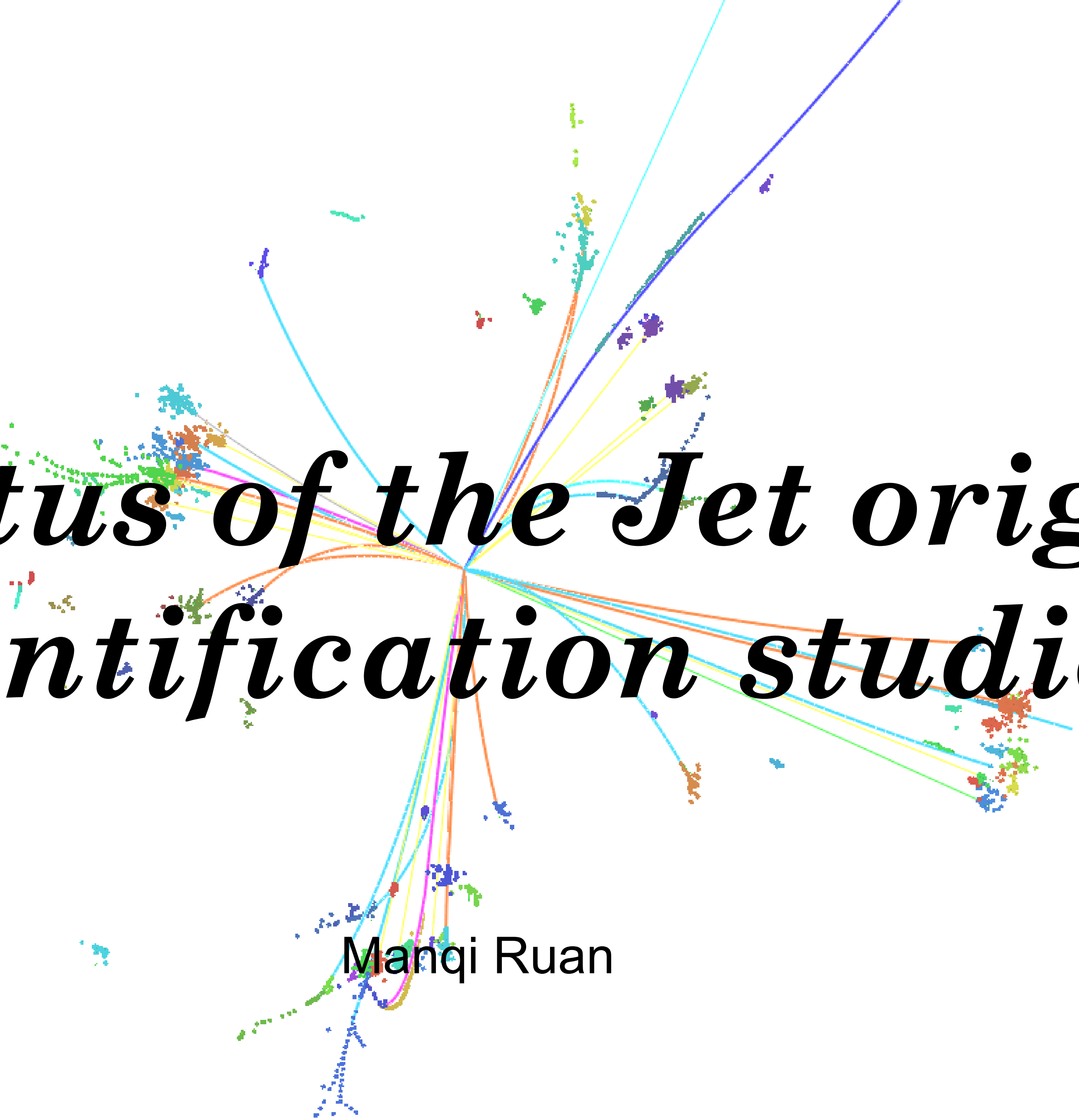
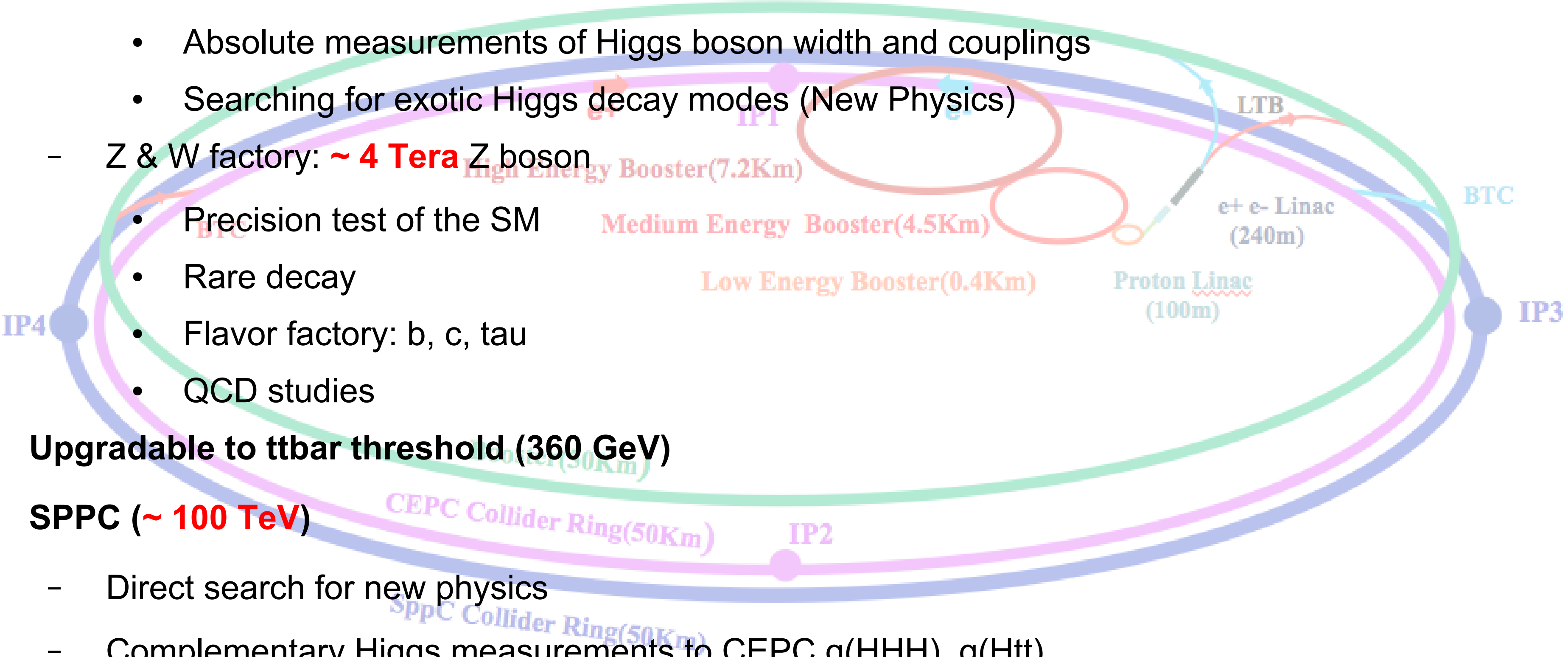


Status of the Jet origin identification studies

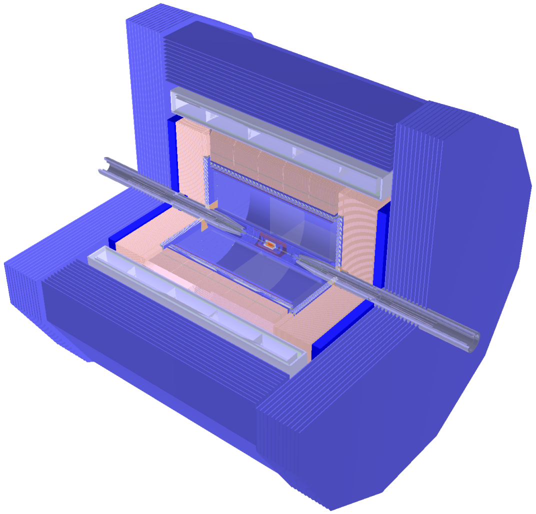
Manqi Ruan

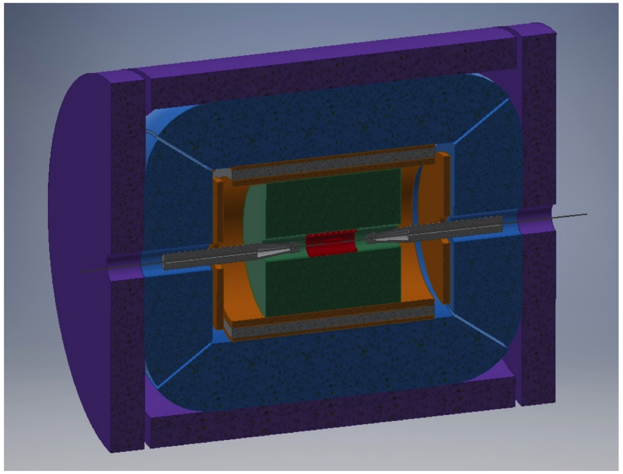
Key figures of the CEPC-SPPC

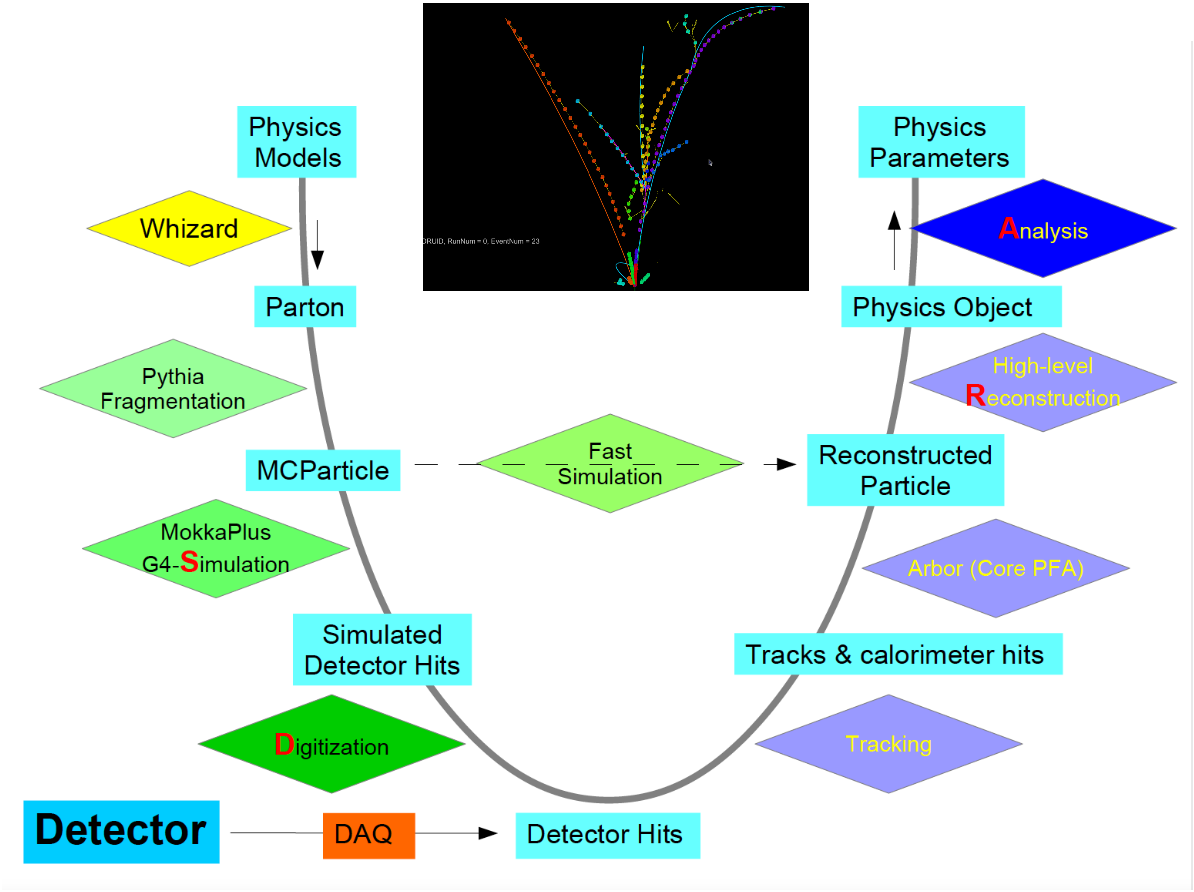


- **Tunnel ~ 100 km**
- **CEPC (90 – 240 GeV)**
  - Higgs factory: **4M** Higgs boson
    - Absolute measurements of Higgs boson width and couplings
    - Searching for exotic Higgs decay modes (New Physics)
  - Z & W factory: **~ 4 Tera** Z boson
    - Precision test of the SM
    - Rare decay
    - Flavor factory: b, c, tau
    - QCD studies
- **Upgradable to ttbar threshold (360 GeV)**
- **SPPC (~ 100 TeV)**
  - Direct search for new physics
  - Complementary Higgs measurements to CEPC g(HHH), g(Htt)
  - ...
- **Heavy ion, e-p collision...**

