# Simulation analyses of $H \rightarrow b\bar{b}/c\bar{c}/gg$ at CEPC & dependency on VTX design; Requirement on PID

Yongfeng Zhu

advisor : Manqi Ruan

CEPC Day, 2022.09.21

# Motivation :

- The measurement of $H \to b\bar{b}/c\bar{c}/gg$ signal strength accuracy is improtant for Higgs coupling studies.
- In the measurement of the relative accuracy of $H \to b\bar{b}/c\bar{c}/gg$ signal strength, the flavor tagging performance has significant impact on the final results.
- The vertex detector design has impact on flavor tagging performance.
- The PID performance is essential to flavor physics.

# The relative accuracy of $H \to b\bar{b}/c\bar{c}/gg$ signal strength.

CEPC baseline detector,
all SM samples corresponding to 240 GeV, 5.6 $ab^{-1}$,
full simulation,
$\ell^+\ell^- Higgs$,
$\nu\bar{\nu} Higgs$,
$q\bar{q} Higgs$,

The analysis procedure can be divided into two steps :
first step : select signal with cut flow
second step : calculate the signal strength accuracy

arXiv : 2203.01469

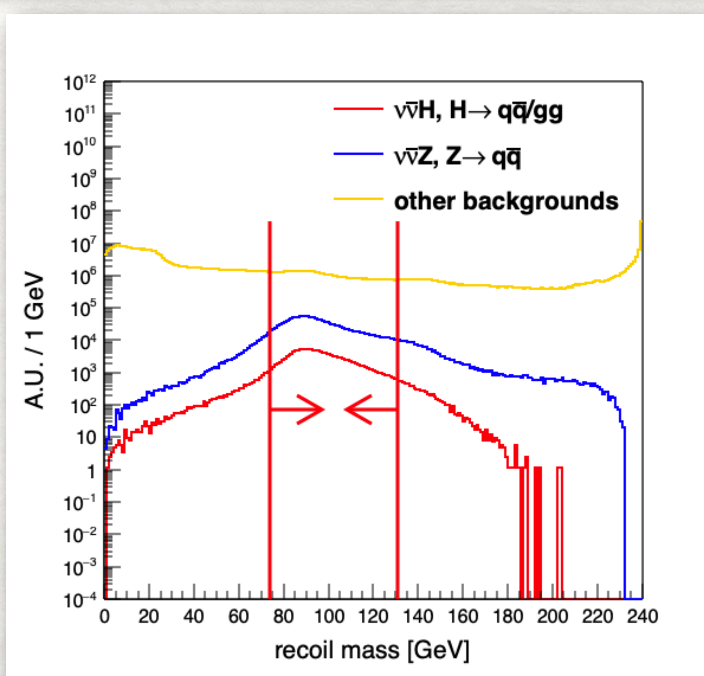
# $\nu\bar{\nu}Higgs$ selection

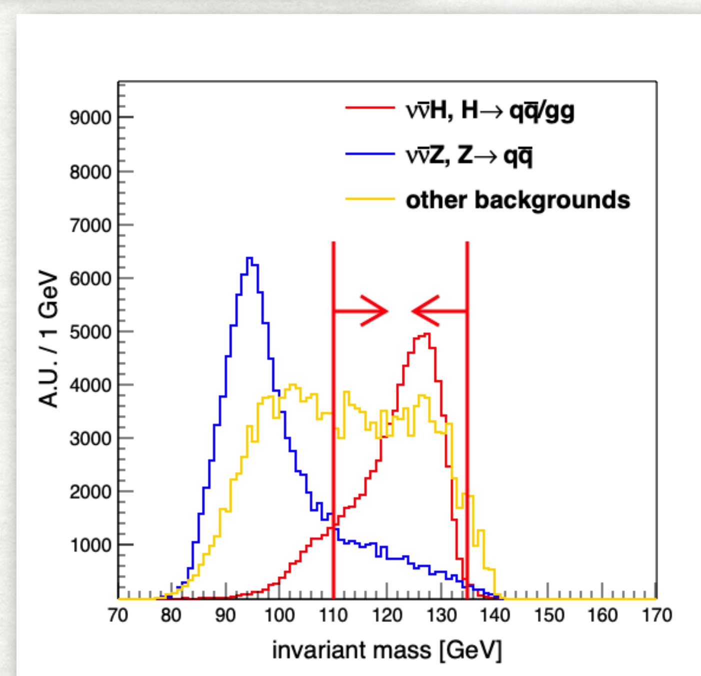
| | $\nu\bar{\nu}Hq\bar{q}/gg$ | 2f | SW | SZ | WW | ZZ | Mixed | ZH | $\frac{\sqrt{S+B}}{S}(\%)$ |
|---|---|---|---|---|---|---|---|---|---|
| total | 178890 | $8.01E8$ | $1.95E7$ | $9.07E6$ | $5.08E7$ | $6.39E6$ | $2.18E7$ | 961606 | 16.86 |
| recoilMass (GeV) $\in(74,131)$ | 157822 | $5.11E7$ | $2.17E6$ | $1.38E6$ | $4.78E6$ | $1.30E6$ | $1.08E6$ | 74991 | 4.99 |
| $visEn$ (GeV) $\in(109,143)$ | 142918 | $2.37E7$ | $1.35E6$ | $8.81E5$ | $3.60E6$ | $1.03E6$ | $6.29E5$ | 50989 | 3.92 |
| $leadLepEn$ (GeV) $\in(0,42)$ | 141926 | $2.08E7$ | $3.65E5$ | $7.24E5$ | $2.81E6$ | $9.72E5$ | $1.34E5$ | 46963 | 3.59 |
| $multiplicity$ $\in(40,130)$ | 139545 | $1.66E7$ | $2.36E5$ | $5.24E5$ | $2.62E6$ | $9.07E5$ | 4977 | 42751 | 3.29 |
| $leadNeuEn$ (GeV) $\in(0,41)$ | 138653 | $1.46E7$ | $2.24E5$ | $4.72E5$ | $2.49E6$ | $8.69E5$ | 4552 | 42303 | 3.12 |
| $Pt$ (GeV) $\in(20,60)$ | 121212 | 248715 | $1.56E5$ | $2.48E5$ | $1.51E6$ | $4.31E5$ | 999 | 35453 | 1.37 |
| $Pl$ (GeV) $\in(0,50)$ | 118109 | 52784 | $1.05E5$ | 74936 | $7.30E5$ | $1.13E5$ | 847 | 34279 | 0.94 |
| -log10(Y23) $\in(3.375,+\infty)$ | 96156 | 40861 | 26088 | 60349 | $2.25E5$ | 82560 | 640 | 10691 | 0.76 |
| InvMass (GeV) $\in(116,134)$ | 71758 | 22200 | 11059 | 6308 | 77912 | 13680 | 248 | 6915 | 0.64 |
| BDT $\in(-0.02,1)$ | 60887 | 9140 | 266 | 2521 | 3761 | 3916 | 58 | 1897 | 0.47 |

