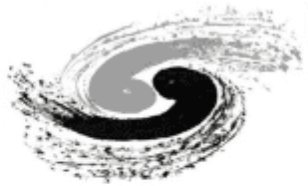


中国科学院高能物理研究所
Institute of High Energy Physics
Chinese Academy of Sciences

Updates on PID and documentation

Geliang Liu

May. 7th, 2025

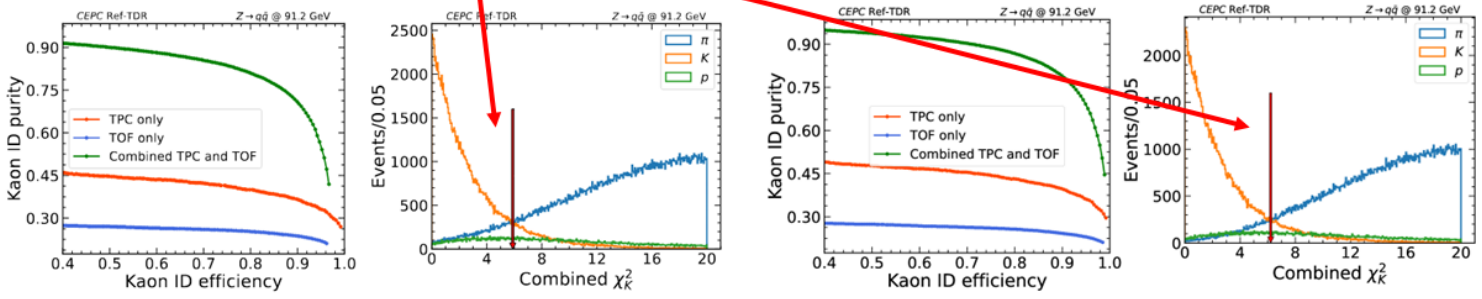
Charged hadron ID

Worse performance compared to Xiaotian's results

	χ^2 method			XGBoost		
	efficiency	purity	eff*pur	efficiency	purity	eff*pur
kaon	83.3%	62.4%	52.0%	70.3%	80.1%	56.7%

Kaon PID Efficiency just among charged π,K,p

Selections	K Ntrack	PID Strategy	K to K Ntrack	PID Eff.	π,K,p to K Ntrack	Purity	Eff*Purity
(a)-(e)	141891	Combined χ^2_K be the minimum	124627	87.83%	153255	81.32%	71.43%
		Maximize Kaon PID eff*purity Combined $\chi^2_K < 5.886$	119230	84.03%	152693	78.08%	65.61%
(a)-(h)	128872	Combined χ^2_K be the minimum	117213	90.95%	135181	86.71%	78.86%
		Maximize Kaon PID eff*purity Combined $\chi^2_K < 6.217$	112700	87.45%	137871	81.74%	71.49%



Need to check the reason, and if there is anything wrong in my codes.

