PID Requirements for Future HTE Factories

Uli Einhaus
ECFA WG3 Workshop on Calorimeters,
Photon Detection and Particle ID
04.05.2023 CERN
ulrich.einhaus@desy.de

ECFA
European Committee for Future Accelerators

DESY.

tl;dr

- What is the PID requirement for future HTE factories?

  → We don’t know (yet)!
  But: By now, a number of studies gives us a good picture of what we should aim for.

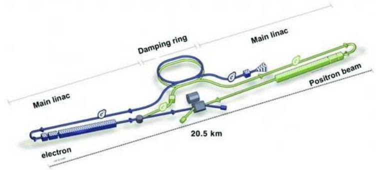
- Disclaimer: aim of overview talk, no claim of completeness, with biased selection!

Content

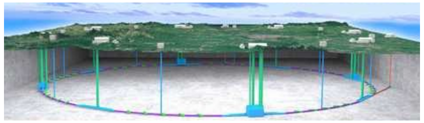
- Introduction and background
- Overview of analyses using PID with focus on ones that connect PID and physics results
- Further considerations
- Summary

The Landscape of Proposed Next-Gen Colliders / Future HTE Factories

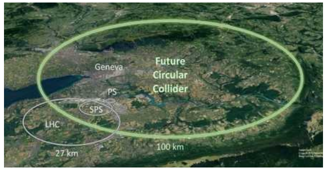
ILC



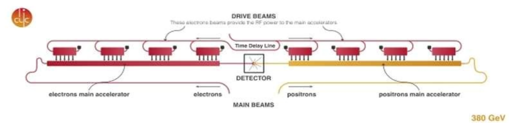
CEPC



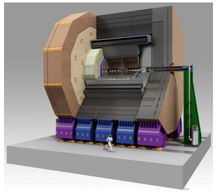
FCC-ee



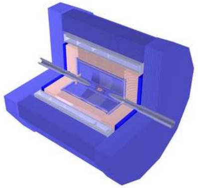
CLIC



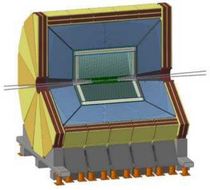
ILD



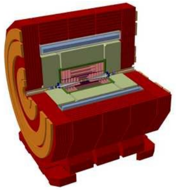
CEPC Baseline



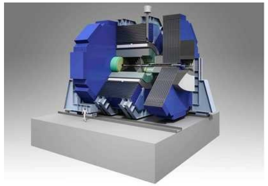
IDEA



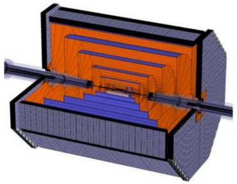
CLICdp



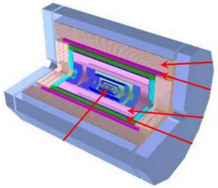
SiD



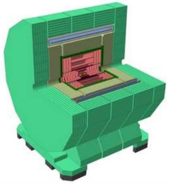
FST



CEPC 4th concept



CLD

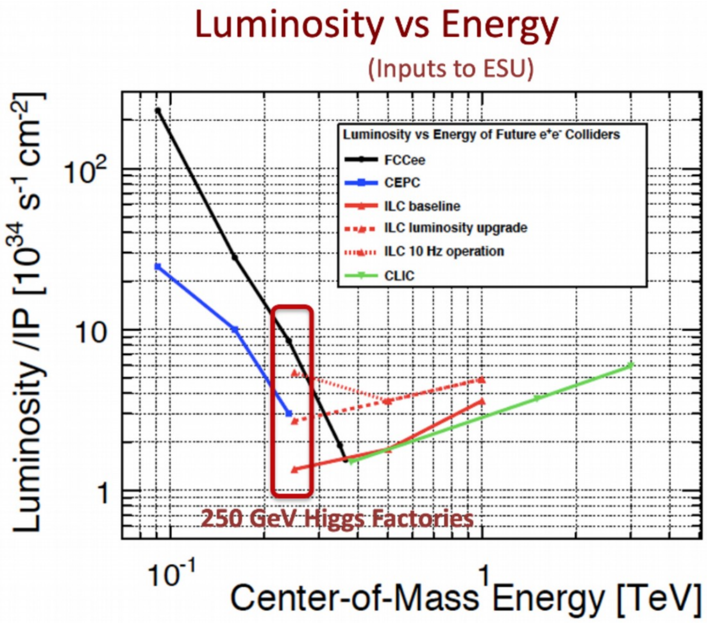


slide stolen from B. Dudar

+ many more concepts, e.g. energy recovery linacs or muon collider

Introduction and Background

- Similar conditions at all these machines: collide e+e- (or muons), produce Higgs
- Biggest difference: linear and circular → energy reach and luminosity
- Energy: lower energies → more precision desired, higher energies → PID more difficult
- Over the last decade, PID has received increasing attention, in particular with the increase of studies at cicular machines
- Overall picture: various projects have been introduced, few have detailed studies, most have similar physics conditions and similar PID requirements



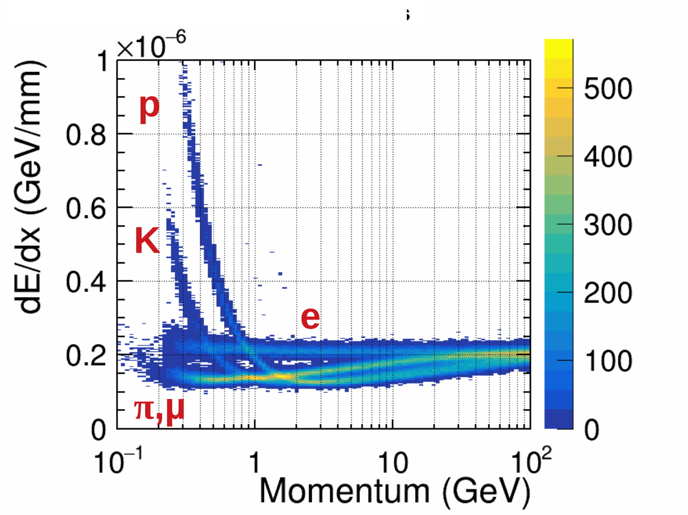
https://arxiv.org/abs/1903.01629

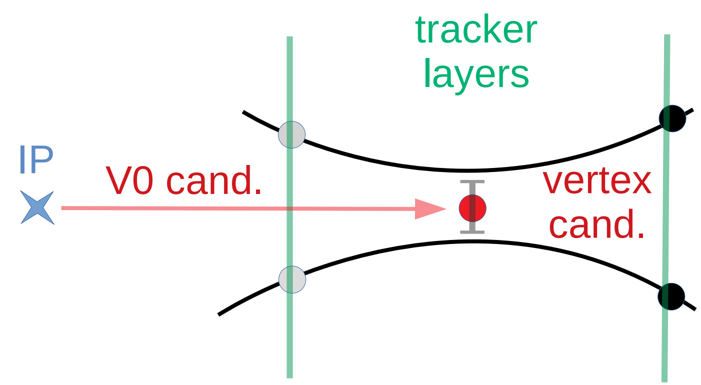
Particle Identification

- Focus: charged hadron ID, mostly pion vs. kaon

  → default performance plot: pion/kaon separation power

- Dedicated electron and muon ID relevant at lower momenta
- Also don't forget V0s: K_S^0, Λ





PID Technology (most prominent)

- Gaseous trackers (Time Projection Chamber, Drift Chamber): specific energy loss dE/dx, via gas ionisation, up to 30 GeV

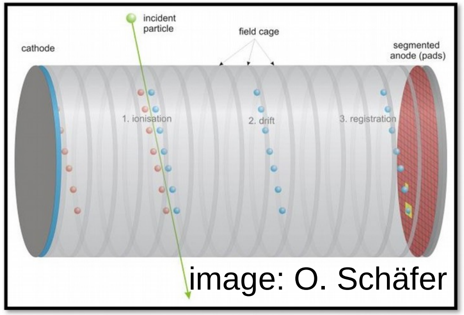
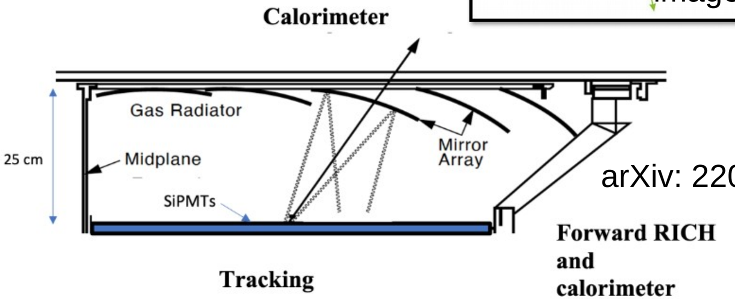


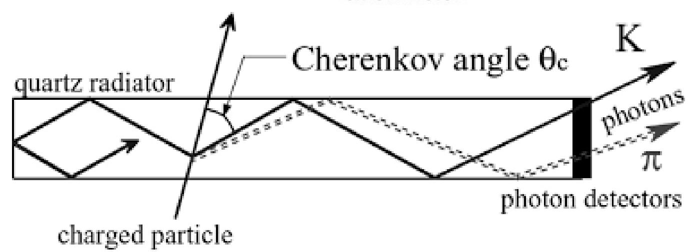
image: O. Schäfer

- Ring Imaging Cherenkov Detectors: Cherenkov angle, via imaging, 10 to 50 GeV



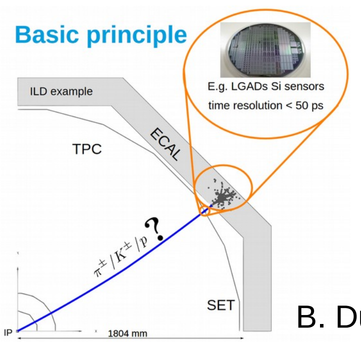
arXiv: 2203:07535

- Time of Propagation Counter: Cherenkov angle, via timing, up to 10 GeV



https://doi.org/10.1016/j.nima.2017.02.045

- Time of Flight: time, via Silicon timing, up to 5 GeV

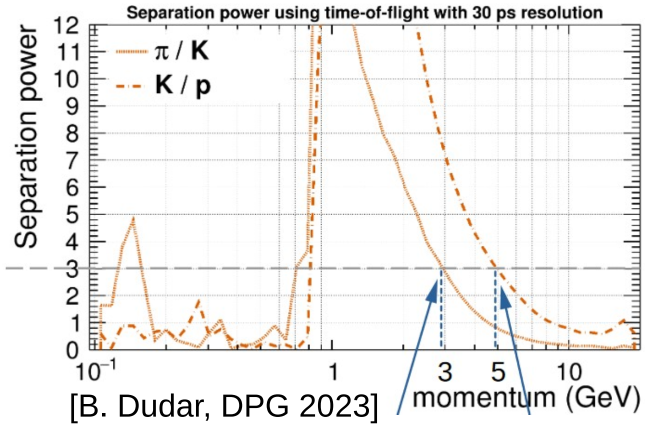


B. Dudar

Typical PID Performance

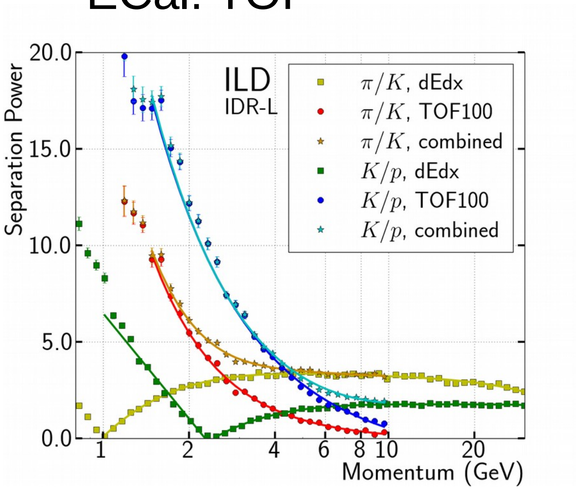
- Systems can cover broad range of momenta

TOF



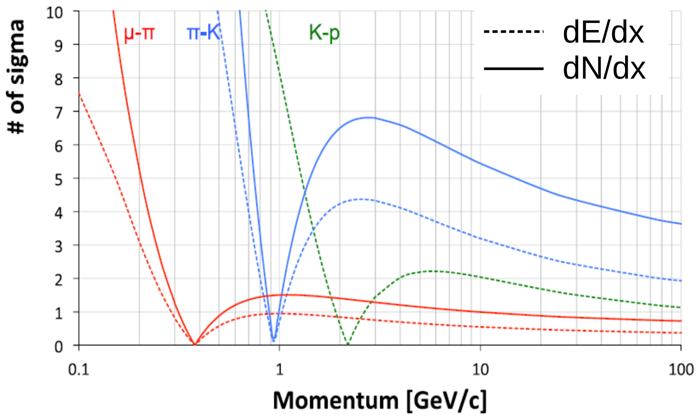
[B. Dudar, DPG 2023]

TPC: dE/dx (dN/dx?)
ECal: TOF



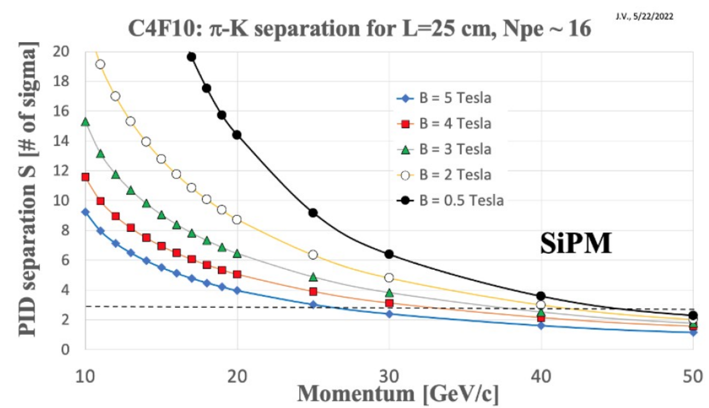
[ILD IDR 2020
https://arxiv.org/abs/2003.01116]