Detector & Software







Physics Models

Whizard

Parton

Pythia Fragmentation

MCParticle

MokkaPlus G4-Simulation

Simulated Detector Hits

Digitization

Detector

DAQ

Detector Hits

Fast Simulation

Reconstructed Particle

Physics Parameters

Analysis

Physics Object

High-level Reconstruction

Arbor (Core PFA)

Tracks & calorimeter hits

Tracking

Full simulation reconstruction Chain with Arbor, iterating/validation with hardware studies

Hadronic events: the main course

240 GeV: 97% of Higgs events has jets final states...

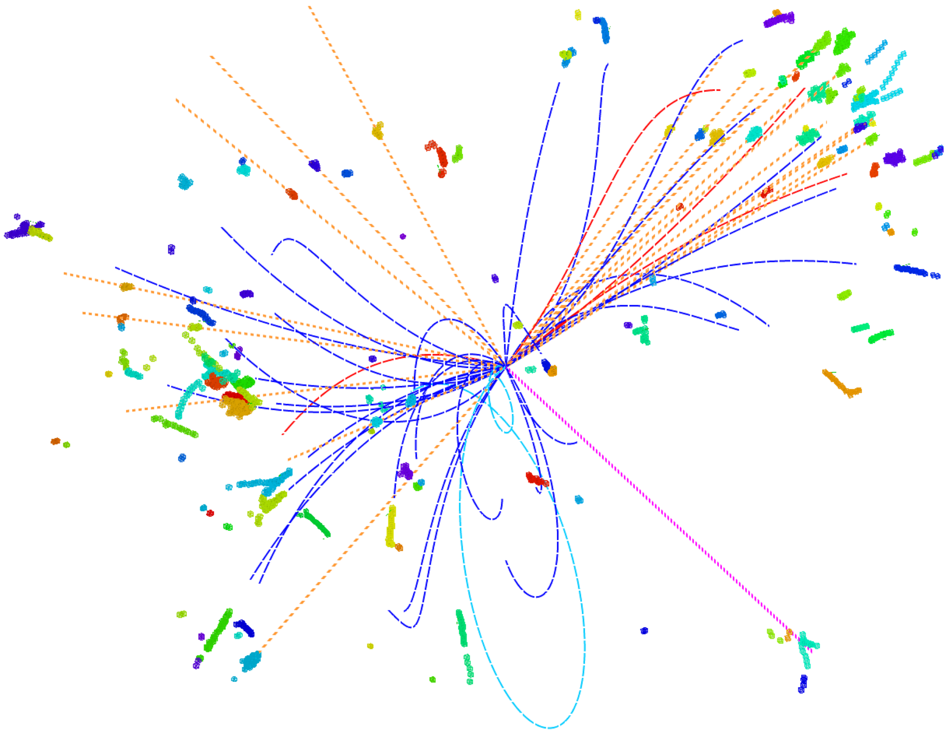
Jet origin id

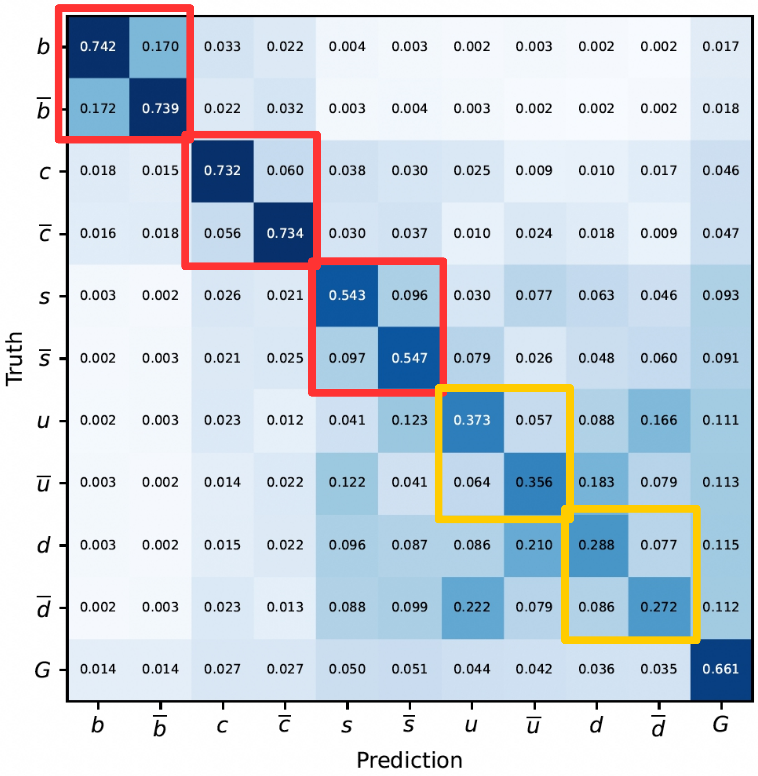
Hao Liang, Yongfeng Zhu, Yuzhi Che, Yuexin Wang, Huiling Qu, Cen Zhou, etc

https://arxiv.org/abs/2310.03440

https://arxiv.org/abs/2309.13231

Recent HL: Jet Origin Identification

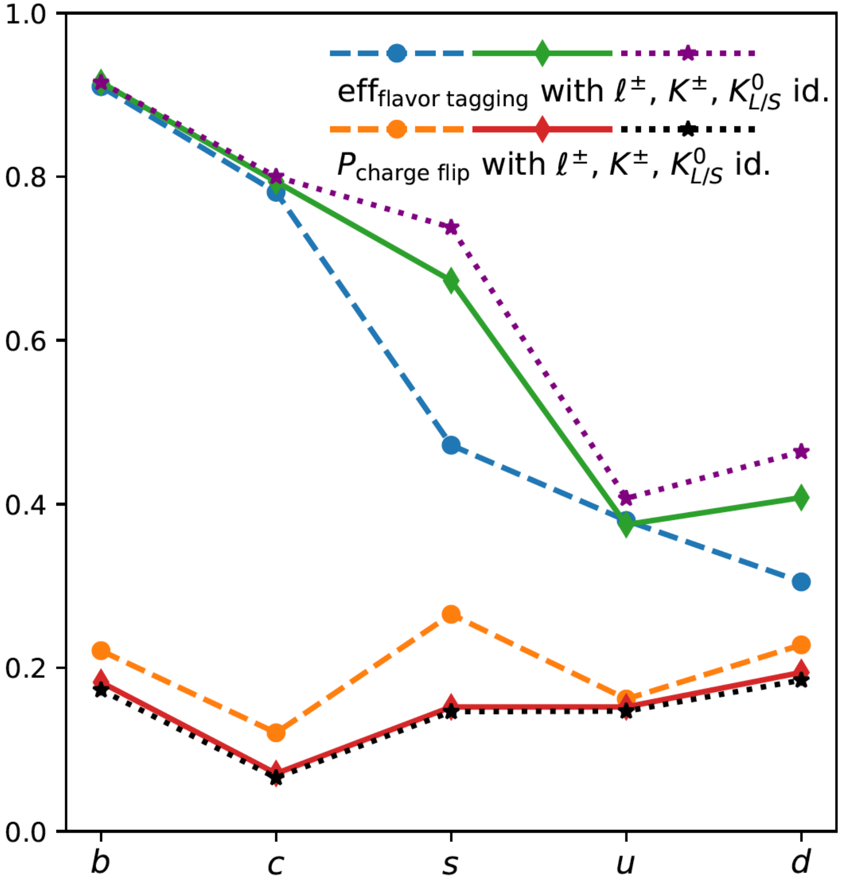


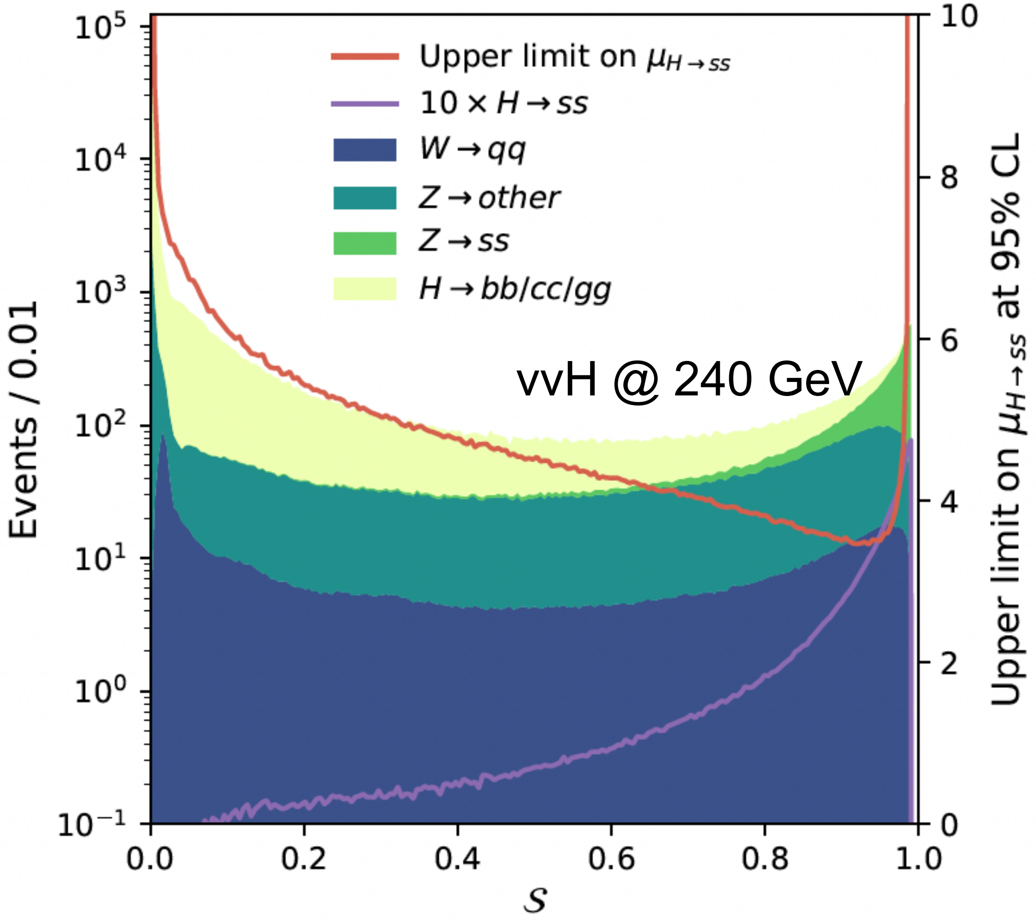


Truth \ Prediction	b	$\bar{b}$	c	$\bar{c}$	s	$\bar{s}$	u	$\bar{u}$	d	$\bar{d}$	G
b	0.742	0.170	0.033	0.022	0.004	0.003	0.002	0.003	0.002	0.002	0.017
$\bar{b}$	0.172	0.739	0.022	0.032	0.003	0.004	0.003	0.002	0.002	0.002	0.018
c	0.018	0.015	0.732	0.060	0.038	0.030	0.025	0.009	0.010	0.017	0.046
$\bar{c}$	0.016	0.018	0.056	0.734	0.030	0.037	0.010	0.024	0.018	0.009	0.047
s	0.003	0.002	0.026	0.021	0.543	0.096	0.030	0.077	0.063	0.046	0.093
$\bar{s}$	0.002	0.003	0.021	0.025	0.097	0.547	0.079	0.026	0.048	0.060	0.091
u	0.002	0.003	0.023	0.012	0.041	0.123	0.373	0.057	0.088	0.166	0.111
$\bar{u}$	0.003	0.002	0.014	0.022	0.122	0.041	0.064	0.356	0.183	0.079	0.113
d	0.003	0.002	0.015	0.022	0.096	0.087	0.086	0.210	0.288	0.077	0.115
$\bar{d}$	0.002	0.003	0.023	0.013	0.088	0.099	0.222	0.079	0.086	0.272	0.112
G	0.014	0.014	0.027	0.027	0.050	0.051	0.044	0.042	0.036	0.035	0.661

