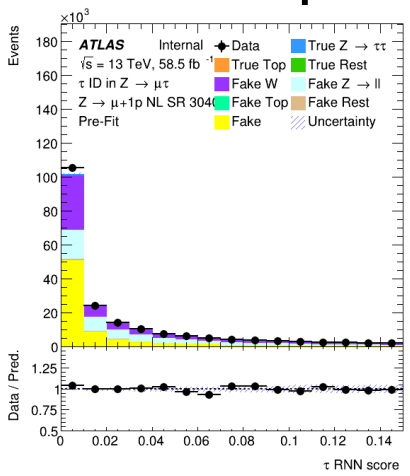


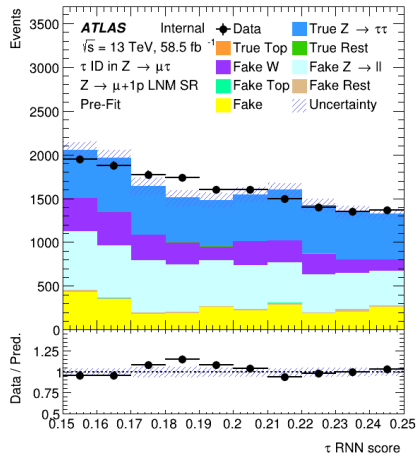


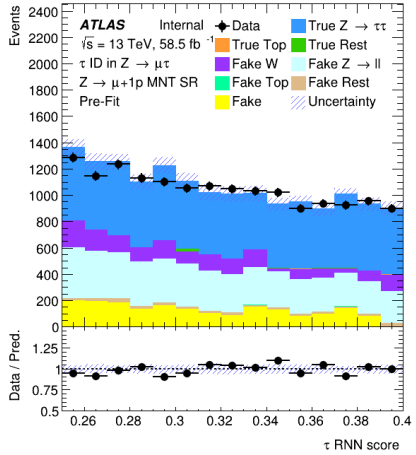
- Modeling looks well

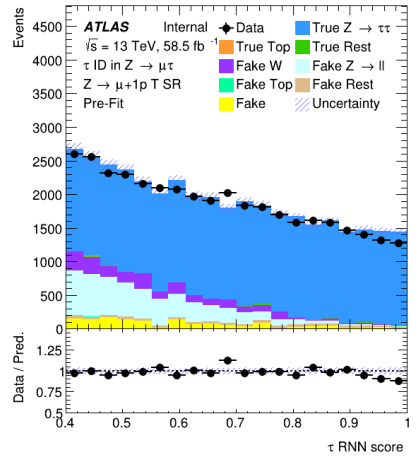
τ ID scale factor: likelihood fits strategy

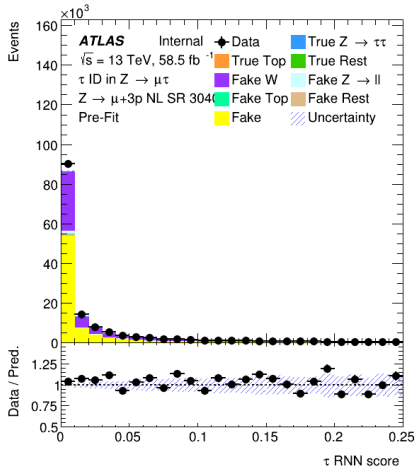
➢ Pre-fit plots in μ channel(pt:30GeV-40GeV)

