

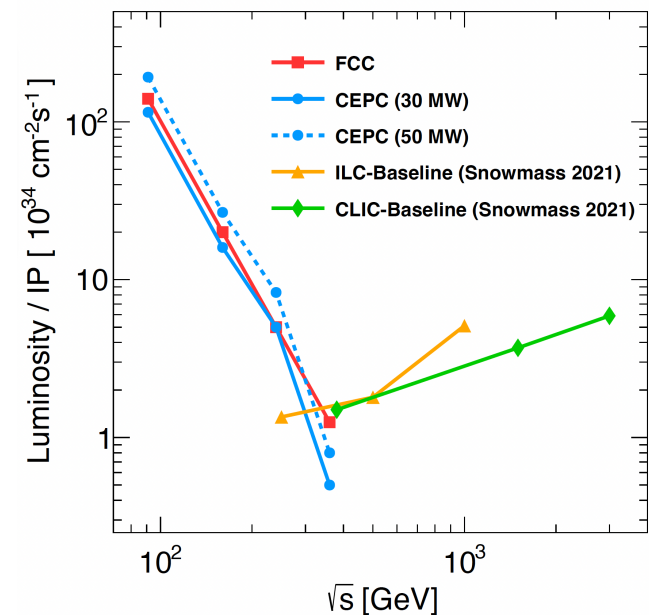
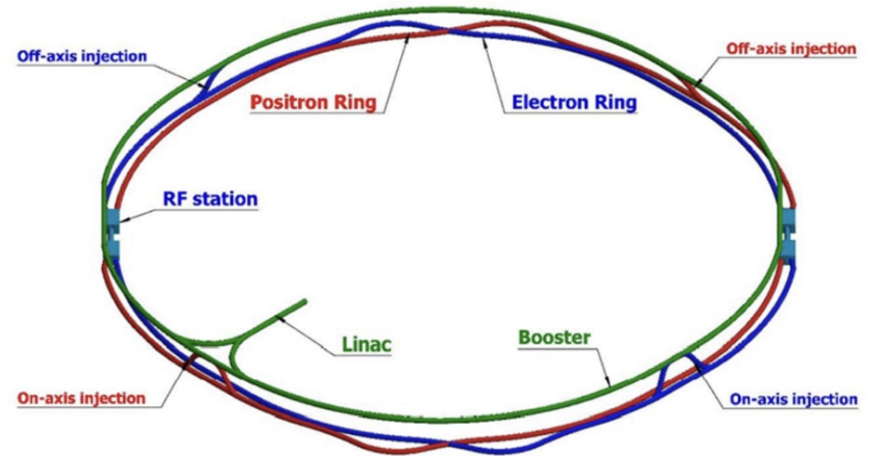


*Trilogy of event  
reconstruction at electron  
positron Higgs factory*

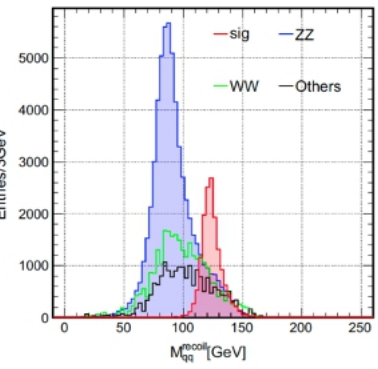
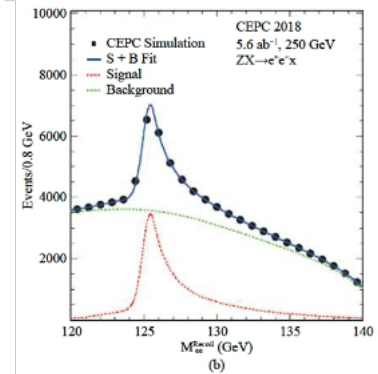
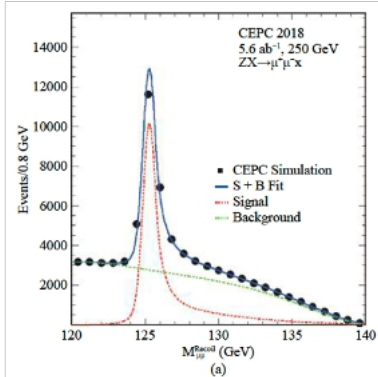
Manqi

# Outline

- CEPC Physics & Requirements
- Jet origin identification & Scaling laws at LLM
- 1-1 correspondence reconstruction
- Color Singlet identification
- Discussion



# CEPC Physics: 4 Million Higgs + 4 Tera Z...



Chinese Physics C Vol. 43, No. 4 (2019) 043002

### Precision Higgs physics at the CEPC\*

Fenfen An(安芬芬)<sup>4,23</sup> Yu Bai(白羽)<sup>6</sup> Chunhui Chen(陈春晖)<sup>23</sup> Xin Chen(陈新)<sup>5</sup> Zhenxing Chen(陈振兴)<sup>8</sup> Joao Guimaraes da Costa<sup>1</sup> Zhenwei Cui(崔振威)<sup>7</sup> Yaquan Fang(方亚泉)<sup>4,6,34,35</sup> Chengdong Fu(付成栋)<sup>1</sup> Jun Gao(高俊)<sup>23</sup> Yanyan Gao(高艳彦)<sup>22</sup> Yuanming Gao(高原宁)<sup>7</sup> Shaofeng Ge(葛韶锋)<sup>12,29</sup> Jiayin Gu(顾嘉茵)<sup>13,20</sup> Fangyi Guo(郭方毅)<sup>4,4</sup> Jun Guo(郭军)<sup>7</sup> Tao Han(韩涛)<sup>3,11</sup> Shuang Han(韩爽)<sup>4</sup> Hongjian He(何建建)<sup>11,18</sup> Xianke He(何显柯)<sup>16</sup> Xiaogang He(何小刚)<sup>11,16,20</sup> Jifeng Hu(胡继峰)<sup>16</sup> Shih-Chieh Hsu(徐士杰)<sup>12</sup> Shan Jin(金山)<sup>8</sup> Maoqiang Jing(荆茂强)<sup>3,5</sup> Susmita JyotiShamti<sup>33</sup> Ryuta Kinoshita<sup>4</sup> Chia-Ming Kuo(郭家裕)<sup>11</sup> Peizhu Lai(赖培筑)<sup>3</sup> Boyang Li(李博扬)<sup>7</sup> Congqiao Li(李聪乔)<sup>7</sup> Gang Li(李刚)<sup>4,34,35</sup> Haifeng Li(李海峰)<sup>12</sup> Liang Li(李亮)<sup>19</sup> Shu Li(李数)<sup>11,19</sup> Tong Li(李通)<sup>12</sup> Qiang Li(李强)<sup>3</sup> Hao Liang(梁浩)<sup>4,6</sup> Zhijun Liang(梁志均)<sup>4</sup> Libo Liao(廖立波)<sup>4</sup> Bo Liu(刘波)<sup>4,23</sup> Jianbei Liu(刘建北)<sup>3</sup> Tao Liu(刘涛)<sup>14</sup> Zhen Liu(刘真)<sup>28,36,40</sup> Xinchou Lou(娄辛丑)<sup>4,43,44</sup> Lianliang Ma(马连良)<sup>12</sup> Bruce Mellado<sup>13,18</sup> Xin Mo(莫欣)<sup>4</sup> Mila Pandurovic<sup>16</sup> Jianming Qian(钱剑明)<sup>34,35</sup> Zhaomou Qian(钱卓昊)<sup>19</sup> Nikolaos Rempotis<sup>22</sup> Manqi Ruan(阮庆奇)<sup>40</sup> Alex Schryr<sup>32</sup> Liangyou Shan(单连友)<sup>7</sup> Jingyuan Shi(史静远)<sup>7</sup> Xin Shi(史欣)<sup>4</sup> Shufang Su(苏淑芳)<sup>7,7</sup> Dayong Wang(王大勇)<sup>7</sup> Jun Wang(王隼)<sup>4</sup> Liantao Wang(王连涛)<sup>7,7,7</sup> Yifeng Wang(王贻芳)<sup>4,6</sup> Yuqian Wei(魏或琦)<sup>4</sup> Yue Xu(许悦)<sup>3</sup> Haijun Yang(杨海军)<sup>30,31</sup> Ying Yang(杨迎)<sup>4</sup> Weiming Yao(姚为民)<sup>38</sup> Dan Yu(于丹)<sup>4</sup> Kaiji Zhang(张凯栗)<sup>4,49</sup> Zhaoru Zhang(张照雷)<sup>4</sup> Mingrui Zhao(赵明锐)<sup>7</sup> Xiaohu Zhao(赵祥虎)<sup>4</sup> Ning Zhou(周宁)<sup>10</sup>

<sup>1</sup>Department of Modern Physics, University of Science and Technology of China, Anhui 230026, China  
<sup>2</sup>China Institute of Atomic Energy, Beijing 102413, China  
<sup>3</sup>School of Physics, Peking University, Beijing 100871, China  
<sup>4</sup>Institute of High Energy Physics, Beijing 100049, China  
<sup>5</sup>Department of Engineering Physics, Physics Department, Tsinghua University, Beijing 100084, China  
<sup>6</sup>University of Chinese Academy of Science (UCAS), Beijing 100049, China  
<sup>7</sup>School of Nuclear Science and Technology, University of South China, Hengyang 421001, China  
<sup>8</sup>Department of Physics, Nanjing University, Nanjing 210093, China  
<sup>9</sup>Department of Physics, Southeast University, Nanjing 210096, China  
<sup>10</sup>School of Physics and Astronomy, Shanghai Jiao Tong University, KLPPAC-MoE, SKLPPC, Shanghai 200240, China  
<sup>11</sup>Tung-Dao Lee Institute, Shanghai 200240, China  
<sup>12</sup>Institute of Frontier and Interdisciplinary Science and Key Laboratory of Particle Physics and Particle Irradiation (MOE), Shandong University, Qingdao 266237, China  
<sup>13</sup>PRISMA Cluster of Excellence & Mainz Institute of Theoretical Physics, Johannes Gutenberg-Universität Mainz, Mainz 55128, Germany  
<sup>14</sup>Department of Physics, Hong Kong University of Science and Technology, Hong Kong  
<sup>15</sup>Kevin IPMU (WPI), UTIAS, The University of Tokyo, Kashiwa, Chiba 277-8583, Japan  
<sup>16</sup>Vincas Institute of Nuclear Sciences, University of Belgrade, Belgrade 11000, Serbia  
<sup>17</sup>School of Physics and Institute for Collider Particle Physics, University of the Witwatersrand, Johannesburg 2050, South Africa

Received 9 November 2018, Revised 21 January 2019, Published online 4 March 2019  
 \* Supported by the National Key Program for S&T Research and Development (2016YFA0400400); CAS Center for Excellence in Particle Physics; Yifeng Wang's Science Studio of the Ten Thousand Talents Project; the CAS-SAFE International Partnership Project for Creative Research Teams (H2150185); BHP Innovation Grant (Y454517072); Key Research Program of Frontier Sciences, CAS (SQZDY-SW-SL3002); Chinese Academy of Science Special Grant for Large Scientific Project (11311KYSB017005); the National Natural Science Foundation of China(11675203); the Hundred Talent Programs of Chinese Academy of Science (Y151540U1); the National 1000 Talents Program of China; From Research Alliance, LLC (DE-AC02-07CH11359); the NSF(PHY1610074); by the Maryland Center for Fundamental Physics (MCFP); Tsinghua University Initiative Scientific Research Program; and the Beijing Municipal Science and Technology Commission project(Z211100004218003)

1) E-mail: fangyi@ihep.ac.cn  
 2) E-mail: jiang@ustc.edu.cn  
 3) E-mail: li.gang@mail.ihep.ac.cn  
 4) E-mail: zhangyhs@ustc.edu.cn  
 5) E-mail: qiong@ustc.edu.cn  
 6) E-mail: manqi.ruan@ihep.ac.cn  
 7) E-mail: liantao@ustc.edu.cn  
 8) E-mail: zhangk1@ihep.ac.cn

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI. Article funded by SCOAP<sup>3</sup> and published under licence by Chinese Physical Society and the Institute of High Energy Physics of the Chinese Academy of Sciences and the Institute of Modern Physics of the Chinese Academy of Sciences and IOP Publishing Ltd

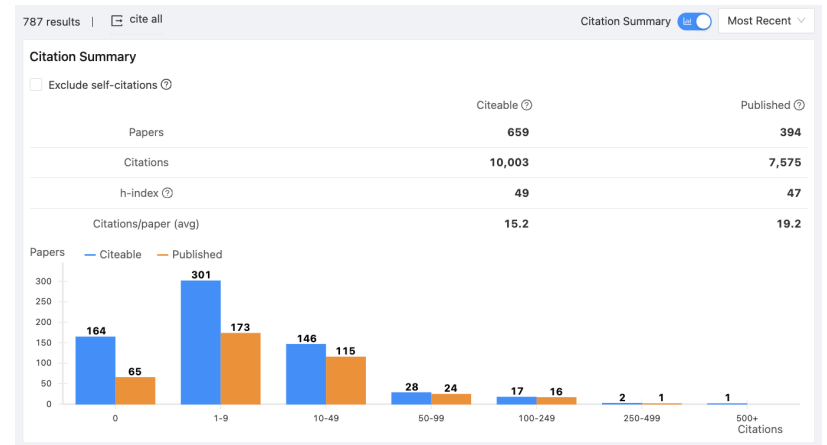


Table 2.1: Precision of the main parameters of interests and observables at the CEPC, from Ref. [1] and the references therein, where the results of Higgs are estimated with a data sample of 20 ab<sup>-1</sup>. The HL-LHC projections of 3000 fb<sup>-1</sup> data are used for comparison. [2]

Observable	Higgs		W, Z and top	
	HL-LHC projections	CEPC precision	Observable	Current precision
$M_H$	20 MeV	3 MeV	$M_W$	9 MeV
$\Gamma_H$	20%	1.7%	$\Gamma_W$	49 MeV
$\sigma(ZH)$	4.2%	0.26%	$M_Z$	2.1 MeV
$B(H \rightarrow bb)$	4.4%	0.14%	$\Gamma_Z$	2.3 MeV
$B(H \rightarrow cc)$	-	2.0%	$M_{top}$	760 MeV
$B(H \rightarrow gg)$	-	0.81%	$\mathcal{O}(10)$	10 MeV
$B(H \rightarrow WW^*)$	2.8%	0.53%	$R_b$	$3 \times 10^{-3}$
$B(H \rightarrow ZZ^*)$	2.9%	4.2%	$R_c$	$1.7 \times 10^{-2}$
$B(H \rightarrow \tau^+\tau^-)$	2.9%	0.42%	$R_\mu$	$2 \times 10^{-3}$
$B(H \rightarrow \gamma\gamma)$	2.6%	3.0%	$R_\tau$	$1.7 \times 10^{-2}$
$B(H \rightarrow \mu^+\mu^-)$	8.2%	6.4%	$A_\mu$	$1.5 \times 10^{-2}$
$B(H \rightarrow Z\gamma)$	20%	8.5%	$A_\tau$	$4.3 \times 10^{-3}$
$B(\text{butterfly } H \rightarrow \text{inv.})$	2.5%	0.07%	$A_b$	$2 \times 10^{-2}$
			$N_\nu$	$2.5 \times 10^{-3}$

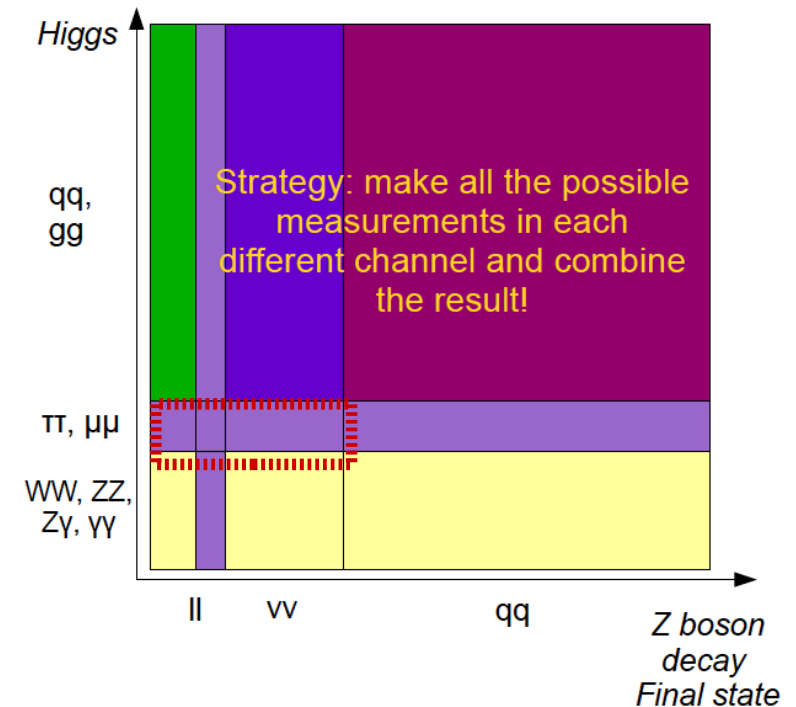
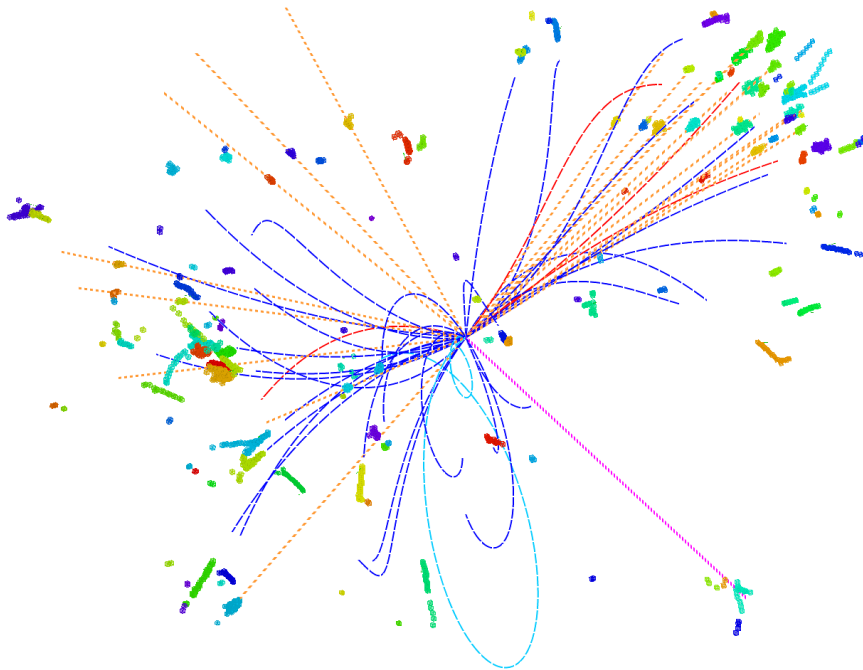
## Scientific Significance quantified by CEPC physics studies, via full simulation/phenomenology studies:

- Higgs: Precisions exceed HL-LHC ~ 1 order of magnitude.
- EW: Precision improved from current limit by 1-2 orders.
- Flavor Physics, sensitive to NP of 10 TeV or even higher.
- Sensitive to varies of NP signal.
- ...

White papers +  
 ~300 Journal/AxXiv citables

# Performance requirements

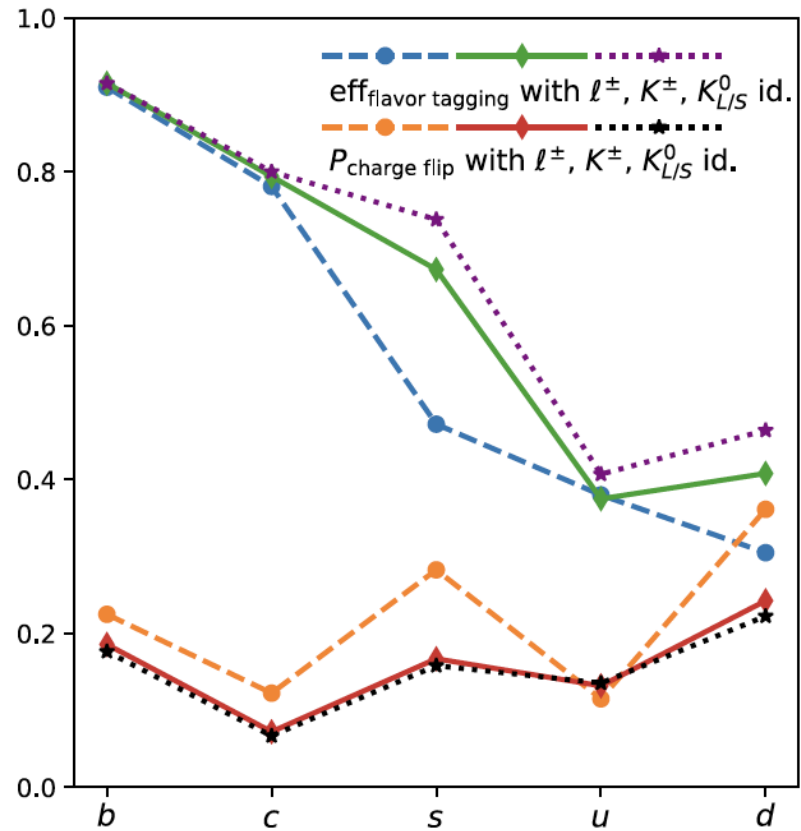
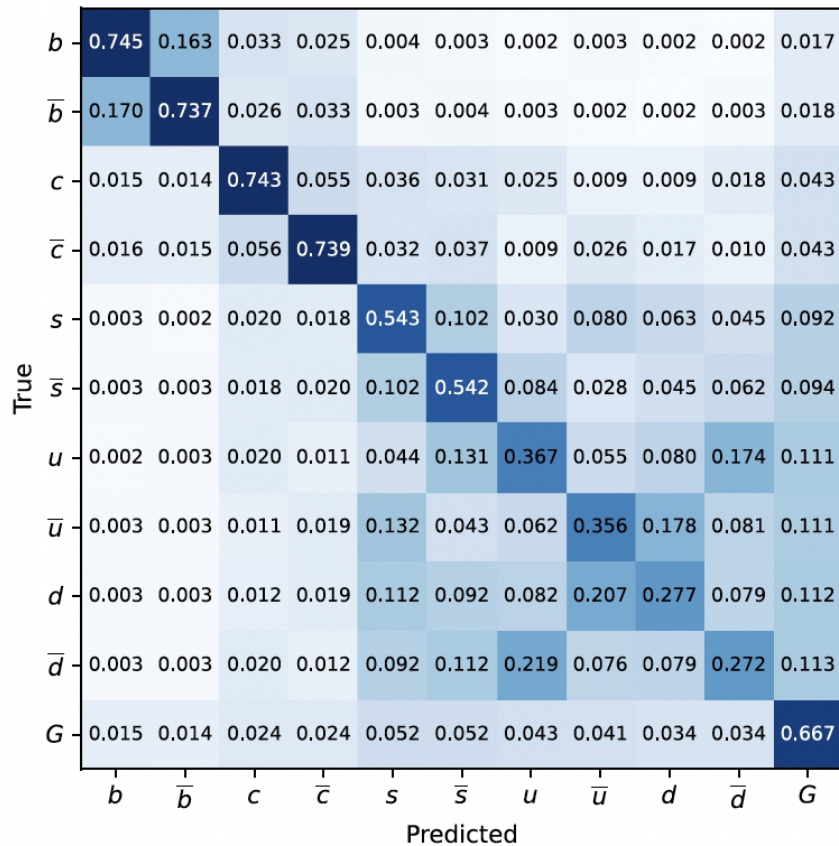
- To reconstruct all Physics Object, especially **Jets**
  - Z & W: ~ 70% goes to a pair of jets
  - Higgs: ~97% final state with jets (ZH events)
  - Top:  $t \rightarrow W + b$



- Look inside the jet: **1-1 correspondence reco.**
  - ~ **confusion free PFA**
  - Larger acceptance...
  - Excellent intrinsic resolutions
  - Extremely stable...
- Be addressed by state-of-art detector design, technology, and **reconstruction algorithm!**



# Jet origin id

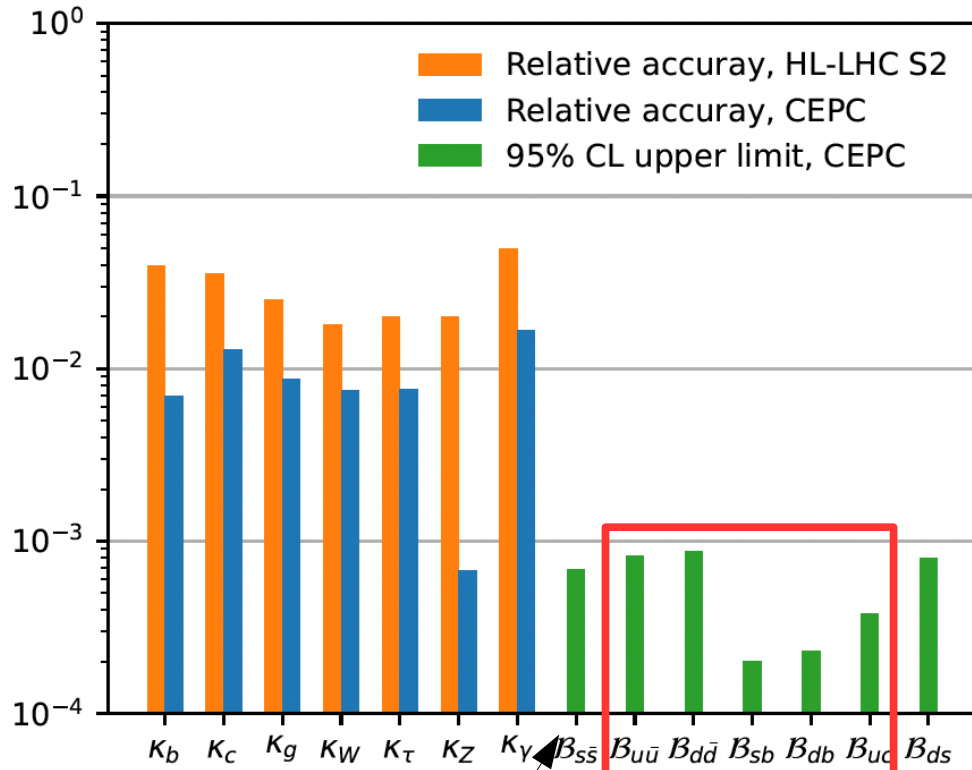


- 11 categories (5 quarks + 5 anti quarks + gluon) identification, realized at Full Simulated di-jet events at CEPC CDR baseline with **Arbor + ParticleNet**
- Published in PRL 132, 221802 (2024). Comment from the referee: *"demonstrate the world-leading performance of tagger", "a "game changer" and opens new horizons for precision flavor studies at all future experiments."*

<https://arxiv.org/abs/2310.03440>

<https://arxiv.org/abs/2309.13231>

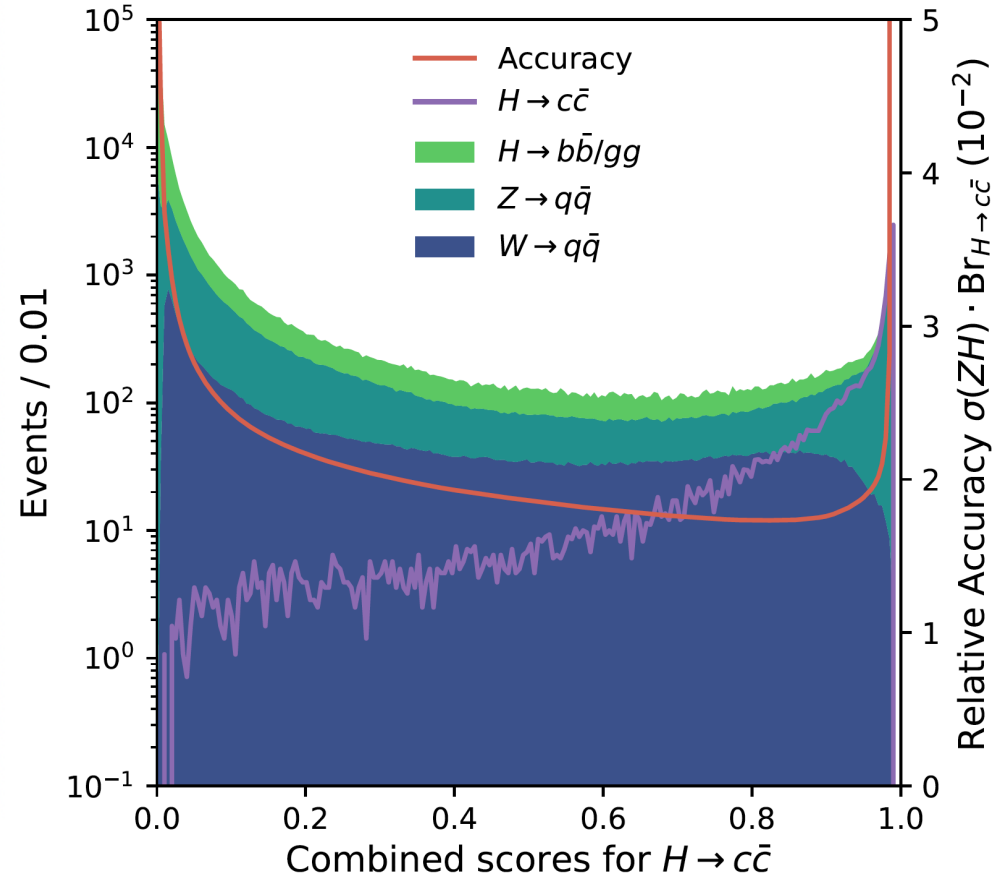
# Impact on Physics: Higgs & W



Improved by ~3 times

Improved by 1-2 orders of magnitudes

Presumably... firstly quantified



• Compared to Conventional :

- $v\bar{v}H, H \rightarrow c\bar{c}$ : 3%  $\rightarrow$  1.7%
- $V_{cb}$ : 0.75%  $\rightarrow$  0.5%
- Applicable to  $V_{cs}, V_{ts}$ , etc.

