Weekly update

FANGYI GUO

Previous review

Higgs to di-photon physics analysis in CEPC:

- CEPC_v4, 240GeV, $\mathcal{L} = 5.6fb^{-1}$
- Whizard 1.95 + Fast simulation MC
- 2 fermion background is the dominant
- FSClasser event reconstruction and cat-based event selection

- Results: $\delta(Br \times \sigma)$ = 6.84% in combined channel, 9.84% in $q\bar{q}\gamma\gamma$ channel

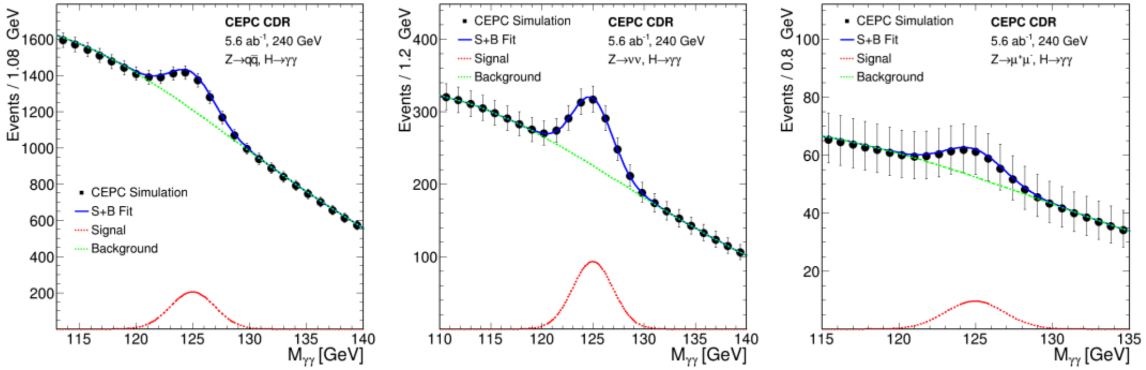


Figure1. di-photon invariant mass distributions in 3 sub-channel

Results

The combination of three sub-channel provides a final result of $\sigma(ZH) \times BR(H \to \gamma\gamma)$ measurement precision

Sub-channel	$q\bar{q}\gamma\gamma$	$l\bar{l}\gamma\gamma$	$\nu\bar{\nu}\gamma\gamma$	combined
precision	9.84%	23.7%	10.5%	6.84%

Event selection

CDR in last year:

$E_{\gamma 1} > 35GeV$

$35GeV < E_{\gamma 2} < 96GeV$

$cos\theta_{\gamma\gamma} >$ -0.95, $cos\theta_{jj}$>-0.95

$pT_{\gamma 1} > 20GeV, pT_{\gamma 1} > 30GeV$

$110GeV < m_{\gamma\gamma} < 140GeV$

$125GeV < E_{\gamma\gamma} < 145GeV$

$min|cos\theta_{\gamma j}|$<0.9

Present:

$E_{\gamma 1} > 25GeV$

$35GeV < E_{\gamma 2} < 96GeV$

$cos\theta_{\gamma\gamma} >$ -0.95, $cos\theta_{jj}$>-0.95

$pT_{\gamma 1} > 20GeV, pT_{\gamma 1} > 30GeV$

$110GeV < m_{\gamma\gamma} < 140GeV$

$E_{\gamma\gamma} < 120GeV$

$\min|cos\theta_{\gamma j}|$<0.9

In order to avoid the bump in background $m_{\gamma\gamma}$ distribution after MVA

MVA method

Considered variables:

- P, E, pT, $cos\theta$ of two photon and 2 jets
- P, E, pT, $cos\theta$, recoil mass, pTt, Pt* of di-photon system
- P, E, mass, recoil mass, $cos\theta$ of jj system
- ΔP, ΔE, $\Delta\phi$ between two photon, $\gamma\gamma$-qq
- Cosine angle between 2 photon, 2 jets, 1 photon and 1 jet, $\gamma\gamma$ and jj system.
- Minimum ΔR between any photon and jet

Totally 42 variables

Pt*: Di-photon P projected perpendicular to the di-photon thrust axis.(similar as pTt but replace pT with P)

