山东大学 SHANDONG UNIVERSITY

Department of Physics, Shandong University

Compressed EWK study(ISRC1N2)

Chengxin Liao
liaocx@ihep.ac.cn

Apr, Wed 15, 2025

Tasklist

- FF method var distribution check
- Multiclass result(failed)
- BSc thesis: https://www.overleaf.com/project/674e7119837a2580151a0868 (need to submit before the end of Apr)

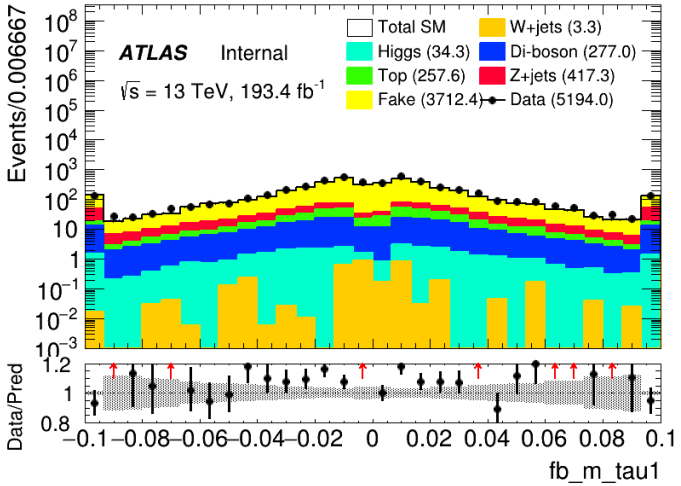
FF method

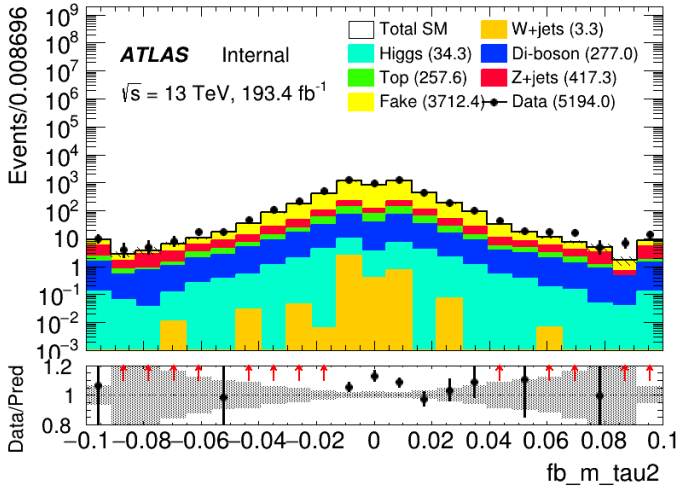
Pre-Selection

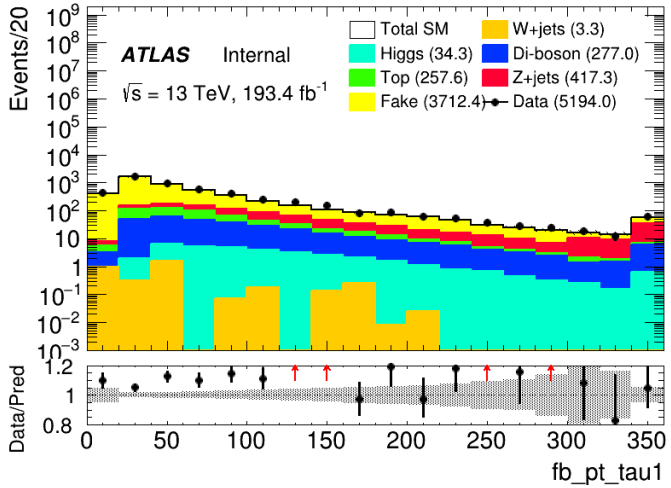
SR(2ID):
nBaseTau == 2, medium tau == 2
nBaseJets >= 1
MET trigger & MET >= 200
OS
bVeto
Mtt_reco <=40 && Mtt_reco >= 130