**Table 1.** The event selection of $\nu\bar{\nu}H(H\to q\bar{q}/gg)$.

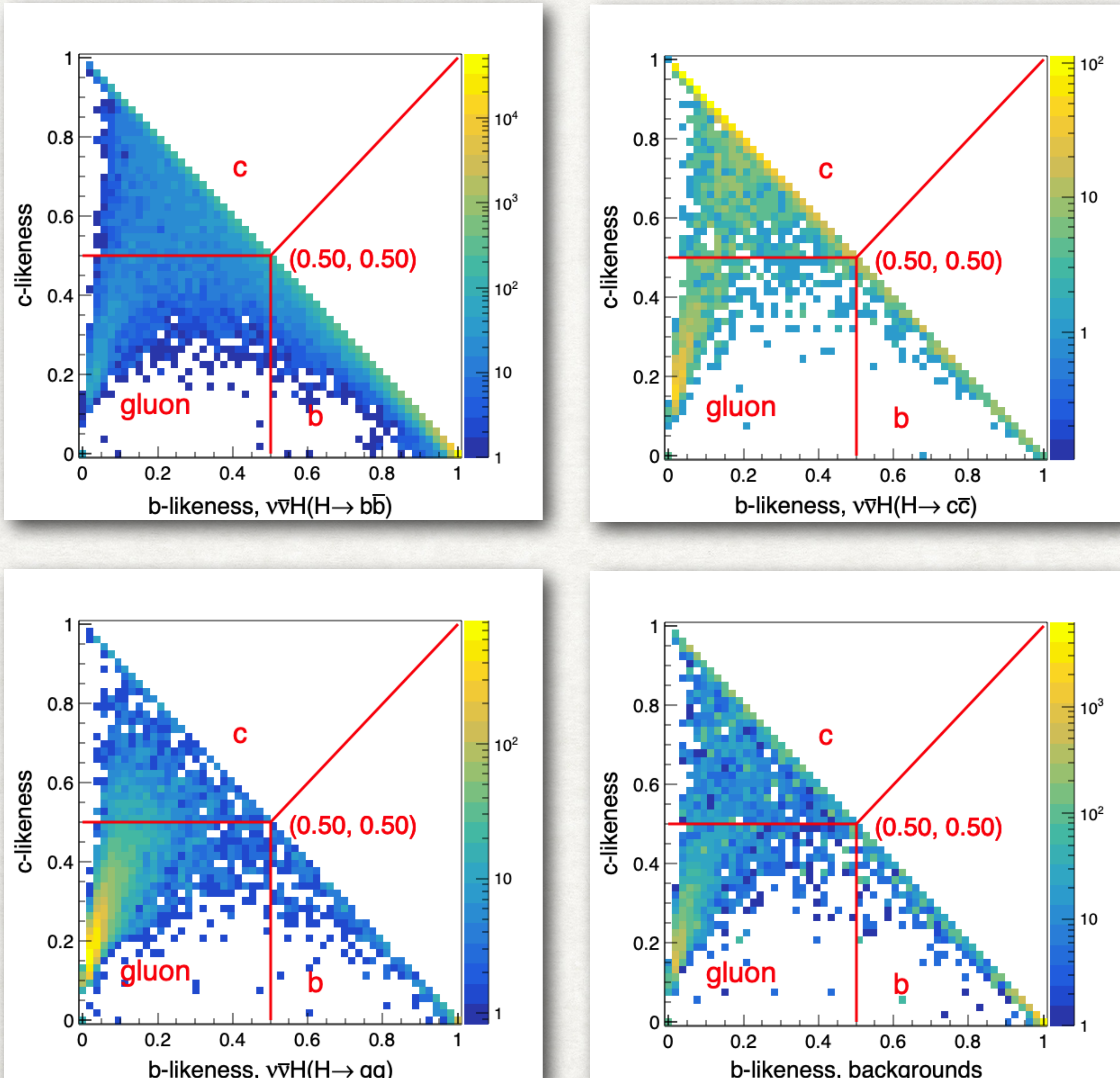
signal-to-background ratio

0.22
0.52
2.91



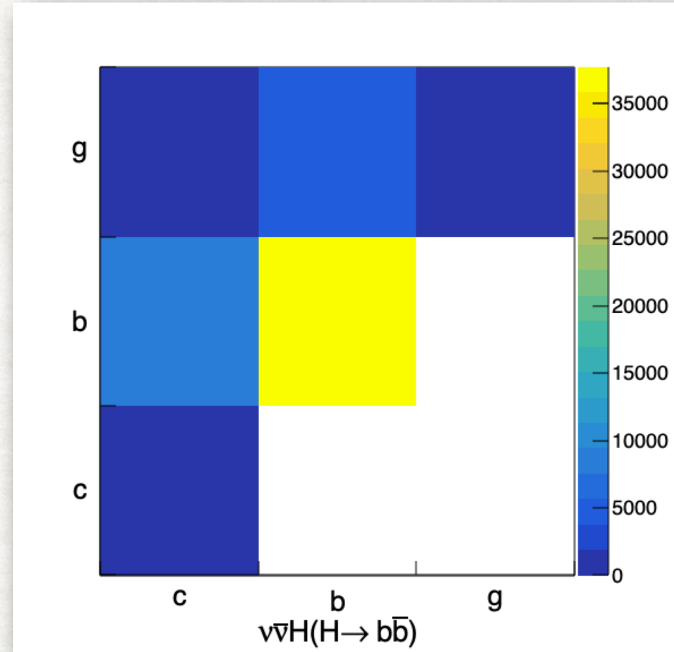



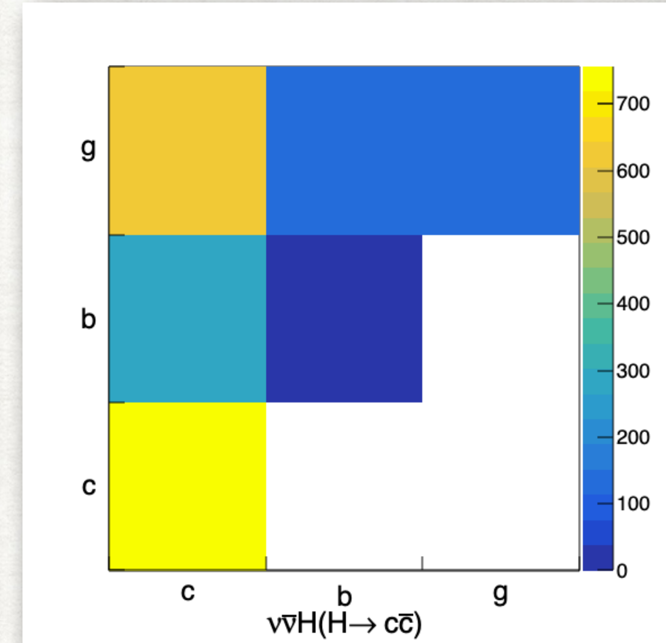
# the second step

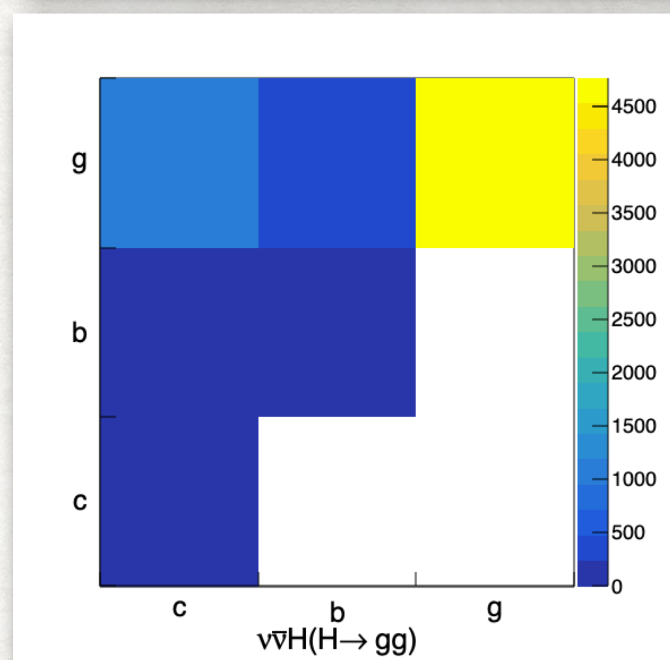


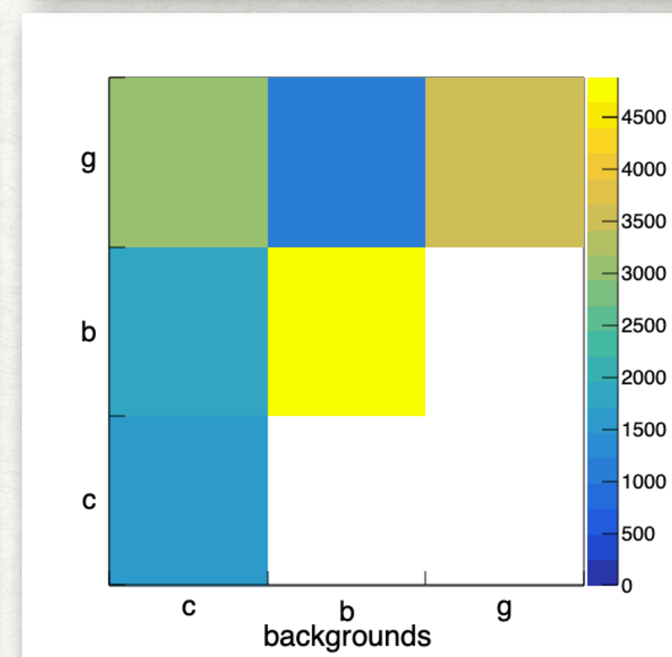

| true \ identified as | b | c | g |
|---|---|---|---|
| b | 0.8675 | 0.0887 | 0.0437 |
| c | 0.1136 | 0.6263 | 0.2601 |
| g | 0.0411 | 0.1007 | 0.8582 |

The identified flavor combinations based on flavor tagging performance matrix.




$$-2 \cdot log(\ell) = \sum_{i=1}^{i=6} \frac{[S_b \cdot N_{b,i} + S_c \cdot N_{c,i} + S_g \cdot N_{g,i} + N_{bkg,i} - N_i]^2}{N_i},$$

# results :

arXiv:1905.12903

| 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.

Extrapolate to CEPC Snowmass, FCC-ee and ILC settings.

| | $\sqrt{s}$ GeV | $\int L$ $ab^{-1}$ | polarization | $H \to b\bar{b}$ | $H \to c\bar{c}$ | $H \to gg$ |
|---|---|---|---|---|---|---|
| CEPC | 240 (CDR) | 5.6 | $\sim$ | 0.27% | 4.03% | 1.56% |
| CEPC | 240 (Snowmass) | 20 | $\sim$ | 0.14% | 2.13% | 0.82% |
| FCC-ee | 240 | 5 | $\sim$ | 0.29% | 4.27% | 1.64% |
| ILC | 250 | 2 | $(e^-, e^+) = (-0.8, +0.3)$ | 0.45% | 6.83% | 2.67% |

**Table 4.** The signal strength accuracy of $H \to b\bar{b}/c\bar{c}/gg$ at the CEPC and extrapolated to the ILC 250 $GeV$ ($2\,ab^{-1}$) and FCC-ee 240 $GeV$ ($5\,ab^{-1}$).

