$qqH\mu\mu$ Analysis on the CEPC Experiment

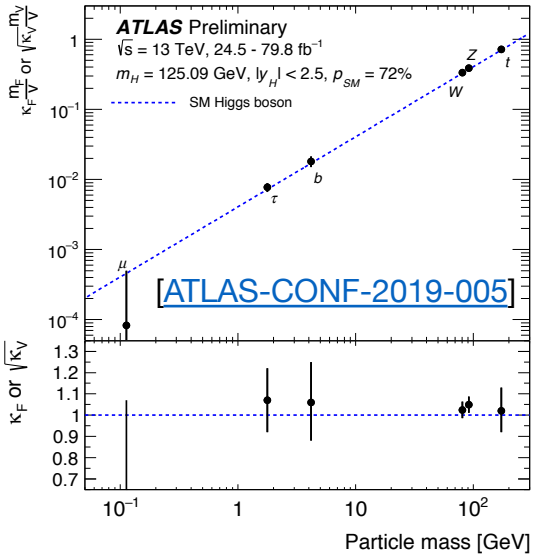
Yanping Huang[1], Shan Jin[2], Gang Li[1], Kunlin Ran[1], Manqi Ruan[1]

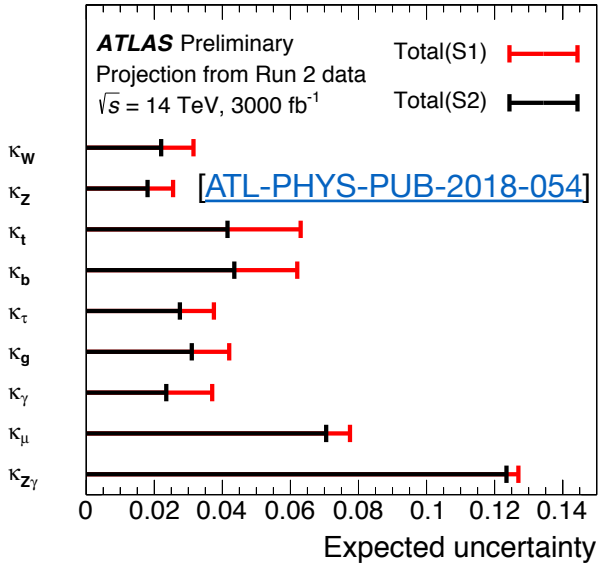
[1]IHEP, CAS

[2]Nanjing University

Introduction

- $H \to \mu\mu$ is important for probing the Higgs Yukawa couplings
  - The interactions of Higgs to the third generation charged fermions have already been observed by both ATLAS and CMS experiment [JHEP 08 (2016) 045]
  - Only upper limits exist on the Higgs interactions with other generations fermions
  - $H \to \mu\mu$ offers the best opportunity to measure Higgs Yukawa couplings to the second generation fermions
- In the ATLAS experiment, the latest (79.8 fb^{-1}) 95% CL of κ_μ is 1.51 [ATLAS-CONF-2019-005]

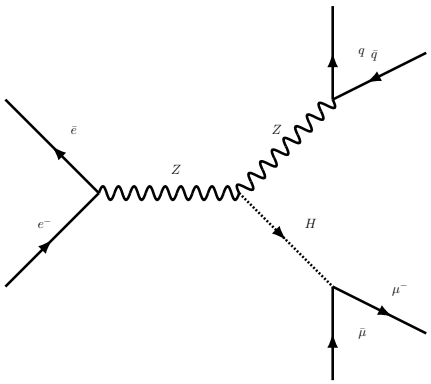


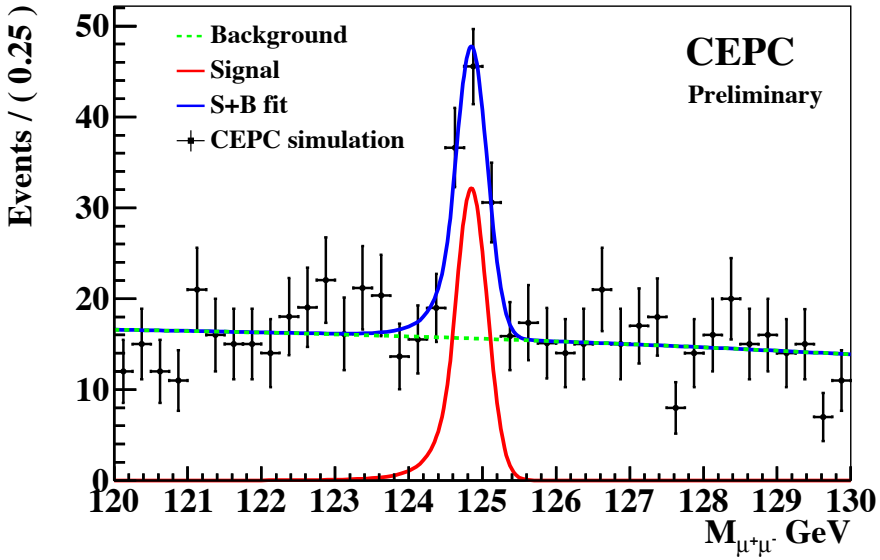


- In the projections with the ATLAS detector at HL-LHC (3000 fb^{-1}), the expected uncertainty of κ_μ is ~7% [ATL-PHYS-PUB-2018-054]

$ee \to Z(qq)H(\mu\mu)$ study in the CEPC

- With electron-positron colliders, we can gain much higher significance due to extremely clean background
- Focus on $ee \to Z(qq)H(\mu\mu)$ channel in the CEPC experiment
- Previous publication [10.1088/1674-1137/42/5/053001] gave counted significance at [124, 125] GeV: 10.8σ





- While the measurement is not perfect enough
  - The simulation didn't consider the Z boson width
  - The significance estimation is a counting result in a peak region
- Try to improve
  - Develop new selection criterial by keeping most signals and suppressing background
  - Use statistical (NLL) method to estimate significance
  - Further make event categories by applying MVA method

Detector and samples

- CEPC detector: v4, $\sqrt{s} = 240$ GeV, 3 T
- Sample (5 ab^{-1})
  - DST data: slimmed reconstruction samples without hits
  - Signal: qqh_e2e2, 100M
    - XS: 136.81 fb, BR: 2.176E-04
  - Background
    - 2 fermions, ~28 M
    - 4 fermions
      - Single W, ~18M
      - Single Z, ~8 M
      - WW, ~46 M
      - ZZ, ~6 M
      - Z or W: ~20 M
- To suppress background and keep most signals, event selections are developed/applied

Event selections

- After several studies, considering both signal/background efficiencies, event selections are finalized
  - Performances of each cut can be found in the backup

Cut	Reason
N_mum > 0, N_mup > 0	H to mumu requires 2 opposite charged muons
105 < M_mumu < 130 GeV	M_mumu should be close to M_H (125 GeV)
25 < N_particle < 115	di-jet system requires more objects than all leptons final stats
55 < M_qq < 125 GeV	M_qq should be close to M_Z (91.2 GeV)
P_qqmumu < 32 GeV, 195 < E_qqmumu < 265 GeV	qqmumu system should has 4 momentum close to (0, 0, 0, En), En=240 GeV
35 < E_mum < 100 GeV, 35 < E_mup < 100 GeV	To suppress WW background
16 < p_mumu < 72 GeV	To suppress hadronic background components with muons in jet clusters
N_em < 6, N_ep < 6, N_e < 9	To suppress semi-leptonic backgrounds; And hadronic backgrounds with electrons in jet clusters
E_em < 10 GeV, E_ep < 10 GeV, E_ee < 19 GeV	

Event selection efficiencies

- After event selections, summarize remaining signals and background components

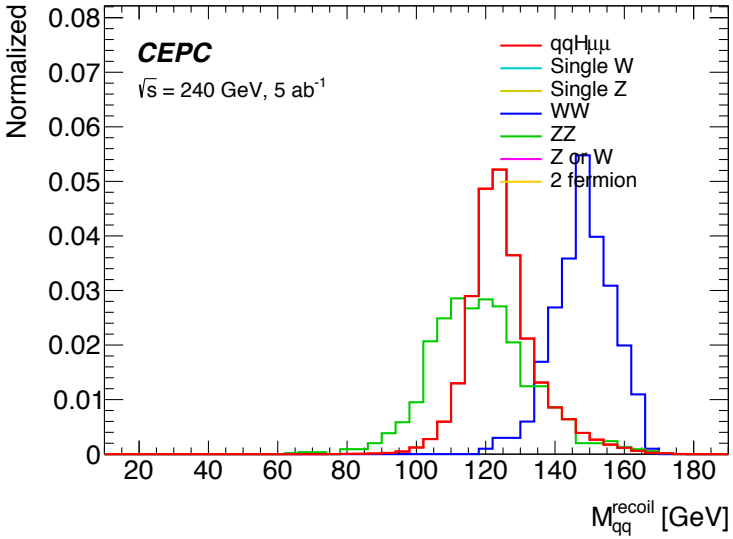
Component	qqh_e2e2			ww_sl0muq			ww_sl0tauq		
	Yield	Eff.	Rel.	Yield	Eff	Rel	Yield	Eff	Rel
Initial	148.85	100	100	12117200	100	100	12117800	100	100
N_mum > 0, N_mup > 0	148	99.43	99.43	623044	5.14	5.14	120233	0.99	0.99
105 < M_mumu < 130 GeV	123.75	83.14	83.62	4284.15	0.04	0.69	110.9	0	0.09
25 < N_particle < 115	123.02	82.64	99.41	4229.7	0.03	98.73	109.91	0	99.11
55 < M_qq < 125 GeV	122.02	81.97	99.19	2879.2	0.02	68.07	48.52	0	44.14
P_qqmumu < 32 GeV, 195 < E_qqmumu < 265 GeV	121.32	81.51	99.43	1947.52	0.02	67.64	20.79	0	42.86
35 < E_mum < 100 GeV, 35 < E_mup < 100 GeV	120.89	81.22	99.65	823.76	0.01	42.3	18.81	0	90.48
16 < p_mumu < 72 GeV	120.31	80.82	99.51	804.95	0.01	97.72	18.81	0	100
N_em < 6, N_ep < 6, N_e < 9	119.33	80.17	99.19	441.58	0	54.86	7.92	0	42.11
E_em < 10 GeV, E_ep < 10 GeV, E_ee < 19 GeV	116	77.93	97.21	246.53	0	55.83	1.98	0	25

Component	zz_l0taumu			zz_sl0mu_down			zz_sl0mu_up			zz_sl0tau_down		
	Yield	Eff	Rel	Yield	Eff	Rel	Yield	Eff	Rel	Yield	Eff	Rel
Initial	93250	100	100	680700	100	100	436950	100	100	336550	100	100
N_mum > 0, N_mup > 0	44626.5	47.86	47.86	328021	48.19	48.19	201081	46.02	46.02	9475.89	2.82	2.82
105 < M_mumu < 130 GeV	1749.25	1.88	3.92	7093.73	1.04	2.16	4290.11	0.98	2.13	3.96	0	0.04
25 < N_particle < 115	33.6	0.04	1.92	7018.41	1.03	98.94	4216.77	0.97	98.29	2.97	0	75
55 < M_qq < 125 GeV	12.13	0.01	36.11	6754.8	0.99	96.24	3601.35	0.82	85.41	2.97	0	100
P_qqmumu < 32 GeV, 195 < E_qqmumu < 265 GeV	10.27	0.01	84.62	6661.65	0.98	98.62	3548.83	0.81	98.54	2.97	0	100
35 < E_mum < 100 GeV, 35 < E_mup < 100 GeV	8.4	0.01	81.82	6177.04	0.91	92.73	3308.01	0.76	93.21	2.97	0	100
16 < p_mumu < 72 GeV	8.4	0.01	100	6079.92	0.89	98.43	3216.84	0.74	97.24	1.98	0	66.67
N_em < 6, N_ep < 6, N_e < 9	8.4	0.01	100	1797.71	0.26	29.57	895.88	0.21	27.85	0.99	0	50
E_em < 10 GeV, E_ep < 10 GeV, E_ee < 19 GeV	1.87	0	22.22	917.69	0.13	51.05	433.07	0.1	48.34	0.99	0	100