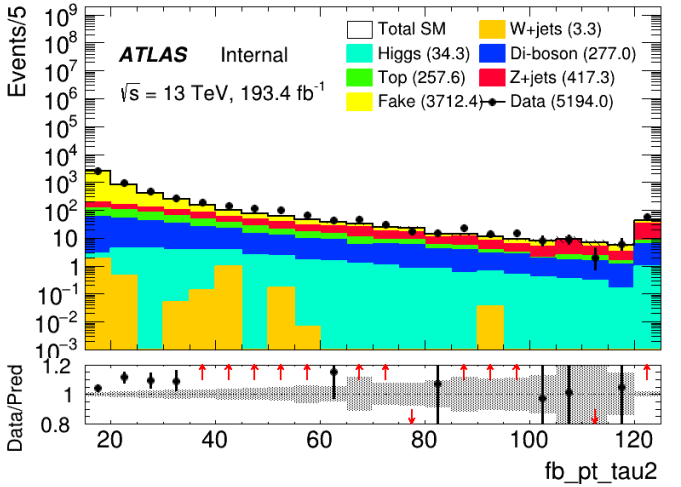
CR(ID):
nBaseTau == 1, medium tau == 1
nLeps >= 1
MET trigger & MET >= 200
bVeto

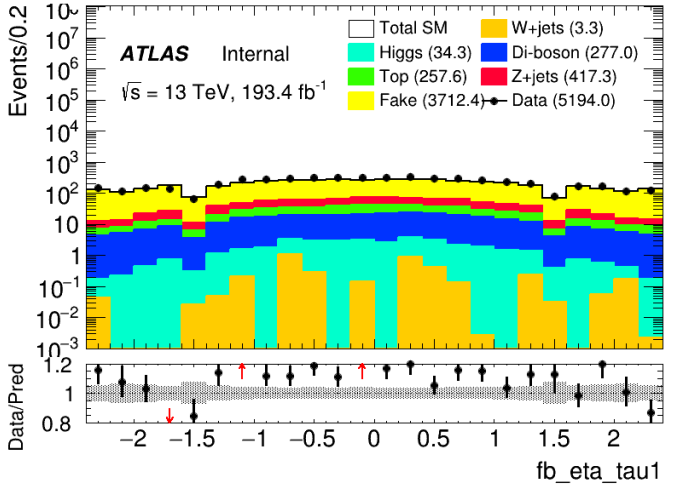
FF method

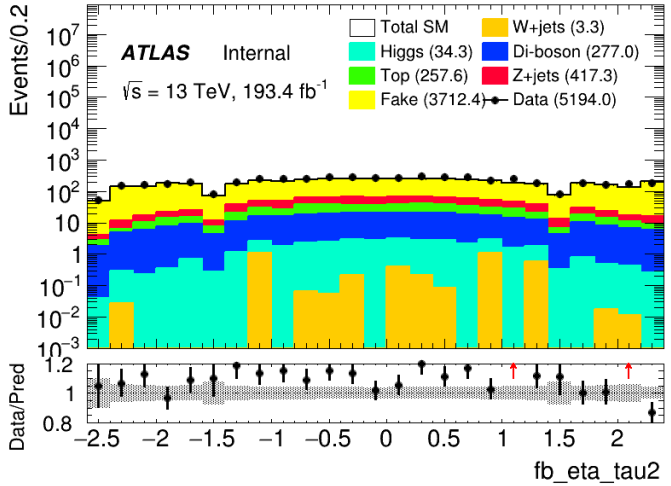

ATLAS Internal
√s = 13 TeV, 193.4 fb-1
Total SM
W+jets (3.3)
Higgs (34.3)
Di-boson (277.0)
Top (257.6)
Z+jets (417.3)
Fake (3712.4)
Data (5194.0)
Events/0.006667
Data/Pred
fb_m_tau1


ATLAS Internal
√s = 13 TeV, 193.4 fb-1
Total SM
W+jets (3.3)
Higgs (34.3)
Di-boson (277.0)
Top (257.6)
Z+jets (417.3)
Fake (3712.4)
Data (5194.0)
Events/0.008696
Data/Pred
fb_m_tau2


ATLAS Internal
√s = 13 TeV, 193.4 fb-1
Total SM
W+jets (3.3)
Higgs (34.3)
Di-boson (277.0)
Top (257.6)
Z+jets (417.3)
Fake (3712.4)
Data (5194.0)
Events/20
Data/Pred
fb_pt_tau1


