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

Yanping Huang[1], Shan Jin[2], Gang Li[1], Qi Liu[3], Kunlin Ran[1], Manqi Ruan[1]

[1]IHEP, CAS

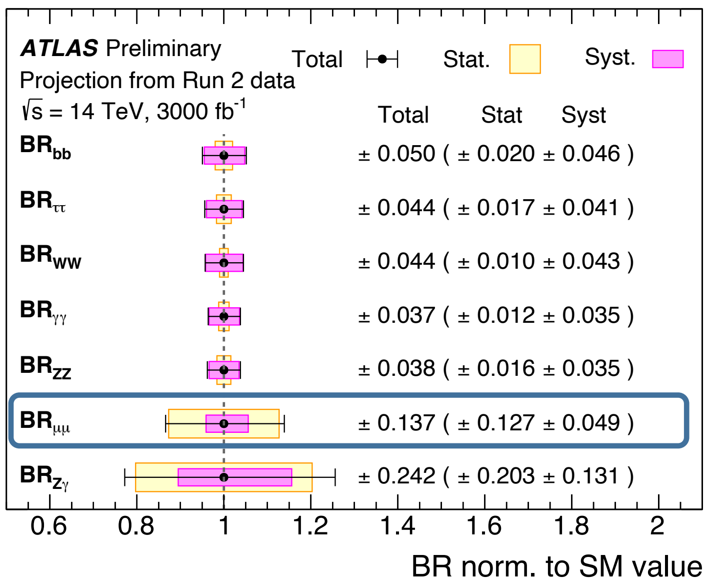
[2]Nanjing University

[3]Wuhan University

Introduction

- $H \rightarrow \mu\mu$ is important for probing the Higgs Yukawa couplings
  - The interactions of Higgs to the third generation charged fermions have been observed in LHC experiments [JHEP 08 (2016) 045]; While the Higgs couplings to other generation fermions haven't
  - $\boldsymbol{H \rightarrow \mu\mu}$ offers the best opportunity to measure Higgs Yukawa couplings to the second generation fermions

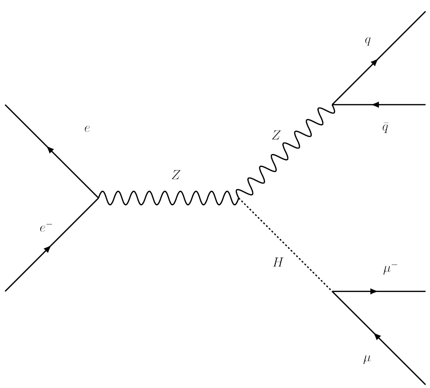
- **LHC result**
  - In the **ATLAS** experiment, the obs. (exp.) significance of the $H \rightarrow \mu\mu$ process is 2.0σ (1.7σ) with **139 fb^{-1} data** [Phys. Lett. B 812 (2021) 135980]
  - In the **CMS** experiment, the obs. (exp.) significance is 3.0σ (2.5σ) with **137 fb^{-1} data** [JHEP 01 (2021) 148]

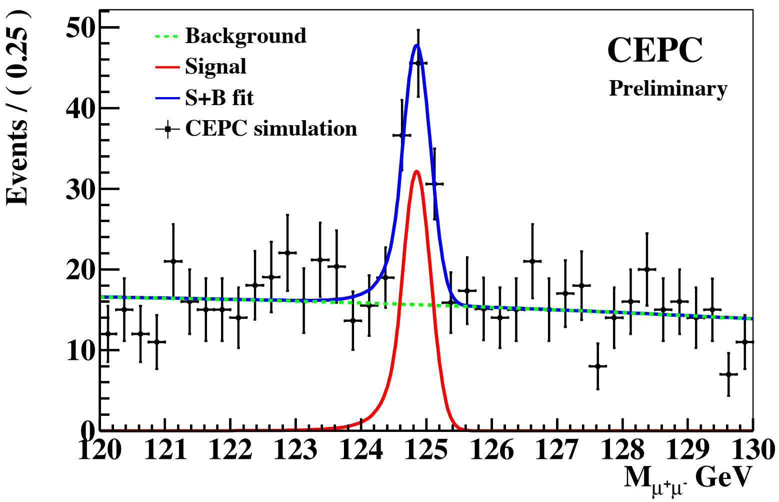


- In the projections with the **ATLAS** detector at **HL-LHC (3000 fb^{-1}),** the exp. precision of $B(H \rightarrow \mu\mu)$ is **14%** [ATL-PHYS-PUB-2018-054]

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

- With electron-positron colliders, we can gain much **higher significance** due to extremely clean background
- Focus on $\boldsymbol{ee \to Z(qq)H(\mu\mu)}$ channel in the CEPC experiment
- Previous publication [10.1088/1674-1137/42/5/053001] gave counted significance at [124.3, 125.2] GeV: 10.8σ, with the precision of ~17%





- While the measurement is not perfect enough
  - The simulation didn't consider the Z boson width
  - Only counting significance was presented

- Try to improve
  - Develop new selection criterial by keeping most signals and suppressing background
  - Use profile likelihood 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
  - PS. The analysis was done with the **obsolete int. lumi of 5 ab^{-1}**, while we'll show the statistical results with the lumi scaling to **5.6 ab^{-1} [**arXiv:1811.10545**]**
  - DST data: slimmed reconstruction samples without hits

  - Signal: $Z(\rightarrow qq)H, H \rightarrow \mu\mu$
    - $m_H = 125$ GeV
    - $\boldsymbol{\sigma(ee \rightarrow Z(\rightarrow qq)H)}$**:** 136.81 fb, $\boldsymbol{B(H \rightarrow \mu\mu)}$**:** 2.176E-04
    - Stat: ~100M

  - Background

Bkg.	Single W	Single Z	WW	ZZ	Z or W	2f
Stat.	~18 M	~8 M	~46 M	~6 M	~20 M	~28 M

- With an int. lumi. of 5.6 ab^{-1} **,** the **exp. bkg yield / statistics** are ~1, besides the 2 fermions (~30)

Event selections

- Taking into account both signal efficiency and signal-to-noise-ratio, event selections are finalized (detailed studies can be found in the backup)

Cut	Purpose
$N_\mu^+ > 0,\ N_\mu^- > 0$	$H \to \mu\mu$ requires 2 opposite charged muons
$115 < m_{\mu\mu} < 135 GeV$	$m_{\mu\mu}$ is closes to the m_H (125 GeV)
$25 < N_{particle} < 115$	di-jet system requires more objects than all lepton final stats
$55 < m_{qq} < 125 GeV$	m_{qq} is closes to the m_Z (91.2 GeV)
$p_{qq\mu\mu} < 32 GeV$, $195 < E_{qq\mu\mu} < 265 GeV$	The 4 momentum of the $qq\mu\mu$ system is closes to $(0, 0, 0, \sqrt{s})$, $\sqrt{s} = 240$ GeV
$35 < E_\mu^- < 100 GeV$, $35 < E_\mu^+ < 100 GeV$	To suppress WW bkg.
$18 < p_{\mu\mu} < 72 GeV$	To suppress hadronic bkg. with muons in jet clusters
$-20 < p_x^{miss},\ p_y^{miss} < 20$ GeV, $\Delta_{\mu^- qq,\mu^+},\ \Delta_{\mu^+ qq,\mu^-} > 2.5$	To suppress WW bkg. with neutrino decays

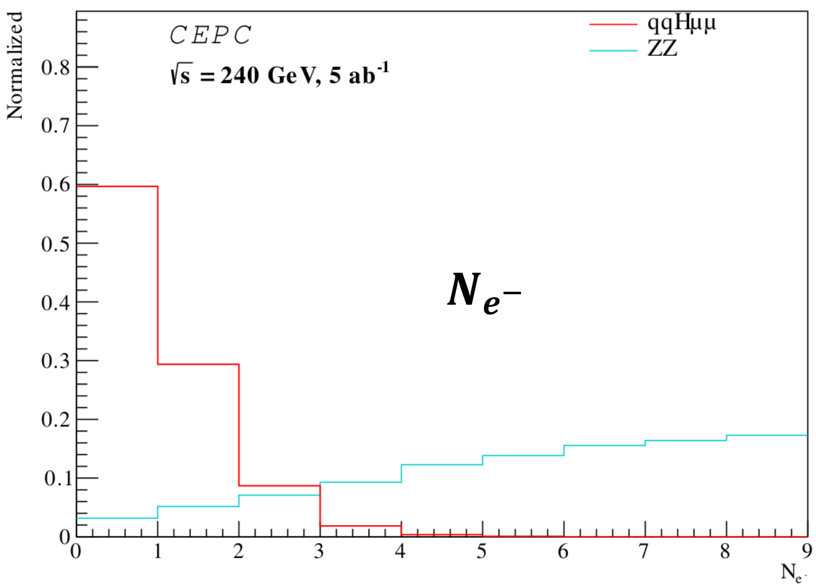
- **Δ** means the **solid angle** of the 2 objects (systems)

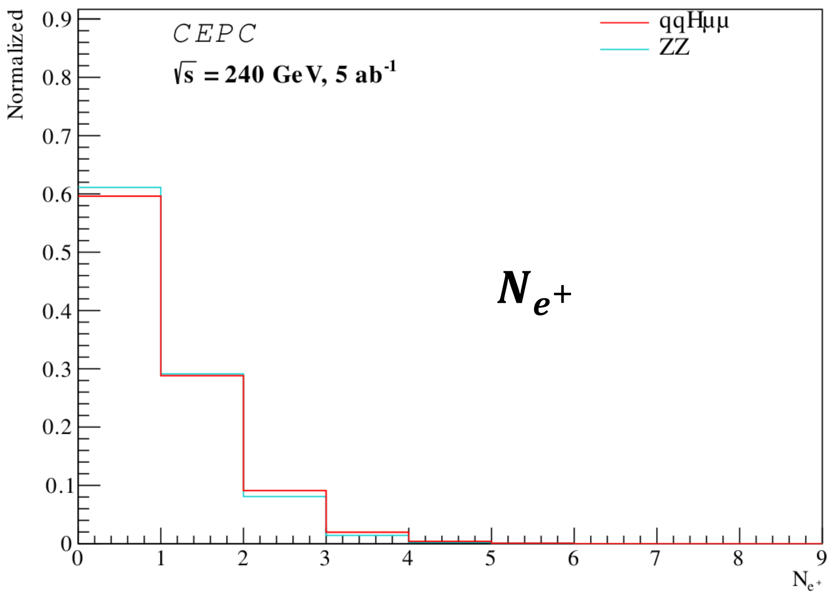
- We previously required **electron** selections, which could largely suppress the background (talk); But there're **mis-charge and truth match issues in electrons** observed in the **bkg. MC.** In the end we **removed** all electron cuts

N_e performances in the sig/bkg MC

- We used **reco-PID=11 or -11** (`getType()` from `ReconstructedParticle`) to identify electrons/positrons in the MC

- N_e **distributions** between **sig.** and ZZ bkg.
  - The distributions were performed **after the event selections** without electron cuts (talk)



