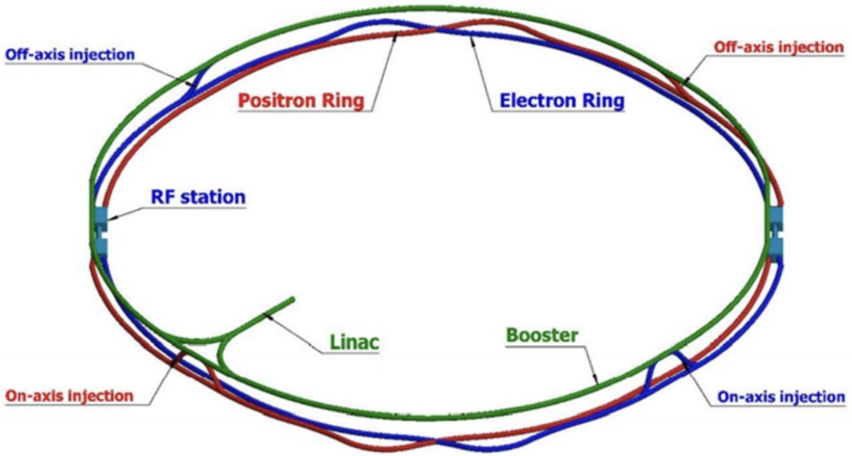
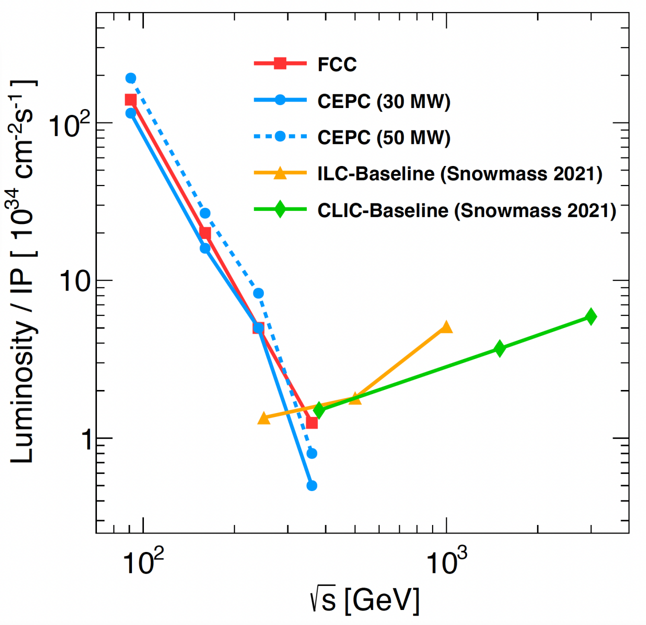
Trilogy of event reconstruction at electron positron Higgs factory

Manqi

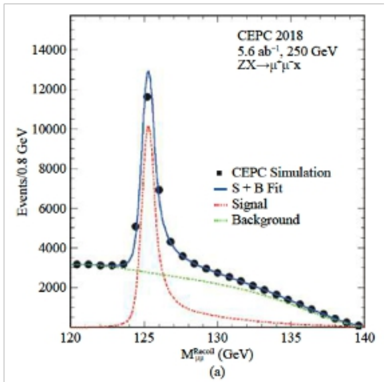
Outline

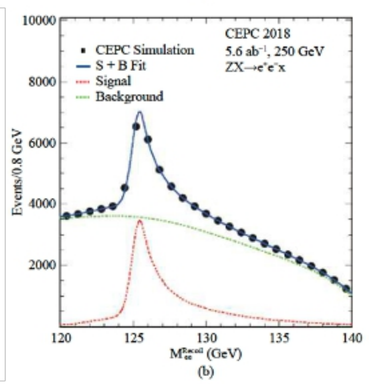
- CEPC Physics & Requirements
- Jet origin identification & Scaling laws at LLM
- 1-1 correspondence reconstruction
- Color Singlet identification
- Discussion

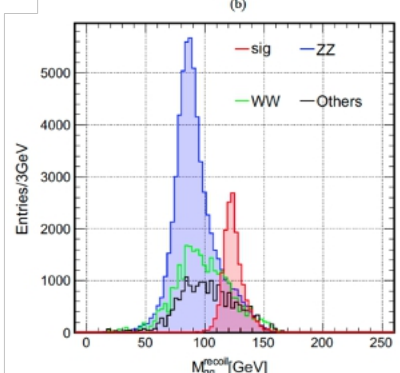




CEPC Physics: 4 Million Higgs + 4 Tera Z...







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

Precision Higgs physics at the CEPC*

Received 9 November 2018, Revised 21 January 2019, Published online 4 March 2019

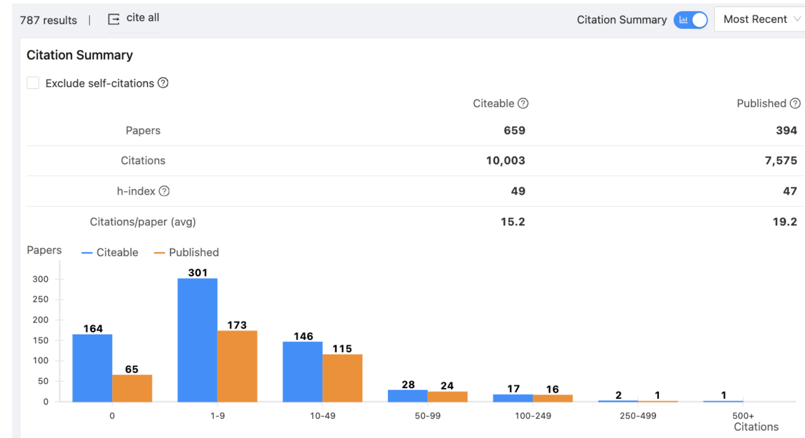


Table 2.1: Precision of the main parameters of interests and observables at the CEPC, from Ref. [1] and the references therein, where the results of Higgs are estimated with a data sample of 20 ab^{-1}. The HL-LHC projections of 3000 fb^{-1} data are used for comparison. [2]

Higgs			W, Z and top		
Observable	HL-LHC projections	CEPC precision	Observable	Current precision	CEPC precision
M_H	20 MeV	3 MeV	M_W	9 MeV	0.5 MeV
Γ_H	20%	1.7%	Γ_W	49 MeV	2 MeV
$\sigma(ZH)$	4.2%	0.26%	M_{top}	760 MeV	$\mathcal{O}(10)$ MeV
$B(H \to bb)$	4.4%	0.14%	M_Z	2.1 MeV	0.1 MeV
$B(H \to cc)$	-	2.0%	Γ_Z	2.3 MeV	0.025 MeV
$B(H \to gg)$	-	0.81%	R_b	3×10^{-3}	2×10^{-4}
$B(H \to WW^*)$	2.8%	0.53%	R_c	1.7×10^{-2}	1×10^{-3}
$B(H \to ZZ^*)$	2.9%	4.2%	R_μ	2×10^{-3}	1×10^{-4}
$B(H \to \tau^+\tau^-)$	2.9%	0.42%	R_τ	1.7×10^{-2}	1×10^{-4}
$B(H \to \gamma\gamma)$	2.6%	3.0%	A_μ	1.5×10^{-2}	3.5×10^{-5}
$B(H \to \mu^+\mu^-)$	8.2%	6.4%	A_τ	4.3×10^{-3}	7×10^{-5}
$B(H \to Z\gamma)$	20%	8.5%	A_b	2×10^{-2}	2×10^{-4}
$Bupper(H \to inv.)$	2.5%	0.07%	N_ν	2.5×10^{-3}	2×10^{-4}

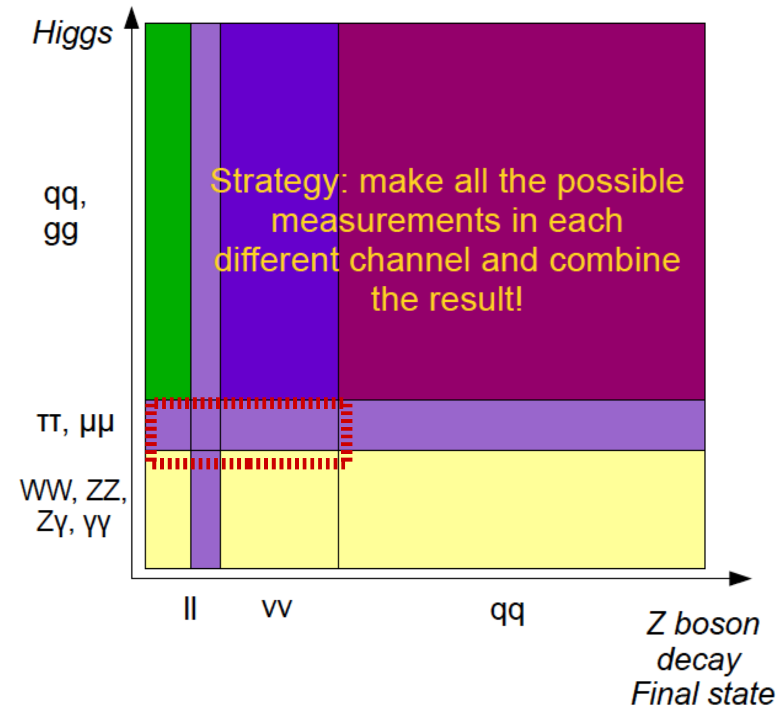
Scientific Significance quantified by **CEPC physics** studies, via full simulation/phenomenology studies:

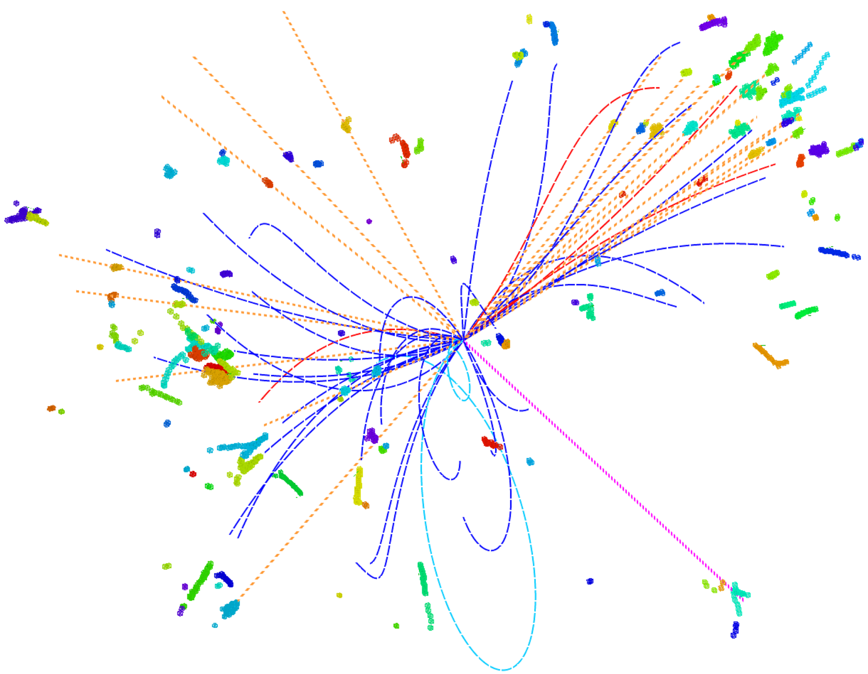
- Higgs: Precisions exceed HL-LHC ~ 1 order of magnitude.
- EW: Precision improved from current limit by 1-2 orders.
- Flavor Physics, sensitive to NP of 10 TeV or even higher.
- Sensitive to varies of NP signal.
- …

White papers +
~300 Journal/AxXiv citables

Performance requirements

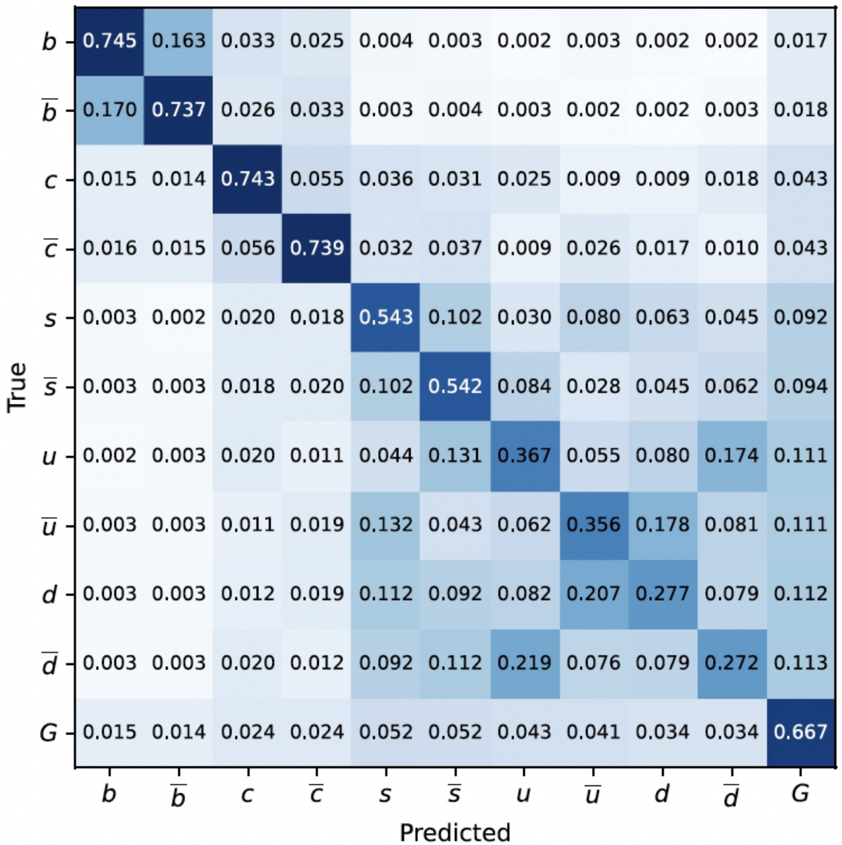
- To reconstruct all Physics Object, especially Jets
  - Z & W: ~ 70% goes to a pair of jets
  - Higgs: ~97% final state with jets (ZH events)
  - Top: t → W + b

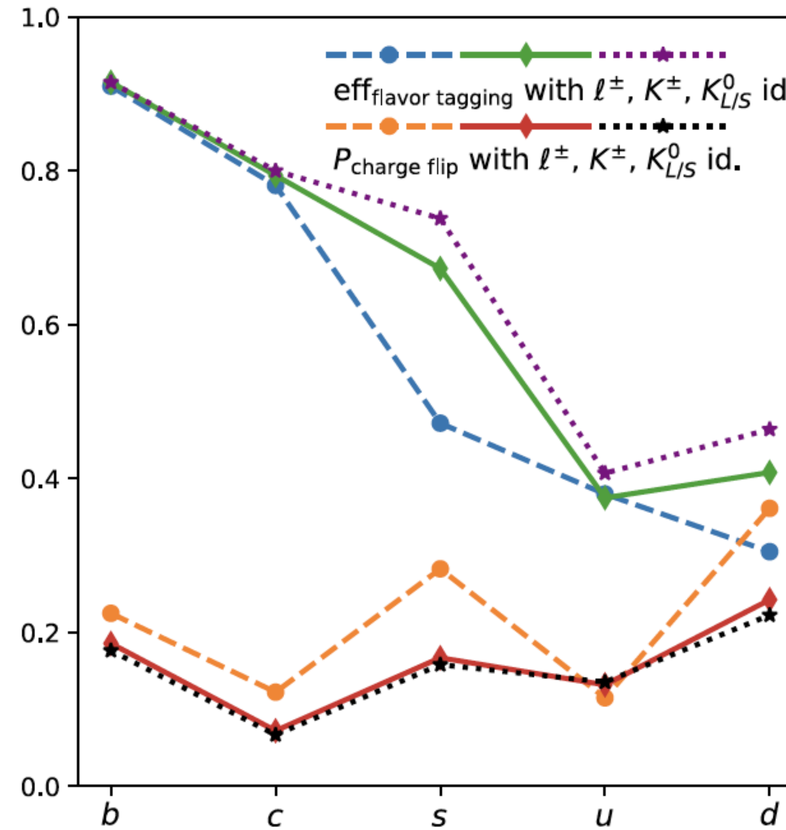




- Look inside the jet: **1-1 correspondence reco.**
  - ~ confusion free PFA
  - Larger acceptance...
  - Excellent intrinsic resolutions
  - Extremely stable...
- Be addressed by state-of-art detector design, technology, and reconstruction algorithm!

Jet origin id

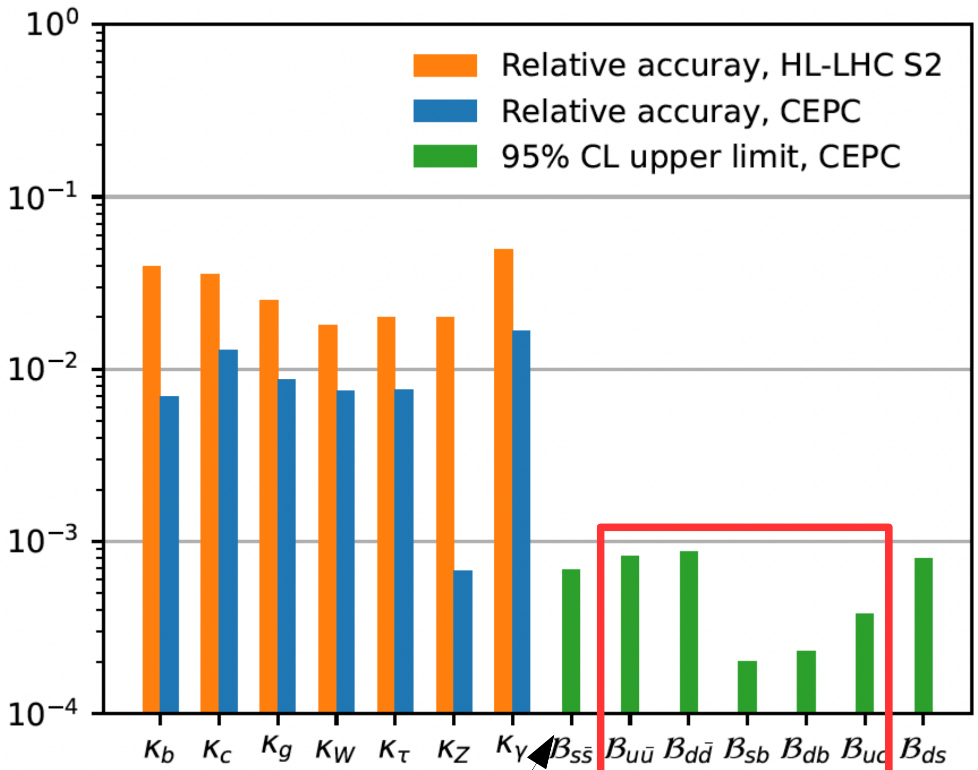


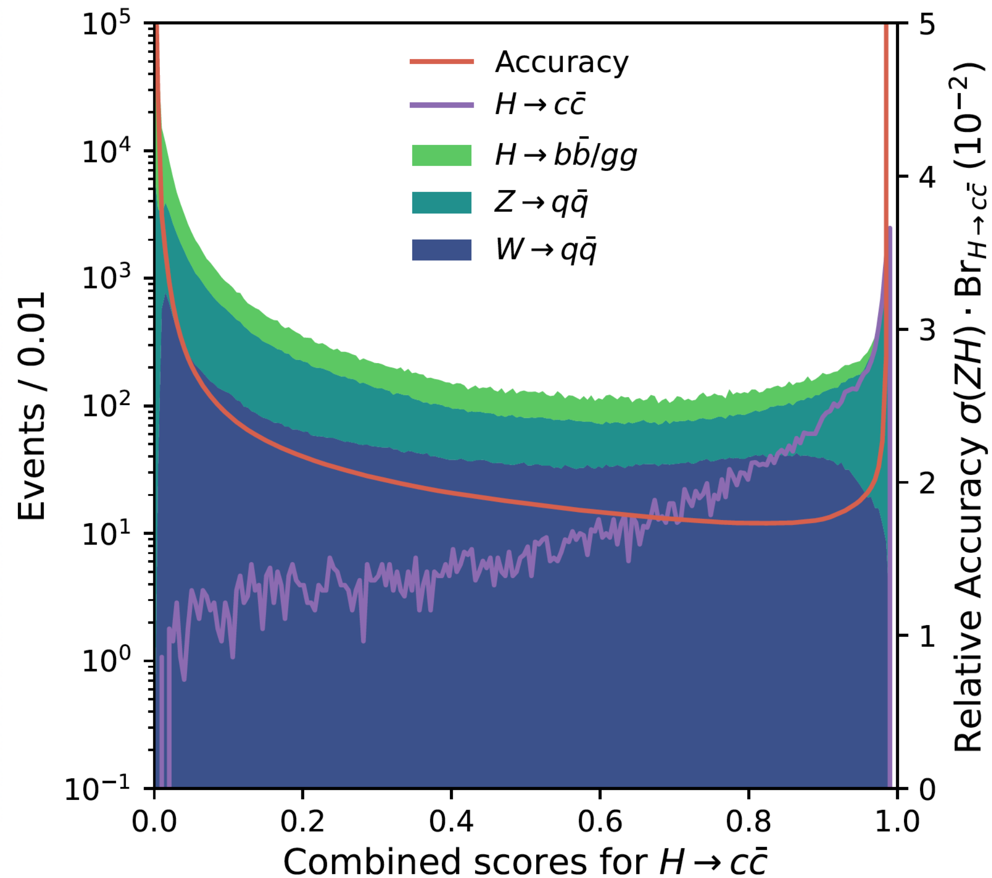


- 11 categories (5 quarks + 5 anti quarks + gluon) identification, realized at Full Simulated di-jet events at CEPC CDR baseline with Arbor + ParticleNet
- Published in PRL *132, 221802 (2024)*. Comment from the referee: *"demonstrate the world-leading performance of tagger", "a "game changer" and opens 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 :
  - vvH, H→cc: 3% → 1.7%
  - Vcb: 0.75% → 0.5%
  - Applicable to Vcs, Vts, etc.

Updated result on $\sin^2\theta^l_{eff}$ measurement

Table 2. Sensitivity S of different final state particles.

$\sqrt{s}$/GeV	S of $A^{e/\mu}_{FB}$	S of A^{d}_{FB}	S of A^{u}_{FB}	S of A^{s}_{FB}	S of A^{c}_{FB}	S of A^{b}_{FB}
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

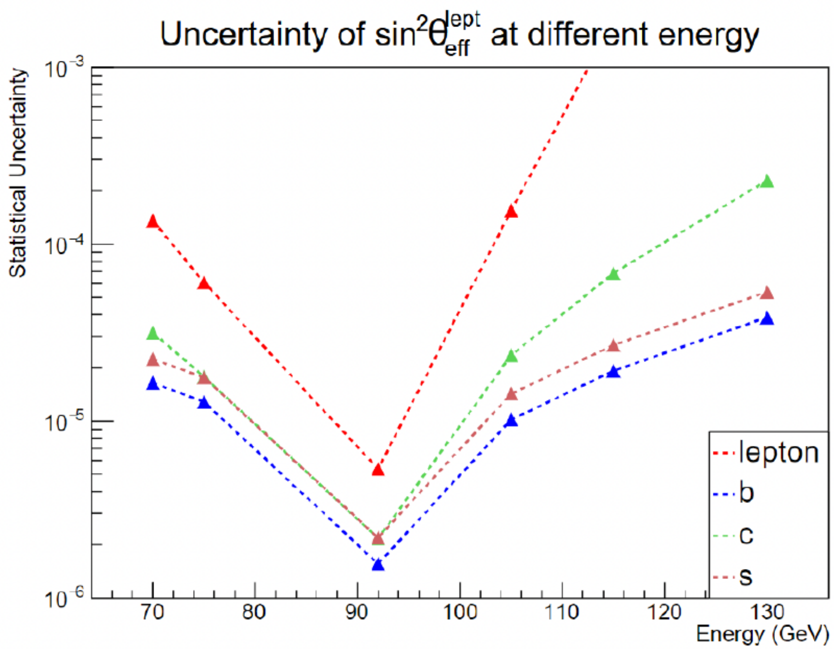
Table 3. Cross section of process $e^+e^- \to f\bar{f}$ calculated using the ZFITTER package. Values of the fundamental parameters are set as $m_Z = 91.1875$ GeV, $m_t = 173.2$ GeV, $m_H = 125$ GeV, $\alpha_s = 0.118$ and $m_W = 80.38$ GeV.

$\sqrt{s}$/GeV	σ_μ/mb	σ_d/mb	σ_u/mb	σ_s/mb	σ_c/mb	σ_b/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 $\sin^2\theta^l_{eff}$ measurement.

