Measurements of decay branching fractions of the Higgs boson to hadronic final states at the CEPC

[1]Yanping Huang, [1]Xiaotian Ma, [1,2]Zuofei Wu

[1]*Institute of High Energy Physics*

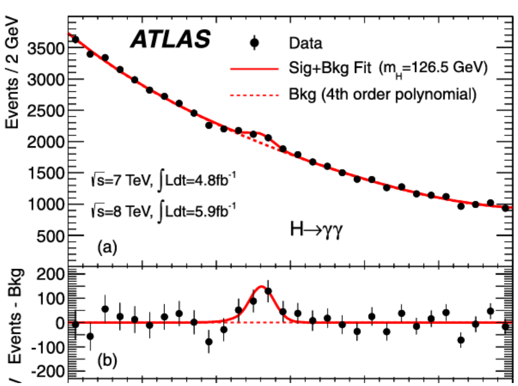
[2]*Nanjing University*

Contents

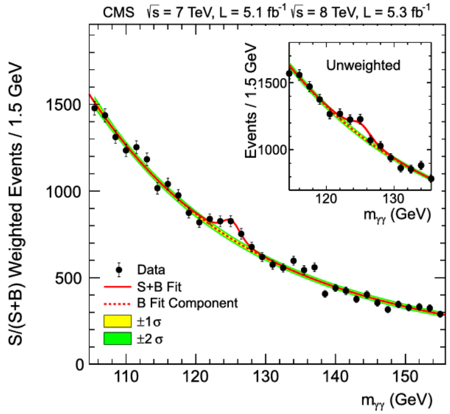
Introduction

- The discovery of the Higgs boson by the ATLAS and CMS collaborations at the Large Hadron Collider (LHC) in July 2012 marked a breakthrough in particle physics, providing deeper insights into the Standard Model (SM).

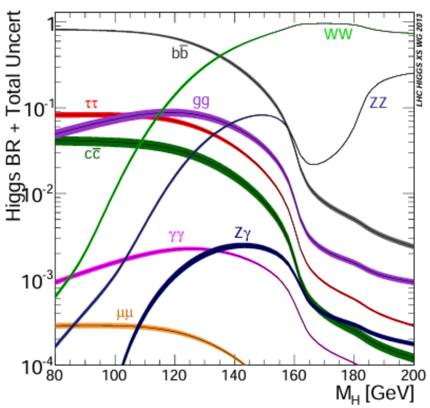
- According to theoretical predictions, the branching fractions for the decay of a 125 GeV Higgs boson into $b\bar{b}$, $c\bar{c}$, gg, WW^*, ZZ^* are 57.7%, 2.91%, 8.57%, 21.5% and 2.64%, respectively.



Physics Letters B 716, 1–29 (2012)



Physics Letters B 716, 30 (2012)

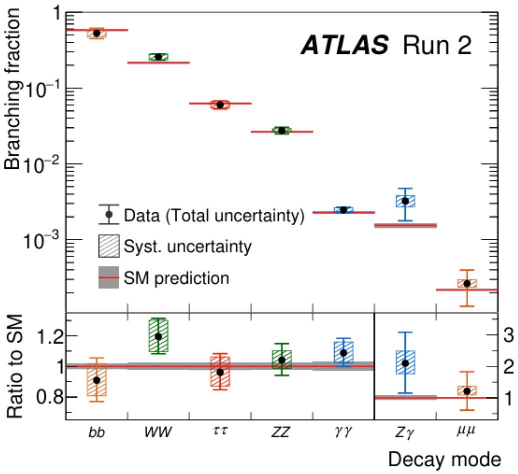


arXiv:1307.1347

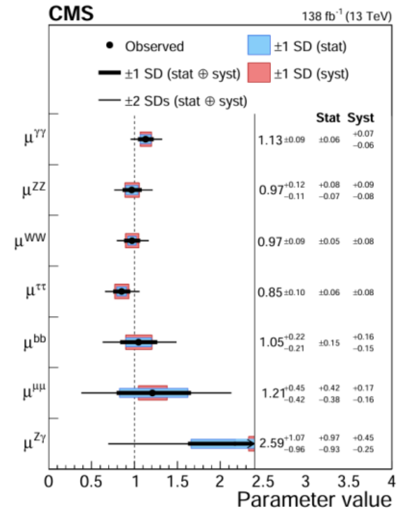
Introduction

- The branching fractions of $H \to b\bar{b}/WW^{*}/ZZ^{*}$ were measured by the ATLAS Collaboration using 139 fb $^{-1}$ of pp collision data at center-of-mass energy of 13 TeV in LHC to be 0.53 ± 0.08, $0.257^{+0.026}_{-0.024}$, 0.028 ± 0.003, respectively.