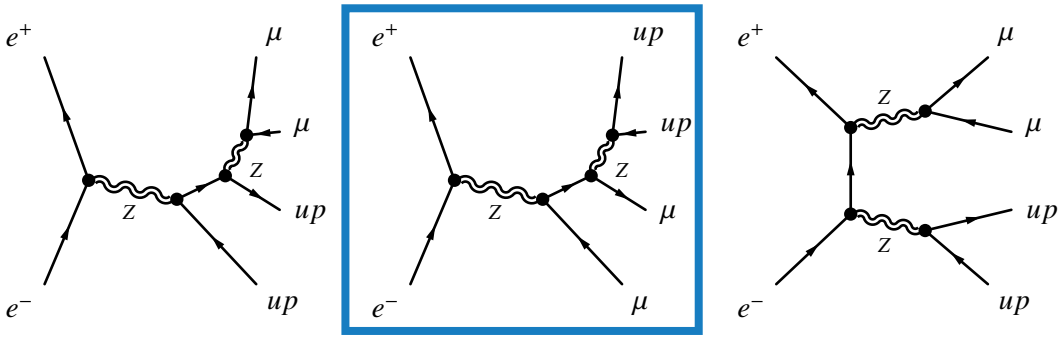
- The signal efficiency: 78%
- The background
  - ZZ (muons/jets, taus/muons, taus/jets): 79%; WW (muon/jets, tau/jets): 21%

Check M_{qq}^{recoil} discrimination

- Essentially thinking M_{qq}^{recoil} would have a good separation power between the signal/ZZ background: $M_{qq}^{recoil} = M\left(p(0,0,0,\sqrt{s}) - p_{qq}\right)$

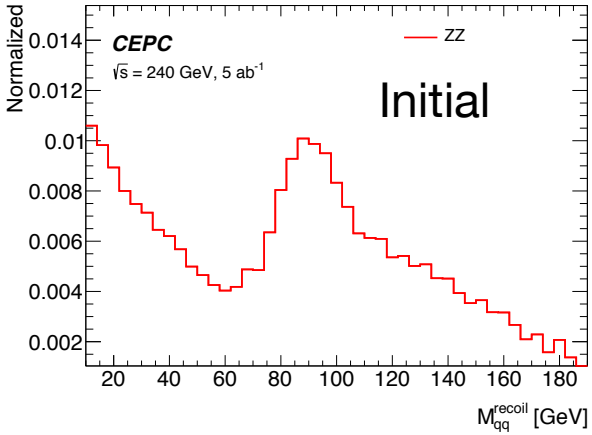


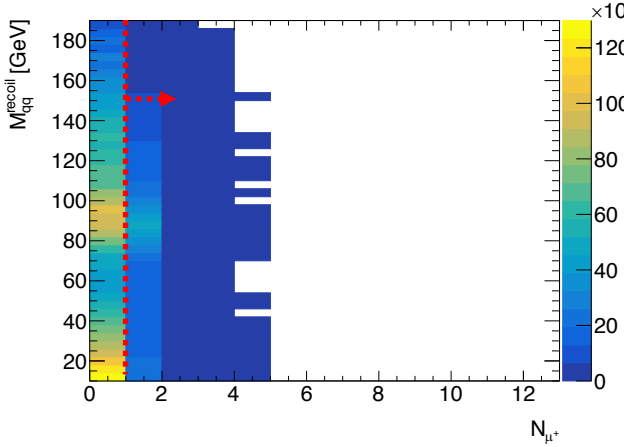
- M_{qq}^{recoil} is expected to be close to M_H in the signal, while close to M_Z in the ZZ background, but not the case in the background distribution after all cuts
- Guess only the middle process dominates in the remaining ZZ events

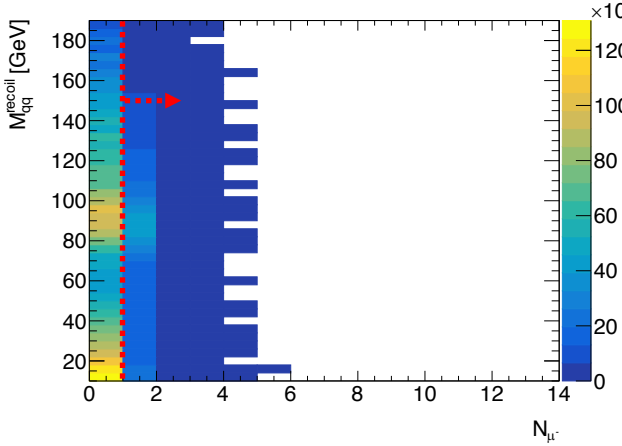


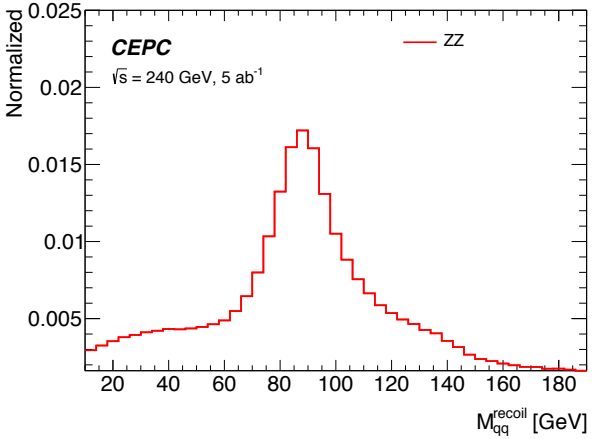
N_μ cuts

- Go over selections step by step to check the M_{qq}^{recoil} changes

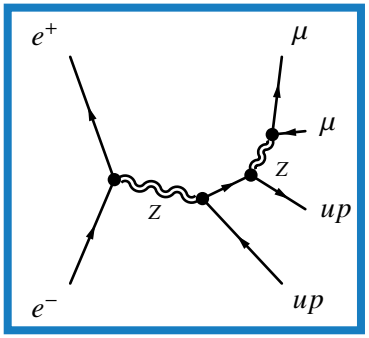


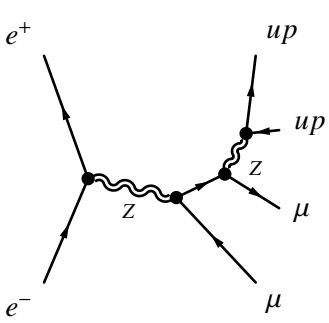


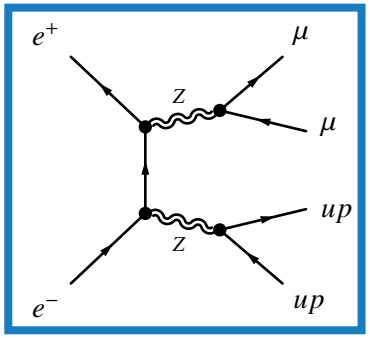




- After N_μ cuts, the remaining ZZ events would include all $Z(qq)Z(\mu\mu)$ processess

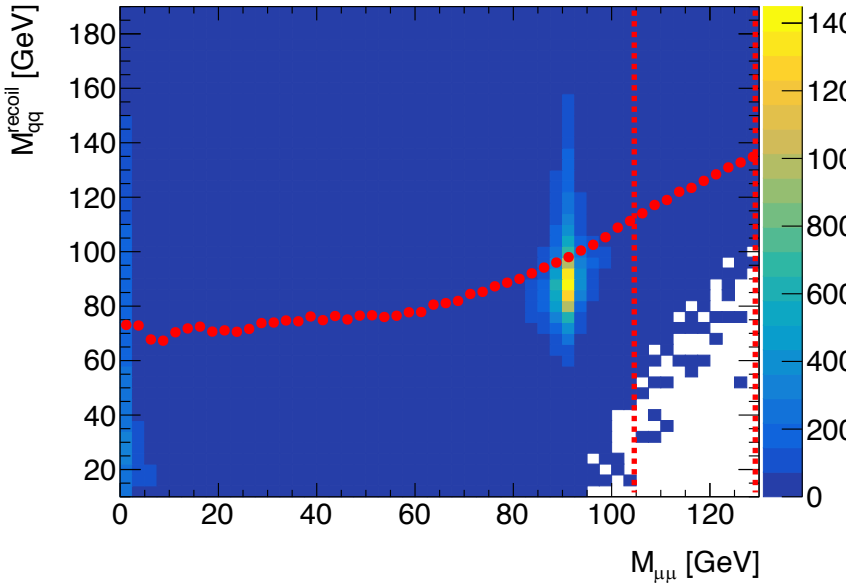


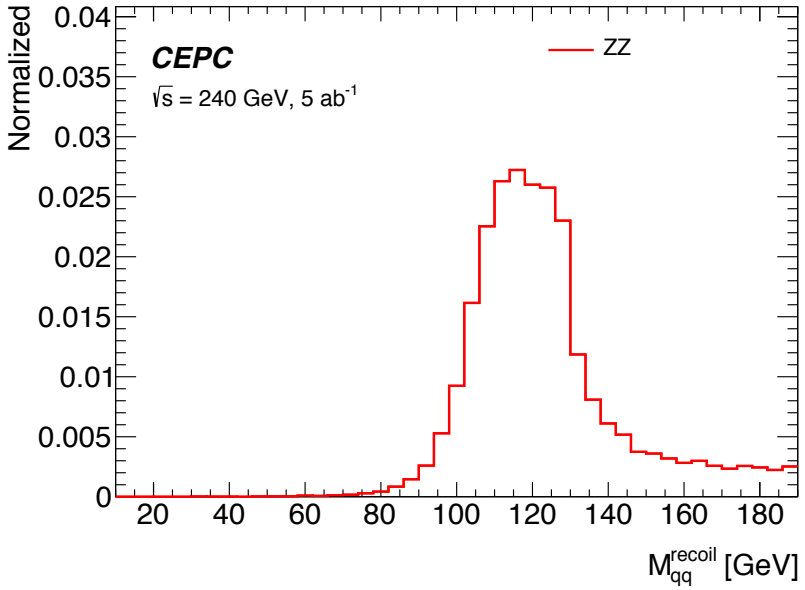




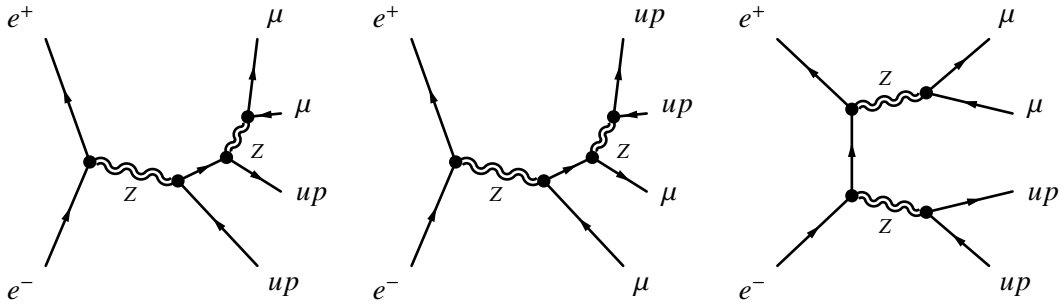
- The first and third processes would have a mass peak close to M_Z in the M_{qq}^{recoil} distribution