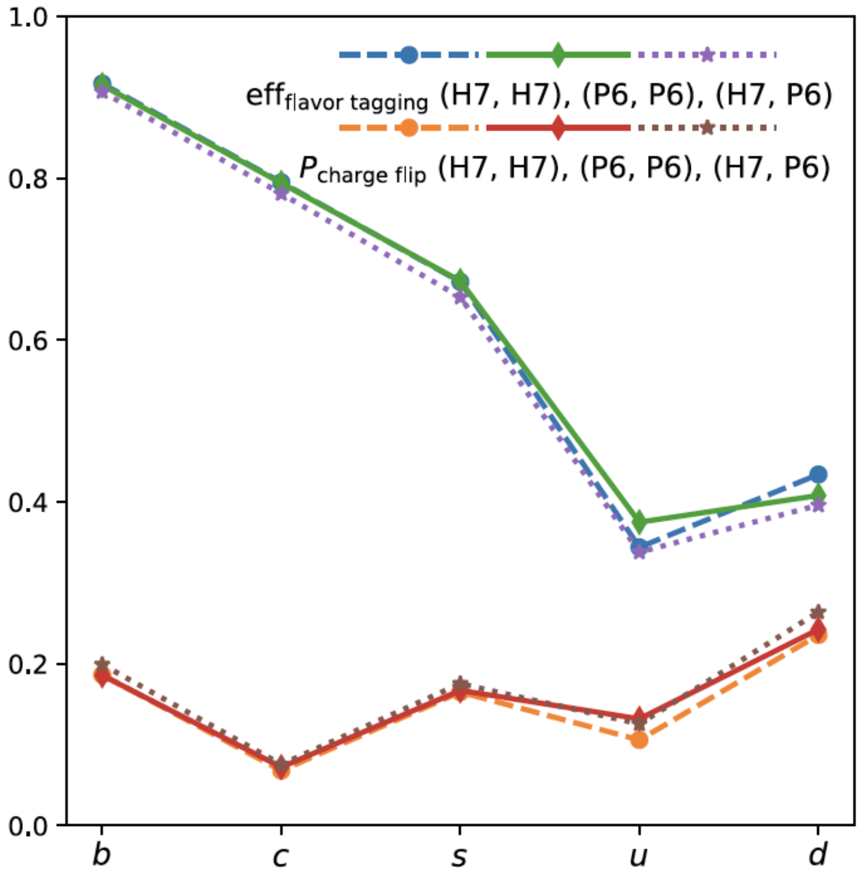
(Using one-month data collection, ~ *4e12/24 Z events* at Z pole)

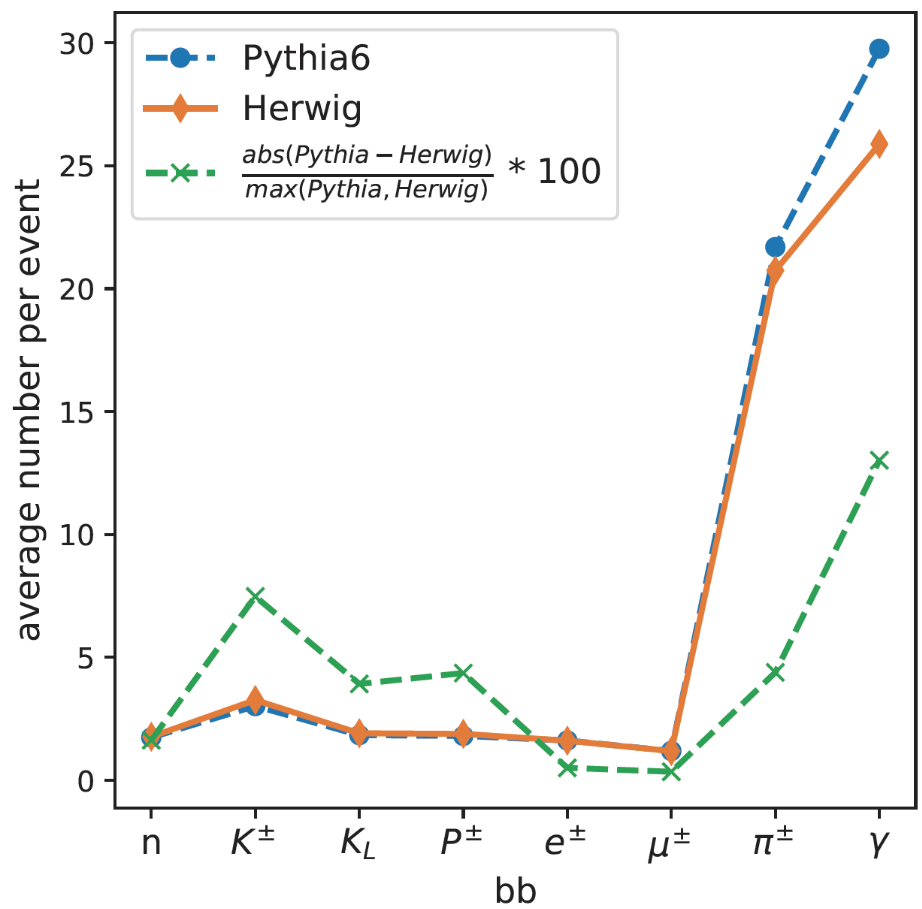


$\sqrt{s}$	b	c	s
70	1.6×10^{-5}	3.2×10^{-5}	2.2×10^{-5}
75	1.3×10^{-5}	1.8×10^{-5}	1.8×10^{-5}
92	1.6×10^{-6}	2.2×10^{-6}	2.2×10^{-6}
105	1.0×10^{-5}	2.4×10^{-5}	1.4×10^{-5}
115	1.9×10^{-5}	6.8×10^{-5}	2.7×10^{-5}
130	3.9×10^{-5}	2.3×10^{-4}	5.4×10^{-5}

...+ Significant impact on Flavor Physics, especially those with Bs oscillation...

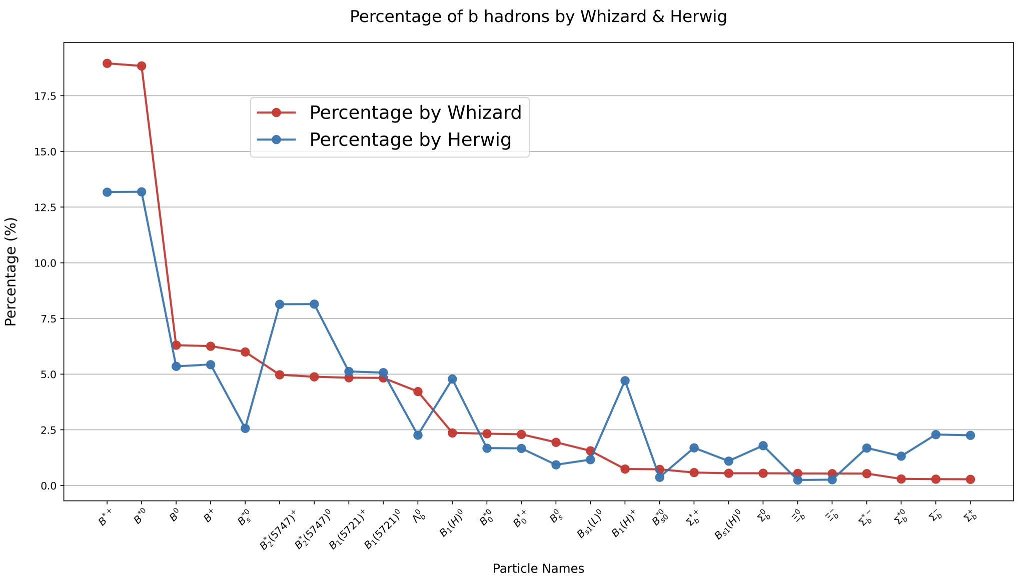
V.S. Hadronization models

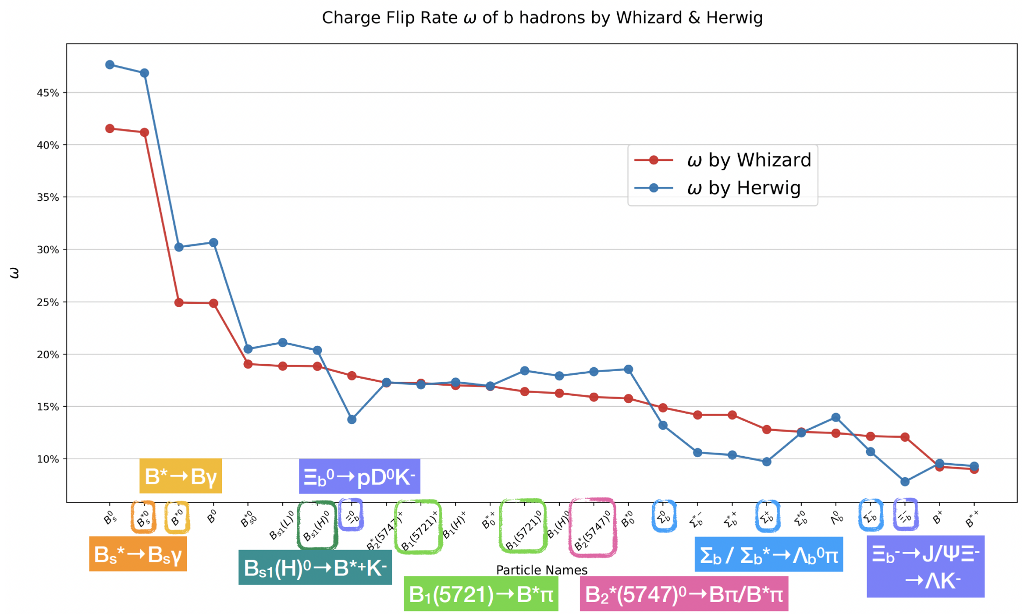


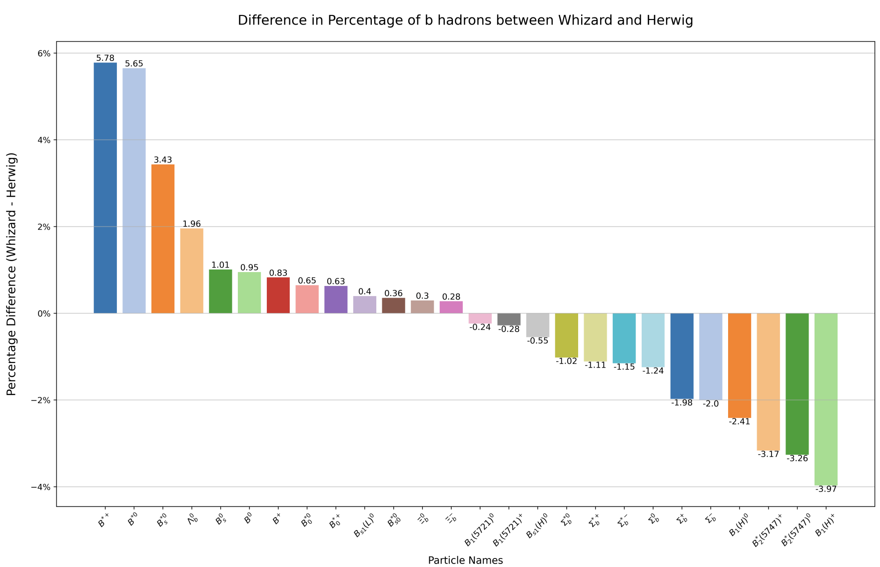


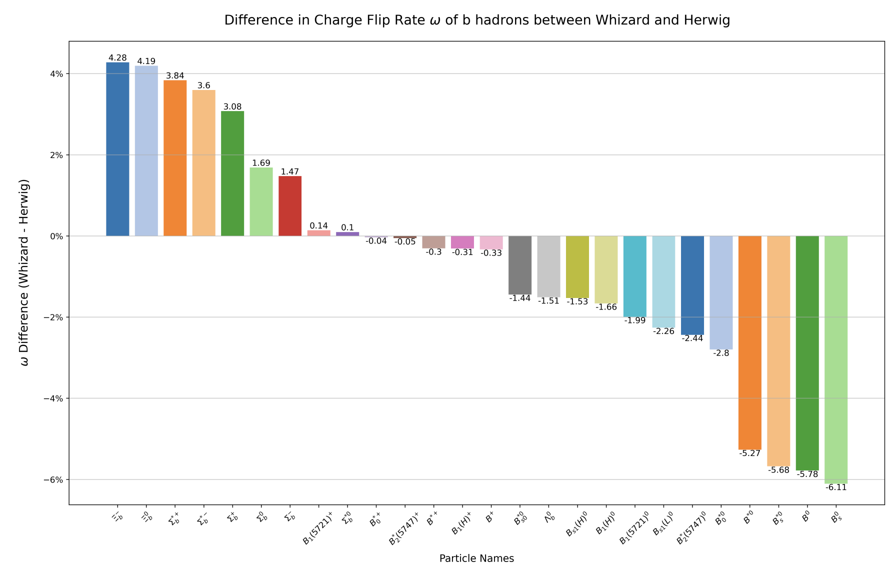
- Different hadronization model have significantly different predictions...

b-jet: leading b-hadrons & flip rates

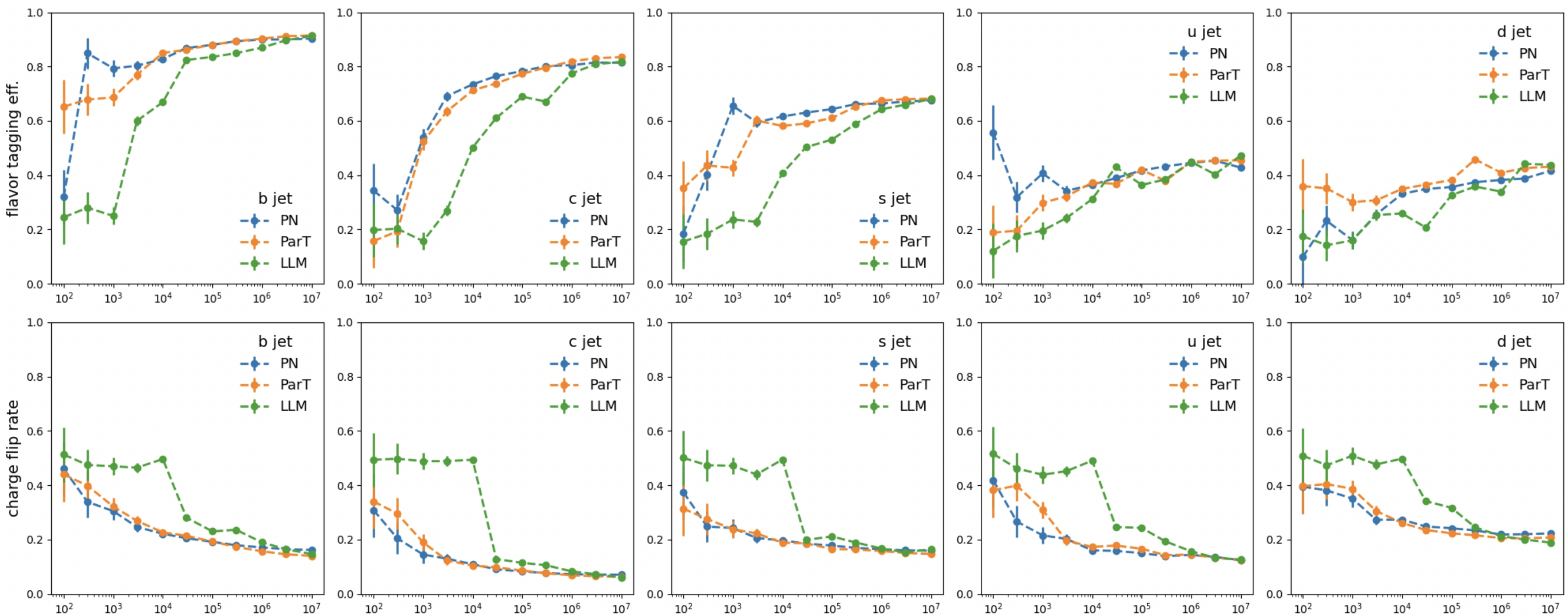








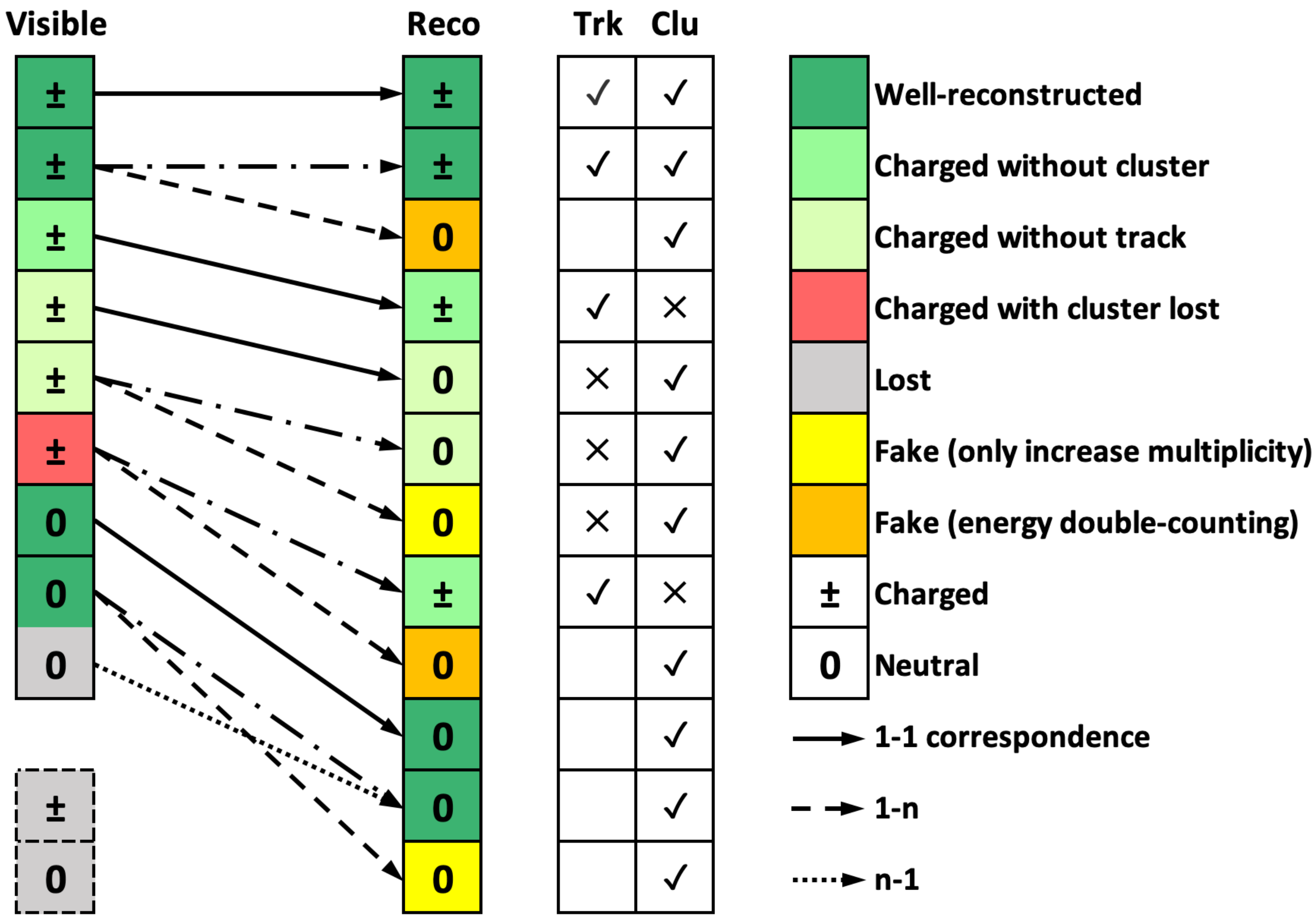
From specialized Models to LLM



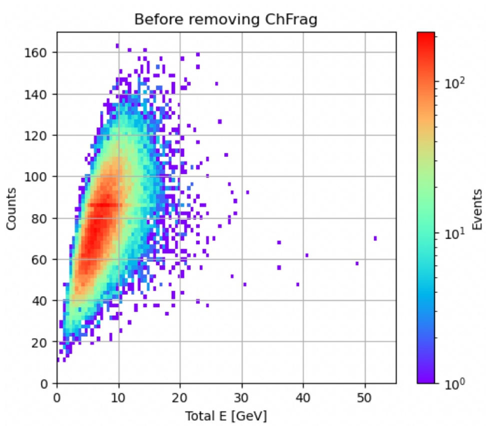
- Comparable result with different scaling behavior
- Para. Numbers: PN 360k, ParT 2.4M, BINBBT(Large Language Base Model) 150 M
- More details at: *https://arxiv.org/pdf/2412.00129*

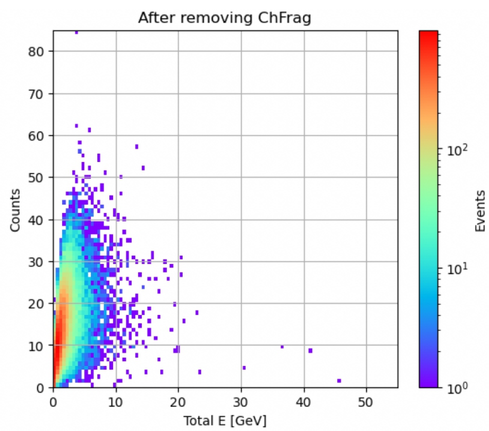
超对称
Super Symmetry Technologies

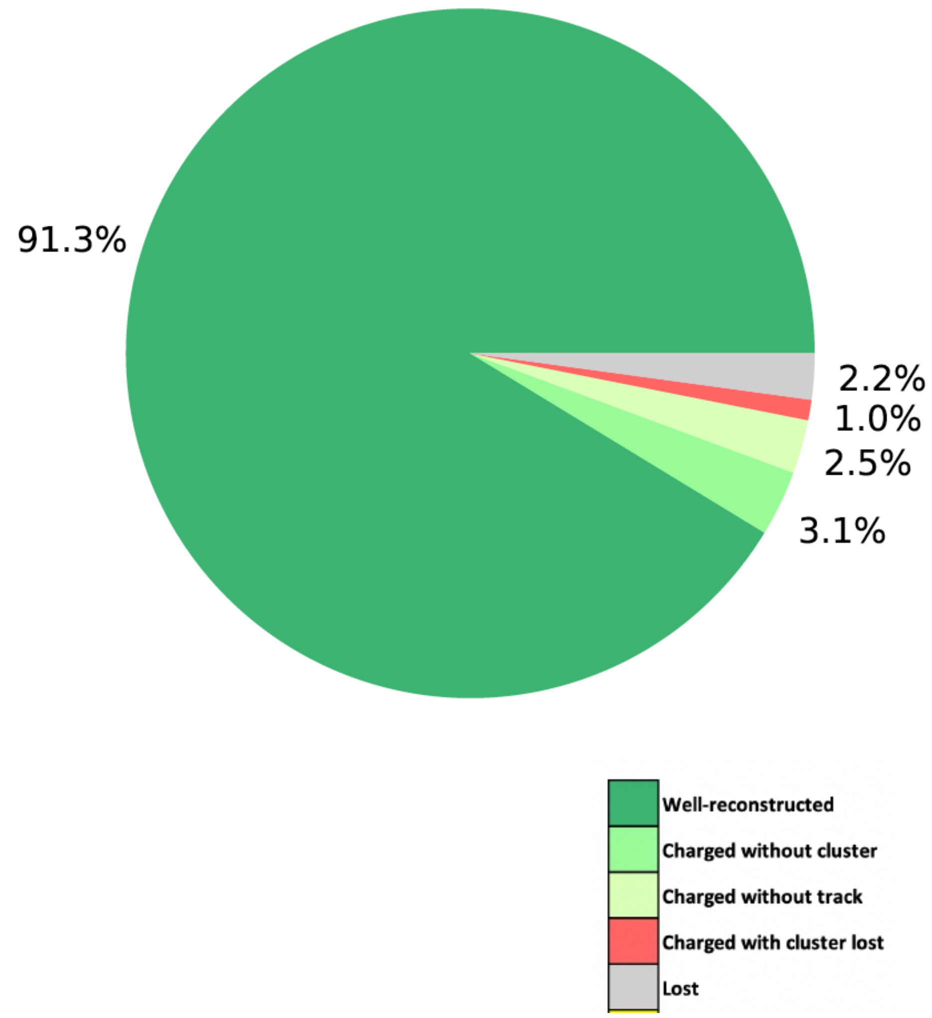
1-1 correspondence: ultimate Mapping between visible & reco



Confusion: frag. Identification & veto



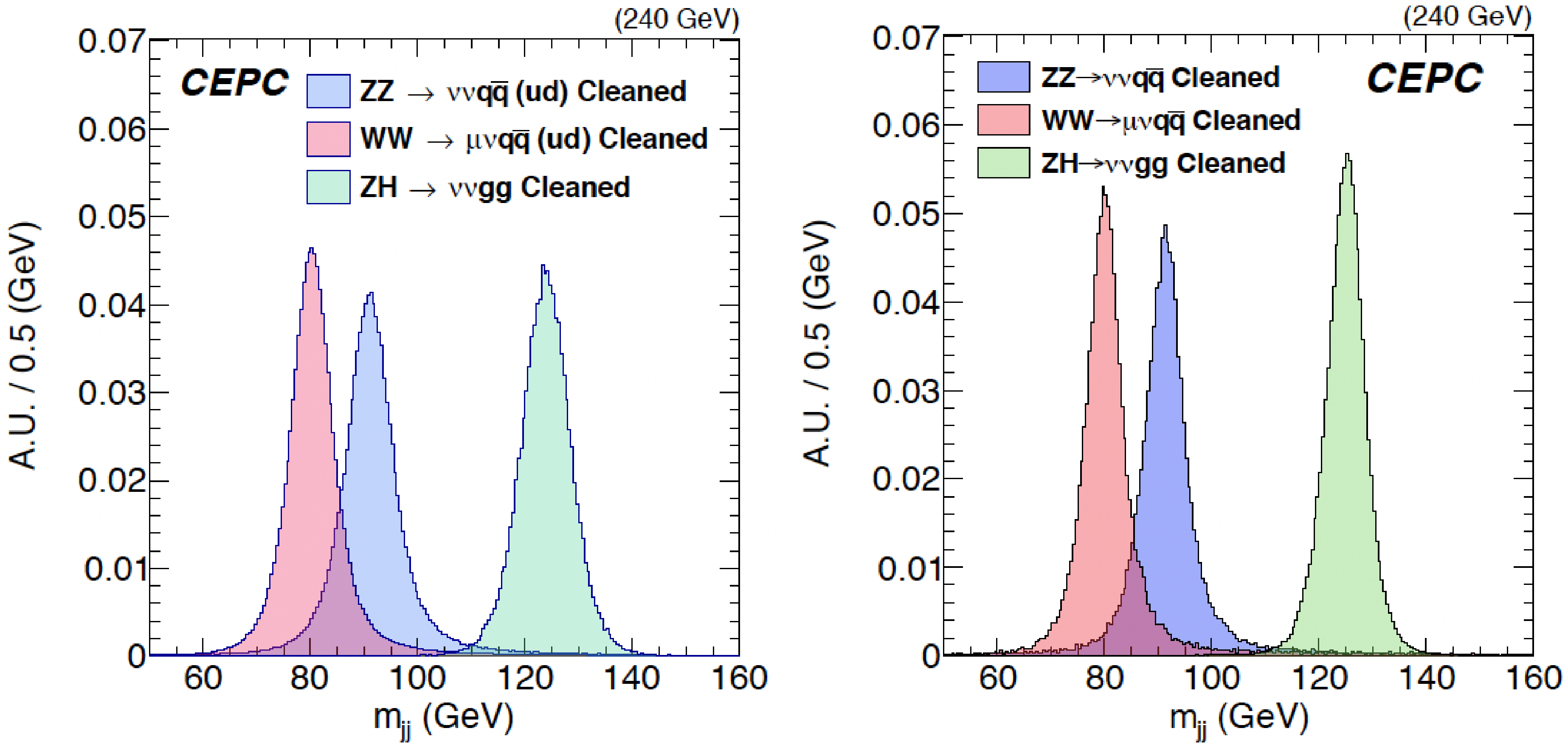




Fake particle originated Confusion reduced by 1 order of magnitude, at nominal ννH, H→gg event

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



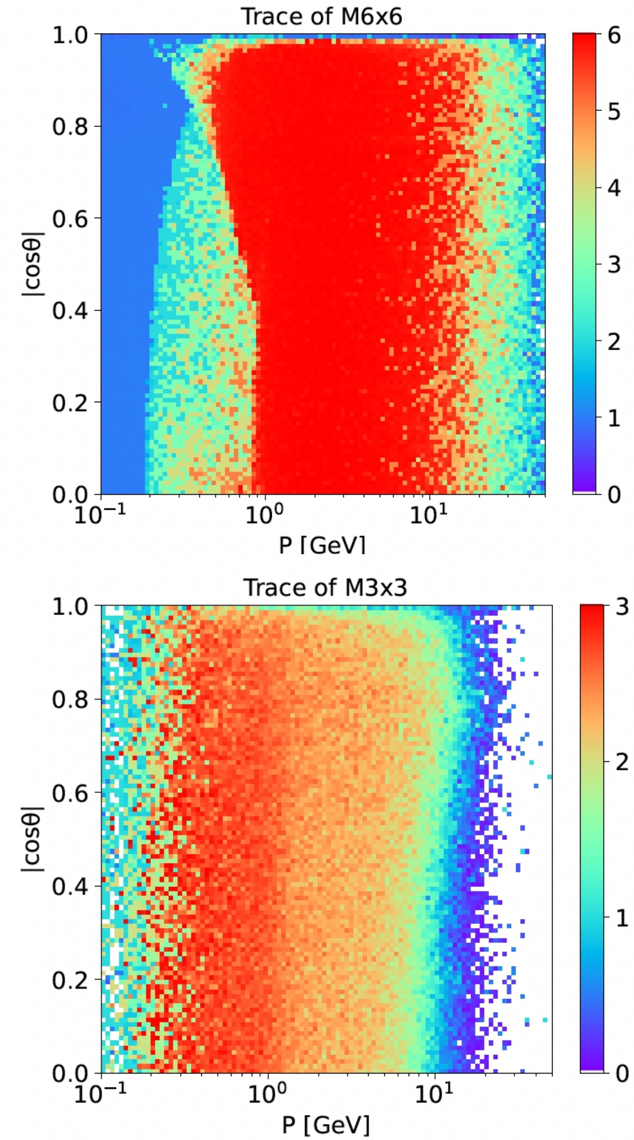
Detector change (usage of high density scintillating glass HCAL): BMR 3.7% → 3.4%;

AI enhanced reconstruction: 3.4% → 2.8%.