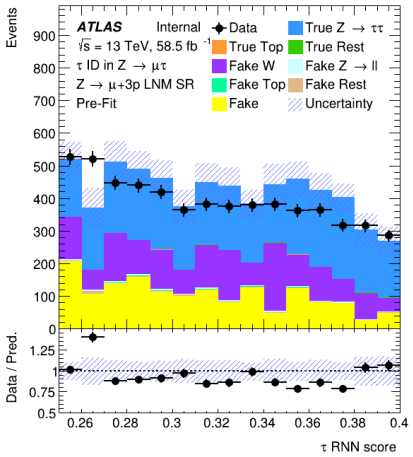


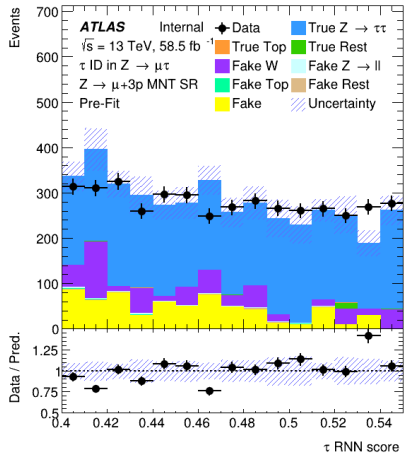


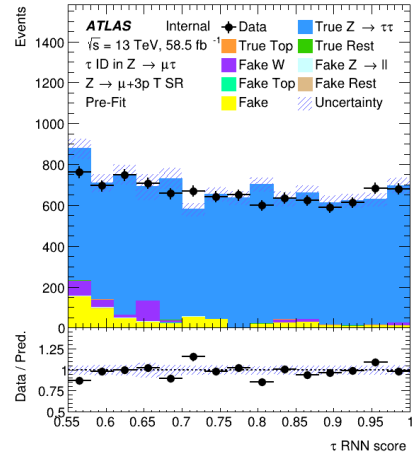






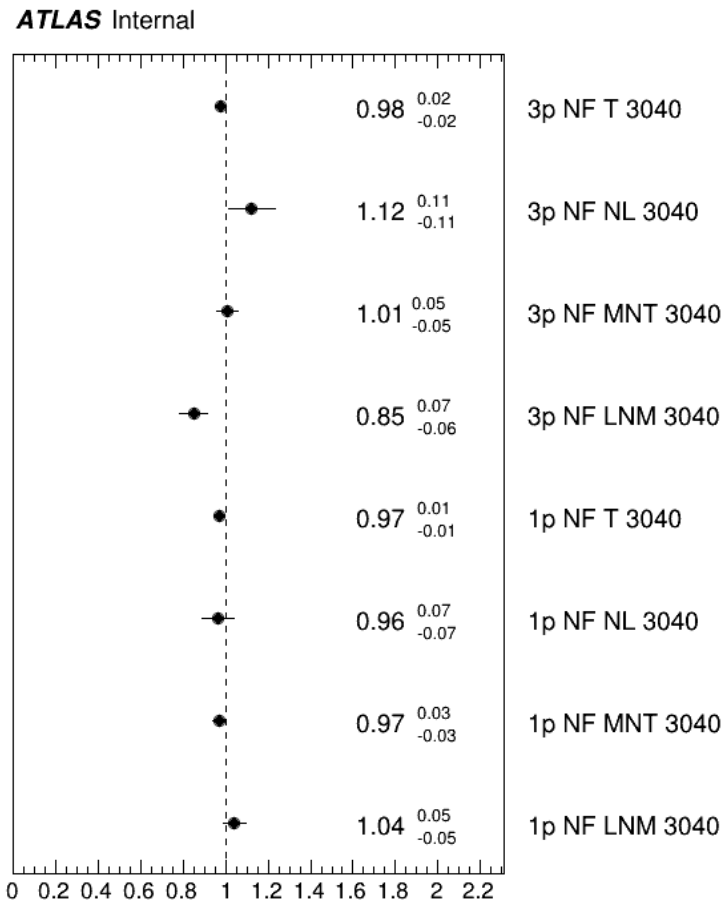






τ ID scale factor: likelihood fits strategy

➢ Fit result(pt:30GeV-40GeV)



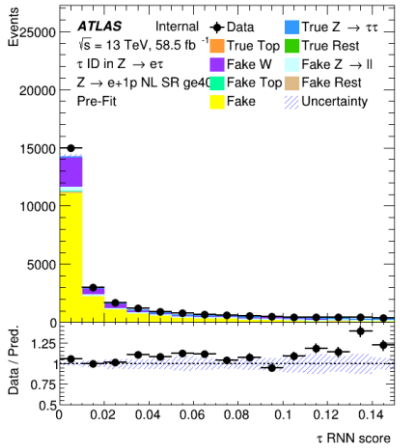
	τ ID scale factor	
	1p	3p
Loose	0.99+-0.01	0.98+-0.01
Medium	0.98+-0.01	0.99+-0.01
Tight	0.99+-0.01	0.98+-0.01

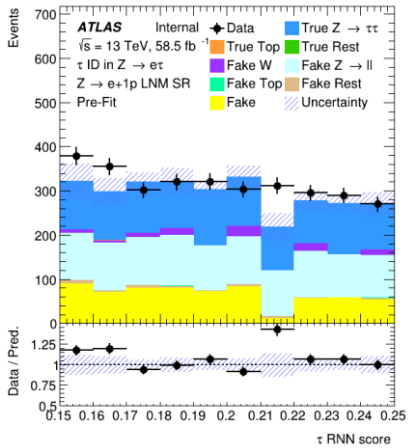
	τ ID scale factor(set NL=1)	
	1p	3p
Loose	1.00+-0.08	1.00+-0.12
Medium	0.99+-0.08	1.01+-0.13
Tight	1.00+-0.08	1.00+-0.12

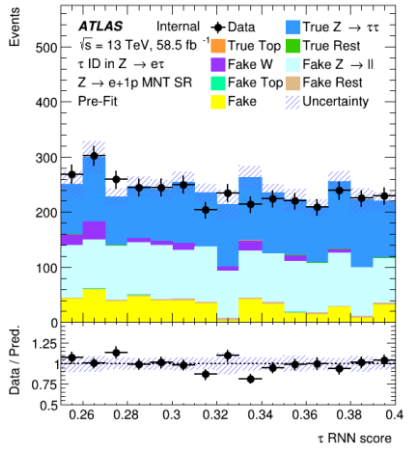
- Norm factors are all close to one.
- τ ID scale factors are close to one for all τ ID both in 1p and 3p.

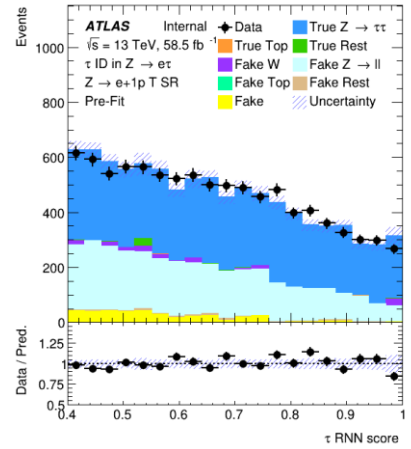
τ ID scale factor: likelihood fits strategy

