Higgs to ZZ measurement at the CEPC

Ryuta Kiuchi [a] , Yanxi Gu [b] , Min Zhong [b] ,
Shih-Chieh Hsu [c] , Xin Shi [a] , Kaili Zhang [a]

[a] IHEP , [b] USTC , [c] UW

April 14th, 2021

Contents

- Status report on HZZ analysis
  - Introduction
  - Brief description of the analysis flow
  - Obtained results
- Future plans
- Summary

Introduction to HZZ channel

・ Since the state has 3 Z bosons, there are multiple combinations of final products.

・ Final states having ($\mu^+\mu^-$, jj, $\nu\nu$) are promising channels, owing to its clear signature. On the other hand, its low statistics could limit the final precision.

・This presentation summarizes the updated results from channels with the decay product combination of ($\mu\mu$, jj, $\nu\nu$).

Table : Promising decay product combinations

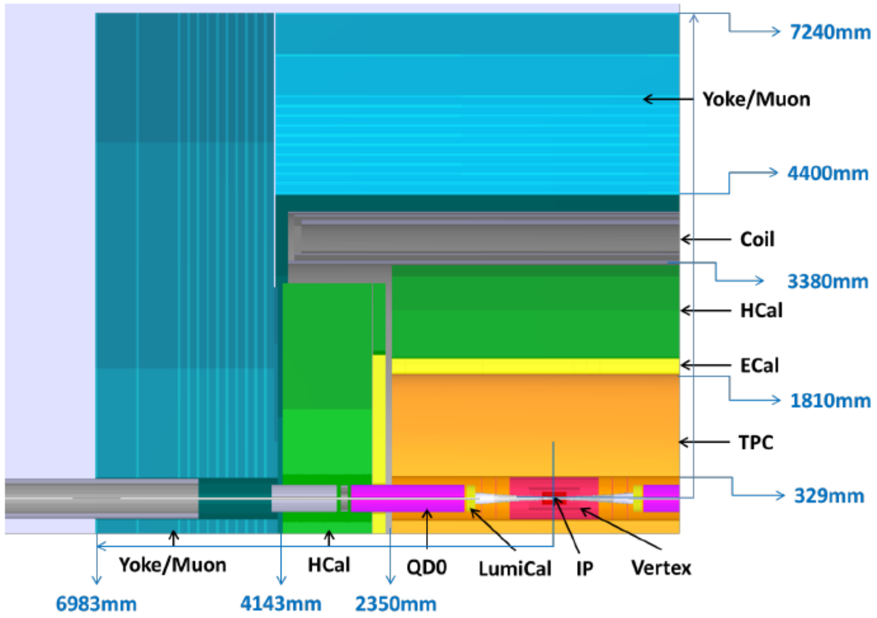
Z / ZZ*				
e^+e^-	$\nu\nu$jj	jj$\nu\nu$		
$\mu^+\mu^-$	$\nu\nu$jj	jj$\nu\nu$		jjjj
$\nu\nu$	e^+e^- jj	jje^+e^-	$\mu^+\mu^-$ jj	jj$\mu^+\mu^-$
jj	e^+e^- $\nu\nu$	$\nu\nu e^+e^-$	$\mu^+\mu^-$ $\nu\nu$	$\nu\nu$ $\mu^+\mu^-$

Monte Carlo Simulation

- CEPC_v4 (240GeV, 3T) configuration
- Generator: Whizard 1.95
  (with ISR, L=5.6 ab^{-1}, M_{higgs}=125 GeV)