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

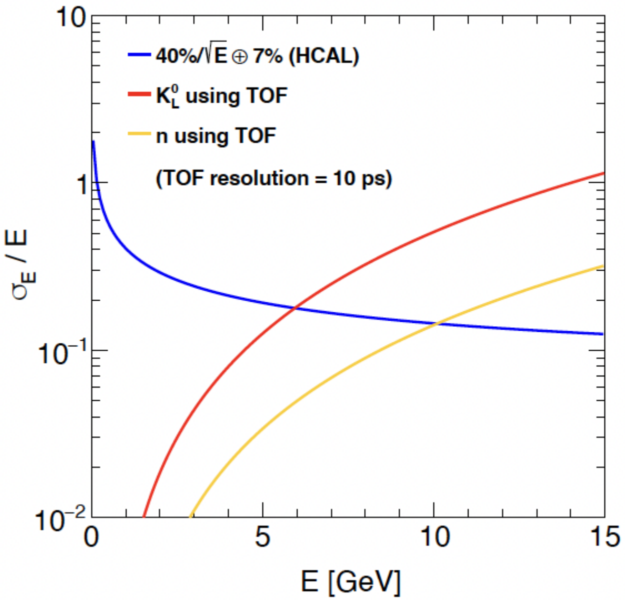
Pid: differential performance

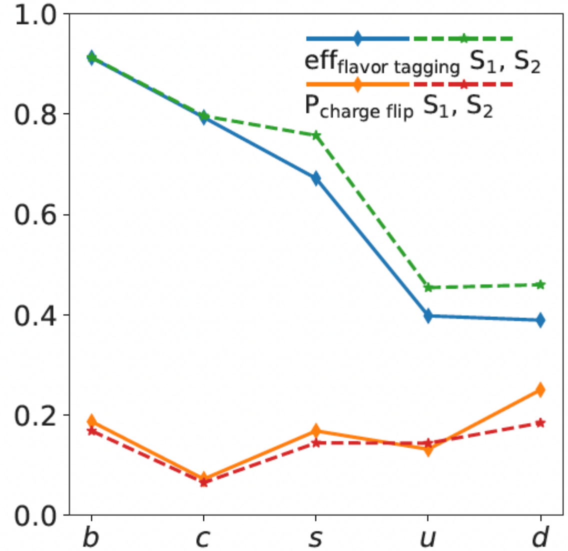
True \ Predicted	$e^{\pm}$	$\mu^{\pm}$	$\pi^{\pm}$	$K^{\pm}$	$p/\bar{p}$	γ	K_L^0	n	$\bar{n}$	others
$e^{\pm}$	0.999	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$\mu^{\pm}$	0.000	0.985	0.015	0.000	0.000	0.000	0.000	0.000	0.000	0.000
$\pi^{\pm}$	0.000	0.019	0.977	0.003	0.000	0.000	0.000	0.000	0.000	0.000
$K^{\pm}$	0.000	0.001	0.003	0.985	0.011	0.000	0.000	0.000	0.000	0.000
$p/\bar{p}$	0.000	0.000	0.001	0.012	0.979	0.000	0.000	0.000	0.000	0.008
γ	0.000	0.000	0.000	0.000	0.000	0.973	0.003	0.004	0.000	0.020
K_L^0	0.000	0.000	0.000	0.000	0.000	0.007	0.756	0.070	0.048	0.119
n	0.000	0.000	0.000	0.000	0.000	0.003	0.083	0.803	0.027	0.084
$\bar{n}$	0.000	0.000	0.000	0.000	0.000	0.003	0.052	0.026	0.796	0.123
others	0.000	0.000	0.000	0.000	0.000	0.014	0.025	0.095	0.011	0.854

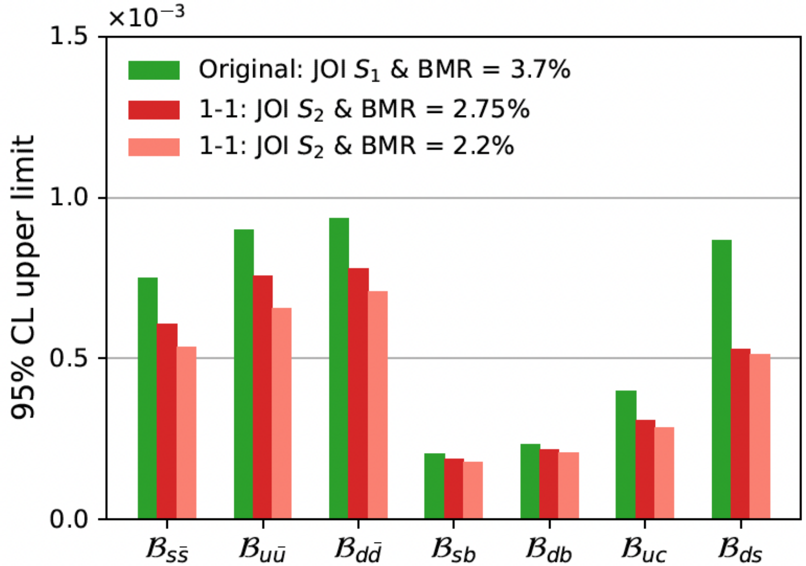


Neutral Hadron ID: excellent Calorimetry with ToF capability (δt~100 ps/hit)

Perspectives with 1-1 correspondence







- ToF enhanced energy measurement: BMR: 2.8 → 2.2-2.4
  - Need excellent CALO + ToF ~ 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

JHEP
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→ $b\bar{b}, c\bar{c}, gg$ measurement at CEPC

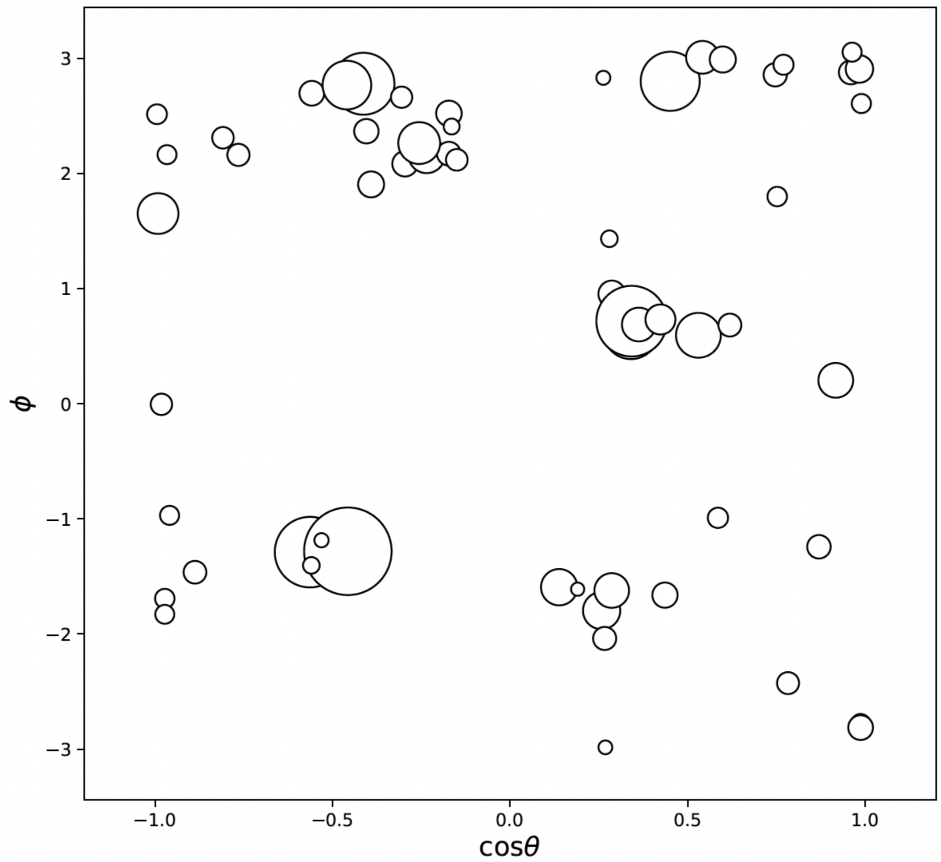
Yongfeng Zhu, Hanhua Cui and Manqi Ruan
Institute of High Energy Physics, Chinese Academy of Sciences,
19B Yuquan Road, Beijing 100049, China
University of Chinese Academy of Sciences,
19A Yuquan Road, Beijing 100049, China
E-mail: ruanmq@ihep.ac.cn

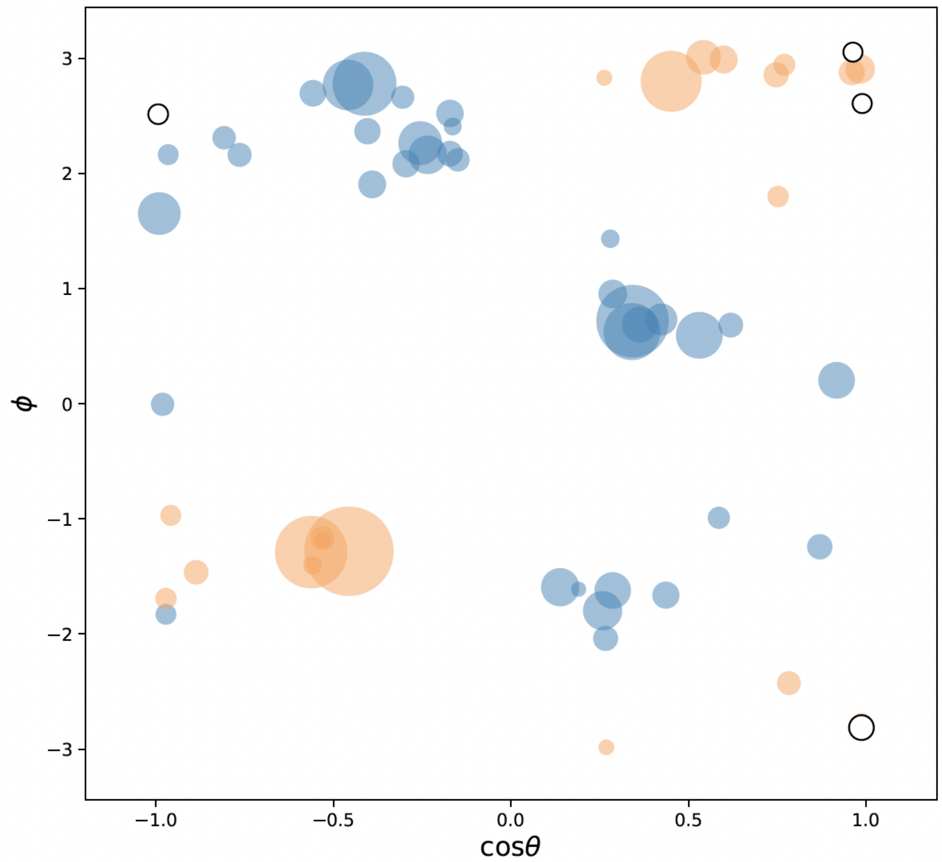
Z decay mode	$H \to b\bar{b}$	$H \to c\bar{c}$	$H \to gg$
$Z \to e^+e^-$	1.57%	14.43%	10.31%
$Z \to \mu^+\mu^-$	1.06%	10.16%	5.23%
$Z \to q\bar{q}$	0.35%	7.74%	3.96%
$Z \to \nu\bar{\nu}$	0.49%	5.75%	1.82%
combination	0.27%	4.03%	1.56%

Table 3. The signal strength accuracies for different channels.

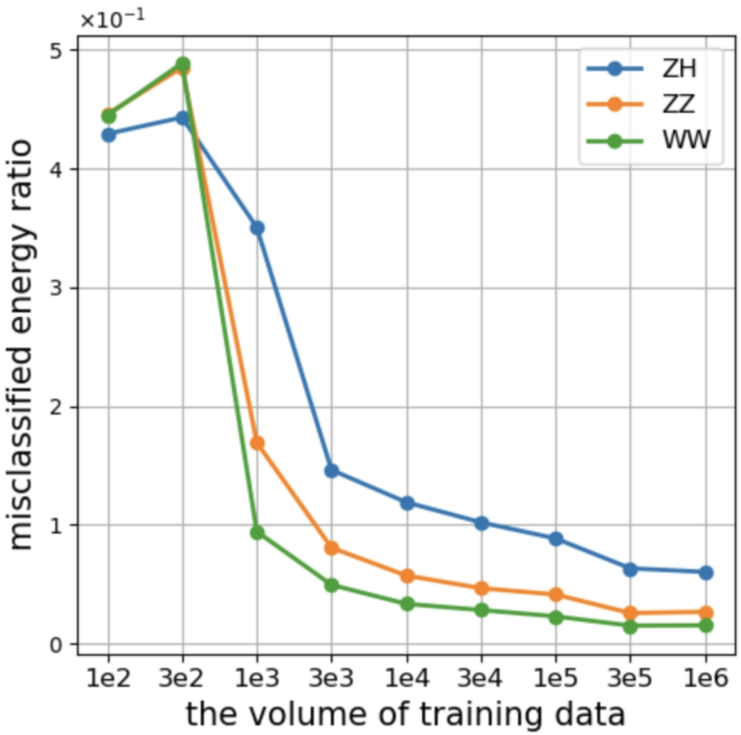
- H→cc & gg measurements at qqH channel is much worse vvH channels, despite the former has 3.5 times more signal statistic
- Reason: Failure of Color Singlet Identification – to distinguish the decay products of each Color Singlet
  - Z & H for 240/250 GeV Higgs factory
  - Which Higgs boson for Higgs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

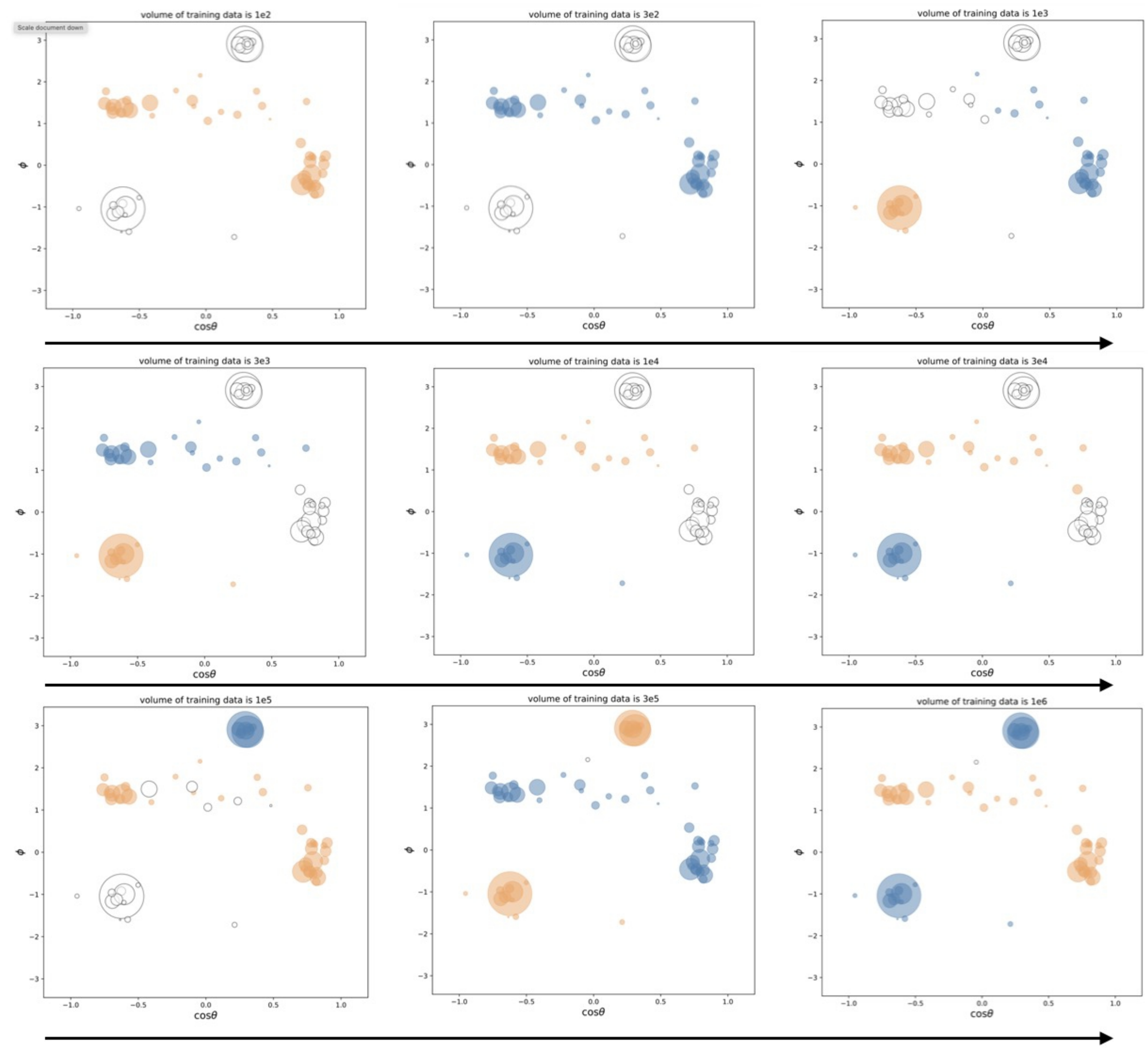
CSI: to group the final state particle





Scaling behavior



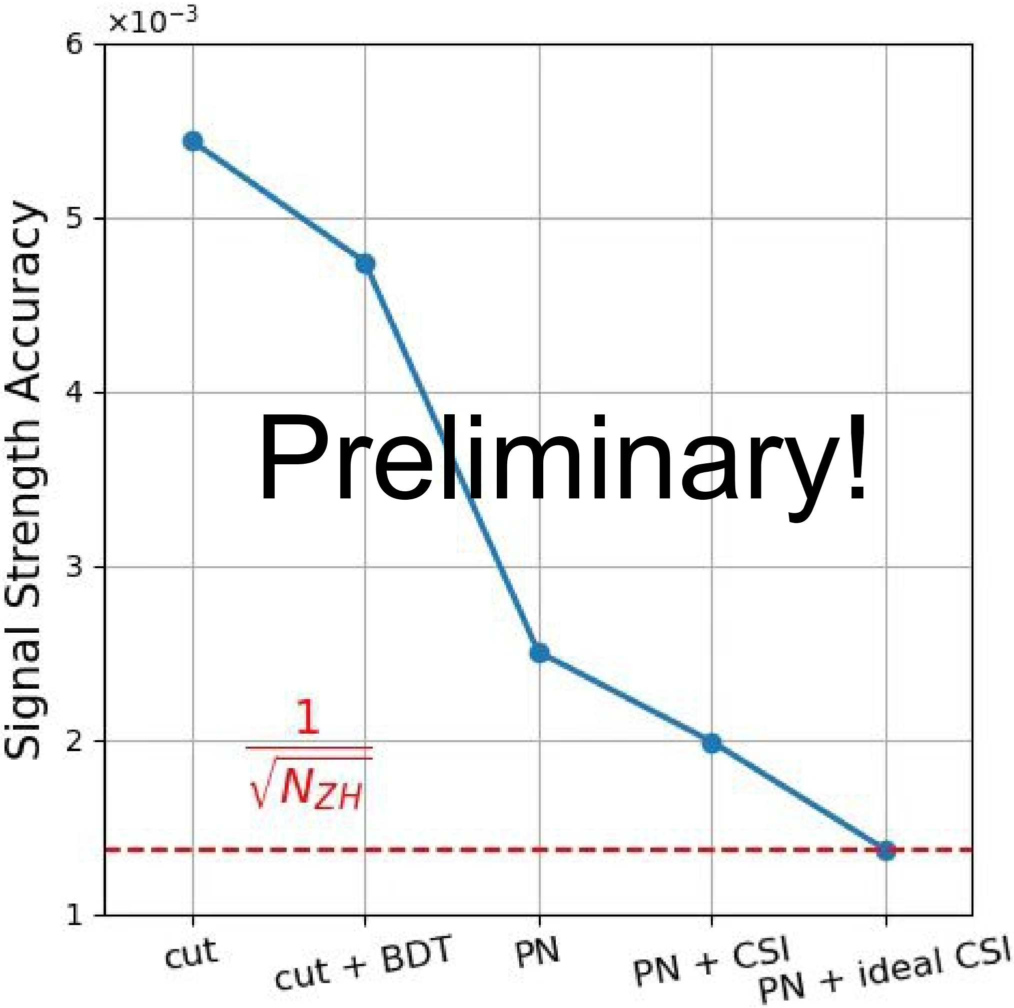


A toy analysis: identify full hadronic ZH signal from ZZ + WW background

- 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



PN ~ holistic event description via 1-1

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

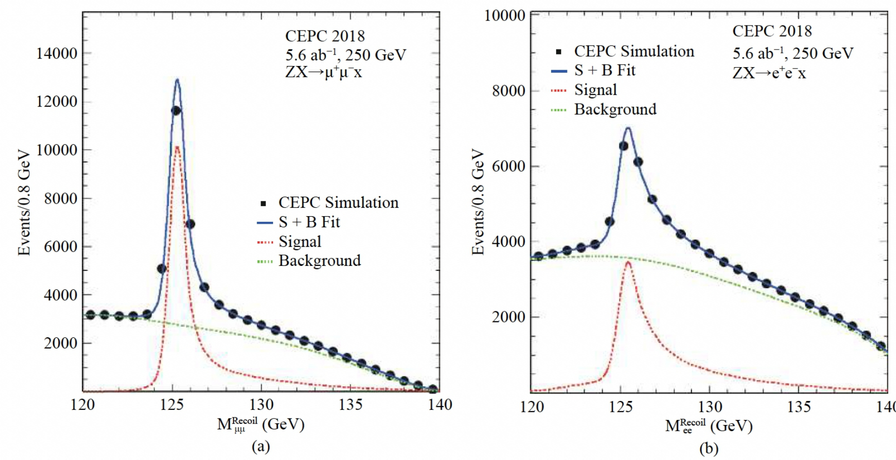
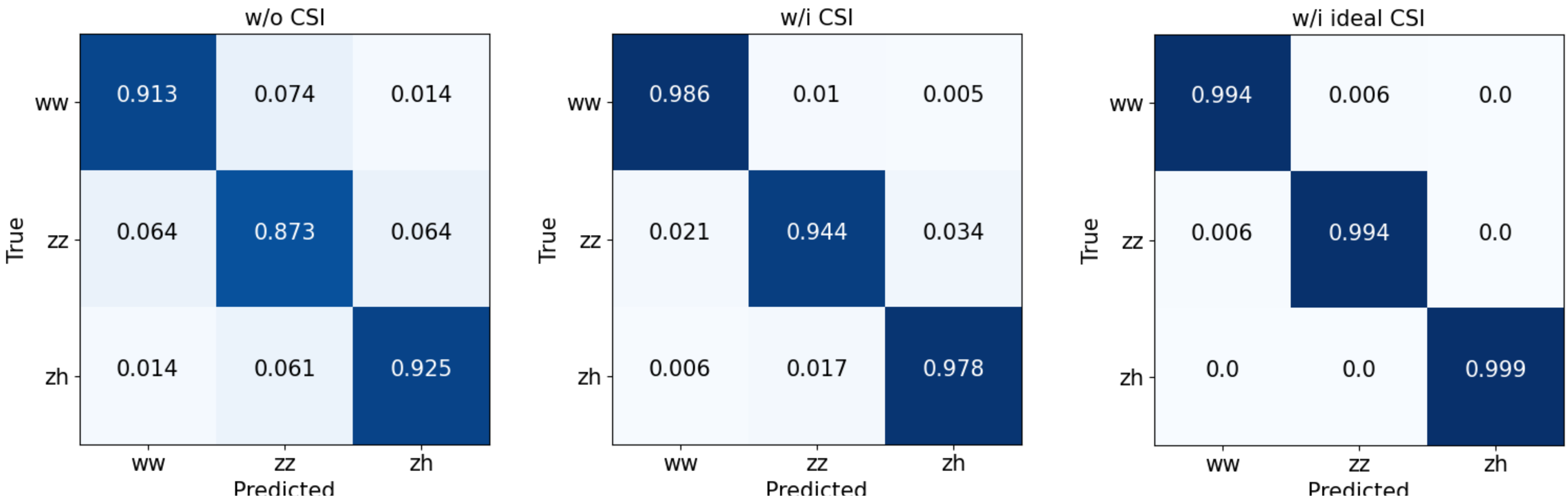


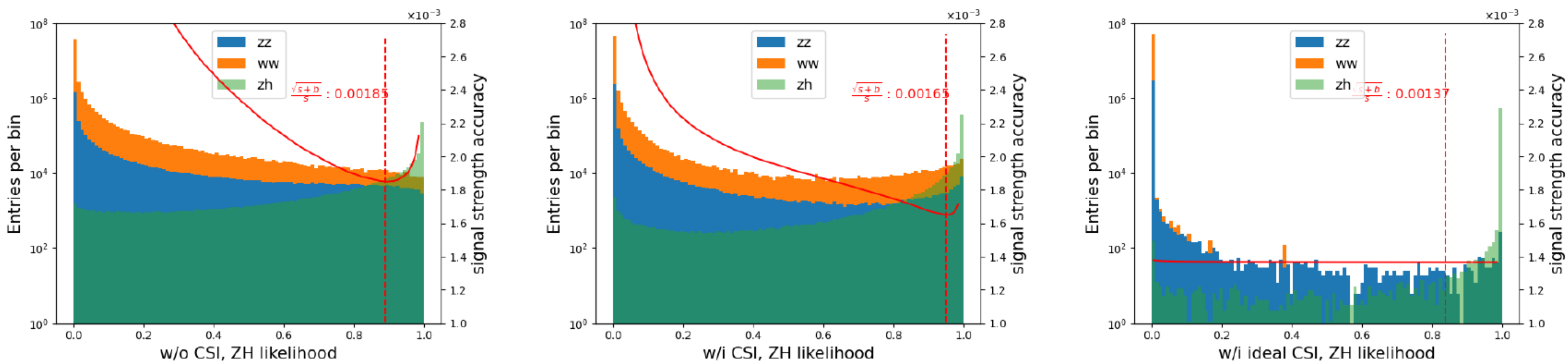
Table 5. Estimated measurement precision for the Higgs boson mass m_H and the $e^+e^- \to ZH$ production cross section $\sigma(ZH)$ from a CEPC dataset of 5.6 ab^{-1}.

