上海交通大学
SHANGHAI JIAO TONG UNIVERSITY

SJTU

Higgs hadronic decay at CEPC

Wang Xinzhu, Zhu Chunxiang, Zhu Yifan
Wang Kun, Yang Haijun

上海交通大学
SHANGHAI JIAO TONG UNIVERSITY

CEPC

饮 水 思 源 • 爱 国 荣 校

TDR draft

- Advanced Tagging algorithm——**ParticleTransformer** deeply used in TDR research

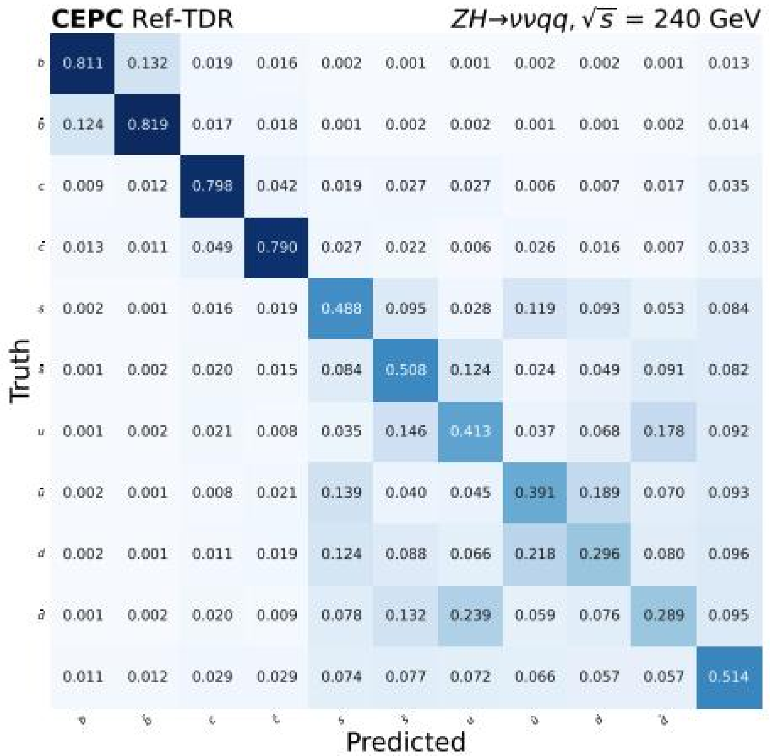


Figure 15.15: The confusion matrix M_{11} of JOI using realistic PID of leptons and charged hadrons for $\nu\bar{\nu}H$, $H \to qq$ events at $\sqrt{s} = 240$ GeV, with the reference detector. The matrix is normalized to unity for each truth label.

Jet Origin Identification (JOI)

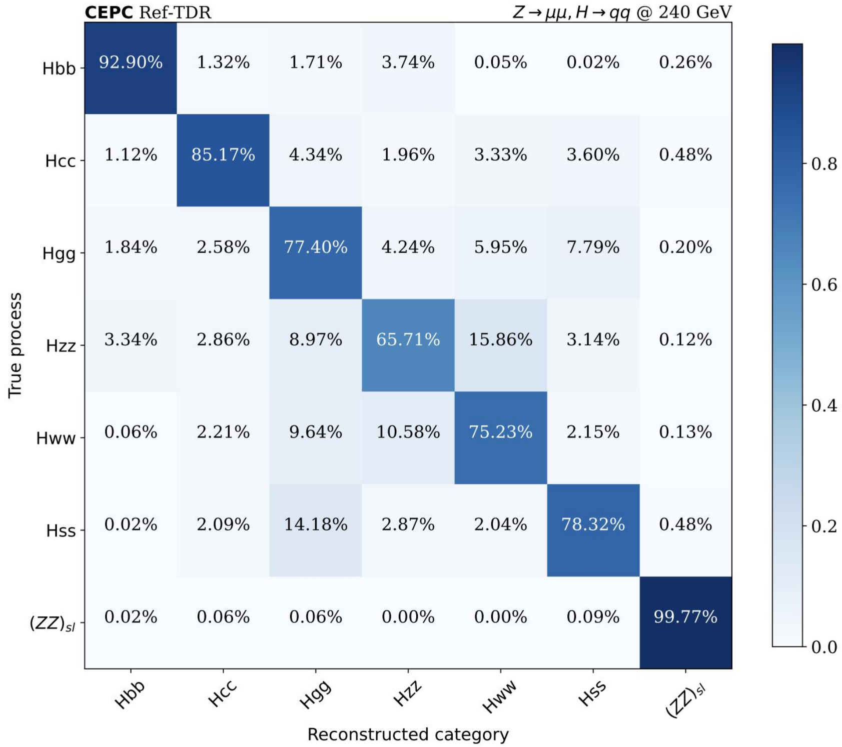
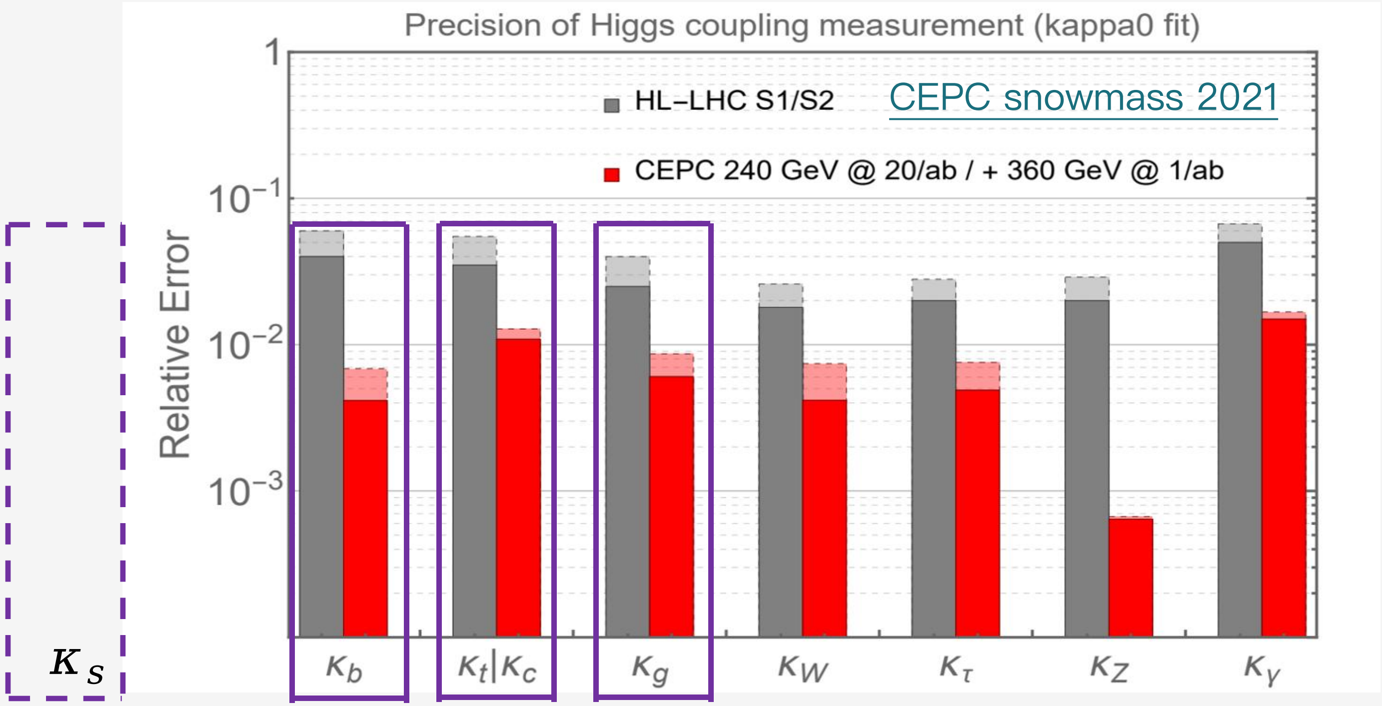


Figure 15.20: The migration matrix for the seven classes is shown. The horizontal axis represents the prediction of the model for each event in the test set, while the vertical axis indicates the true labels. The sum of values in each row equals 1.

Higgs hadronic decay

Motivation

- Better flavor ID enhances sensitivity in rare Higgs decay channels
- JOI –> physics analysis
- Test Higgs couplings across generations



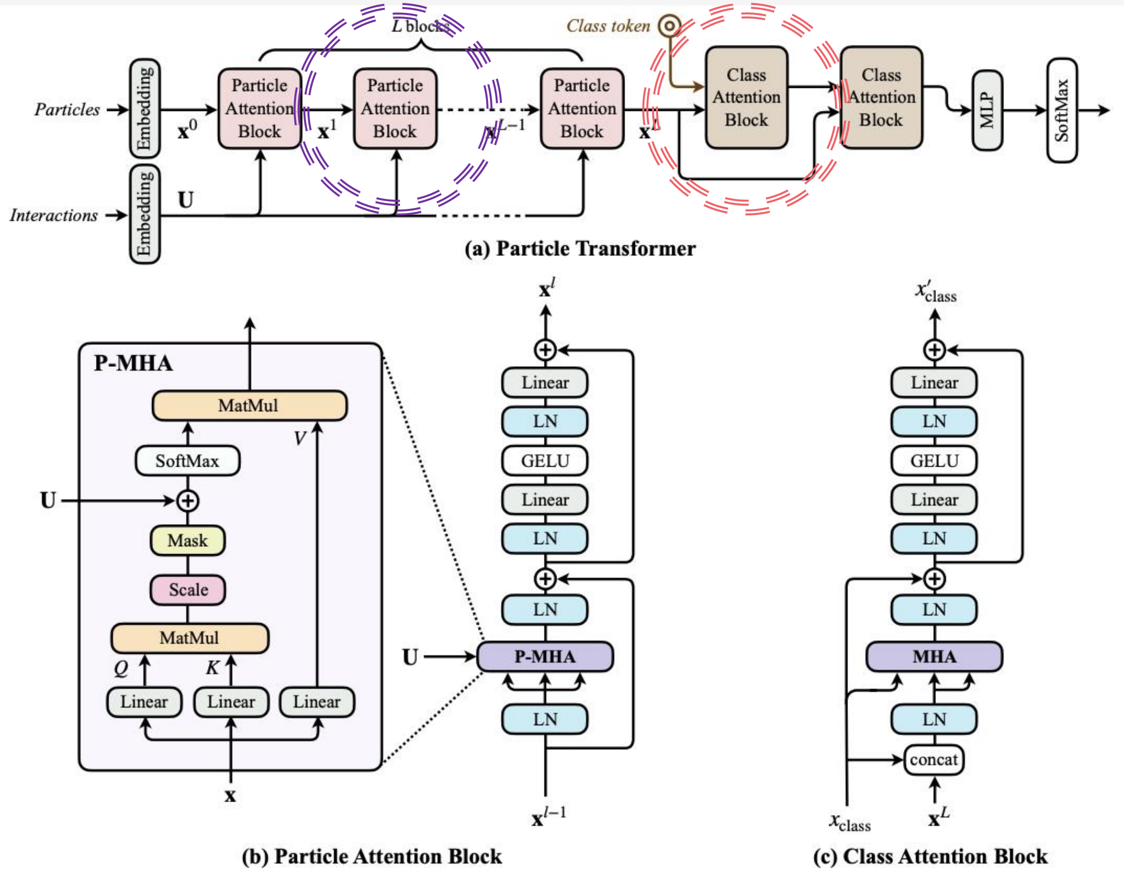
Advanced Tagging: Particle Transformer

- **Particle transformer**: attention on particle and their "interactions"