Pt* = $|(\overrightarrow{P_1} + \overrightarrow{P_2}) \times \frac{\overrightarrow{P_1} - \overrightarrow{P_2}}{|\overrightarrow{P_1} - \overrightarrow{P_2}|}|$

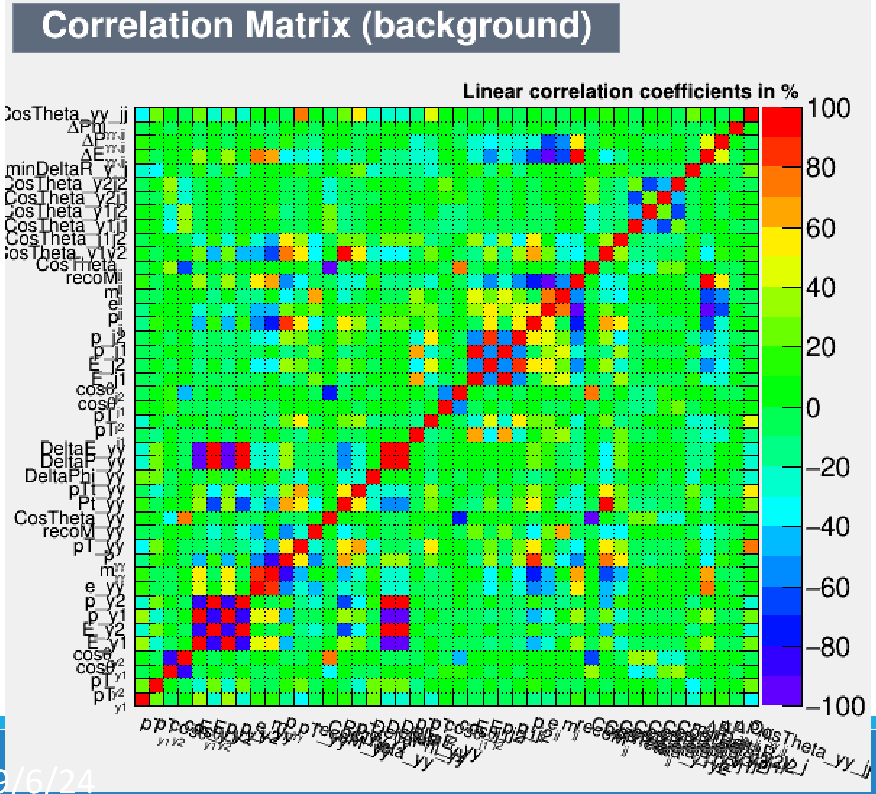
Separation power:

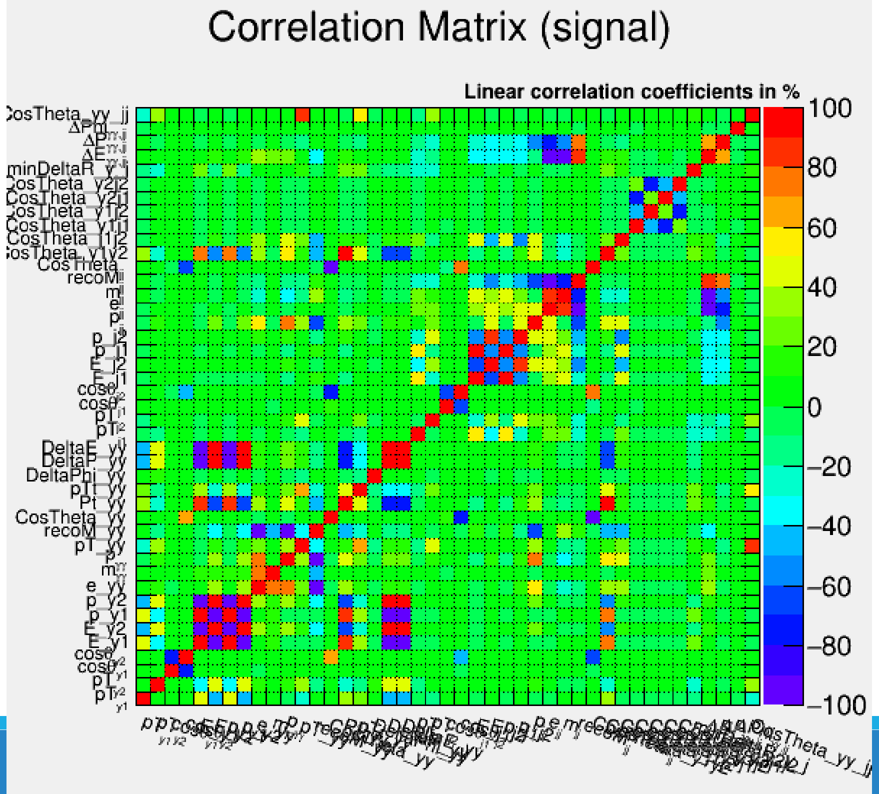
$$< S^2 >= \frac{1}{2} \int \frac{(\hat{y}_s(y) - \hat{y}_b(y))^2}{\hat{y}_s(y) + \hat{y}_b(y)} dy.$$

