$e^-e^+ \rightarrow s\bar{s}$ at $\sqrt{s}$ = 250 GeV in future linear colliders

Jesús P. Márquez Hernández (IJCLAB, CNRS)

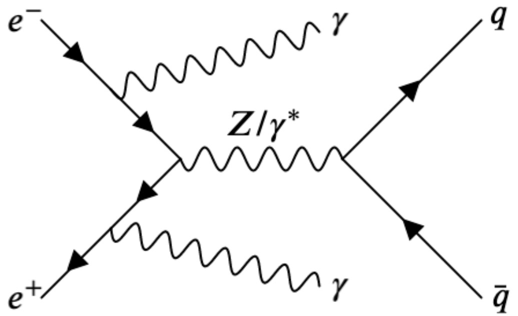
cnrs IJCLab Irène Joliot-Curie Laboratoire de Physique des 2 Infinis ILD CALO5D A new dimension in shower reconstruction

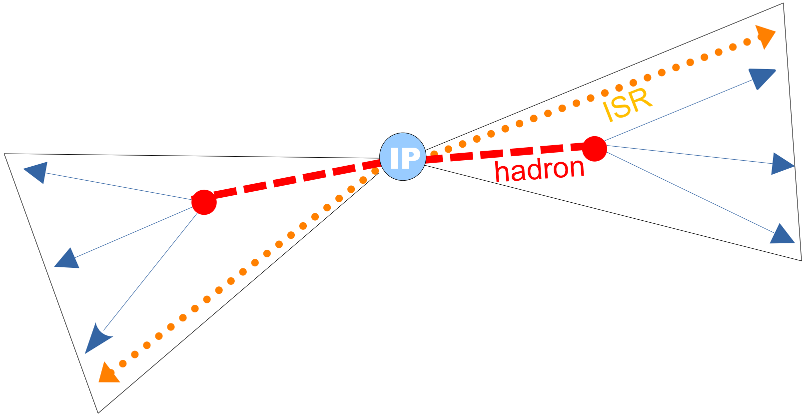
(b & c) diquark production in $e^- e^+$ collisions

- Topology: Two back-to-back jets
- MC simulations at 250 GeV
  - **ILC/LCF** run plan
  - **Full simulation** of the International Large Detector (ILD)
- Procedure:
  - 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)
  - Charge measurement → Quark-Antiquark identification
    - Double charge

R_q

A_{FB}

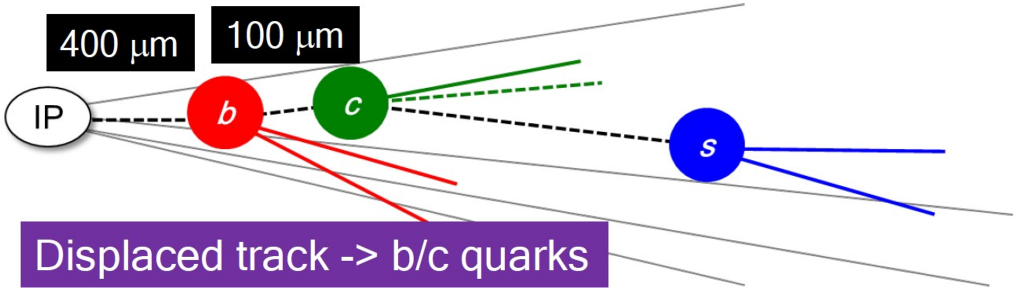




How can we move from here to strange quarks?
Can we get ‰-level uncertainties like for the b & c quarks?

From b/c to strange quark

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



- Strange quark mostly 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 s-quark/ud-quark signals (Modification of Y. Okugawa's analysis)
  - After the $q\bar{q}$ selection

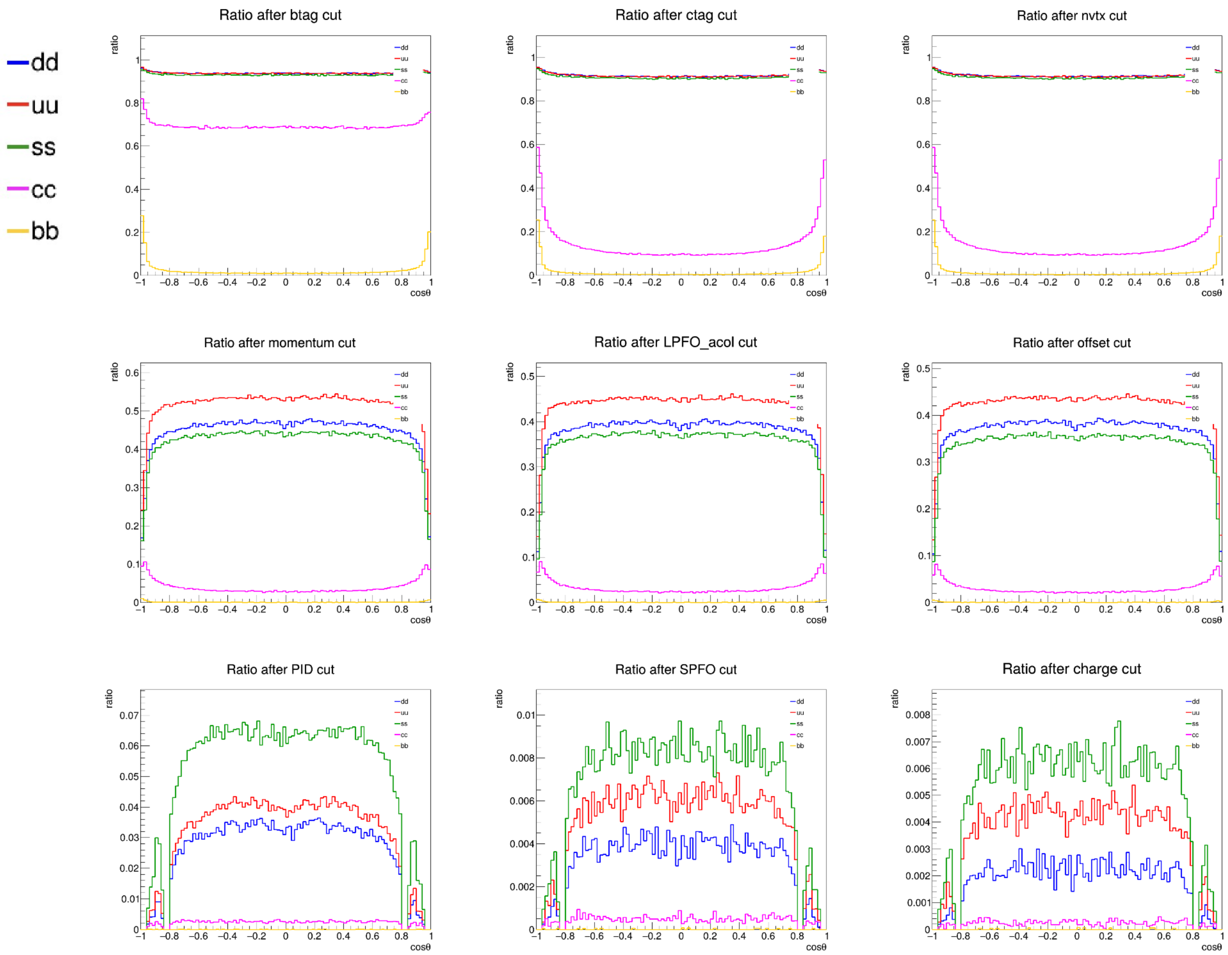
	#	Name	Quantity	Description
uds selection	1	b-tag	$btag < 0.3$	Reject events with b-like jets
	2	c-tag	$ctag < 0.65$	Reject events with c-like jets
	3	nvtx	$nvtx = 1$	Jets should have only PV as vertex
Cut-based s-tag (or ud-tag)	4	Leading momentum	$p_{LPFO} > 15\,\mathrm{GeV}$	Leading momentum cut
	5	LPFO acollinearity	$\cos\theta_{LPFO_{1,2}} > 0.97$	LPFOs should be back-to-back
	6	Offset	$V_0 = \sqrt{d_0^2 + z_0^2} < 1\,\mathrm{mm}$	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 $p_{SPFO} > 10\,\mathrm{GeV}$ 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

Cuts visualization (K selection for s-jets)

- Results for P(e-,e+)=(-0.8,+0.3)
- Flat when |cos(θ)| < 0.8

Efficiency (%)

	dd	uu	ss	cc	bb
+ Cut 1	93.9	93.9	93.1	69.3	2.11
+ Cut 2	91.8	91.6	90.9	14.0	1.35
+ Cut 3	91.8	91.6	90.9	14.0	1.35
+ Cut 4	44.9	51.8	42.3	4.01	0.07
+ Cut 5	38.2	43.9	35.9	3.35	0.06
+ Cut 6	36.8	42.4	34.0	3.10	0.05
+ Cut 7	2.38	2.91	4.80	0.22	<0.01
+ Cut 8	0.29	0.46	0.63	0.04	<0.01
+ Cut 9	0.17	0.33	0.48	0.02	<0.01



Reconstruction of A_{FB}

- The signal data is estimated by resting the expected angular distributions of backgrounds and doing a set of corrections to the selected signal:
  - Efficiency estimation
  - Kaon PID stability
  - Charge migration (p-q method)

$$A_{FB} = \frac{\int_0^1 d\sigma_\theta d\cos\theta - \int_{-1}^0 d\sigma_\theta d\cos\theta}{\int_{-1}^1 d\sigma_\theta d\cos\theta}$$

$$A_{FB}^{exp} = \frac{N_F - N_B}{N_F + N_B}$$

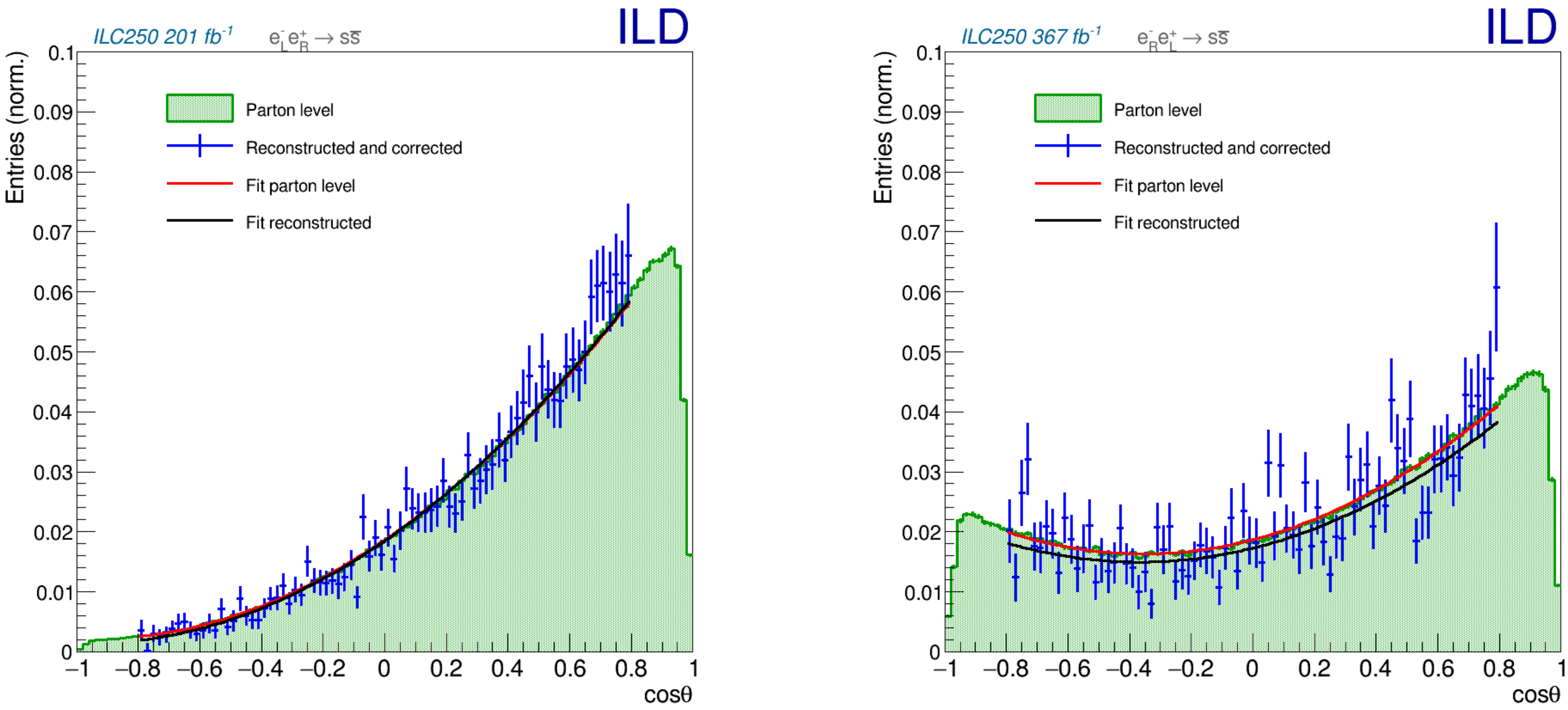
- A fit is performed to the corrected signal:

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

- Pseudo-experiments are performed for an estimation of systematical uncertainties due to the “tagging and correction” process (impact of q = u, d, b, c backgrounds)
  - 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 |cosθ| < 0.8 shows good agreement