- **Particle attention**: particle <–> particle
- **Class attention**: particles –> class/jet flavor

$$
\begin{aligned}
\Delta &= \sqrt{(y_a - y_b)^2 + (\phi_a - \phi_b)^2},\\
k_\mathrm{T} &= \min(p_{\mathrm{T},a}, p_{\mathrm{T},b})\Delta,\\
z &= \min(p_{\mathrm{T},a}, p_{\mathrm{T},b})/(p_{\mathrm{T},a} + p_{\mathrm{T},b}),\\
m^2 &= (E_a + E_b)^2 - \|\mathbf{p}_a + \mathbf{p}_b\|^2,
\end{aligned}
$$

interactions



https://arxiv.org/abs/2202.03772

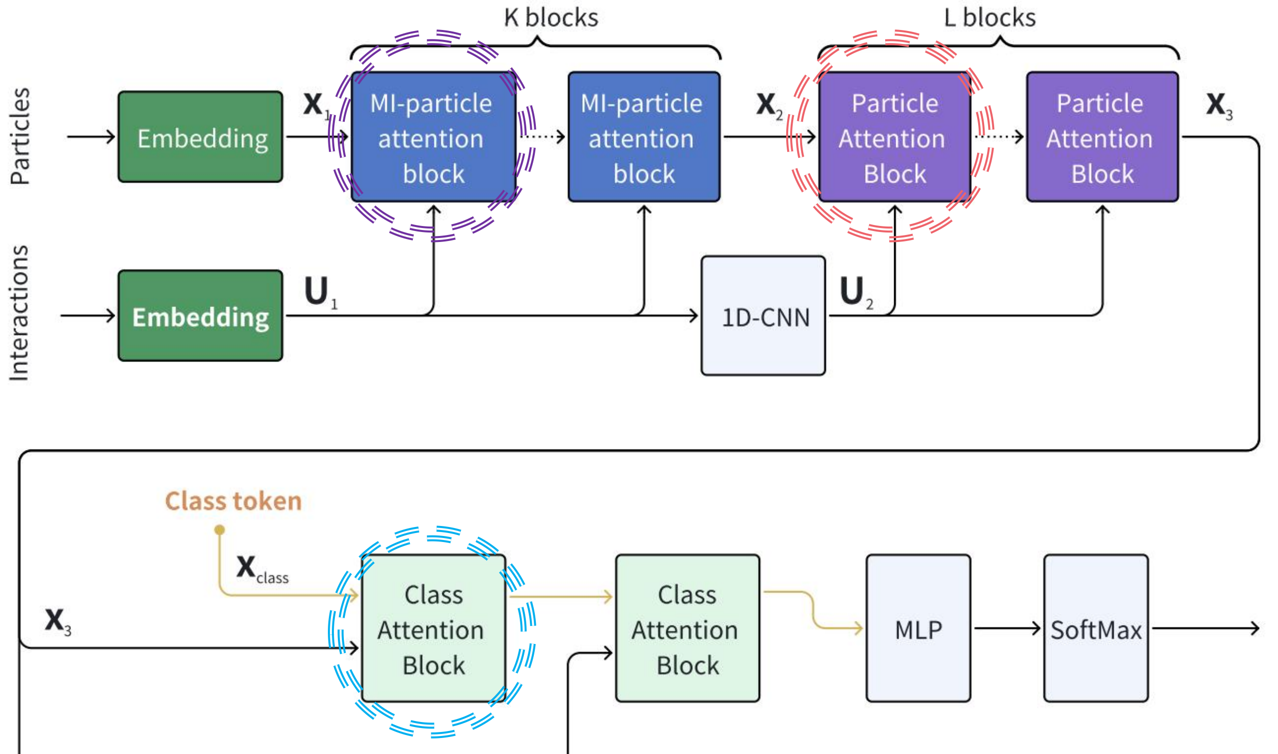
Advanced Tagging: More-Interaction Particle Transformer

- **Particle transformer**: attention on particle and their "interactions"

- **Particle attention**: particle <-> particle
- **Class attention**: particles -> class/jet flavor

- **More-Interaction attention**: reduces model complexity compared with particle attention

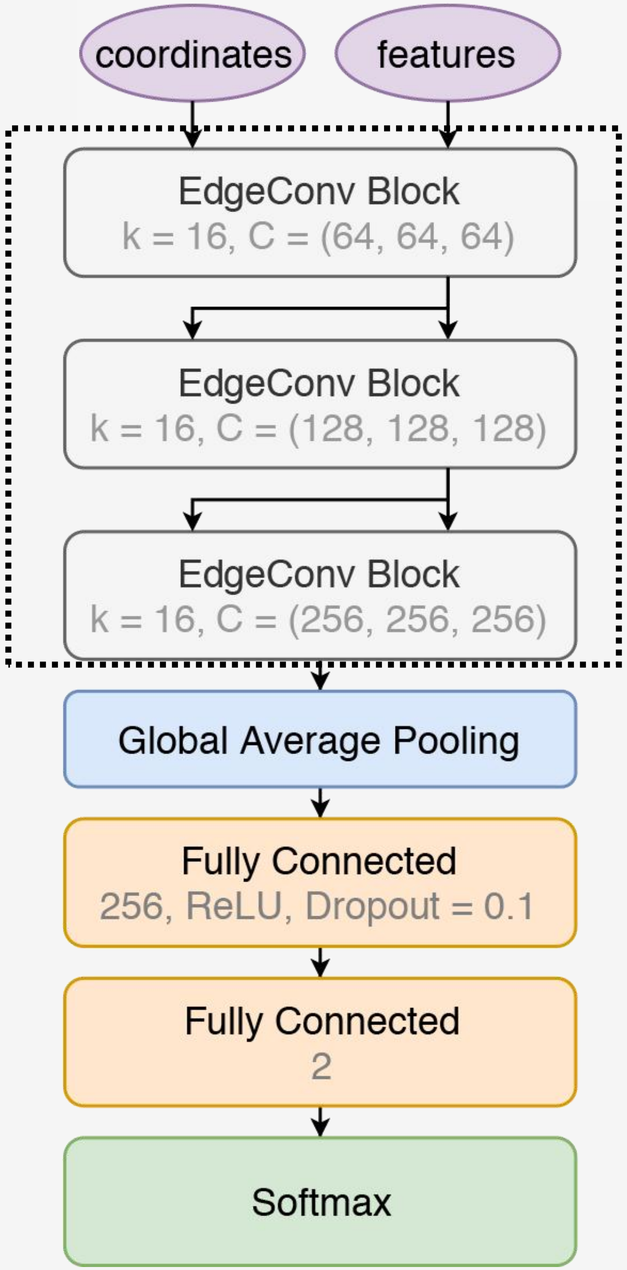
model	Params
ParticleNet	370k
ParT	2.14M
MIParT	720.9k



arXiv:2407.08682

Advanced Tagging: ParticleNet

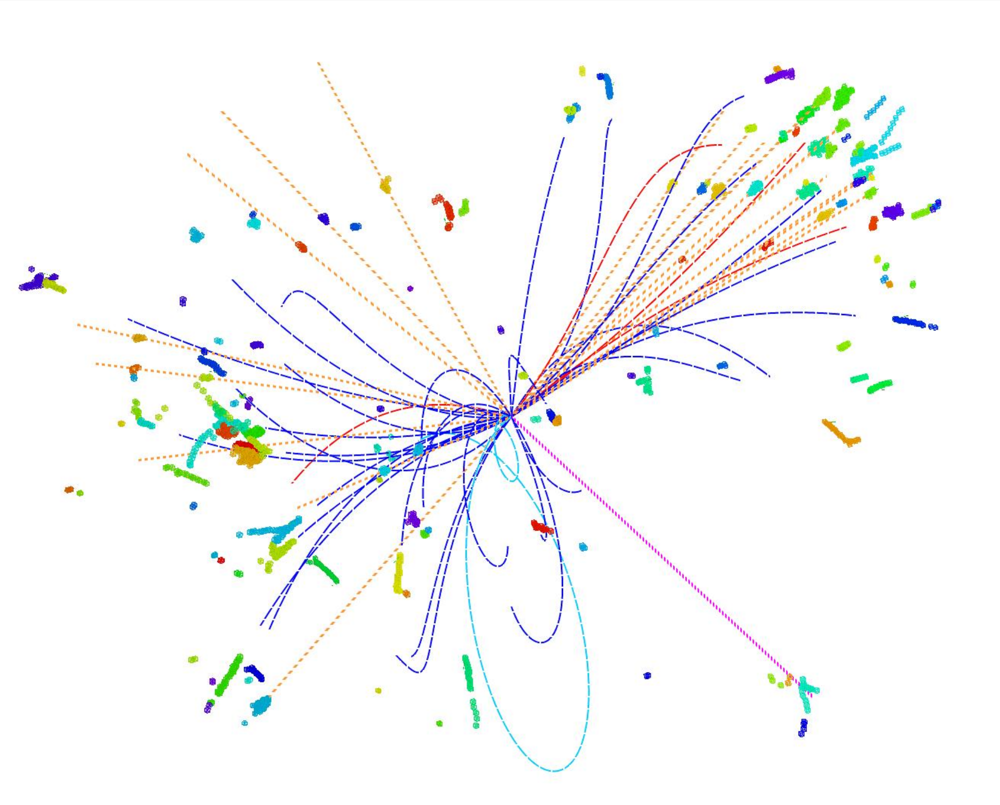
- **Dynamic Graph Convolutional Neural Network**
  - Dynamically updates the graph to reflect the changes in the edges
- **Permutationally symmetric** operation
  - Particles are connected to neighbors
- **Stackability**
  - Graph can evolve with more blocks



Variable	Definition
$\Delta\eta$	Difference in pseudorapidity between the particle and the jet axis
$\Delta\phi$	Difference in azimuthal angle between the particle and the jet axis
$\log p_T$	Logarithm of the particle's p_T
$\log E$	Logarithm of the particle's energy
$\log \frac{p_T}{p_T(\text{jet})}$	Logarithm of the particle's p_T relative to the jet p_T
$\log \frac{E}{E(\text{jet})}$	Logarithm of the particle's energy relative to the jet energy
ΔR	Angular separation between the particle and the jet axis $\left(\sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}\right)$
q	Electric charge of the particle

PhysRevD.101.056019

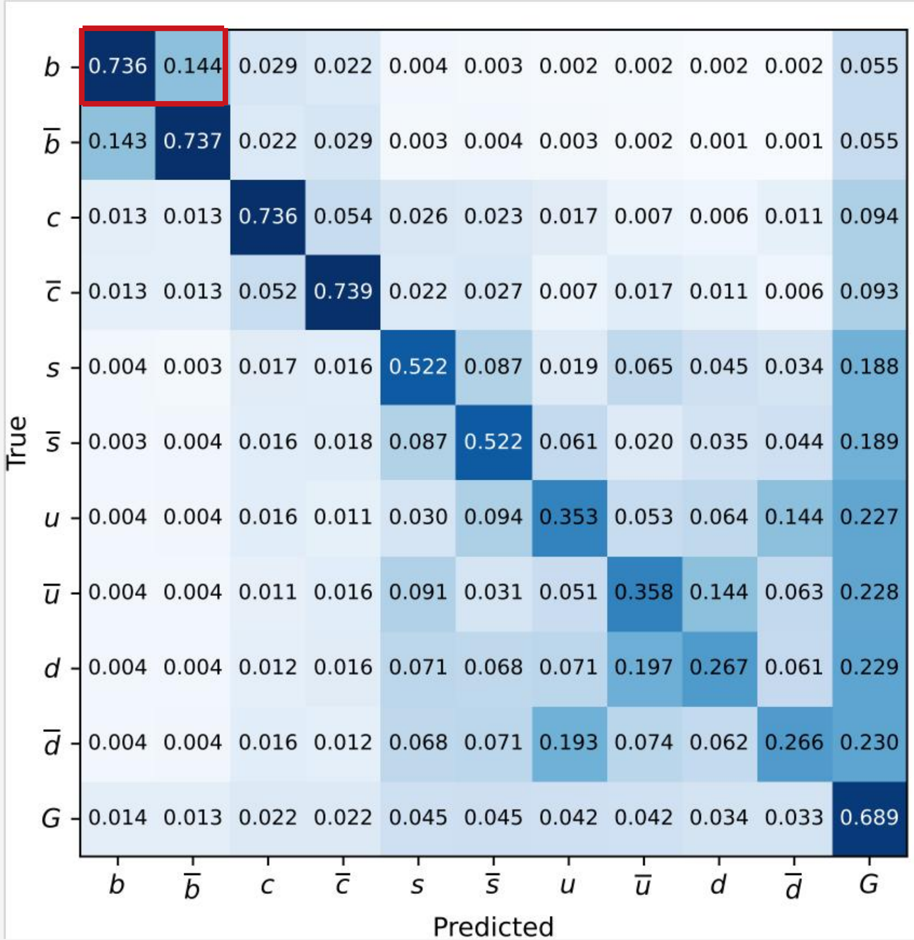
JOI sample



- $e^+e^- -> \nu\bar{\nu}H$, *H–>qq/gg at* 240GeV, fast simulation, CDR card
- Whizard(1.95+pythia6.4)
- 1 Million samples each, 0.8M for tagging, 0.2M for event selection