Charged hadron ID

Overall performance with Z→qq @91.2 GeV

New results

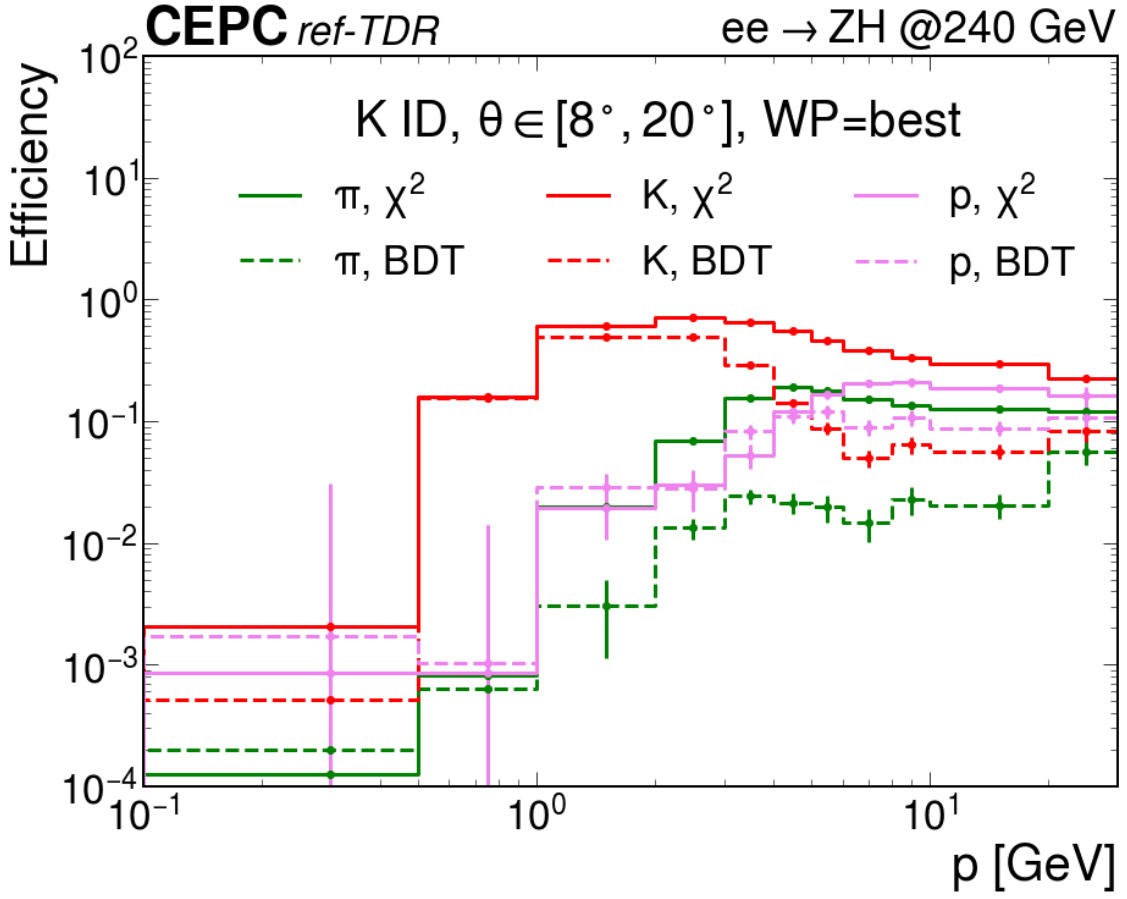
- Conditions a-e and h are applied.
- Set noLep WP.
- Use χ^2 method.
- **Overall kaon efficiency = 89.9%**
- **Overall kaon purity = 85.9%**
- **eff*purity = 77.3%**
- Close to Xiaotian's result.
  - Difference caused by f, g, and potentially statistical fluctuations.