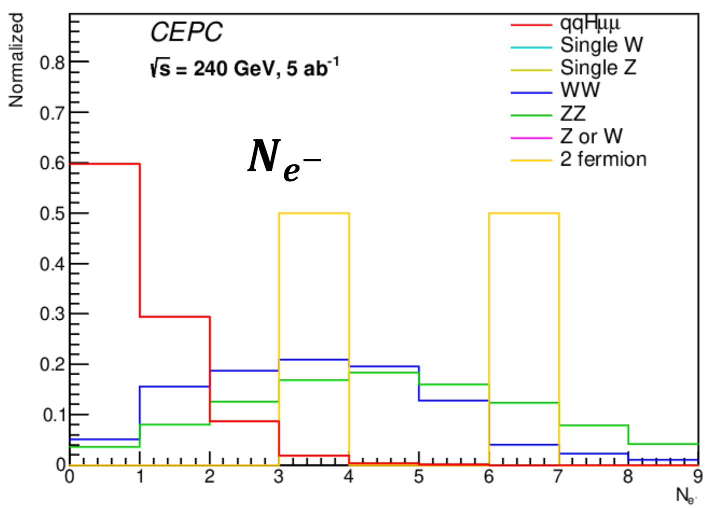


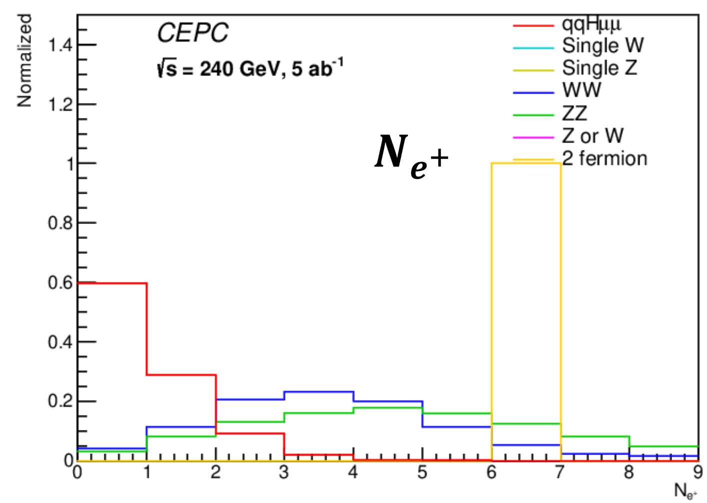
- In the phase space of the signal process, **electrons** can be **radiated from jets**

- It's **strange** that
  1. Why **sig** and **bkg** MC behave **different in N_{e^-}?** Why N_{e^-} and N_{e^+} are different in the **bkg MC**?
  2. Why there're so **many electrons in the bkg?**

Two steps to identify electrons/positrons

- To make validation, we tried another way to identify the **electrons/positrons**
    1. Select $(N_{e^-} + N_{e^+})$ by requiring **|reco-PID|=11**
    2. Separate the electrons/positrons by their **charge (**`getCharge()` from `ReconstructedParticle`**)**
- The **sum number** of electrons, positrons in the **2 steps method** is equal to the **one step method (previous slide)**
- Perform N_e **distributions**





- The reason N_{e^-} and N_{e^+} behave different in the **bkg MC** is due to the **charge mis-identification** in the PID

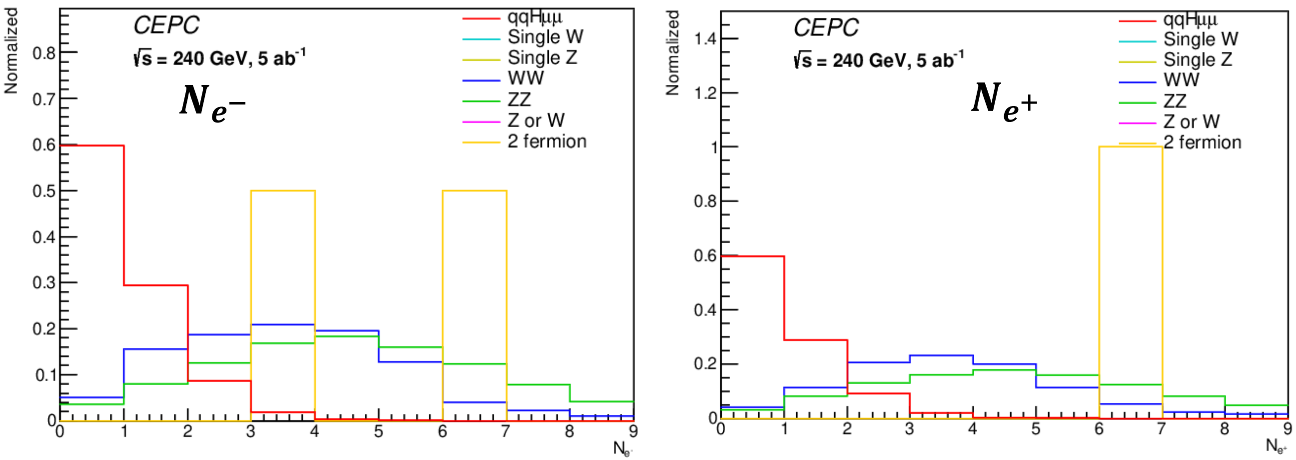
- Still 2 questions need to be understood
    1. Why **sig** and **bkg** MC behave **different?**
    2. Why there're so **many electrons in the bkg?**

Truth match efficiency

- We decided to look after the **truth information** of electrons
- In the `RecoMCTruthLink` collection, reco-particle can be linked to a truth-particle, estimate the **truth-match efficiency** in the sig/bkg MC

numbers	qqh_e2e2			zz			ww		
	reco	linked	eff	reco	linked	eff	reco	linked	eff
N_{e^-}	43106	31691	73.5%	39232	1599	4.07%	2425	71	4.07%
N_{e^+}	43801	30962	70.7%	39699	3207	8.07%	2488	228	9.164%

- The truth match efficiencies of bkg MC are **small**, thus many reco-electrons in the bkg are **fake particles**
- $\Longrightarrow$ That's why **the number of reco-electrons** in the **bkg MC** are greater than **sig**



- Since there're potential issues in the electron ID and truth match algorithms, we decided to **remove all electron cuts** in the analysis

Cut flow

- After event selections, all other bkg. are excluded, except for **semi-lep decay of ZZ/WW** and 2 fermions

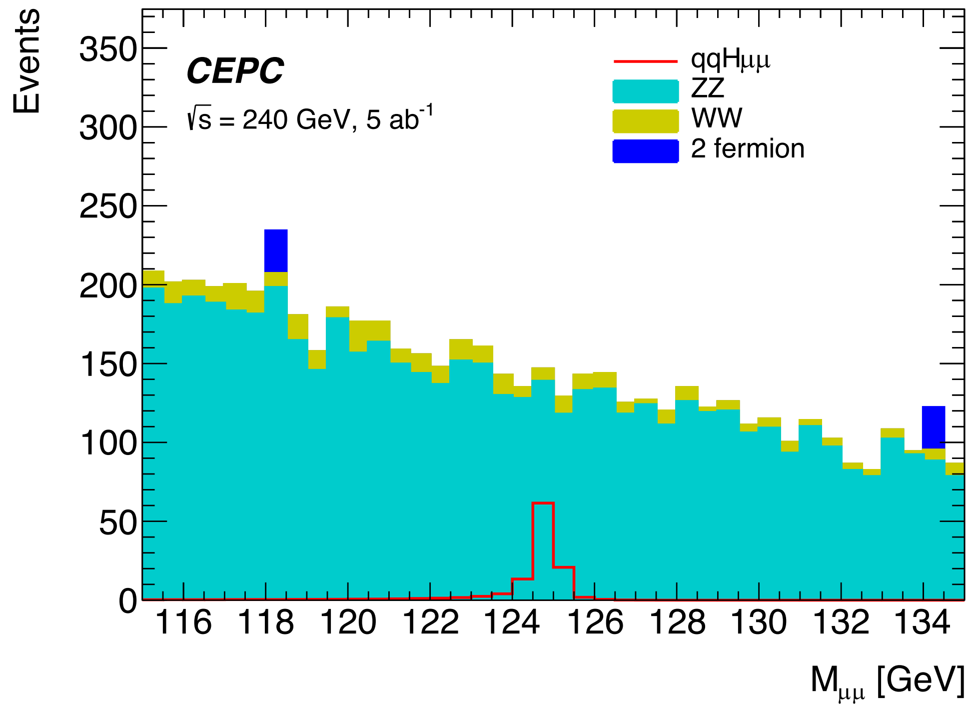
Cut	$H \to \mu\mu$		ZZ	WW	$2f$
	Event	Eff.	Event	Event	Event
Initial	148.849	1	5.70e+06	4.53e+07	7.15e+08
$N_\mu^+ > 0,\ N_\mu^- > 0$	147.917	0.993	665833	948669	3.52e+06
$115 < m_{\mu\mu} < 135 GeV$	118.544	0.796	11062.9	2781.69	48700
$25 < N_{particle} < 115$	117.927	0.792	6175.83	2236.64	1328.93
$55 < m_{qq} < 125 GeV$	117.15	0.787	5616.92	1264.36	216.449
$p_{qq\mu\mu} < 32 GeV$, $195 < E_{qq\mu\mu} < 265 GeV$	116.522	0.783	5557.46	948.514	216.449
$35 < E_\mu^- < 100 GeV$, $35 < E_\mu^+ < 100 GeV$	116.263	0.781	5518.81	658.416	216.449
$18 < p_{\mu\mu} < 72 GeV$	115.5	0.776	5428.63	654.455	54.1121
$-20 < p_x^{miss},\ p_y^{miss} < 20$ GeV, $\Delta_{\mu^- qq,\mu^+},\ \Delta_{\mu^+ qq,\mu^-} > 2.5$	114.952	0.772	5402.87	359.406	54.1121

- Signal efficiency: 77%

Bkg.	**ZZ**	WW	2f
Fraction	**93%**	6%	1%

$m_{\mu\mu}$ distribution after the event selections

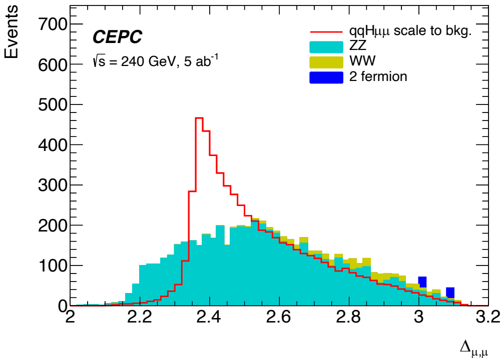
- After the event selections, perform the $\boldsymbol{m_{\mu\mu}}$ distribution