Jet tagging output

- Capable of tagging light jets even charged jets



ParticleNet, Fast simulation, CDR

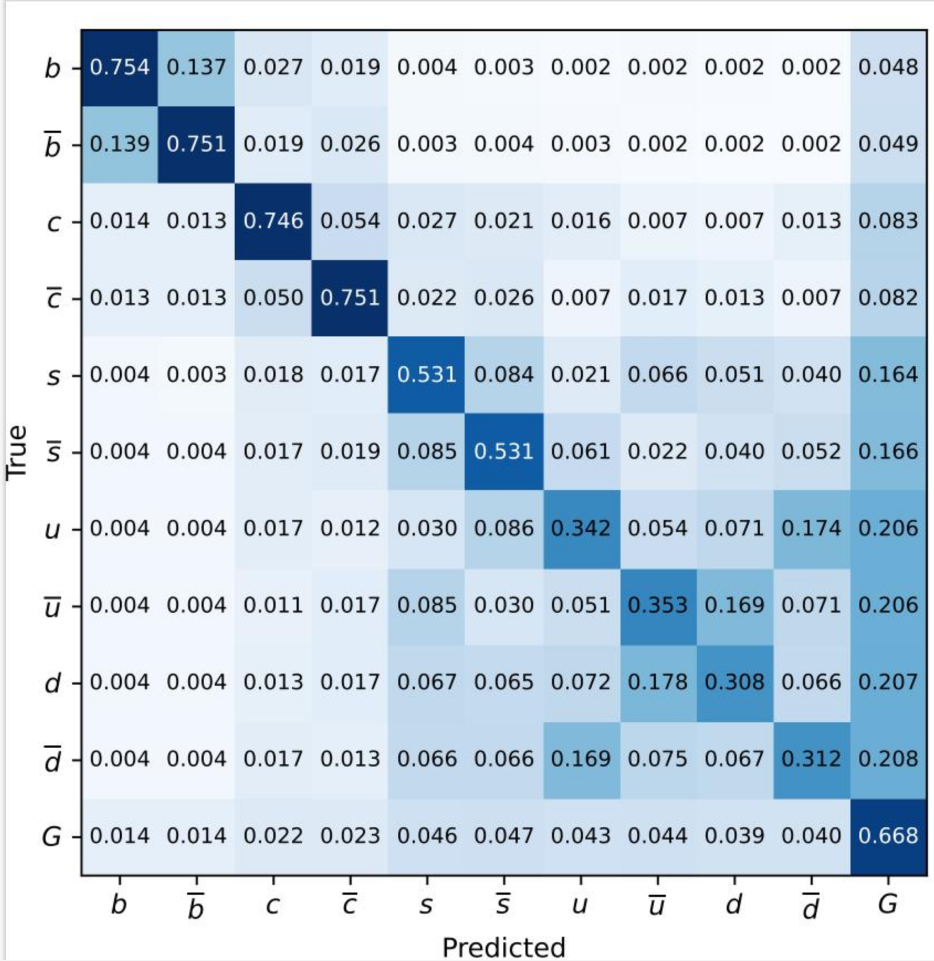
- diag = $\frac{1}{Num\ of\ classes}$ trace(matrix)

- Each model trained **3** times, the below tabel shows average value

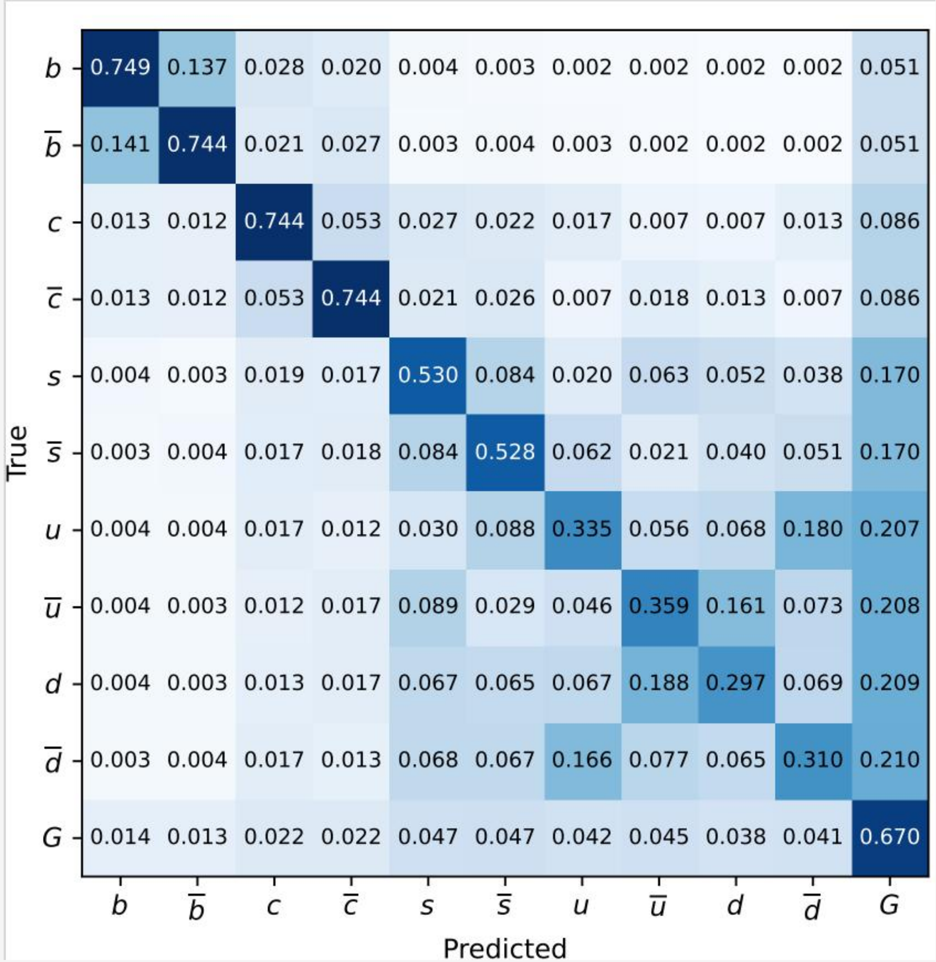
	ParticleNet	ParT	MIParT
diag	0.537	0.550	0.546
b tag as b/$\bar{b}$	0.878	**0.891**	0.885
c tag as c/$\bar{c}$	0.790	**0.800**	0.797
s tag as s/$\bar{s}$	0.608	0.616	**0.618**
g tag as g	**0.686**	0.670	0.664

Jet tagging output

- Comparable results from ParT & MIParT



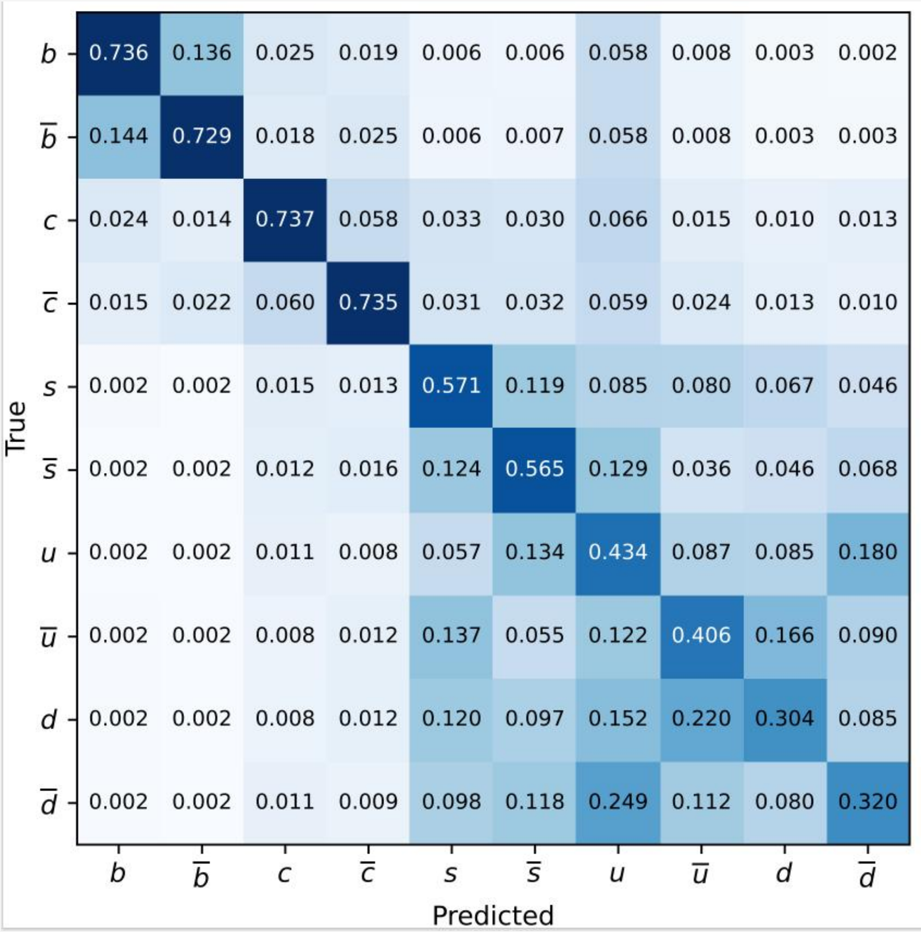
ParT,Fast simulation, CDR



MIParT, Fast simulation, CDR

Jet tagging infer

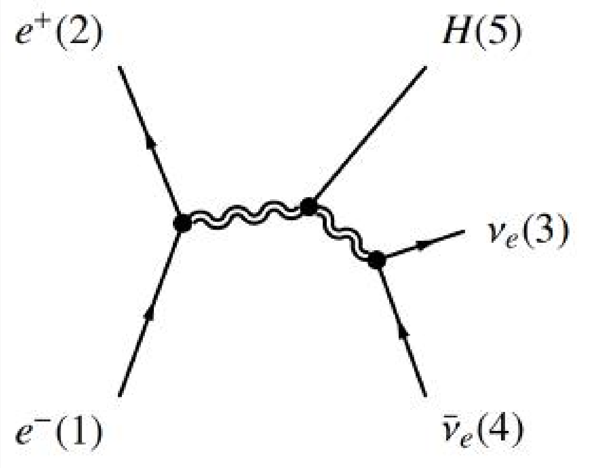
- Inference on $2f$
- Similar flavor score distributions
  –> has good generalization ability



ParT,Fast simulation, TDR

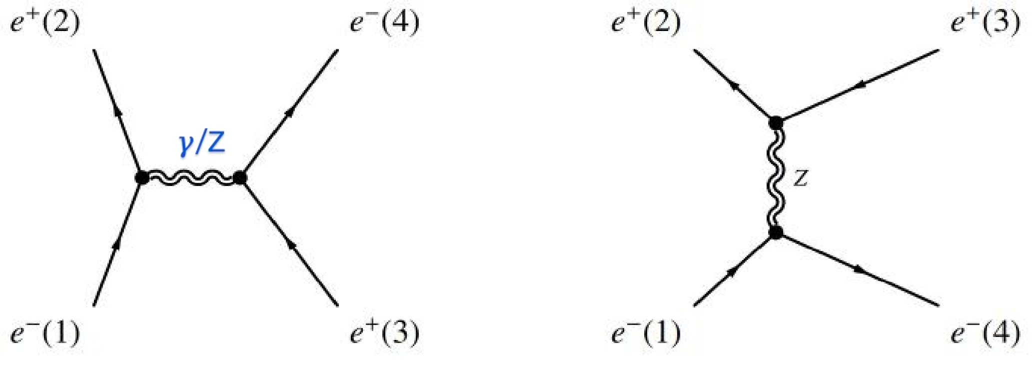
Input channels

- ZH, Z→νν, H→qq/gg/ZZ

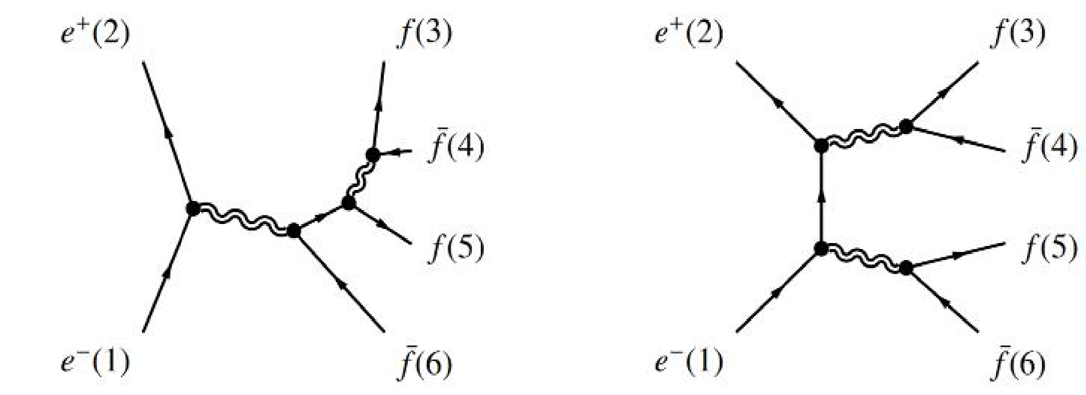


for H–>ZZ, Z–>νν, Z–>qq

- 2 fermions process



- 4 fermions_nu process



Feynman diagram from the CEPC note

Input channels

- Fast simulation: whizard (1.95+pythia6.4, CDR card)
  - H–>ZZ sample used TDR card

for event selection

channel	cross section [fb]	expected events [M]	simulated events [M]	scale factor
ZH, Z→νν, H–>bb	26.71	0.5342	0.2	2.671
ZH, Z→νν, H–>cc	1.35	0.027	0.2	0.135
ZH, Z→νν, H–>ss	0.01	0.0002	0.2	0.001
ZH, Z→νν, H–>gg	3.97	0.0794	0.2	0.397
ZH, Z→νν, H–>ZZ; Z–>νν, Z–>qq	0.34	0.0068	10	0.0007
ZZ, Z→νν, Z→dd/ss/bb	139.71	2.7942	1	2.80
ZZ, Z→νν, Z→uu/cc	84.38	1.6876	1	1.69
Single Z, νν, Z→dd/ss/bb	90.03	1.8006	1	1.80
Single Z, νν, Z→uu/cc	55.59	1.1118	1	1.11
2f, qq	54106.86	1082.1372	4	270.53

https://arxiv.org/pdf/2203.01469

Input features

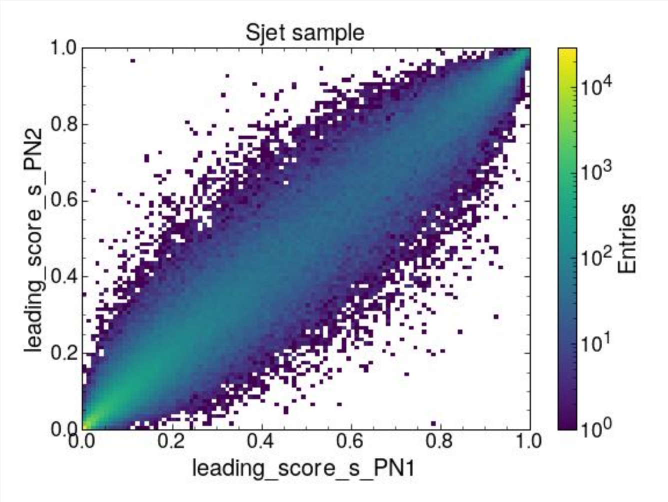
- Jet constructed by eekt algorithm, each event has 2 jets
  - each jet-level feature, both leading and subleading jet are contained