- **Jet origin identification: 11 categories (5 quarks + 5 anti quarks + gluon)**
  - Jet Flavor Tagging + Jet Charge measurements + s-tagging + gluon tagging...
- Full Simulated vvH, Higgs to two jets sample at CEPC baseline configuration: CEPC-v4 detector, reconstructed with **Arbor + ParticleNet (Deep Learning Tech.)**
- 1 Million samples each, 60/20/20% for training, validation & test

Performance with different PID scenarios & H→ss measurements

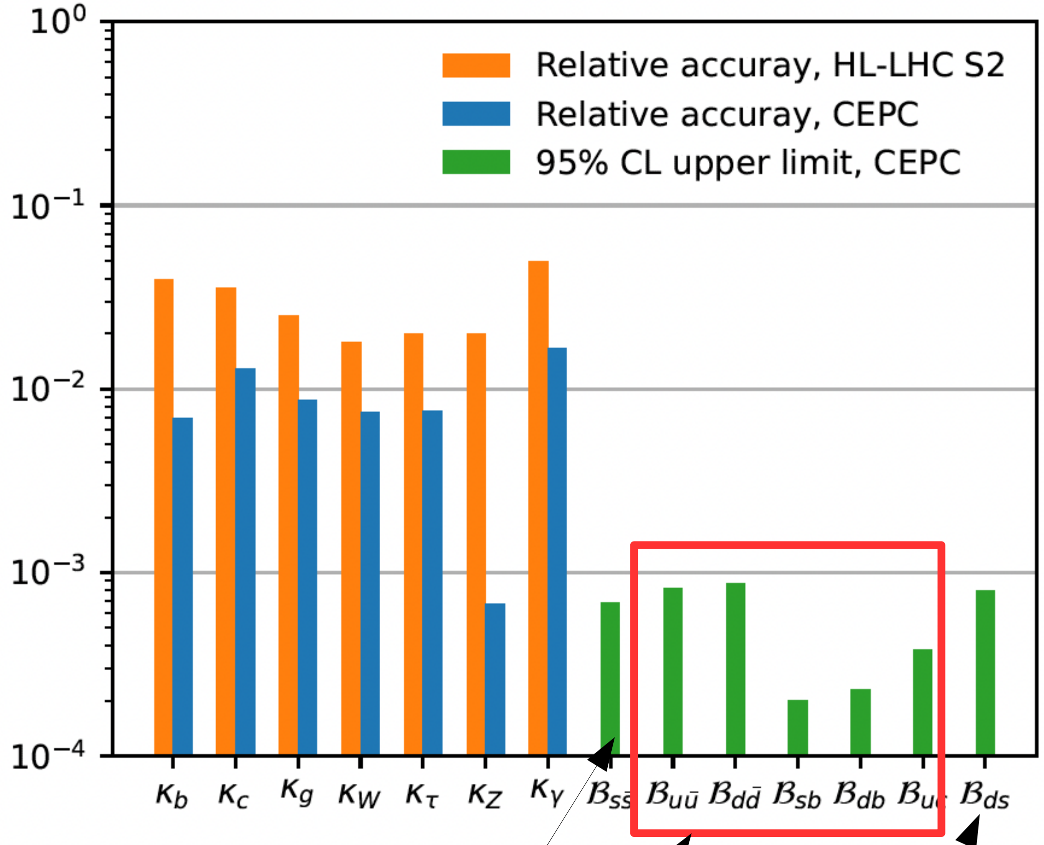




Flavor tagging: type that maximize {L_q + L_q_bar, L_g}

If quark jet: jet charge ~ compare {L_q, L_q_bar}

Benchmark analyses: Higgs rare/FCNC



Improved by ~3 times

Improved by 1-2 orders of magnitudes

Presumably... firstly quantified

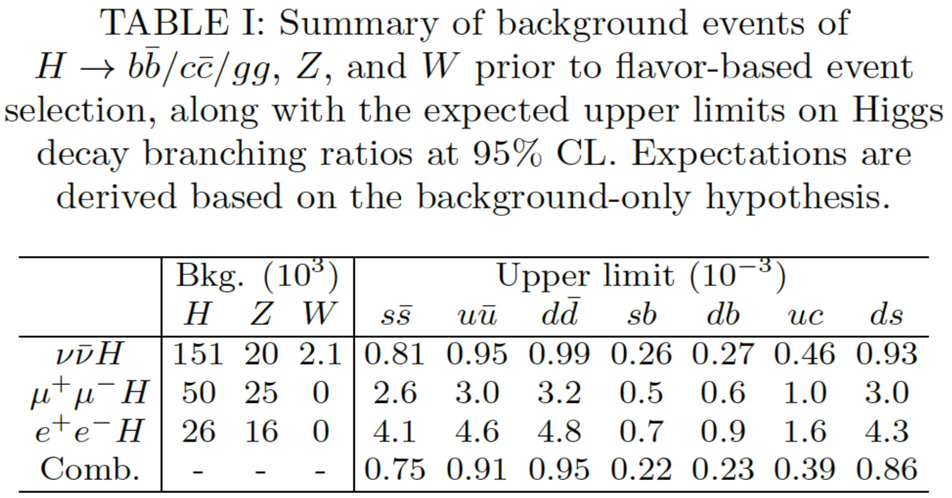
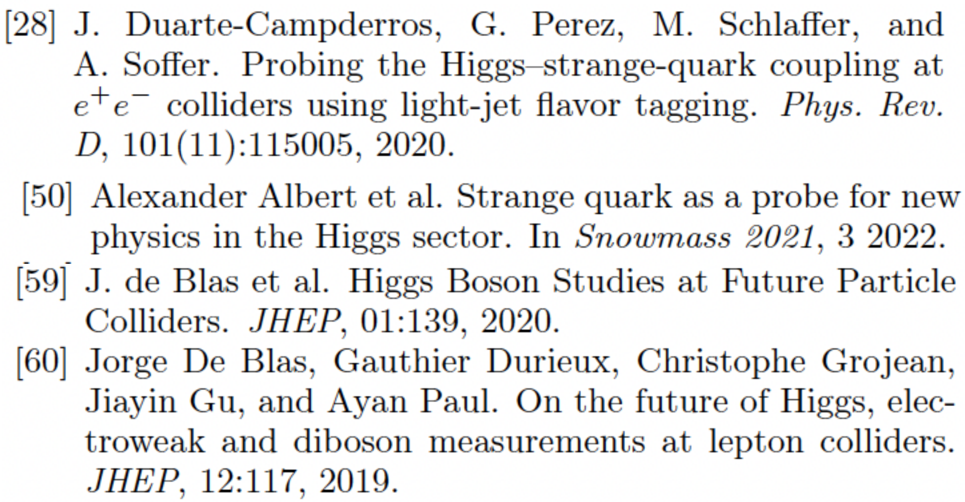


TABLE I: Summary of background events of $H \to b\bar{b}/c\bar{c}/gg$, Z, and W prior to flavor-based event selection, along with the expected upper limits on Higgs decay branching ratios at 95% CL. Expectations are derived based on the background-only hypothesis.

	Bkg. (10^3)			Upper limit (10^{-3})						
	H	Z	W	$s\bar{s}$	$u\bar{u}$	$d\bar{d}$	sb	db	uc	ds
$\nu\bar{\nu}H$	151	20	2.1	0.81	0.95	0.99	0.26	0.27	0.46	0.93
$\mu^+\mu^-H$	50	25	0	2.6	3.0	3.2	0.5	0.6	1.0	3.0
e^+e^-H	26	16	0	4.1	4.6	4.8	0.7	0.9	1.6	4.3
Comb.	-	-	-	0.75	0.91	0.95	0.22	0.23	0.39	0.86



[28] J. Duarte-Campderros, G. Perez, M. Schlaffer, and A. Soffer. Probing the Higgs–strange-quark coupling at e^+e^- colliders using light-jet flavor tagging. *Phys. Rev. D*, 101(11):115005, 2020.

[50] Alexander Albert et al. Strange quark as a probe for new physics in the Higgs sector. In *Snowmass 2021*, 3 2022.

[59] J. de Blas et al. Higgs Boson Studies at Future Particle Colliders. *JHEP*, 01:139, 2020.

[60] Jorge De Blas, Gauthier Durieux, Christophe Grojean, Jiayin Gu, and Ayan Paul. On the future of Higgs, electroweak and diboson measurements at lepton colliders. *JHEP*, 12:117, 2019.

For H->bb, cc, gg: results in 20 – 40% improvement in relative accuracies (preliminary)...

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

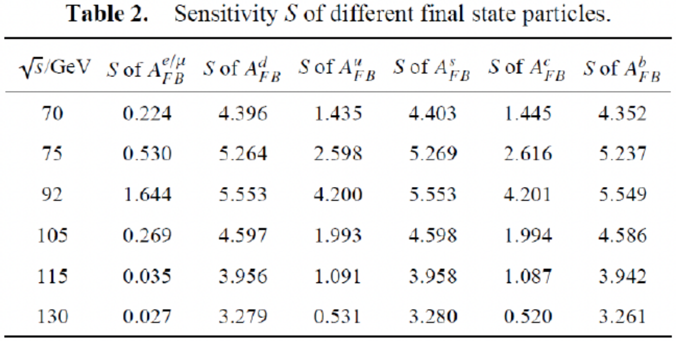


Table 2. Sensitivity S of different final state particles.