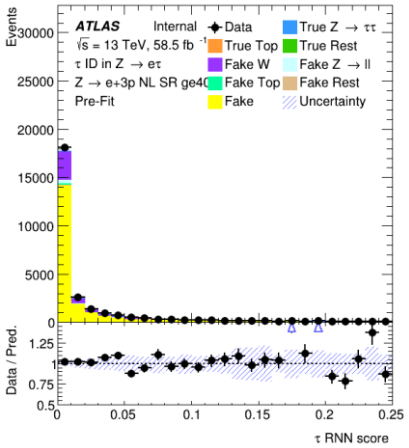
➢ Pre-fit plots(pt:>40GeV) in e channel

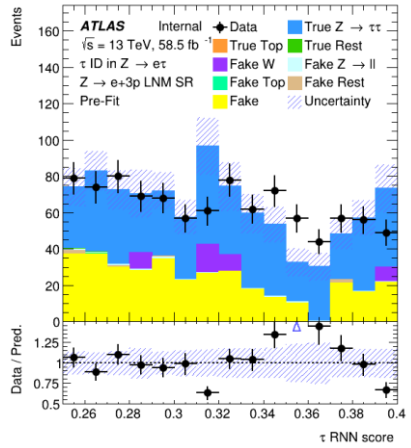


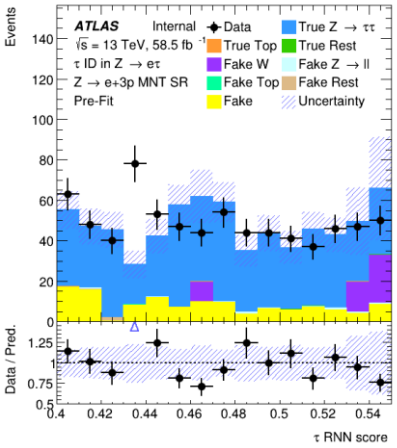


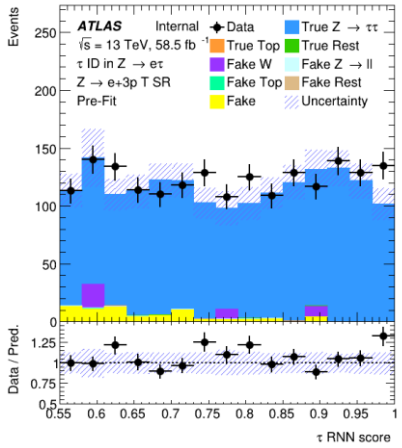








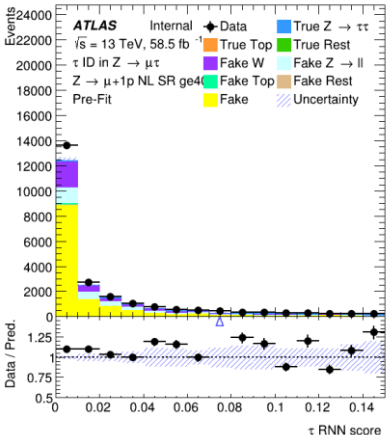


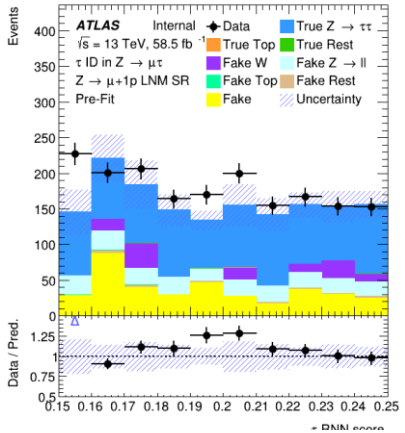


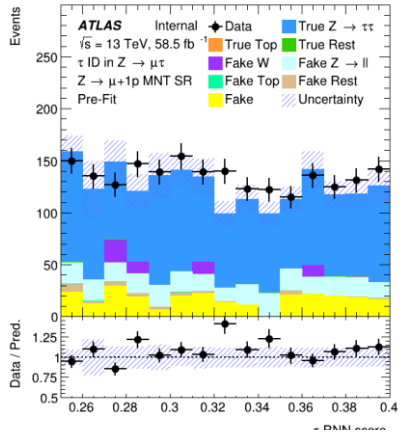
- Statistic in 3prong channel is a little limited.

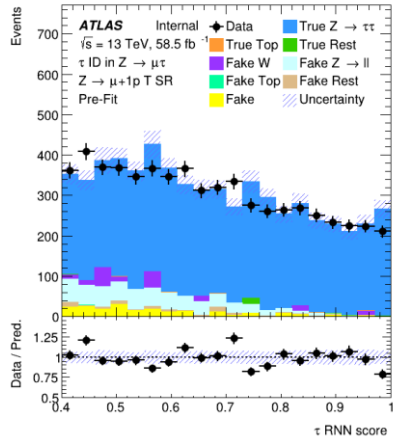
τ ID scale factor: likelihood fits strategy