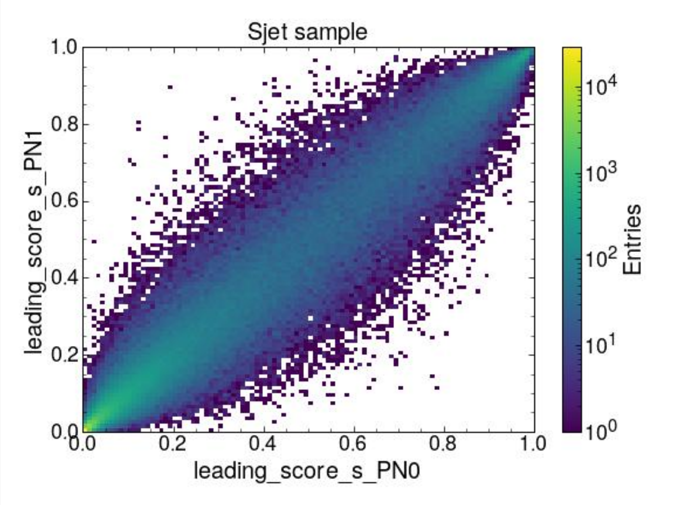
- For PN/ParT/MIParT, each trained three times
  - **3 scores per type of tagging model per jet**

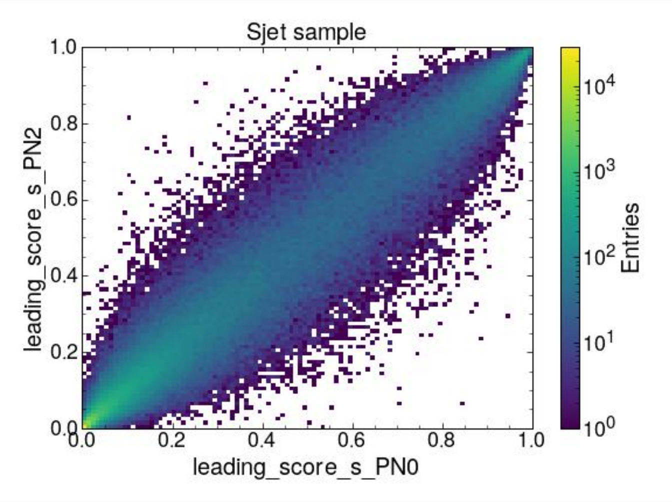
- Single Jet Kinematics: jet_pt, **jet_pz**, jet_eta, **jet_theta**, jet_phi, jet_energy
- Jet Shape & Composition: jet_nParticles, jet_dR, jet_dPT
- Jet Pair Observables: mjj, detajj, dthetajj, dphijj
- Missing Energy: MET, ME_eta, ME_theta, ME_phi, MEZ, METOHT
- Jet—MET Angular Correlations: jet_ME_deta/dphi/dtheta
- **Jet Flavor Tagging Scores: jet_flavor_score(PN/ParT/MIParT)**

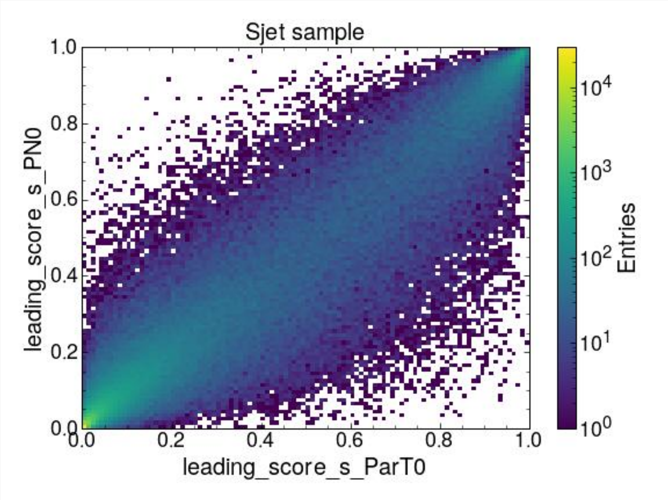
Jet score distribution

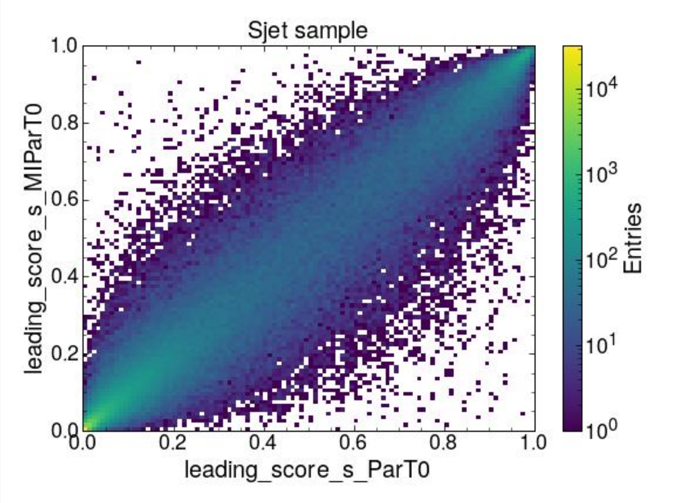
- Distribution for different models' score
- the same model was used to perform inference three times

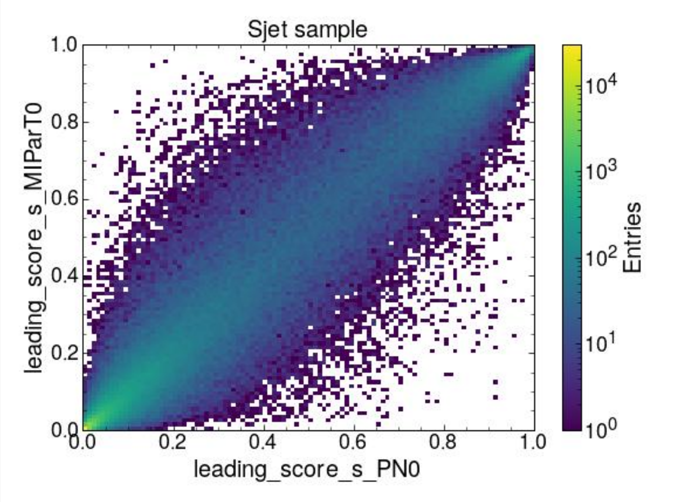




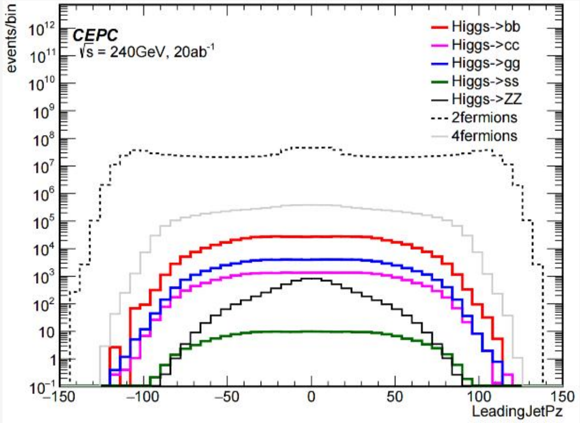




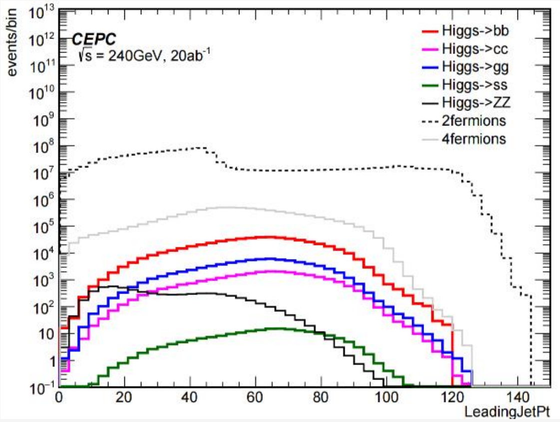




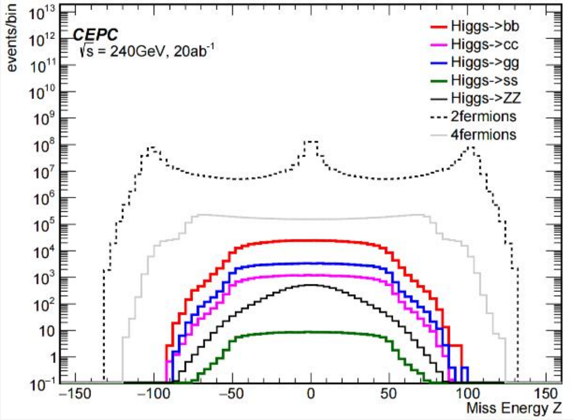
Raw distribution



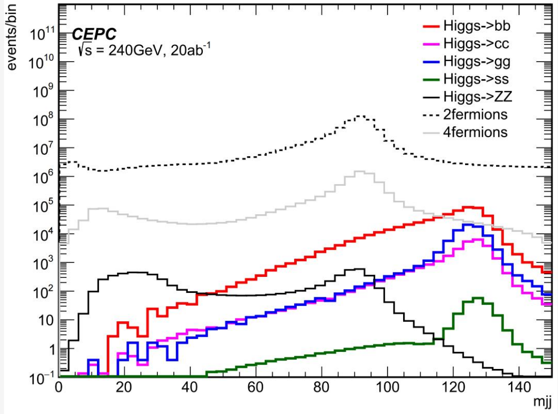
Longitudinal momentum of the leading jet



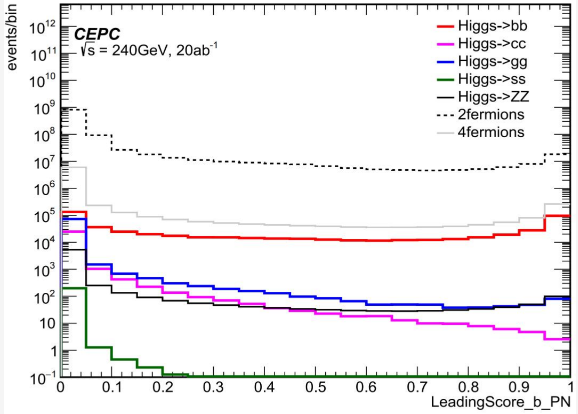
Transverse momentum of the leading jet



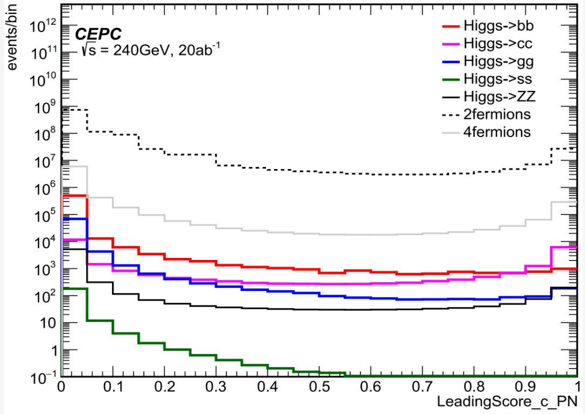
Missing energy along the beam axis (MEZ)