## the dependency of $H \to b\bar{b}/c\bar{c}/gg$ accuracy on flavor tagging performance:

| true \ identified as | b | c | g |
|---|---|---|---|
| b | 0.8675 | 0.0887 | 0.0437 |
| c | 0.1136 | 0.6263 | 0.2601 |
| g | 0.0411 | 0.1007 | 0.8582 |

$\nu\bar{\nu}H$

| true | b | c | g |
|---|---|---|---|
| b | 1 | 0 | 0 |
| c | 0 | 1 | 0 |
| g | 0 | 0 | 1 |

identified as
perfect

| true | b | c | g |
|---|---|---|---|
| b | 1/3 | 1/3 | 1/3 |
| c | 1/3 | 1/3 | 1/3 |
| g | 1/3 | 1/3 | 1/3 |

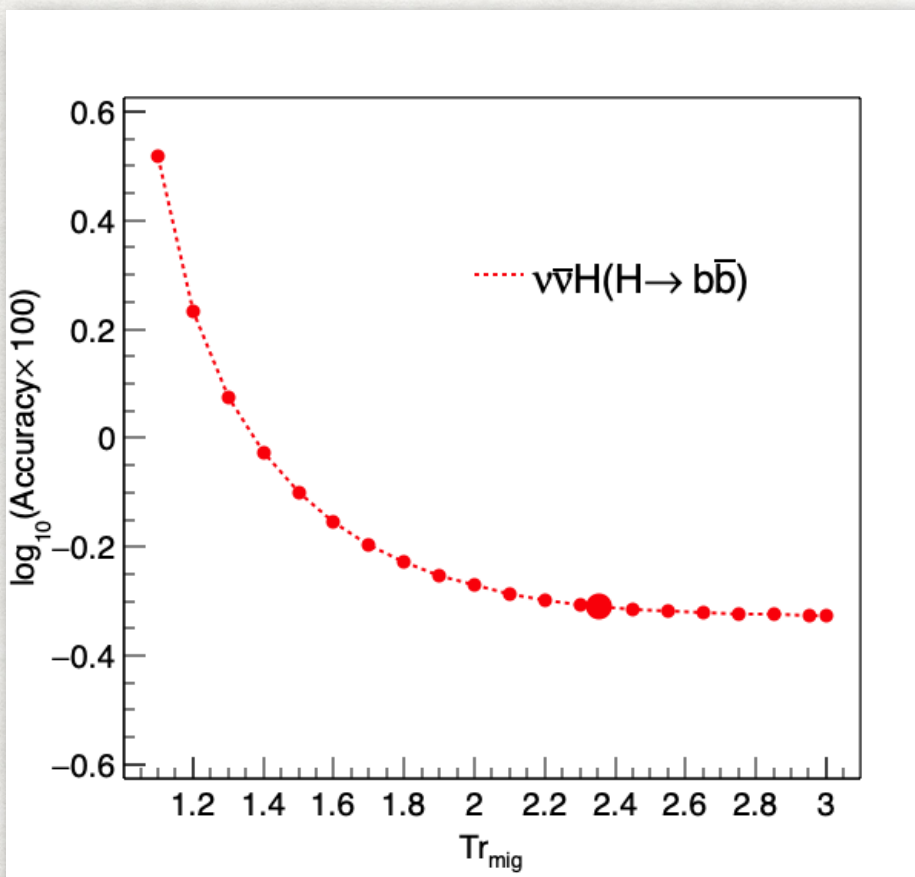
identified as
none

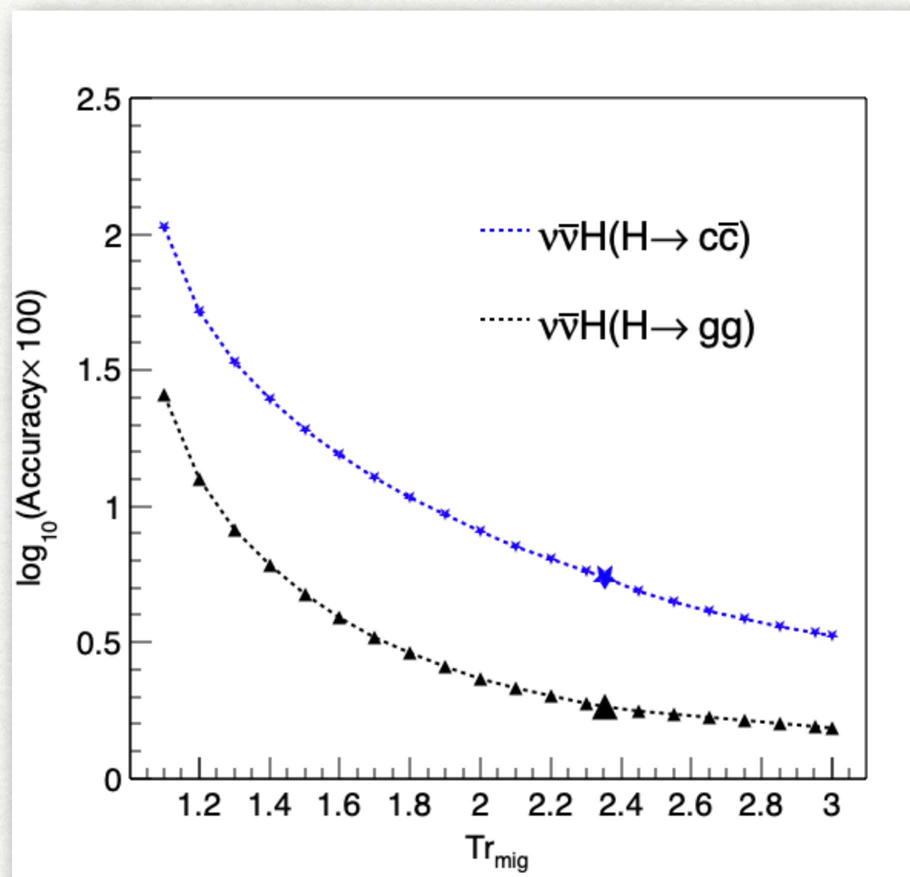
Obtain different flavor tagging performance matrices.

$$M_{mig} = \frac{Tr_{mig} - Tr_{opt}}{Tr_I - Tr_{opt}} \cdot (M_I - M_{opt}) + M_{opt}$$

$$M_{mig} = \frac{Tr_{mig} - Tr_{opt}}{Tr_{1/3} - Tr_{opt}} \cdot (M_{1/3} - M_{opt}) + M_{opt}$$

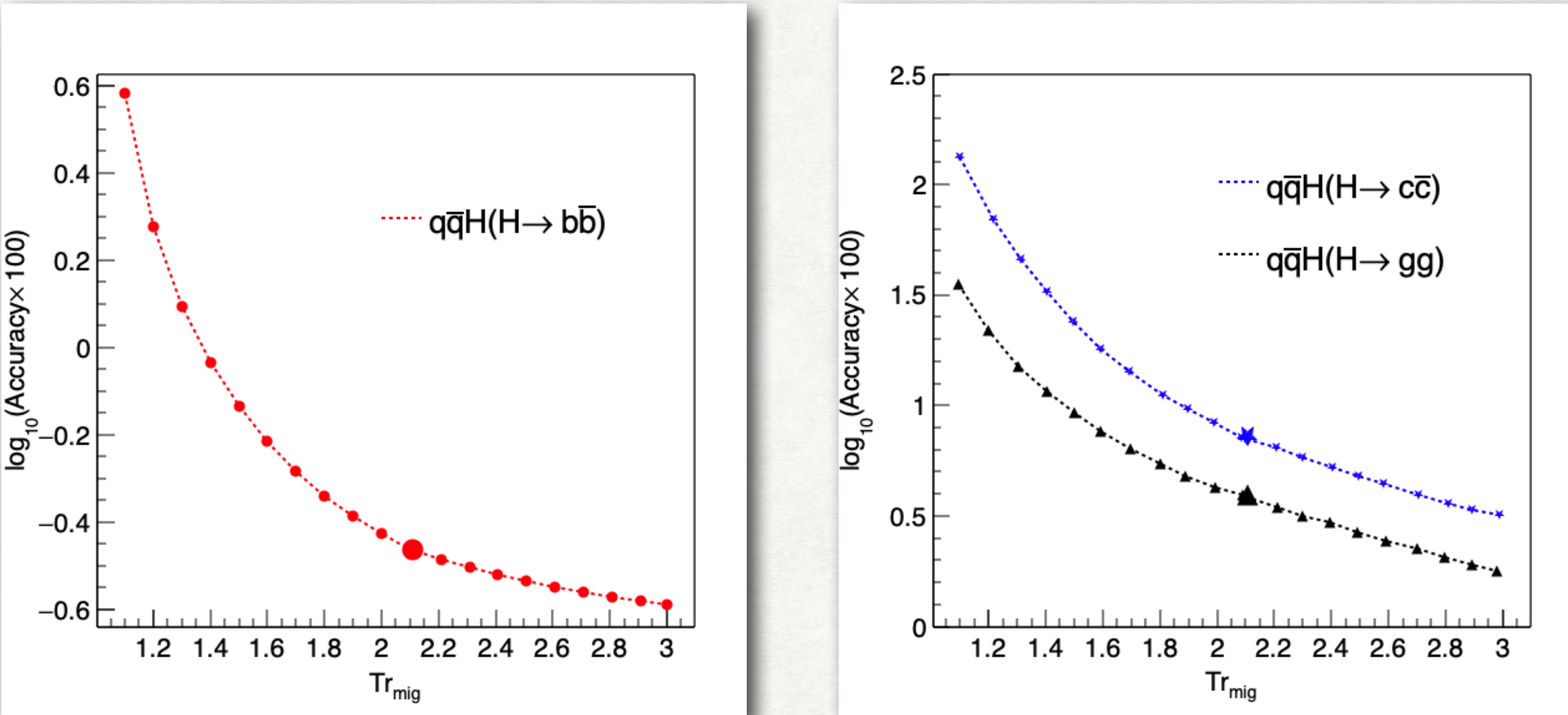
relative accuracy at different migration matrices