DC: dE/dx or dN/dx



[IDEA, FCC-ee CDR 2019
https://doi.org/10.1140/epjst/e2019-900045-4]

RICH

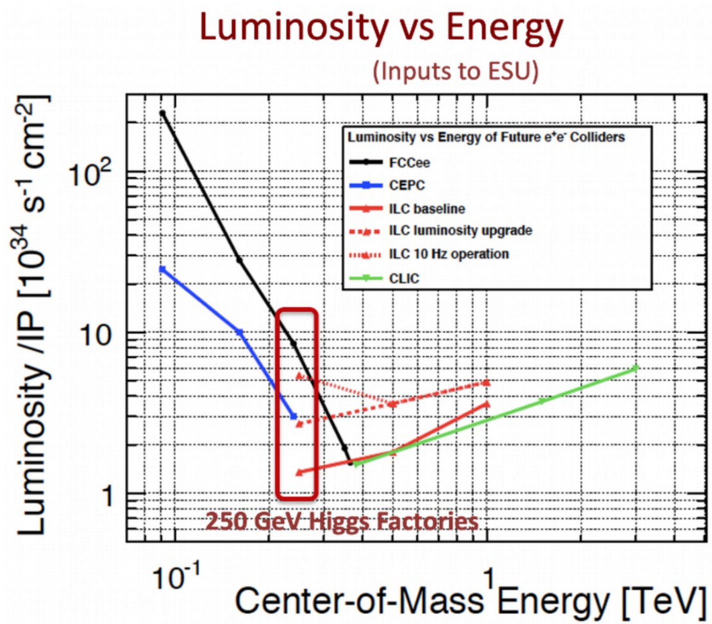


[https://arxiv.org/abs/2203.07535]

Introduction and Background

- Individual PID systems are fairly interchangable between colliders, though not completely independent and some notable dependencies exist:
  - Luminosity: rate capability of PID system, e.g. TPC ion backflow at Z-pole (study ongoing)
  - Effect of magnetic field e.g. on readout components, TPC drift, RICH resolution
  - …
- Much less interchangable between detector concepts, often highly integrated in overall design

→ Detailed studies necessary, many variations possible

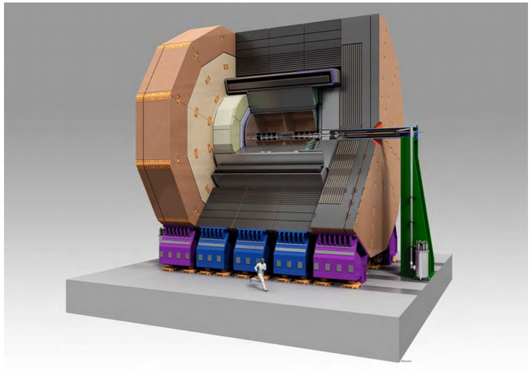


https://arxiv.org/abs/1903.01629

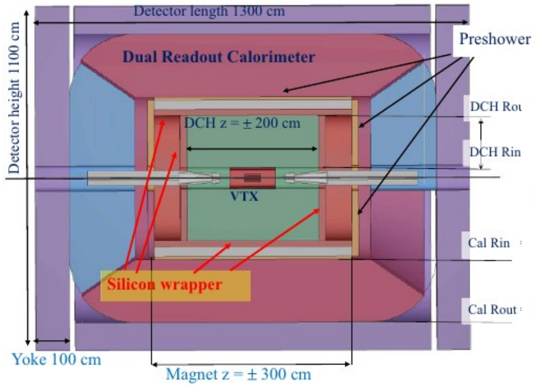
Examples List

- Flavour tagging
  - $H \rightarrow q\bar{q}$
  - $W \rightarrow q\bar{q}'$
  - $Z \rightarrow q\bar{q}$
  - non-SM top decays
- Vertex charge reconstruction
- B physics
- Track refitting
- Electron & muon ID
- TOF

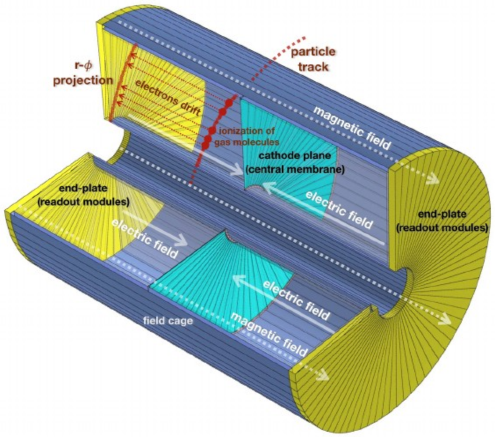
Most studies from three detector concepts with gaseous trackers, i.e. integrated PID systems (dE/dx / dN/dx)



ILD



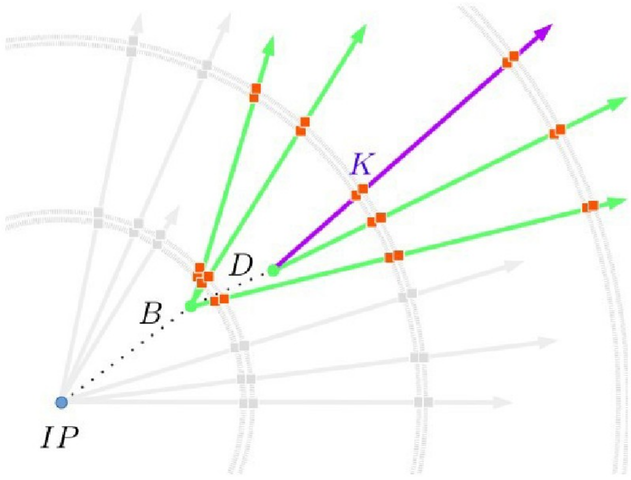
IDEA



CEPC-det

Flavour Tagging

- Basic jet flavour tagging via vertices and leading particles:
  - u, d: no secondary vertex, no hard strange particle
  - s: no secondary vertex, one hard strange particle
  - c: one secondary vertex, one hard strange particle
  - b: two secondary vertices, one hard strange particle
  - g: more average jet constituents than quarks;
    possibly tag the resulting $q\bar{q}$ system separately



https://indico.desy.de/event/33640/contributions/127531/

- Need excellent vertex resolution
- Need PID for charged kaons vs. pions & protons, K^0_S, Λ

→ tagging s vs. d/u is **only** possible via PID!

Higgs decays

- Measure Yukawa couplings, put limits on H-s coupling
- Several groups for FCC, one Hss analysis for ILD, one b/c/g for CEPC

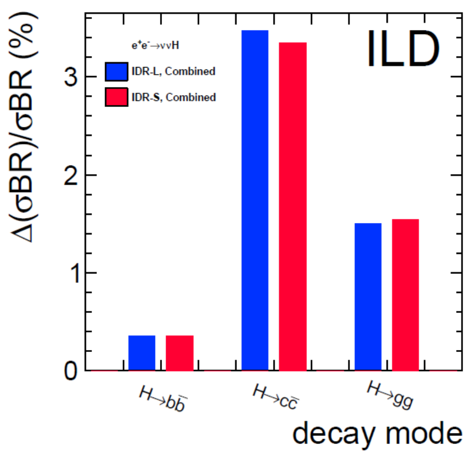
- **At HL-LHC**, projected
  - **bb ~4%**
  - **cc ~100%**
  - **ss, gg: none**

FCC-ee: $\kappa_s < 1.5$, 240 GeV, 5 ab^{-1}
https://indico.cern.ch/event/1176398/contributions/5208222

- 0.4% for Hbb, 1.2% for Hgg, 2.8% for Hcc, 160% for Hss in the neutrino channel
- 0.8% for Hbb, 3% for Hgg, 5.3% for Hcc, O(400%) for Hss in the leptonic channel
  - 0.4% / 1.2% / 2.7% / 150% when combining the two

CEPC: 240 GeV, 5.6 ab^{-1}
https://arxiv.org/abs/2203.01469

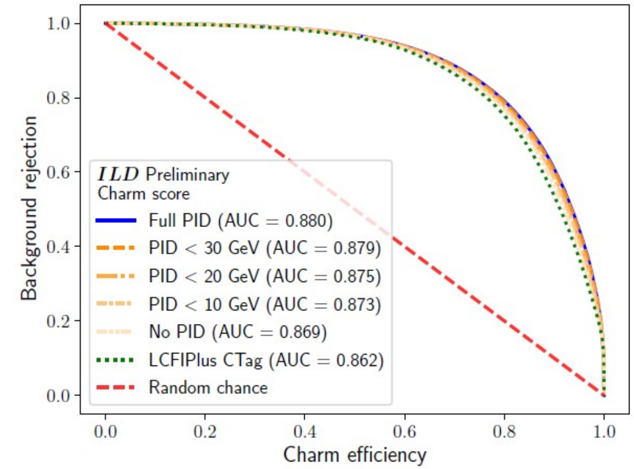
Z decay mode	$H \to b\bar{b}$	$H \to c\bar{c}$	$H \to gg$
$Z \to e^+e^-$	1.57%	14.43%	10.31%
$Z \to \mu^+\mu^-$	1.06%	10.16%	5.23%
$Z \to q\bar{q}$	0.35%	7.74%	3.96%
$Z \to \nu\bar{\nu}$	0.49%	5.75%	1.82%
combination	0.27%	4.03%	1.56%

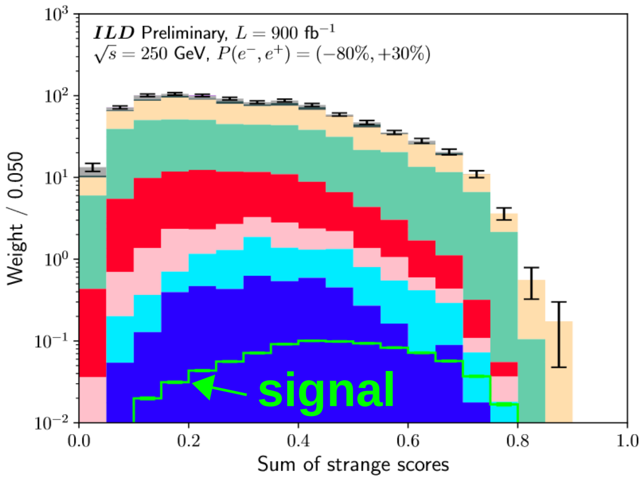


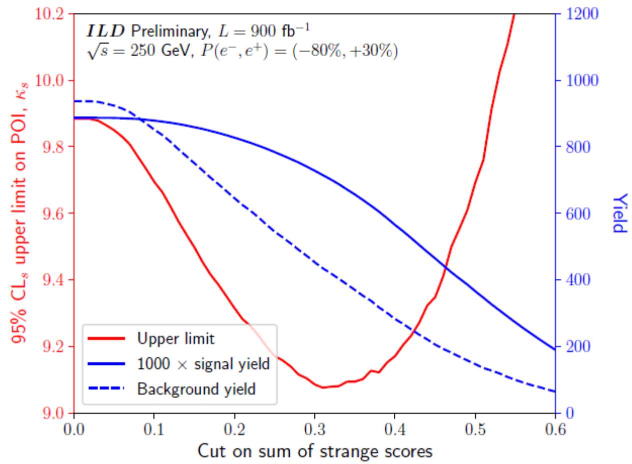
ILC: $\kappa_s < 7$, 250 GeV, 0.9 ab^{-1}
https://arxiv.org/abs/2203.07535

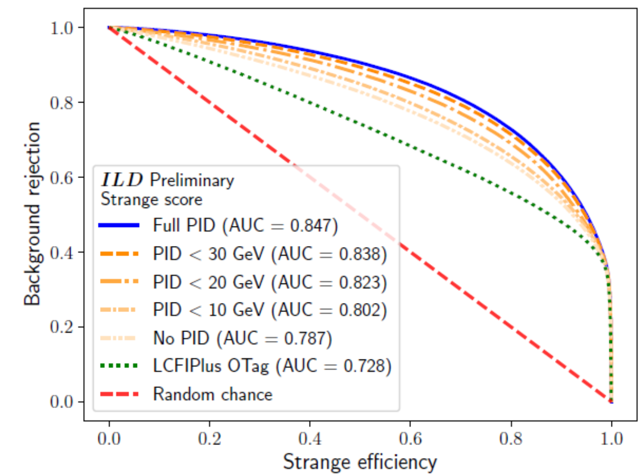
Higgs → $s\bar{s}$

- ILD 250 GeV, 900 fb^{-1} : κ_s < 7
  M. Basso, V. Cairo e.a. https://arxiv.org/abs/2203.07535
- Dedicated strange tagger (on top of LCFIPlus)
  → cuts background in half, with 85% of signal



