variables for VBF category

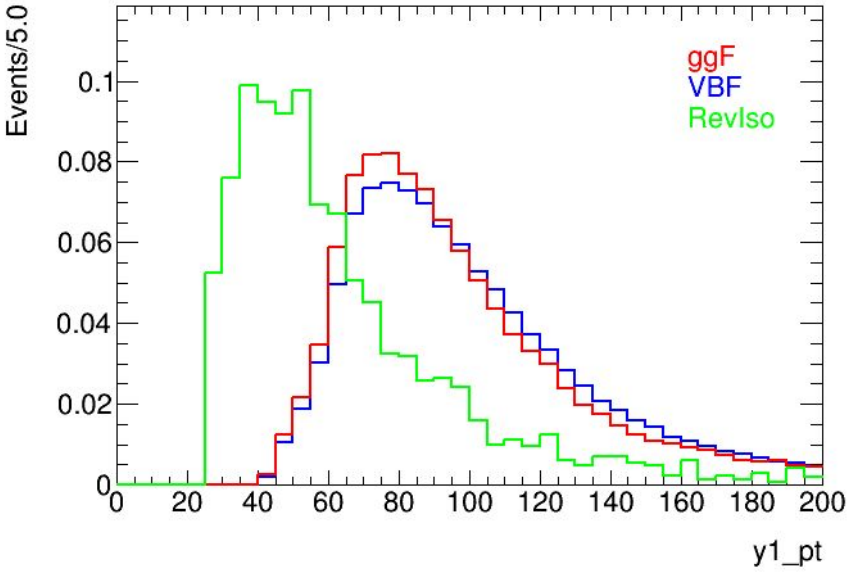
Yu Zhang

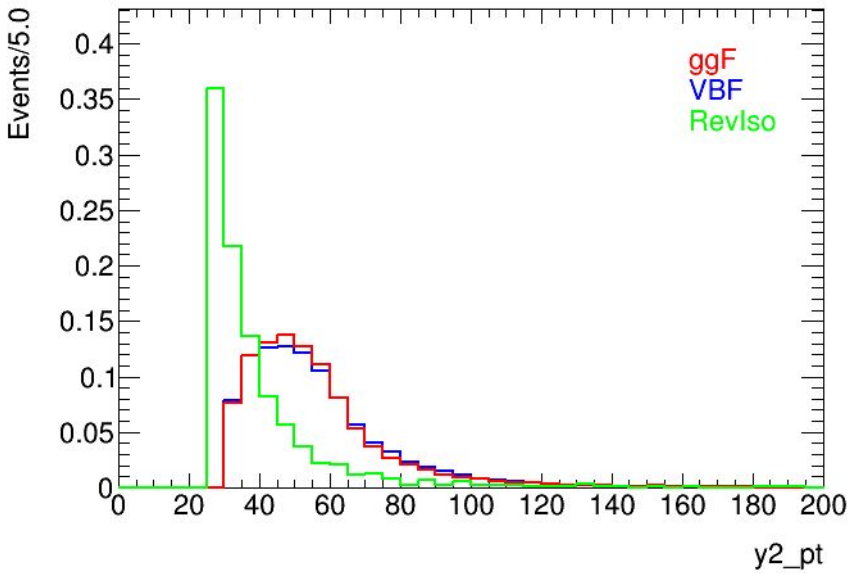
08-31

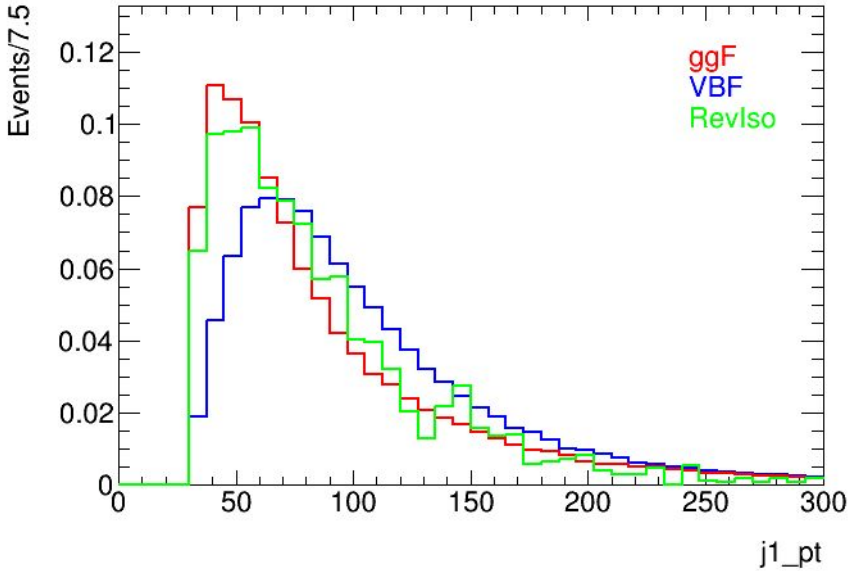
introduction

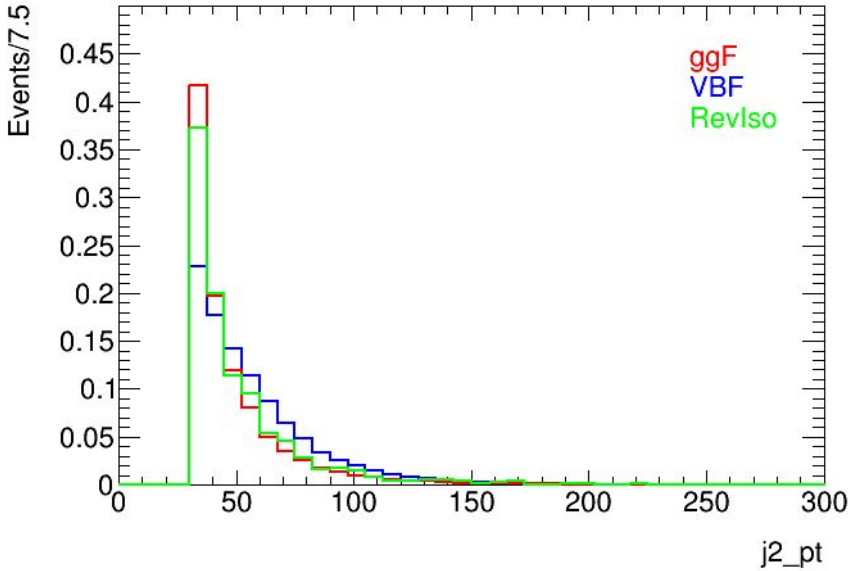
- move to official MxAOD
- selection:Pre,PID,Iso,RelPt,Mass (jet_pt>25GeV,run1 is 30GeV)
- ggF,VBF:use 500k events
- RevIso:samples in data fail Isolation :2610(1737,30GeV)
- data after Njets>2:179 (114 if 30GeV)
- not use diphoton pythia sample,gamgam+jets sherpa sample is not ready(only 25ns AOD ready,please help me check that)
- the following plot do not show data passing all selection(due to few events) and the jet pt threshold is 30GeV
- jj_m , jj_DeltaEta , yy_jj_DeltaPhi , yy_pTt , y_j_DeltaR_min , eta_Zeppenfeld

