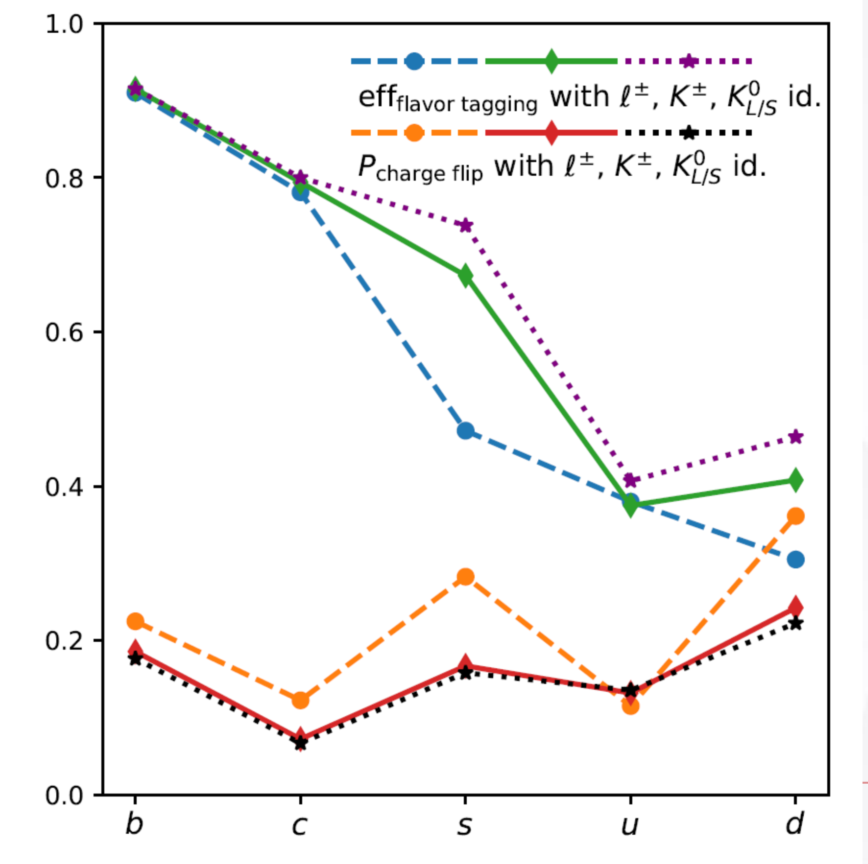
JOI on vertex detector optimization

Jialin Li

2025/8/20

Introduction

- Jet Origin Identification(JOI): determine from which colored particle a jet is generated.
  - 11 classes: b, $\bar{b}$, c, $\bar{c}$, s, $\bar{s}$, u, $\bar{u}$, d, $\bar{d}$, Gluon
- Test the performance of JOI and the measurement of H->CC/SS in different CEPC vertex detector parameters.



Setup

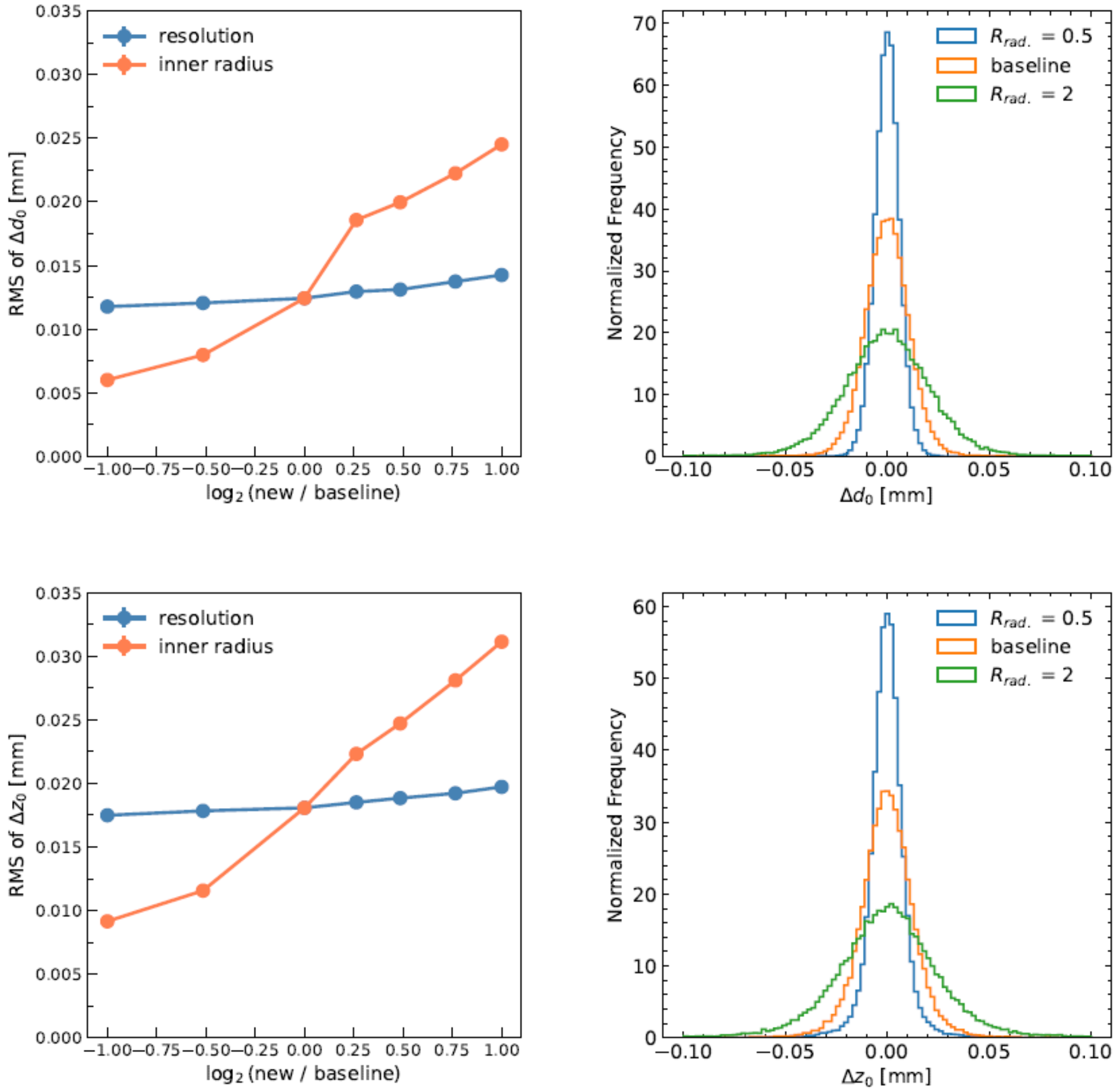
- Baseline: CEPC_CDR
- Generator: Whizard + Pythia6 + Fast simulation(Delphes)
- Training: ParN
- Parameters: inner radius(**R**) and resolution(**δ**) range [0.5, 2] vs baseline