- The signal strengths of $H \to b\bar{b}/WW^{*}/ZZ^{*}$ were measured by the CMS Collaboration using 138 fb $^{-1}$ of pp collision data at center-of-mass energy of 13 TeV in LHC to be $1.05^{+0.22}_{-0.21}$, 0.97 ± 0.09, $0.97^{+0.12}_{-0.11}$, respectively.



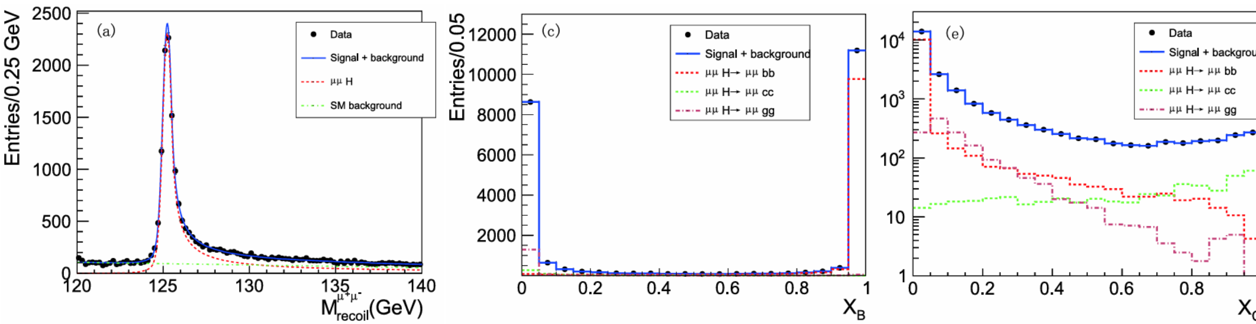
Nature 607, 52 (2022)



Nature 607, 60 (2022)

Introduction

- Previous study performed measurements of decay branch fractions of $H \rightarrow b\bar{b}/c\bar{c}/gg$ in associated $(e^+e^-/\mu^+\mu^-)H$ production at the CEPC with a center-of-mass energy of 250 GeV and integrated luminosity of 5000 fb $^{-1}$.



Higgs boson production	$\mu^+\mu^-H$			e^+e^-H		
Higgs boson decay	$H\rightarrow b\bar{b}$	$H\rightarrow c\bar{c}$	$H\rightarrow gg$	$H\rightarrow b\bar{b}$	$H\rightarrow c\bar{c}$	$H\rightarrow gg$
statistic uncertainty	1.1%	10.5%	5.4%	1.6%	14.7%	10.5%

- b-jet and c-jet tagging training models for calculating X_B and X_C
- Combined fit with M_{recoil}^{ll}

Chinese Phys. C 44 013001

- This study focuses on the determination of the branching fractions of $H \rightarrow b\bar{b}/c\bar{c}/gg/WW^*/ZZ^*$ in associated $Z(\mu^+\mu^-)H$ production at the CEPC with a center-of-mass energy of 240 GeV and integrated luminosity of 5600 fb $^{-1}$. The Particle Flow Network is applied to separate all decay channels simultaneously with high accuracy.

arXiv:2410.04465

Simulation samples

- Using Whizard 1.95 and Pythia6 for the fragmentation and hadronization.
- Using a Delphes-based software suite for fast detector simulation.
- Signal process: Z decays to a pair of muons and H decays in pairs of $b\bar{b}/c\bar{c}/gg/WW^*/ZZ^*$.
- Backgrounds: processes with two-fermion and four-fermion final states.
- Generated to the expected yields in data with an integrated luminosity of 5600 fb $^{-1}$.

Signal process

Process	Higgs decays	Cross section/fb
ZH process	$H \to b\bar{b}$	3.91
	$H \to c\bar{c}$	0.20
	$H \to gg$	0.58
	$H \to WW^*$	1.46
	$H \to ZZ^*$	0.18

Two-fermion background process

Category	Name	Decay modes	Cross section/fb
Two-fermion background	$l\bar{l}$	$e^+e^- \to e^+e^-$	24992.21
		$e^+e^- \to \mu^+\mu^-$	4991.91
		$e^+e^- \to \tau^+\tau^-$	4432.18
	$\nu\bar{\nu}$	$e^+e^- \to \nu_e\bar{\nu}_e$	45390.79
		$e^+e^- \to \nu_\mu\bar{\nu}_\mu$	4416.30
		$e^+e^- \to \nu_\tau\bar{\nu}_\tau$	4410.26
	$q\bar{q}$	$e^+e^- \to u\bar{u}$	10110.43
		$e^+e^- \to d\bar{d}$	10010.07
		$e^+e^- \to c\bar{c}$	10102.75
		$e^+e^- \to s\bar{s}$	9924.40
		$e^+e^- \to b\bar{b}$	9957.70