The $M_{\mu\mu}$ cut





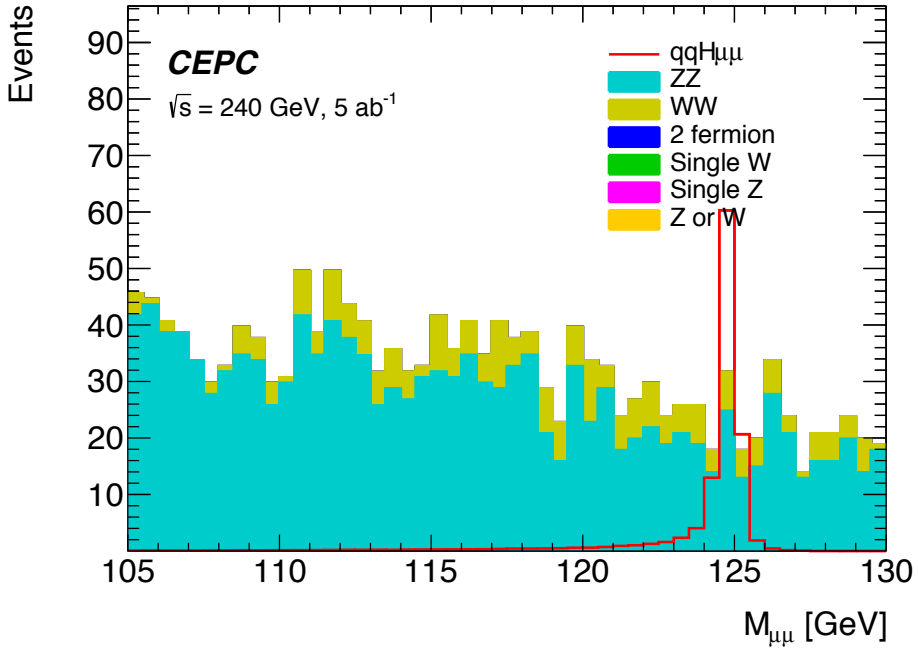
- After $M_{\mu\mu}$ cut ([105, 130] GeV), the 91.2 GeV M_{qq}^{recoil} peak disappear



- Since di-muon mass window (M_H) is required, the first and third processes is largely suppressed, thus the Z boson resonant mass disappear
- Intend not to apply M_{qq}^{recoil} cut in the selection criterial

The mass spectrum

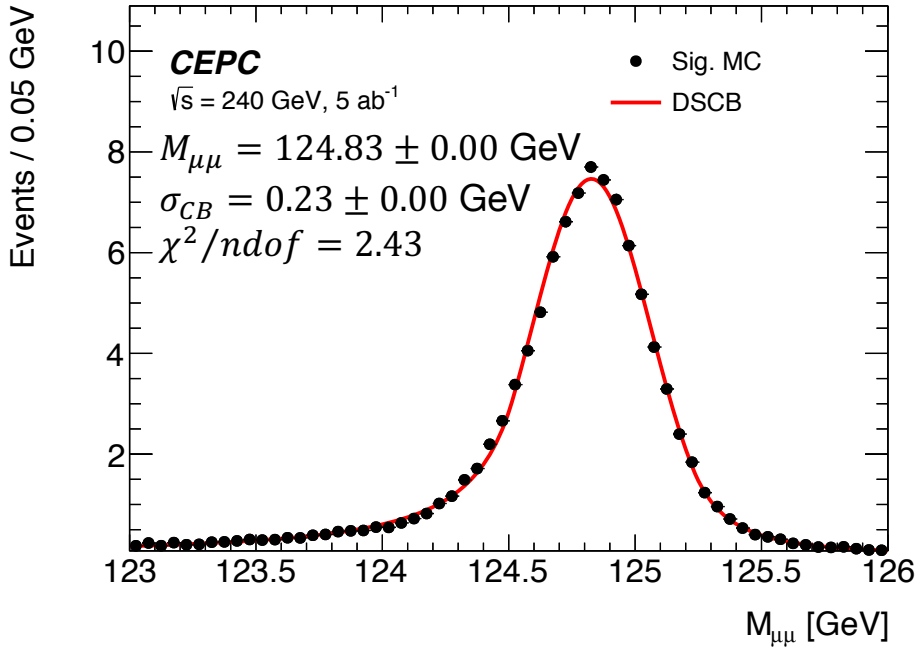
- After selection criterial, the signal shows high significance in the $\mu\mu$ mass spectrum



- Estimate cut-based signal sensitivity in [124,125] GeV (~63% of selected signal)
  - $\sqrt{2\left[(s+b)\cdot\ln\left(1+\frac{s}{b}\right)-s\right]}=8.74\sigma$

Signal modelling

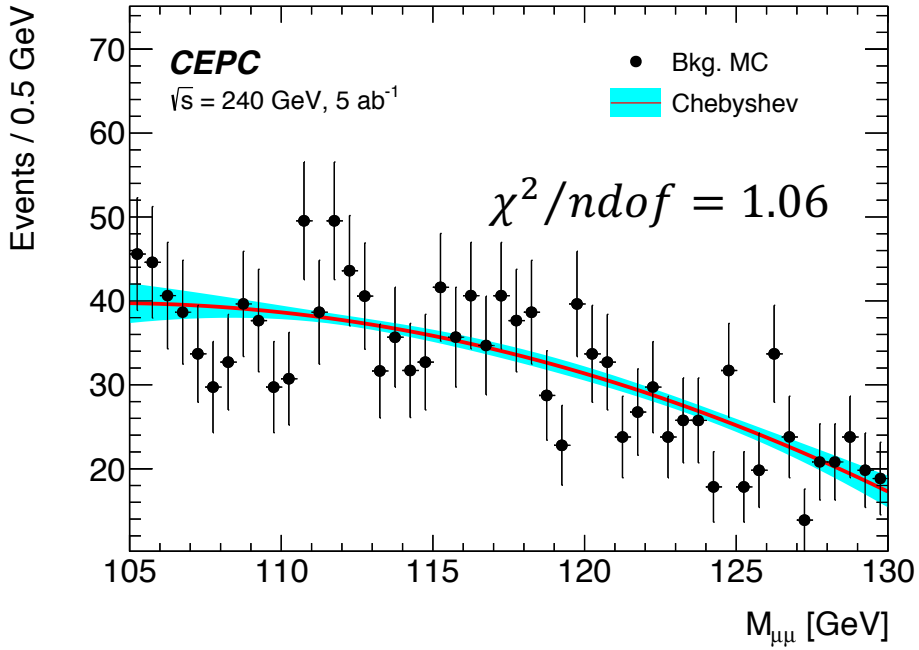
- Fitting function: DSCB
- $$f(t) = N \cdot \begin{cases} e^{-0.5t^2}, \ if \ -\alpha_{low} \le t \le \alpha_{high} \\ e^{-0.5\alpha_{low}^2} \left[\frac{\alpha_{low}}{n_{low}} \left(\frac{n_{low}}{\alpha_{low}} - \alpha_{low} - t\right)\right]^{-n_{low}}, \ if \ t < -\alpha_{low} \\ e^{-0.5\alpha_{high}^2} \left[\frac{\alpha_{high}}{n_{high}} \left(\frac{n_{high}}{\alpha_{high}} - \alpha_{high} + t\right)\right]^{-n_{high}}, \ if \ t > \alpha_{high} \end{cases}$$
- $t = (m_{\mu\mu} - \mu_{CB})/\sigma_{CB}$



- DSCB shows great agreement with the signal MC

Background modelling

- Fitting function: second order Chebyshev polynomial
- $f(m_{\mu\mu}) = N \cdot [1 + a_0 m_{\mu\mu} + a_1(2m_{\mu\mu}^2 - 1)]$



- The statistics of background is a bit low: 1617, may introduce bias in choosing background model
- Could investigate further more about background models in the future

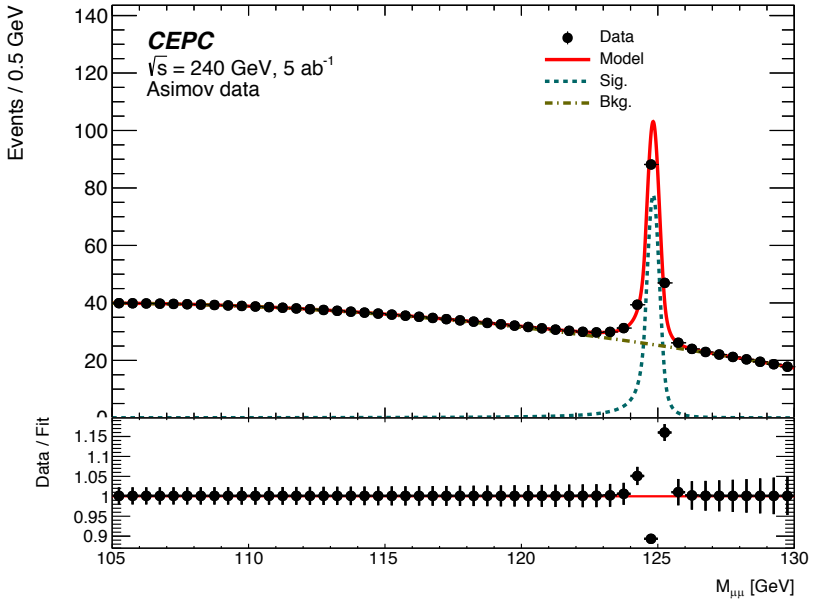
Statistical procedure

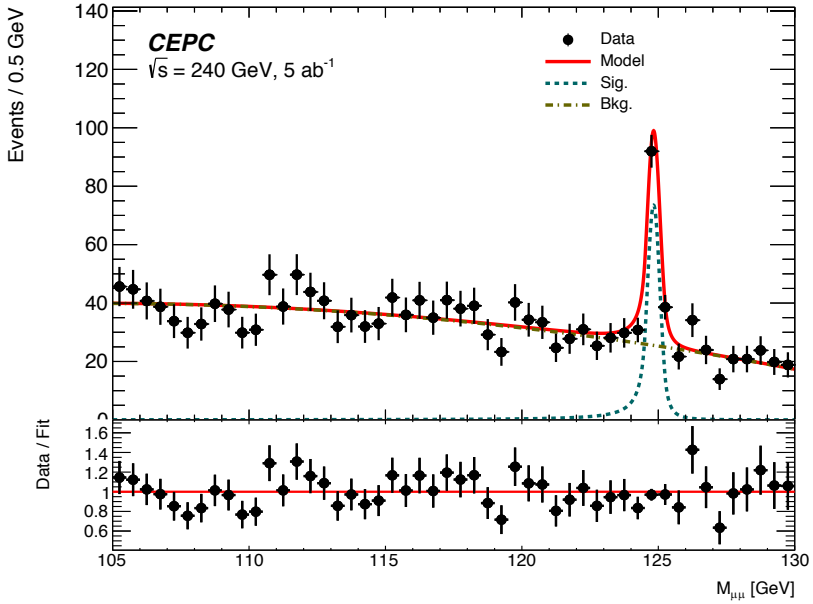
- Use statistical methods to get significance
- Use signal/background models to fit MC events and also generate Asimov dataset to smooth the background
- ML method is applied