$\sqrt{s}$/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

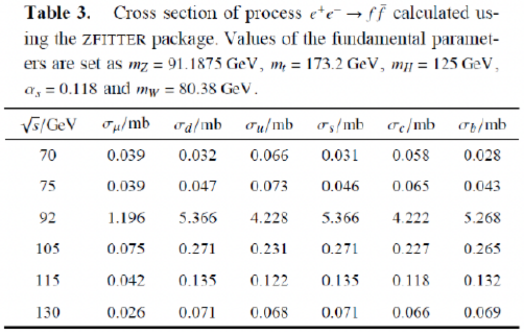


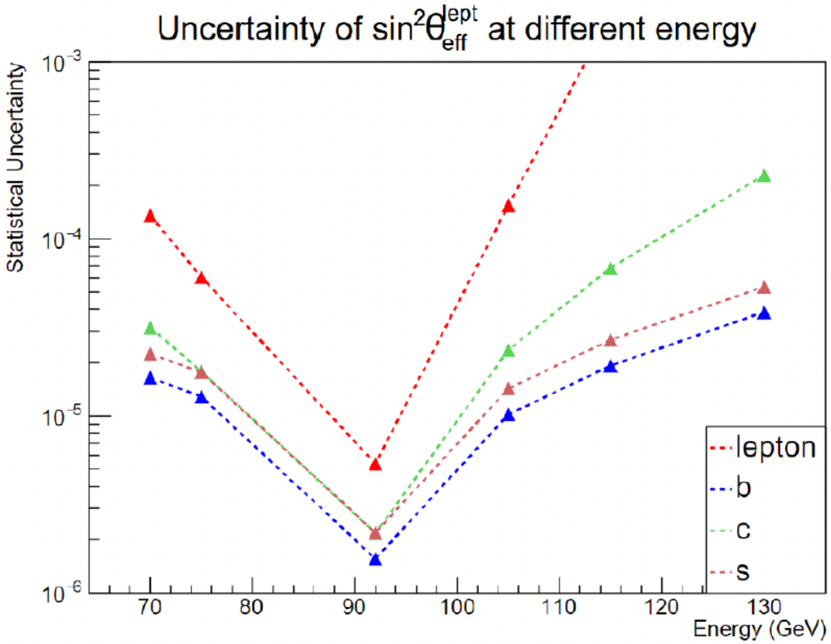
Table 3. Cross section of process $e^+e^- \rightarrow f\bar{f}$ calculated using the ZFITTER package. Values of the fundamental parameters are set as $m_Z = 91.1875$ GeV, $m_t = 173.2$ GeV, $m_H = 125$ GeV, $\alpha_s = 0.118$ and $m_W = 80.38$ GeV.

$\sqrt{s}$/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

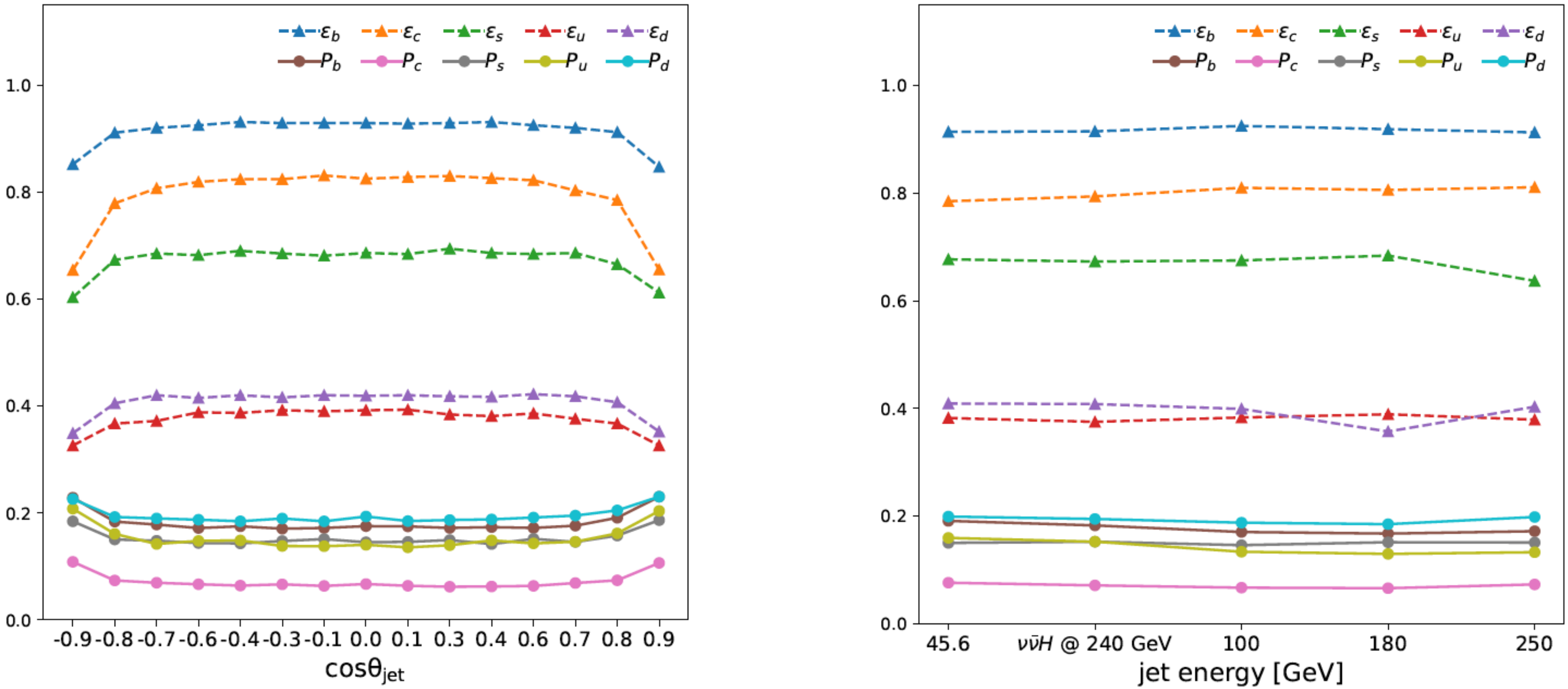
Expected statistical uncertainties on $\sin^2\theta^l_{eff}$ measurement.

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

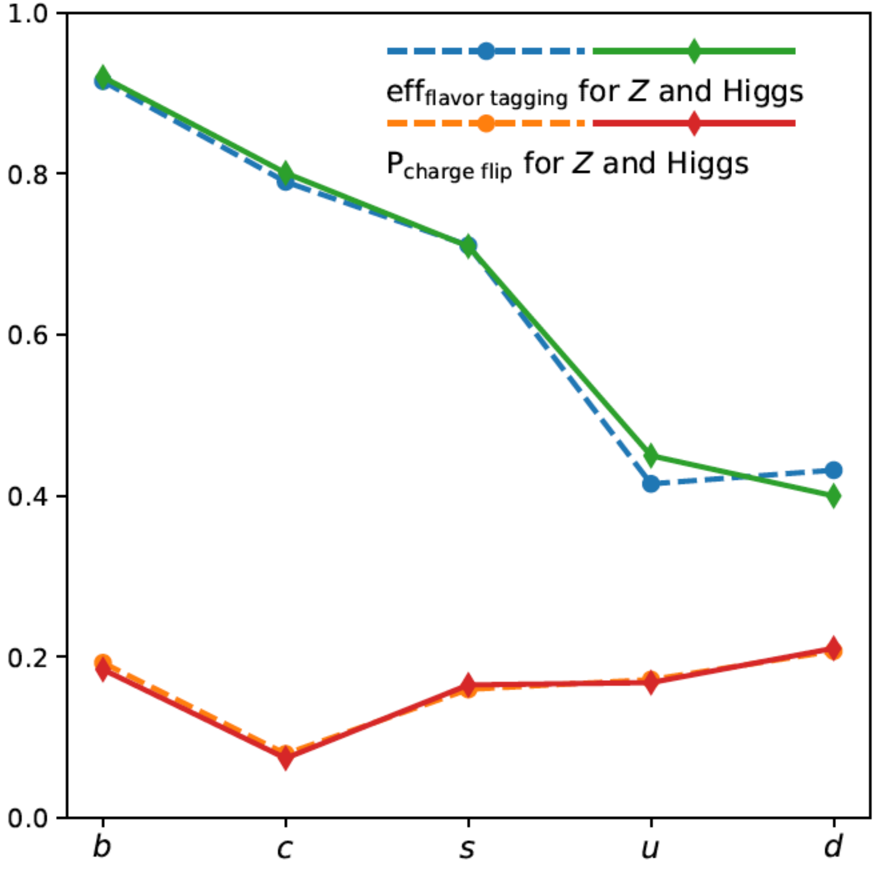


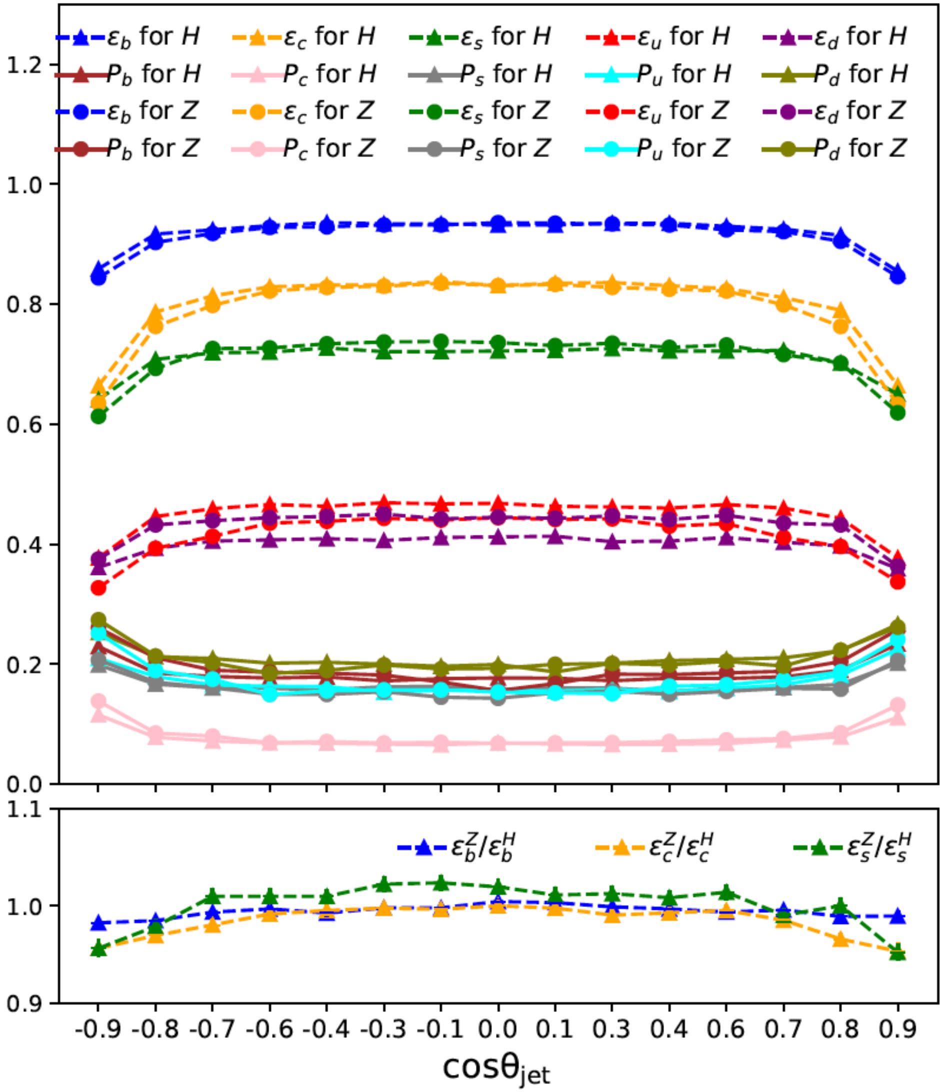
$\sqrt{s}$	b	c	s
70	1.6×10^{-5}	3.2×10^{-5}	2.2×10^{-5}
75	1.3×10^{-5}	1.8×10^{-5}	1.8×10^{-5}
92	1.6×10^{-6}	2.2×10^{-6}	2.2×10^{-6}
105	1.0×10^{-5}	2.4×10^{-5}	1.4×10^{-5}
115	1.9×10^{-5}	6.8×10^{-5}	2.7×10^{-5}
130	3.9×10^{-5}	2.3×10^{-4}	5.4×10^{-5}

