



# AFB studies at 500 GeV

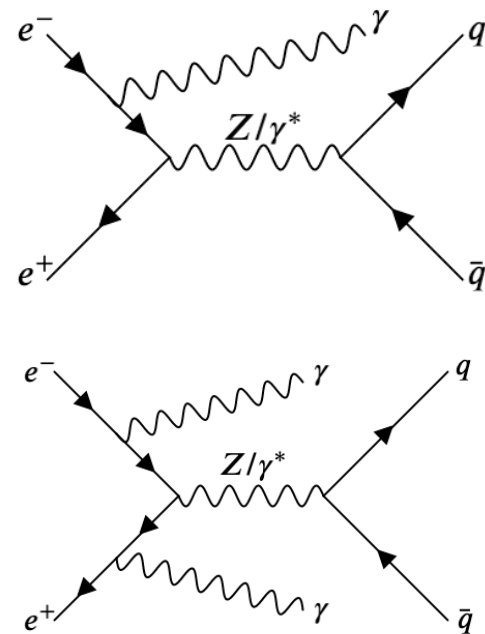
500 GeV  $q\bar{q}$  sample validation

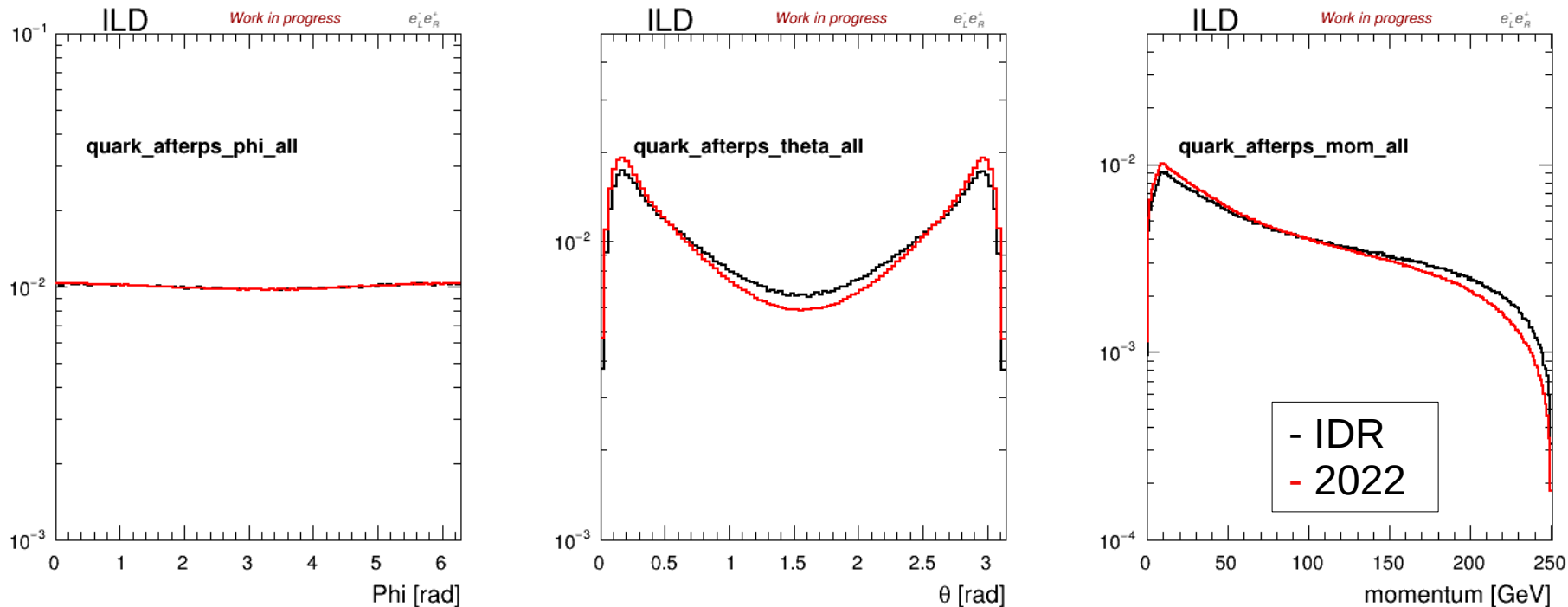
*ILD Analysis/Software Meeting*  
*26/10/22*