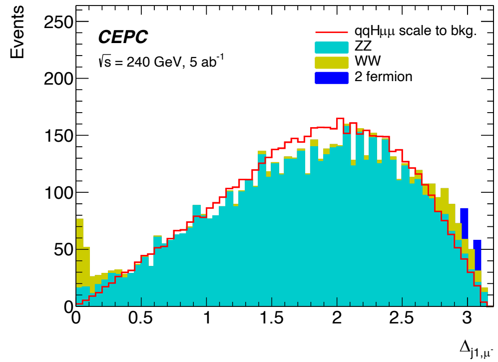


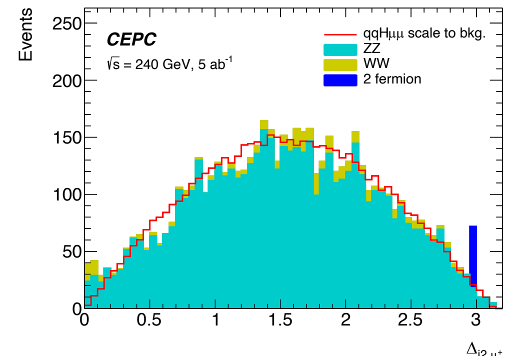
- $\boldsymbol{ZZ}$ background is **dominant**
- Estimate the **counting significance**: $Z = \sqrt{2\left((s+b)\ln\left(1+\frac{s}{b}\right)-s\right)} = \mathbf{4.9}\boldsymbol{\sigma}$, in the $\boldsymbol{m_{\mu\mu}}$ region [124.1, 125.5] GeV (**3$\boldsymbol{\sigma}$ width** of the sig)
  - The resolution $\boldsymbol{\sigma}$ is obtained by fitting the sig MC with DSCB function, detailed in later slides

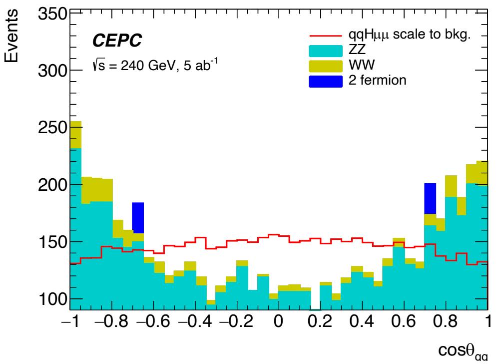
MVA optimizations

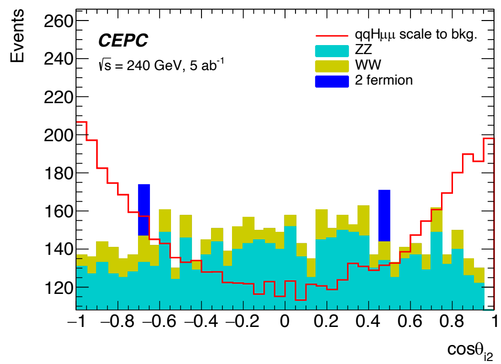
- Next to apply discriminant variables into MVA (BDTG) to optimize the signal
- Input discriminant variables: Δ_{q1,μ^-}, Δ_{q2,μ^+}, $\Delta_{\mu,\mu}$, $\cos\theta_{qq}$, $\cos\theta_{q2}$, m_{qq}, p_x^{q1}, p_y^{q2}

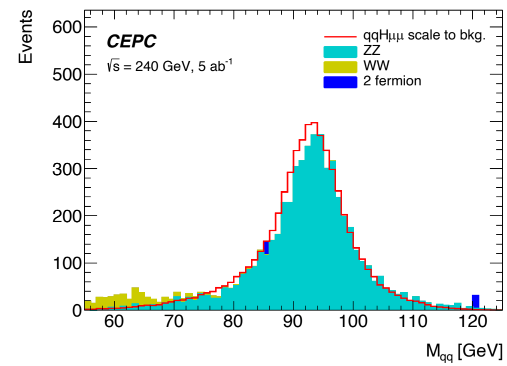


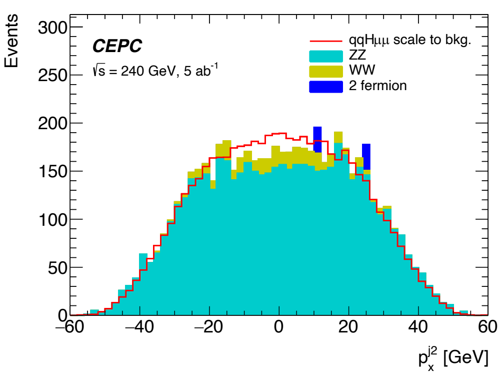


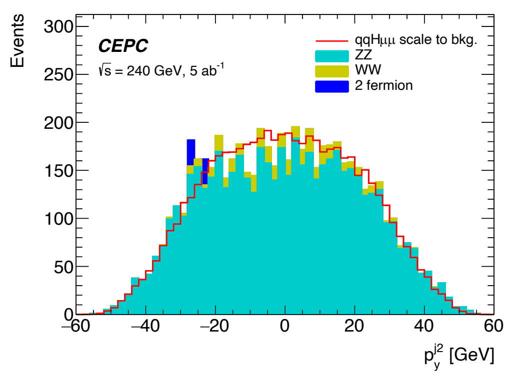










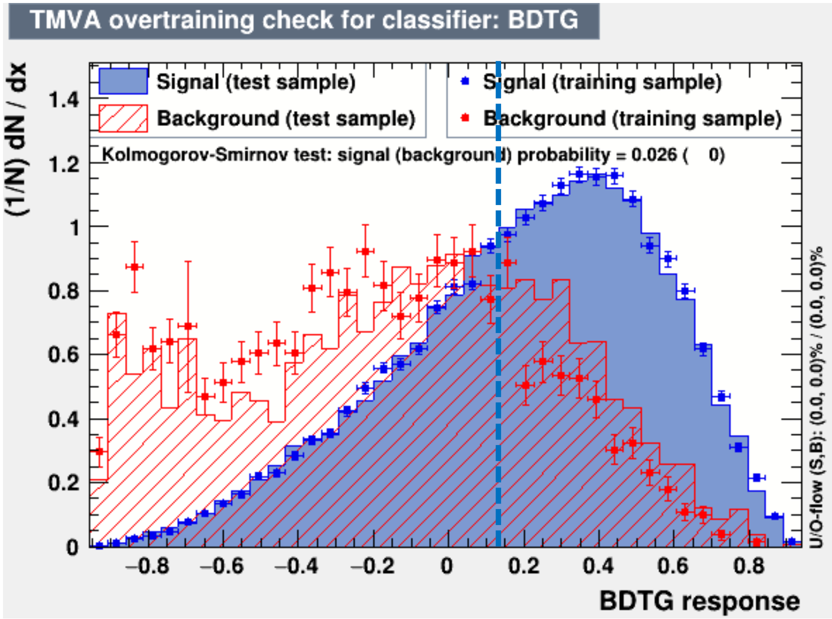


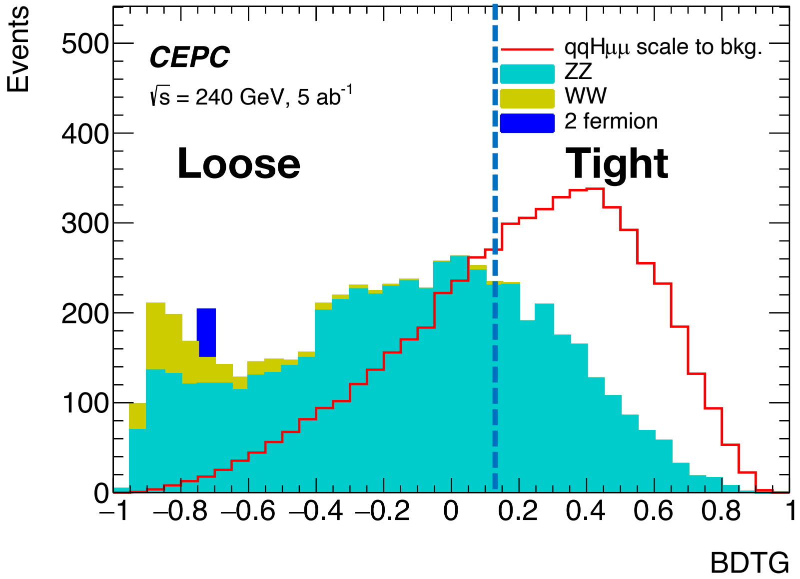
- The input variables are **optimized:**
- To **not highly correlated** with $m_{\mu\mu}$
- To have **separation power** between sig/bkg

- **Top rank** variable in the BDTG training: $\cos\theta_{q2}$

BDTG

- After training, perform the **BDTG response**



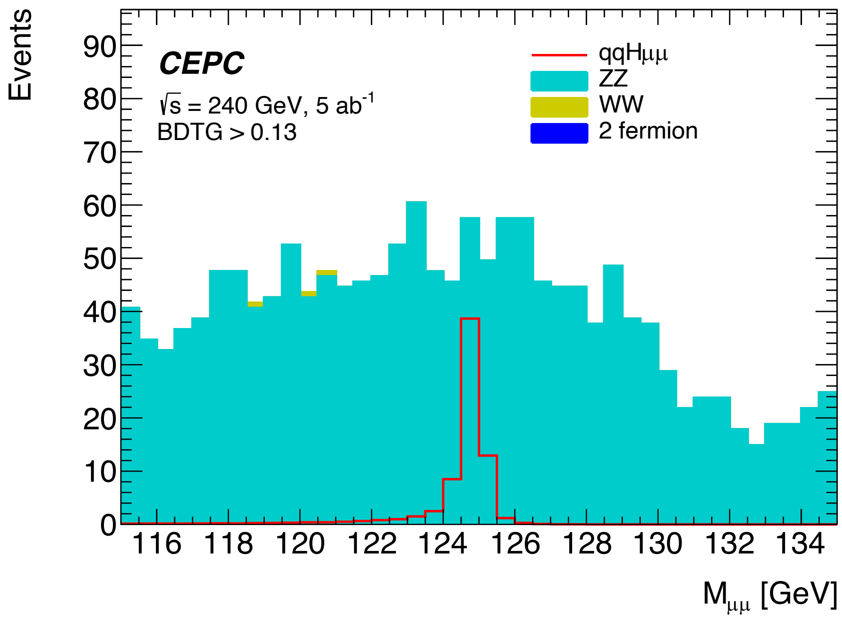


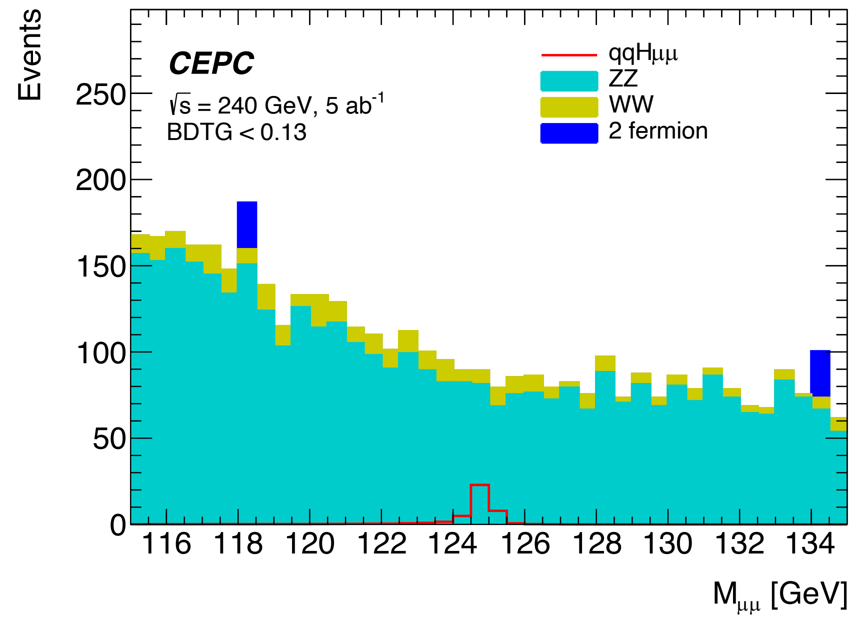
- There is **overtraining** in the bkg. due to **poor statistics**
- The BDTG algorithm can't make deeper optimization with the same reason