$$R_{rad} = \frac{R_{inner}}{R_{baseline}}, \quad R_{res} = \frac{\delta_{inner}}{\delta_{baseline}}$$

	R (mm)	$\lvert z \rvert$ (mm)	$\lvert \cos\theta \rvert$	σ (μm)
Layer 1	16	62.5	0.97	2.8
Layer 2	18	62.5	0.96	6
Layer 3	37	125.0	0.96	4
Layer 4	39	125.0	0.95	4
Layer 5	58	125.0	0.91	4
Layer 6	60	125.0	0.90	4

Effect on impact parameter

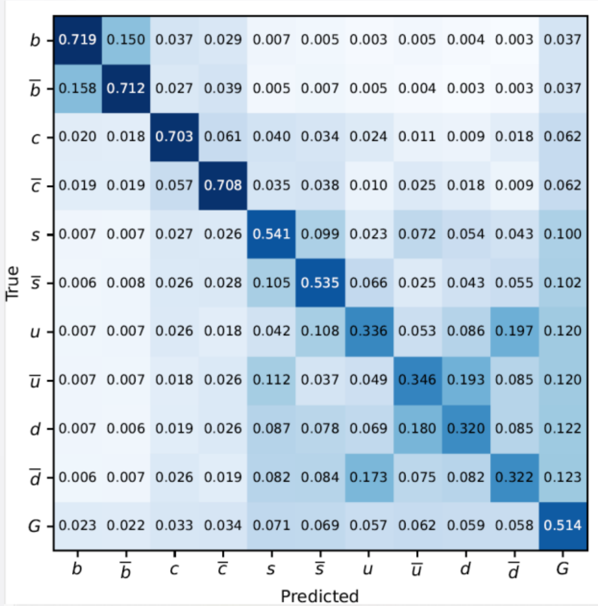
- Both d_0 and z_0 resolutions increase linearly with smaller inner radius and worse spatial resolution.
- Inner radius has a much stronger impact than resolution.
- Distributions show narrower peaks for smaller R_rad(right panel), confirming improved tracking precision.