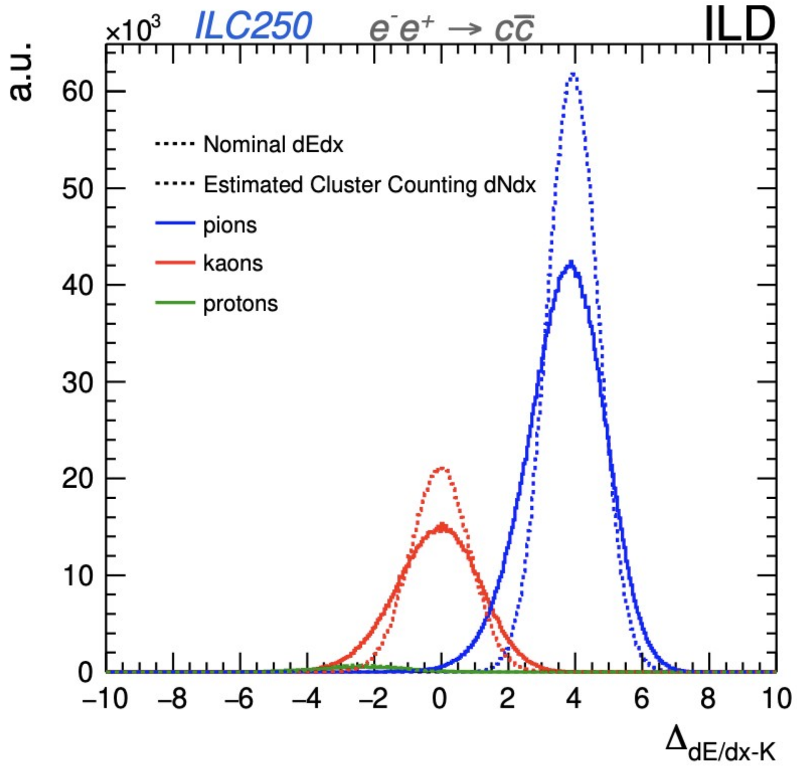


PID hardware prospects

- 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 a Perfect PID case

- Caveat: Only PFOs with PID available are improved

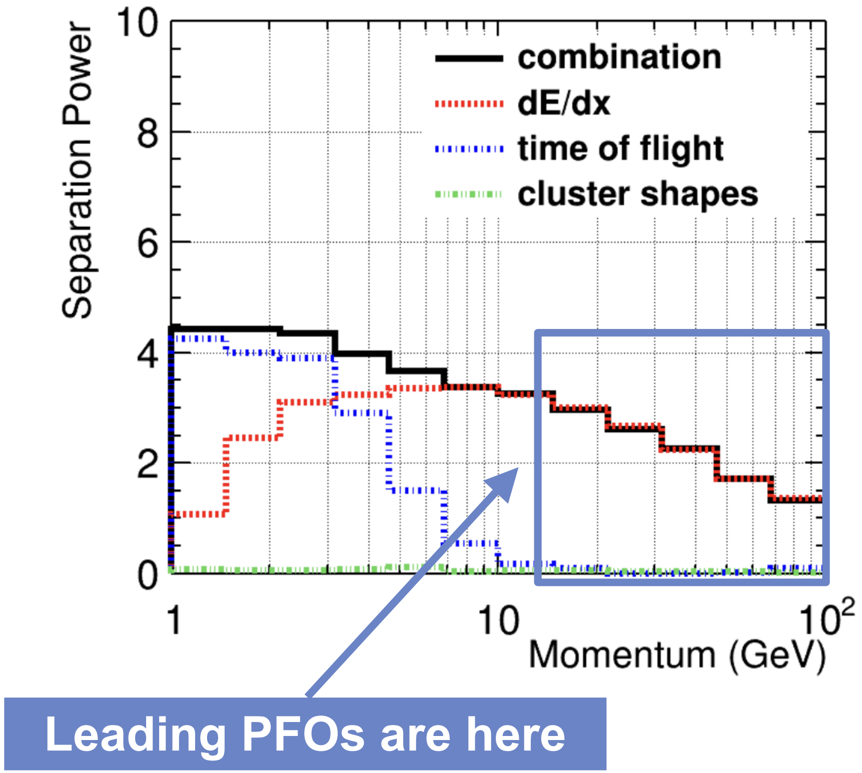


https://github.com/QQbarAnalysis/CheatdEdxDist

PID software improvement: CPID

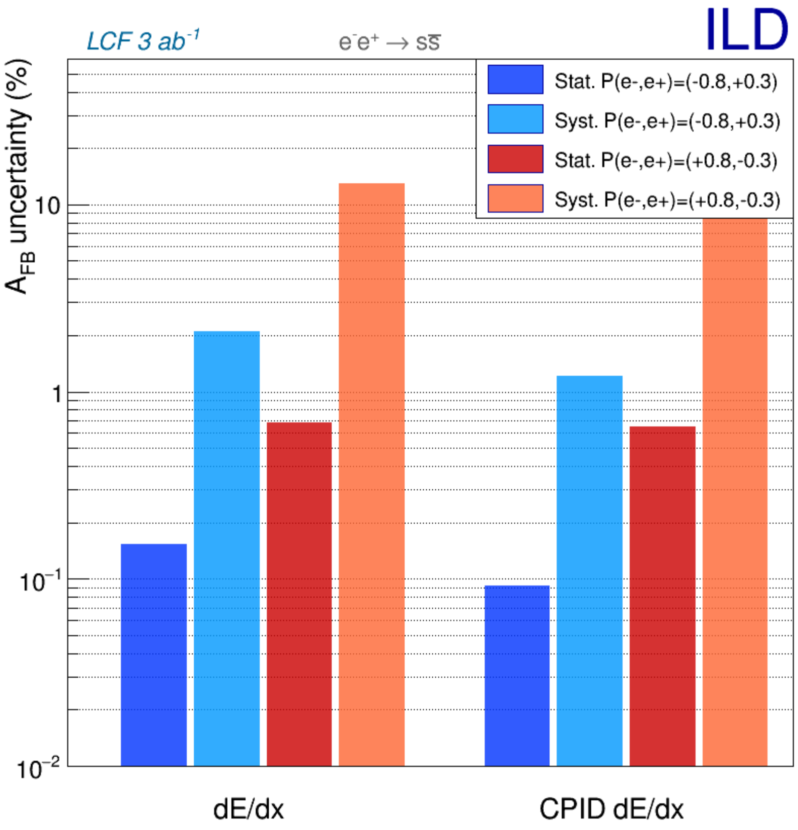
- **C**omprehensive **P**article **ID** Marlin processor:
  - Uses different PID inputs (dE/dx, TOF, etc.)
  - Uses a BDT-based ML algorithm for classification
  - Easy to adapt to different MC ids or PID info

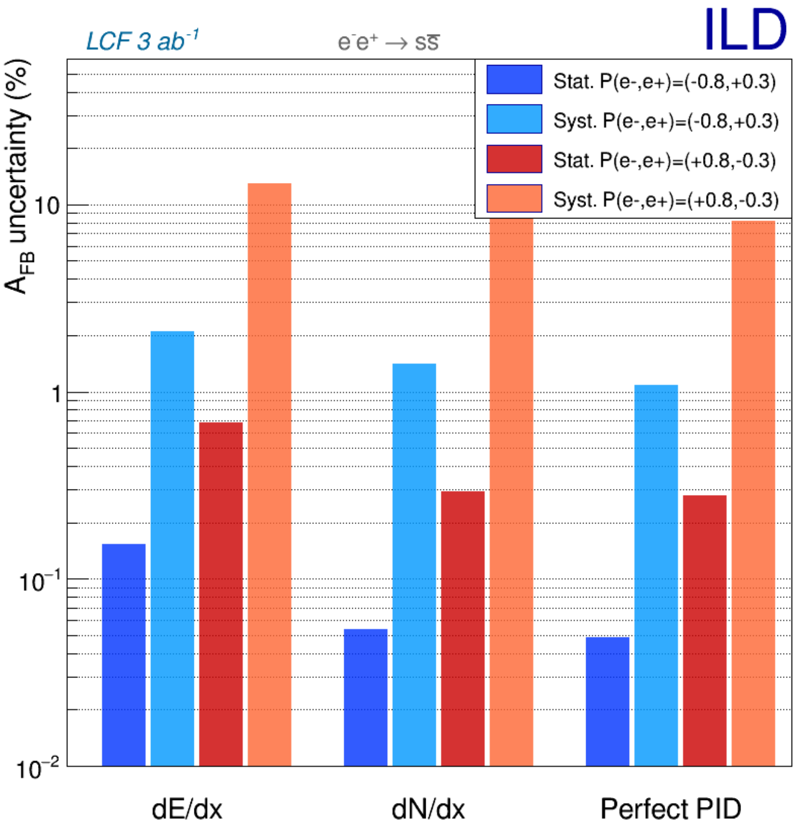
- In our case, the CPID was trained tackling our leading PFOS:
  - Only Kaon/Pion/Proton separation
  - 3 GeV < Momentum < 100 GeV



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

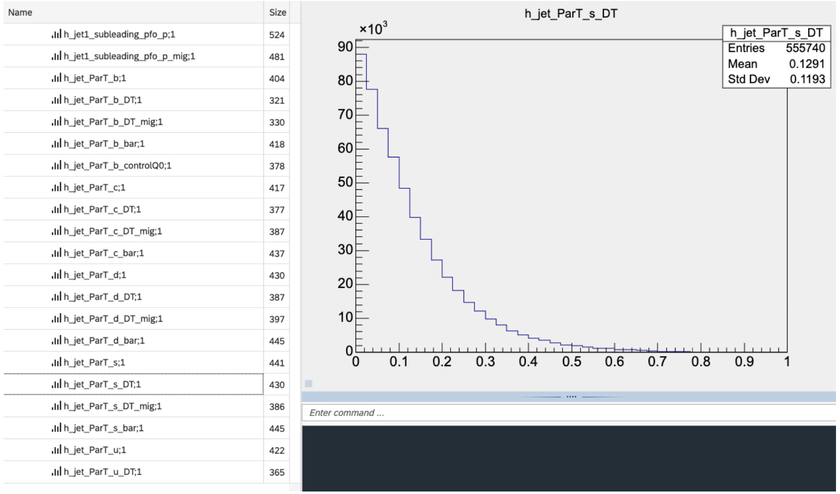
Results (LCF250)





Selection using Particle Transformer (ParT) tagging

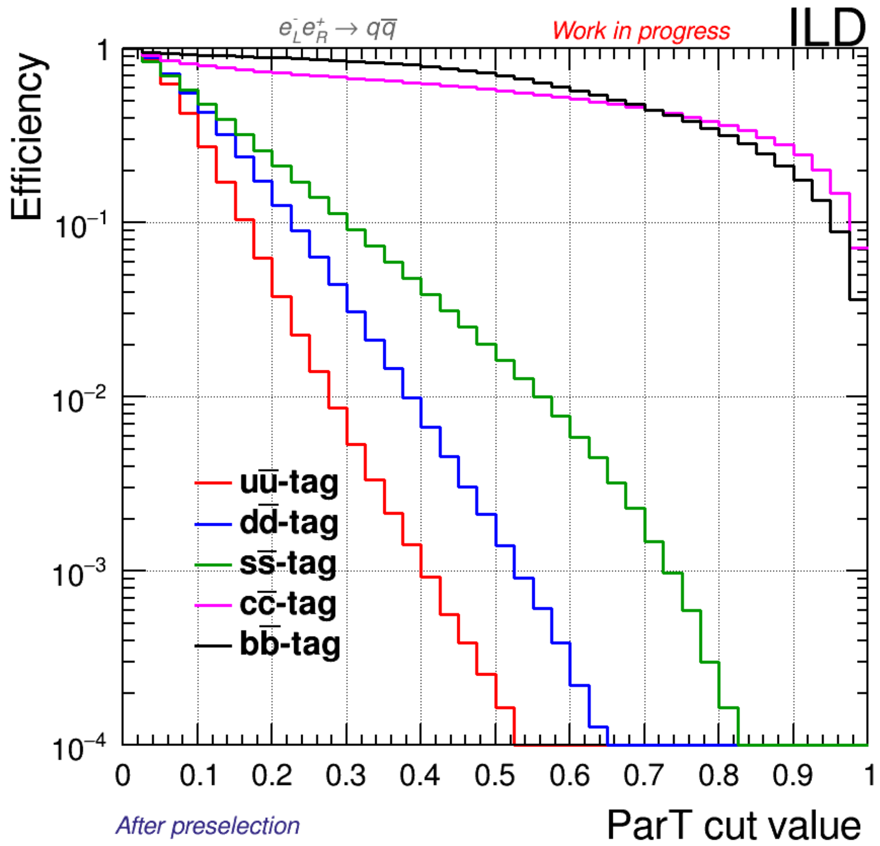
- New approach with a direct $q\bar{q}$ combined score
  - ParT_{DT} defined as: $\max(j^1_q \cdot j^2_{\bar{q}},\ j^1_{\bar{q}} \cdot j^2_q)$
- Easy to cut off events with both jets having similar tag (qq or $\overline{qq}$)
- Now the Double Tag and Double Charge happens at once
  - No PID corrections
  - No p-q method
  - Only an efficiency correction is applied to the data!

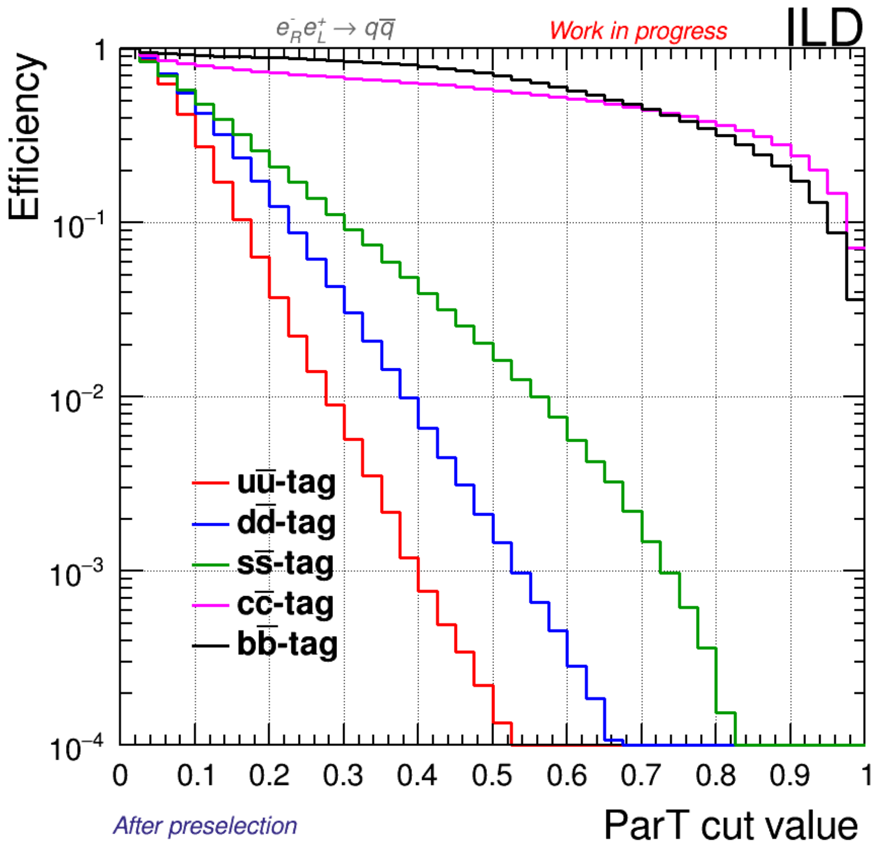


Tag with 11 categories (q, $\bar{q}$, g)

Signal efficiency for different ParT tagging cuts

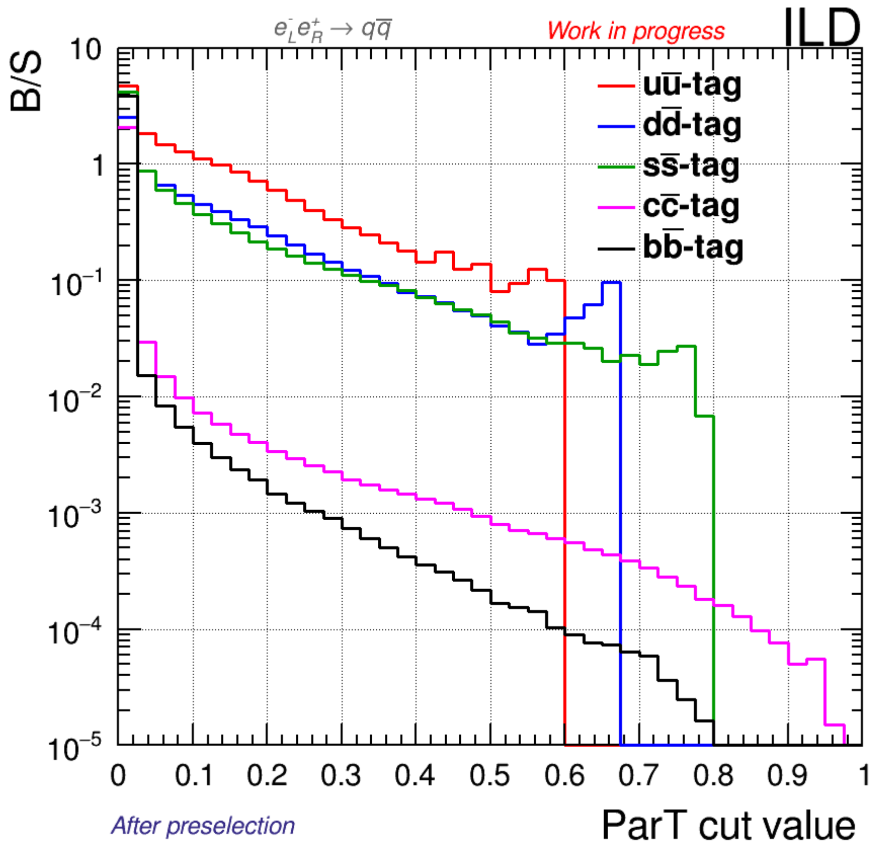
- I set different WP for each flavor (0.4 for s, 0.6 for b/c)