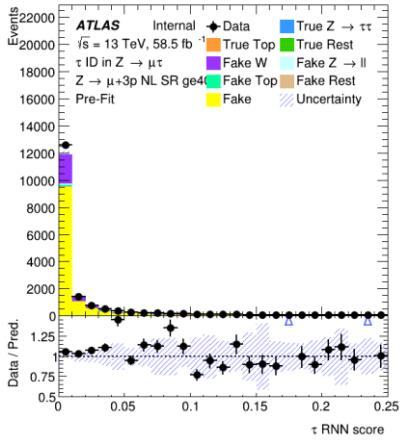
➢ Pre-fit plots(pt:>40GeV) in μ channel

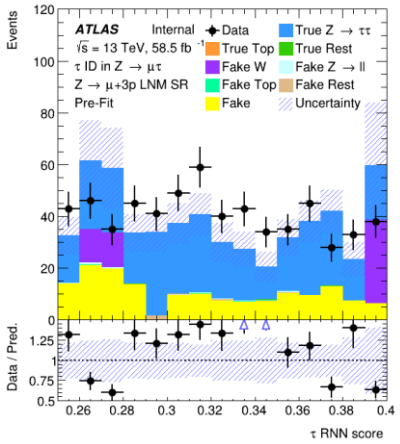


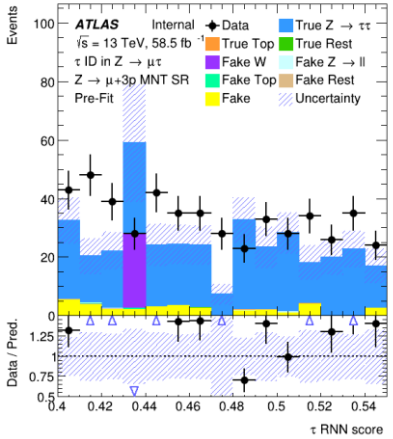


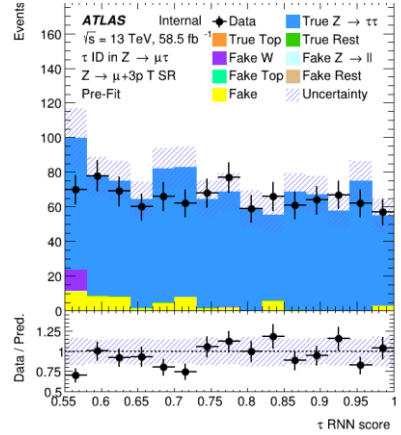








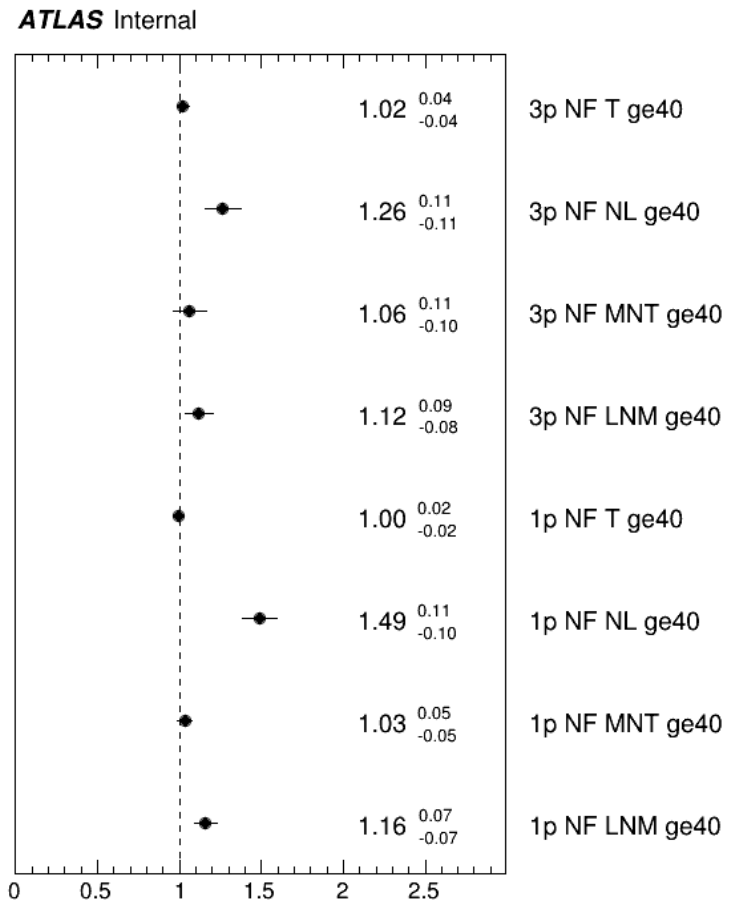




- Statistic is limited especially in 3prong channel.
- Underestimation in NL, LNM, MNT selection

τ ID scale factor: likelihood fits strategy

➢ Fit result(pt:>40GeV)



	τ ID scale factor	
	1p	3p
Loose	0.93+-0.01	0.95+-0.02
Medium	0.91+-0.01	0.94+-0.02
Tight	0.90+-0.01	0.91+-0.03

	τ ID scale factor(set NL=1)	
	1p	3p
Loose	1.00+-0.08	1.02+-0.12
Medium	0.98+-0.08	1.00+-0.12
Tight	0.97+-0.07	0.97+-0.12

- τ ID scale factors are less than one due to larger-than-one NL norm factors. Underestimation occur in NL selection.

τ ID scale factor: background subtraction strategy

- Background subtraction strategy to extract tau ID scale factor in Z->τ τ channel

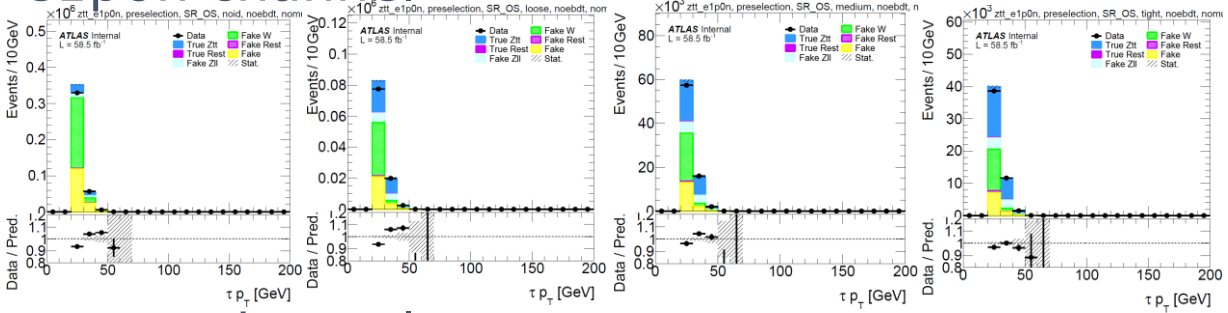
$$\text{scale factor} = \frac{eff_{data}}{eff_{MC}} \frac{\frac{(\text{data} - \text{background})^{\text{pass the ID}}}{(\text{data} - \text{background})^{\text{all ID}}}}{\frac{\text{true}\,\tau^{\text{pass the ID}}}{\text{true}\,\tau^{\text{all ID}}}}$$