Event selection

- At least two muons with opposite charge
- Isolation cut: $E_{\text{cone}}^2 < 4E_\mu + 12.2\text{GeV}$, where E_{cone} is the sum of energy within a cone ($\cos\theta_{\text{cone}} > 0.98$) around the muon
- $M_{\mu\mu}$ in Z-mass window [75 GeV, 105 GeV], if there are more than two muons, choose the muon pair closest to the Z boson mass
- $M_{\mu\mu}^{\text{recoil}}$ in H-mass window [110 GeV, 150 GeV] $M_{\mu\mu}^{\text{recoil}} = \sqrt{(\sqrt{s} - E_{\mu^+} - E_{\mu^-})^2 - (\overrightarrow{P_{\mu^+}} + \overrightarrow{P_{\mu^-}})^2}$
- $|\cos\theta_{\mu^+\mu^-}| < 0.996$: to further reduce the two-fermion backgrounds

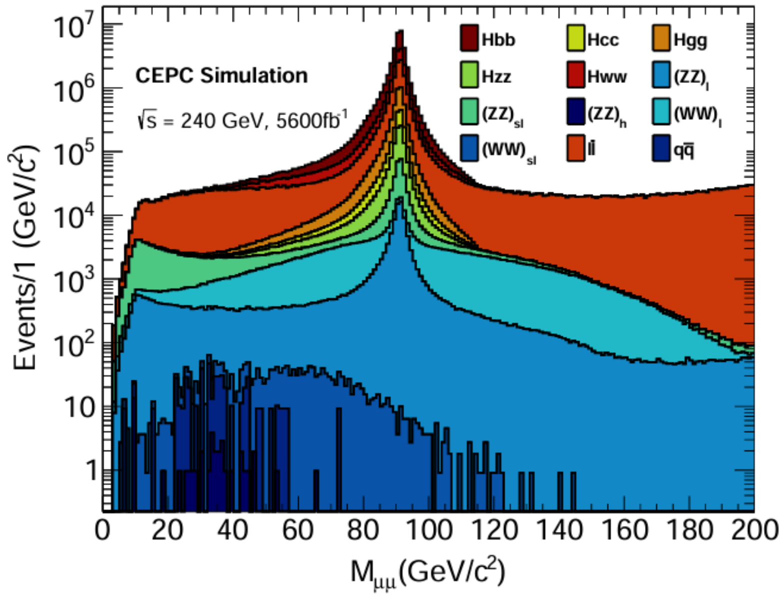
The cutflow selection efficiency

	$H \to b\bar{b}$	$H \to c\bar{c}$	$H \to gg$	$H \to WW^*$	$H \to ZZ^*$
Muon pair	94.45%	94.24%	94.17%	94.91%	94.43%
Isolation	91.52%	92.81%	93.37%	93.83%	94.04%
Z-mass window	95.34%	95.46%	95.49%	91.89%	94.29%
H-mass window	99.73%	99.74%	99.74%	99.05%	99.46%
$\lvert\cos\theta_{\mu^+\mu^-}\rvert < 0.996$	99.65%	99.65%	99.67%	99.65%	99.64%
Total efficiency	81.90%	82.98%	83.46%	80.77%	82.98%

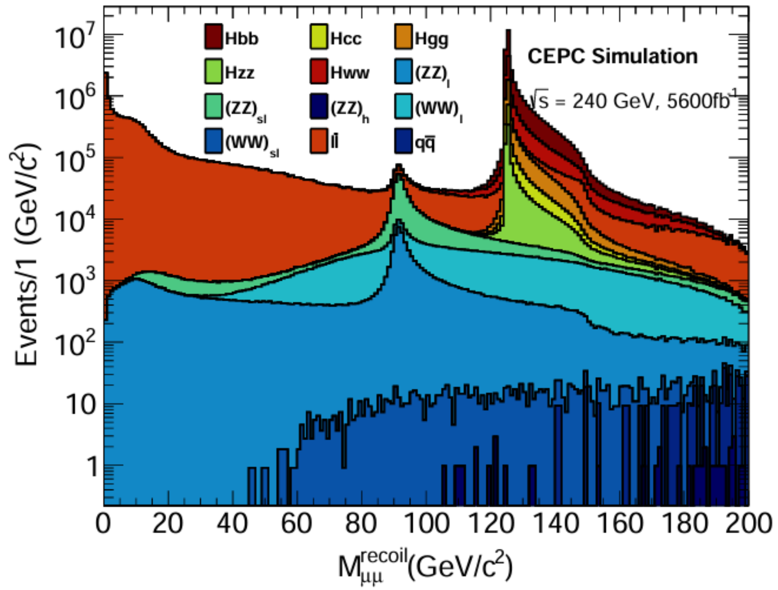
	$l\bar{l}$	$\nu\bar{\nu}$	$q\bar{q}$	$(ZZ)_h$	$(ZZ)_l$	$(ZZ)_{sl}$	$(WW)_h$	$(WW)_l$	$(WW)_{sl}$
Muon pair	11.95%	0	0.05%	0.08%	46.20%	18.91%	0.00%	11.02%	0.16%
Isolation	91.63%	0	0.40%	2.60%	74.16%	66.47%	0	96.46%	3.75%
Z-mass window	41.28%	0	0	0	66.29%	70.41%	0	30.86%	16.67%
H-mass window	6.42%	0	0	0	14.22%	14.80%	0	57.32%	0.35%
$\lvert\cos\theta_{\mu^+\mu^-}\rvert < 0.996$	93.10%	0	0	0	99.07%	99.24%	0	98.40%	98.94%
Total efficiency	0.27%	0.00%	0.00%	0.00%	3.20%	1.30%	0.00%	1.85%	0.00%