Performance V.S. Jet Kinematics



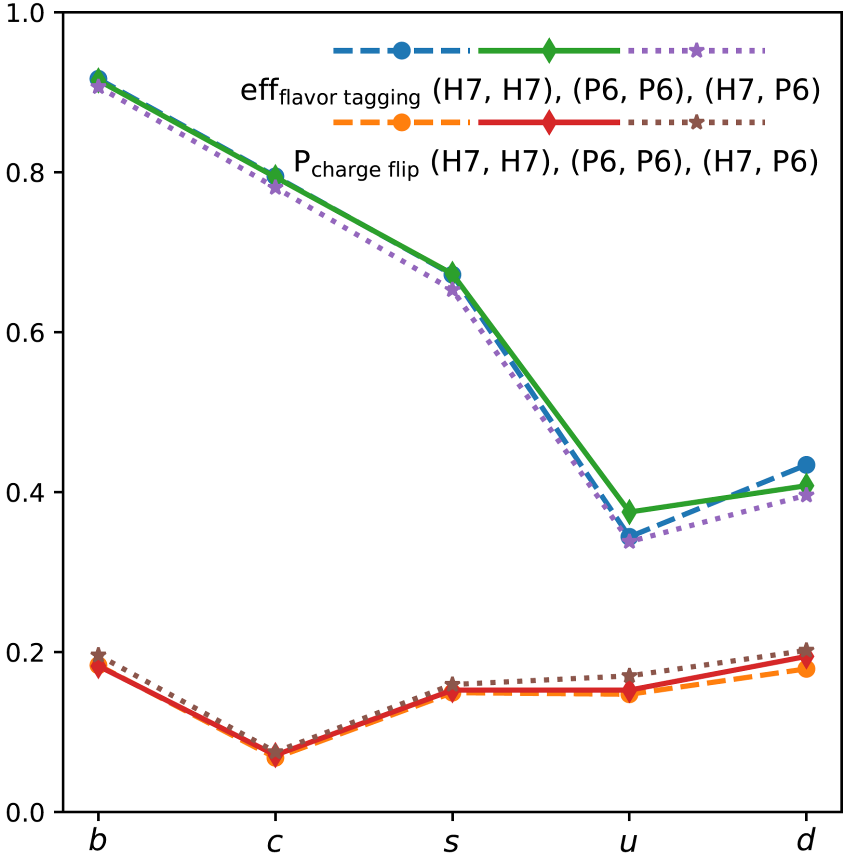
Performance @ Z and Higgs

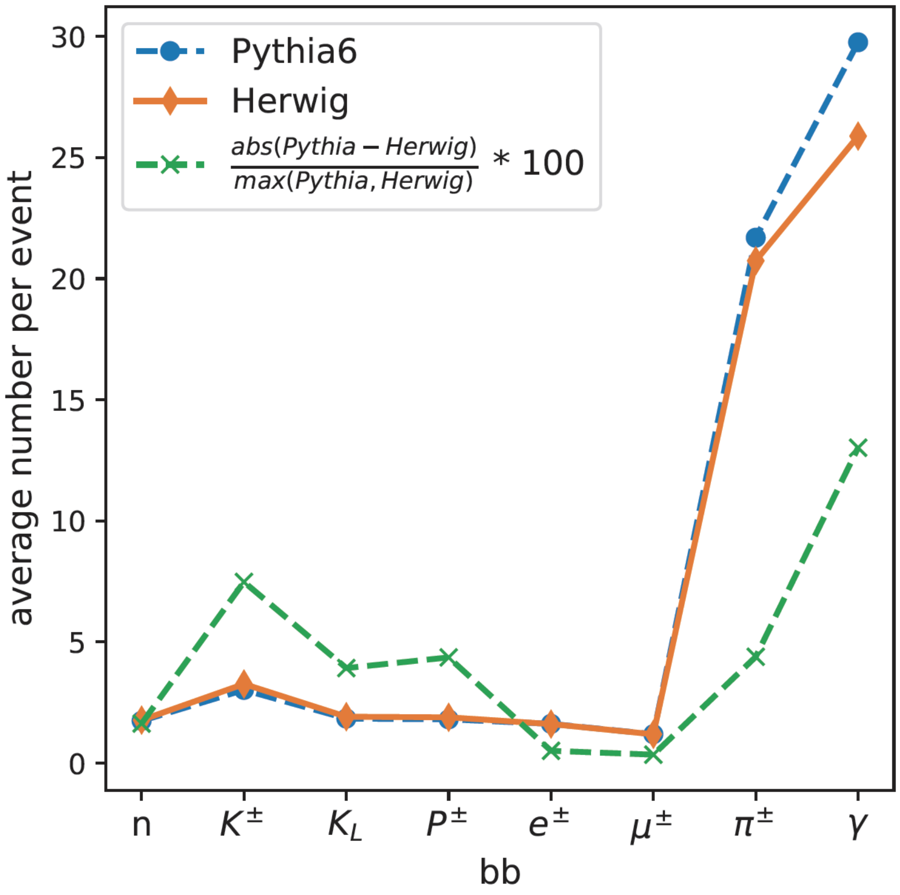




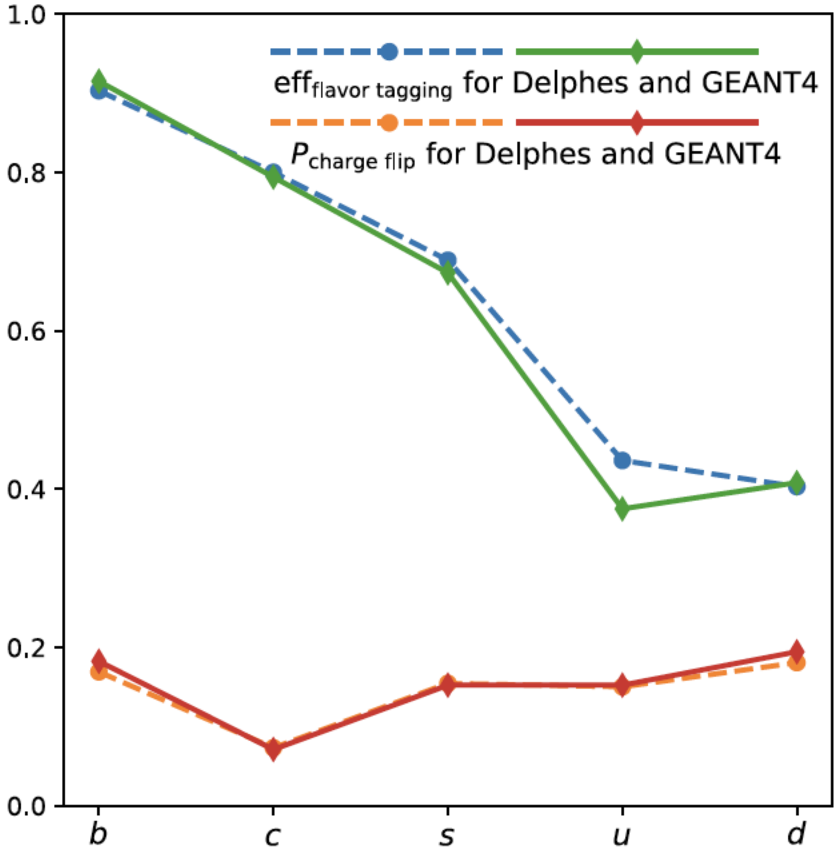
- *M10 instead of M11*

V.S. Hadronization models

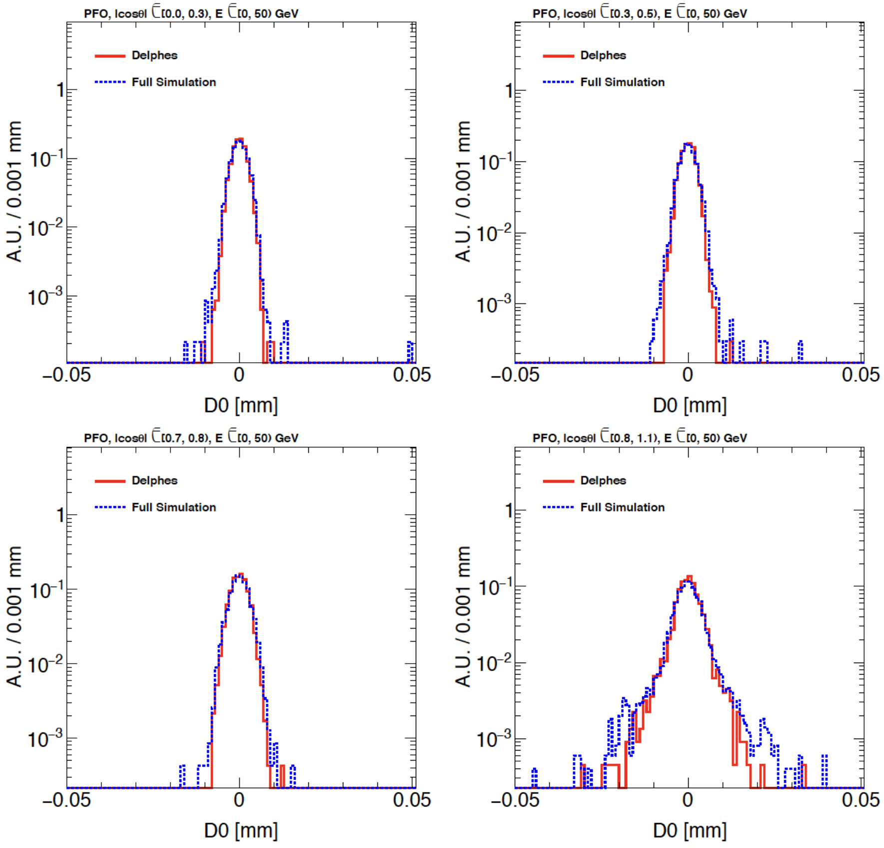




Fast/Full Simulation

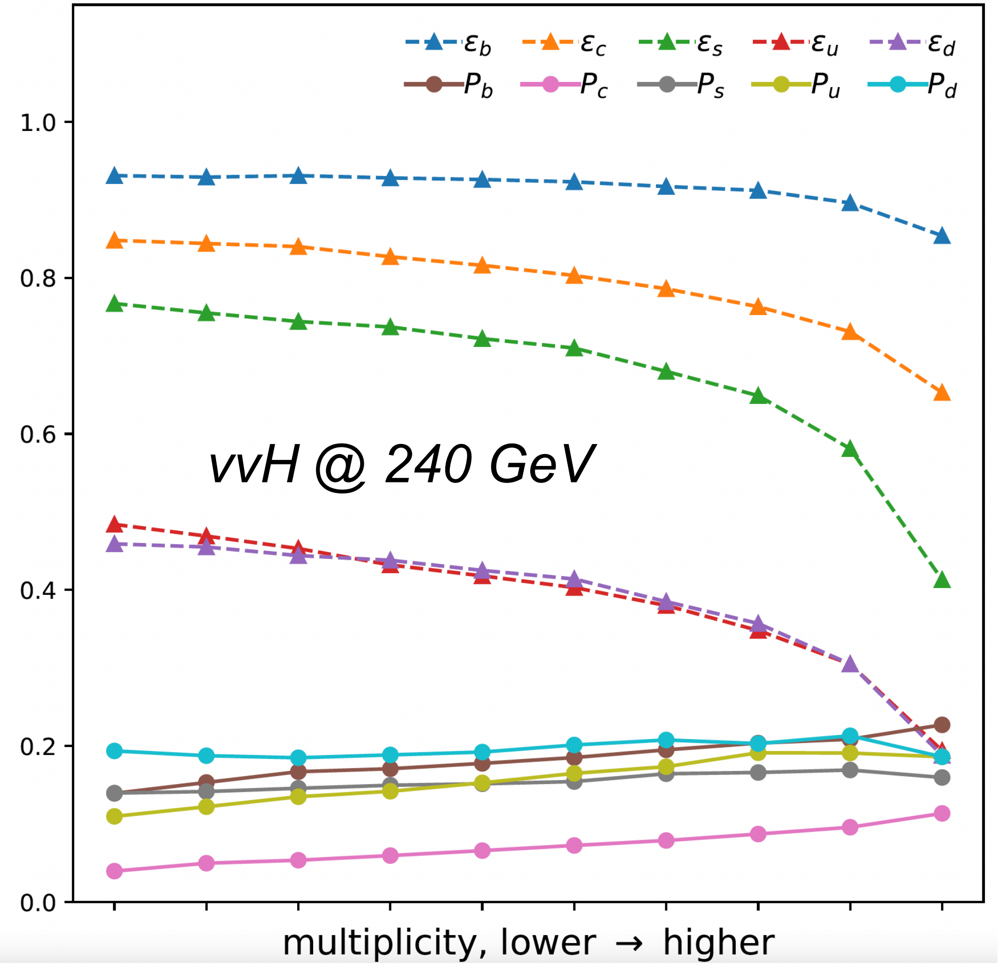


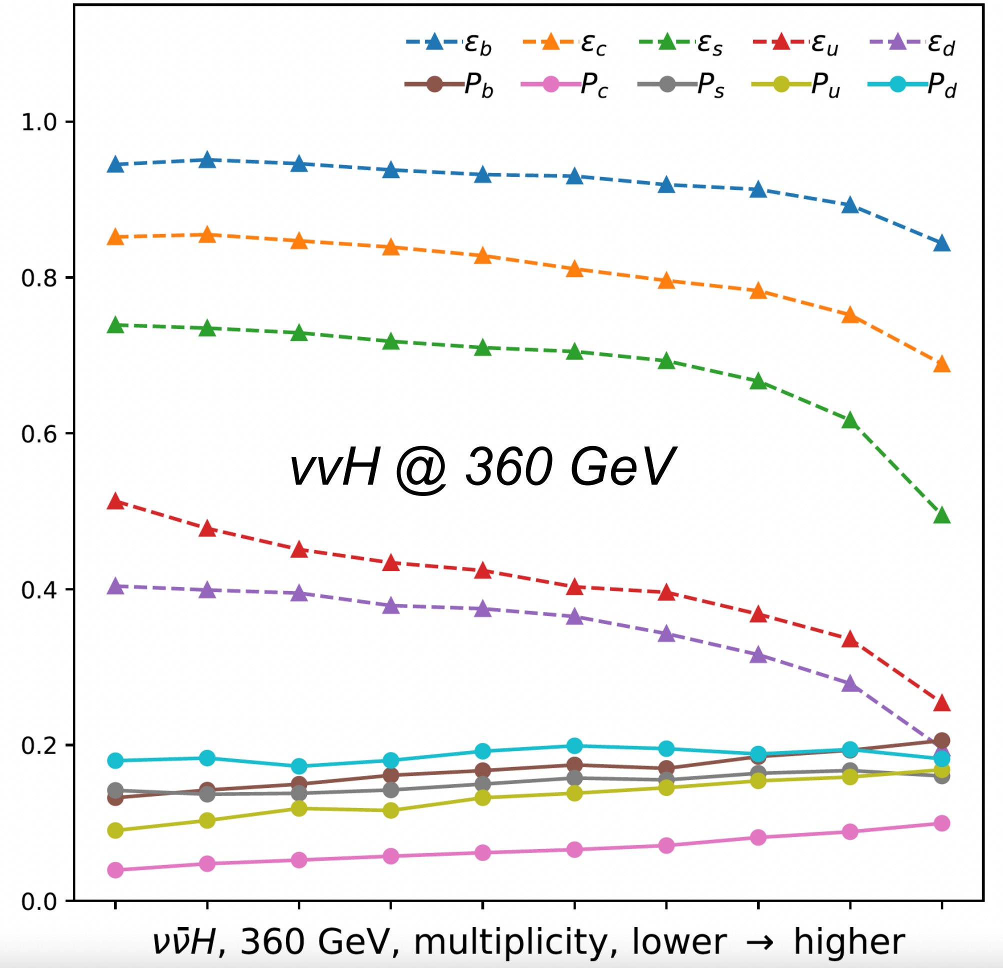
Z->μμ (91.2 GeV)



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

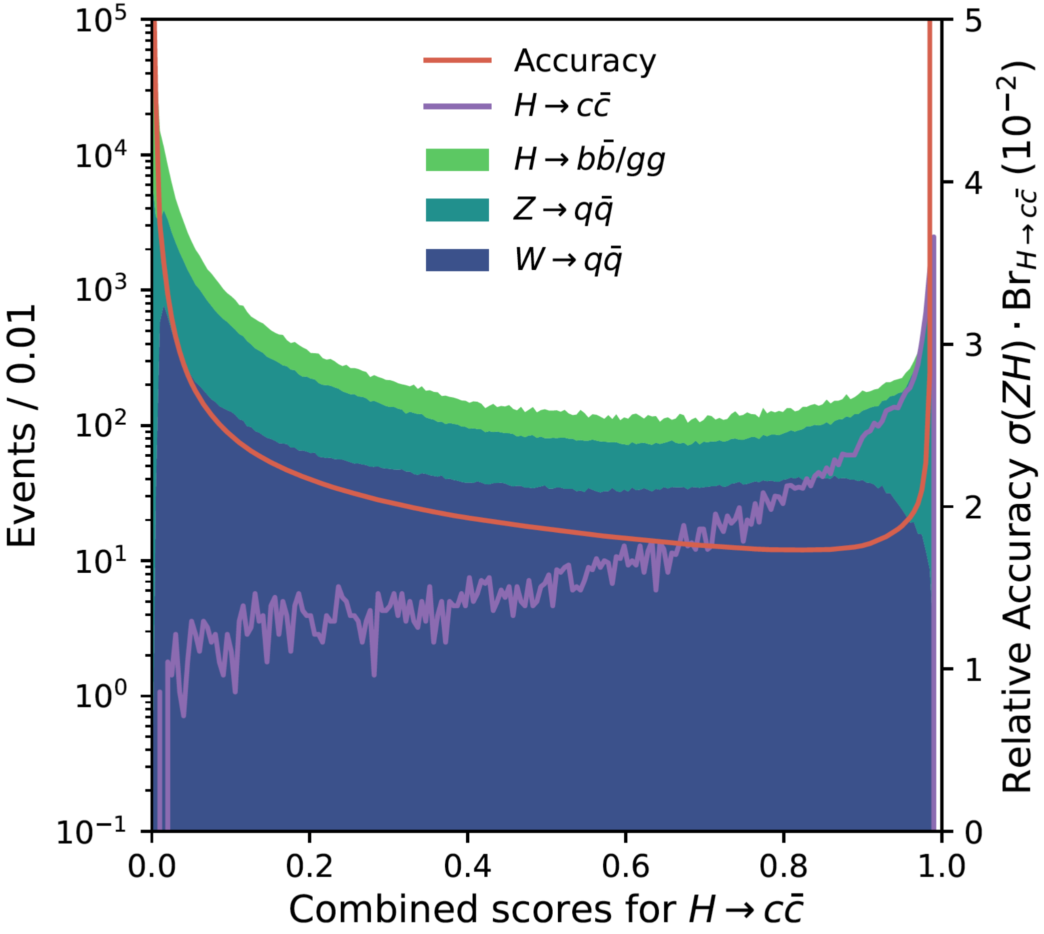
V.S. Multiplicity

