

$e^-e^+ \rightarrow ss$ at ILC 250 GeV



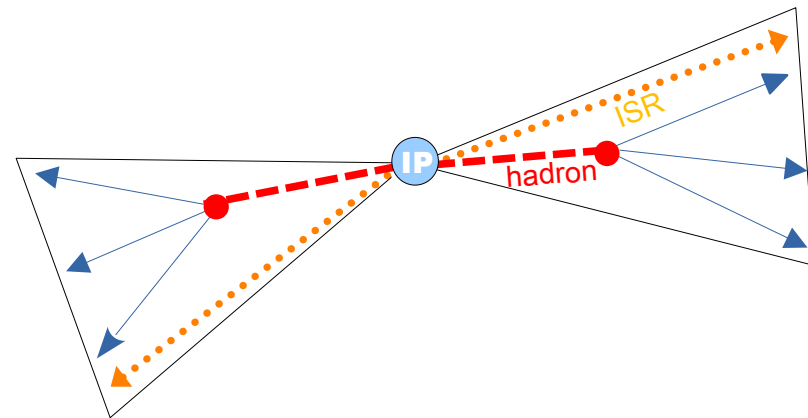
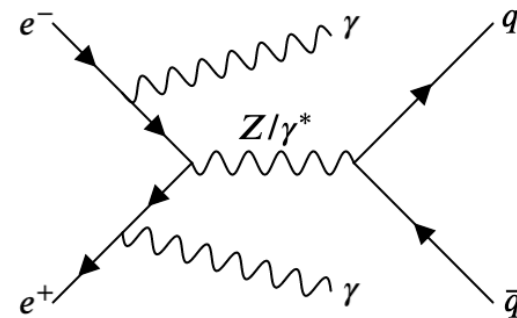
Jesús P. Márquez Hernández

15/10/25

(b & c) diquark production in e^-e^+ collisions

- MC simulations at 250.
 - International Linear Collider (ILC) run plan.
 - Full simulation of the International Large Detector (ILD).
- Topology: Two back-to-back jets.
- Procedure:

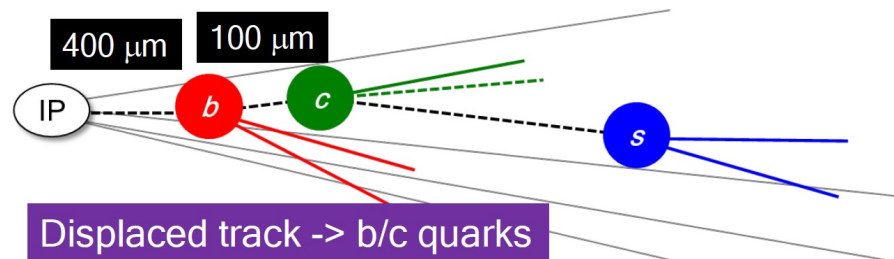
- R_q
 - Background suppression → Selection of $q\bar{q}$ events.
 - Flavor tagging → Selection of $b\bar{b}$ & $c\bar{c}$ events.
 - Double tagging (b-tag, c-tag).
- A_{FB}
 - Charge measurement → Quark-Antiquark identification.
 - Double charge.



**How can we move from here to strange quarks (or u/d quarks)?
Can we get %o-level uncertainties like for the b & c quarks?**

From b/c to strange quark

- Flavor tagging of b and c jets is “easy”:
 - Decay of b and c hadrons: Displaced vertexes from de IP, at a distance ($\tau_q \cdot c$)



- But the strange quark produce kaons... no decays in the tracker to be used
 - We need to build/use an s-tag relying on **kaon PID**
 - Our first attempt is a “classic” cut-based analysis
 - I worked on top of the previous analysis done by Y. Okugawa in his thesis, (directed by R. Poeschl)

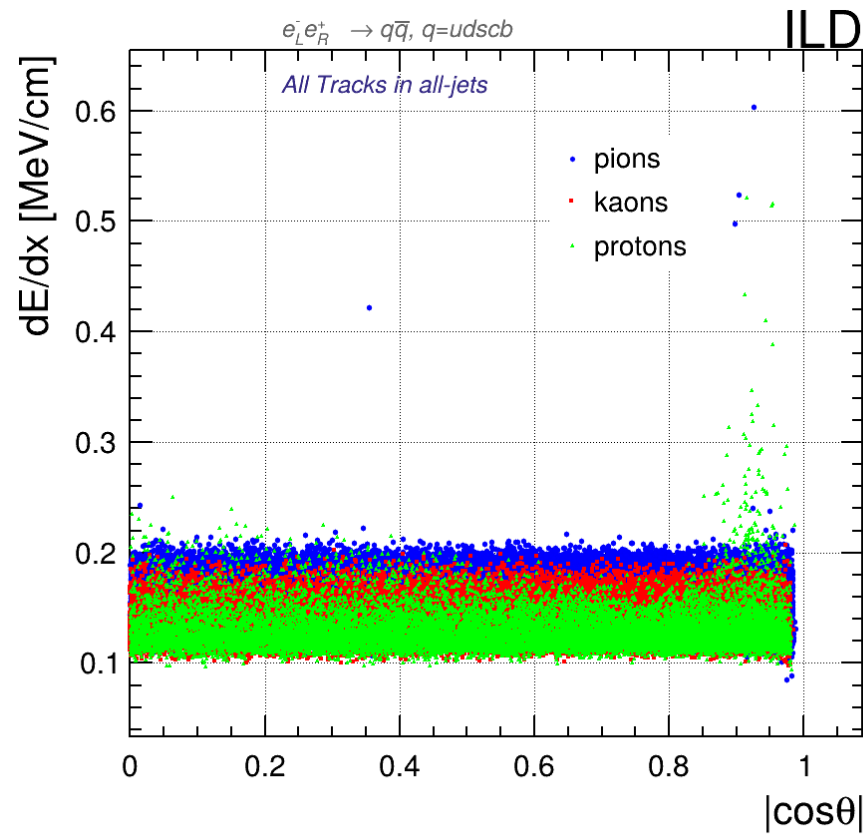
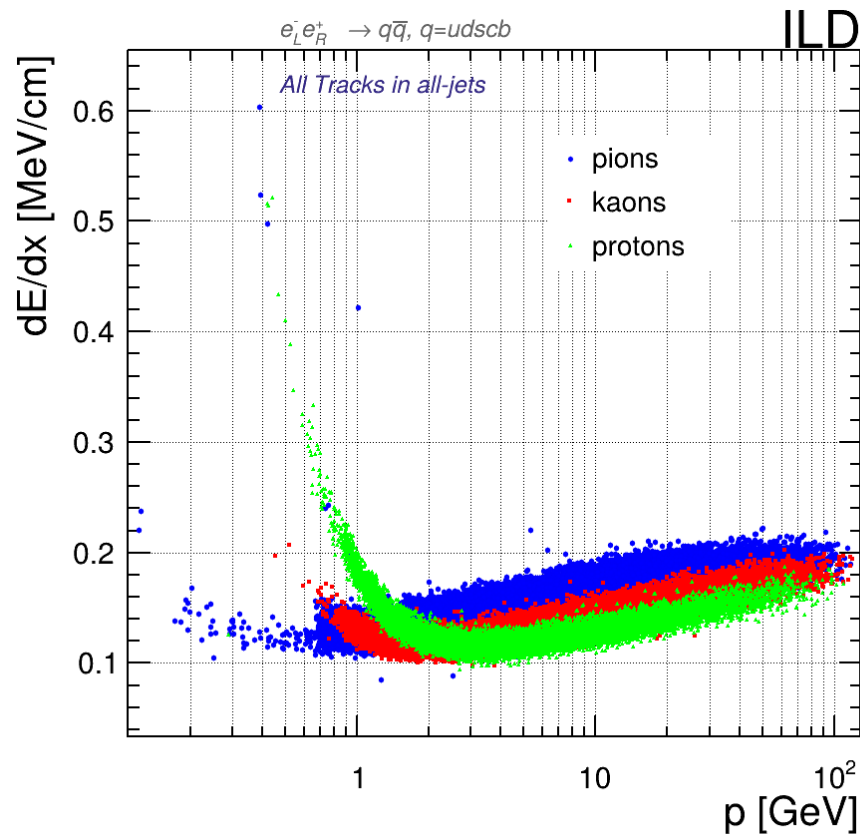
Redoing of the ssbarAnalysis

- Preselection of the diquark signals (Modification of Yuichi Okugawa's analysis)

	#	Name	Quantity	Description
uds selection	1	b -tag	$btag < 0.3$	Reject events with b-like jets
	2	c -tag	$ctag < 0.65$	Reject events with c-like jets
	3	nvtx	$nvtx = 1$	Jets should have only PV as vertex
Cut-based s-tag (or ud-tag)	4	Leading momentum	$p_{LPFO} > 15 \text{ GeV}$	Leading momentum cut
	5	LPFO acollinearity	$\cos \theta_{LPFO1,2} > 0.97$	LPFOs should be back-to-back
	6	Offset	$V_0 = \sqrt{d_0^2 + z_0^2} < 1 \text{ mm}$	Offset cut to reject Λ_0 contribution
	7a	dE/dx PID (π)	New angular k-distance cuts	$\pi^\pm$ identification
	7b	dE/dx PID (K)		$K^\pm$ identification
Migration correction	8	SPFO	Veto $p_{SPFO} > 10 \text{ GeV}$ and charge opposite to LPFO.	Attenuate the charge migration by rejecting oppositely charge LPFO competitor
	9	Charge	$Q_{LPFO1} \times Q_{LPFO2} < 0$ opposite charge.	Charge of LPFOs from both sides has

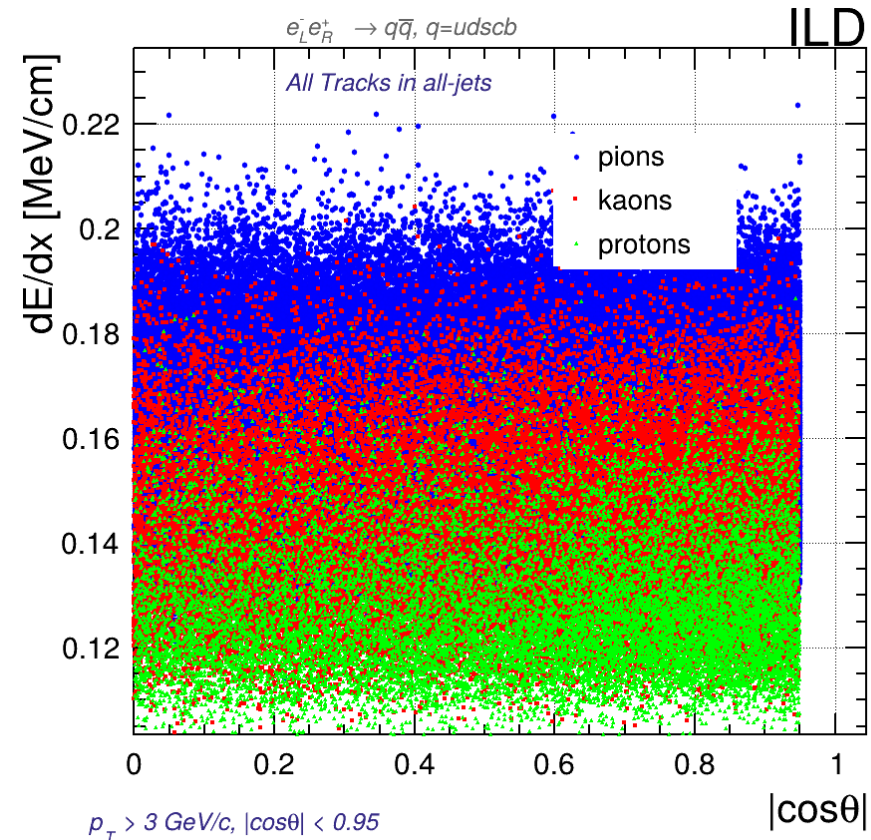
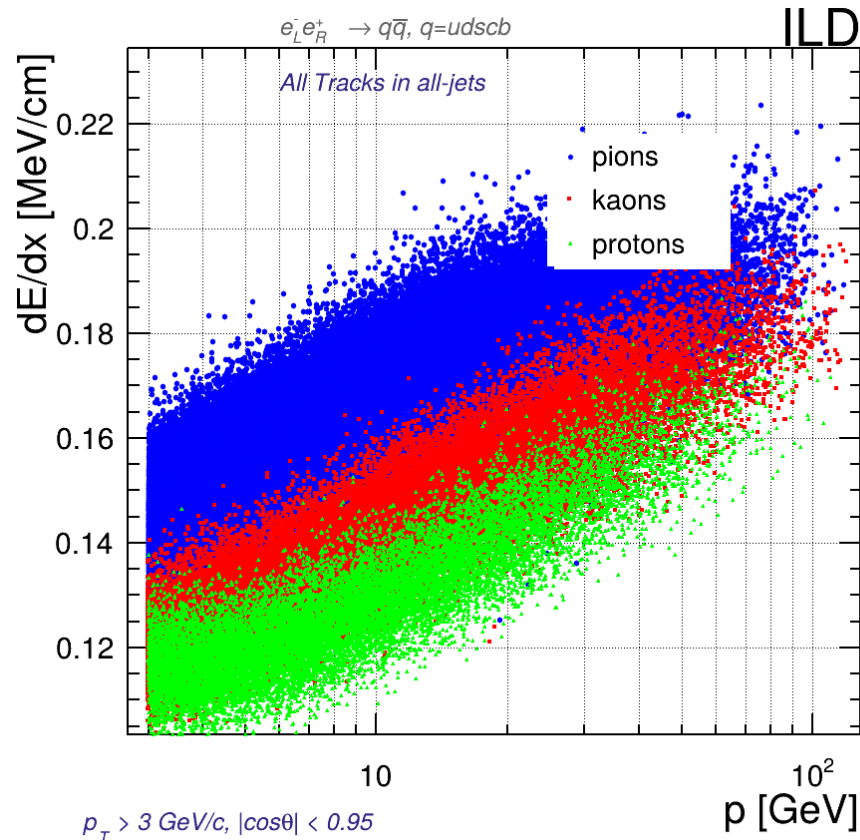
PID via dE/dx: Starting point