- Simulation :
  Geant4 and Mokka with ISR and bremsstrahlung effects

- Reconstruction:
  Marlin and ArborPFA



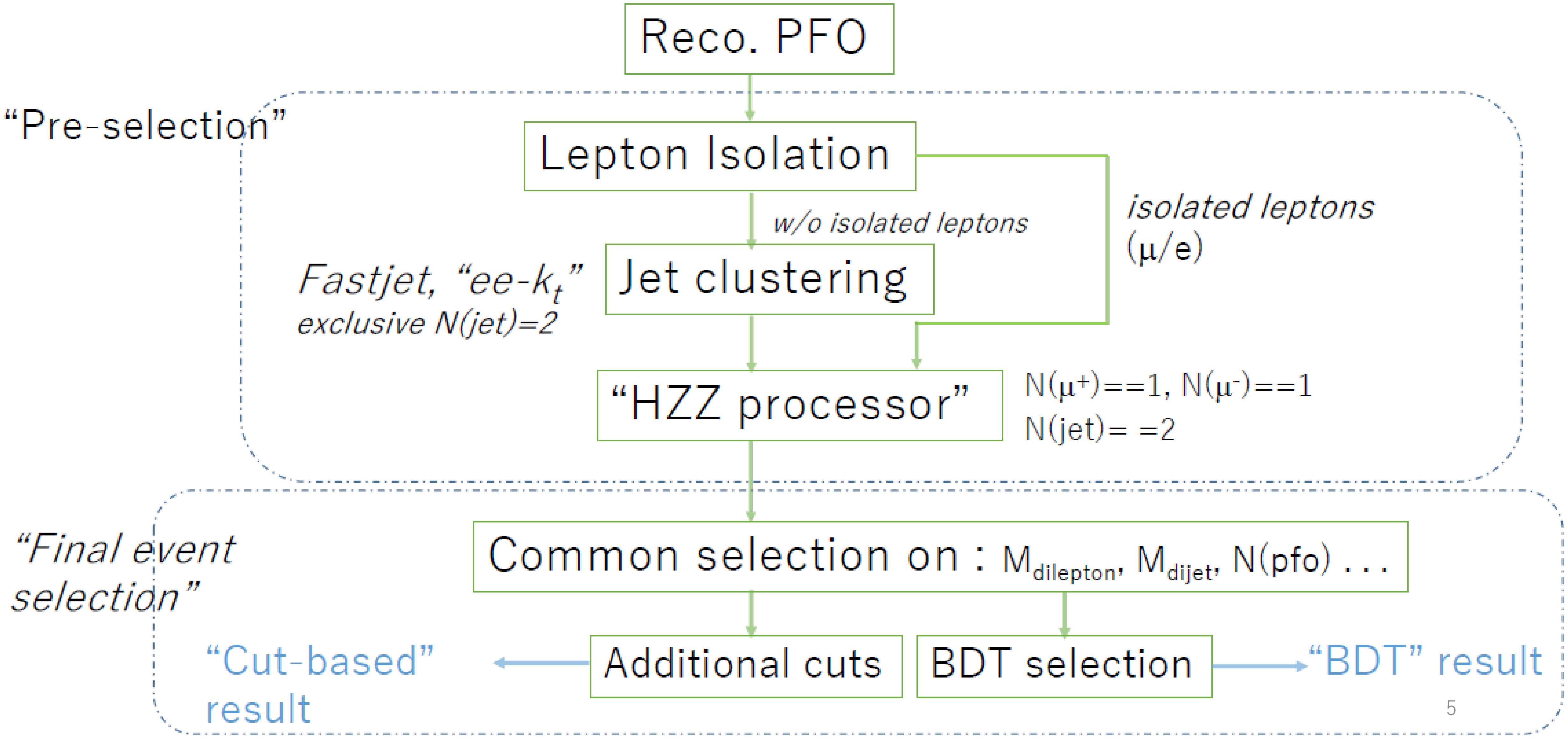
<u>Signal Sample</u>

-- $Z\to\mu\mu$, $H\to ZZ^*\to\nu\nu qq$

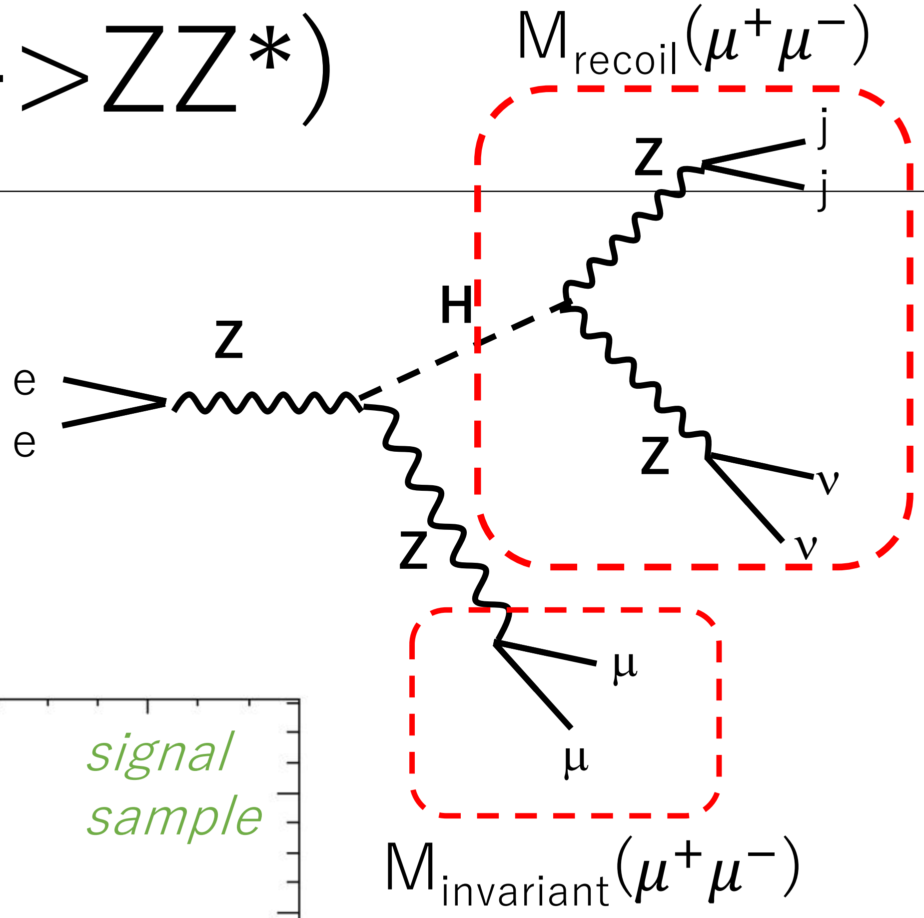
-- $Z\to\nu\nu$, $H\to ZZ^*\to\mu\mu qq$

-- $Z\to qq$, $H\to ZZ^*\to\nu\nu\mu\mu$

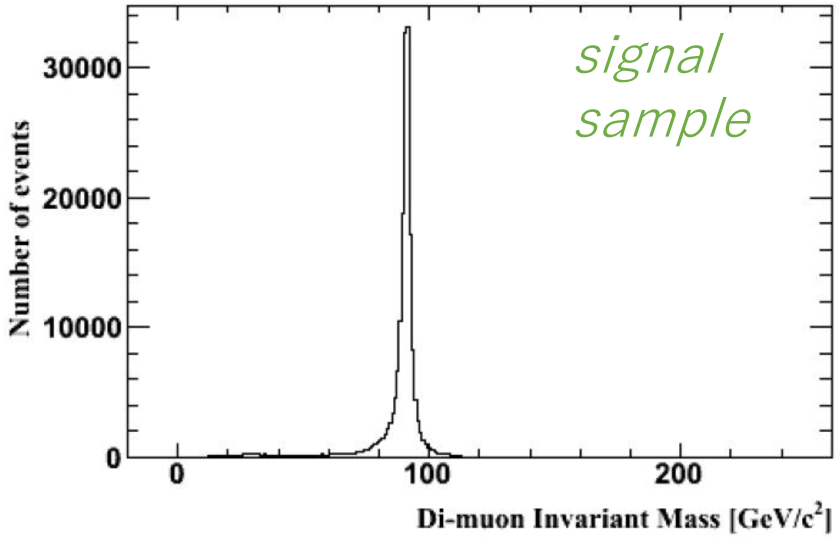
Analysis flow chart



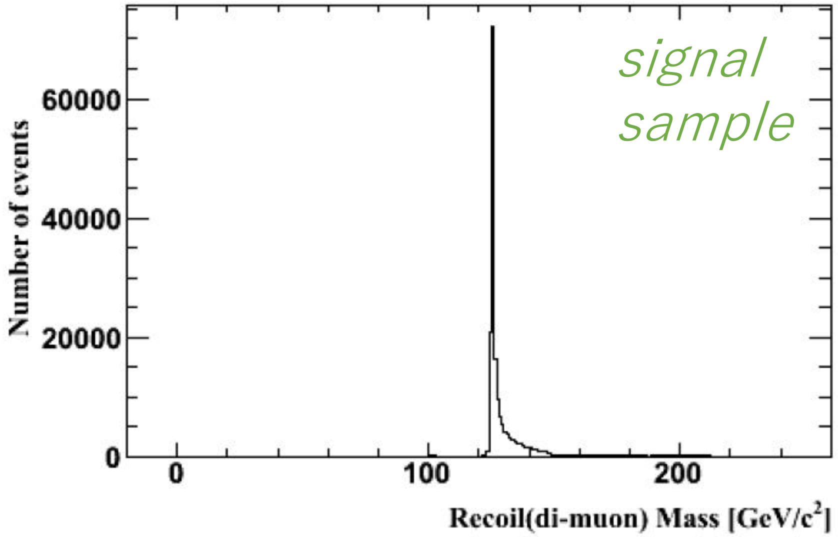
Signature of Z(->$\mu^+\mu^-$)H(->ZZ*)