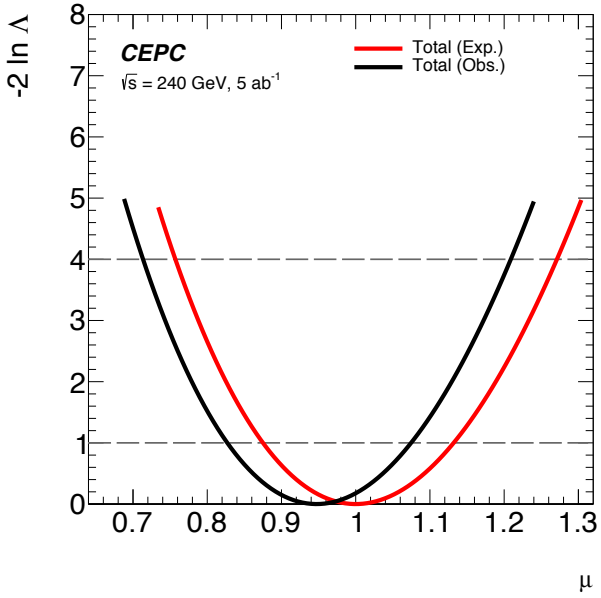
- Likelihood model
  - $L = \exp[-(N_S + N_B)] \prod_i^n [N_s f_s(x_i) + N_B f_B(x_i)]$
  - $N_S = \mu \cdot \mathcal{L} \cdot \sigma_{Z(qq)H}^{SM} \cdot B_{\mu\mu}^{SM} \cdot Acc$

  - POI: μ (signal strength of $ee \to Z(qq)H(\mu\mu)$)
  - Observable: $m_{\mu\mu}$
  - $\mathcal{L} = 5000$ fb^{-1}, $\sigma_{Z(qq)H}^{SM} = 136.81$ fb, $B_{\mu\mu}^{SM} = 2.176E-04$
  - f_s, f_B are from the signal/background modellings, all parameters are fixed since the fitted dataset is the MC; It's already used for choosing signal/background functions
  - Acc is from the event selections (77.93%)

Fitting results



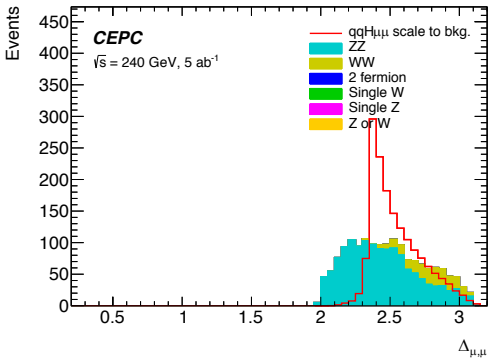


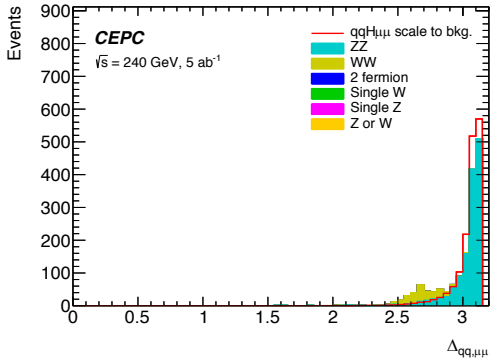


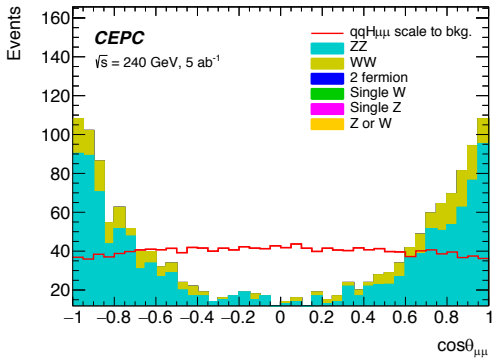
- $\mu = 0.95^{+0.13}_{-0.12}(1.00 \pm 0.13)$
- Significance: 10.59σ (10.78σ)

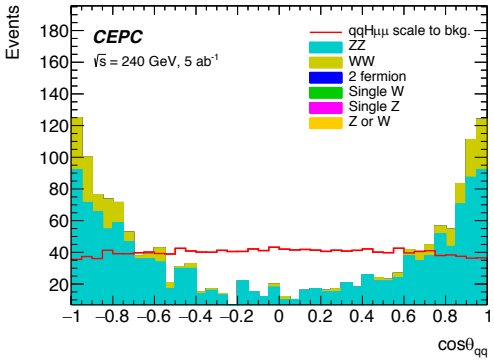
MVA optimizations

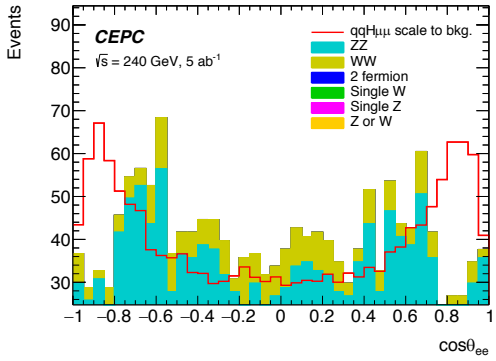
- Next to apply discriminant variables into MVA (BDTG) to optimize significances
- Search for discriminant variables

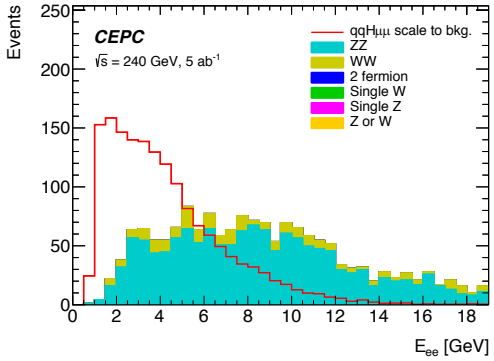


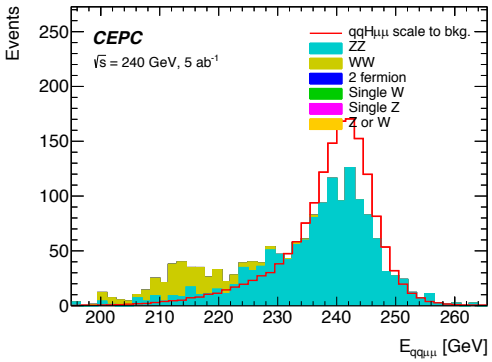


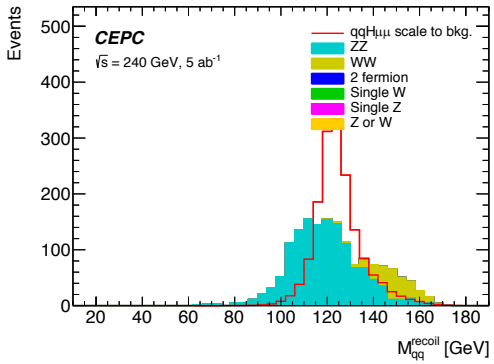


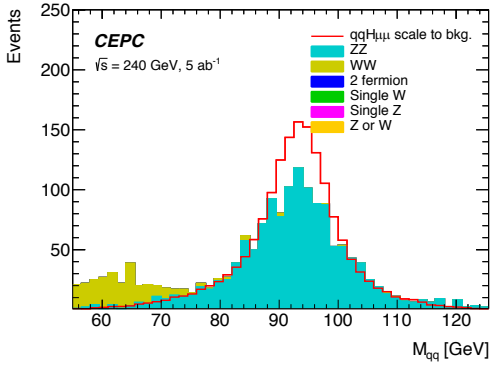






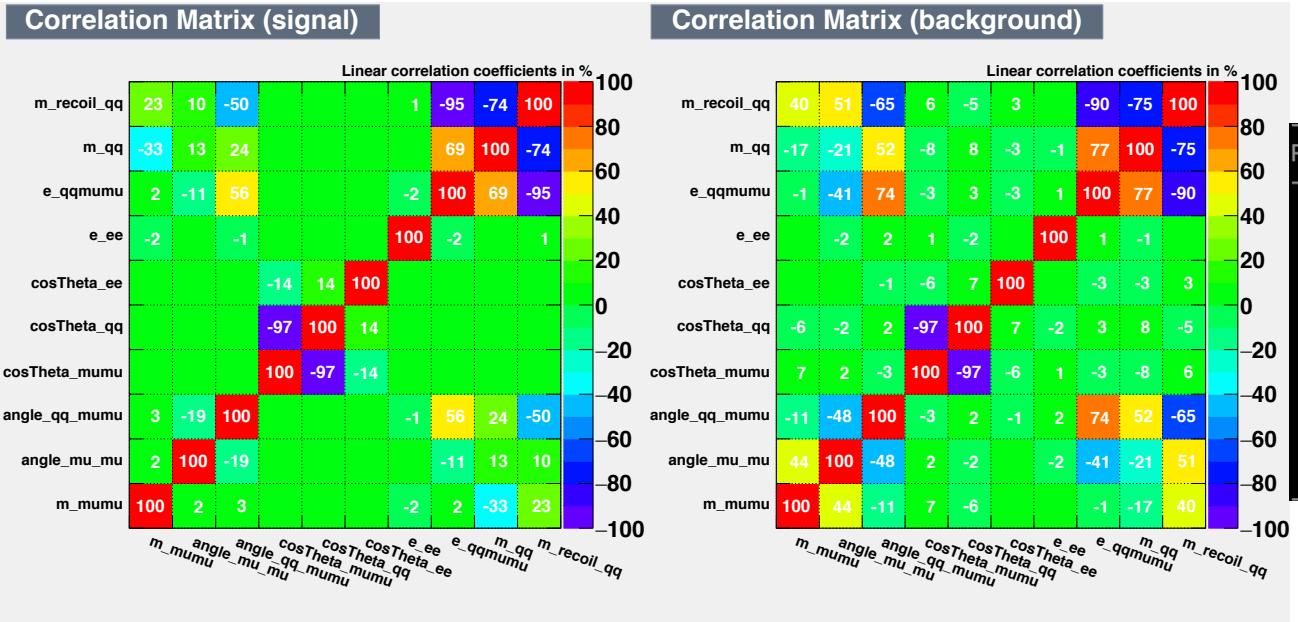






Select training variables

- Training variables selection criterial
  - Not highly correlated with $M_{\mu\mu}$
  - Not strongly anti-correlated with each-other