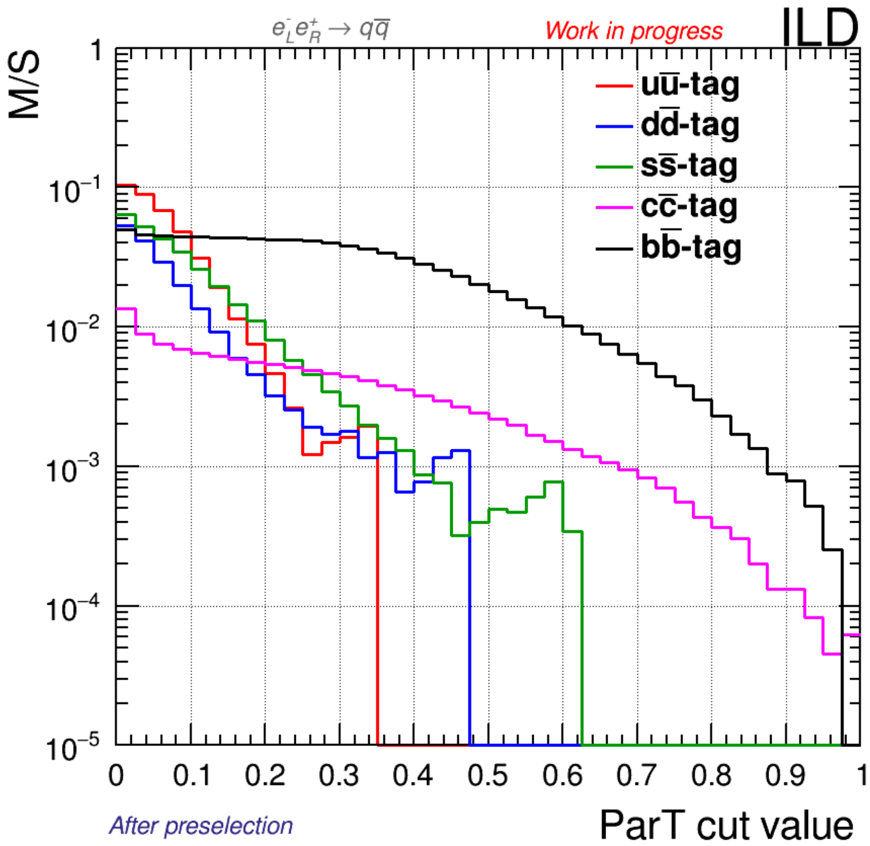




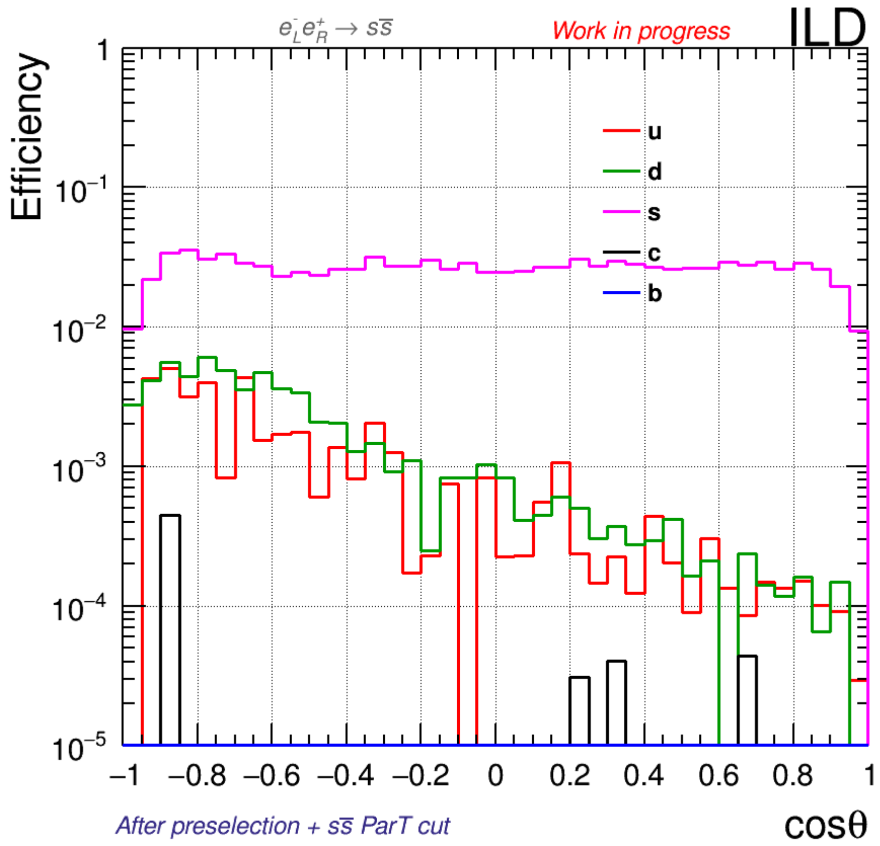
Background/signal and Migrations/signal ($e^-_L e^+_R$)

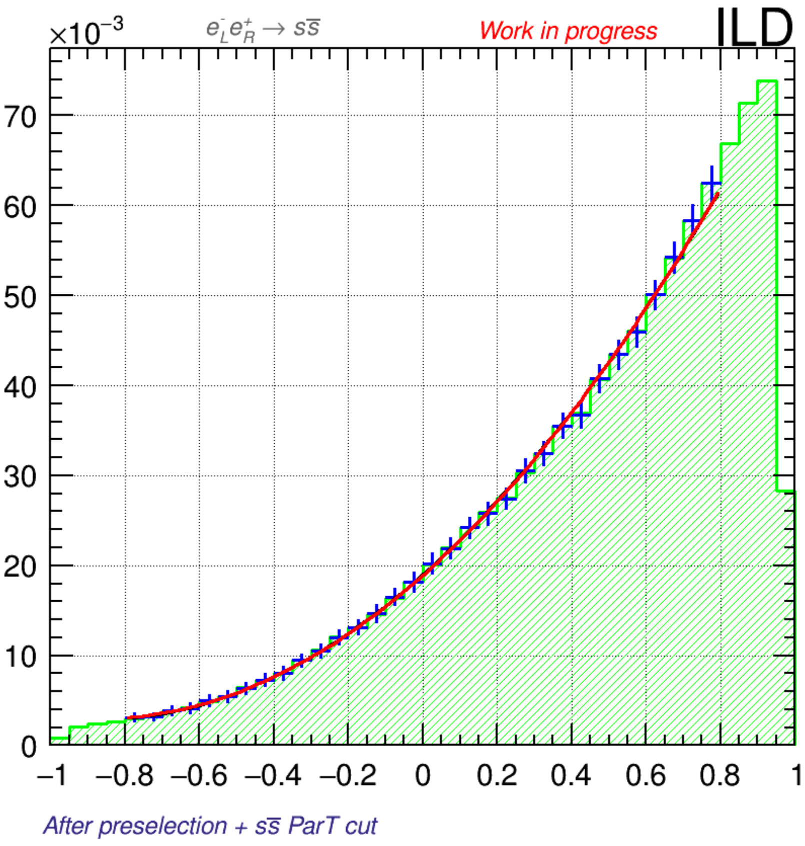
- I set different WP for each flavor (0.4 for s, 0.6 for b/c)