Jesús P. Márquez Hernández

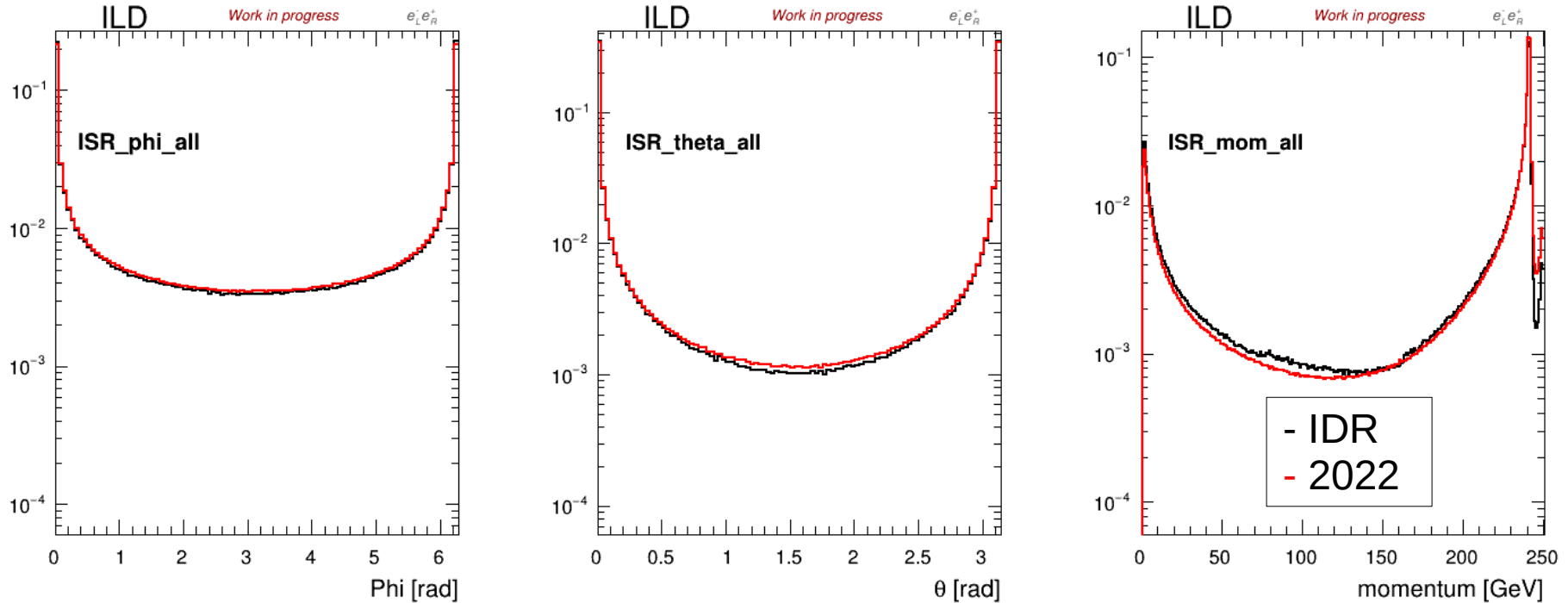


- We define “signal” as those events whose ISR total energy is below 50 GeV.
  - We define “radiative events” as those with  $E_{\text{ISR}} > 50$  GeV.
- Samples:
  - Old sample (IDR samples):
    - 46.4 and 47.0  $\text{fb}^{-1}$  ( $e_{\text{L}}p_{\text{R}}$  &  $e_{\text{R}}p_{\text{L}}$ ).
    - Whizard 1.9.5.
    - ILD\_I5\_v02
    - ILCSoft v02-02-01.
  - New sample (2022 samples):
    - 241.1 and 429.6  $\text{fb}^{-1}$  ( $e_{\text{L}}p_{\text{R}}$  &  $e_{\text{R}}p_{\text{L}}$ ).
    - Whizard 2.8.5.
    - ILD\_I5\_v02
    - ILCSoft v02-02-03.



