TYL FJPPN / FKPPN 2026

Hamamatsu

Download this slides

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

Jesús P. Márquez Hernández (IJCLAB, CNRS)

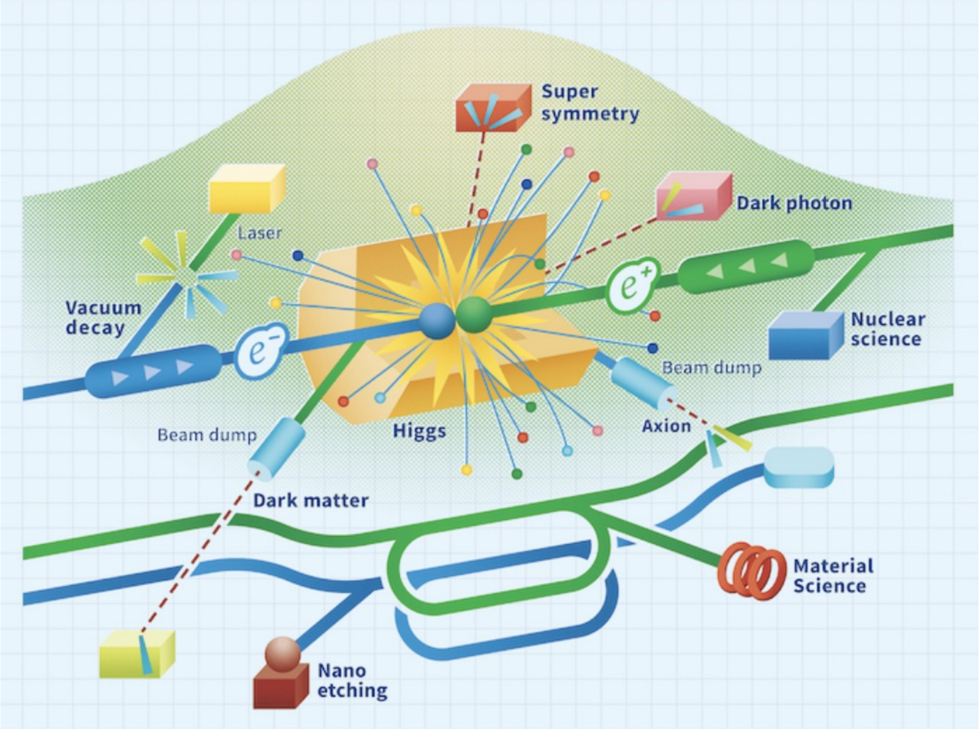
cnrs IJCLab Irène Joliot-Curie Laboratoire de Physique des 2 Infinis ILD CALO5D A new dimension in shower reconstruction

Outline

- **Context for future linear colliders**
  - International Linear Collider (ILC) → Linear Collider Facility at CERN (LCF, LCF@CERN)
  - International Large Detector (ILD)
- Diquark analyses ($q\bar{q} \leftrightarrow s\bar{s}$)
  - Cut-based analysis
  - New ParT preliminary results
- Application to phenomenology
  - GHU B-models

ILC → Linear Collider Facility at CERN

- Update of the International Linear Collider:
  - $P(e^-,e^+) = (\pm0.8, \pm0.3)$
    - Possibility of improve $P(e^+) = (\pm0.6)$
  - 2 IPs
  - $\sqrt{s}$ = (91.2), **250**, (350), 550 GeV



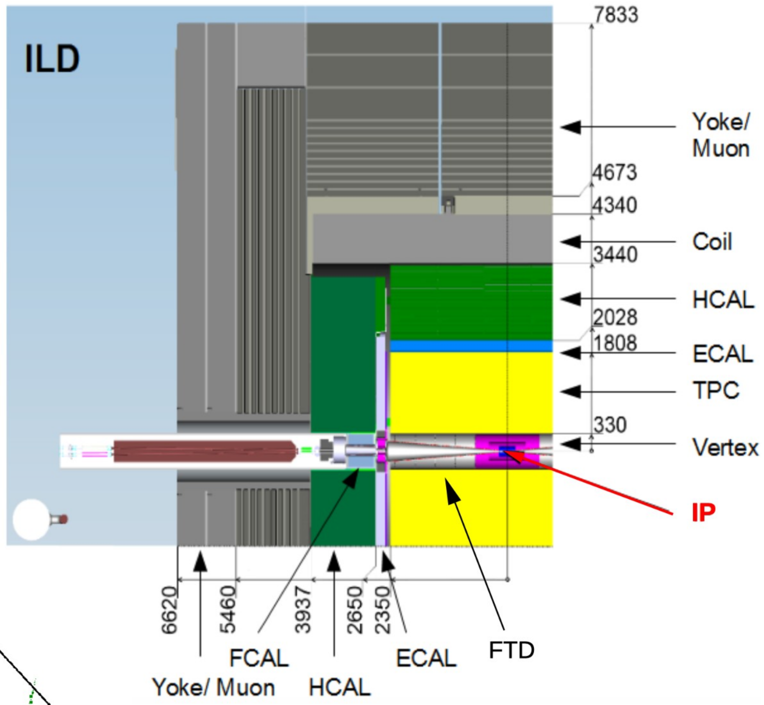
- Flexibility to adopt CLIC-like or C^3-like acceleration technology in the future
  - Reaching 1-3 TeV or even more!

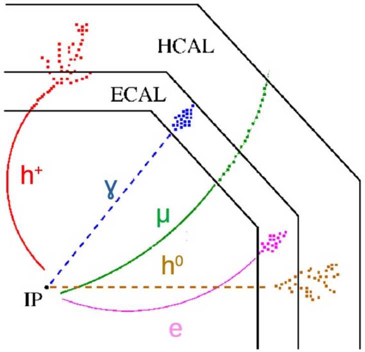
Polarization configurations used:
$P(e^-,e^+) = (-0.8, +0.3) \to e_L p_R$ or $e^-_L e^+_R$
$P(e^-,e^+) = (+0.8, -0.3) \to e_R p_L$ or or $e^-_R e^+_L$

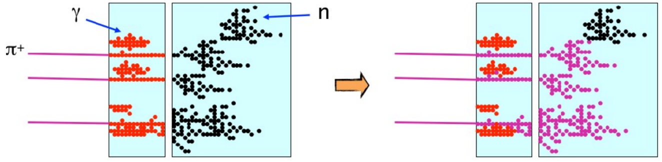
For a deeper introduction: Check this *introductory talk (by J. List)* or this Linear Collider Vision paper (2503.24049)

International Large Detector (ILD)

- Excellent tracking, vertexing, and IP constraining capabilities
- High granularity, compact and hermetic calorimetry system
- Full simulation available
- Optimised for **Particle Flow Algorithms**:
  - Determination of single particles
  - Powerful Particle identification (PID) tools
    - Jet energy measurement, flavour tagging, etc.







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

Outline

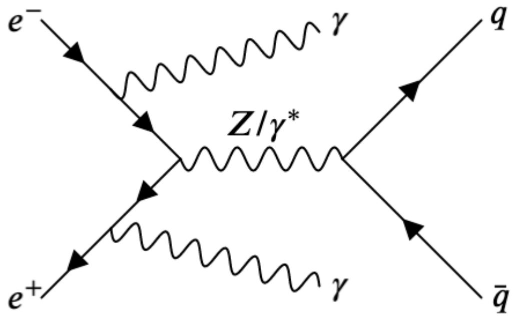
- Context for future linear colliders
  - International Linear Collider (ILC) → Linear Collider Facility at CERN (LCF, LCF@CERN)
  - International Large Detector (ILD)

- **Diquark analyses ($q\bar{q} \leftrightarrow s\bar{s}$)**
  - Cut-based analysis
  - New ParT preliminary results

- Application to phenomenology
  - GHU B-models

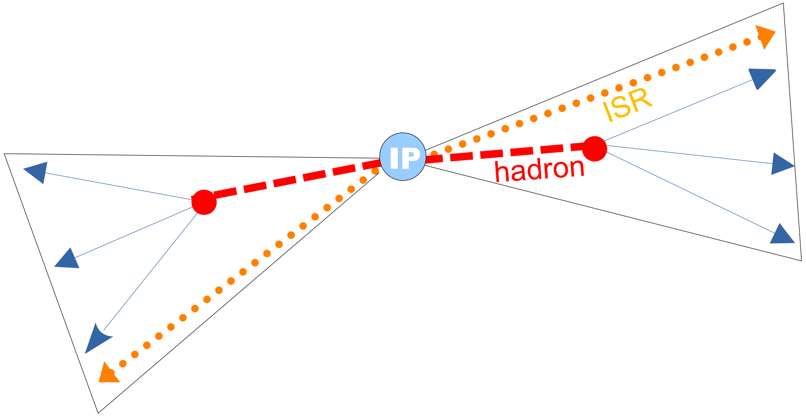
Diquark production in $e^- e^+$ collisions

- Topology: Two back-to-back jets
- MC simulations at 250 GeV
  - **ILC/LCF** run plan
  - **Full simulation** of the International Large Detector (ILD)
- Procedure:



R_q {
  - Background suppression → Selection of $q\bar{q}$ events
  - Flavor tagging → Selection of $b\bar{b}$, $c\bar{c}$, $s\bar{s}$ events
    - Double tagging (b-tag, c-tag, s-tag?)
}

A_{FB} {
  - Background suppression, Flavor tagging (as above)
  - Charge measurement → Quark-Antiquark identification
    - Double charge
}



At first we didn't have a s-tagging so we had to build our own

Redoing of the ssbarAnalysis

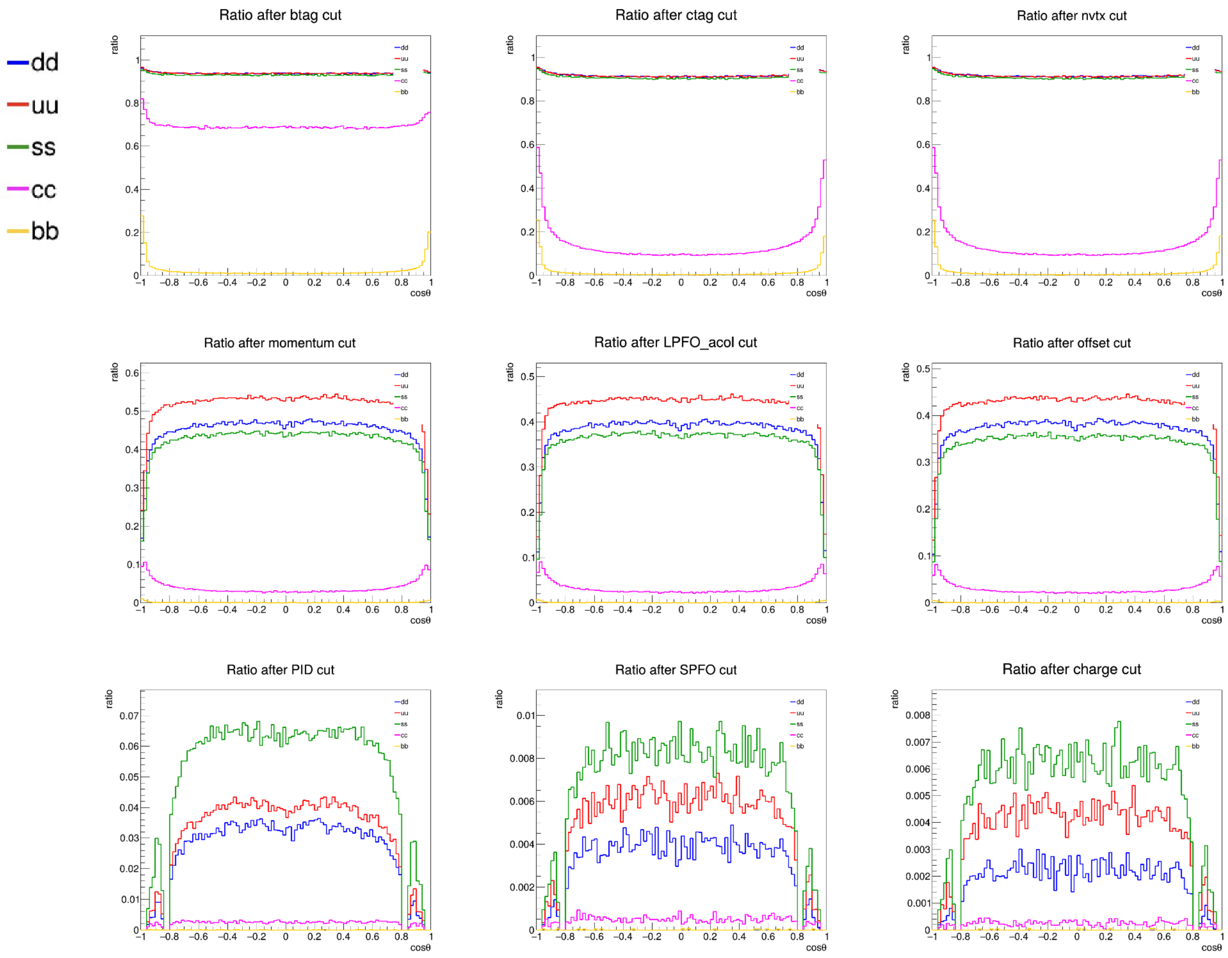
- Cut-based s-tagging (Modification of Y. Okugawa's analysis)
  - After the $q\bar{q}$ selection

	#	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 attenuation	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	$Q_{LPFO1} \times Q_{LPFO2} < 0$	Charge of LPFOs from both sides has opposite charge.

Cuts visualization (K selection for s-jets)

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

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

We integrate the angular fit!

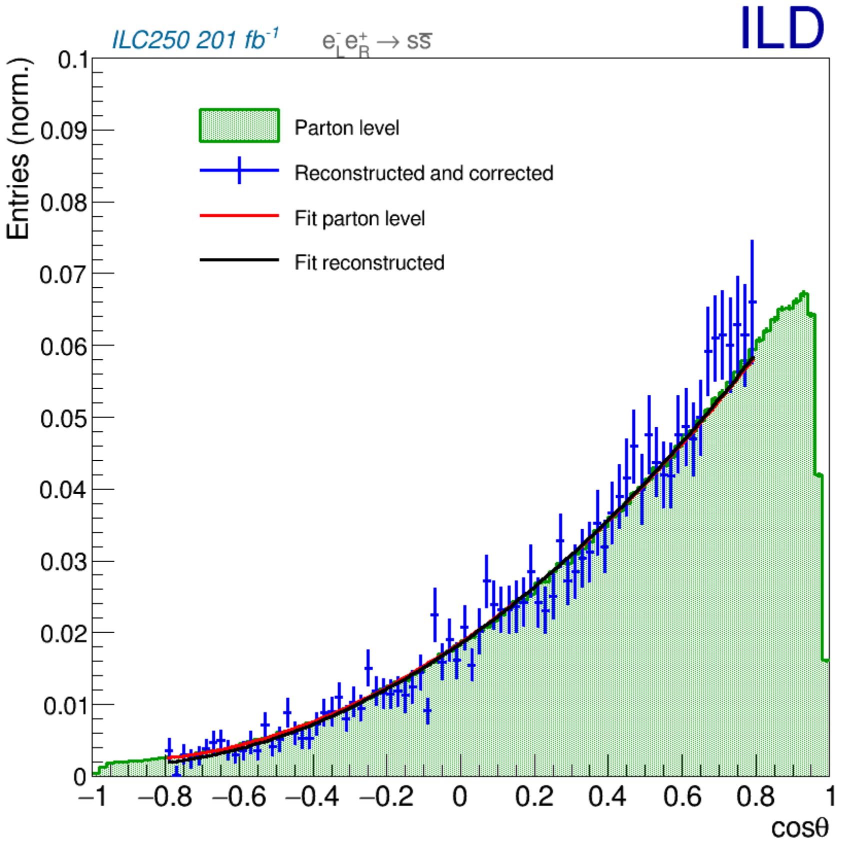
- A fit is performed to the corrected signal:

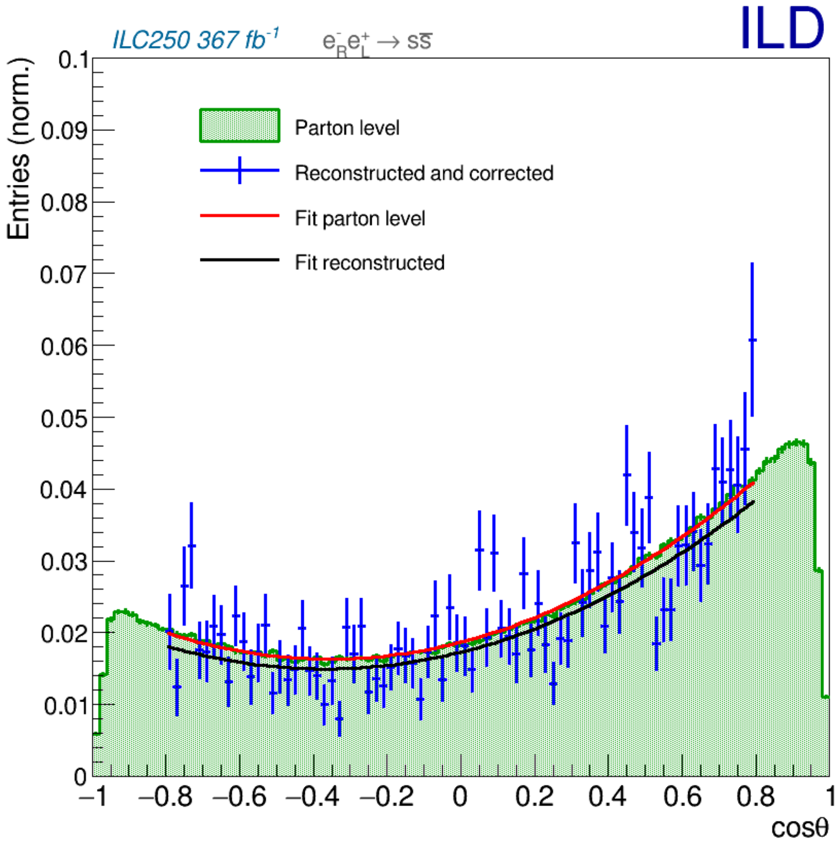
$$\frac{d\sigma}{d\cos\theta} = S\left(1+\cos^2\theta\right) + A\cos\theta$$