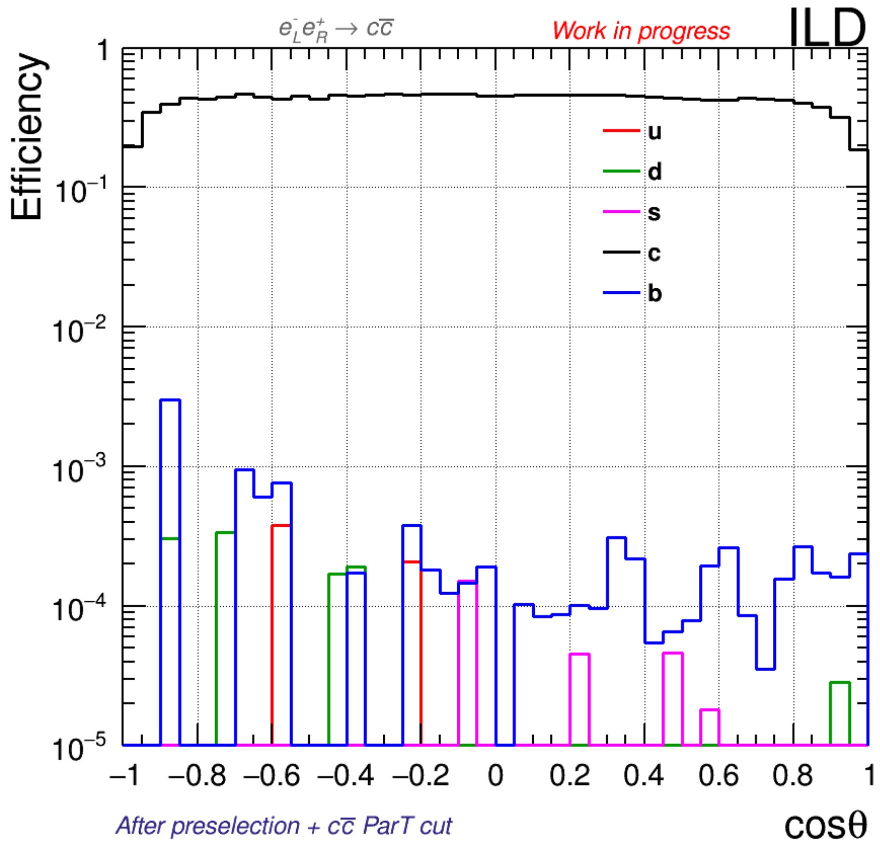


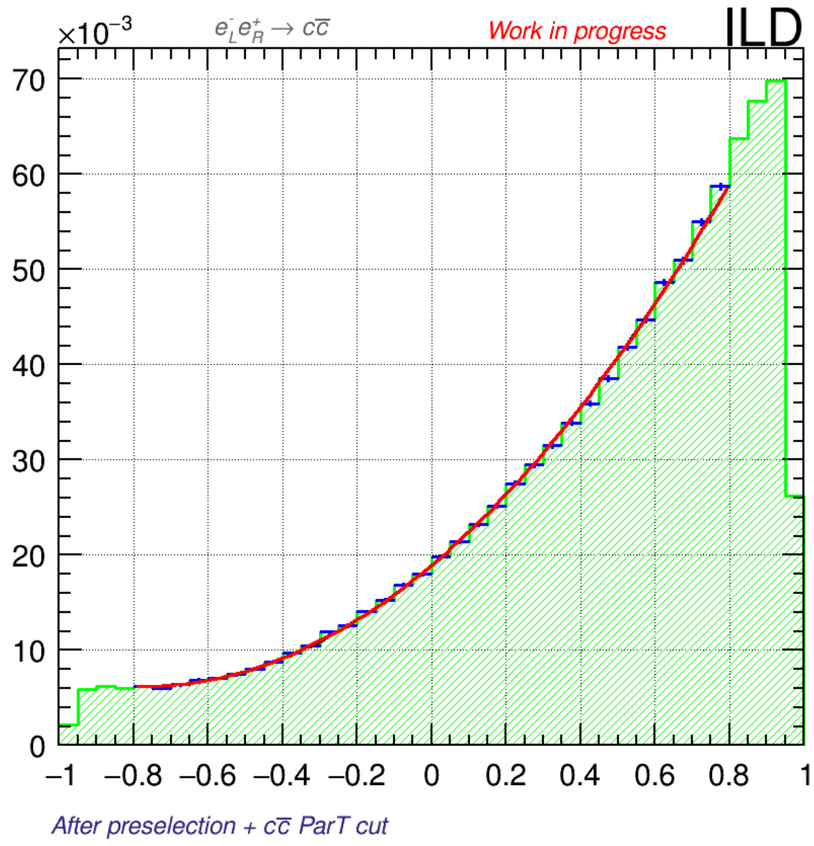
Selection and fit for $s\bar{s}$



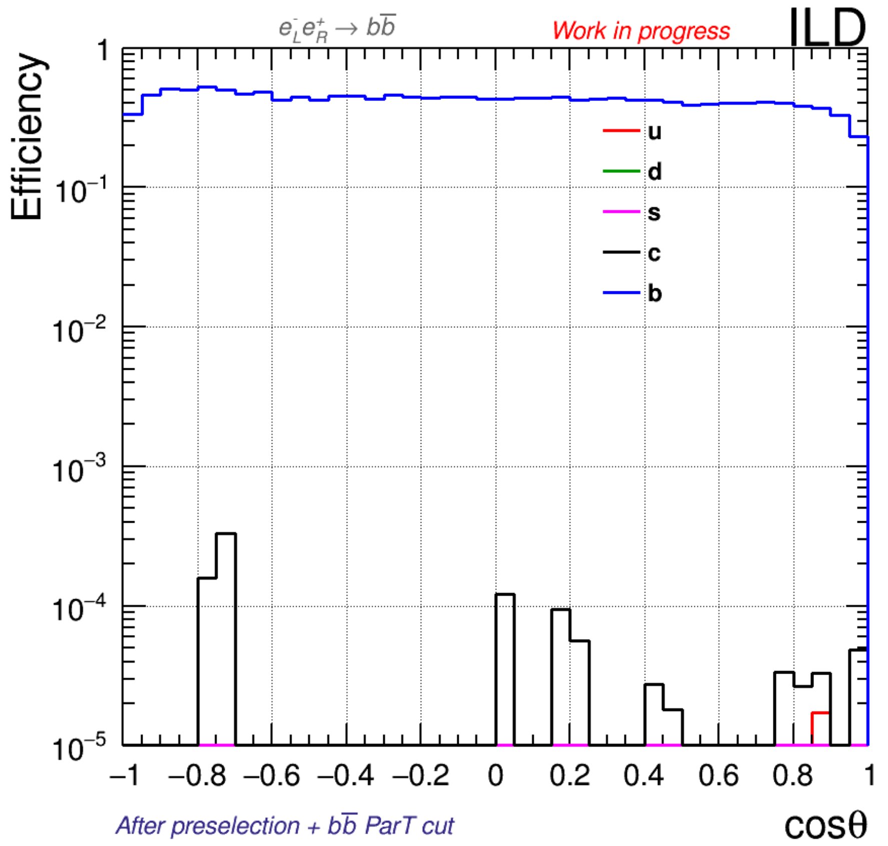


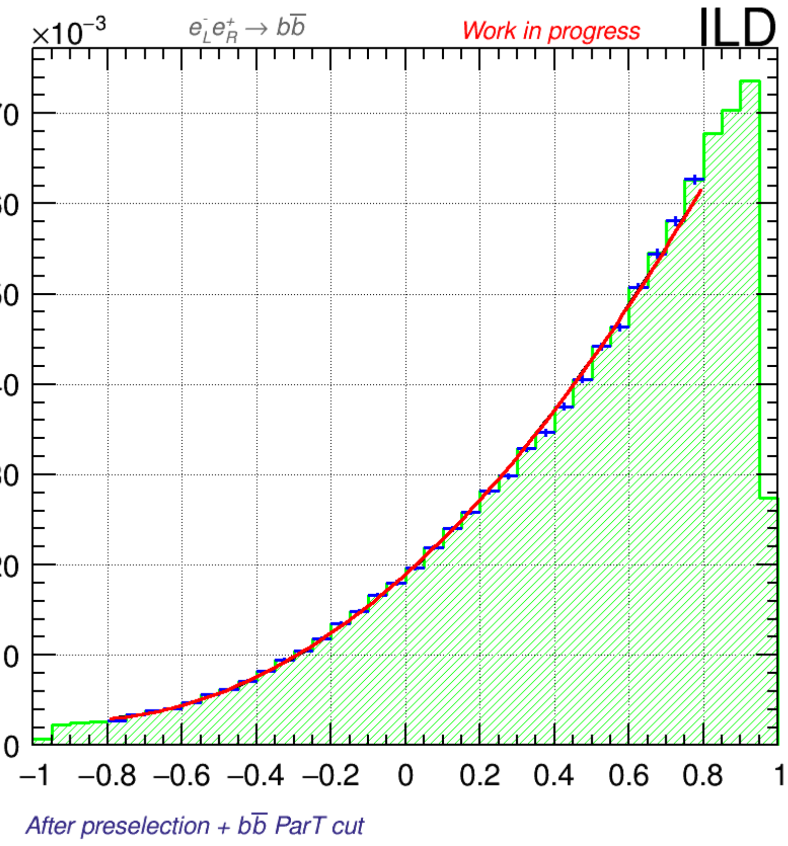
Selection and fit for $c\bar{c}$





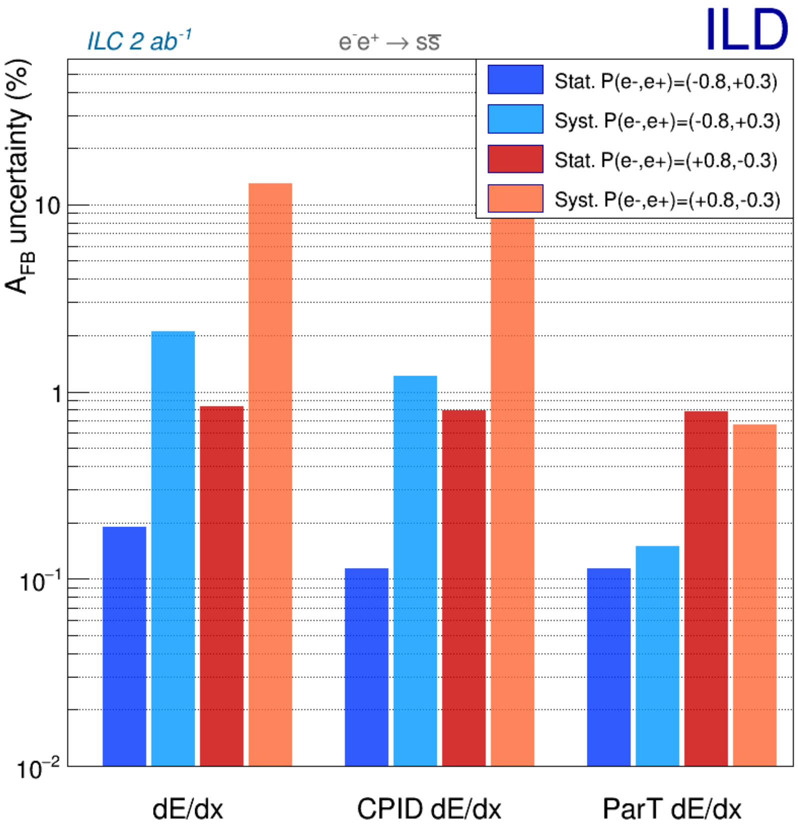
Selection and fit for $b\bar{b}$

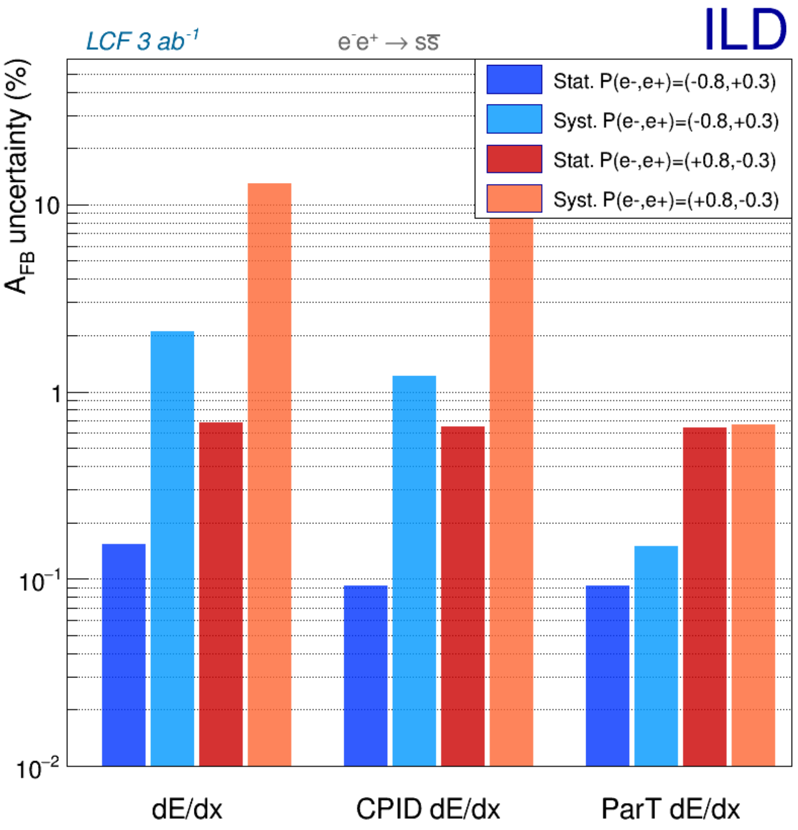




Results for s-quark (ILC250 & LCF250)

- O(10) improvement of systematic uncertainties!

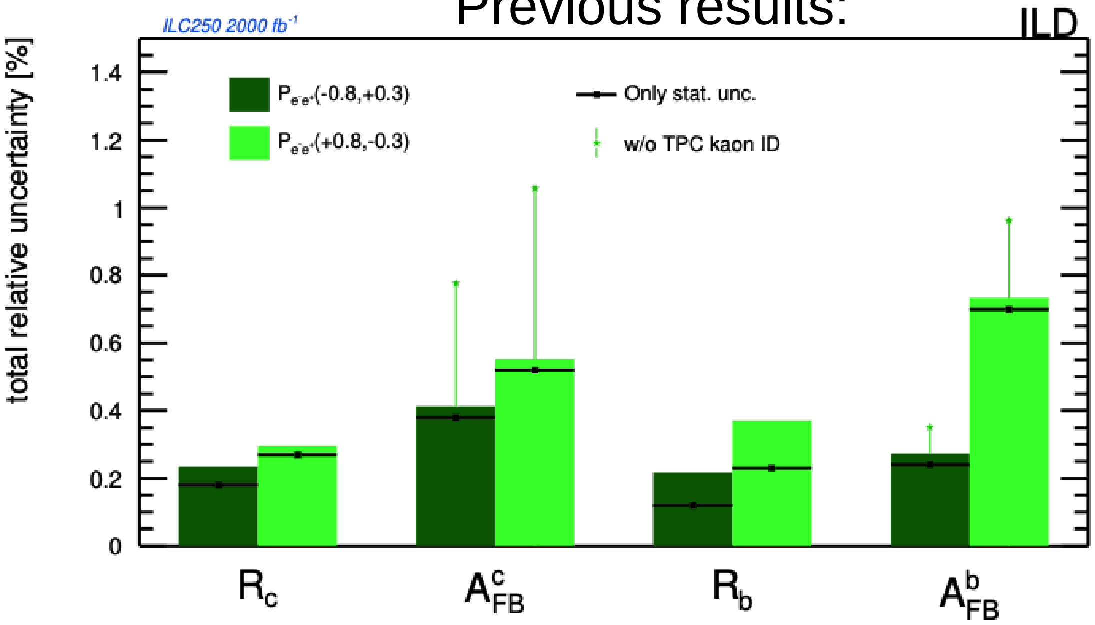




Preliminary uncertainties for b & c quarks

- Relative statistical uncertainties:
  - B-quark
    - For $e^-_L e^+_R$ ~ 0.04 %
    - For $e^-_R e^+_L$ ~ 0.5 %
  - C-quark
    - For $e^-_L e^+_R$ ~ 0.06 %
    - For $e^-_R e^+_L$ ~ 0.05 %

Previous results:



Source	$e^-e^+ \to c\bar{c}$				$e^-e^+ \to b\bar{b}$			
	$P_{e^-e^+}(-0.8,+0.3)$		$P_{e^-e^+}(+0.8,-0.3)$		$P_{e^-e^+}(-0.8,+0.3)$		$P_{e^-e^+}(+0.8,-0.3)$	
	R_c	$A_{FB}^{c\bar{c}}$	R_c	$A_{FB}^{c\bar{c}}$	R_b	$A_{FB}^{b\bar{b}}$	R_b	$A_{FB}^{b\bar{b}}$
Statistics	**0.18%**	**0.38%**	**0.27%**	**0.52%**	**0.12%**	**0.24%**	**0.23%**	**0.70%**
Preselection eff.	<0.01%	0.12%	0.02%	0.16%	<0.01%	0.08%	0.06%	0.12%
Background	0.01%	0.01%	0.02%	0.02%	0.01%	0.01%	0.06%	<0.01%
heavy quark mistag	0.11%	<0.01%	0.06%	<0.01%	0.12%	<0.01%	0.22%	<0.01%
uds mistag	0.03%	<0.01%	0.02%	<0.01%	0.08%	<0.01%	0.14%	<0.01%
Angular correlations	0.10%	0.10%	0.10%	0.10%	0.10%	0.10%	0.10%	0.10%
Beam Polarisation	<0.01%	<0.01%	0.02%	0.01%	<0.01%	0.01%	0.03%	0.15%
Systematics	**0.15%**	**0.16%**	**0.12%**	**0.19%**	**0.18%**	**0.13%**	**0.29%**	**0.22%**
Total	**0.24%**	**0.41%**	**0.30%**	**0.55%**	**0.21%**	**0.27%**	**0.37%**	**0.73%**

- Systematic uncertainties *from background subtraction*:
  - B-quark, both pol. < 0.01 %
  - C-quark, both pol. < 0.01 %

Prospects for GHU phenomenology

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

[Current ILD (dNdx PID) for b & c quarks + s quark with 1% stat. unc.]

ILD

GHU vs SM (separation power)

$m_{Z'}$ [TeV]		ILC250♦ (no pol.) B	ILC250♦ (no pol.) C	ILC250♦ (no pol.) Q	ILC250 B	ILC250 C	ILC250 Q	ILC250 +500 B	ILC250 +500 C	ILC250 +500 Q	ILC250 +500 +1000* B	ILC250 +500 +1000* C	ILC250 +500 +1000* Q
19.6	B_3^+	0.3	0.0	0.4	0.5	0.0	0.8	0.6	0.4	1.3	0.8	1.4	2.3
19.6	B_3^-	0.1	0.2	0.3	0.3	0.4	0.7	0.6	1.7	1.9	1.3	4.5	4.9
14.9	B_2^+	0.5	0.1	0.6	1.0	0.2	1.5	1.2	0.8	2.1	1.6	2.6	3.6
14.9	B_2^-	0.0	0.5	0.5	0.4	0.8	1.1	0.8	3.2	3.5	1.9	9.2	9.6
10.2	B_1^+	1.1	0.3	1.3	2.3	0.5	3.2	2.9	1.3	4.4	2.9	3.8	6.6
10.2	B_1^-	0.0	0.9	1.0	0.8	1.6	2.0	1.6	6.8	7.1	4.0	>10	>10
8.5	A_2	2.5	0.4	2.6	3.8	0.4	4.3	>10	2.6	>10	>10	9.9	>10
7.2	A_1	2.9	0.6	3.1	3.9	0.6	4.8	>10	3.9	>10	>10	>10	>10

Quark flavor
- *B: b-quark*
- *C: c-quark*
- *Q: scb*

$< 3\sigma$ | 3-4 σ | 4-5 σ | $> 5\sigma$

[Prospects for s, c, b quarks with 0.1% stat. unc.]

ILD

GHU vs SM (separation power)

$m_{Z'}$ [TeV]		ILC250♦ (no pol.) B	ILC250♦ (no pol.) C	ILC250♦ (no pol.) Q	ILC250 B	ILC250 C	ILC250 Q	ILC250 +500 B	ILC250 +500 C	ILC250 +500 Q	ILC250 +500 +1000* B	ILC250 +500 +1000* C	ILC250 +500 +1000* Q
19.6	B_3^+	0.6	0.1	1.1	3.1	0.1	5.2	3.5	2.0	7.2	4.2	>10	>10
19.6	B_3^-	0.2	0.7	0.9	1.5	1.3	2.6	2.0	6.1	6.8	5.3	>10	>10
14.9	B_2^+	1.0	0.3	1.8	6.2	0.6	9.9	6.6	3.9	>10	9.9	>10	>10
14.9	B_2^-	0.1	1.5	1.7	1.9	2.5	3.8	2.6	>10	>10	9.1	>10	>10
10.2	B_1^+	2.3	0.9	3.8	>10	1.6	>10	>10	5.6	>10	>10	>10	>10
10.2	B_1^-	0.1	3.1	3.4	3.8	5.2	7.2	5.3	>10	>10	>10	>10	>10
8.5	A_2	5.2	1.3	5.6	>10	1.6	>10	>10	>10	>10	>10	>10	>10
7.2	A_1	6.1	1.9	6.8	>10	2.4	>10	>10	>10	>10	>10	>10	>10

Quark flavor
- *B: b-quark*
- *C: c-quark*
- *Q: scb*

$< 3\sigma$ | 3-4 σ | 4-5 σ | $> 5\sigma$

Conclusions & next steps

- ParT shows great potential to reduce uncertainties
  - O(10) improvement in most observables!