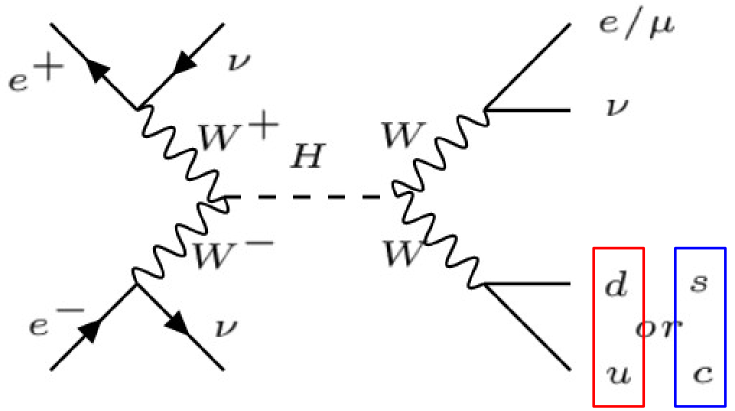


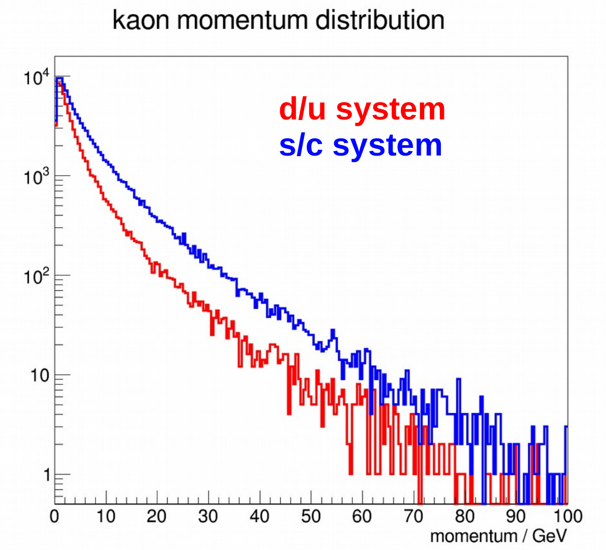




$W \to q\bar{q}$

- W decays mostly same-generationally
- CKM matrix, in particular V_{cs} = 0.97320 ± 0.00011
- Analogous to DELPHI analysis (120 W bosons),
  but with 10^8 Ws (ILC) → ± 0.00003 possible, independent of other CKM elements
- UE, https://ediss.sub.uni-hamburg.de/handle/ediss/9928



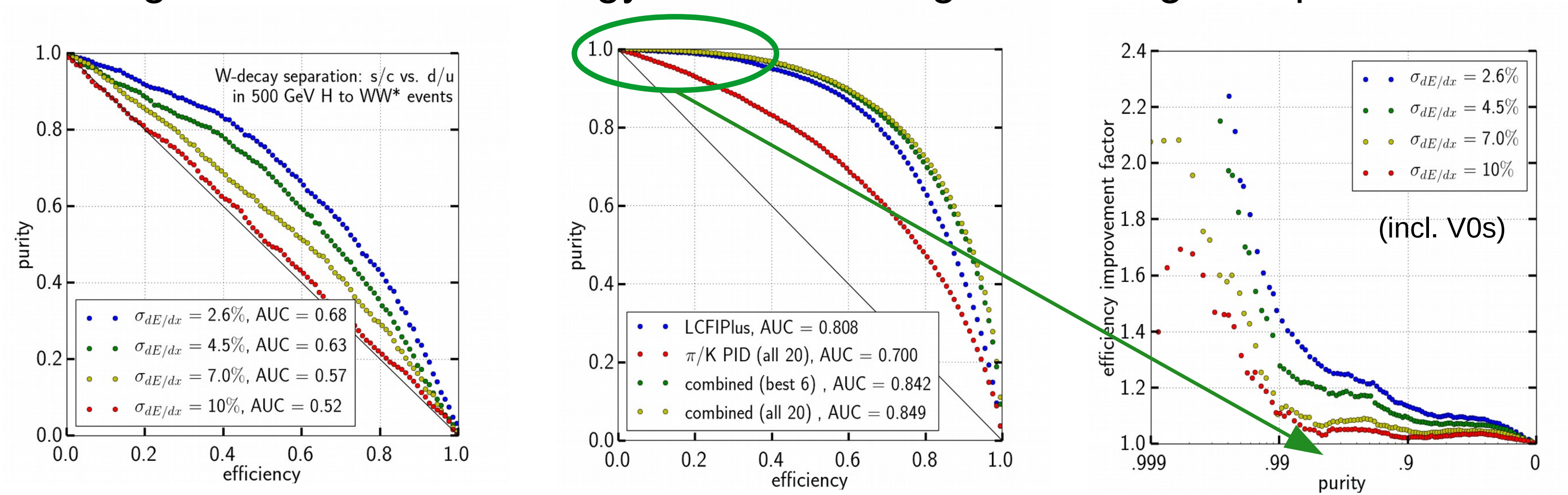


ILC 500 GeV

W → $q\bar{q}$

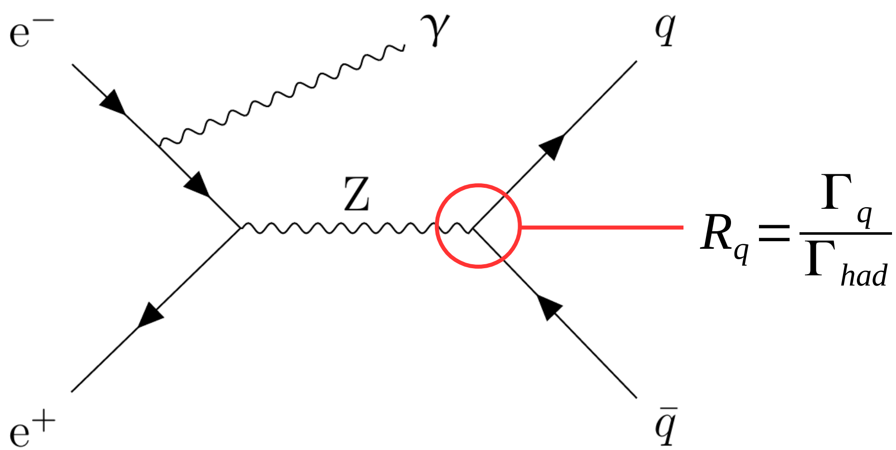
- Separation of W → d/u vs c/s via BDT, compare situation of flavour tag without PID (LCFIPlus) and with PID
- Direct depence on dE/dx resolution → 10% is pointless, 2.6% delivers substantial gains compared to default 4.5% (ILD)

  → At high center of mass energy also desire larger PID range of up to 100 GeV



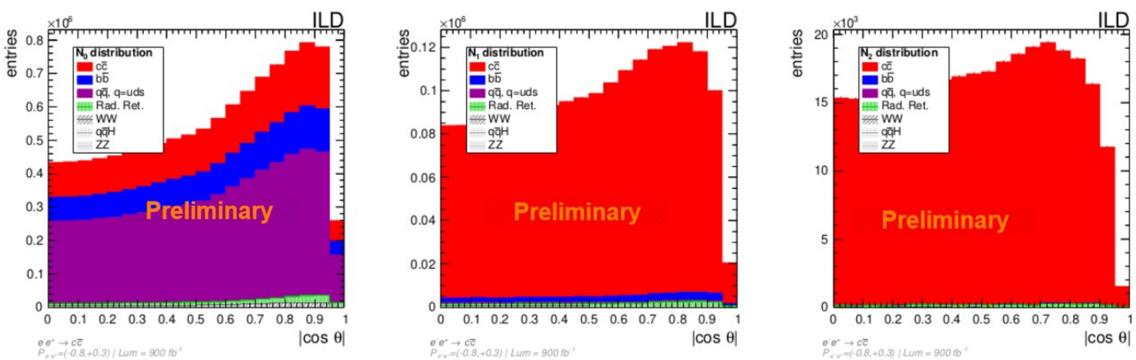
$Z \rightarrow q\bar{q}$

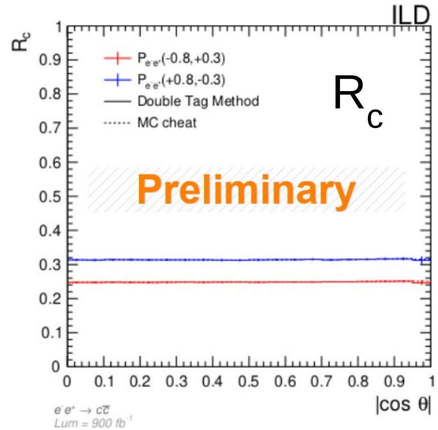
- Measure Z-q couplings R_q
- Via fractional cross section of tagged jets
- Pöschl, Irles, Marquez, e.a. (see backup) https://indico.desy.de/event/33640/contributions/127531
- Get very clean sample of quark species via double tagging

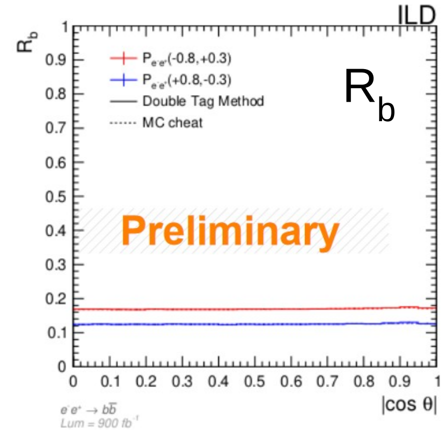


ILC 250 GeV

N0, N1, N2 for e+e- → $c\bar{c}$





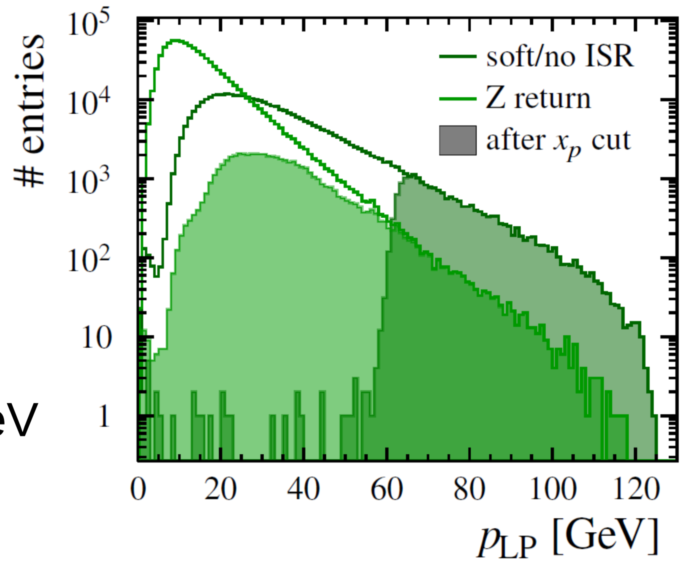


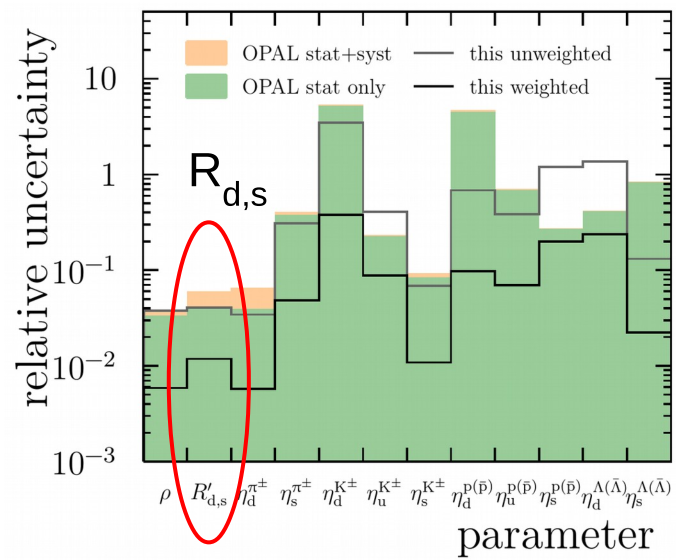
$Z \rightarrow q\bar{q}$

- Inclusive correlation between jet flavour and constituents via single and double tagging of individual leading particles (and assuming symmetries between q-to-contituent probabilities)
- P. Malek: https://ediss.sub.uni-hamburg.de/handle/ediss/9634

hadron type	single tags	double-tagged events $\pi^\pm$	$K^\pm$	$p(\bar{p})$	K^0_S	$\Lambda(\bar{\Lambda})$
$\pi^\pm$	1 148 510	8434	6370	4253	1160	426
$K^\pm$	632 973		3004	2749	1931	773
$p(\bar{p})$	307 653			460	492	659
K^0_S	160 730				176	311
$\Lambda(\bar{\Lambda})$	79 053					49

ILC 250 GeV



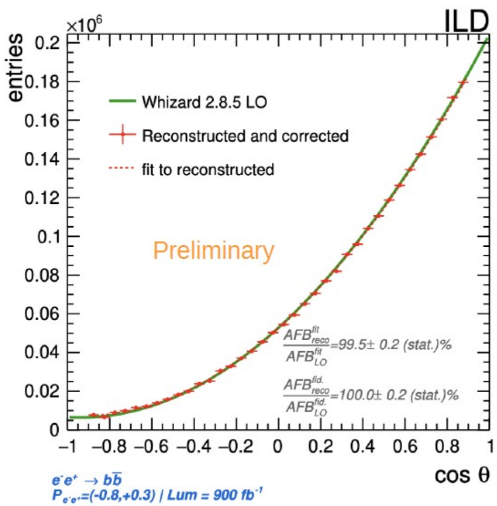


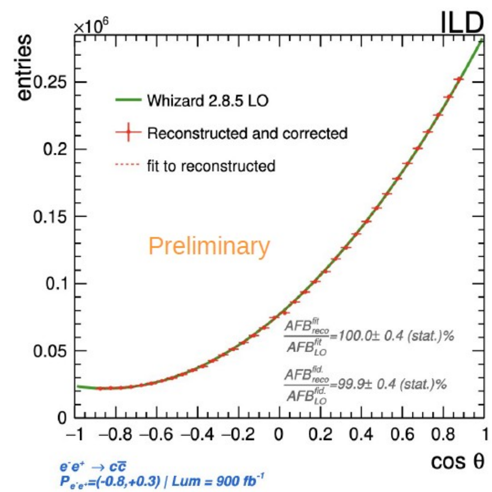
Z → qq A_{FB}

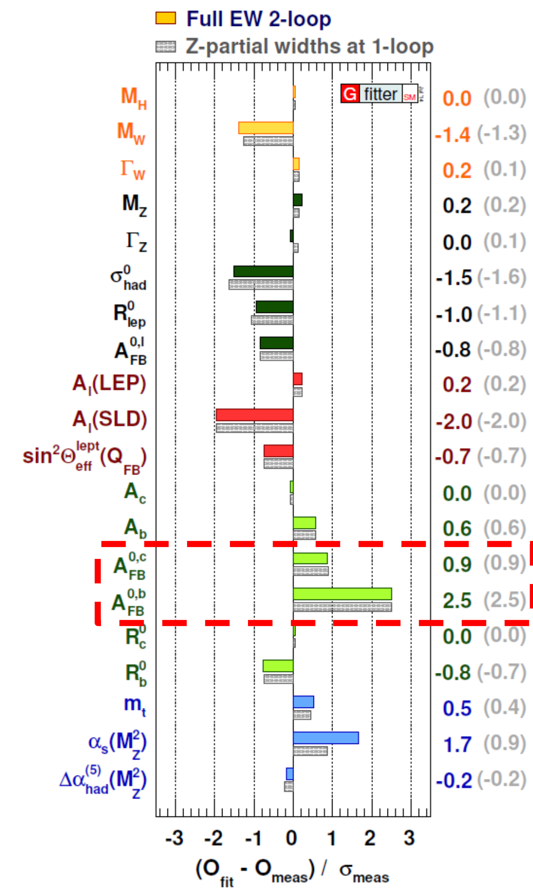
- Forward/backward asymmetry A_{FB} of s-channel exchange sensitive to new physics

$$A_{FB}^{q\bar{q}} = \frac{\sigma_F^{q\bar{q}} - \sigma_B^{q\bar{q}}}{\sigma_F^{q\bar{q}} + \sigma_B^{q\bar{q}}}$$

