

# Global Analysis of Higgs Aided by Machine Learning

Quantum Computing & Machine Learning Workshop,  
August 11-14, 2023,  
Shandong University

Shudong Wang, Gang Li

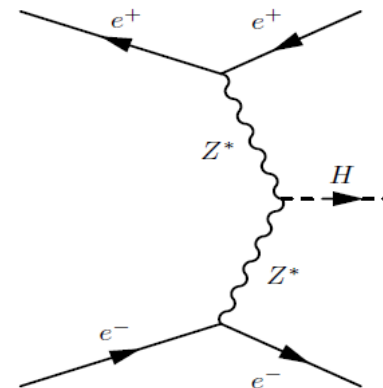
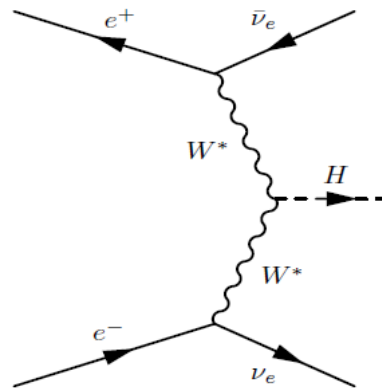
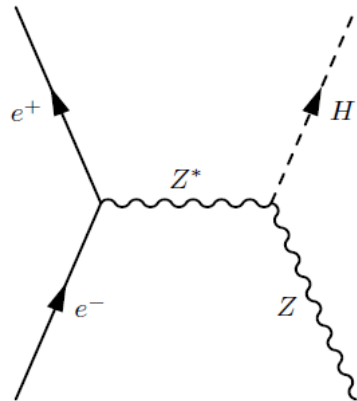
Institute of High Energy Physics, CAS

# Outline

- Introduction
- Determining efficiency (confusion) matrix with ML
- Preliminary results
- Summary

# Introduction

- Higgs decay branching ratios at future Higgs factories can be measured by directly using counting method.
- Events can be categorized into classes using event properties linked to the expected Higgs decay modes. The counts per class are used to fit the Higgs branching ratios in a model independent way.



# Introduction

Personal ranking of the difficulty of Higgs analysis at e-e+ colliders

4 x 9 modes in this study, [ 5 production and 13 (9) decay modes in SM ]

Prod/decay	cc	bb	$\mu\mu$	$\tau\tau$	$\gamma\gamma$	gg	WW	ZZ	$\gamma Z$	ee, uu, dd, ss
eeH (incl. Z fusion)	3	1	5	2	4	1	2	3	5	Not covered yet
$\mu\mu$ H	3	1	5	2	4	1	2	3	5	
$\tau\tau$ H	3	1	5	2	4	1	2	3	5	
qqH	4	1	2	1	2	5	5	5	3	
$\nu\nu$ H (incl. W fusion)	5	1	3	2	3	5	4	2	4	

According to production rate, signal signature, backgrounds, complexity of analysis, ...

# Introduction

## Current estimation of Higgs precision

CEPC: 2205.08553

FCC-ee

	240 GeV, 20 $ab^{-1}$		360 GeV, 1 $ab^{-1}$		
	ZH	vvH	ZH	vvH	eeH
any	<b>0.26%</b>		<b>1.40%</b>	\	\
H→bb	<b>0.14%</b>	<b>1.59%</b>	<b>0.90%</b>	<b>1.10%</b>	<b>4.30%</b>
H→cc	<b>2.02%</b>		<b>8.80%</b>	<b>16%</b>	<b>20%</b>
H→gg	<b>0.81%</b>		<b>3.40%</b>	<b>4.50%</b>	<b>12%</b>
H→WW	<b>0.53%</b>		<b>2.80%</b>	<b>4.40%</b>	<b>6.50%</b>
H→ZZ	<b>4.17%</b>		<b>20%</b>	<b>21%</b>	
$H \rightarrow \tau\tau$	<b>0.42%</b>		<b>2.10%</b>	<b>4.20%</b>	<b>7.50%</b>
$H \rightarrow \gamma\gamma$	<b>3.02%</b>		<b>11%</b>	<b>16%</b>	
$H \rightarrow \mu\mu$	<b>6.36%</b>		<b>41%</b>	<b>57%</b>	
$Br_{upper}(H \rightarrow inv.)$	<b>0.07%</b>		\	\	
$H \rightarrow Z\gamma$	<b>8.50%</b>		<b>35%</b>	\	
Width	<b>1.65%</b>		<b>1.10%</b>		