Identify two muons from initial Z boson using invariant & recoil mass as usual



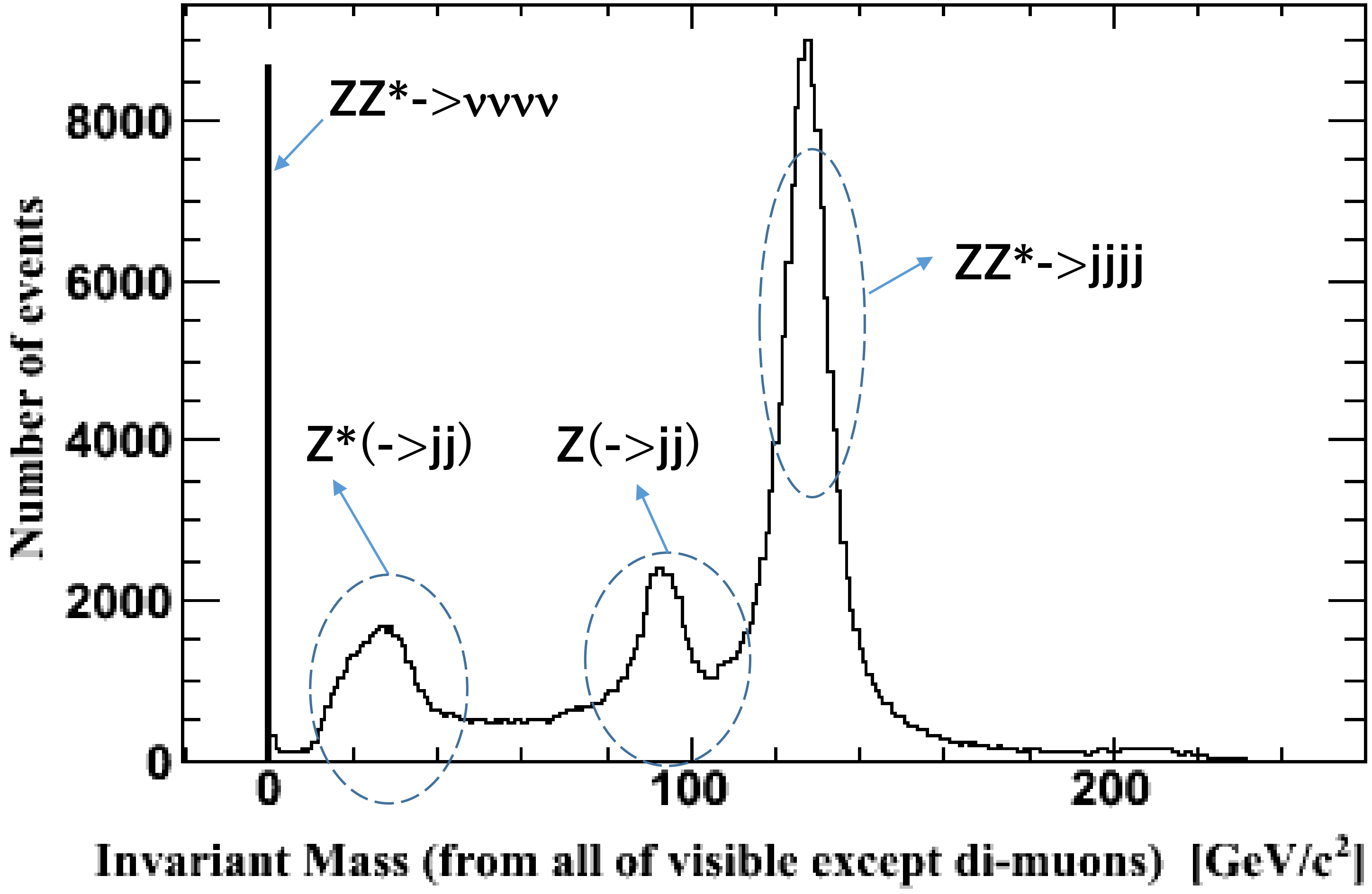
$M_{Invatiant}(\mu^+\mu^-)$



$M_{Recoil}(\mu^+\mu^-)$

Signature of Z(->$\mu^+\mu^-$)H(->ZZ*)

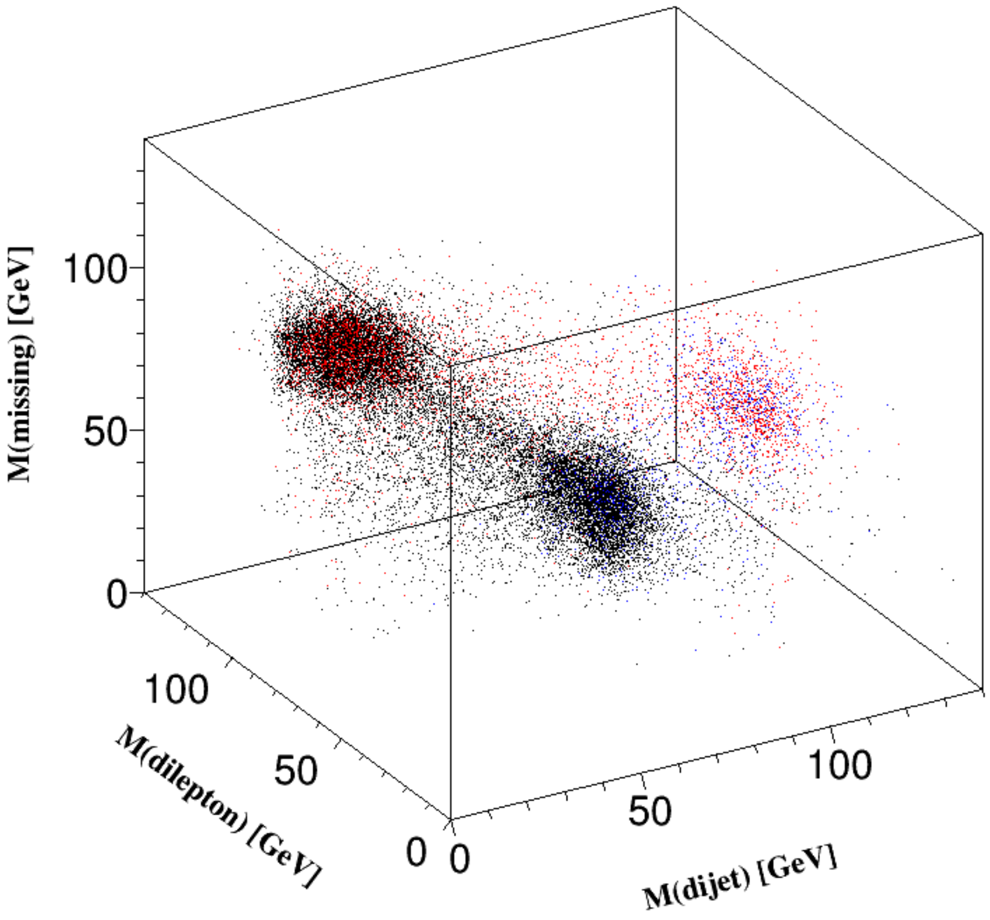
Distribution of invariant mass except two muons clearly shows each decay mode.