- 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

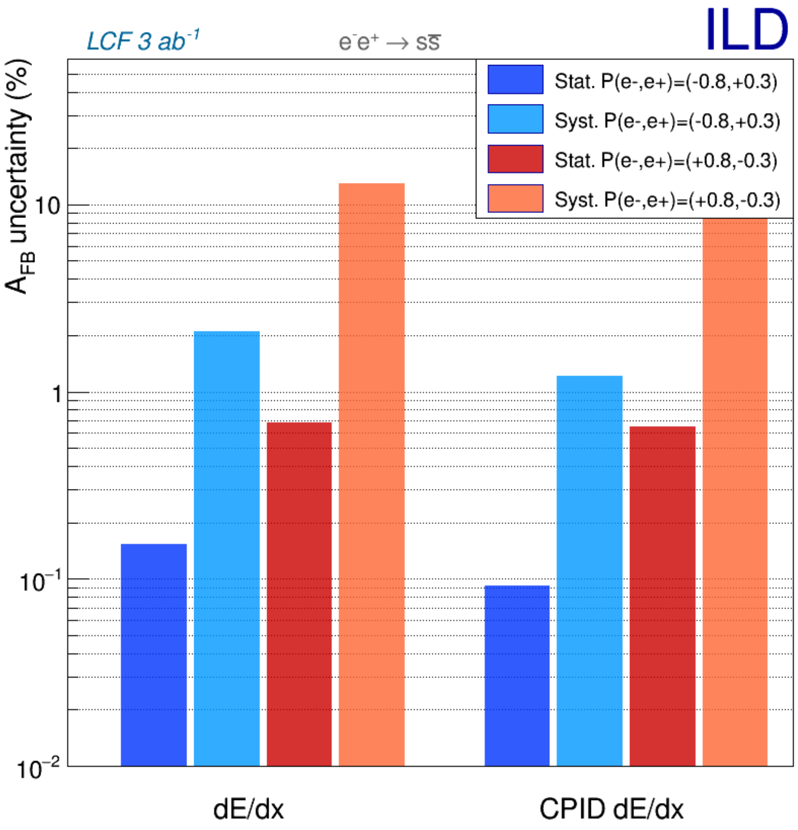
- Fit constrained to |cosθ| < 0.8 shows good agreement

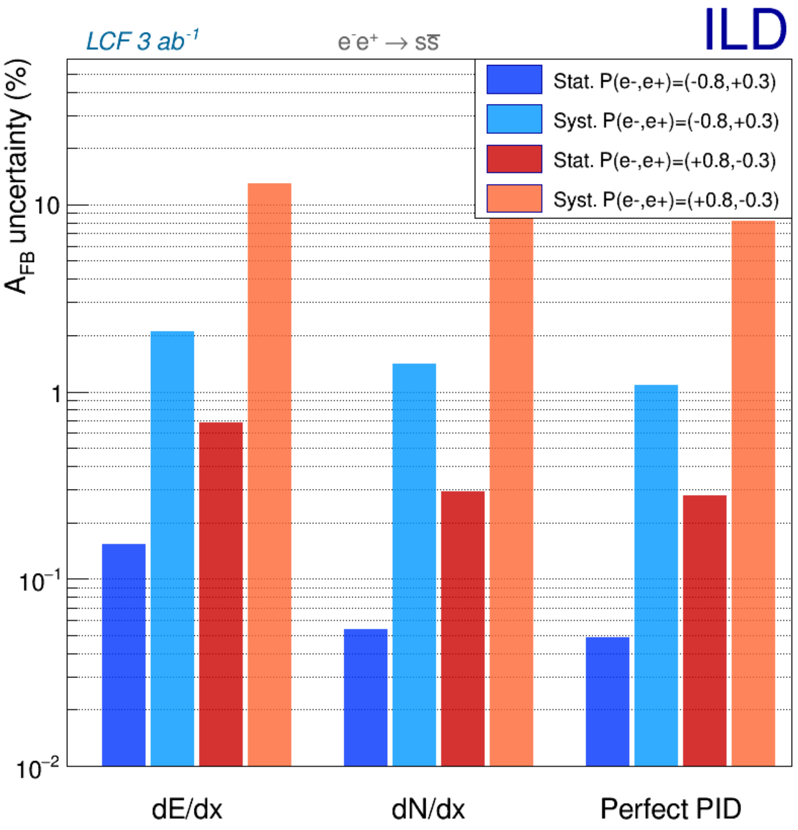




Results (LCF250)

- Prospects for software (left) & hardware (right) improvements within the same analysis



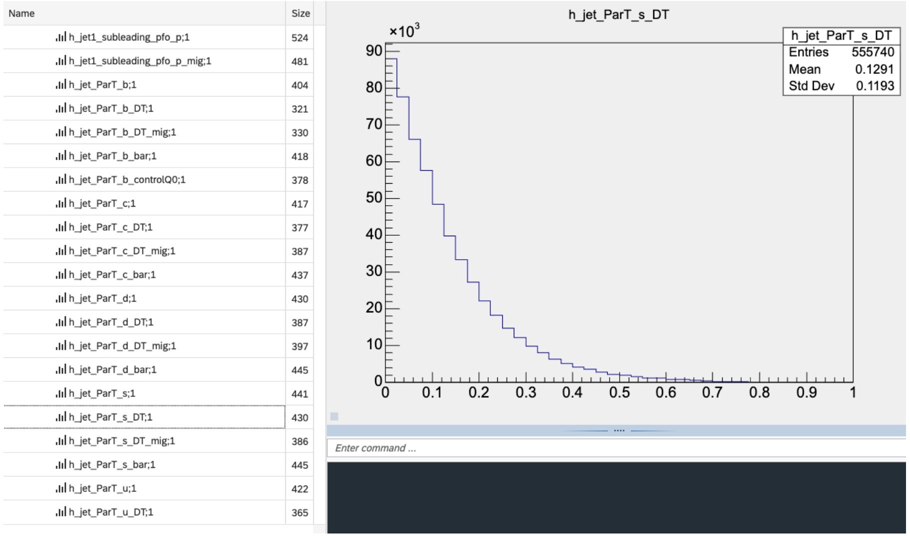


Selection using Particle Transformer (ParT) tagging

- 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 $\overline{qq}$)

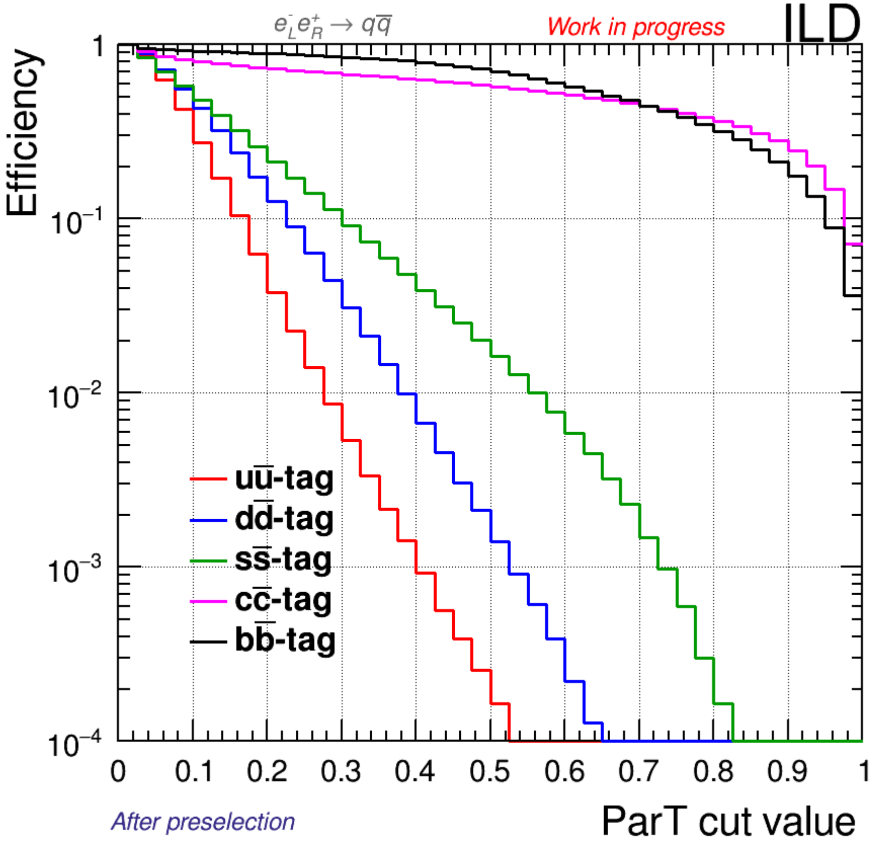
- Now the Double Tag and Double Charge happens at once
  - No PID corrections
  - No p-q method
  - Only an efficiency correction is applied to the data!

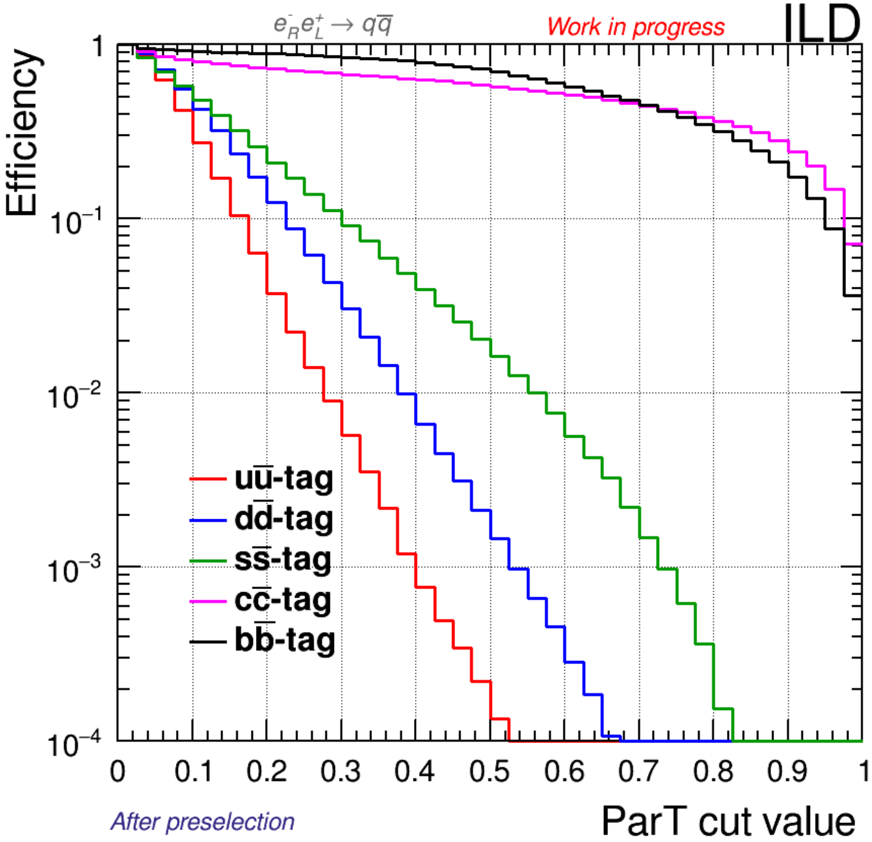


Tag with 11 categories (q, $\bar{q}$, g)

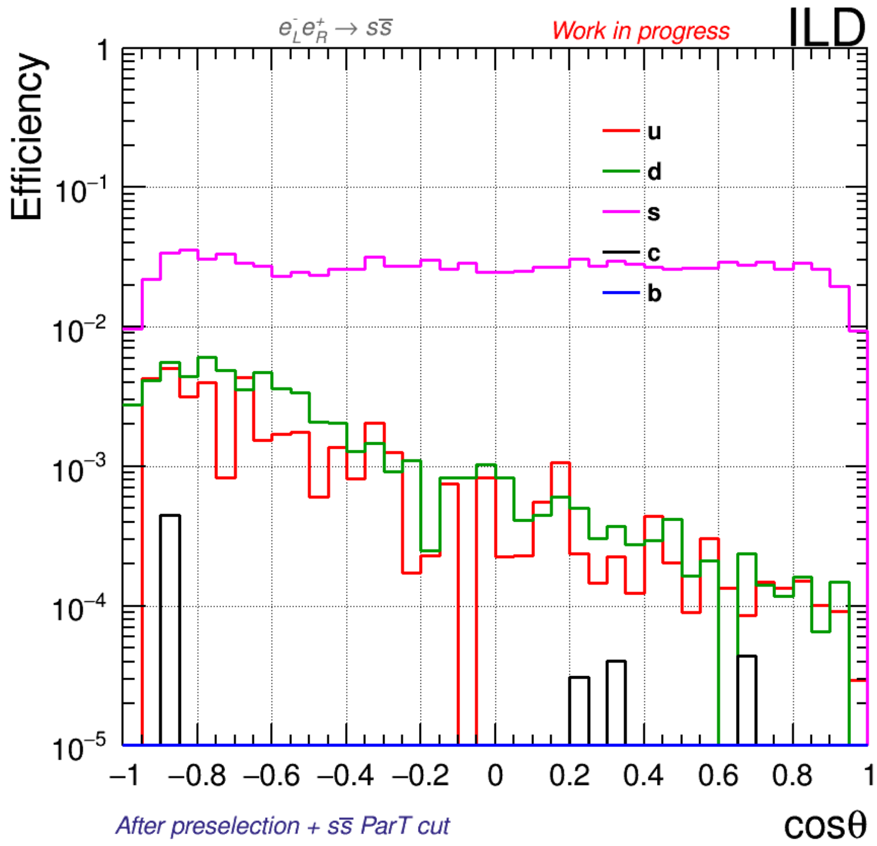
Signal efficiency for different ParT tagging cuts

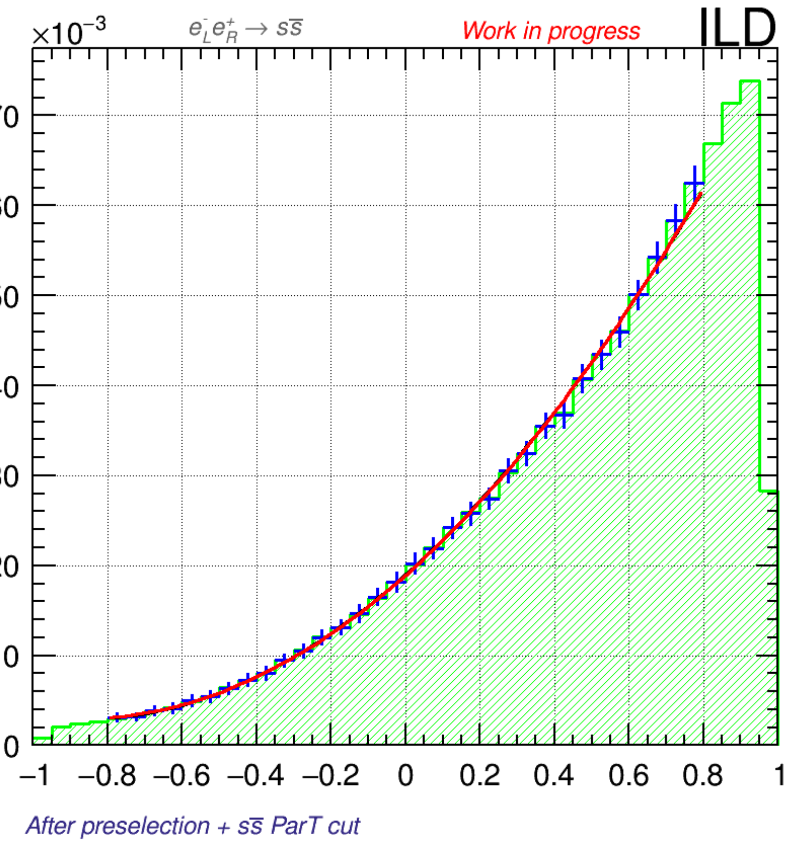
- I set different WP for each flavor (0.4 for s, 0.6 for b/c)