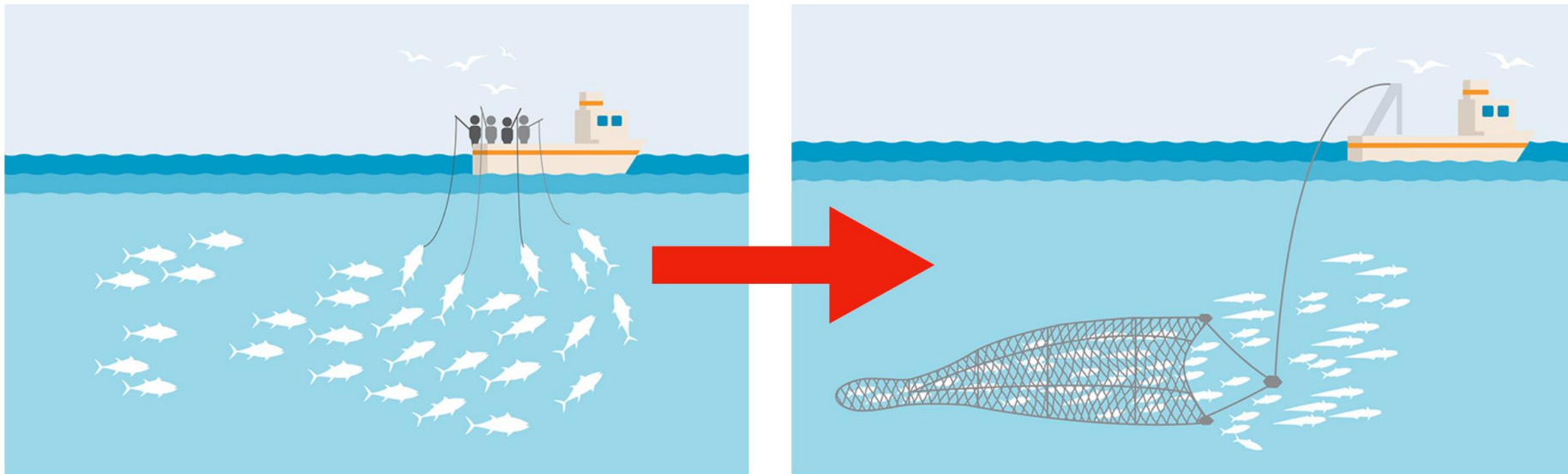
$\sqrt{s}$ (GeV)	240		365	
Luminosity ( $ab^{-1}$ )	5		1.5	
$\delta(\sigma BR)/\sigma BR$ (%)	HZ	$\nu\bar{\nu} H$	HZ	$\nu\bar{\nu} H$
H → any	±0.5		±0.9	
H → $b\bar{b}$	±0.3	±3.1	±0.5	±0.9
H → $c\bar{c}$	±2.2		±6.5	±10
H → gg	±1.9		±3.5	±4.5
H → $W^+W^-$	±1.2		±2.6	±3.0
H → ZZ	±4.4		±12	±10
H → $\tau\tau$	±0.9		±1.8	±8
H → $\gamma\gamma$	±9.0		±18	±22
H → $\mu^+\mu^-$	±19		±40	
H → invisible	< 0.3		< 0.6	

- Results of CEPC and FCC-ee based individual analysis
- Comparable precision

**A lot of efforts**

# Introduction

- Signal of an individual analysis is the background of another and vice versa.
- Combine efforts in a global analysis style?



With the help of ML/DL techniques

# Introduction

- Individual style
- $N_s$ : signal
- $N_b$ : backgrounds, could be different types, higgs crosstalks and higgs-free events
- $Br_s$  to be measured
- Events selection:

$$Br_s = \frac{N_s}{N_s + N_{b_{higgs}}}$$

$$\begin{pmatrix} n_s \\ n_b \end{pmatrix} = \begin{pmatrix} \epsilon_{ss} & \epsilon_{sb} \\ \epsilon_{bs} & \epsilon_{bb} \end{pmatrix} \times \begin{pmatrix} N_s \\ N_b \end{pmatrix}$$

