





# MVA method in $ZH \rightarrow qq\gamma\gamma$ analysis in CEPC

<u>Fangyi Guo</u>, Yaquan Fang, Xinchou Lou Institute of High Energy Physics, CAS

CEPC Physics Workshop 3rd July, 2019, Peking University

#### Content

Review of CEPC CDR

MC sample and simulation

Selection adjustment

MVA method

2-D signal and background fit

Results

Further work

Conclusion

## Review of CEPC CDR

CEPC CDR release: Nov 2018, IHEP

Higgs  $\rightarrow \gamma \gamma$  physics analysis:

- $^{\circ}\,$  Design point at CEPC\_v4,  $\sqrt{s}=240 GeV, \mathcal{L}=5.6 a b^{-1}$
- Whizard 1.95 + MoccaC generator, dedicated fast simulation based on parametrized detector response.
- Considered  $H \rightarrow \gamma \gamma$  signal and 2 fermion dominant background
- Result:  $\delta(Br(H \to \gamma \gamma) \times \sigma(ZH))=6.84\%$  in 3 combined channel, and 9.84% in  $q\bar{q}\gamma\gamma$  channel.
- Could be improved by applying MVA method



**Results** 

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

Sub-channel	$q\overline{q}\gamma\gamma$	ΙĪγγ	ν⊽γγ	combined
precision	9.84%	23.7%	10.5%	6.84%

Arxiv: <u>1810.09037</u> <u>Report in 2018 CEPC WS</u>

## MC sample and simulation

In order to keep consistent with CDR:

MC sample:

- Signal:  $ee \rightarrow ZH \rightarrow q\bar{q}\gamma\gamma$ , 100k events.
- Background:  $ee \rightarrow qq$ +radiation photons, 20M events.

#### Simulation and event reconstruction

- Fast simulation: smear the objects with the resolution and efficiency with parametrized detector response,  $\frac{\Delta E}{E} \sim \frac{16\%}{\sqrt{E}} \bigoplus 1\%$
- Event reconstruction: FSClasser FastJetClustering processer.

reconstructed 2 on-shell photon with  $m_{\gamma\gamma}$ ~125GeV and 2 jets define:  $\gamma_1/j_1$  as photon/jet with lower energy, and  $\gamma_2/j_2$  as higher energy one.

## Event selection adjustment

Release the event selection criteria for the further MVA method

Selection in CDR:  $E_{\gamma 1} > 35 GeV$   $35 GeV < E_{\gamma 2} < 96 GeV$   $cos \theta_{\gamma \gamma} > -0.95, cos \theta_{jj} > -0.95$   $pT_{\gamma 1} > 20 GeV, pT_{\gamma 1} > 30 GeV$   $110 GeV < m_{\gamma \gamma} < 140 GeV$   $125 GeV < E_{\gamma \gamma} < 145 GeV$  $min |cos \theta_{\gamma j}| < 0.9$ 



Old	signal	background
Total efficiency	53.08%	0.010%
Scaled to 5.6 $ab^{-1}$	923.13	29875.6

Present:

 $E_{\gamma 1} > 25 GeV$   $35 GeV < E_{\gamma 2} < 96 GeV$   $cos \theta_{\gamma \gamma} > -0.95, cos \theta_{jj} > -0.95$   $pT_{\gamma 1} > 20 GeV, pT_{\gamma 1} > 30 GeV$   $110 GeV < m_{\gamma \gamma} < 140 GeV$   $E_{\gamma \gamma} > 120 GeV$   $min |cos \theta_{\gamma j}| < 0.9$ 

New	signal	background
Total efficiency	53.34%	0.010%
Scaled to 5.6 $ab^{-1}$	927.65	31587.6

Considered variables:

- P, E, pT, *cosθ* 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
- $\Delta P$ ,  $\Delta E$ ,  $\Delta \phi$  between two photon,  $\gamma \gamma$ -qq
- Cosine angle between 2 photon, 2 jets, 1 photon and 1 jet,  $\gamma\gamma$  and jj system.
- $\,\circ\,$  Minimum  $\Delta R$  between any photon and jet

Totally 42 variables

Separation power:

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

Pt\*: Di-photon P projected perpendicular to the diphoton thrust axis.(similar as pTt but replace pT with P) pTt =  $|(\overrightarrow{P_1} + \overrightarrow{P_2}) \times \frac{\overrightarrow{P_1} - \overrightarrow{P_2}}{|\overrightarrow{P_1} - \overrightarrow{P_2}|}|$ 

y: discriminating variable  $\hat{y}_s(y)$  and  $\hat{y}_B(y)$ : the distributions of the variable for signal and background samples

#### Variable correlation matrix





First step: remove high  $m_{\gamma\gamma}$ -related variable Second step: remove high co-related variables