Jet clustering
N(jet)=2

Phase space overwrap

ee→ZH→ZH(ZZ*)

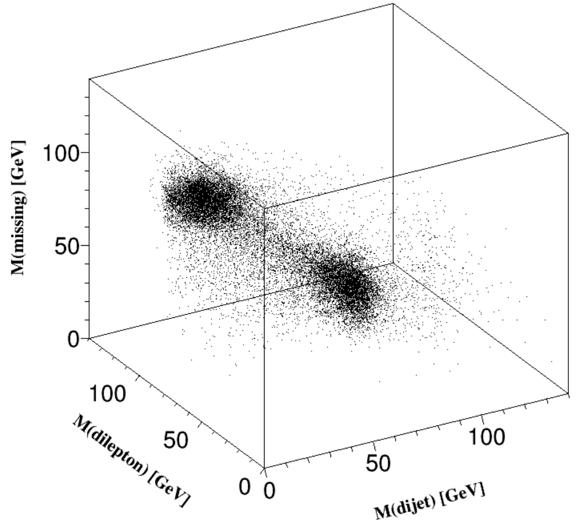


-- μμHννqq , ννHμμqq

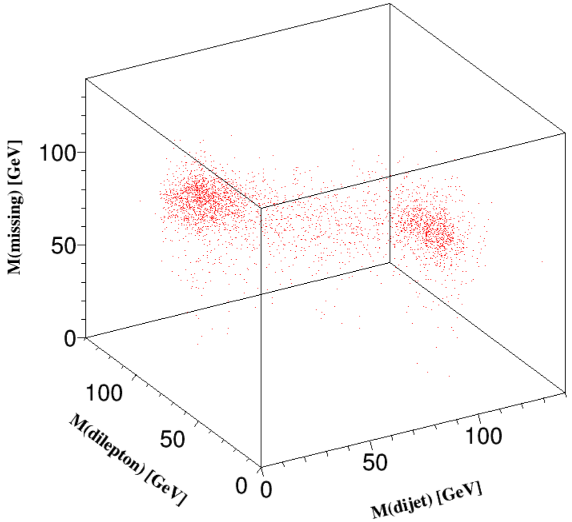
-- ννHqqμμ , qqHννμμ

-- qqHμμνν , μμHqqνν

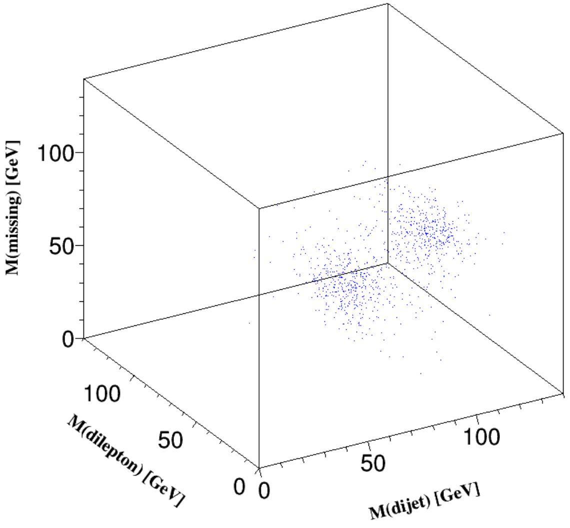
μμHZZ



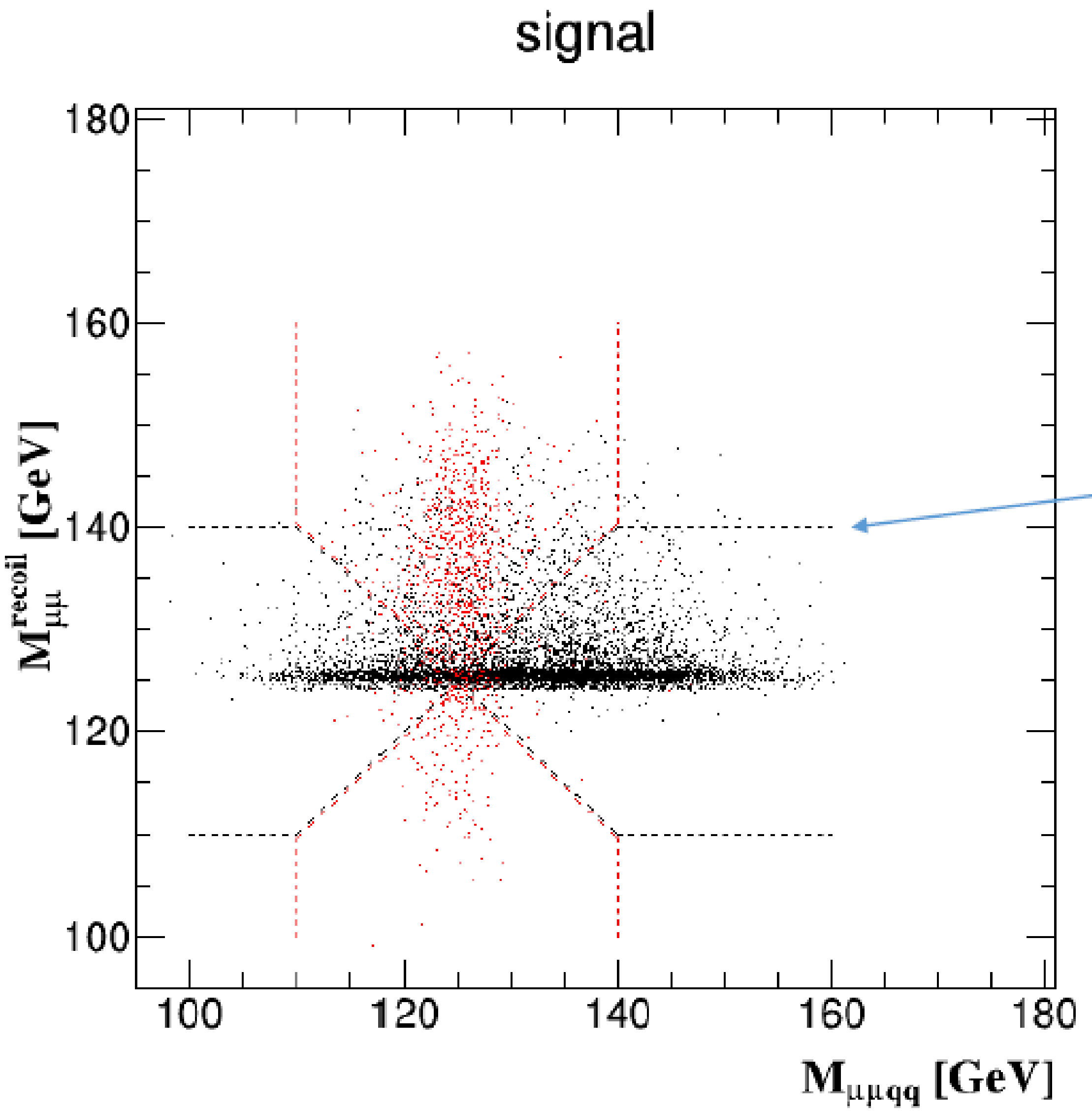
ννHZZ



qqHZZ



Separation on recoil mass distribution



・2D Recoil mass distribution shows overwrap of two signal channels

Black : $Z\rightarrow\mu\mu$, $H\rightarrow ZZ^*\rightarrow\nu\nu qq$

Red : $Z\rightarrow\nu\nu$, $H\rightarrow ZZ^*\rightarrow\mu\mu qq$

・To make each category exclusive, in other words, no double counting of events, "separation" has been performed on the recoil mass distribution.