- Signal: All true τ
- Background Estimation
  - background= QCD+ W + other fake process MC
  - $QCD=R_{QCD}*(SS^{data}-SS^{MC}_{W}*k^{SS}_{W}\text{-other MC})$
  - $W=k^{OS}_{W}*OS^{MC}_{W}$
  - R_{QCD} is from QCD CR, $R_{QCD}=\frac{(\text{data} - \text{MC})^{OS}}{(\text{data} - \text{MC})^{ss}}$, $k^{OS}_{W}(k^{SS}_{W})$ is from WCR OS(SS), $k=\frac{\text{data} - \text{MC}_{other}}{\text{MC}_{W}}$

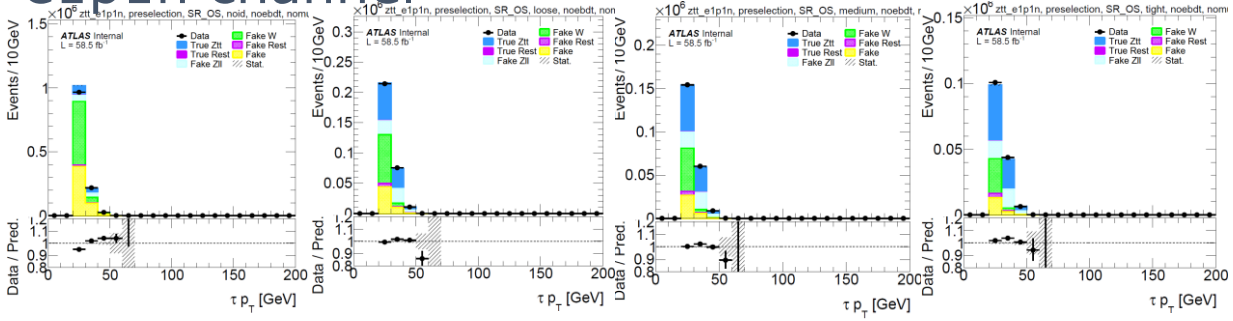
τ ID scale factor: background subtraction strategy