- Scan the **significance VS BDTG** to find the optimal cut point to split events in 2 categories: **tight/loose**, to obtain the **maximum combined significance**: $Z = \sqrt{Z_{tight}^2 + Z_{loose}^2}$; Optimal cut point: 0.13
- In the **tight** category, the **bkg** components are almost $\boldsymbol{ZZ}$, the WW and 2f are excluded

$m_{\mu\mu}$ in tight/loose categories

- Perform $\boldsymbol{m_{\gamma\gamma}}$ between sig/bkg MC in 2 categories





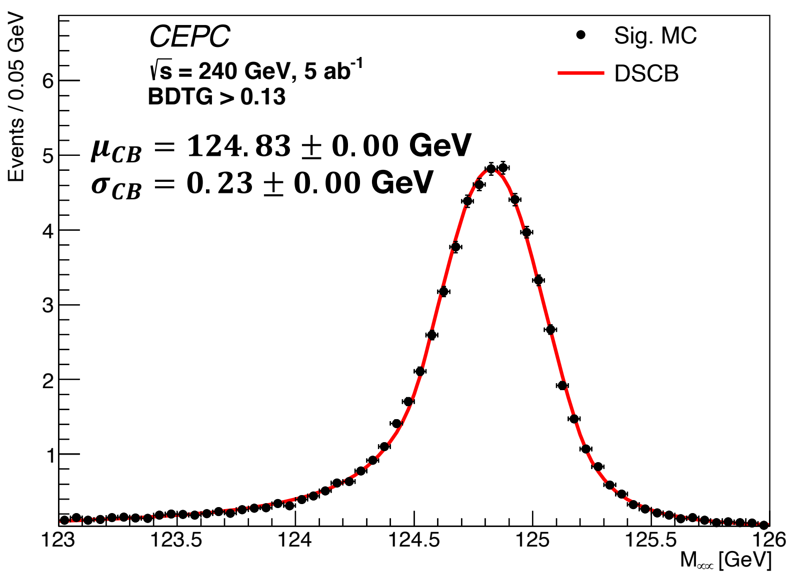
- Summarize the **event yield and counting significance** in each category

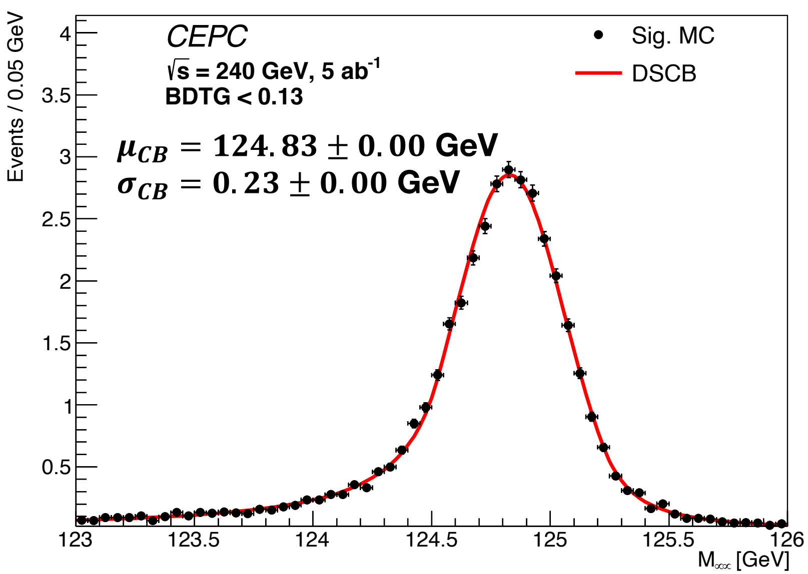
Category	$H \to \mu\mu$	ZZ	WW	2 f	Significance
Tight	80.09	1770.72	3.36	0	5.0σ
Loose	48.66	4280.64	398.72	60.48	2.3σ
Combined	128.74	6051.36	402.08	60.48	5.5σ

- The **combined significance $\boldsymbol{Z = 5.5\sigma}$**, which is 11% better than inclusive case

Signal modelling

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

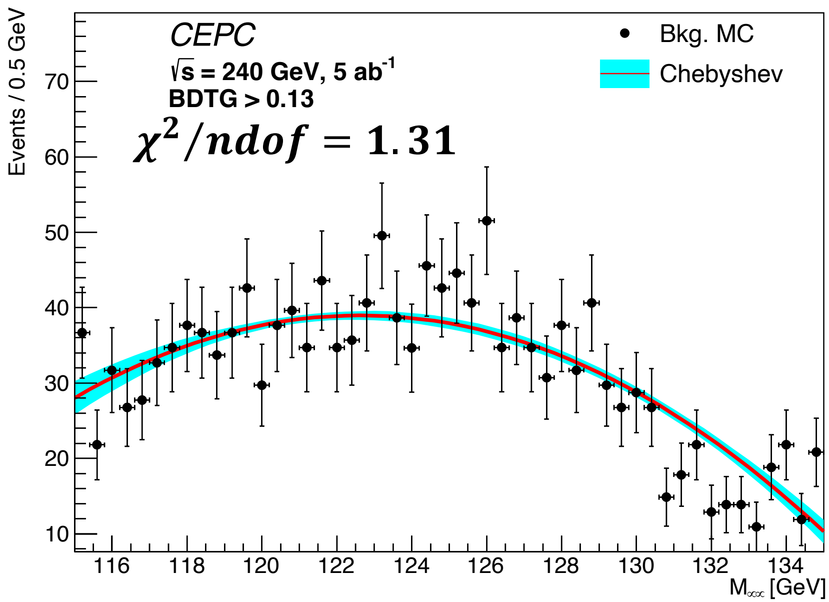


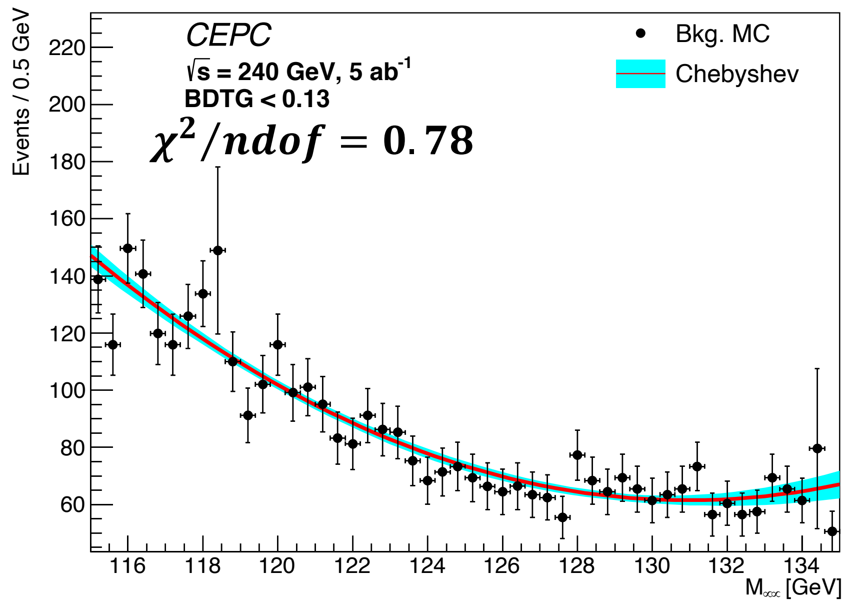


- DSCB shows great agreement with the sig. MC
- In fitting the final **pseudo-data (sig. + bkg. MC),** the parameters of the sig model are **fixed from fitting the sig. MC alone**

Background modelling

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





- Tested **different functions** for fitting the bkg: $\exp(-am_{\mu\mu})$, polynomials, etc
- Finally select the second order Chebyshev polynomial with the **best** χ^2

- In fitting the final **pseudo-data (sig. + bkg. MC),** the parameters of the bkg model are **floated**

Statistical method

- Make the **pseudo-data** with **sig.+bkg. MC**

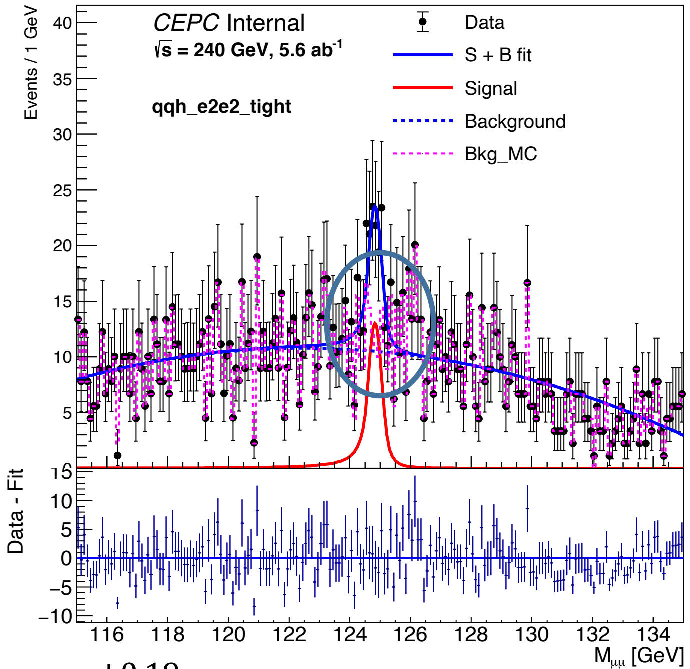
- Apply the unbinnd **likelihood** fit on the $m_{\mu\mu}$ (observable), fit on all categories simultaneously
- $\mathcal{L} = \prod_c \left(\text{Pois}(n_c | \mu S_c + B_c) \prod_1^{n_c} \frac{\mu S_c f_{S,c} + B_c f_{B,c}}{\mu S_c + B_c} \right)$
- $\boldsymbol{c}$**:** event category; $\boldsymbol{n_c}$**:** event number in category c
- $\boldsymbol{S_c}$**:** expected **signal** number in category c; $\boldsymbol{B_c}$**:** expected **background** number in category c