Event selection

- $M_{\mu\mu}$ and $M_{\mu\mu}^{\text{recoil}}$ distributions for signal and background events, following the muon pair and isolation selection criteria.



The invariant mass distributions of the muon pair

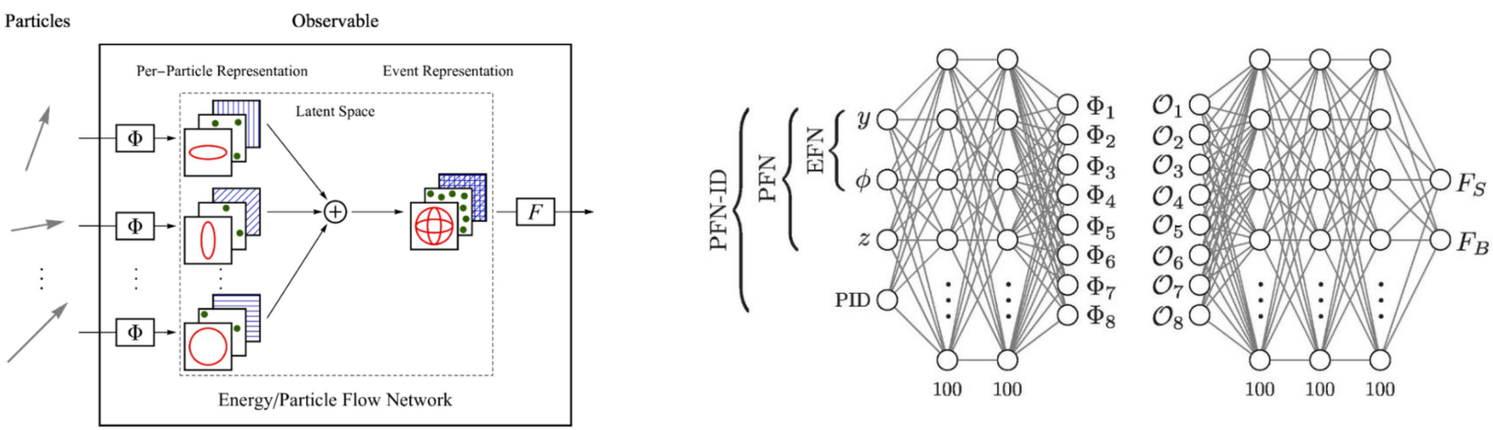


The invariant mass distributions of the muon pair recoil system

- The signal is well preserved while background contributions are significantly suppressed.

Particle Flow Networks

- An end-to-end learning approach, which eliminates the dependency on jet clustering and e/γ isolation.
- Defines a mapping for encoding events $F(\sum_i \Phi(p_i))$, where p represents particle features, and $\Phi(p)$ is a latent space representation of those features. The function F maps the encoded representations to the network's output.
- The architecture of the PFN model is defined by the number of layers and neurons within both F and Φ.