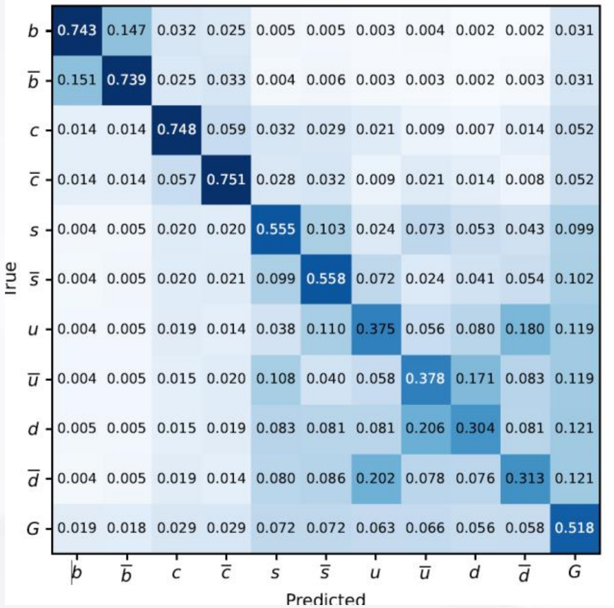
JOI Confusion matrix of Rrad

- 0.5ir : 0.5*baseline inner radius
- Trace: 5.76, 5.98, 6.20, 6.35

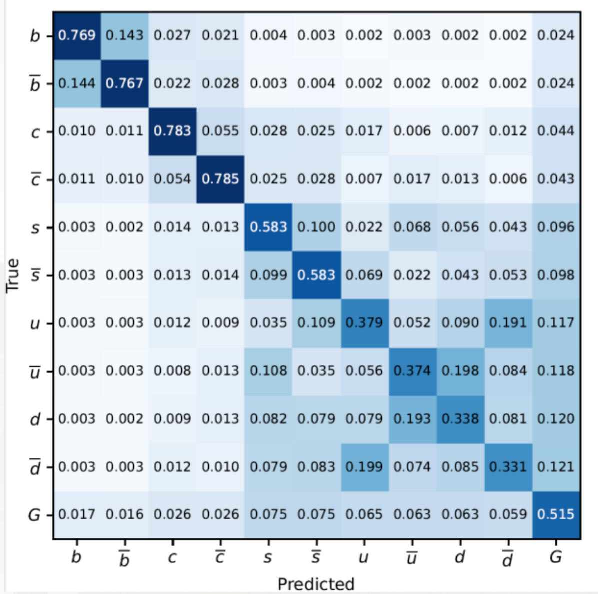
2ir



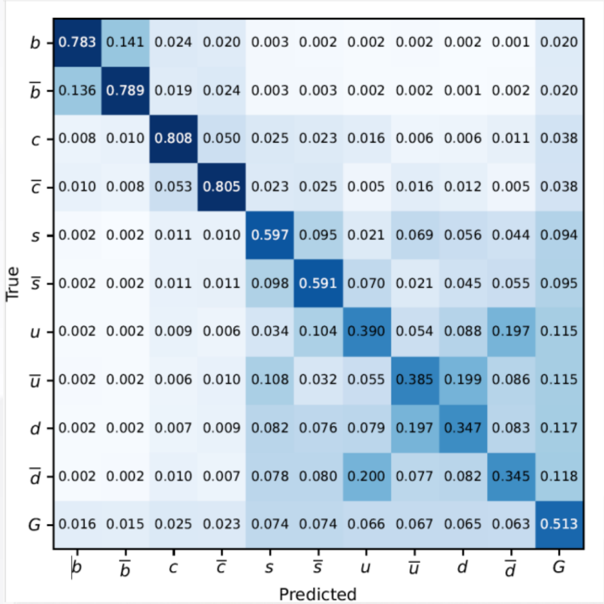
Baseline



0.7ir



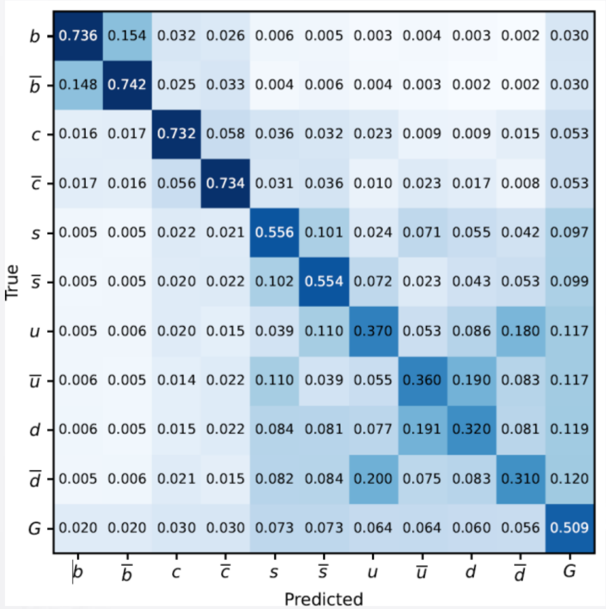
0.5ir