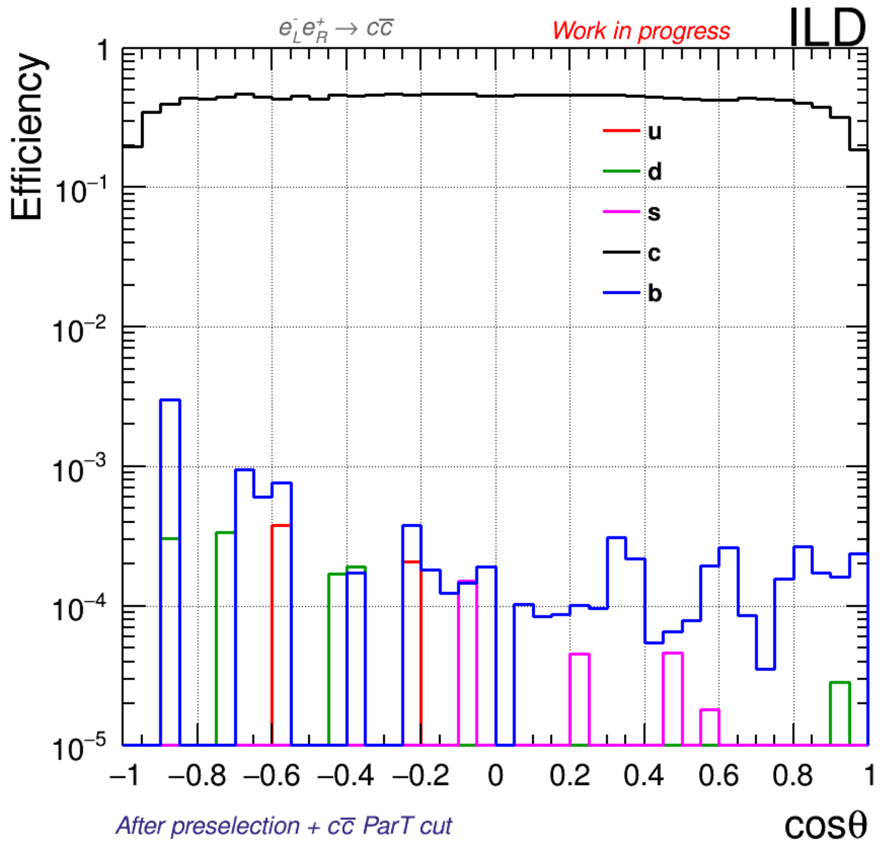


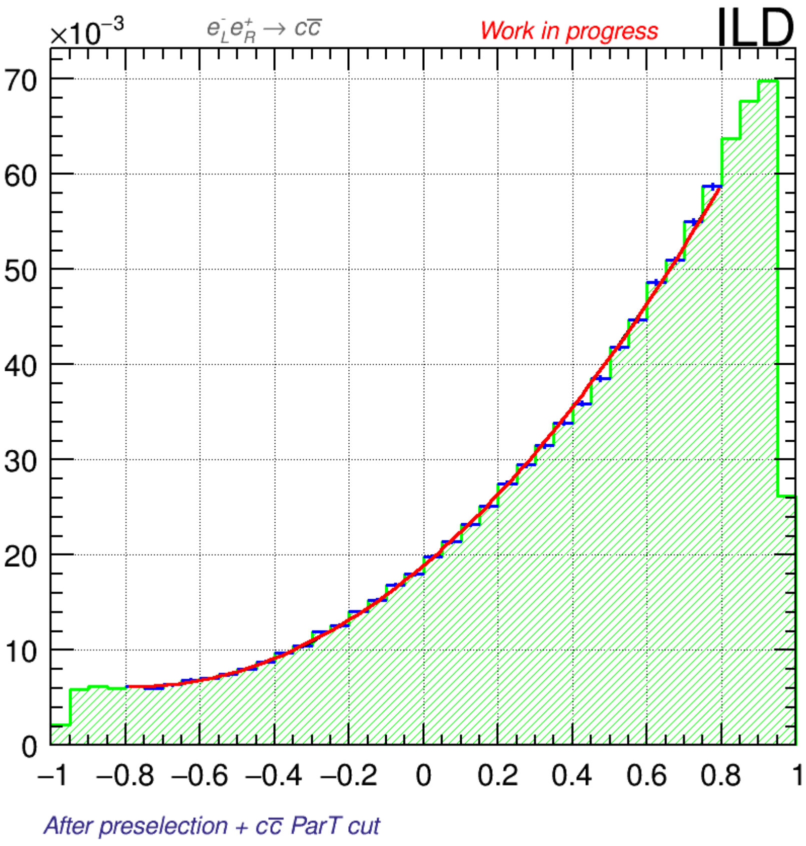
Selection and fit for $s\bar{s}$



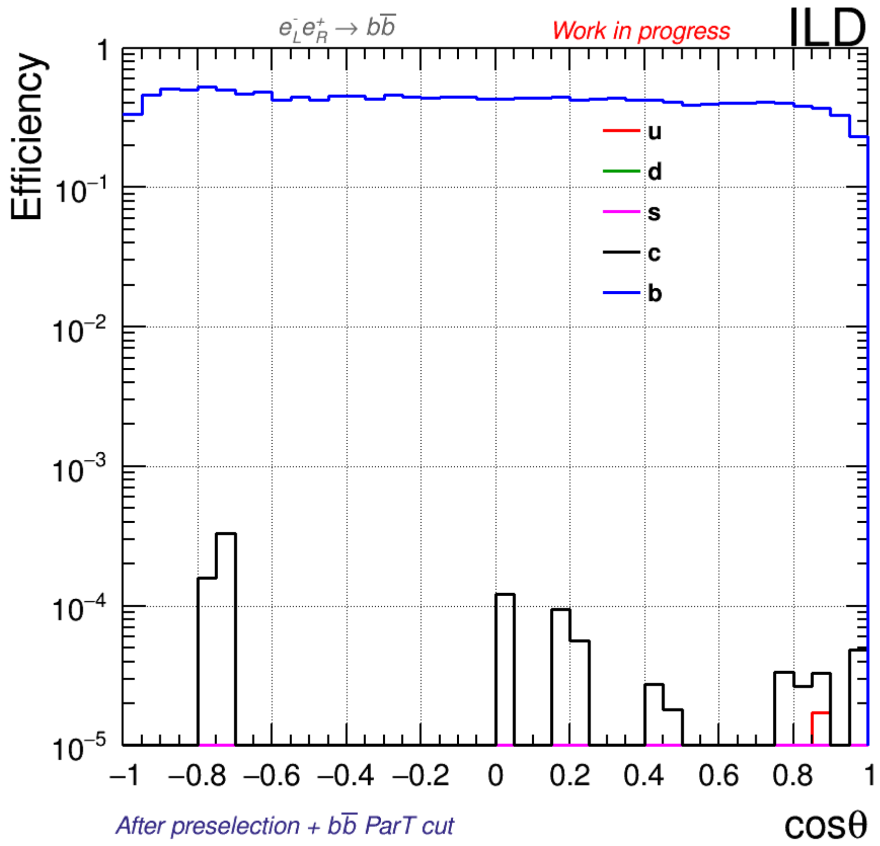


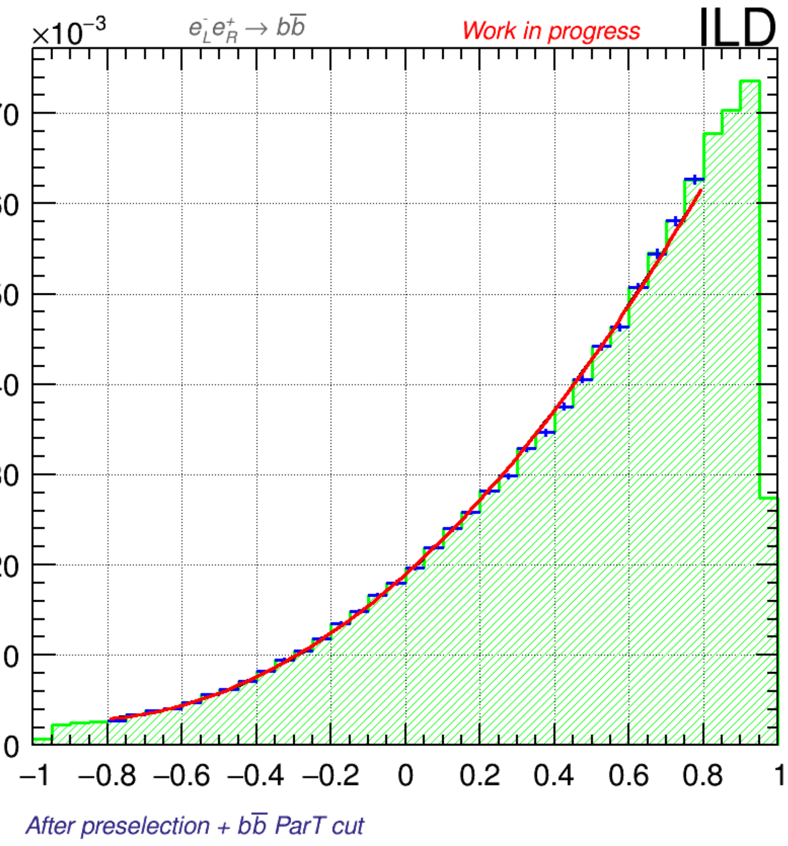
Selection and fit for $c\bar{c}$





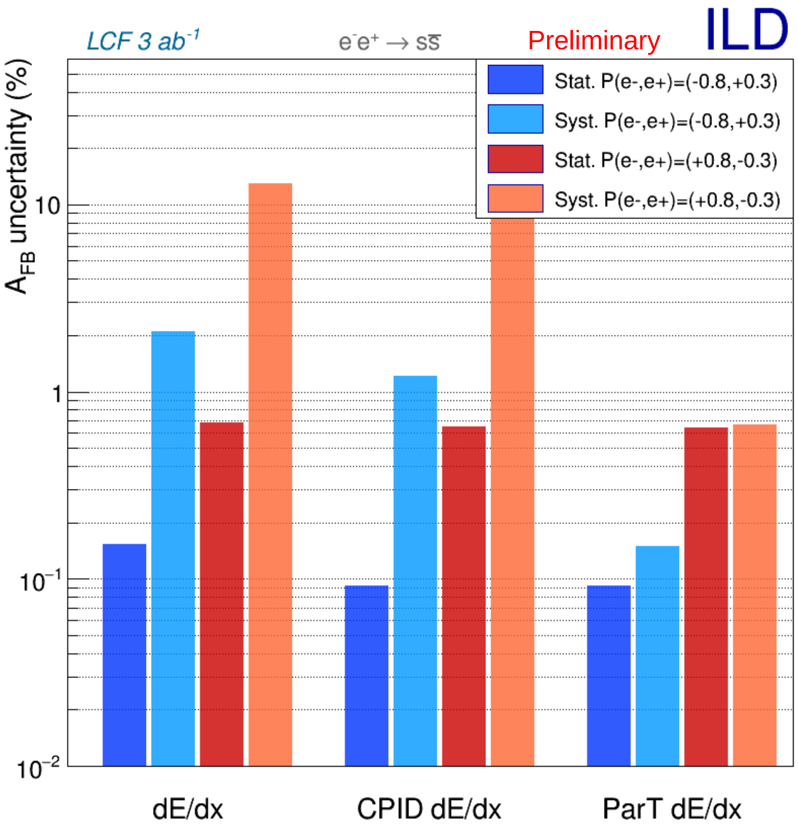
Selection and fit for $b\bar{b}$





Results for s-quark (ILC250 & LCF250)

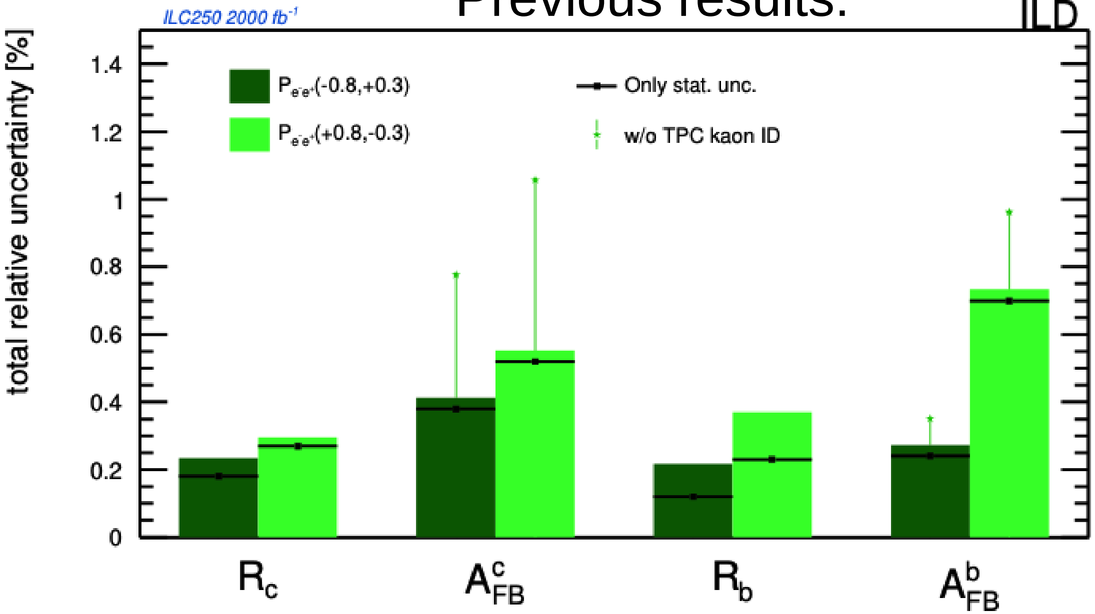
- O(10) improvement of systematic uncertainties!



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	$A^{c\bar{c}}_{FB}$	R_c	$A^{c\bar{c}}_{FB}$	R_b	$A^{b\bar{b}}_{FB}$	R_b	$A^{b\bar{b}}_{FB}$
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 %

Outline

- Context for future linear colliders
  - International Linear Collider (ILC) → Linear Collider Facility at CERN (LCF, LCF@CERN)
  - International Large Detector (ILD)

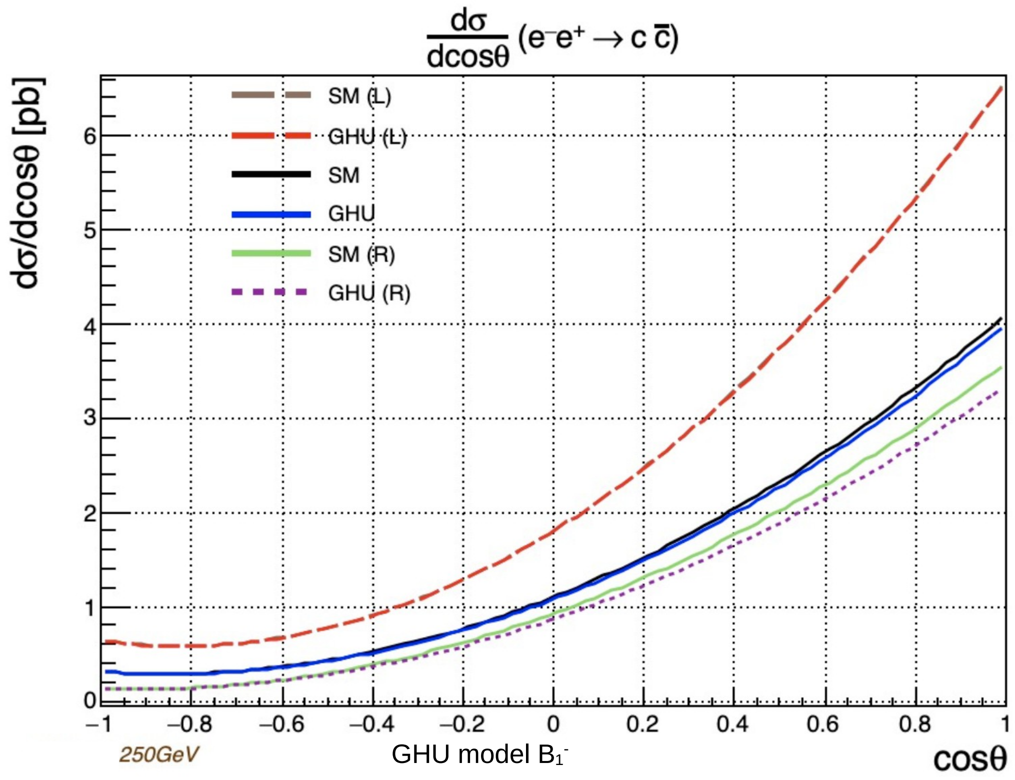
- Diquark analyses ($q\bar{q} \leftrightarrow s\bar{s}$)
  - Cut-based analysis
  - New ParT preliminary results

- **Application to phenomenology**
  - GHU B-models

Gauge-Higgs Unification Models (GHU)

- Randall-Sundrum metric (5D)

- The symmetry breaking pattern is different than in the SM and features the Hosotani mechanism:
  - Masses are generated dynamically from the extra-dimension properties
  - Hosotani's angle θ_H determines the projection of the 5D fields, fixing most physical effects:
    - Resonances of Z/γ with $m_{kk} \sim O(10)$ TeV.
    - Modifications and new EW couplings/helicity amplitudes.
      - ALL fermions are affected!



Deviations already visible at 250 GeV!