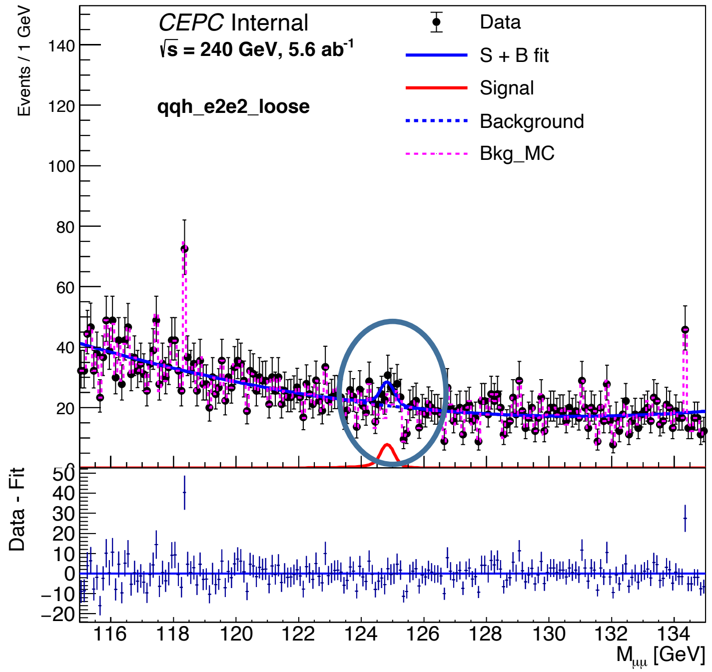
- **POI:** signal strength $\mu = \frac{\sigma(Z(qq)H) \cdot B(H \to \mu\mu)}{\left(\sigma(Z(qq)H) \cdot B(H \to \mu\mu)\right)_{SM}}$
- **Signal model $\boldsymbol{f_S}$:** DSCB; **Bkg model $\boldsymbol{f_B}$:** Second order Chebyshev polynomials

- Also fit on the **Asimov** dataset to avoid statistical fluctuations
- Only **statistical** uncertainties are considered

Fit on the pseudo-data

- Simutaneous fit on $m_{\mu\mu}$ in 2 categories



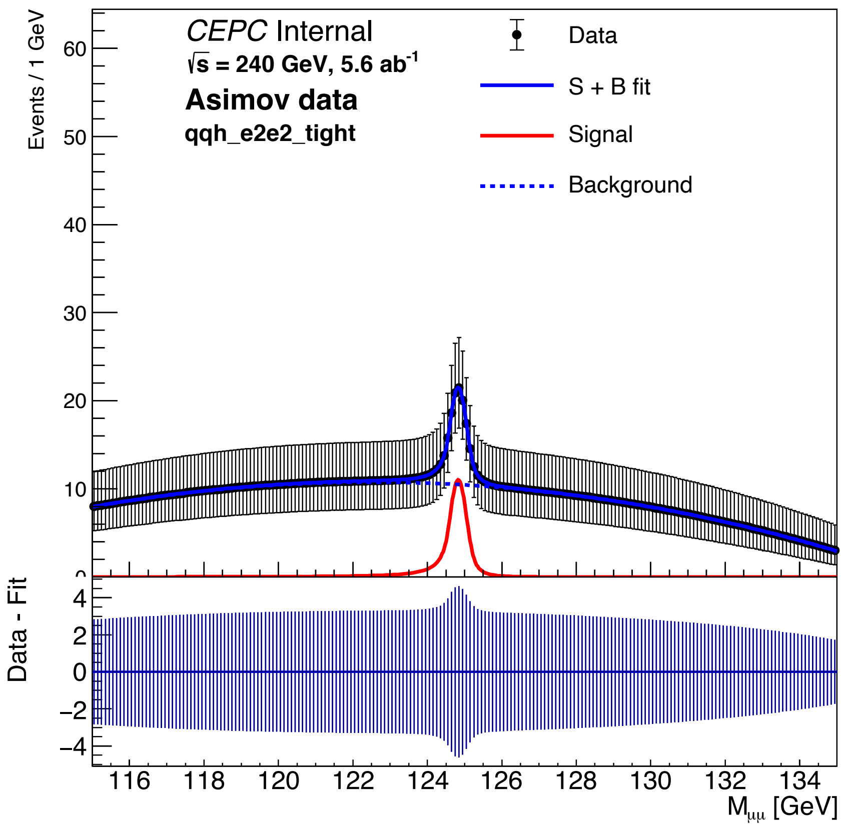


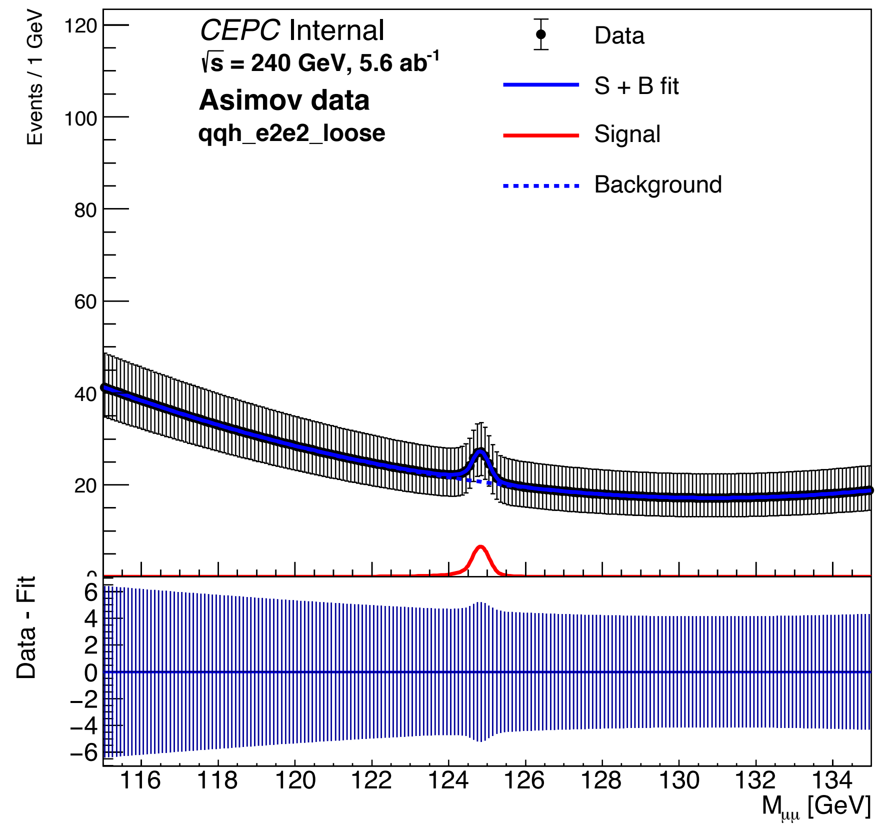
- $\mu = 1.18^{+0.19}_{-0.18}$, **significance:** 7.4σ

- The **fitted bkg. component (dashed blue curve)** is lower than the **bkg. MC (dashed pink curve)** in the peak region, which is due to
  - Large **statistical fluctuations** in the bkg. MC
  - **Imperfect bkg. function** to model the MC bkg. shape (can further study the bkg. model in the future if possible)
- Thus the fitted **sig.** is **over-estimated,** perform the **Asimov** fit to avoid fluctuations and give the **nominal results**

Asimov results

- Fit on the **Asimov** dataset

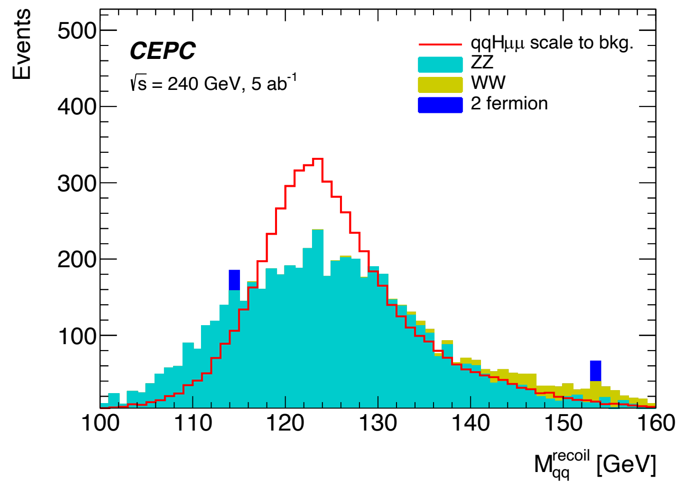


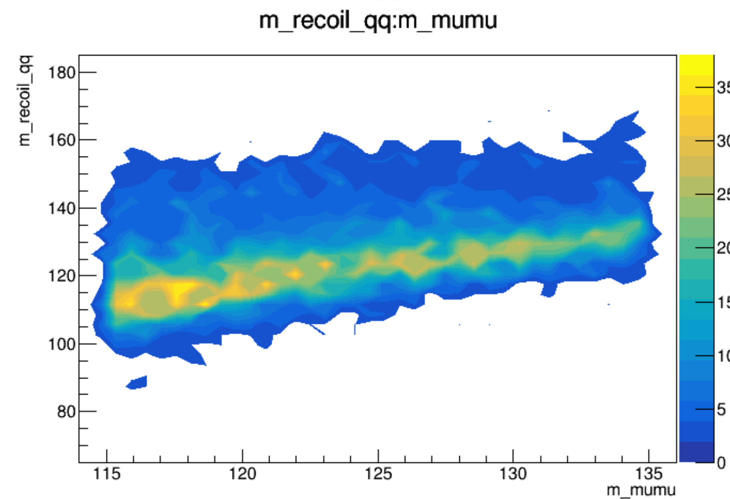


- $\mu = 1.00 \pm \mathbf{0.18}$, **significance:** 6.4σ

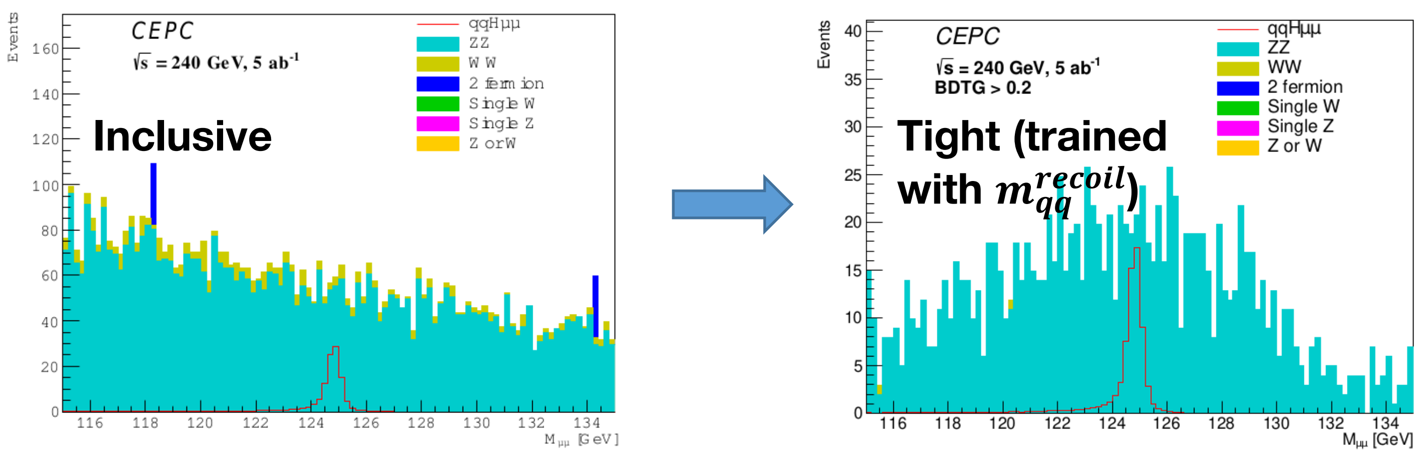
Results by adding m_{qq}^{recoil} in categorization