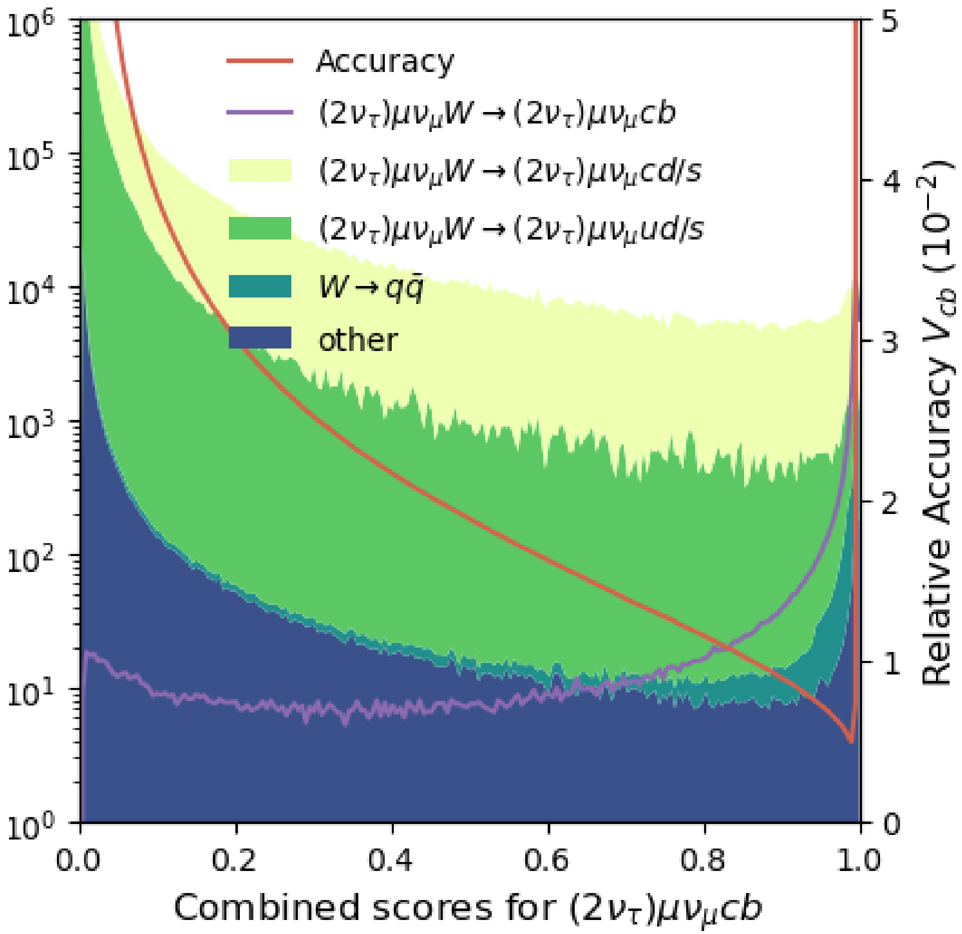




- *...many patterns need further understanding & towards further optimization...*

Recent update at more benchmarks





- From Jet Flavor Tagging to Jet Origin ID (**Preliminary**):
  - vvH, H→cc: 3% → 1.7%
  - Vcb: 0.75% → 0.5%

Summary

- Jet origin id: efficiently separate different species of colored SM particle
  - Stable & Smooth V.S.
    - Jet kinematic & Physics Processes: Calibration
    - Hadronization models: tools for QCD
    - Det. Geometry, Fast & Full Sim: reference for det. Optimization