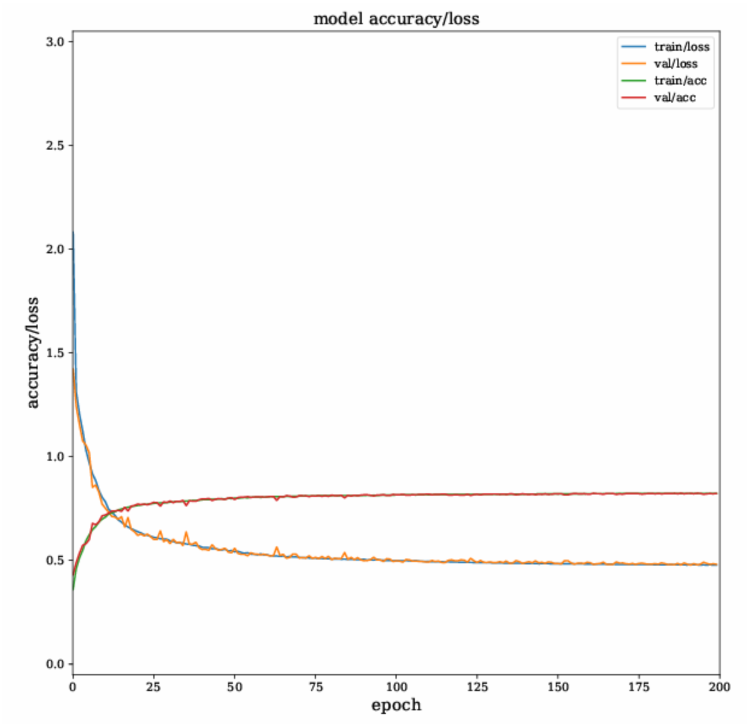
The PFN architecture

Particle Flow Networks

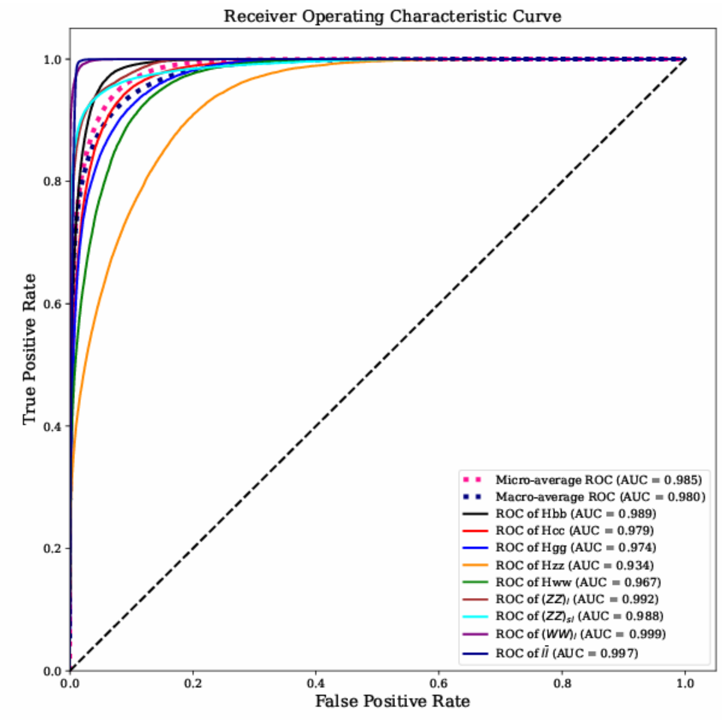
- Samples：
  - 300k for each category, (training: validation: test sets) = (8:1:1)
  - Signal: $H \to b\bar{b}$、$H \to c\bar{c}$、$H \to gg$、$H \to WW^*$、$H \to ZZ^*$
  - Background: $(ZZ)_l$、$(ZZ)_{sl}$、$(WW)_l$、$l\bar{l}$
- Training variables：
  - Energy, momentum, $\cos\theta$, ϕ, PID, D_0, Z_0
- Training parameters：
  - Φ_sizes: (64, 64, 50), F_sizes: (64, 64, 40)
  - Fully connected layer: ReLU activation function and adam optimizer
  - Output layer: SoftMax activation function
  - loss function: cross-entropy
  - Epoch: 200, Learning rate: 0.001, Batchsize: 1000