- Can be determined about one order better than before
- J. Márquez, https://agenda.linearcollider.org/event/10037/

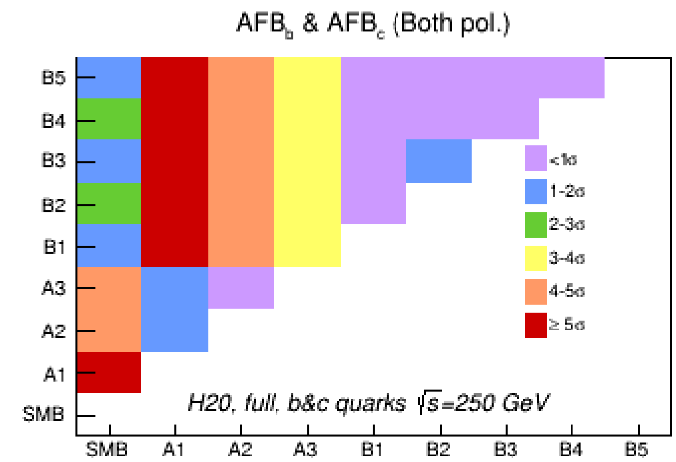


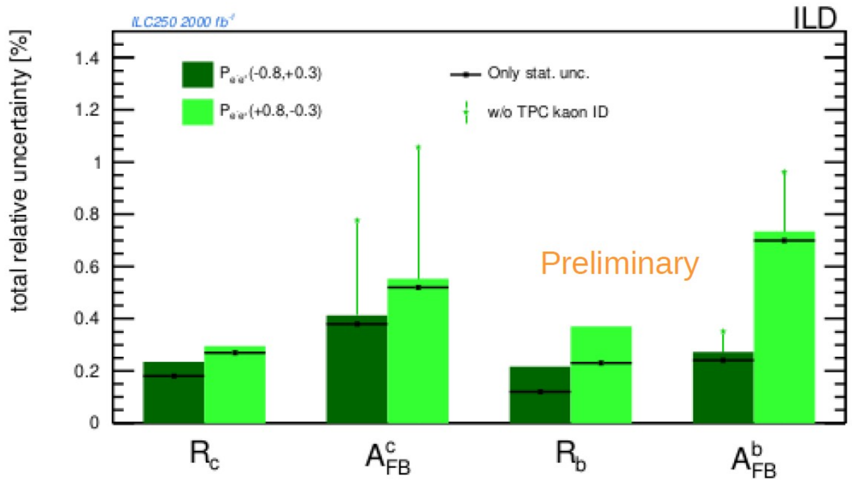


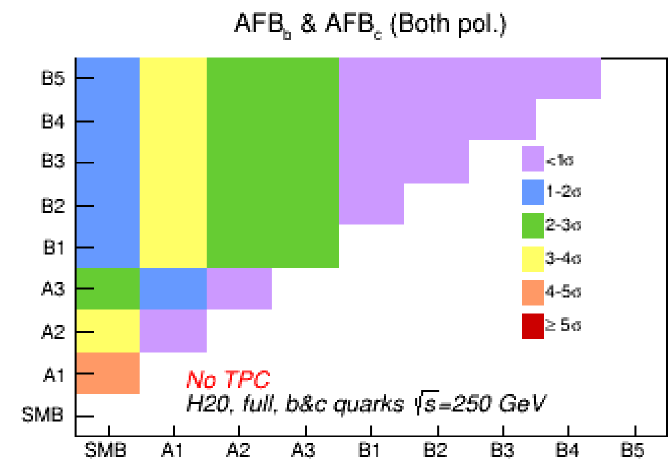


$Z \to qq$ A_{FB}

- With ILD TPC:
  - factor 2 better A_{FB} measurement
  - better separation of different BSM models wrt. each other in case of deviations from SM
- J. Márquez, https://agenda.linearcollider.org/event/10037/

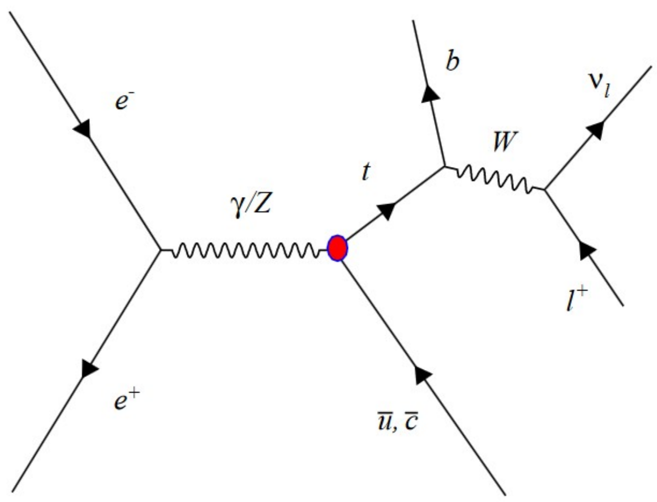






Flavour changing neutral currents / flavour violating processes

- New physics in rare cross-generational, non-SM top decays/couplings
- $t \rightarrow cH$, $t \rightarrow c\gamma$ @CLIC 380 GeV, 1 ab^{-1}: https://arxiv.org/abs/1807.02441
  - BR($t \rightarrow c\gamma$) < 2.6 10-5
  - BR($t \rightarrow cH(bb)$) < 8.8 10-5
- $tq\gamma$ and tqZ @FCC-ee 350 GeV, 3 ab^{-1}: https://arxiv.org/abs/1408.2090
  - BR($t \rightarrow q\gamma$) < 9.86 10^{-6}
  - BR($t \rightarrow qZ$ (vector)) < 1.41 10^{-5}
  - BR($t \rightarrow qZ$ (tensor)) < 5.27 10^{-5}



B Physics

- Z pole running allows for broad programme on B physics for CKM matrix and unitarity triangle(s)

  → all B (and D) species available, not all at Belle II

  → clean, low background environment, compared LHCb

  → only really relevant at 91 GeV, though

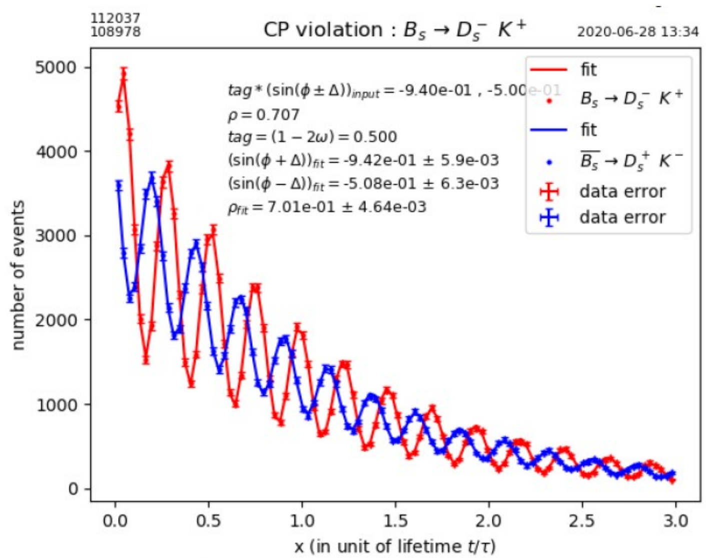
- Study rare decays with high accuracy and reconstruct full decay chain

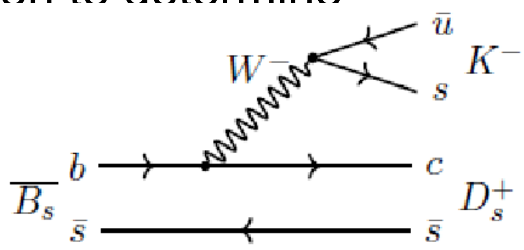
  → need to ID all decay products, i.e. often multiple charged hadrons, with high efficiency and purity!

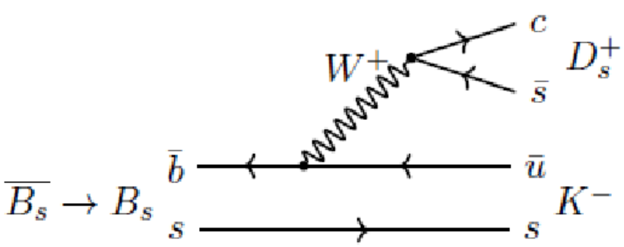
B Physics

- B_s-$\overline{B}_s$ oscillations in $B_s \rightarrow D_s K$
- R. Aleksan
  https://indico.in2p3.fr/event/23012/contributions/89990/
- Sensitive to (time dependent) CP violation and unitarity triangle angle ɣ
- Need to separate from nearby B decay modes, in particular $B_s \rightarrow D_s \pi$
- Need to tag opposite side B-meson to determine oscillation state at creation

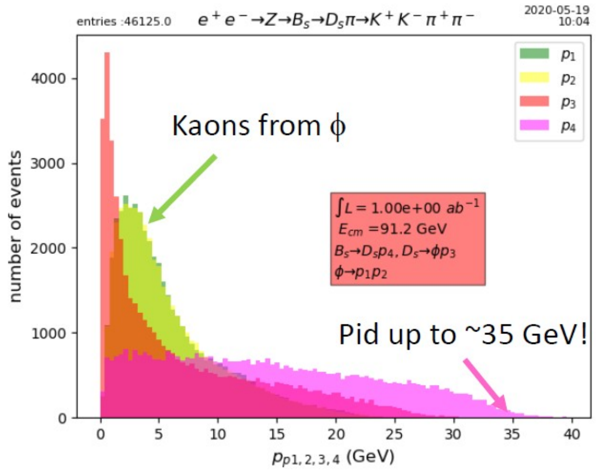
  → need high momentum resolution, JER and PID

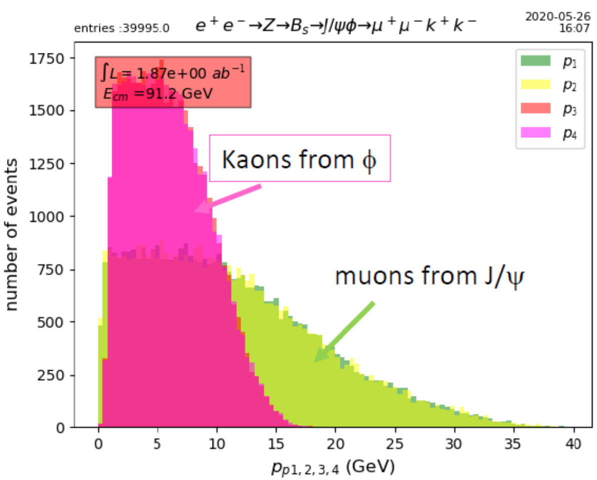




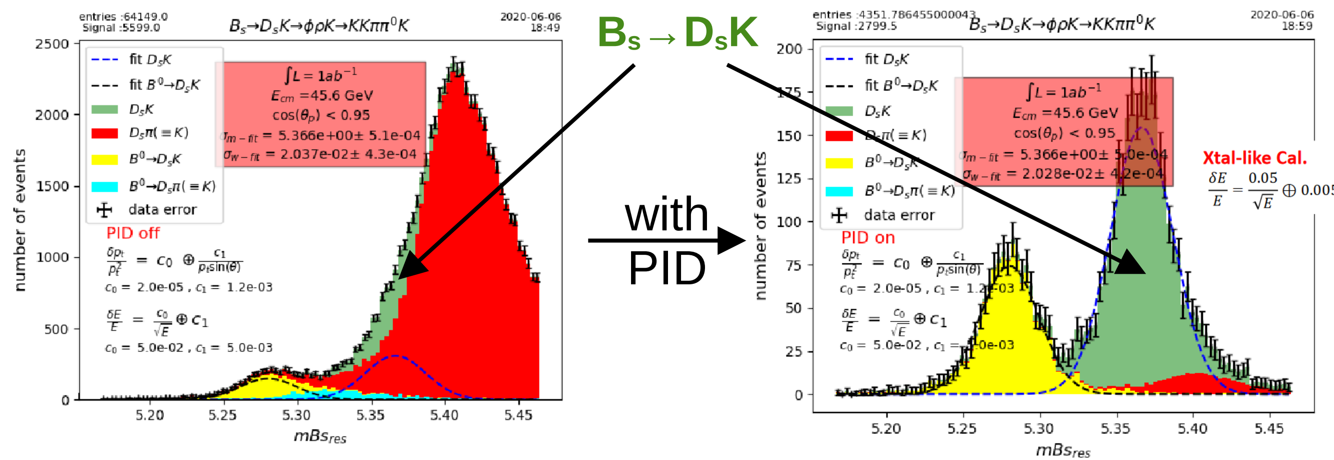


B Physics





R. Aleksan https://indico.in2p3.fr/event/23012/contributions/89990/



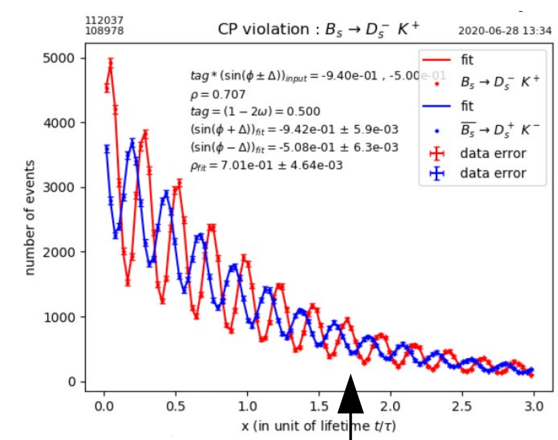
> 3 σ K/π separation up to 25 GeV (covers also K tagging), Ideally up to 35 GeV