# Updated result on $\sin^2 \theta_{eff}^l$ measurement

**Table 2.** Sensitivity  $S$  of different final state particles.

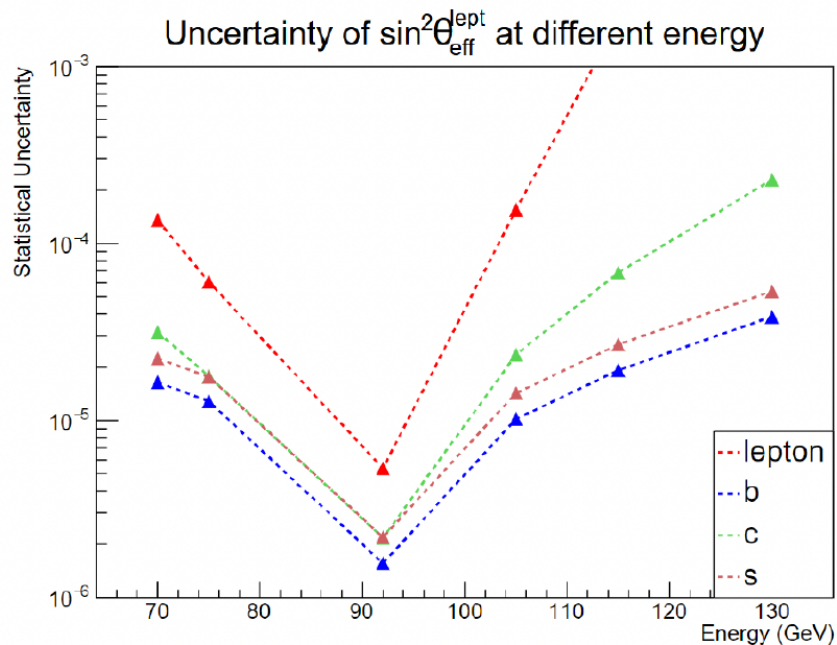
$\sqrt{s}/\text{GeV}$	$S$ of $A_{FB}^{e/\mu}$	$S$ of $A_{FB}^d$	$S$ of $A_{FB}^u$	$S$ of $A_{FB}^s$	$S$ of $A_{FB}^c$	$S$ of $A_{FB}^b$
70	0.224	4.396	1.435	4.403	1.445	4.352
75	0.530	5.264	2.598	5.269	2.616	5.237
92	1.644	5.553	4.200	5.553	4.201	5.549
105	0.269	4.597	1.993	4.598	1.994	4.586
115	0.035	3.956	1.091	3.958	1.087	3.942
130	0.027	3.279	0.531	3.280	0.520	3.261

**Table 3.** Cross section of process  $e^+e^- \rightarrow f\bar{f}$  calculated using the ZFITTER package. Values of the fundamental parameters are set as  $m_Z = 91.1875 \text{ GeV}$ ,  $m_t = 173.2 \text{ GeV}$ ,  $m_H = 125 \text{ GeV}$ ,  $\alpha_s = 0.118$  and  $m_W = 80.38 \text{ GeV}$ .

$\sqrt{s}/\text{GeV}$	$\sigma_\mu/\text{mb}$	$\sigma_d/\text{mb}$	$\sigma_u/\text{mb}$	$\sigma_s/\text{mb}$	$\sigma_c/\text{mb}$	$\sigma_b/\text{mb}$
70	0.039	0.032	0.066	0.031	0.058	0.028
75	0.039	0.047	0.073	0.046	0.065	0.043
92	1.196	5.366	4.228	5.366	4.222	5.268
105	0.075	0.271	0.231	0.271	0.227	0.265
115	0.042	0.135	0.122	0.135	0.118	0.132
130	0.026	0.071	0.068	0.071	0.066	0.069

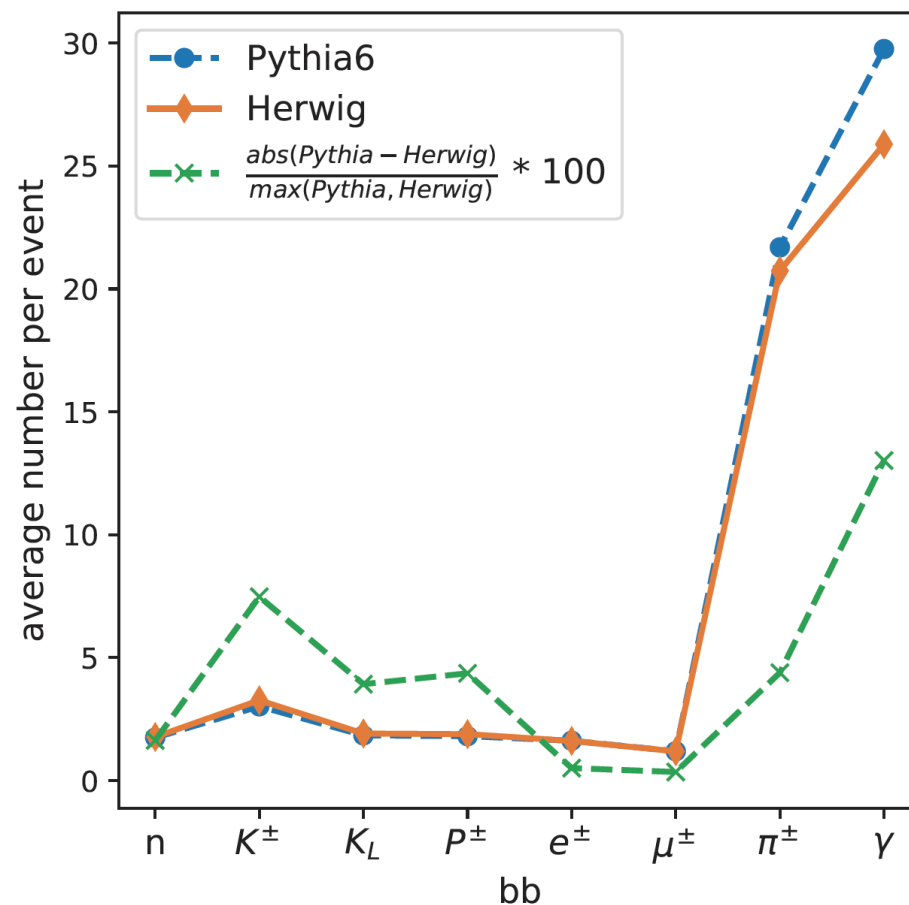
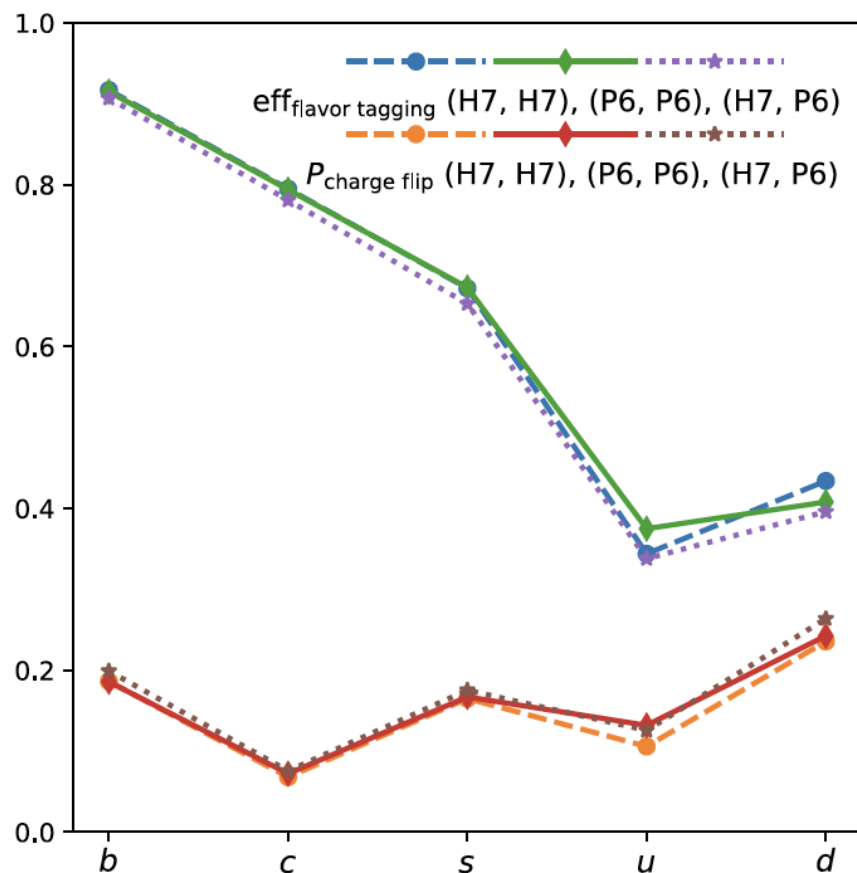
Verify the RG behavior... using  
~1 month of data taking

**Expected statistical uncertainties on  $\sin^2 \theta_{eff}^l$  measurement.**  
(Using one-month data collection, ~ **4e12/24 Z events** at Z pole)



$\sqrt{s}$	$b$	$c$	$s$
70	$1.6 \times 10^{-5}$	$3.2 \times 10^{-5}$	$2.2 \times 10^{-5}$
75	$1.3 \times 10^{-5}$	$1.8 \times 10^{-5}$	$1.8 \times 10^{-5}$
92	$1.6 \times 10^{-6}$	$2.2 \times 10^{-6}$	$2.2 \times 10^{-6}$
105	$1.0 \times 10^{-5}$	$2.4 \times 10^{-5}$	$1.4 \times 10^{-5}$
115	$1.9 \times 10^{-5}$	$6.8 \times 10^{-5}$	$2.7 \times 10^{-5}$
130	$3.9 \times 10^{-5}$	$2.3 \times 10^{-4}$	$5.4 \times 10^{-5}$

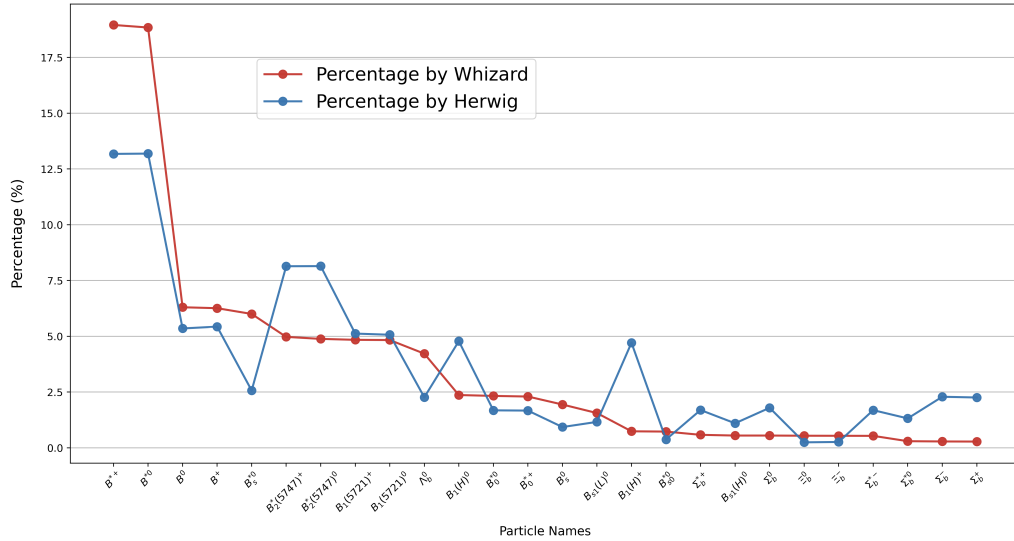
# V.S. Hadronization models



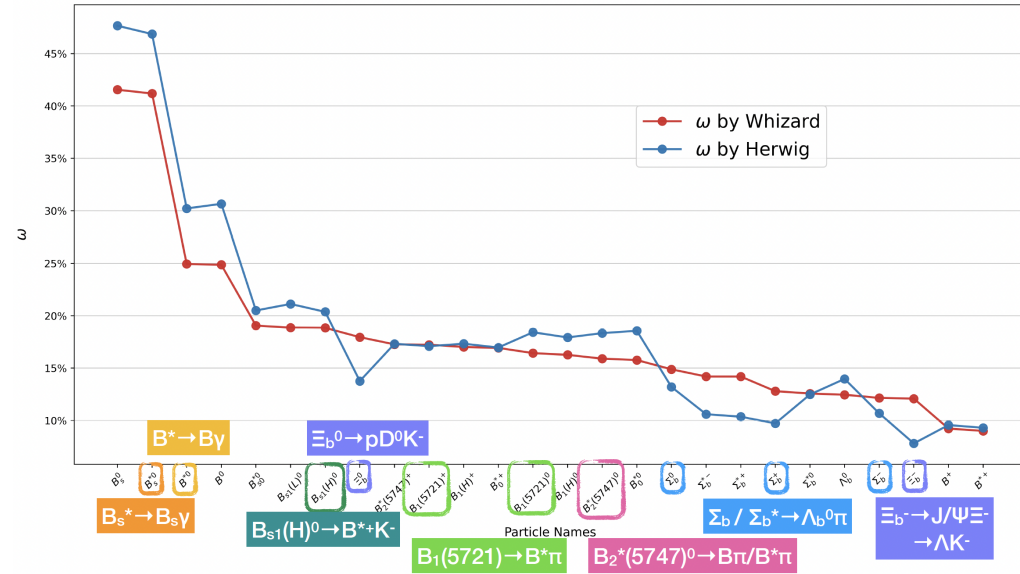
- Different hadronization model have significantly different predictions...

# b-jet: leading b-hadrons & flip rates

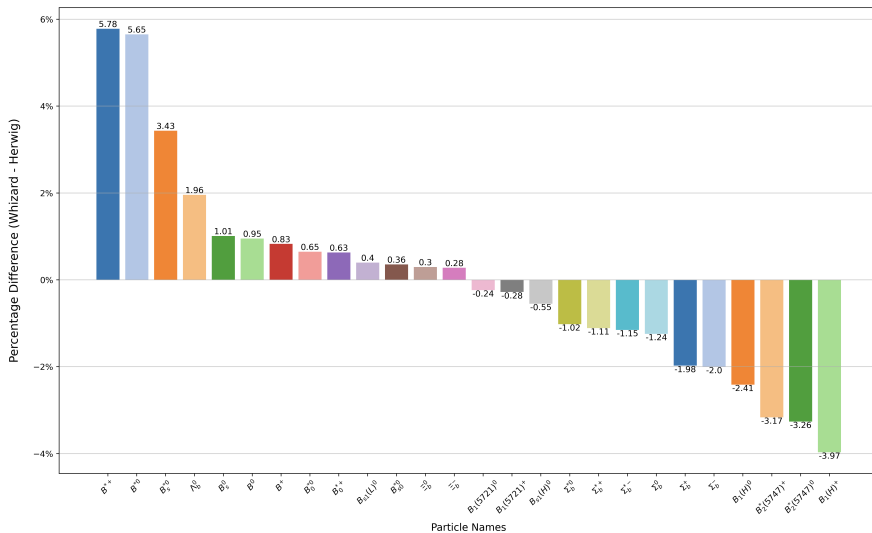
Percentage of b hadrons by Whizard & Herwig



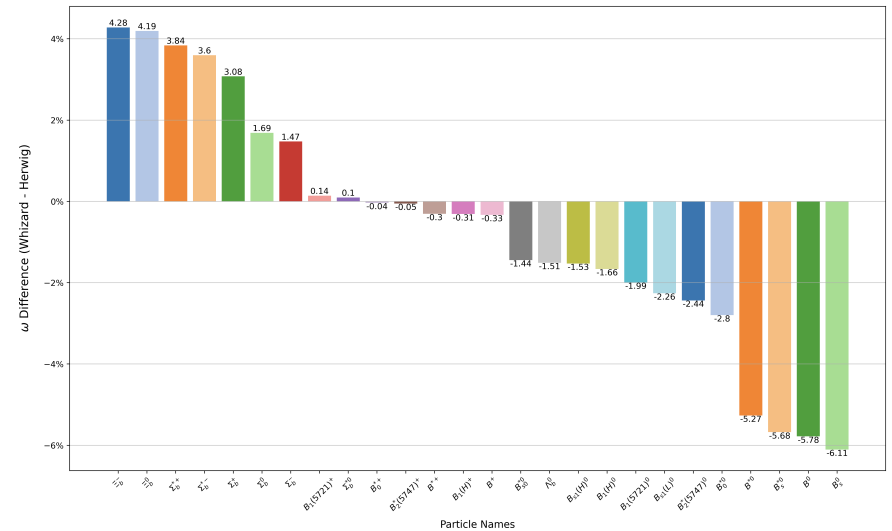
Charge Flip Rate ω of b hadrons by Whizard & Herwig



Difference in Percentage of b hadrons between Whizard and Herwig

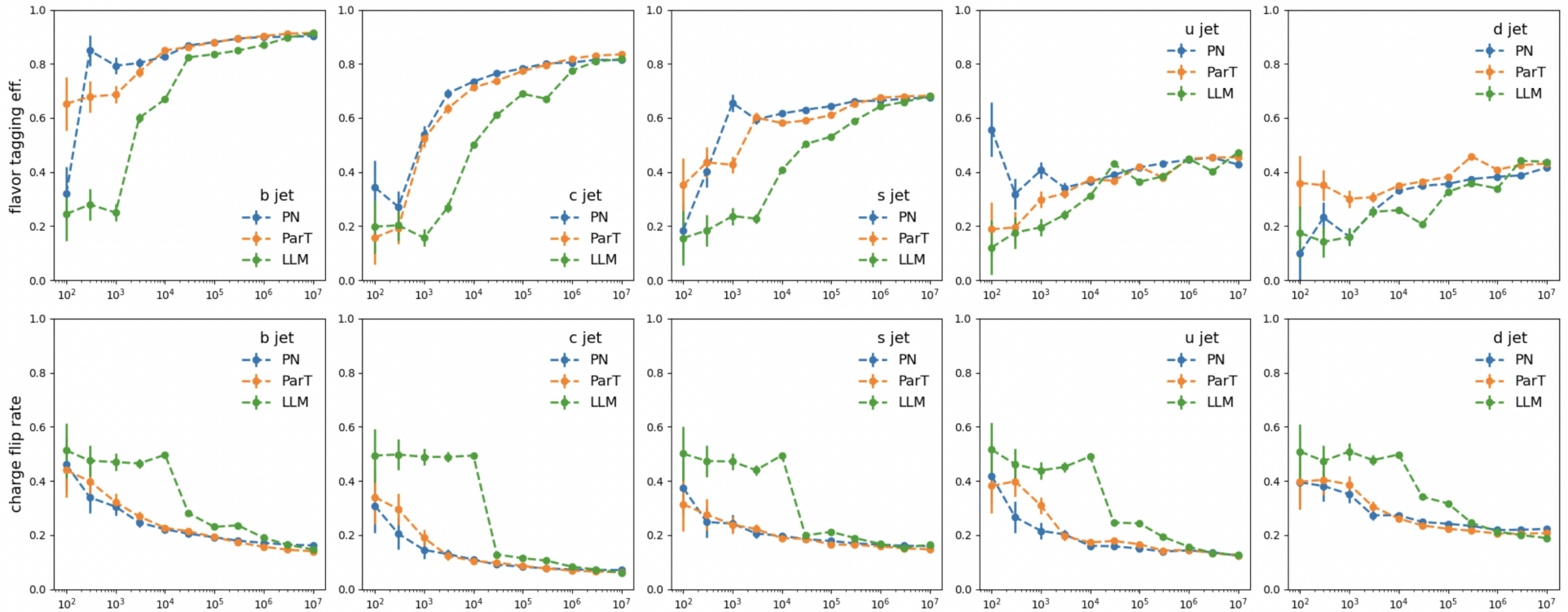


Difference in Charge Flip Rate ω of b hadrons between Whizard and Herwig





# From specialized Models to LLM

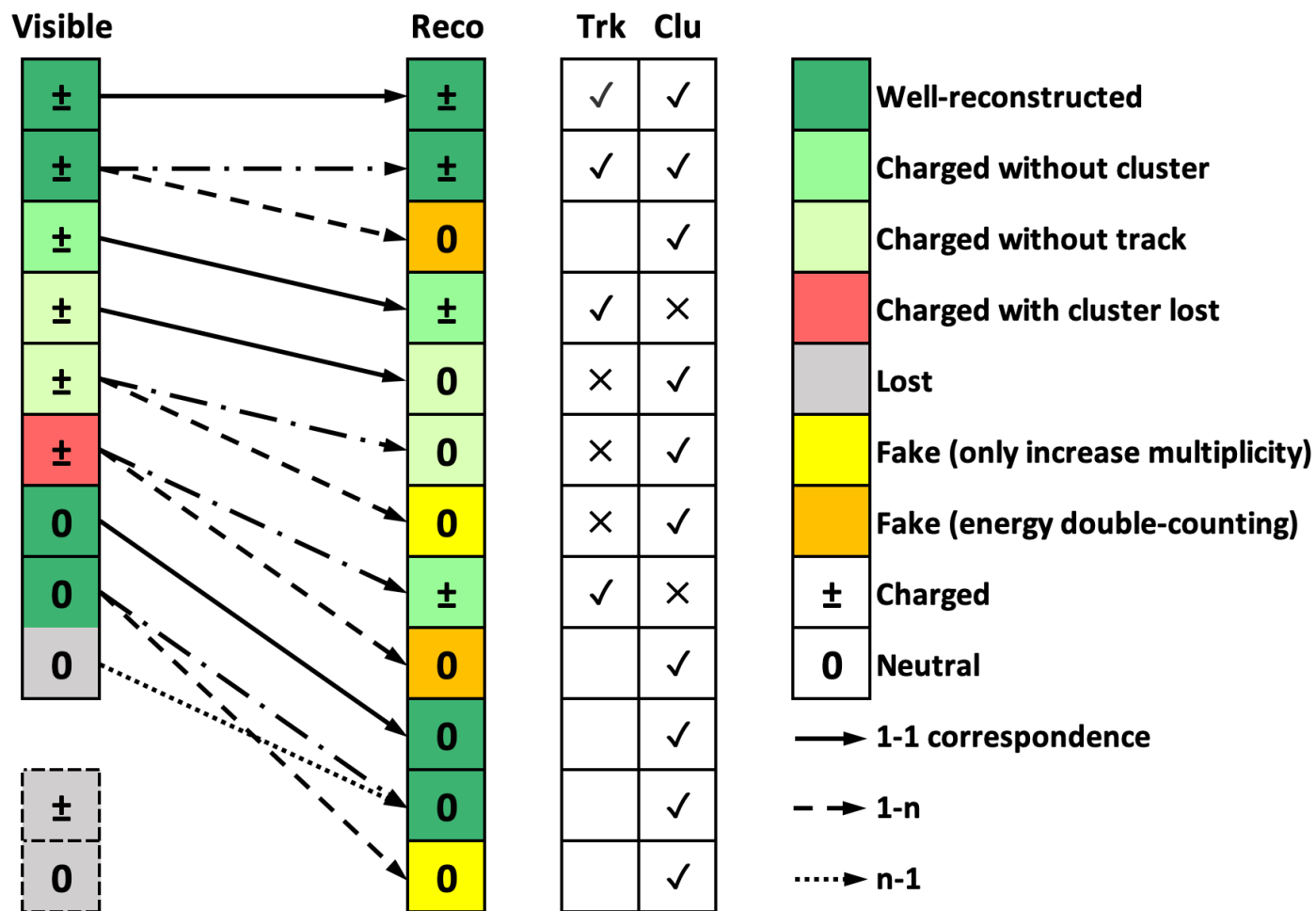


- Comparable result with different scaling behavior
- Para. Numbers: PN 360k, ParT 2.4M, BINBBT(Large Language Base Model) 150 M
- More details at: <https://arxiv.org/pdf/2412.00129>

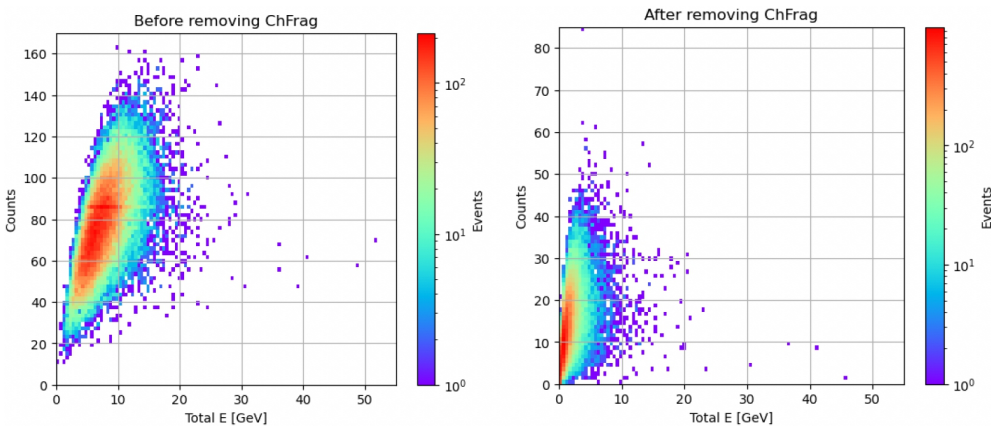


超对称  
Super Symmetry  
Technologies

# 1-1 correspondence: ultimate Mapping between visible & reco

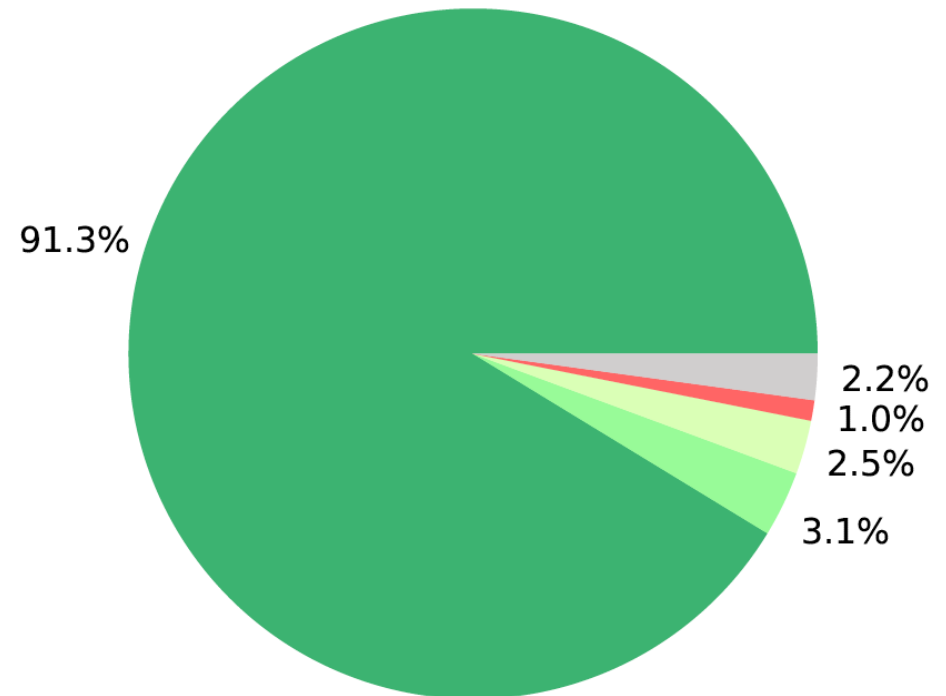


# Confusion: frag. Identification & veto

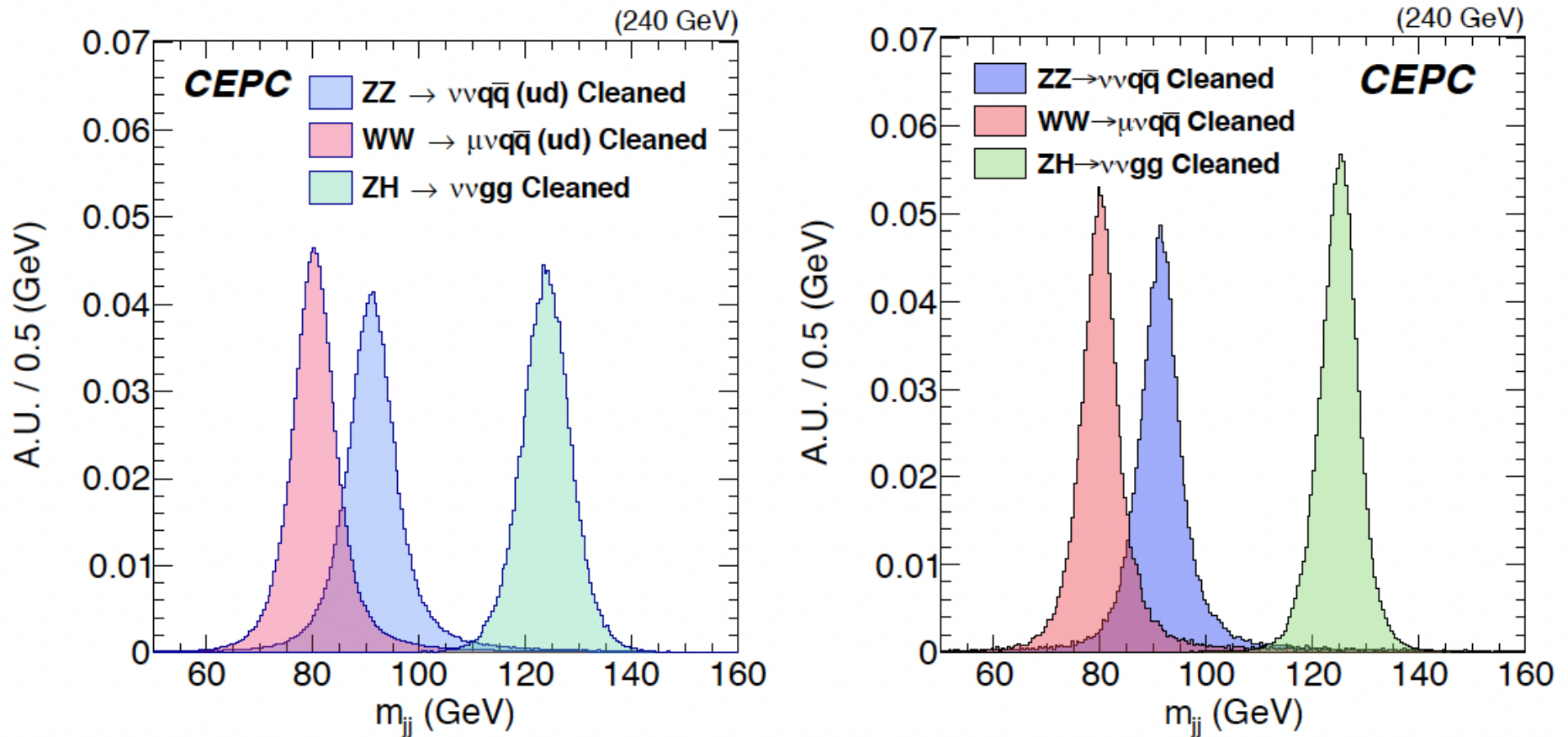


Fake particle originated Confusion reduced by 1 order of magnitude, at nominal  $\nu\nu H$ ,  $H \rightarrow gg$  event