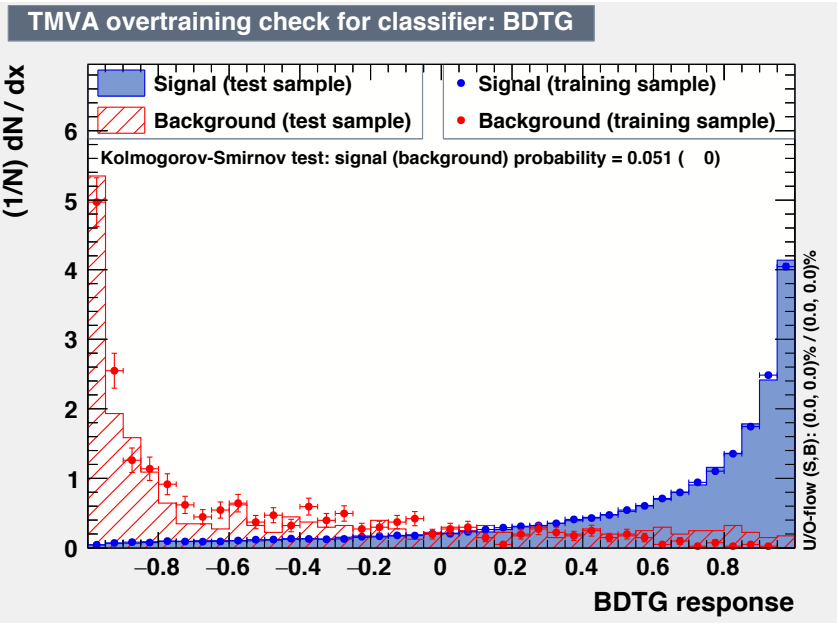


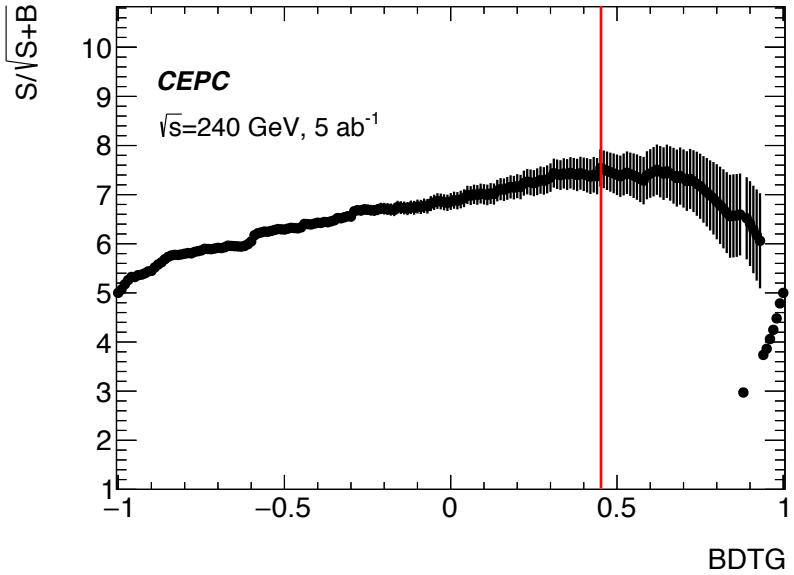
Rank	Variable	Variable Importance
1	m_qq	1.500e-01
2	m_mumu	1.234e-01
3	angle_mu_mu	1.217e-01
4	cosTheta_ee	1.168e-01
5	cosTheta_mumu	9.340e-02
6	e_ee	9.250e-02
7	cosTheta_qq	9.059e-02
8	e_qqmumu	8.207e-02
9	m_recoil_qq	7.172e-02
10	angle_qq_mumu	5.781e-02

- Drop one of the strongly correlated variables referring to the ranking (bottom to top)
  - Eventually keep $cos\theta_{ee}$, $cos\theta_{\mu\mu}$, $\Delta_{\mu,\mu}$, M_{qq}, E_{ee}, $E_{qq\mu\mu}$

BDTG training results

- After training with 6 variables: $cos\theta_{ee}$, $cos\theta_{\mu\mu}$, $\Delta_{\mu,\mu}$, M_{qq}, E_{ee}, $E_{qq\mu\mu}$, get the BDTG response





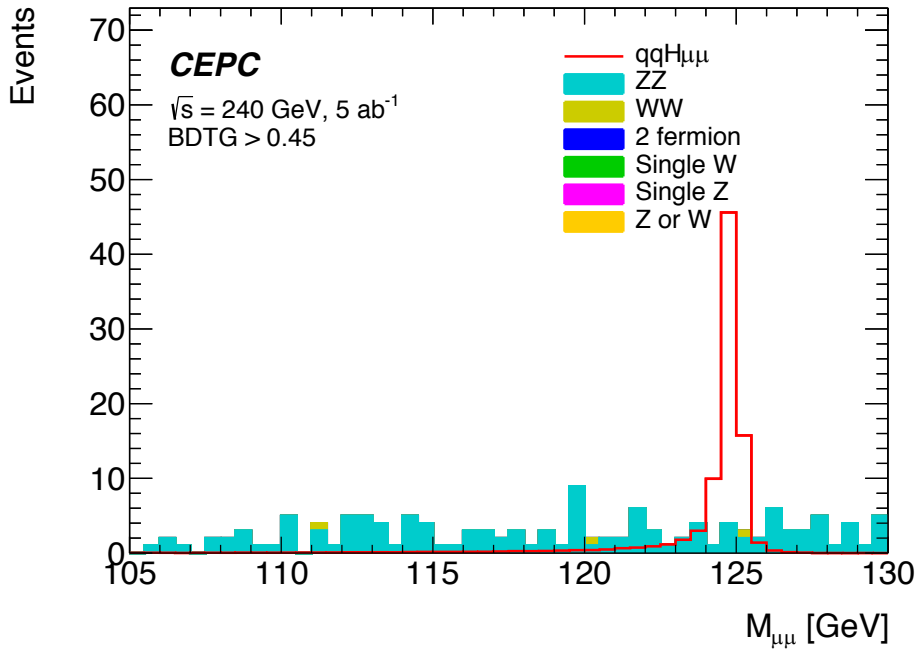
- There is a overtraining in the background due to poor statistics: ~1600
- Scan the total sensitivity ($S/\sqrt{S+B}$) vs BDTG to find the optimal BDTG point
- The sensitivity is estimated in the 90% signal coverage region

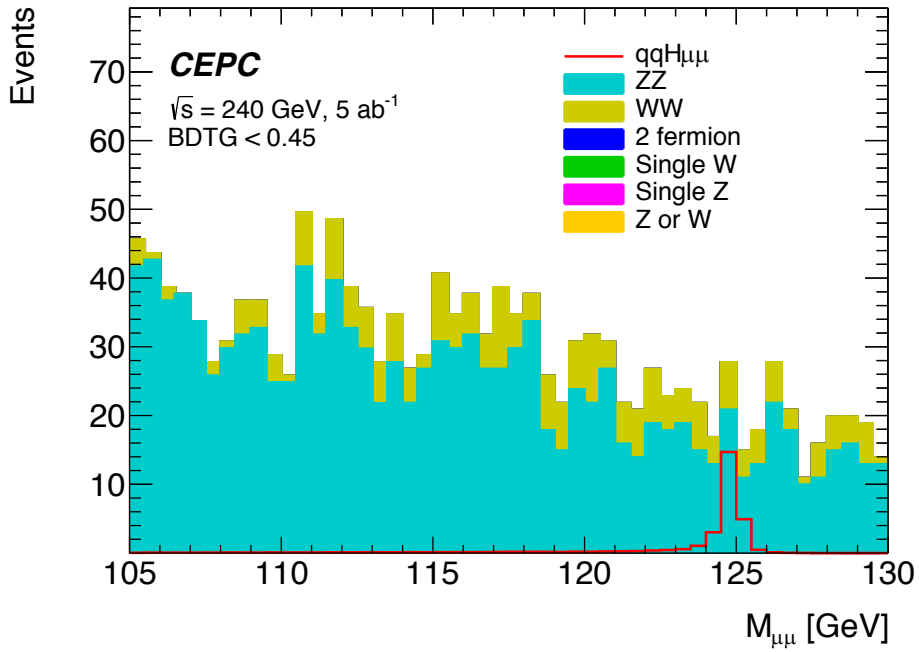
	Sig yield	Bkg yield	Sensitivity	Mass range (GeV)
BDTG > 0.45	86.20 +/- 0.51	198.20 +/- 19.82	7.46 +/- 0.27	[120.78 - 125.33]
BDTG < 0.45	29.77 +/- 0.30	1402.95 +/- 52.73	1.08 +/- 0.03	[114.08 - 125.28]
Total	115.97 +/- 0.59	1601.15 +/- 56.33	7.54 +/- 0.38	

- Get the maximum sensitivity cutting BDTG at 0.45, split events into tight/loose categories

Categories signal/background fractions

- Perform $M_{\mu\mu}$ distributions after BDTG cuts



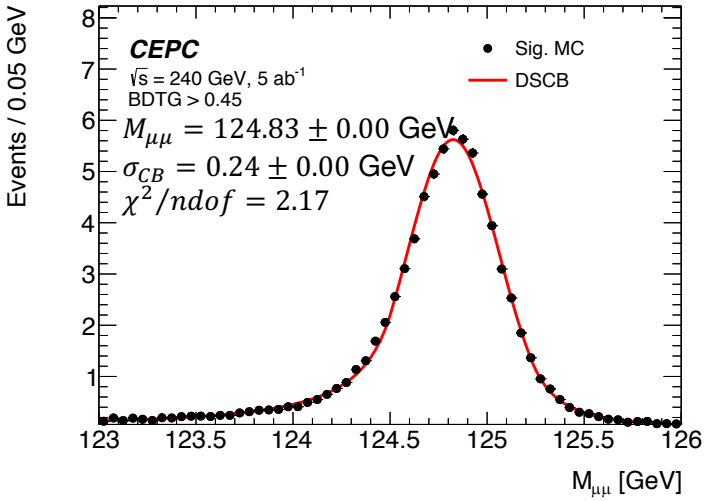


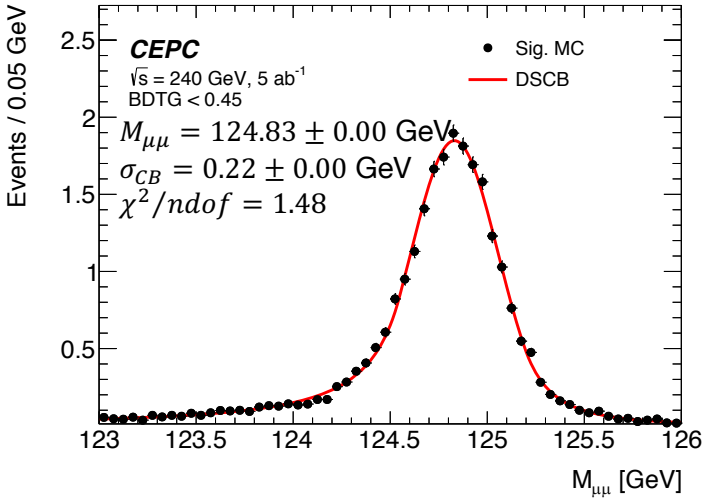
Fraction	BDTG > 0.45	BDTG < 0.45	Total yield
Sig	0.74	0.26	115.97
ZZ	0.1	0.9	1353.62
WW	0.01	0.99	248.52