Model performance

The Loss-accuracy vs epochs curves



ROC curves for signal and background processes

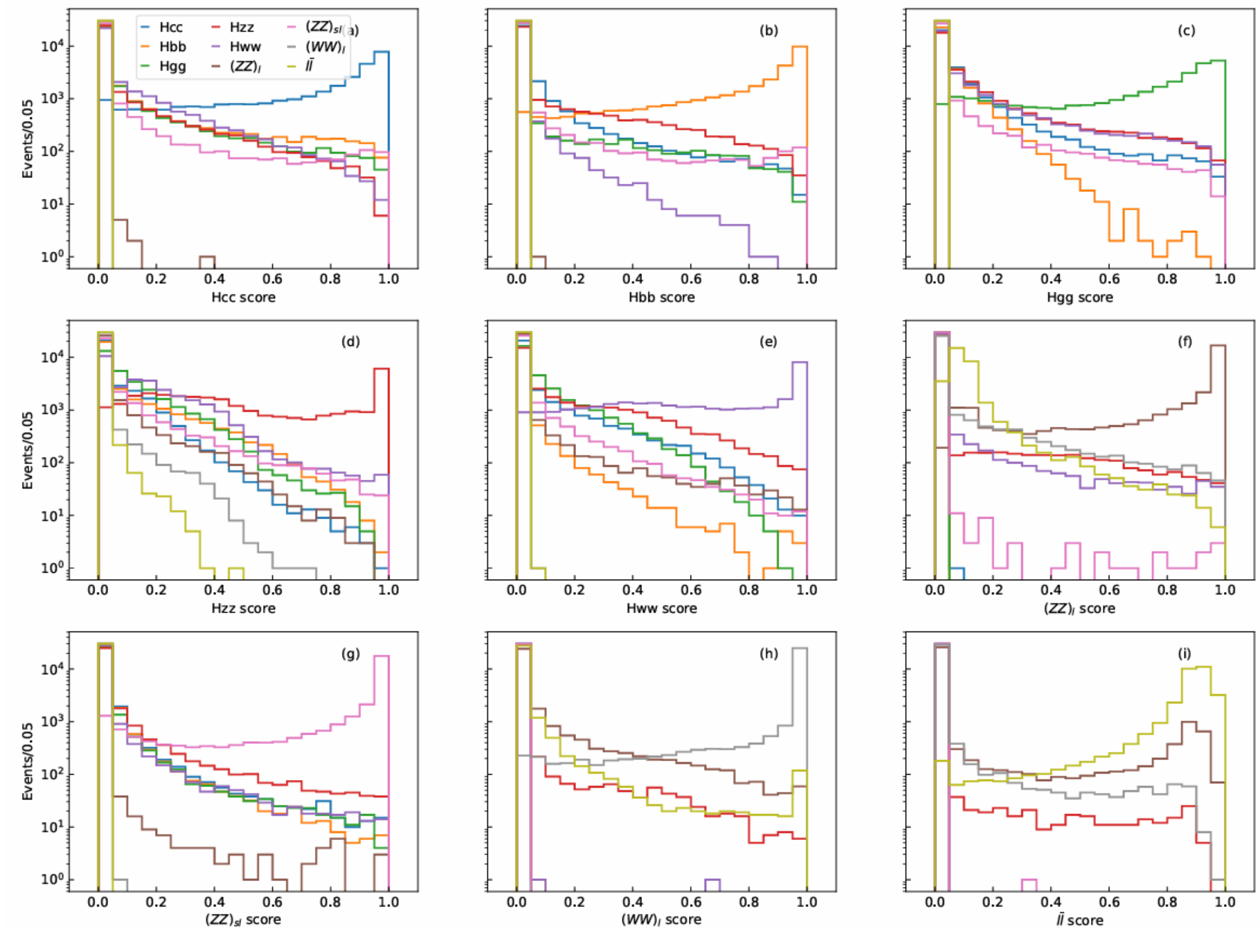


- the loss and accuracy curves converge
  - training and validation set curves overlap highly
  - the model has strong generalization capabilities

- AUC value for each class is above 0.93
  - strong classification performance
  - the model effectively distinguish between classes