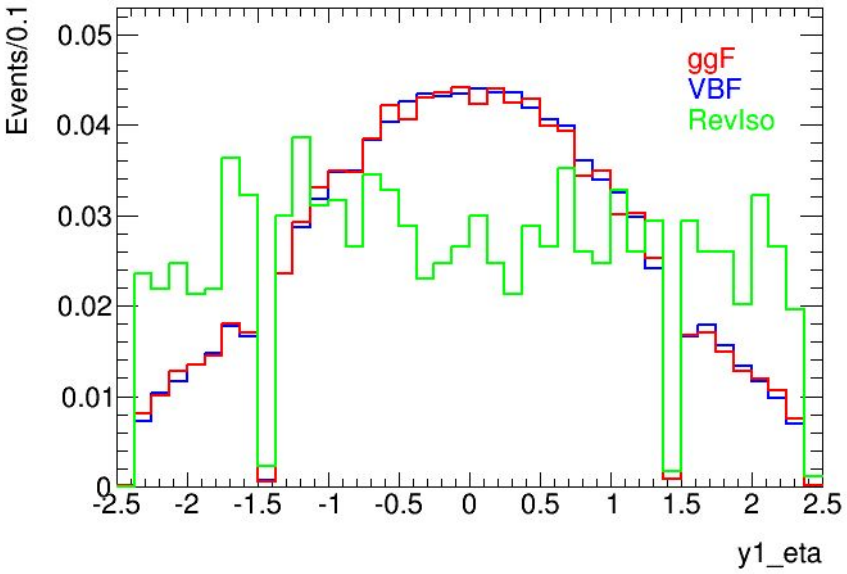
kinematics

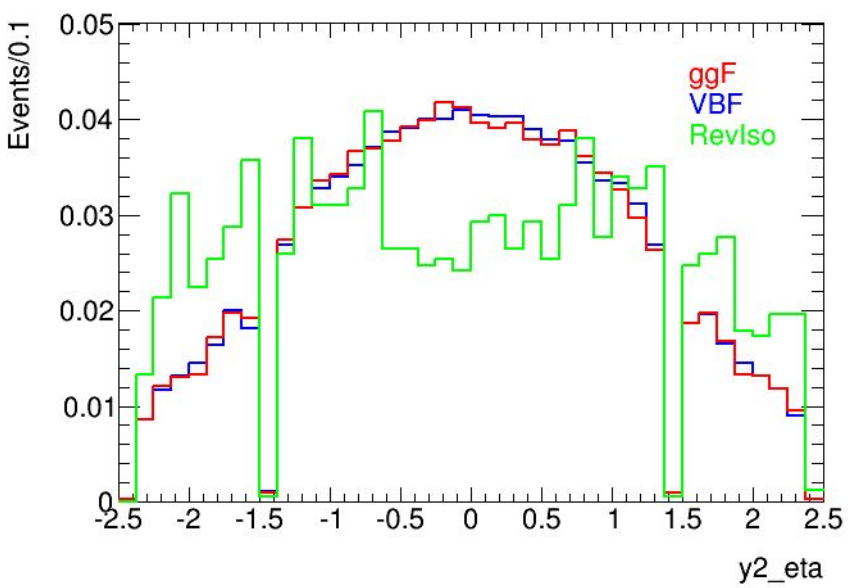

Events/5.0
ggF
VBF
RevIso
y1_pt

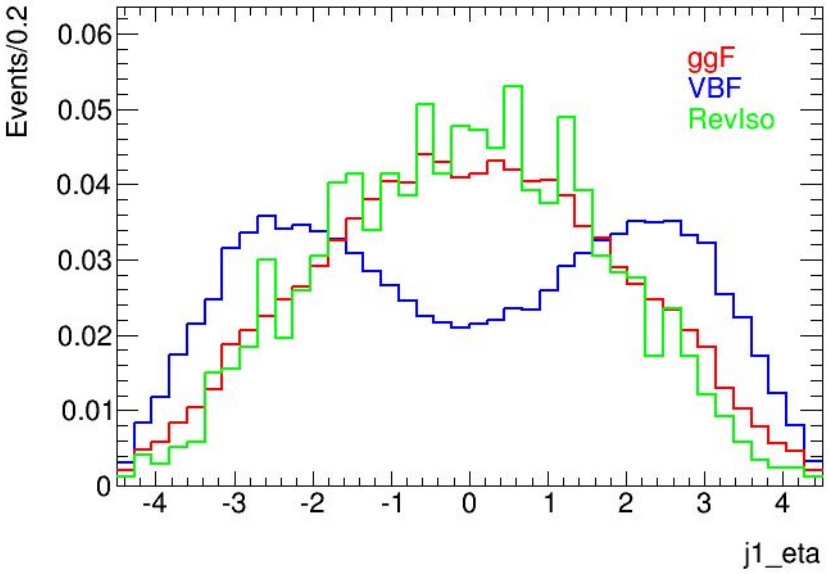

Events/5.0
ggF
VBF
RevIso