GHU Phenomenology

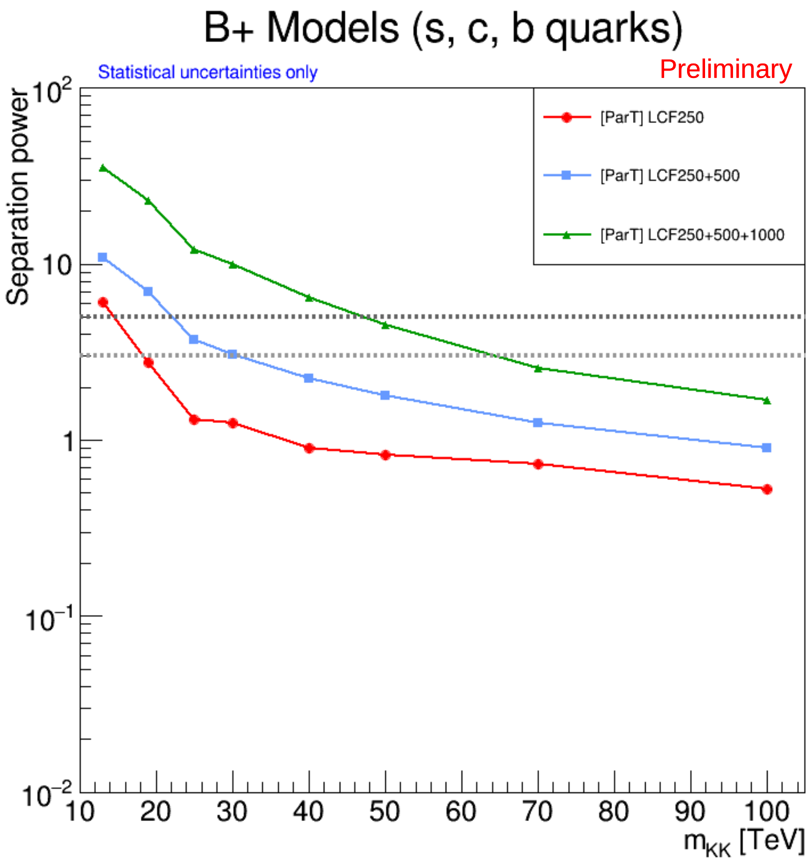
- Two types of models (B-, B+)
  - The ± sign comes from the fermionic term of the lagrangian
    - Opposite angular shape/deviation from SM

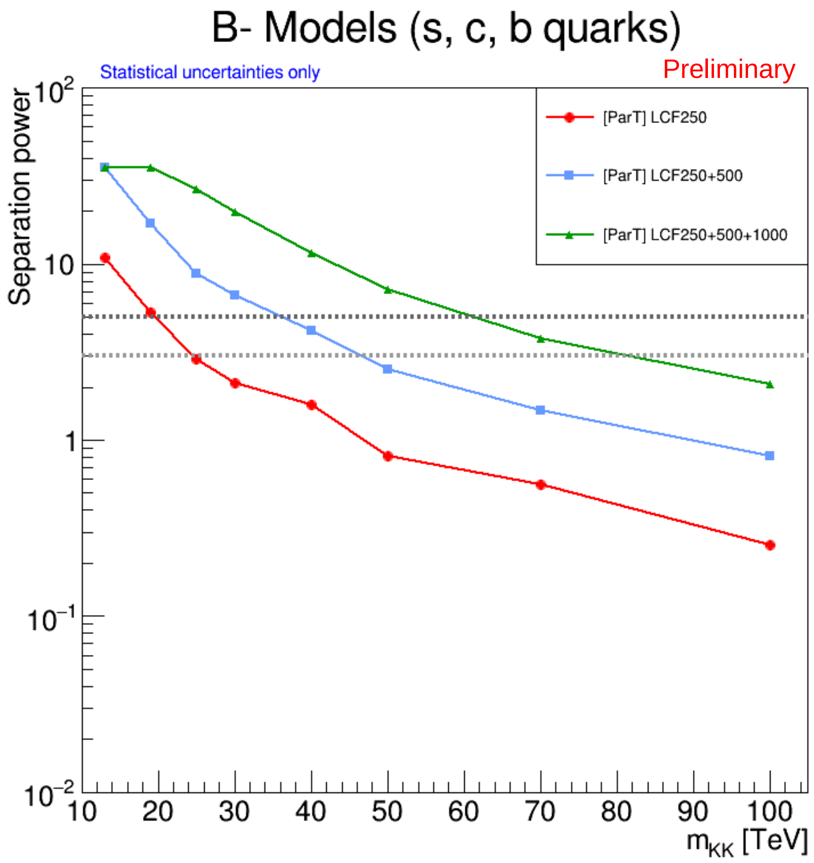
- As **Benchmark**, we use the [Funatsu, Hatanaka, Hosotani, Orikasa, Yamatsu] models
  - Best Indirect limits expected for LHC (HL-LHC) ~18 (22) TeV (2111.05624)
  - Naoki Yamatsu provided us with extra models (not yet explored in the bibliography)
    - Masses of KK resonances tested: 13, 19, 25, 30, 40, 50, 70, 100 TeV

- We test the models using a multivariate gaussian
  - We plot results as the total combination of channels available in each case
    - Different $\sqrt{s}$, polarization, and quark flavours are combined
    - The number plotted is the separation power when comparing each model wrt the SM

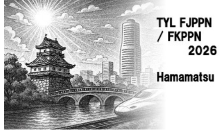
Prospects for GHU phenomenology (LCF@CERN)

- Extrapolating statistical uncertainties from 250 to 500 & 1000 GeV





Conclusions & next steps



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

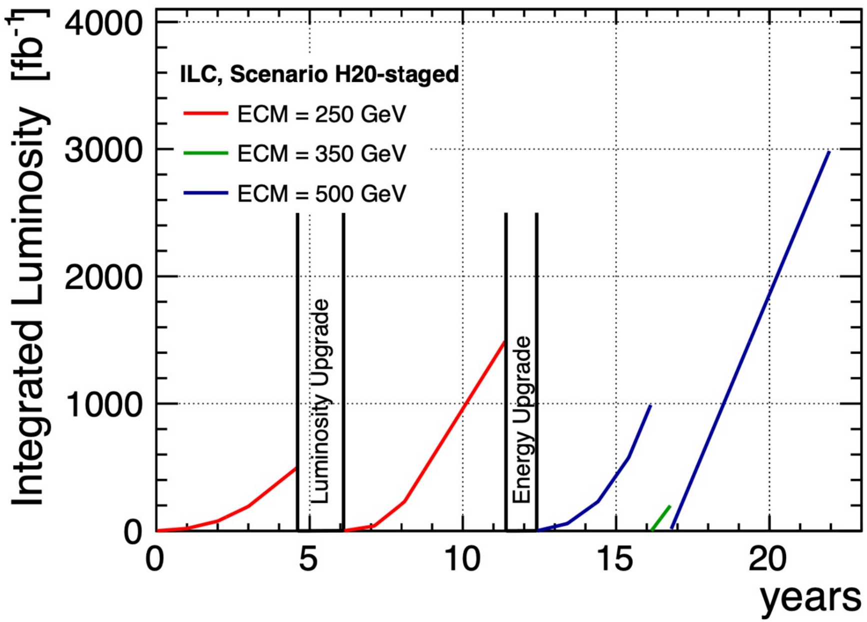
- There’s still work to be done within this analysis:
  - Optimize WP
  - Check migrations, preselection,…, revision of systematics (WIP)
  - Add other leptonic channels

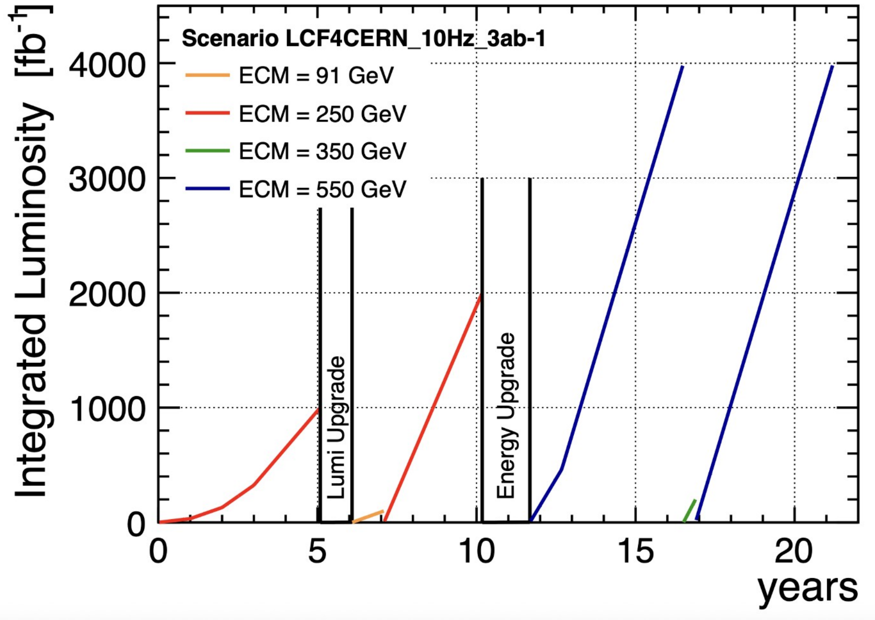
- Future plans:
  - Extension to 500-550 GeV analysis
  - Calculating final GHU phenomenology prospects
    - Considering systematics may be mandatory now

THANKS FOR YOUR ATTENTION

BACK-UP

ILC H20-staged & LCF4CERN run plans



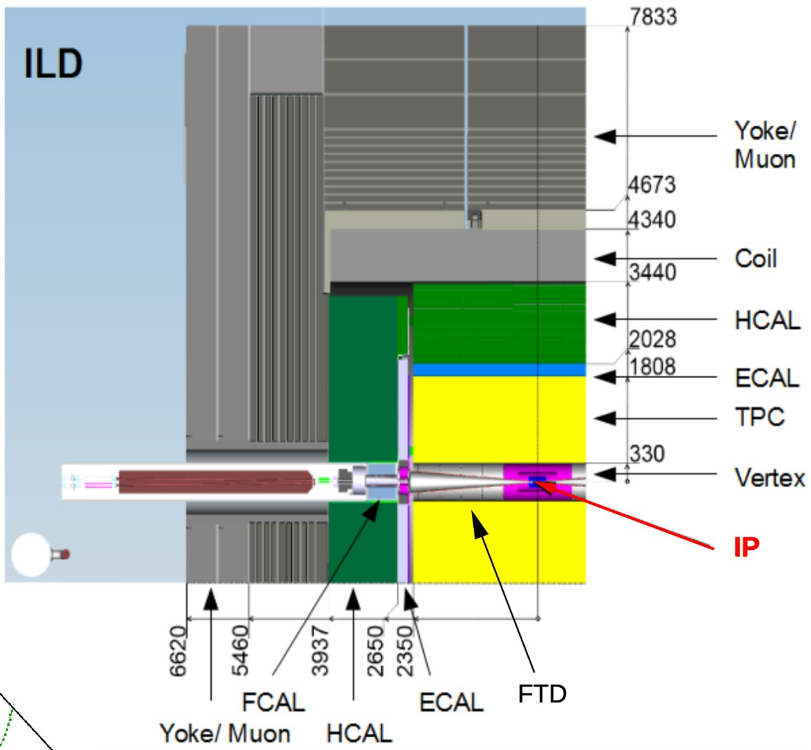


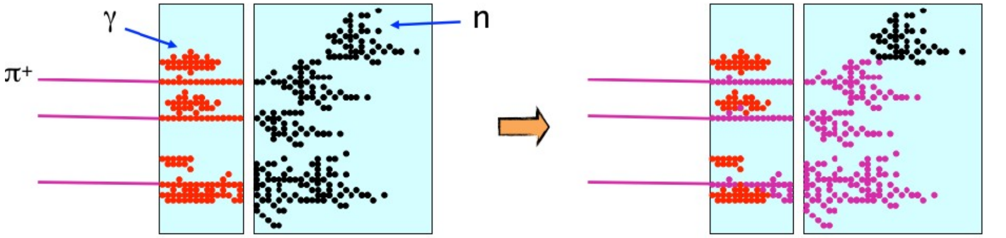
	91 GeV	250 GeV	350 GeV	500 GeV	1000 GeV
$\int \mathcal{L}$ (ab^{-1})	0.1	2	0.2	4	8
duration (yr)	1.5	11	0.75	9	10
beam polarization (e^-/e^+; %)	80/30	80/30	80/30	80/30	80/20
(–, -+, +-, ++) (%)	(10,40,40,10)	(5,45,45,5)	(5,68,22,5)	(10,40,40,10)	(10,40,40,10)
δ_{ISR} (%)	10.8	11.7	12.0	12.4	13.0
δ_{BS}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

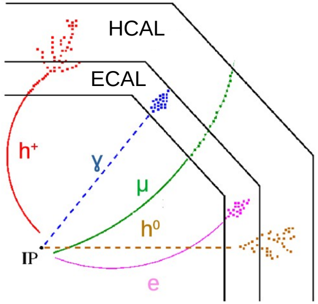
- ►It features excellent tracking, vertexing and IP constraining capabilities with minimal material budget
- ►**High granularity, compact** and **hermetic calorimetry system**
- ►**Full simulation available**: detailed geometry, materials, reconstruction chain, etc.
- ►**Optimized for Particle Flow**: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





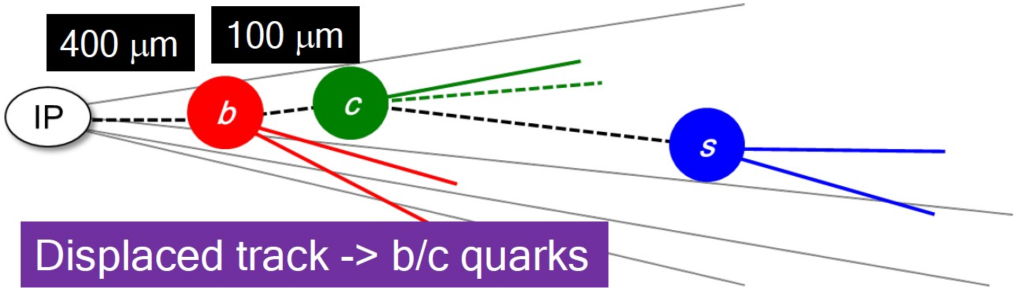
$E_{JET} = E_{ECAL} + E_{HCAL}$

$E_{JET} = E_{TRACK} + E_{\gamma} + E_{n}$



From b/c to strange quark

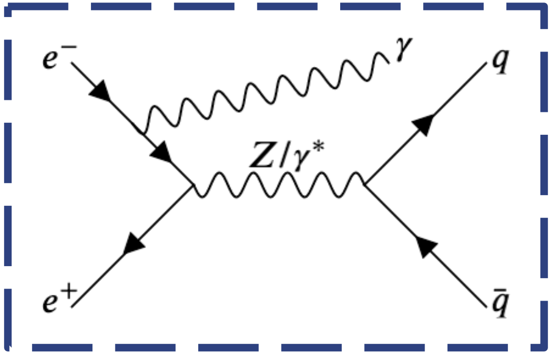
- Flavor tagging of b and c jets is “easy”:
  - Decay of b/c hadrons: displaced vertexes at a distance ($\tau_q \cdot c$) from de IP

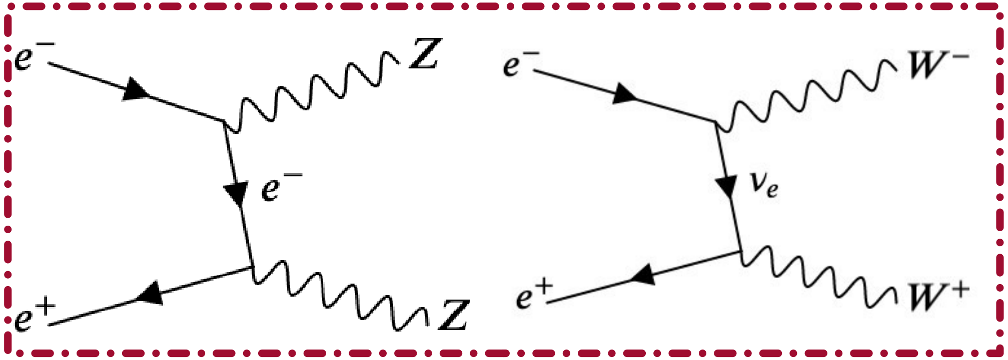


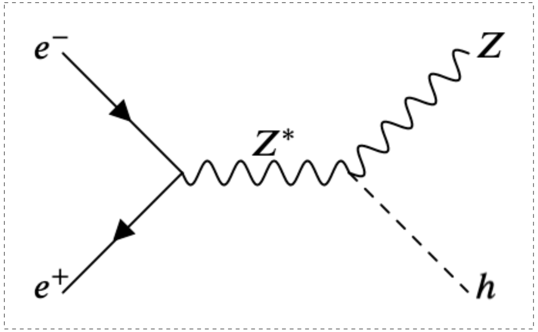
- Strange quark mostly produce kaons… no decays in the tracker to be used!
  - We need to build/use an s-tag **relying on kaon PID**
    - Our first attempt is a “classic” cut-based analysis
      - Previous analysis done by Y. Okugawa in his thesis, directed by R. Poeschl