Ignoring the remaining fragments with total  $E < 1$  GeV, more than 95% of the visible energy preserves 1-1 correspondence



# BMR of 2.75% reached

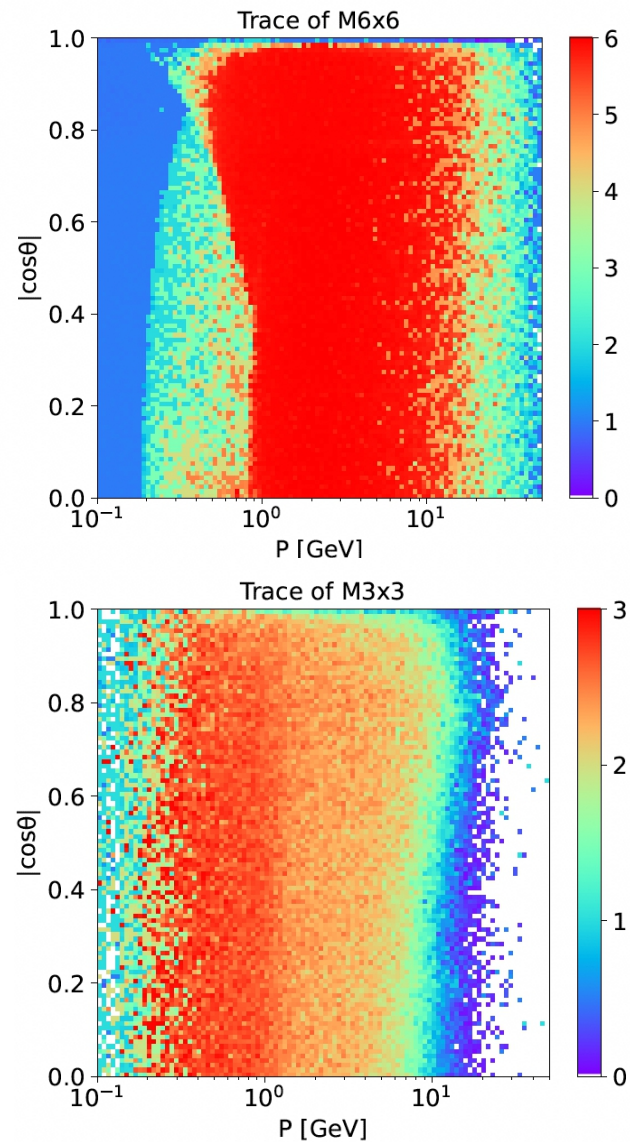
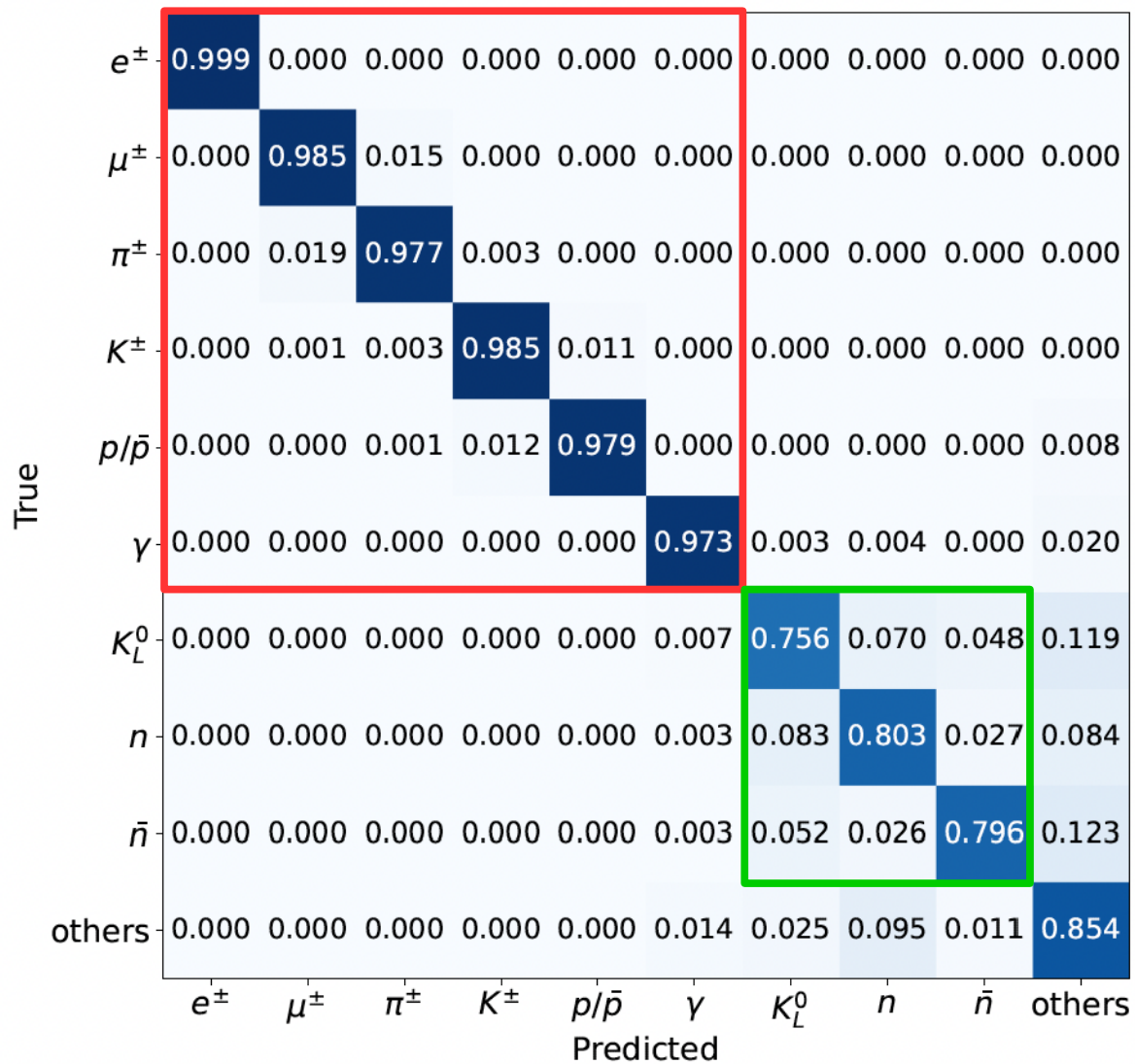


Detector change (usage of high density scintillating glass HCAL): BMR 3.7%  $\rightarrow$  3.4%;

AI enhanced reconstruction: 3.4%  $\rightarrow$  2.8%.

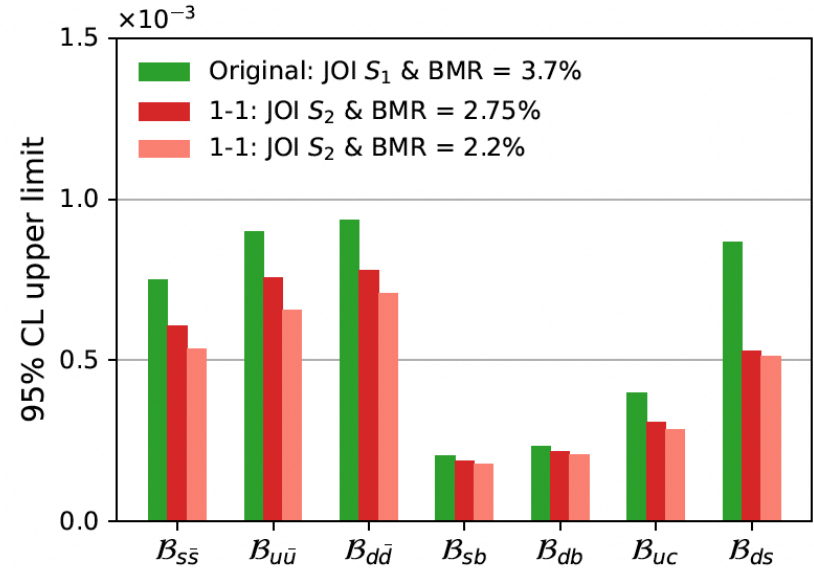
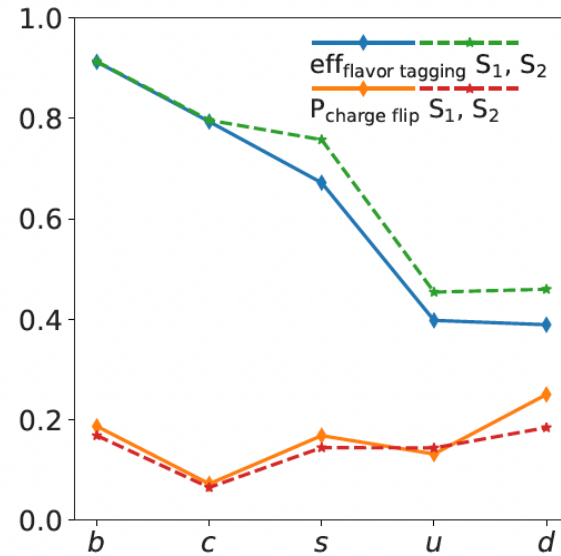
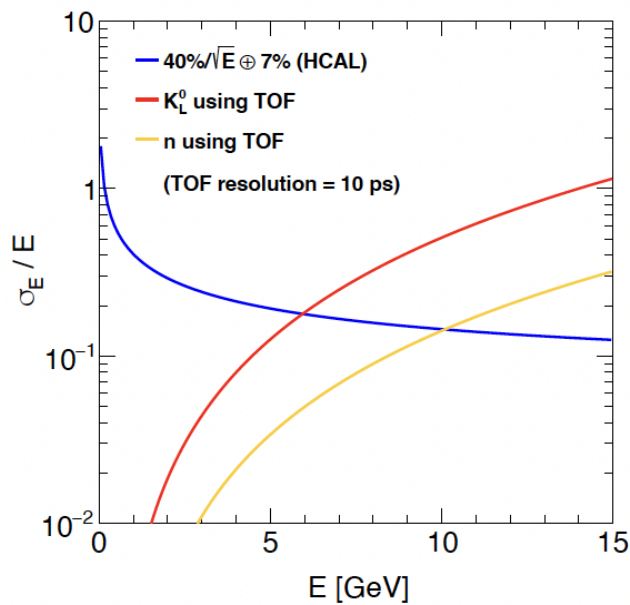
Impact from Beam induced background + impact on objects inside jet reco: to be evaluated.

# Pid: differential performance





# Perspectives with 1-1 correspondence



- ToF enhanced energy measurement: BMR: 2.8  $\rightarrow$  2.2-2.4
  - Need excellent CALO + ToF  $\sim$  o(10 ps)
  - Assume Low energy neutrons & secondary particles can be tamed... still very challenge...
- Strongly Boost the light quark ID.
- Benchmark precision improved... up to nearly two times.

# Color Singlet Identification



PUBLISHED FOR SISSA BY SPRINGER

RECEIVED: March 11, 2022

REVISED: September 9, 2022

ACCEPTED: November 11, 2022

PUBLISHED: November 16, 2022

JHEP11(2022)100

## The Higgs $\rightarrow b\bar{b}, c\bar{c}, gg$ measurement at CEPC

Yongfeng Zhu, Hanhua Cui and Manqi Ruan

*Institute of High Energy Physics, Chinese Academy of Sciences,  
19B Yuquan Road, Beijing 100049, China*

*University of Chinese Academy of Sciences,  
19A Yuquan Road, Beijing 100049, China*

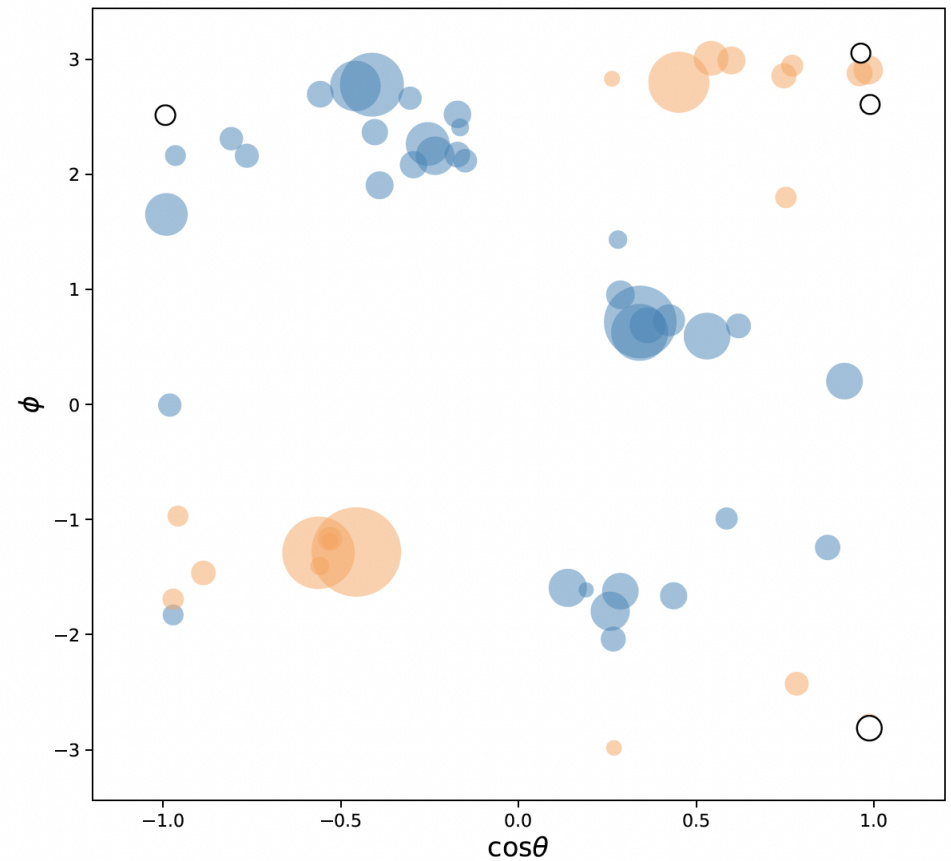
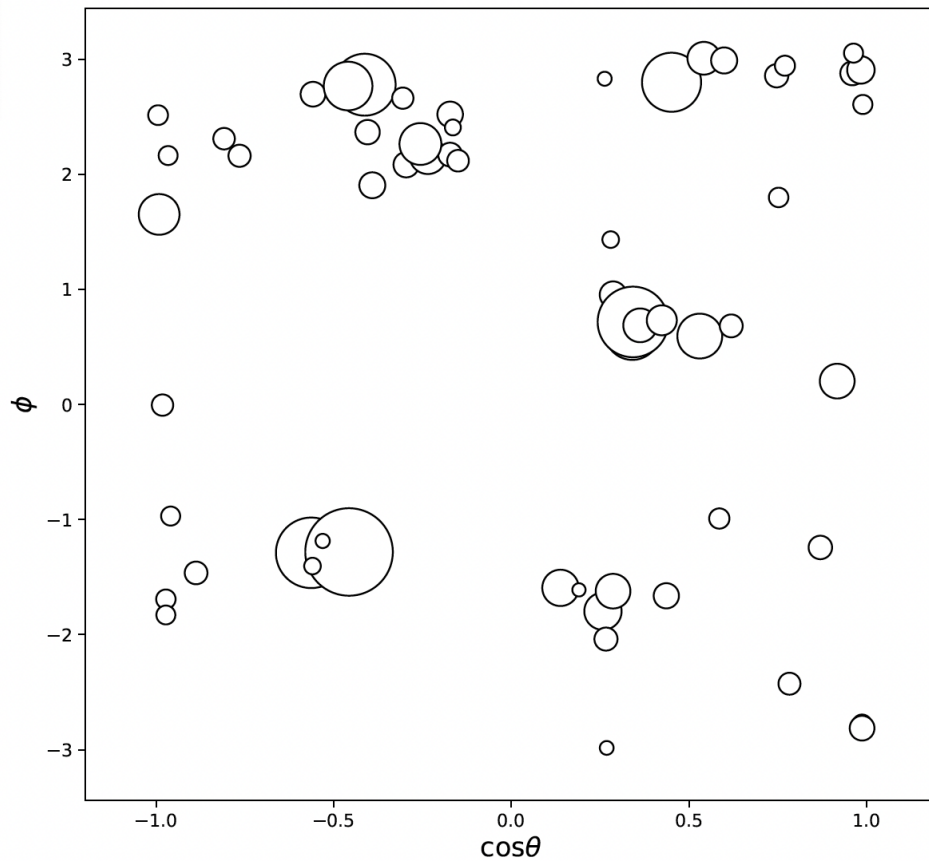
E-mail: [ruanmq@ihep.ac.cn](mailto:ruanmq@ihep.ac.cn)

Z decay mode	$H \rightarrow b\bar{b}$	$H \rightarrow c\bar{c}$	$H \rightarrow gg$
$Z \rightarrow e^+e^-$	1.57%	14.43%	10.31%
$Z \rightarrow \mu^+\mu^-$	1.06%	10.16%	5.23%
$Z \rightarrow q\bar{q}$	0.35%	7.74%	3.96%
$Z \rightarrow \nu\bar{\nu}$	0.49%	5.75%	1.82%
combination	0.27%	4.03%	1.56%

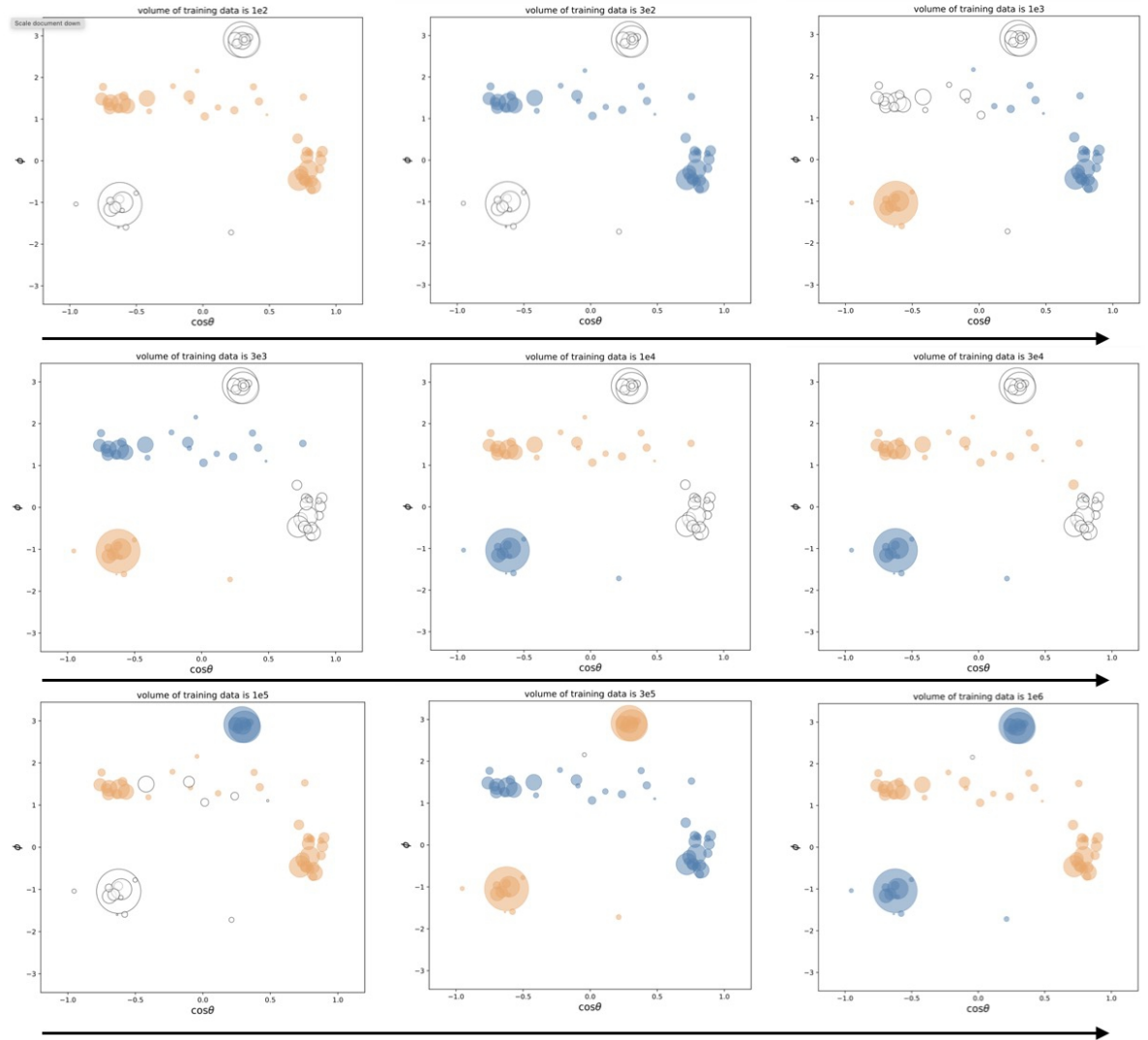
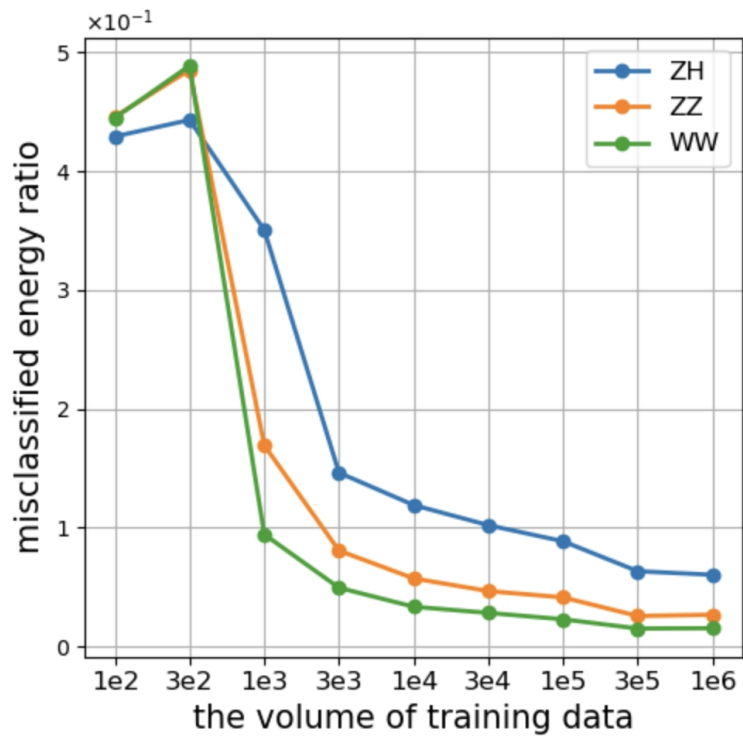
**Table 3.** The signal strength accuracies for different channels.

- $H \rightarrow cc$  &  $gg$  measurements at  $qqH$  channel is much worse  $\nu\nu H$  channels, despite the former has 3.5 times more signal statistic
- Reason: Failure of Color Singlet Identification – to distinguish the decay products of each Color Singlet
  - Z & H for 240/250 GeV Higgs factory
  - Which Higgs boson for Higgs self-coupling measurements (i.e., at  $\nu\nu HH$  events at 500 GeV, etc)

# CSI: to group the final state particle



# Scaling behavior



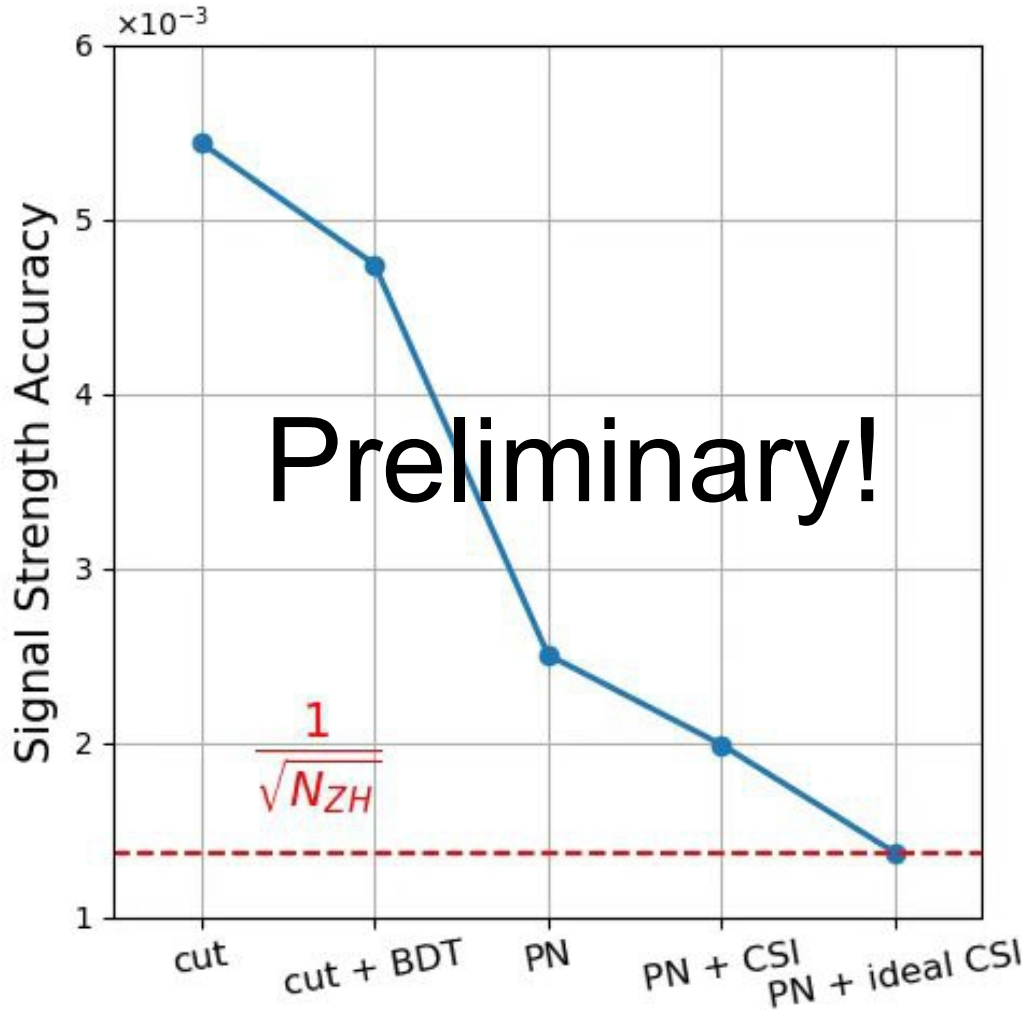
# A toy analysis: identify full hadronic ZH signal from ZZ + WW background

- Cut based
- BDT
- 1-1 correspondence
- 1-1 correspondence with reconstructed CSI
- 1-1 correspondence with truth level CSI
  
- 5.6 iab: 540k ZH + 3.1M ZZ + 47 M WW full hadronic events



# Comparison of different analysis methods

Chinese Physics C Vol. 43, No. 4 (2019) 043002



*PN ~ holistic event description via 1-1*

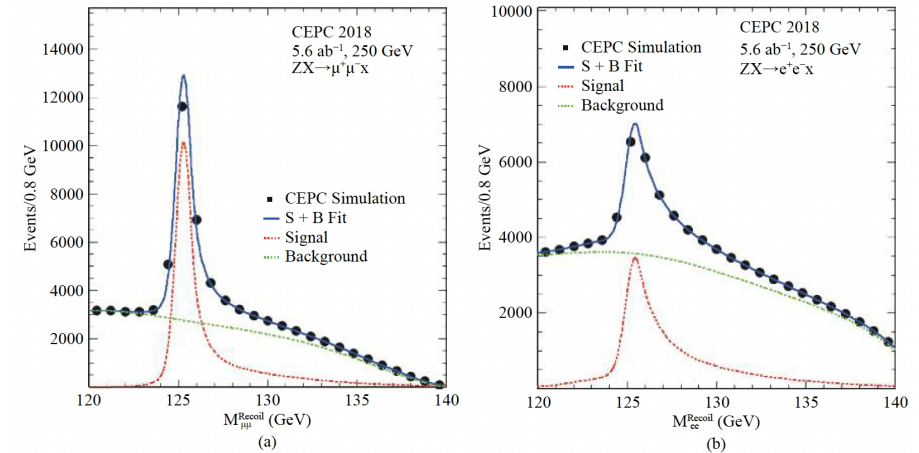


Table 5. Estimated measurement precision for the Higgs boson mass  $m_H$  and the  $e^+e^- \rightarrow ZH$  production cross section  $\sigma(ZH)$  from a CEPC dataset of  $5.6 \text{ ab}^{-1}$ .