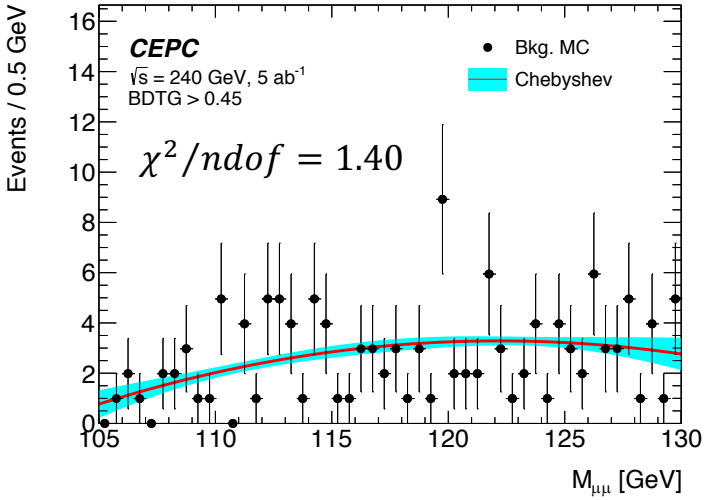
- In the BDTG tight category, nearly all WW backgrounds are rejected

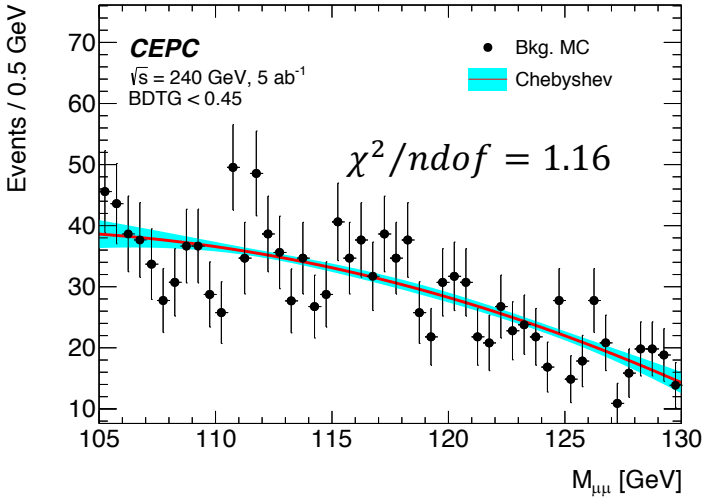
Signal/background modellings in categories

- Making signal/background models in each category
  - Signal: DSCB
  - Background: second order Chebyshev polynomial





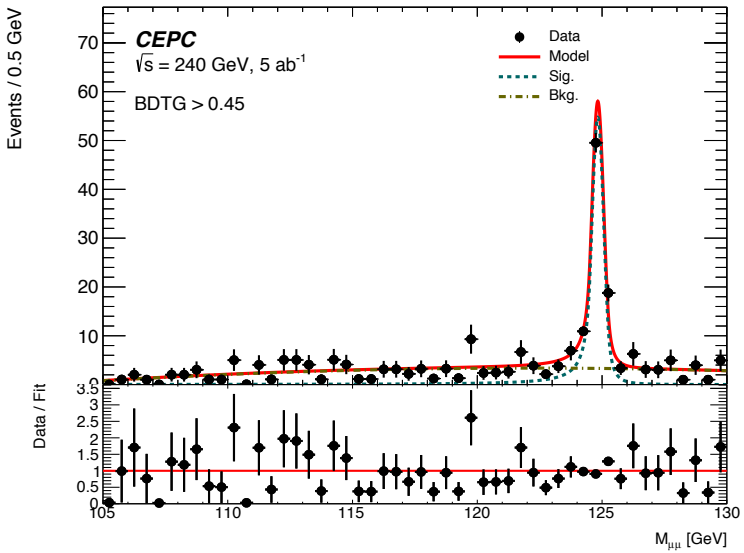


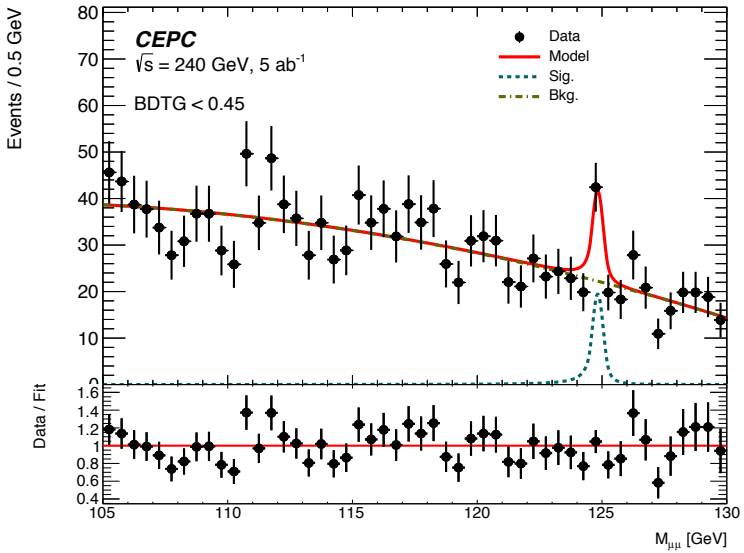


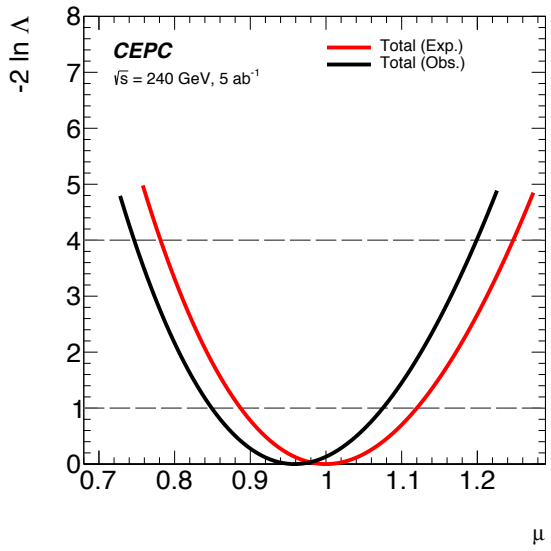
- DSCB still works well for the signal modelling
- Since the statistic fluctuation is large in the background, hard to tell whether the model is perfect, could make further study in the future

The total significance

- Make a simultaneous fit on 2 categories to get the total signal strength







- $\mu = 0.96^{+0.12}_{-0.11}(1.00^{+0.12}_{-0.11})$
- Significance: 14.41σ (14.57σ)

- Compare to the results with 1 inclusive category: 10.59σ (10.78σ), the significance increases ~35%

Summary

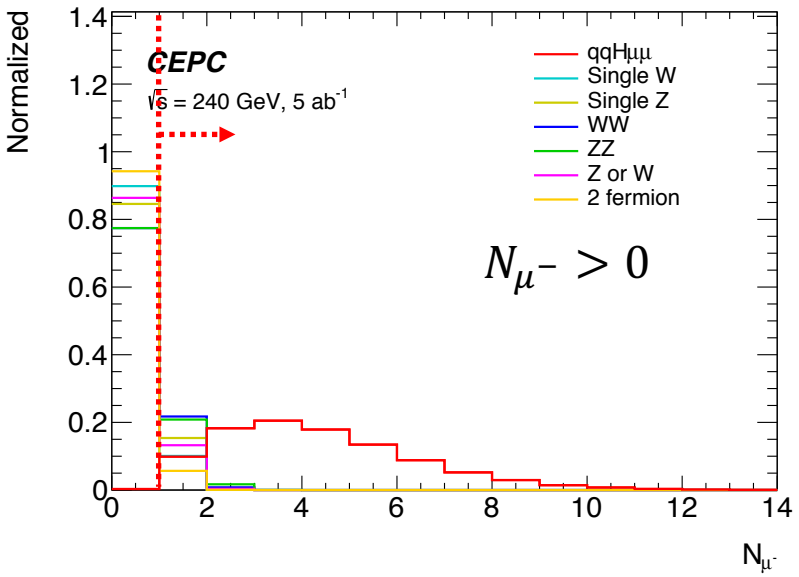
- Study $ee \to Z(qq)H(\mu\mu)$ on the CEPC experiment
  - Develop selection criterial to try to keep most signals and reject backgrounds
  - After event selections, the signal efficiency is 78%; the dominant background is ZZ to muons/jets

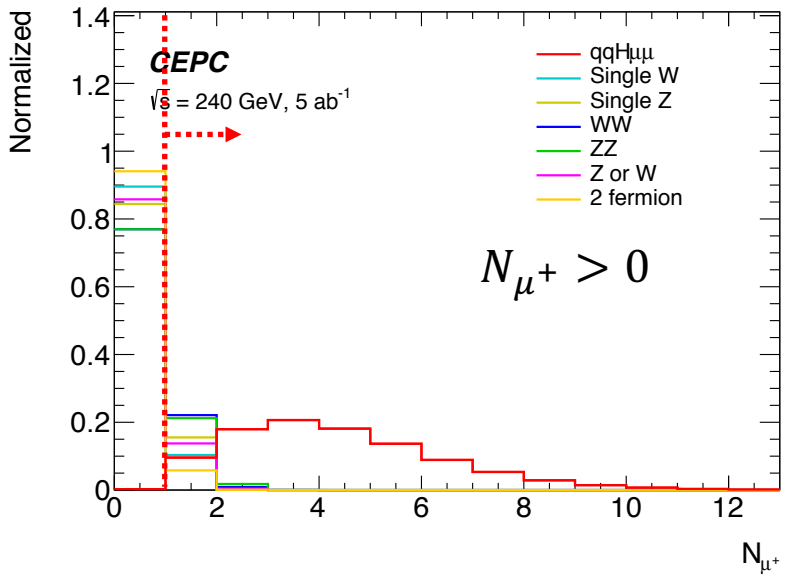
  - Use DSCB as the signal model, second order Chebyshev polynomial as the background model
  - To fit the mass spectrum of dimuon to estimate the $ee \to Z(qq)H(\mu\mu)$ significance: 10.59σ (10.78σ)

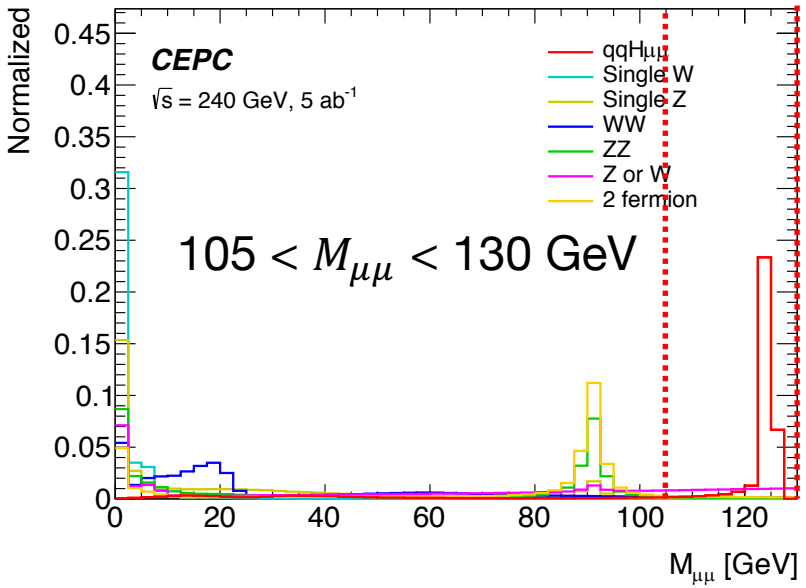
  - Further choose appropriate variables for the MVA optimization: $cos\theta_{ee}$, $cos\theta_{\mu\mu}$, $\Delta_{\mu,\mu}$, M_{qq}, E_{ee}, $E_{qq\mu\mu}$
  - Apply the optimal BDTG cut to split events into 2 categories
  - Still use the same signal/background models as inclusive case, make a simultaneous fit on 2 categories
  - Get the significance: 14.41σ (14.57σ), which has 35% improvement w.r.t the inclusive case