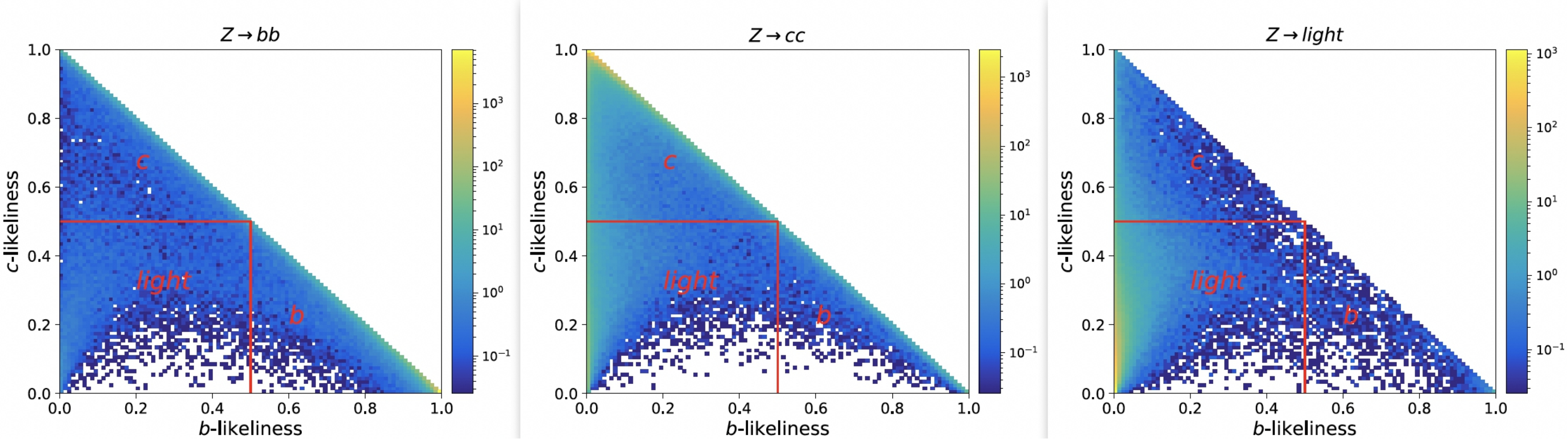
- Significantly impact on physics
  - Boost the access to g(Hss) and Higgs exotic/FCNC with jet final state (3 – 100 times)
  - **Preliminary**: Improve the H→cc/Vcb precision by ~ 100%/50%
  - Weak mixing angle
    - Measure to 1E-6 level precision (at 92 GeV) using 1 month data taking.
    - Access Afb_f with different flavors.
    - Verify RG behavior of Weak mixing angle at different c.m.s. Energy
  - *Time dependent CP measurements...*

- Long term version: identify jet origin... as we identify final state particle.

Back up

Comparison to Conventional Algo.

Three categories: b, c, & light



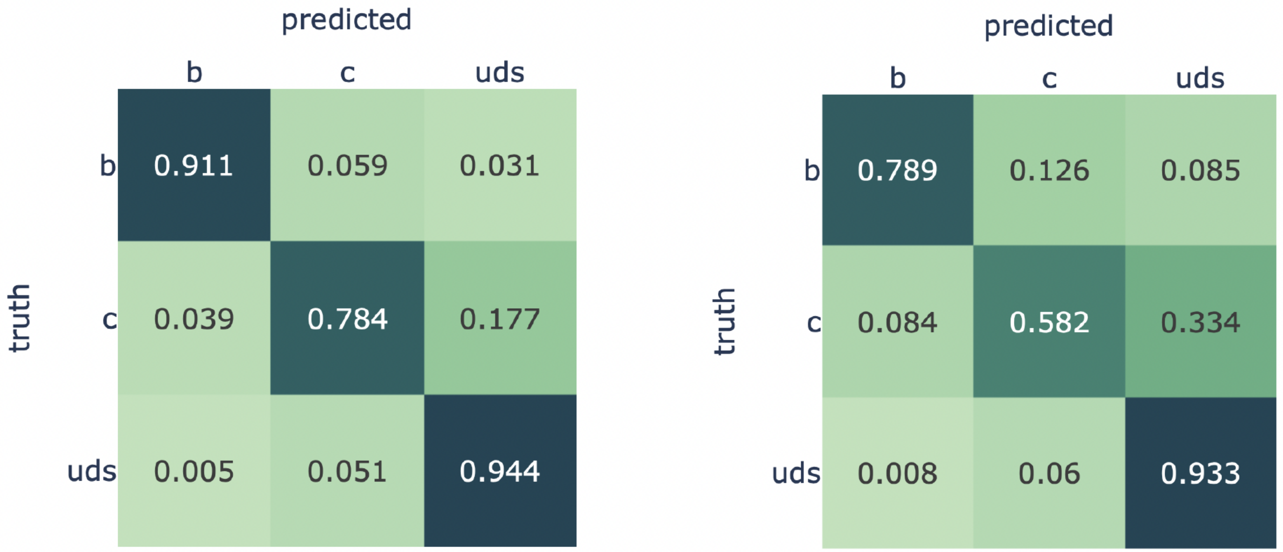
Hadronic Z pole sample

1 M Z→bb, cc, (uds) each

60/20/20% for

training/validating/testing.

Result on Testing sample

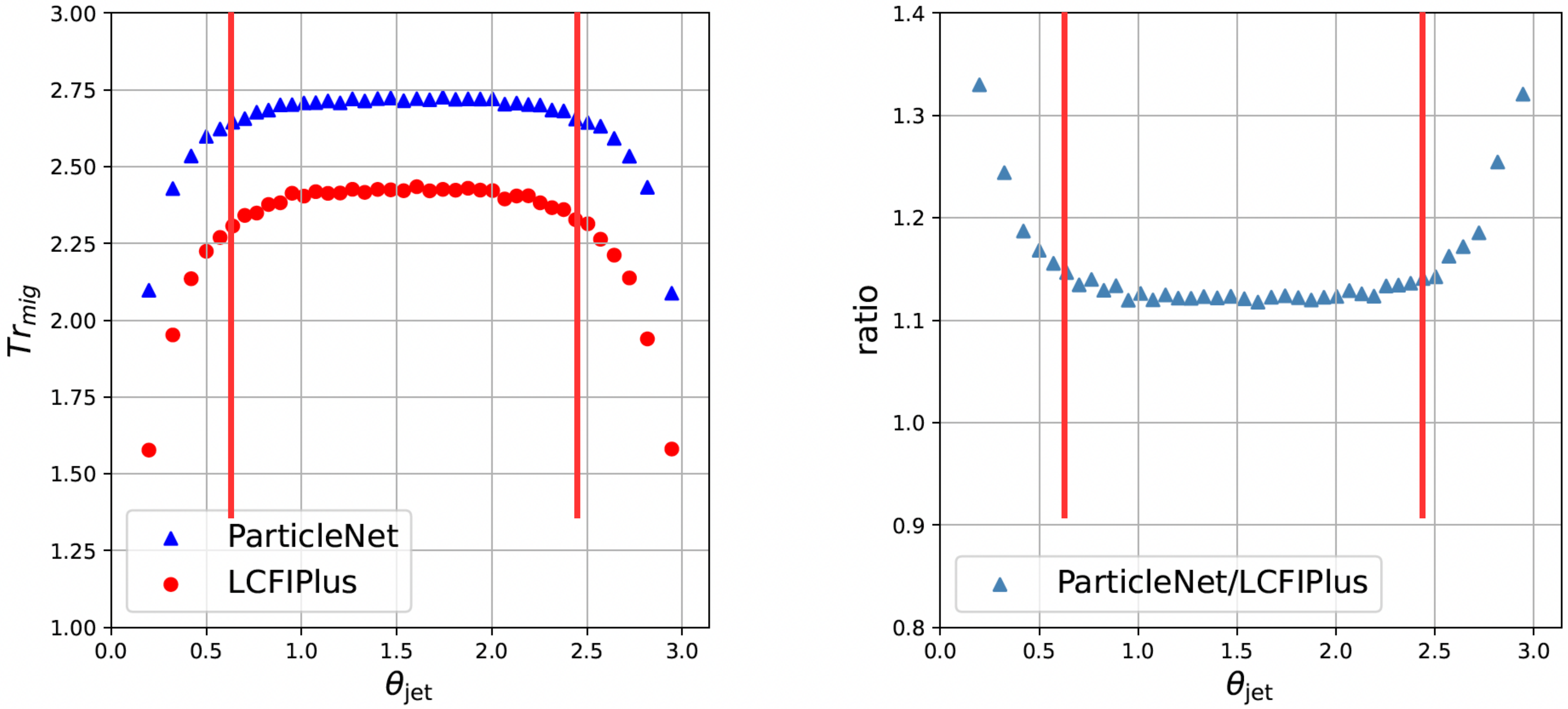


truth \ predicted	b	c	uds
b	0.911	0.059	0.031
c	0.039	0.784	0.177
uds	0.005	0.051	0.944

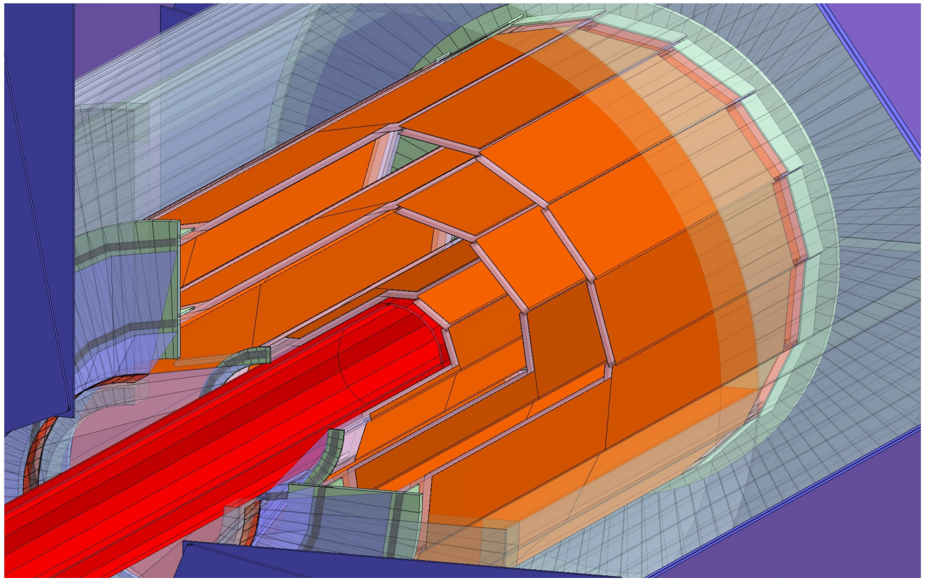
truth \ predicted	b	c	uds
b	0.789	0.126	0.085
c	0.084	0.582	0.334
uds	0.008	0.06	0.933

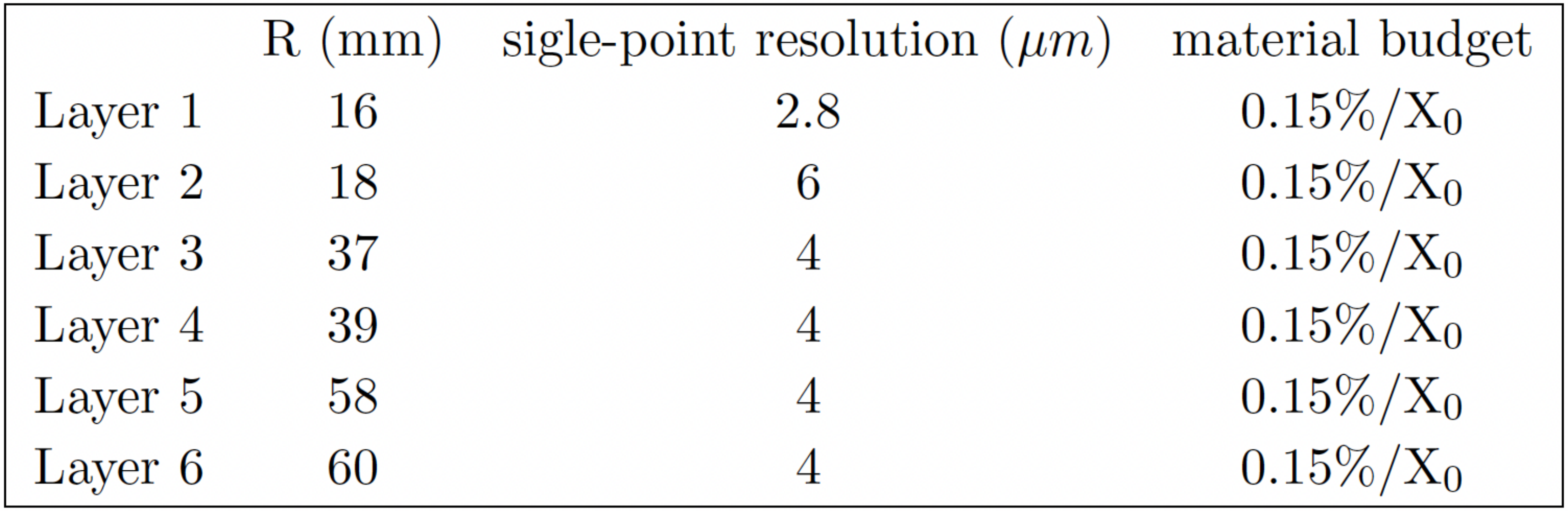
Figure 7. The migration matrix of ParticleNet (left) and LCFIPlus (right) at the CEPC.