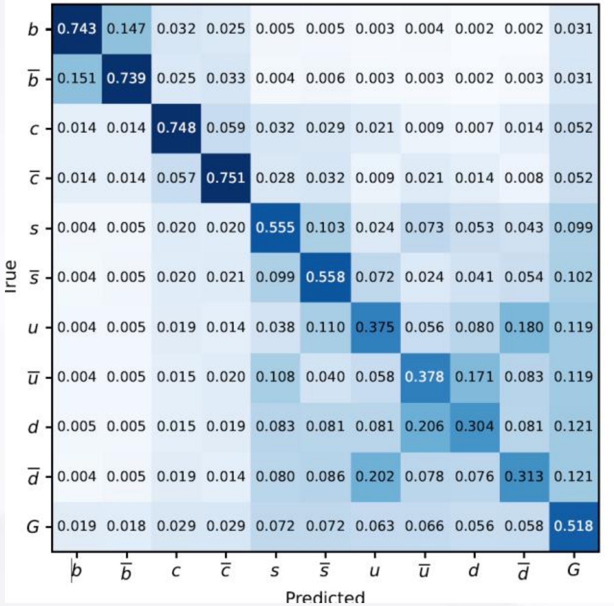
JOI Confusion matrix of Res

- Trace: 5.92, 5.98, 6.00, 6.01

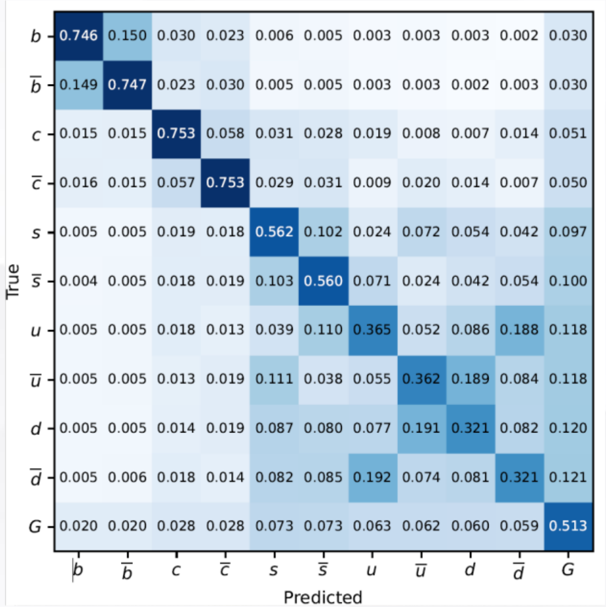
2vtx



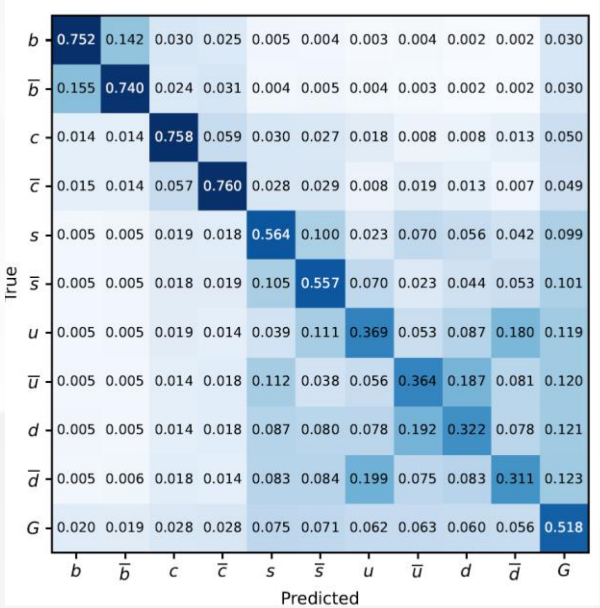
Baseline



0.7 vtx

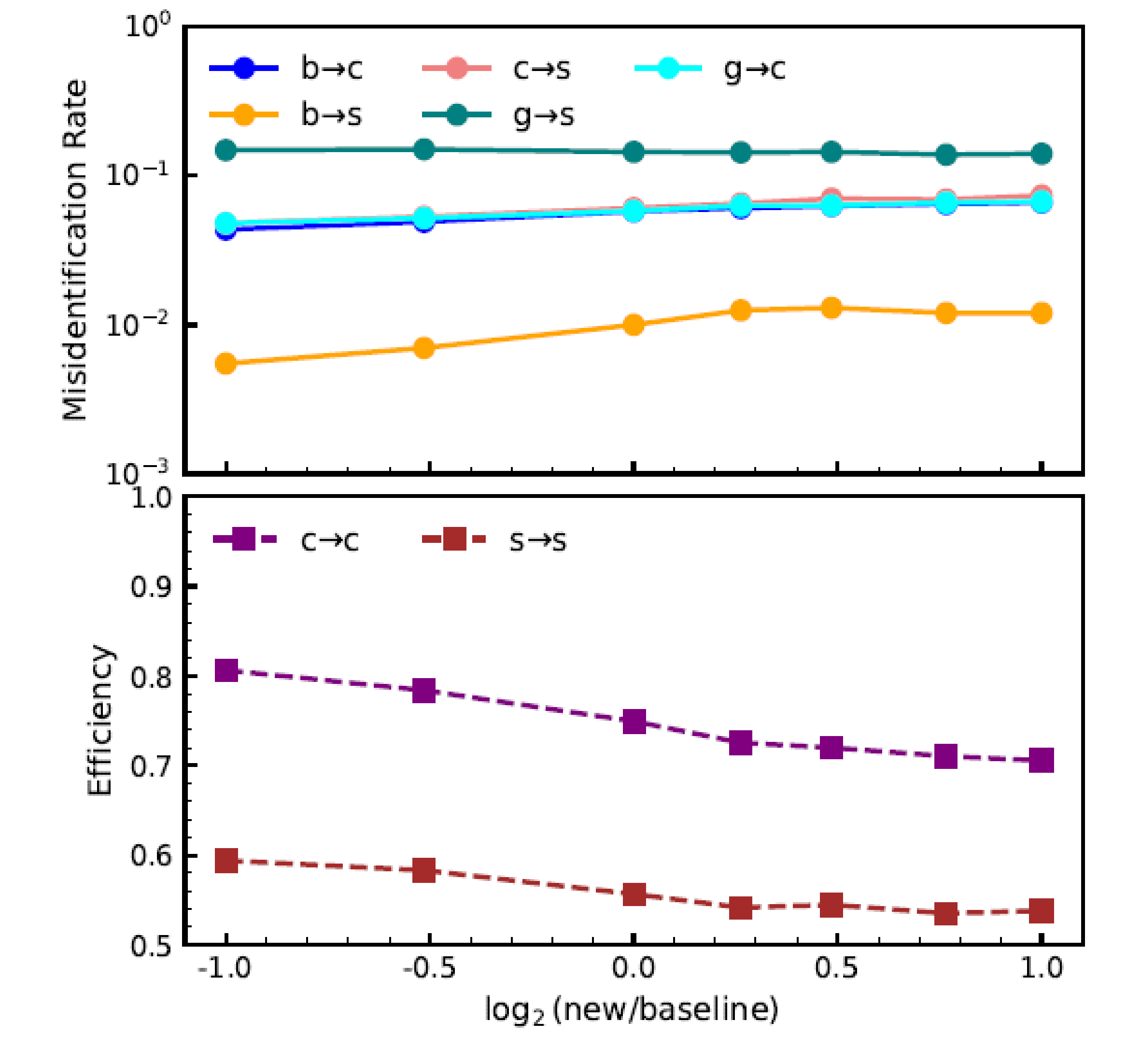


0.5 vtx



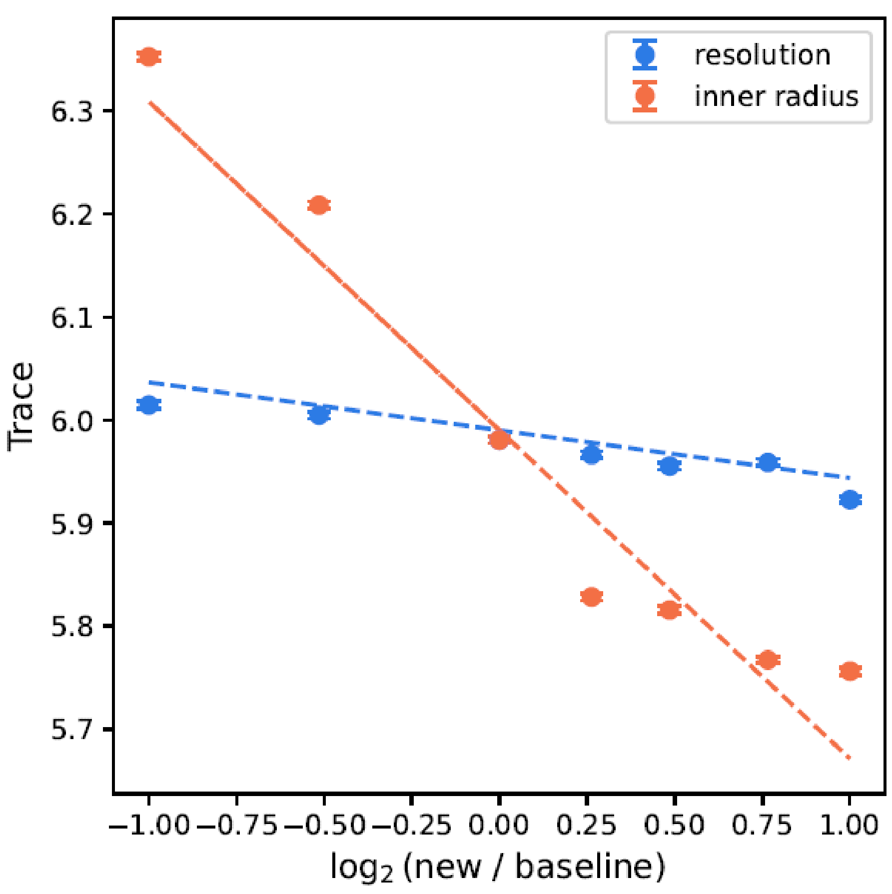
Tagging Efficiency and mis-id rate

- Misidentification rates (top):
  - b→s increases significantly with larger inner radius;
  - g→s stay stable. Gluon originate from the primary vertex; the misidentification is unaffected clear.
- Efficiency (bottom):
  - c-jet and s-jet tagging efficiencies drop as radius increases.