- Not all tracks/PFOs are valid for dE/dx



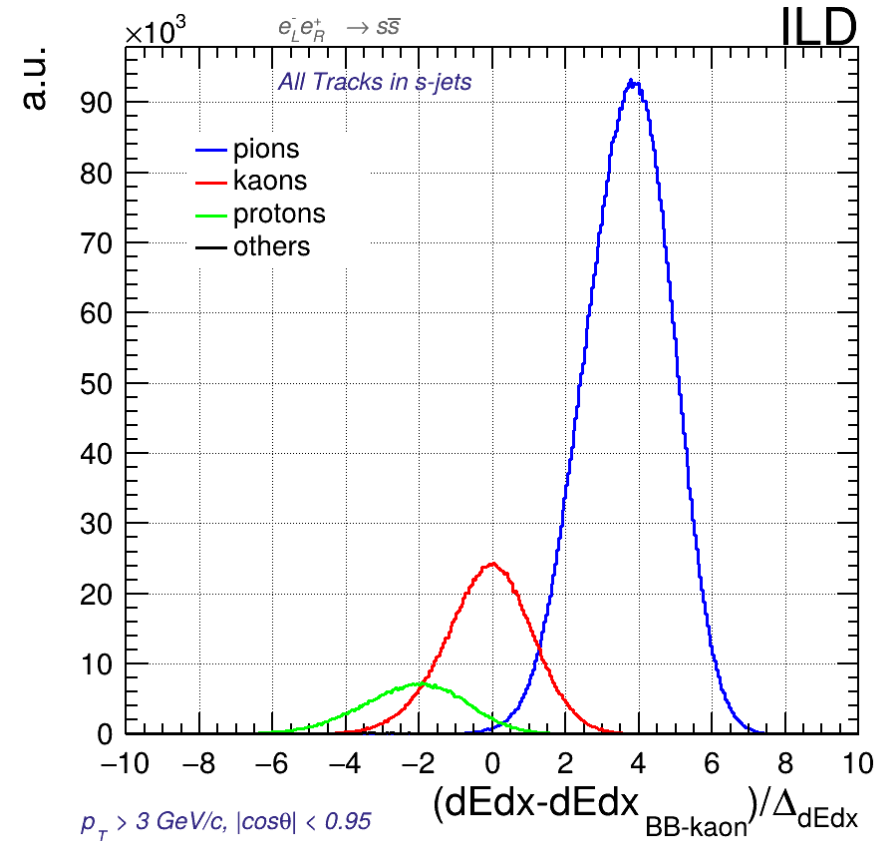
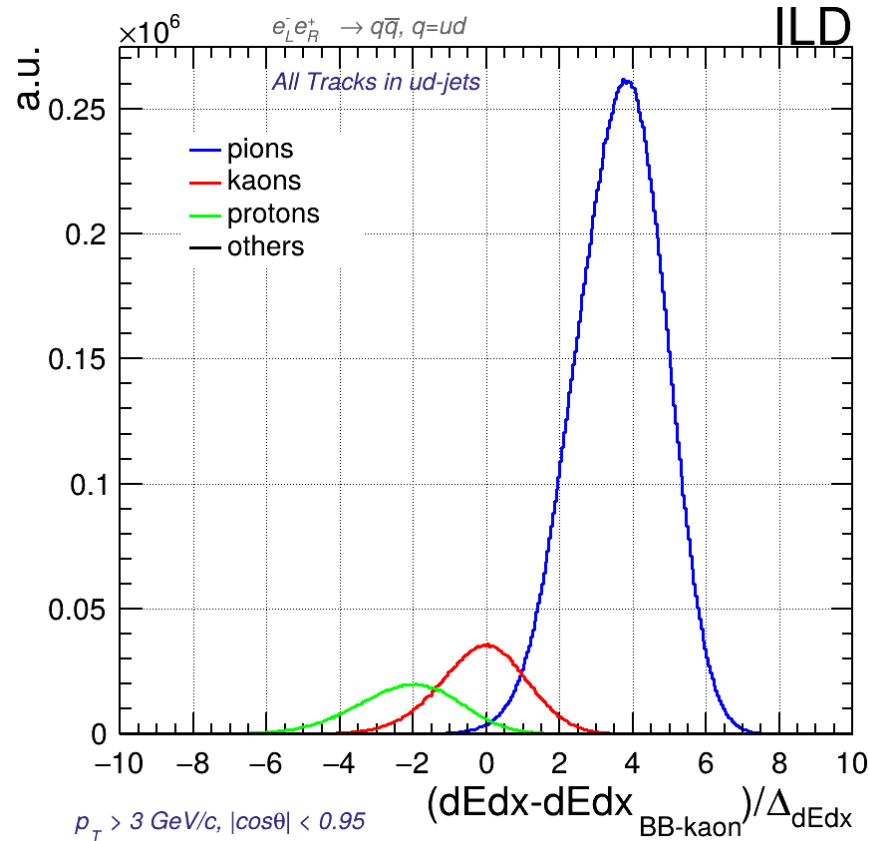
PID: Preselection

- These three bands can be used to measure an statistical distance



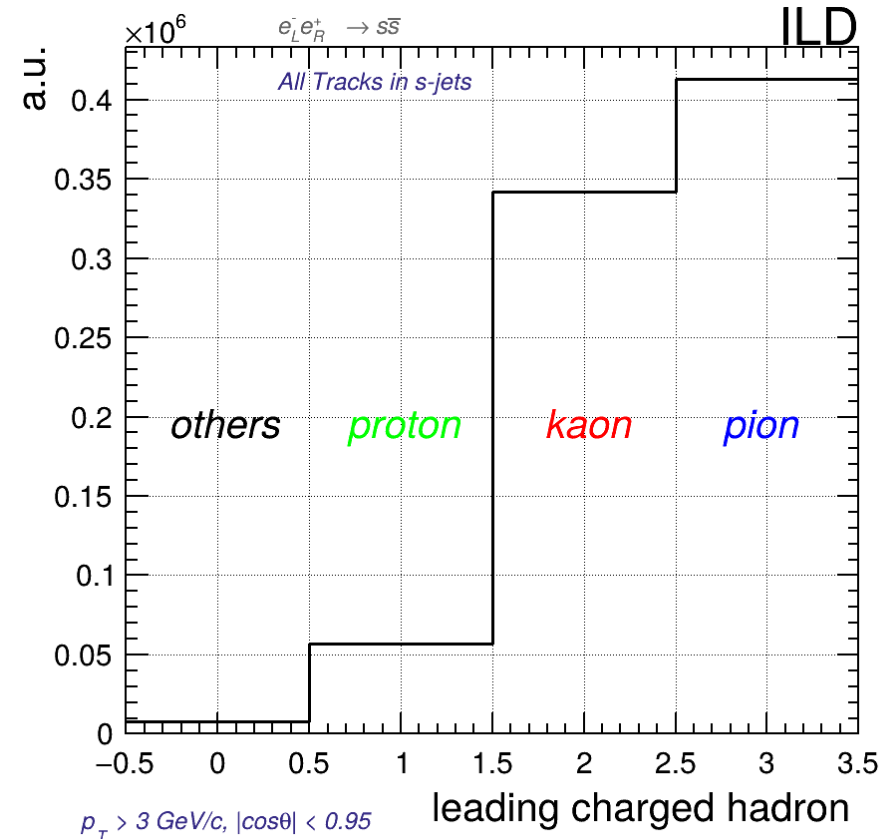
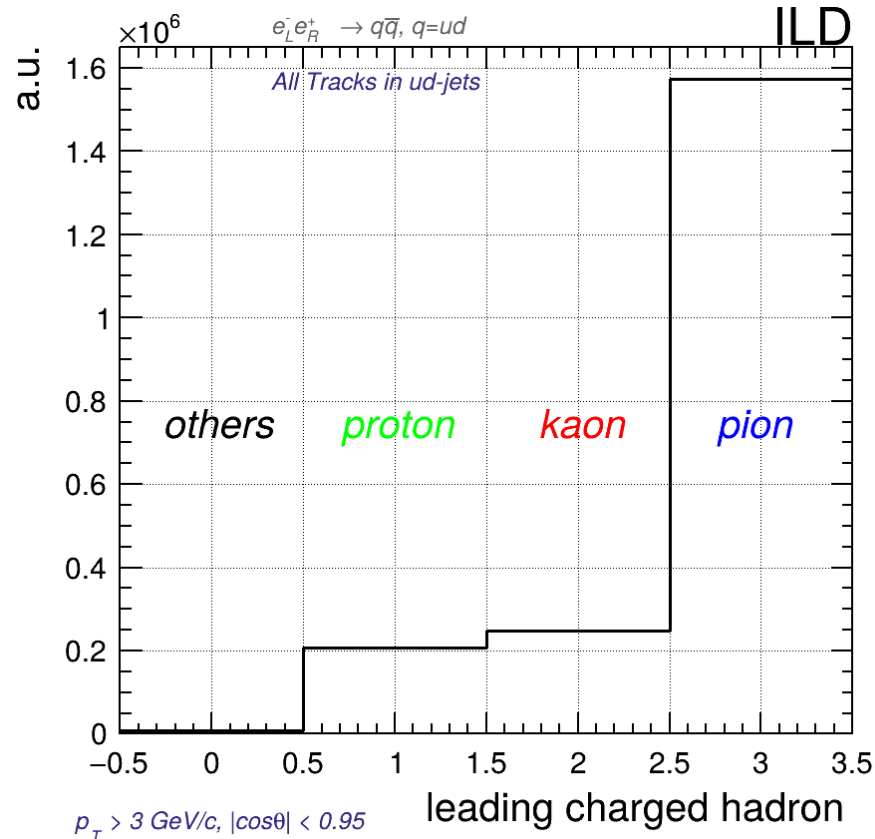
s vs ud: k-distance of tracks