MVA method

Variable correlation:

- Remove high $m_{\gamma\gamma}$-relative variables $|Corr_{v-m_{\gamma\gamma}}|$<30
- Remove high co-related variables $|Corr_{v_1-v_2}|$<50





MVA method

Remaining variable: 11 variables

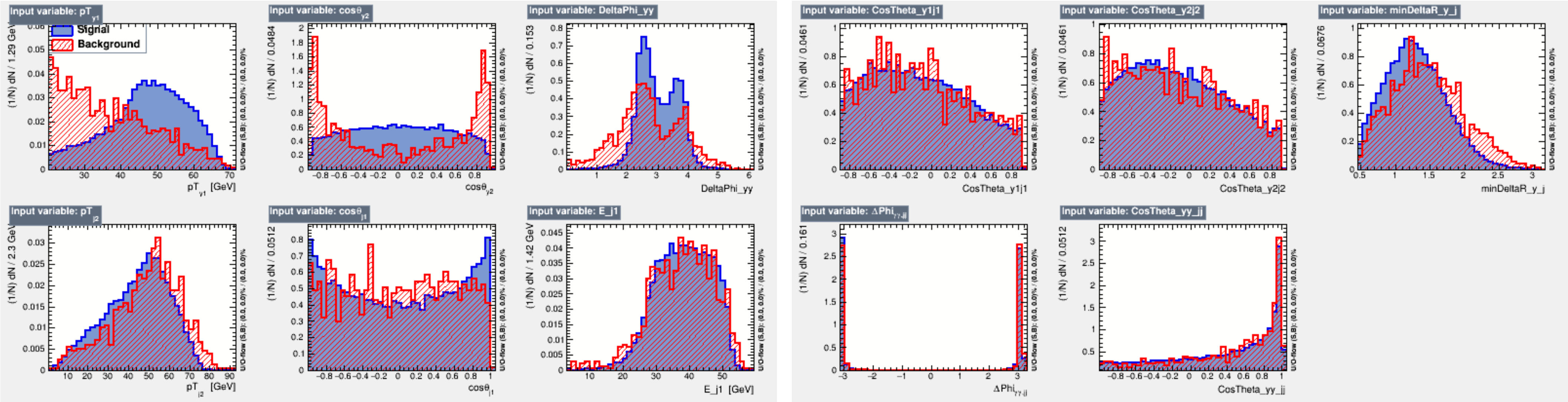
variable	Separation power
pT_y1	0.391071
cosTheta_y2:	0.388952
DeltaPhi_yy:	0.304471
minDeltaR:	0.089791
cosTheta_j1:	0.083385

pT_j2	0.077177
e_j1:	0.030657
costheta_y2j2:	0.026474
costheta_y1j1:	0.022066
DeltaPhi_yy_jj:	0.012697
DeltaCosTheta_yy_jj:	0.011999