Overview table I. Cut-based analysis

Table 1 Overview of the requirements applied when selecting events (cut-based).

Pre-selections						
$N(l) = 2$, where leptons(l) should pass the isolation criteria $N(\mu^+) = 1$, $N(\mu^-) = 1$ with $E(\mu^\pm) > 3$ GeV $N(jet) = 2$						
Selection (Cut-based)	$\mu\mu$H$\nu\nu qq$	$\mu\mu$H$qq\nu\nu$	$\nu\nu$H$\mu\mu qq$	$\nu\nu$H$qq\mu\mu$	qqH$\nu\nu\mu\mu$	qqH$\mu\mu\nu\nu$
Mass order	$M_{\text{miss}} > M_{jj}$	$M_{\text{miss}} < M_{jj}$	$M_{\mu\mu} > M_{jj}$	$M_{\mu\mu} < M_{jj}$	$M_{\text{miss}} > M_{\mu\mu}$	$M_{\text{miss}} < M_{\mu\mu}$
$M_{\mu\mu}$ (GeV)	[80, 100]		[60, 100]	[10, 60]	[15, 55]	[75, 100]
M_{jj} (GeV)	[15, 60]	[60, 105]	[10, 55]	[60, 100]	[75, 105]	
M_{miss} (GeV)	[75, 105]	[10, 55]	[75, 110]		[70, 110]	[10, 50]
$M_{\mu\mu}^{\text{recoil}}$ (GeV)	[110, 140]		-	-	[175, 215]	[115, 155]
M_{vis} (GeV)	-	[175, 215]	[110, 140]		[115, 155]	[185, 215]
M_{jj}^{recoil} (GeV)	[185, 220]	-	-	-	[110, 140]	
N_{PFO}	[20, 90]	[30, 100]	[20, 60]	[30, 100]	[40, 95]	[40, 95]
$\lvert\cos\theta_{\text{vis}}\rvert$	< 0.95					
$\Delta\phi_{ZZ}$ (degree)	[60, 170]	[60, 170]	< 135	< 135	-	[120, 170]
Region masking	*not-$\nu\nu$HZZ & not-qqHZZ*		*not-$\mu\mu$HZZ & not-qqHZZ*		*not-$\nu\nu$HZZ & not-$\mu\mu$HZZ*	

Overview table II. BDT-based analysis

Table 3 Overview of the requirements applied when selecting events (BDT-based).

Pre-selections						
$N(l) = 2$, where leptons(l) should pass the isolation criteria $N(\mu^+) = 1$, $N(\mu^-) = 1$ with $E(\mu^{\pm}) > 3$ GeV $N(jet) = 2$						
Selection (MVA)	$\mu\mu$H$\nu\nu qq$	$\mu\mu$H$qq\nu\nu$	$\nu\nu$H$\mu\mu qq$	$\nu\nu$H$qq\mu\mu$	qqH$\nu\nu\mu\mu$	qqH$\mu\mu\nu\nu$
Mass order	$M_{\mathrm{miss}} > M_{jj}$	$M_{\mathrm{miss}} < M_{jj}$	$M_{\mu\mu} > M_{jj}$	$M_{\mu\mu} < M_{jj}$	$M_{\mathrm{miss}} > M_{\mu\mu}$	$M_{\mathrm{miss}} < M_{\mu\mu}$
$M_{\mu\mu}$ (GeV)	[80,100]		-	-	-	-
M_{jj} (GeV)	-	-	-	-	[75, 105]	
M_{miss} (GeV)	-	-	[75, 110]		-	-
$M_{\mu\mu}^{\mathrm{recoil}}$ (GeV)	[110, 140]		-	-	-	-
M_{vis} (GeV)	-	-	[110, 140]		-	-
M_{jj}^{recoil} (GeV)	-	-	-	-	[110, 140]	
N_{PFO}	[20, 90]	[30, 100]	[20, 60]	[30, 100]	[40, 95]	[40, 95]
$\lvert\cos\theta_{\mathrm{vis}}\rvert$	< 0.95					
Region masking	*not-$\nu\nu$HZZ & not-qqHZZ*		*not-$\mu\mu$HZZ & not-qqHZZ*		*not-$\nu\nu$HZZ & not-$\mu\mu$HZZ*	
BDT score	> 0.14	> 0.01	> -0.01	> -0.01	> -0.04	> -0.01

Number of expected events & efficiency

Table 2 Summary of the selection efficiency ϵ and the number of expected events $N_{evt.}$ for each category after the final event selection in the cut-based analysis..

(Cut-based)	$\mu\mu H\nu\nu qq$		$\mu\mu Hqq\nu\nu$		$\nu\nu H\mu\mu qq$	
Process	ϵ [%]	$N_{evt.}$	ϵ [%]	$N_{evt.}$	ϵ [%]	$N_{evt.}$
Signal ("dominant")	38	53	36	50	54	76
Signal ("sub")	6	8	10	14	6	9
Higgs decays Bg.	$2.2 \cdot 10^{-3}$	25	$7.0 \cdot 10^{-2}$	794	$5.3 \cdot 10^{-4}$	6
SM four-fermion Bg.	$3.7 \cdot 10^{-6}$	4	$4.9 \cdot 10^{-4}$	520	$5.6 \cdot 10^{-6}$	6
SM two-fermion Bg.	0	0	0	0	0	0
	$\nu\nu Hqq\mu\mu$		$qqH\nu\nu\mu\mu$		$qqH\mu\mu\nu\nu$	
Process	ϵ [%]	$N_{evt.}$	ϵ [%]	$N_{evt.}$	ϵ [%]	$N_{evt.}$
Signal ("dominant")	36	51	26	37	23	32
Signal ("sub")	8	11	7	10	4	6
Higgs decays Bg.	$1.0 \cdot 10^{-2}$	114	$2.4 \cdot 10^{-2}$	275	$1.4 \cdot 10^{-2}$	160
SM four-fermion Bg.	$4.3 \cdot 10^{-5}$	46	$1.5 \cdot 10^{-4}$	157	$1.8 \cdot 10^{-4}$	190
SM two-fermion Bg.	0	0	0	0	0	0