- Example of distance from tracks dE/dx and the theoretical values for kaons

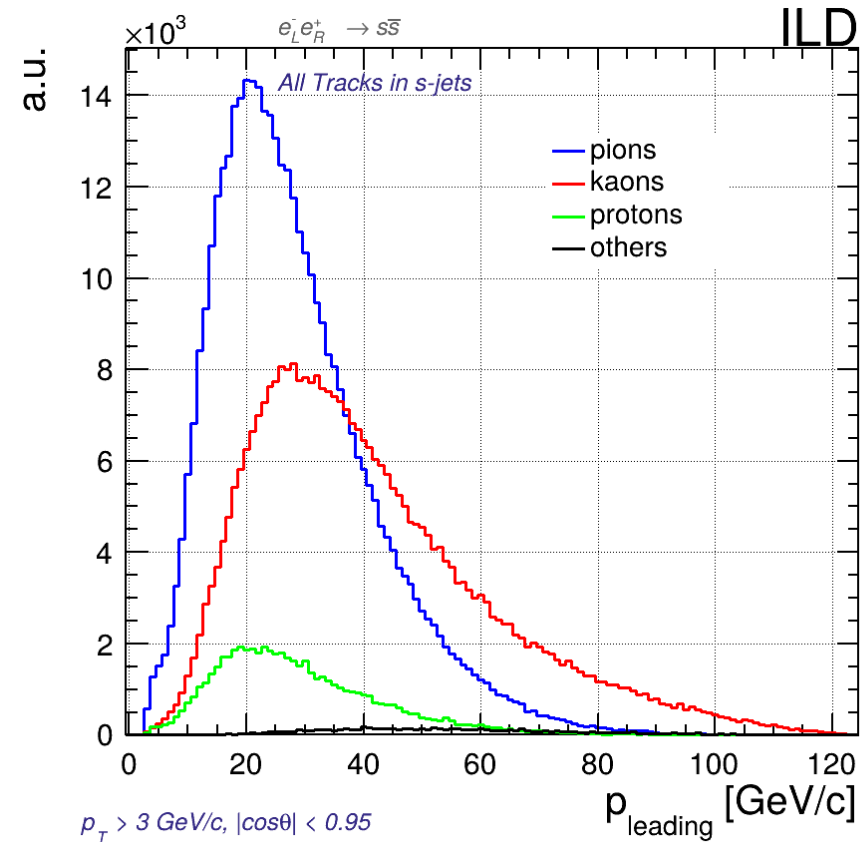
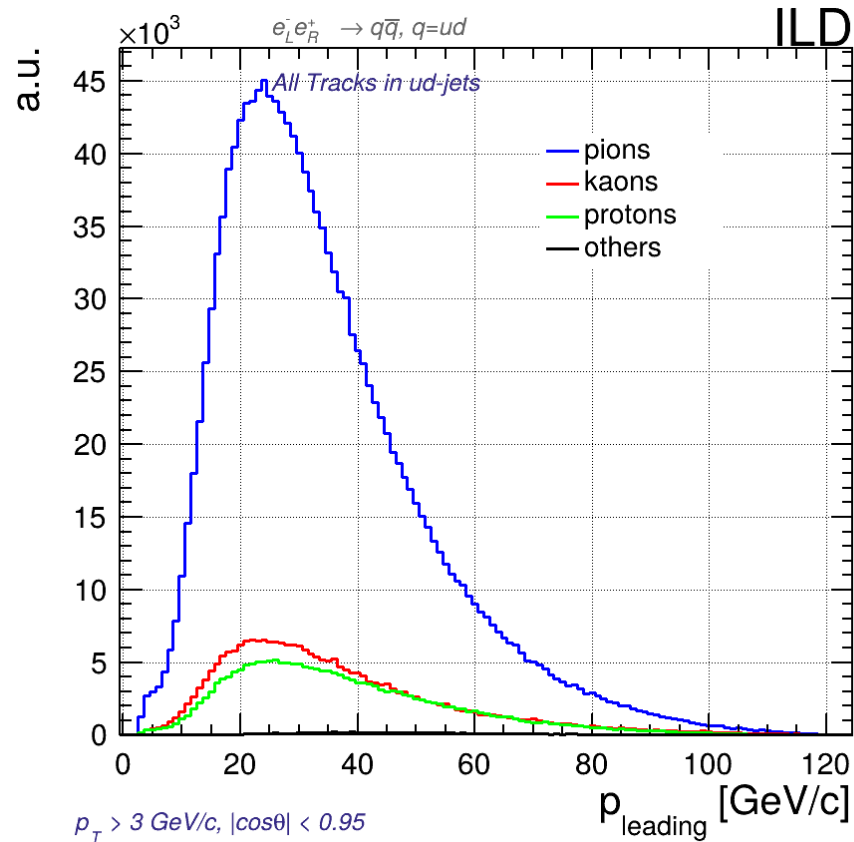


s vs ud: leading charged hadrons

- Inspecting only the leading track shows the difference between s-jets and u/d-jets

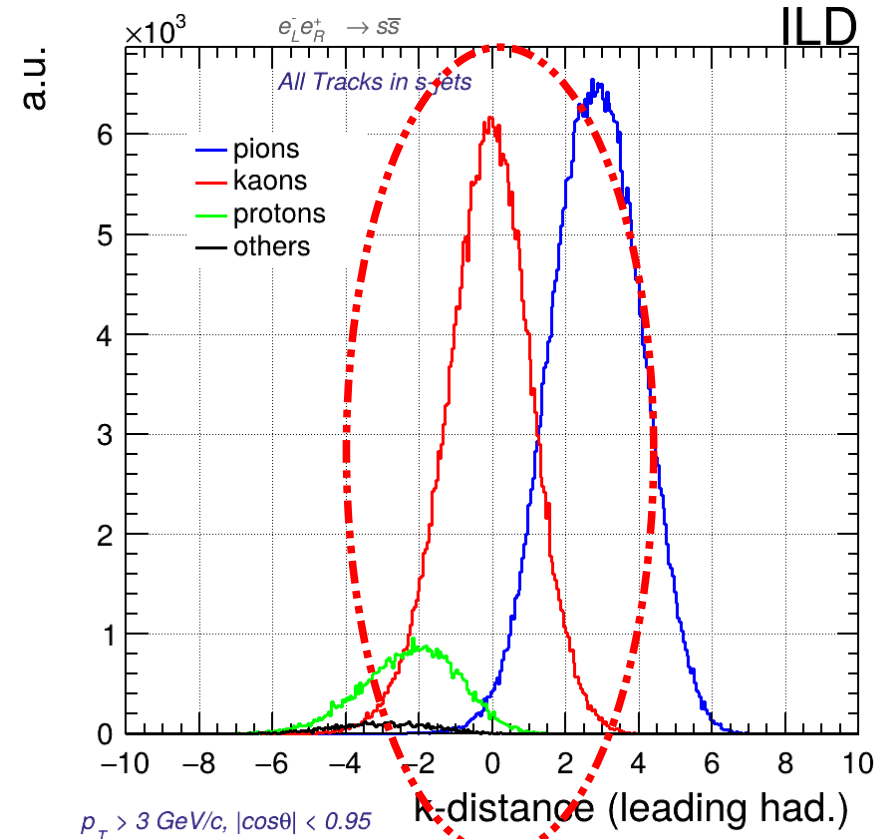
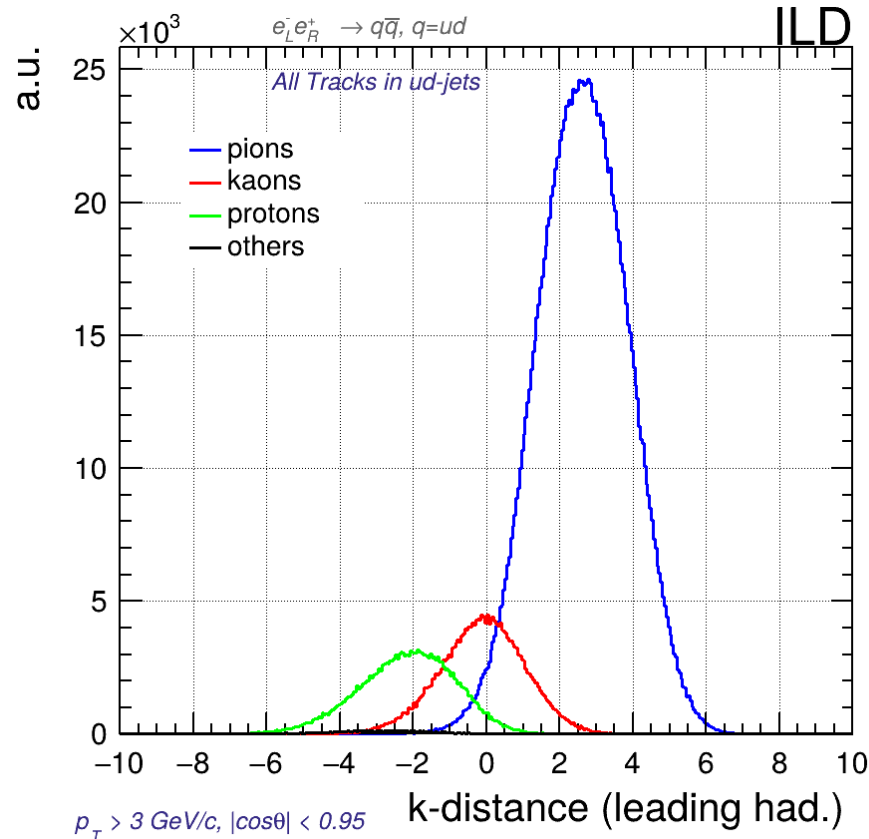


s vs ud: leading charged hadrons



s vs ud: k-dist of leading charged hadrons

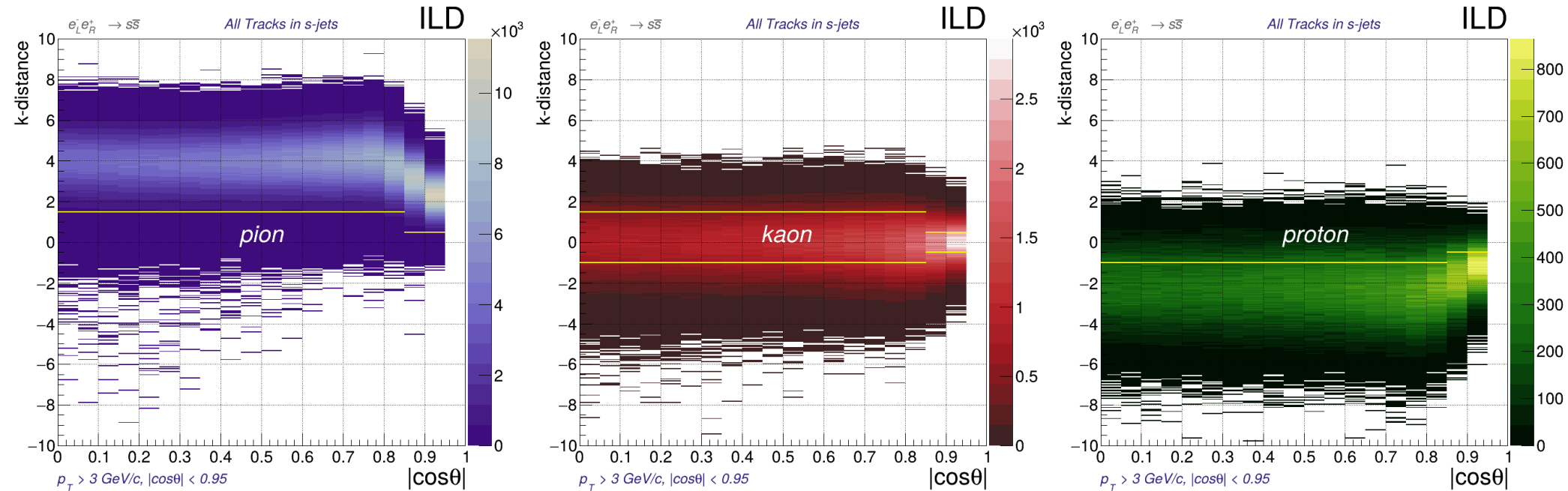
- Similar review was done via pi-distance, but shows similar behavior (as expected)



Our target for s-tagging!

2d view of k-distance (s quarks)

- Angular cuts are performed in these distributions for selection of pions/kaons

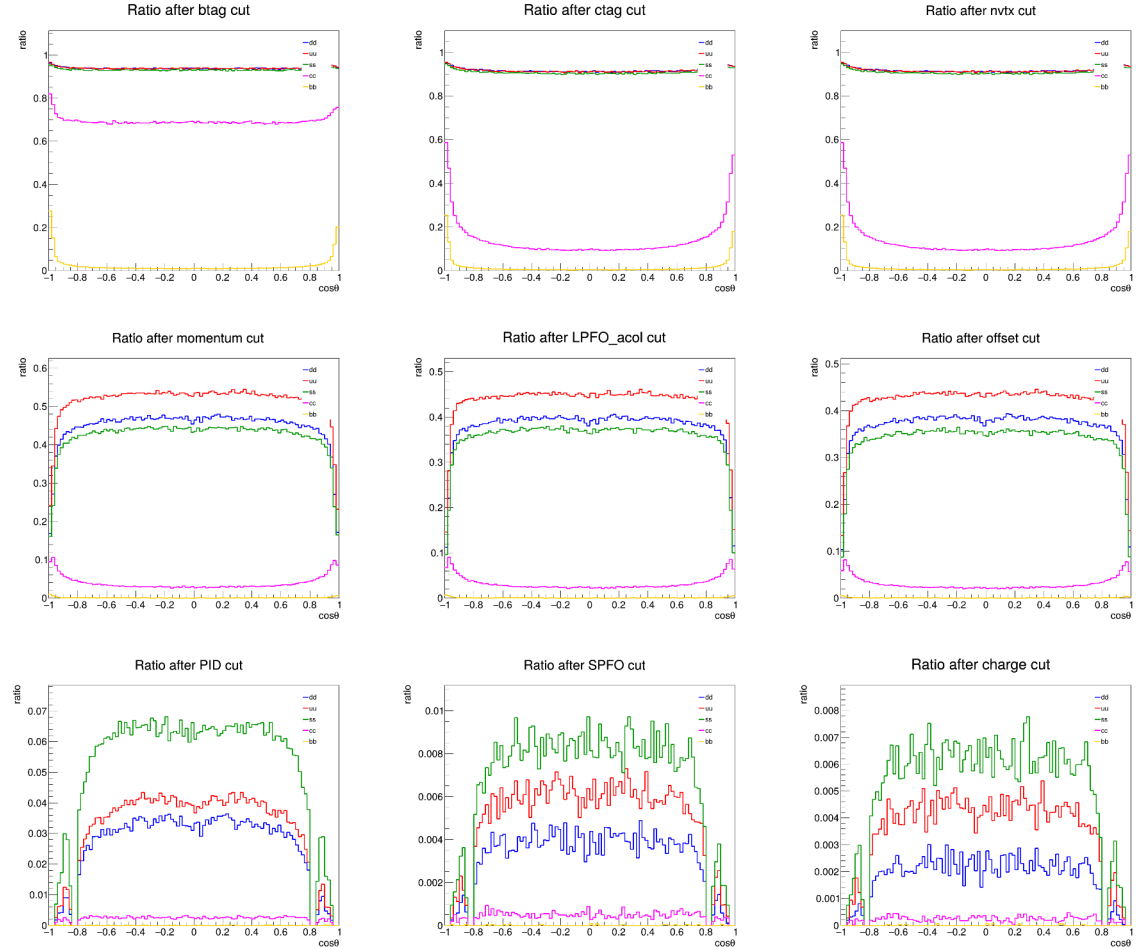


Cuts visualization (K for s-quark selection)