- Global style

$$\mathbf{n} = \mathbf{E}\mathbf{N}$$

$$\begin{pmatrix} n_1 \\ n_2 \\ \vdots \\ n_9 \\ n_{b_1} \\ n_{b_2} \\ \vdots \end{pmatrix} = \begin{pmatrix} \epsilon_{11} & \epsilon_{12} & \cdots & \epsilon_{19} \\ \epsilon_{21} & \epsilon_{22} & \cdots & \epsilon_{29} \\ \vdots & \vdots & \ddots & \vdots \\ \epsilon_{91} & \epsilon_{92} & \cdots & \epsilon_{99} \\ \vdots & \cdots & & \vdots \\ \vdots & \ddots & & \vdots \\ \vdots & \cdots & & \vdots \end{pmatrix} \begin{pmatrix} N_1 \\ N_2 \\ \vdots \\ N_9 \\ N_{b_1} \\ N_{b_2} \\ \vdots \end{pmatrix}$$

$$Br_i = \frac{N_i}{N_1 + N_2 + \cdots + N_9}$$

A “one-stop” approach:  
measure all branching ratios simultaneously,  
more efficient and hopefully better precision

# Introduction

Key ingredient: efficiency matrix ( $E$ ) also challenging

$$\begin{pmatrix} \epsilon_{11} & \epsilon_{12} & \cdots & \epsilon_{19} \\ \epsilon_{21} & \epsilon_{22} & \cdots & \epsilon_{29} \\ \vdots & \vdots & \ddots & \vdots \\ \epsilon_{91} & \epsilon_{92} & \cdots & \epsilon_{99} \end{pmatrix}$$

**A multiple classification problem**

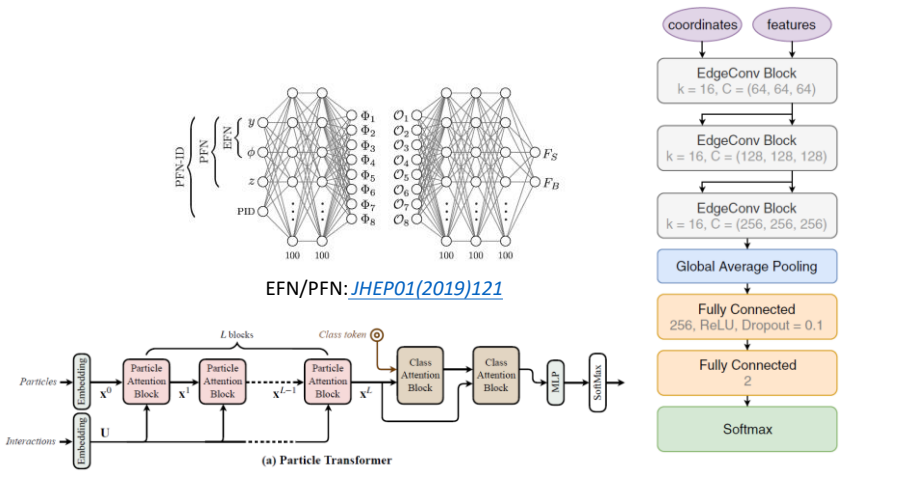
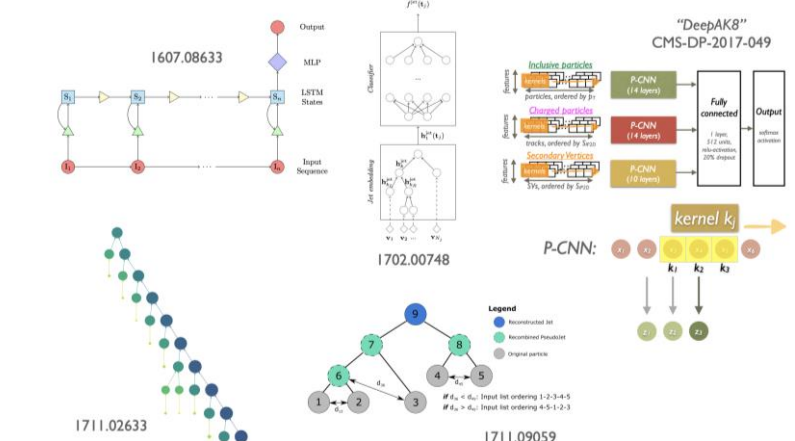
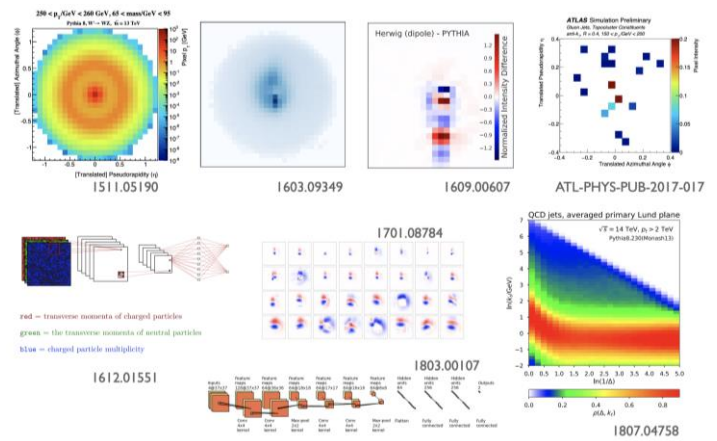
**In 2020s, try machine learning/deep learning?**