MVA method: Boost decision tree

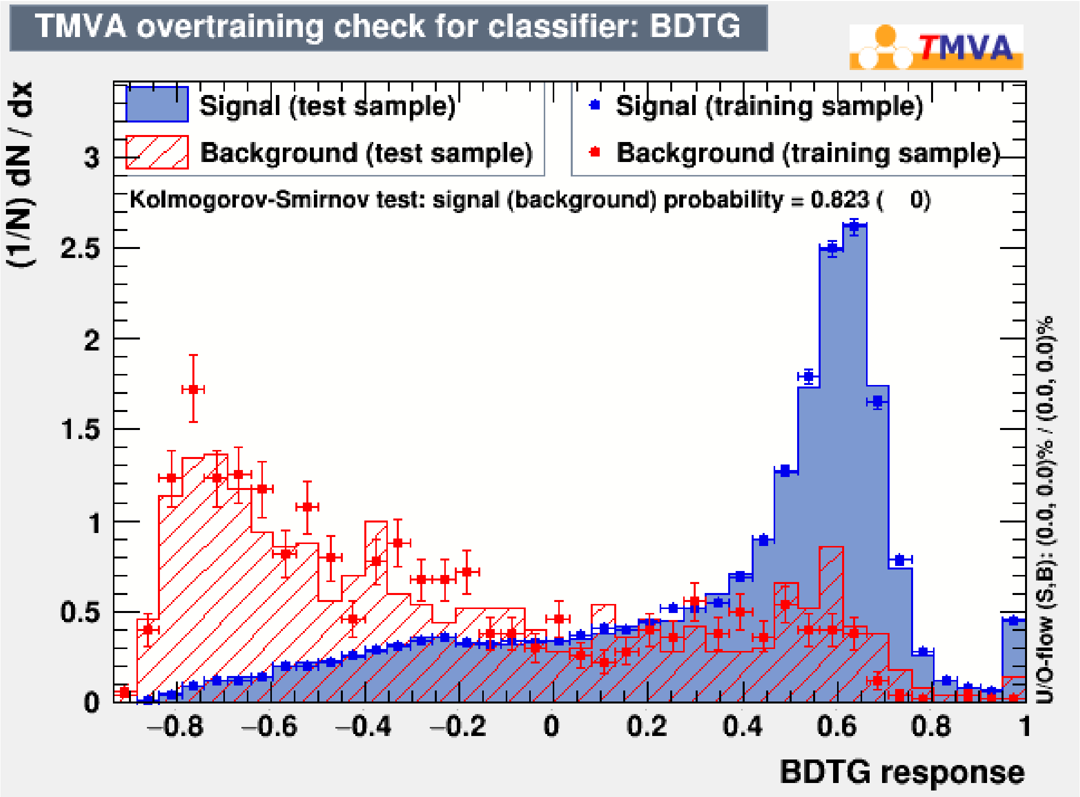
```
factory->BookMethod( TMVA::Types::kBDT,"BDTG",
"!H:!V:NTrees=900:nEventsMin=50:BoostType=Grad:Shrinkage=0.06:UseBaggedGrad:GradBaggingFraction=0.6:nCuts=20:MaxDepth=3" );
```