- There's still some refinement to be done within this analysis:
  - Optimize WP
  - Flatness of selection for b-quark
  - Check migrations impact

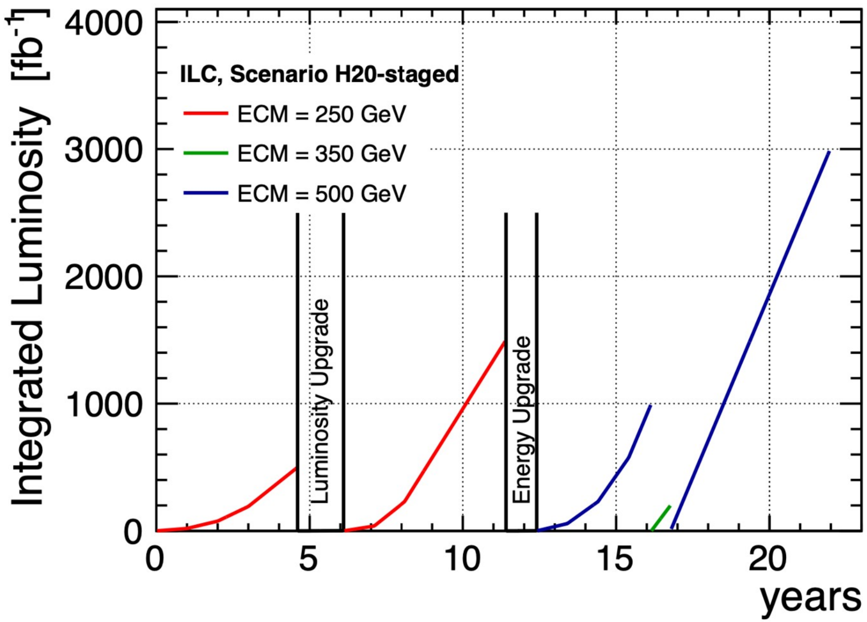
- Future plans:
  - Extension to 500 GeV analysis
  - Apply final results to GHU phenomenology
    - We have a set of new models with up to **100 TeV** Z' resonances to be tested!

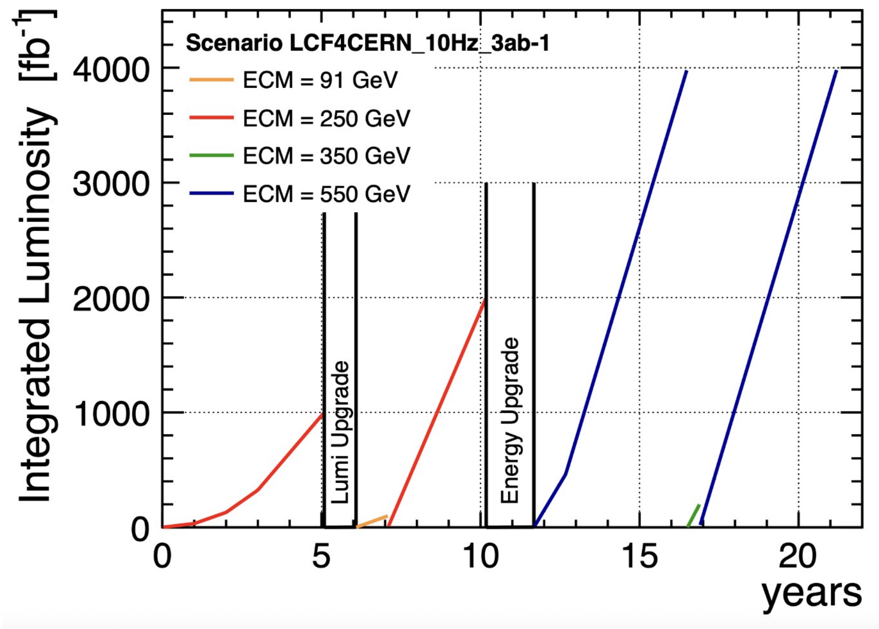
THANKS FOR YOUR ATTENTION

BACK-UP

BACK-UP

ILC H20-staged & LCF4CERN run plans



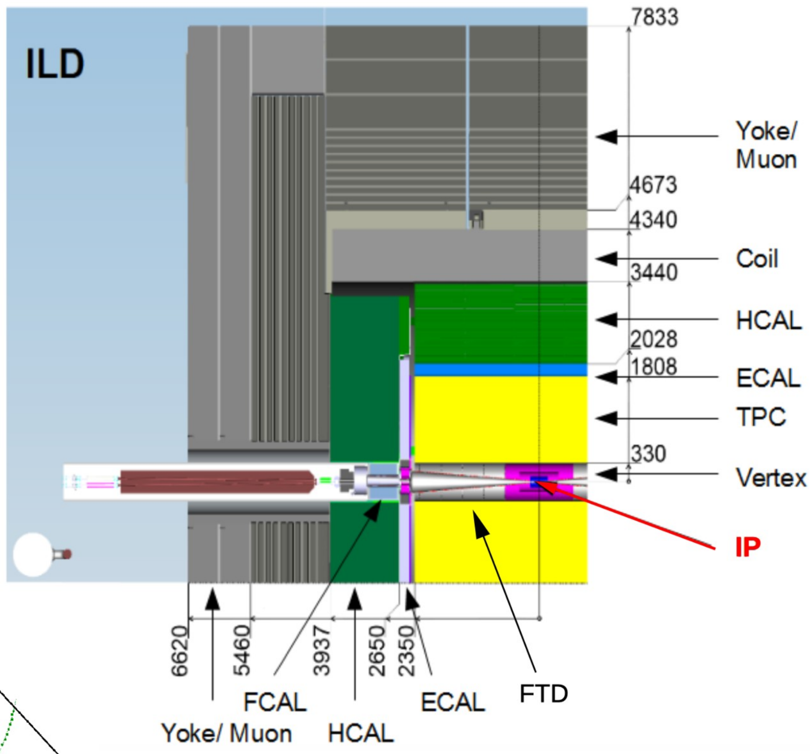


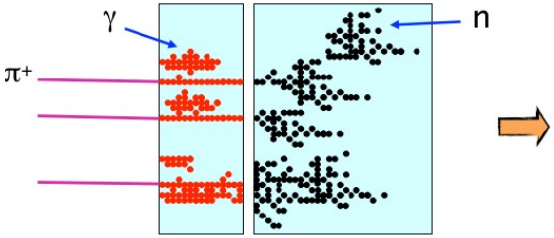
	91 GeV	250 GeV	350 GeV	500 GeV	1000 GeV
$\int \mathcal{L}$ (ab^{-1})	0.1	2	0.2	4	8
duration (yr)	1.5	11	0.75	9	10
beam polarization (e^-/e^+; %)	80/30	80/30	80/30	80/30	80/20
(–, -+, +-, ++) (%)	(10,40,40,10)	(5,45,45,5)	(5,68,22,5)	(10,40,40,10)	(10,40,40,10)
δ_{ISR} (%)	10.8	11.7	12.0	12.4	13.0
δ_{BS} (%)	0.16	2.6	1.9	4.5	10.5

Quantity	Symbol	Unit	Initial-250	Upgrades		Initial-550	Upgrade
Centre-of-mass energy	$\sqrt{s}$	GeV	250	250	550	550	550
Inst. Luminosity	$\mathscr{L}$	(10^{34}cm^{-2}s^{-1})	2.7	5.4	7.7	3.9	7.7
Polarisation	$\|P(e^-)\|/\|P(e^+)\|$	(%)	80 / 30	80 / 30	80 / 60	80 / 30	80 / 60
Bunches per pulse	n_{bunch}	1	1312	2625	2625	1312	2625
Average beam power	P_{ave}	MW	10.5	21	46	23	46
Site AC power	P_{site}	MW	143	182	322	250	322
Construction cost		BCHF	8.29	+0.77	+5.46	13.13	+1.40
Operation & maintenance		MCHF/y	170	196	342	291	342
Electricity		MCHF/y	66	77	142	115	142
Operating Personnel		FTE	640	640	850	850	850

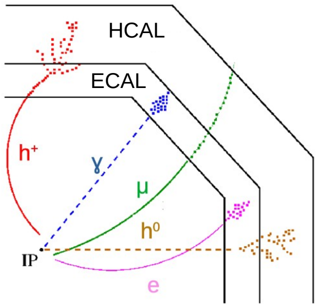
International Large Detector (ILD)

- ▶ It features excellent tracking, vertexing and IP constraining capabilities with minimal material budget
- ▶ **High granularity, compact** and **hermetic calorimetry system**
- ▶ **Full simulation available**: detailed geometry, materials, reconstruction chain, etc.
- ▶ **Optimized for Particle Flow**:
  - ▷ Determination of single particles
  - ▷ Based on Particle Flow Algorithms (PFA)
  - ▷ Powerful Particle identification (PID) tools
    - Jet energy measurement, flavor tagging, etc.



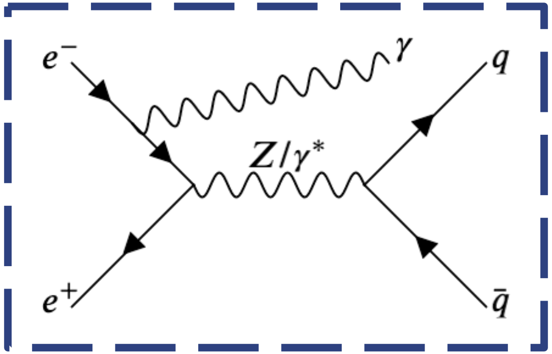


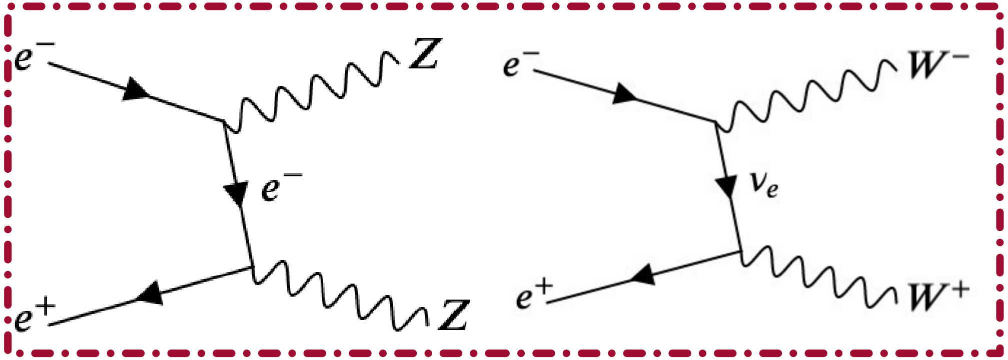
$E_{JET} = E_{ECAL} + E_{HCAL}$ → $E_{JET} = E_{TRACK} + E_{\gamma} + E_{n}$

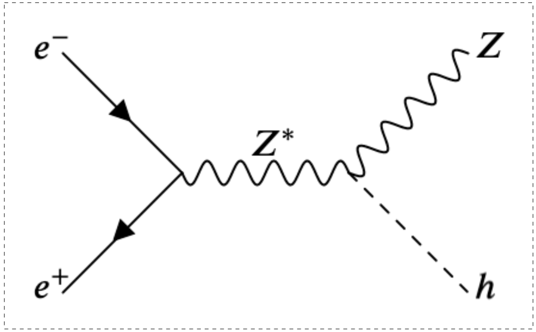


Preselection of $q\bar{q}$ signals

- 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

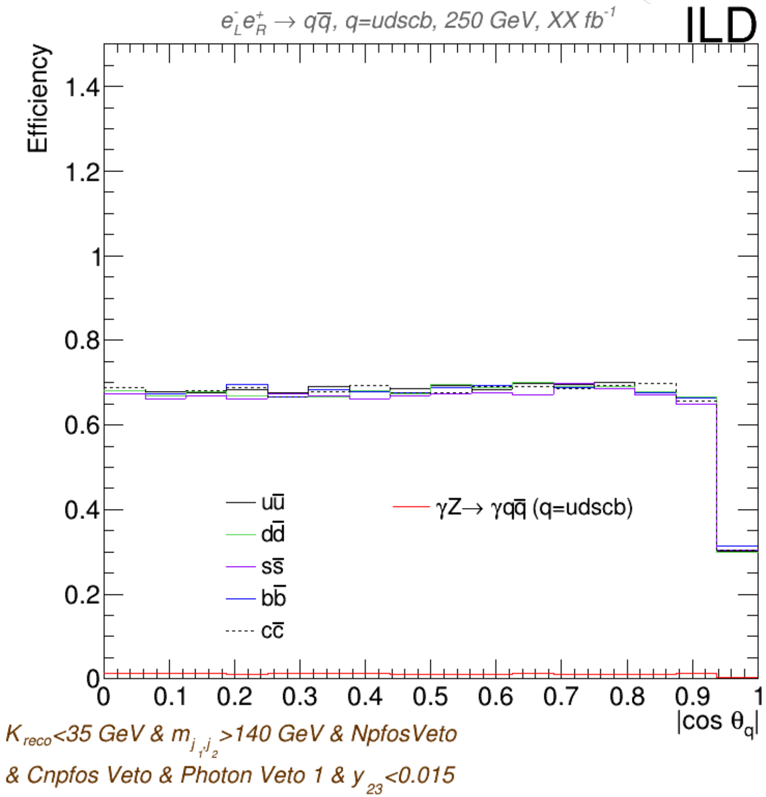
VLC Algorithm parameters:

- R = 1.0
- γ = 0.0
- β = 1.0

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

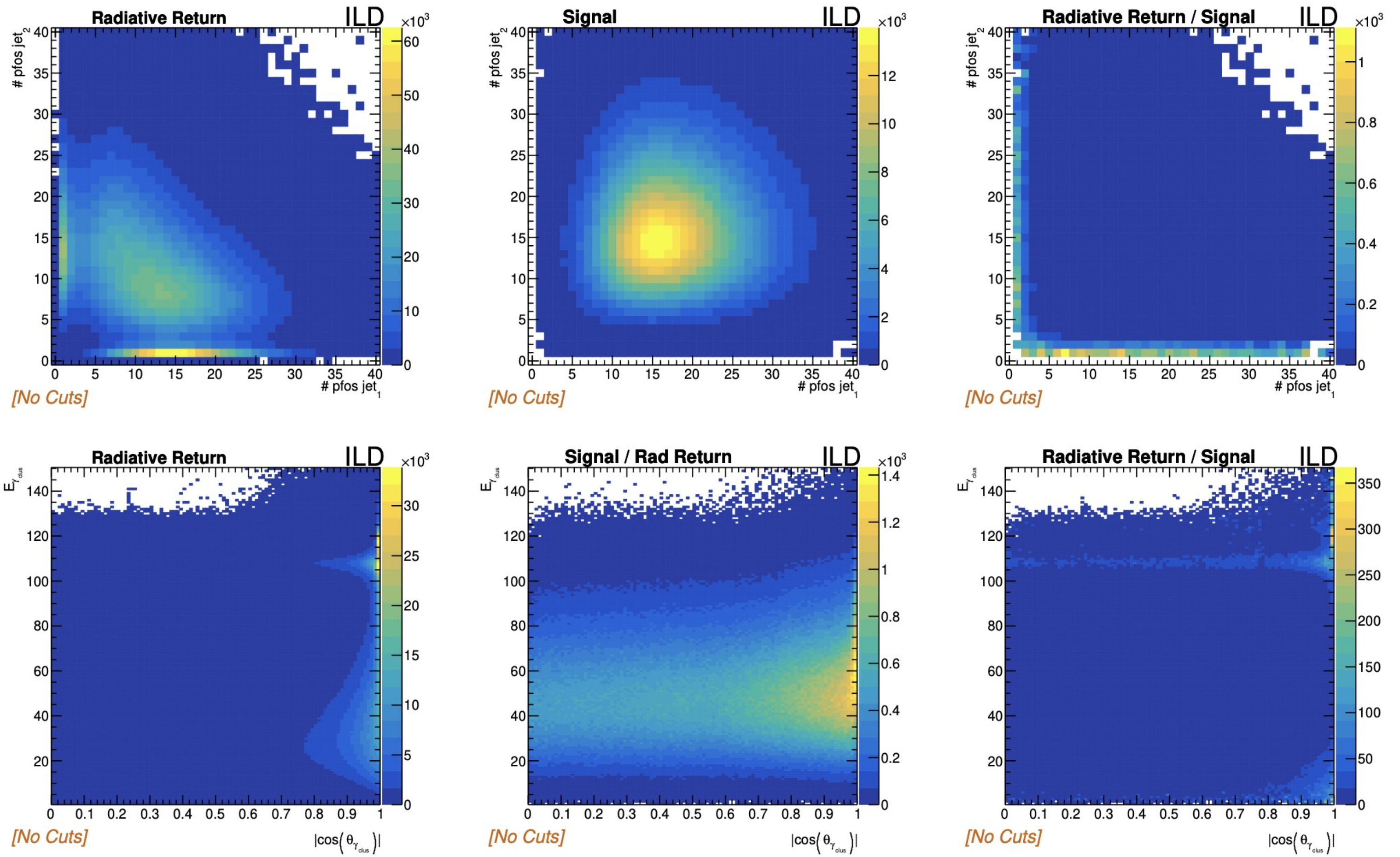
← |cosθ| < 0.9

Total efficiency of the preselection for the different quark flavours and radiative return for the chosen configuration (γ=0). The second row is for |cosθ| < 0.9