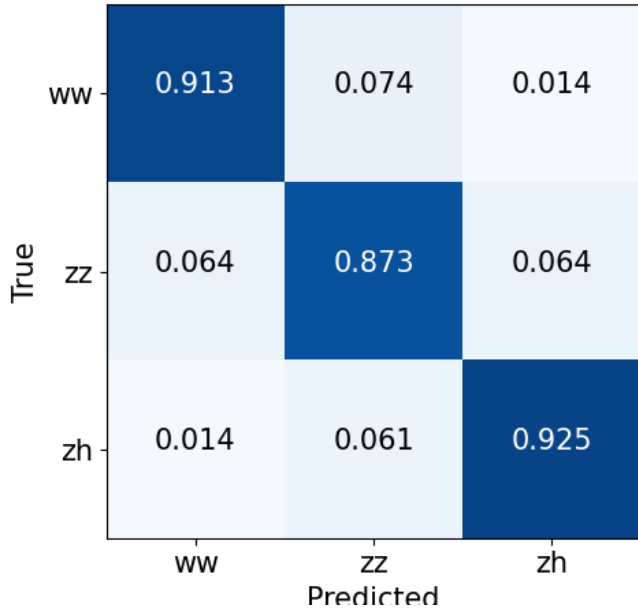
Z decay mode	$\Delta m_H/\text{MeV}$	$\Delta\sigma(ZH)/\sigma(ZH)$
$e^+e^-$	14	1.4%
$\mu^+\mu^-$	6.5	0.9%
$q\bar{q}$	—	0.6%
combination	5.9	0.5%

**Optimistic Guesses:**

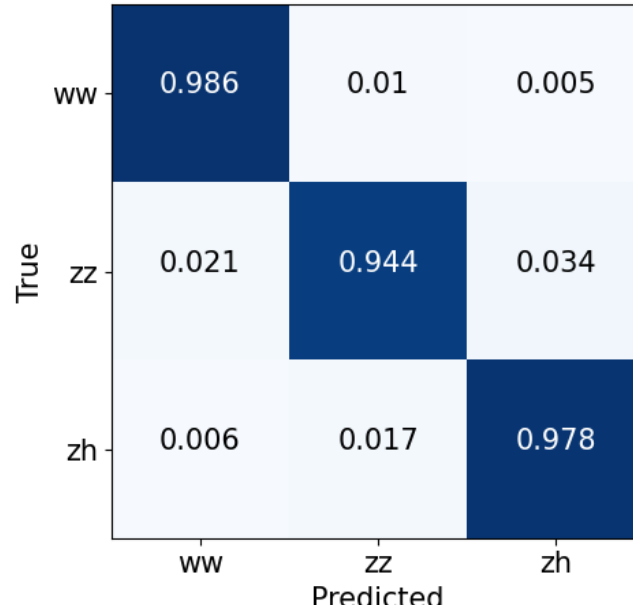
Applied to Higgs recoil analysis with  $qqH$  channel,  $\sigma(ZH)$  measurements could be improved by more than 2 times...

# Analysis with 1-1 & CSI: Preliminary!

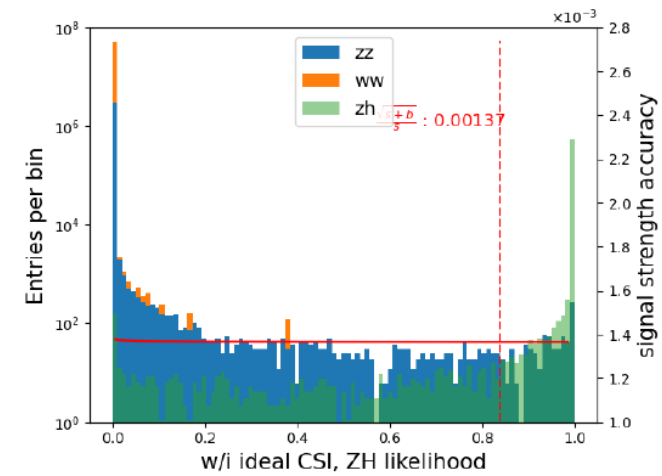
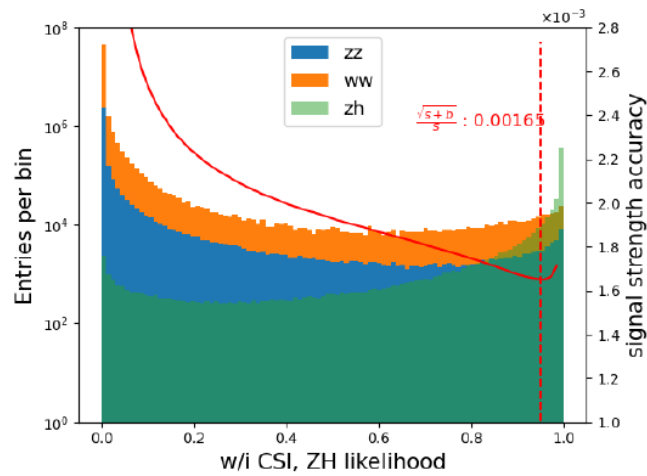
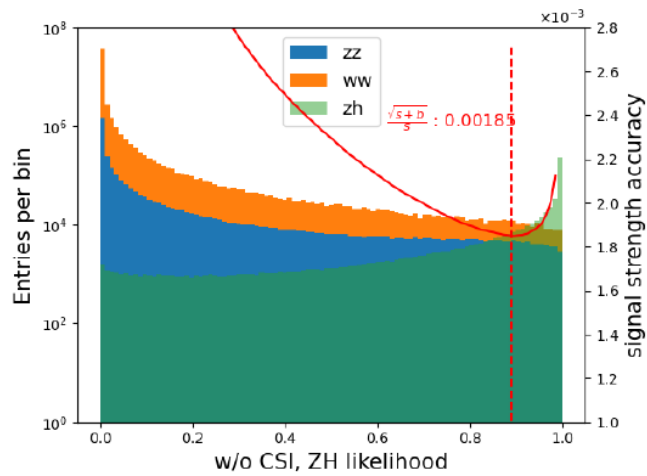
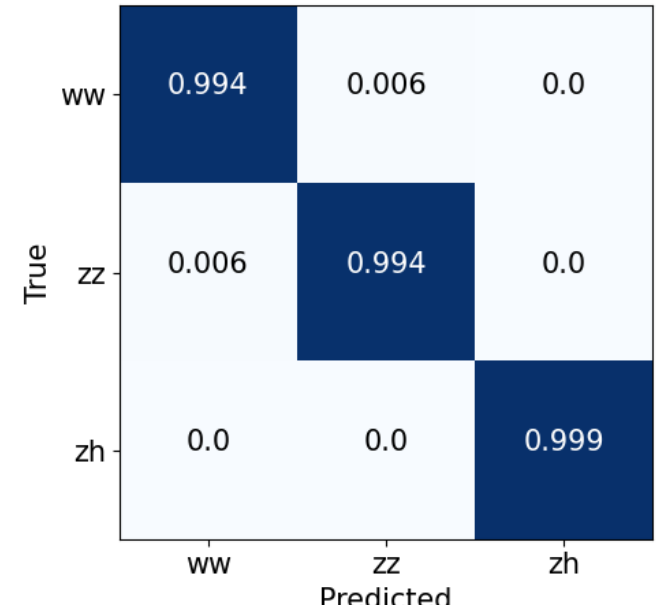
w/o CSI



w/i CSI



w/i ideal CSI



# Meta questions

- Problem categorization
  - Identification problem: Jol, Pid, 1-1 correspondence (from Arbor)
  - Grouping problem: Color singlet id, tracking, clustering, ...
  - Assessment/regression problem: such as energy/momentum/time estimation, fitting
  - What's the most suited corresponding AI architecture, or general AI, and Why?
- AI for HEP, and HEP for AI (HEP → Science)
  - HEP, as a mature & vivid field, has the potential to impact the AI development, i.e., interpretability analysis
- Be relax, and have fun!...

# Summary

- Higgs factory: extremely rich physics requires excellent performance
- **Trilogy**: Significantly enhance the discovery power & alter the experiments design
  - Jet Origin ID: 'see' quark & gluon as lepton & photon
    - *...A “game changer” and opens new horizon for precise flavor studies at all future experiments...*
  - 1-1 correspondence, at least at Higgs factory: **Should & Could**
    - New paradigm for analyses:
      - Forget about artificial variable definition – feed all the reconstructable
      - Provide much more detailed info for system monitoring & systematic control
  - Color Singlet Id: decently addressed
- Bottleneck Shifts & Lots to be explored
  - Confusion → Det. Acceptance
  - Clever variable selection → High Quality MC: better QCD modeling, high precision calculation, detector calibration – monitoring, event building...
  - Particle Physics Provides excellent benchmarks to quantify the AI performance & interpretability study...

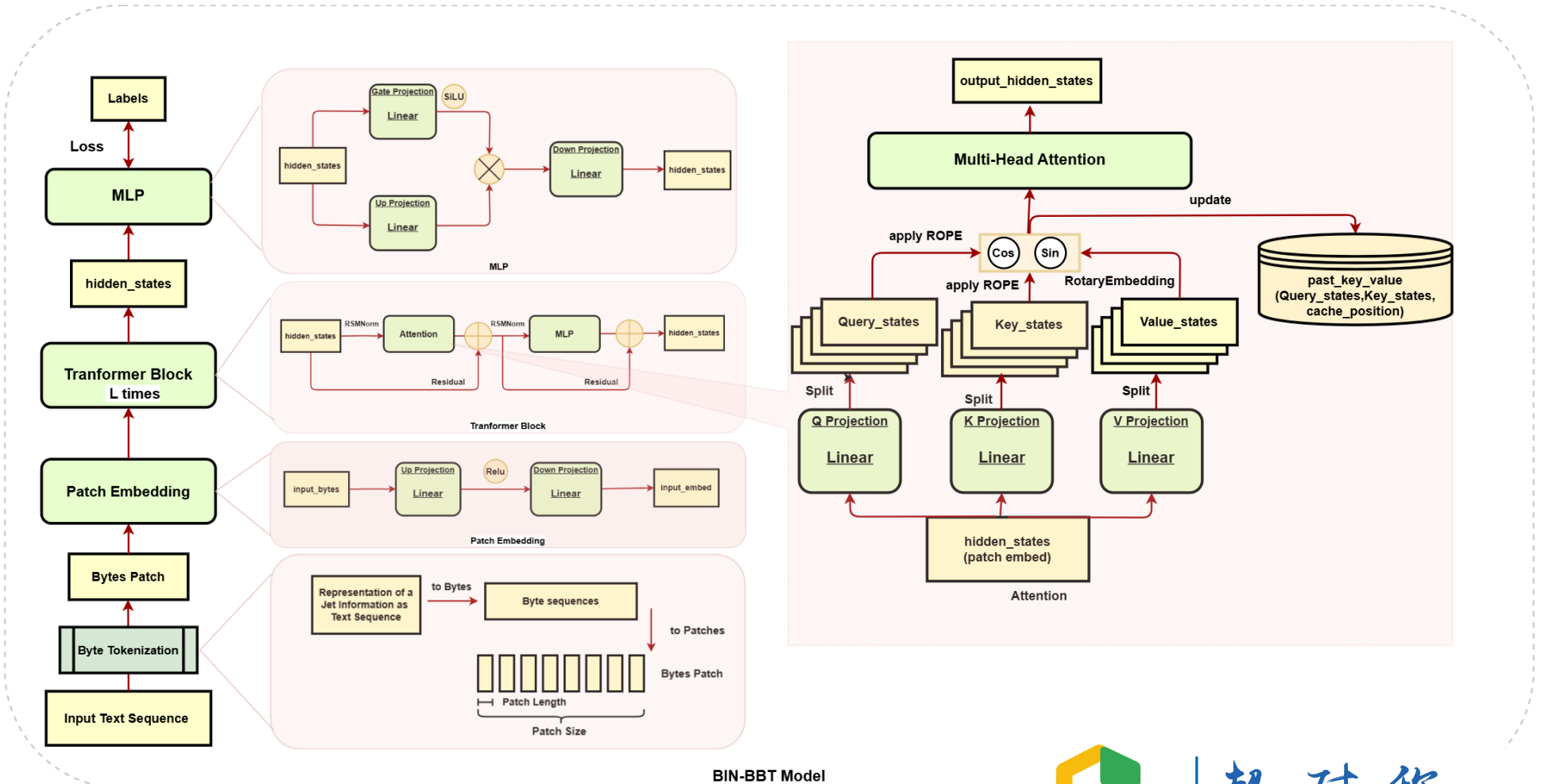
# Resources

- Jet Origin id: [zhuyongfeng@pku.edu.cn](mailto:zhuyongfeng@pku.edu.cn)
- <https://github.com/ZHUZYFgit/CEPC-Jet-Origin-Identification/tree/main>
- LLM (BINBBT):
- <https://github.com/supersymmetry-technologies/bbt-neutron>
- 1-1 Correspondence: [wangyuexin@ihep.ac.cn](mailto:wangyuexin@ihep.ac.cn)
- Source deposition in construction...



# Back up

# Recent update: from specialized Models to LLM



- New tokenization method to address numeric problems at LLM

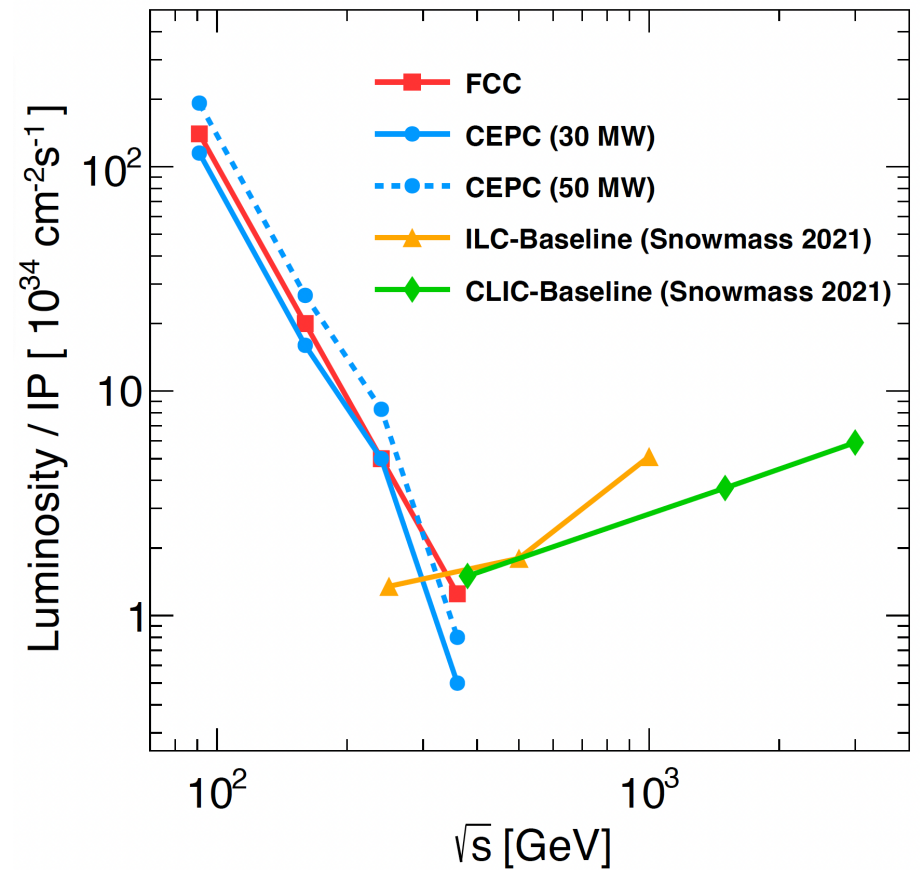
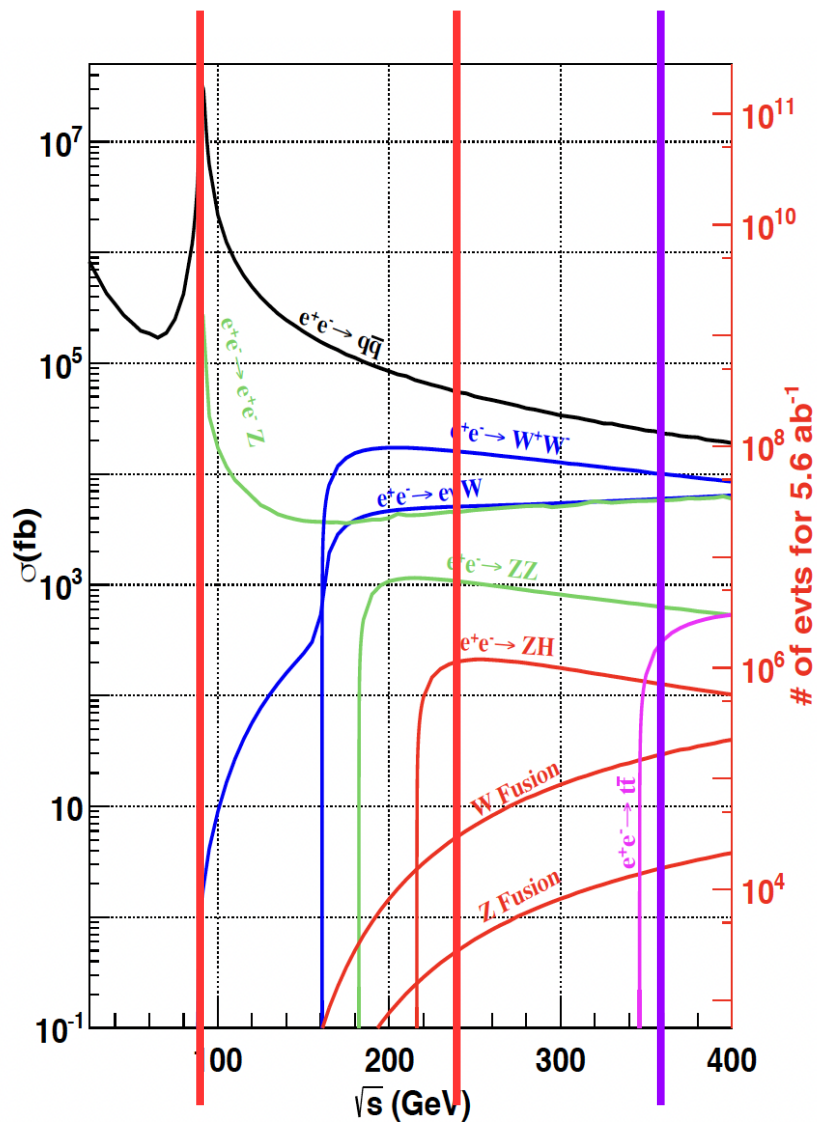
15/01/25

ILD Analysis



超对称  
Super Symmetry  
Technologies

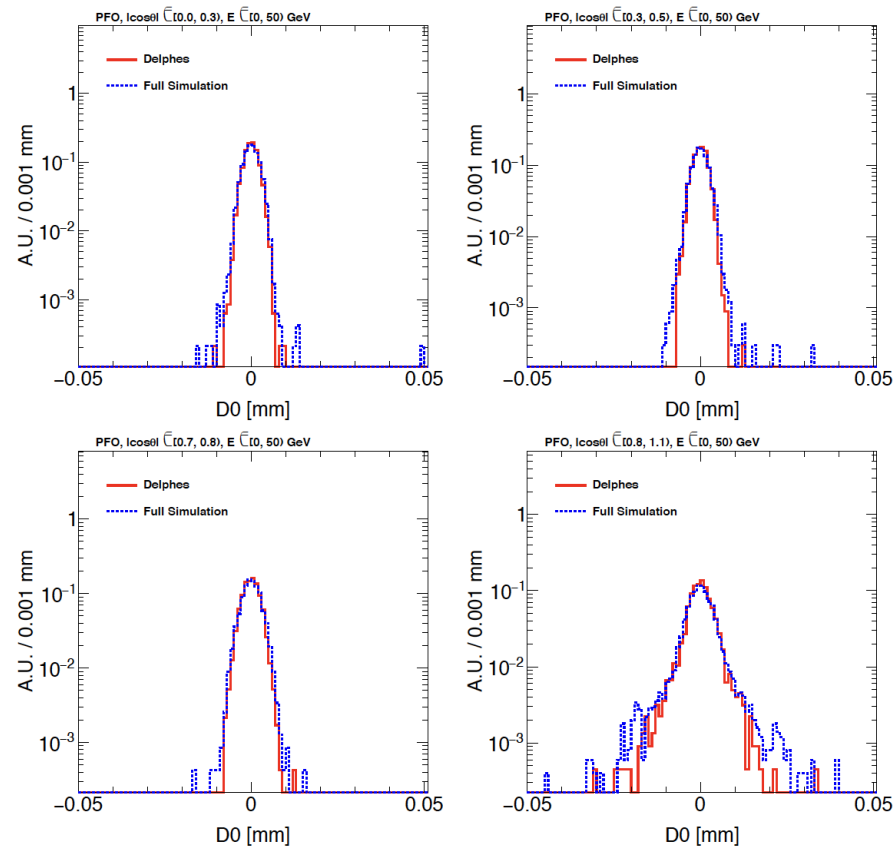
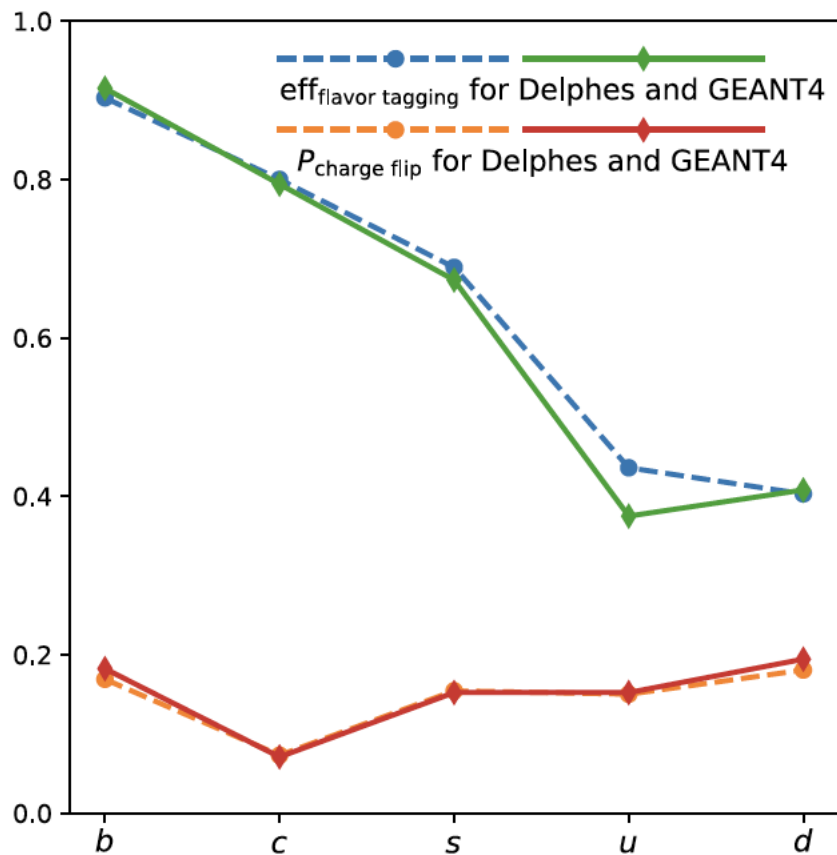
# Yields $\sim$ Xsec \* Lumi \* Time



- 4 Million Higgs (10 years)
- ~ 1 Giga W (1 year) + 4 Tera Z (2 years)
- Upgradable: Top factory (500 k ttbar)

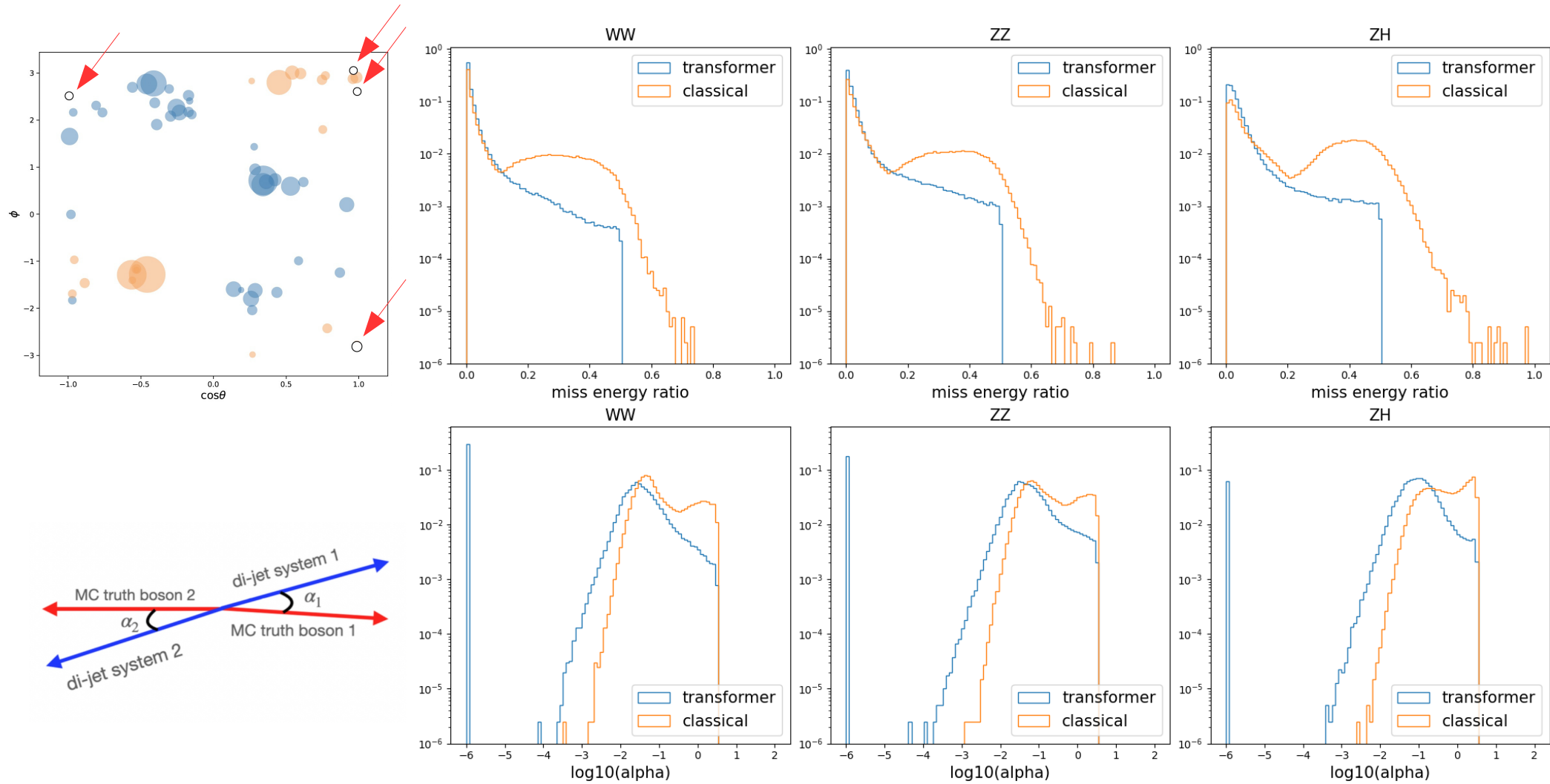
# Fast/Full Simulation

Z → μμ (91.2 GeV)