MVA method



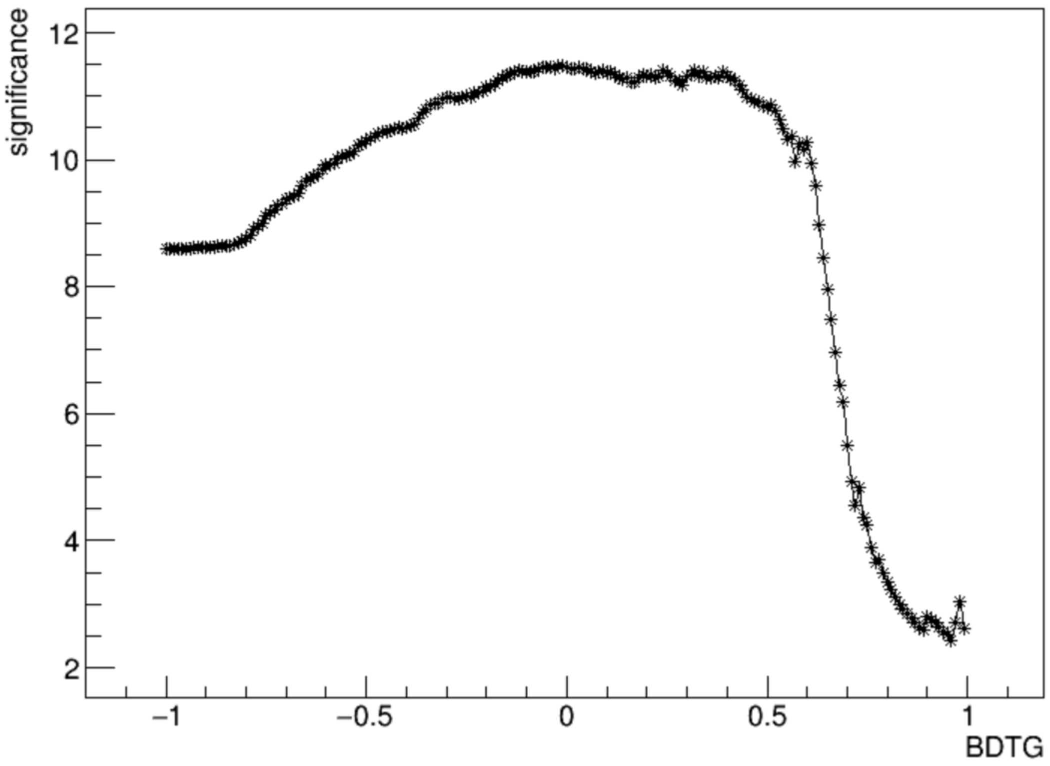
Input variable: pT_γ1
Signal
Background
(1/N) dN / 1.29 GeV
pT_γ1 [GeV]
Input variable: cosθ_γ2
(1/N) dN / 0.0484
cosθ_γ2
Input variable: DeltaPhi_yy
(1/N) dN / 0.153
DeltaPhi_yy
Input variable: CosTheta_y1j1
(1/N) dN / 0.0461
CosTheta_y1j1
Input variable: CosTheta_y2j2
(1/N) dN / 0.0461
CosTheta_y2j2
Input variable: minDeltaR_y_j
(1/N) dN / 0.0676
minDeltaR_y_j
Input variable: pT_j2
(1/N) dN / 2.3 GeV
pT_j2 [GeV]
Input variable: cosθ_j1
(1/N) dN / 0.0512
cosθ_j1
Input variable: E_j1
(1/N) dN / 1.42 GeV
E_j1 [GeV]
Input variable: ΔPhi_γγ-jj
(1/N) dN / 0.161
ΔPhi_γγ-jj
Input variable: CosTheta_yy_jj
(1/N) dN / 0.0512
CosTheta_yy_jj

MVA method


TMVA overtraining check for classifier: BDTG
TMVA
Signal (test sample)
Background (test sample)
Signal (training sample)
Background (training sample)
Kolmogorov-Smirnov test: signal (background) probability = 0.823 (0)
(1/N) dN / dx
3
2.5
2
1.5
1
0.5
0
−0.8
−0.6
−0.4
−0.2
0
0.2
0.4
0.6
0.8
1
BDTG response
U/O-flow (S,B): (0.0, 0.0)% / (0.0, 0.0)%

MVA categorization

Categorization: maximum $\sigma = N_{sig}/\sqrt{N_{sig} + N_{bkg}}$ count N in $m_{\gamma\gamma} \in [120, 130]GeV$