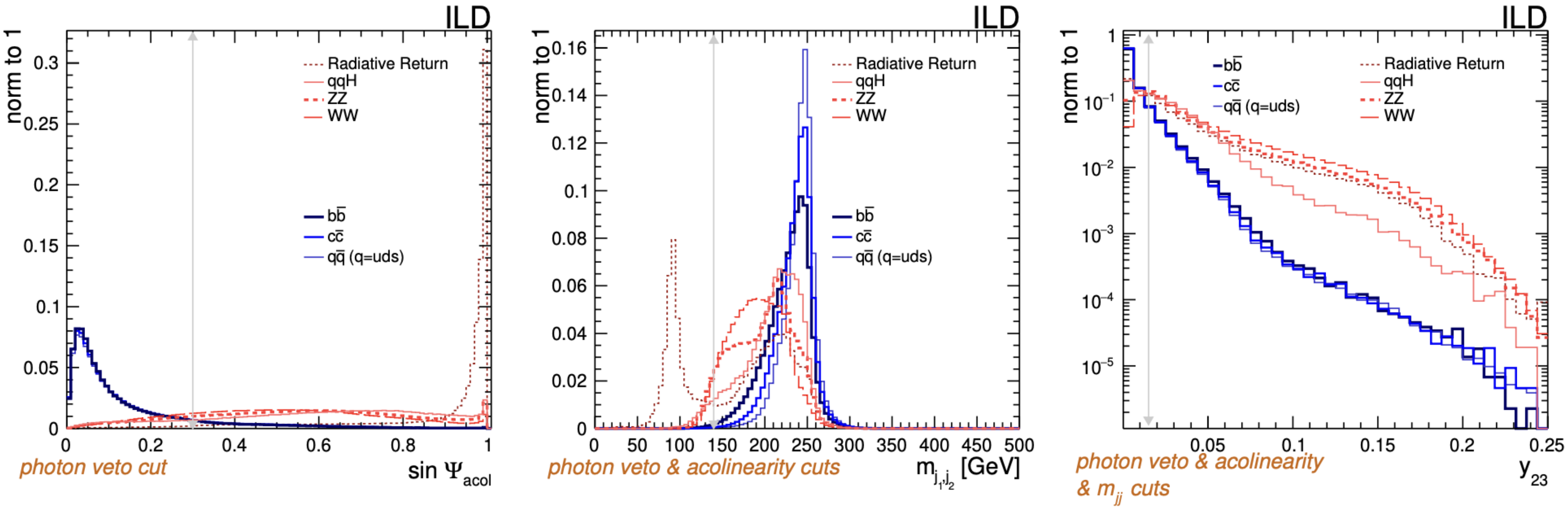


Efficiency of the preselection for the different quark flavours vs the angular distribution of the two jet system (new samples, final configuration)

Radiative return event rejection

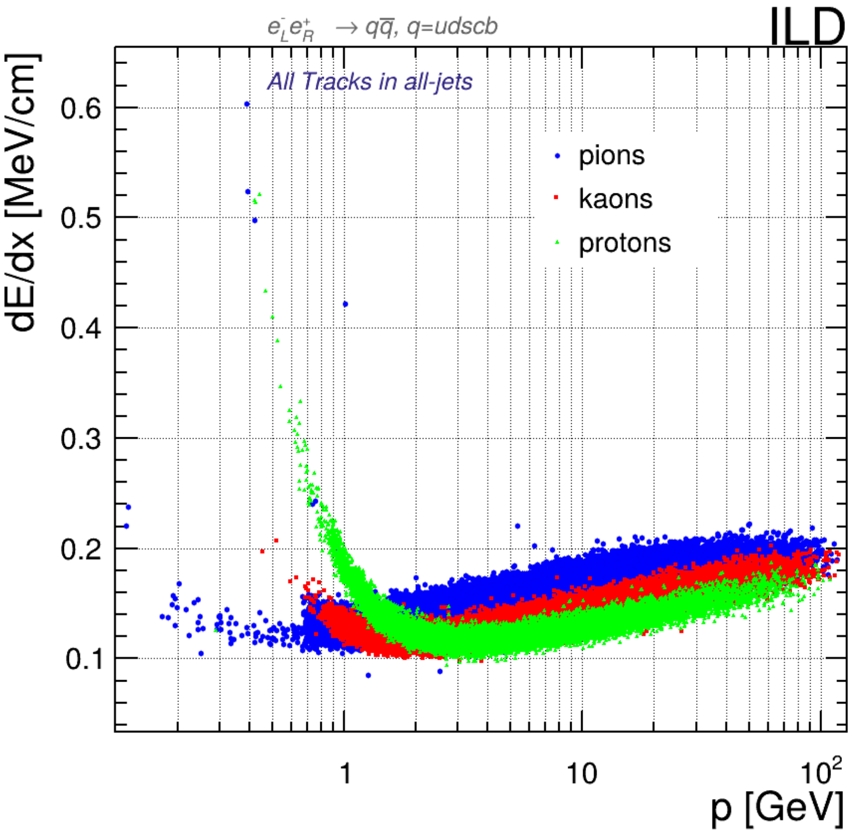


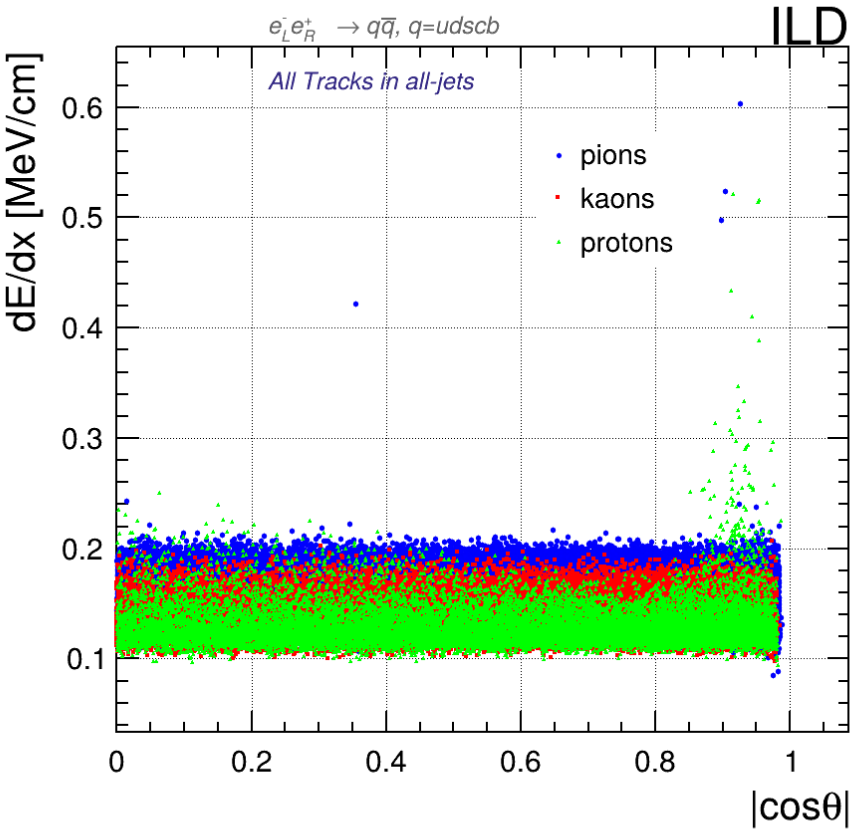
Background rejection



PID via dE/dx: Starting point

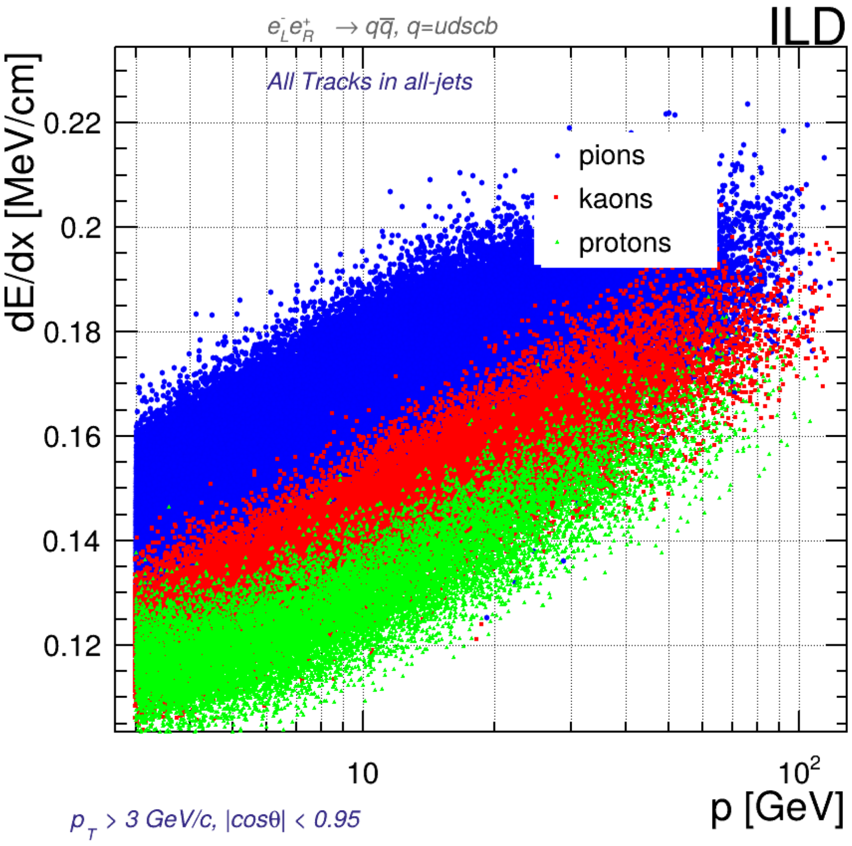
- Not all tracks/PFOs are valid for dE/dx

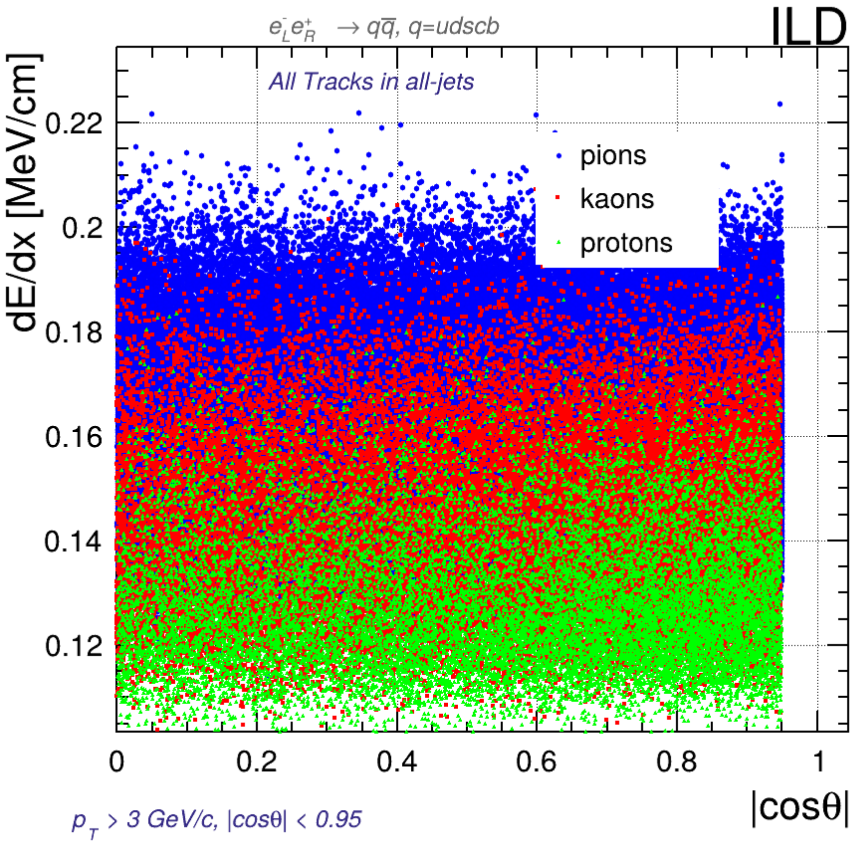




PID: Preselection

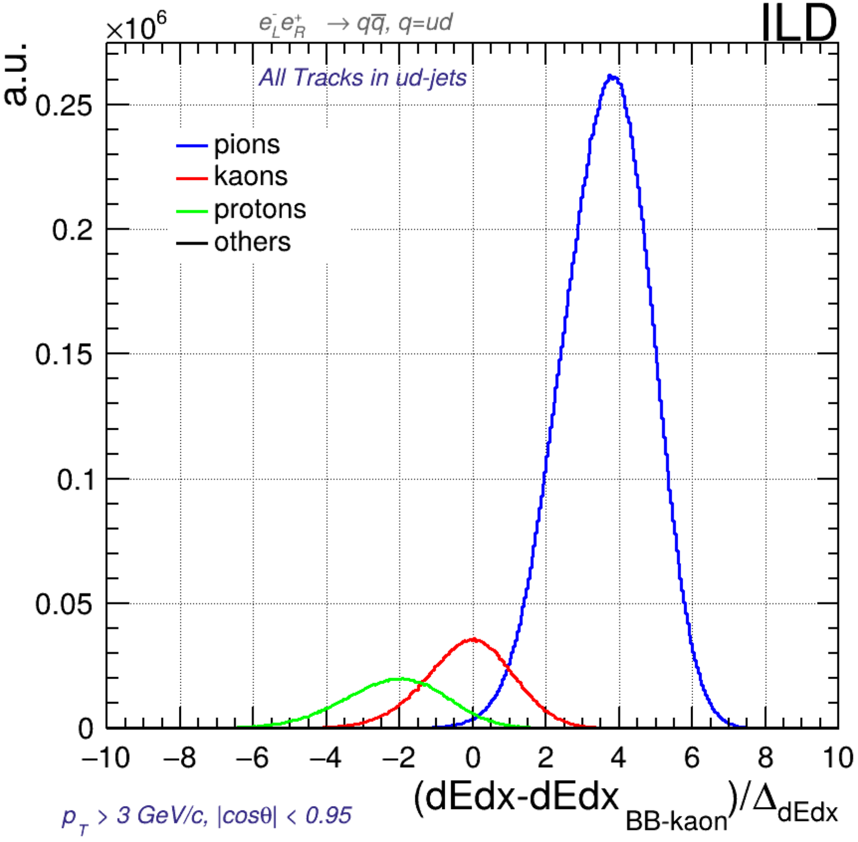
- These three bands can be used to measure an statistical distance