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

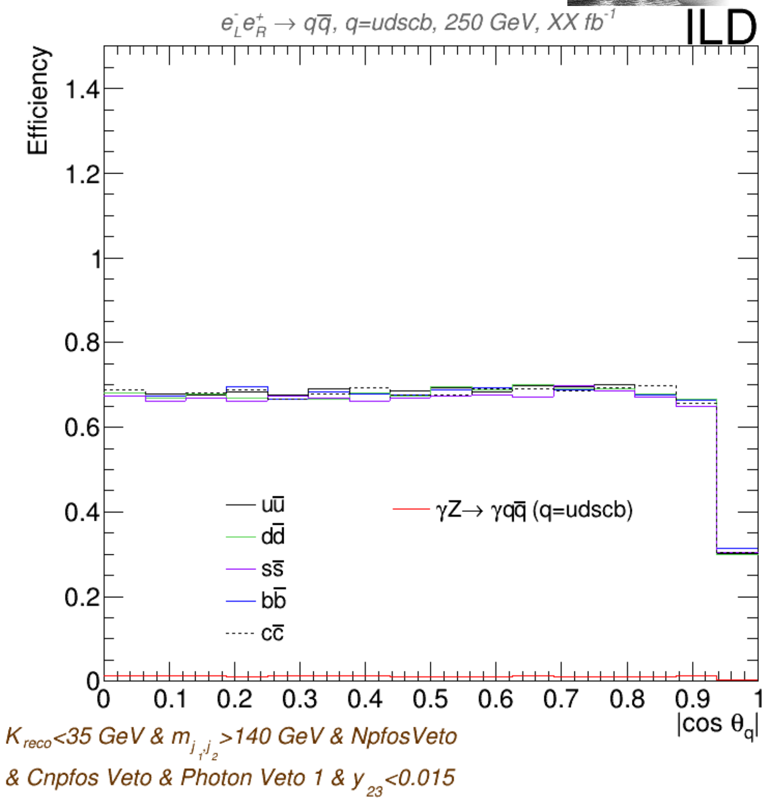
Cuts:
- K_{reco} < 35 GeV
- m_{2jets} > 140 GeV
- Charged N pfos
- Photon veto
- Y_{23} < 0.015

VLC Algorithm parameters:
- R = 1.0
- γ = 0.0
- β = 1.0

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

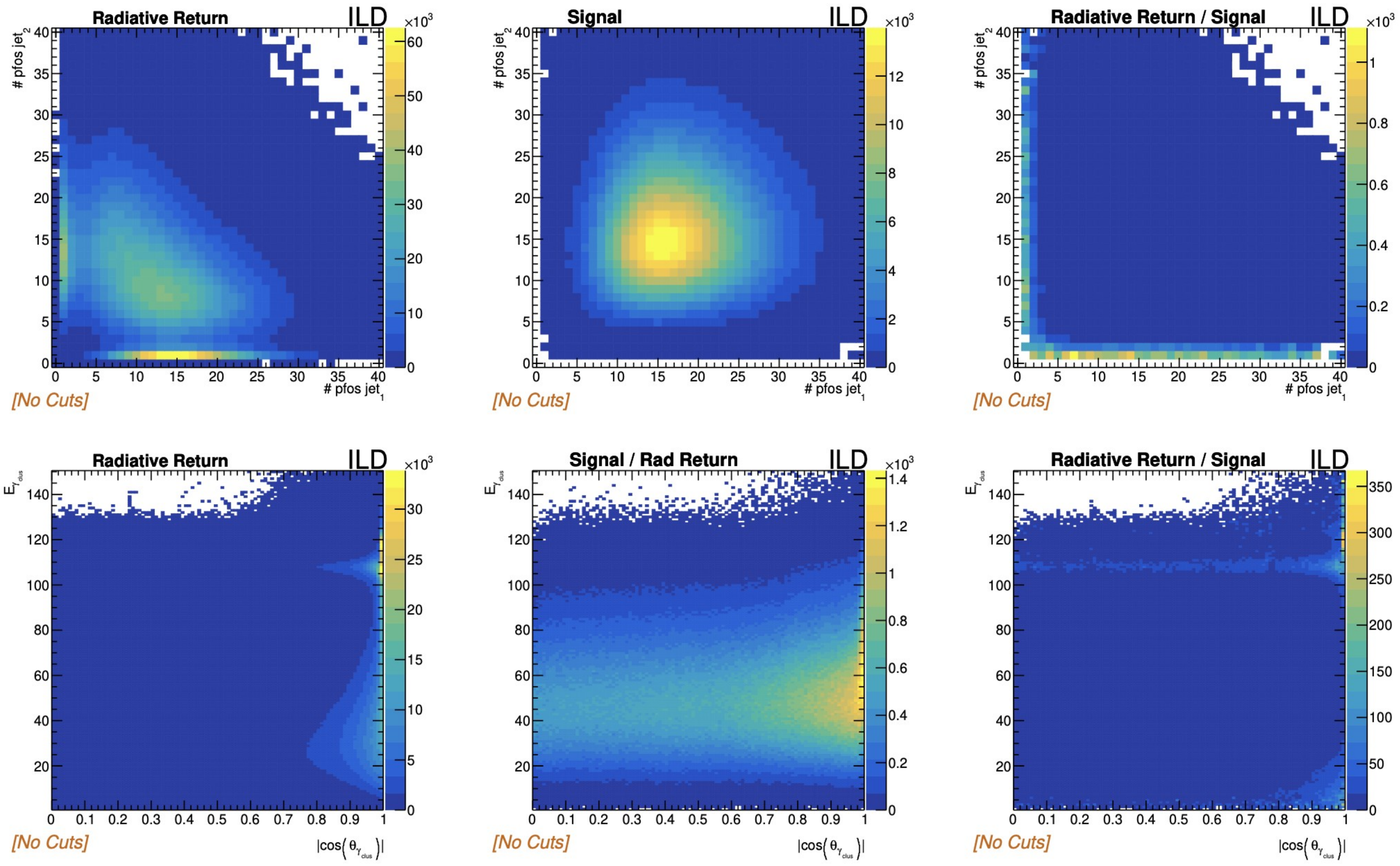
← |cosθ| < 0.9

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

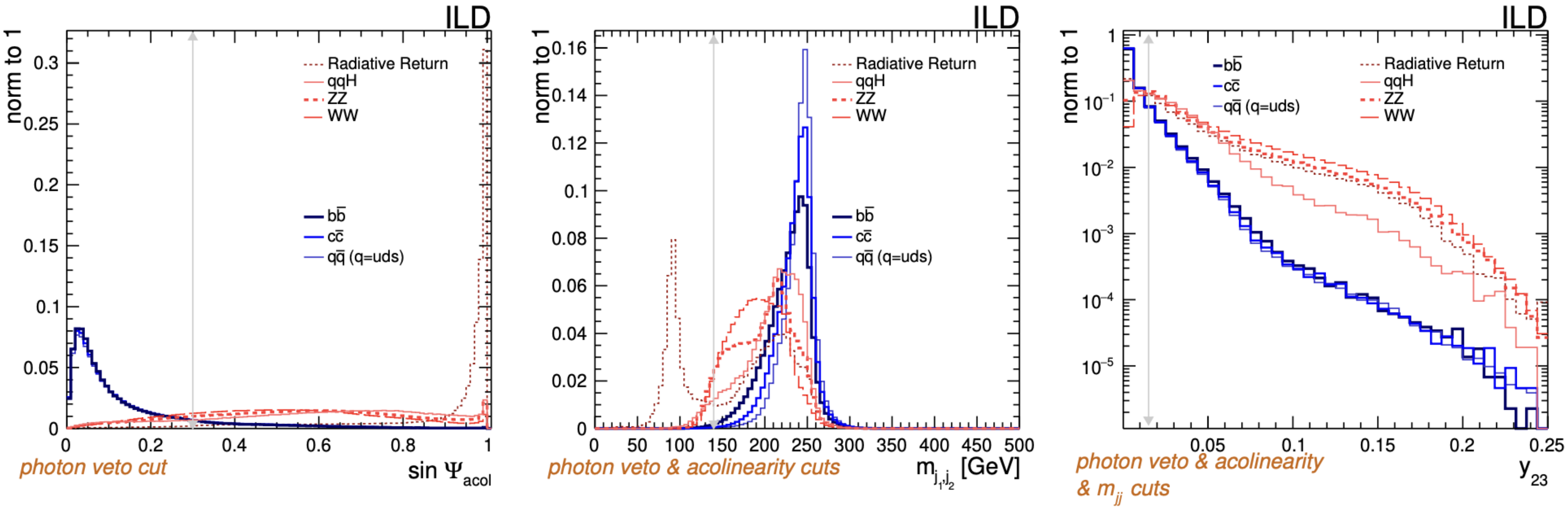


Efficiency of the preselection for the different quark flavours vs the angular distribution of the two jet system (new samples, final configuration)

Radiative return event rejection

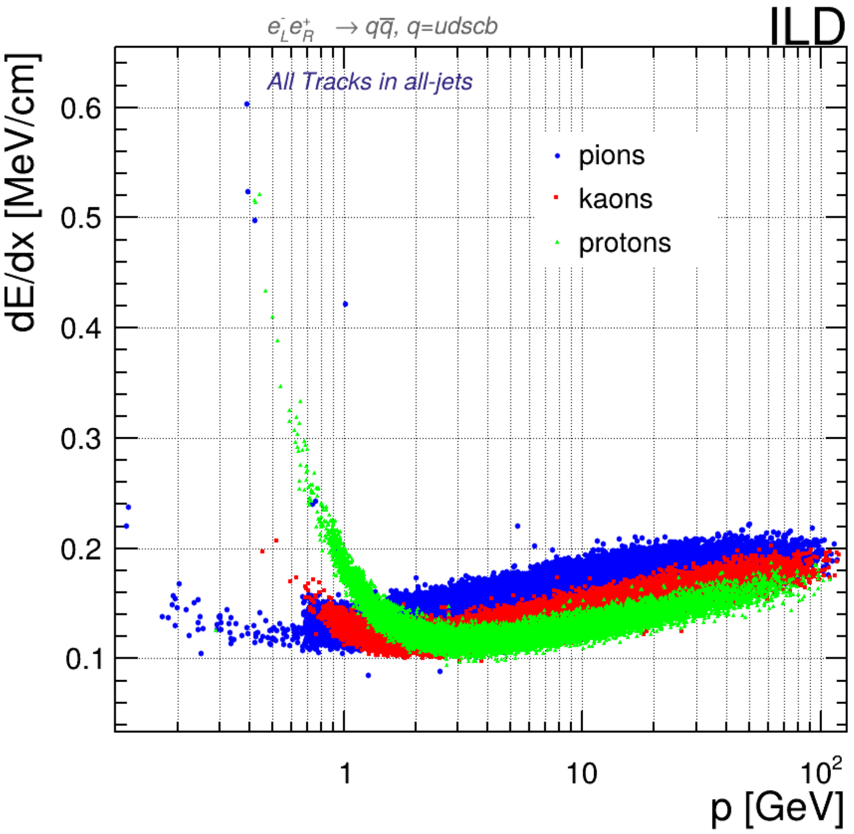


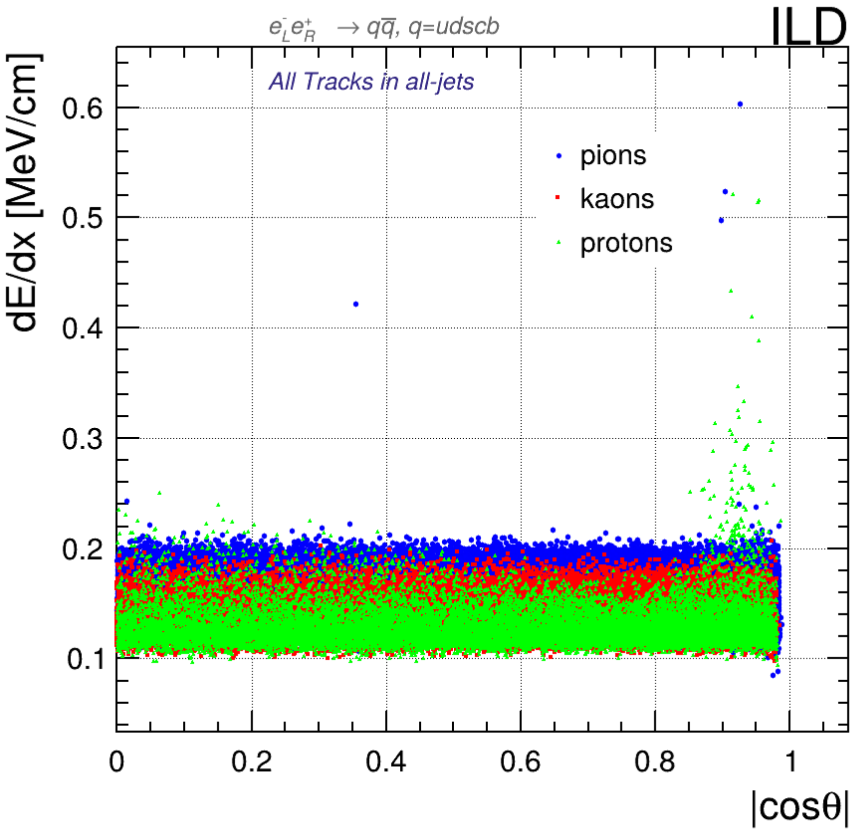
Background rejection



PID via dE/dx: Starting point

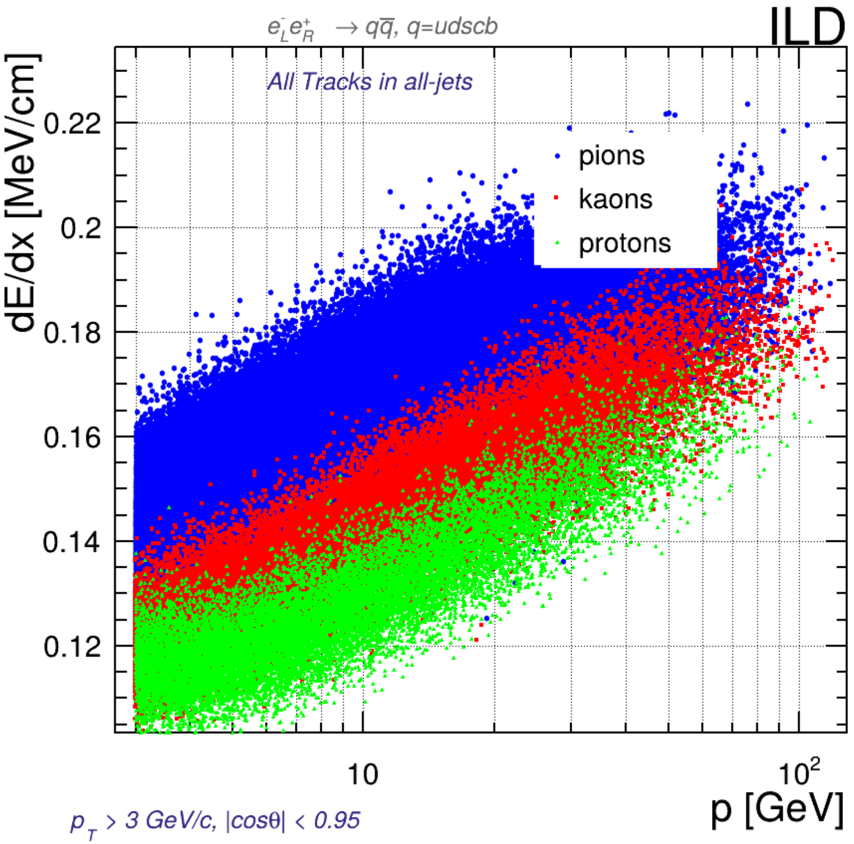
- Not all tracks/PFOs are valid for dE/dx