Z decay mode	Δm_H/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!





Meta questions

- Problem categorization
  - Identification problem: JoI, 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

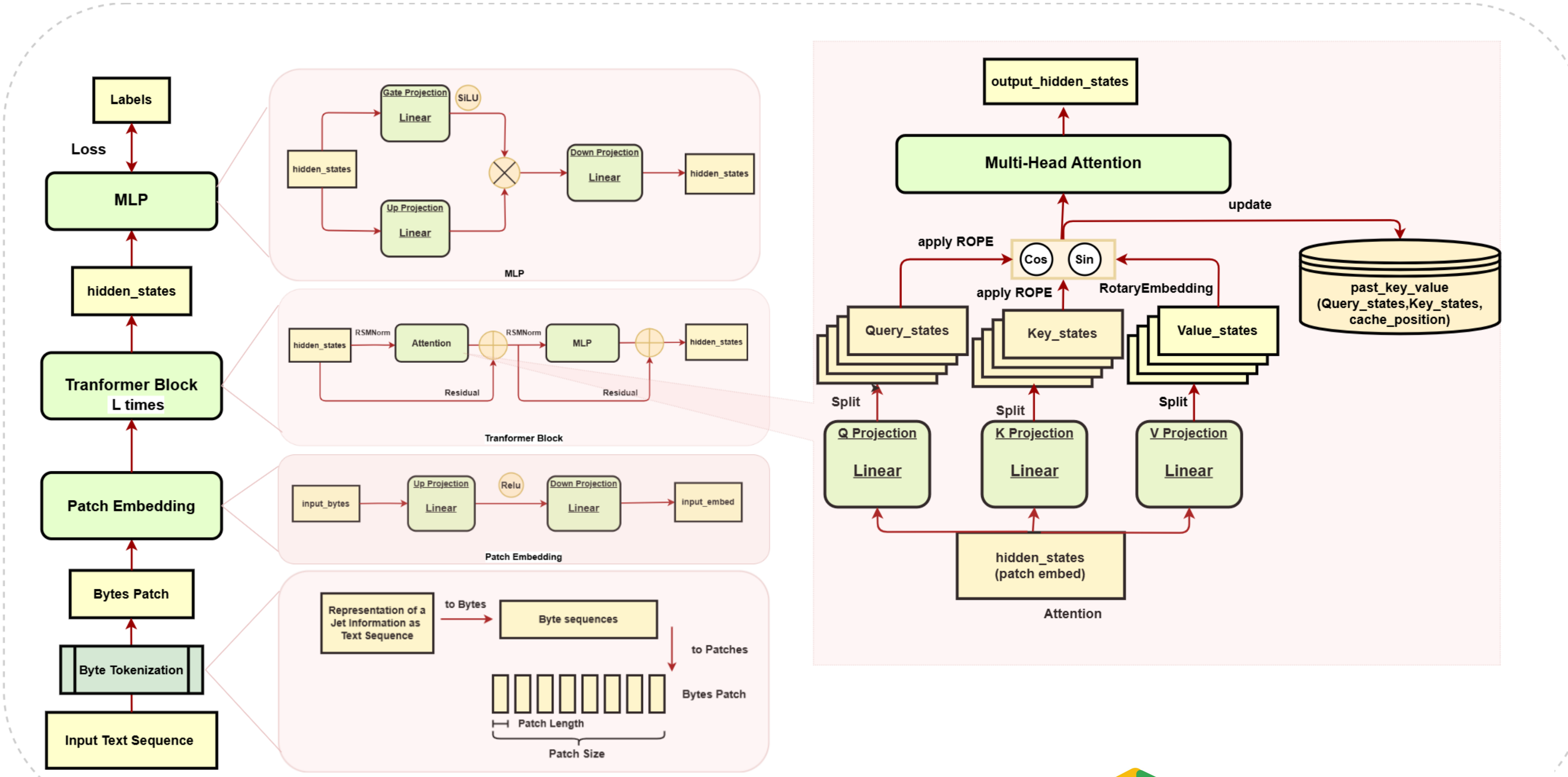
- Jet Origin id: zhuyongfeng@pku.edu.cn
- https://github.com/ZHUYFgit/CEPC-Jet-Origin-Identification/tree/main

- LLM (BINBBT):
- https://github.com/supersymmetry-technologies/bbt-neutron

- 1-1 Correspondence: wangyuexin@ihep.ac.cn
- Source deposition in construction...

Back up

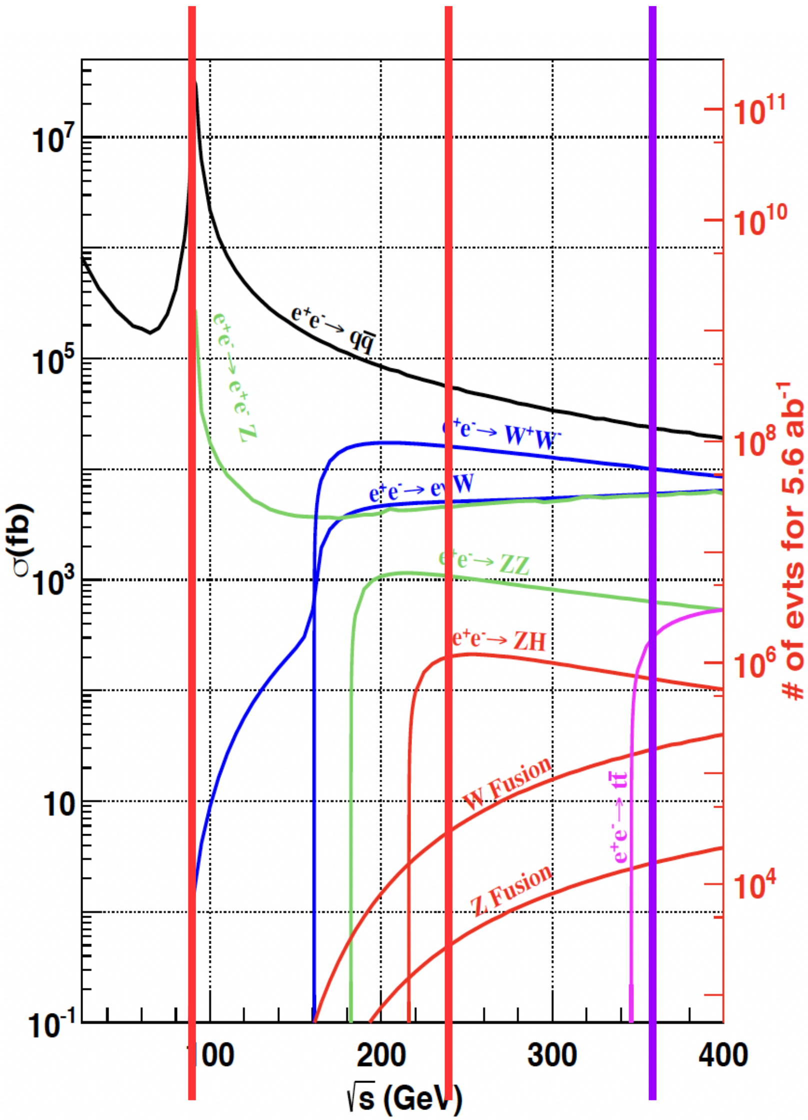
Recent update: from specialized Models to LLM

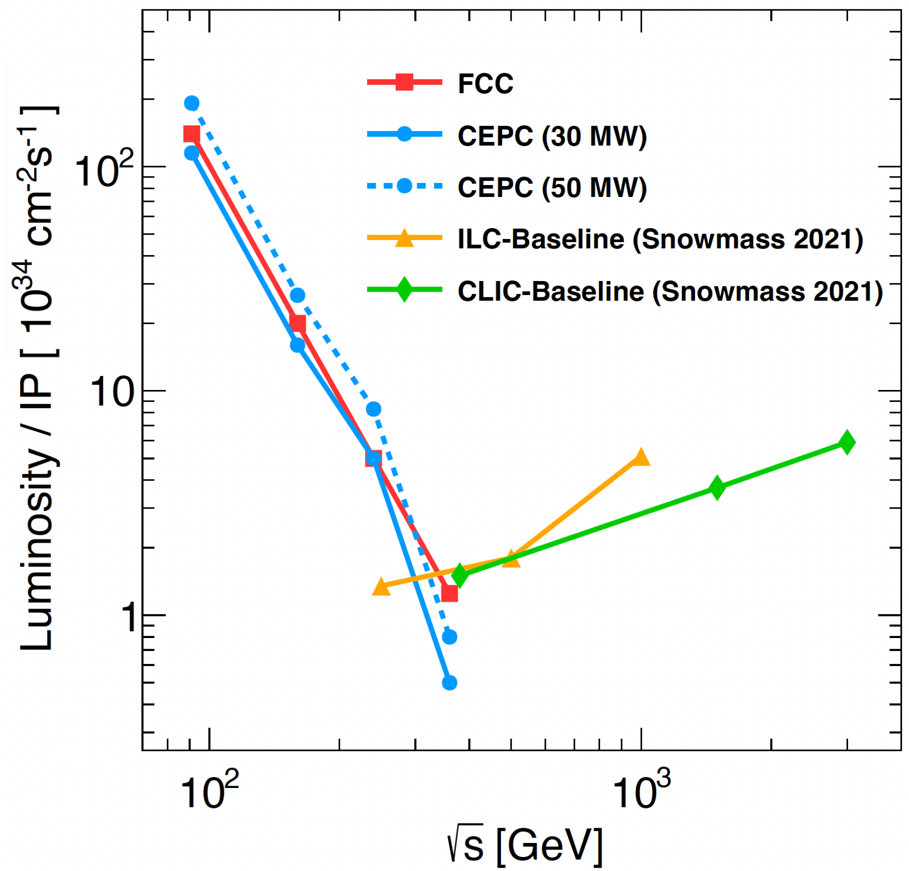


BIN-BBT Model

- New tokenization method to address numeric problems at LLM

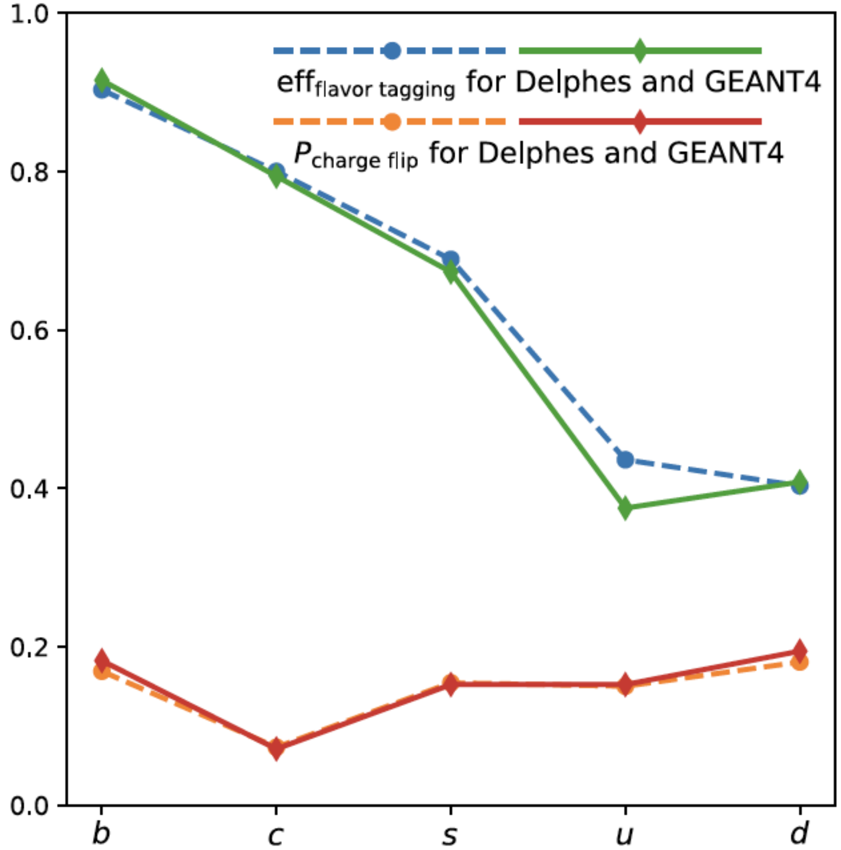
Yields ~ Xsec * Lumi * Time



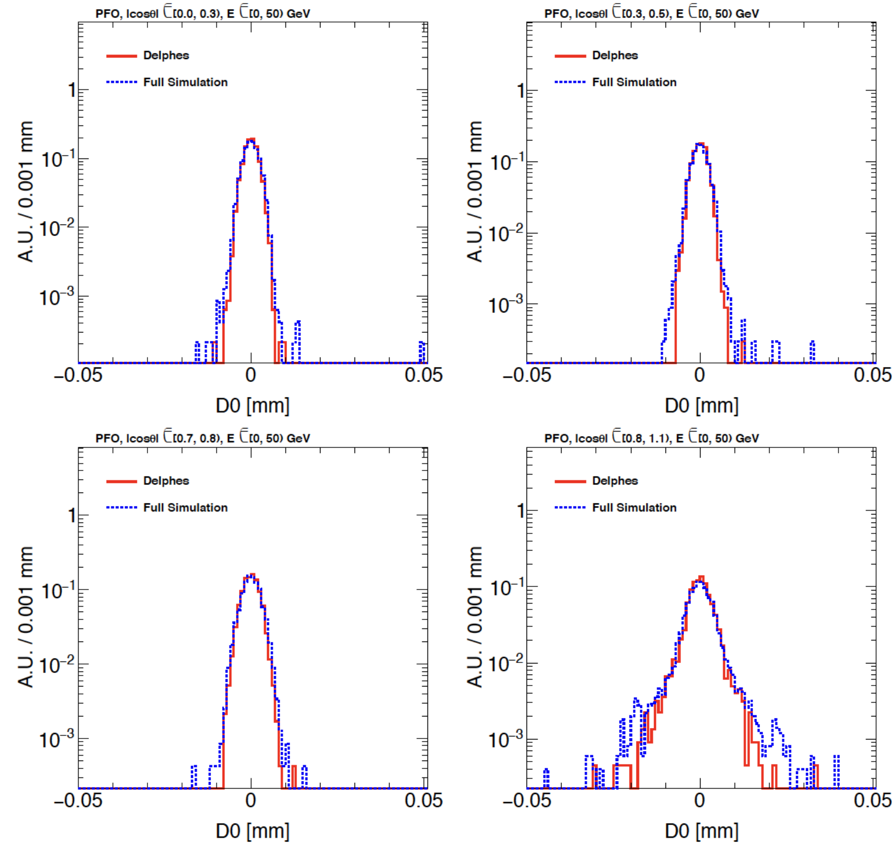


- 4 Million Higgs (10 years)
- ~ 1 Giga W (1 year) + 4 Tera Z (2 years)
- Upgradable: Top factory (500 k ttbar)

Fast/Full Simulation

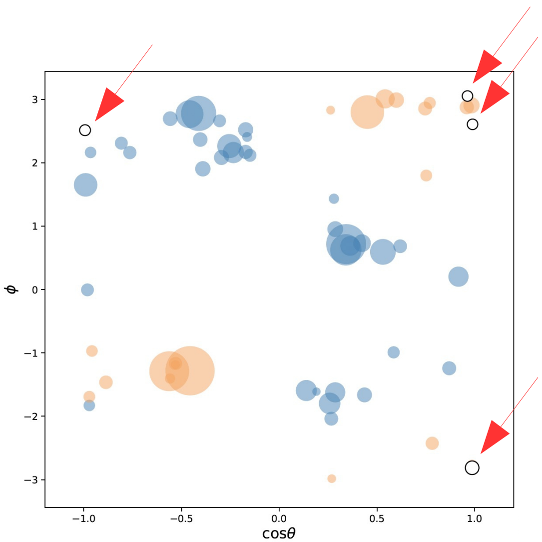


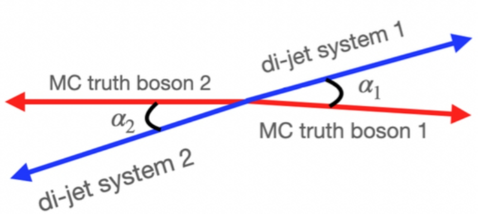
Z->μμ (91.2 GeV)

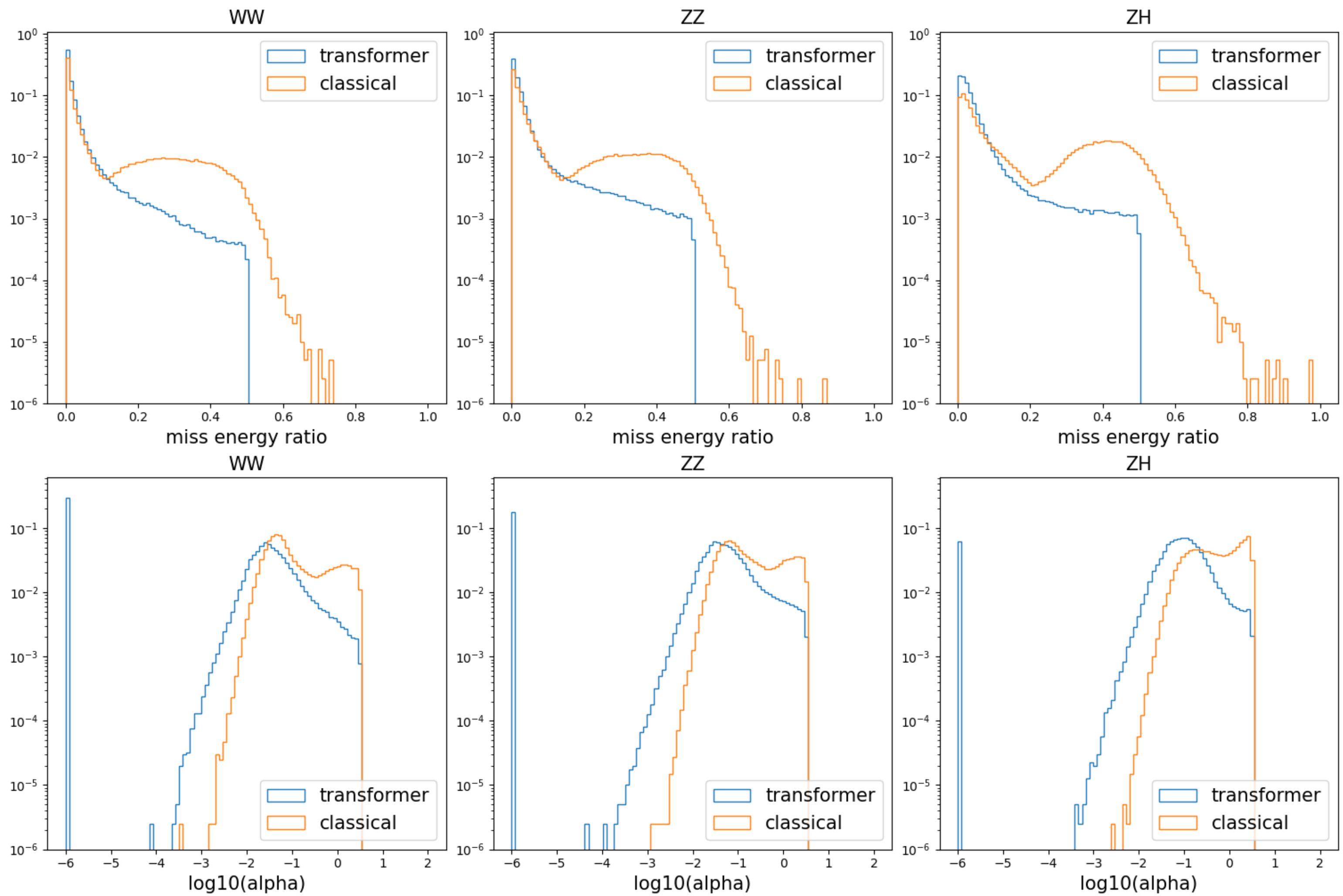


- Delphes ~ Perfect PFA (1 – 1 correspondence..)

CSI: classical VS AI (Transformer)

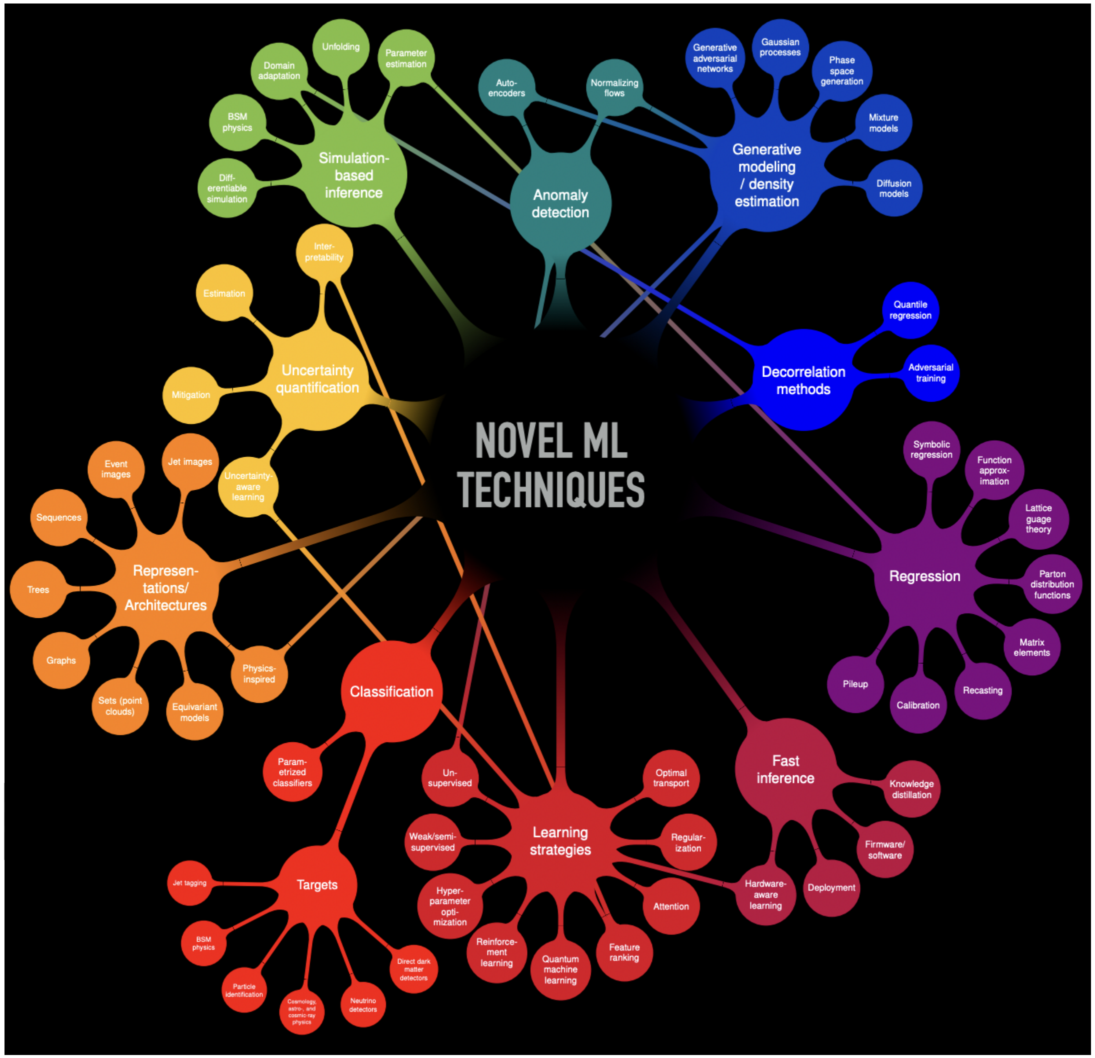






Classical: Jet Clustering + Matching with min(Chi-2)

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: (o(1E5) Hits → 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 | Adva

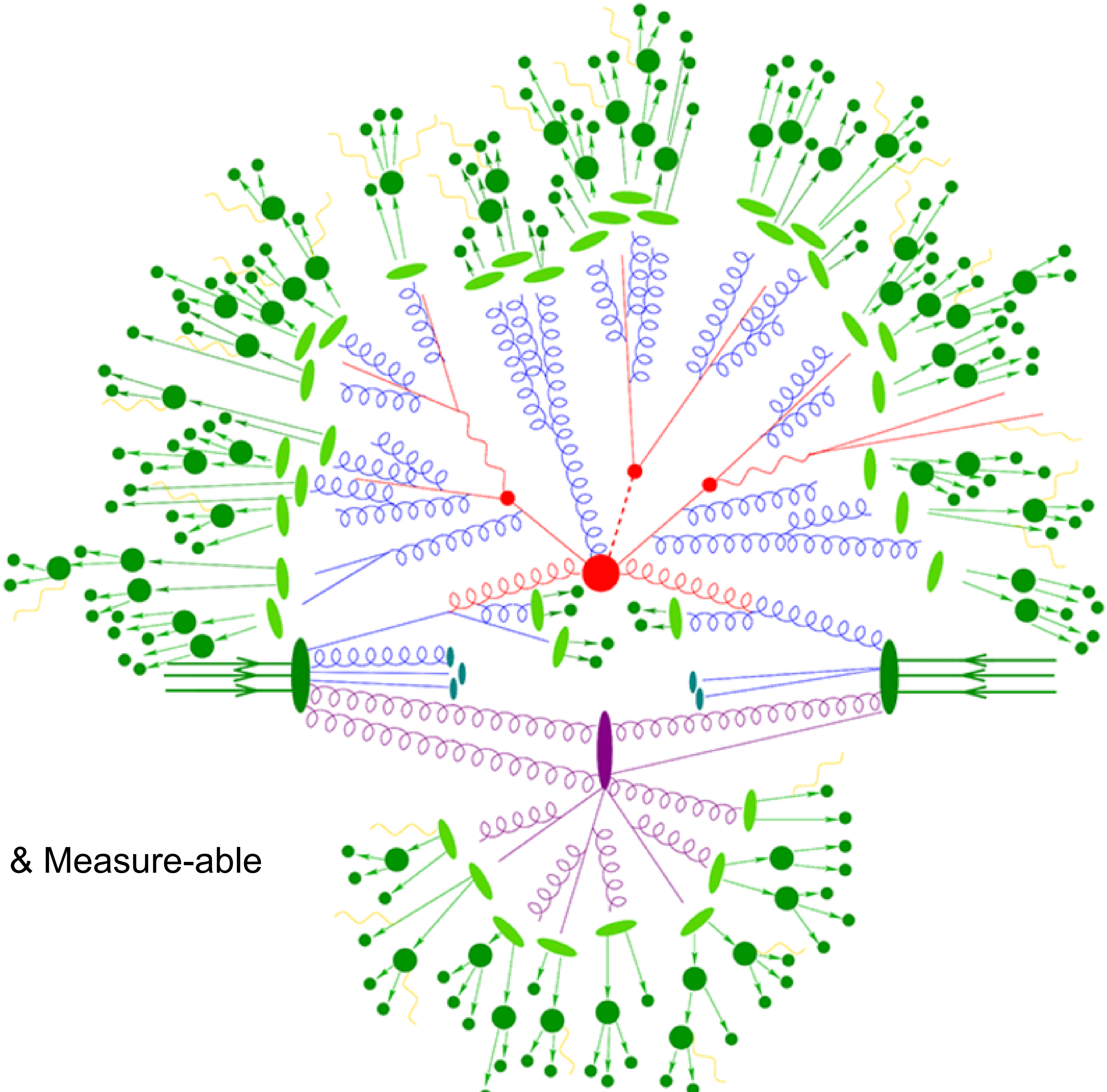
High Energy Physics – Experiment

[Submitted on 11 Nov 2024]