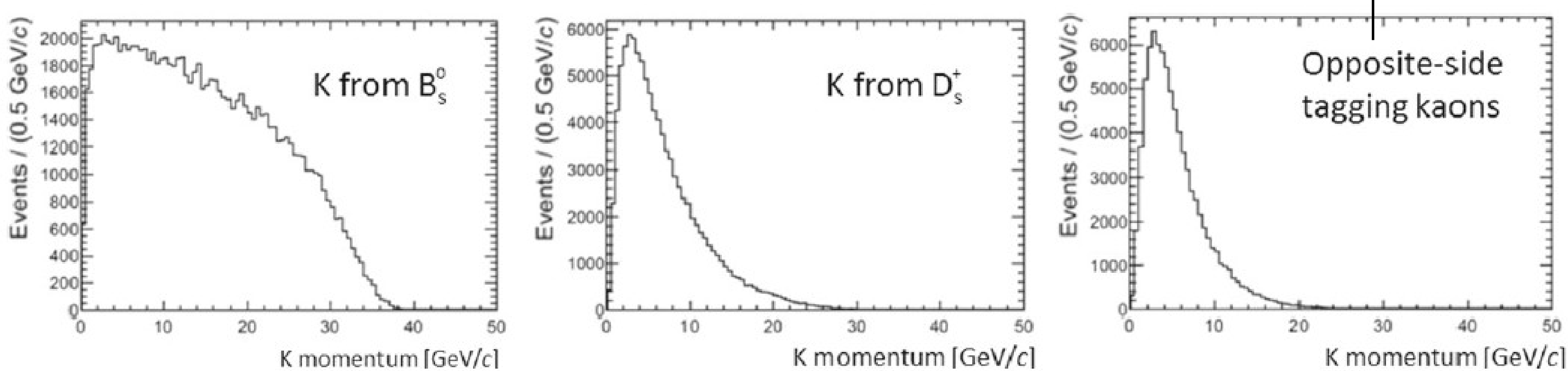
- Also need PID down to 0 GeV!

B Physics

- Similar analysis from G. Wilkinson, kaon spectra from near-0 up to 35 GeV

  https://doi.org/10.1140/epjp/s13360-021-01810-4



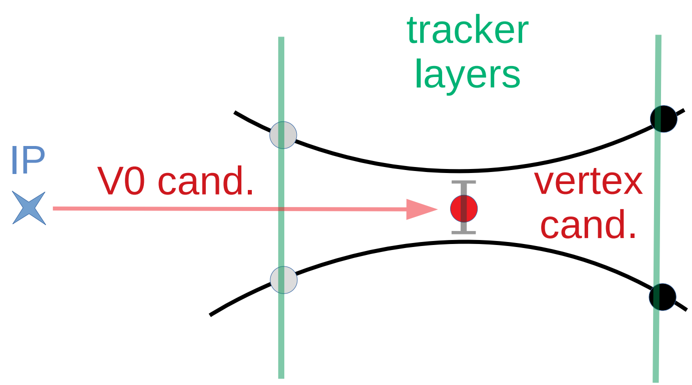


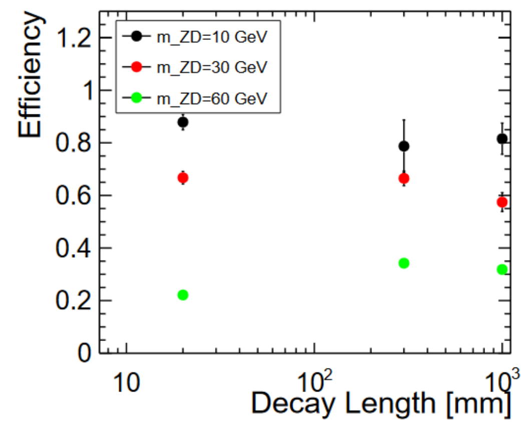
Momentum spectra for kaons occurring in Z^0 events containing a $B_s^0 \to D_s^{\pm} K^{\mp}$ decay.

Honorable mentions

- τ decays, e.g. BR(τ → ντ) to BR (τ → νK) for V_{us}

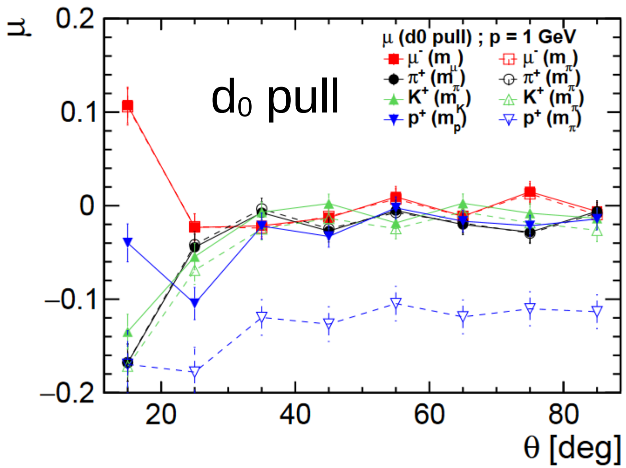
- LLPs → charged LLPs seen via ionisation (dE/dx / dN/dx),
  neutral LLPs have V0-like signature (e.g. T. Reisch)

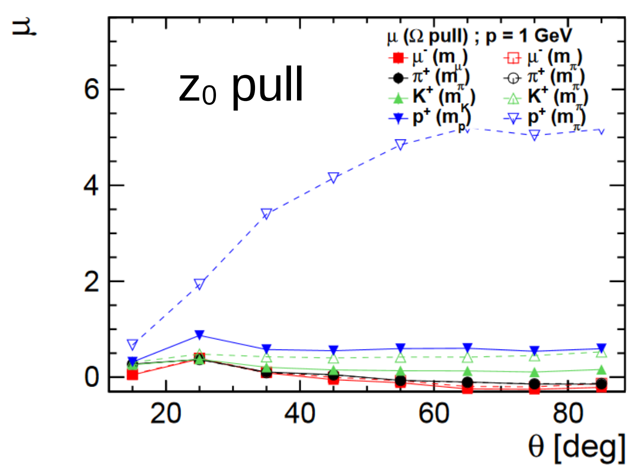


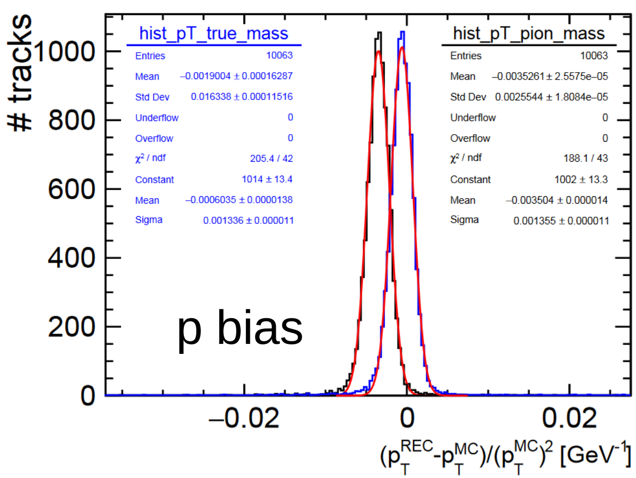


Track refitting with PID mass

- Y. Radkhorrami, https://agenda.linearcollider.org/event/8498
- All tracks are usually fitted assuming pion mass
- Fit with correct mass, i.e. correct energy loss from ionisation, gives
  - better estimates of track parameters
  - better estimates of momentum and impact prameters
- No tangible impact on vertex reconstruction found so far

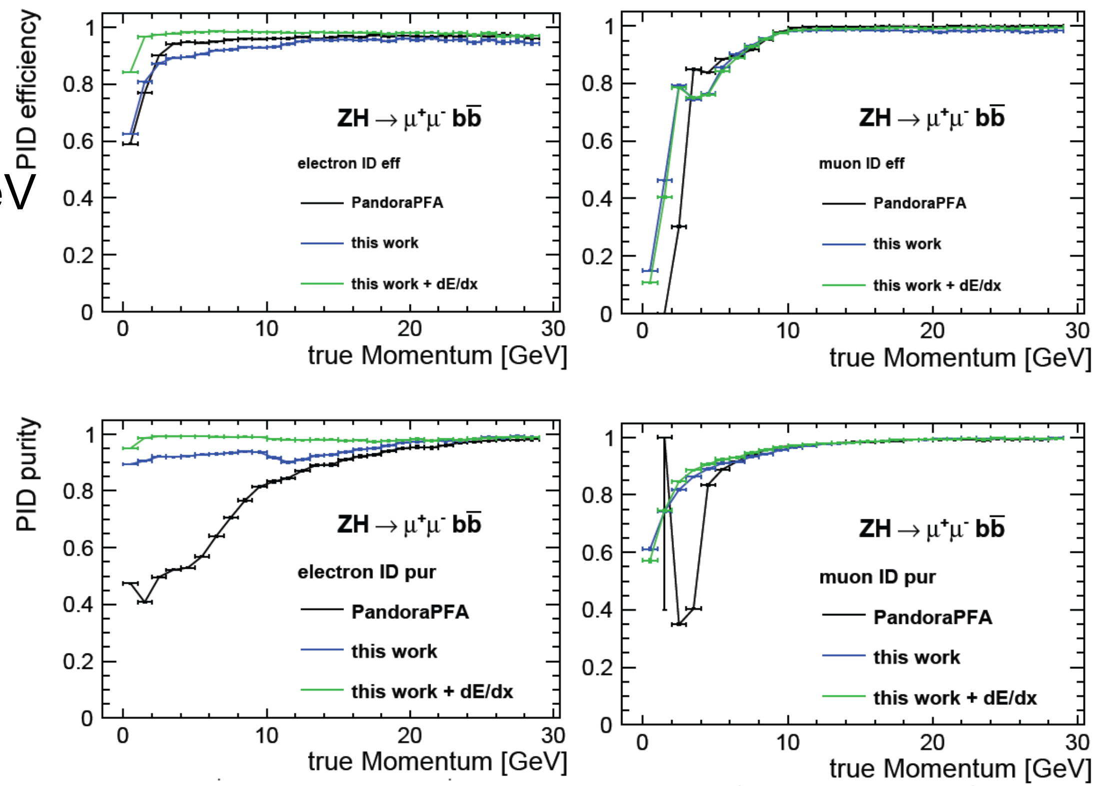


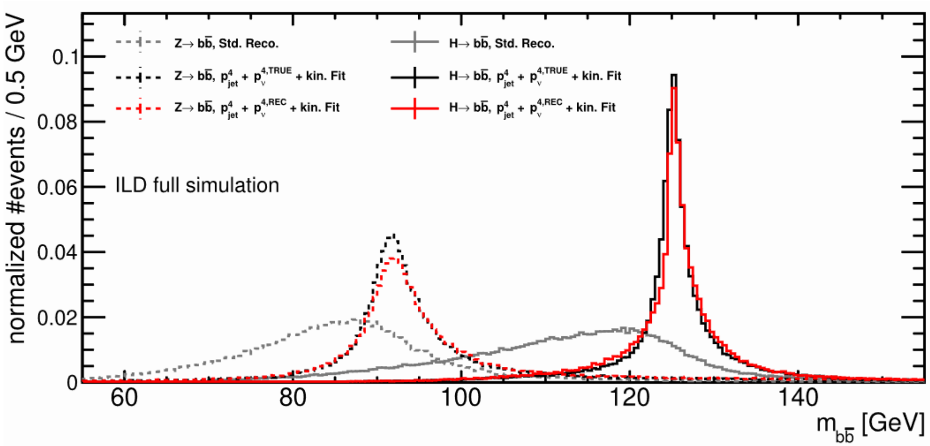




Electron & Muon ID

- Basic e and μ ID implemented in Particle Flow (e.g. Pandora), but room for improvement, in particular with dedicated PID at low momenta
- MSc Thesis L. Reichenbach
  PhD Thesis Y. Radkhorrami
- Here: applied to ZH→ μμ bb at 250 GeV with semileptonic B decays
- Used to separate Z and H peaks