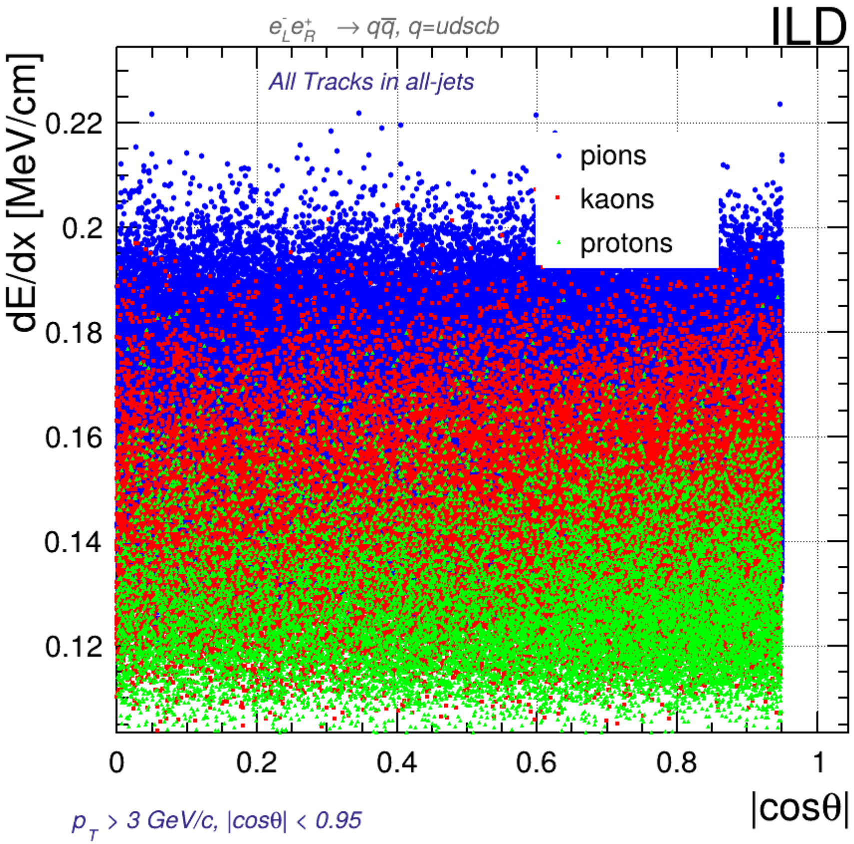




PID: Preselection

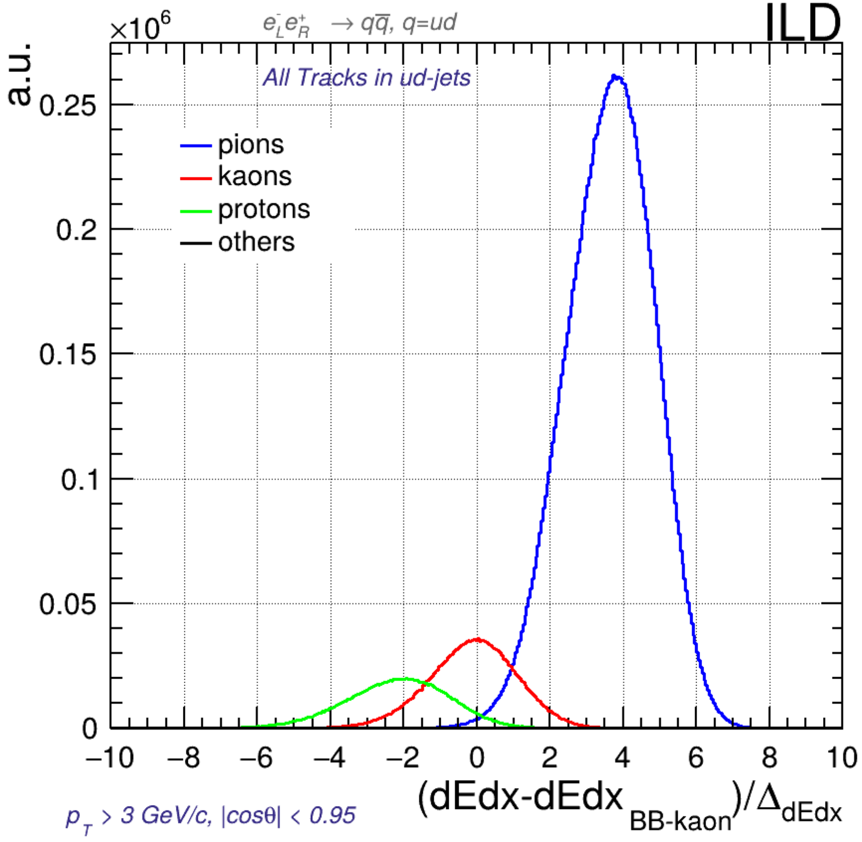
- These three bands can be used to measure an statistical distance

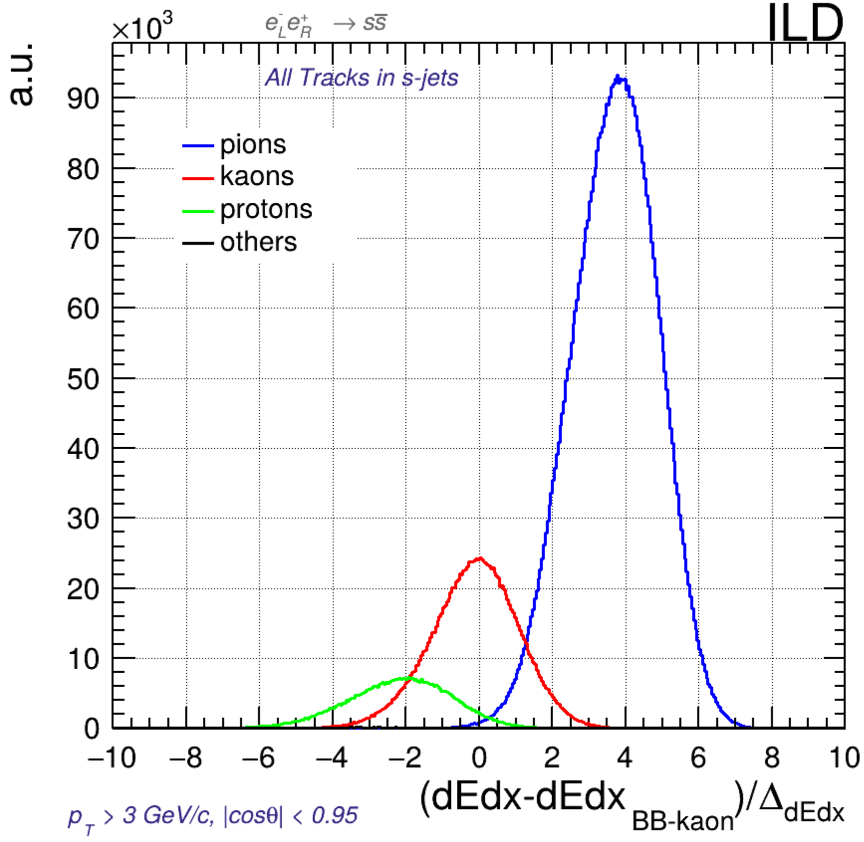




s vs ud: k-distance of tracks

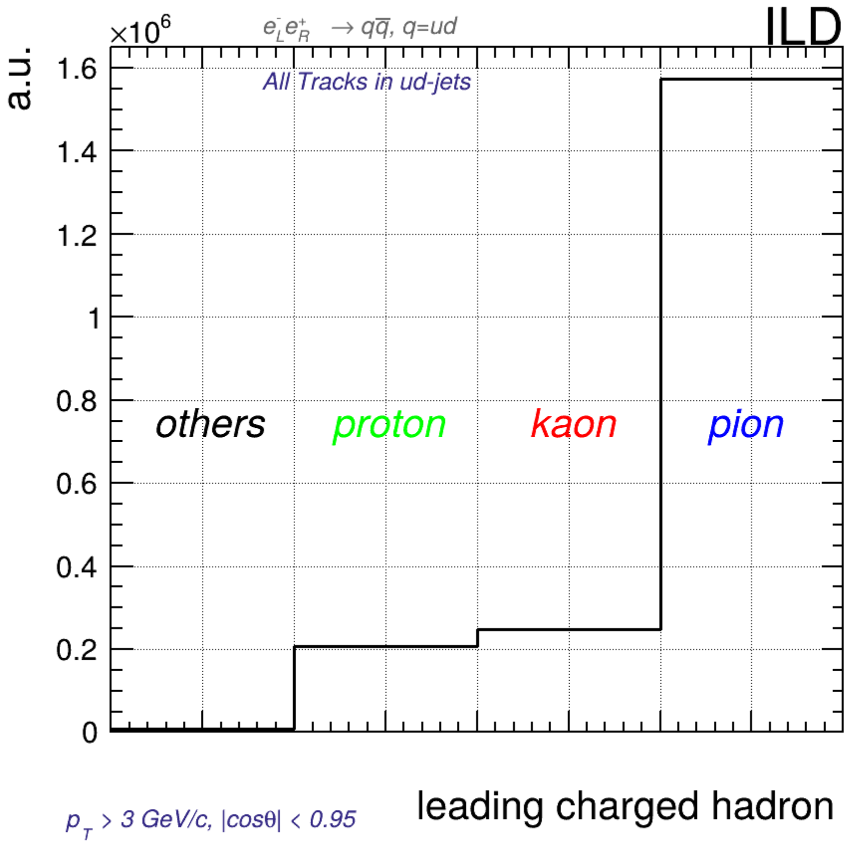
- Example of distance from tracks dE/dx and the theoretical values for kaons

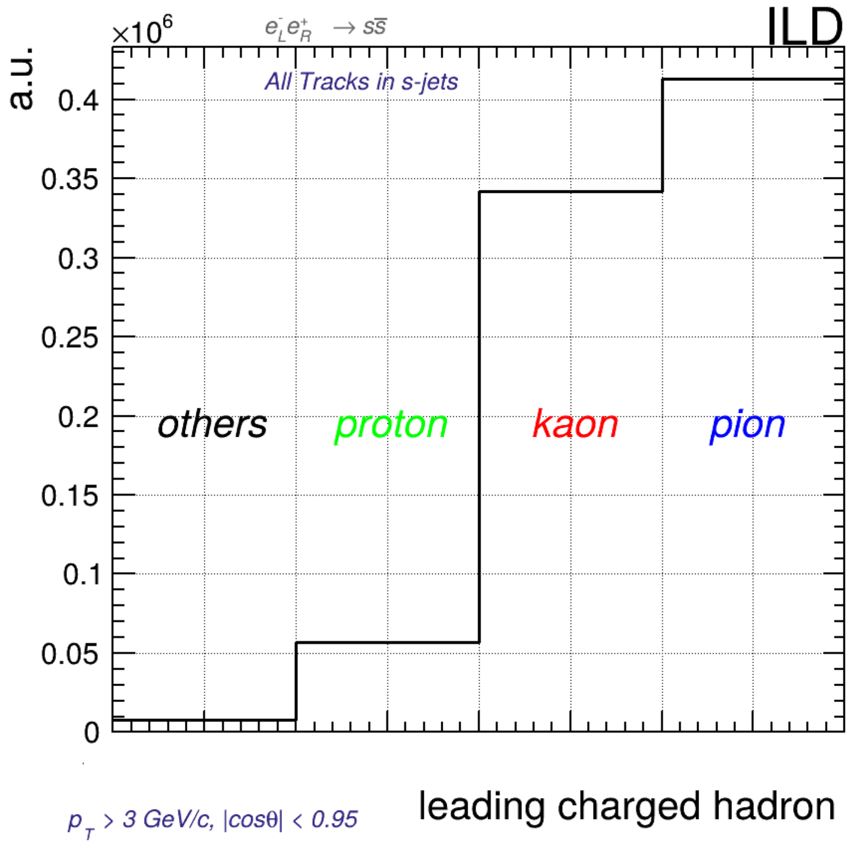




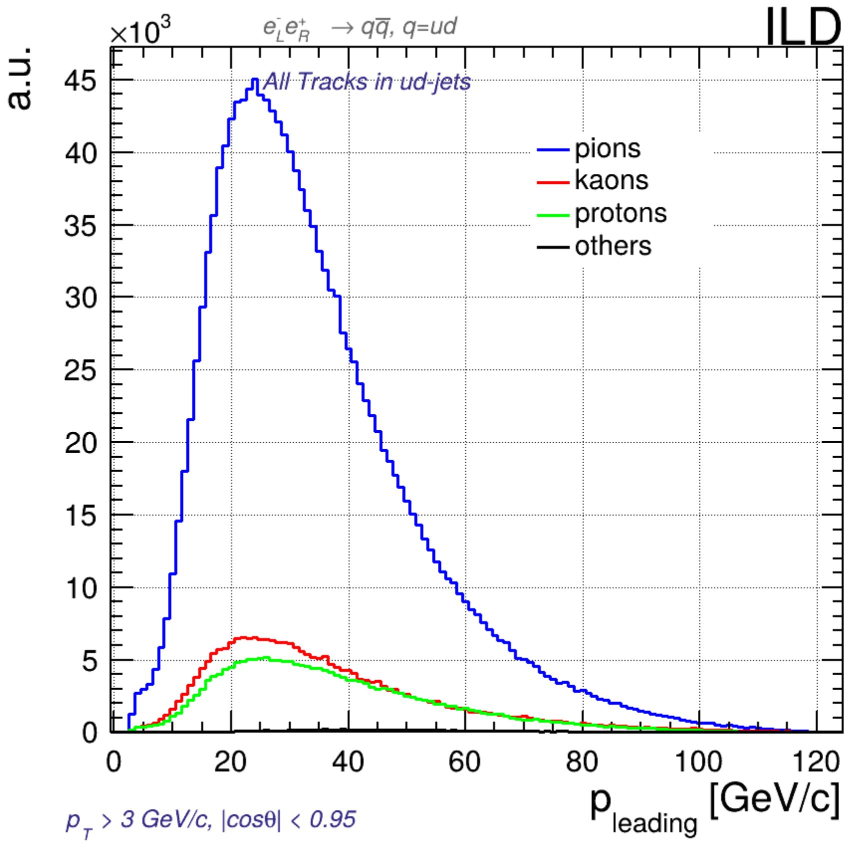
s vs ud: leading charged hadrons

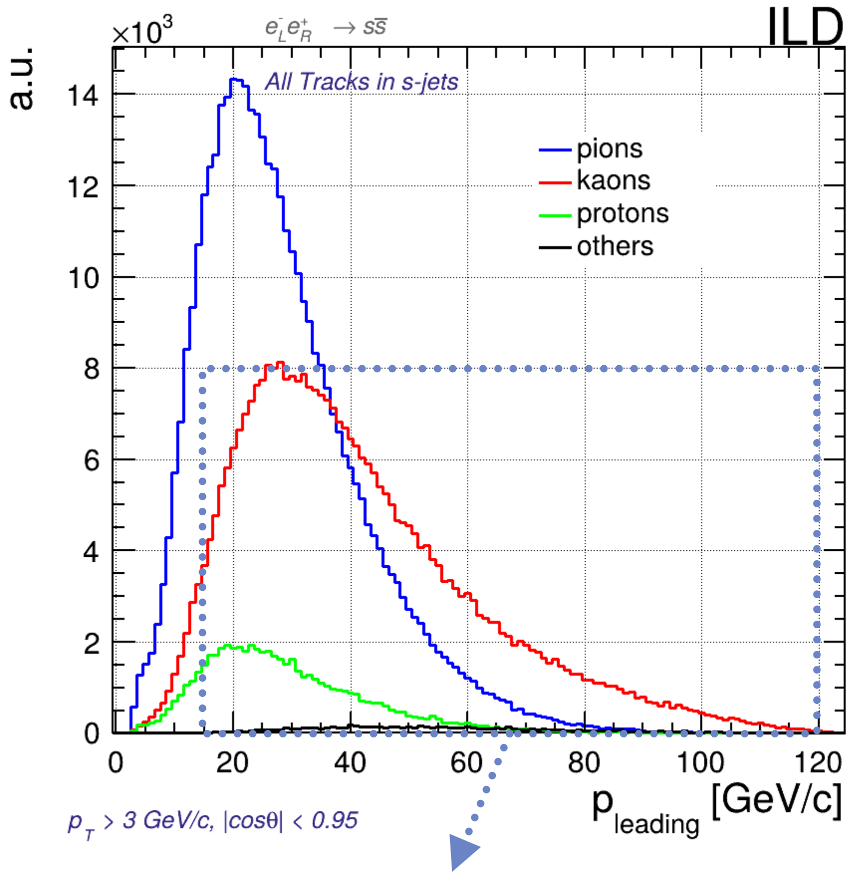
- Different leading track population between s-jets and u/d-jets