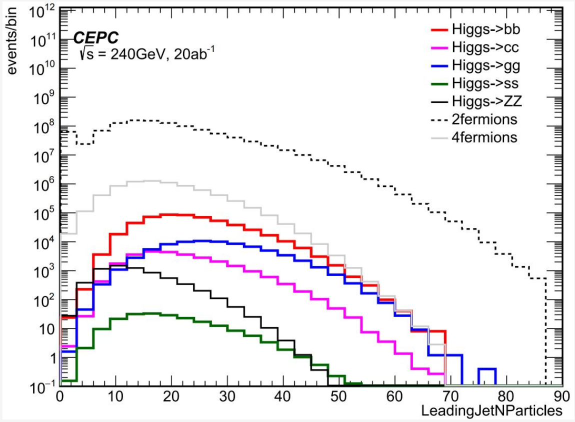
mjj



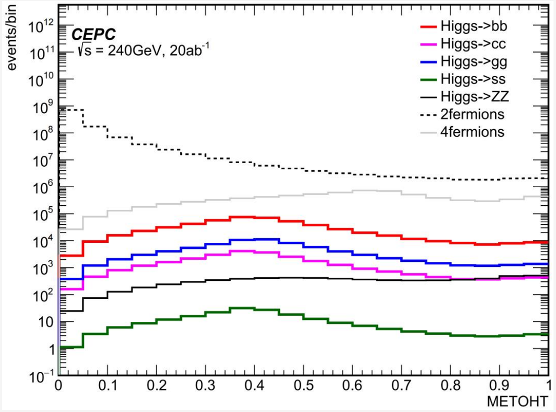
b score predicted by ParticleNet



c score predicted by ParticleNet



LeadingJetNParticles



METOHT

Region

- bbccgg/ss region

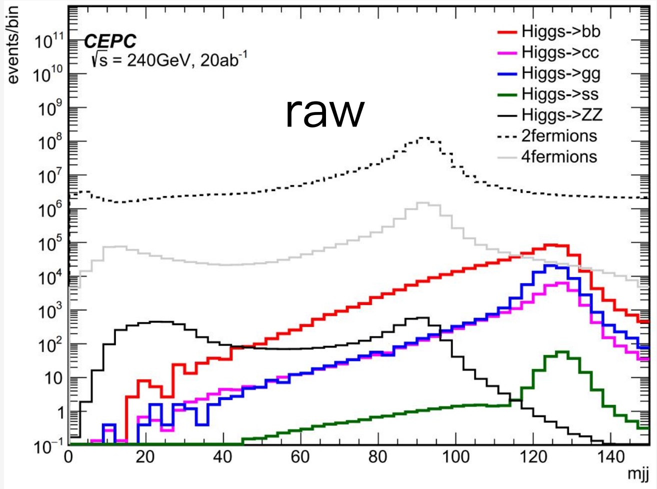
MC Sample

leading jet pz: (−95~95)GeV
leading jet pt: (15~100)GeV
MEZ: (−55,55)GeV
mass peak: (110, 140)GeV

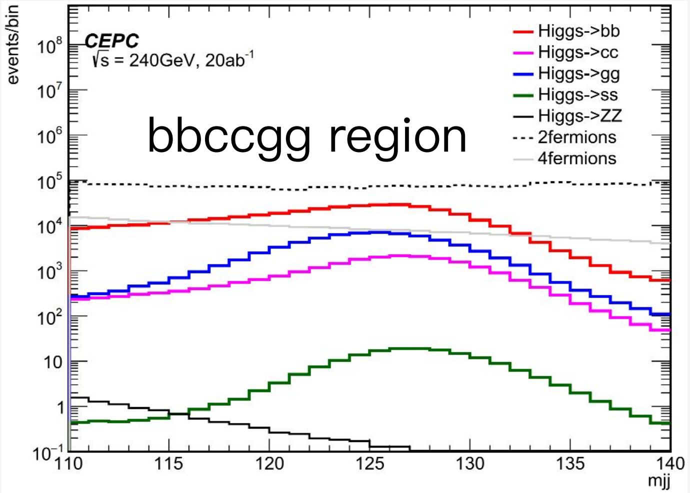
bbccgg region

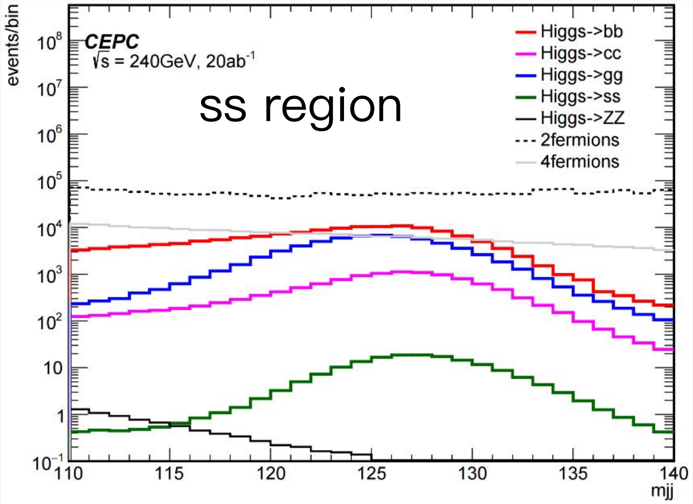
same as above, plus
leading_score_b_PN<0.2
leading_score_c_PN<0.2

ss region



mjj





mjj after region cut

set ss region: Suppress bkg variation from other processes

Pre Cutflow table

- bbccgg region
- Keep enough events/topology for ML training
- Suppress 2f, 4f and HZZ process

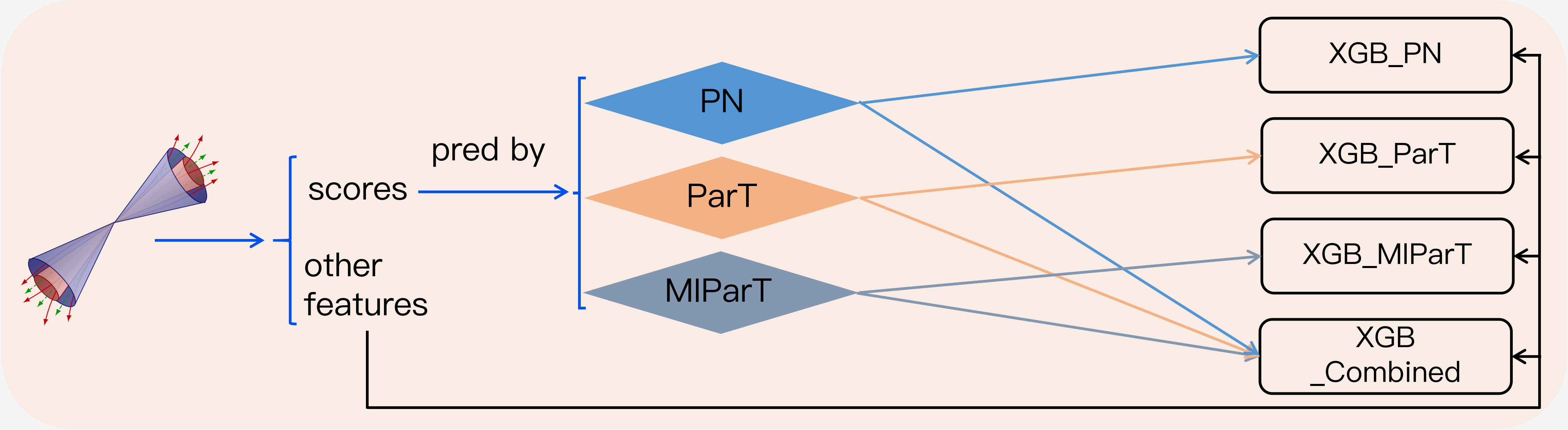
cut	H->bb	H->cc	H->gg	H->ss	H->ZZ	4f_nu	2f
No cut	200,000	200,000	200,000	200,000	10M	4,000,000	4,000,000
leading jet pz (-95~95)GeV	99.93%	99.91%	99.95%	99.89%	99.99%	99.67%	78.70%
leading jet pt (15~100)GeV	99.39%	99.42%	99.52%	99.38%	82.38%	97.04%	78.85%
MEZ (-55,55)GeV	97.81%	97.67%	97.81%	97.56%	98.69%	63.21%	32.18%
mass peak (110,140)GeV	77.76%	89.70%	95.54%	90.85%	0.20%	3.51%	1.06%
pre-cut eff	75.54%	87.03%	92.96%	87.99%	0.16%	24.31%	0.21%
remaining	151082	174061	180418	175984	15621	140277	8491

Pre Cutflow table

- ss region

cut	H->bb	H->cc	H->gg	H->ss	H->ZZ	4f_nu	2f
No cut	200,000	200,000	200,000	200,000	10M	4,000,000	4,000,000
leading jet pz (-95~95)GeV	99.93%	99.91%	99.95%	99.89%	99.99%	99.67%	78.70%
leading jet pt (15~100)GeV	99.39%	99.42%	99.52%	99.38%	82.38%	97.04%	78.85%
MEZ (-55~55)GeV	97.81%	97.67%	97.81%	97.56%	98.69%	63.21%	32.18%
mass peak (110~140)GeV	77.76%	89.70%	95.54%	90.85%	0.20%	5.74%	1.06%
lead_score_b (0~0.2)	39.76%	98.10%	98.00%	99.80%	94.48%	91.57%	83.75%
lead_score_c (0~0.2)	94.80%	53.46%	97.16%	98.37%	87.66%	86.45%	86.04%
pre-cut eff	28.48%	45.64%	88.51%	86.38%	0.13%	2.78%	0.15%
remaining	56950	91289	177020	172761	12109	111052	6118