# Determining efficiency (confusion) matrix with ML

- From jet tagging to event tagging
- Jets are important physics objects, >90% of ZH events contain jets
- Jet tagging with BDT: feature engineering
  - Lots of works
- Novel ML/DL algorithms: jets treated as images, sequences, sets/point clouds
  - Performance boost, less human labor

nvtx=0	trk1d0sig trk2d0sig trk1z0sig trk2z0sig trk1pt_jete trk2pt_jete jprobr5sigma jprobz5sigma d0bprob d0cprob d0qprob z0bprob z0cprob z0qprob nmuon nelectron trkmass(17)
nvtx=1&&nvtxall=1	trk1d0sig trk2d0sig trk1z0sig trk2z0sig trk1pt_jete trk2pt_jete jprobr jprobz vtxlen1_jete vtxsig1_jete vtxdirang1_jete vtxmom1_jete vtxmass1 vtxmult1 vtxmasspc vtxprob d0bprob d0cprob d0qprob z0bprob z0cprob z0qprob trkmass nelectron nmuon(25)
nvtx=1&&nvtxall=2	trk1d0sig trk2d0sig trk1z0sig trk2z0sig trk1pt_jete trk2pt_jete jprobr jprobz vtxlen1_jete vtxsig1_jete vtxdirang1_jete vtxmom1_jete vtxmass1 vtxmult1 vtxmasspc vtxprob 1vtxprob vtxlen12all_jete vtxmassall(19)
Nvtx>=2	trk1d0sig trk2d0sig trk1z0sig trk2z0sig trk1pt_jete trk2pt_jete jprobr jprobz vtxlen1_jete vtxsig1_jete vtxdirang1_jete vtxmom1_jete vtxmass1 vtxmult1 vtxmasspc vtxprob vtxlen2_jete vtxsig2_jete vtxdirang2_jete vtxmom2_jete vtxmass2 vtxmult2 vtxlen12_jete vtxsig12_jete vtxdirang12_jete vtxmom_jete vtxmass vtxmult 1vtxprob(29)

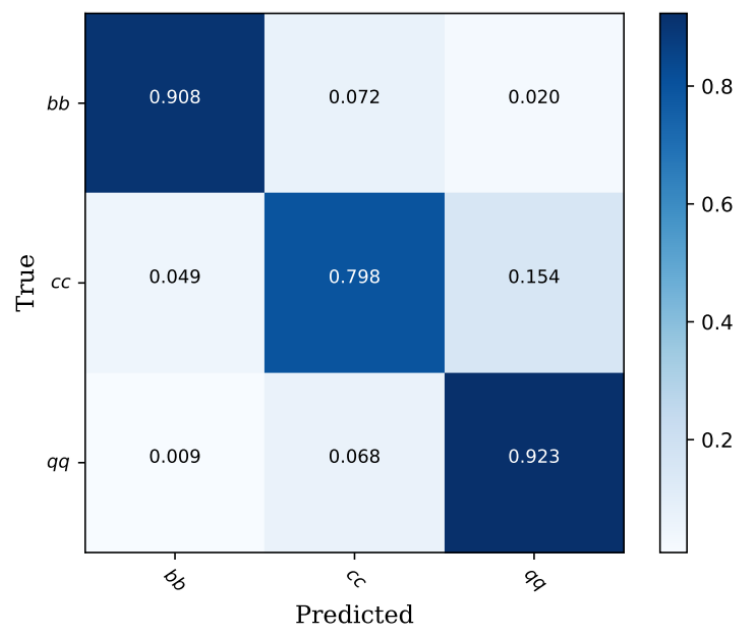


ParticleTransformer:  
[2202.03772](https://arxiv.org/abs/2202.03772)

ParticleNet:  
[Phys.Rev.D 101 \(2020\) 5, 056019](https://arxiv.org/abs/1909.02843)

# Determining efficiency (confusion) matrix with ML

- From jet tagging to event tagging
- Novel ML/DL algorithms:
  - Performance on jet flavor tagging in a realistic scenario (CEPC fullsim) is pretty good