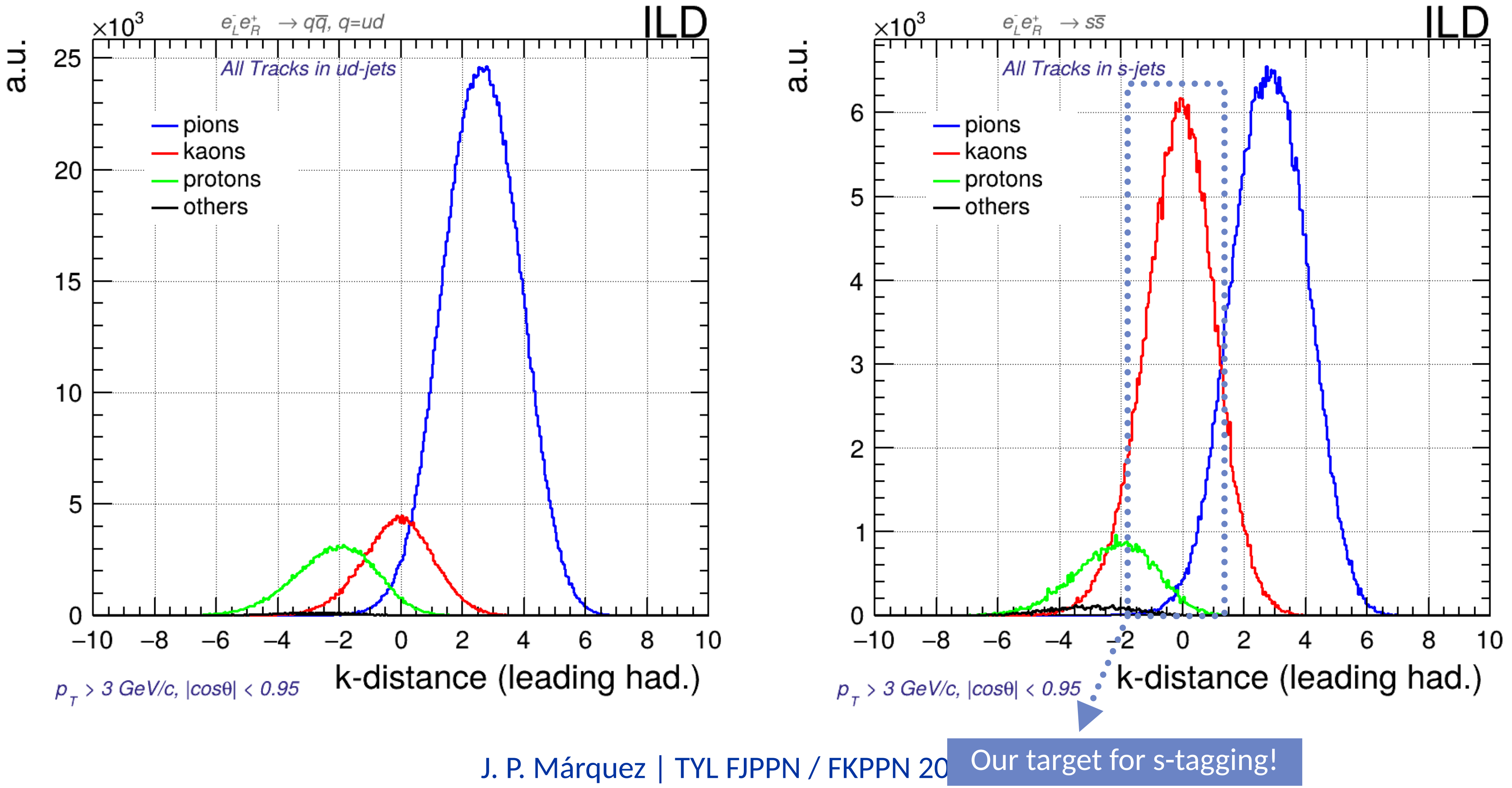
s vs ud: leading charged hadrons





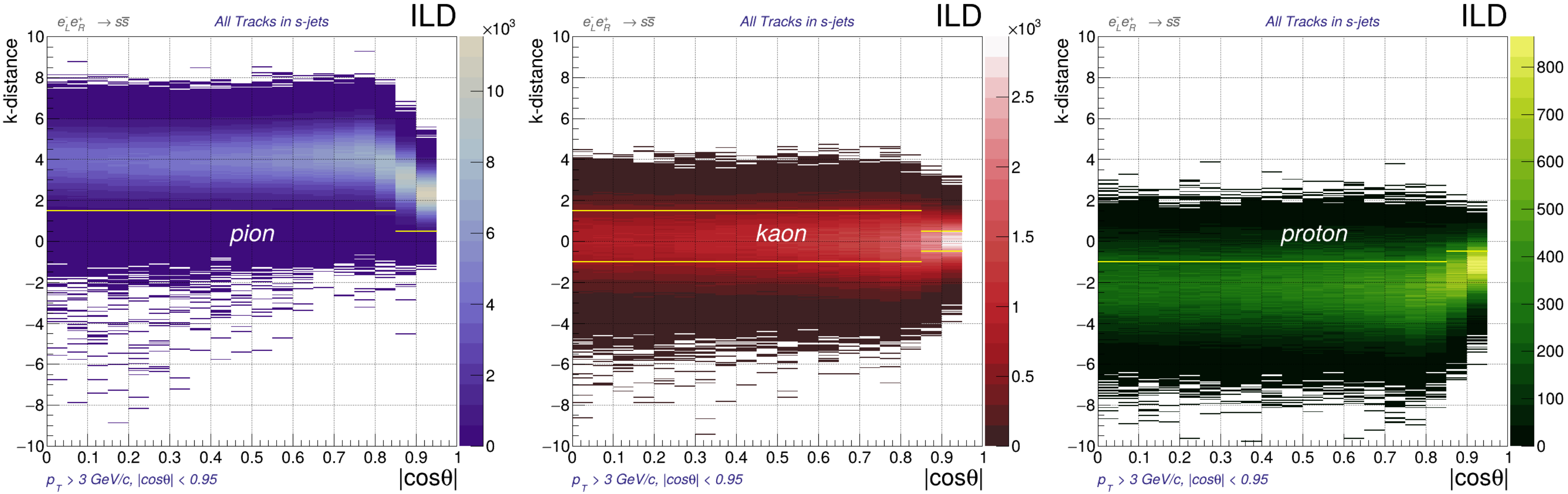
Our target for s-tagging!

s vs ud: k-dist of leading charged hadrons



2d view of k-distance (s quarks)

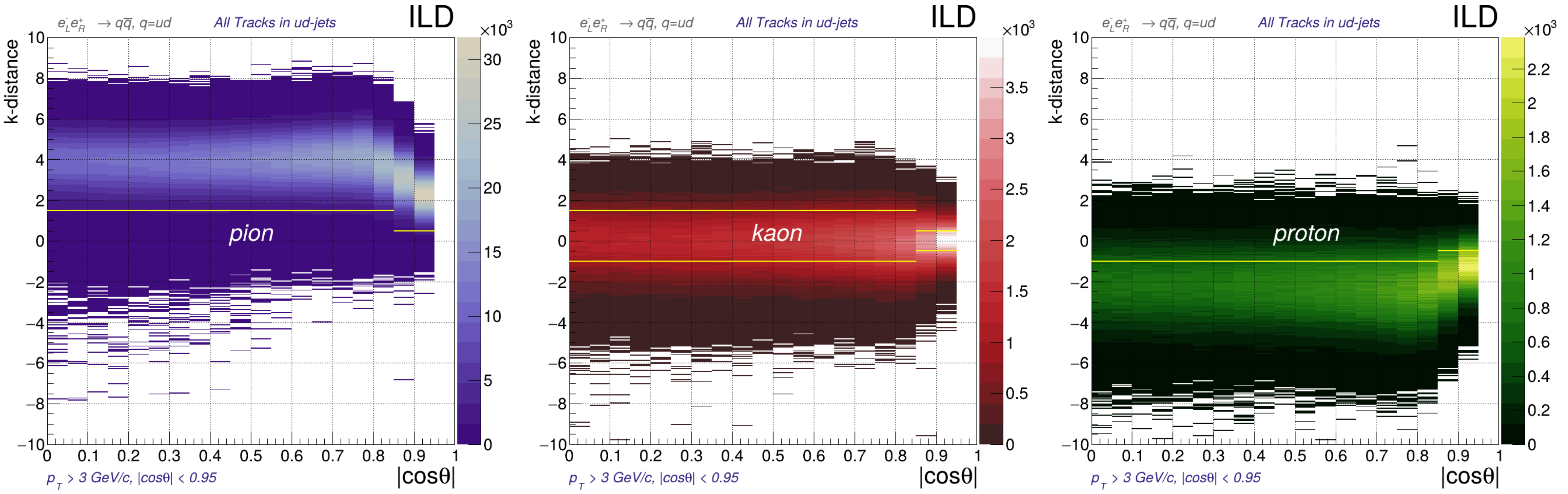
- Angular cuts are performed in these distributions for selection of pions/kaons



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

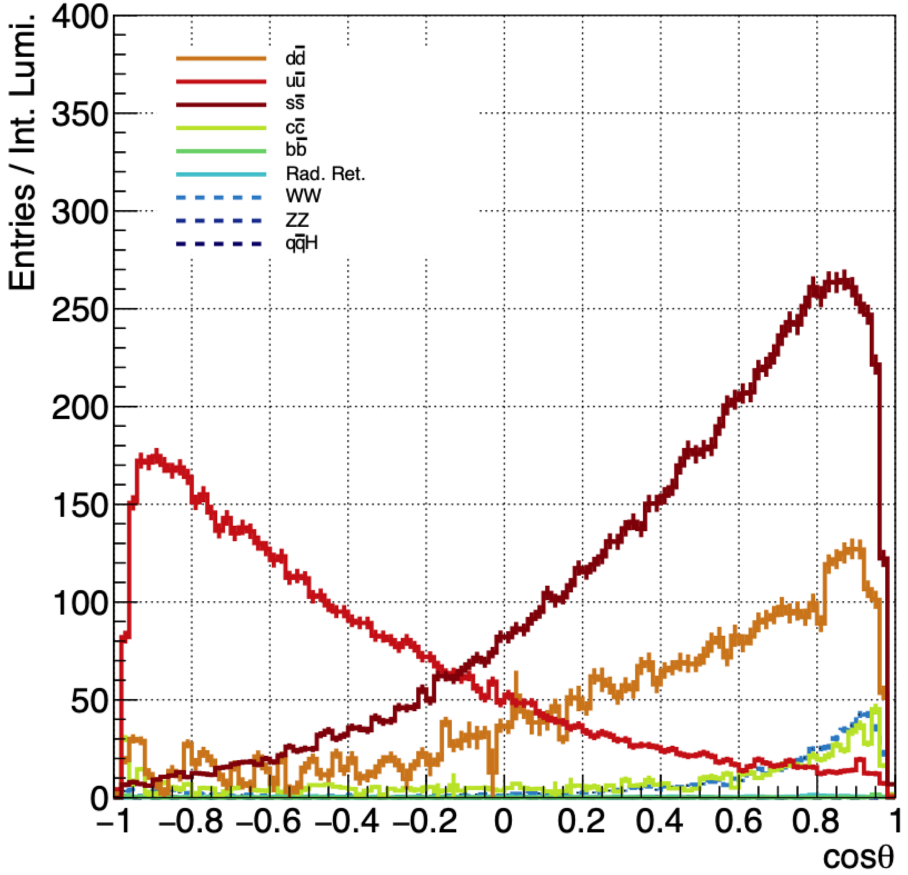
2d view of k-distance (ud quarks)

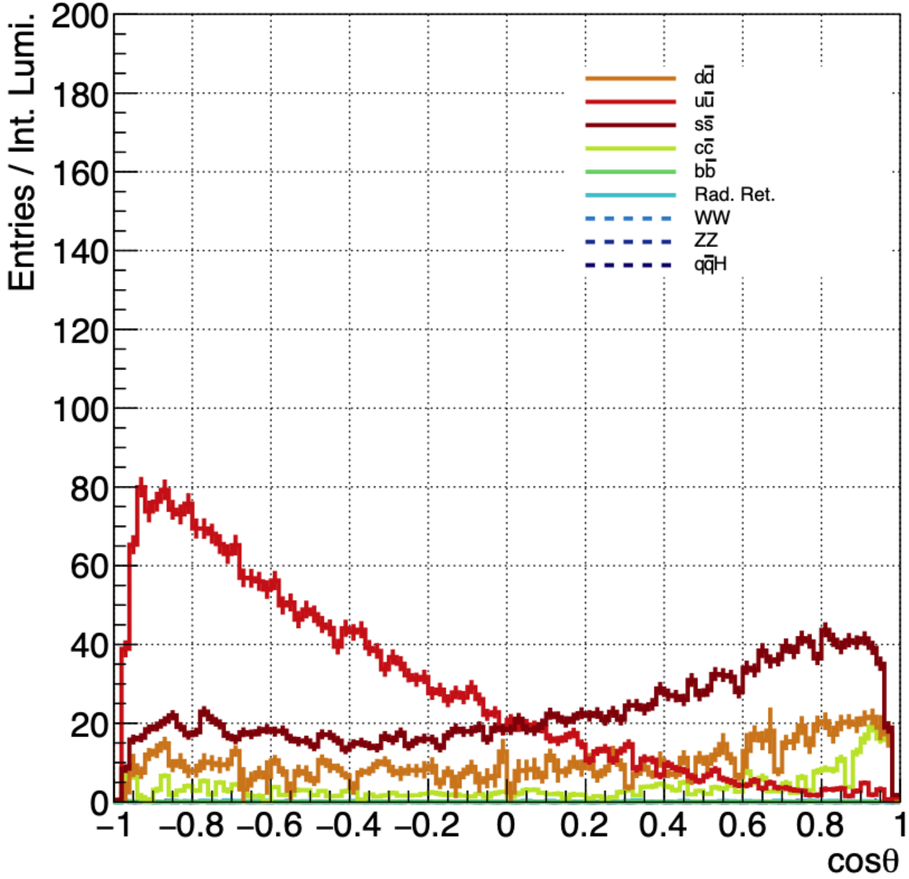
▶ Angular cuts are performed in these distributions for selection of pions



Contributions after preselection

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



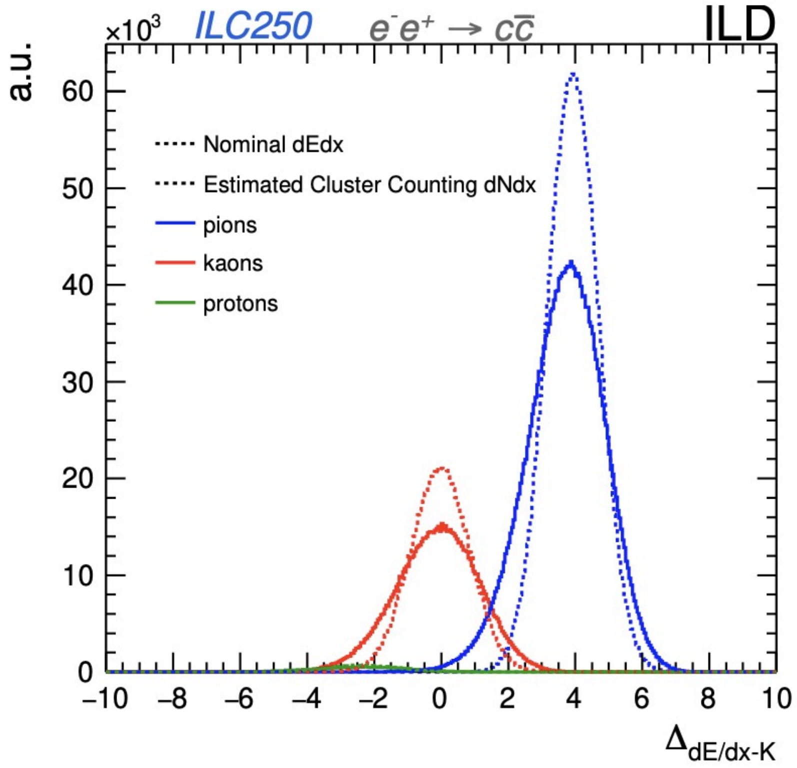


PID hardware prospects

- ► 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

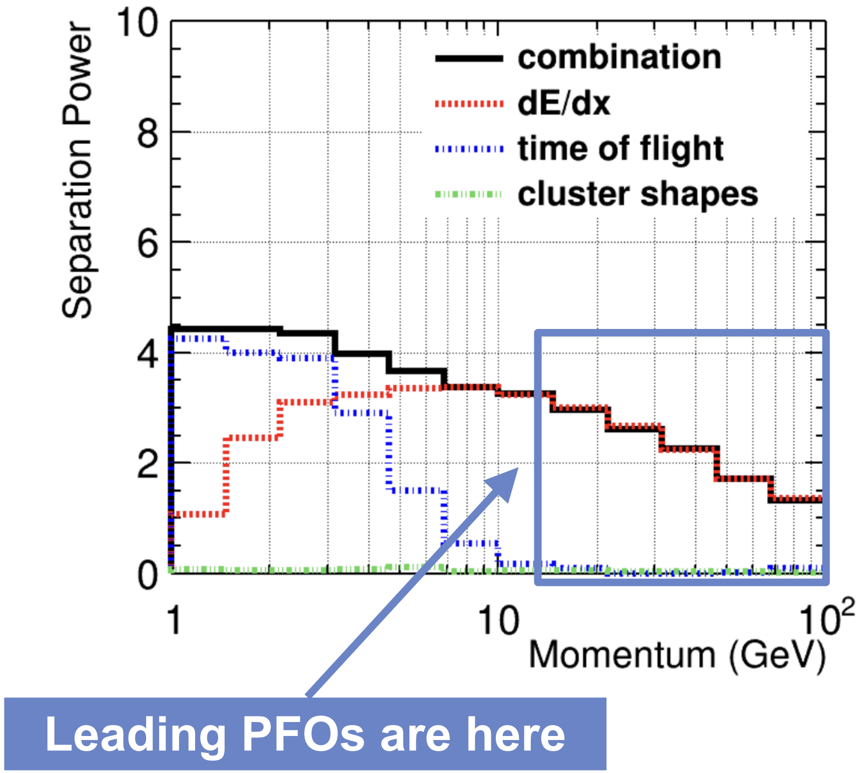


https://github.com/QQbarAnalysis/CheatdEdxDist

PID software improvement: CPID

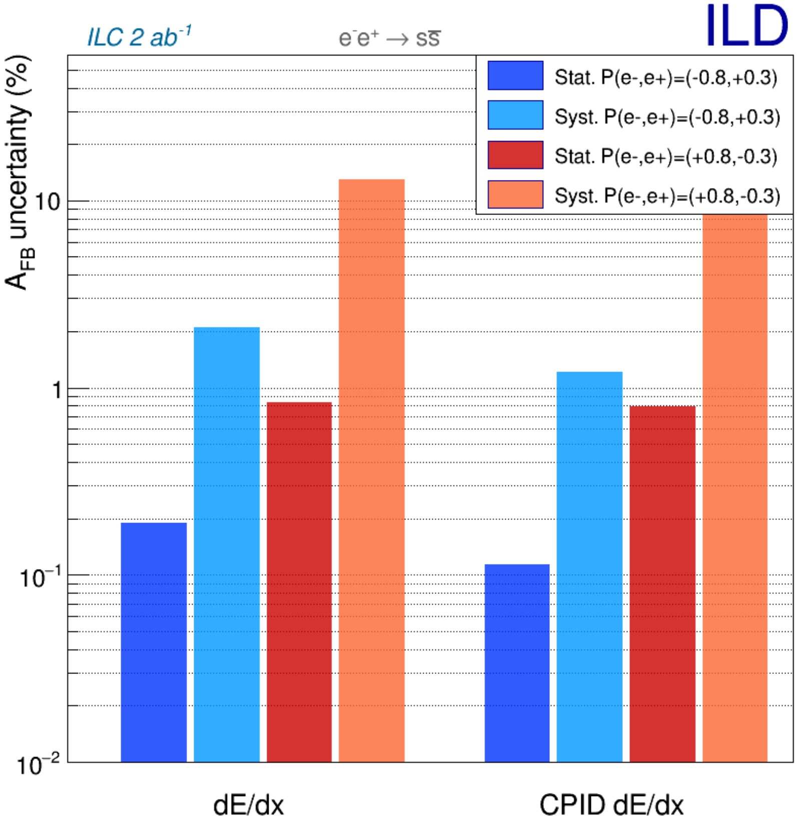
- **C**omprehensive **P**article **ID** Marlin processor:
  - Uses different PID inputs (dE/dx, TOF, etc.)
  - Uses a BDT-based ML algorithm for classification
  - Easy to adapt to different MC ids or PID info