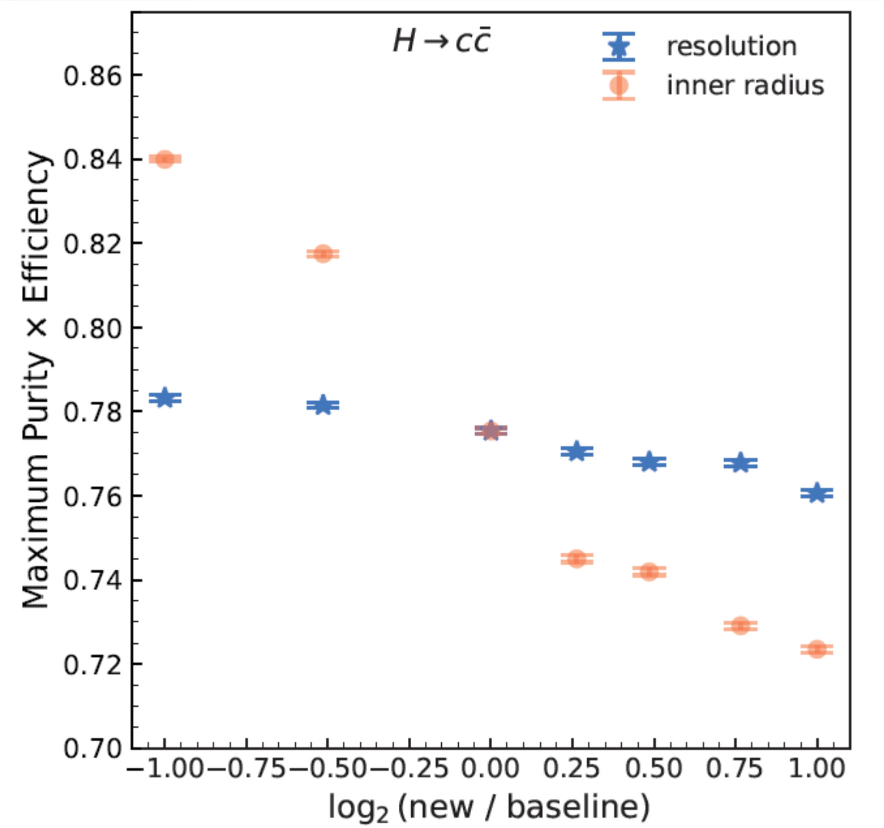
Trace of JOI

- vtx: 5.98 -0.05 * $\log2(\frac{new}{baseline})$
- ir: 5.98 -0.27 * $\log2(\frac{new}{baseline})$
- Orange line drops sharply with increasing R_{rad} confirming radius dominance. (5 times of R_{res})



Measurement of H->CC

- Inner radius has a greater impact than spatial resolution.
- The best configuration improves the branching ratio uncertainty by **~3.8%**, while the worst worsens it by **~3.5%** compared to baseline.

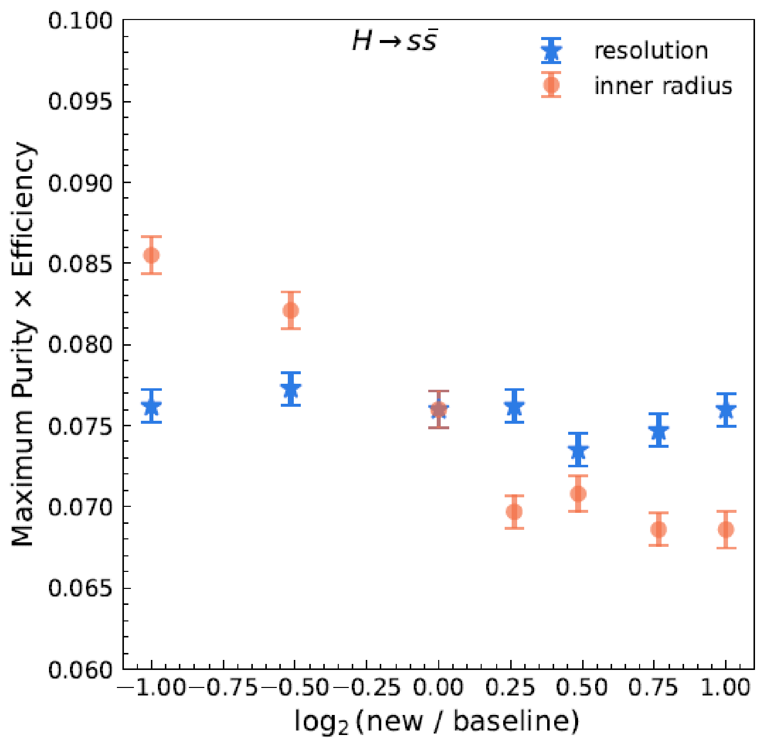


$$\frac{\delta \mathcal{B}}{\mathcal{B}} = \frac{\sqrt{s+b}}{s}$$

Configuration	$c\bar{c}\,(10^3)$	$b\bar{b}\,(10^3)$	$s\bar{s}$	$gg\,(10^3)$	Uncertainty(10^{-3})
baseline	24.4	1.26	11.0	1.92	6.81
$R_{\rm res} = 0.5$	24.5	1.21	8.95	1.85	6.78
$R_{\rm res} = 2$	24.0	1.21	12.9	1.99	6.88
$R_{\rm rad} = 0.5$	25.6	0.75	5.54	1.75	6.55
$R_{\rm rad} = 2$	23.2	1.54	1.61	2.04	7.05

Measurement of H->SS

- Backgrounds (especially gg) dominate.
- Significance reaches 4.76σ with Rrad=0.5
- The best configuration enhances the significance by **~7.7%**, while the worst reduces it by **~4.3%** relative to baseline.