y2_pt

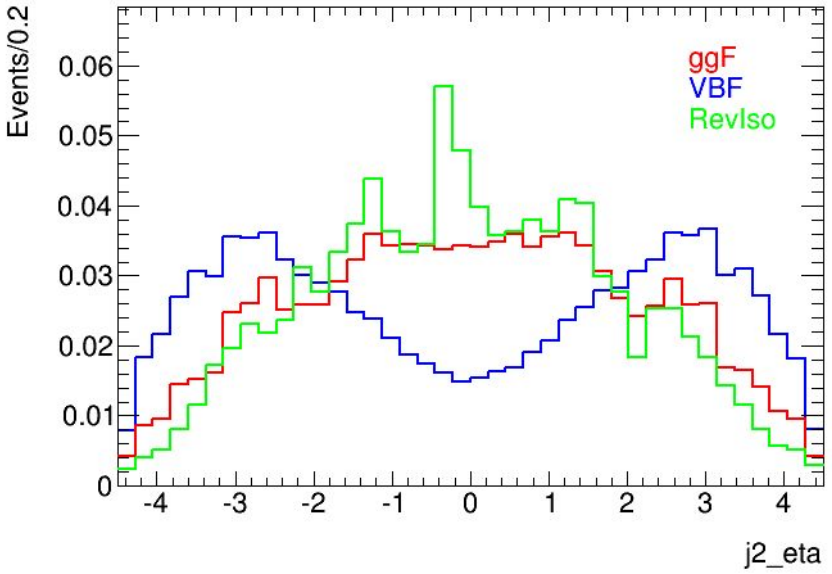

Events/7.5
ggF
VBF
RevIso
j1_pt

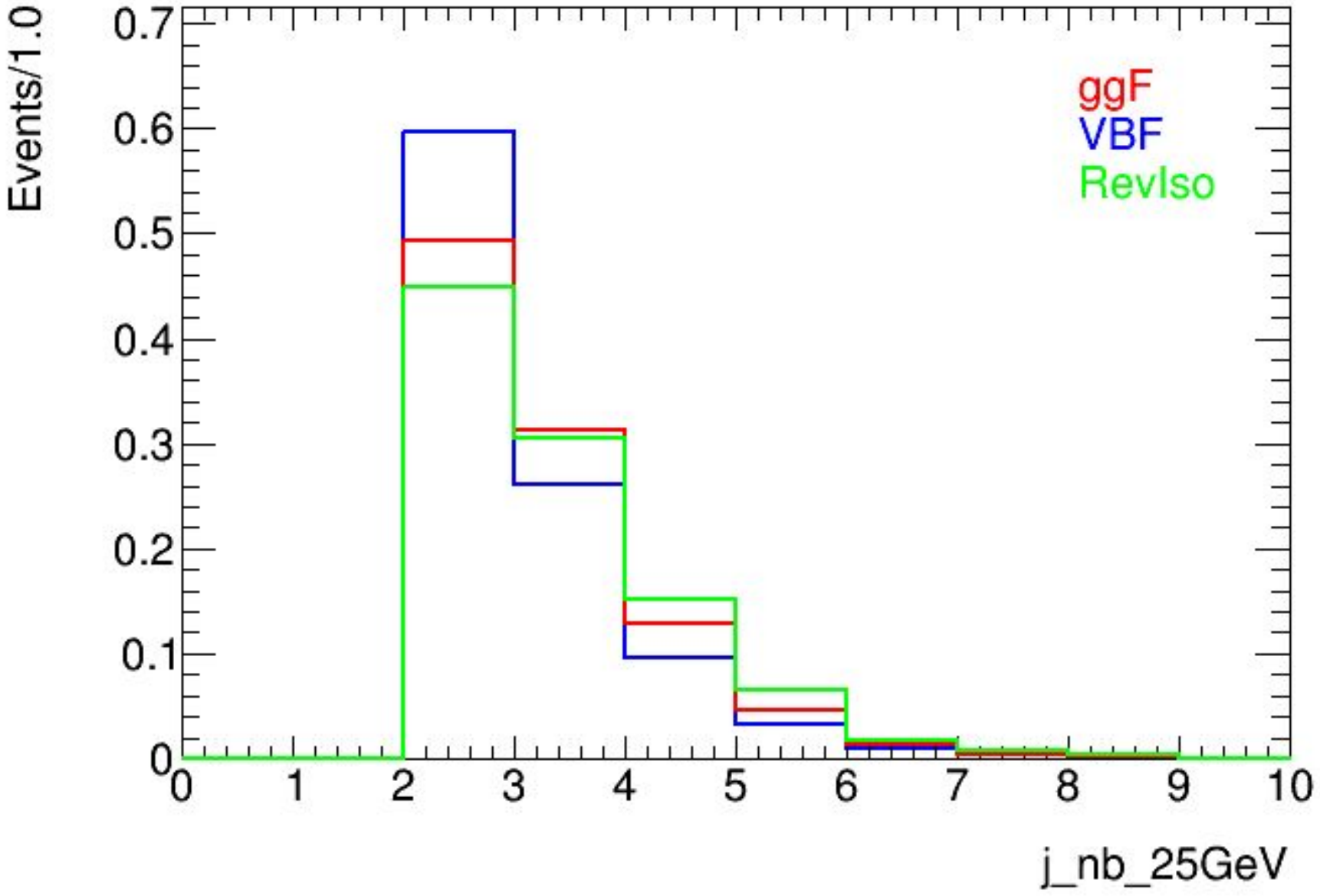

Events/7.5
ggF
VBF
RevIso
j2_pt

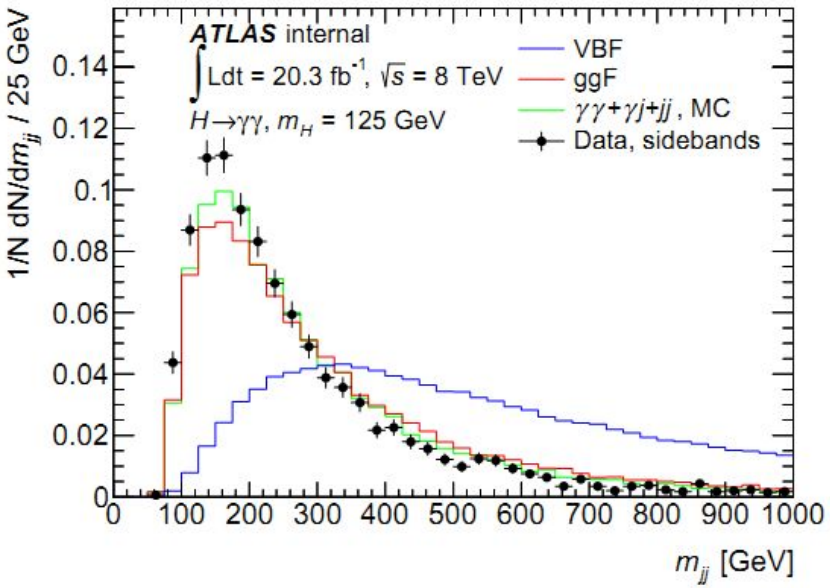