perfect flavor tagging performance can improve $H \to b\bar{b}/c\bar{c}/gg$ signal strength accuracy by 2%/63%/13%

# $q\bar{q}H$



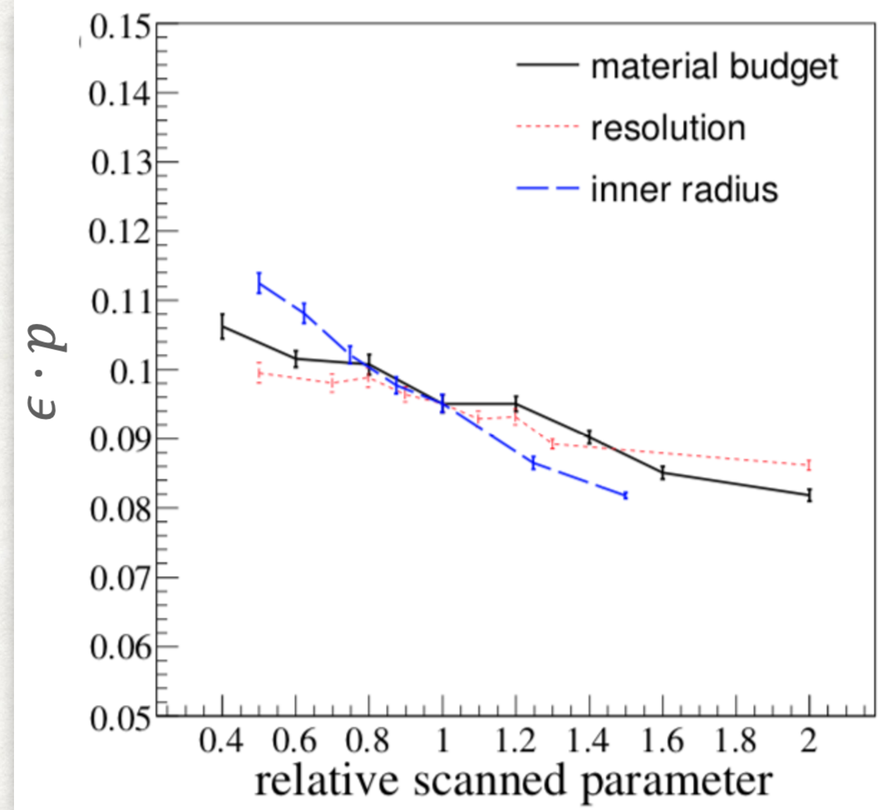

perfect flavor tagging performance can improve $H \to b\bar{b}/c\bar{c}/gg$ signal strength accuracy by 35%/122%/181%.

# The dependency of flavor tagging performance on vertex detector parameters

Z. Wu *et al* 2018 *JINST* **13** T09002

**Table 1.** The baseline design parameters of the CEPC vertex system.

| | R(mm) | Z(mm) | single-point resolution($\mu m$) | material budget |
|---|---|---|---|---|
| Layer 1 | 16 | 62.5 | 2.8 | $0.15\%/X_0$ |
| Layer 2 | 18 | 62.5 | 6 | $0.15\%/X_0$ |
| Layer 3 | 37 | 125.0 | 4 | $0.15\%/X_0$ |
| Layer 4 | 39 | 125.0 | 4 | $0.15\%/X_0$ |
| Layer 5 | 58 | 125.0 | 4 | $0.15\%/X_0$ |
| Layer 6 | 60 | 125.0 | 4 | $0.15\%/X_0$ |




$$\epsilon \cdot p = 0.095(1 - 0.14\frac{\Delta x_{\text{material}}}{x_{\text{material}}})(1 - 0.09\frac{\Delta x_{\text{resolution}}}{x_{\text{resolution}}})(1 - 0.23\frac{\Delta x_{\text{radius}}}{x_{\text{radius}}})$$

$$Tr_{mig} = 2.118 + 0.054 \cdot log_2 \frac{R^0_{material}}{R_{material}} + 0.040 \cdot log_2 \frac{R^0_{resolution}}{R_{resolution}} + 0.098 \cdot log_2 \frac{R^0_{radius}}{R_{radius}}$$

$R^0_{material}$ : the default material budget of vertex detector

$R_{material}$ : the changed material budget

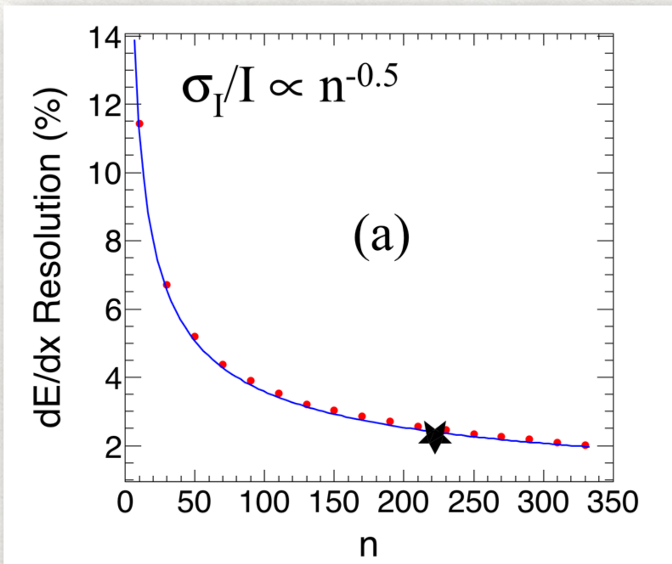
**Table 2.** Reference geometries.

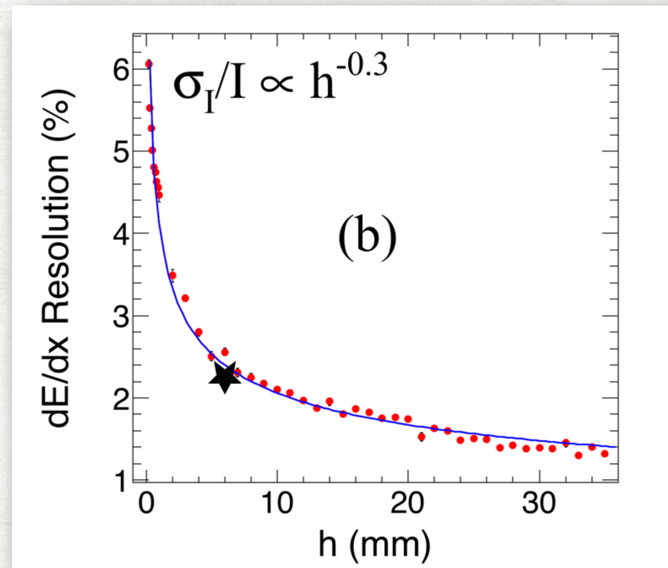
| | Scenario A (Aggressive) | Scenario B (Baseline) | Scenario C (Conservative) |
|---|---|---|---|
| Material per layer/$X_0$ | 0.075 | 0.15 | 0.3 |
| Spatial resolution/μm | 1.4 - 3 | 2.8 - 6 | 5 - 10.7 |
| $R_{in}$/mm | 8 | 16 | 23 |
| trace | 2.3 | 2.1 | 1.9 |

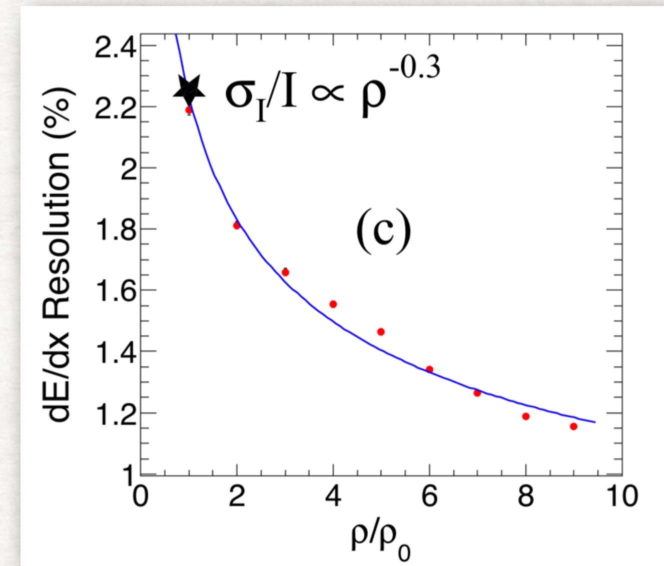
If the inner radius changes from 16 mm to 10 mm, the trace changes from 2.10 to 2.18.