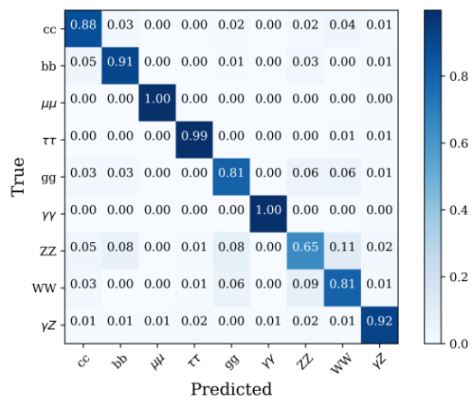
2208.13503

# Determining efficiency (confusion) matrix with ML

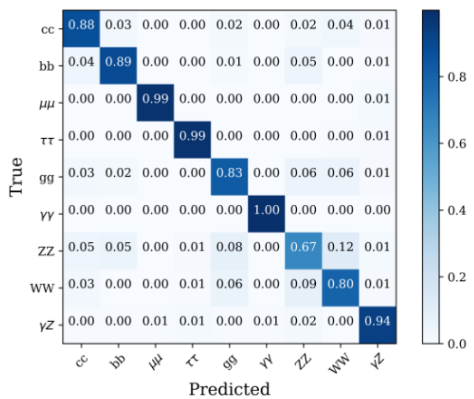
- From jet tagging to event tagging
- Number of particles
  - Jets:  $O(10) \sim O(100)$
  - Events on a ee collider:  $O(10) \sim O(100)$
- Classify events (event tagging) using ML algorithms?
  - Most important objective :  $e^+e^- \rightarrow ZH$
  - 4 Z decay modes  
 $ee, \mu\mu, \tau\tau, q\bar{q}$
  - 9 Higgs decay modes  
 $c\bar{c}, b\bar{b}, \mu^+\mu^-, \tau^+\tau^-, \gamma\gamma, gg, WW^*, ZZ^*, \gamma Z$
  - 36 in total
  - 9 background processes:  $\nu\nu H, l^+l^-, q\bar{q}, W^+W_l^-, W^+W_{sl}^-, W^+W_h^-, ZZ_l, ZZ_{sl}, ZZ_h$
  - 400k events for each class, DELPHES fast simulation
  - Particle level information as input

# Determining efficiency (confusion) matrix with ML

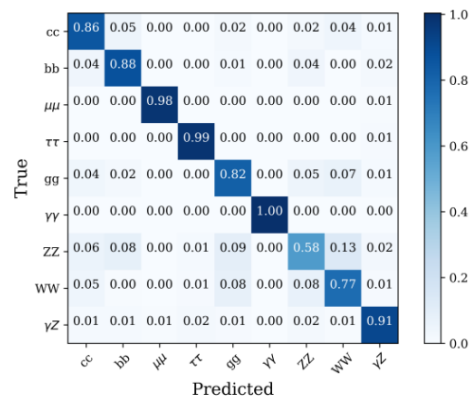
- From jet tagging to event tagging
- First attempt: classify higgs decay (9-category classification) using ParticleNet