- Signal efficiency : 27 - 60 %

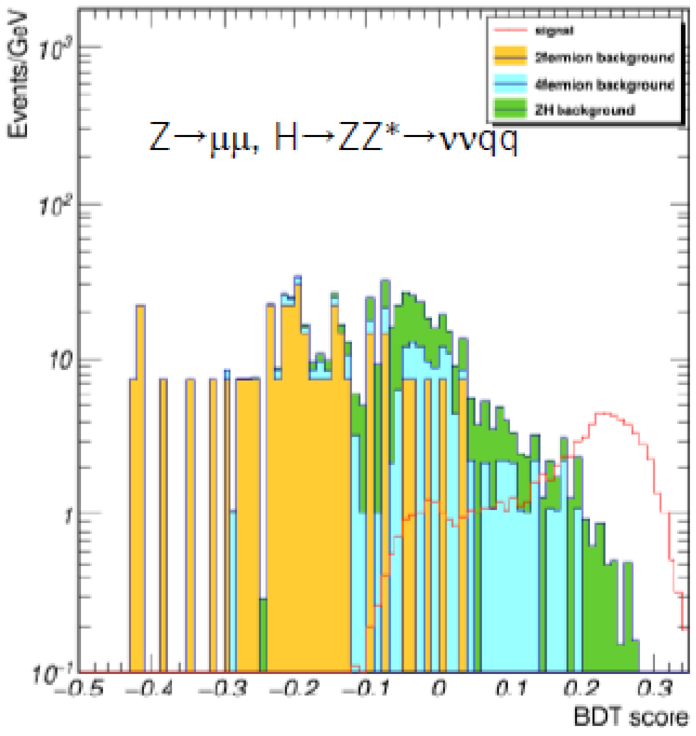
BDT analysis

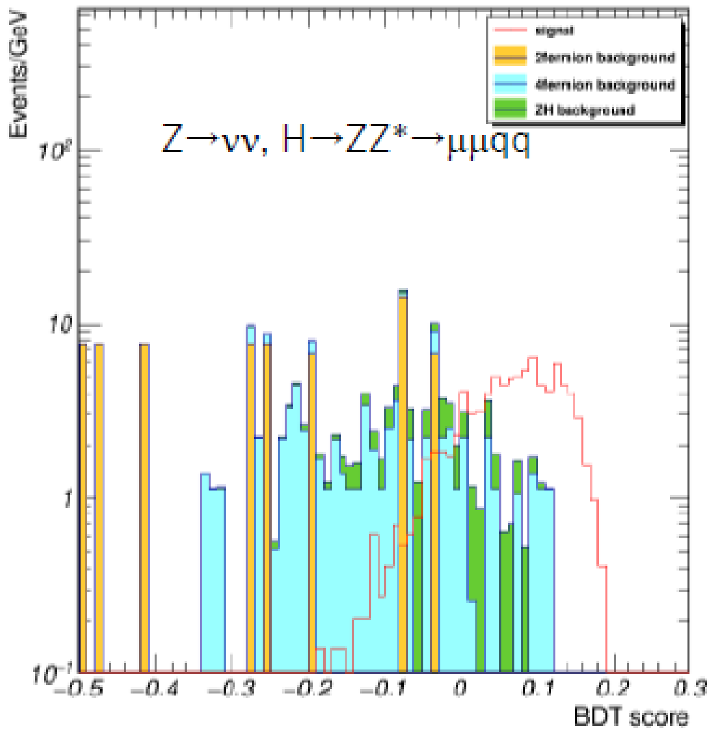
・ BDT training and cuts are applied after several simple cuts : M_{xx}, M^{recoil}_{xx}, N_{pfo}, cosθ, etc.

・ BDT cut position is optimized based on the measure, $S/\sqrt{(S+B)}$, which to be maximized.

Variables used in the BDT

$P_{\text{all visible}}$	$E_{\text{leading-jet}}$	Cos(θ)
$Pt_{\text{all visible}}$	$E_{\text{sub-leading-jet}}$	($\text{RecoilM}_{\text{dimuon}}$)
M_{dijet}	N(pfo)	($M_{\text{all visible}}$)
M_{dimuon}	$\text{Angle}_{(\text{dijet-dimuon})}$	

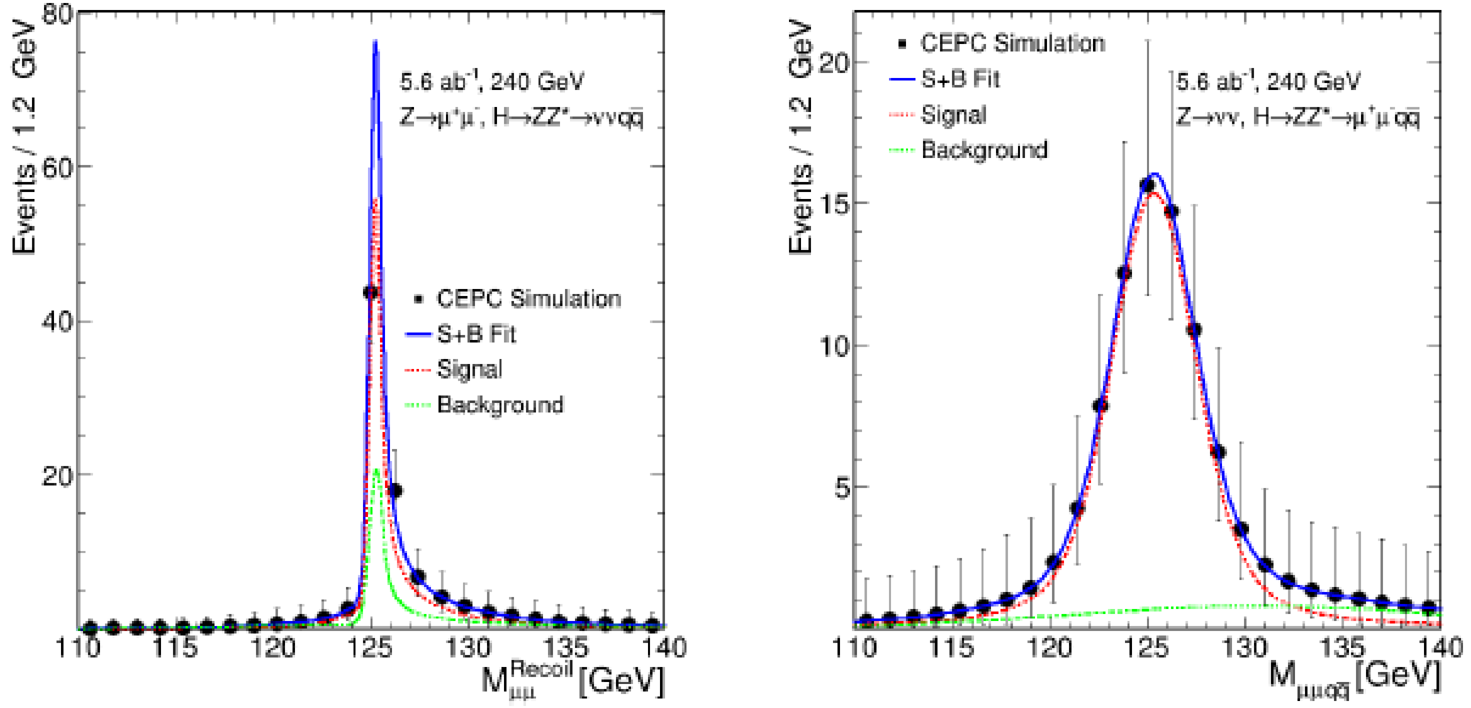




Obtained precision

Obtained uncertainty from each category. both of cut-based and BDT-based results are shown together.