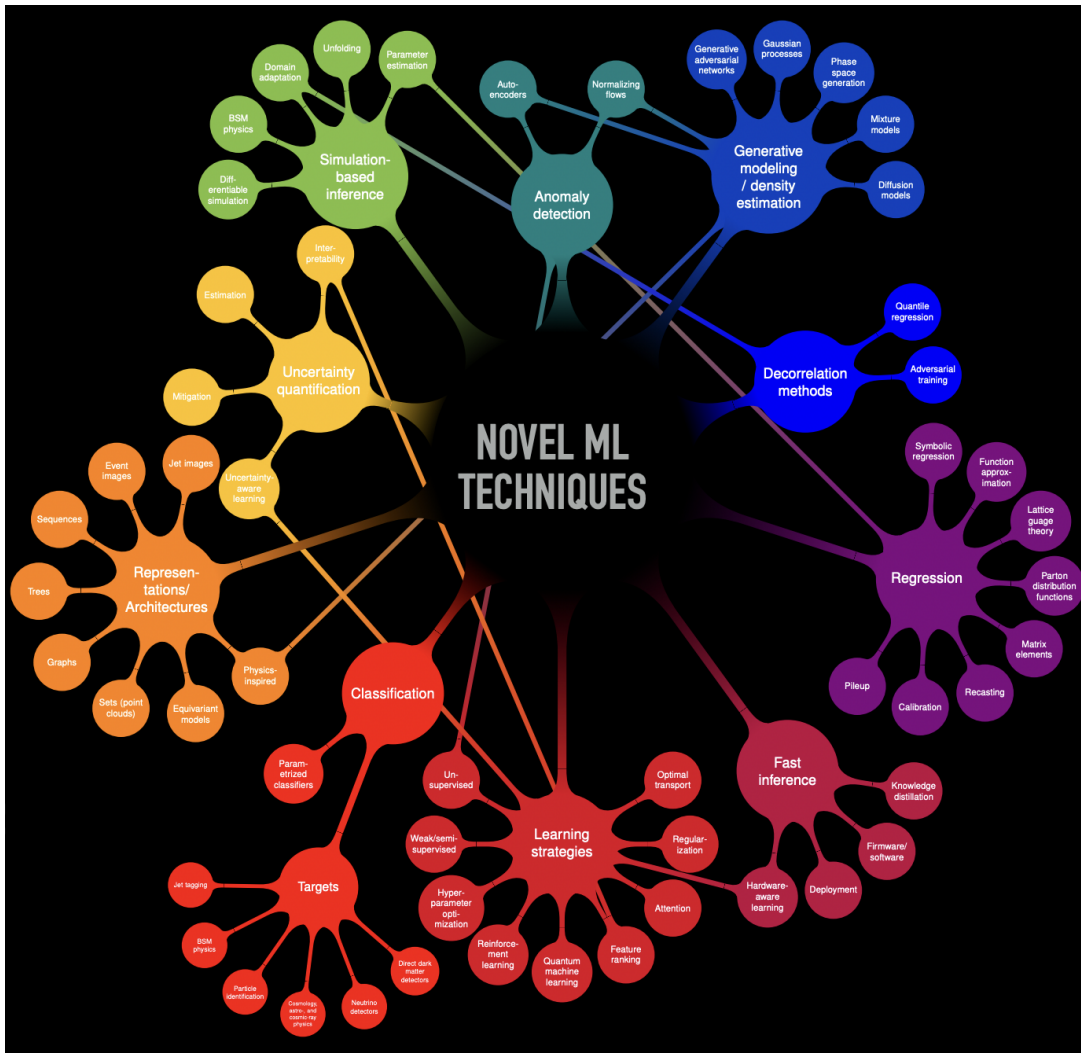
- Delphes ~ Perfect PFA (1 – 1 correspondence.. )

# CSI: classical VS AI (Transformer)



Classical: Jet Clustering + Matching with min( $\chi^2$ )

# AI & HEP



Plot from Javier Mauricio Duarte's talk at ICHEP 2024

- HEP: data intensive + clear, meaningful & interpretable processing
  - Pioneering for neural network application, i.e., in tracking in 1980s
- An irresistible trend:
  - 17/48 parallel talks of Computing & Data handling session at ICHEP 2024 are relevant to AI: 11 machine learning, 3 deep learning, 3 neural network
  - Many domestic discussions & efforts
- CEPC is actively implementing AI to its data processing:
  - Trigger + DAQ...
  - Simulation: Fast sim.
  - Reconstruction: PFA, Jet Origin id, etc
  - Analysis



# 1-1 Correspondence

Holistic description of physics events

Efficient & interpretable information compression: ( $\mathcal{O}(1E5)$  Hits  $\rightarrow$   $\mathcal{O}(100)$  reco particles)

~ Confusion Free PFA + Excellent Particle identification

~ New method for the detector monitoring & measurements

arXiv > hep-ex > arXiv:2411.06939

Search...  
Help | Adv

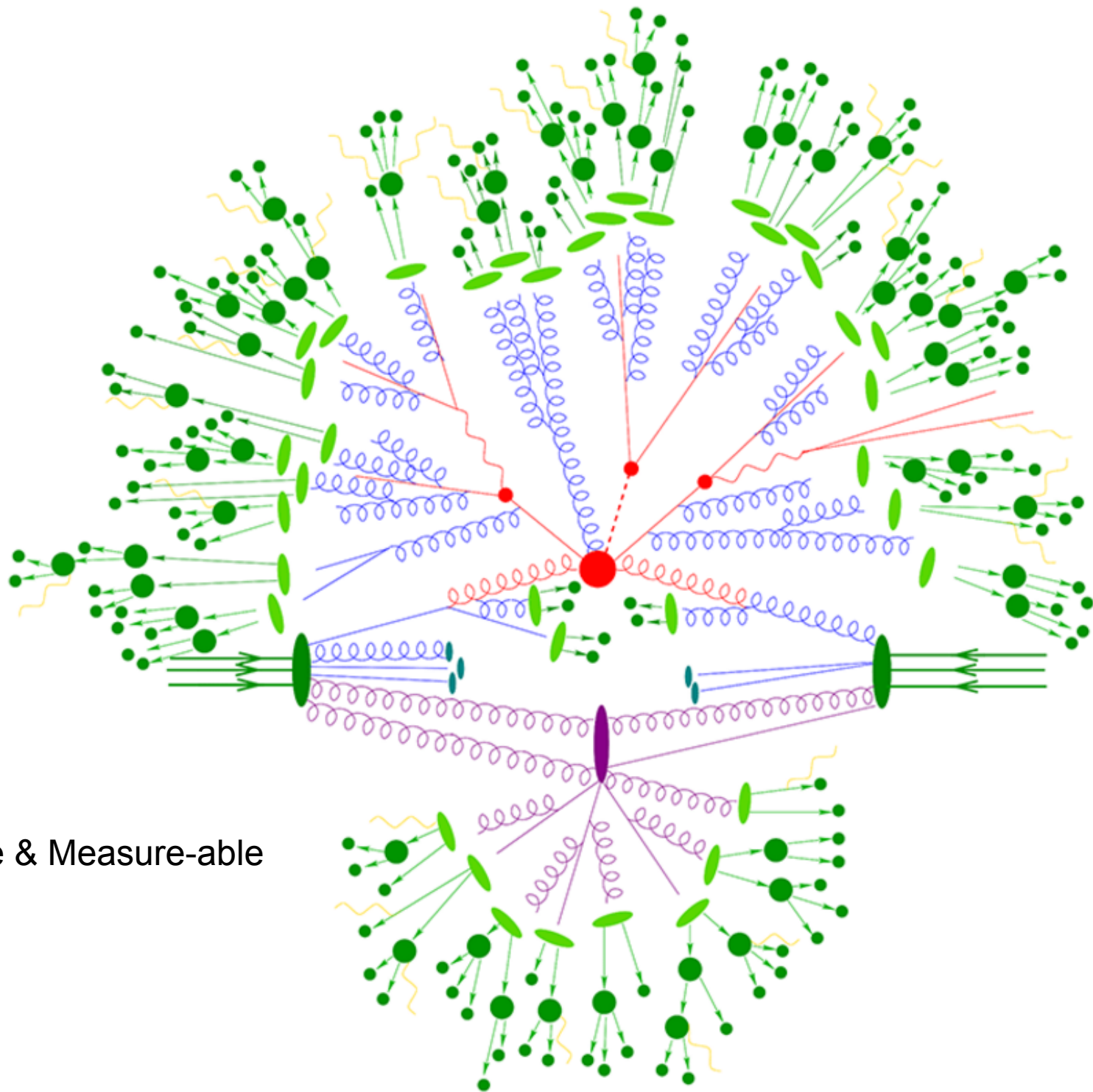
High Energy Physics – Experiment

[Submitted on 11 Nov 2024]

## One-to-one correspondence reconstruction at the electron-positron Higgs factory

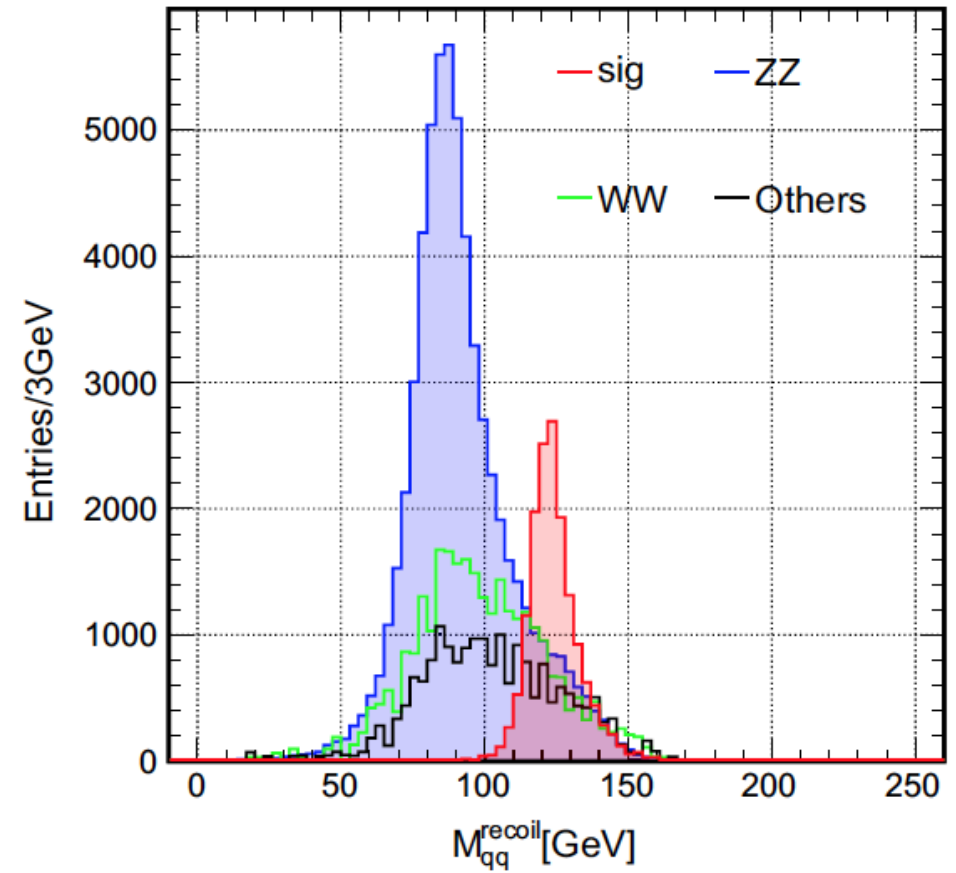
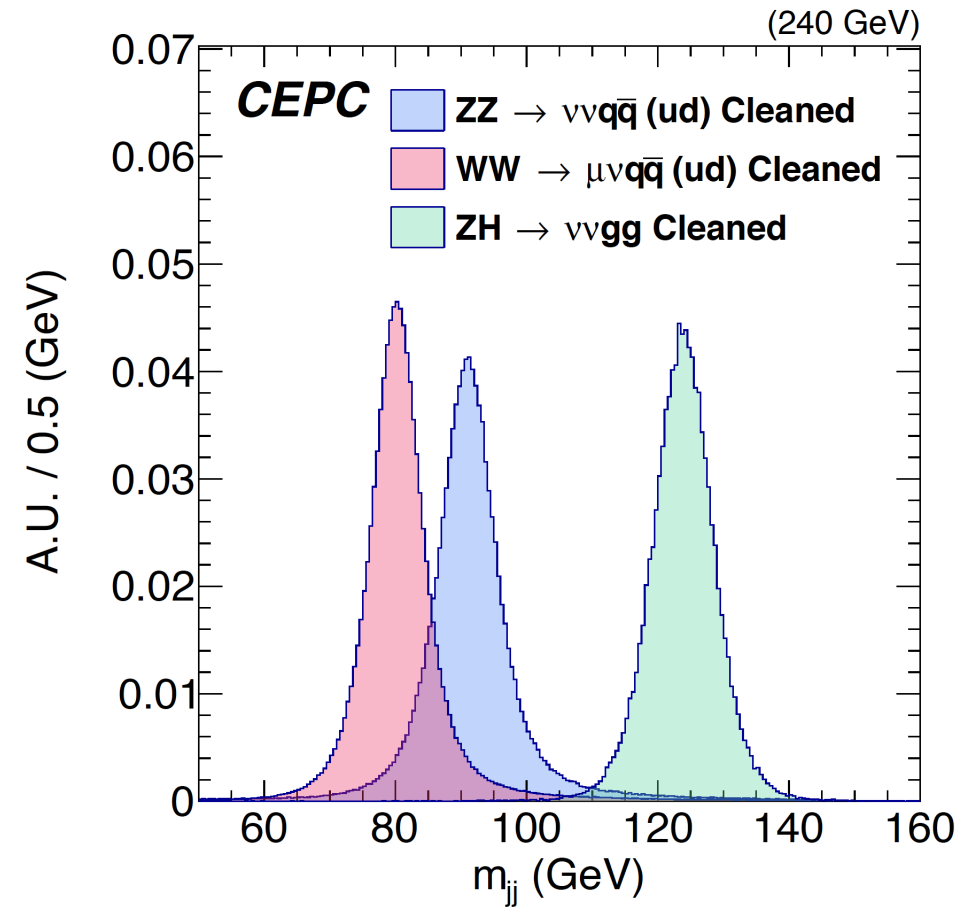
Yuexin Wang, Hao Liang, Yongfeng Zhu, Yuzhi Che, Xin Xia, Huilin Qu, Chen Zhou, Xuai Zhuang, Manqi Ruan

We propose one-to-one correspondence reconstruction for electron-positron Higgs factories. For each visible particle, one-to-one correspondence aims to associate relevant detector hits with only one reconstructed particle and accurately identify its species. To achieve this goal, we develop a novel detector concept featuring 5-dimensional calorimetry that provides spatial, energy, and time measurements for each hit, and a reconstruction framework that combines state-of-the-art particle flow and artificial intelligence algorithms. In the benchmark process of Higgs to di-jets, over 90% of visible energy can be successfully mapped into well-reconstructed particles that not only maintain a one-to-one correspondence relationship but also associate with the correct combination of cluster and track, improving the invariant mass resolution of hadronically decayed Higgs bosons by 25%. Performing simultaneous identification on these well-reconstructed particles, we observe efficiencies of 97% to nearly 100% for charged particles ( $e^\pm$ ,  $\mu^\pm$ ,  $\pi^\pm$ ,  $K^\pm$ ,  $p/\bar{p}$ ) and photons ( $\gamma$ ), and 75% to 80% for neutral hadrons ( $K_L^0$ ,  $n$ ,  $\bar{n}$ ). For physics measurements of Higgs to invisible and exotic decays, golden channels to probe new physics, one-to-one correspondence could enhance discovery power by 10% to up to a factor of two. This study demonstrates the necessity and feasibility of one-to-one correspondence reconstruction at electron-positron Higgs factories.



Calculable & Measure-able

# Boson Mass Resolution: Key Per. Para

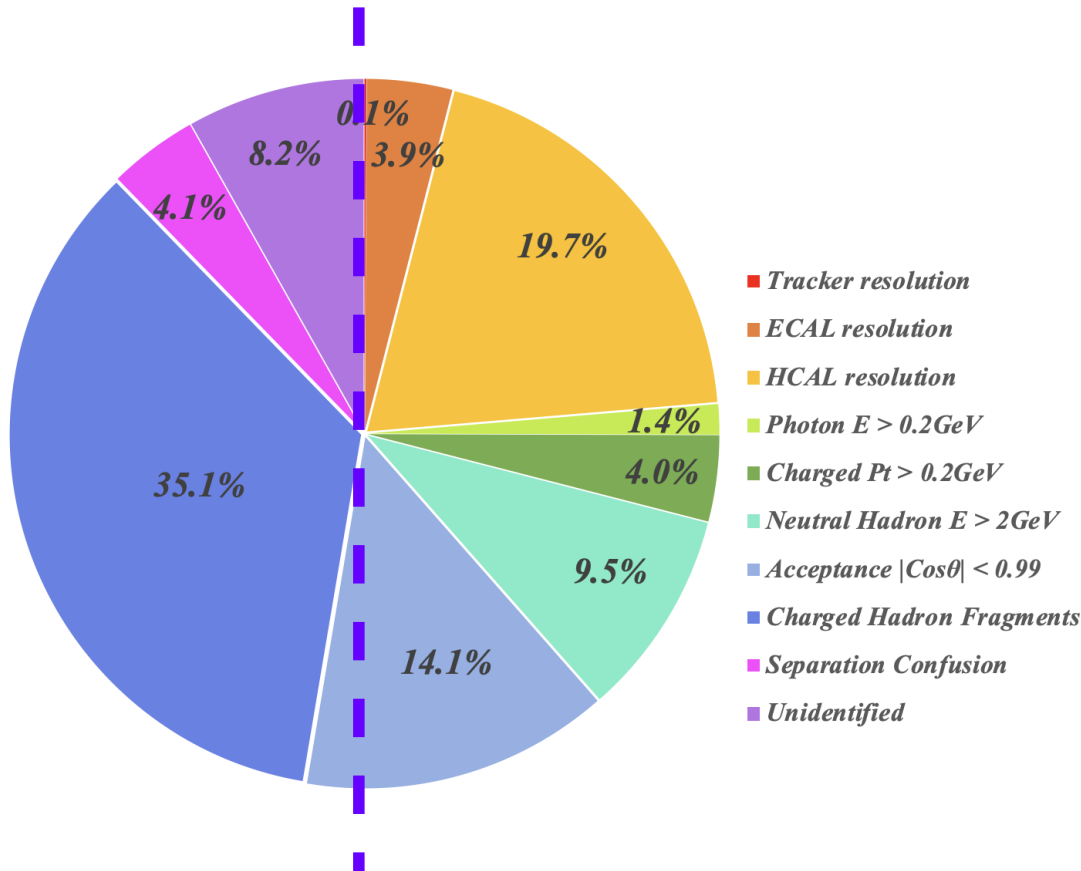


Higgs factory: need BMR < 4% (critical for qqH & qqZ separation using recoil mass to di-jet)

Strongly motivated to improve BMR to 3% or even lower, especially for NP & Flavor

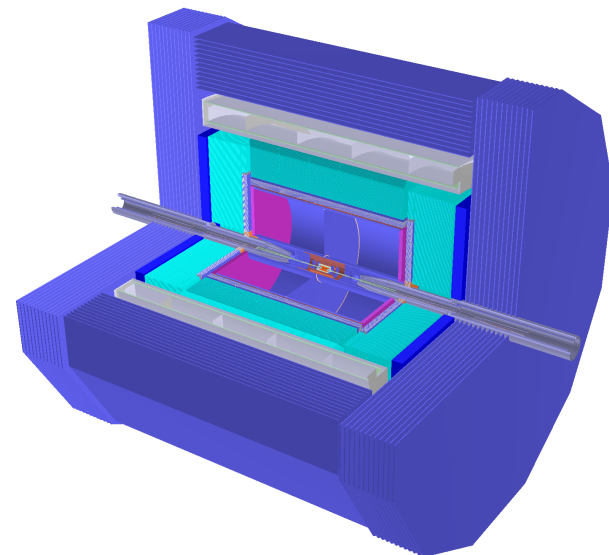
CDR baseline (left plot): BMR = 3.75%

# BMR decomposition @ CDR baseline

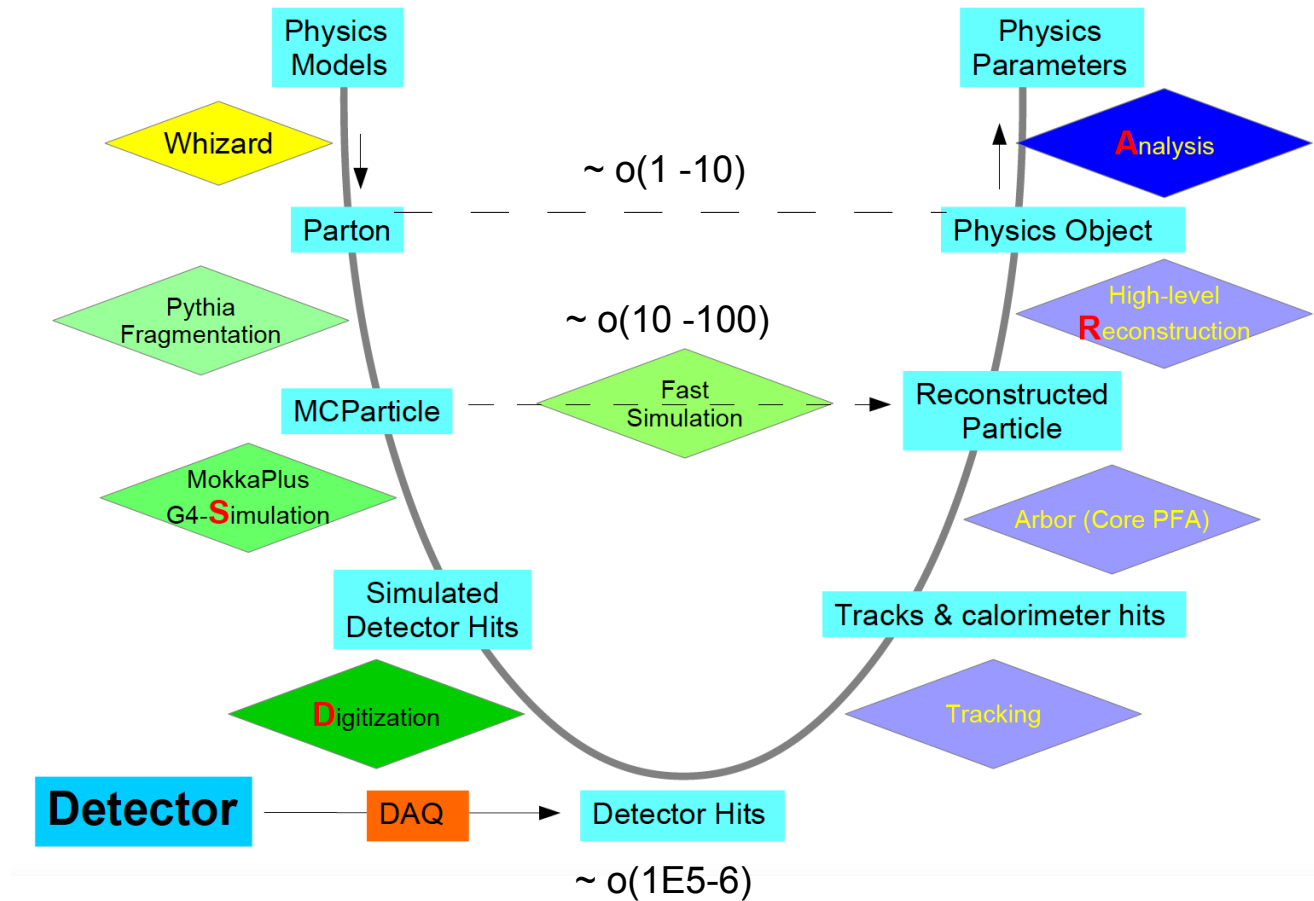
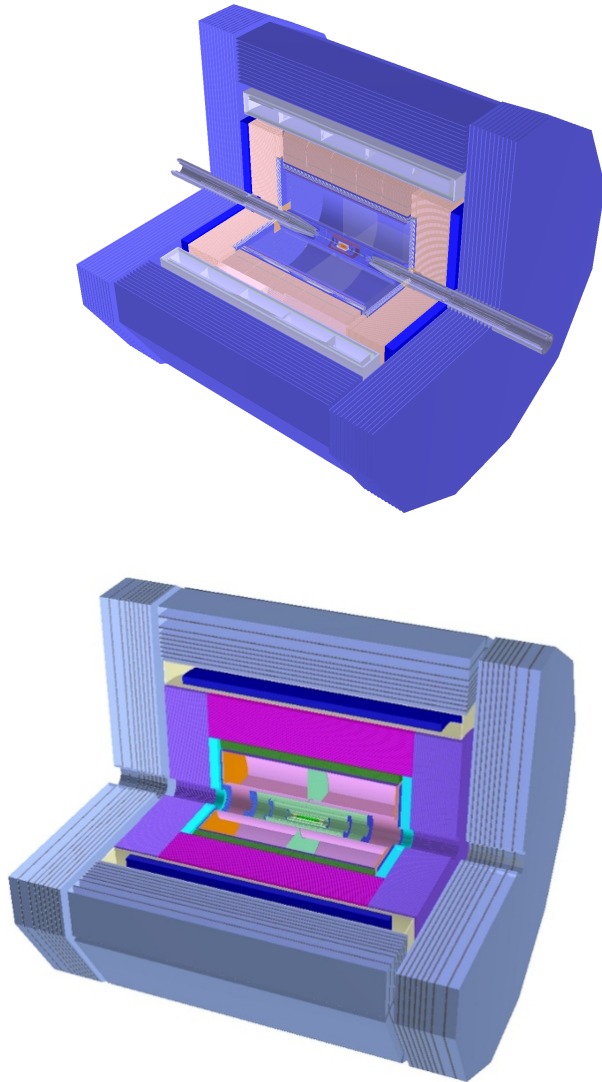


- 1<sup>st</sup> HCAL resolution dominant the uncertainties from intrinsic detector resolution: *need better HCAL → usage of GSHCAL*
- 2<sup>nd</sup> Leading contribution: *Confusion from shower Fragments (fake particles), need better Pattern Reco.*