Dependence on polar angle



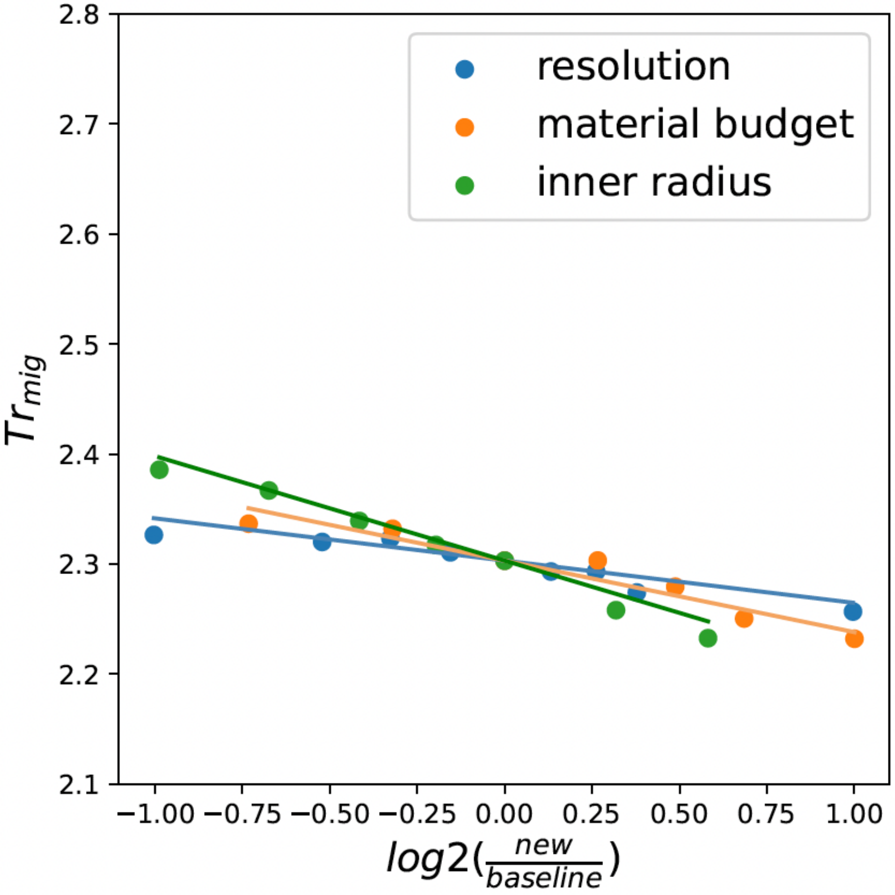
Comparison on Det. Optimization

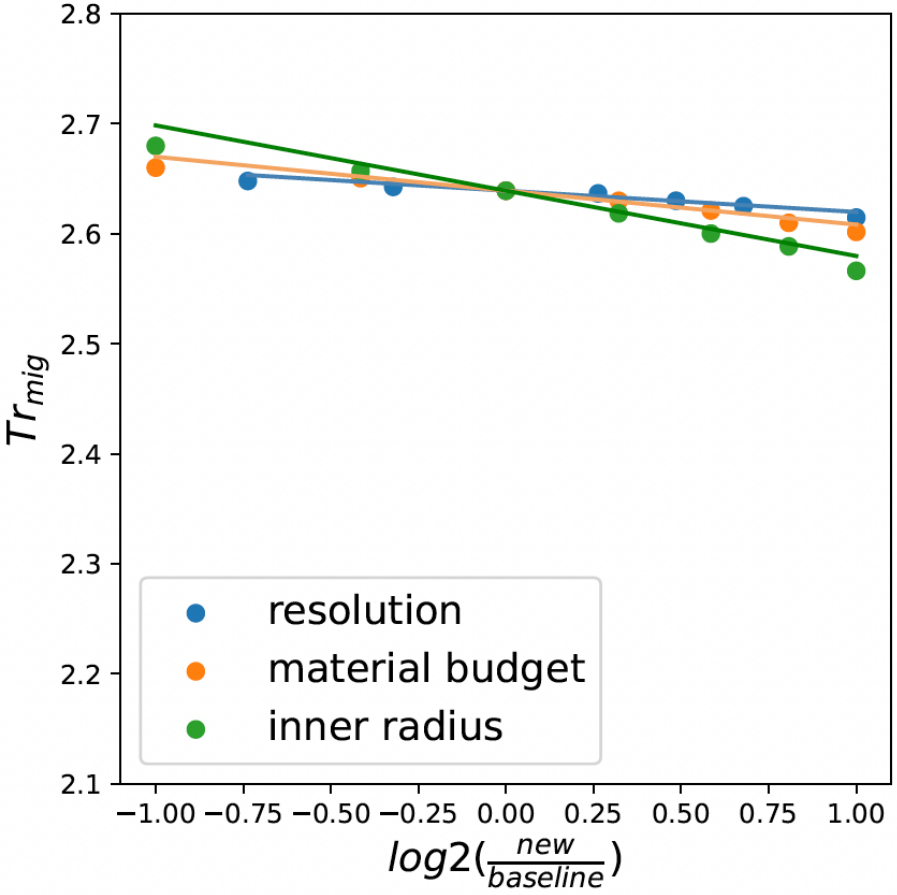




	R (mm)	sigle-point resolution (μm)	material budget
Layer 1	16	2.8	$0.15\%/X_0$
Layer 2	18	6	$0.15\%/X_0$
Layer 3	37	4	$0.15\%/X_0$
Layer 4	39	4	$0.15\%/X_0$
Layer 5	58	4	$0.15\%/X_0$
Layer 6	60	4	$0.15\%/X_0$

Comparison on Det. Optimization

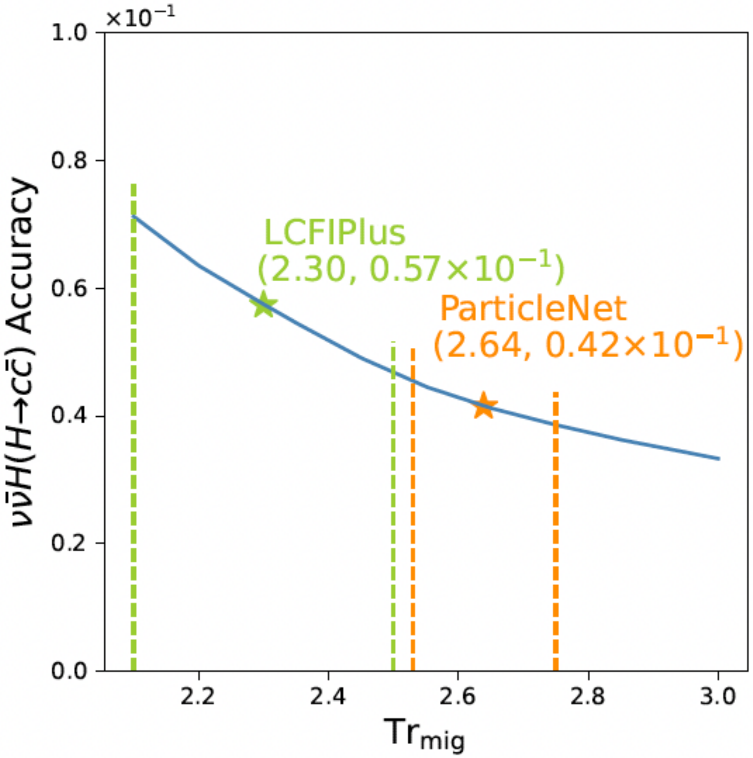


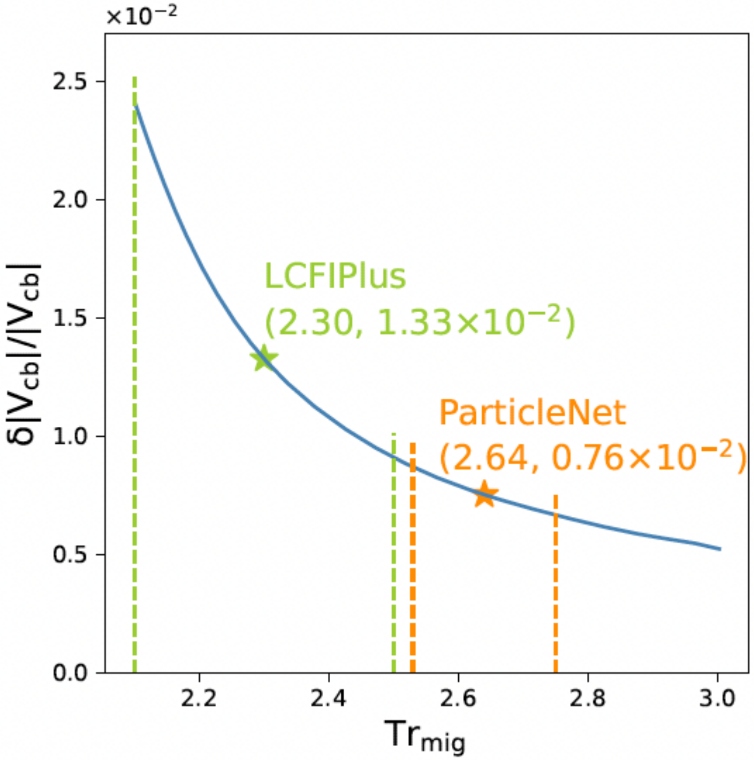


$$Tr_{mig} = 2.30 + 0.06 \cdot log_2 \frac{R^0_{material}}{R_{material}} + 0.04 \cdot log_2 \frac{R^0_{resolution}}{R_{resolution}} + 0.10 \cdot log_2 \frac{R^0_{radius}}{R_{radius}} \qquad (4.1)$$

$$Tr_{mig} = 2.64 + 0.03 \cdot log_2 \frac{R^0_{material}}{R_{material}} + 0.02 \cdot log_2 \frac{R^0_{resolution}}{R_{resolution}} + 0.06 \cdot log_2 \frac{R^0_{radius}}{R_{radius}} \qquad (4.2)$$

Impact on physics benchmarks





Conservative/Aggressive:

all three parameters 2/0.5*Baseline

		conservative	baseline	optimal
$\nu\nu Hc\bar{c}$	LCFIPlus	0.071	0.057	0.047
	ParticleNet	0.045	0.042	0.038
	$\frac{\text{LCFIPlus}}{\text{ParticleNet}}$	1.58	1.38	1.26
$\lvert V_{cb}\rvert$	LCFIPlus	0.0241	0.0133	0.0091
	ParticleNet	0.0086	0.0076	0.0067
	$\frac{\text{LCFIPlus}}{\text{ParticleNet}}$	2.80	1.75	1.36