Events/0.1
ggF
VBF
RevIso
y1_eta

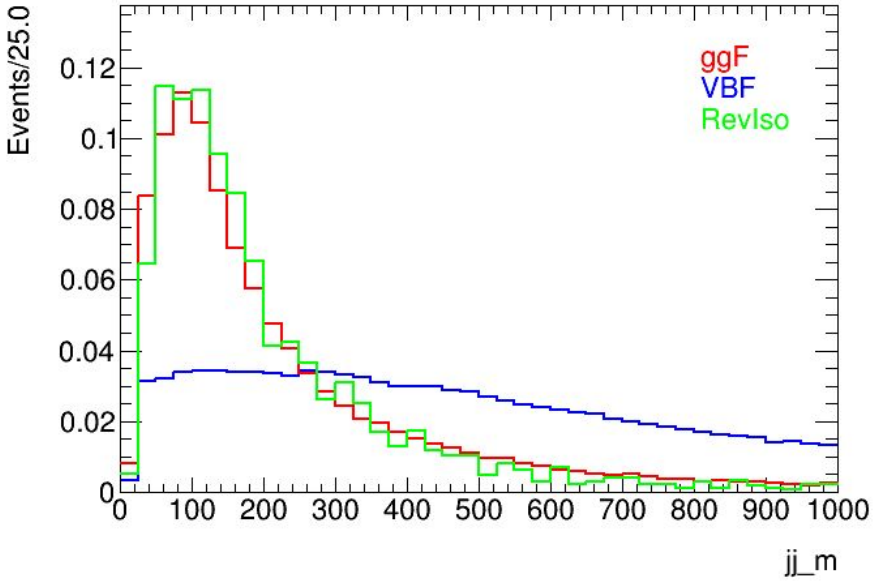

Events/0.1
ggF
VBF
RevIso
y2_eta


Events/0.2
ggF
VBF
RevIso
j1_eta

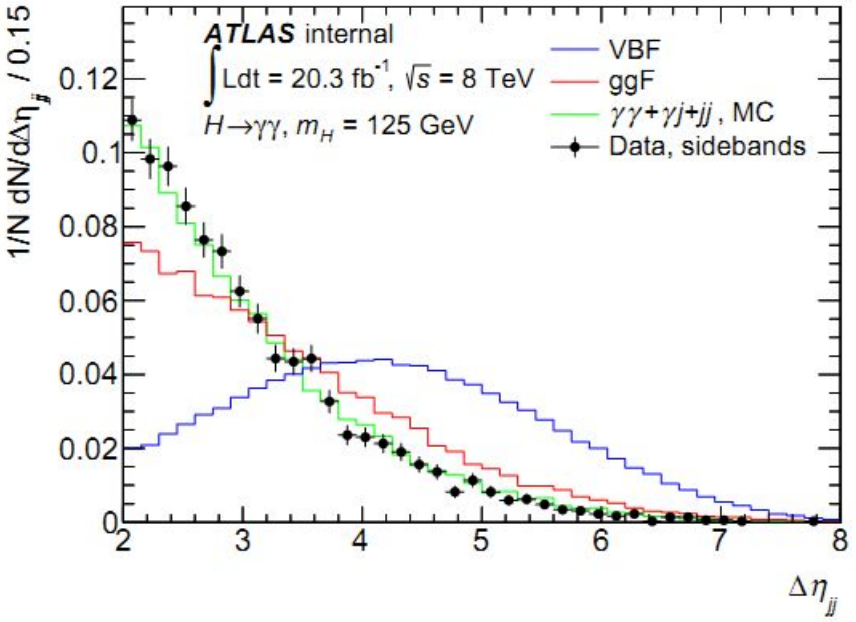

Events/0.2
ggF
VBF
RevIso
j2_eta

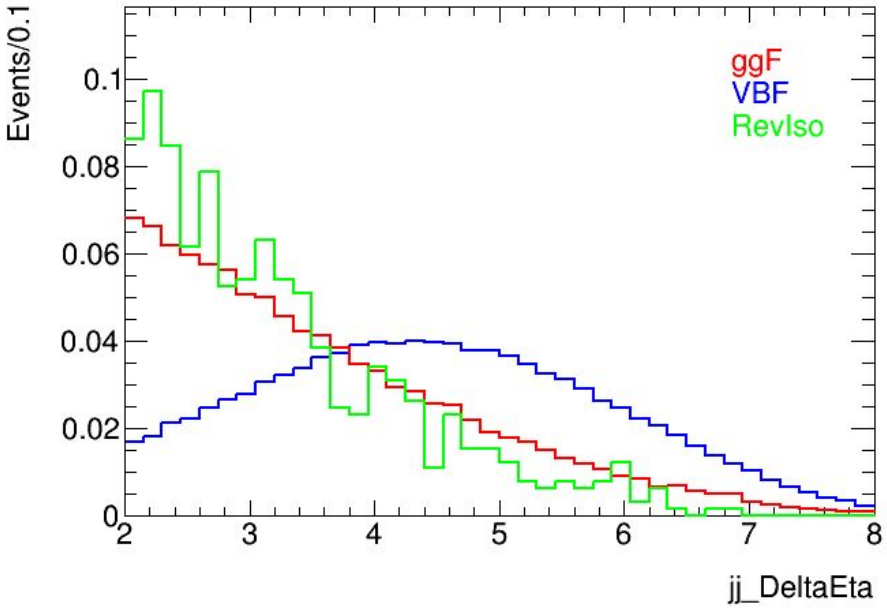

Events/1.0
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
1
2
3
4
5
6
7
8
9
10
ggF
VBF
RevIso
j_nb_25GeV



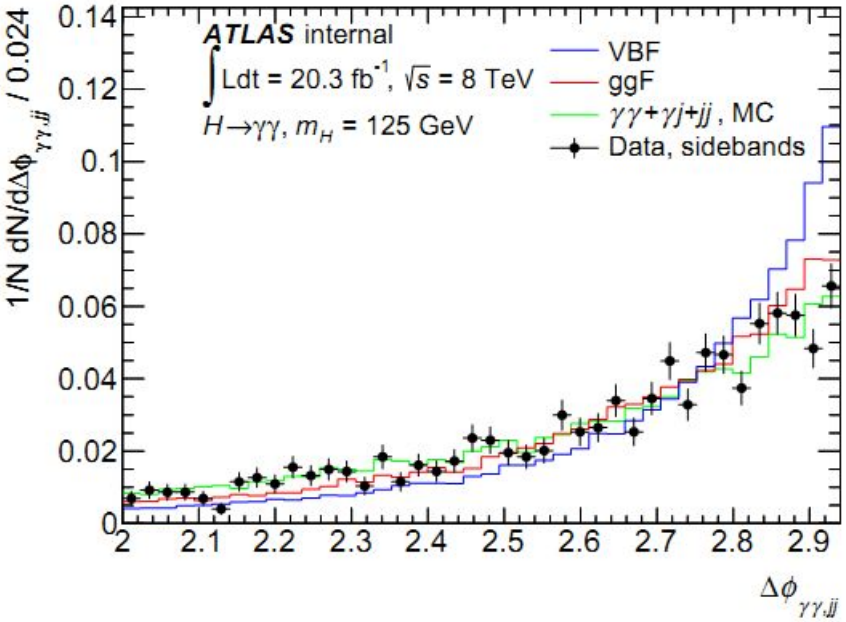


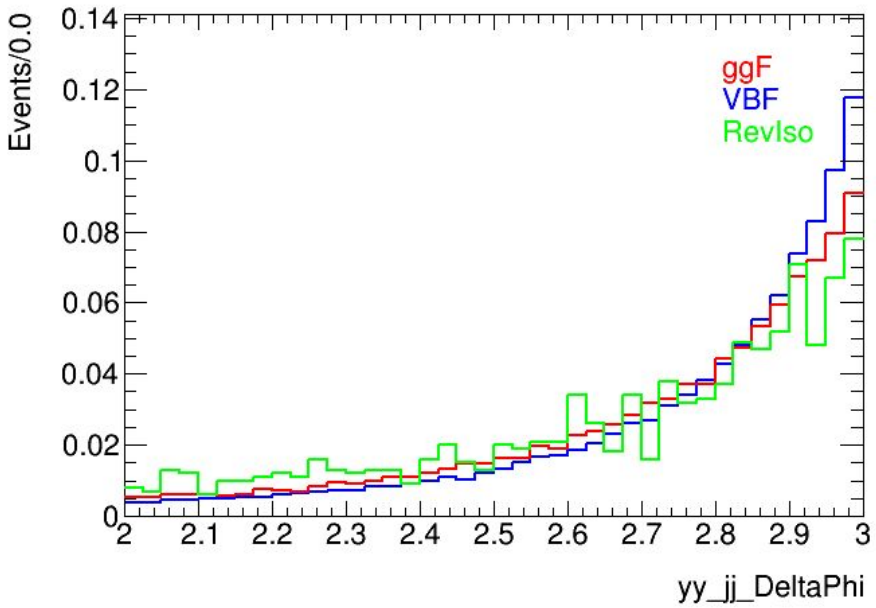
- VBF is more flat and peak position is lower
- distribution of run1 does not start from 0,right does.(resonable ?)
- ggF and RevIso jj_m value is also lower than run1
- run1 selection:jj_m>400(520)GeV



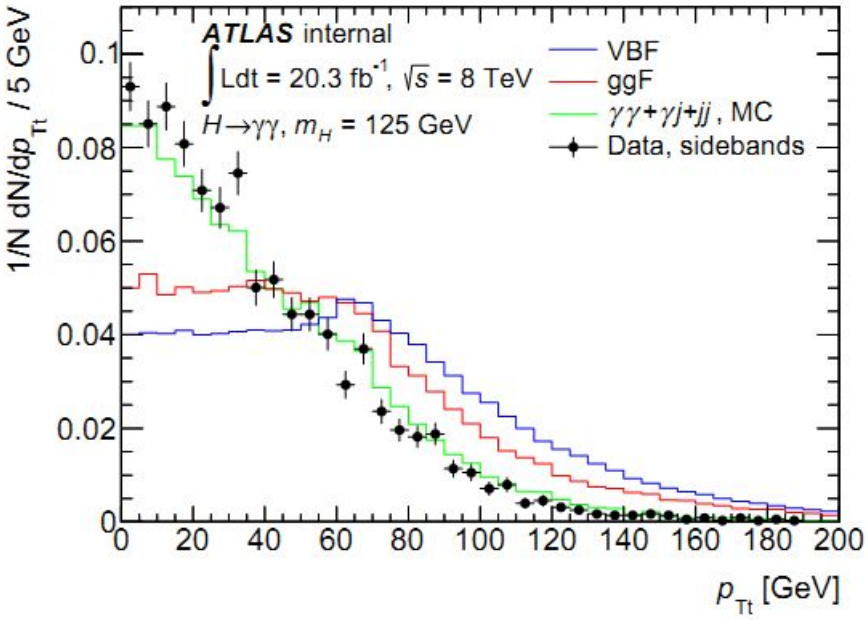


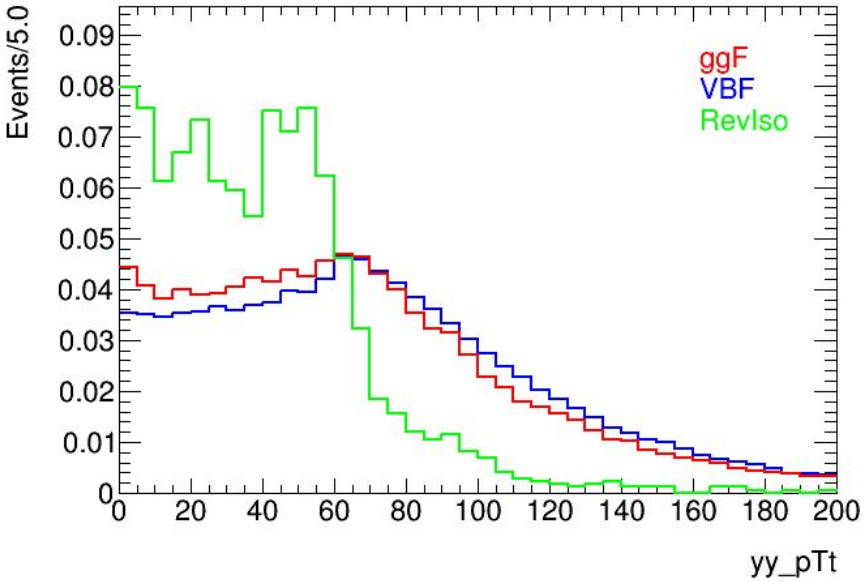
- look similar
- RevIso fluctuation is a little huge
- we don't care about jj_DeltaEta<2 region?
- run1 selection: jj_DeltaEta>2.8



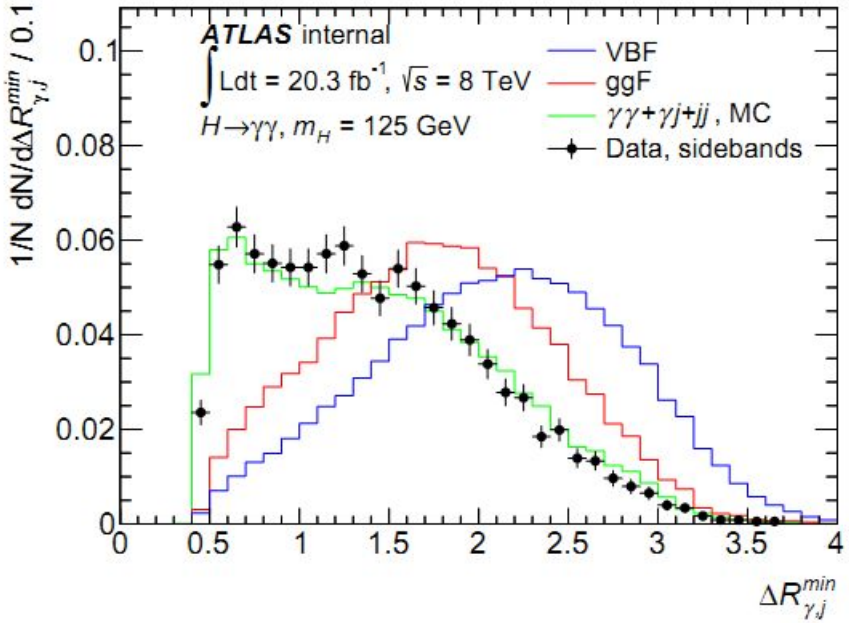


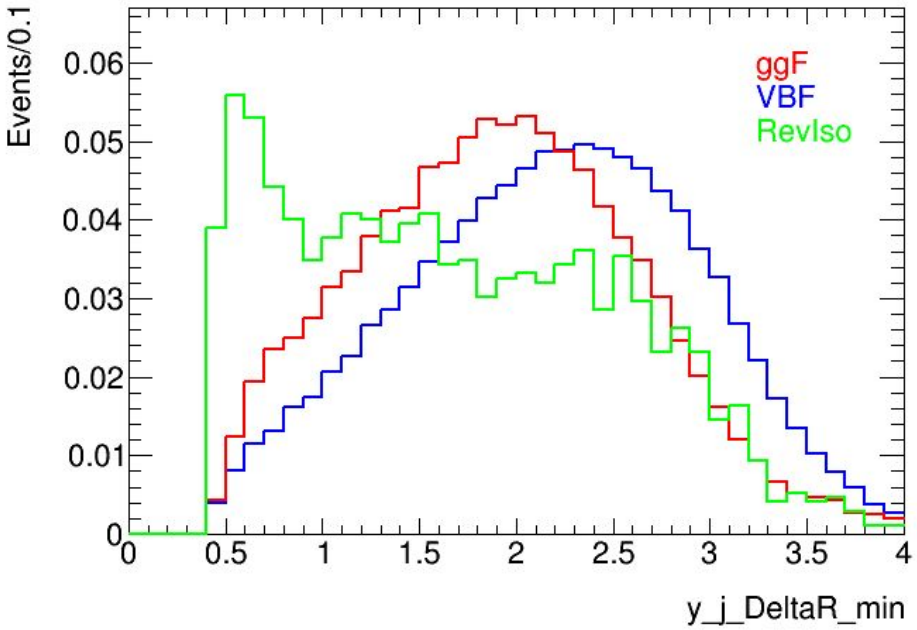
- look similar
- run1 tight:yy_jj_DeltaPhi>2.6



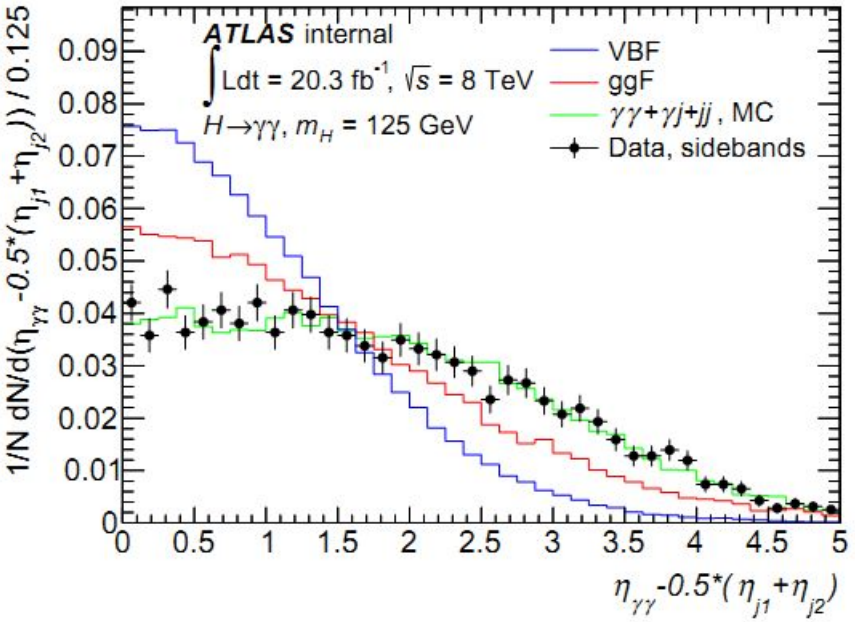


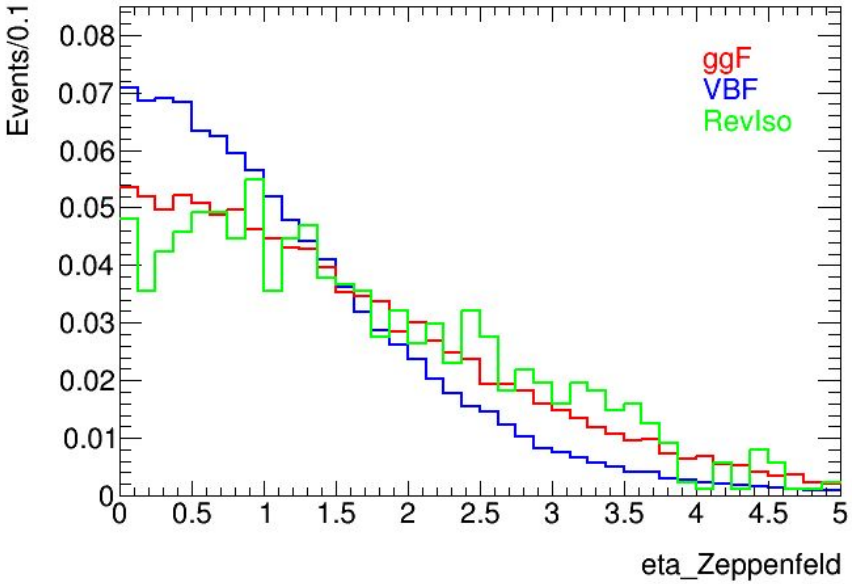
- RevIso fluctuation
- only for MVA method





- min DeltaR($\gamma_{lead/sub}$,$jet_{lead/sub}$)
- tight: y_j_DeltaR_min>2





- eta_Zeppenfeld=yy_eta-0.5(j1_eta+j2_eta)
- tight:eta_Zeppenfeld<2.4

cut efficiency

	data	RevIso	VBF	ggF
total:	72200		506394	500000
pass:	668(0.00925208	9040	324325(0.64046	323336(0.646672
Njet:	179(0.00247922	2610(0.288717	221417(0.437243	79034(0.158068
pT30:	114(0.00157895	1737(0.192146	184609(0.364556	47582(0.095164
pT50_j1:	88(0.00121884	1359(0.150332	165777(0.327368	35637(0.071274

method

- cut-based
- MVA,BDT(boosted decision tree)

$$\sigma_{VBF} = \sqrt{2 \times (N_{VBF} + N_{ggF} + N_{Background}) \times ln\left(1 + \frac{N_{VBF}}{N_{ggF} + N_{Background}}\right) - N_{VBF}}$$

- scan the selection and get the highest significance

try to do cut-based analysis

- estimate the ggF,VBF numbers after cut

  cut efficiency,lumi,XS,br
- use the data number as nBkg(few events)

  fit the sideband in run1
- scan the selection value to see the significance

jj_DeltaEta	2	2.5	3	3.5	4	4.5
sigma	1.918	1.872	1.814	1.714	1.580	1.433

summary

- due to low statistic,cut-based seems failing to improve significance now.
- MVA
- any suggestion for new variables?
- optimize and automize the code