- Results for $P(e^-, e^+) = (-0.8, +0.3)$
- Flat when $|\cos(\theta)| < 0.8$

	dd	uu	ss	cc	bb
+ Cut 1	93.9%	93.9%	93.1%	69.3%	2.12%
+ Cut 2	91.7%	91.6%	90.9%	14.1%	1.37%
+ Cut 3	91.7%	91.6%	90.9%	14.1%	1.37%
+ Cut 4	44.9%	51.7%	42.3%	4.02%	0.0755%
+ Cut 5	38.2%	43.9%	35.9%	3.37%	0.0589%
+ Cut 6	36.8%	42.3%	34.1%	3.12%	0.0489%
+ Cut 7	2.37%	2.9%	4.8%	0.218%	0.00191%
+ Cut 8	0.285%	0.464%	0.634%	0.0432%	0.00115%
+ Cut 9	0.163%	0.329%	0.481%	0.0207%	0.000573%

Preliminary results





Reconstruction of A_{FB}

- The signal data is estimated by resting the expected backgrounds to the count of events and doing several corrections to it:
 - Efficiency estimation
 - Kaon PID stability
 - Charge migration (p-q method)

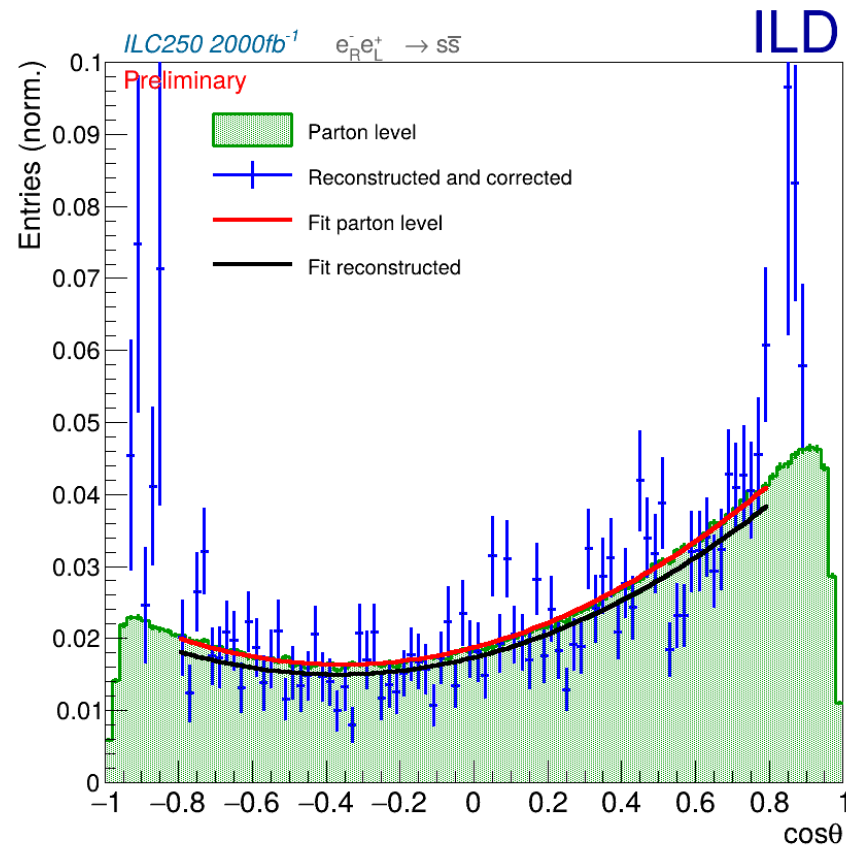
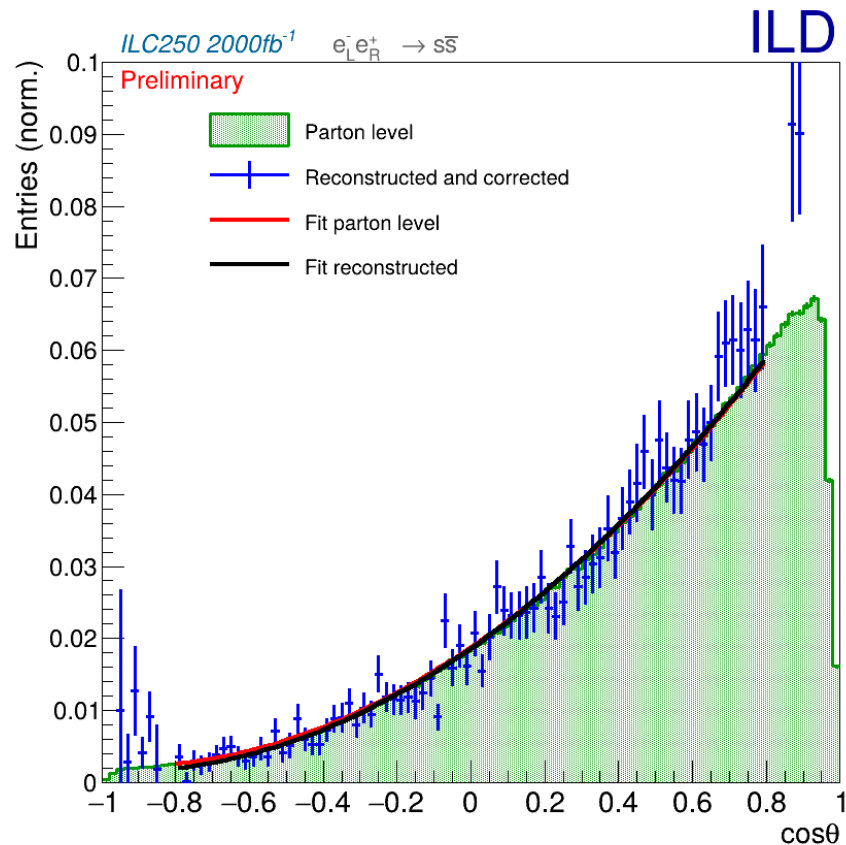
- A fit is performed to the corrected signal, following the function:

$$\frac{d\sigma}{d\cos\theta} = S \left(1 + \cos^2\theta \right) + A \cos\theta$$

- Pseudo-experiments are performed for an estimation of systematical uncertainties due to the “tagging and correction” process
 - Other systematical uncertainties are not yet consider (beam polarization, diboson backgrounds, angular correlations, etc.), but minor contributions are expected

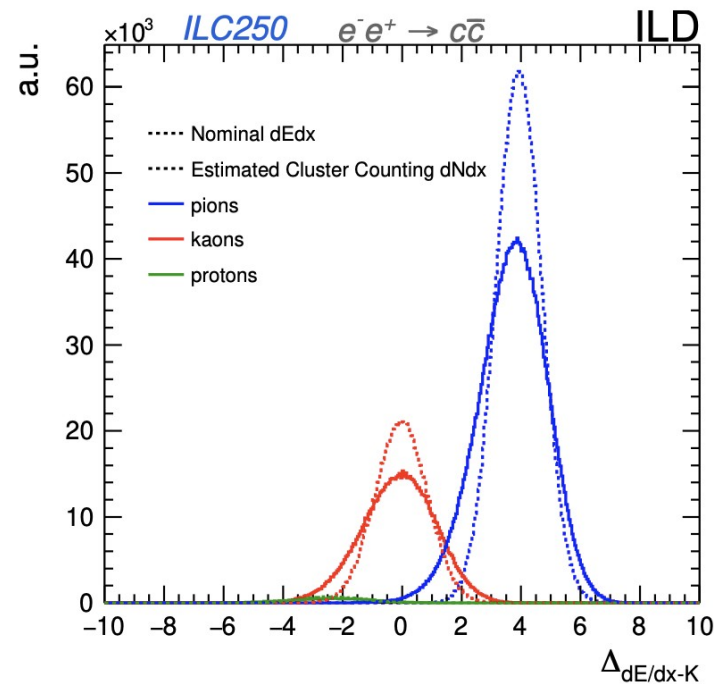
Fit to reconstructed signal

- Fit constrained to $|\cos\theta| < 0.8$ shows good agreement



PID hardware prospects

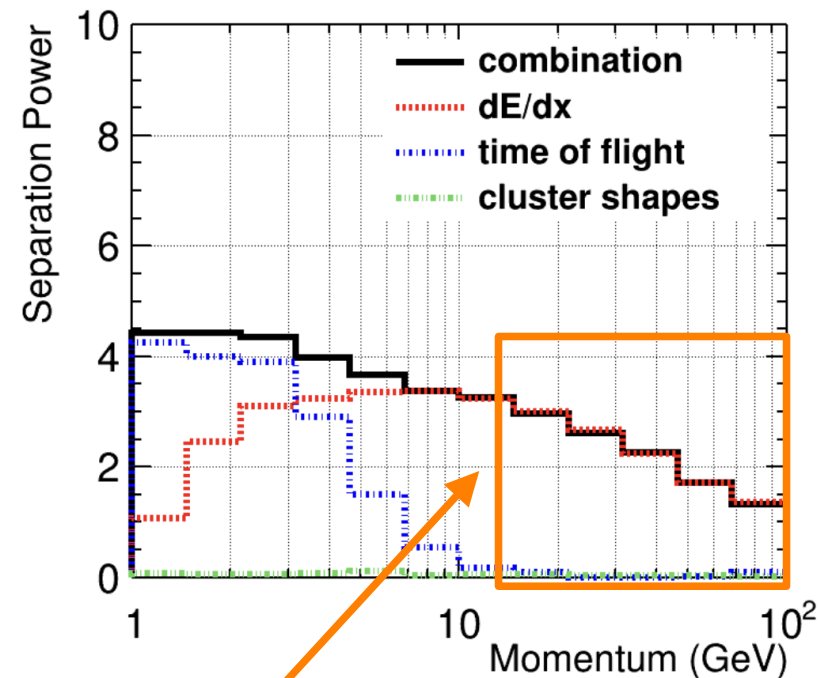
- A Marlin processor “CheatdEdxProcessor” is used for estimates of better PID cases
 - It uses fits to the bins of the 2D k-distance distribution
 - Then narrows those fits and rewrites the PFO info
- We consider two different cases:
 - 30% improvement for a pixel TPC PID case (dN/dx)
 - 99% improvement for an ideal TPC PID case
- Caveat: Only PFOs with PID available are improved



CPID for Kaon/Pion/Proton ID

- Comprehensive **P**article **ID** Marlin processor:
 - Uses different PID inputs (dE/dx, TOF, etc.)
 - Uses a BDT-based ML algorithm for classification
 - Easy to adapt to different MC ids or PID info
- In our case, the CPID was trained tackling our leading PFOS:
 - Only Kaon/Pion/Proton separation
 - $3 \text{ GeV} < \text{Momentum} < 100 \text{ GeV}$

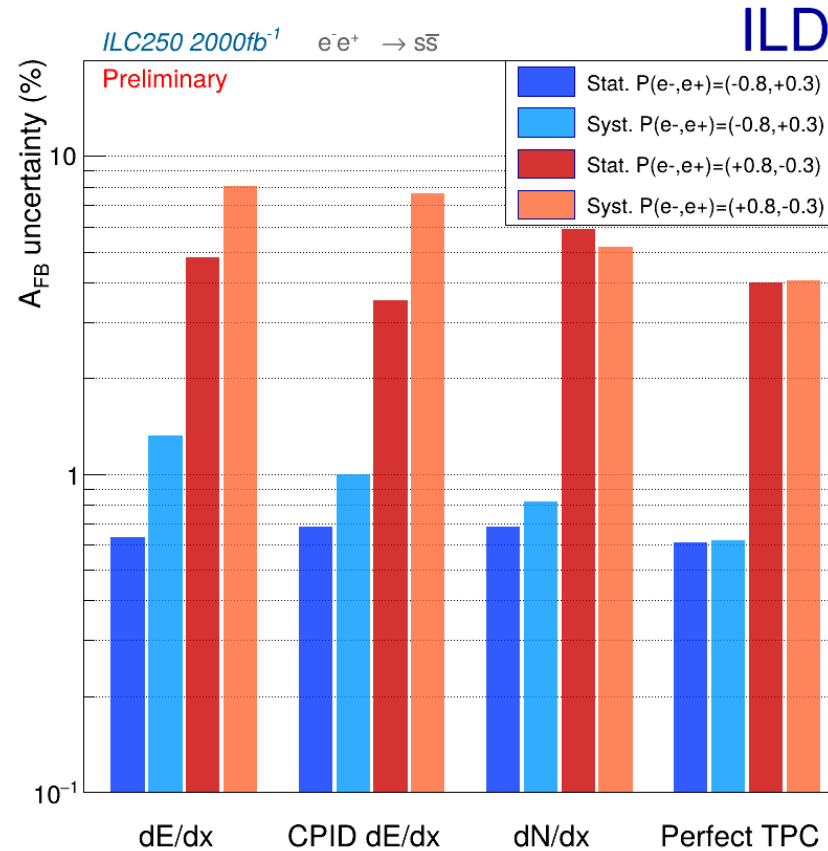
<https://arxiv.org/abs/2307.15635> (U. Einhaus)