 $|Corr_{v-m_{\gamma\gamma}}| < 30\%$  $|Corr_{v_1-v_2}| < 40\%$ 

#### 2019/7/3

#### Remaining variables:

Variable	Definition	$\langle S^2 \rangle$
$pT_{\gamma 1}$ / $pT_{\gamma 2}$	pT of $\gamma 1$ or $\gamma 2$	0.39 / 0.30
$cos\theta_{\gamma 2}$	Cosine polar angle of $\gamma 2$	0.39
$ \Delta \Phi_{\gamma\gamma} $	$ \Delta \Phi $ between 2 photon	0.30
$min\Delta R_{\gamma,j}$	Minimum $\Delta R$ between photon and jet	0.09
$cos \theta_{j1}$	Cosine polar angle of <i>j</i> 1	0.08
$pT_{j2}$	pT of <i>j</i> 2	0.08
$E_{j1}$	Energy of <i>j</i> 1	0.03
$cos\theta_{\gamma 2-j2}$	Cosine value of the angle between $\gamma 2$ and $j 2$	0.03
$cos\theta_{\gamma 1-j1}$	Cosine value of the angle between $\gamma 1$ and $j 1$	0.02
$ \Delta \Phi_{\gamma\gamma-jj} $	$ \Delta \Phi $ between $\gamma \gamma$ and <i>jj</i> system	0.01
$cos\theta_{\gamma\gamma-jj}$	Cosine value of the angle between $\gamma\gamma$ and $jj$ system	0.01

BDT training parameter:

"BDTG","NTrees=900:nEventsMin=50:BoostType=Grad:Shrinkage=0.06:UseBaggedGrad:GradBaggingFracti on=0.6:nCuts=20:MaxDepth=3"



MVA variables distribution for signal(blue) and background(red)



MVA variables distribution(left) and BDT response(right) for signal(blue) and background(red)

 $|Corr_{m_{\gamma\gamma}-BDTout}|$  is 5.4% in signal and 17% in background

## 2-D signal and background fit

Fit the sample in 2 dimension

• Di-photon invariant mass  $m_{\gamma\gamma}$ 

Signal: Gaussian PDF

Background: 2<sup>nd</sup> polynomial exponential PDF Р

$$PDF_{m_{\gamma\gamma}} = \mu \times N_{sig}^{SM} \times PDF_{sig} + N_{bkg} \times PDF_{bkg}$$

• BDT response (
$$|Corr_{m_{\gamma\gamma}-BDT}| < 20\%$$
)  
Binned PDF for signal and background  
 $PDF_{BDT} = \mu \times N_{sig}^{SM} \times PDF_{sig} + N_{bkg} \times PDF_{bkg}$ 

• 
$$PDF_{2D} = PDF_{m_{\gamma\gamma}} \times PDF_{BDT}$$



#### Results

Fit result in  $ee \rightarrow ZH \rightarrow qq\gamma\gamma$  channel

 $\mu = 1.000 \pm 0.066$ 

	$qq\gamma\gamma$ channel	3 Combined channel
Pre MVA (CDR result)	9.84%	6.80%
After MVA	6.56%	5.39%
Improvement	33%	21%

\*combined results are based on MVA  $qq\gamma\gamma$  + no MVA  $(ll\gamma\gamma + \nu\nu\gamma\gamma)$ 

FCC-ee case:

- 3% @240GeV, 10ab<sup>-1</sup>, based on CMS ECal resolution, <u>TLEP physics, 2013</u>
- 9% @240GeV, 5ab<sup>-1</sup>, <u>FCC-ee CDR,</u> 2018

#### Future work

MVA method in  $ZH \rightarrow ll\gamma\gamma$ ,  $ZH \rightarrow \nu\nu\gamma\gamma$  channel

• Estimation: if the same improvement(30%) could be reached in these two channel

*ll* $\gamma\gamma$  channel: 23.7%  $\implies$  16.6%  $\nu\nu\gamma\gamma$  channel: 10.5%  $\implies$  7.4%

combined results after MVA: ~4.7%

Simulation sample

- Update the fast simulation sample to full simulation
  - Signal: broaden the  $m_{\gamma\gamma}$  distribution, lose ~7.5% events
  - Background: include more high energy photon, might be excluded by reconstruction algorithm.

#### Conclusion

MVA method improvement in CEPC  $H \rightarrow \gamma \gamma$  analysis

- $qq\gamma\gamma$  channel has been tested: 33% improvement in one sub-channel, reaching the result  $\delta(Br \times \sigma) = 6.56\%$
- Prospect: 4.7% combined precision could be reached after MVA in 3 channels

Next step towards TDR

- Full simulation samples are ready, but a accurate reconstruction algorithm is necessary
- Based on previous study, a ~20% decrease would appear after converting to full simulation