Kcut: BDTG=-0.01

Tight category: BDTG>-0.01
- Nsig: 738.90
- Nbkg: 3408.73
- significance: 11.47

Loose category: BDTG<-0.01
- Nsig: 178.71
- Nbkg: 7044.71
- significance: 2.10

Combined significance: 11.93

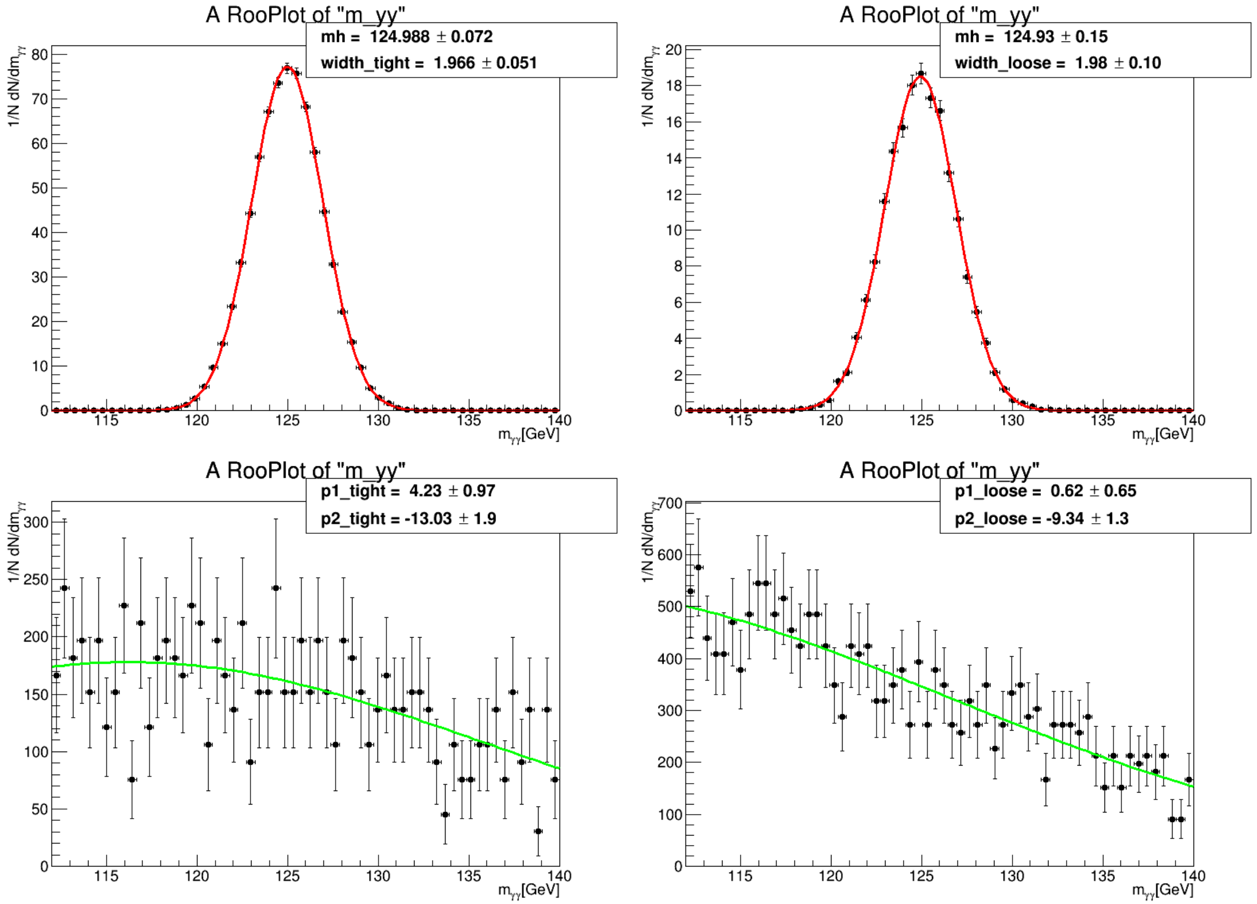
Fit sample

Signal: gauss function

Background: 2nd polynomial exponential PDF

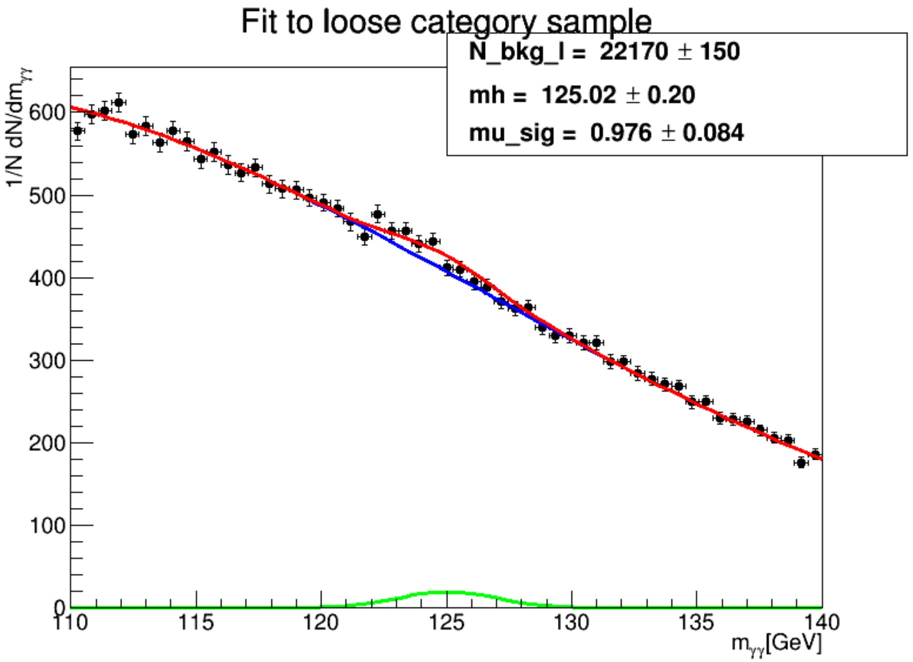
Combine signal and background sample into $PDF_{sum} = \mu \times N_{sig}^{SM} \times PDF_{sig} + N_{bkg} \times PDF_{bkg}$ in 2 categories

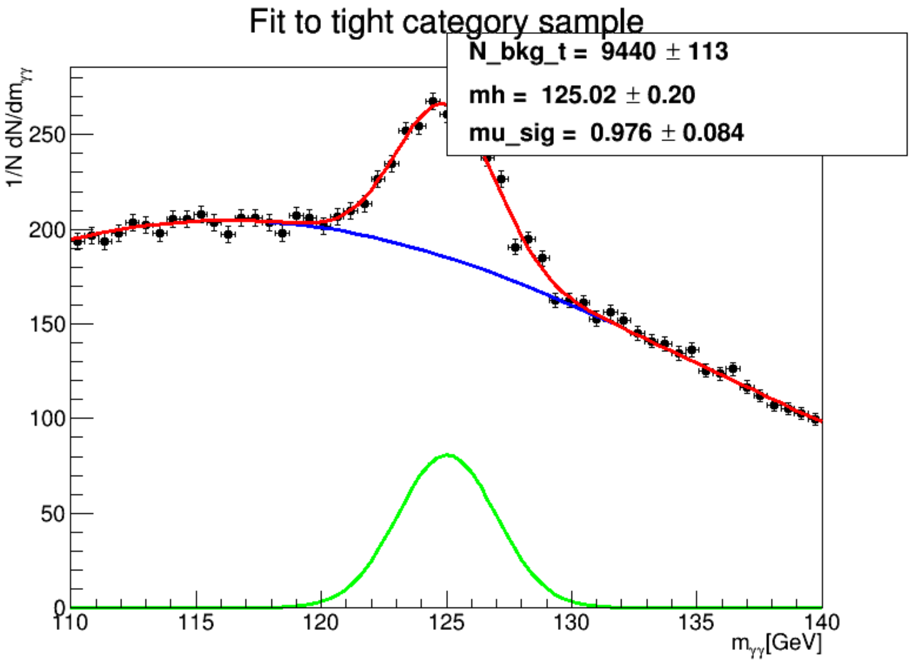
Plot: fit in tight signal(left up), loose signal(right up), tight background(left b), loose background(right b)



Fit result

Simultaneous fit in 2 categories:





$\mu = 0.97 \pm 0.084$

Conclusion

MVA could improve the performance in Higgs to di-photon channel by ~20%

- Significance:

  8.60 before applying MVA method (mass region [120, 130]GeV)

  11.93 after the BDT categorization

- Precision:

  9.84% in CDR ← Done by Kaili last year

  11% after adjusting the event selection ← Fitted by myself. Need to cross check with him

  8.4% after BDT