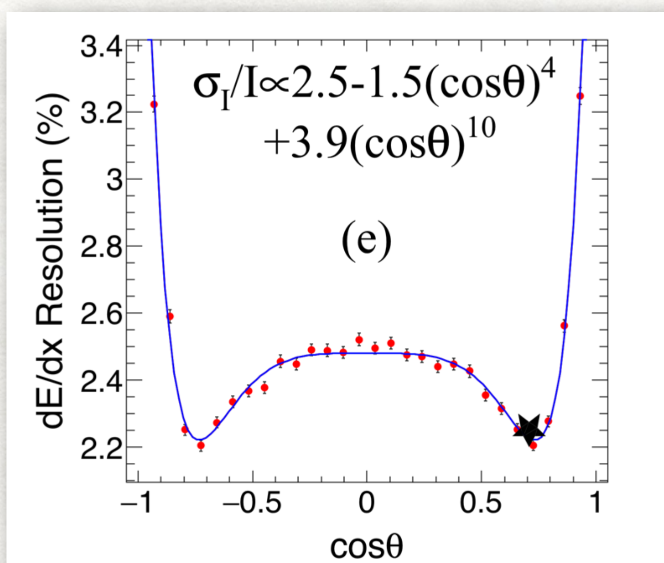
# PID performance based on dE/dx & TOF information

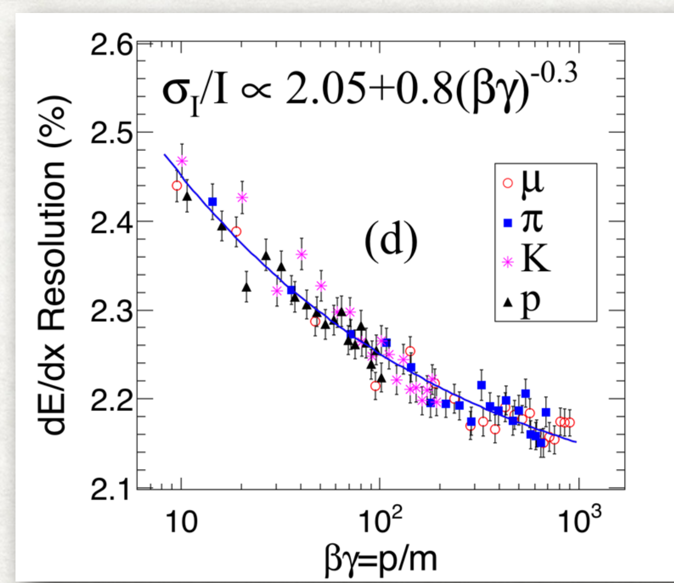
# intrinsic dE/dx resolution



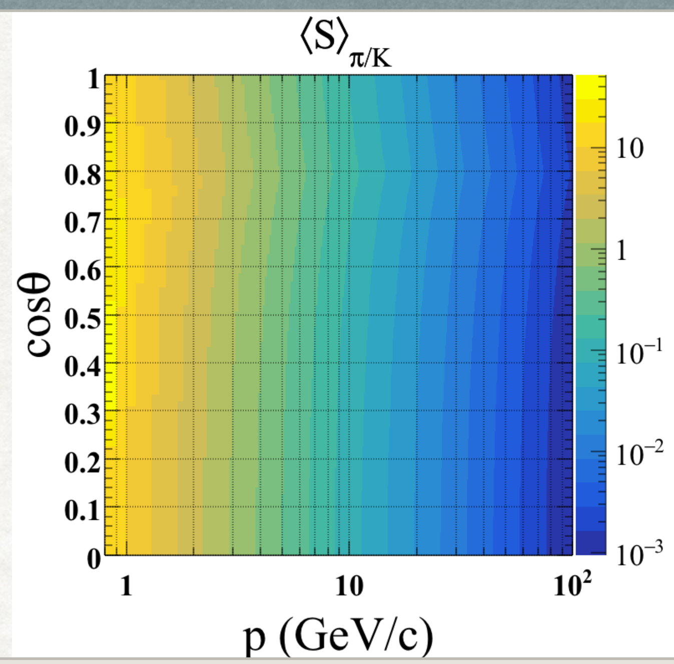
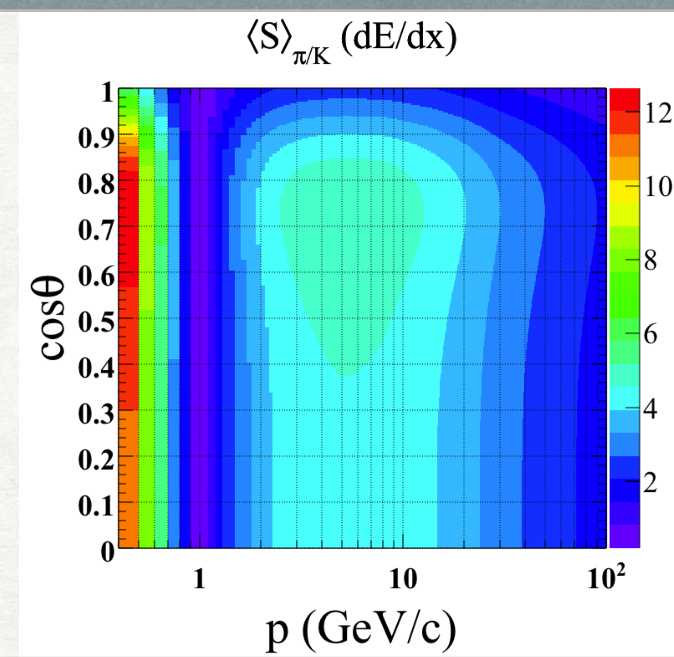
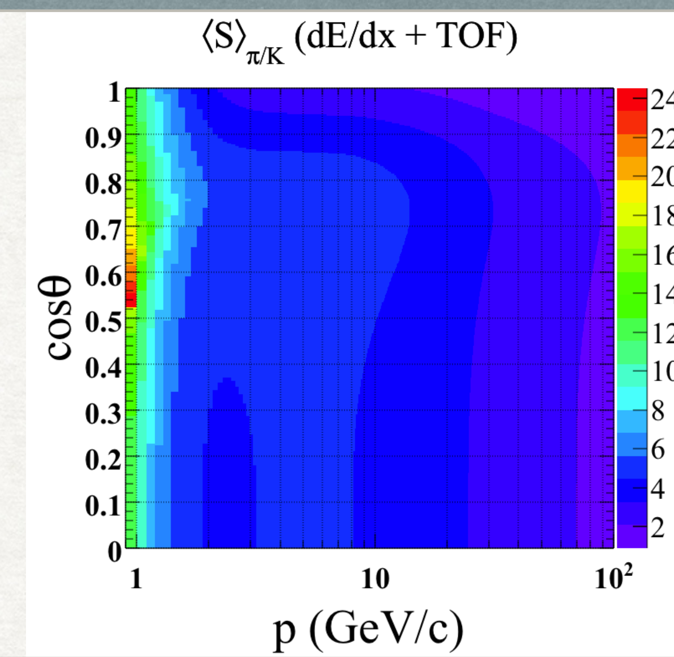
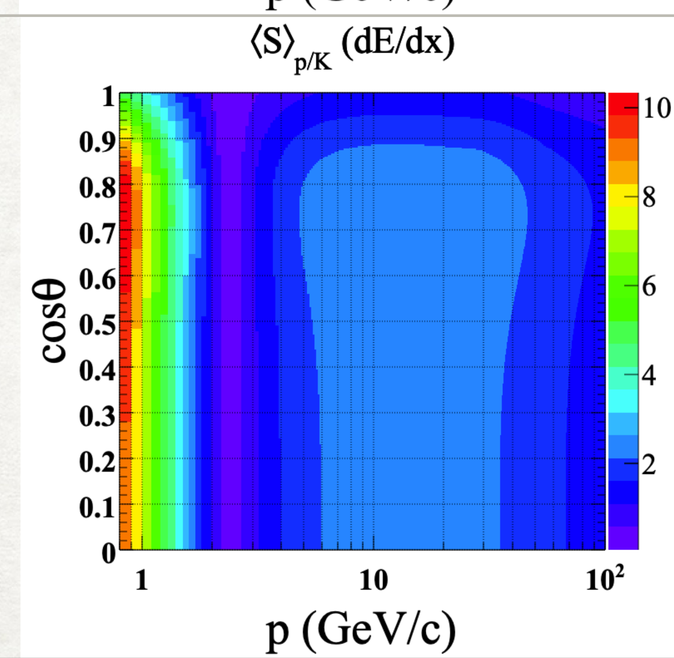
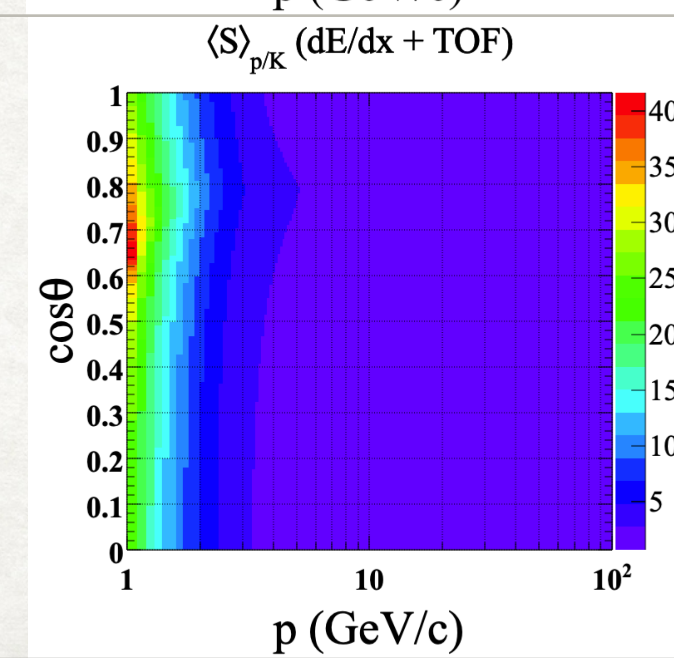