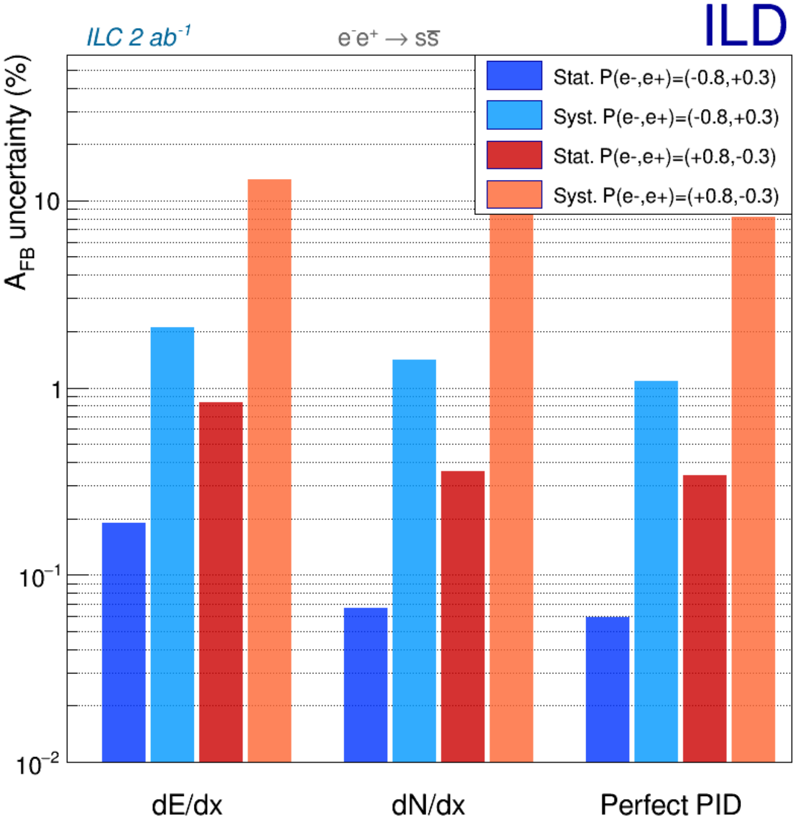
- In our case, the CPID was trained tackling our leading PFOS:
  - Only Kaon/Pion/Proton separation
  - 3 GeV < Momentum < 100 GeV



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

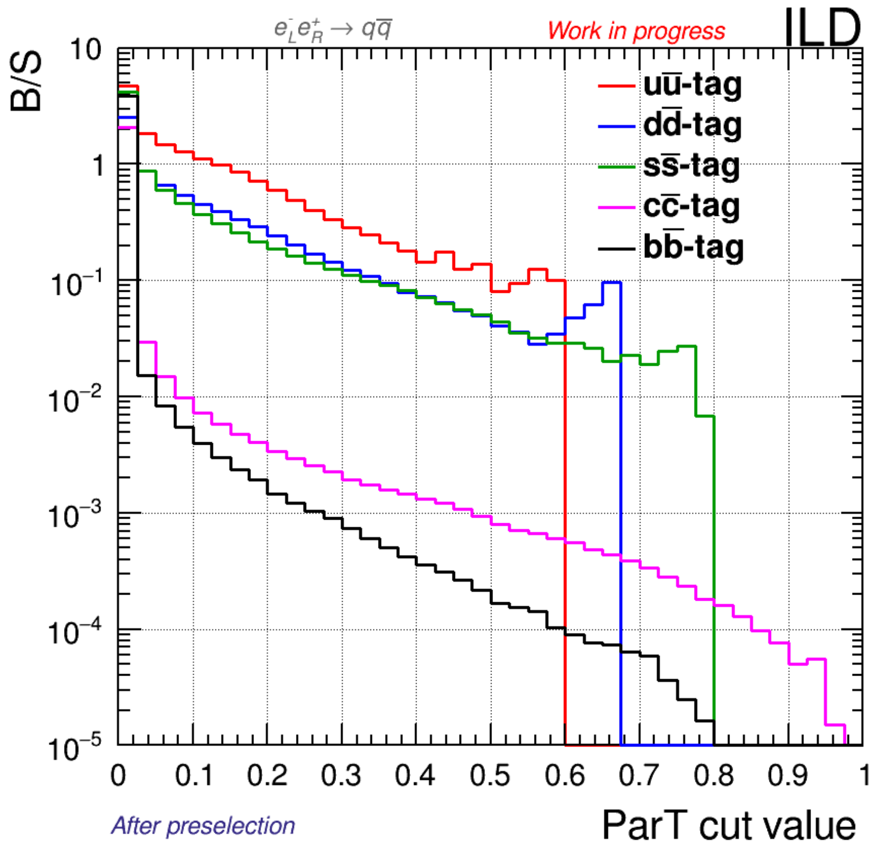
Results (ILC250)

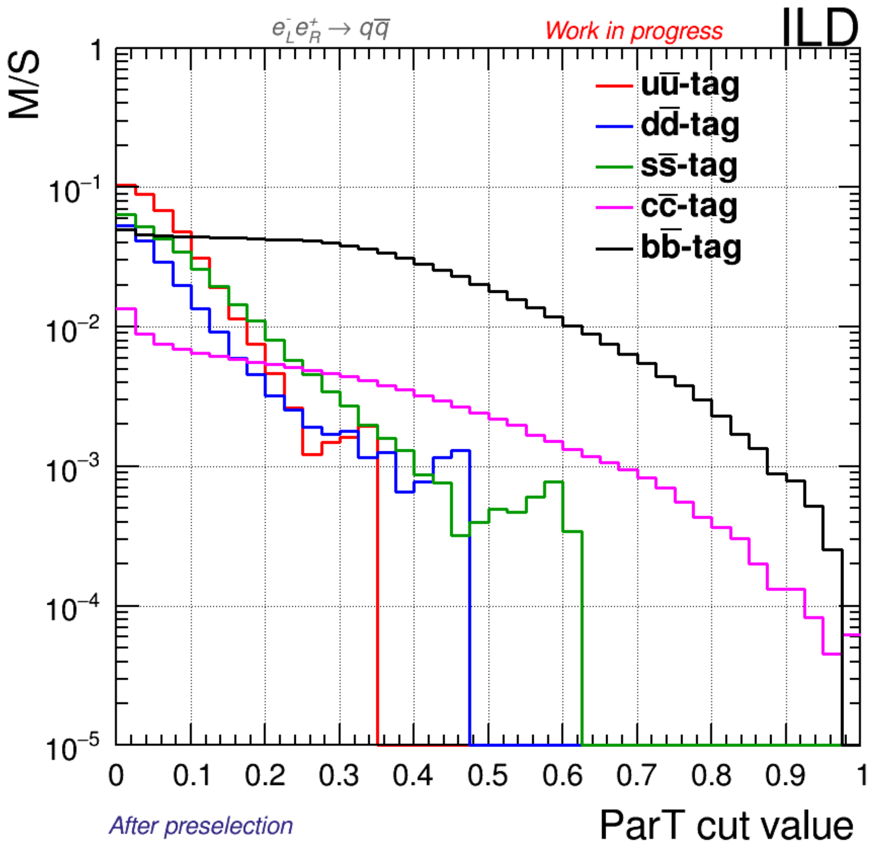




Background/signal and Migrations/signal ($e^-_L e^+_R$)

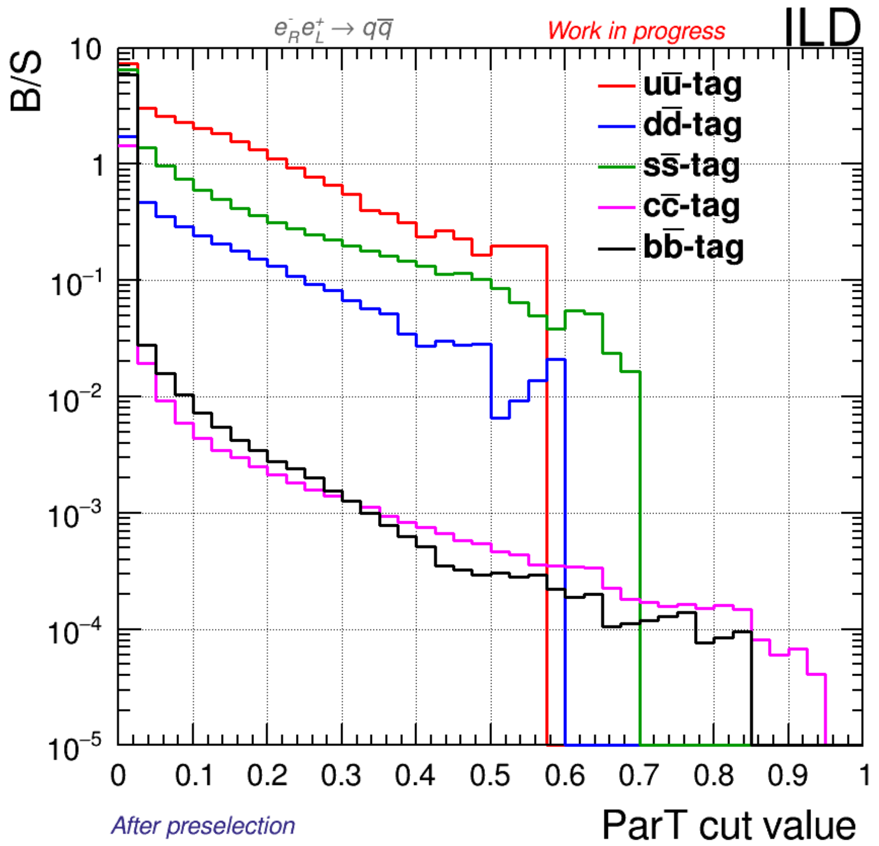
- I set different WP for each flavor (0.4 for s, 0.6 for b/c)

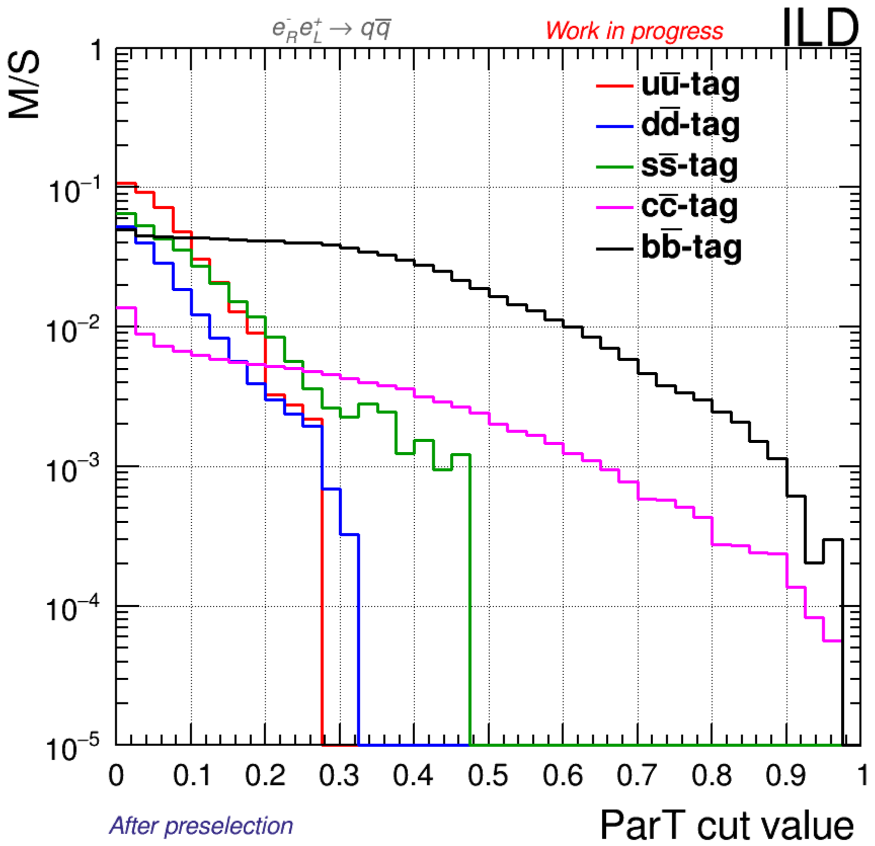




Background/signal and Migrations/signal ($e^-_R e^+_L$)

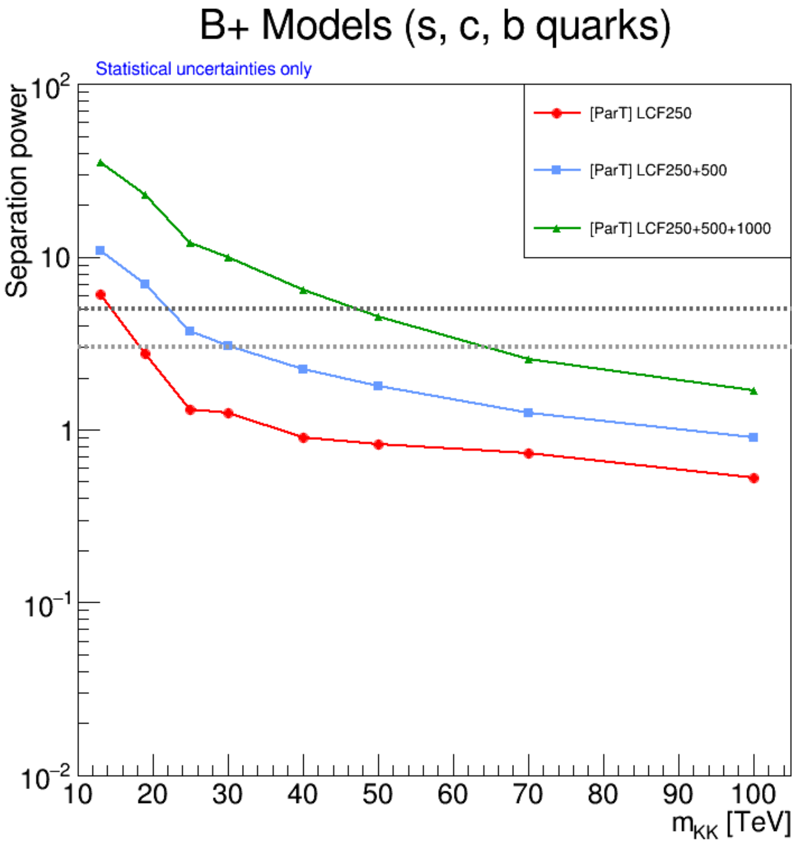
- I set different WP for each flavor (0.5 for s, 0.6 for b/c)

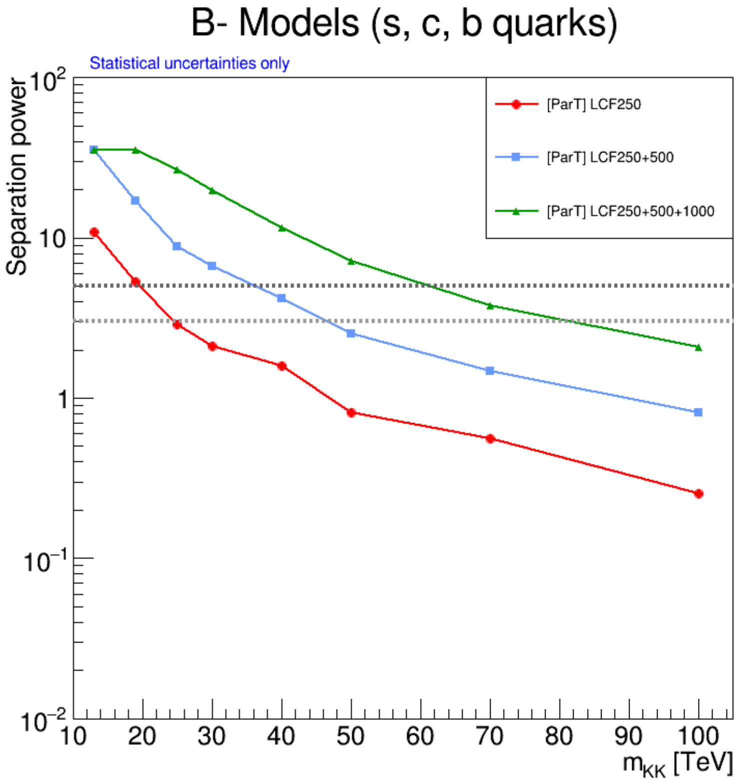




Prospects for GHU phenomenology (LCF@CERN)

- Extrapolating statistical uncertainties from 250 to 500 & 1000 GeV





Up to ~45 TeV mass range for B+ and ~60 TeV for B-