- CDR baseline - GRPC HCAL



# CEPC Detector & Reconstruction

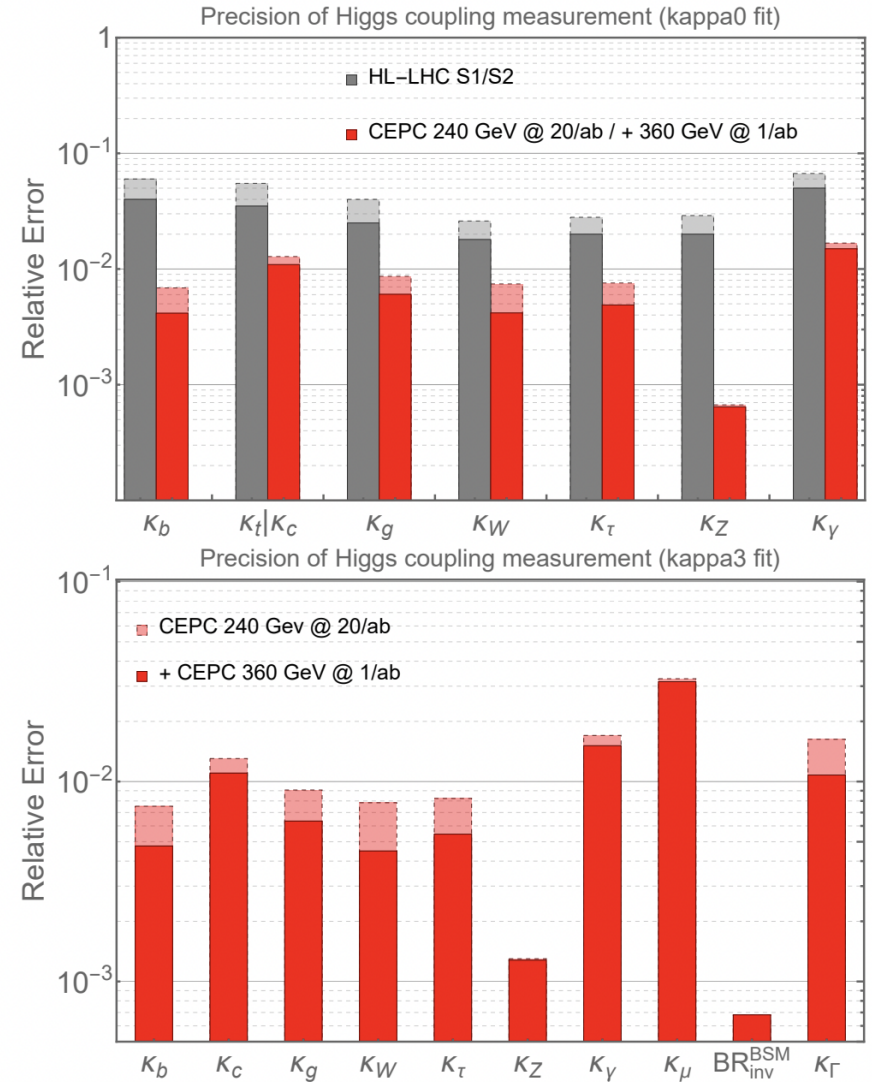


Full simulation reconstruction Chain with **Arbor, etc**



# Higgs & Snowmass White Paper

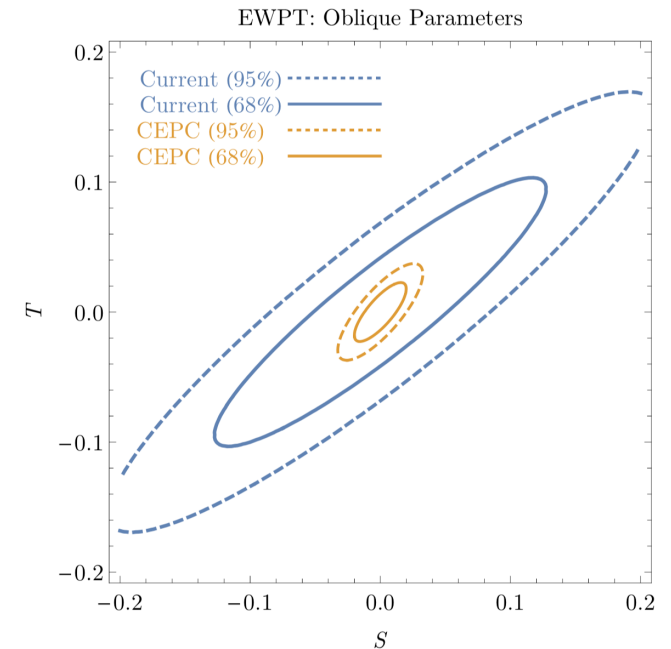
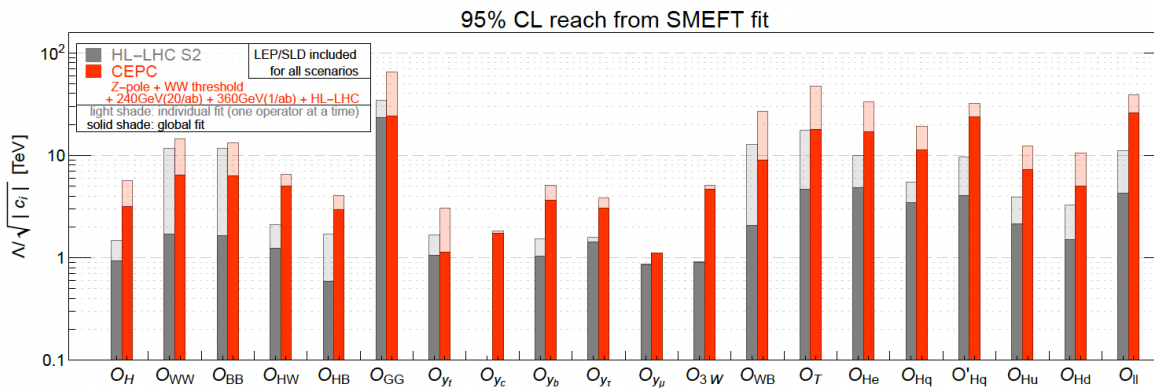
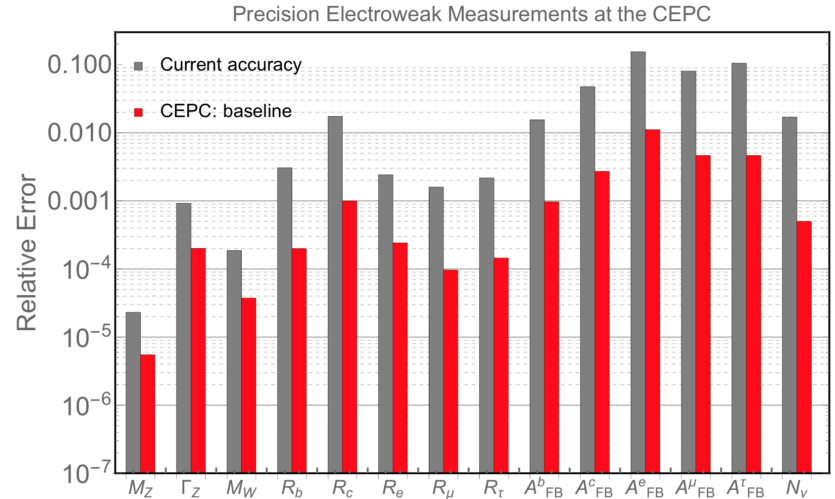
	240 GeV, 20 ab <sup>-1</sup>		360 GeV, 1 ab <sup>-1</sup>		
	ZH	vvH	ZH	vvH	eeH
inclusive	<b>0.26%</b>		<b>1.40%</b>	\	\
H→bb	<b>0.14%</b>	<b>1.59%</b>	<b>0.90%</b>	<b>1.10%</b>	<b>4.30%</b>
H→cc	<b>2.02%</b>		<b>8.80%</b>	<b>16%</b>	<b>20%</b>
H→gg	<b>0.81%</b>		<b>3.40%</b>	<b>4.50%</b>	<b>12%</b>
H→WW	<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 → ττ	<b>0.42%</b>		<b>2.10%</b>	<b>4.20%</b>	<b>7.50%</b>
H → γγ	<b>3.02%</b>		<b>11%</b>	<b>16%</b>	
H → μμ	<b>6.36%</b>		<b>41%</b>	<b>57%</b>	
H → Zγ	<b>8.50%</b>		<b>35%</b>		
Br <sub>upper</sub> (H → inv.)	<b>0.07%</b>				
Γ <sub>H</sub>	<b>1.65%</b>		<b>1.10%</b>		



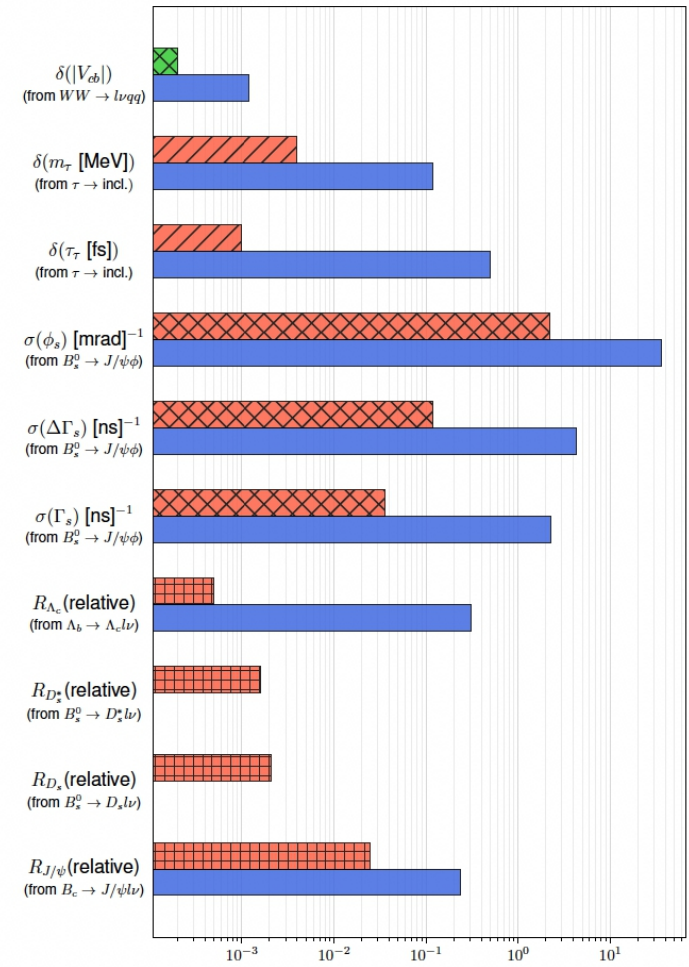
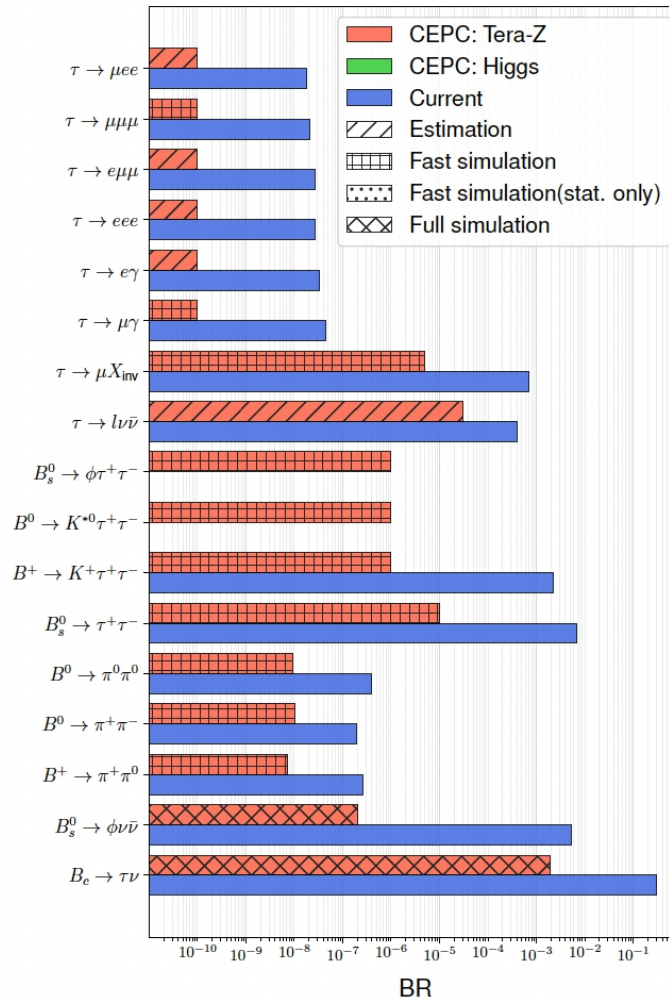
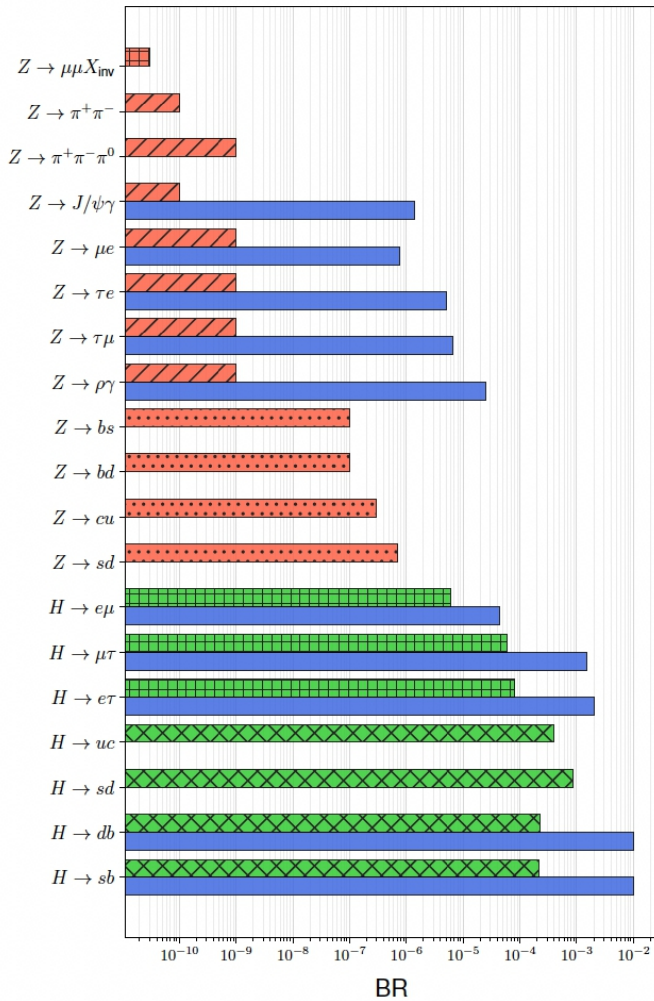


# EW measurements & SMEFT

Observable	current precision	CEPC precision (Stat. Unc.)	CEPC runs	main systematic
$\Delta m_Z$	2.1 MeV [37–41]	0.1 MeV (0.005 MeV)	Z threshold	$E_{beam}$
$\Delta \Gamma_Z$	2.3 MeV [37–41]	0.025 MeV (0.005 MeV)	Z threshold	$E_{beam}$
$\Delta m_W$	9 MeV [42–46]	0.5 MeV (0.35 MeV)	WW threshold	$E_{beam}$
$\Delta \Gamma_W$	49 MeV [46–49]	2.0 MeV (1.8 MeV)	WW threshold	$E_{beam}$
$\Delta m_t$	0.76 GeV [50]	$\mathcal{O}(10)$ MeV <sup>a</sup>	$t\bar{t}$ threshold	
$\Delta A_e$	$4.9 \times 10^{-3}$ [37, 51–55]	$1.5 \times 10^{-5}$ ( $1.5 \times 10^{-5}$ )	Z pole ( $Z \rightarrow \tau\tau$ )	Stat. Unc.
$\Delta A_\mu$	0.015 [37, 53]	$3.5 \times 10^{-5}$ ( $3.0 \times 10^{-5}$ )	Z pole ( $Z \rightarrow \mu\mu$ )	point-to-point Unc.
$\Delta A_\tau$	$4.3 \times 10^{-3}$ [37, 51–55]	$7.0 \times 10^{-5}$ ( $1.2 \times 10^{-5}$ )	Z pole ( $Z \rightarrow \tau\tau$ )	tau decay model
$\Delta A_b$	0.02 [37, 56]	$20 \times 10^{-5}$ ( $3 \times 10^{-5}$ )	Z pole	QCD effects
$\Delta A_c$	0.027 [37, 56]	$30 \times 10^{-5}$ ( $6 \times 10^{-5}$ )	Z pole	QCD effects
$\Delta \sigma_{had}$	37 pb [37–41]	2 pb (0.05 pb)	Z pole	luminosity
$\delta R_b^0$	0.003 [37, 57–61]	0.0002 ( $5 \times 10^{-6}$ )	Z pole	gluon splitting
$\delta R_c^0$	0.017 [37, 57, 62–65]	0.001 ( $2 \times 10^{-5}$ )	Z pole	gluon splitting
$\delta R_e^0$	0.0012 [37–41]	$2 \times 10^{-4}$ ( $3 \times 10^{-6}$ )	Z pole	$E_{beam}$ and t channel
$\delta R_\mu^0$	0.002 [37–41]	$1 \times 10^{-4}$ ( $3 \times 10^{-6}$ )	Z pole	$E_{beam}$
$\delta R_\tau^0$	0.017 [37–41]	$1 \times 10^{-4}$ ( $3 \times 10^{-6}$ )	Z pole	$E_{beam}$
$\delta N_\nu$	0.0025 [37, 66]	$2 \times 10^{-4}$ ( $3 \times 10^{-5}$ )	ZH run ( $\nu\nu\gamma$ )	Calo energy scale



# Flavor Physics



See the non-seen: i.e,  $B_c \rightarrow \tau \nu$ ,  $B_s \rightarrow \text{Phiv}$   
 Orders of magnitudes improvements (1 – 2.5 orders...)  
 Access New Physics with energy scale of 10 TeV, or even above

# New Physics white paper

	5		
VIII. Flavor Portal NP(Lingfeng, Xinqiang)	28	4. Prospects of heavy neutrinos in $U(1)$ models	151
IX. Electroweak phase transition and gravitational wave (Kepan Xie, Sai Wang, Fa Peng Huang)	28	5. Prospects of heavy neutrinos in the LRSM	153
A. Electroweak phase transition in standard model effective field theory	28	B. Non-standard neutrino interactions	155
B. Electroweak phase transition in well-motivated new physics models	28	C. Active-sterile neutrino transition magnetic moments	155
1. singlet model	28	D. Neutral and doubly-charged scalars in seesaw models	157
2. doublet model	28	E. Connection to Leptogenesis and Dark Matter	162
C. Cosmological implication and complementary test with gravitational wave	28	F. Summary	164
1. electroweak baryogenesis	28	XI. More Exotics (Yu, Zuowei)	165
2. dark matter	28	A. Axion-like particles	165
3. primordial black hole	28	B. Lepton form factors	169
4. Complementary test with gravitational wave	28	1. General remarks on $\mu/e$ $g-2$	169
X. More Exotics (Yu, Zuowei)	28	2. $\mu/e$ dipole moments in SUSY	171
A. Heavy neutrinos	29	3. $\tau$ weak-electric dipole moments	172
B. Axion-like particles	31	C. Emergent Hadron Mass	174
C. Axion-like particles 2 (from Kingman & Ouseph)	31	D. Exotic lepton mass models	175
D. Axion-like particles 3 (Chih-Ting Lu)	32	E. Spin entanglement	177
E. Emergent Hadron Mass (Roberts Craig)	33	XII. Global Fits (Jiayin, Yang, Yong Du)	182
F. Active-sterile neutrino transition magnetic moments(Yu Zhang)	35	A. SMEFT global fits	182
G. tau-lepton weak electric dipole moment (Long Chen)	37	B. 2HDM global fits (Tao Han, Shufang Su, Wei Su, Yongcheng Wu)	183
H. Nonstandard neutrino interactions (Jiajun Liao & Yu Zhang )	39	C. SUSY global fits	187
I. Lepton mass relation models (Zheng Sun)	40	XIII. Conclusion (Jia LIU)	190
XI. Gloable Fits (Jiayin, Yang)	40	Acknowledgements (Manqi?)	195
A. SUSY global fits	40	Glossary (Xuai)	196
XII. Conclusion (Liantao, Xuai, Manqi,Jia, Zhen, ...)	42	References	197
References	42		

2023

2024



- Credit: hanhua Cui, Yu Gao, Xuai Zhuang

Contents extends from 40 pages → 200 pages...

# Arbor

## Tree topology of particle shower

*Ori. Idea from Henri Videau @ ALEPH*

Eur. Phys. J. C (2018) 78:426  
<https://doi.org/10.1140/epjc/s10052-018-5876-z>

THE EUROPEAN  
PHYSICAL JOURNAL C



Special Article - Tools for Experiment and Theory

### Reconstruction of physics objects at the Circular Electron Positron Collider with Arbor

Manqi Ruan<sup>1,a</sup>, Hang Zhao<sup>1</sup>, Gang Li<sup>1</sup>, Chengdong Fu<sup>1</sup>, Zhigang Wang<sup>1</sup>, Xinchou Lou<sup>6,7,8</sup>, Dan Yu<sup>1,2</sup>, Vincent Boudry<sup>2</sup>, Henri Videau<sup>2</sup>, Vladislav Balagura<sup>2</sup>, Jean-Claude Brient<sup>2</sup>, Peizhu Lai<sup>3</sup>, Chia-Ming Kuo<sup>3</sup>, Bo Liu<sup>1,4</sup>, Fenfen An<sup>1,4</sup>, Chunhui Chen<sup>4</sup>, Soeren Prell<sup>4</sup>, Bo Li<sup>5</sup>, Imad Laketineh<sup>5</sup>

<sup>1</sup> Institute of High Energy Physics, Beijing, China

<sup>2</sup> Laboratoire Leprince-Ringuet, Ecole Polytechnique, Palaiseau, France

<sup>3</sup> Department of Physics and Center of high energy and high field physics, National Central University, Taoyuan City, Taiwan

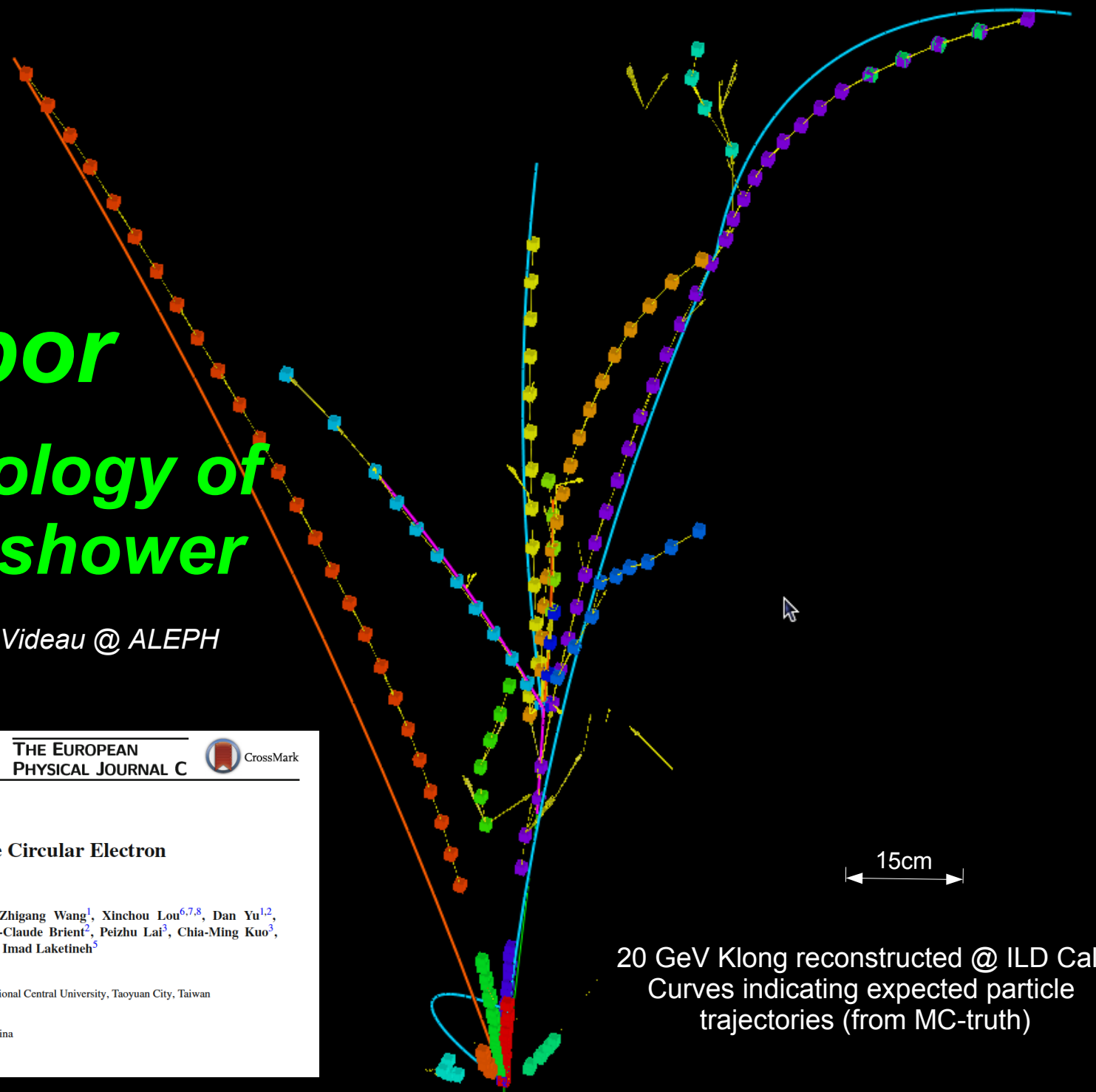
<sup>4</sup> Iowa State University, Ames, USA

<sup>5</sup> Institut de Physique Nucleaire de Lyon, Lyon, France

<sup>6</sup> Institute of High Energy Physics, Chinese Academy of Sciences, Beijing, China

<sup>7</sup> Physics Department, University of Texas at Dallas, Richardson, TX, USA

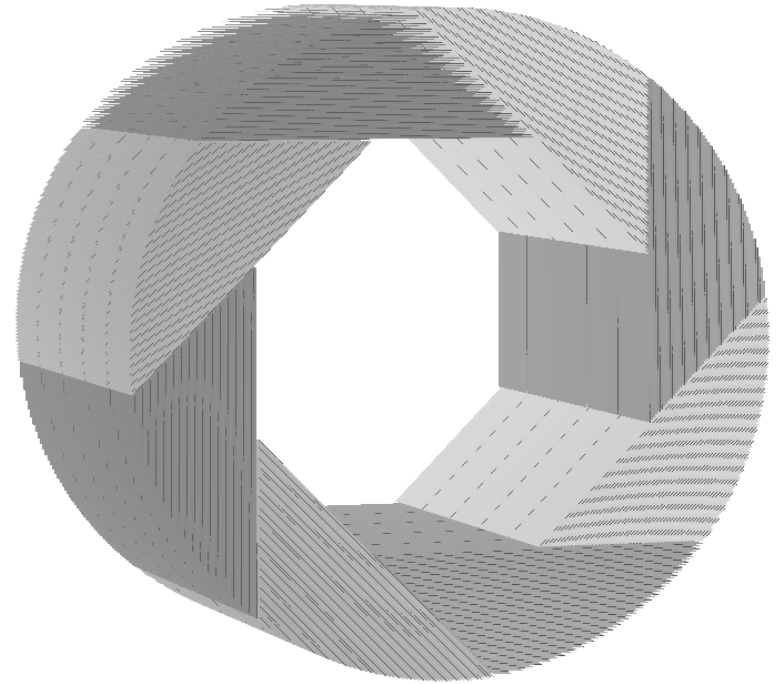
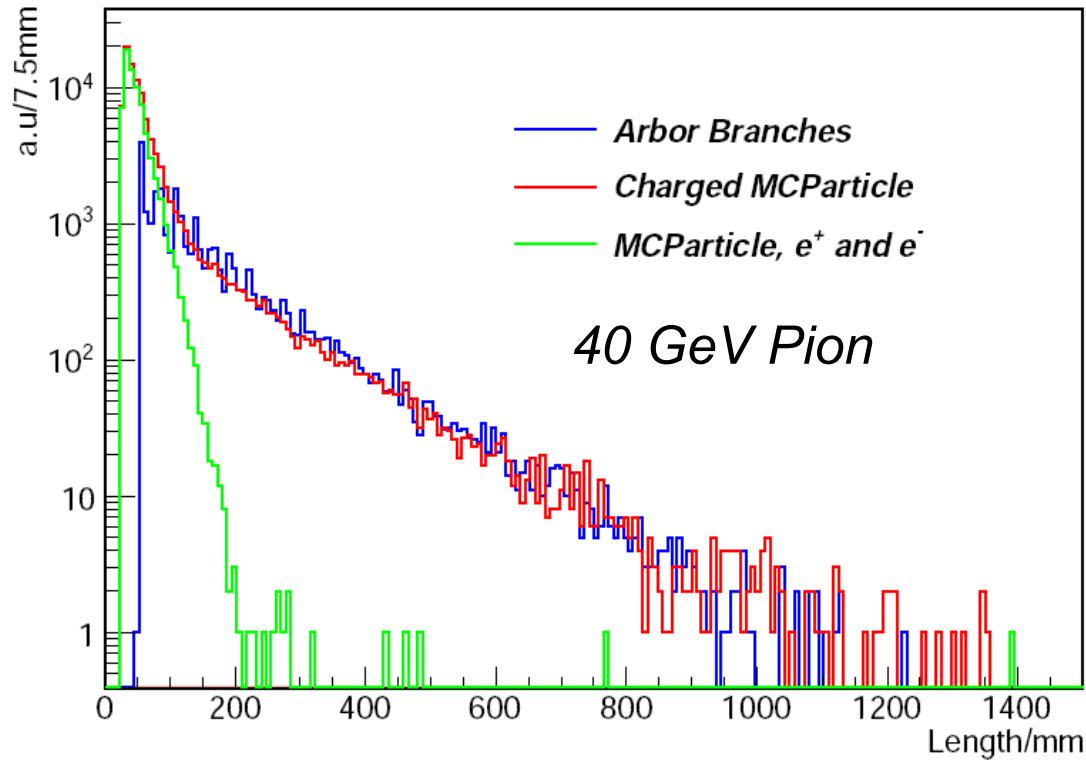
<sup>8</sup> University of Chinese Academy of Sciences (UCAS), Beijing, China



20 GeV Klong reconstructed @ ILD Calo  
Curves indicating expected particle trajectories (from MC-truth)



# Validation: Arbor Branch Length Vs MC Truth



Arbor: successfully **tag** sub-shower structure

*Samples: Particle gun event at ILD HCAL (readout granularity  $1\text{cm}^2$  & layer thickness  $2.65\text{cm}$ )*

*Length:*

*Charged MCParticle: spatial distance between generation/end points*

*Arbor branch: sum of distance between neighboring cells*

$Z \rightarrow 2 \text{ muon}$ ,  
 $H \rightarrow 2 \text{ b}$   
 $\sim 2\%$

$Z \rightarrow 2 \text{ jet}$ ,  
 $H \rightarrow 2 \text{ tau}$   
 $\sim 5\%$

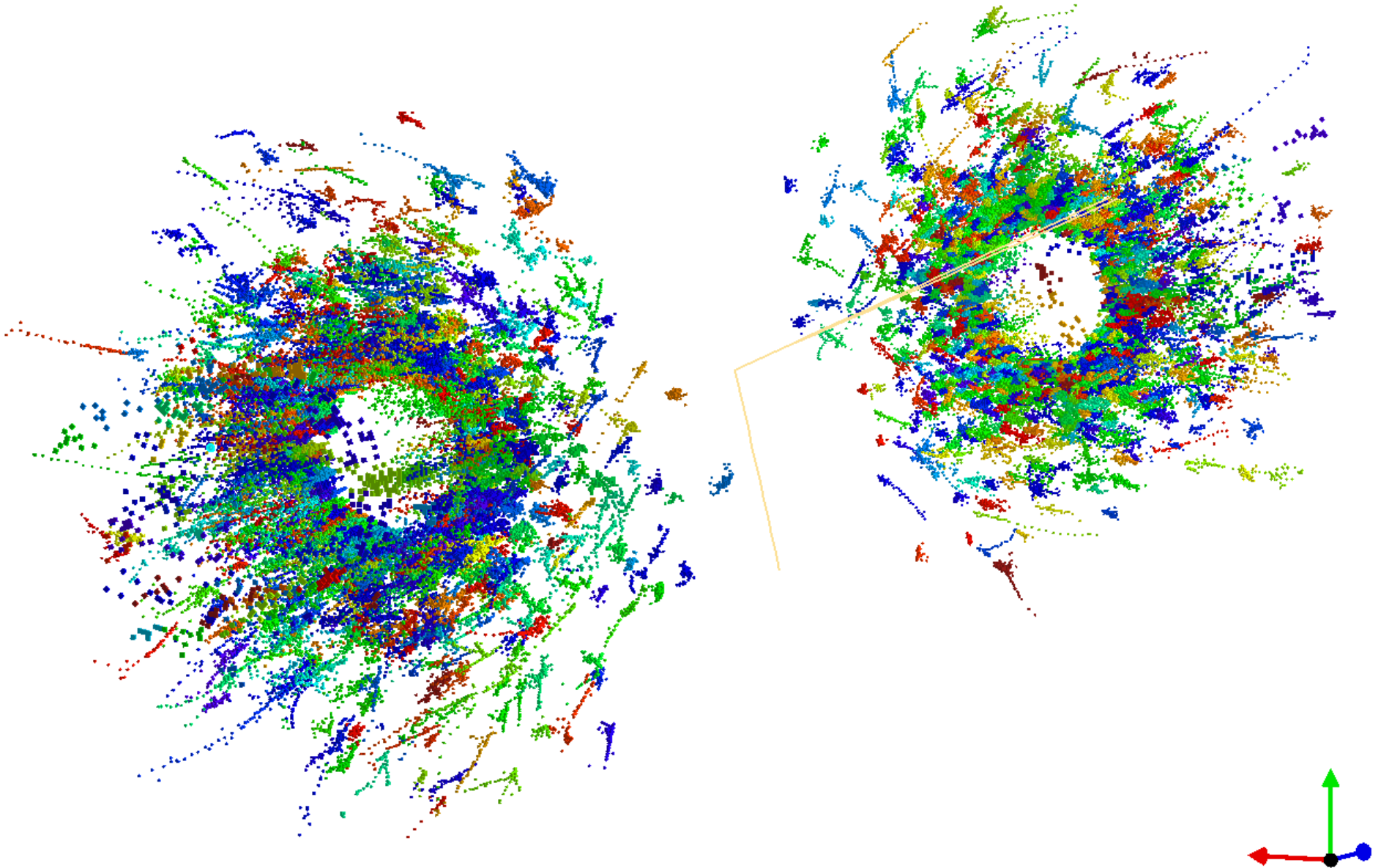
$ZH \rightarrow 4 \text{ jets}$   
 $\sim 50\%$