- In the analysis, we didn't use the m_{qq}^{recoil} for event selections and category optimizations
- After the nominal selections, hard to determine the **cut point on m_{qq}^{recoil}** to separate **sig./bkg.** events



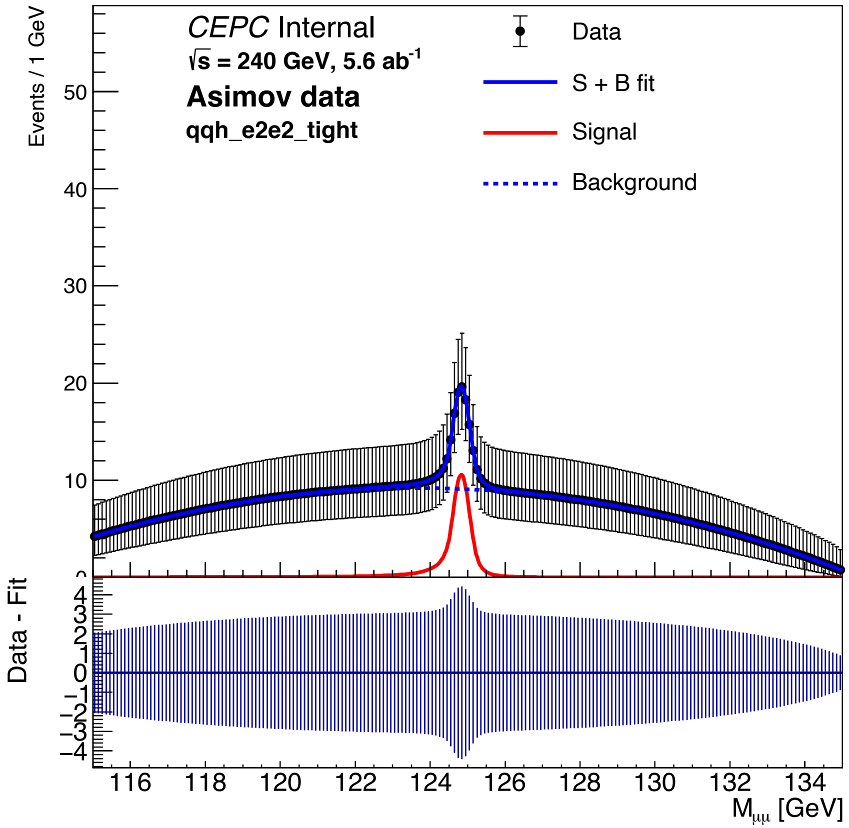


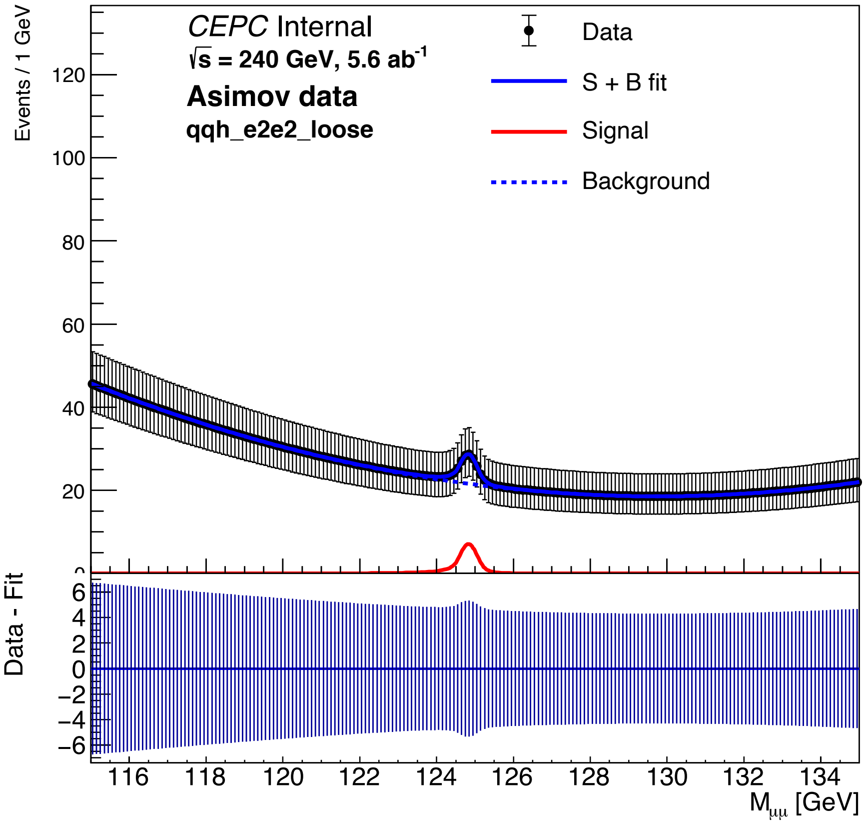
- In addition, there is **large linear correlation** between $m_{\mu\mu}$ and m_{qq}^{recoil} in the **bkg. (42%);** If we put the variable in the MVA training, we'll largely change the mass shape



Exp. results by adding m_{qq}^{recoil}

- Re-define event categories (tight/loose) with the BDTG trained with an additional $\boldsymbol{m_{qq}^{recoil}}$
- Estimate the **exp. signal strength** by fitting on the Asimov dataset





- $\mu = 1.00^{+0.18}_{-0.17}$, **significance: 6.5$\sigma$**
- Compared with the **nominal case (6.4σ),** the improvement isn't sizable

Momentum smearing study

- To test the **performance** of the **CEPC detector**, we **smear the momentum resolution** of the muon ($\sigma = (p_\mu^{reco} - p_\mu^{truth})/p_\mu^{truth}$) by 25%, 50% and 100%
  - $\sigma_{smear} = \sigma_{nom}(1+\delta)$

- Apply event selections, categorizations and sig./bkg. modelling based on the **smeared variables,** re-estimate **precision** and **significance** of the signal

Smearing	0 (nominal)	25%	50%	100%
μ	1.00 ± 0.18	$1.00^{+0.20}_{-0.19}$	$1.00^{+0.21}_{-0.20}$	$1.00^{+0.23}_{-0.22}$
Significance	6.4σ	5.8σ	5.3σ	4.8σ
Reduction in significance		-9%	-16%	-24%

- If the performance of the detector for calculating the resolution of the muon is 100% worse than the designed parameters, we'll have 24% reduction in the signal significance

Summary

- Studied $ee \rightarrow Z(qq)H(\mu\mu)$ in the CEPC experiment
  - Develop the **selection criteria** to select the $qqH\mu\mu$ sig.
  - After event selections, the **sig. efficiency is 77%;** the **dominant background** is **ZZ to muons/jets**

  - Apply the **BDTG method for categorization** to further improve the significance
  - Divide the events in 2 categories, the counting significance **improved by 11%** than the inclusive case

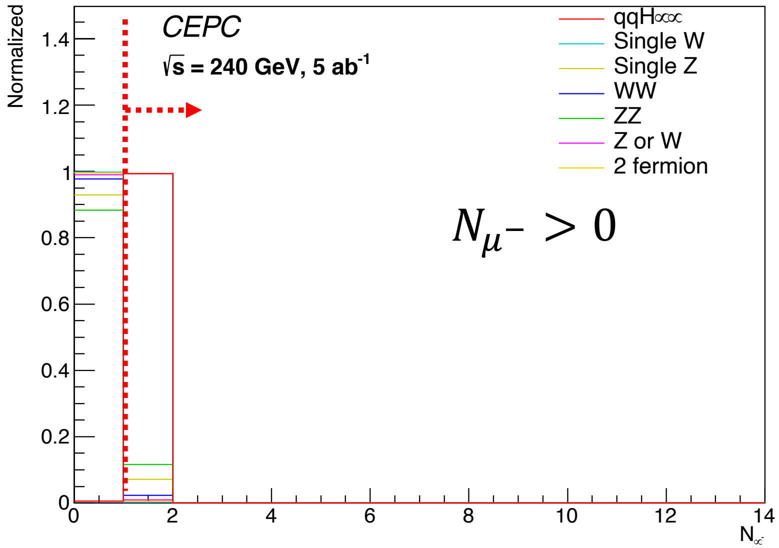
  - Choose **DSCB** as the **sig.** model and **second order Chebyshev polynomial** as the **bkg.** model for final statistical analysis
  - By simultaneously fitting on the $\boldsymbol{m_{\mu\mu}}$ in 2 categories with the Asimov data, the **exp. precision is 18%,** with the **significance of 6.4$\boldsymbol{\sigma}$**

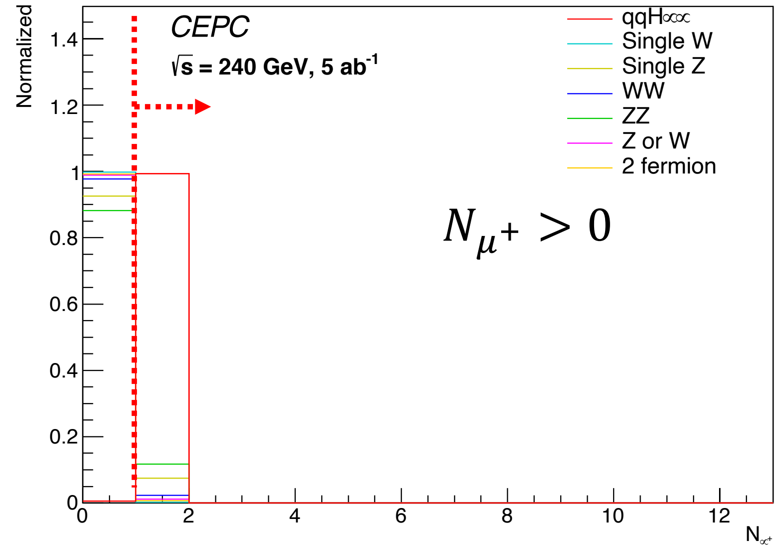
  - In **bkg.** $\boldsymbol{m_{qq}^{recoil}}$ is largely **correlated** with $\boldsymbol{m_{\mu\mu}}$ after selections (42%); The **mass shape** will be largely changed if adding the variable in the BDTG training
  - Besides, the improvement in significance **isn't sizable** (6.4σ → 6.5σ)