➢ tau pt distribution in e channel

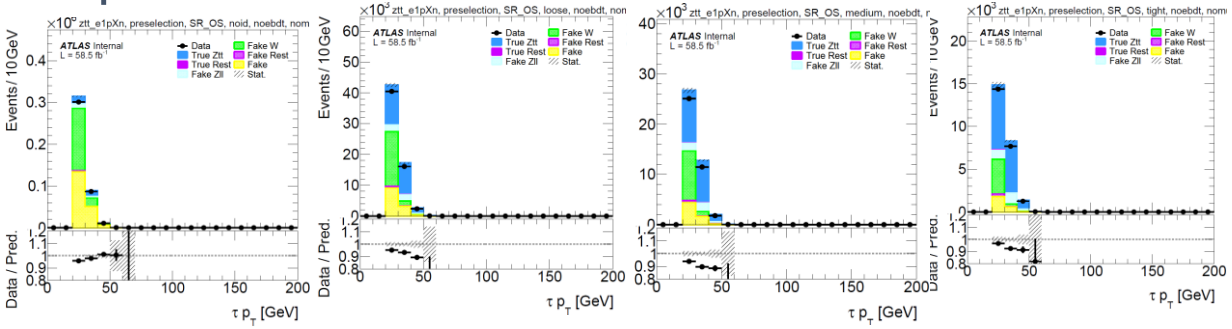
e1p0n channel



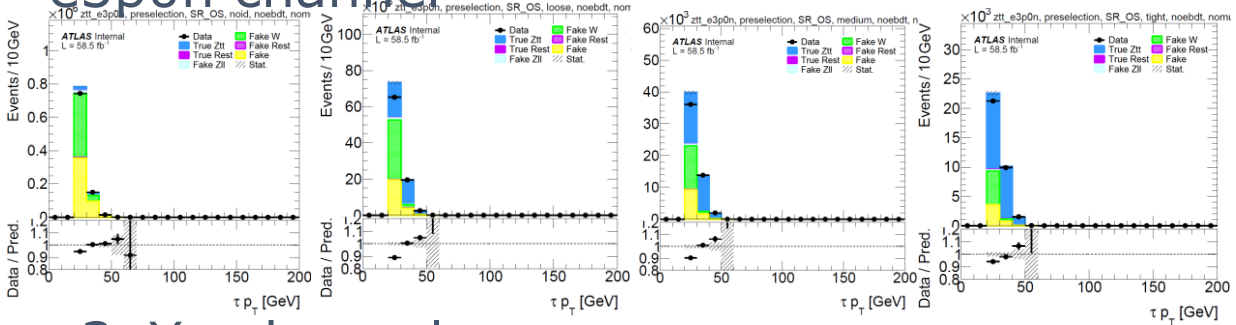
e1p1n channel



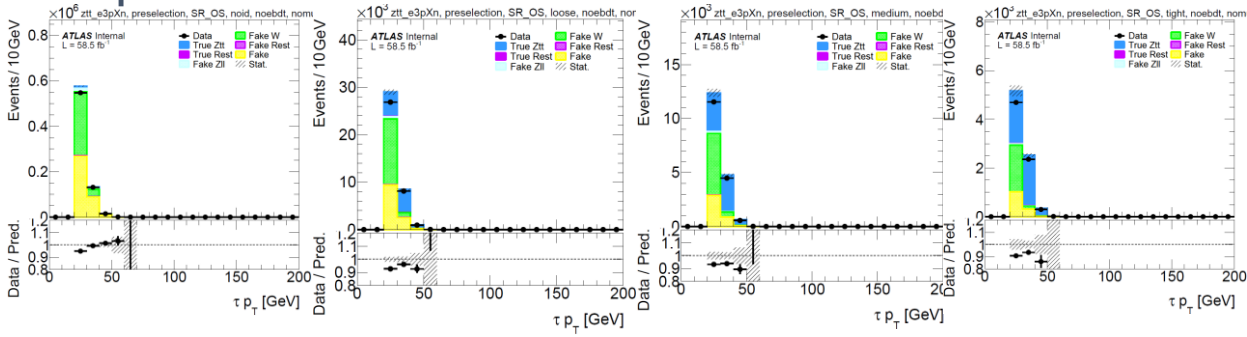
e1pXn channel



e3p0n channel



e3pXn channel



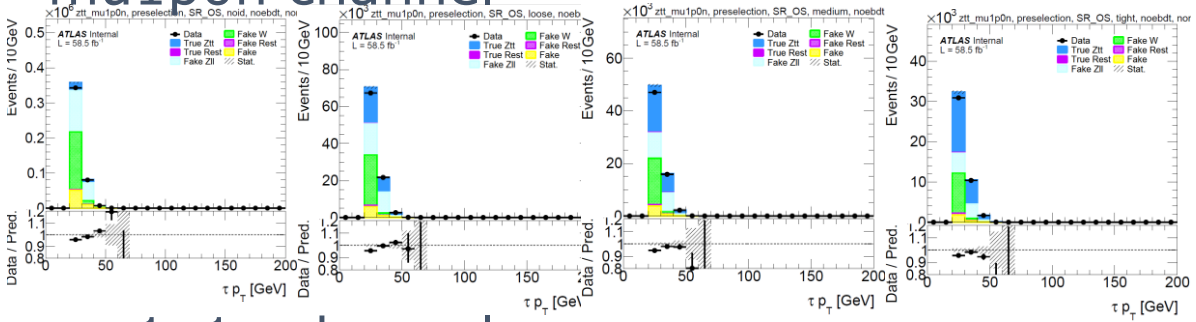
- The data/total-predicted-process ratio increase as τ pt increase in e1p channel.
- Over-estimation in e3p channel.

- more kinamatic distributions are shown in back-up

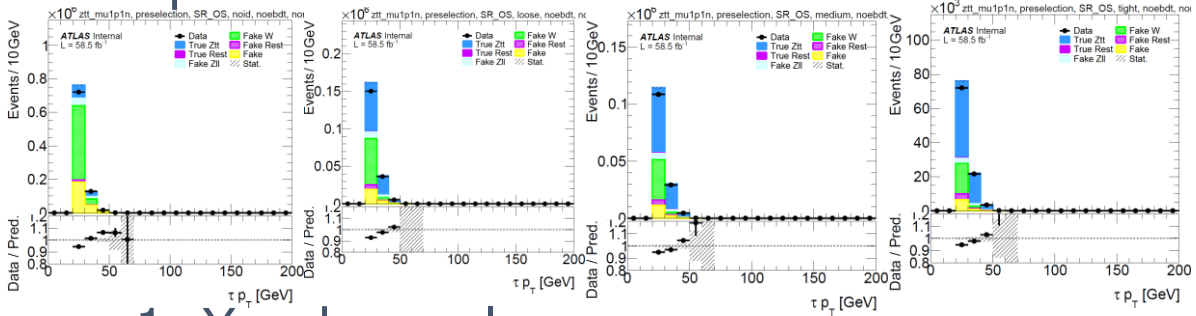
τ ID scale factor: background subtraction strategy