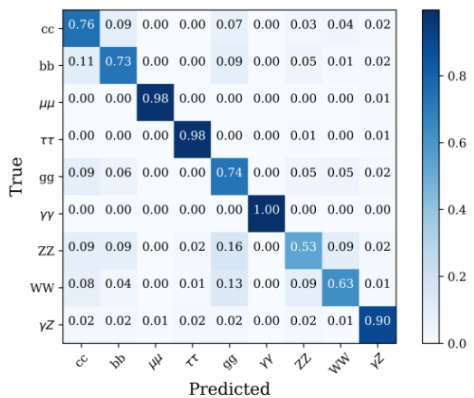
eeH



$\mu\mu H$



$\tau\tau H$



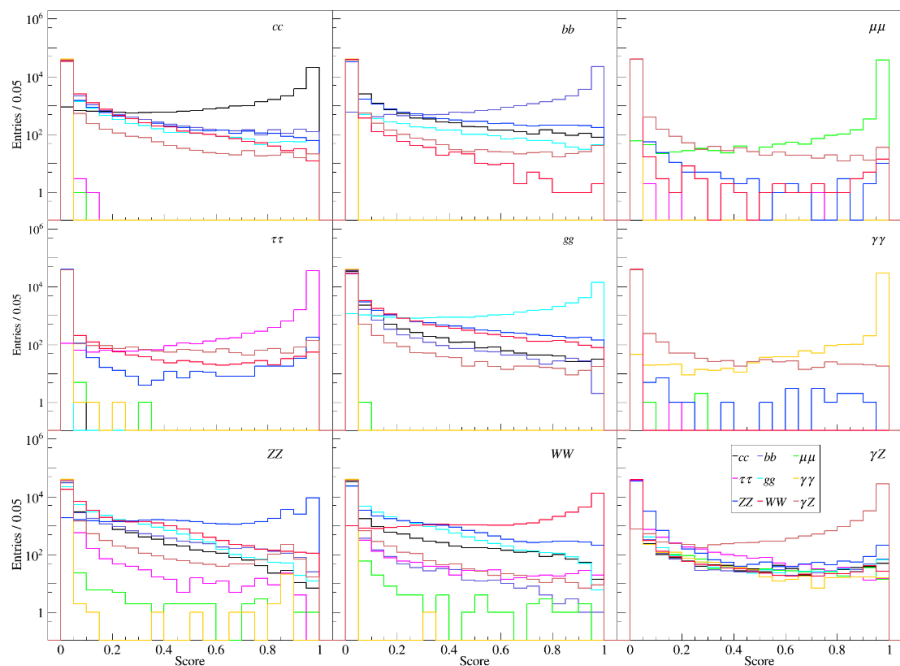
$q\bar{q}H$

- Sufficiently good performance
- Average Accuracy  $\sim 87\%$   
(11% for random guess)

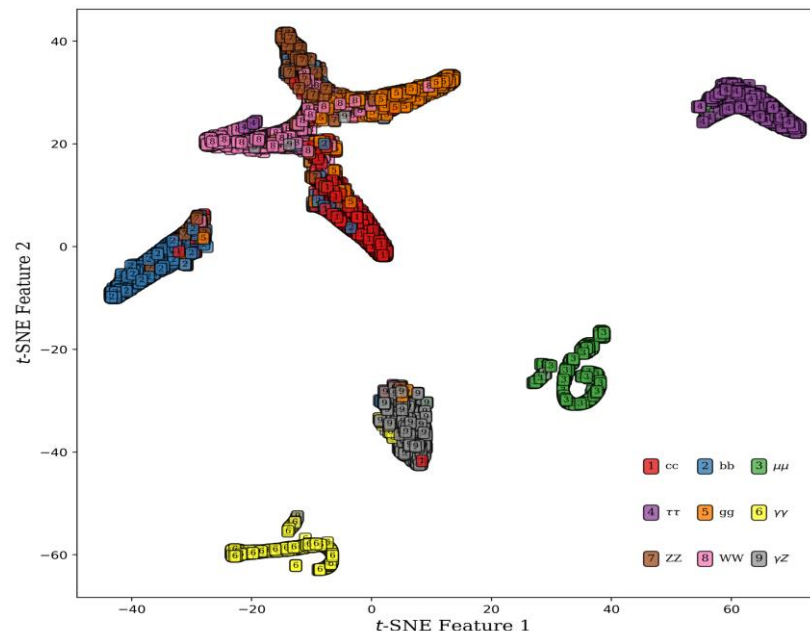
# Determining efficiency (confusion) matrix with ML

- From jet tagging to event tagging
- First attempt: classify higgs decay (9-category classification) using ParticleNet