Backup

Selection criterial

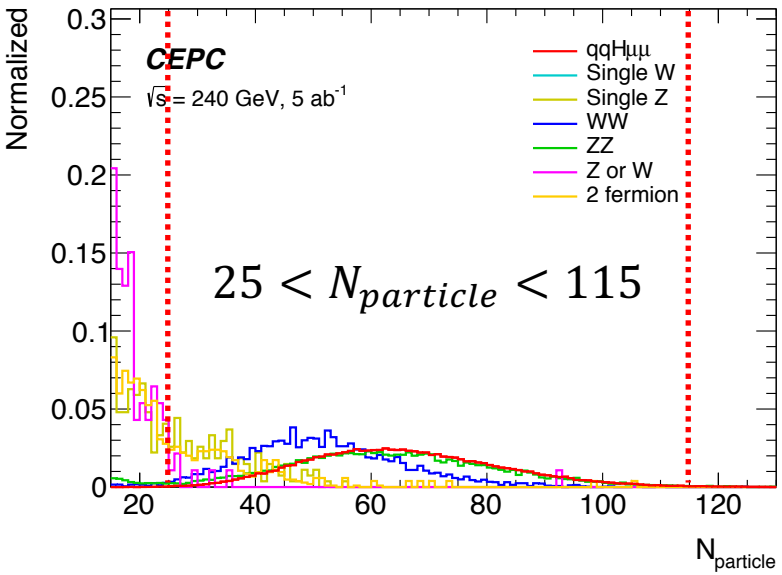




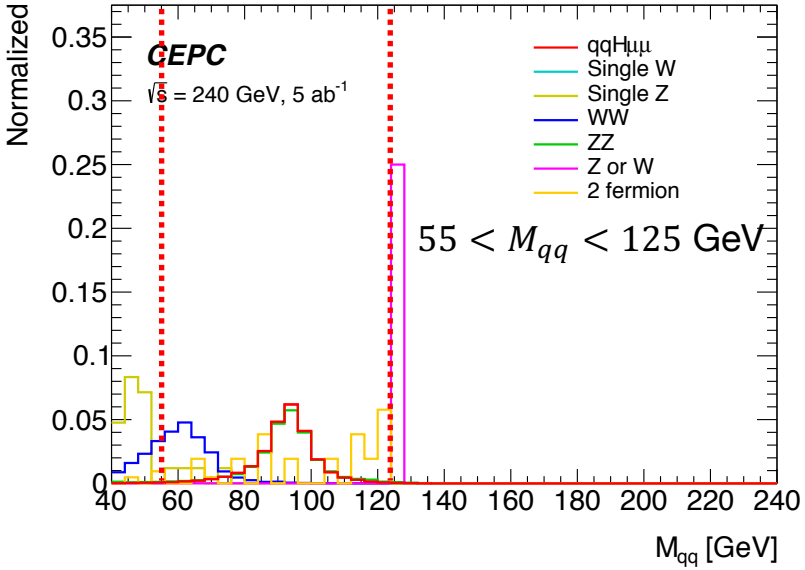


- N_μ cuts: $H \to \mu\mu$ requires 2 opposite charged muons
- $M_{\mu\mu}$ cut: $M_{\mu\mu}$ should close to M_H (125 GeV)

Selection criterial



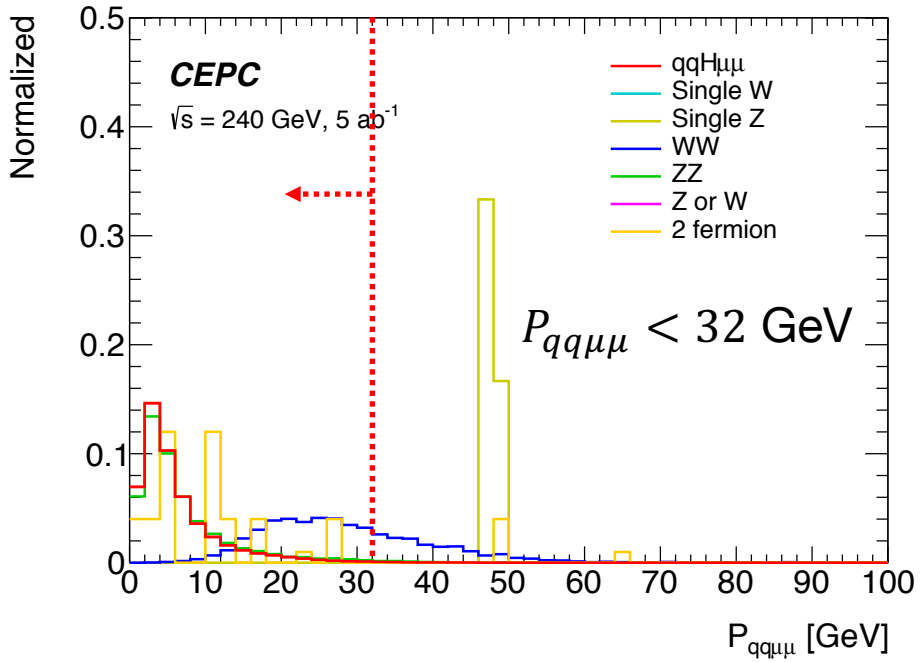
- All reconstructed visible objects are grouped as “particle”
- $N_{particle}$ cut: di-jet system requires more objects
- Suppress background components with lepton final stats
  - Single Z to muons/electrons
  - ZZ/WW to muons/taus
  - 2 fermions (muons/taus)

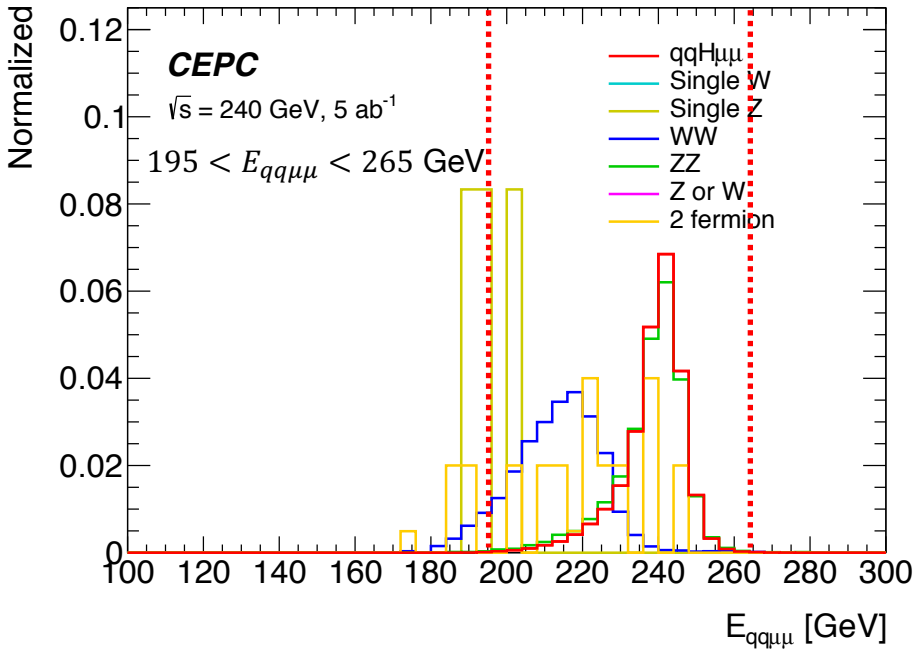


- 4 momentum of all visible particles: $p_{particle}$; 4 momentum of di-muon: $p_{\mu\mu}$
- Didn’t apply jet-cluster algorithm, while define di-jet system: $p_{qq} = p_{particle} - p_{\mu\mu}$
- Since the dominant background would be $Z(\mu\mu)Z(qq)$ (see later), the rough definition of di-jet system makes sense for the specific channel
- M_{qq} cut: M_{qq} should be close to M_Z (91.2 GeV)

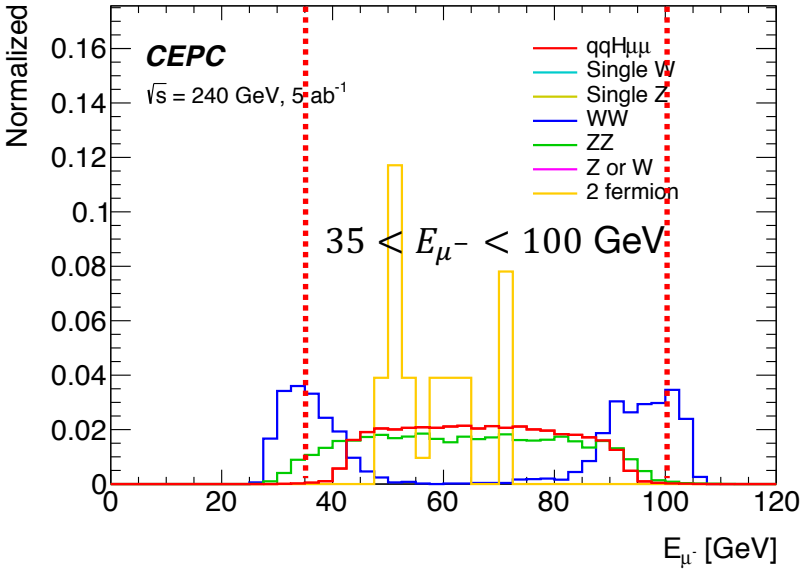
Selection criterial

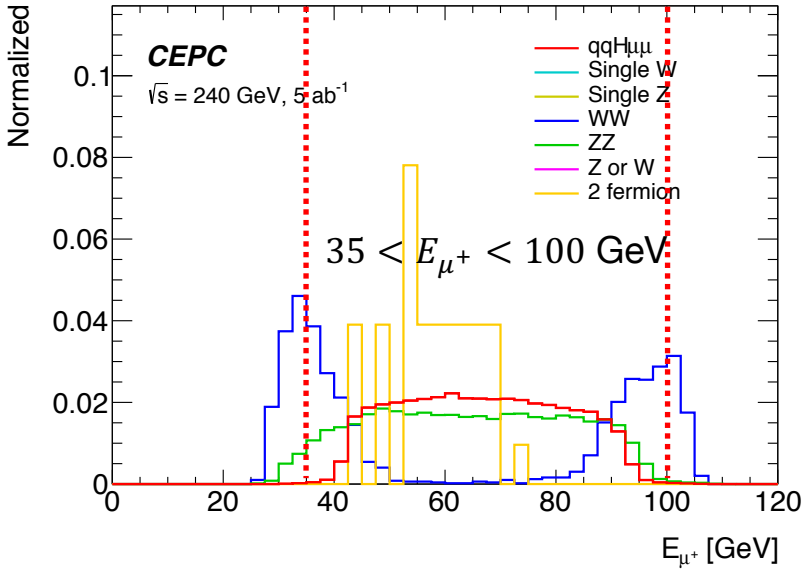
- $p_{qq\mu\mu}$ cut: $ee \to Z(qq)H(\mu\mu)$ system should has 4 momentum close to (0, 0, 0, $\sqrt{s}$), $\sqrt{s} = 240$ GeV