Leading PFOs are here

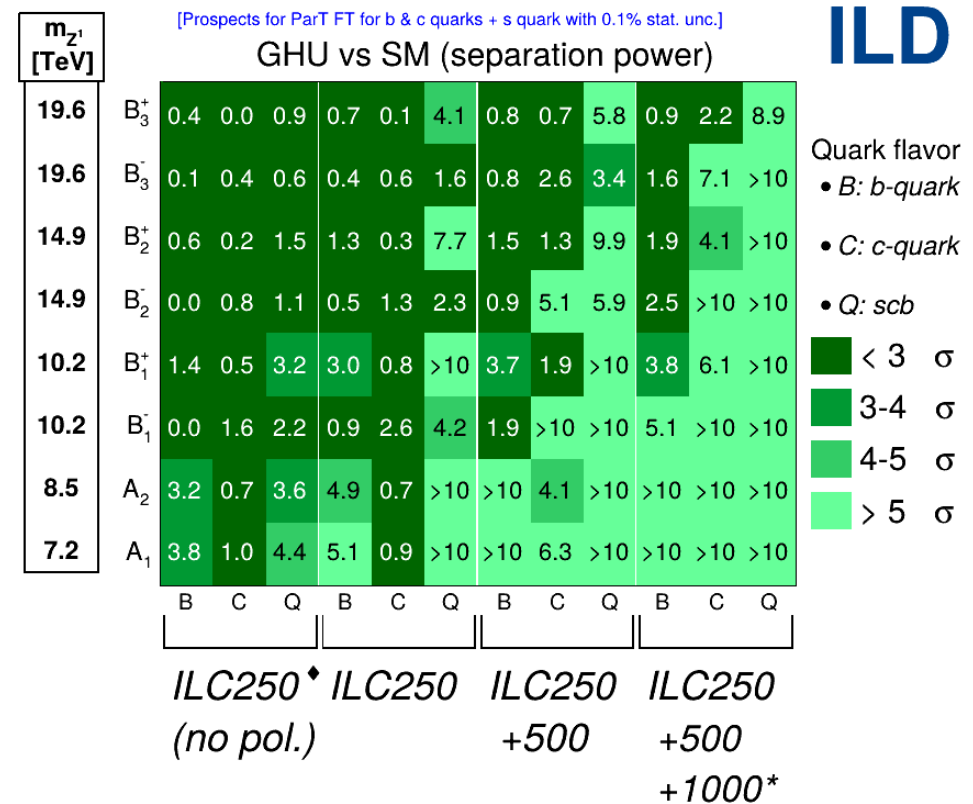
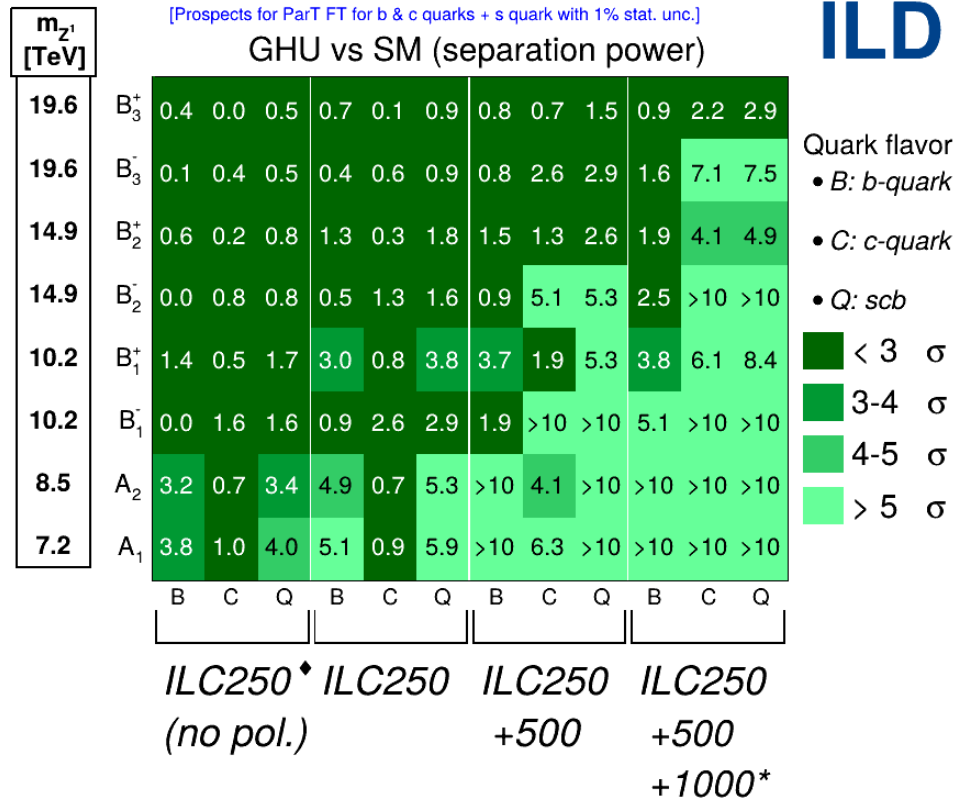
Preliminary results

- Purer signal as we move from left to right



Application into GHU phenomenology (TBD)

- 1% stat. unc. for s-quark A_{FB} (left) vs 1‰ stat. unc. for s-quark A_{FB} (right)



We have a new set of heavier models to explore this further (Thanks to N. Yamatsu)

Conclusions/Overview

- A cut-based analysis has been re-tested and improved considering:
 - Software improvements: Using CPID for optimal dE/dx handling
 - Hardware prospects: A pixel TPC (dN/dx) or a “perfect” TPC
- There seem to be a bottleneck of statistics due to Kaon selection
 - But the ML approach shows potential for reducing systematics
- There are many plans for the future of this analysis:
 - Using more MC simulation data
 - Using Particle Transformer (ParT) s-tagging
 - Running at 500 GeV
 - Re-scaling results to the LCF@CERN
 - Applying the obtained precisions to BSM (GHU prospects)

THANKS FOR YOUR ATTENTION

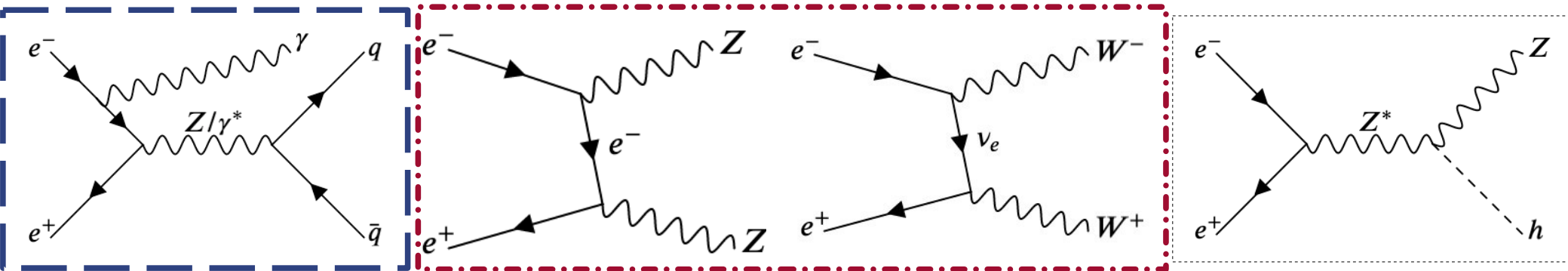


BACK-UP



Preselection of $q\bar{q}$ signals

- Once we have the reconstructed pfos of the events with different targets:
 - We cluster the signal in jets (VLC algorithm):
 - The algorithm packs together the PFOs into two jets.
 - Signal is expected in a back-to-back topology (but not the backgrounds!)
 - Most of the background is **radiative return ($\gamma q\bar{q}$)**
 - And most of the data is background!
 - x3 for $e^-_L e^+_R$ and x6 for $e^-_R e^+_L$ at 250 GeV
 - x4 for $e^-_L e^+_R$ and x7 for $e^-_R e^+_L$ at 500 GeV
 - Then we apply different cuts to the signal to remove the background processes



Preselection for 250 GeV

Cuts:

- $K_{reco} < 35$ GeV
- $m_{2jets} > 140$ GeV
- Charged N pfos
- Photon veto
- $Y_{23} < 0.015$

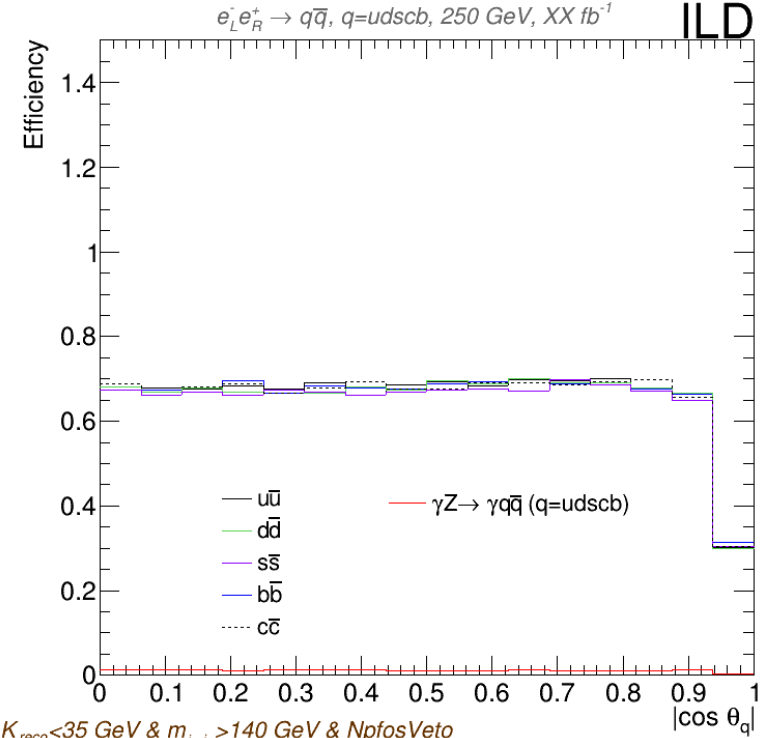
VLC Algorithm parameters:

- $R = 1.0$
- $\gamma = 0.0$
- $\beta = 1.0$

R	Efficiencies (%)				S/B
	$b\bar{b}$	$c\bar{c}$	$q\bar{q}$ (uds)	ISR	
1.0	64.7	64.6	64.3	0.9	23.7
	68.3	68.5	68.1	1.1	28.1

← $|\cos\theta| < 0.9$

Total efficiency of the preselection for the different quark flavours and radiative return for the chosen configuration ($\gamma=0$). The second row is for $|\cos\theta| < 0.9$



$K_{reco} < 35$ GeV & $m_{j_1, j_2} > 140$ GeV & NpfosVeto
& Cnpfos Veto & Photon Veto 1 & $y_{23} < 0.015$

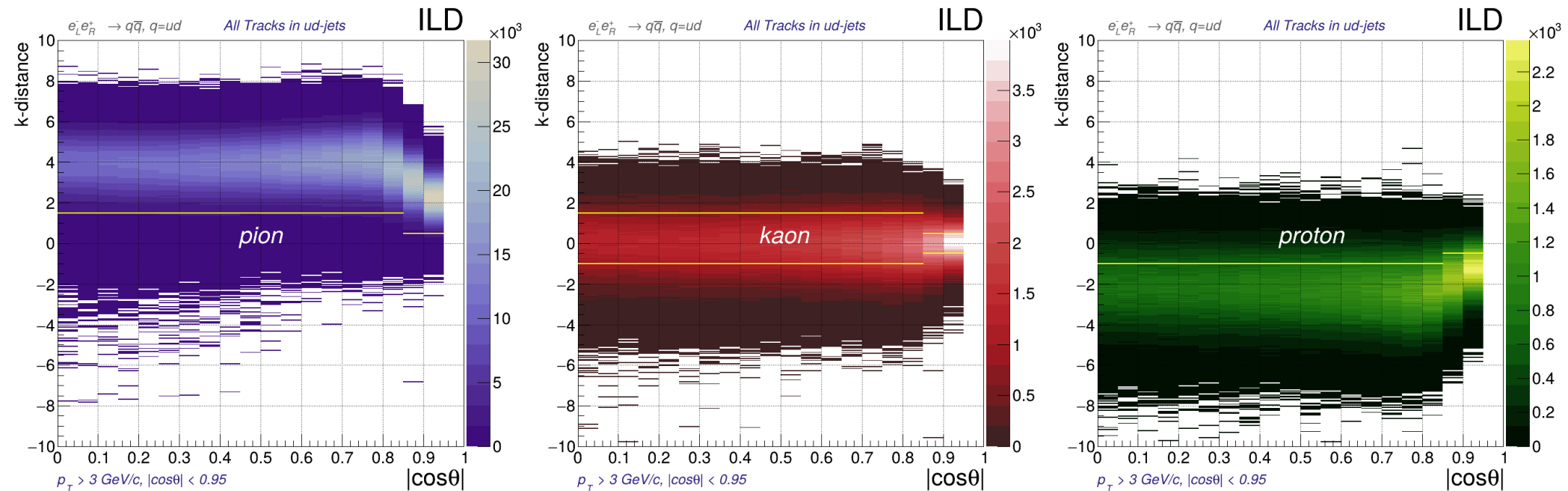
Efficiency of the preselection for the different quark flavours vs the angular distribution of the two jet system (new samples, final configuration)