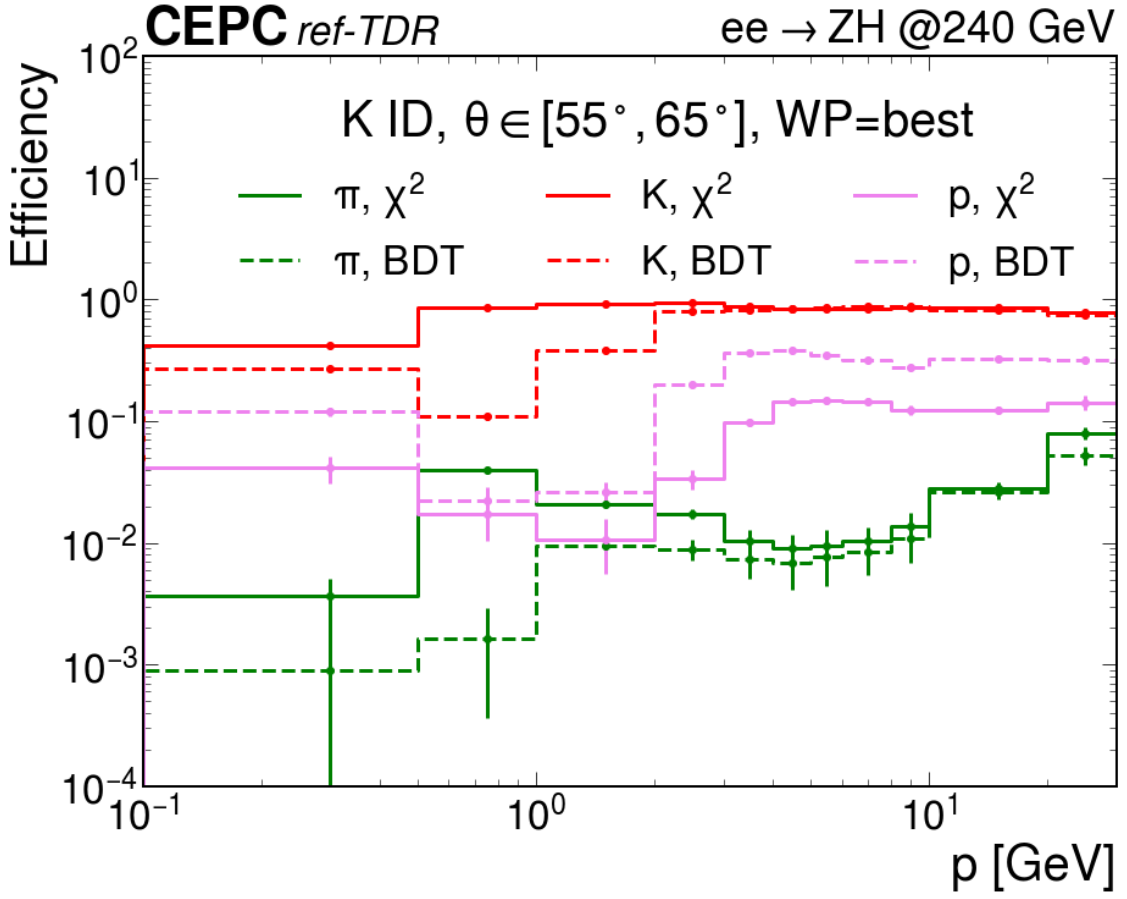
- However there is nothing wrong in the XGBoost method. Therefore, its performance is worse.

$Z \to q\bar{q}$ 100000 events (truth MCParticle charged stable e,μ,π,K,p: 2288676)				
Track selection criteria		Ntrack	Absolute Efficiency	Relative Efficiency
(a) CompleteTracks	…	2268408	100%	100%
(b) Track-MCP Match: CompleteTracks ParticleAssociation	Track hits match with stable MCParticle Choose MCParticle with max weight as truth particle	2268408	100%	100%
	Match with charged e,μ,π,K,p	2184944	96.32%	96.32%
	Match with charged K	230785	10.17%	10.56%
(c) No decays	Veto isdecayintrker	169639	7.48%	73.51%
(d) Hit TPC	matchedtpc	157234	6.93%	92.69%
(e) Hit TOF	matchedtof	141891	6.26%	90.24%
(f) No daughters	daughtersize==0	137638	6.07%	97.00%
(g) Track fitting	$\chi^2/ndof < 2$	136439	6.01%	99.13%
(h) Veto regions	Not in (p<0.5GeV && $\lvert\cos\theta\rvert < 0.55$) (p<1GeV && $\lvert\cos\theta\rvert > 0.55$) (p>4GeV && $\lvert\cos\theta\rvert > 0.9$)	128872	5.68%	94.45%