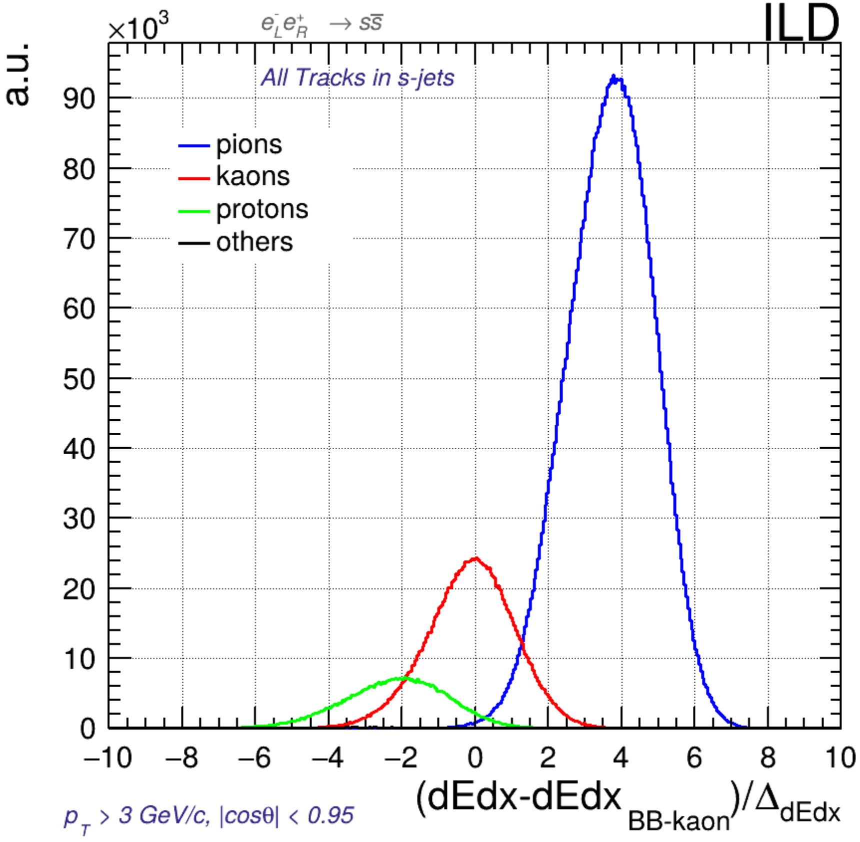




s vs ud: k-distance of tracks

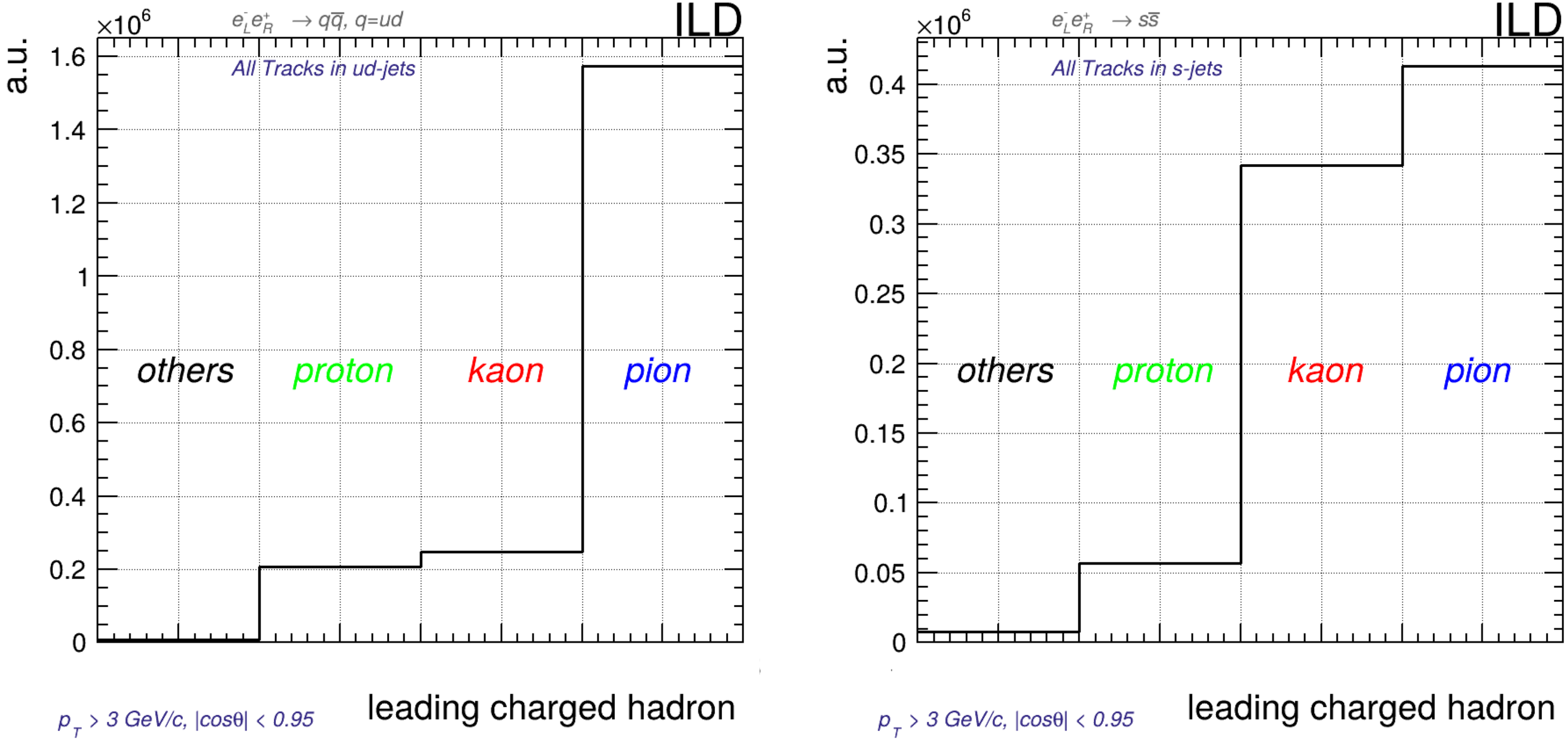
- Example of distance from tracks dE/dx and the theoretical values for kaons



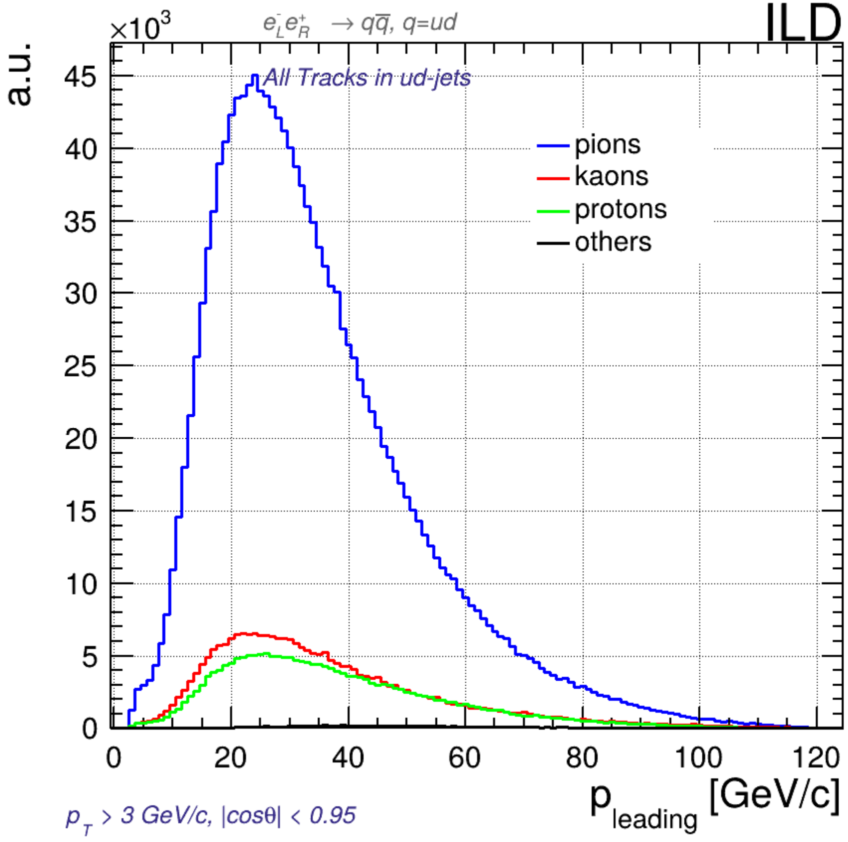


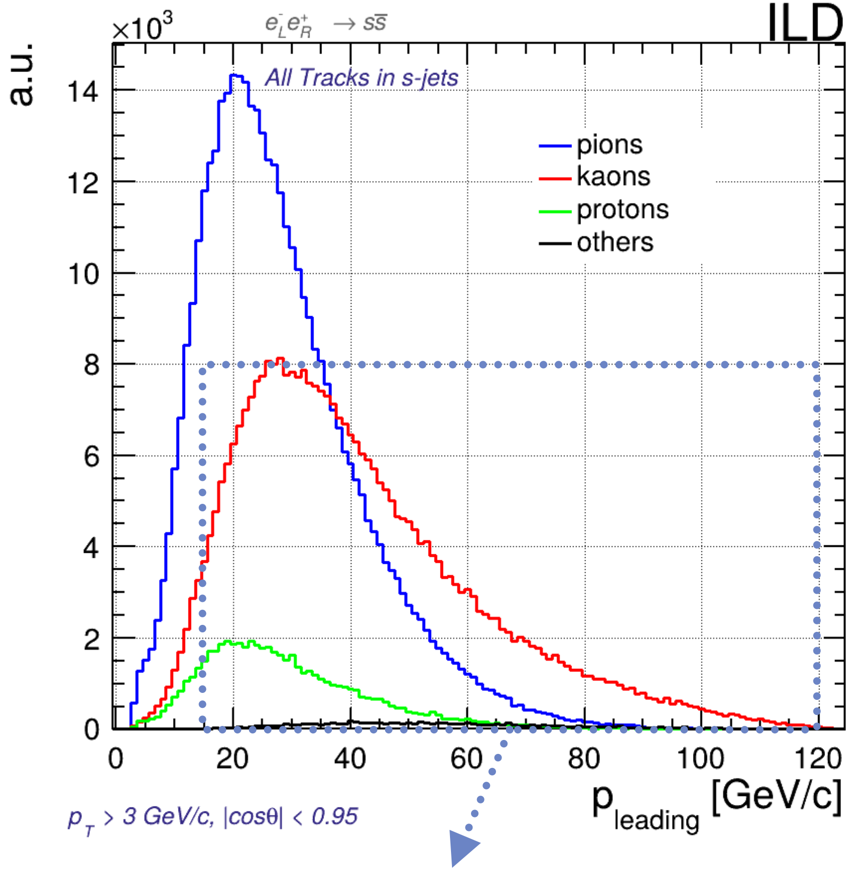
s vs ud: leading charged hadrons

- Different leading track population between s-jets and u/d-jets



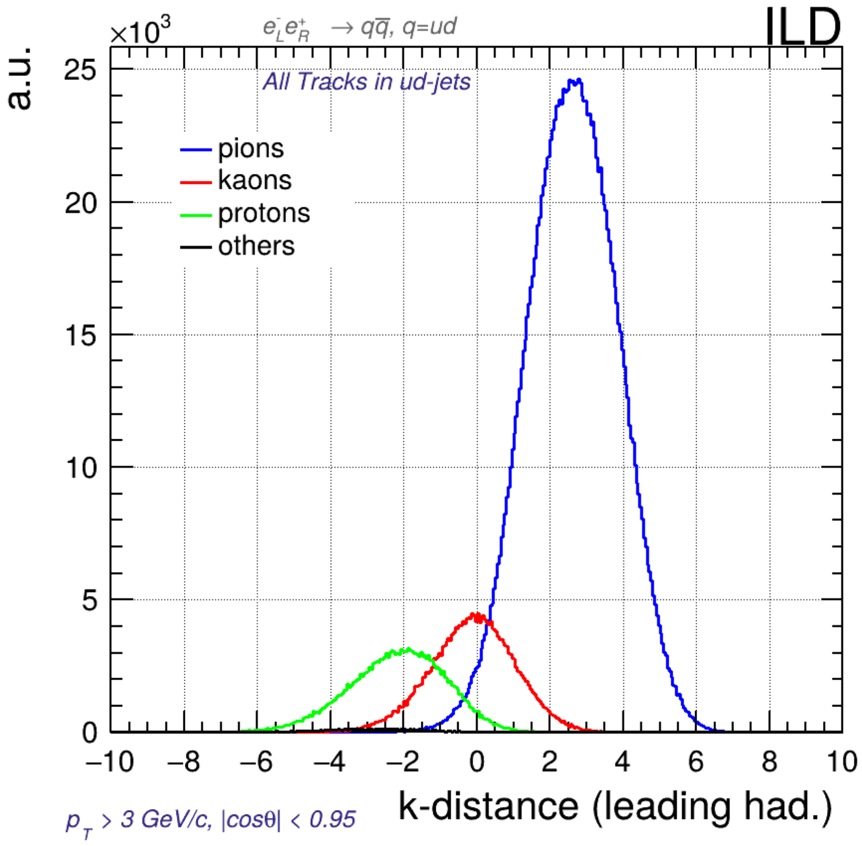
s vs ud: leading charged hadrons

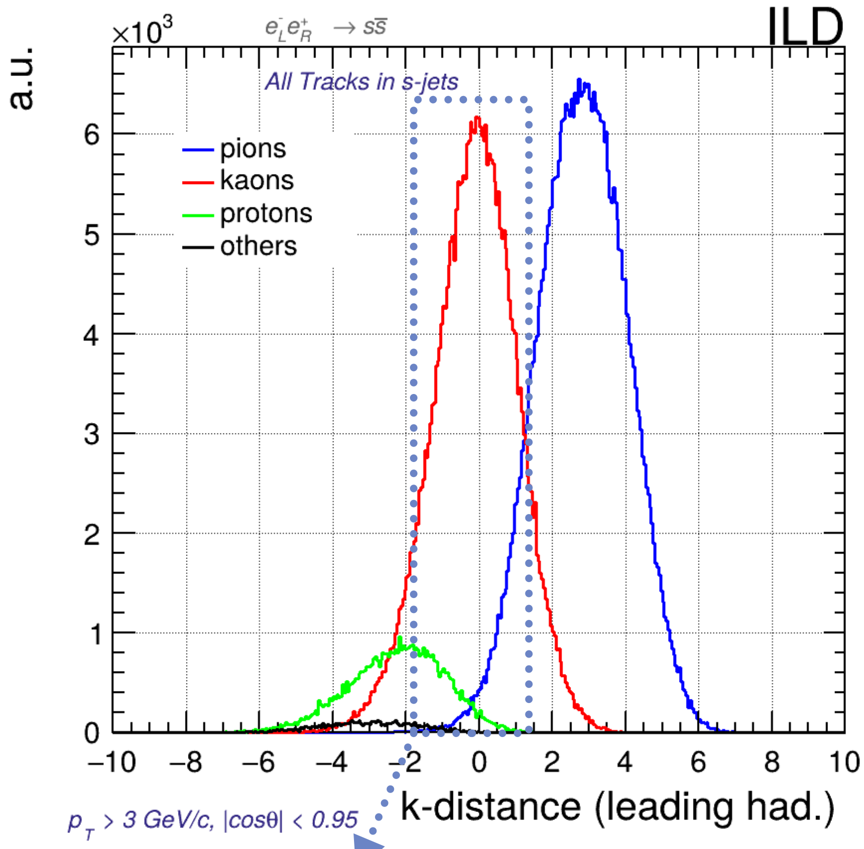




Our target for s-tagging!