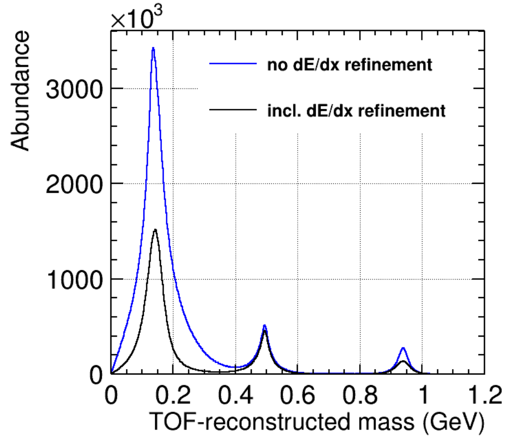
Electron & Muon ID

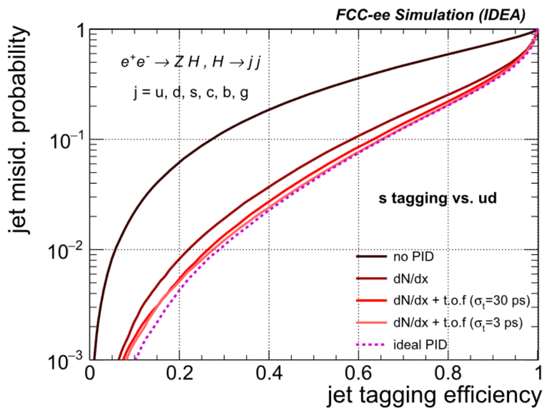
- Worthwhile to make sure, all charged particles are covered, not just pion/kaon
- Largely get this ‘for free’ with a dedicated charged-hadron ID system,

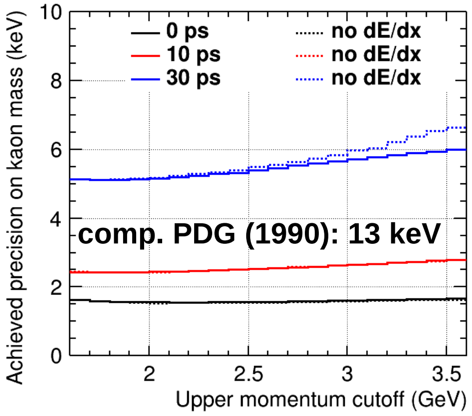
  BUT: muons often indistinguishable from pions, electron ID via TOF negligible

TOF

- Applicability to any existing detector concept, but works only at low momenta
- Semi-serious separate physics case: charged kaon mass
- Recent work for a proper implementation by B. Dudar, (e.g. https://indico.desy.de/event/33640/contributions/128388/) showed
  - limits in purity
  - no smoking gun physics case for high center of mass energies
- M. Selvaggi may disagree
- Need to know track length as well as time:
  - L = 1.8 m, t = 6 ns with σ_t = 60 ps → σ_L = 18 mm







Further Considerations

- Common performance assessment: separation power definition
- PID software
- Flavour tagging software

PID Performance Assessment

- Should use a coherent observable, mostly done with pion/kaon seapration power
- But: What is the exact definition?
  How are the underlying spectra defined?
  → sim / detailed sim / measurements
  Are they sufficiently Gaussian?
  → dE/dx, dN/dx: yes, TOF: no
- Should use a unified approach for comparability!
  Dedicated discussion desired?
- Proposal: p-value assessment

$$S = \frac{\mu_K - \mu_\pi}{\sigma_\pi}$$

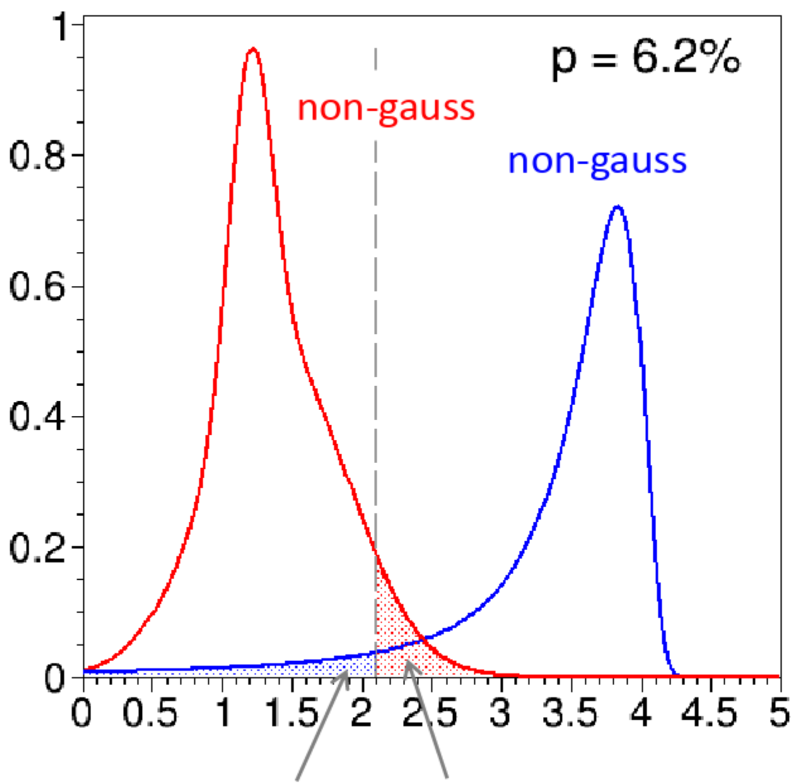
$$S = \frac{\mu_K - \mu_\pi}{1/2\left(\sigma_K + \sigma_\pi\right)}$$

$$S = \frac{\mu_K - \mu_\pi}{\sqrt{1/2\left(\sigma_K^2 + \sigma_\pi^2\right)}}$$

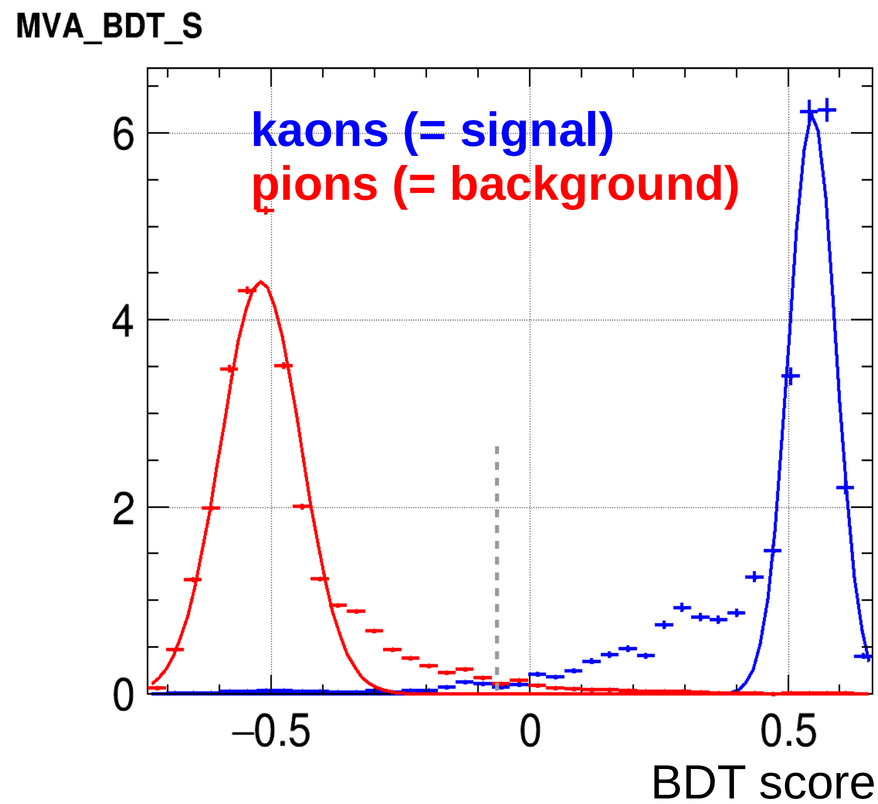
$$S = \frac{\mu_K - \mu_\pi}{\sqrt{\sigma_K^2 + \sigma_\pi^2}}$$

p-value Assessment

- Find cut with mis-ID = 1 - efficiency = p-value → find Gaussian quantile → compute Z = 2 · quantile of standard Gauss



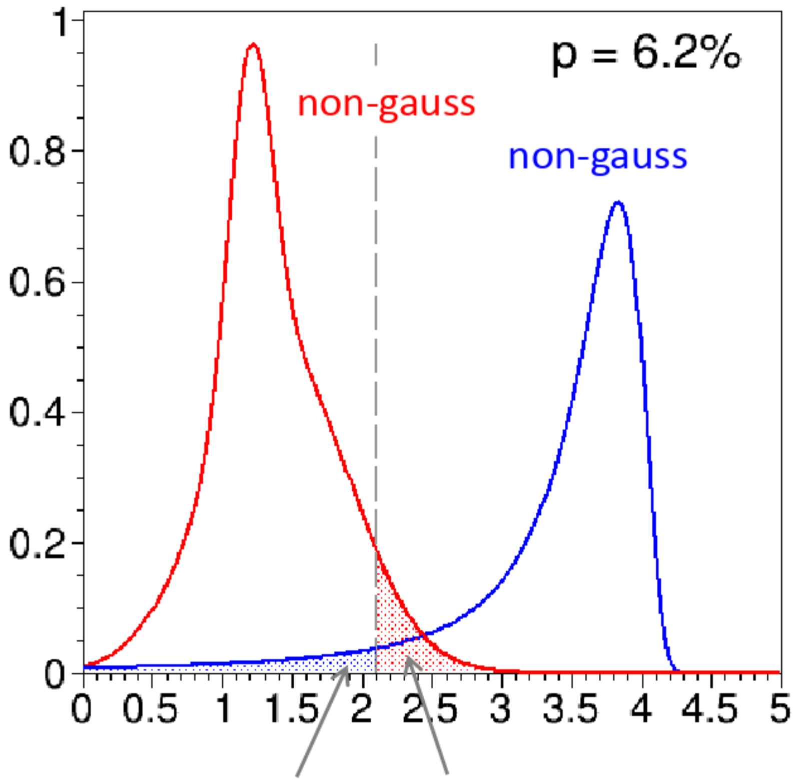
mis-ID = 1 - efficiency = 6.2%



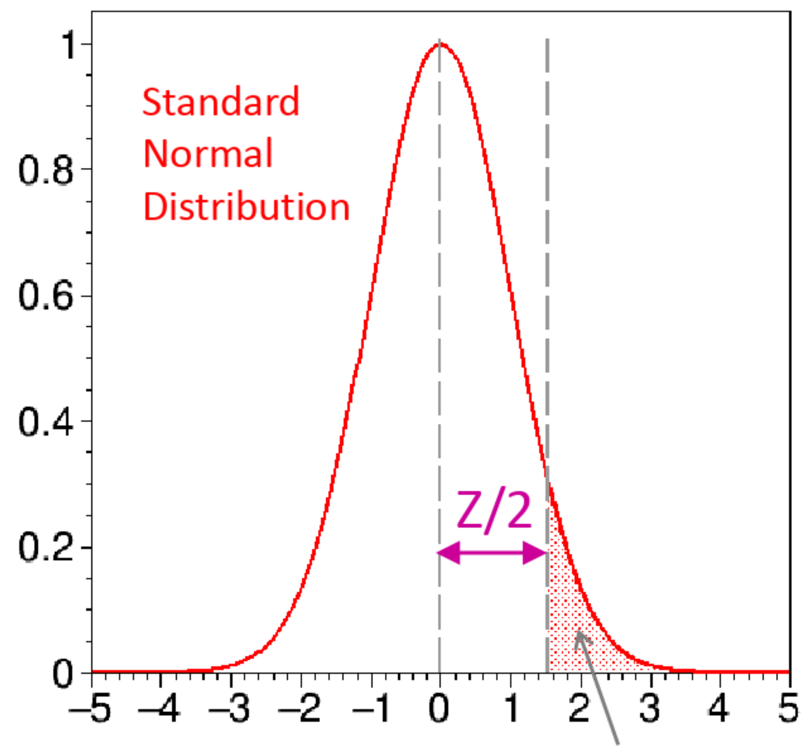
K. Götzen: https://indico.gsi.de/event/7080/contributions/31950/attachments/22952/28789/pid_kgoetzen_separationpower.pdf

p-value Assessment

- Find cut with mis-ID = 1 - efficiency = p-value → find Gaussian quantile → compute Z = 2 · quantile of standard Gauss