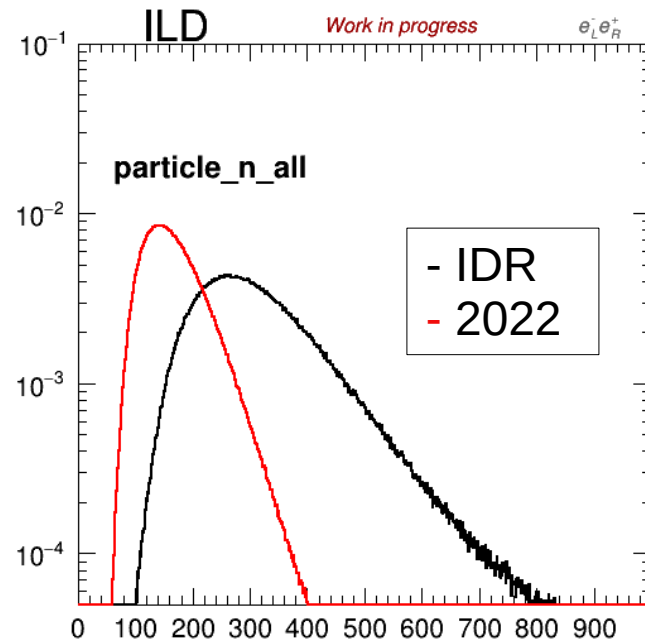
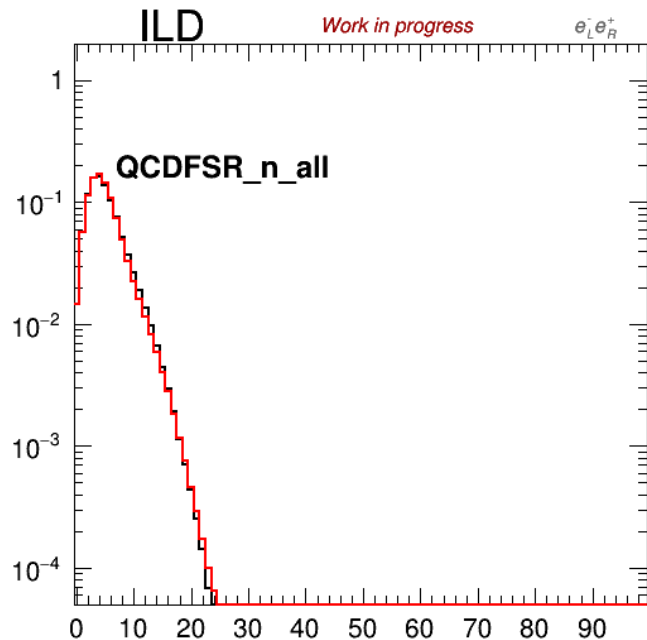


- Kinematics of the quarks after QCD PS.