Model performance

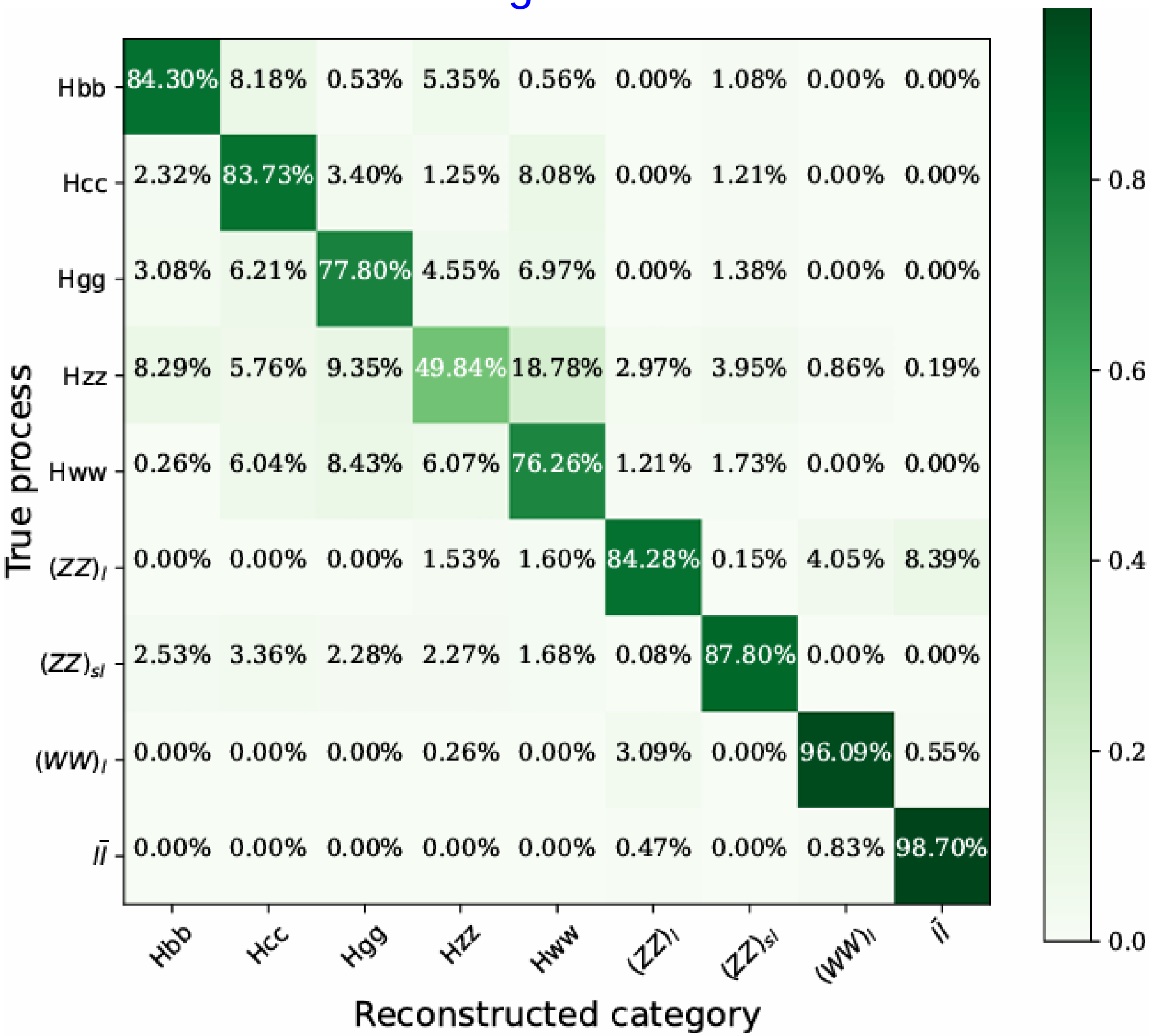
The distributions of classifier outputs for nine categories



- In the region where the score exceeds 0.8, very few events originate from other processes
- The classification is pretty good

Model performance

The migration matrix



- The reconstructed category refers to the process with the highest score for a given event
- The purity of each category is above 76% except for $H \rightarrow ZZ^*$ process
- The migration matrix reflects the overall high accuracy of the model

Measurements of branching fractions

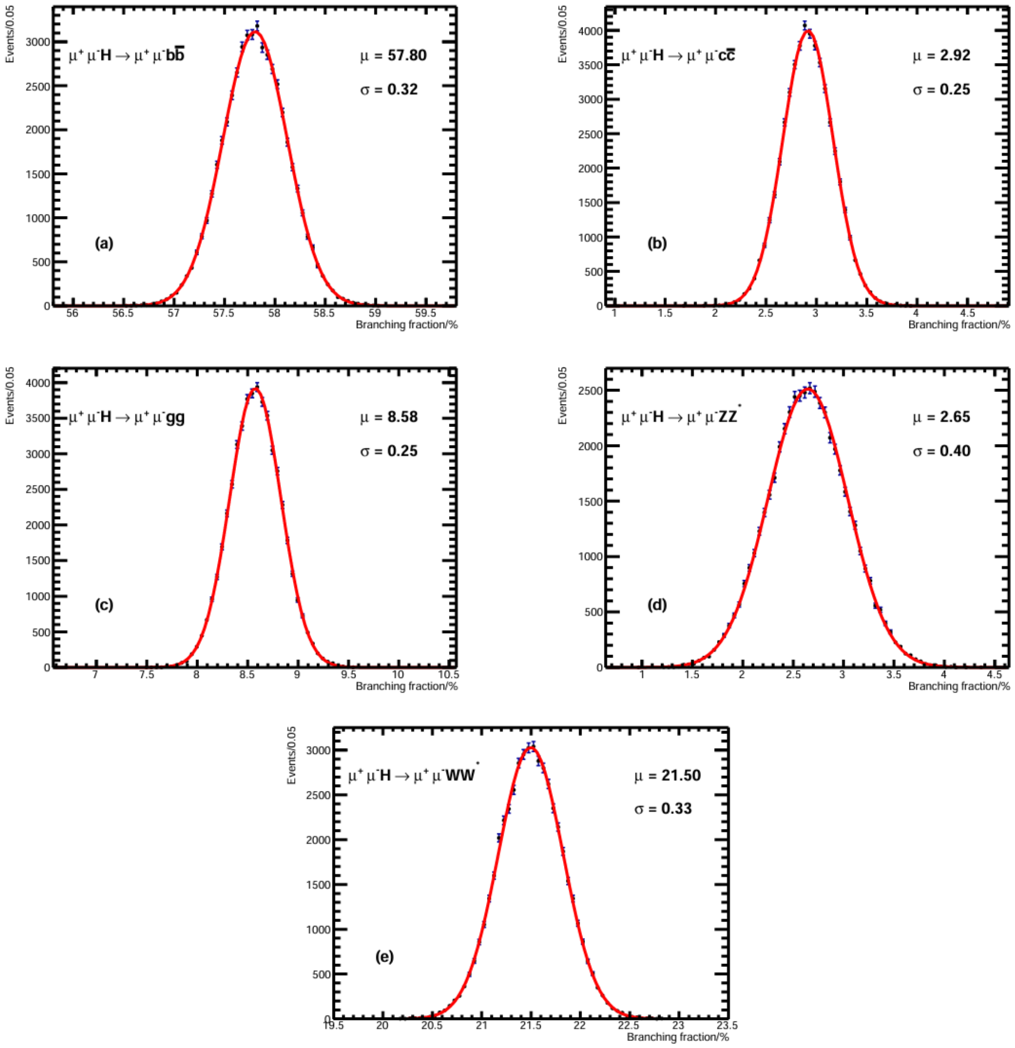
- Use the migration matrix method
  - Can be unfolded to represent the generated number of signals
  - Calculated as follows:

$$\begin{bmatrix} N_{s1} \\ N_{s2} \\ \cdots \\ N_{b1} \\ N_{b2} \\ \cdots \end{bmatrix} = \left(M_{mig}^{T} M_s\right)^{-1} \times \begin{bmatrix} n_{s1} \\ n_{s2} \\ \cdots \\ n_{b1} \\ n_{b2} \\ \cdots \end{bmatrix}$$

  - n_i and N_i are the expected and generated number of events of class i, where n_i is obtained from simulation samples processed by the PFN model
  - M_s is a diagonal matrix containing the selection efficiencies, M_{mig}^{T} denotes the transposed migration matrix

Measurements of branching fractions

ToyMC test result



- Use toyMC method to estimate statistical uncertainties
  - Poisson distribution according to the number of events
  - Multinomial distribution according to the migration matrix
  - Least square fit for 50k times

$$Q^2 = \sum_{i=1}^{N} \left(\frac{Y_i - \eta_i}{\sigma_i} \right)^2$$

  - Fit with gaussian function

Measurements of branching fractions

- Results of the measured branching fractions with statistical uncertainties:

Higgs boson decay	$H \to b\bar{b}$	$H \to c\bar{c}$	$H \to gg$	$H \to WW^*$	$H \to ZZ^*$
branching fraction	0.5780	0.0292	0.0858	0.2150	0.0265
statistical uncertainty	$\pm 0.58\%$	$\pm 8.41\%$	$\pm 2.99\%$	$\pm 2.32\%$	$\pm 9.81\%$

- To account for the systematic uncertainty, the resolution of the detector was adjusted by increasing it by 2%.

$$\sigma_{\frac{1}{p_T}} = 2 \times 10^{-5} \oplus \frac{1 \times 10^{-3}}{p \sin^{3/2} \theta} \mathrm{GeV}^{-1}$$

- The systematic uncertainties for the branching fractions are estimated to be 1.29%, 9.69%, 0.67%, 2.61% and 1.11% for $b\bar{b}/c\bar{c}/gg$ and WW^*/ZZ^* final states, respectively.

Summary

- The Higgs boson branching fractions into $b\bar{b}/c\bar{c}/gg$ and WW^*/ZZ^*, where the W or Z bosons decay hadronically, via the $Z(\mu^+\mu^-)H$ process are studied using the PFN method at a center-of-mass energy of 240 GeV and a luminosity of 5600 fb $^{-1}$ at the CEPC.
- The statistical uncertainty of branching fractions of $H \rightarrow b\bar{b}/c\bar{c}/gg$/$WW^*$/ ZZ^* processes are estimated to be 0.58%, 8.41%, 2.99%, 2.32% and 9.81%, respectively.
- Compared to a previous analysis which reported statistical uncertainties of 1.1%, 10.5% and 5.4% for the branching fractions of $H \rightarrow b\bar{b}/c\bar{c}/gg$ process, the PFN method achieves higher precision in a single execution, due to its better performance and deeper data exploitation.