mis-ID = 1 - efficiency = 6.2%

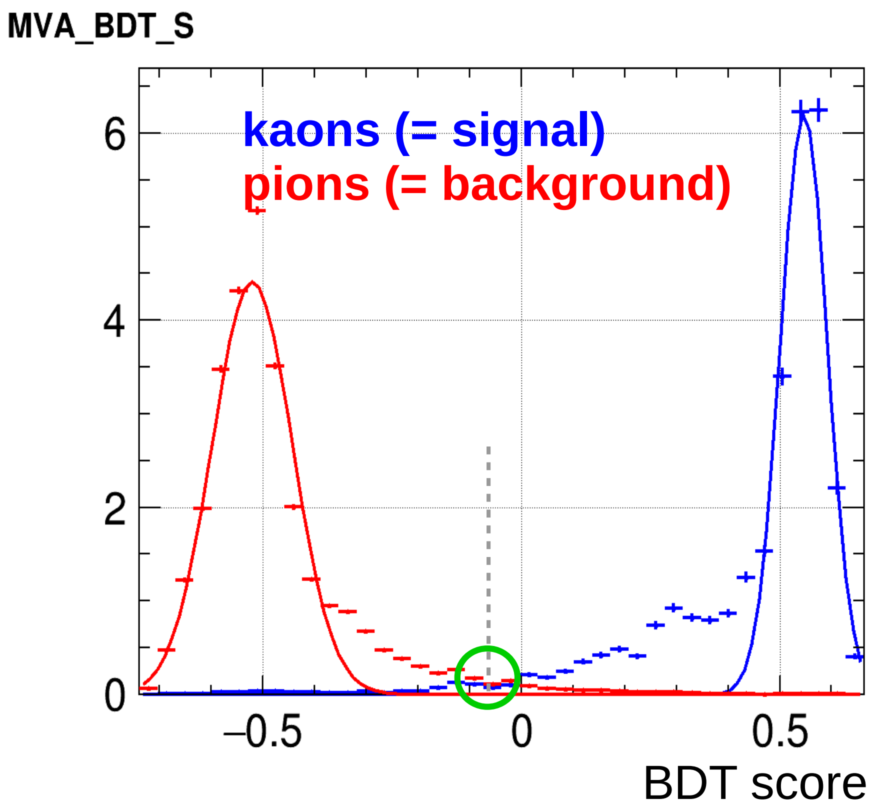


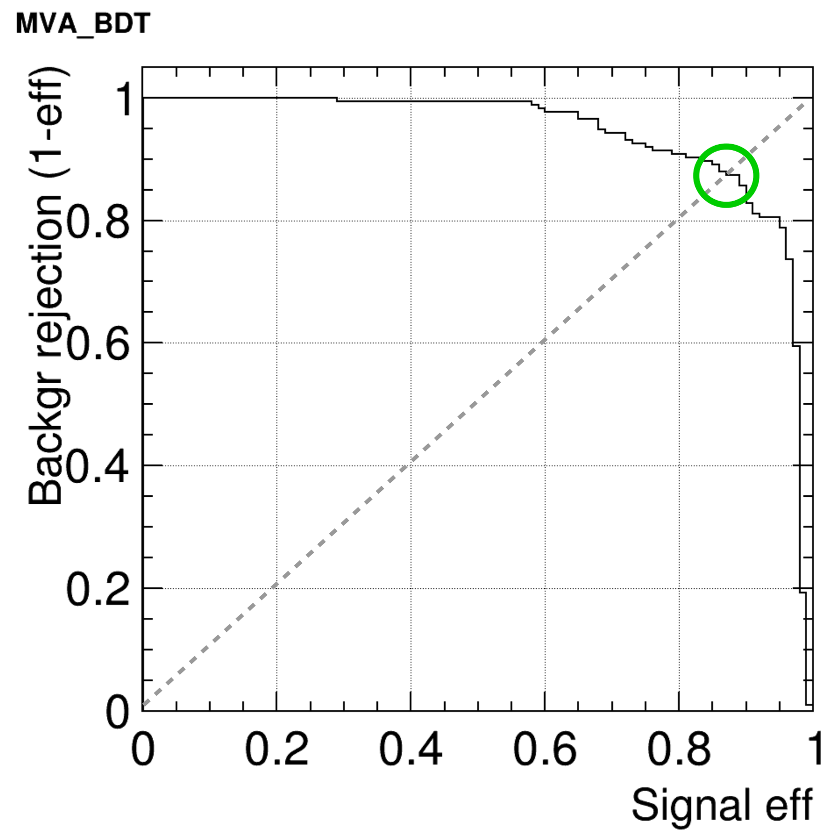
p = 6.2%

K. Götzen: https://indico.gsi.de/event/7080/contributions/31950/attachments/22952/28789/pid_kgoetzen_separationpower.pdf

p-value Assessment

- ‘Central tail split’ of BDT score is equivalent to crossing point of ROC curve with x=y line

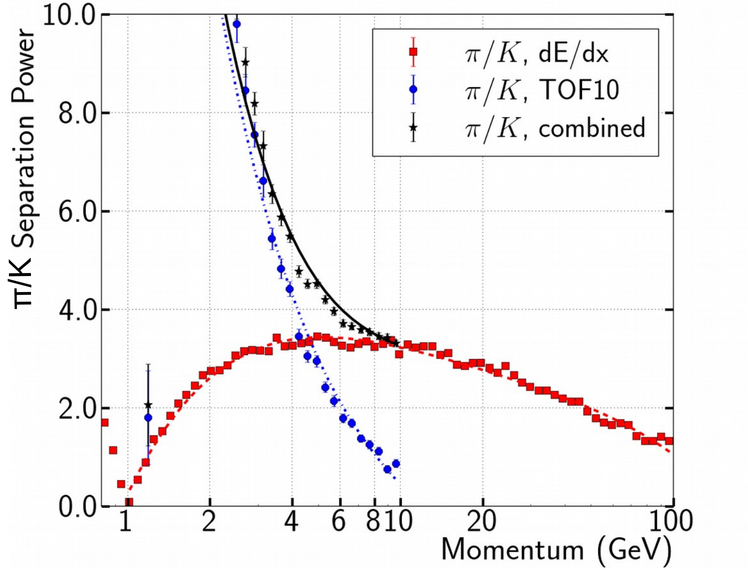




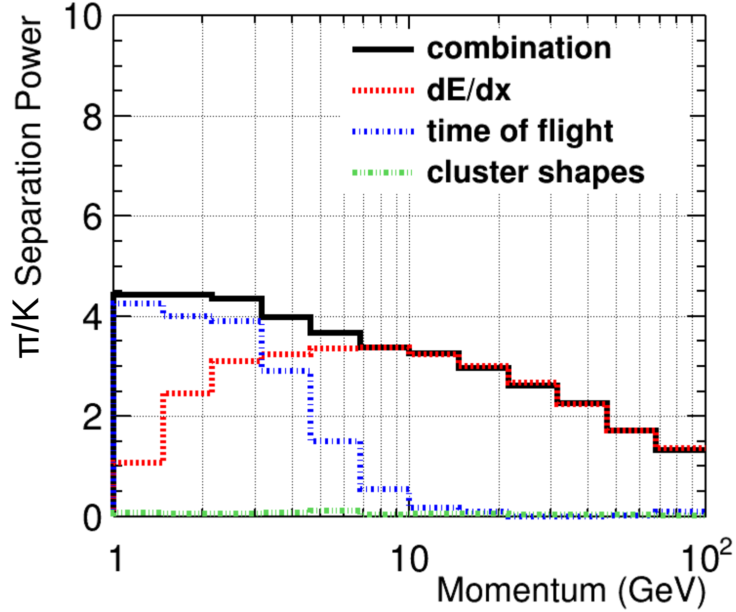
p-value Assessment

- Combination of dE/dx + TOF in detailed simulation shows that TOF ‘flattens out’ at low momenta → misinterpreted tracks create remaining background, which limits S
- ‘Arbitrary’ separation power at p < 3 GeV not realistic, but still very much good enough with S > 4, covers the dE/dx blind spot → less fancy, more honest

analogue to ILD Interim Design Report

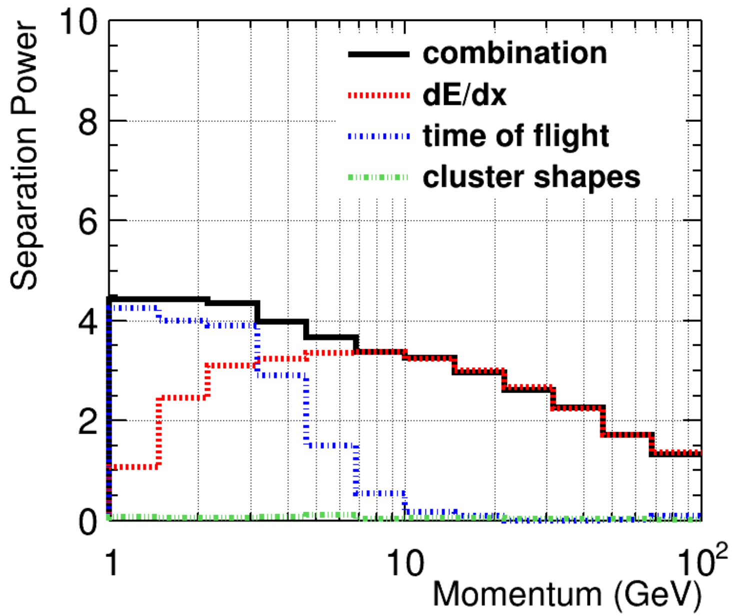


CPID, $q\bar{q}$ events, BDT result



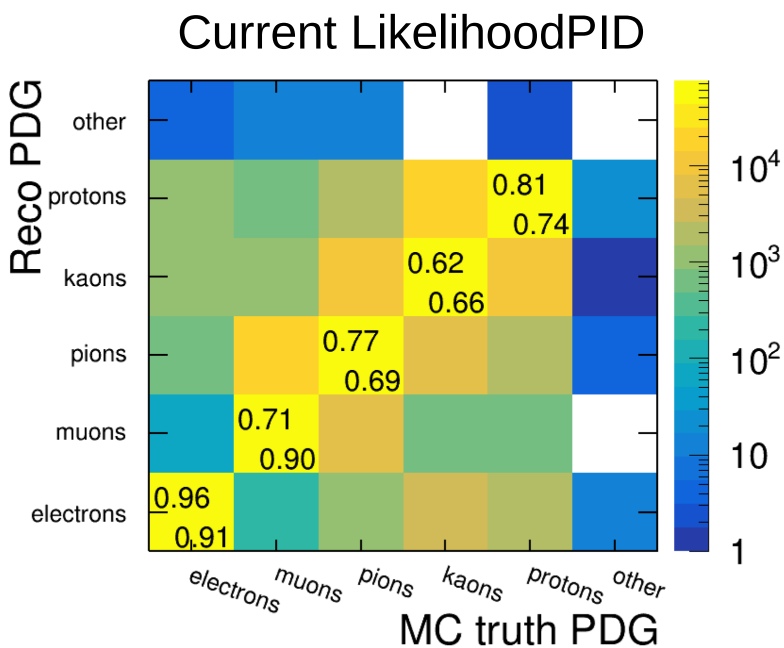
PID Software

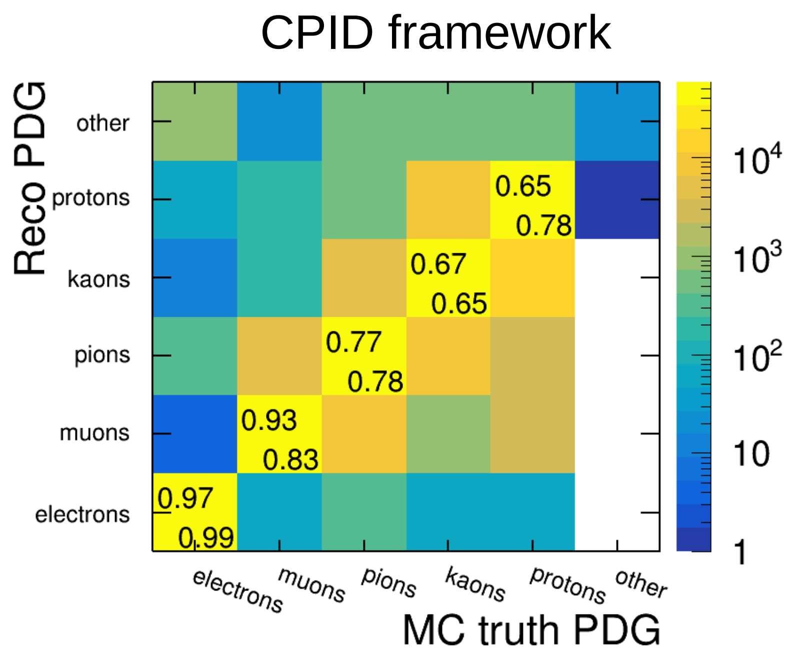
- New modular framework “Comprehensive PID” allows to select different PID systems as well as different MVA methods via steering file
- Work in progress, atm ILD (iLCSoft), soon Key4HEP
- Target: easy to use, easy to compare different PID situations



PID Software

- Here: multiclass BDT; confusion matrix with $^{eff}/_{pur}$ on diagonal
- Simple BDT already generates similar performance to current LikelihoodPID

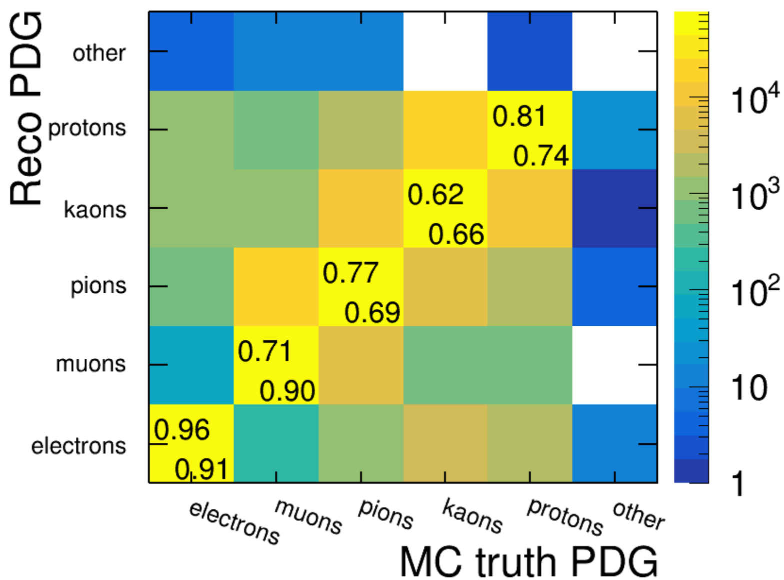




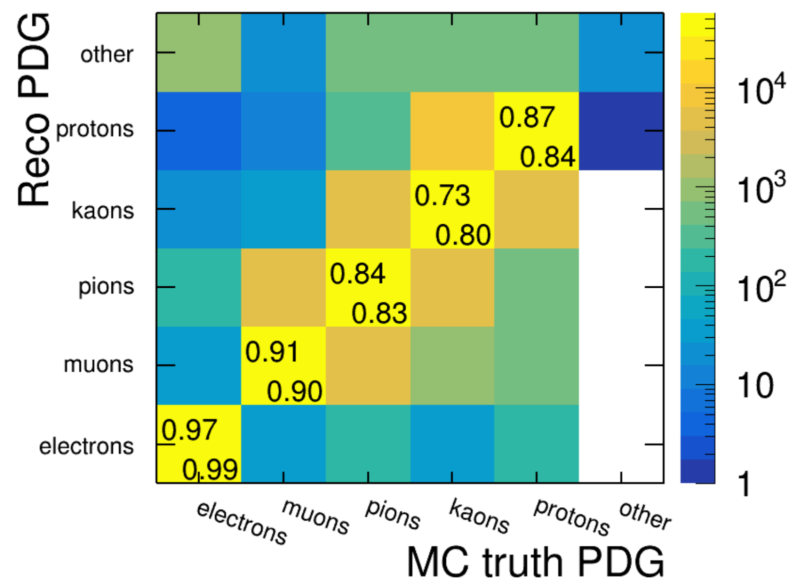
PID Software

- Here: multiclass BDT; confusion matrix with eff/pur on diagonal
- Simple BDT already generates similar performance to current LikelihoodPID
- Addition of TOF gives immediately better result – previously hard, easy in CPID