$Z \rightarrow 2 \text{ muon}$   
 $H \rightarrow WW^* \rightarrow eevv$   
 $\sim 1\%$



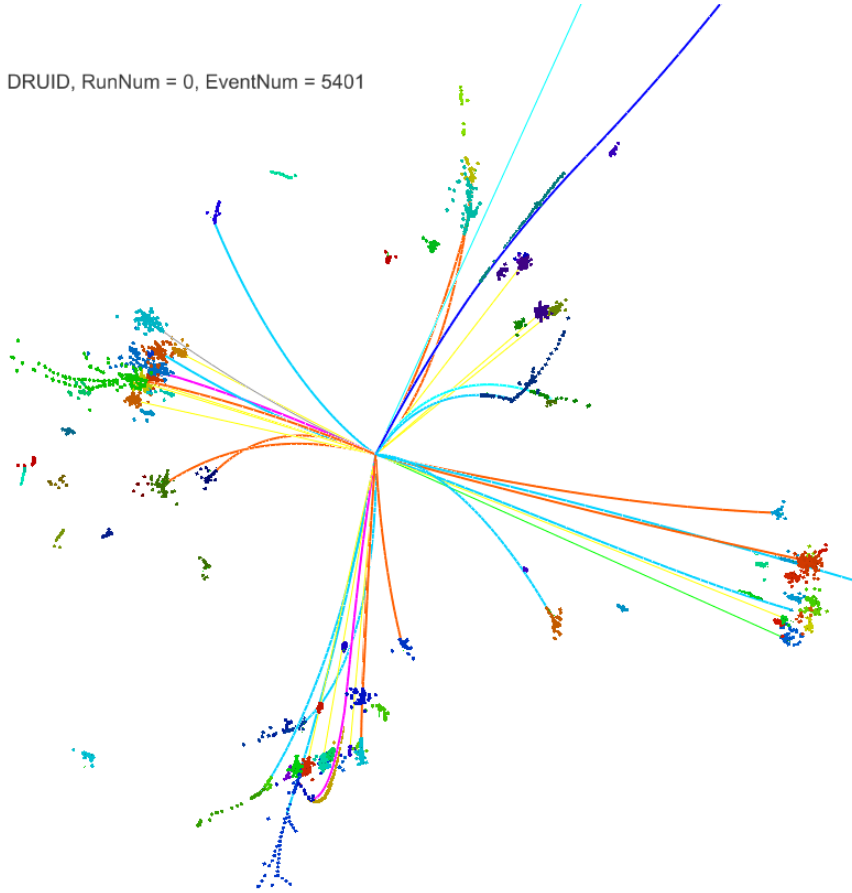


CMS Experiment at LHC, CERN  
Data recorded: Thu Jan 1 01:00:00 1970 CEST  
Run/Event: 1 / 1201  
Lumi section: 13

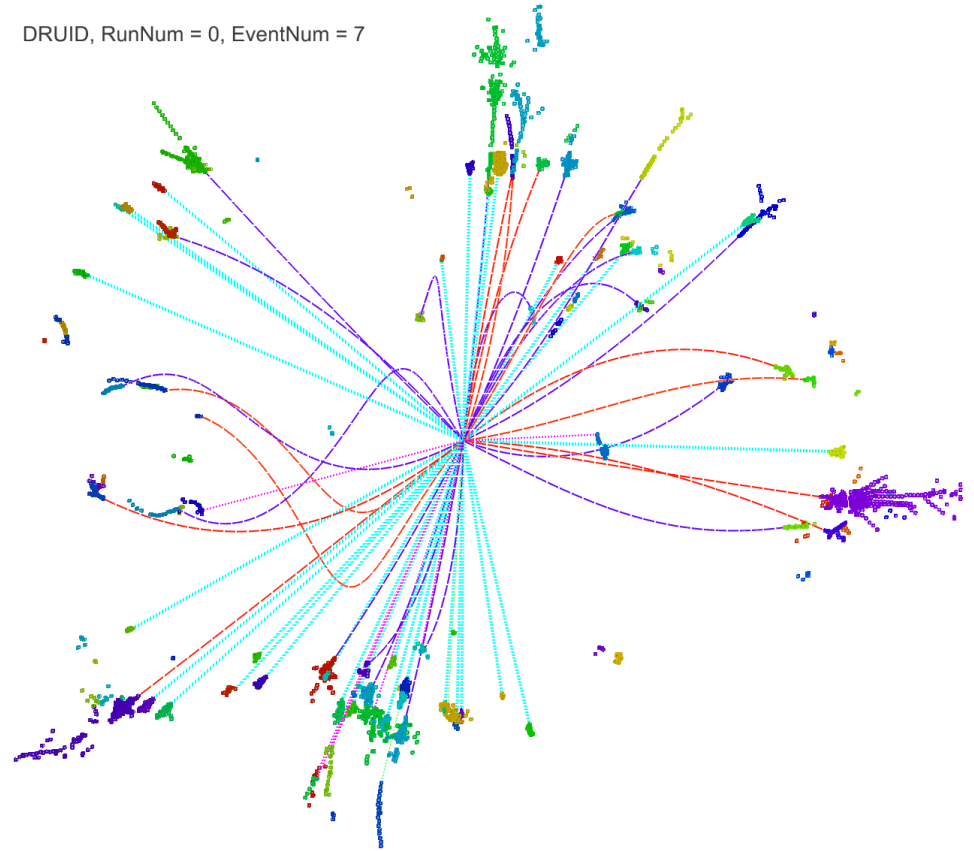


# Color Singlet Identification

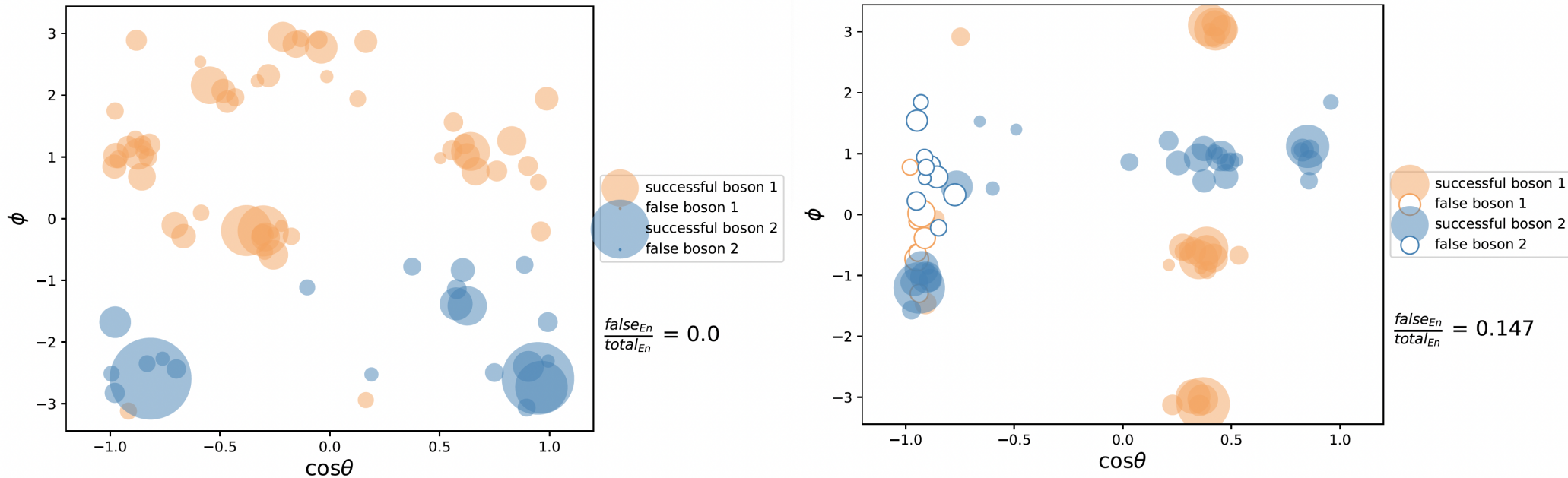
DRUID, RunNum = 0, EventNum = 5401



DRUID, RunNum = 0, EventNum = 7

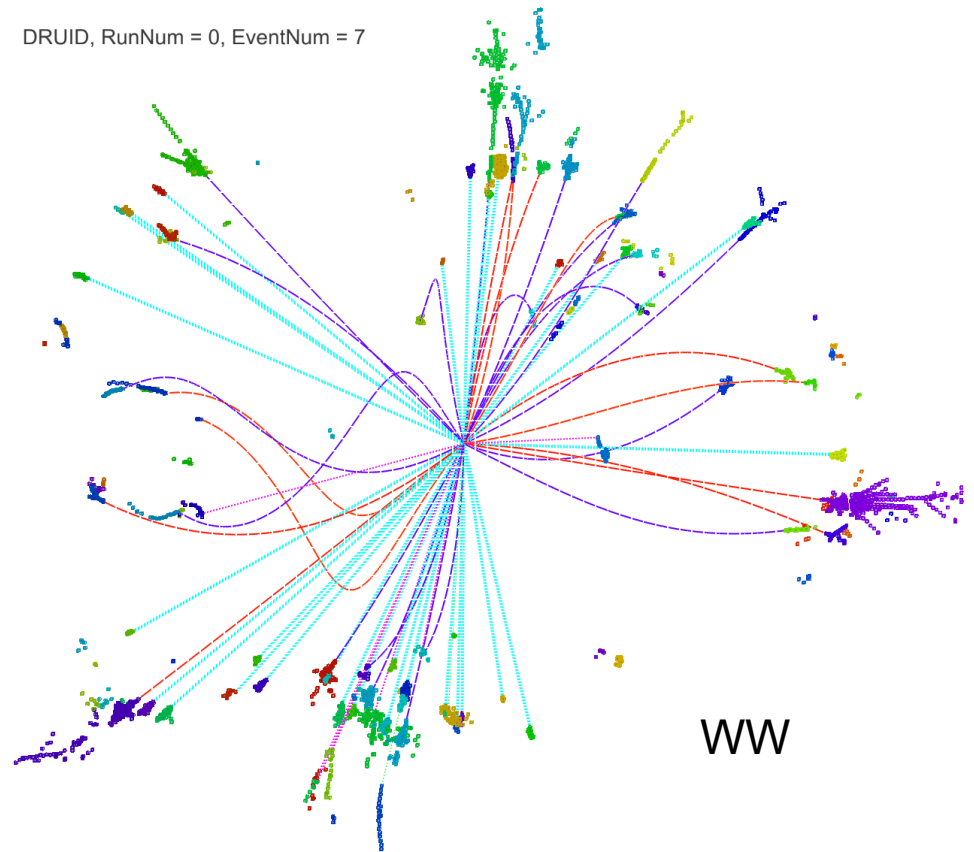
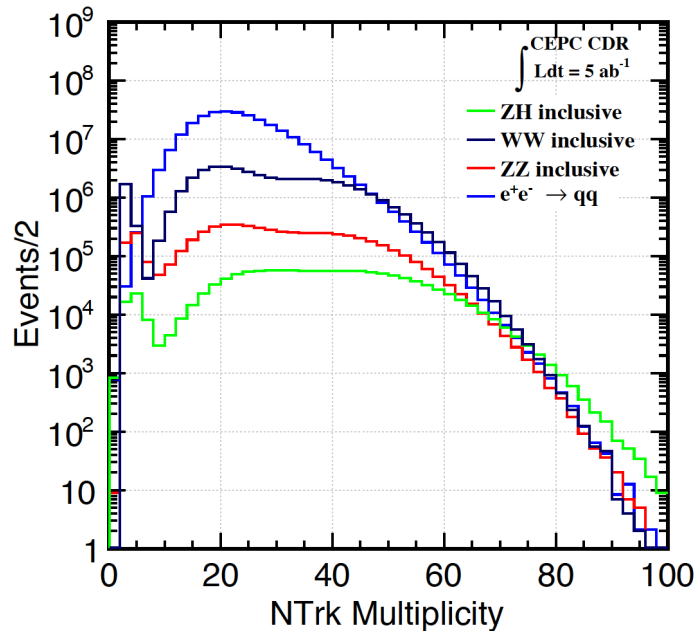


# Color Singlet Identification



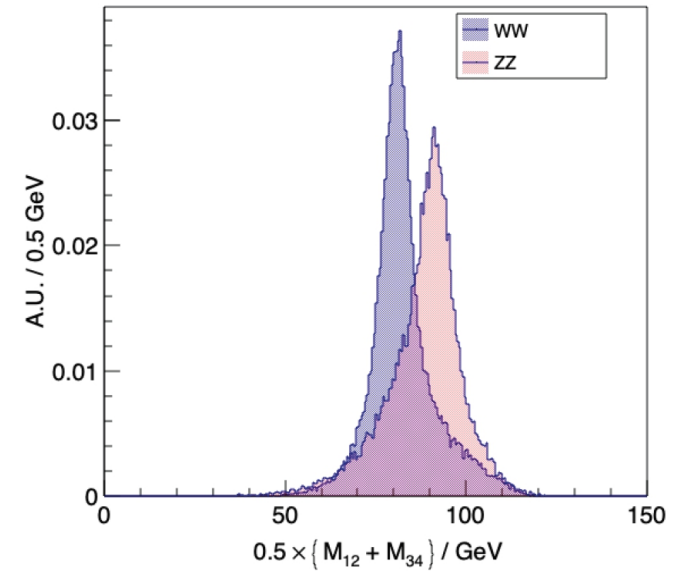
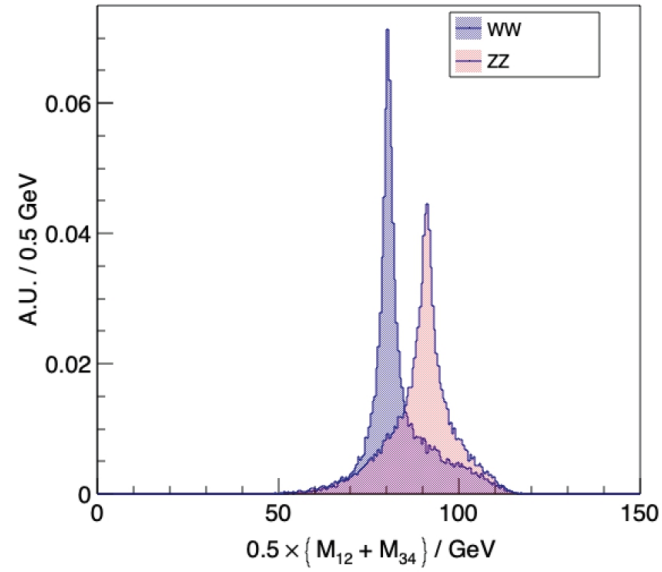
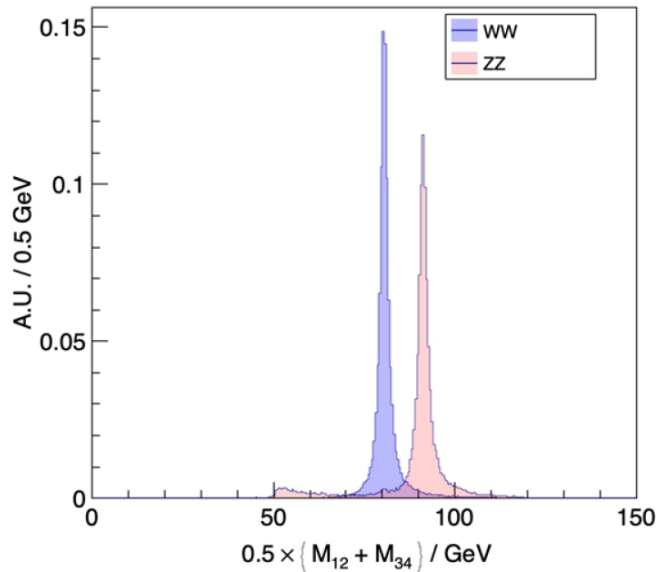
- CSI: identify the color single origin of each final state particle
- Grouping problem: essential for the physics measurements with multi-jet events, i.e., measurements with full hadronic ZH events
- AI might well strongly enhance its performance: compared to conventional jet clustering & matching

# BM-III: full hadronic WW-ZZ separation



- Low energy jets! (20 – 120 GeV)
- Typical multiplicity ~ o(100)
- WW-ZZ Separation: determined by
  - Intrinsic boson mass/width
  - Jet confusion from color single reconstruction – jet clustering & pairing
  - Detector response

# Jet confusion: the leading term



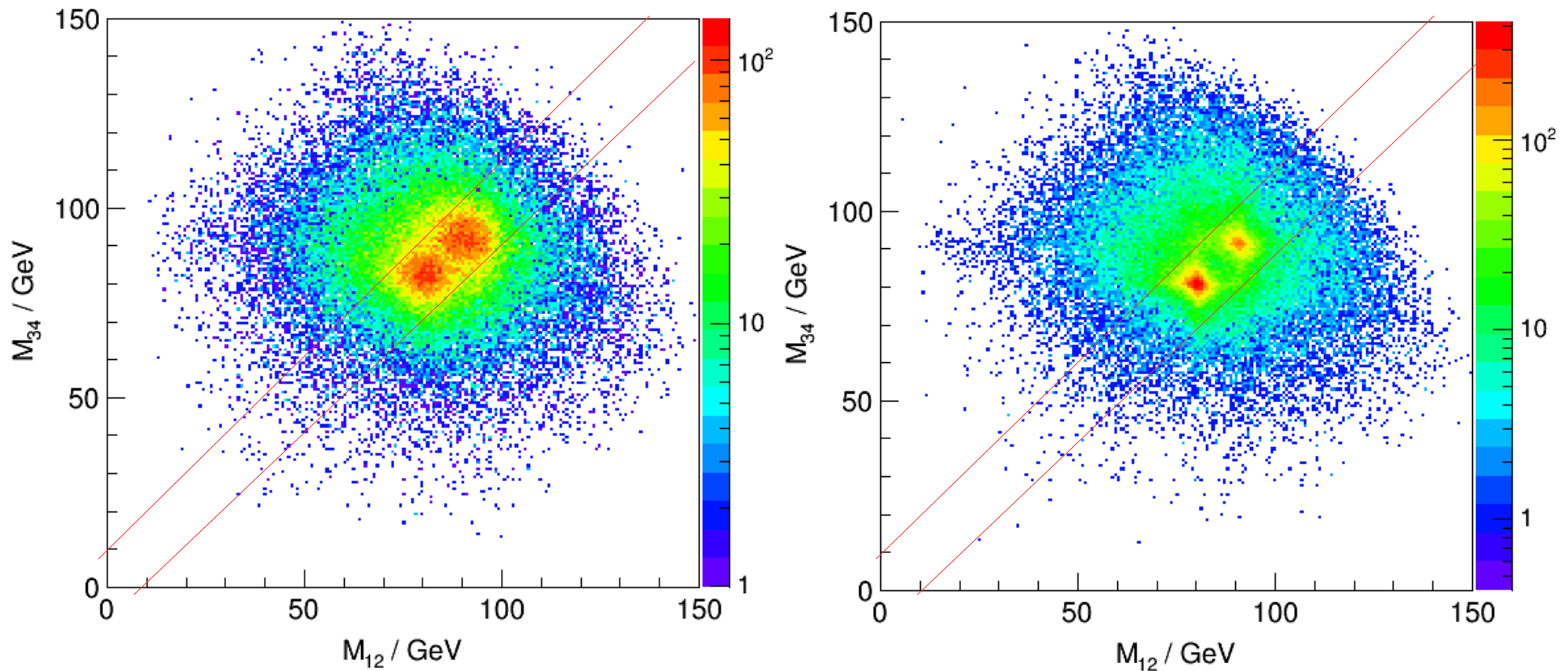
- Separation be characterized by
- Final state/MC particles are clustered into Reco/Genjet with ee-kt, and paired according to chi2
- WW-ZZ Separation at the inclusive sample:
  - Intrinsic boson mass/width - lower limit: Overlapping ratio of 13%
  - + Jet confusion – Genjet: Overlapping ratio of **53%**
  - + Detector response – Recojet: Overlapping ratio of 58%

$$\text{overlapping ratio} = \sum_{bins} \min(a_i, b_i)$$

$$\chi^2 = \frac{(M_{12} - M_B)^2 + (M_{34} - M_B)^2}{\sigma_B^2}$$



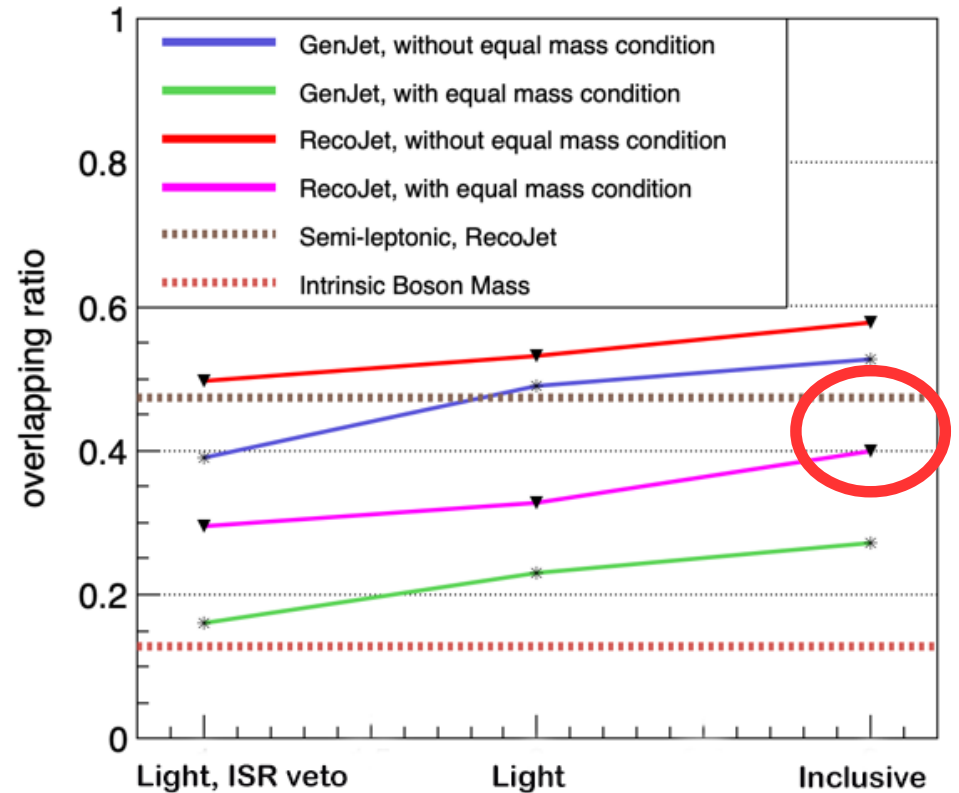
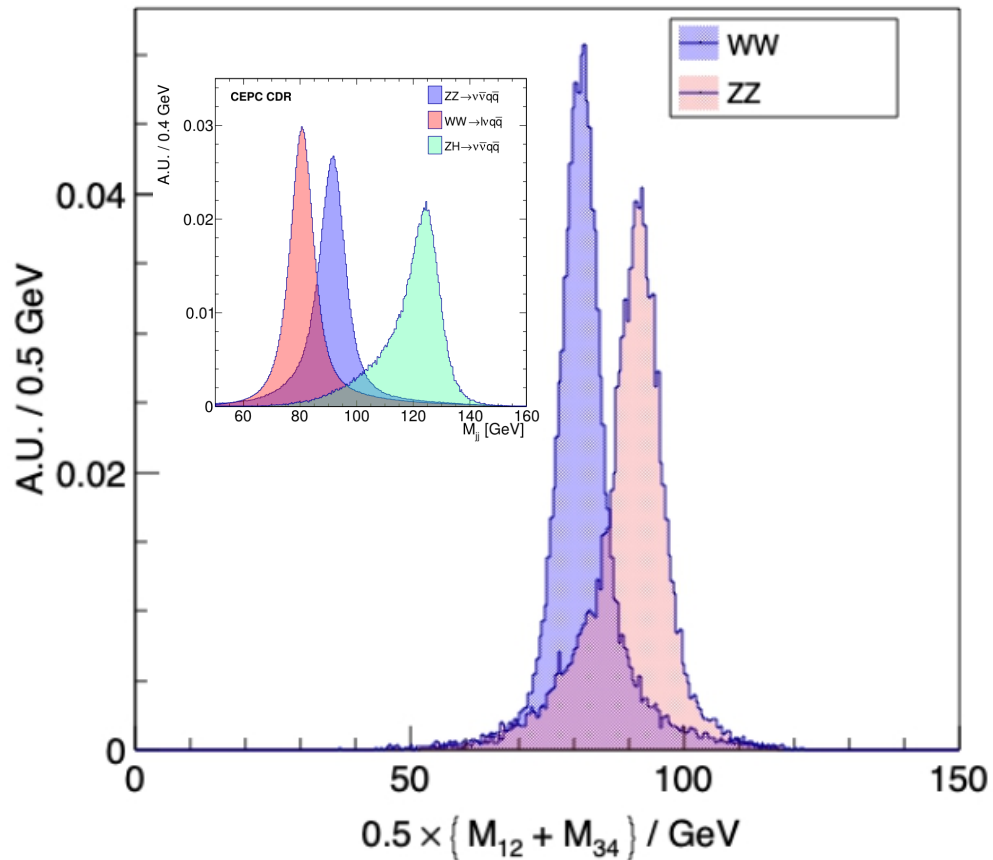
# Reconstructed mass of the two di-jet system



Equal mass condition  $|M_{12} - M_{34}| < 10$  GeV: At the cost of half the statistic, the overlapping ratio can be reduced from 58%/53% to 40%/27% for the Reco/Genjet

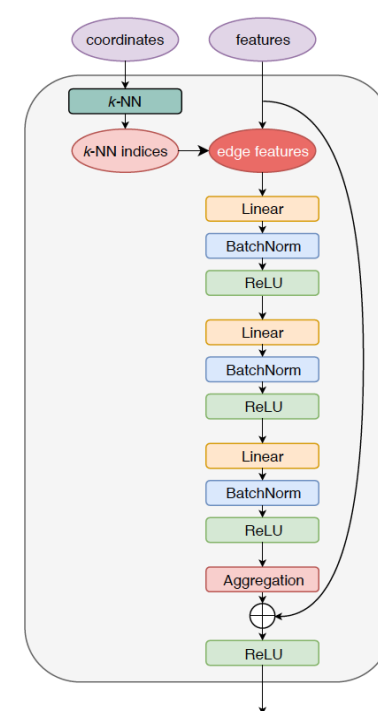
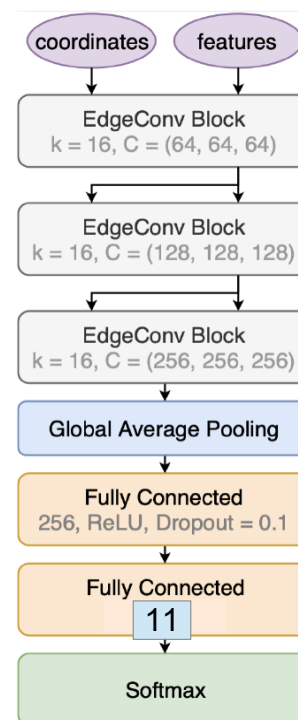
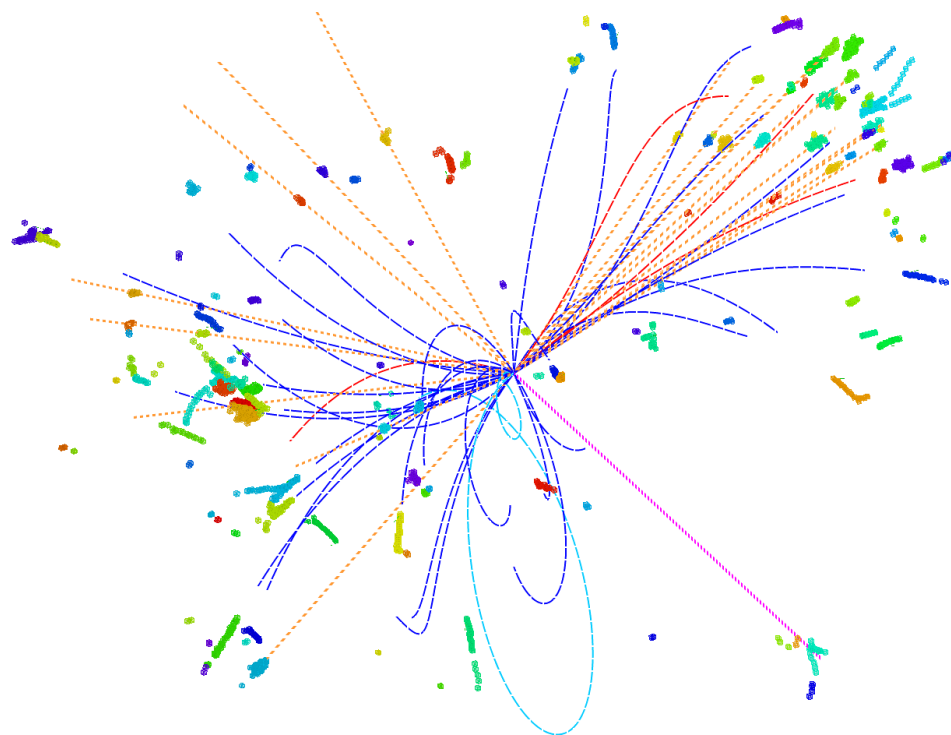