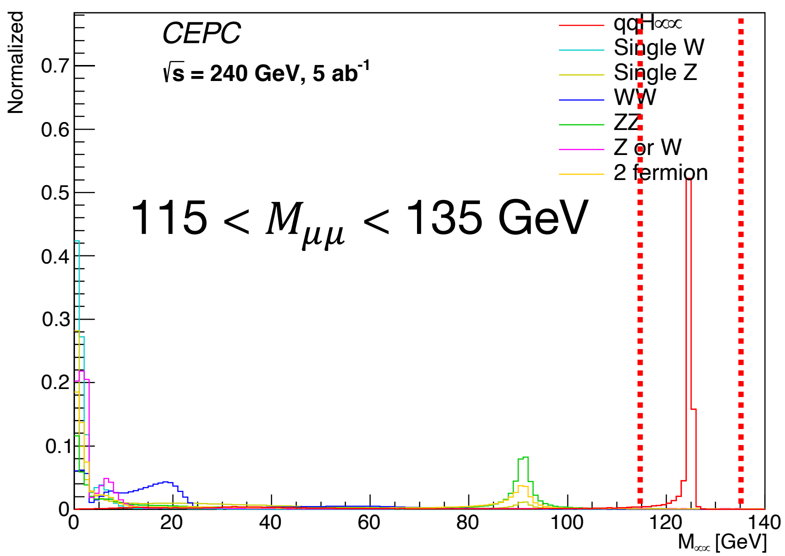
  - To test the performance of the CEPC detector, **the momentum resolution** of the **muon** is smeared by 25%, 50% and 100%; In the worst case, the significance reduces by 24%

Backup

Selection criterial

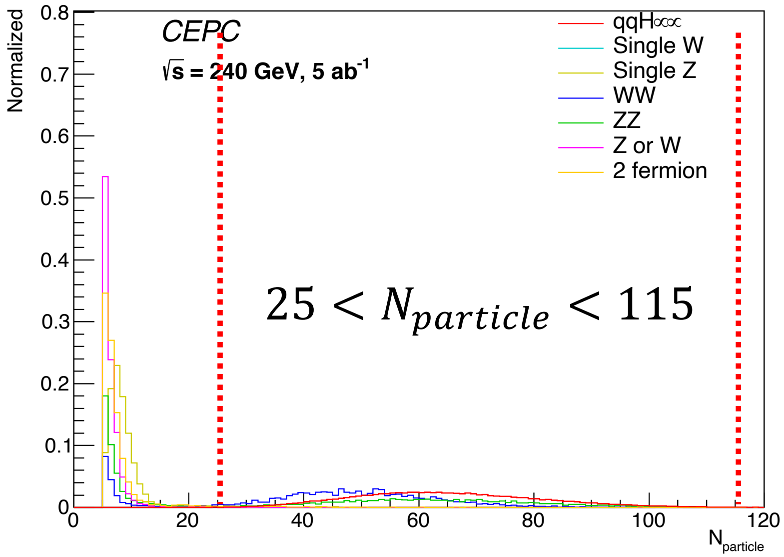




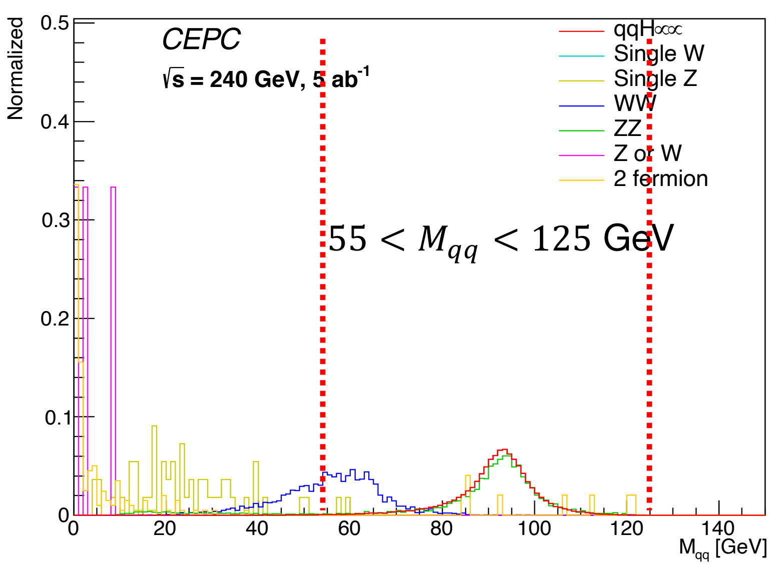


- N_{μ} cuts: $H \rightarrow \mu\mu$ requires 2 opposite charged muons
- $M_{\mu\mu}$ cut: $M_{\mu\mu}$ should close to M_H (125 GeV)

Selection criterial



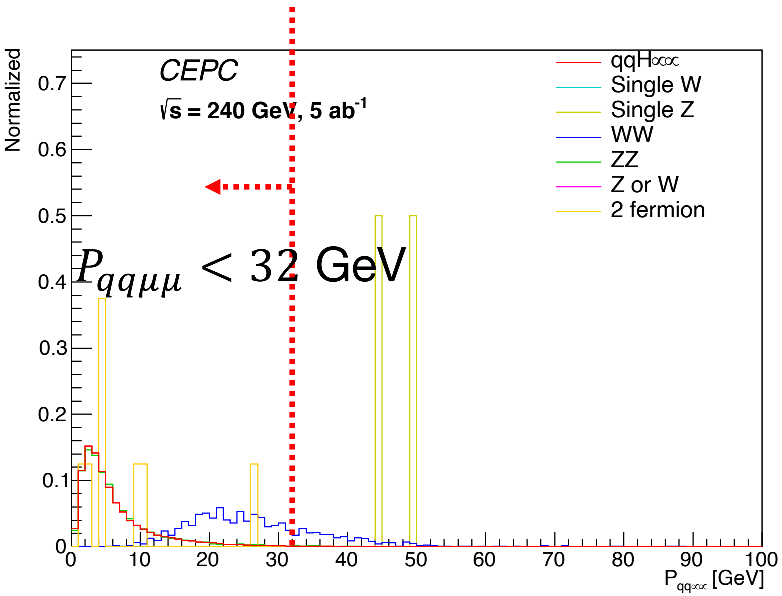
- All reconstructed visible objects are grouped as "particle"
- $N_{particle}$ cut: di-jet system requires more objects
- Suppress background components with lepton final stats
  - Single Z to muons/electrons
  - ZZ/WW to muons/taus
  - 2 fermions (muons/taus)

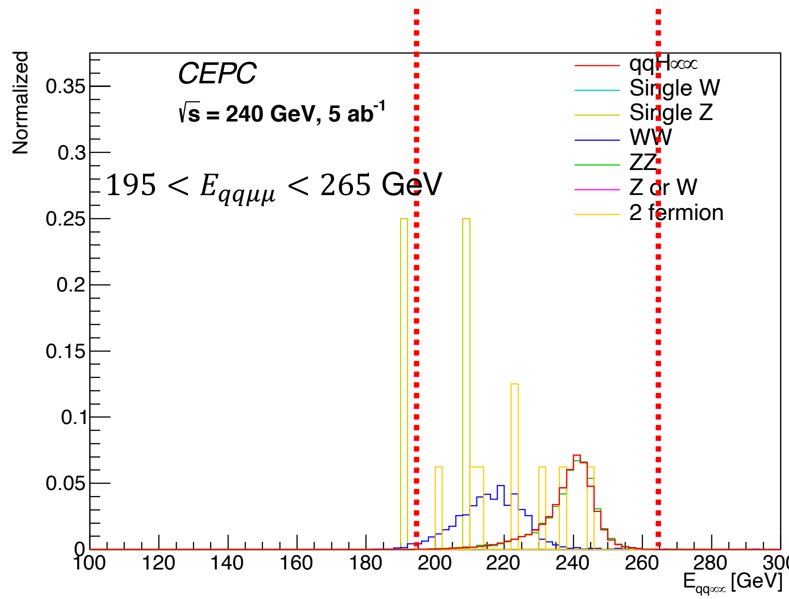


- 4 momentum of all visible particles: $p_{particle}$; 4 momentum of di-muon: $p_{\mu\mu}$
- Didn't apply jet-cluster algorithm, while define di-jet system: $p_{qq} = p_{particle} - p_{\mu\mu}$
- Since the dominant background would be $Z(\mu\mu)Z(qq)$ (see later), the rough definition of di-jet system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

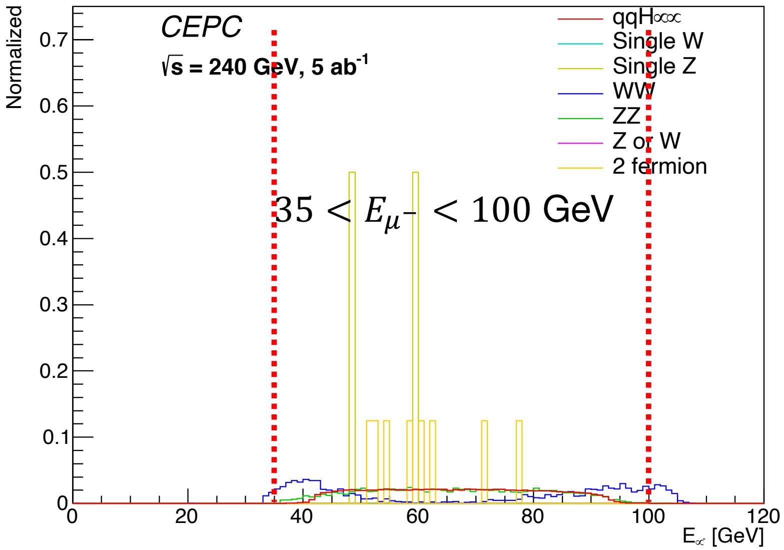
Selection criterial

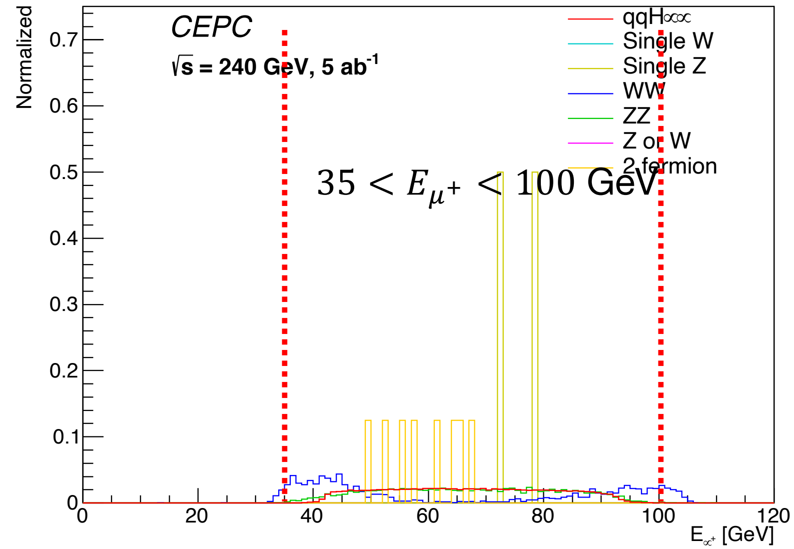
- $p_{qq\mu\mu}$ cut: $ee \to Z(qq)H(\mu\mu)$ system should has 4 momentum close to (0, 0, 0, $\sqrt{s}$), $\sqrt{s} = 240$ GeV