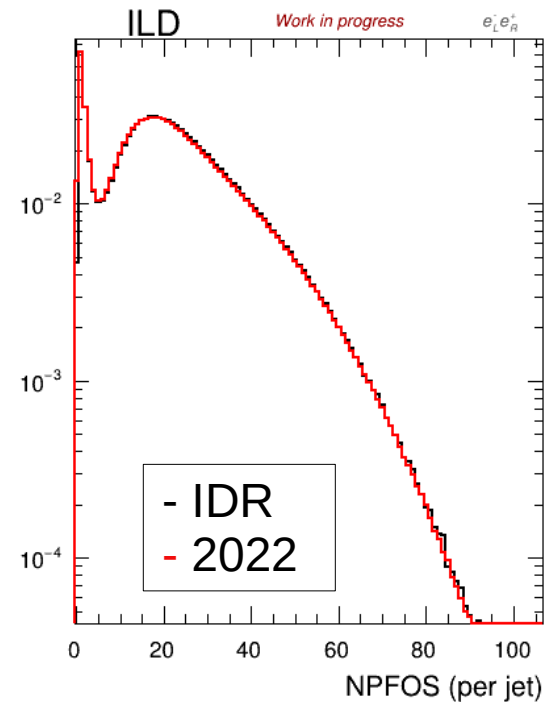
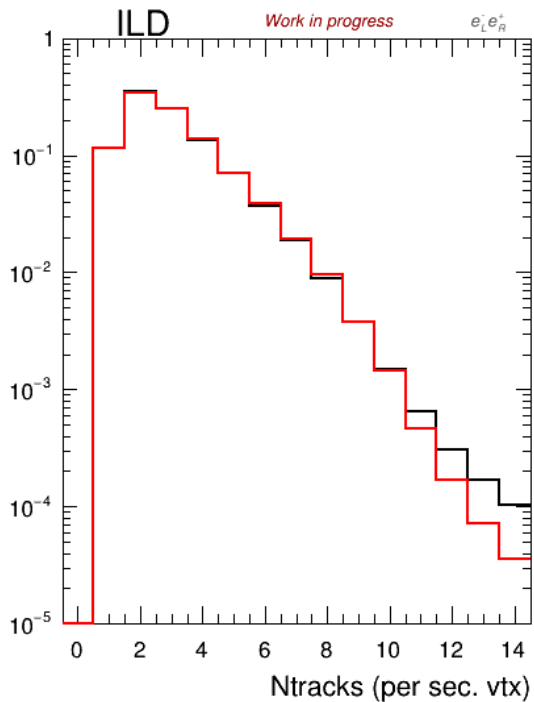
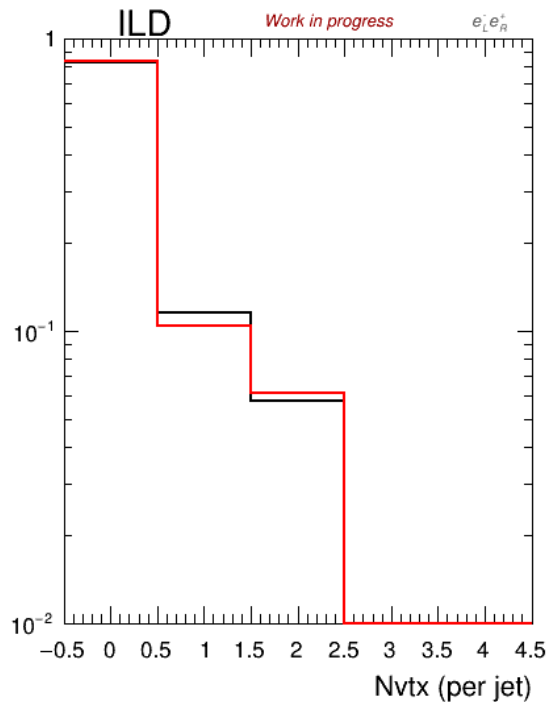
- Kinematics of photon ISR (from both incoming particles).