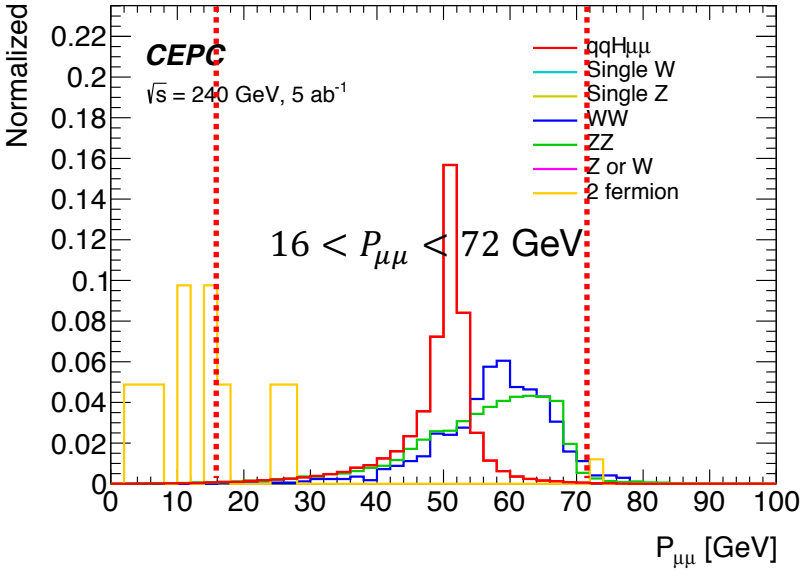




Selection criterial

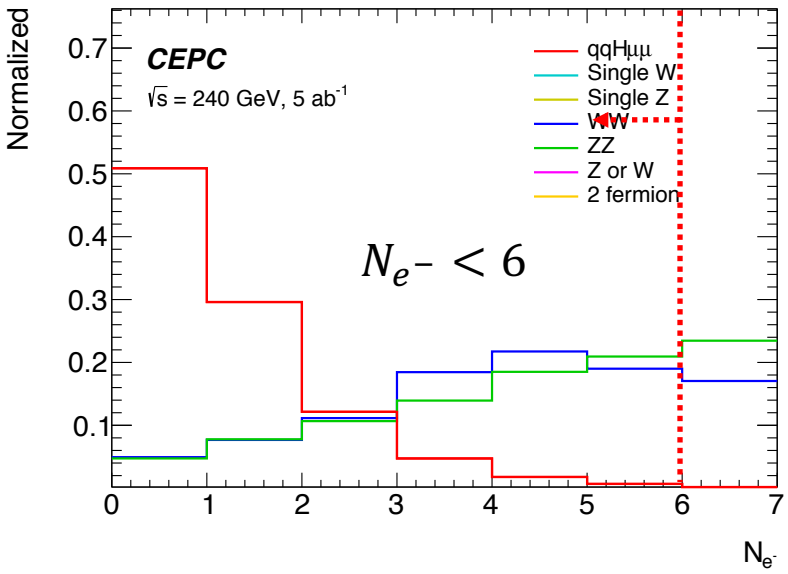


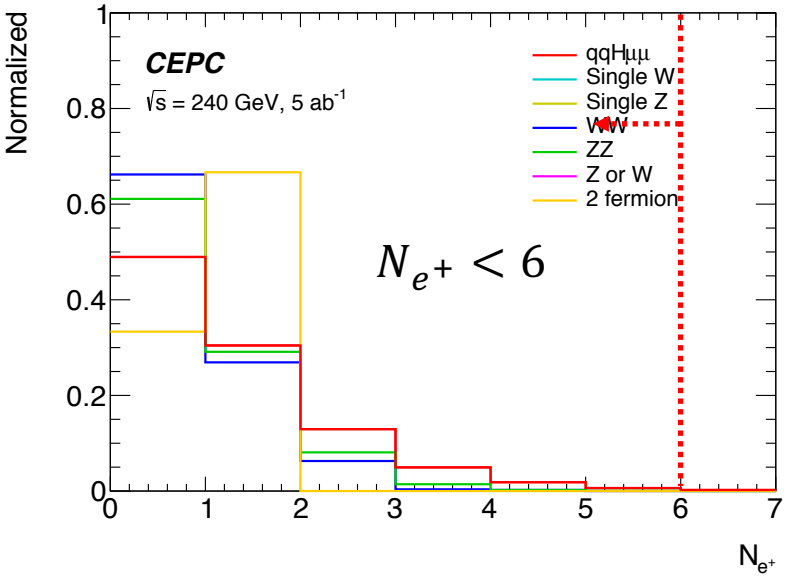


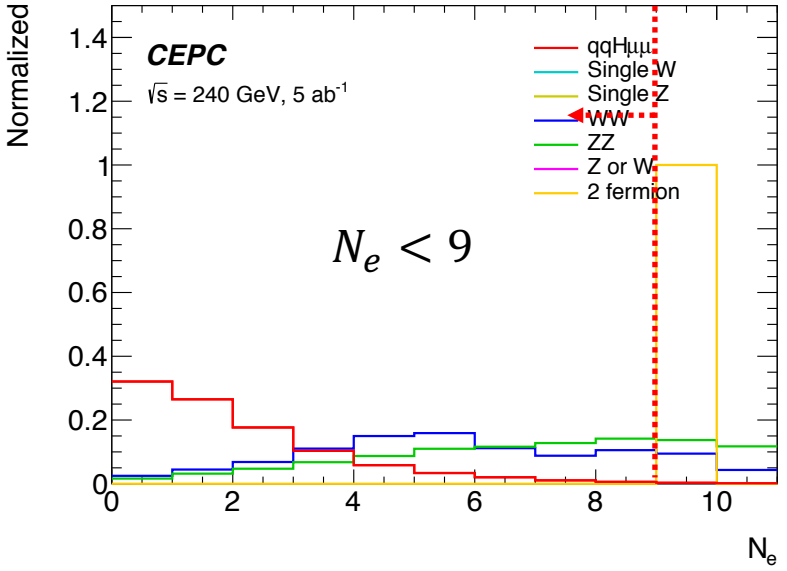


- E_μ cuts: To suppress WW background
  - WW to muon/jets
- $p_{\mu\mu}$ cut: To suppress hadronic background components with muons in jet clusters
  - Di-jet background
  - ZZ to taus/jets

Selection criterial

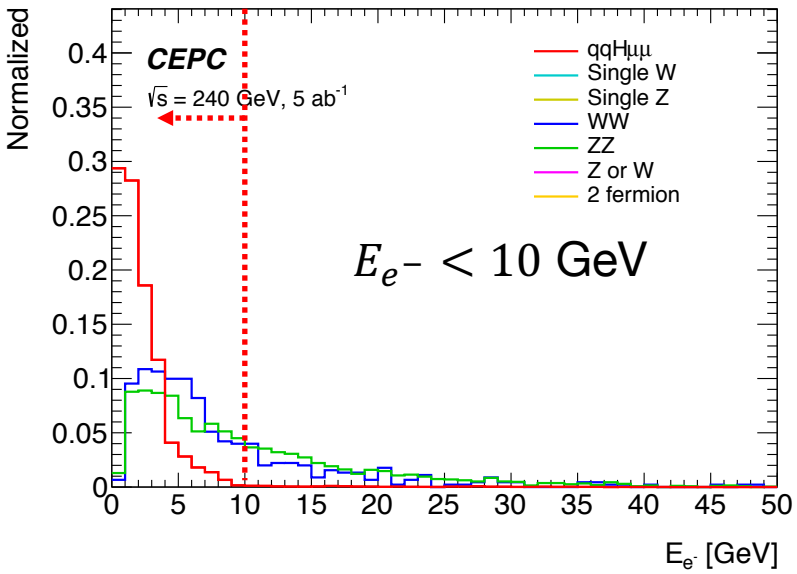


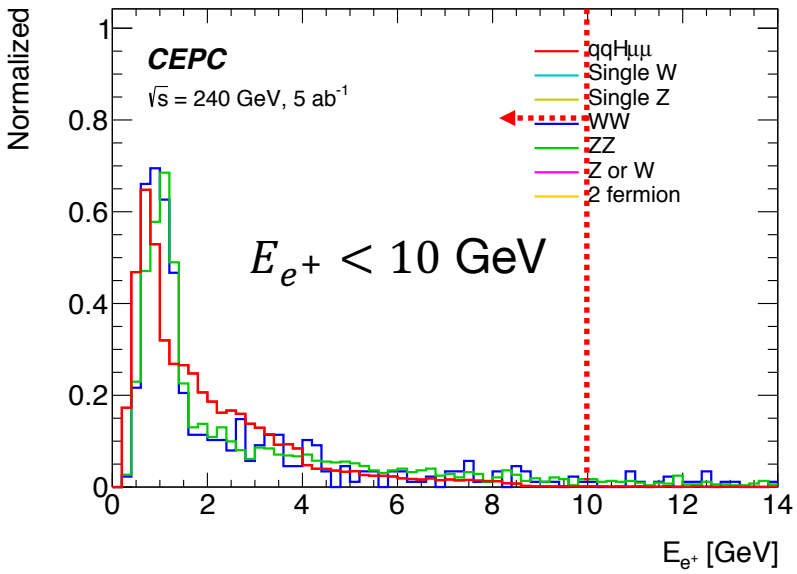


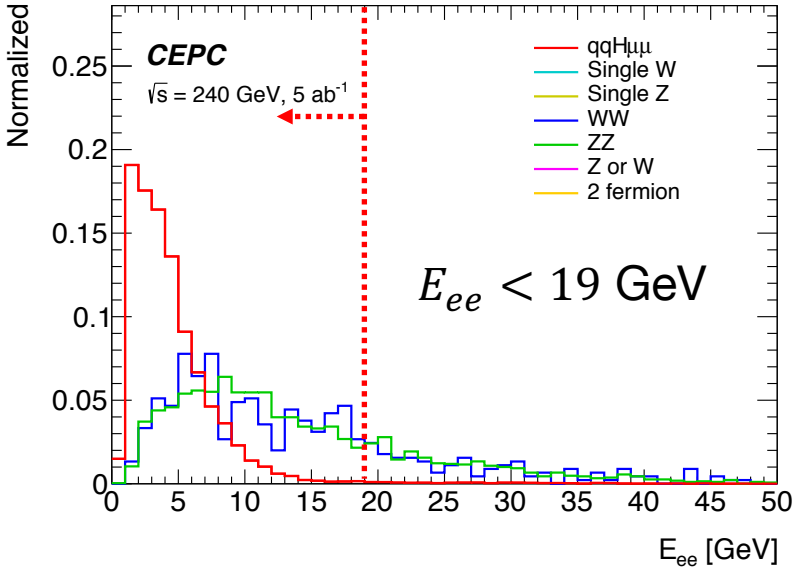


- N_e cuts: To suppress semi-leptonic backgrounds; And hadronic backgrounds with electrons in jet clusters
  - WW/ZZ to jets
  - WW/ZZ to muons/jets
  - Di-jet background

Selection criterial







- Electrons energy cuts: Also to suppress semi-leptonic backgrounds; And hadronic backgrounds with electrons in jet clusters
  - WW/ZZ to muons/jets
  - ZZ to muons/taus

Asimov fit results