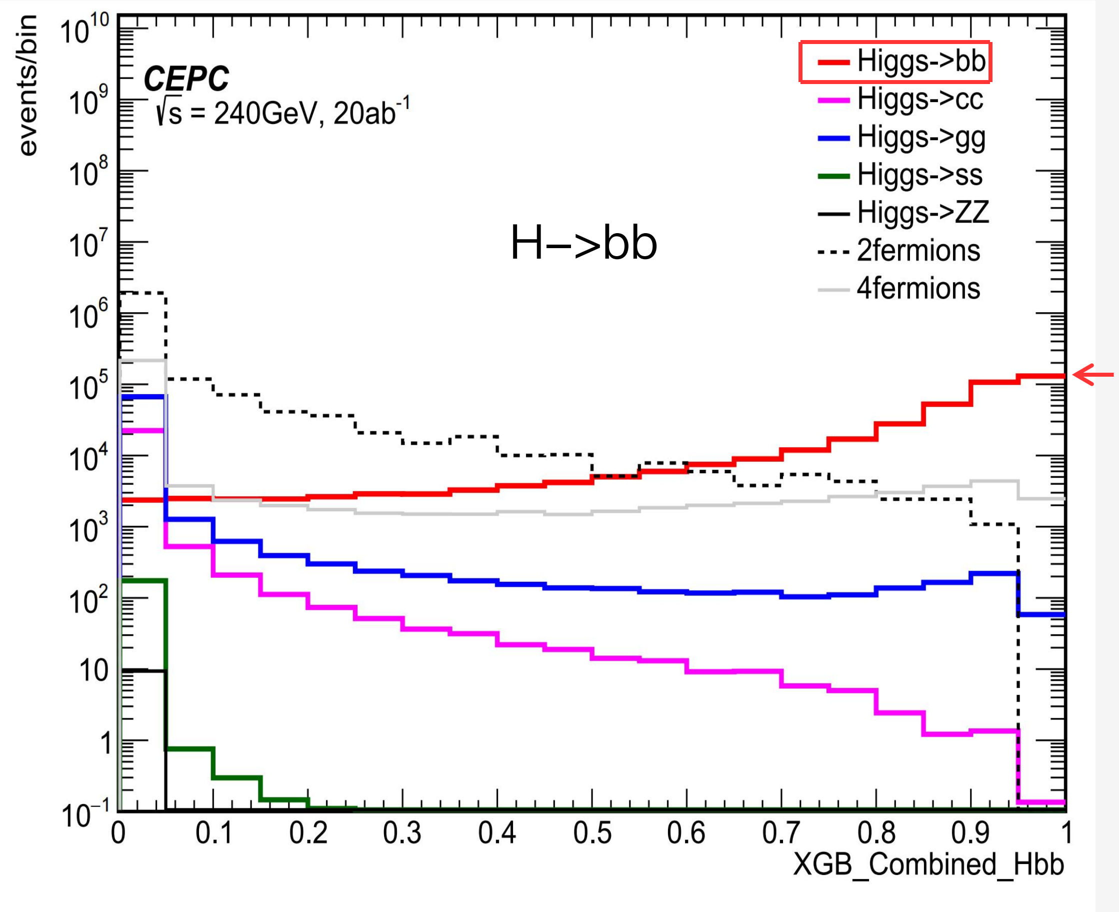
BDT setting

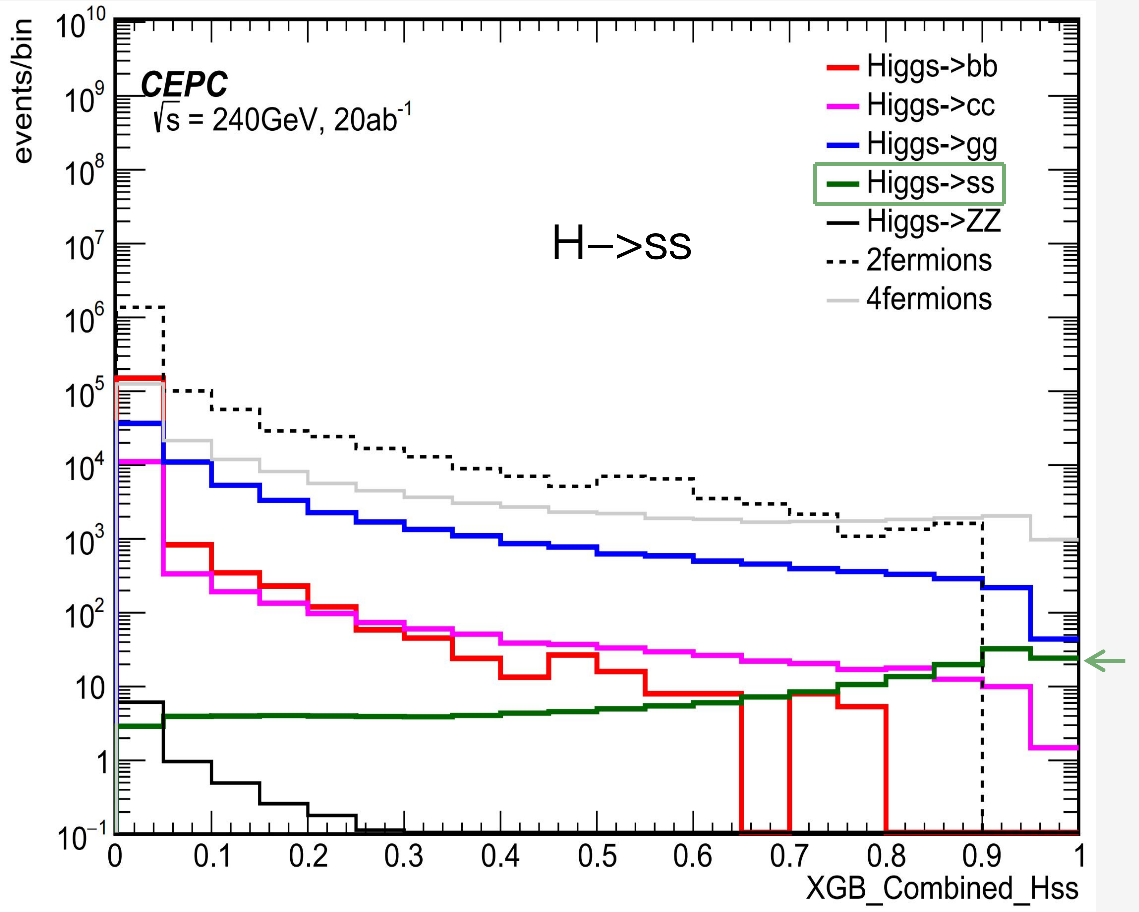


- main model: XGBoost
  - num_boost_round = **300**,
  - learning_rate = **0.1**
  - max_depth = **3**
  - train:val:test = **6:2:2**
- For PN/ParT/MlParT, each trained three times
  → 3 scores per type of tagging model per jet
- 6 classes: h–>bb, h–>cc, h–>gg, h–>ss, 2fermions, 4fermions

BDT output

- Strong discriminative power for main classes





Cutflow table

- H->bb/cc/gg as signal
- XGB Combined
- ML-cut at the point which makes the highest significance of signal

signal	eff & remain	H->bb	H->cc	H->gg	H->ss	H->ZZ	4f_nu	2f
H->bb	total-cut	**58.40%**	1.95E-04	0.75%	1.20E-04	1.09E-05	0.19%	5.94E-06
	Events after ML-cut (20/ab)	**311979.35**	5.26	592.39	0.02	0.08	14408.67	6423.24
H->cc	total-cut	1.45E-04	**36.34%**	0.30%	6.50E-05	3.40E-06	0.08%	2.47E-07
	Events after ML-cut (20/ab)	77.46	**9810.80**	240.55	0.01	0.02	6161.97	267.63
H->gg	total-cut	0.13%	0.35%	**39.70%**	1.06%	1.20E-06	0.02%	2.47E-07
	Events after ML-cut (20/ab)	675.75	95.85	**31523.73**	2.12	0.01	1736.05	267.63

Cutflow table

- H−>ss as signal
- XGB Combined
- ML−cut at the point which makes the highest significance of signal

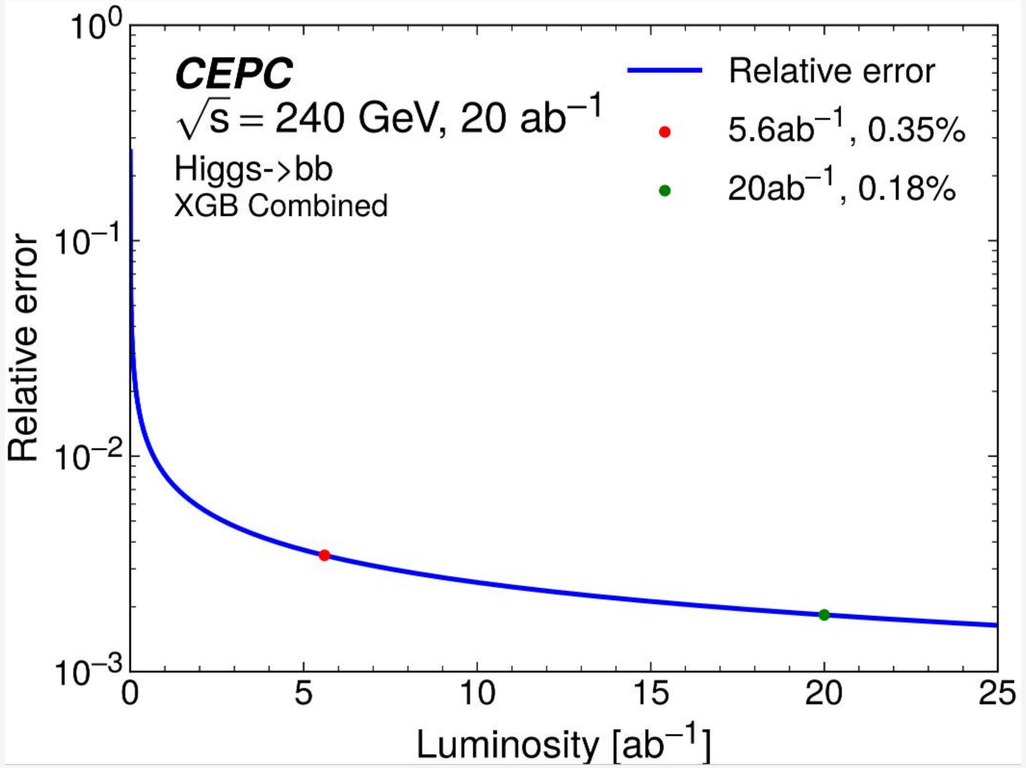
Signal

cut	H−>bb	H−>cc	H−>gg	**H−>ss**	H−>ZZ	4f_nu	2f
No cut	100%	100%	100%	**100%**	100%	100%	100%
pre−cut	28.48%	45.64%	88.51%	**86.38%**	0.13%	2.78%	0.15%
ML−cut	0	0.11%	0.46%	**35.55%**	0.10%	1.34%	0
total−cut	0	0.05%	0.41%	**30.70%**	1.29E−06	0.04%	0
σ[fb]	26.71	1.35	3.97	**0.01**	0.36	369.71	54106.86
Events after pre−cut (20/ab)	152140.16	12322.80	70276.94	**172.76**	9.36	205558.76	1623205.80
Events after ML−cut (20/ab)	0	13.90	324.83	**61.41**	0.01	**2745.05**	0

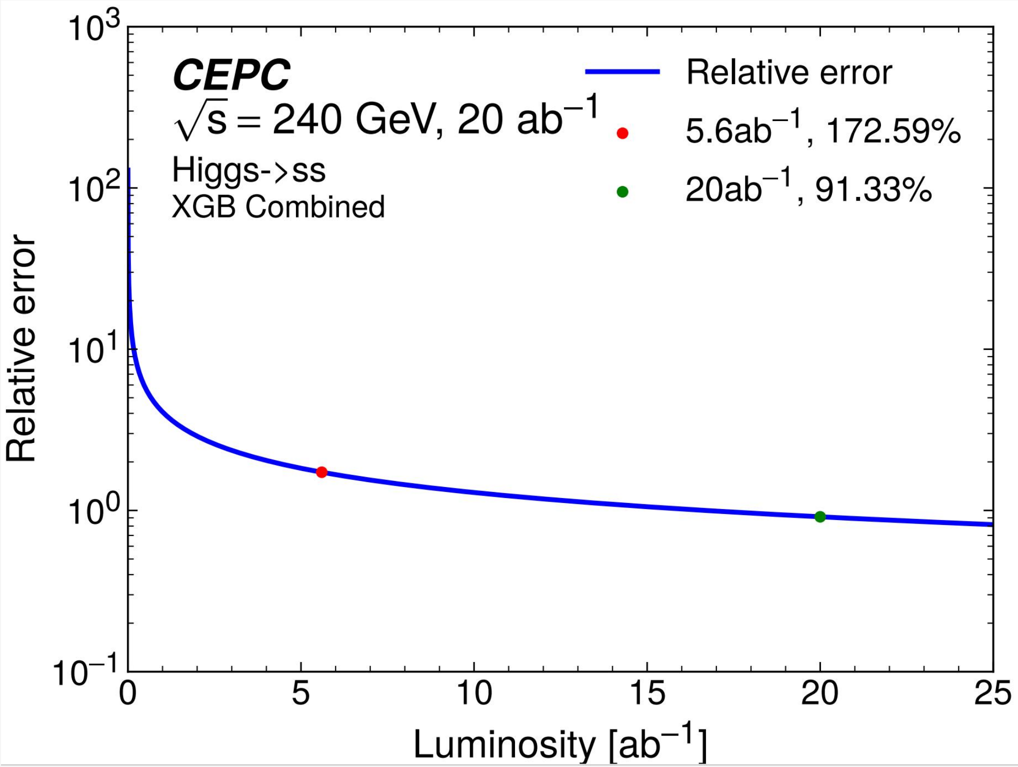
Relative Error

- Relative error = $\frac{\sqrt{S+B}}{S}$

- Lower relative error implies better precision in measuring signal yield



H–>bb



H–>ss

Relative Error

- Relative error compare with published paper, at 20ab^{-1} (5.6ab^{-1})
  - encourage result
  - precision of "Z →νν, H–>cc" has significant improvement (~35%)

	Z →νν, H–>bb %	Z →νν, H–>cc %	Z →νν, H–>gg %	Z →νν, H–>ss %
published	0.2 (0.4)	1.85 (3.7)	0.7 (1.4)	–
XGB_PN	0.18 (0.35)	1.24 (2.35)	0.61 (1.14)	93.34(176.40)
XGB_ParT	0.19 (0.35)	1.23 (2.33)	0.59 (1.11)	93.62(176.93)
XGB_MIParT	0.19 (0.35)	1.22 (2.31)	0.59 (1.11)	92.10(174.05)
XGB_Combined	**0.18 (0.35)**	**1.22 (2.31)**	**0.58 (1.09)**	**91.33(172.59)**

Compare with the lastest result

上海交通大學 SHANGHAI JIAO TONG UNIVERSITY

- Compare with the lastest holistic result

https://arxiv.org/pdf/2506.11783

- Only considered process H–>bb/cc/gg/ss
- Beyond "cut+BDT"
- Near the holistic

$20\ \text{ab}^{-1}$

	$\nu\bar{\nu}H$			
	$H \to b\bar{b}$	$H \to c\bar{c}$	$H \to gg$	$H \to s\bar{s}$
cut + BDT	0.26%[21]	3.04%[21]	0.96%[21]	190.00%[19]
holistic	0.14%	0.72%	0.46%	29.34%
holistic with CSI	-	-	-	-
holistic with ideal CSI	-	-	-	-
statistical limit	0.14%	0.61%	0.36%	6.91%

model	Z →νν, H–>bb %	Z →νν, H–>cc %	Z →νν, H–>gg %	Z →νν, H–>ss %
XGB_PN	0.17	0.84	0.49	33.09
XGB_ParT	0.17	0.84	0.49	32.08
XGB_MIParT	0.17	0.86	0.48	32.58
XGB_Combined	0.17	0.82	0.48	32.30

Summary & todo

- Three advanced tagging algorithms have been evaluated for jet tagging task
- Jet–based study of Higgs hadronic decay considered main background
- Relative error:

  0.18% for $H \rightarrow$ bb, **1.22%** for cc, **0.58%** for gg, and **91.33%** for ss at $20\mathrm{ab}^{-1}$