Distribution of score



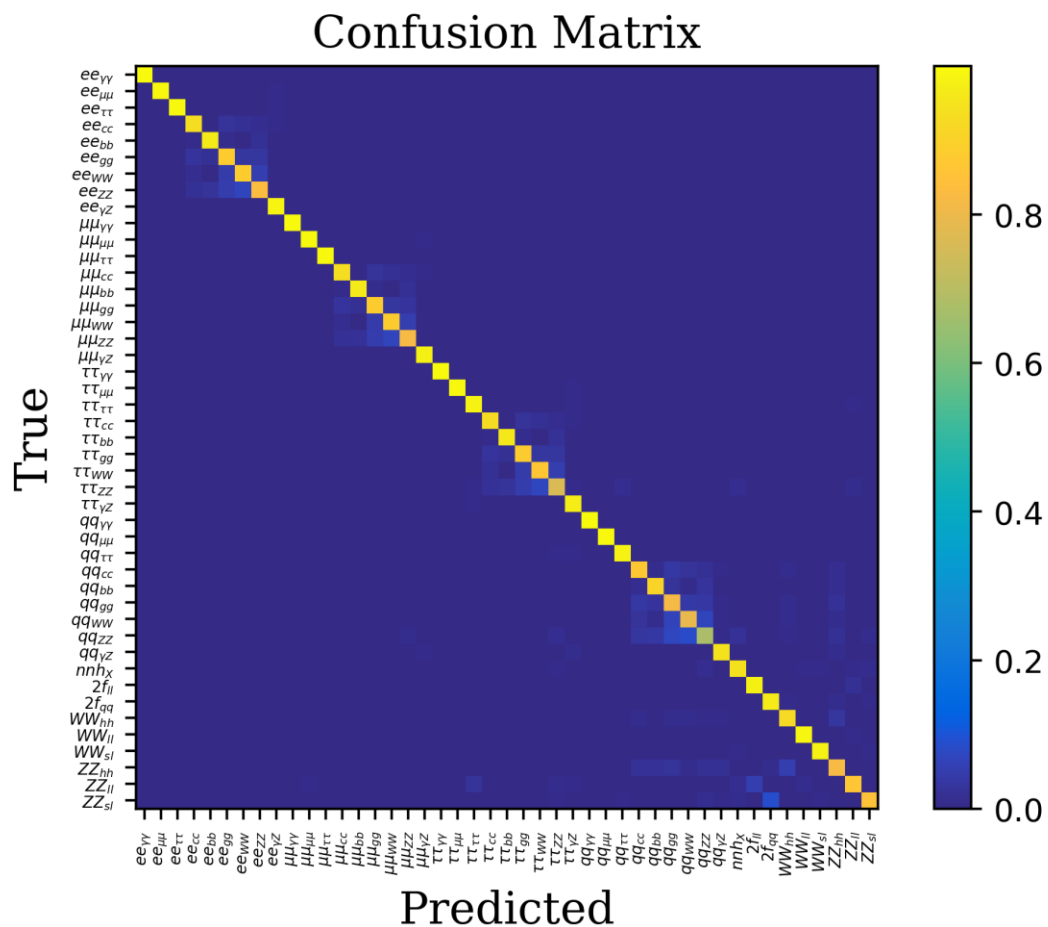
Dimensionality reduction (t-SNE)



- ✓  $\mu\mu$ ,  $\gamma\gamma$ ,  $\tau\tau$  well classified as expected
- ✓  $bb$  and  $\gamma Z$  also good
- ✓  $cc$ ,  $gg$ ,  $WW$ , and  $ZZ$  fake each other, but under control

# Determining efficiency (confusion) matrix with ML

- From jet tagging to event tagging
- More ambitious: classify 36 signals + 9 bkgs using ParticleTransformer