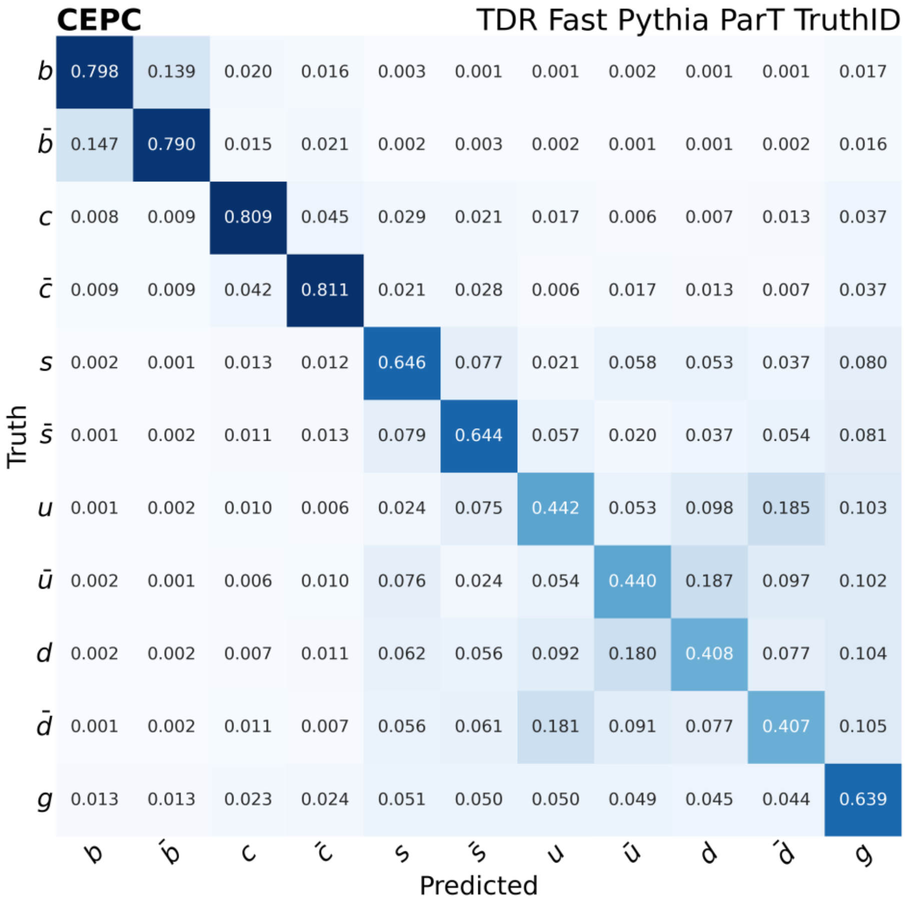


$$S = \sqrt{2\left(\left(s+b\right)\ln\left(1+\frac{s}{b}\right)-s\right)}$$

Configuration	$s\bar{s}$	$b\bar{b}$	$c\bar{c}$	gg	Significance
baseline	70.9	0	10.0	224.6	4.42 σ
$R_{\text{res}} = 0.5$	64.1	0	7.6	179.4	4.45 σ
$R_{\text{res}} = 2$	64.4	0	10.3	181.4	4.41 σ
$R_{\text{rad}} = 0.5$	70.2	0	6.2	189.4	4.76 σ
$R_{\text{rad}} = 2$	64.8	0	11.5	203.2	4.23 σ

Best result @kaili

- Setup: TDR+Fast+P8+ParT
- Trace is 6.834
- 14% improvement vs. my baseline and 7% vs. my best result(0.5ir)



Truth \ Predicted	b	b̄	c	c̄	s	s̄	u	ū	d	d̄	g
b	0.798	0.139	0.020	0.016	0.003	0.001	0.001	0.002	0.001	0.001	0.017
b̄	0.147	0.790	0.015	0.021	0.002	0.003	0.002	0.001	0.001	0.002	0.016
c	0.008	0.009	0.809	0.045	0.029	0.021	0.017	0.006	0.007	0.013	0.037
c̄	0.009	0.009	0.042	0.811	0.021	0.028	0.006	0.017	0.013	0.007	0.037
s	0.002	0.001	0.013	0.012	0.646	0.077	0.021	0.058	0.053	0.037	0.080
s̄	0.001	0.002	0.011	0.013	0.079	0.644	0.057	0.020	0.037	0.054	0.081
u	0.001	0.002	0.010	0.006	0.024	0.075	0.442	0.053	0.098	0.185	0.103
ū	0.002	0.001	0.006	0.010	0.076	0.024	0.054	0.440	0.187	0.097	0.102
d	0.002	0.002	0.007	0.011	0.062	0.056	0.092	0.180	0.408	0.077	0.104
d̄	0.001	0.002	0.011	0.007	0.056	0.061	0.181	0.091	0.077	0.407	0.105
g	0.013	0.013	0.023	0.024	0.051	0.050	0.050	0.049	0.045	0.044	0.639