# Separation of full hadronic WW-ZZ event



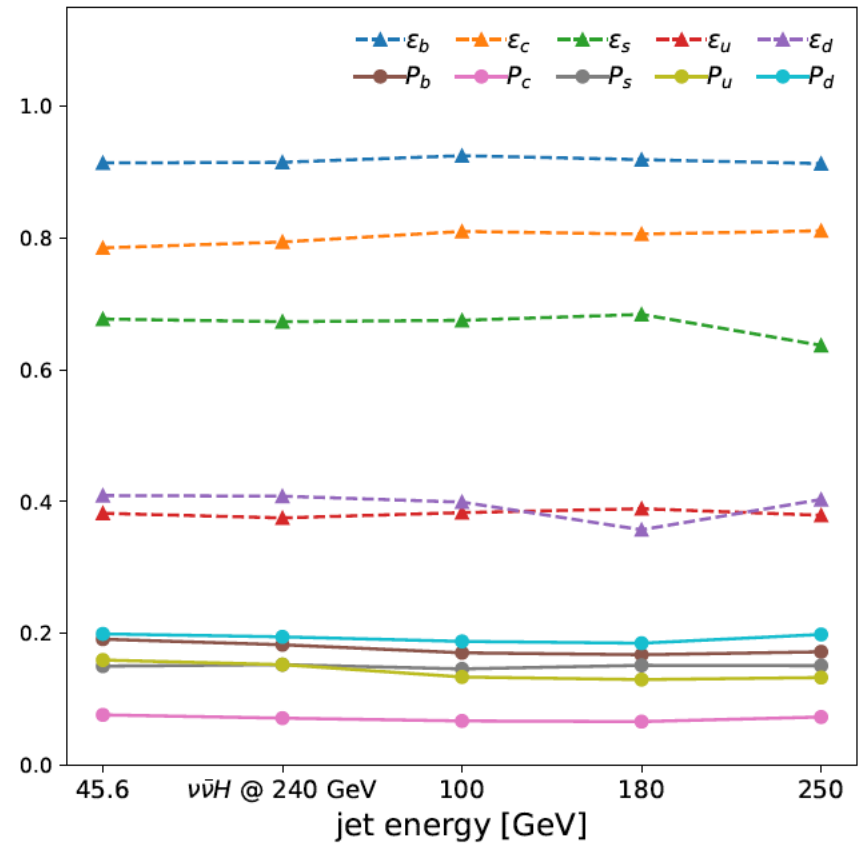
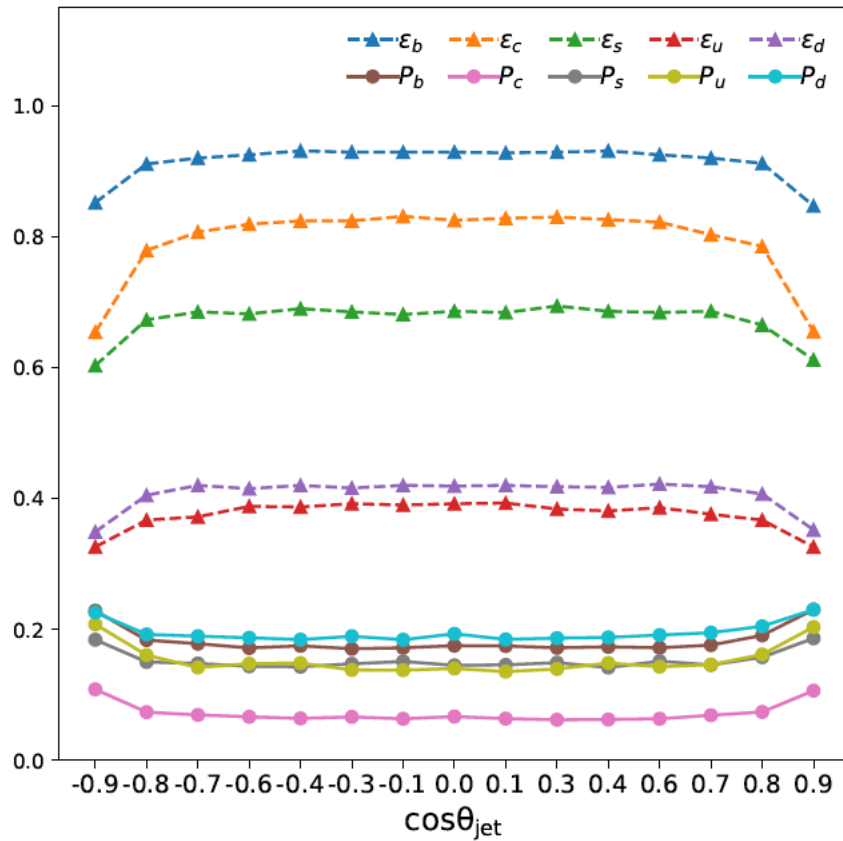
The CEPC Baseline could separate efficiently the WW-ZZ with full hadronic final state.  
 Critical to develop color singlet reconstruction: improve from the naive Jet clustering & pairing.  
*Quantified by differential overlapping ratio.*  
 Control of ISR photon/neutrinos from heavy flavor jet is important.

# Geo. & Tools

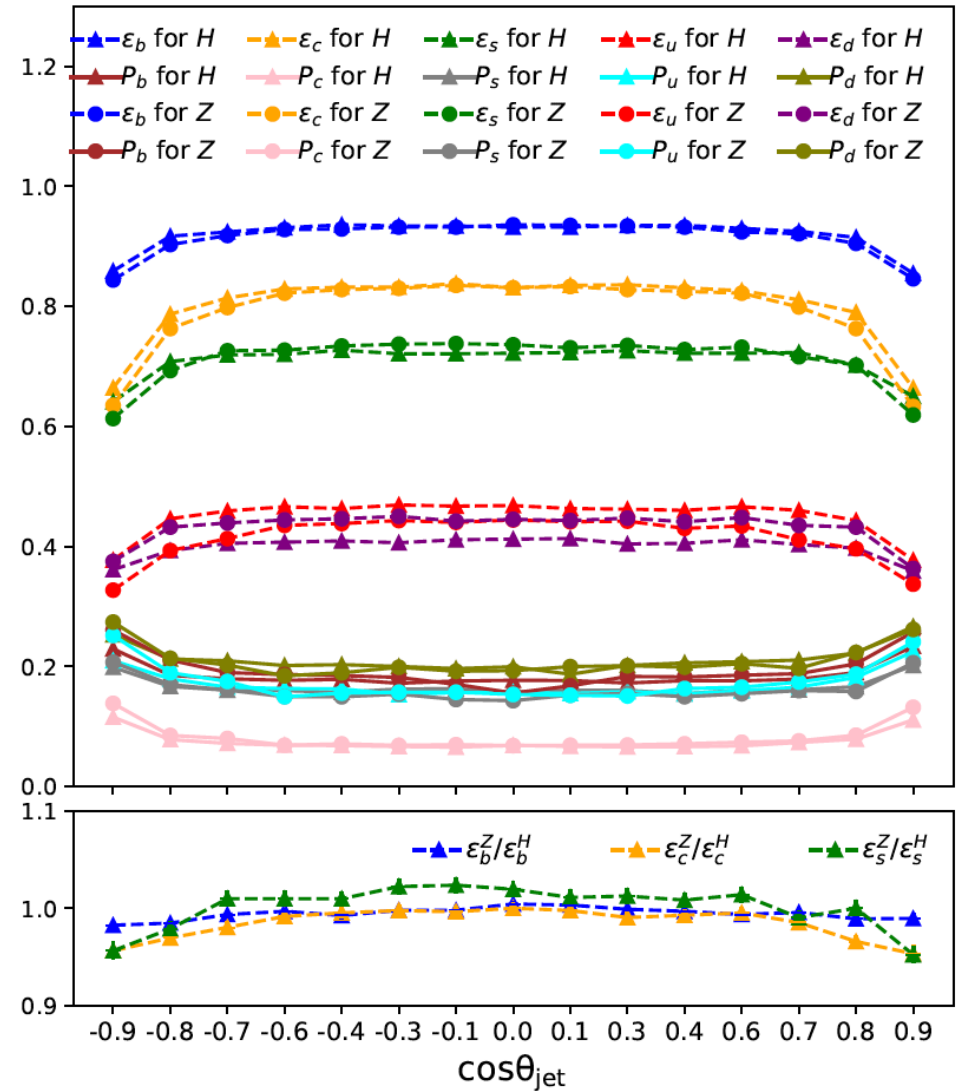
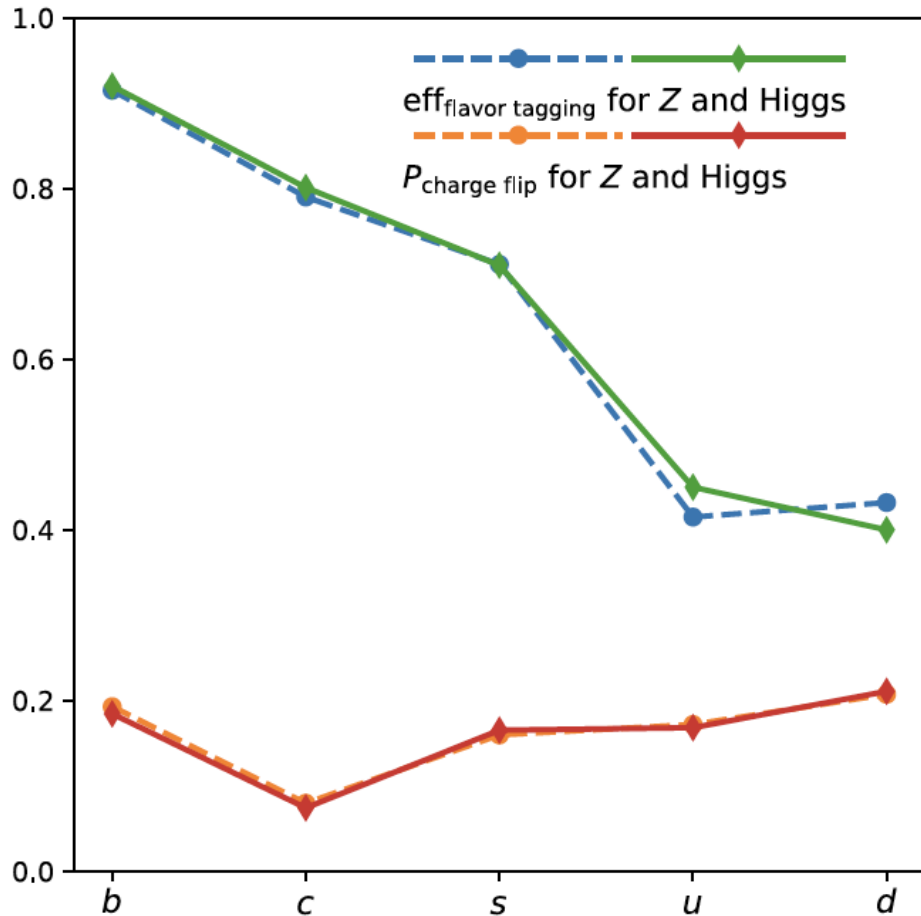


- **Jet origin identification: 11 categories (5 quarks + 5 anti quarks + gluon)**
- Full Simulated  $\nu\nu H$ , Higgs to two jets sample at CEPC baseline configuration: CEPC-v4 detector, reconstructed with **Arbor + ParticleNet (Deep Learning Tech.)**
  - Input: measurable information of all reconstructed jet particles ( $\sim 10$  float)
  - Output: 10(11)-likelihoods to different categories
- 1 Million samples each, 60/20/20% for training, validation & test

# Performance V.S. Jet Kinematics

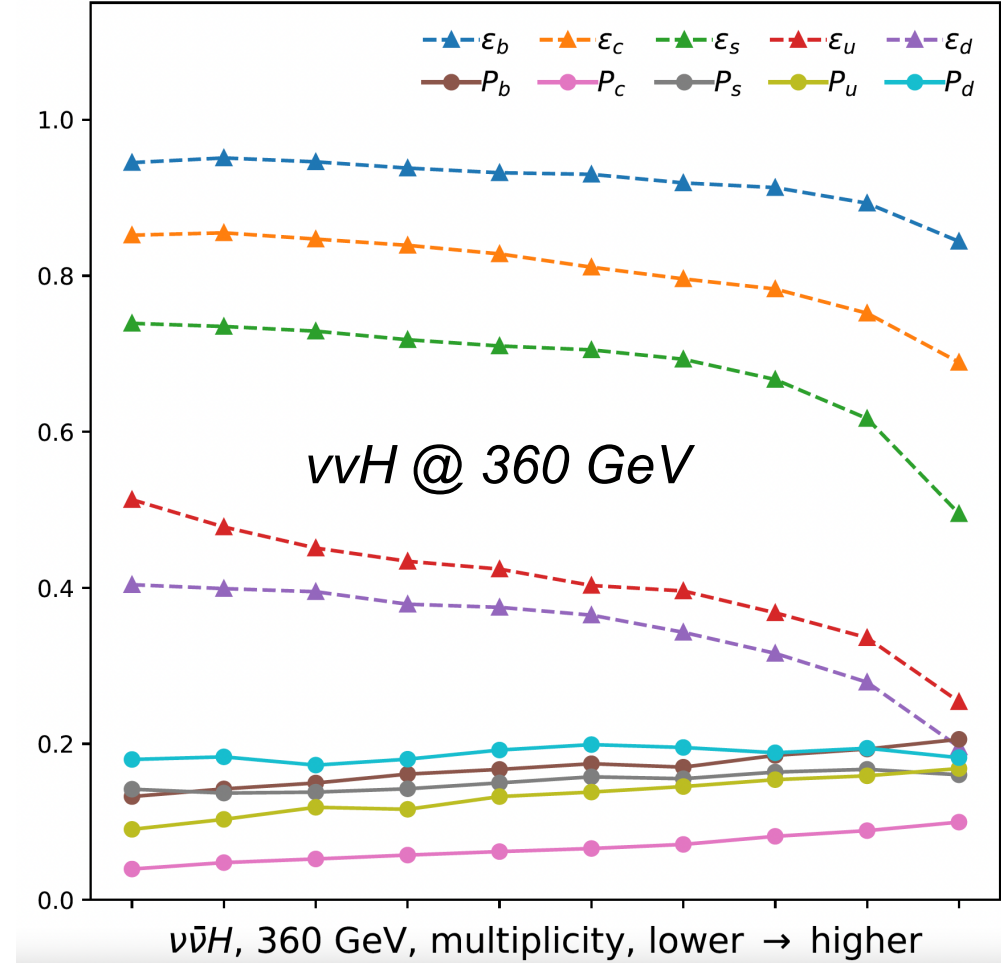
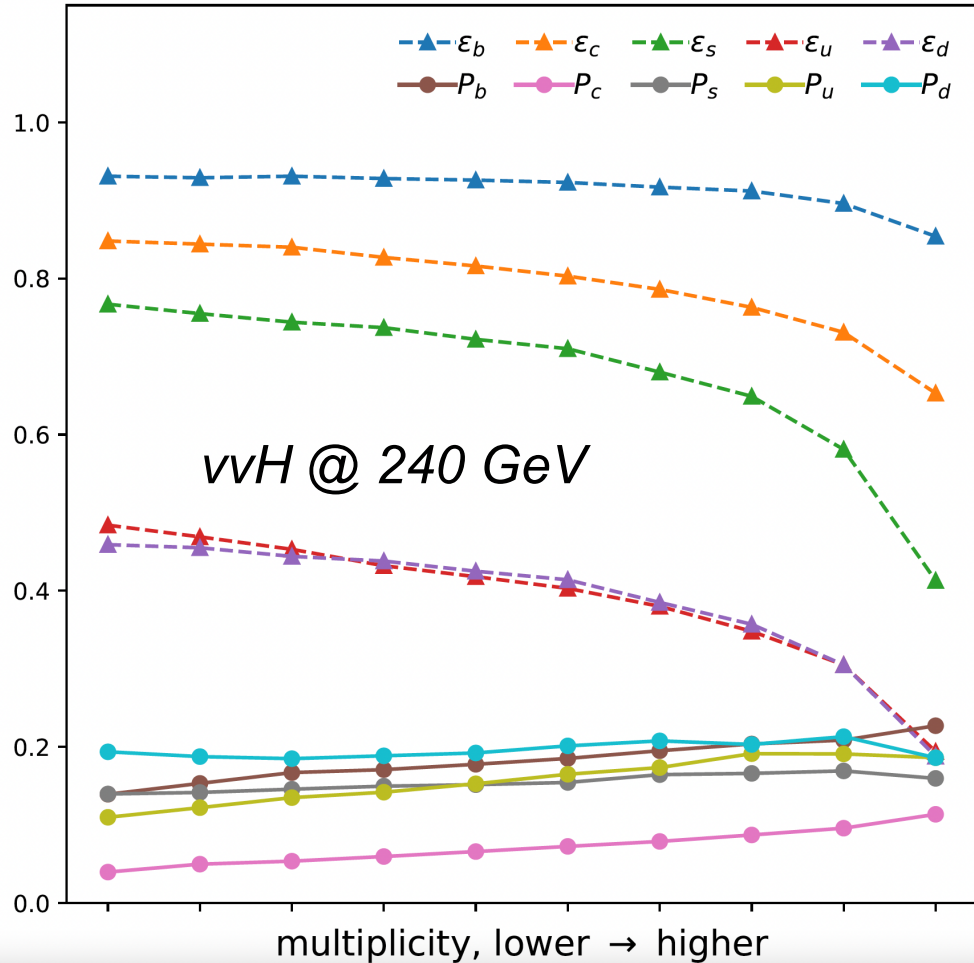


# Performance @ Z and Higgs



- *M10 instead of M11*

# V.S. Multiplicity

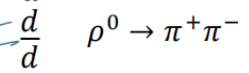
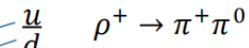
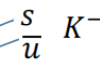
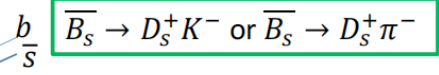


- ...many patterns need further understanding & towards further optimization...*

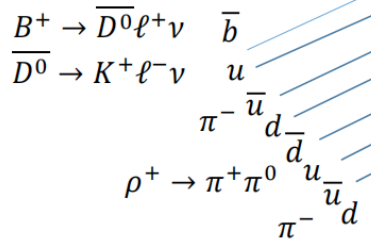
# B-charge flip rate: Bs oscillations

Opposite side

- p charged Leptons with impact param.
- p charged Kaons with impact param.
- p charged pions with impact param.
- p protons with impact param. ?

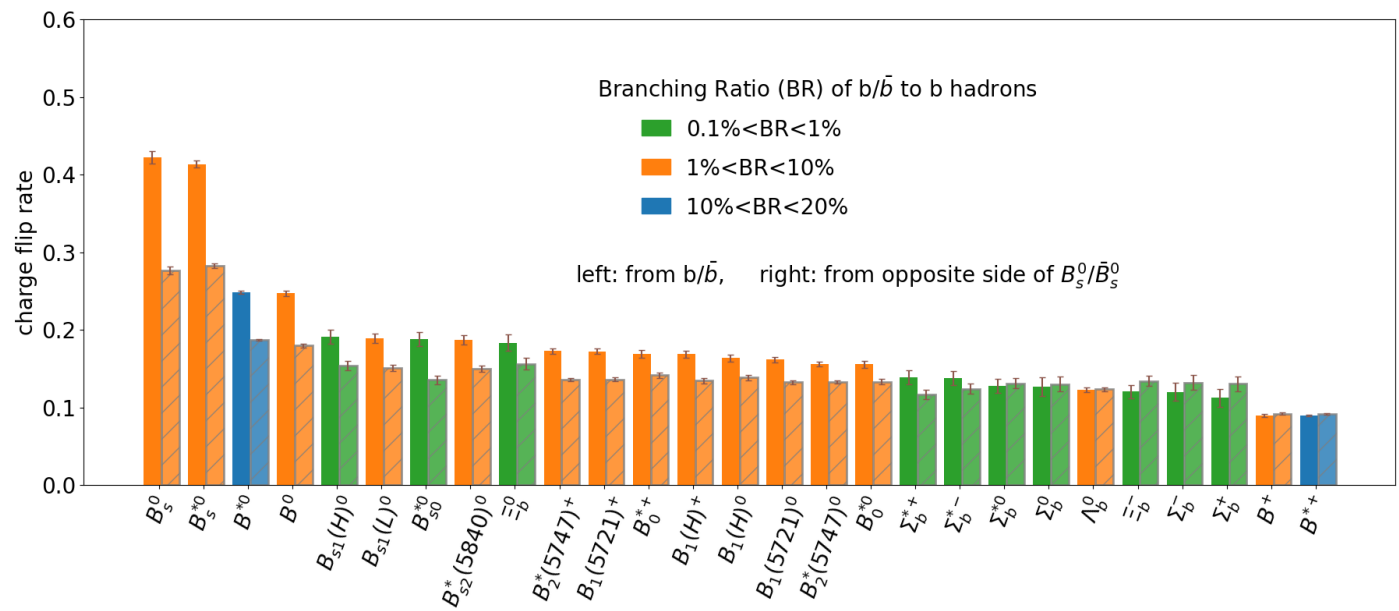


- Using all reco P (exc. Bs decay final state):
- Flip rate ~ 15%, Eff. Tagging power > 40%



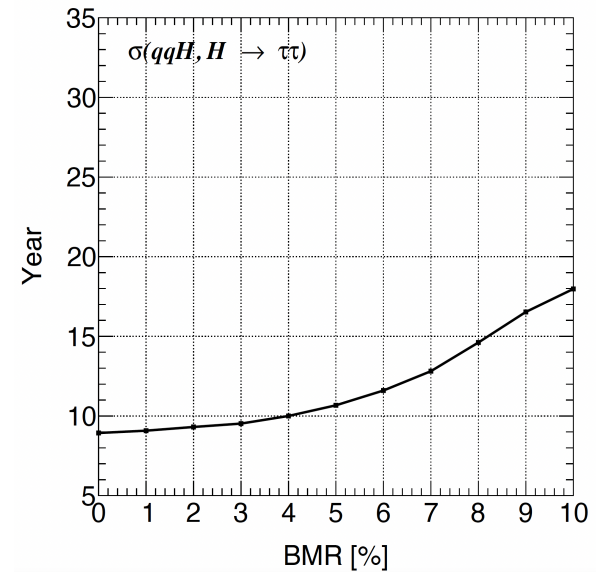
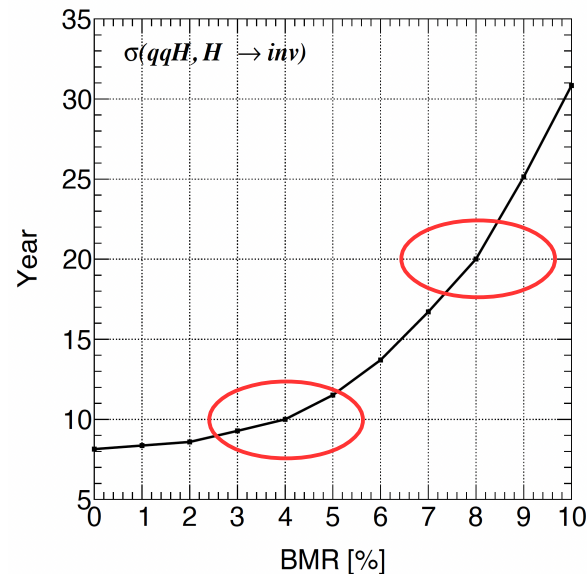
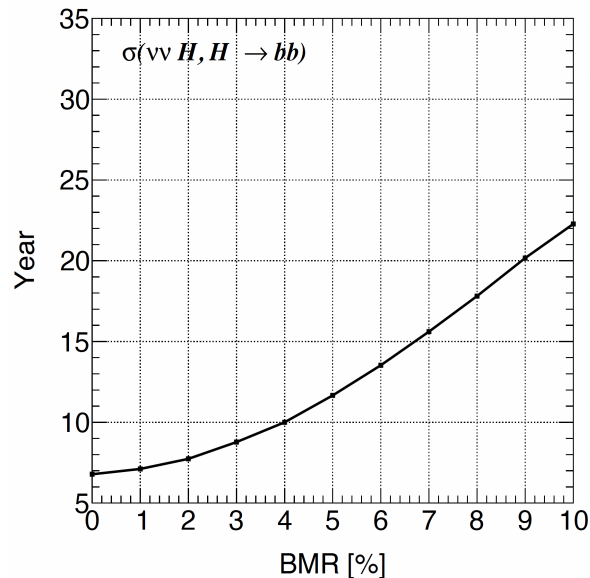
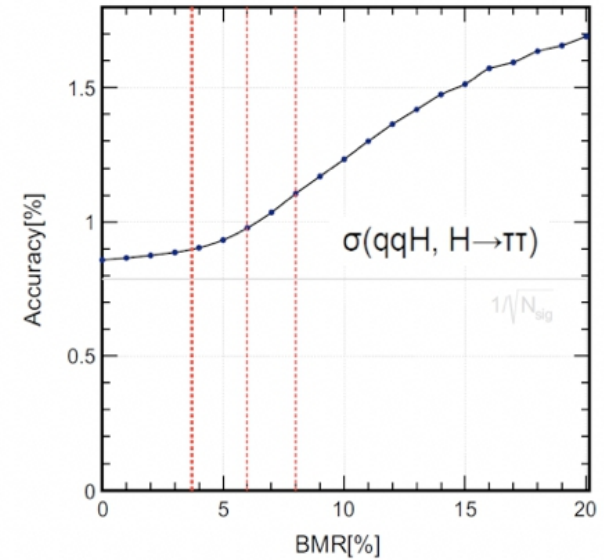
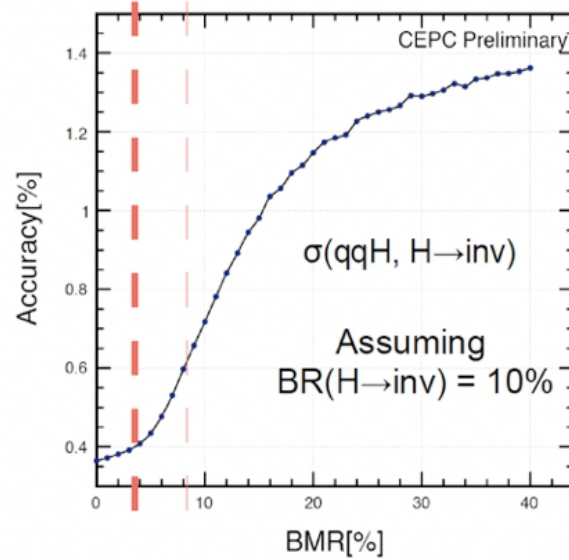
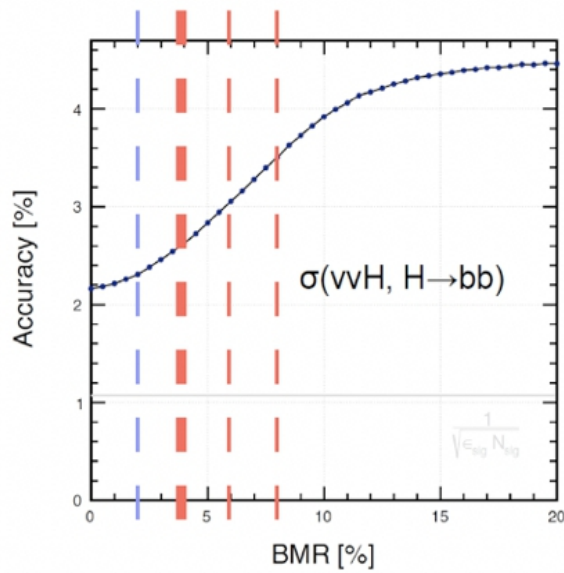
Same side

- p charged Kaons with impact param.
- p charged pions with impact param.





# BMR: impact on critical measurements



# s-jets: dependency on Leading hadron