Current LikelihoodPID



CPID framework; + TOF



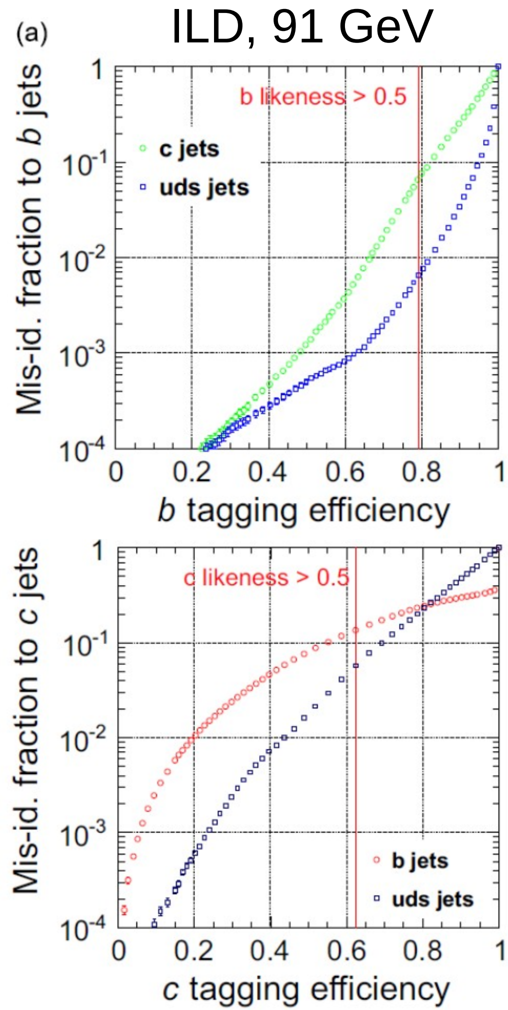
Flavour Tagging Software

- LCFIPlus: work horse in LC community for many years

  https://github.com/lcfiplus/LCFIPlus

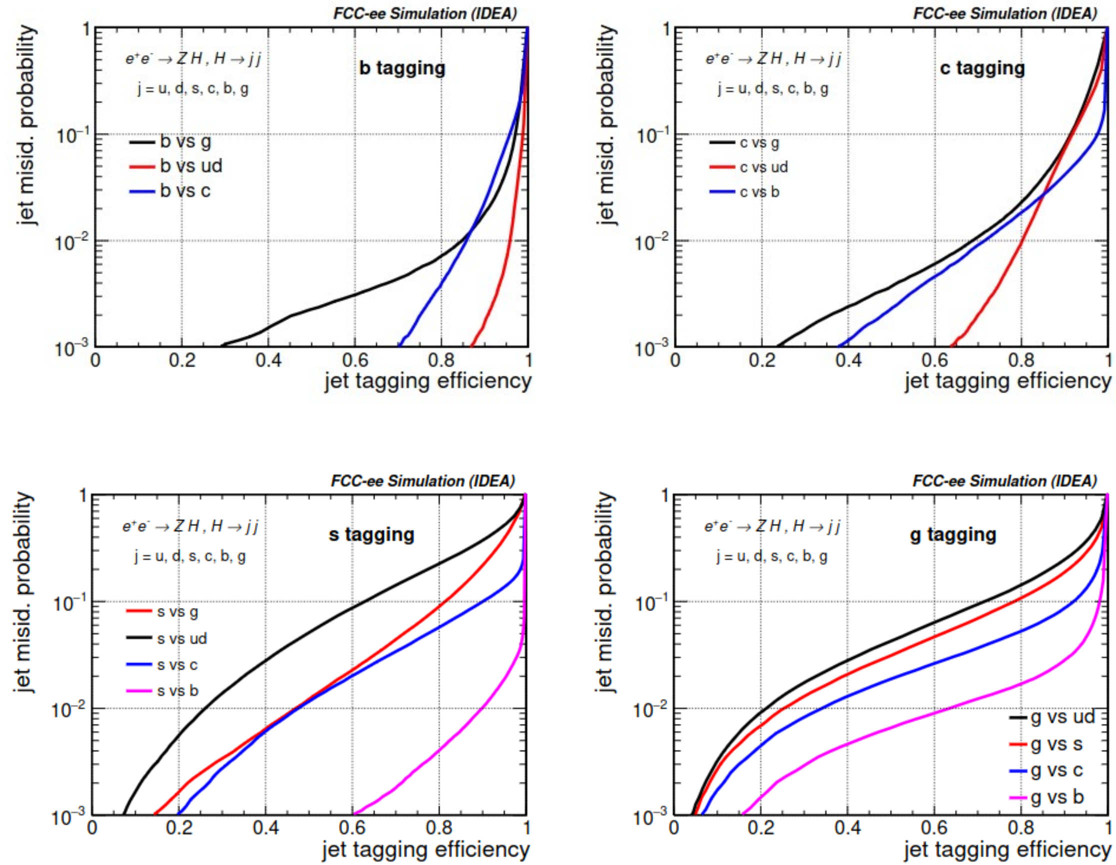
  https://indico.desy.de/event/33640/contributions/128011/

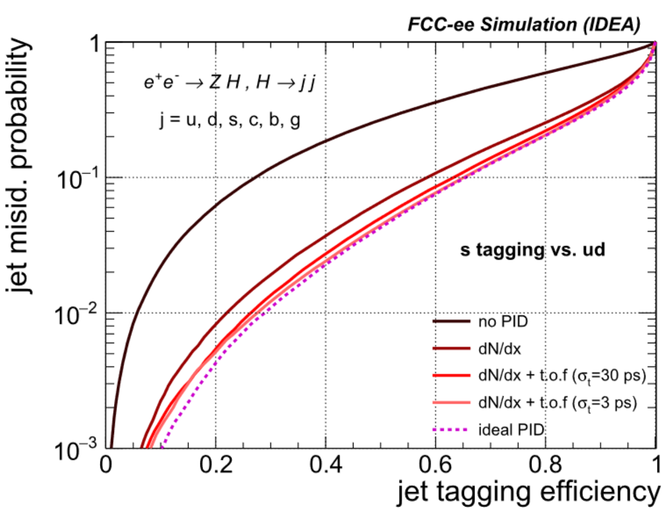
- ILD, SiD, CLICdp, e.a.; full sim
- BDT-based; DNN and GNN under study; no s-tag for now
- Proposal to include dE/dx-derived PID observables



Flavour Tagging Software

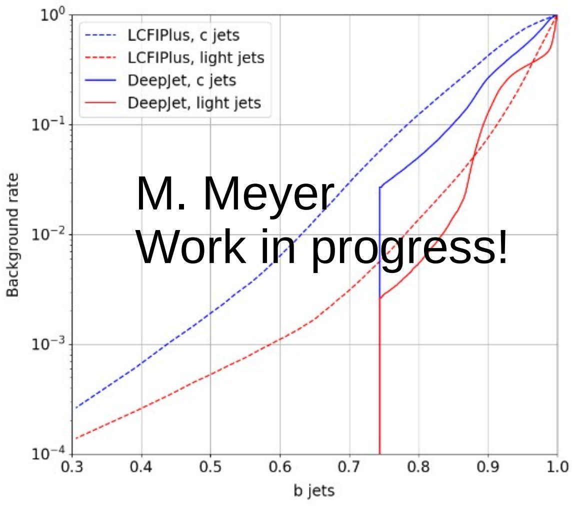
- ‘FCC’-tagger:
  https://arxiv.org/abs/2202.03285
- IDEA; DELPHES sim
- ParticleNet-based
- Using dN/dx and TOF mass
- Excellent tagging performance based on (more advanced) IDEA





PID & Flavour Tagging Software

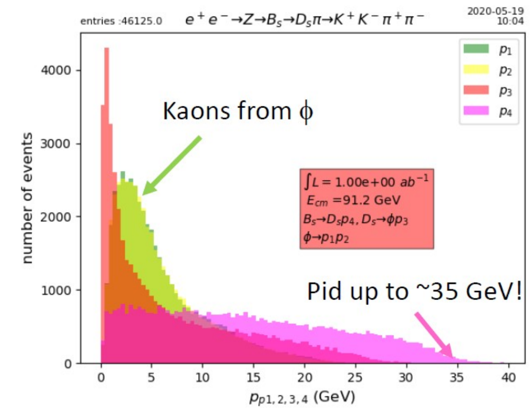
- Potentially more PID & flavour taggers under development
- Individual taggers (s-tagger, c/s-tagger, etc. prevalent)

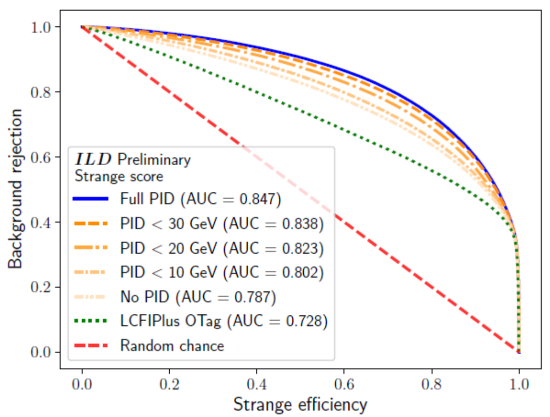


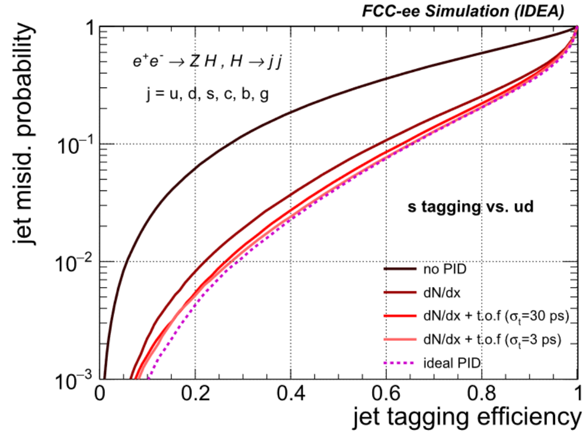
- Make sure to coordinate and not reinvent the same product multiple times
- Make sure frameworks are broadly available (→ Key4HEP!), easy to use and easy to re-train for studies with alternative detector layouts!

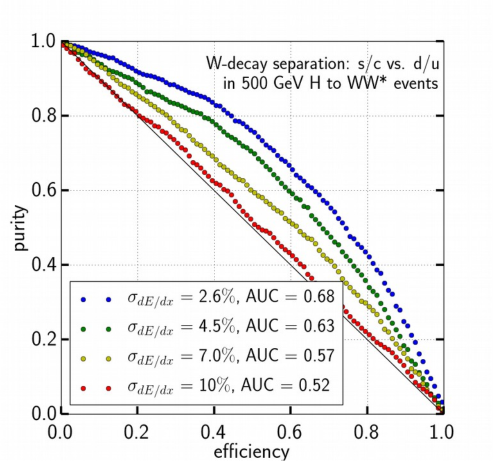
Conclusions I

- Various studies done, overall picture:
  - need to cover sub-1-GeV up to 30 GeV with $S_{\pi/K} > 3$
  - sky's the limit!
  - larger range desired for larger c.o.m. energies









Conclusions II

- More studies needed
- With detailed simulation to cover correlations and intricacies
- Define benchmark points
  - physics cases
  - high level performance observables
- Use common software
- In an analysis, make a plot with and without PID, ideally with varying performance
- Open question: hermeticity?

Thanks

- Emmanuel Perez, Patrizia Azzi, Philipp Roloff
- Adrián Irles, Jesùs Márquez
- And to all the contributions used in this talk!

Backup

$Z \rightarrow q\bar{q}$ A_{fb} & R_q References

- S. Bilokin's PhD thesis https://tel.archives-ouvertes.fr/tel-01826535
- $e^+e^- \rightarrow bb$, 2019 https://agenda.linearcollider.org/event/8147
- $e^+e^- \rightarrow tt$, bb 2019 https://confluence.desy.de/download/attachments/42357928/ILD-PHYS-PUB-2019-007.pdf
- $e^+e^- \rightarrow cc$, 2020 https://arxiv.org/abs/2002.05805
- $e^+e^- \rightarrow bb/cc$, ss 2021 https://agenda.linearcollider.org/event/9440, https://agenda.linearcollider.org/event/9285
- $e^+e^- \rightarrow bb/cc$ 2021 https://agenda.linearcollider.org/event/9211/contributions/49358/
- overview, all quarks: https://indico.desy.de/event/33640/contributions/127531