s vs ud: k-dist of leading charged hadrons

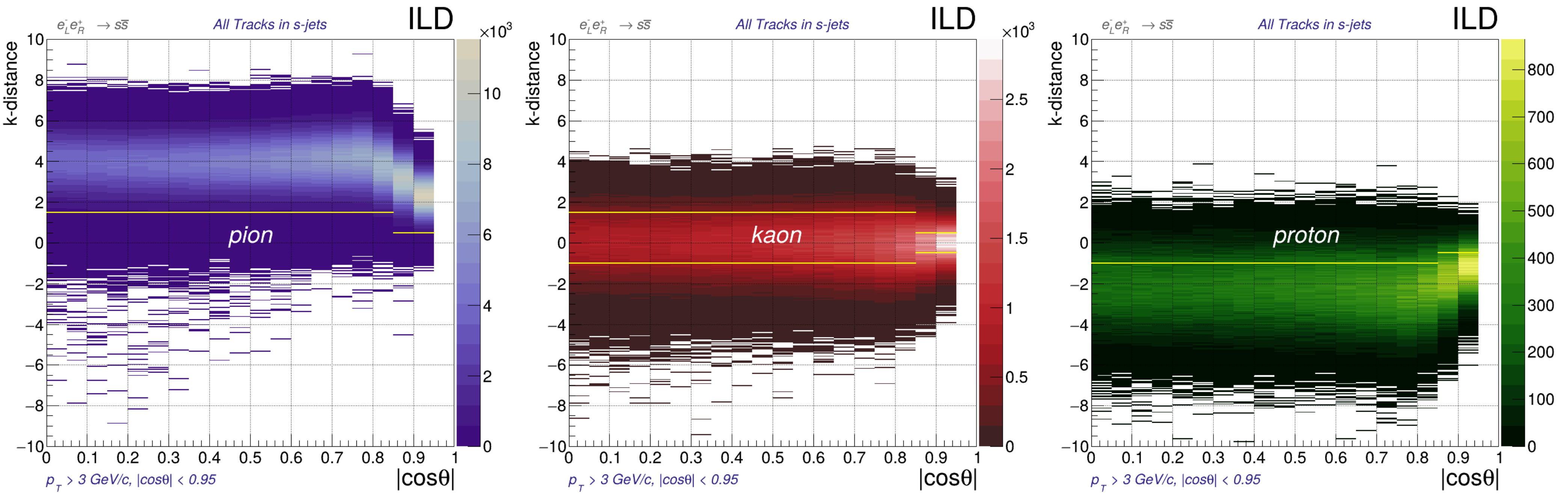




Our target for s-tagging!

2d view of k-distance (s quarks)

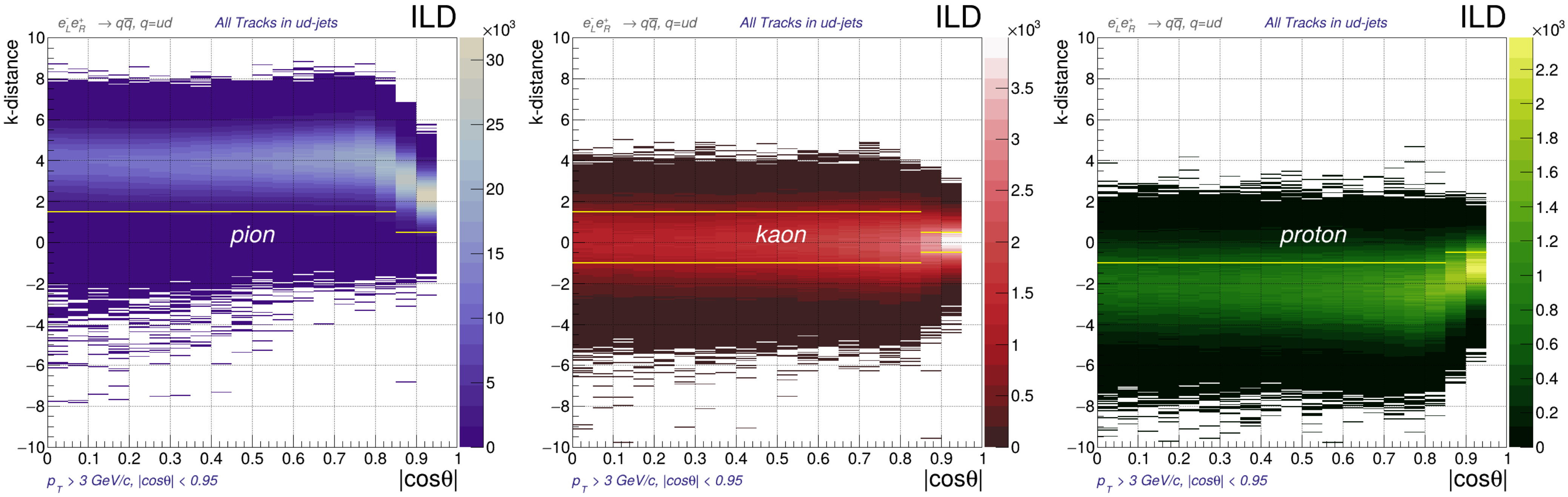
- Angular cuts are performed in these distributions for selection of pions/kaons



k-distance: statistical significance of the deviation between the measured dE/dx for a given track and the theoretical Bethe–Bloch value for kaons.

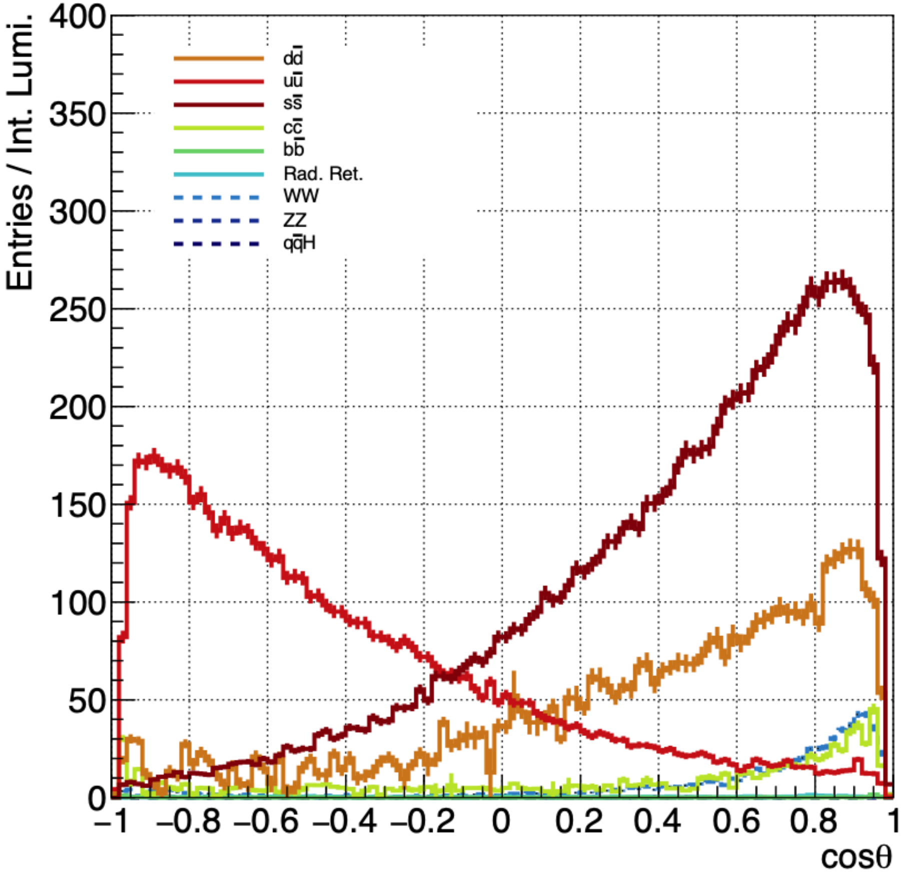
2d view of k-distance (ud quarks)

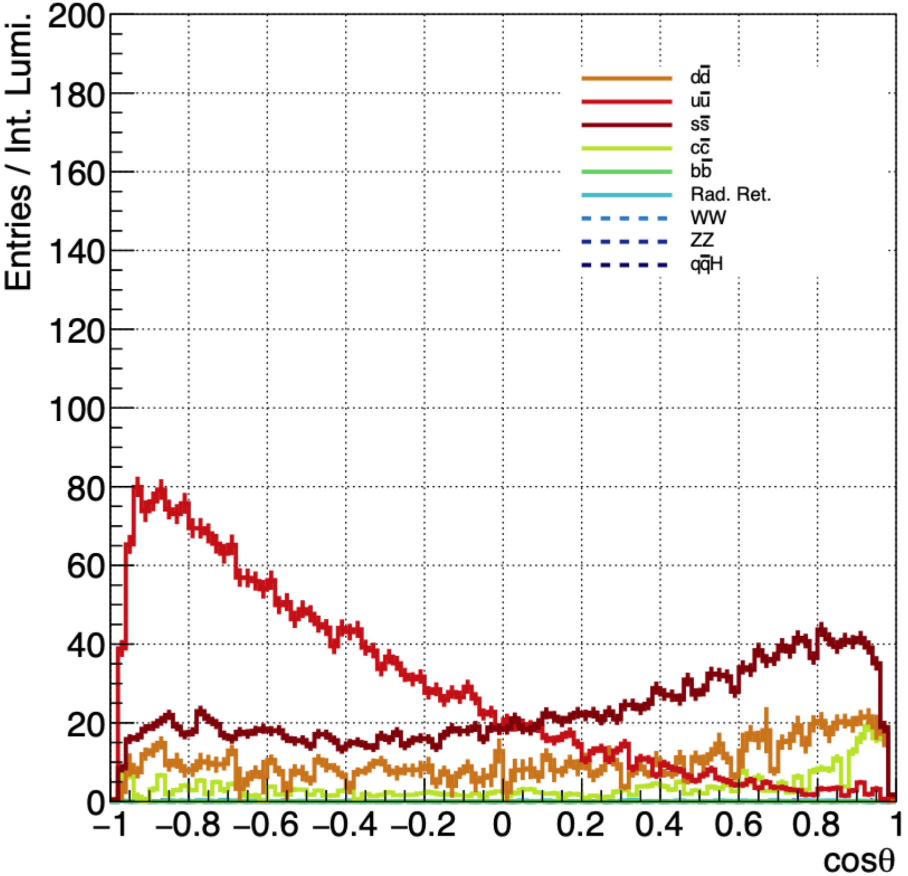
- Angular cuts are performed in these distributions for selection of pions



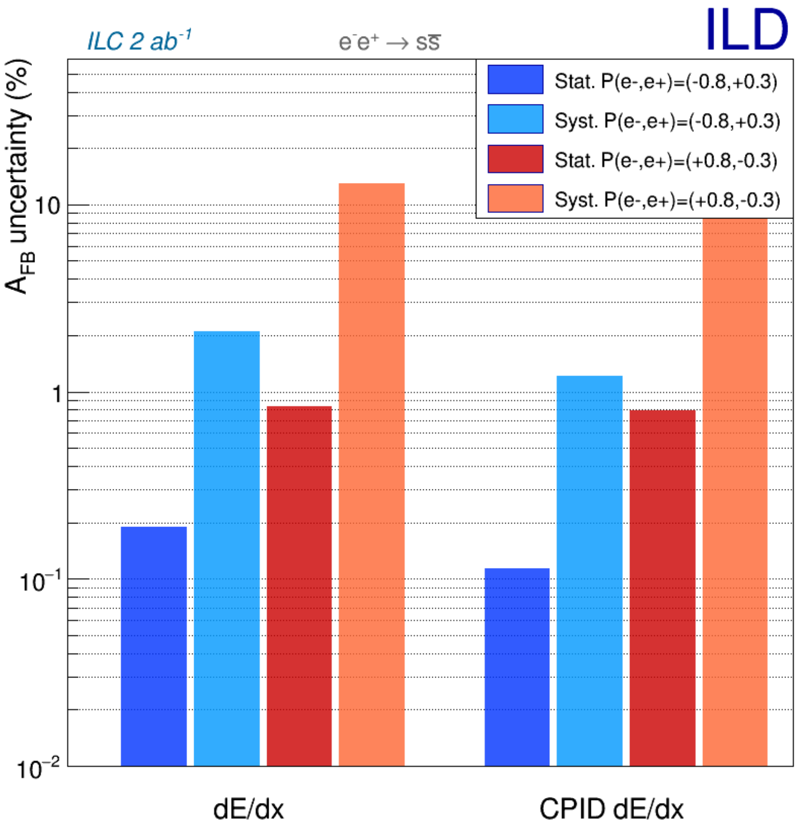
Contributions after preselection

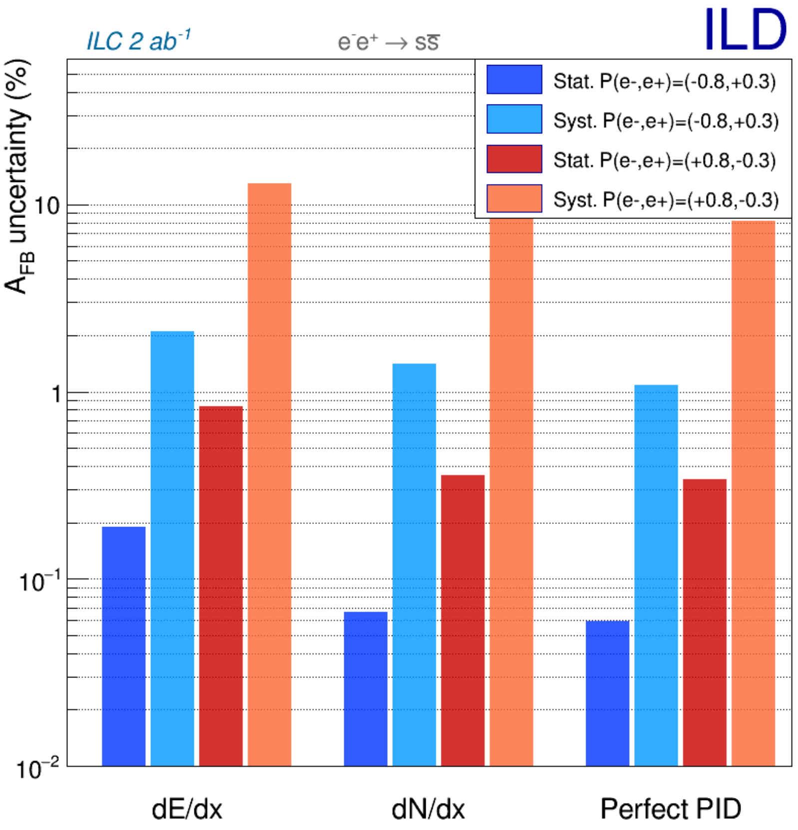
- After K LPFO selection (Plots from Yuichi's analysis)





Results (ILC250)





Background/signal and Migrations/signal ($e^-_R e^+_L$)

- I set different WP for each flavor (0.5 for s, 0.6 for b/c)