Re-run of previous analysis



2d view of k-distance (ud quarks)

- Angular cuts are performed in these distributions for selection of pions

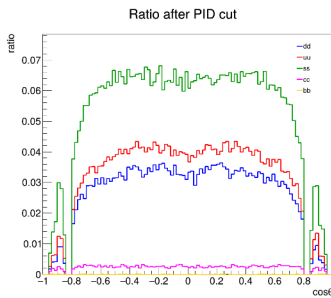
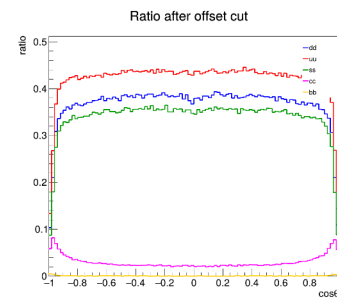
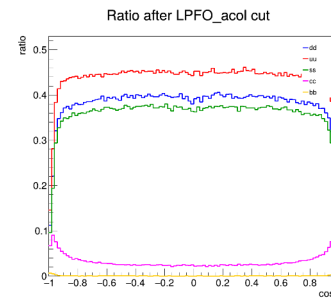
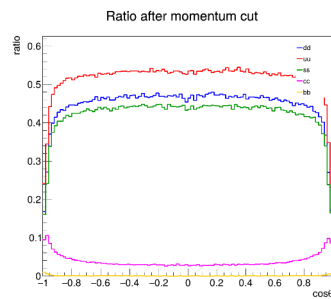
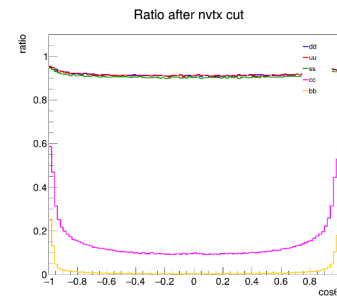
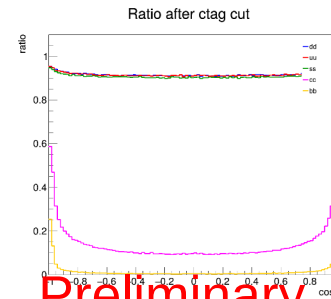
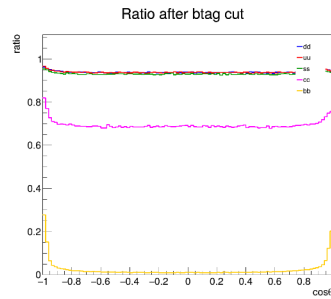


Cuts visualization (K for s-quark selection)



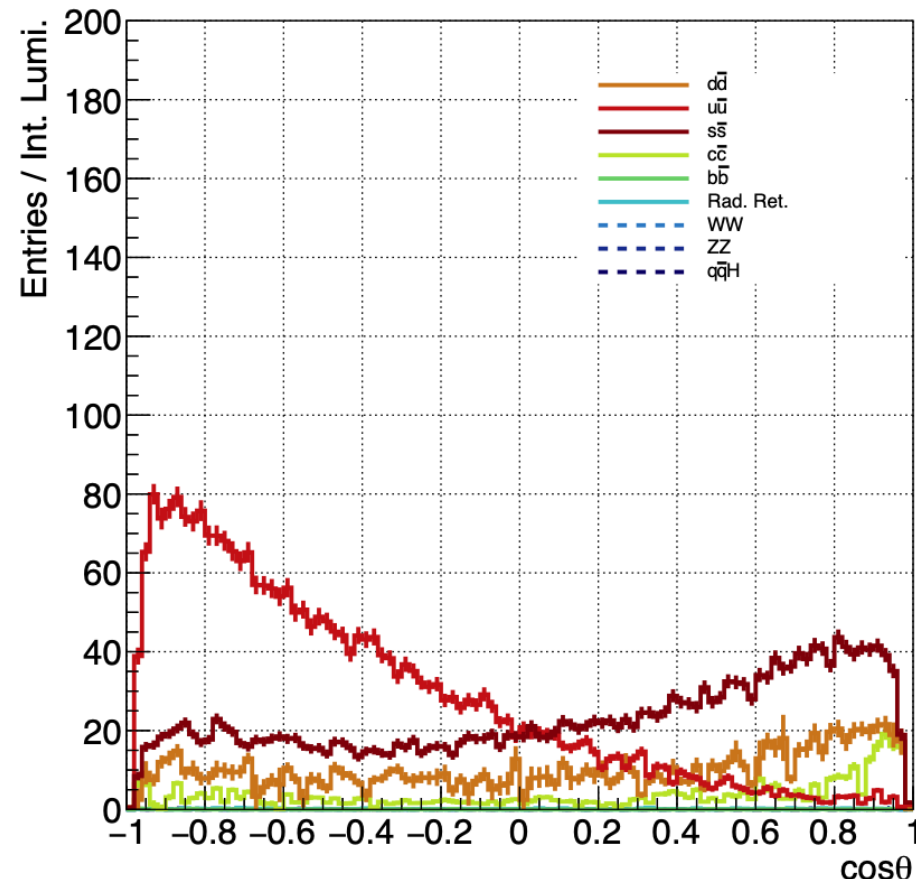
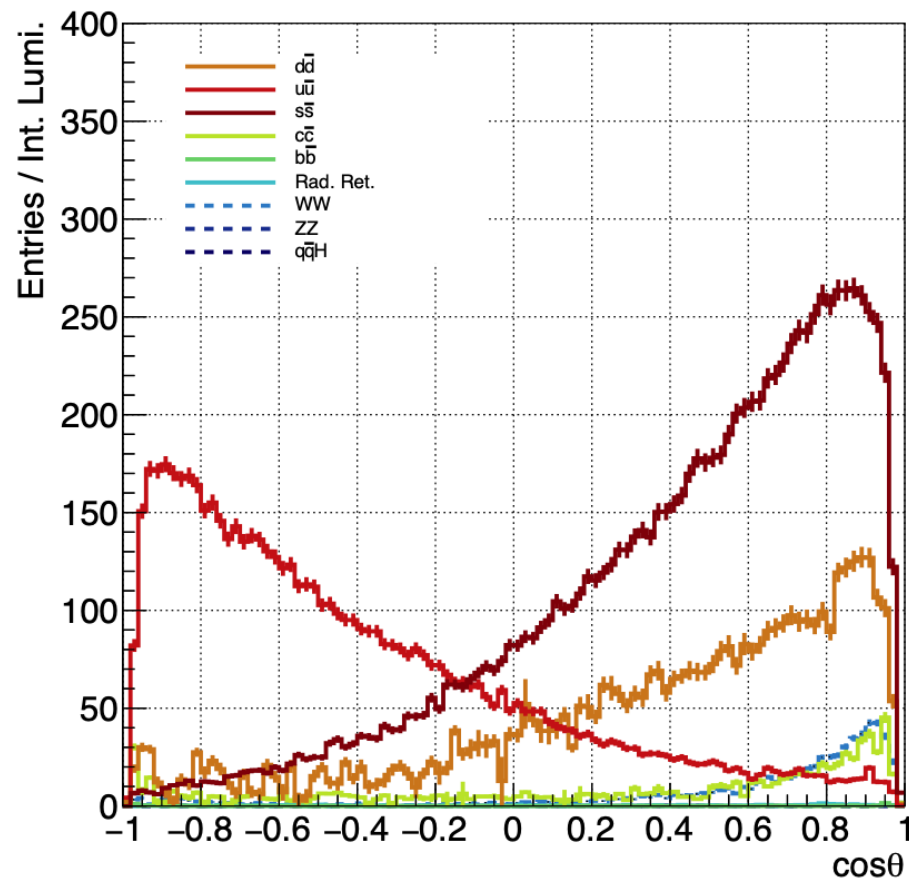
	dd	uu	ss	cc	bb
+ Cut 1	93.9%	93.9%	93.1%	69.3%	2.12%
+ Cut 2	91.7%	91.6%	90.9%	14.1%	1.37%
+ Cut 3	91.7%	91.6%	90.9%	14.1%	1.37%
+ Cut 4	44.9%	51.7%	42.3%	4.02%	0.0755%
+ Cut 5	38.2%	43.9%	35.9%	3.37%	0.0589%
+ Cut 6	36.8%	42.3%	34.1%	3.12%	0.0489%
+ Cut 7	2.37%	2.9%	4.8%	0.218%	0.00191%
+ Cut 8	0.285%	0.464%	0.634%	0.0432%	0.00115%
+ Cut 9	0.163%	0.329%	0.481%	0.0207%	0.000573%

Preliminary



Contributions after preselection

- After K LPFO selection (Plots from Yuichi's analysis)





Preliminary results (π mode for u/d selection)

- Selecting u/d quarks

- Results for $e^-_L e^+_R$

- (Left) New dE/dx analysis vs (right) CPID dE/dx

CPID
dE/dx

B/S=0.23

	dd	uu	ss	cc	bb		dd	uu	ss	cc	bb
+ Cut 1	93.9%	93.9%	93.1%	69.3%	2.12%	+ Cut 1	93.9%	93.9%	93.1%	69.3%	2.12%
+ Cut 2	91.7%	91.6%	90.9%	14.1%	1.37%	+ Cut 2	91.7%	91.6%	90.9%	14.1%	1.37%
+ Cut 3	91.7%	91.6%	90.9%	14.1%	1.37%	+ Cut 3	91.7%	91.6%	90.9%	14.1%	1.37%
+ Cut 4	44.9%	51.7%	42.3%	4.02%	0.0755%	+ Cut 4	44.9%	51.7%	42.3%	4.02%	0.0758%
+ Cut 5	38.2%	43.9%	35.9%	3.37%	0.0589%	+ Cut 5	38.2%	43.9%	35.9%	3.37%	0.0589%
+ Cut 6	36.8%	42.3%	34.1%	3.12%	0.0489%	+ Cut 6	36.8%	42.3%	34.1%	3.12%	0.0485%
+ Cut 7	12.8%	12.9%	5.61%	0.667%	0.0105%	+ Cut 7	19.4%	19.8%	8.74%	1.02%	0.0196%
+ Cut 8	1.72%	1.93%	0.86%	0.162%	0.00287%	+ Cut 8	2.59%	2.97%	1.33%	0.247%	0.00546%
+ Cut 9	1%	1.22%	0.503%	0.0823%	0.00172%	+ Cut 9	1.51%	1.89%	0.778%	0.126%	0.00222%

From 0.23 to 0.23 B/S



Preliminary results (K mode for s selection)

- Selecting s quark

- Results for $e^-_L e^+_R$

- (Left) New dE/dx analysis vs (right) CPID dE/dx

CPID
dE/dx

B/S=0.78

	dd	uu	ss	cc	bb		dd	uu	ss	cc	bb
+ Cut 1	93.9%	93.9%	93.1%	69.3%	2.12%	+ Cut 1	93.9%	93.9%	93.1%	69.3%	2.12%
+ Cut 2	91.7%	91.6%	90.9%	14.1%	1.37%	+ Cut 2	91.7%	91.6%	90.9%	14.1%	1.37%
+ Cut 3	91.7%	91.6%	90.9%	14.1%	1.37%	+ Cut 3	91.7%	91.6%	90.9%	14.1%	1.37%
+ Cut 4	44.9%	51.7%	42.3%	4.02%	0.0755%	+ Cut 4	44.9%	51.7%	42.3%	4.02%	0.0758%
+ Cut 5	38.2%	43.9%	35.9%	3.37%	0.0589%	+ Cut 5	38.2%	43.9%	35.9%	3.37%	0.0589%
+ Cut 6	36.8%	42.3%	34.1%	3.12%	0.0489%	+ Cut 6	36.8%	42.3%	34.1%	3.12%	0.0485%
+ Cut 7	2.37%	2.9%	4.8%	0.218%	0.00191%	+ Cut 7	0.991%	1.43%	4.21%	0.267%	0.00364%
+ Cut 8	0.285%	0.464%	0.634%	0.0432%	0.00115%	+ Cut 8	0.13%	0.228%	0.548%	0.0495%	0.00142%
+ Cut 9	0.163%	0.329%	0.481%	0.0207%	0.000573%	+ Cut 9	0.0674%	0.162%	0.421%	0.0262%	0.000607%