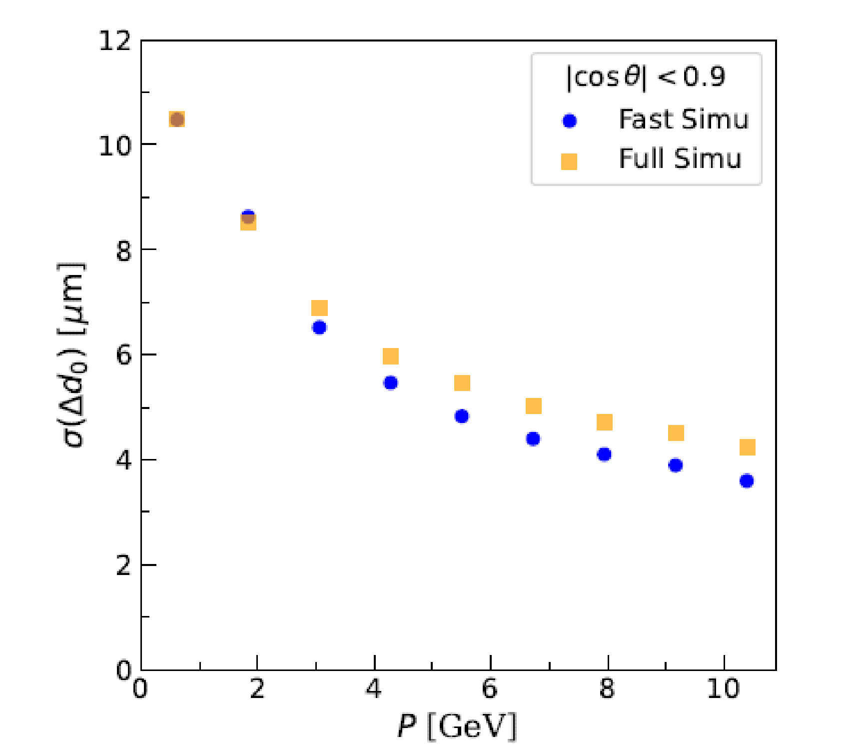
Conclusion and Next plan (submitted to JHEP)

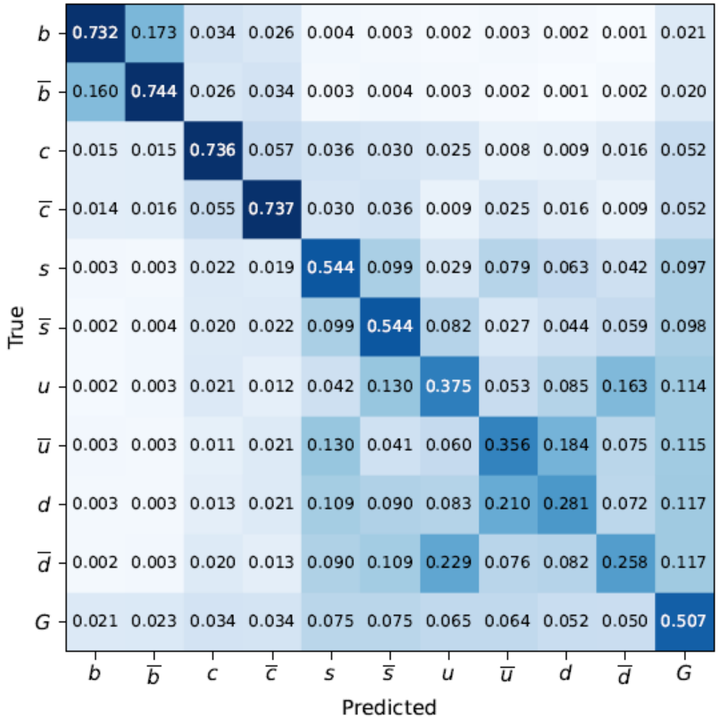
- Vertex detector parameters, especially inner radius, play a critical role in flavor tagging performance.
- For H→ $c\bar{c}$: Best setup improves branching ratio uncertainty by ~3.8%,Worst case worsens it by ~3.5%, compared to baseline.
- For H→ $s\bar{s}$: Best configuration yields ~7.7% gain in significance,Worst case results in ~4.3% loss.
- 4.76σ for H→ $s\bar{s}$ brings CEPC closer to discovering this rare Higgs decay mode.
- According to Kaili's latest results, the improvment of JOI suggests that H→$s\bar{s}$ measurement may achieve ≥5σ significance.

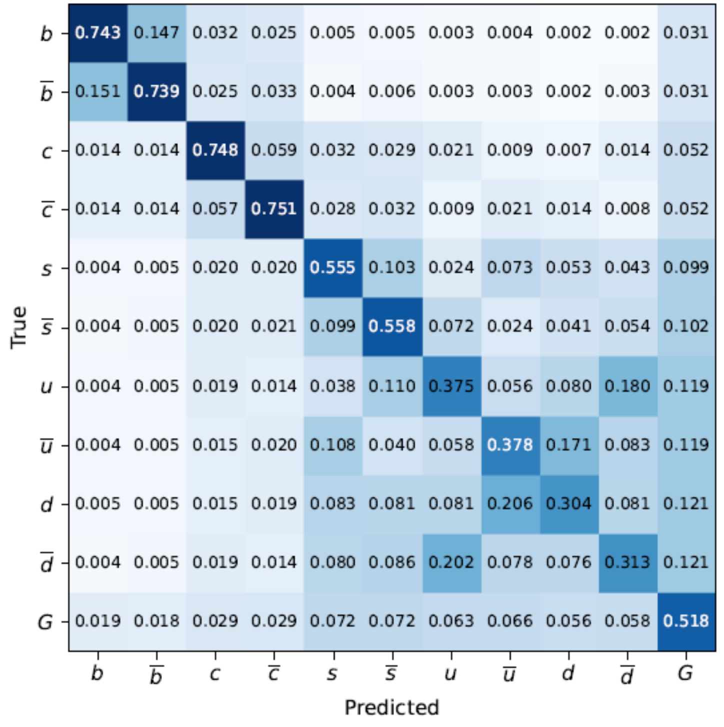
SJTU

Backup

Simulation Validation (Delphes vs. Geant4)







3*3 Full simu

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