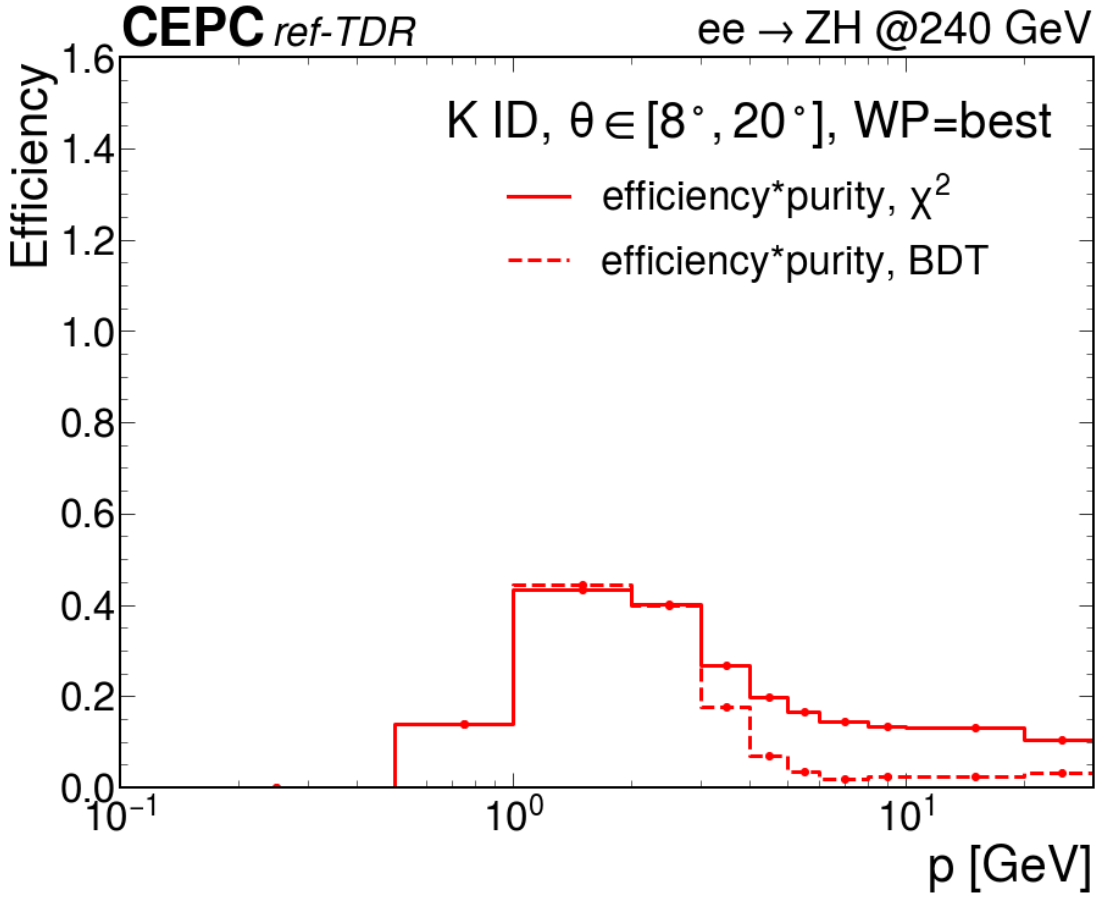
Kaon PID Efficiency just among charged π,K,p							
Selections	K Ntrack	PID Strategy	K to K Ntrack	PID Eff.	π,K,p to K Ntrack	Purity	Eff*Purity
(a)-(e)	141891	Combined χ^2_K be the minimum	124627	87.83%	153255	81.32%	71.43%
		Maximize Kaon PID eff*purity Combined $\chi^2_K < 5.886$	119230	84.03%	152693	78.08%	65.61%
(a)-(h)	128872	Combined χ^2_K be the minimum	117213	90.95%	135181	86.71%	78.86%
		Maximize Kaon PID eff*purity Combined $\chi^2_K < 6.217$	112700	87.45%	137871	81.74%	71.49%