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

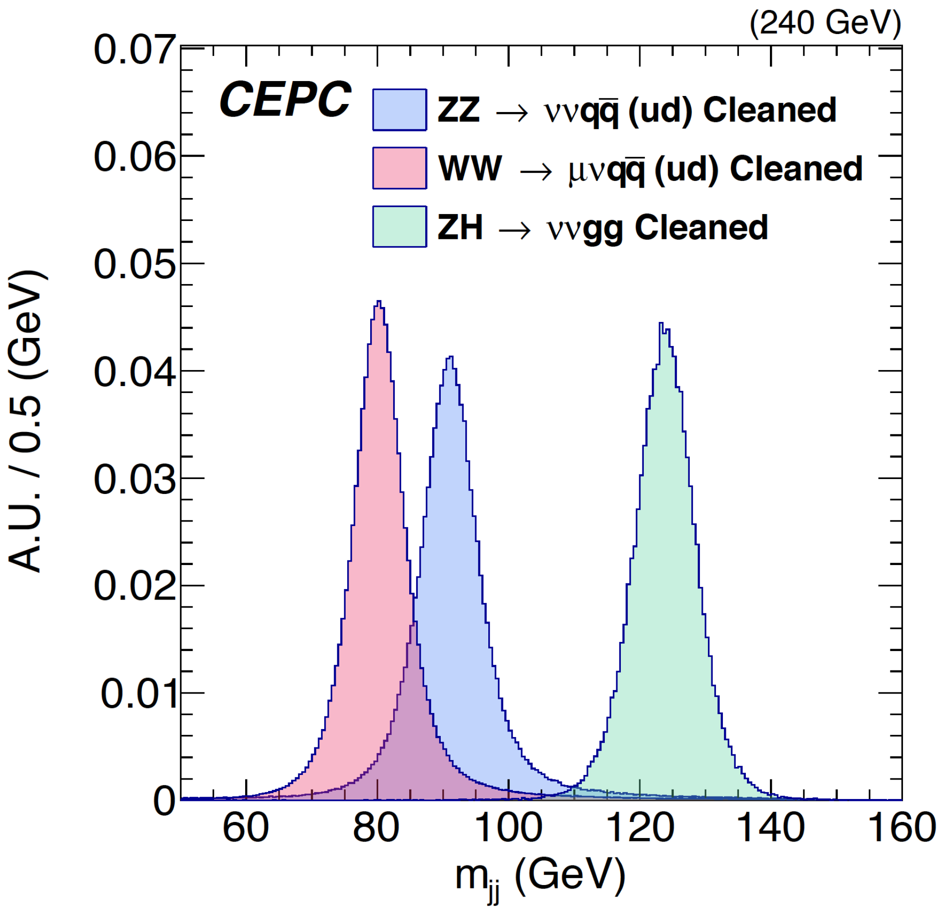
Yuexin Wang, Hao Liang, Yongfeng Zhu, Yuzhi Che, Xin Xia, Huilin Qu, Chen Zhou, Xuai Zhuang, Manqi Ruan

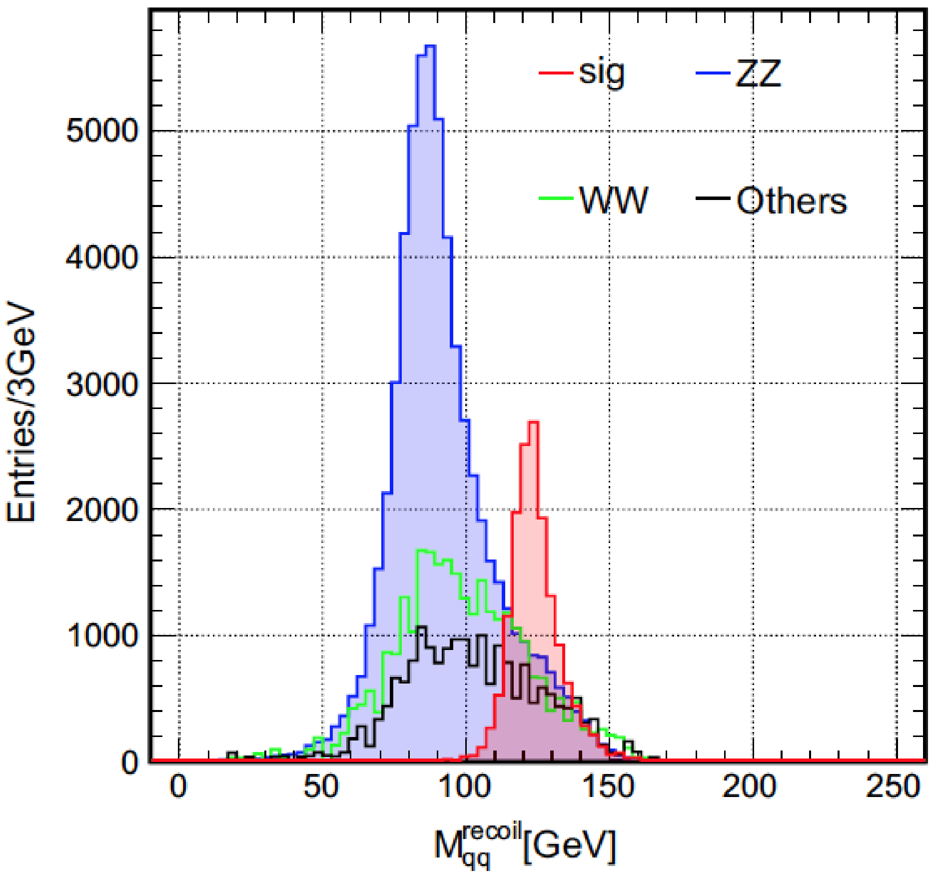
We propose one-to-one correspondence reconstruction for electron-positron Higgs factories. For each visible particle, one-to-one correspondence aims to associate relevant detector hits with only one reconstructed particle and accurately identify its species. To achieve this goal, we develop a novel detector 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 (γ), 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



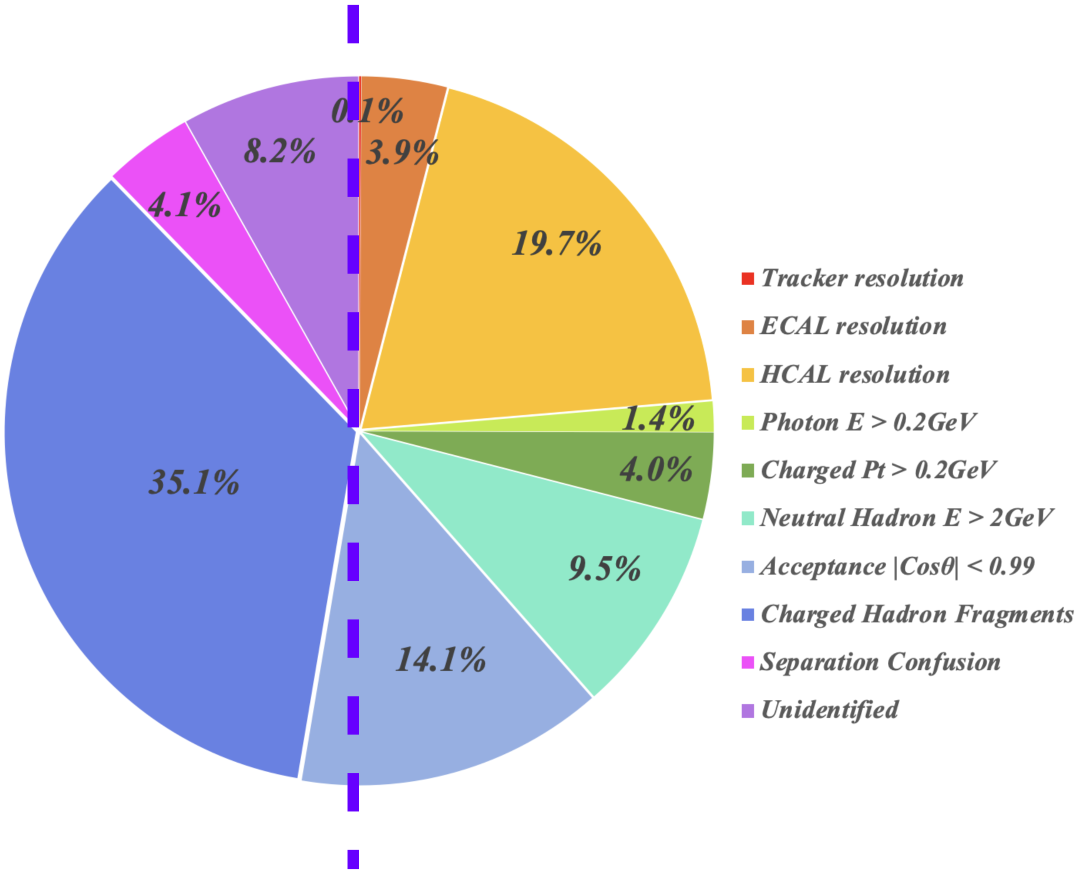


Higgs factory: need BMR < 4% (critical for qqH & qqZ separation using recoil mass to di-jet)

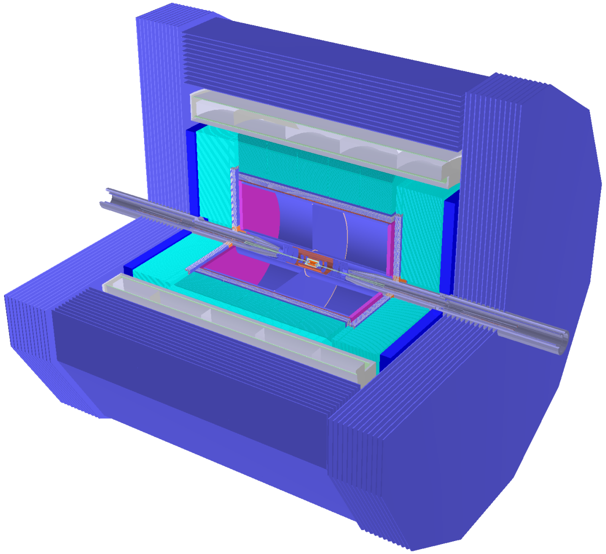
Strongly motivated to improve BMR to 3% or even lower, especially for NP & Flavor

CDR baseline (left plot): BMR = 3.75%

BMR decomposition @ CDR baseline

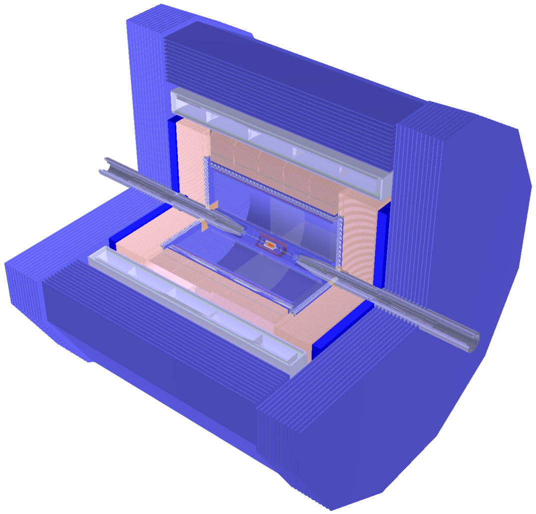


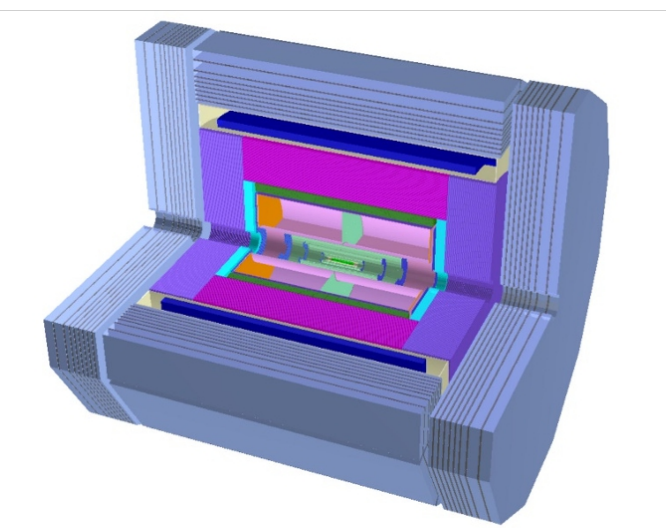
- *1st HCAL resolution dominant the uncertainties from intrinsic detector resolution: need better HCAL → usage of GSHCAL*
- *2nd Leading contribution: Confusion from shower Fragments (fake particles), need better Pattern Reco.*

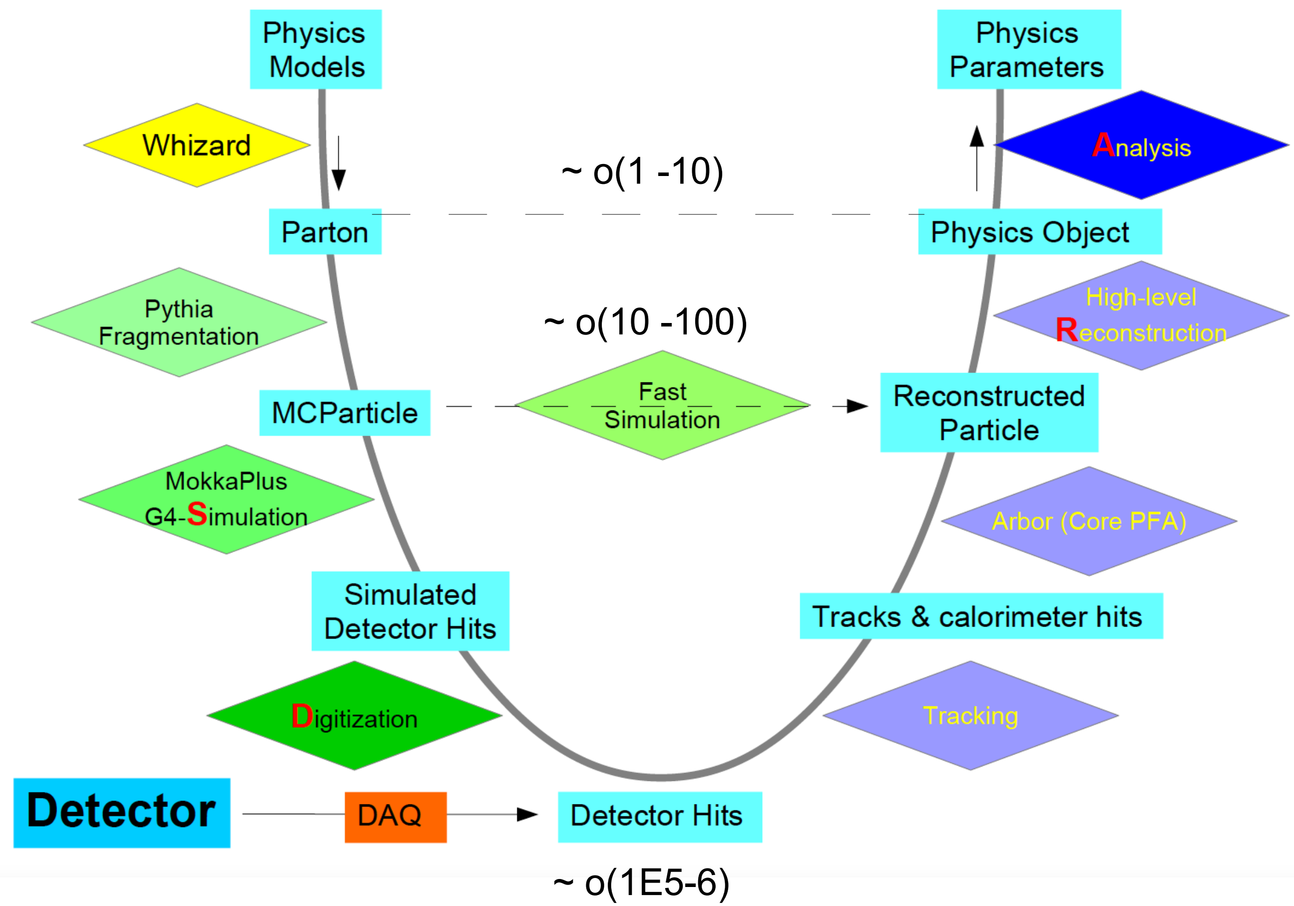


- *CDR baseline - GRPC HCAL*

CEPC Detector & Reconstruction



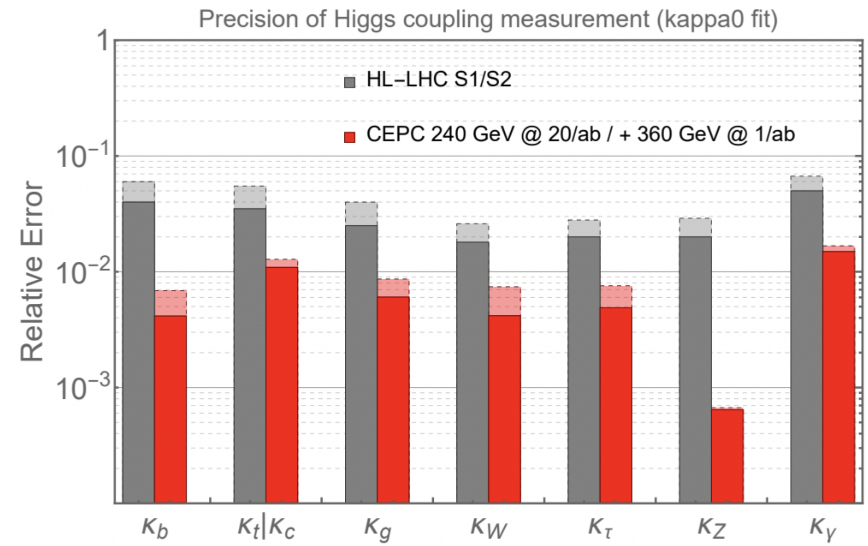


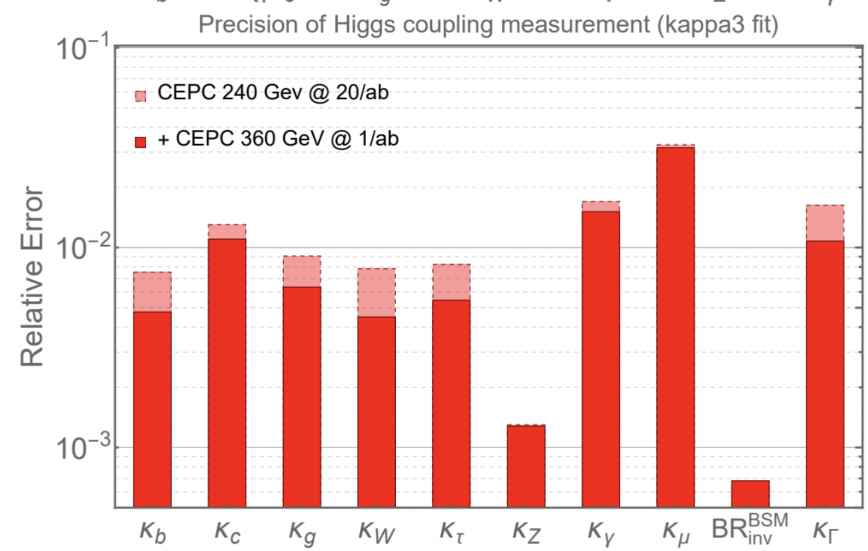


Full simulation reconstruction Chain with **Arbor, etc**

Higgs & Snowmass White Paper

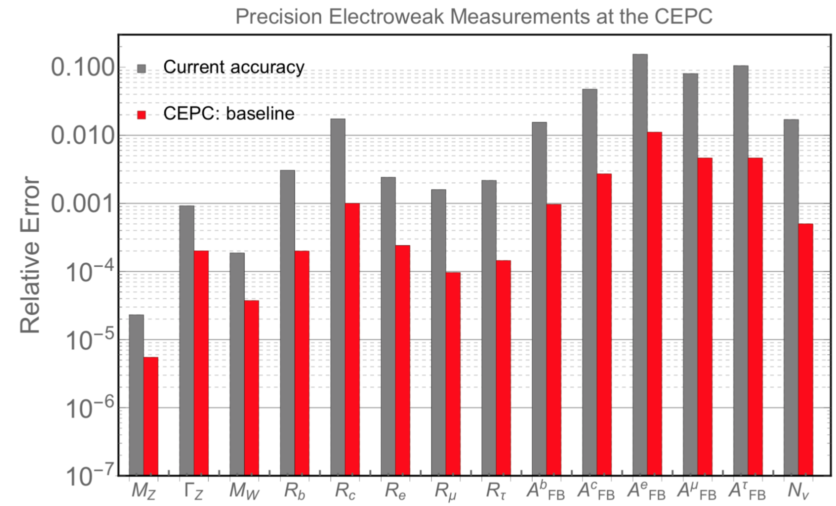
	240 GeV, 20 ab^{-1}		360 GeV, 1 ab^{-1}		
	ZH	**vvH**	ZH	**vvH**	**eeH**
inclusive	**0.26%**		**1.40%**	\	\
H→bb	**0.14%**	**1.59%**	**0.90%**	**1.10%**	**4.30%**
H→cc	**2.02%**		**8.80%**	**16%**	**20%**
H→gg	**0.81%**		**3.40%**	**4.50%**	**12%**
H→WW	**0.53%**		**2.80%**	**4.40%**	**6.50%**
H→ZZ	**4.17%**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