# Generator level: Particles produced

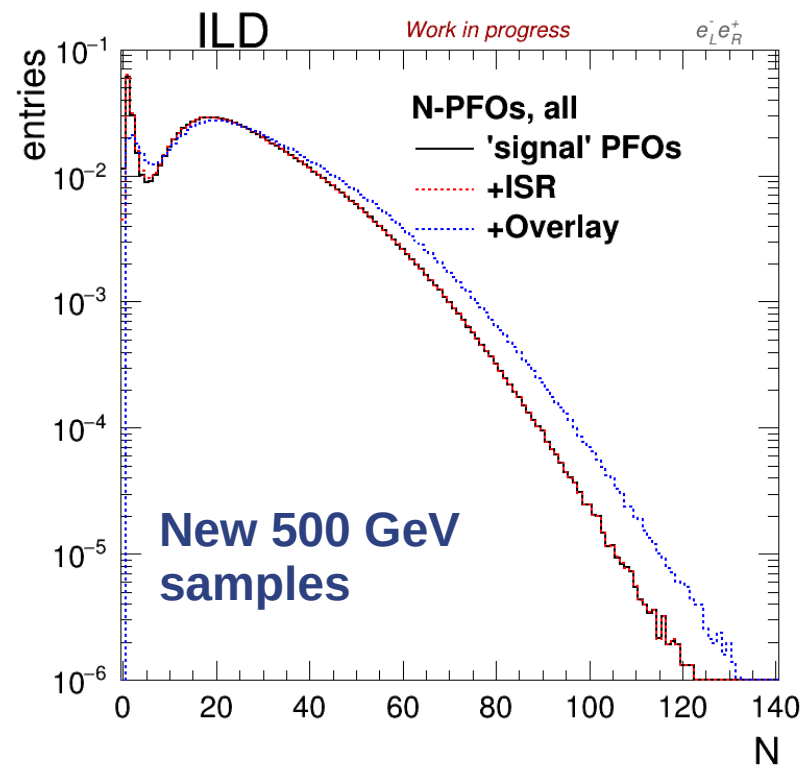
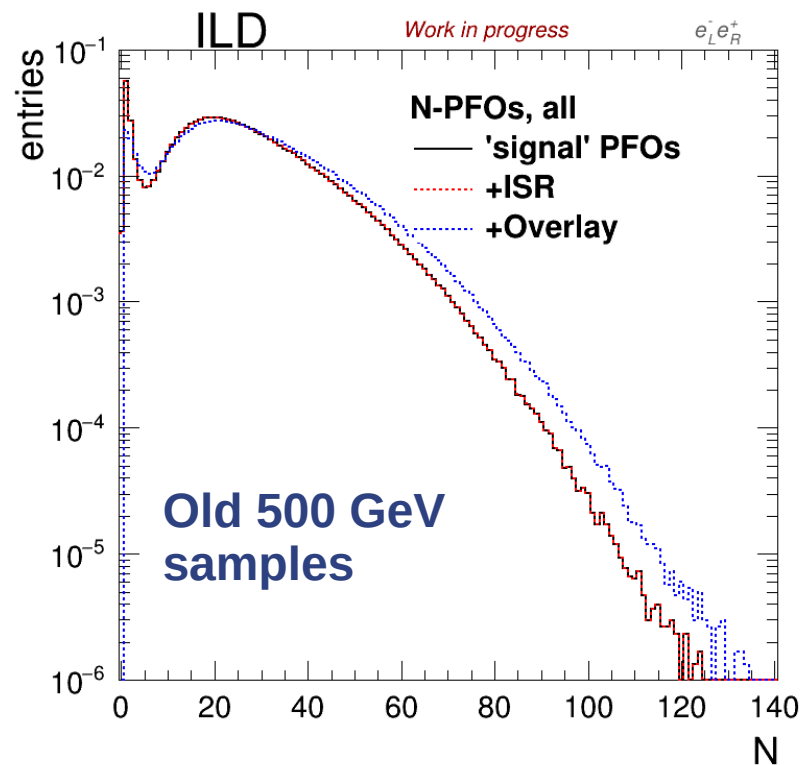