From 1.36 to 0.78 B/S



Preliminary results (π mode for u/d selection)

- Selecting u/d quarks

dN/dx

- Results for $e^-_L e^+_R$

- (Left) New dE/dx analysis vs (right) dN/dx

B/S=0.20

	dd	uu	ss	cc	bb		dd	uu	ss	cc	bb
+ Cut 1	93.9%	93.9%	93.1%	69.3%	2.12%	+ Cut 1	93.9%	93.8%	93.1%	69.5%	2.67%
+ Cut 2	91.7%	91.6%	90.9%	14.1%	1.37%	+ Cut 2	91.7%	91.6%	90.9%	14.7%	1.83%
+ Cut 3	91.7%	91.6%	90.9%	14.1%	1.37%	+ Cut 3	91.7%	91.6%	90.9%	14.7%	1.83%
+ Cut 4	44.9%	51.7%	42.3%	4.02%	0.0755%	+ Cut 4	45.1%	52.6%	42.4%	4.76%	0.166%
+ Cut 5	38.2%	43.9%	35.9%	3.37%	0.0589%	+ Cut 5	38.1%	44.4%	35.7%	3.94%	0.119%
+ Cut 6	36.8%	42.3%	34.1%	3.12%	0.0489%	+ Cut 6	36.2%	42.2%	33.4%	3.5%	0.0776%
+ Cut 7	12.8%	12.9%	5.61%	0.667%	0.0105%	+ Cut 7	20.2%	20.6%	7.65%	0.95%	0.0184%
+ Cut 8	1.72%	1.93%	0.86%	0.162%	0.00287%	+ Cut 8	2.5%	2.85%	1.07%	0.215%	0.00482%
+ Cut 9	1%	1.22%	0.503%	0.0823%	0.00172%	+ Cut 9	1.49%	1.85%	0.632%	0.116%	0.00219%

From 0.23 to 0.20 B/S



Preliminary results (K mode for s selection)

- Selecting s quark

dN/dx

- Results for $e^-_L e^+_R$

- (Left) New dE/dx analysis vs (right) dN/dx

B/S=0.34

	dd	uu	ss	cc	bb		dd	uu	ss	cc	bb
+ Cut 1	93.9%	93.9%	93.1%	69.3%	2.12%	+ Cut 1	93.9%	93.8%	93.1%	69.5%	2.67%
+ Cut 2	91.7%	91.6%	90.9%	14.1%	1.37%	+ Cut 2	91.7%	91.6%	90.9%	14.7%	1.83%
+ Cut 3	91.7%	91.6%	90.9%	14.1%	1.37%	+ Cut 3	91.7%	91.6%	90.9%	14.7%	1.83%
+ Cut 4	44.9%	51.7%	42.3%	4.02%	0.0755%	+ Cut 4	45.1%	52.6%	42.4%	4.76%	0.166%
+ Cut 5	38.2%	43.9%	35.9%	3.37%	0.0589%	+ Cut 5	38.1%	44.4%	35.7%	3.94%	0.119%
+ Cut 6	36.8%	42.3%	34.1%	3.12%	0.0489%	+ Cut 6	36.2%	42.2%	33.4%	3.5%	0.0776%
+ Cut 7	2.37%	2.9%	4.8%	0.218%	0.00191%	+ Cut 7	0.348%	0.524%	3.55%	0.144%	0.0011%
+ Cut 8	0.285%	0.464%	0.634%	0.0432%	0.00115%	+ Cut 8	0.0389%	0.0865%	0.457%	0.022%	0.000438%
+ Cut 9	0.163%	0.329%	0.481%	0.0207%	0.000573%	+ Cut 9	0.0214%	0.0629%	0.366%	0.0109%	0.000219%

From 1.36 to 0.34 B/S

Preliminary results (π mode for u/d selection)

- Selecting u/d quarks

- Results for $e^-_L e^+_R$

- (Left) dNdx vs (right) Perfect TPC PID

Perfect
TPC PID

B/S=0.19

	dd	uu	ss	cc	bb		dd	uu	ss	cc	bb
+ Cut 1	93.9%	93.8%	93.1%	69.5%	2.67%	+ Cut 1	93.9%	93.9%	93.1%	69.5%	2.66%
+ Cut 2	91.7%	91.6%	90.9%	14.7%	1.83%	+ Cut 2	91.7%	91.6%	90.9%	14.7%	1.82%
+ Cut 3	91.7%	91.6%	90.9%	14.7%	1.83%	+ Cut 3	91.7%	91.6%	90.9%	14.7%	1.82%
+ Cut 4	45.1%	52.6%	42.4%	4.76%	0.166%	+ Cut 4	45.1%	52.5%	42.3%	4.76%	0.165%
+ Cut 5	38.1%	44.4%	35.7%	3.94%	0.119%	+ Cut 5	38.2%	44.4%	35.7%	3.94%	0.115%
+ Cut 6	36.2%	42.2%	33.4%	3.5%	0.0776%	+ Cut 6	36.2%	42.2%	33.4%	3.5%	0.0743%
+ Cut 7	20.2%	20.6%	7.65%	0.95%	0.0184%	+ Cut 7	20.2%	20.6%	7.41%	0.907%	0.017%
+ Cut 8	2.5%	2.85%	1.07%	0.215%	0.00482%	+ Cut 8	2.5%	2.84%	1.04%	0.206%	0.00503%
+ Cut 9	1.49%	1.85%	0.632%	0.116%	0.00219%	+ Cut 9	1.49%	1.85%	0.608%	0.113%	0.00262%

From 0.20 to 0.19 B/S

Preliminary results (K mode for s selection)

- Selecting s quark

- Results for $e^-_L e^+_R$

- (Left) dNdx vs (right) Perfect TPC PID

Perfect
TPC PID

B/S=0.28

	dd	uu	ss	cc	bb		dd	uu	ss	cc	bb
+ Cut 1	93.9%	93.8%	93.1%	69.5%	2.67%	+ Cut 1	93.9%	93.9%	93.1%	69.5%	2.66%
+ Cut 2	91.7%	91.6%	90.9%	14.7%	1.83%	+ Cut 2	91.7%	91.6%	90.9%	14.7%	1.82%
+ Cut 3	91.7%	91.6%	90.9%	14.7%	1.83%	+ Cut 3	91.7%	91.6%	90.9%	14.7%	1.82%
+ Cut 4	45.1%	52.6%	42.4%	4.76%	0.166%	+ Cut 4	45.1%	52.5%	42.3%	4.76%	0.165%
+ Cut 5	38.1%	44.4%	35.7%	3.94%	0.119%	+ Cut 5	38.2%	44.4%	35.7%	3.94%	0.115%
+ Cut 6	36.2%	42.2%	33.4%	3.5%	0.0776%	+ Cut 6	36.2%	42.2%	33.4%	3.5%	0.0743%
+ Cut 7	0.348%	0.524%	3.55%	0.144%	0.0011%	+ Cut 7	0.385%	0.571%	4.85%	0.185%	0.00175%
+ Cut 8	0.0389%	0.0865%	0.457%	0.022%	0.000438%	+ Cut 8	0.0472%	0.0924%	0.601%	0.0263%	0.000437%
+ Cut 9	0.0214%	0.0629%	0.366%	0.0109%	0.000219%	+ Cut 9	0.0258%	0.0683%	0.497%	0.0146%	0.000218%

From 0.34 to 0.28 B/S

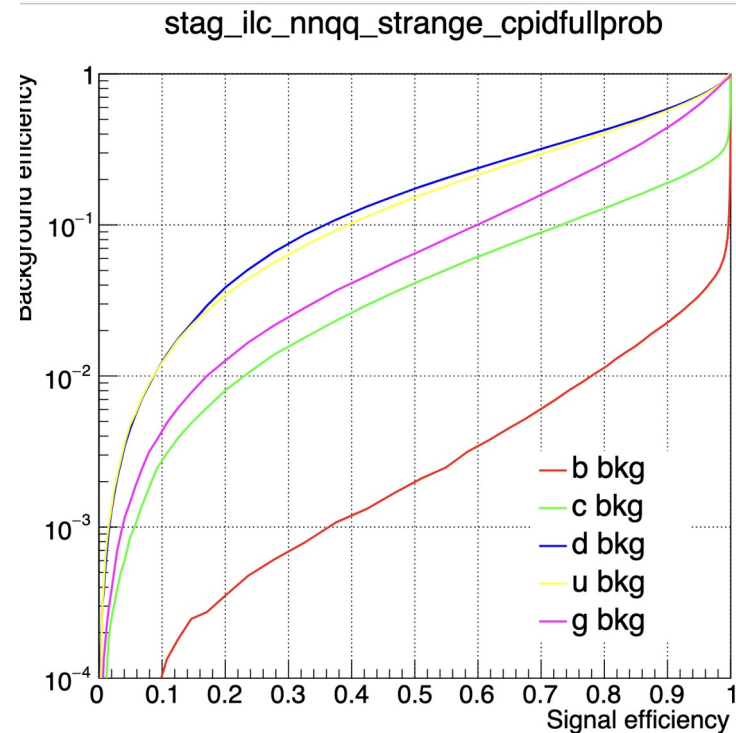
Preliminary results (numbers)



	$e^-e^+ \rightarrow s\bar{s}$					
	$P_{e^-e^+}(-0.8, +0.3)$			$P_{e^-e^+}(+0.8, -0.3)$		
	$\delta A_{FB}^{stat}(\%)$	$\delta A_{FB}^{syst}(\%)$	B/S	$\delta A_{FB}^{stat}(\%)$	$\delta A_{FB}^{syst}(\%)$	B/S
dE/dx	0.63	1.32	1.38	4.80	8.03	2.32
CPID (dE/dx)	0.68	1.00	0.79	3.53	7.62	1.54
dN/dx	0.68	0.82	0.37	5.91	5.19	0.58
Perfect TPC	0.61	0.62	0.28	4.01	4.05	0.48

The “holy grail”: ParT s-tagging

- Particle Transformers is state-of-the-art ML software
- It uses *CPID* for the tracks PID
- It can be 10x better than the cut-based approach
 - But how? Is this code available? Trying to get access to it to incorporate it into a chain of analysis
 - Can reduce the cuts in the analysis into:
 - B-tag
 - C-tag
 - S-tag → Much more powerful than kaon ID
 - Migration cuts:
 - Secondary PFO candidate cut
 - Opposite charge LPFO cut