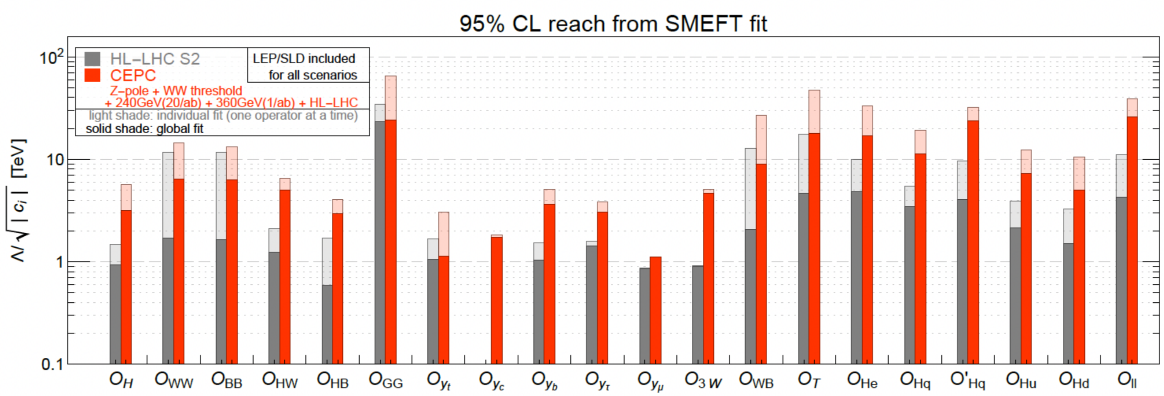


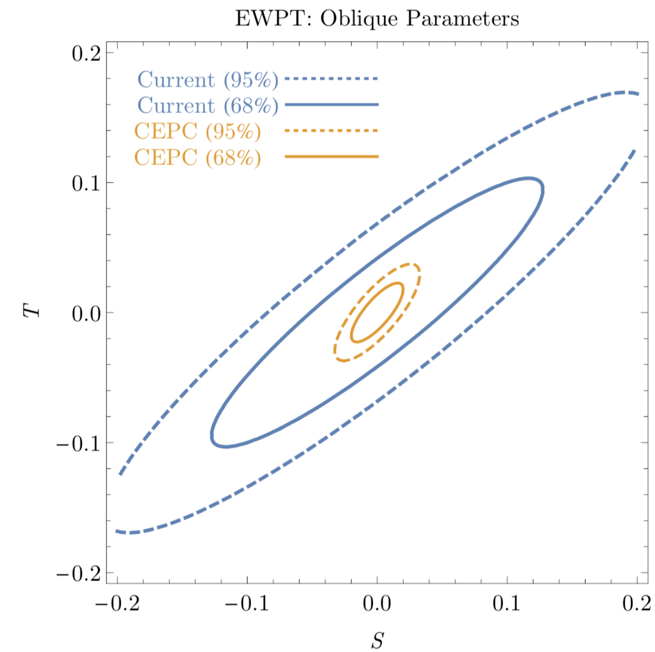


EW measurements & SMEFT

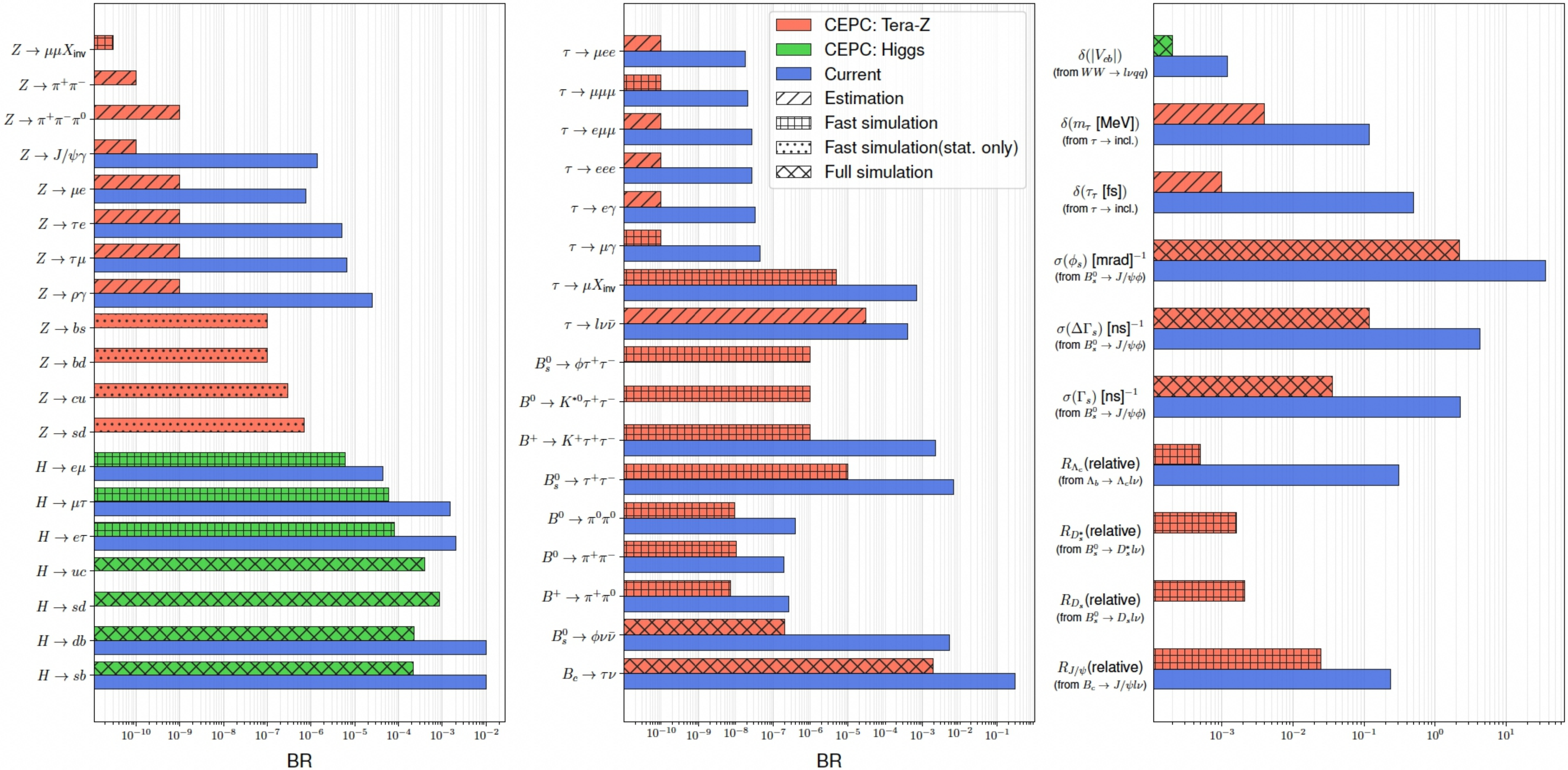
Observable	current precision	CEPC precision (Stat. Unc.)	CEPC runs	main systematic
Δm_Z	2.1 MeV [37–41]	0.1 MeV (0.005 MeV)	Z threshold	E_{beam}
$\Delta \Gamma_Z$	2.3 MeV [37–41]	0.025 MeV (0.005 MeV)	Z threshold	E_{beam}
Δm_W	9 MeV [42–46]	0.5 MeV (0.35 MeV)	WW threshold	E_{beam}
$\Delta \Gamma_W$	49 MeV [46–49]	2.0 MeV (1.8 MeV)	WW threshold	E_{beam}
Δm_t	0.76 GeV [50]	$\mathcal{O}(10)$ MeV[a]	$t\bar{t}$ threshold	
ΔA_e	4.9×10^{-3} [37, 51–55]	1.5×10^{-5} (1.5×10^{-5})	Z pole ($Z \to \tau\tau$)	Stat. Unc.
ΔA_μ	0.015 [37, 53]	3.5×10^{-5} (3.0×10^{-5})	Z pole ($Z \to \mu\mu$)	point-to-point Unc.
ΔA_τ	4.3×10^{-3} [37, 51–55]	7.0×10^{-5} (1.2×10^{-5})	Z pole ($Z \to \tau\tau$)	tau decay model
ΔA_b	0.02 [37, 56]	20×10^{-5} (3×10^{-5})	Z pole	QCD effects
ΔA_c	0.027 [37, 56]	30×10^{-5} (6×10^{-5})	Z pole	QCD effects
$\Delta \sigma_{had}$	37 pb [37–41]	2 pb (0.05 pb)	Z pole	lumiosity
δR_b^0	0.003 [37, 57–61]	0.0002 (5×10^{-6})	Z pole	gluon splitting
δR_c^0	0.017 [37, 57, 62–65]	0.001 (2×10^{-5})	Z pole	gluon splitting
δR_e^0	0.0012 [37–41]	2×10^{-4} (3×10^{-6})	Z pole	E_{beam} and t channel
δR_μ^0	0.002 [37–41]	1×10^{-4} (3×10^{-6})	Z pole	E_{beam}
δR_τ^0	0.017 [37–41]	1×10^{-4} (3×10^{-6})	Z pole	E_{beam}
δN_ν	0.0025 [37, 66]	2×10^{-4} (3×10^{-5})	ZH run ($\nu\nu\gamma$)	Calo energy scale







Flavor Physics



See the non-seen: i.e, Bc→tauv, Bs→Phivv
Orders of magnitudes improvements (1 – 2.5 orders...).
Access New Physics with energy scale of 10 TeV, or even above

New Physics white paper

2023

2024

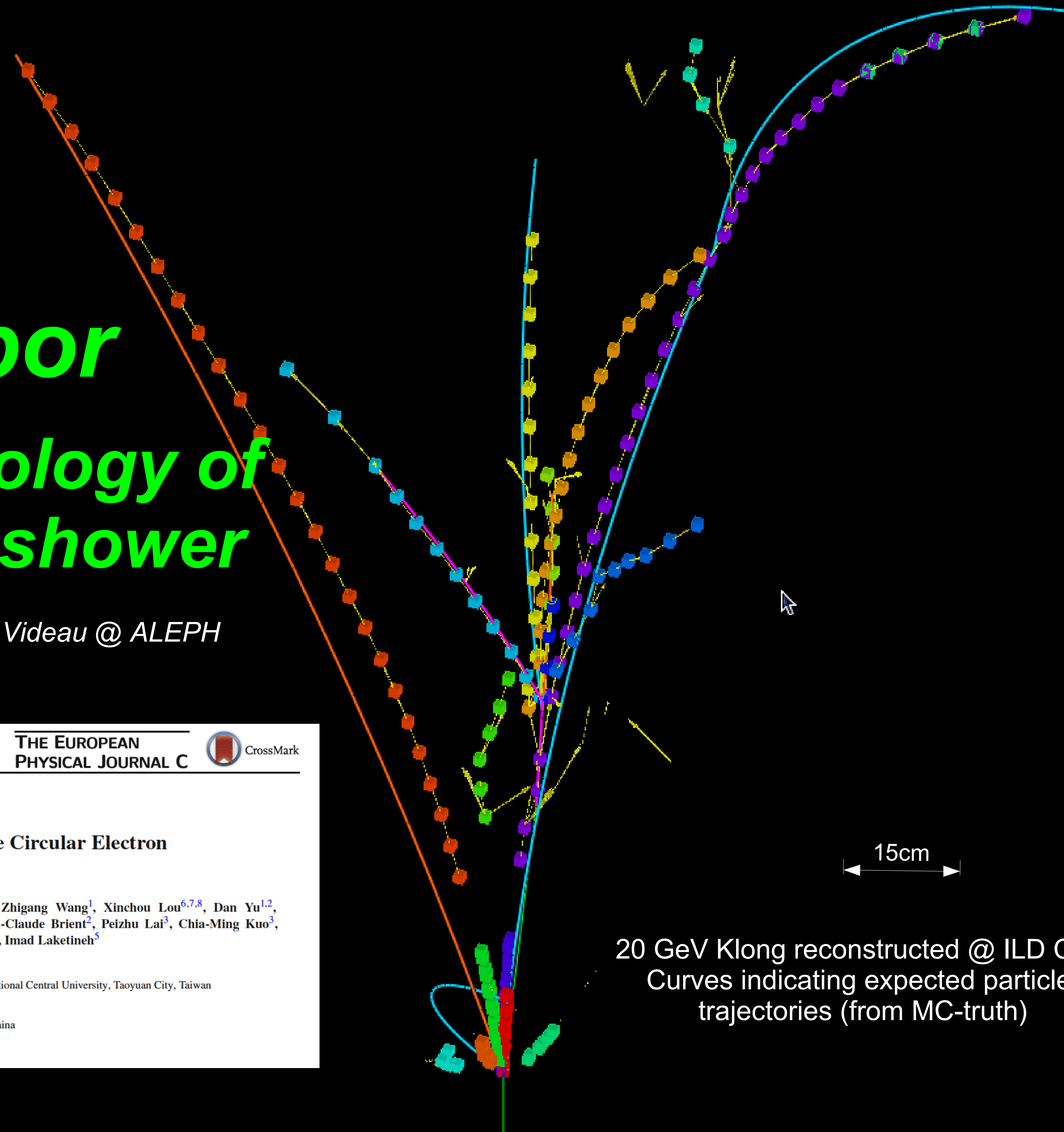
Contents extends from 40 pages → 200 pages...

- Credit: hanhua Cui, Yu Gao, Xuai Zhuang

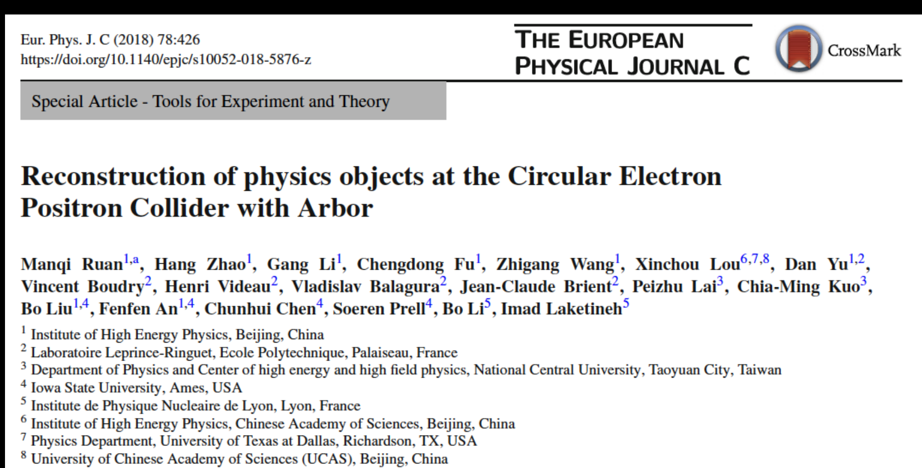
Arbor

Tree topology of particle shower

Ori. Idea from Henri Videau @ ALEPH



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



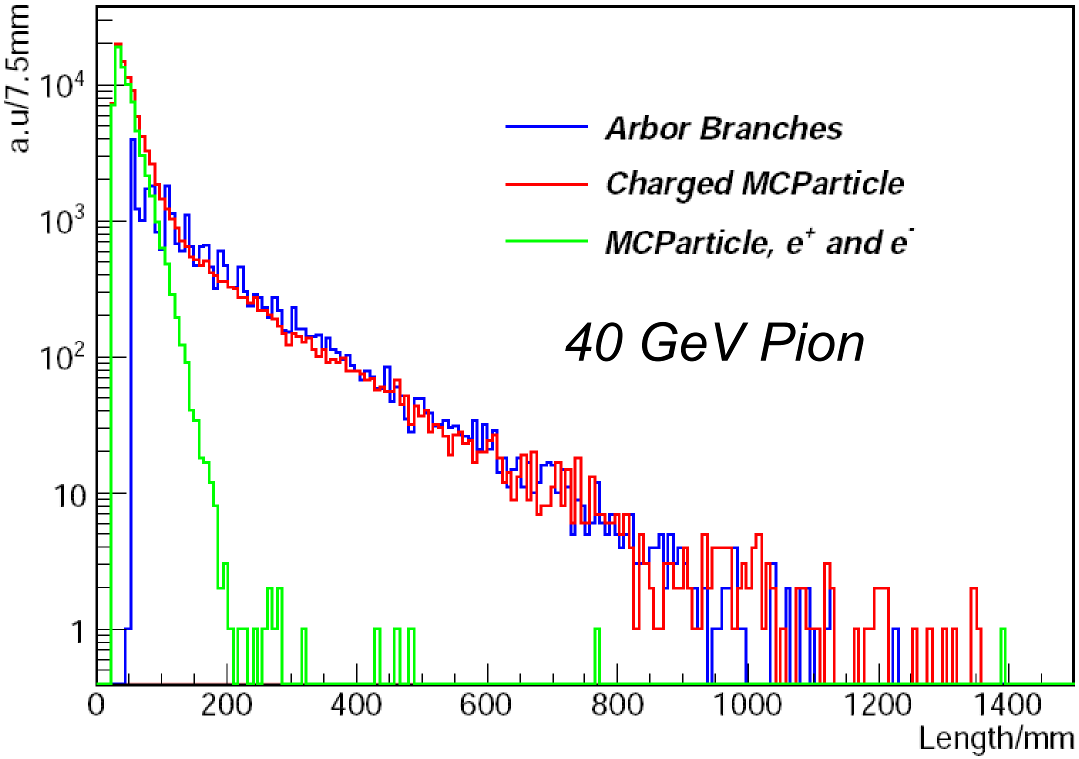
Eur. Phys. J. C (2018) 78:426
https://doi.org/10.1140/epjc/s10052-018-5876-z
THE EUROPEAN PHYSICAL JOURNAL C
CrossMark
Special Article - Tools for Experiment and Theory

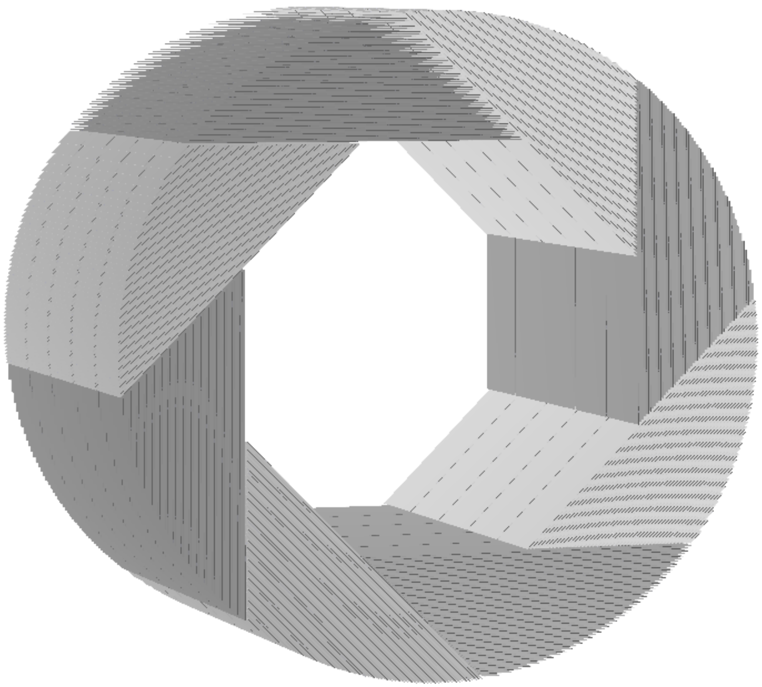
Reconstruction of physics objects at the Circular Electron Positron Collider with Arbor

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

1 Institute of High Energy Physics, Beijing, China
2 Laboratoire Leprince-Ringuet, Ecole Polytechnique, Palaiseau, France
3 Department of Physics and Center of high energy and high field physics, National Central University, Taoyuan City, Taiwan
4 Iowa State University, Ames, USA
5 Institute de Physique Nucleaire de Lyon, Lyon, France
6 Institute of High Energy Physics, Chinese Academy of Sciences, Beijing, China
7 Physics Department, University of Texas at Dallas, Richardson, TX, USA
8 University of Chinese Academy of Sciences (UCAS), Beijing, China

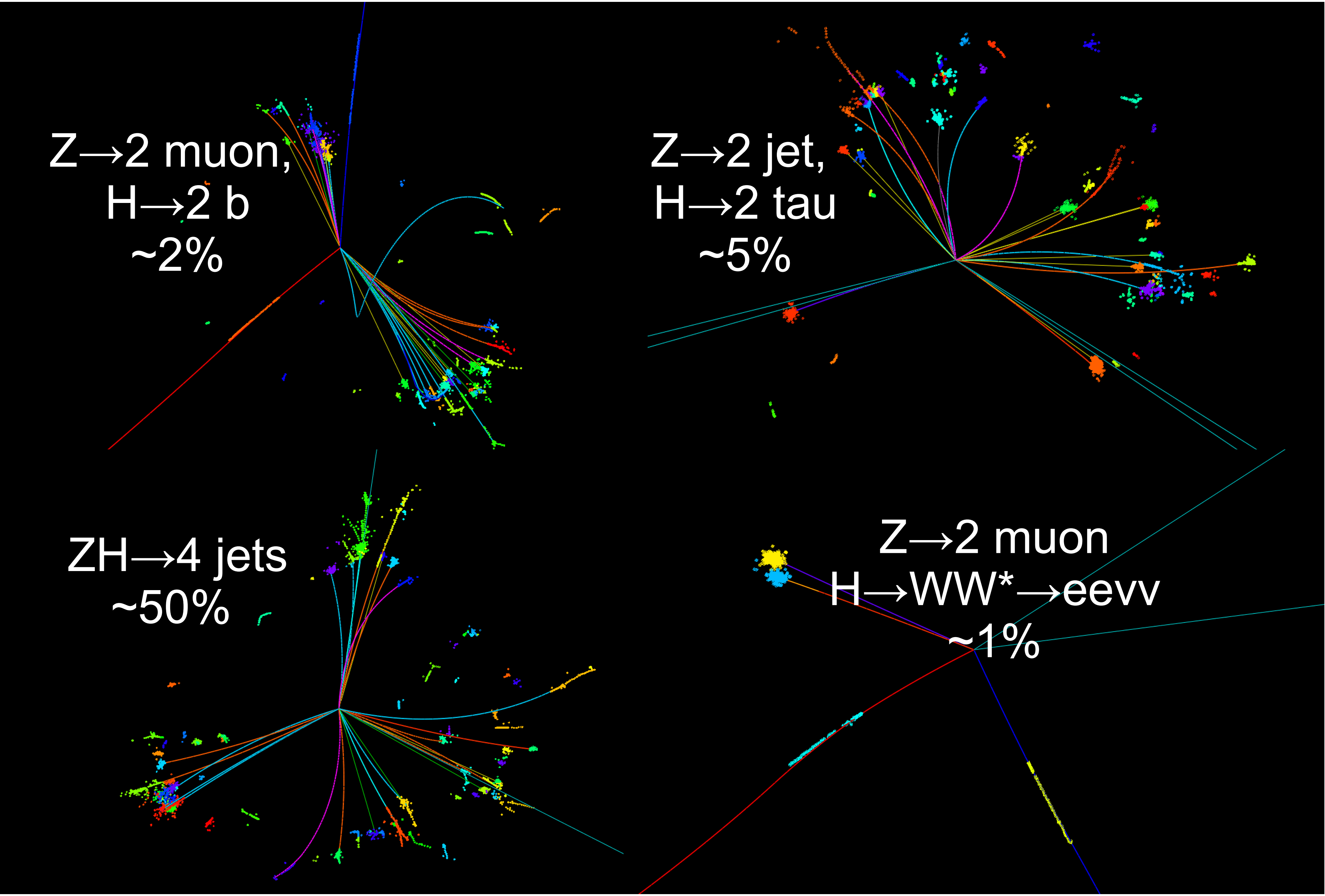
Validation: Arbor Branch Length Vs MC Truth



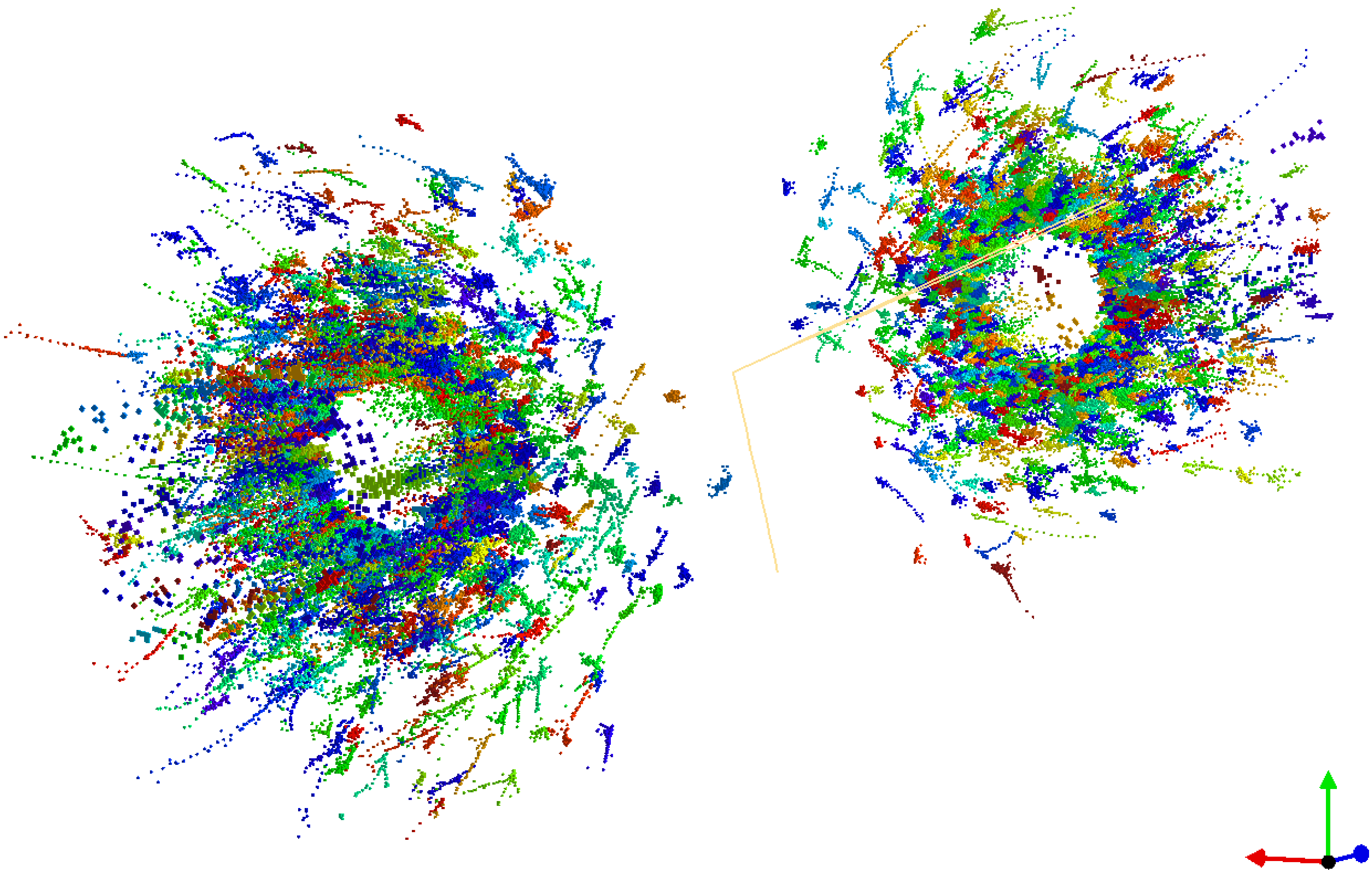


Arbor: successfully tag sub-shower structure

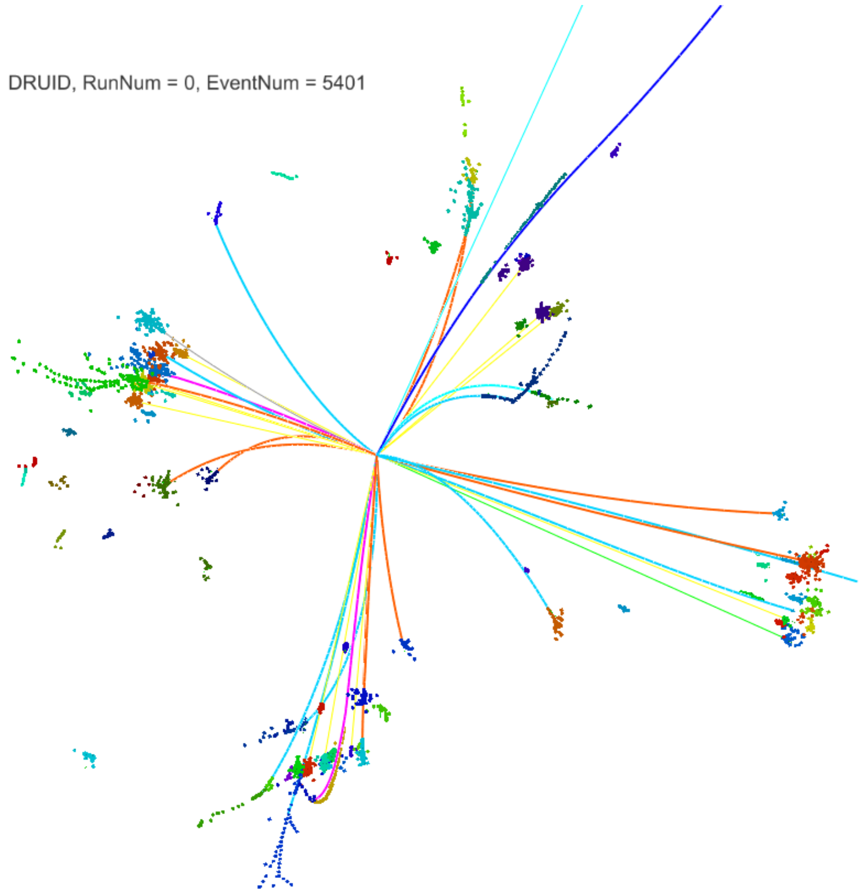
Samples: Particle gun event at ILD HCAL (readout granularity 1cm^2 & layer thickness 2.65cm)
Length:
Charged MCParticle: spatial distance between generation/end points
Arbor branch: sum of distance between neighboring cells