Fenfen An
arXiv:1803.05134

The dE/dx resolution could reach 2.5% in the barrel region with particle's energy ranges from 0.1 GeV to 100 GeV.

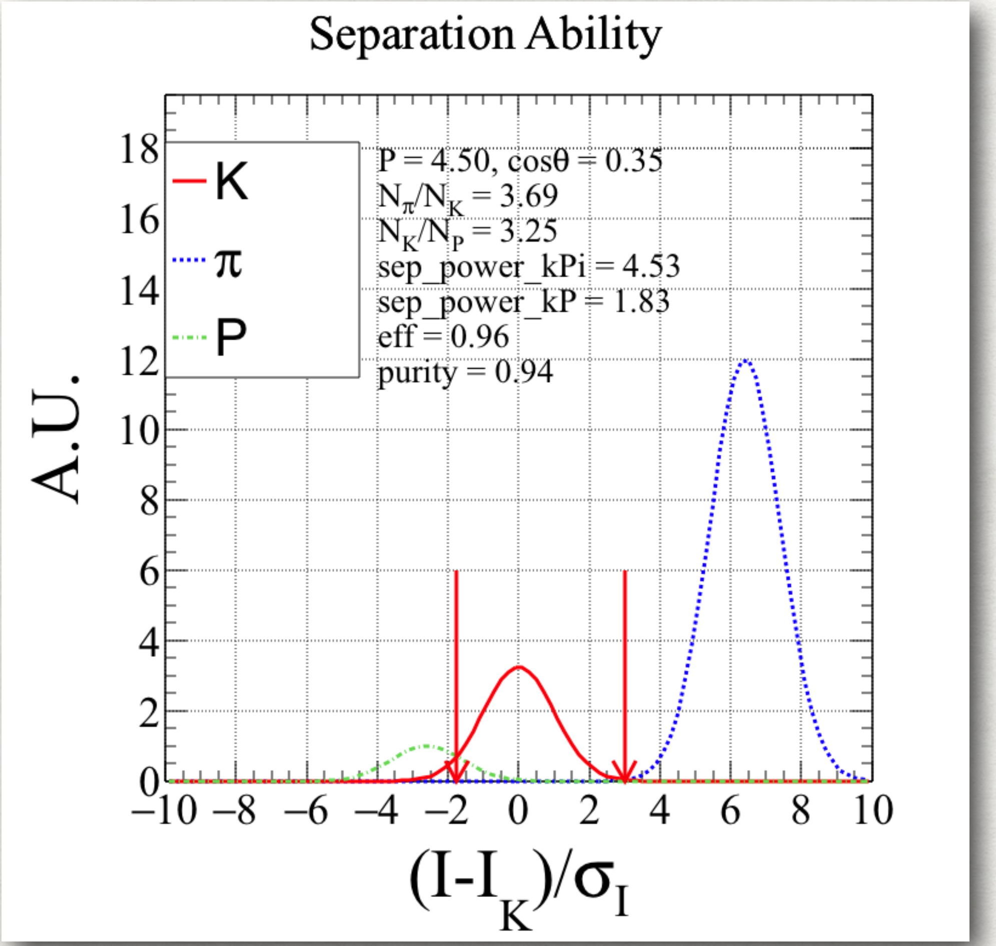
# Separation power

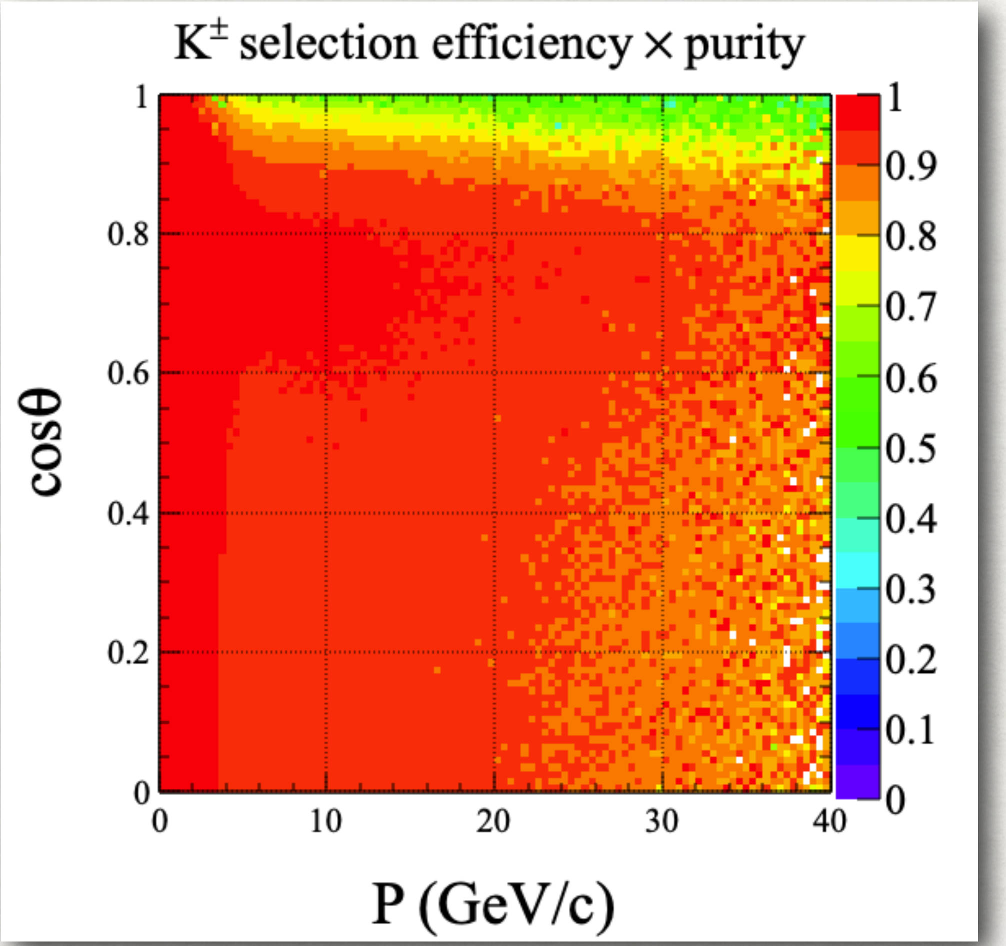
$$S_{A/B} = \frac{|t_A - t_B|}{\sqrt{\sigma_A^2 + \sigma_B^2}} \qquad S_{A/B} = \frac{|dE/dx_A - dE/dx_B|}{\sqrt{\sigma_A^2 + \sigma_B^2}}$$

| | TOF | dE/dx | combine |
|---|---|---|---|
| Kaon/ Pion |  |  |  |
| Kaon/ Proton | |  |  |