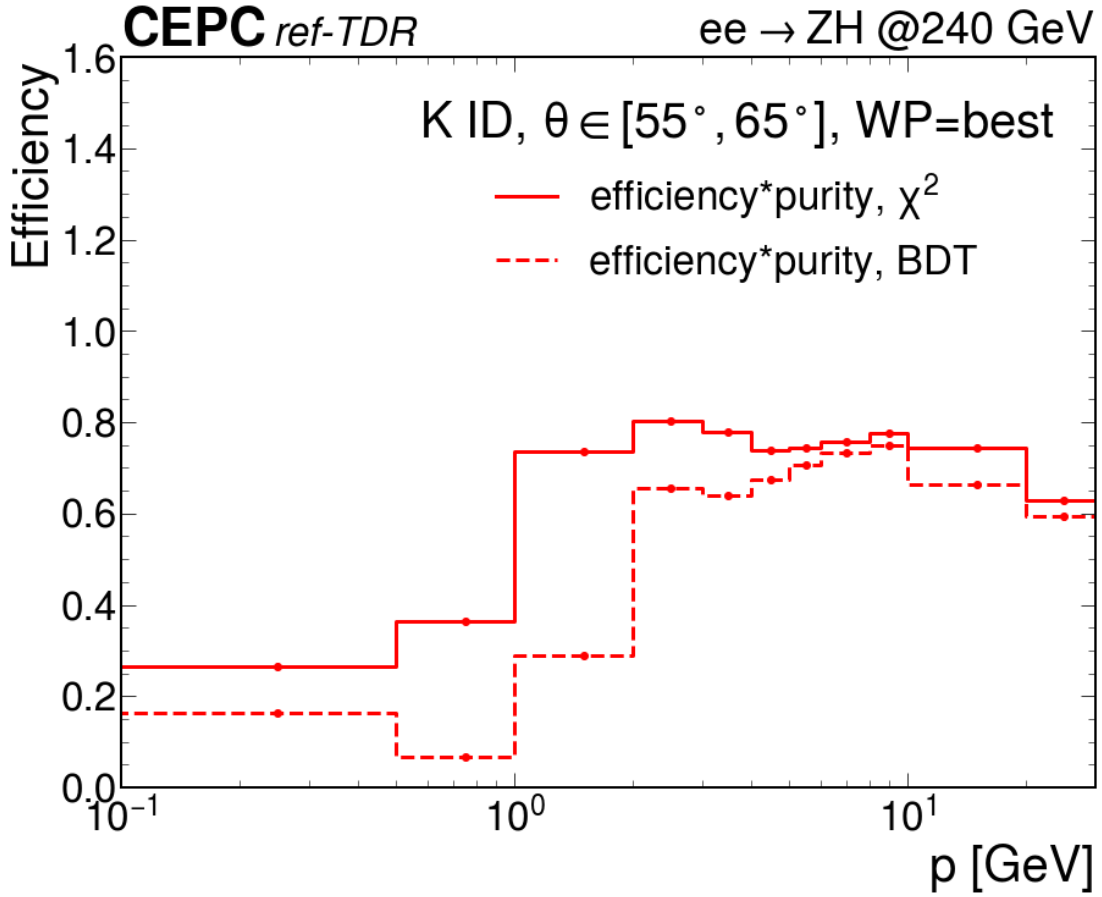
Efficiency in different phase space





Efficiency * purity





Overall better performances with the χ^2 method.

- The p / theta binning of the XGBoost model is not fine enough: Dndx feature changes very much with them.
- Too many unuseful features put in the model regarding distinguishment of charged hadrons.

XGBoost implemented in SW

In a new branch (XGBoost)

- https://code.ihep.ac.cn/glliu/FinalPIDSvc/-/tree/XGBoost?ref_type=heads
- The user interface is the same: SetCollections, MatchMuonHitsToTracks, Set WPs, LoadPFO, and GetType.
- Six WPs for leptons are provided: 50%, 70%, 90%, 98%, Best, noLep.
- For charged hadrons, the traditional χ^2 method is used.

Tests

- Kaili is using the new PID in JOI. Based on the confusion matrices and eff*purity, the results are much better than the χ^2 method.
- Shuo is checking if the 98% WP does give a 98% efficiency in Z→μμ.

Documentation

To well save our studies, and give hands-on instructions for newcomers, a git repository is created:

- https://code.ihep.ac.cn/glliu/CEPC_PhyPerf_Intro
- Composed of a few markdowns.

Convenience

- Online edit: click Edit and then Web IDE Quickly and easily edit multiple files in your project.
- Markdown is better to be written or read than twiki or plain texts.
- Easy to track the history.
- Easy for cooperation editing.

I will begin to fill in more information.

- Some needs the help of others, especially Sample production.

Instructions on physics performance studies at CEPC

This project provides solely the instroductions on performing physics analyses using the CEPC and offline reconstruction based on the CEPC reference detector. It will not dig into the softwa information about what it does, how to use it, what we have done so far, and what you can dire

The instructions are divided into a few parts:

Overall introduction of CEPCSW

See 1.CEPCSW.md.

Sample production

Full simulation

See 2.1.FullSimulation.md.

Fast simulation

See 2.2.FastSimulation.md.

Gen filter

See 2.3.GenFilter.md.

Object performance

See 3.ObjectPerformance.md.

Ongoing physics analyses