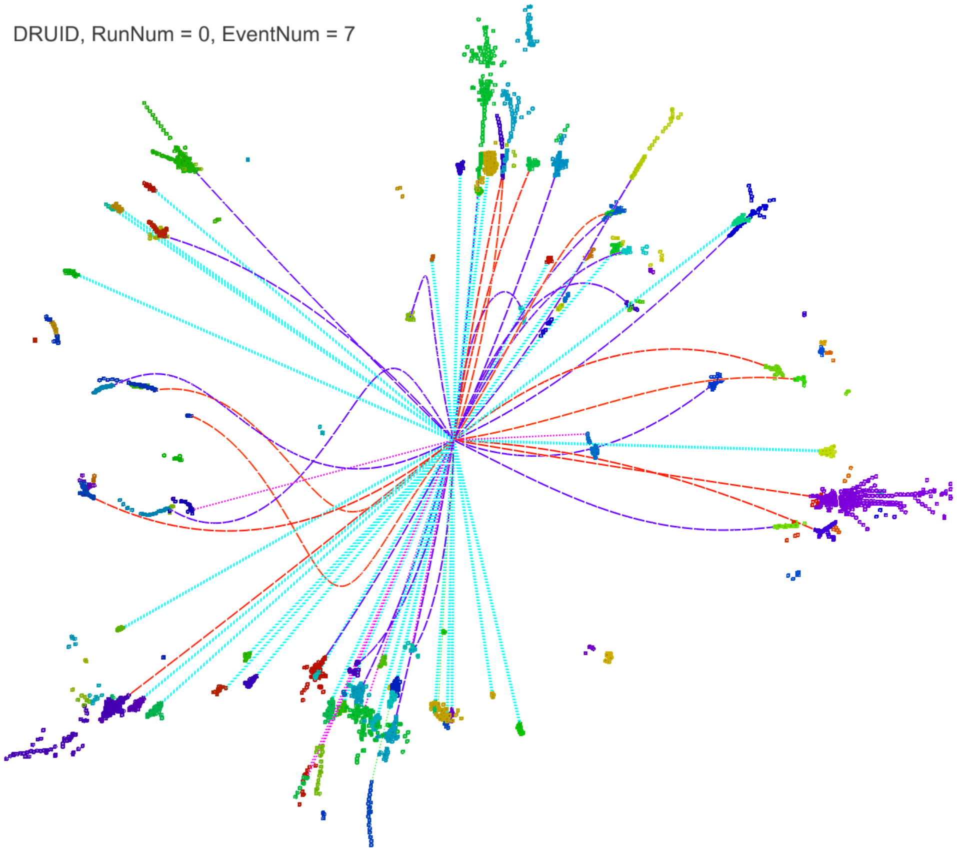

Z→2 muon,
H→2 b
~2%
Z→2 jet,
H→2 tau
~5%
ZH→4 jets
~50%
Z→2 muon
H→WW*→eevv
~1%

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

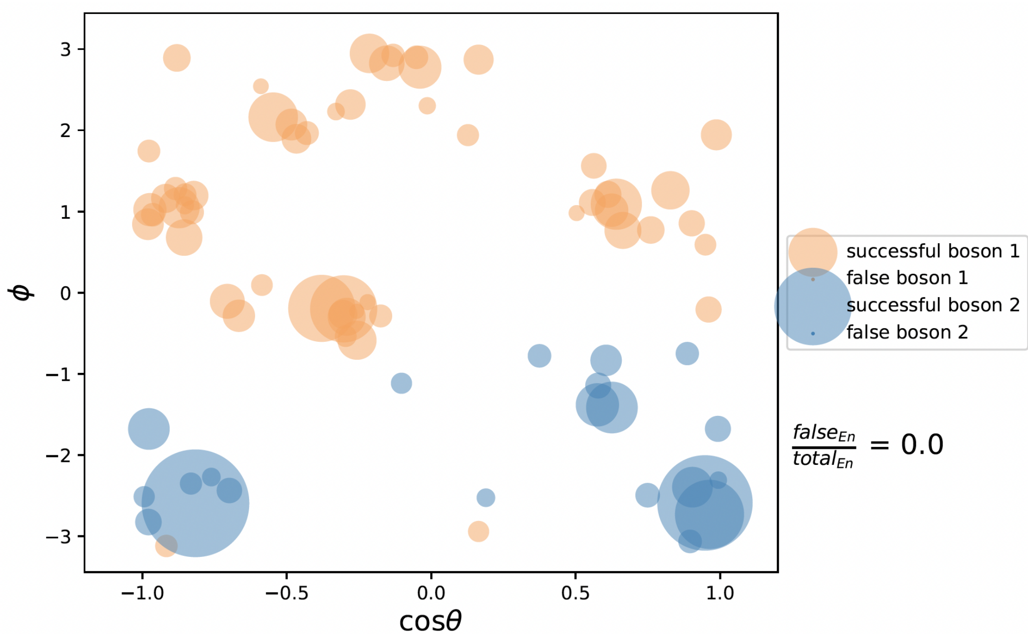


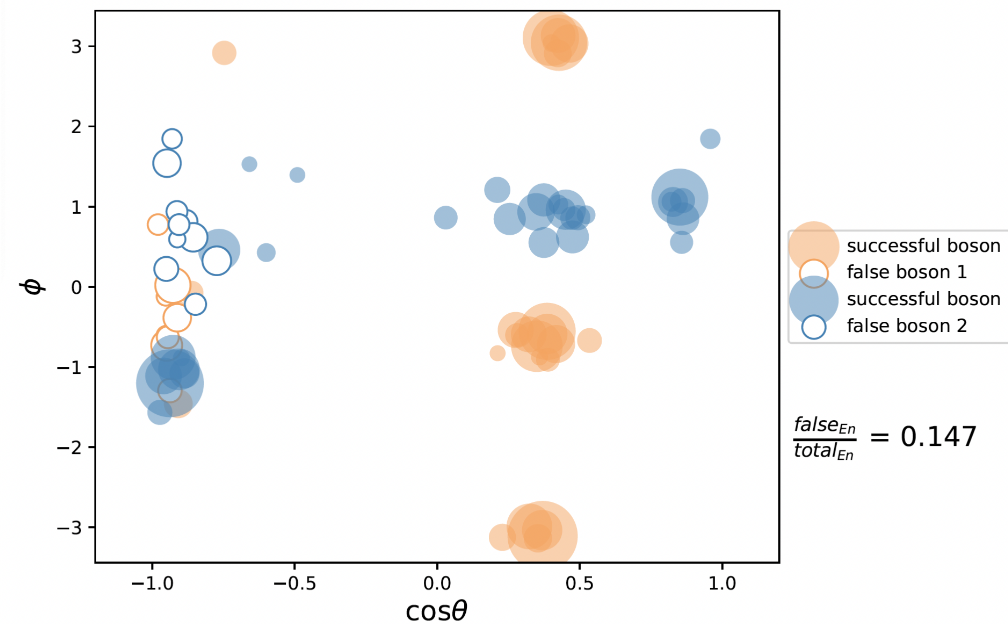
Color Singlet Identification





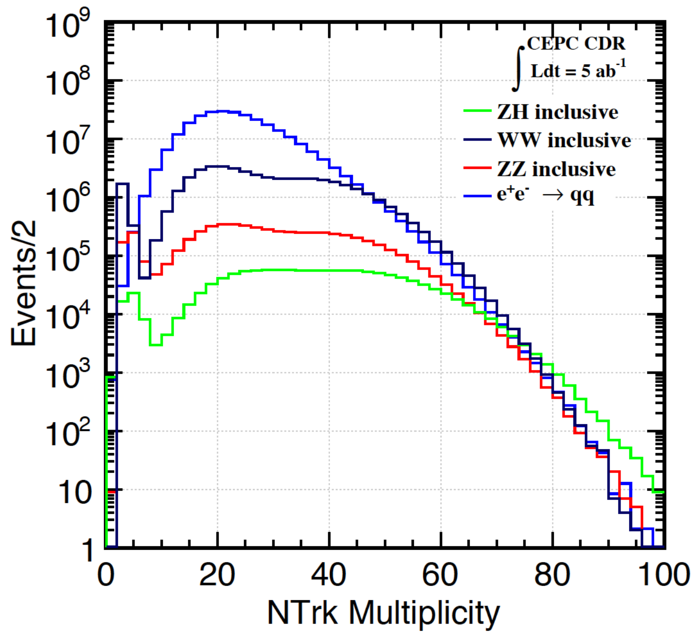
Color Singlet Identification

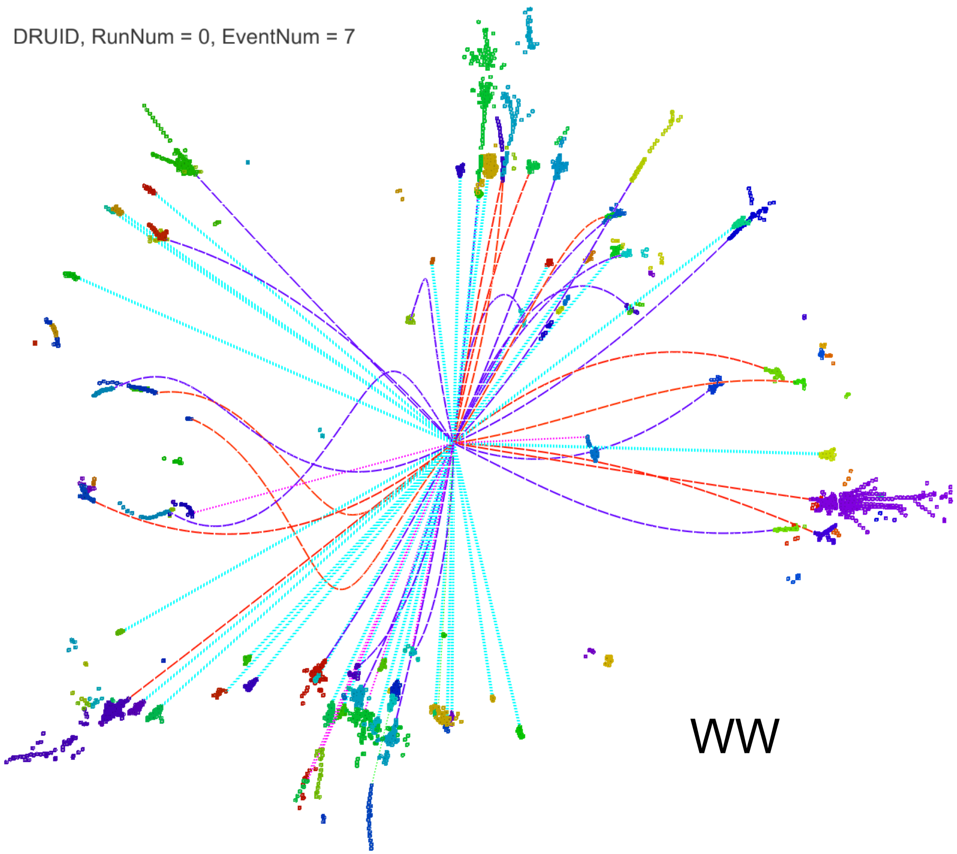




- CSI: identify the color single origin of each final state particle
- Grouping problem: essential for the physics measurements with multi-jet events, i.e., measurements with full hadronic ZH events
- AI might well strongly enhance its performance: compared to conventional jet clustering & matching

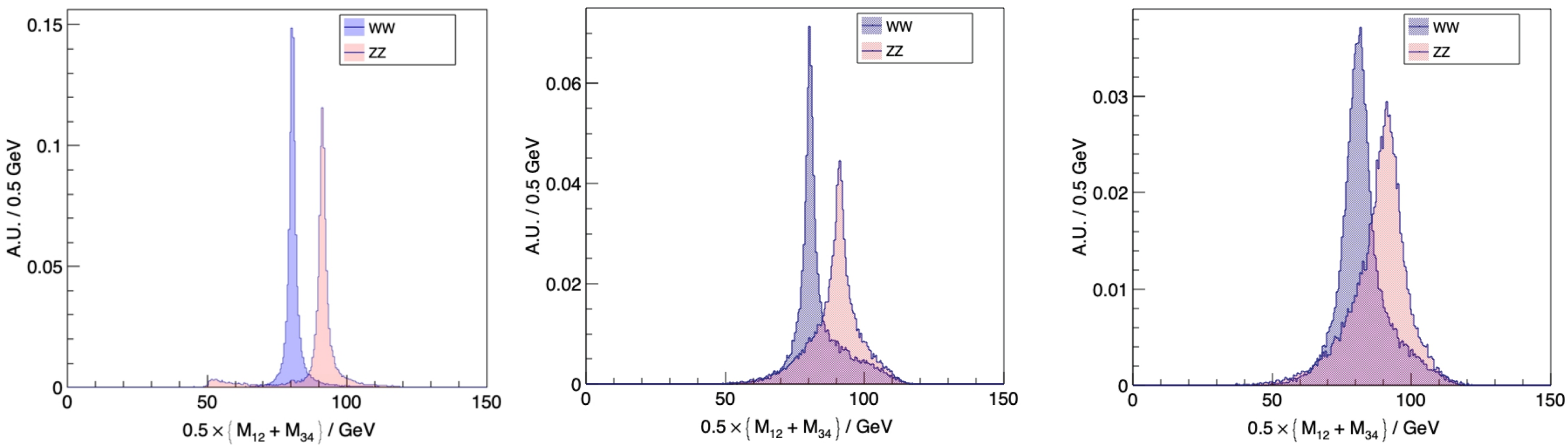
BM-III: full hadronic WW-ZZ separation





- 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

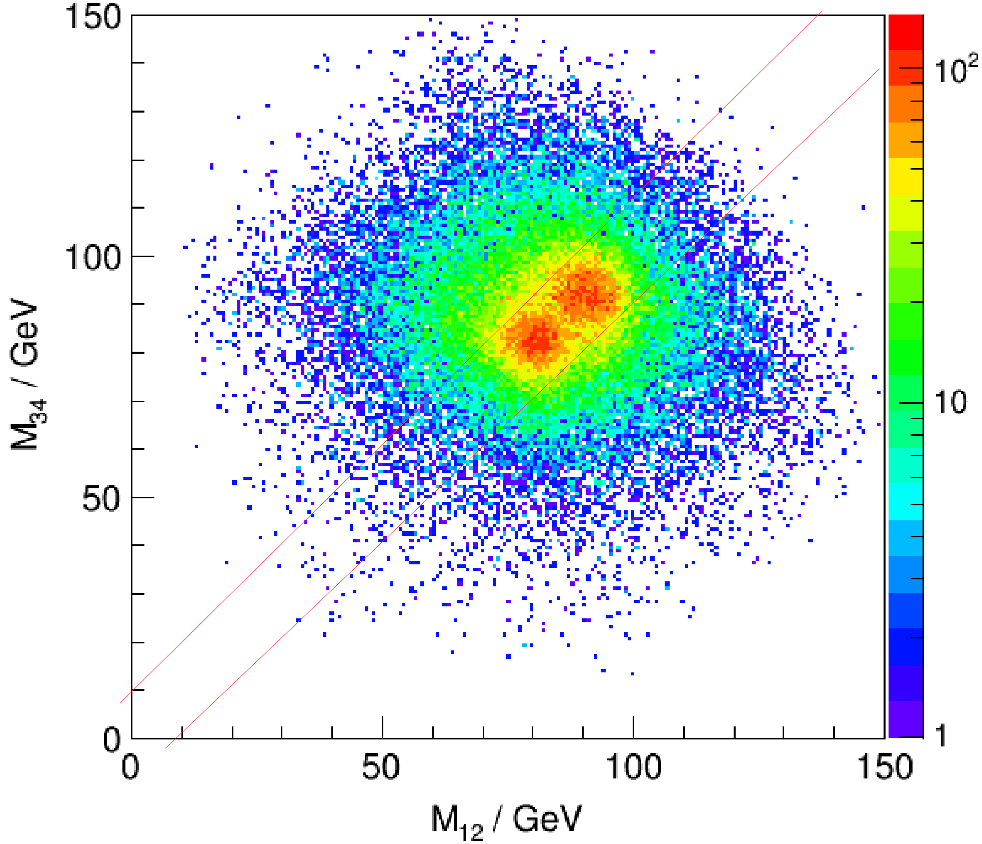
$$overlapping\ ratio = \sum_{bins} min(a_i, b_i)$$

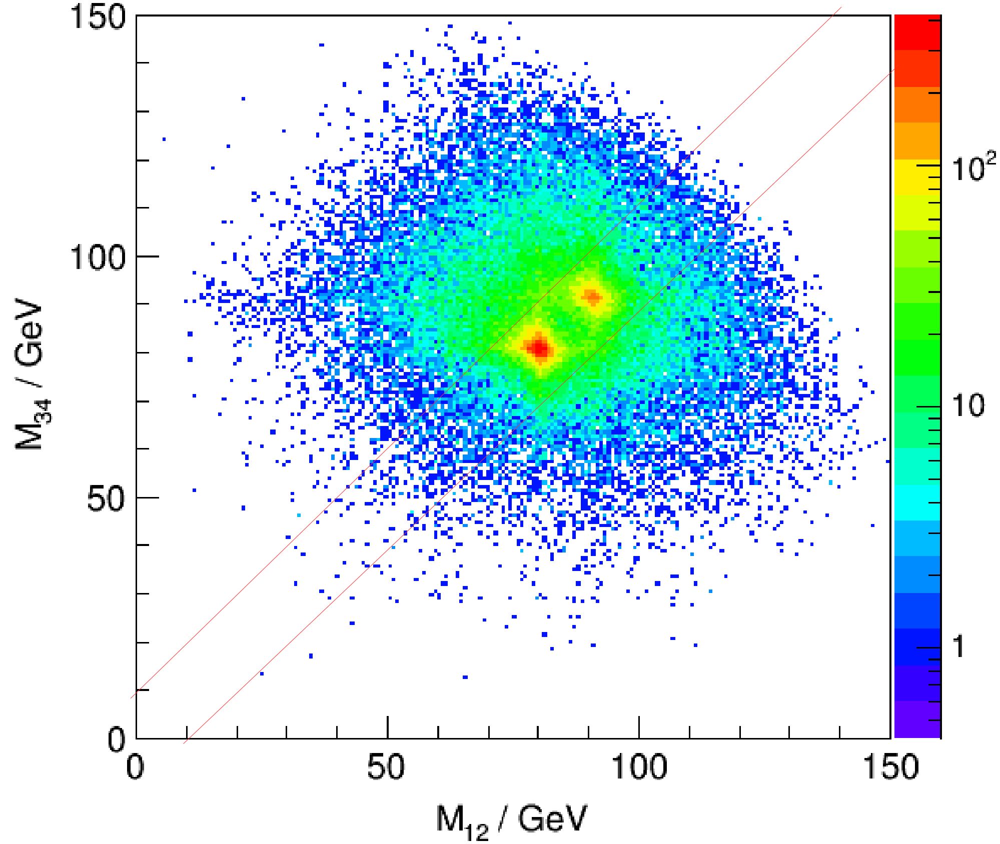
- Final state/MC particles are clustered into Reco/Genjet with ee-kt, and paired according to chi2

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

- 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%

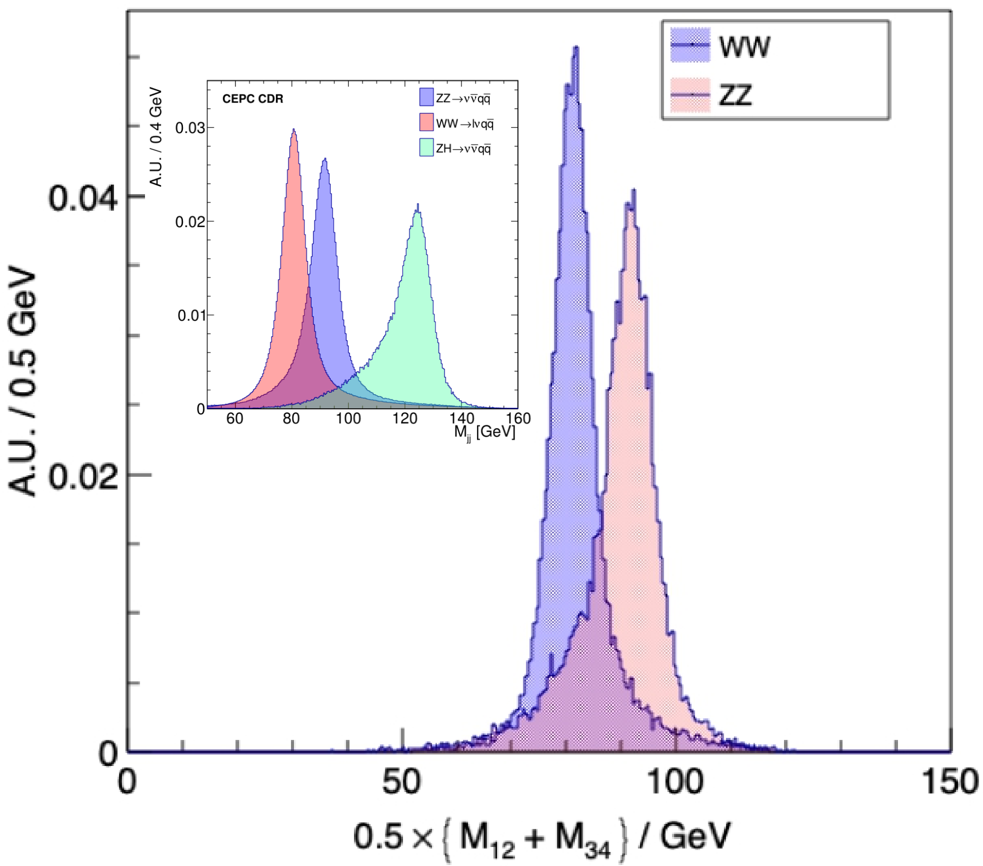
Reconstructed mass of the two di-jet system

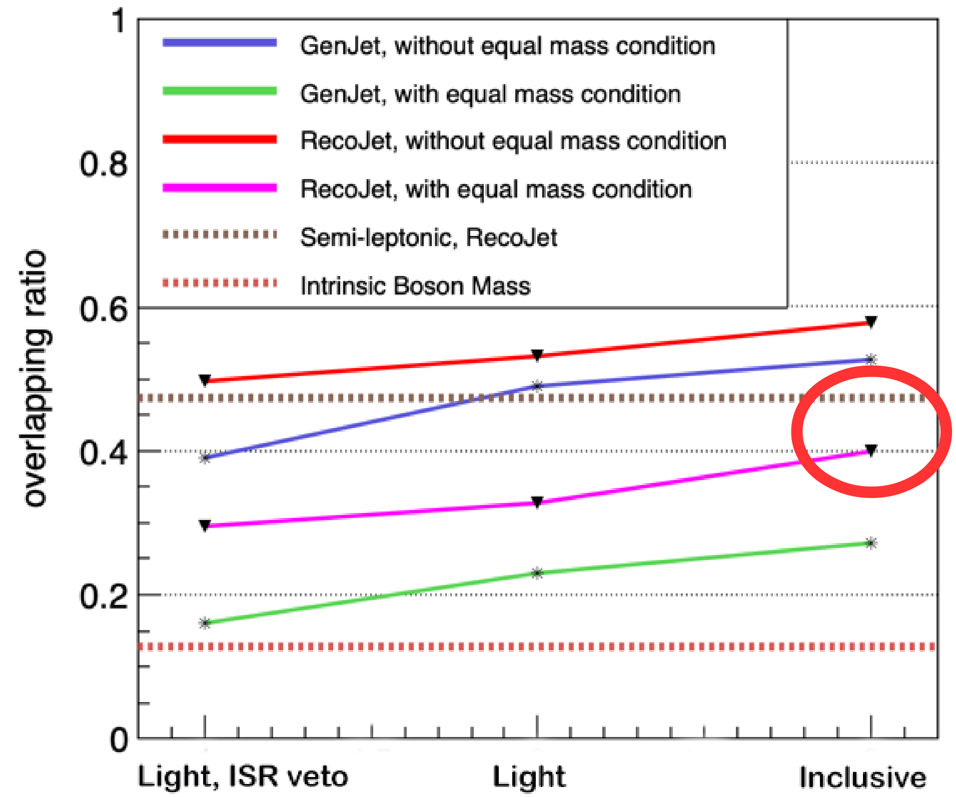




Equal mass condition |M12 – M34| < 10 GeV: At the cost of half the statistic, the overlapping ratio can be reduced from 58%/53% to 40%/27% for the Reco/Genjet

Separation of full hadronic WW-ZZ event





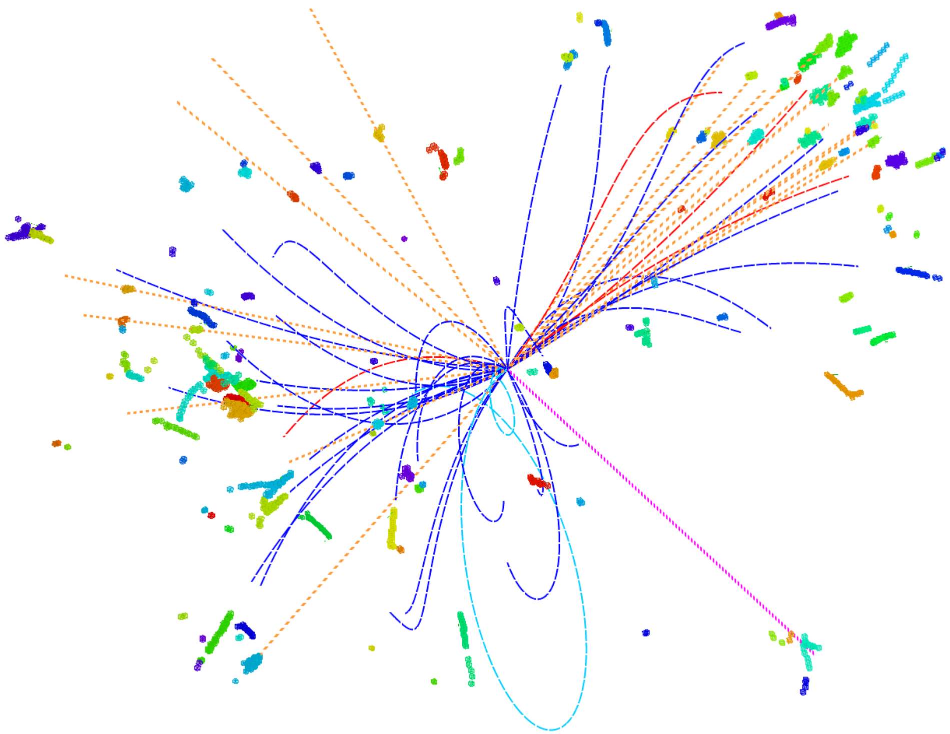
The CEPC Baseline could separate efficiently the WW-ZZ with full hadronic final state.

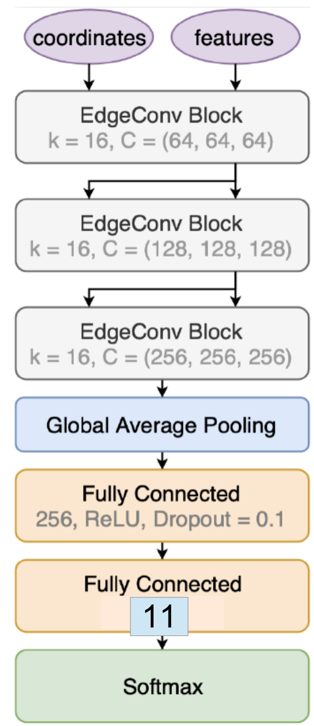
Critical to develop color singlet reconstruction: improve from the naive Jet clustering & pairing.

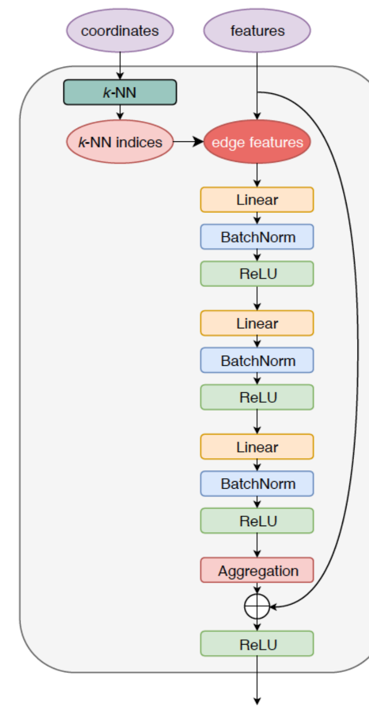
Quantified by differential overlapping ratio.