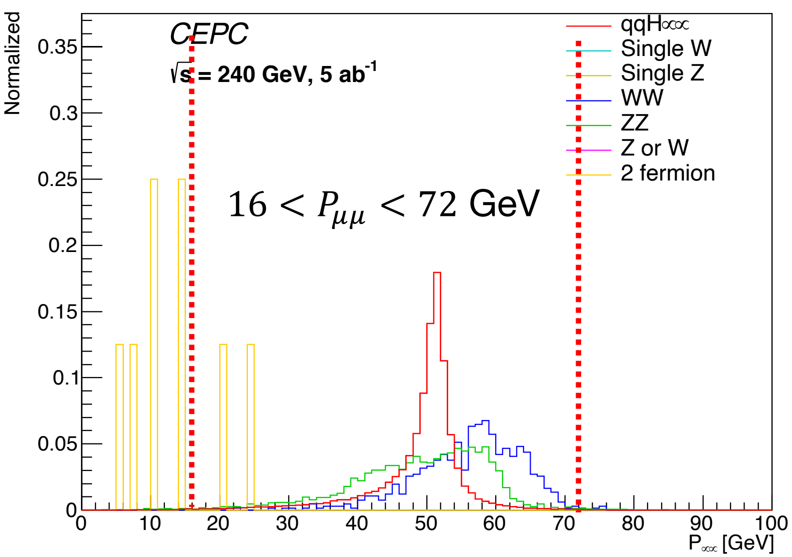




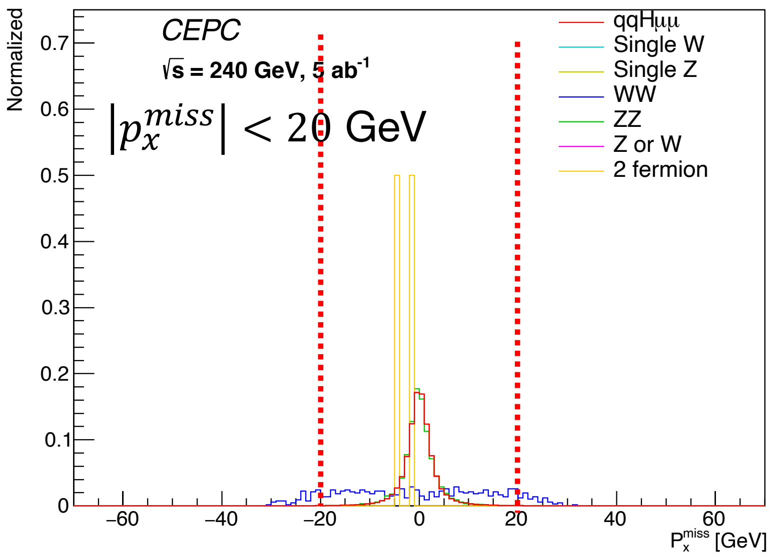
Selection criterial

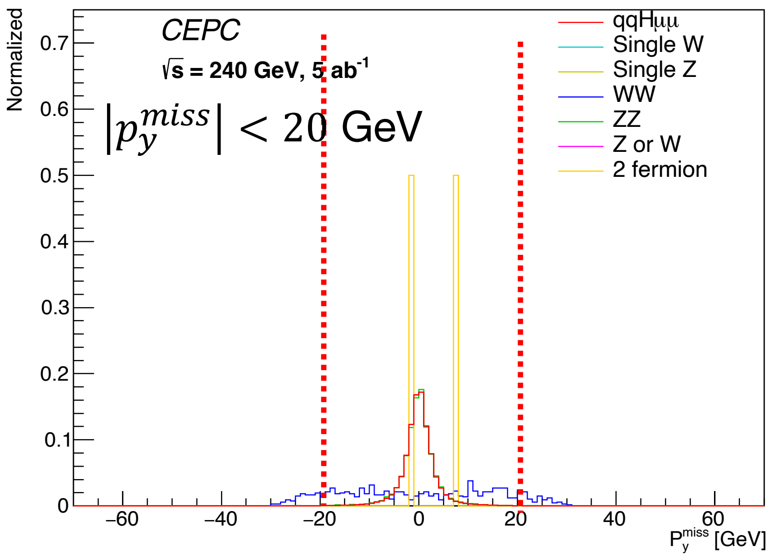


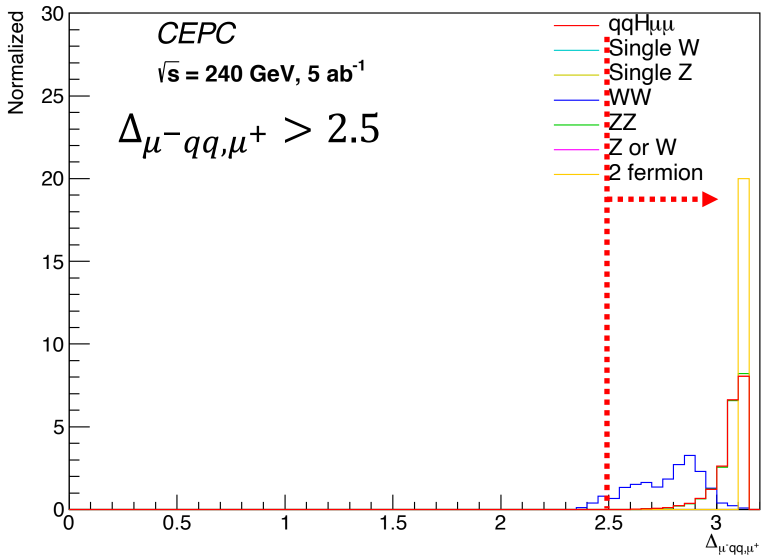


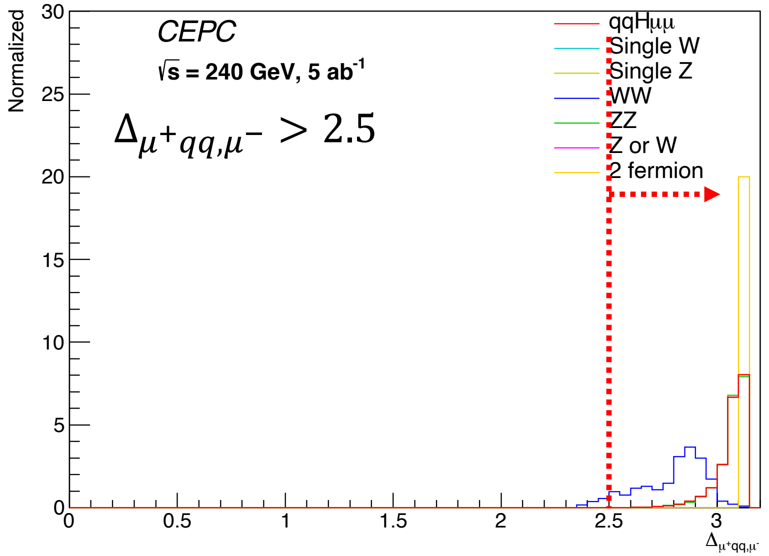


- E_{μ} cuts: To suppress WW background
  - WW to muon/jets
- $p_{\mu\mu}$ cut: To suppress hadronic background components with muons in jet clusters
  - Di-jet background
  - ZZ to taus/jets


CEPC
$\sqrt{s}$ = 240 GeV, 5 ab^{-1}
$|p_x^{miss}| < 20$ GeV
qqHμμ
Single W
Single Z
WW
ZZ
Z or W
2 fermion
Normalized
P_x^{miss} [GeV]

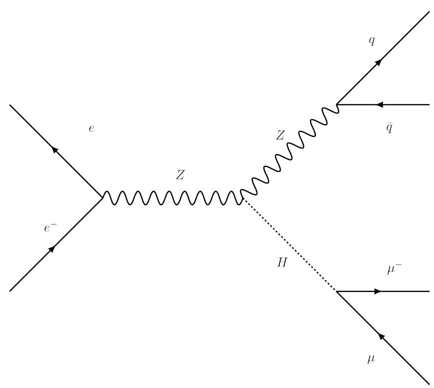

CEPC
$\sqrt{s}$ = 240 GeV, 5 ab^{-1}
$|p_y^{miss}| < 20$ GeV
qqHμμ
Single W
Single Z
WW
ZZ
Z or W
2 fermion
Normalized
P_y^{miss} [GeV]

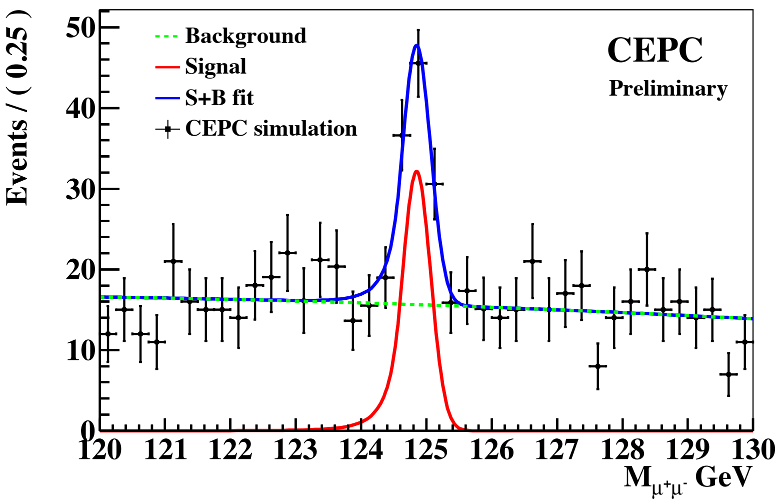

CEPC
$\sqrt{s}$ = 240 GeV, 5 ab^{-1}
$\Delta_{\mu^- qq,\mu^+} > 2.5$
qqHμμ
Single W
Single Z
WW
ZZ
Z or W
2 fermion
Normalized
$\Delta_{\mu^- qq,\mu^+}$


CEPC
$\sqrt{s}$ = 240 GeV, 5 ab^{-1}
$\Delta_{\mu^+ qq,\mu^-} > 2.5$
qqHμμ
Single W
Single Z
WW
ZZ
Z or W
2 fermion
Normalized
$\Delta_{\mu^+ qq,\mu^-}$

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

- With electron-positron colliders, we can gain much **higher significance** due to extremely clean background
- Focus on $\boldsymbol{ee \to Z(qq)H(\mu\mu)}$ channel in the CEPC experiment
- Previous publication [10.1088/1674-1137/42/5/053001] gave counted significance at [124.3, 125.2] GeV: 10.8σ, with the precision of ~17%





- While the measurement is not perfect enough
  - The simulation didn't consider the Z boson width
  - A typo observed in the formula for calculating significance
    - $\sqrt{2(s+b)\ln\left(1+\frac{s}{b}\right)-s} \Longrightarrow \sqrt{2\left((s+b)\ln\left(1+\frac{s}{b}\right)-s\right)}$
- Try to improve
  - Develop new selection criterial by keeping most signals and suppressing background
  - Use profile likelihood method to estimate significance
  - Further make event categories by applying MVA method