ATLAS Internal
√s = 13 TeV, 193.4 fb-1
Total SM
W+jets (3.3)
Higgs (34.3)
Di-boson (277.0)
Top (257.6)
Z+jets (417.3)
Fake (3712.4)
Data (5194.0)
Events/5
Data/Pred
fb_pt_tau2


ATLAS Internal
√s = 13 TeV, 193.4 fb-1
Total SM
W+jets (3.3)
Higgs (34.3)
Di-boson (277.0)
Top (257.6)
Z+jets (417.3)
Fake (3712.4)
Data (5194.0)
Events/0.2
Data/Pred
fb_eta_tau1


ATLAS Internal
√s = 13 TeV, 193.4 fb-1
Total SM
W+jets (3.3)
Higgs (34.3)
Di-boson (277.0)
Top (257.6)
Z+jets (417.3)
Fake (3712.4)
Data (5194.0)
Events/0.2
Data/Pred
fb_eta_tau2

Multiclass

Try to use **optuna** to auto-optimize, but also failed

Test_auc = \sum{Test_auc_class}
Train_auc = \sum{Train_auc_class}

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)

hyperparameter setting(Grid Search)

Ntrees: [100, 200, 300, 400, 500]
MaxDepth: [4, 6, 8, 10]
MinNode: [1, 3, 5, 7]
Learning rate: [0.01, 0.03, 0.05, 0.8, 0.1]
400 models

If we don't require AUC constraint, so model will be like this

the accuracy for that even lower than threw coin

Severe overfit

Dataset Name:	MVA Method / Class:	ROC AUC test (train)	Sig eff@B=0.01 test (train)	Sig eff@B=0.10 test (train)	Sig eff@B=0.30 test (train)
mva_bkg_est	BDT0_500_8_1_01				
	Other_bkg	0.898 (0.947)	0.597 (0.332)	0.770 (0.853)	0.799 (0.972)
	Top	0.582 (0.895)	0.017 (0.166)	0.185 (0.669)	0.407 (0.916)
	VV	0.598 (0.802)	0.064 (0.111)	0.266 (0.446)	0.480 (0.750)
	Fake	0.428 (0.749)	0.003 (0.062)	0.034 (0.326)	0.154 (0.671)
	Zjets	0.947 (0.844)	0.341 (0.092)	0.832 (0.501)	0.978 (0.829)
	C1N2	0.895 (0.963)	0.349 (0.550)	0.707 (0.889)	0.888 (0.984)

```
=== Best Trial ===
Params: {'NTrees': 500, 'MaxDepth': 8, 'MinNodeSize': 1, 'Shrinkage': 0.1}
Value: 0.6821346409864086
Number of finished trials: 400
Results saved to grid_search_results.csv
```

Dataset Name:	MVA Method / Class:	ROC AUC test (train)	Sig eff@B=0.01 test (train)	Sig eff@B=0.10 test (train)	Sig eff@B=0.30 test (train)
mva_bkg_est	BDT0_100_4_7_003				
	Other_bkg	0.457 (0.500)	0.009 (0.013)	0.099 (0.104)	0.279 (0.301)
	Top	0.500 (0.510)	0.008 (0.008)	0.101 (0.103)	0.311 (0.319)
	VV	0.497 (0.497)	0.010 (0.008)	0.099 (0.094)	0.296 (0.297)
	Fake	0.503 (0.494)	0.010 (0.010)	0.102 (0.100)	0.300 (0.292)
	Zjets	0.501 (0.503)	0.011 (0.009)	0.099 (0.095)	0.304 (0.296)
	C1N2	0.492 (0.496)	0.014 (0.013)	0.094 (0.103)	0.289 (0.295)

Backup

Bkg decay mode

Wjets: W->e/muon + nu
W->tau+nu(can contribute true tau_had)
jet misidentified to a fake tau

Zjets: Z->ll/tautau
jet misindentified to fake tau

Top: top->W+b, W can contribute a true tau_had
b-quark is a source of fake

VV: W/Z

LH channel: $\geq 1tau,\ \geq 1lep$
Wjets: W contribute lep, jets misidentified to fake
Zjets:
SingleTop: W contribute lep, b-quark misidentified to fake
VV:

HH channel: $\geq 2tau,\ == 0lep$
Wjets: W contribute tau_had, plus a fake tau
Zjets: Z->tautau(had) or 2 fake tau
SingleTop: W contribute a tau_had, plus a fake tau
VV: