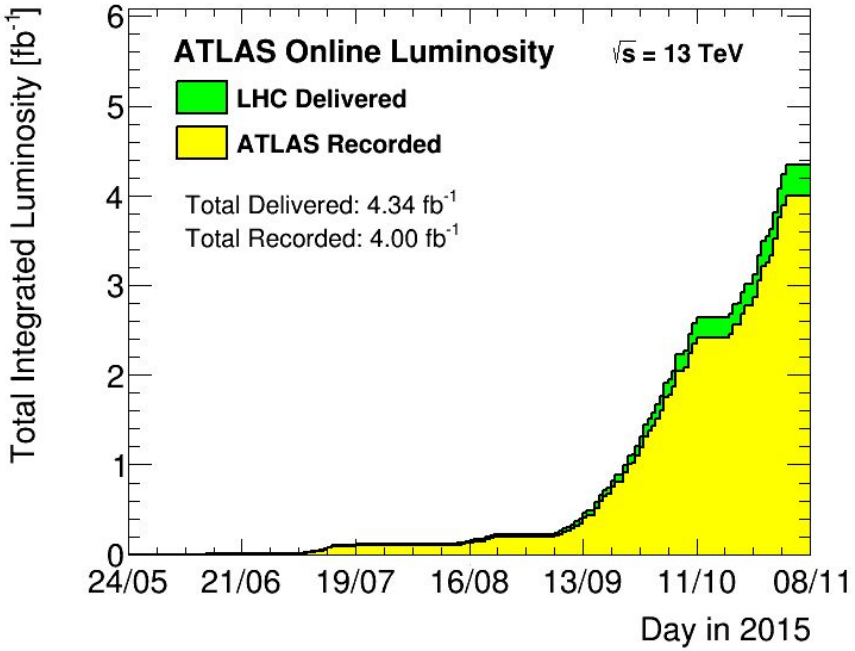
13TeV VBF H->$\gamma\gamma$ Analysis

Yu Zhang

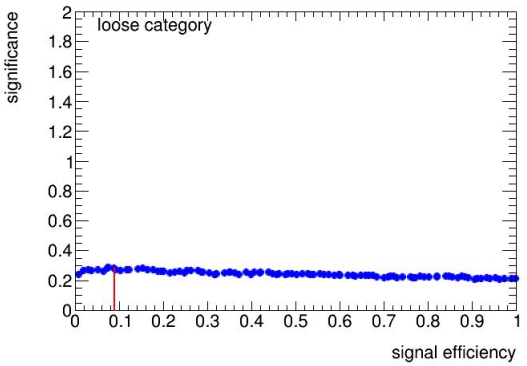
11.09

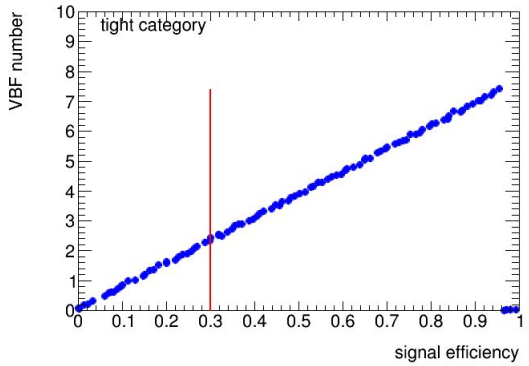
introduction

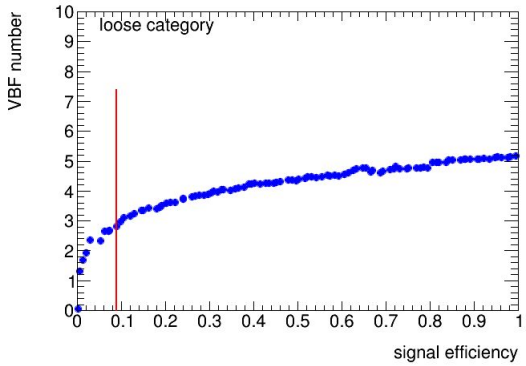
- cut-based
  - flat
- MVA
  - new variables
  - comments from HGam Coupling meeting
- data
  - all:4fb^{-1}
  - for physics:~3.5fb^{-1}
  - GRL:2.7fb^{-1}
  - MxAOD:2fb^{-1}
- document



Cut-based







- a flat line in work line of cut-based loose category
- check the VBF number of events in each category after selection
- the tight category is as expected:proportional
- the line loose category is bad
- maybe there is something wrong with the weight file
- under investigation...

new variables

- related systems

	γ1	γ2	γ1γ2
j1	γ1 j1	γ2 j1	γ1 γ2 j1
j2	γ1 j2	γ2 j2	γ1 γ2 j2
j1 j2	γ1 j1 j2	γ2 j1 j2	γ1 γ2 j1 j2

- related variables : system pT,sum of pT,sum of P , mass,Eta,ΔpT,ΔΦ,ΔR,pT/P of system,E,E/M
- addtional variables:$pT_{\gamma 1}/m_{\gamma\gamma}$,$pT_{\gamma 2}/m_{\gamma\gamma}$,$pT_{\gamma\gamma}/m_{\gamma\gamma}$,$\eta_{j1}*\eta_{j2}$
- variable selection
  - loop all variables ,rank them by separation power
  - keep the most powerful one if varialbes are correlated
  - remove variables that correlates to mγγ
  - find the best combination
- more than 100 variables ,many are highly correlated

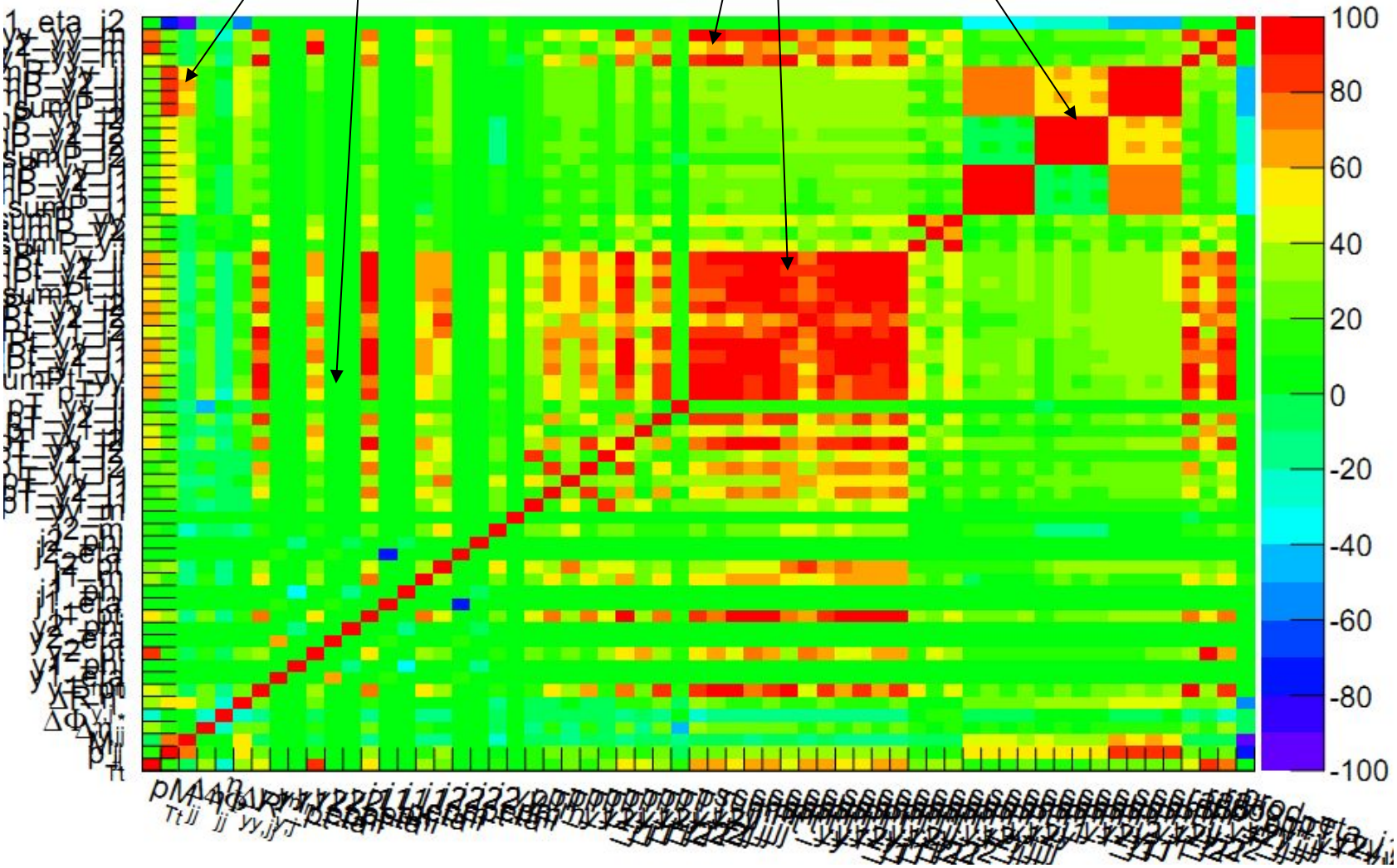
correlation

60 vars

1

2

3



correlation

- 1:
  - pTt is correlated to photon pT
  - mjj &Δηjj are correlated to sumP of some systems
- 2:
  - pT_y1/myy,pT_y2/myy,pT_yy/myy correlated to sumPt
  - pT of systems correlated to sumPt
- 3:
  - sumP of different systems are correlated
- all these variables are low correlated to myy
- the one most correlated to myy is y2_pt: 0.041

- The **separation** $\langle S^2 \rangle$ of a classifier y, defined by the integral [7]

$$\langle S^2 \rangle = \frac{1}{2} \int \frac{(\hat{y}_S(y) - \hat{y}_B(y))^2}{\hat{y}_S(y) + \hat{y}_B(y)} dy \,,$$

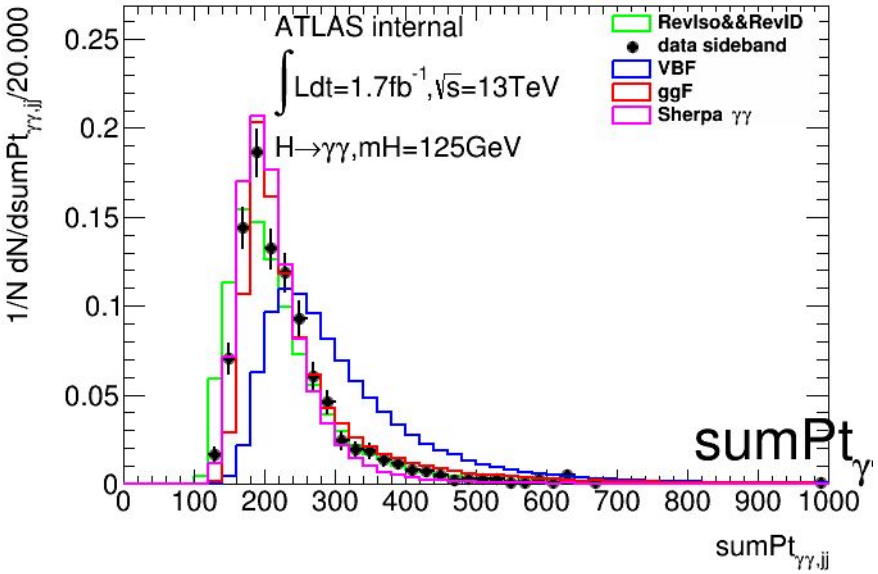
- Rank : Variable : Separation
- ----------
- 1 : yy_m : 6.820e-01
- 2 : sumPt_yy_jj : 2.627e-01
- 3 : sumPt_yy_j1 : 2.565e-01
- 4 : sumPt_yy_j2 : 2.565e-01
- 5 : mjj : 2.540e-01
- 6 : pT_yy/myy : 2.496e-01
- 7 : sumPt_y1_jj : 2.423e-01
- 8 : sumPt_y2_jj : 2.338e-01
- 9 : sumPt_y1_j1 : 2.334e-01
- 10 : pTt_yy : 2.285e-01
- 11 : sumPt_y1_j2 : 2.246e-01
- 12 : pT_yy_j2 : 2.238e-01
- 13 : sumPt_y2_j1 : 2.229e-01
- 14 : sumPt_yy : 2.190e-01
- 15 : sumPt_jj : 2.105e-01
- 16 : ΔΦyy,jj : 2.103e-01
- 17 : pT_j1 : 1.964e-01
- 18 : pT_y1/myy : 1.848e-01
- 19 : sumPt_y2_j2 : 1.820e-01
- 20 : min(ΔR_y_j) : 1.807e-01

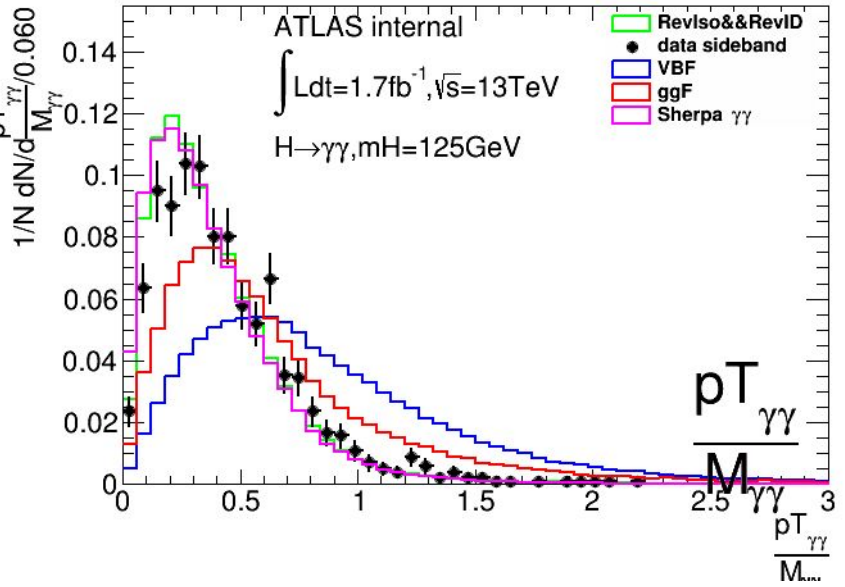
21 : pT_y1 : 1.766e-01
22 : sumP_y1_jj : 1.747e-01
23 : sumP_yy_jj : 1.739e-01
24 : sumP_jj : 1.727e-01
25 : sumP_y2_jj : 1.722e-01
26 : pT_jj : 1.509e-01
27 : pT_j2 : 1.365e-01
28 : η Zepp : 1.355e-01
29 : sumP_y1_j1 : 1.298e-01
30 : sumP_yy_j1 : 1.298e-01
31 : sumP_y2_j1 : 1.245e-01
32 : Δηjj : 1.226e-01
33 : pT_y2/myy : 1.210e-01
34 : ηj1*ηj2 : 1.199e-01
35 : sumP_j1 : 1.195e-01

Variable	Separation	Importance
m_{jj}	25.6%	17.0%
$\Delta\eta_{jj}$	22.8%	17.2%
$\min(\Delta R_{j\gamma})$	18.6%	14.5%
$\eta^{Zepp}_{\gamma\gamma}$	12.4%	14.6%
p_{Tt}	9.8%	9.7%
$\Delta\phi_{\gamma\gamma,jj}$	9.1%	14.6%
η_{jet2}	7.8%	5.4%
η_{jet1}	4.7%	7.0%

from run1

discriminating variables





8 vars ,BDT [0.77,0.95,1]

few events

	MVA tight	mva loose
VBF	1.43	3.71
ggF	0.33	1.54
background	0.52	8.92
VBF purity	0.77	0.59
significance	1.02	0.65
combined significance	1.21	

feedback

- $\Delta\Phi\gamma\gamma$,jj is sensitive to systematics in high $\Delta\Phi$ region ,modify the denfinition:if $\Delta\Phi\gamma\gamma$,jj>2.92,then $\Delta\Phi\gamma\gamma$,jj=2.92
- train with ggF sample,$\gamma\gamma$,γj and jj as background
- check the systematic of different combination of variables

other variables

- ---
- Rank : Variable : Separation
- ---
- 1 : yy_m : 6.790e-01
- 2 : DeltaR_yy : 3.000e-01
- 3 : M_{jj} : 2.558e-01
- 4 : yy_pt : 2.475e-01
- 5 : p_{Tt} : 2.354e-01
- 6 : mass_y2_jj : 2.218e-01
- 7 : mass_y1_jj : 2.139e-01
- 8 : DeltaPhi_yy : 1.966e-01
- 9 : mass_yy_jj : 1.958e-01
- 10 : #Delta#Phi_{yy,jj} : 1.952e-01
- 11 : DeltaPt_yy_j2 : 1.899e-01
- 12 : #DeltaR^{min}_{y,j} : 1.853e-01
- 13 : mass_y1_j1 : 1.608e-01
- 14 : mass_yy_j1 : 1.608e-01
- 15 : mass_y2_j1 : 1.444e-01
- 16 : eta_yy_jj : 1.394e-01
- 17 : centrality_yy_jj : 1.332e-01
- 18 : #Delta#eta_{jj} : 1.299e-01
- 19 : #eta_{*} : 1.261e-01
- 20 : DeltaR_jj : 1.070e-01
- 21 : DeltaPt_yy_jj : 1.010e-01
- 22 : DeltaPt_yy_j1 : 9.851e-02

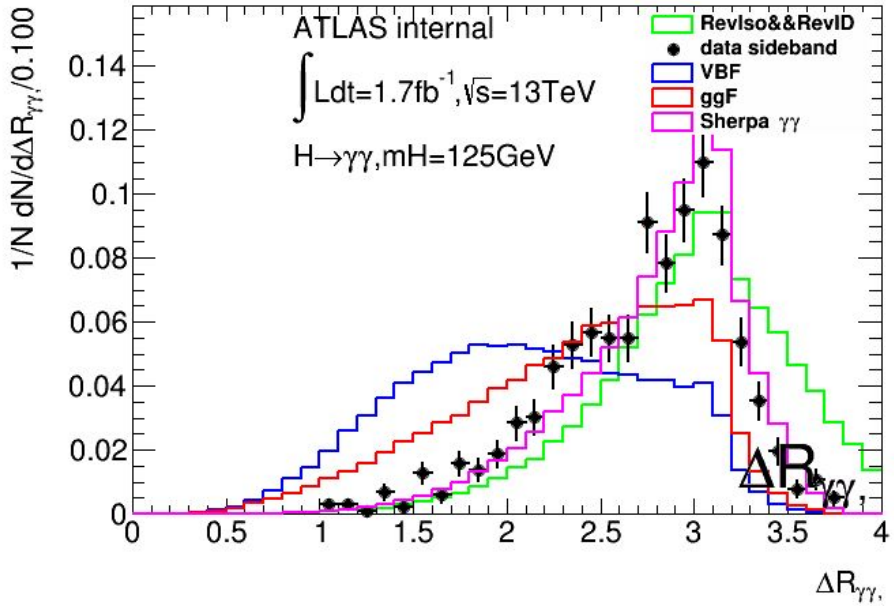
- 22 : DeltaPt_yy_j1 : 9.851e-02

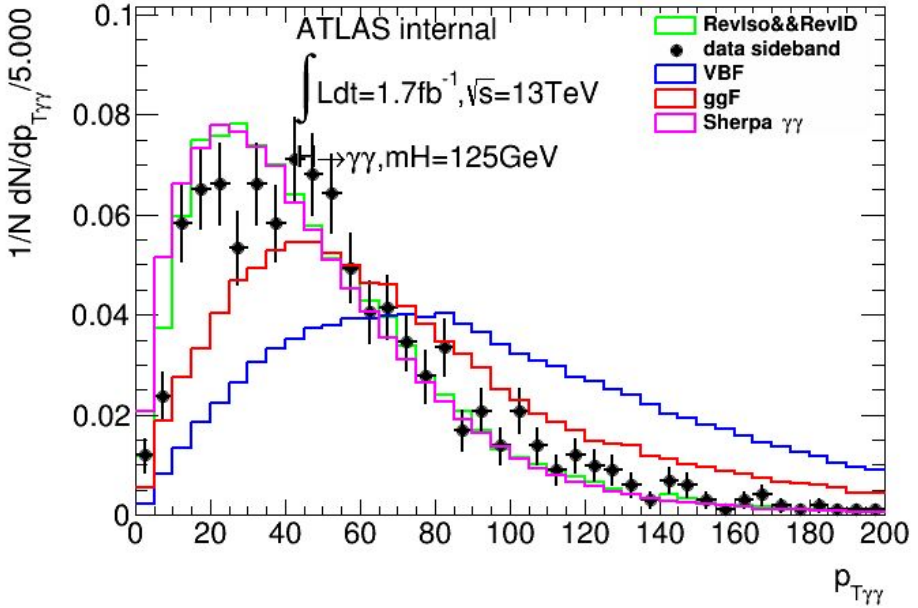
- 23 : DeltaPt_y2_j1 : 9.652e-02
- 24 : DeltaPt_min_y_j : 9.650e-02
- 25 : DeltaPt_jj : 9.514e-02

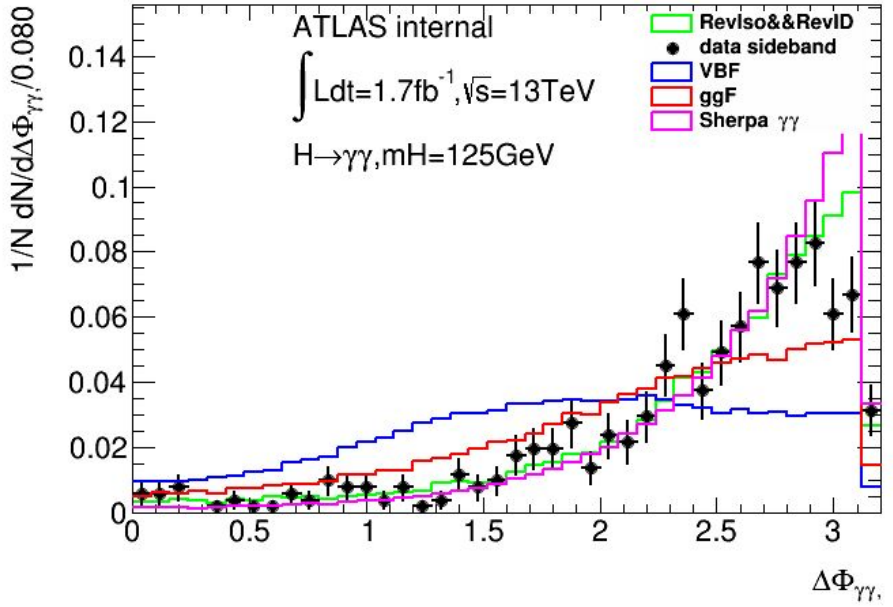
- The **separation** $\langle S^2 \rangle$ of a classifier y, defined by the integral [7]

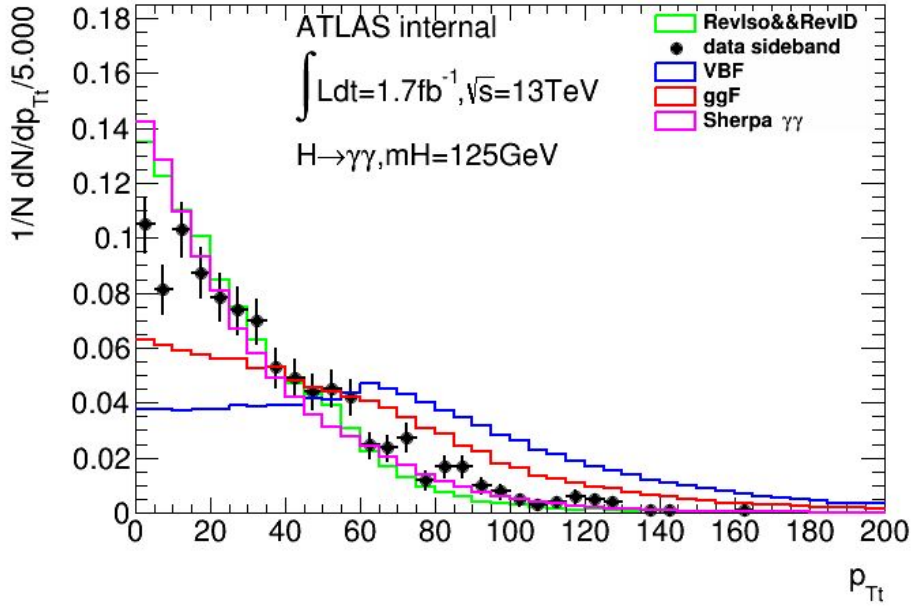
$$\langle S^2 \rangle = \frac{1}{2} \int \frac{\left(\hat{y}_S(y) - \hat{y}_B(y)\right)^2}{\hat{y}_S(y) + \hat{y}_B(y)} dy \,,$$

some variables









document

- a complete indepent VBF note or part of coupling note?
- focus on it these days
- hope to give a first draft tomorrow
- submit to Coupling group before Friday
- not the final result

derivation

- xAOD->DAOD
- skimming and thinning
- skimming :a loose selection ,keep yy,1e1y,ee,1mu1y
- thinning :remove some track information
- help to ask for derivation of pythia yy samples

summary

- cut-based
  - check the loose category
- MVA
  - check the systematic of different combination
  - find the best variable combination
- document
  - the most important work for now!!!