- Results are close to the lastest performance

- Todo
  - Large data: 1M–>10M
  - H–>ss further
  - new model structure

上海交通大學

SHANGHAI JIAO TONG UNIVERSITY

Thanks

饮 水 思 源 爱 国 荣 校

SHANGHAI JIAO TONG UNIVERSITY

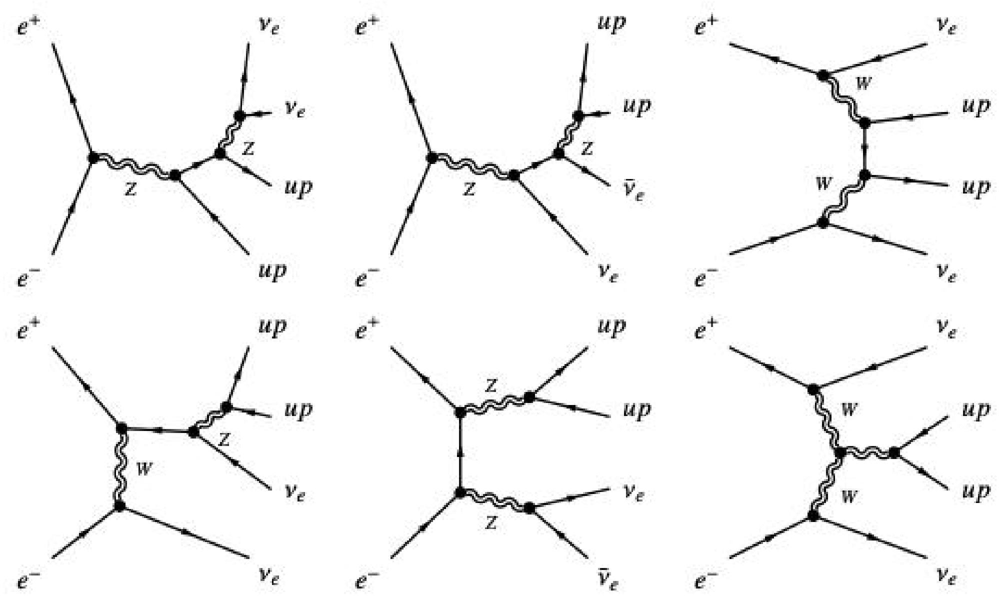
Back up

饮 水 思 源　爱 国 荣 校

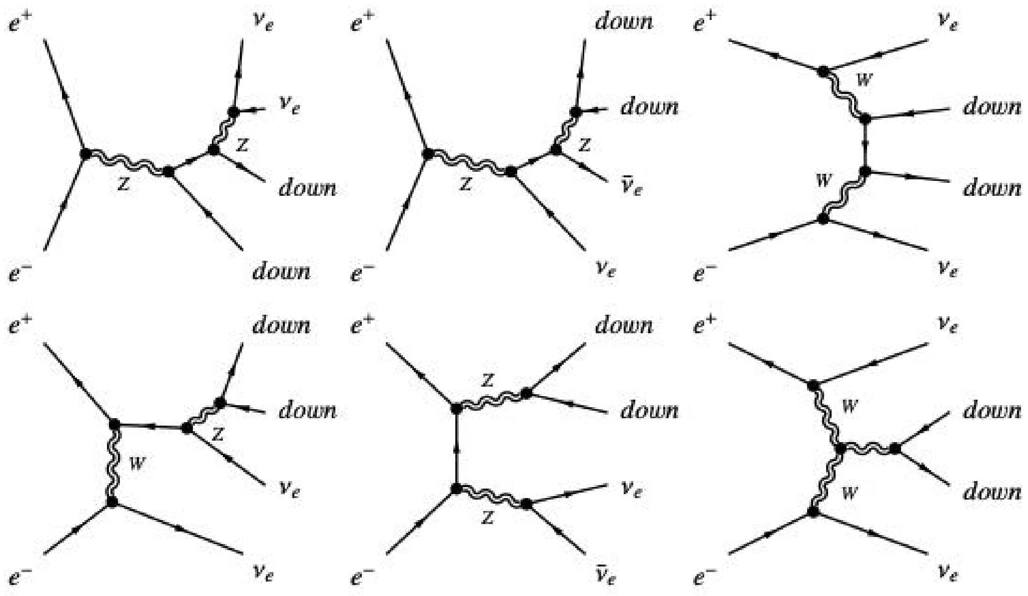
Input channels 4f

single Z_nu

215 **6.35 sznu_sl0nu_up**



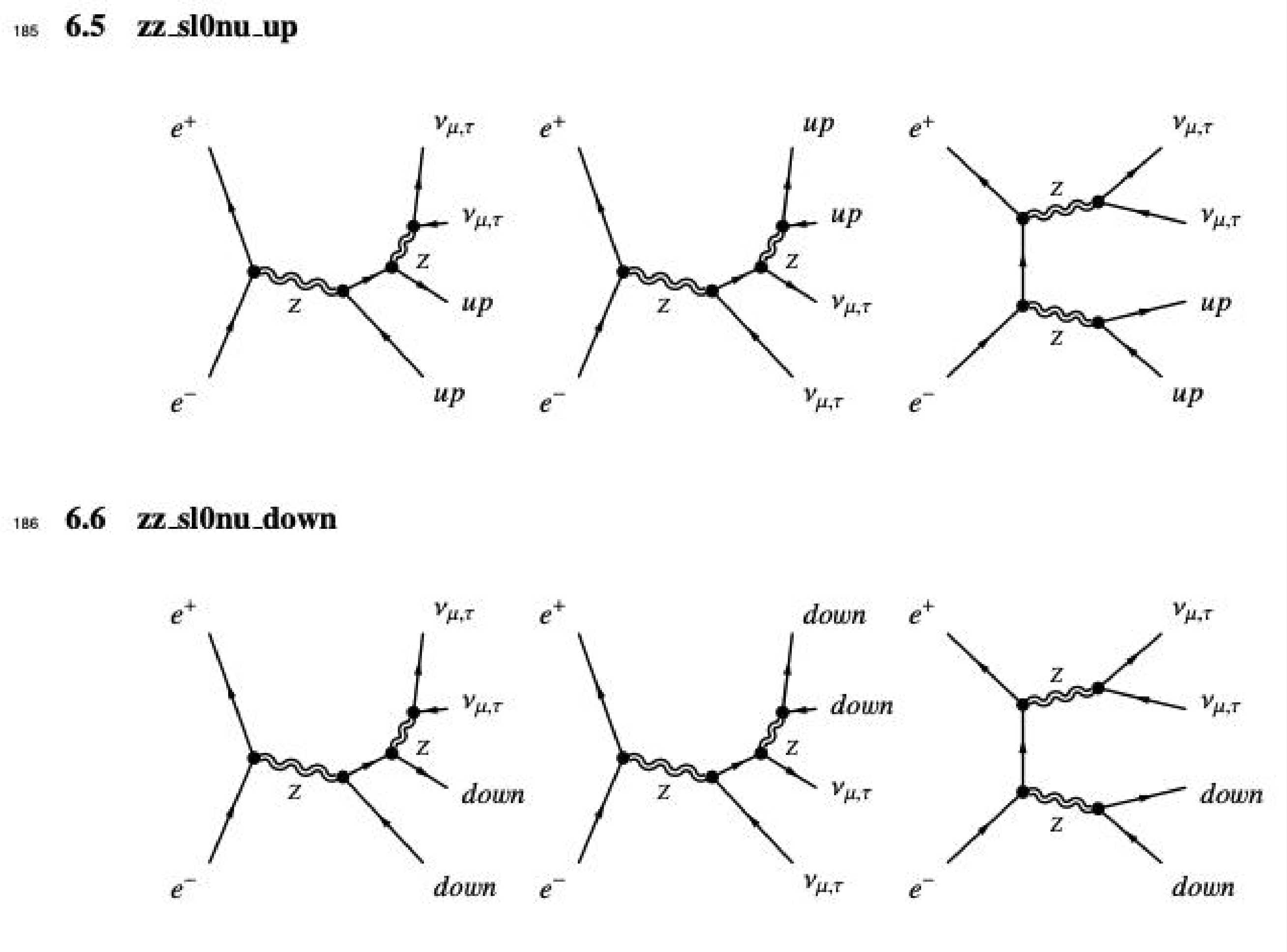
216 **6.36 sznu_sl0nu_down**



Feynman diagram from the CEPC note

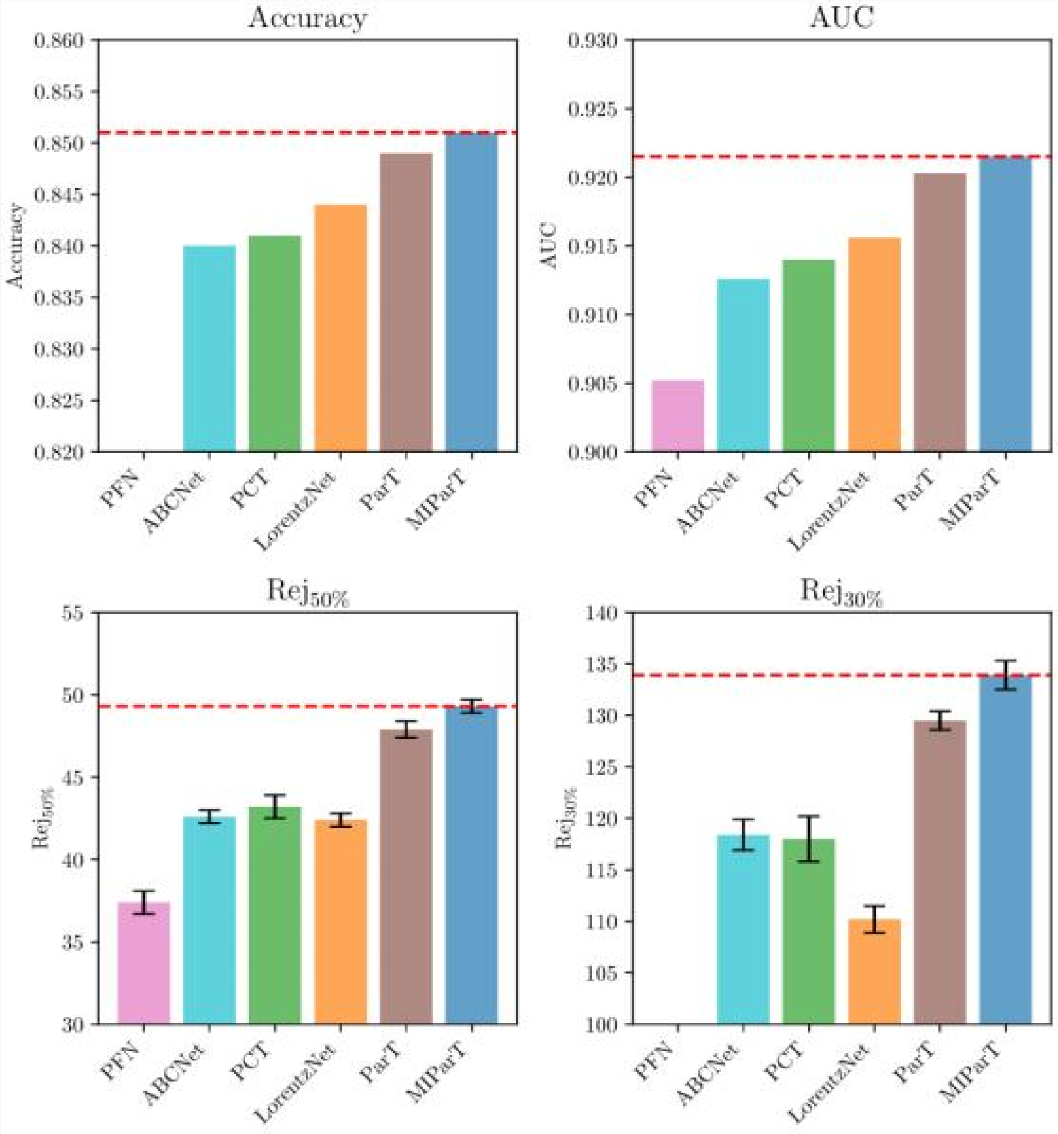
Input channels 4f_nu

ZZ_nu



Feynman diagram from the CEPC note

Performance

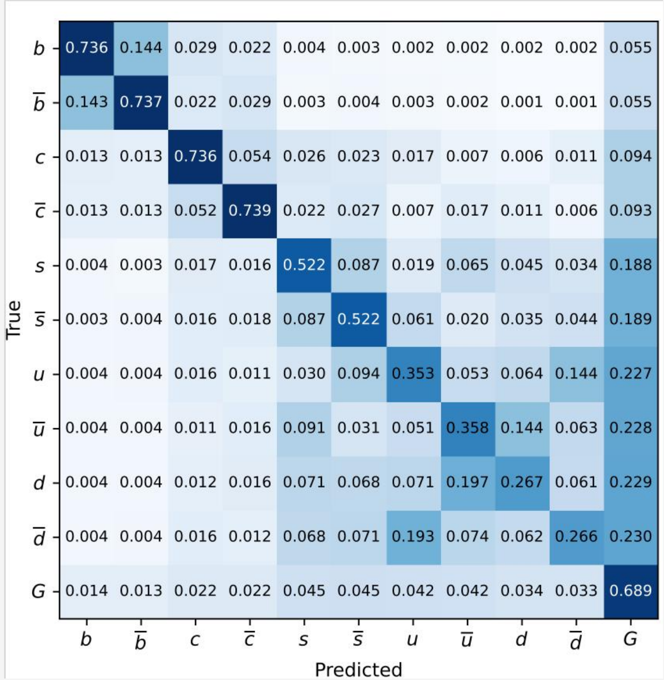


- Jet level kinematics + PID from Quark–gluon tagging dataset
- Representation suit with model structure

$$\mathrm{Rej}_{50\%} = \frac{1}{\text{bkg mis–id rate}}\Big|_{\text{signal efficiency [\%]}}$$