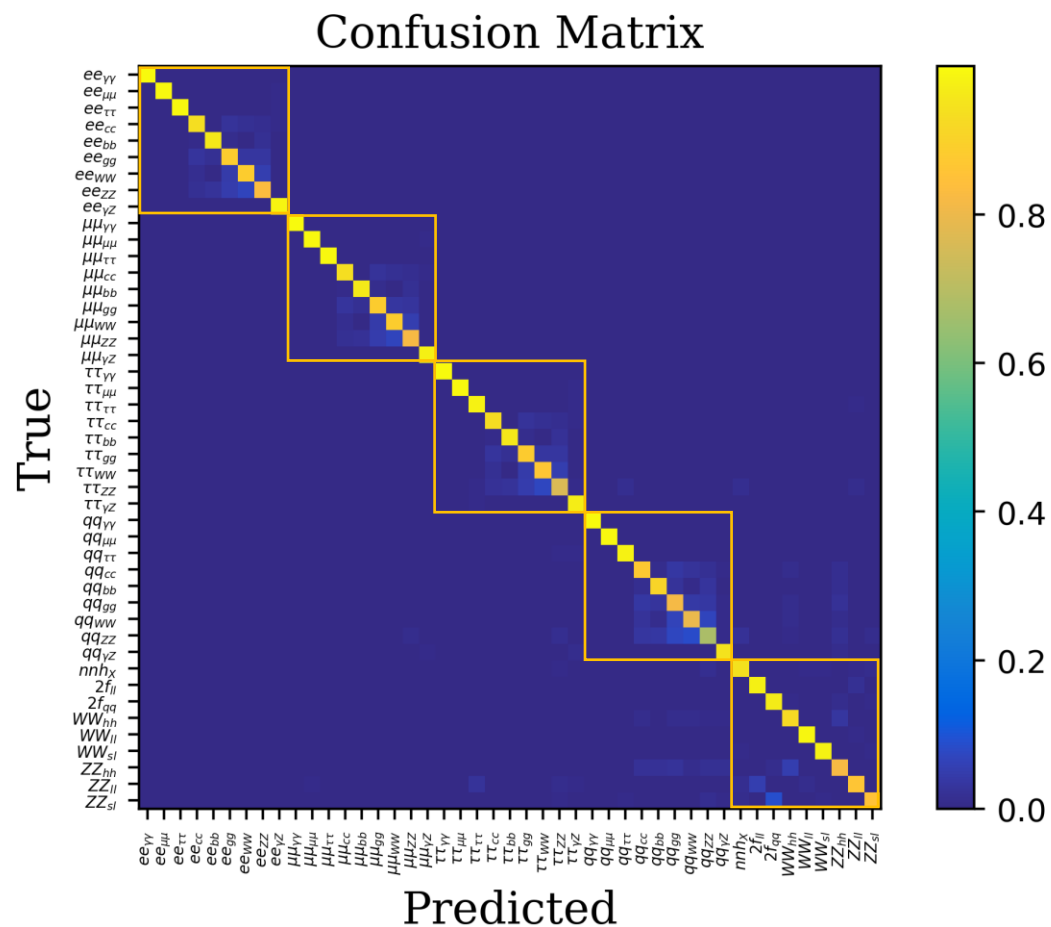


- Average Accuracy  $\sim 92\%$
- Appealing performance
- Use this for global analysis...

1M events for each class

# Determining efficiency (confusion) matrix with ML

- From jet tagging to event tagging
- More ambitious: classify 36 signals + 9 bkgs using ParticleTransformer



- Average Accuracy  $\sim 92\%$
- Appealing performance
- Use this for global analysis...

1M events for each class

# Preliminary results

- Extract  $Br$ s and their uncertainties using toy method.
  - Integrated luminosity:  $20 \text{ ab}^{-1}$

- Minimizing  $\chi^2 = \sum \frac{(N_{obs} - N_{exp})^2}{N_{exp}} + N_{ZH} * (\sum_i Br_i - 1)^2$

(assuming Higgs have and only have the 9 decay modes we studied)

- Hope to improve the uncertainties of Higgs decay branching ratios, but... 😞

Decay Mode	Rel.Unc.
$H \rightarrow c\bar{c}$	1.63%
$H \rightarrow b\bar{b}$	0.16%
$H \rightarrow \mu^+\mu^-$	19.53%
$H \rightarrow \tau^+\tau^-$	0.37%
$H \rightarrow gg$	0.73%
$H \rightarrow \gamma\gamma$	28.45%
$H \rightarrow ZZ$	2.7%
$H \rightarrow W^+W^-$	0.28%
$H \rightarrow \gamma Z$	38.5%



- Uncertainties of some branching ratios are even worse than results in CDR ( $5.6 \text{ ab}^{-1}$ )
- The branching ratios of those decay modes are relatively small
- Contaminated by backgrounds, even if the proportion is small, the cross-section of the backgrounds are large -> large contamination, large uncertainty



# Preliminary results

- Hope to improve the uncertainties of higgs decay branching ratios, but... 😞
- No much pre-selection before events classification.
  - Only require #Particles > 2,  $E_{\text{vis}} > 20$  GeV
  - Too optimistic, too radical? 😬

Decay Mode	Rel.Unc.
$H \rightarrow c\bar{c}$	1.63%
$H \rightarrow b\bar{b}$	0.16%
$H \rightarrow \mu^+\mu^-$	19.53%
$H \rightarrow \tau^+\tau^-$	0.37%
$H \rightarrow gg$	0.73%
$H \rightarrow \gamma\gamma$	28.45%
$H \rightarrow ZZ$	2.7%
$H \rightarrow W^+W^-$	0.28%
$H \rightarrow \gamma Z$	38.5%

- **Next step:** 😬
  - Fall back to a classical style, try to add few more pre-selections.
    - Problem solved, precision improved?
    - or
    - Worsen the performance of classifier? The loss outweighs the gain.

Based on “Shape variables”?

- Thrust

$$T = \frac{\sum_i |\hat{T} \cdot \vec{p}_i|}{\sum_i |\vec{p}_i|}$$

- Sphericity, Aplanarity

$$S^{\alpha,\beta} = \frac{\sum_i p_i^\alpha p_i^\beta}{\sum_i |\vec{p}_i|^2}$$

- Fox-Wolfram moments

$$H_l = \sum_{i,j} \frac{|\vec{p}_i| |\vec{p}_j|}{E^2} P_l(\cos\theta_{ij}),$$

- ...

Still in progress...

# Summary

- Higgs physics is the top priority in future Higgs factories.
- Feasibility study shows that novel ML/DL algorithms could classify events to a very good extent.
- A new analysis paradigm
  - Multi-classification: accurate enough and fast enough.
  - Only particle level information as input, no dependence on jet algorithm, ... less software work.
  - Preliminary results not as good as expected, need more study.
  - Systematics will be very challenging, need more study.