Category	$\frac{\Delta(\sigma \cdot BR)}{(\sigma \cdot BR)}$ [%] cut-based	BDT
$\mu\mu H\nu\nu qq^{\mathrm{cut/mva}}$	15	14
$\mu\mu Hqq\nu\nu^{\mathrm{cut/mva}}$	48	42
$\nu\nu H\mu\mu qq^{\mathrm{cut/mva}}$	12	12
$\nu\nu Hqq\mu\mu^{\mathrm{cut/mva}}$	23	20
$qqH\nu\nu\mu\mu^{\mathrm{cut/mva}}$	45	37
$qqH\mu\mu\nu\nu^{\mathrm{cut/mva}}$	52	44
Combined	8.3	7.9



Figs. Recoil mass distribution from two best categories

Submission

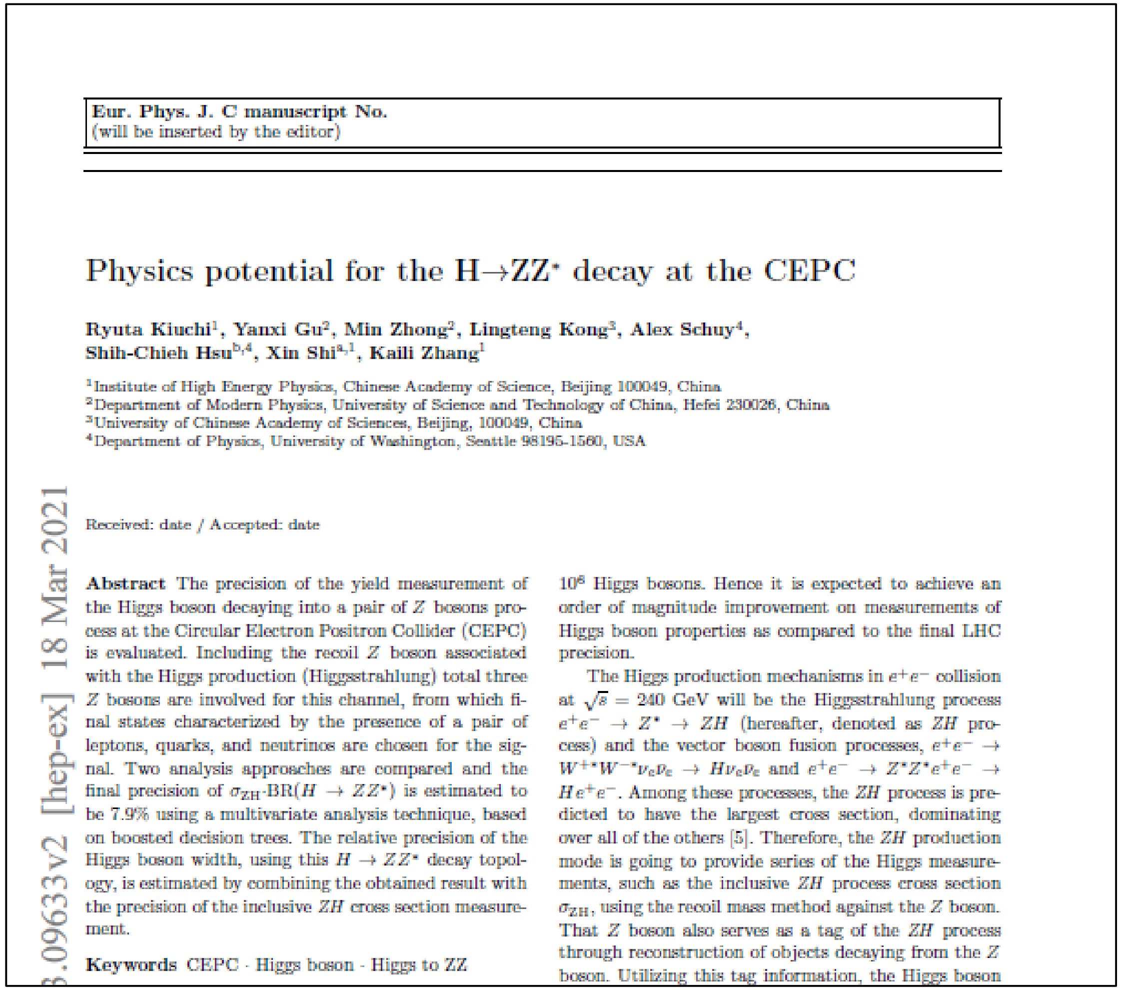
- Uploaded on the arXiv

  -- https://arxiv.org/abs/2103.09633

- Submitted to EPJC (last month)

  -- have received reply when we are traveling to Yangzhou …

- Start to consider necessary steps and action items.


Eur. Phys. J. C manuscript No.
(will be inserted by the editor)

Physics potential for the H→ZZ* decay at the CEPC

Ryuta Kiuchi[1], Yanxi Gu[2], Min Zhong[2], Lingteng Kong[3], Alex Schuy[4], Shih-Chieh Hsu[b,4], Xin Shi[a,1], Kaili Zhang[1]

[1]Institute of High Energy Physics, Chinese Academy of Science, Beijing 100049, China
[2]Department of Modern Physics, University of Science and Technology of China, Hefei 230026, China
[3]University of Chinese Academy of Sciences, Beijing, 100049, China
[4]Department of Physics, University of Washington, Seattle 98195-1560, USA

arXiv:2103.09633v2 [hep-ex] 18 Mar 2021

Received: date / Accepted: date

Abstract The precision of the yield measurement of the Higgs boson decaying into a pair of Z bosons process at the Circular Electron Positron Collider (CEPC) is evaluated. Including the recoil Z boson associated with the Higgs production (Higgsstrahlung) total three Z bosons are involved for this channel, from which final states characterized by the presence of a pair of leptons, quarks, and neutrinos are chosen for the signal. Two analysis approaches are compared and the final precision of $\sigma_{ZH}\cdot\mathrm{BR}(H \to ZZ^*)$ is estimated to be 7.9% using a multivariate analysis technique, based on boosted decision trees. The relative precision of the Higgs boson width, using this $H \to ZZ^*$ decay topology, is estimated by combining the obtained result with the precision of the inclusive ZH cross section measurement.