# PID performance on Z-pole samples

## $K^{\pm}$ identification




$$\text{overall } efficiency = \frac{\sum_{bins} \epsilon_K * N_k}{\sum_{bins} N_k}$$ 95.97%

$$\text{overall } purity = \frac{\sum_{bins} p_K * N_k}{\sum_{bins} N_k}$$ 81.56%

the dependency of $K^{\pm}$ identification performance on dE/dx resolution

$\sigma_{actual} = factor \cdot \sigma_{intrinsic}$ ,

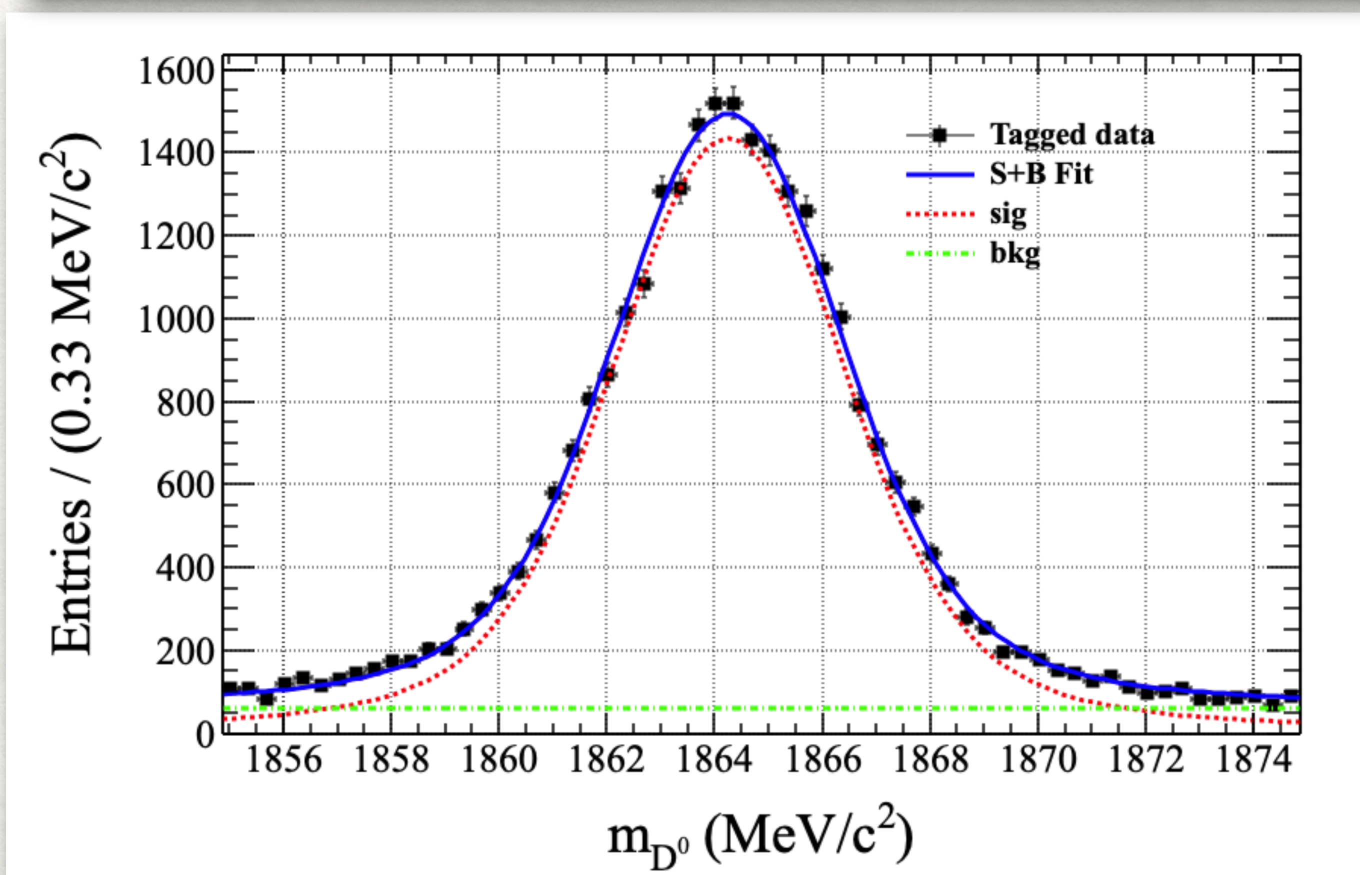
$\sigma$ : dE/dx resolution,

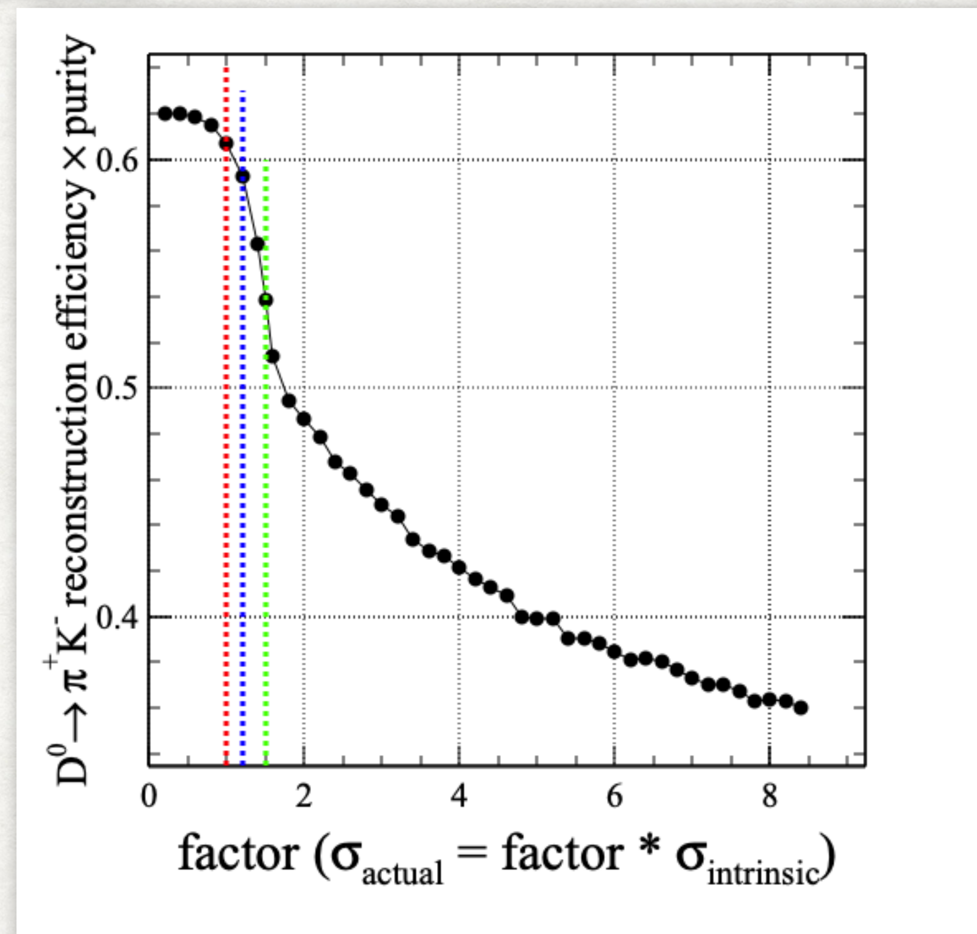
The factors are selected according to other detectors, such as PEP-4, TOPAZ, DELPHI, ALEPH, ALICE.

| | factor | 1. | 1.2 | 1.5 |
|---|---|---|---|---|
| | $\varepsilon_K$ (%) | 95.97 | 94.09 | 91.19 |
| dE/dx | $purity_K$ (%) | 81.56 | 78.17 | 71.85 |
| | $\varepsilon_K$ (%) | 98.43 | 97.41 | 95.52 |
| dE/dx& TOF | $purity_K$ (%) | 97.89 | 96.31 | 93.25 |