➢ tau pt distribution in mu channel

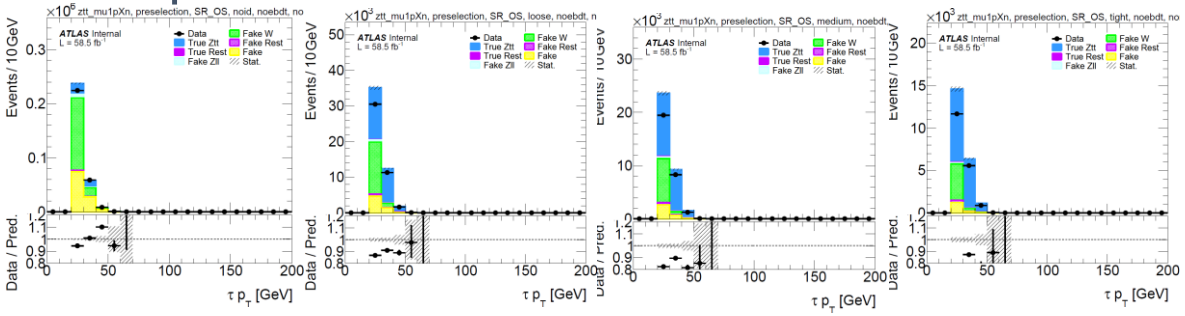
mu1p0n channel



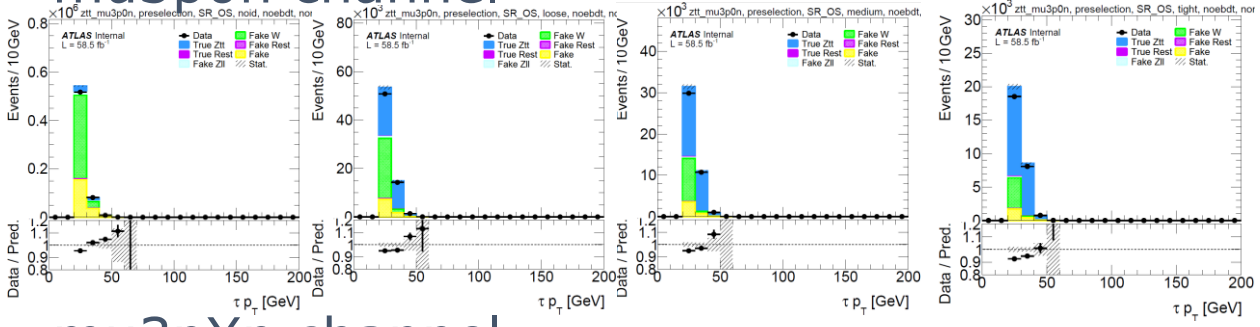
mu1p1n channel



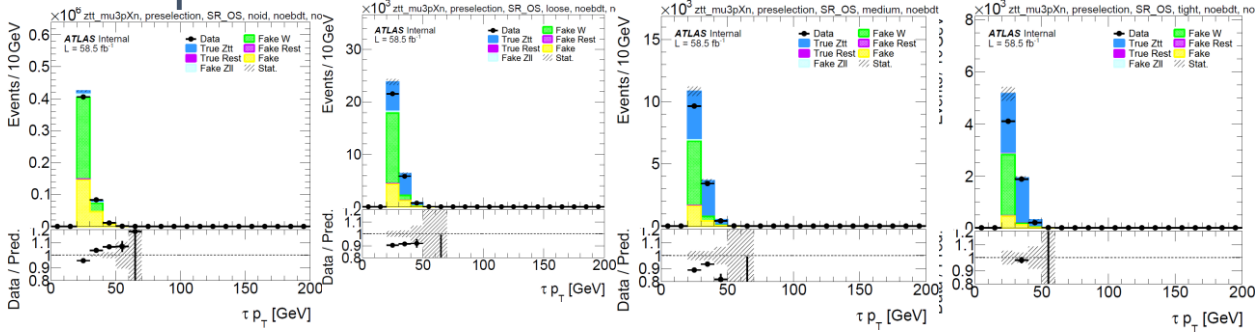
mu1pXn channel



mu3p0n channel



mu3pXn channel



- Modeling show a bit over-estimation
- Over-estimation increase as tau-ID get tighter in mu1pXn and mu3pXn channel.

τ ID scale factor: background subtraction strategy

- Divide process into signal/background
- Signal: “True Ztt”,”True Rest”
- Background: ”Fake Zll ”,”Fake W”,”Fake Rest”,”Fake”

- Combine the channel into e1p,e3p,mu1p,mu3p：

  e1p: e1p0n, e1p1n, e1pXn

  e3p: e3p0n, e3pXn

  mu1p: mu1p0n, mu1p1n, mu1pXn

  mu3p: mu3p0n, mu3pXn

τ ID scale factor: background subtraction strategy

- Signal, background yields, data in e1p,e3p,mu1p,mu3p channel

sig	bkg	data	data-bkg	sig_err	bkg_err	tauID	channel
190547.87	1894271.49	2001394.00	107122.51	605.15	4616.77	noid	e1p
157303.94	310172.87	460081.00	149908.13	547.31	1676.72	loose	e1p
136933.88	203731.35	337573.00	133841.65	511.19	1285.37	medium	e1p
107549.55	117636.39	226119.00	108482.61	452.82	939.10	tight	e1p
61778.58	1609888.18	1603684.00	-6204.18	346.00	4119.48	noid	e3p
45868.99	87927.46	123375.00	35447.54	296.61	1066.67	loose	e3p
36860.63	36430.21	68574.00	32143.79	264.09	684.78	medium	e3p
27924.53	14252.64	40061.00	25808.36	230.11	435.82	tight	e3p
184047.33	1472187.19	1589110.00	116922.81	592.31	4281.99	noid	mu1p
151532.86	197298.63	327640.00	130341.37	535.73	1458.08	loose	mu1p
132069.03	119394.66	236557.00	117162.34	499.67	1105.70	medium	mu1p
104487.29	63846.21	158323.00	94476.79	444.26	777.37	tight	mu1p
57456.26	1086940.02	1107803.00	20862.98	331.77	3929.78	noid	mu3p
43456.33	57049.65	94452.00	37402.35	286.89	1010.67	loose	mu3p
35000.91	23465.75	55139.00	31673.25	257.53	675.58	medium	mu3p
26439.86	10244.80	33566.00	23321.20	223.97	461.74	tight	mu3p