Control of ISR photon/neutrinos from heavy flavor jet is important.

Geo. & Tools

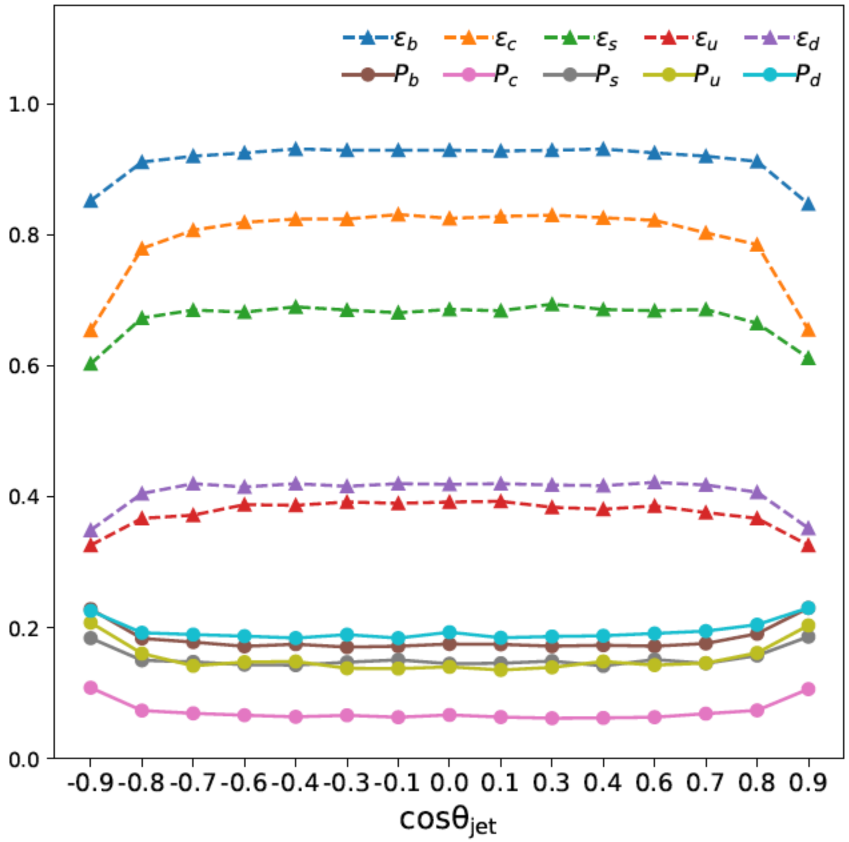


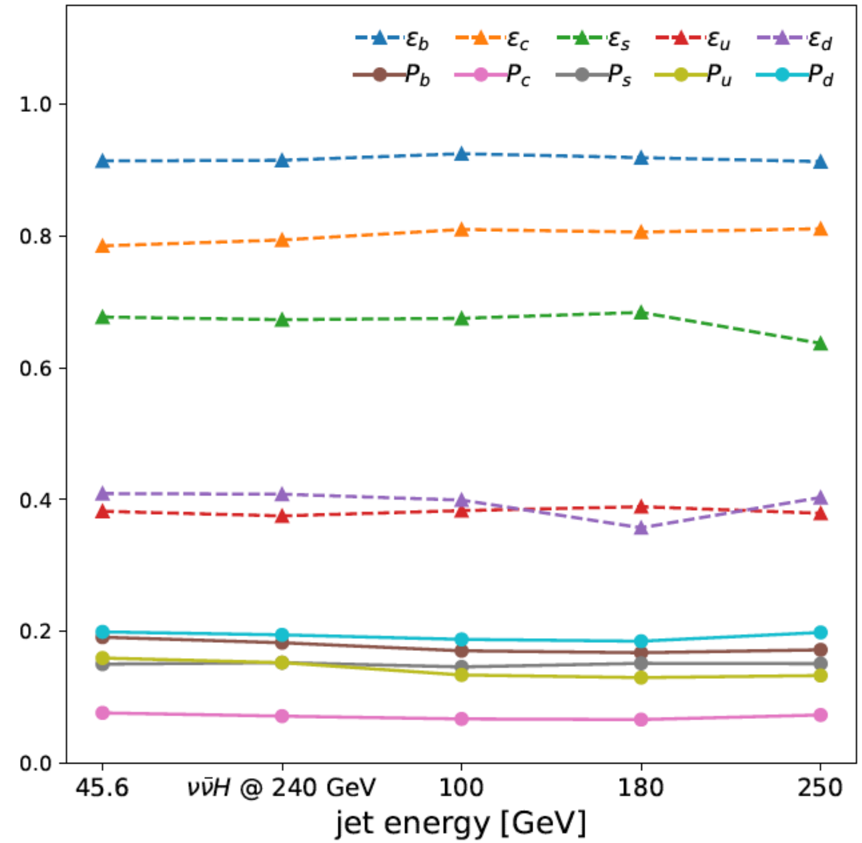




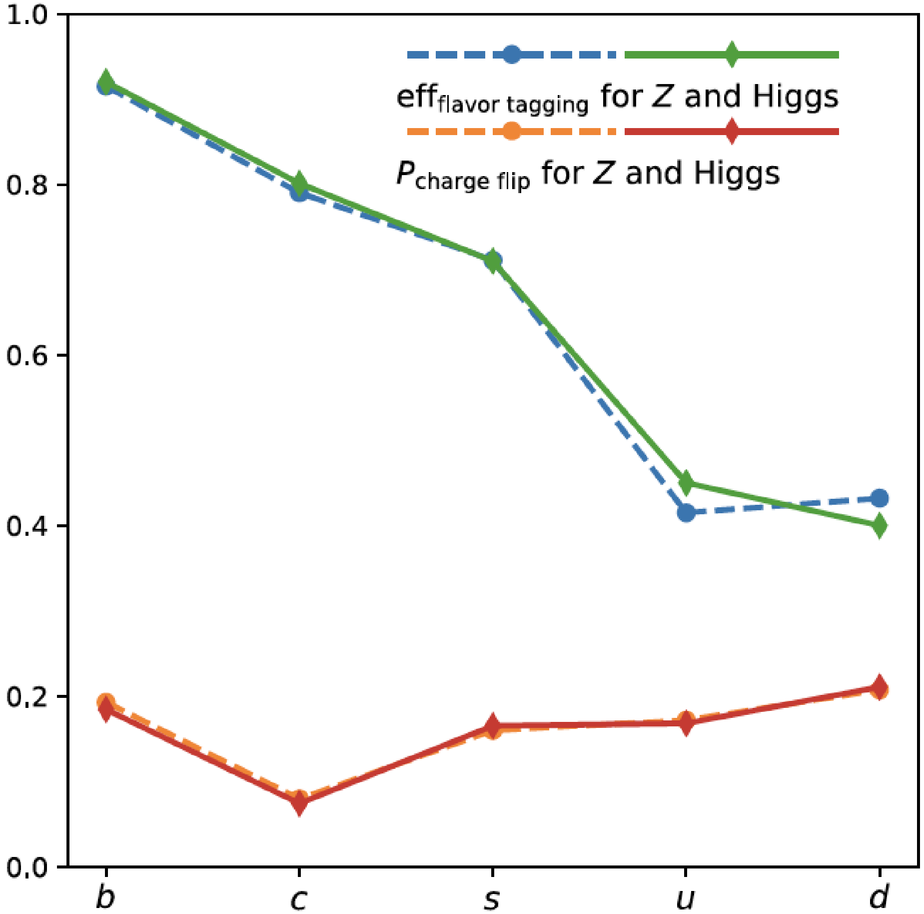
- **Jet origin identification: 11 categories (5 quarks + 5 anti quarks + gluon)**
- Full Simulated vvH, Higgs to two jets sample at CEPC baseline configuration: CEPC-v4 detector, reconstructed with **Arbor + ParticleNet (Deep Learning Tech.)**
  - Input: measurable information of all reconstructed jet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

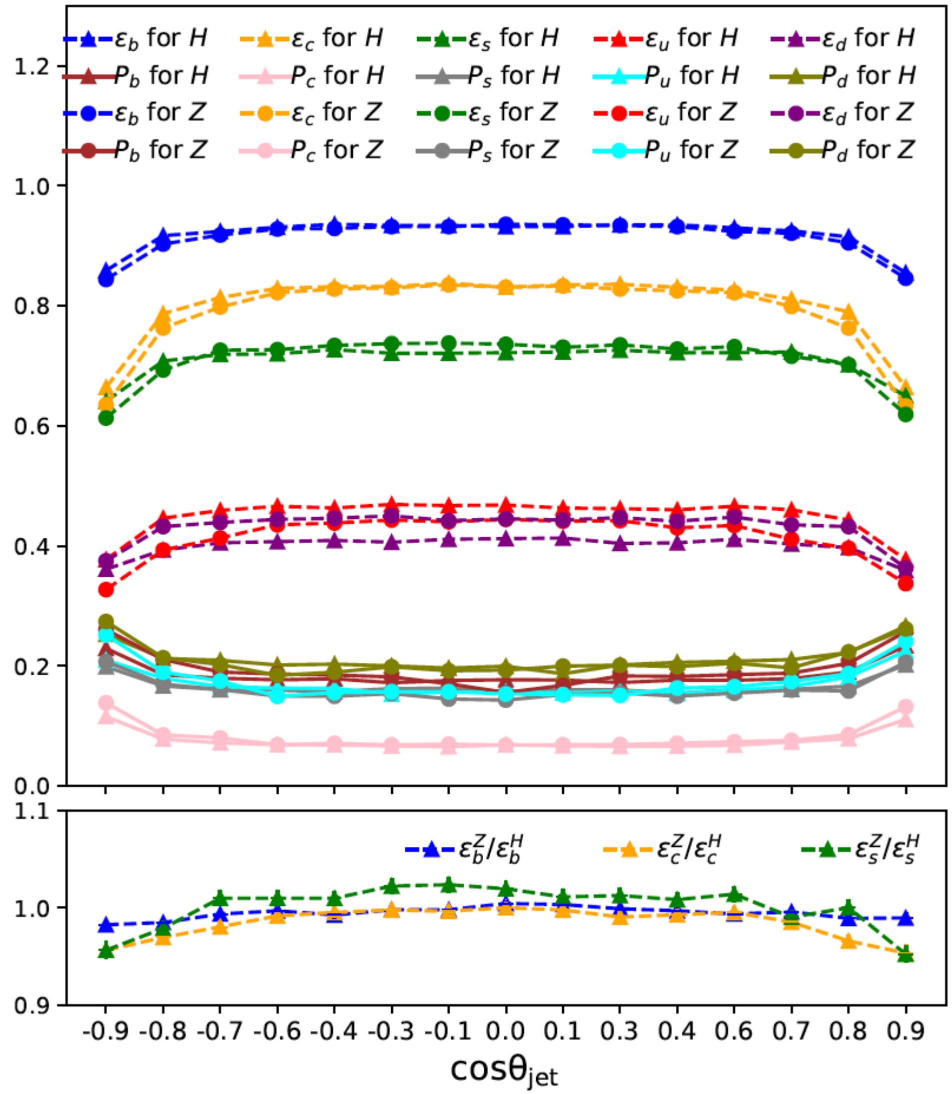
Performance V.S. Jet Kinematics





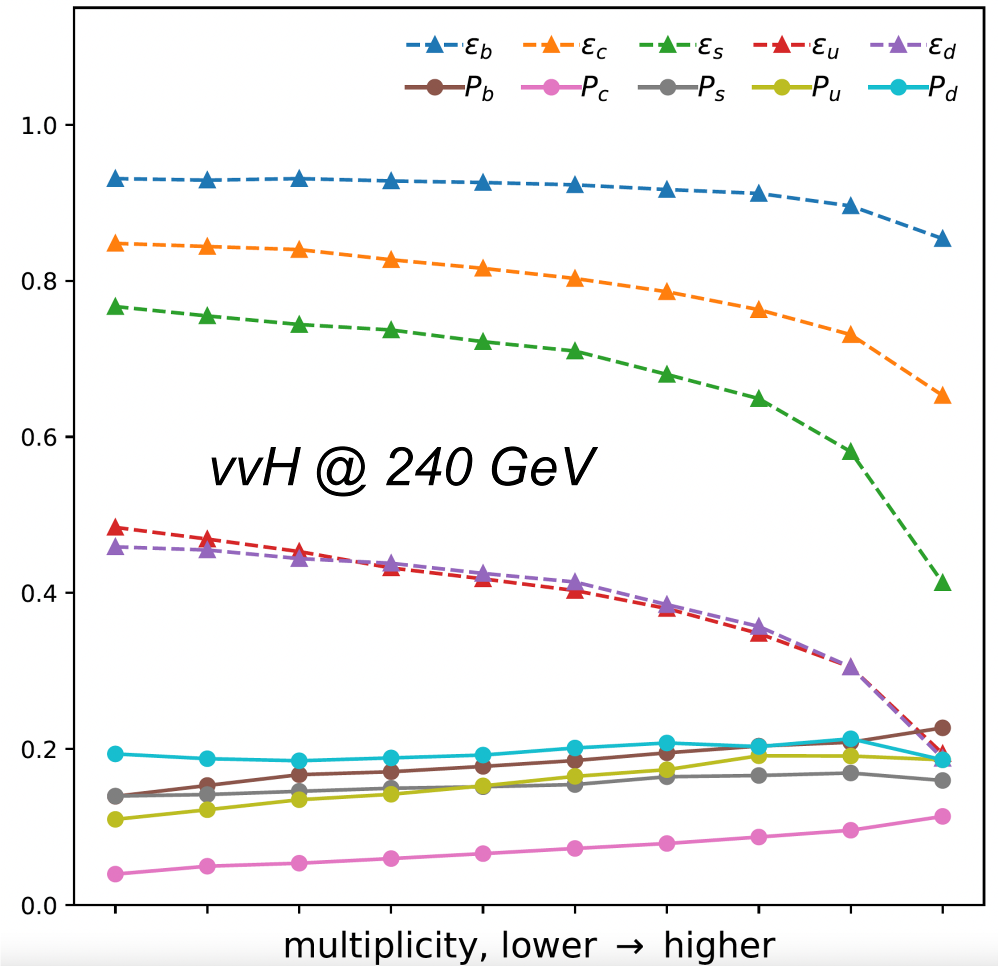
Performance @ Z and Higgs

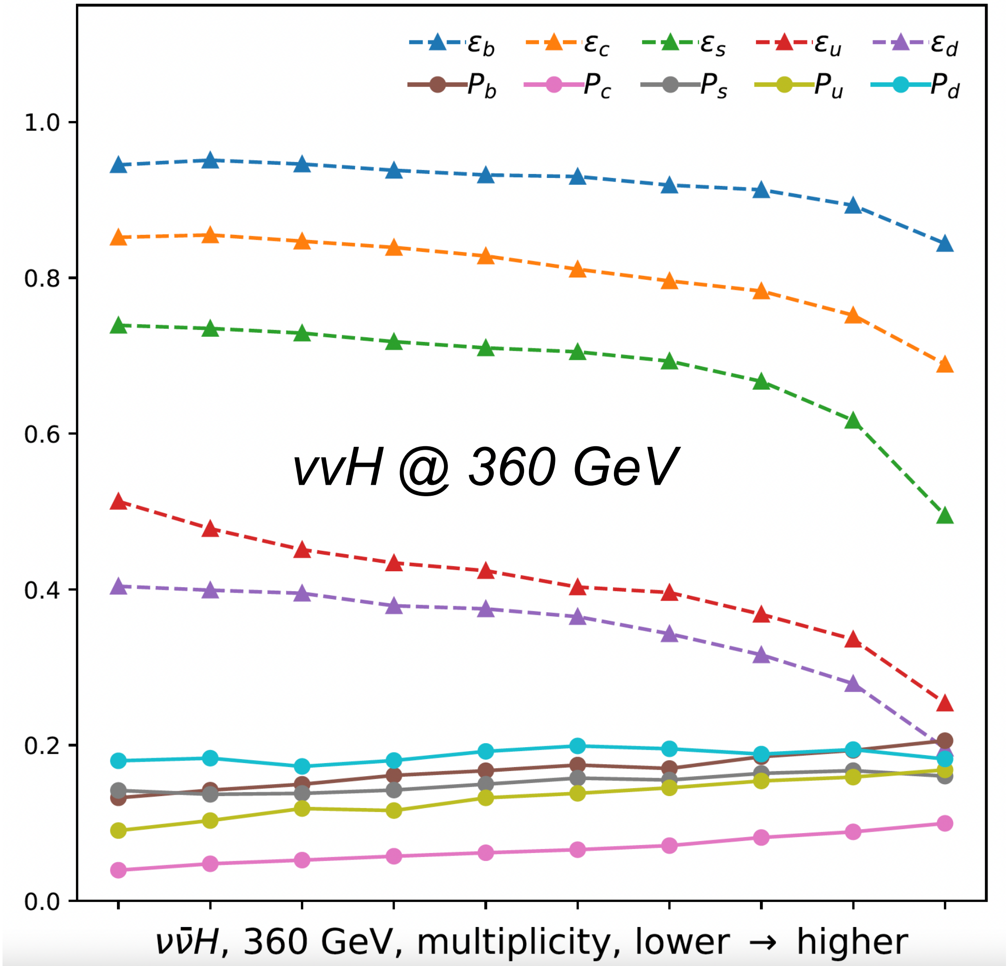




- *M10 instead of M11*

V.S. Multiplicity

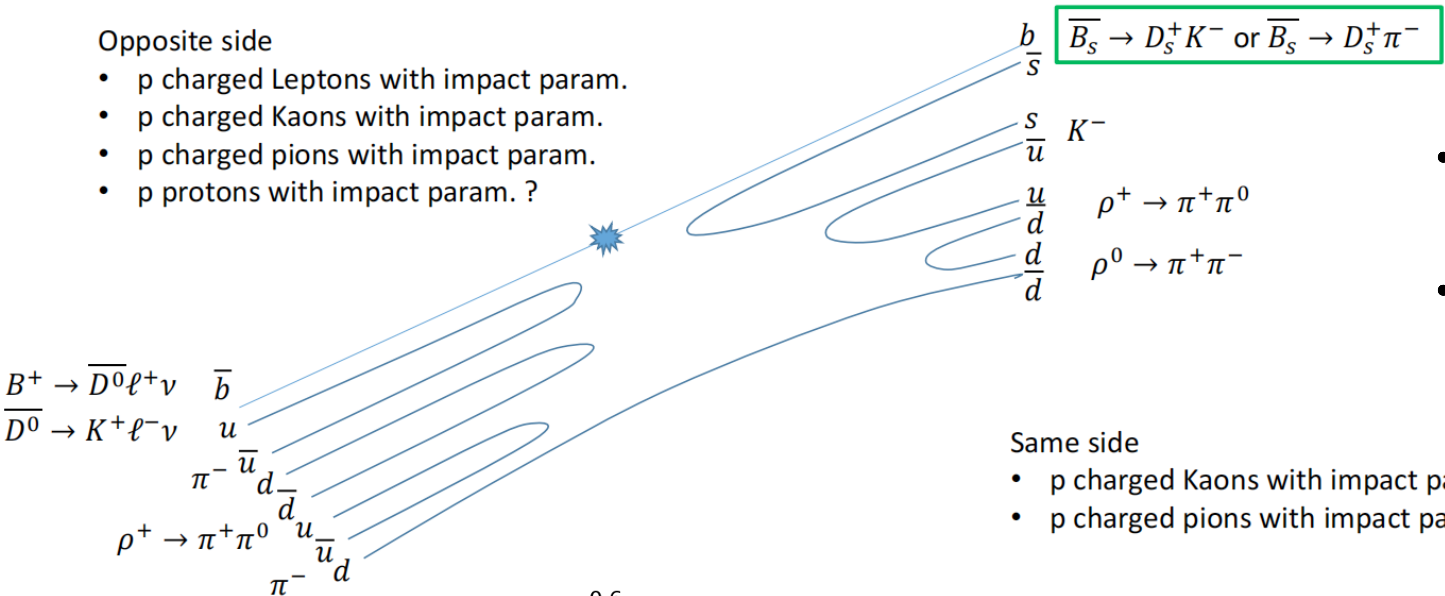




- *...many patterns need further understanding & towards further optimization...*

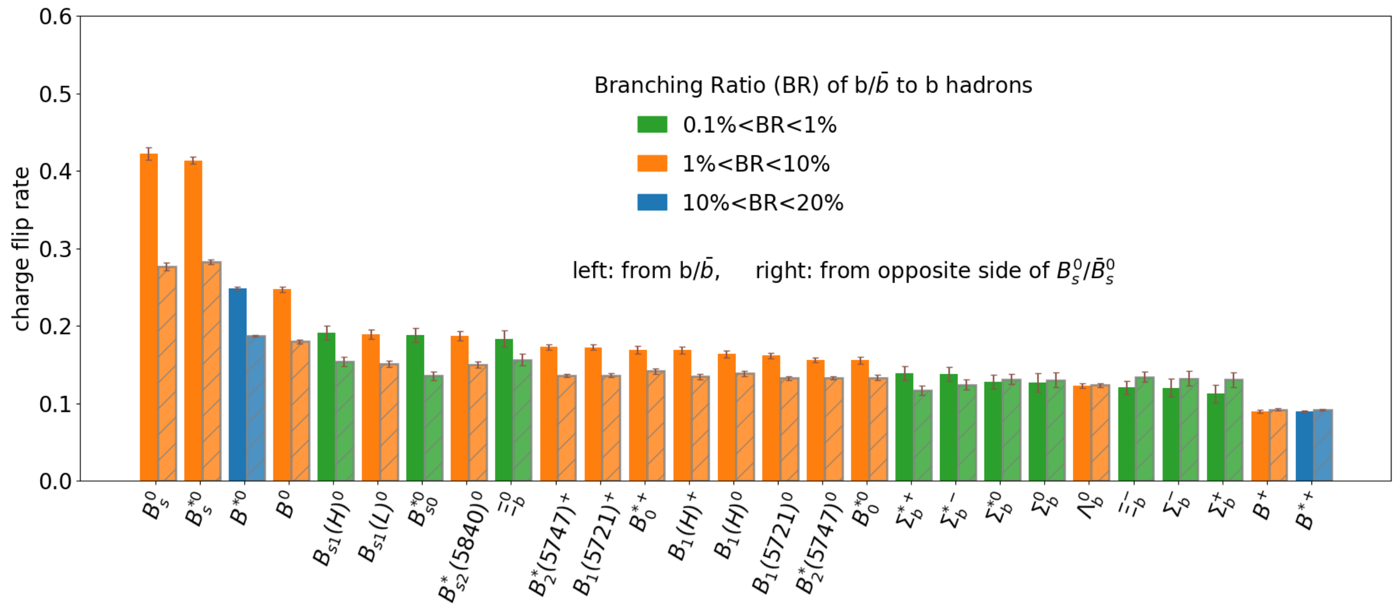
B-charge flip rate: Bs oscillations

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. ?

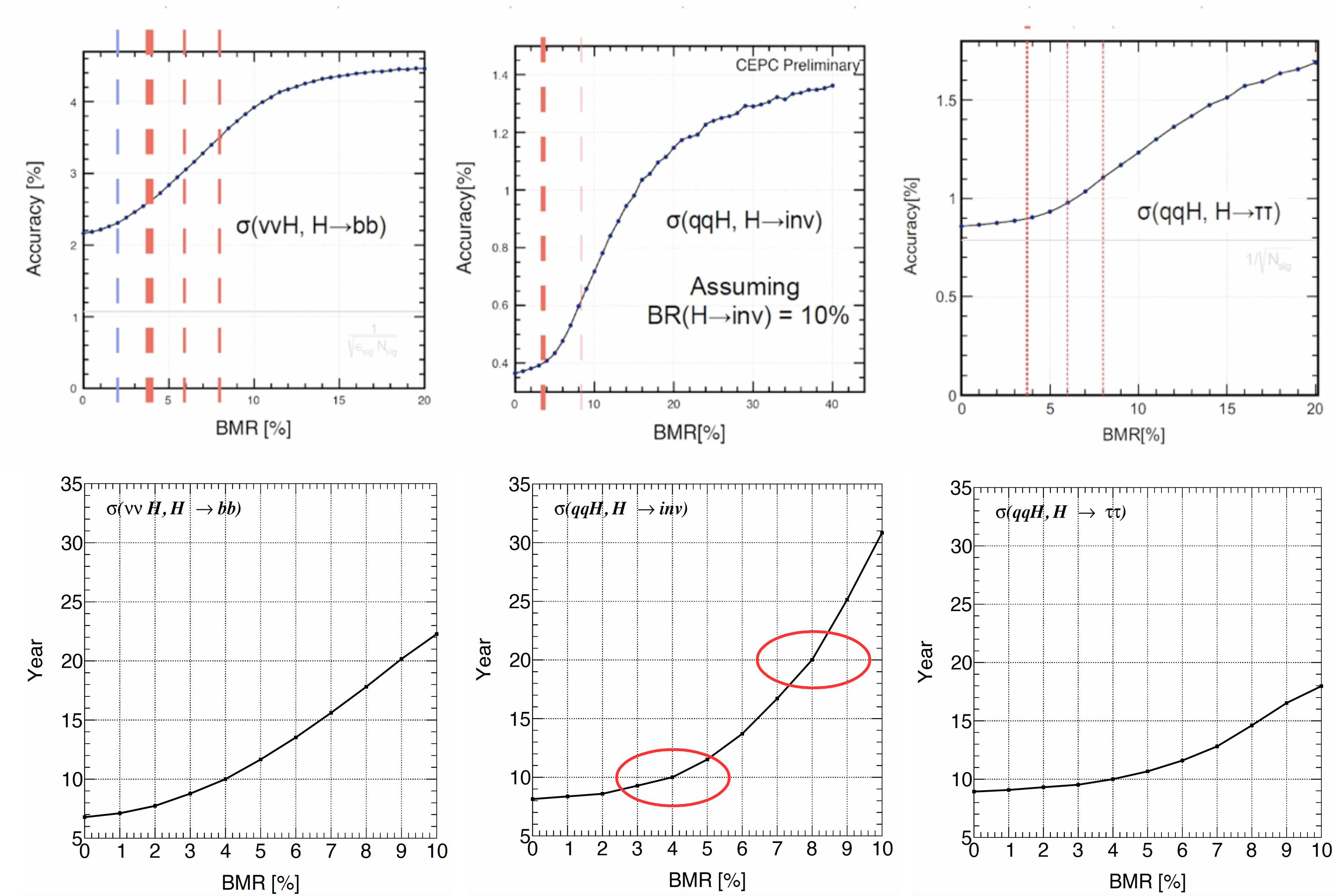


- Using all reco P (exc. Bs decay final state):
- Flip rate ~ 15%, Eff. Tagging power > 40%

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