Keywords CEPC · Higgs boson · Higgs to ZZ

10^6 Higgs bosons. Hence it is expected to achieve an order of magnitude improvement on measurements of Higgs boson properties as compared to the final LHC precision.

The Higgs production mechanisms in e^+e^- collision at $\sqrt{s}$ = 240 GeV will be the Higgsstrahlung process $e^+e^- \to Z^* \to ZH$ (hereafter, denoted as ZH process) and the vector boson fusion processes, $e^+e^- \to W^{+*}W^{-*}\nu_e\bar{\nu}_e \to H\nu_e\bar{\nu}_e$ and $e^+e^- \to Z^*Z^*e^+e^- \to He^+e^-$. Among these processes, the ZH process is predicted to have the largest cross section, dominating over all of the others [5]. Therefore, the ZH production mode is going to provide series of the Higgs measurements, such as the inclusive ZH process cross section σ_{ZH}, using the recoil mass method against the Z boson. That Z boson also serves as a tag of the ZH process through reconstruction of objects decaying from the Z boson. Utilizing this tag information, the Higgs boson

Signal yield and remaining background

・ Related with the difference of the signal yield in the white paper and current analysis, would need explanation as well as the other information such as background reduction rate and component.

$Z\rightarrow\mu\mu$, $H\rightarrow ZZ^*\rightarrow\nu\nu qq$

- $M_{\mu\mu}$: 80-100 GeV
- $M^{recoil}_{\mu\mu}$: 120-160 GeV
- M_{jj} : 10-38 GeV
- P_t : > 10 GeV

$Z\rightarrow\nu\nu$, $H\rightarrow ZZ^*\rightarrow l^+l^-qq$

- Visible Energy : < 180 GeV
- $M_{missing}$: 58-138 GeV
- Mass & pt cuts on lepton/jet pairs

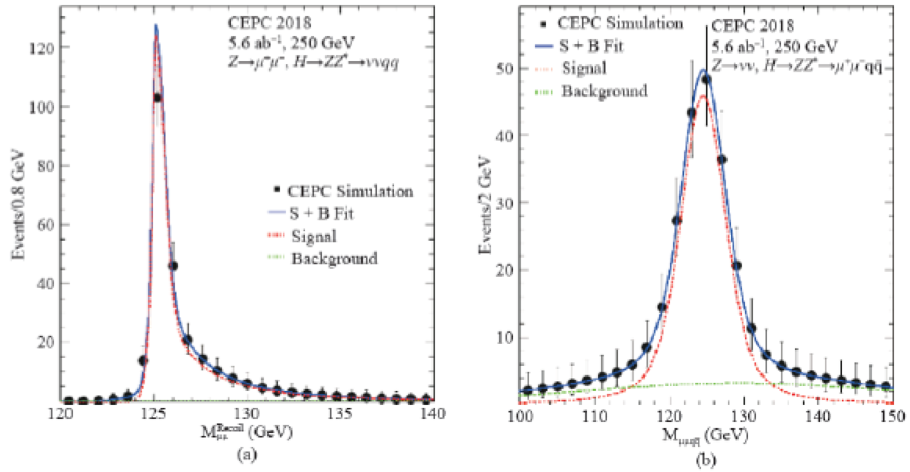


Table 8. Expected relative precision for the $\sigma(ZH)\times\mathrm{BR}(H\rightarrow ZZ^*)$ measurement with an integrated luminosity 5.6 ab^{-1}.

ZH final state		precision
$Z\rightarrow\mu^+\mu^-$	$H\rightarrow ZZ^*\rightarrow\nu\bar{\nu}q\bar{q}$	7.2%
$Z\rightarrow\nu\bar{\nu}$	$H\rightarrow ZZ^*\rightarrow\ell^+\ell^-q\bar{q}$	7.9%
combination		4.9%

16

from the white paper

Some action items

・ B-tagging

Current analysis does not utilize the vertex tagging information. This is asked specifically how much improvement could be achieved if it is utilized.

・ Investigate background components whether further effective reduction can be achieved or not.

(・ Systematics associated with

-- signal cross-talk

-- BDT cut position)

Beyond current analysis

・ Current analysis only looking at the HZZ final states having di-muon, di-jets, and di-neutrino (missing) , which is less than 3% of total HZZ decay

・ Final states having “4q”, such as Z(→qq)H(Z→ll, Z*→qq), Z(→ll)H(Z→qq, Z*→qq) … which has larger statistics.

・ What about “6q” status ?

Summary

・ We have analyzed 3 combinations of Z boson decays for HZZ measurement. ($\mu\mu H\nu\nu qq/qq\nu\nu$, $\nu\nu H\mu\mu qq/qq\mu\mu$, $qqH\mu\mu\nu\nu/\nu\nu\mu\mu$) The obtained combined final precision is ~7.9%.

・It is submitted to EPJC. We will perform actions triggered by the inquiry.

・ Any recommendations are highly appreciated !

Thank you very much !

Backup

Ref: From a past discussion on introducing b-tagging

・ H->bb background, namely "e2e2h_bb" ($Z(\rightarrow\mu\mu)H(\rightarrow bb)$), is a dominant background in following channels

-- $Z(\rightarrow qq)H(Z\rightarrow\nu\nu, Z^*\rightarrow\mu\mu)$
-- $Z(\rightarrow\mu\mu)H(Z\rightarrow\nu\nu, Z^*\rightarrow qq)$

・A rough estimation about how much improvement could be achieved

<u>Assuming following scenario for a comparison</u>

-- $Z(\rightarrow\mu\mu)H(\rightarrow bb)$ event is completely cut by using the b-tagging information

-- Since, the signal and the dominant channels in remaining four-fermion bg. ("zz_sl0mu_up/down) include a decay of Z->bb, it is assumed that their yield becomes 80% by b-tagging.

Ref: From a past discussion on introducing b-tagging

- Comparison of the numbers between the original & w. b-tagging

Case for the channel
Z(→qq)H(Z→νν, Z*→μμ)

method	N_{event}(signal)	N_{event}(zh)	N_{event}(4F)	√(S+B)/S
Original	35	206	305	0.667
w B-tagging	28	86	245	0.677

Case for the channel
Z(→μμ)H(Z→νν, Z*→qq)

method	N_{event}(signal)	N_{event}(zh)	N_{event}(4F)	√(S+B)/S
Original	48	774	659	0.802
w B-tagging	38	355	530	0.799

-- From this "coarse" comparison, the improvement might not be so huge.

-- But of course, estimation is very rough. (not consider b-tagging eff. , as well as the reduction on HWW bg. events)