## Back up

#### Correlation matrix (after removing $m_{\gamma\gamma}$ -related variables)

#### Linear correlation coefficients in % 100 12-8 100 osTheta\_yy\_jj 2931 5 5 -192039 5 5 - 3 $\Delta Phi_{\gamma\gamma,jj} = 1$ $\Delta P_{\gamma\gamma,jj} = 2 - 3$ ninDeltaR\_y\_j = 1328 -100 1 31522 -242354 1 1 2 100-1-8 80 -82821 -2 1 -2 1 -1919 3 4 - 7 -3-3-3-300 44 66 osTheta\_y2j2 -2 2 29734900-3 60 osTheta\_y2j1 3 2 -72280049-3 2 55 44 osTheta\_y1j2 3 2 1-3 1 -4-4 2 -49002873-3 osTheta\_y1j1 2 2 1-4 2-1 -6-6 3 100497229-3 40 CosTheta -97 -165 100 -1576 m; -3 -1010 2 454200 3 2 1 3-754 1-3 22912 20 p\_j1 -3 E\_j1 -3 5 5920 940042 -6-4-4-4 4<mark>2</mark>3 15721 1009445 -6-4-5-4 324 -1-1 cos<sub>θ</sub>, -646 -73 -7000 0 76 cost 9 1 -112 10970 -**15**-1 pT. -6-500 139 -20 44 2 1 -20 pT<sup>P</sup> 100-5 575929 -4-3-2-22815 7-11 DeltaPhi\_yy 100 -6 612 55 2 -8 3 -10 CosTheta\_yy -97 1273 -40 100 166 m<sub>γγ</sub> 131300 -1 7 5 2 p\_y2 4348 1000013 4 5 6 419 -10 -60 E\_y2 1000013 -10 4 5 6 4 1 9 1-46 9 -6 cos0 -65 cosť -80 pT 4343 5 -3 300 1211 7 2 2 2 2 2 2 2 8 - 3 pT<sup>2</sup> 00 3 -4343 7 6-7-18 -3-3 -2-3-3-213 2 -100PTPTCose & Balta Cose ATPTCose & P. /7, Cose & Cose hpsTheta yy

Correlation Matrix (signal)

#### Correlation Matrix (background)

Linear correlation coefficient	ents in %	100
osTheta_yy_jj 3121 4 - 42525 5 - 221945 1 - 2 - 2 - 2 1 - 6 - 3 20 - 8	100	100
ΔPhi <sub>yyj</sub> -2 -1 1 4 4 4 2 -2 3 -1 3 3 -3 -2 -2 1 -1 -1 1	00	
ΔP <sub>γγ,i</sub> 10 4 2 5 510 1-2119 4-6-272660-4 3-3-3 1-6 <mark>00</mark>	-8-	80
ninDeltaR_y_i 2935 2 - 21919 1 - 221222-5 3 1717 3 - 5 -8 -4 -4 00-6	-120	
osTheta_y2j2 -2 3030 5 5 222 6-3 20 7 -8 -7 3 212 664 500 -4 1	-1-3	60
osTheta_y2j1 -2 -12630 4 4 626 -1 31015 -4 -5 72667240045 -4 -3	1	00
osTheta_y1j2 32931 8 8 1124 1-5 41411-6-7 52343002466-8-3		40
osiheta_yiji 13131 7 7 723 4-6-21311-4-3 -2200436721-5 3	-2 -6	40
	-2 1	
	-3-2	20
	3-2	
	<u> - 2</u>	0
$\cos \theta_2$ -7 1440 5 5 -74-5 5 -10000 4 5 7 -111 5 7 5-0 $\cos \theta_1$ A 11 7 1 1 2 9 2 3 00060 9 8 2 012141020 5 A		U
	45	
PT <sup>2</sup> 211 4-3 40017 3 3696530 1 6 5 31001	210	-20
DeltaPhi vy 2425-2 2121212 300 415 2 3 1 1 -3 4 1 1 623 1	-2.21	
CosTheta vy 5 3170-5-5 100 3 1 74 2 4 39823242622	2.2	_40
$m_{\gamma\gamma}$ 12-9-2 1-3-3 00 12 -16-3 2 2 7 711 6 2 110	-4 5	-10
p y2-2721 -10000-3-512 -9 -1 5-3-215 5 7 8 4 519 5	425	
E y2-2721 -10000-3-512 -9-1 5-3-215 5 7 8 4 519 5	425	-60
cosθ <sub>y2</sub> 5 - 8700-1 -1 170 2 -3 1 -748 -6831313930-2	1-4	
cosθ <sub>v1</sub> -4 30987 -231-2 4-21117 1 2 3131292630 2 2	-14 -	-80
PI 2000 3 -12121-9 2511 8 -5 -310 1 3 -1 -35 4	21	
PI 0020-4 5272712 524-223 4-7 -3 -3 -6 -5 2 -2 -22910	-2-31	100
PTPTCOGOG P. M. C.D. RTPTCOGOG D. M. COC. C.C. DMA	P1PCon	-100
N V2 VI V2 V	Bagy Th	eta
- YYY 11 - P4994	ESPERIE_j	NY.