## $D^0 \to \pi^+ K^-$ reconstruction with Z-pole samples

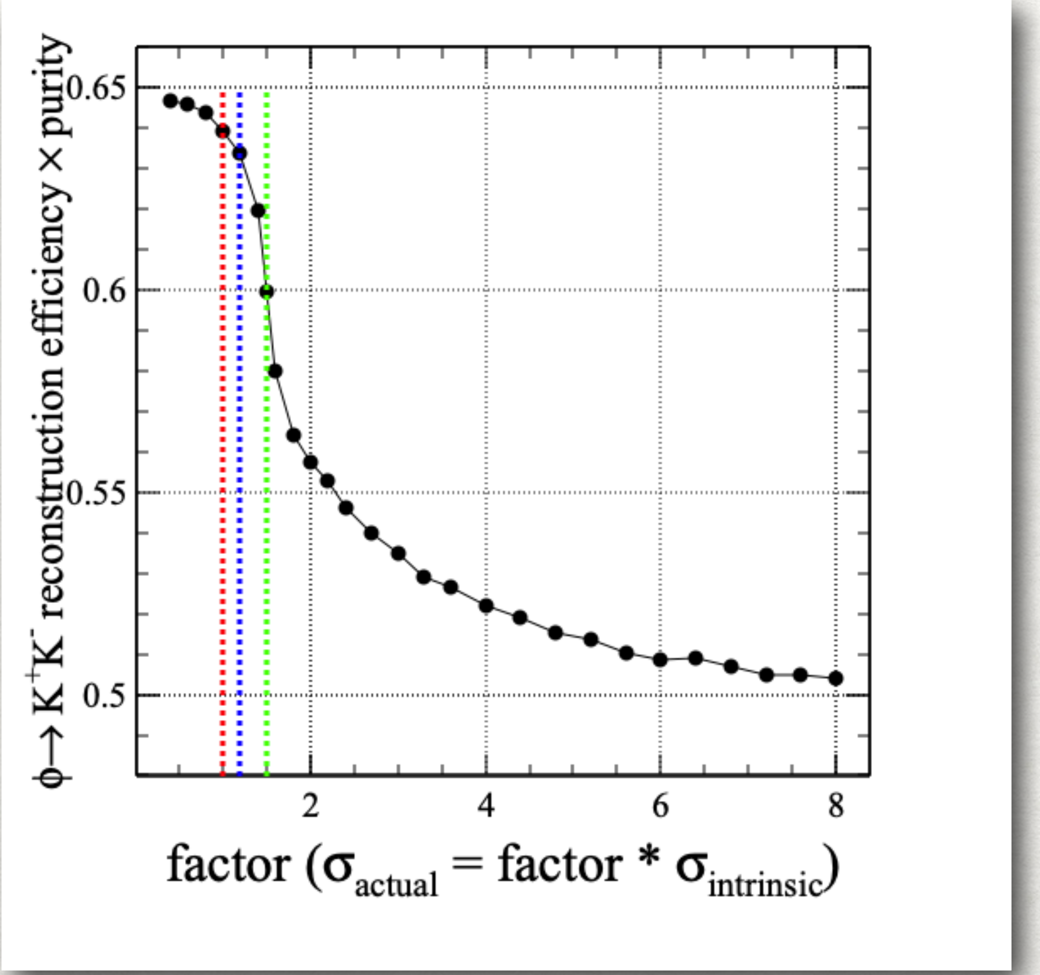
| | $\epsilon$ (%) | p (%) |
|---|---|---|
| $\lvert mass - mass_{D0} \rvert < 0.01\,\mathrm{GeV/c^2}$ | $90.39 \pm 0.24$ | $2.16 \pm 0.07$ |
| $\mathrm{IMP} > 0.02\,\mathrm{mm}^2$ | $79.12 \pm 0.21$ | $5.04 \pm 0.11$ |
| vertex fitted $\chi^2 < 5.15$ | $72.62 \pm 0.23$ | $15.36 \pm 0.18$ |
| dis of vertex to IP $> 0.305$ mm | $69.24 \pm 0.24$ | $\boxed{28.41 \pm 0.23}$ |
| PID | $68.19 \pm 0.24$ | $\boxed{89.05 \pm 0.16}$ |



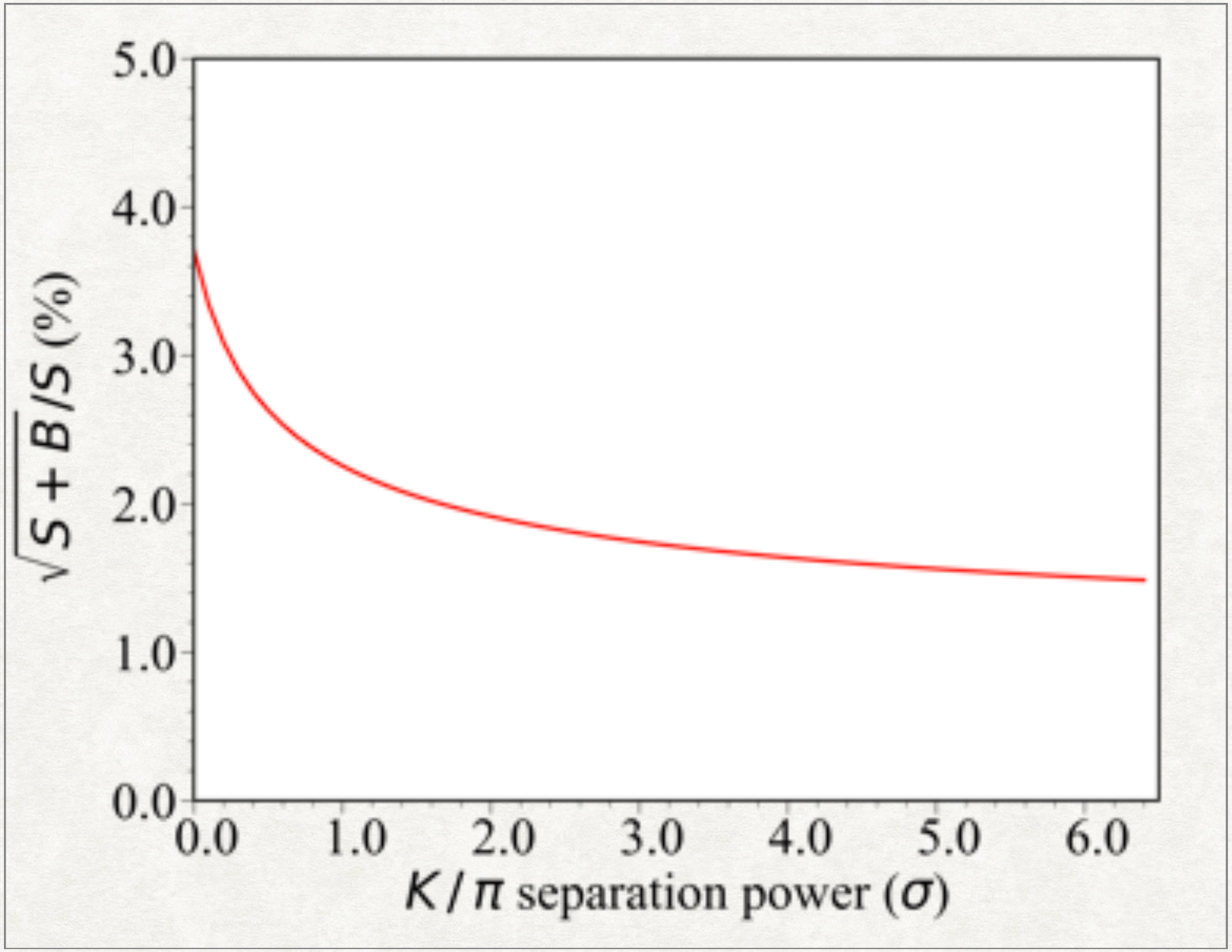
$\phi \to K^+K^-$ reconstruction

| | $\epsilon(\%)$ | p (%) |
|---|---|---|
| $\lvert mass - mass_{D0} \rvert < 8\,\mathrm{MeV/c^2}$ | $89.42 \pm 0.07$ | $6.43 \pm 0.06$ |
| vertex fitted $\chi^2 < 9.95$ | $83.15 \pm 0.09$ | $9.14 \pm 0.07$ |
| PID | $82.26 \pm 0.09$ | $77.70 \pm 0.10$ |



# $B_s \to \phi\nu\nu$ measurement

arXiv:2201.07374



# Summary :

- The relative accuracy of $H \to b\bar{b}/c\bar{c}/gg$ signal strength has been measured at the CEPC with full simulation of all SM samples corresponding to 240 GeV.

| | $H \to b\bar{b}$ | $H \to c\bar{c}$ | $H \to gg$ |
|---|---|---|---|
| $5.6ab^{-1}$ | 0.27% | 4.03% | 1.56% |
| $20ab^{-1}$ | 0.14% | 2.13% | 0.82% |

- The vertex detector has great impact on flavor tagging performance,

$$Tr_{mig} = 2.118 + 0.054 \cdot log_2\frac{R^0_{material}}{R_{material}} + 0.040 \cdot log_2\frac{R^0_{resolution}}{R_{resolution}} + 0.098 \cdot log_2\frac{R^0_{radius}}{R_{radius}}$$

- The charged kaon identification and $D^0 \to \pi^+K^-/\phi \to K^+K^-/B_s \to \phi\nu\nu$ reconstruction on hadronic Z-pole samples require the degradation of dE/dx resolution to be less than 20%, corresponding to dE/dx resolution better than 3%.

# Backup

The signal strength accuracy of H->bb/cc/gg at CEPC, ILC and FCC-ee.

| | bb | cc | gg |
|---|---|---|---|
| CEPC | 0.27% | 4.03% | 1.56% |
| ILC, 250 GeV, $2ab^{-1}$ | 0.7% | 4% | |
| FCC-ee, 240 GeV, $5ab^{-1}$ | 0.3% | 2.2% | 1.9% |