Negative (data-bkg) in e3p channels when no tau-ID requirement

τ ID scale factor: background subtraction strategy

➢ Scale factors results

scale factor	eff_MC	eff_Data	tauID	channel
1.70 +- 0.08	0.83	1.40	loose	e1p
1.74 +- 0.08	0.72	1.25	medium	
1.79 +- 0.08	0.56	1.01	tight	
-7.70 +- 5.35	0.74	-5.71	loose	e3p
-8.68 +- 6.04	0.60	-5.18	medium	
-9.20 +- 6.40	0.45	-4.16	tight	
1.35 +- 0.05	0.82	1.11	loose	mu1p
1.40 +- 0.06	0.72	1.00	medium	
1.42 +- 0.06	0.57	0.81	tight	
2.37 +- 0.47	0.76	1.79	loose	mu3p
2.49 +- 0.49	0.61	1.52	medium	
2.43 +- 0.48	0.46	1.12	tight	

- Negative tau ID scale factors in e3p channels
- Tau ID scale factors in other channel are all > 1.
- tau ID scale factors in mu3p get large uncertainty because of lack of statistic

$$\text{SF error} = \sqrt{\frac{err(eff_{data})^2}{eff_{MC}^2} + err(eff_{MC})^2 * \frac{eff_{data}^2}{eff_{MC}^4}}$$

τ ID scale factor: background subtraction strategy (set $\frac{Data}{Total\,Predicted\,Process}$ =1for NL)

- Signal, background yields, data in e1p,e3p,mu1p,mu3p channel (set Data - Bkg = Signal for not_loose case)

sig	bkg	data	data-bkg	sig_err	bkg_err	tauID	channel
190547.87	1894271.49	2077423.55	183152.06	605.1474547	4616.766525	noid	e1p
157303.94	310172.87	460081	149908.13	547.3130892	1676.716782	loose	e1p
136933.88	203731.35	337573	133841.65	511.189469	1285.374651	medium	e1p
107549.55	117636.39	226119	108482.61	452.8240291	939.1023706	tight	e1p
61778.58	1609888.18	1661245.31	51357.13	346.0022003	4119.483225	noid	e3p
45868.99	87927.46	123375	35447.54	296.6068007	1066.665651	loose	e3p
36860.63	36430.21	68574	32143.79	264.087229	684.7818799	medium	e3p
27924.53	14252.64	40061	25808.36	230.1133334	435.8154223	tight	e3p
184047.33	1472187.19	1635043.03	162855.84	592.3072065	4281.992014	noid	mu1p
151532.86	197298.63	327640	130341.37	535.7315148	1458.07797	loose	mu1p
132069.03	119394.66	236557	117162.34	499.6700381	1105.69823	medium	mu1p
104487.29	63846.21	158323	94476.79	444.2641245	777.368384	tight	mu1p
57456.26	1086940.02	1138342.3	51402.28	331.7747325	3929.780227	noid	mu3p
43456.33	57049.65	94452	37402.35	286.8918078	1010.673013	loose	mu3p
35000.91	23465.75	55139	31673.25	257.5311672	675.5799335	medium	mu3p
26439.86	10244.8	33566	23321.2	223.9650305	461.7367606	tight	mu3p

τ ID scale factor: background subtraction strategy(set $\frac{Data}{TotalPredictedProcess}$ =1for NL)

➢ Scale factors results(set Data - Bkg = Signal for not_loose case)

scale factor	eff_MC	eff_Data	tauID	channel
0.99 +- 0.03	0.83	0.82	loose	
1.02 +- 0.03	0.72	0.73	medium	e1p
1.05 +- 0.03	0.56	0.59	tight	
0.93 +- 0.08	0.74	0.69	loose	
1.05 +- 0.09	0.60	0.63	medium	e3p
1.11 +- 0.10	0.45	0.50	tight	
0.97 +- 0.03	0.82	0.80	loose	
1.00 +- 0.03	0.72	0.72	medium	mu1p
1.02 +- 0.03	0.57	0.58	tight	
0.96 +- 0.08	0.76	0.73	loose	
1.01 +- 0.08	0.61	0.62	medium	mu3p
0.99 +- 0.08	0.46	0.45	tight	

- τ ID scale factors are close to one after fixing NL-data-TotalPredictedProcess to 1.
- τ ID scale factors are quite close to likelihood-fits-results.
- Still got large uncertainty in 3 prong channel.
- The modeling did affect τ ID scale factors in background subtraction strategy.

Separating e,m u channel does not improve the precision of τ ID scale factors.

- Combine e,mu channel to get τ ID scale factors is better?

Likelihood fits results

tau ID SF	e1p	e3p	mu1p	mu3p	e+mu combined 1p	e+mu combined 3p
Loose	1.19+-0.02	1.35+-0.02	1.12+-0.02	1.07+-0.04	1.15+-0.01	1.21+-0.02
Medium	1.22+-0.02	1.45+-0.03	1.15+-0.02	1.13+-0.04	1.18+-0.01	1.29+-0.03
Tight	1.25+-0.02	1.52+-0.04	1.17+-0.02	1.12+-0.04	1.21+-0.01	1.31+-0.03

Likelihood fits results(set NL_NF=1)

tau ID SF	e1p	e3p	mu1p	mu3p	e+mu combined 1p	e+mu combined 3p
Loose	1.00+-0.09	0.96+-0.14	0.98+-0.09	0.97+-0.13	0.99+-0.09	0.96+-0.13
Medium	1.03+-0.09	1.03+-0.15	1.00+-0.10	1.02+-0.14	1.01+-0.09	1.03+-0.14
Tight	1.05+-0.09	1.08+-0.16	1.02+-0.10	1.01+-0.14	1.04+-0.10	1.05+-0.15