**1% ss signal with
0.02% u/d backgrounds?
Expected B/S = 0.33**

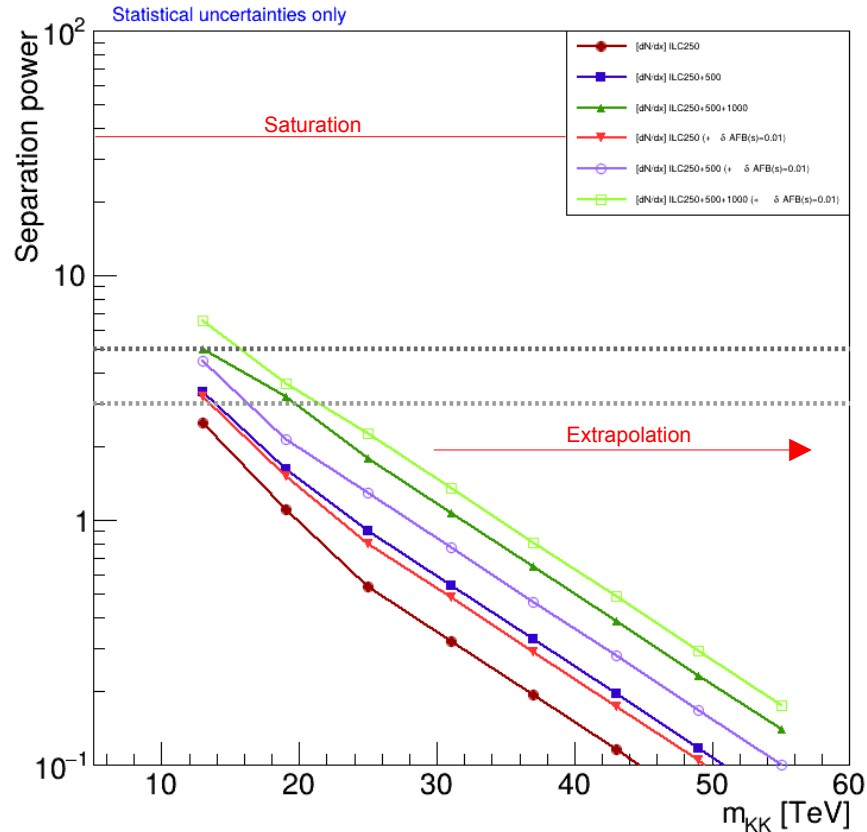
GHU phenomenology



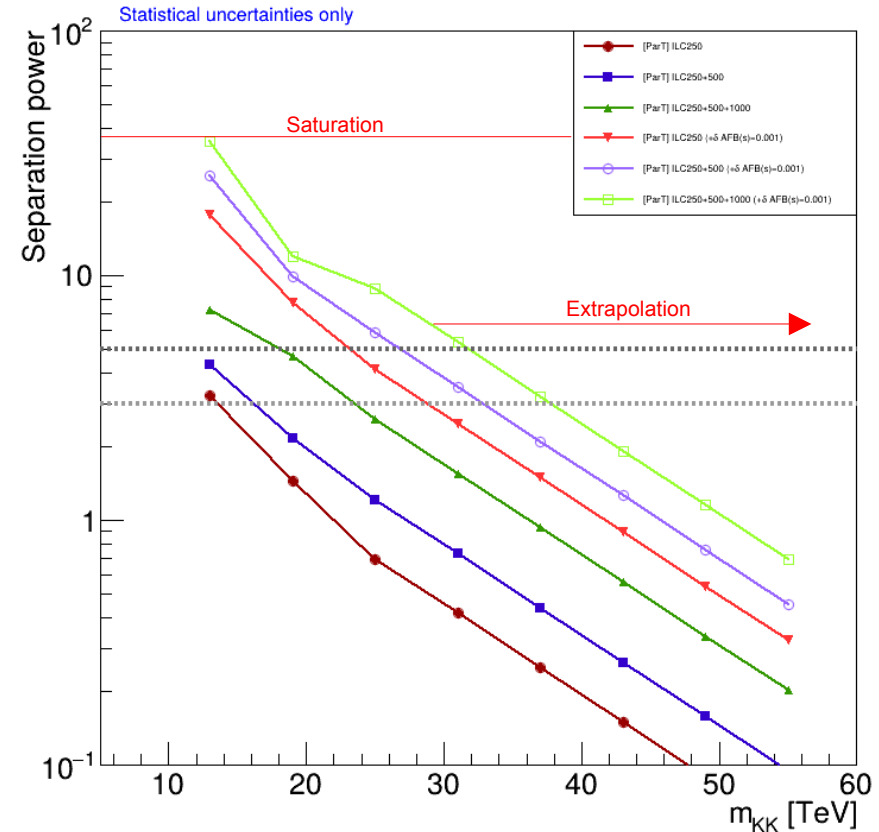
B+ models mass scale

- Worst ($dN/dx + 1\% \delta A_{FB}$ for s-quark) vs best (ParT + $1\% \delta A_{FB}$ for s-quark) prospects

B+ Models (b & c quarks)



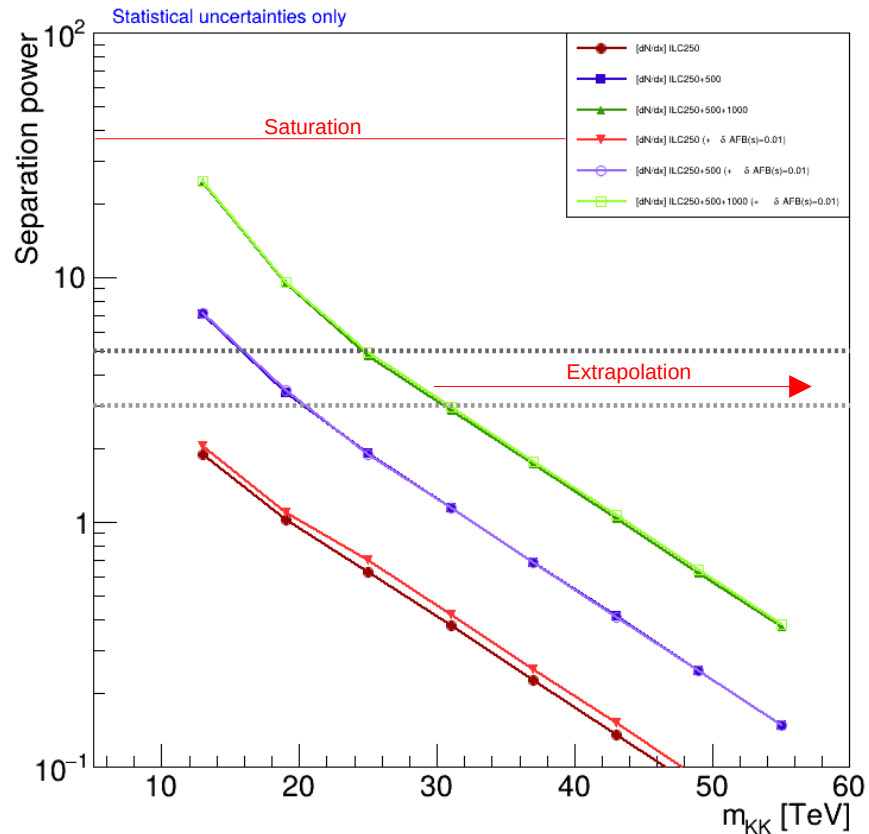
B+ Models (b & c quarks)



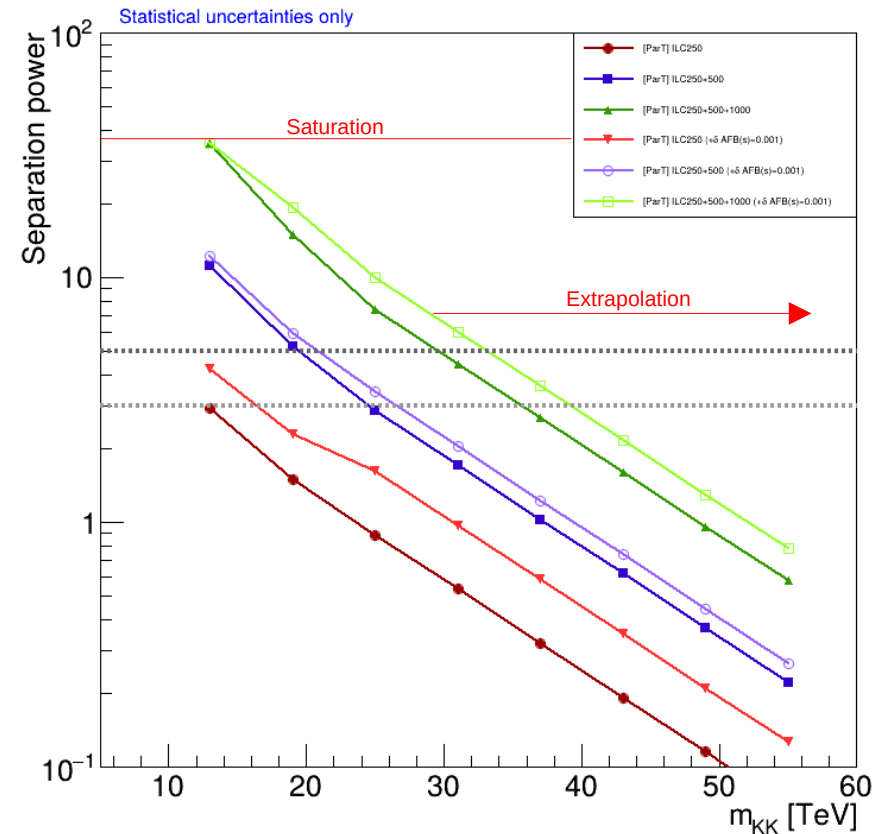
B- models mass scale

- Worst ($dN/dx + 1\% \delta A_{FB}$ for s-quark) vs best (ParT + $1\% \delta A_{FB}$ for s-quark) prospects

B- Models (b & c quarks)

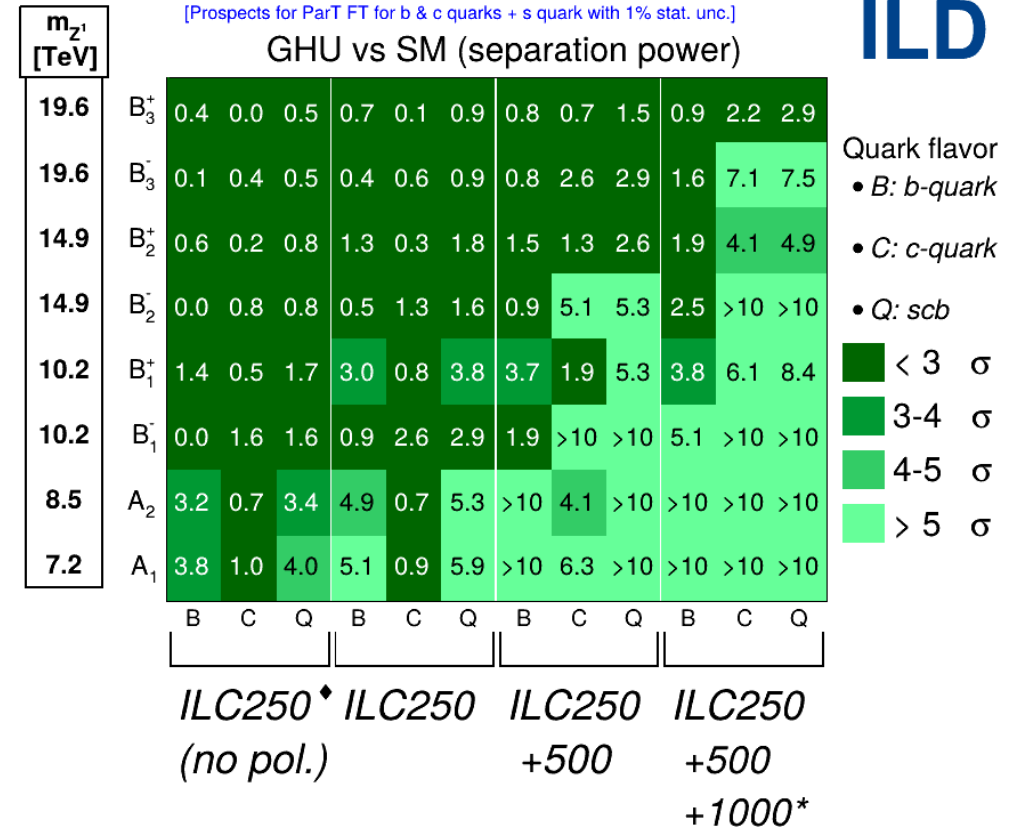
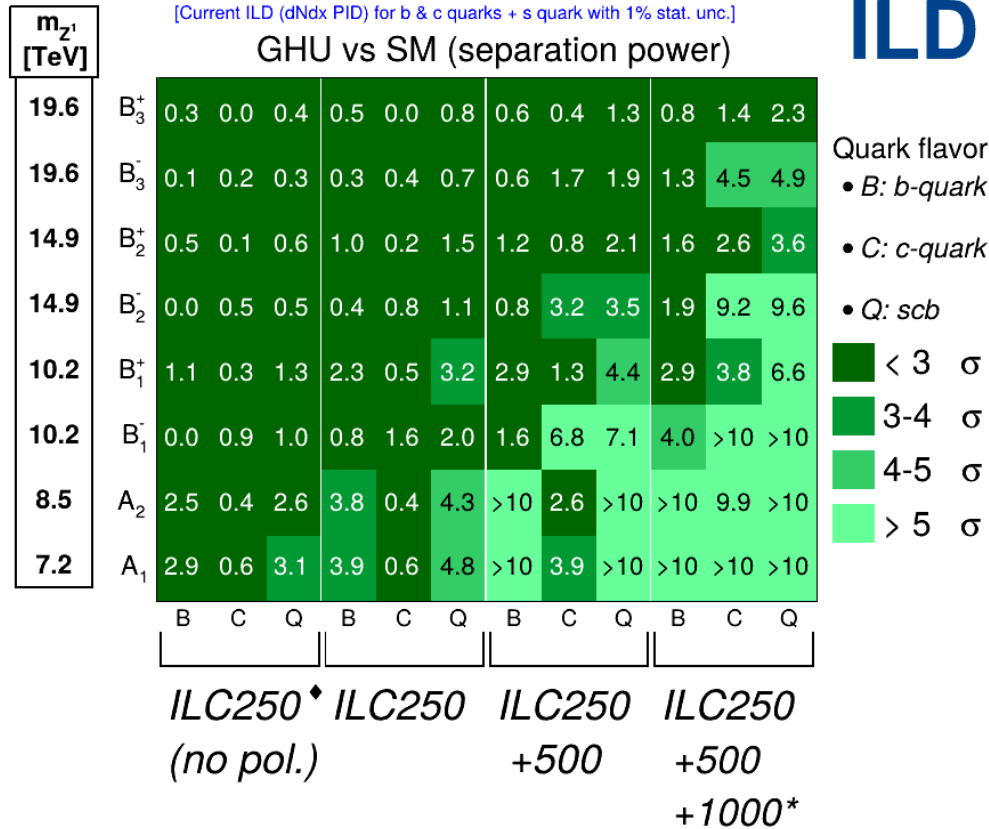


B- Models (b & c quarks)



Adding s quark (1% relative error)

- ILC with pixel TPC (dN/dx for PID) || ILC with prospects using ParT flavour tagging



Adding s quark (1‰ relative error)

- ILC with pixel TPC (dN/dx for PID) || ILC with prospects using ParT flavour tagging