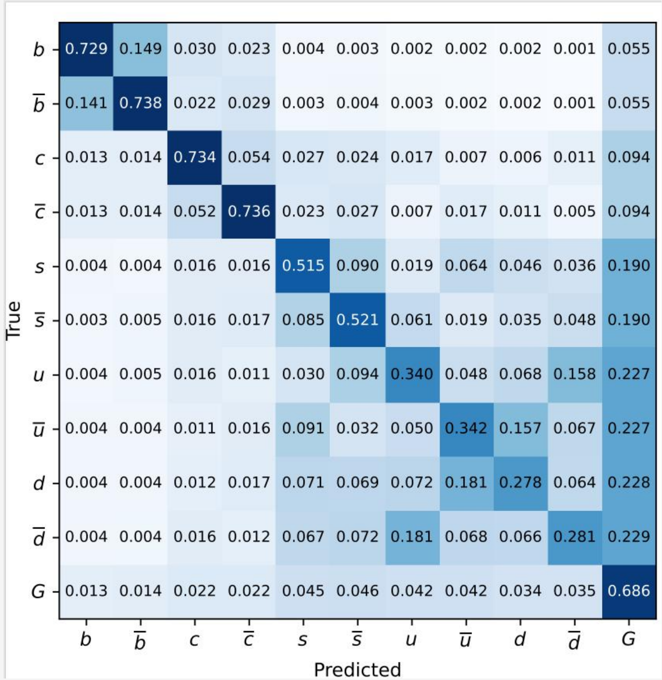
	Accuracy	AUC	$\mathrm{Rej}_{50\%}$
PFN	—	0.9052	37.4±0.7
ABCNet	0.840	0.9126	42.6±0.4
PCT	0.841	0.9140	43.2±0.7
LorentzNet	0.844	0.9156	42.4±0.4
ParT	0.849	0.9203	47.9±0.5
MIParT (ours)	**0.851**	**0.9215**	**49.3±0.4**
ParT f.t.	0.852	0.9230	50.6±0.2
MIParT-L f.t. (ours)	**0.853**	**0.9237**	**51.9±0.5**

PN CDR



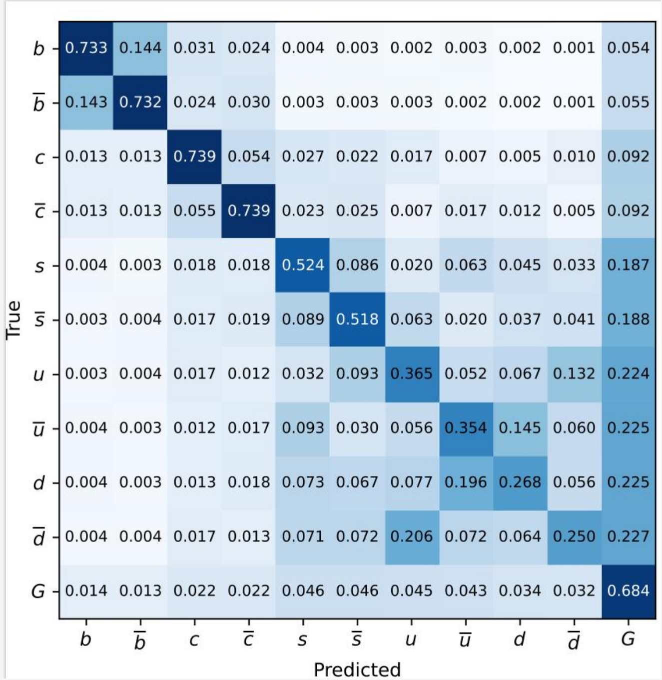
5.925

random0



5.902

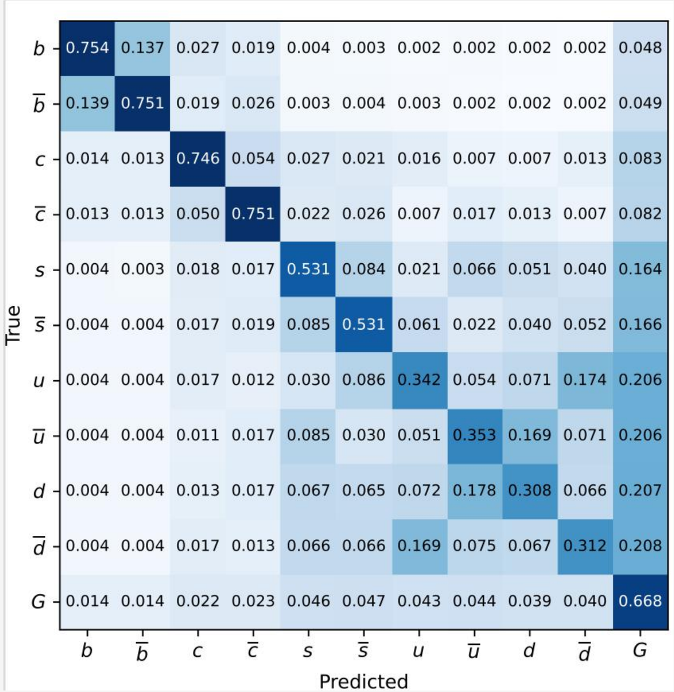
random1

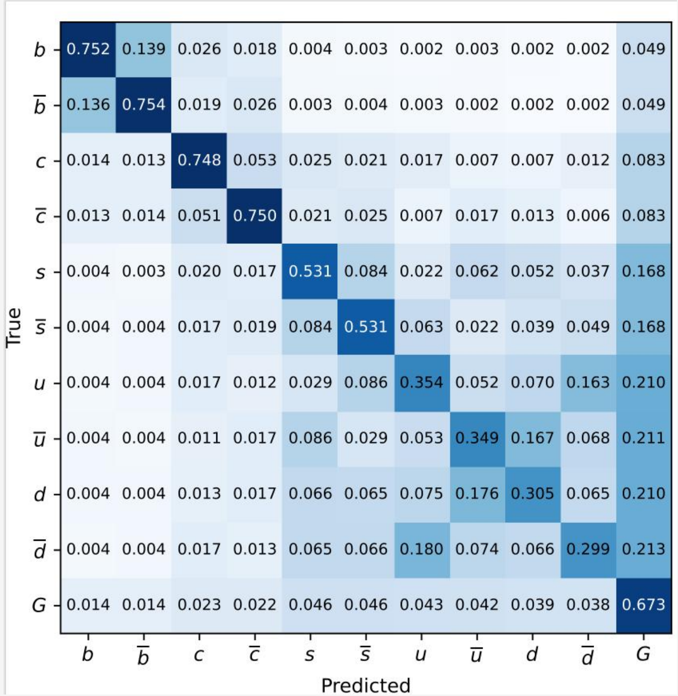


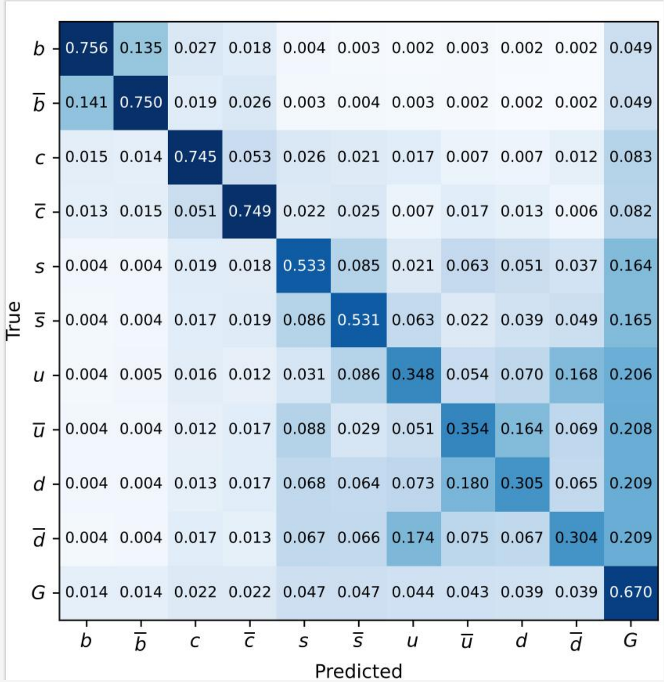
5.907

random2

PN CDR







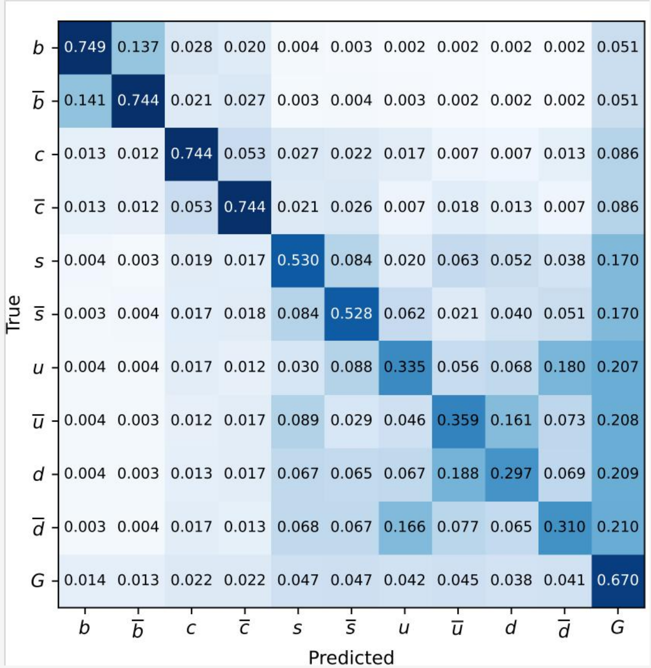
6.047
random0

6.046
random1

6.044
random2

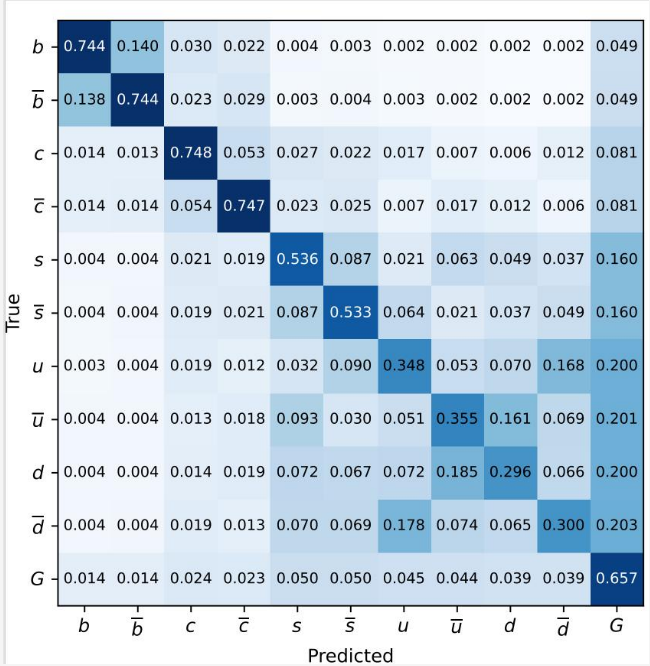
PN CDR

上海交通大学
SHANGHAI JIAO TONG UNIVERSITY



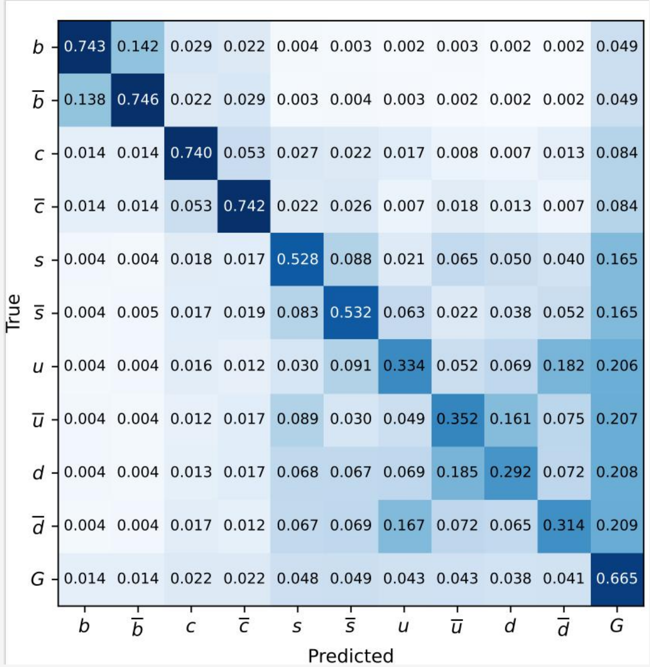
6.011

random0



6.007

random1



5.989

random2