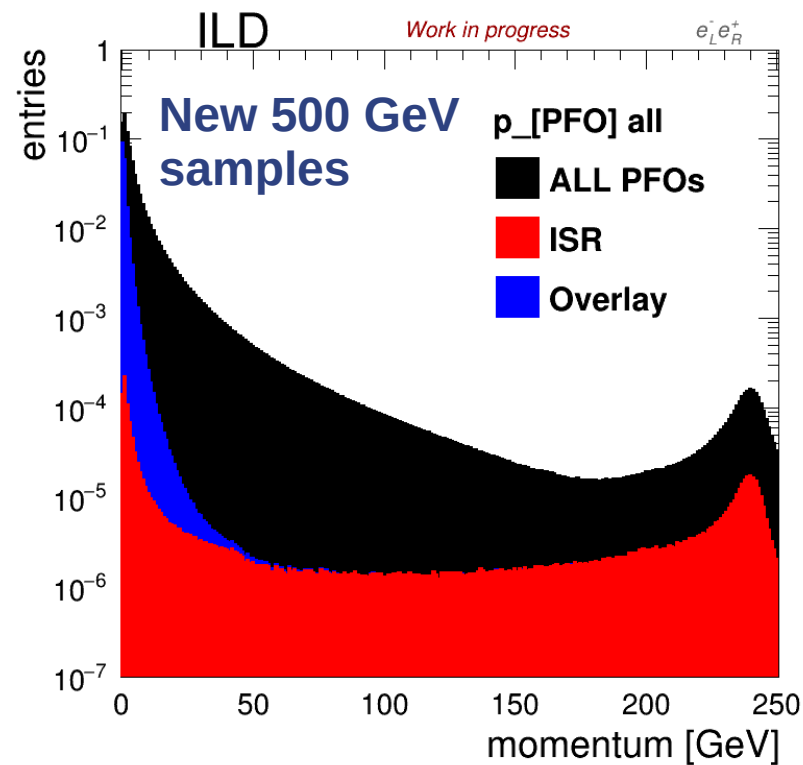
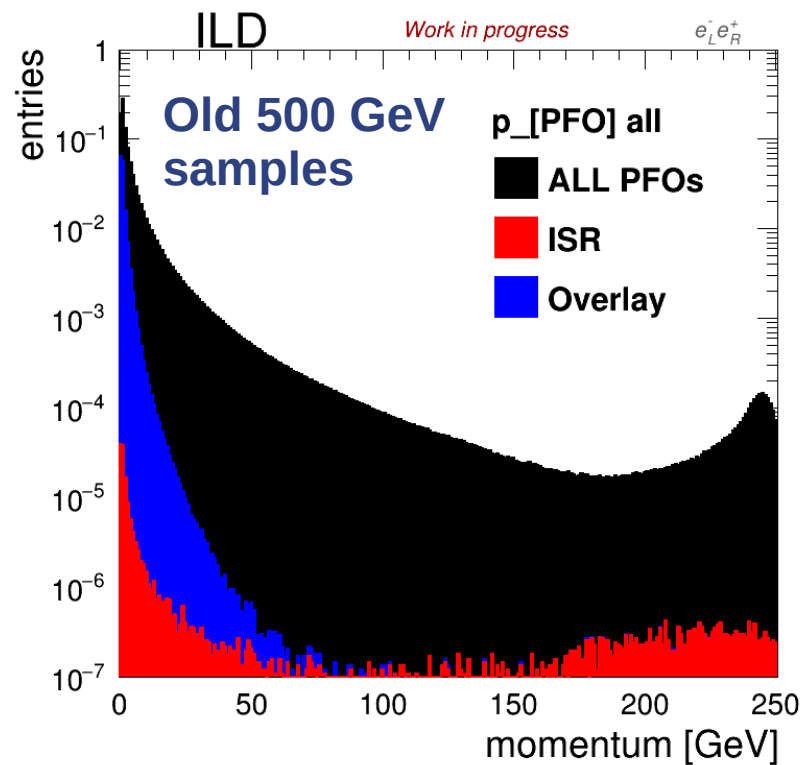
- Partons produced during PS (Left plot).
- Stable particles before the detector (Right plot).





# Reco. level: Overlay BKG (Signal)

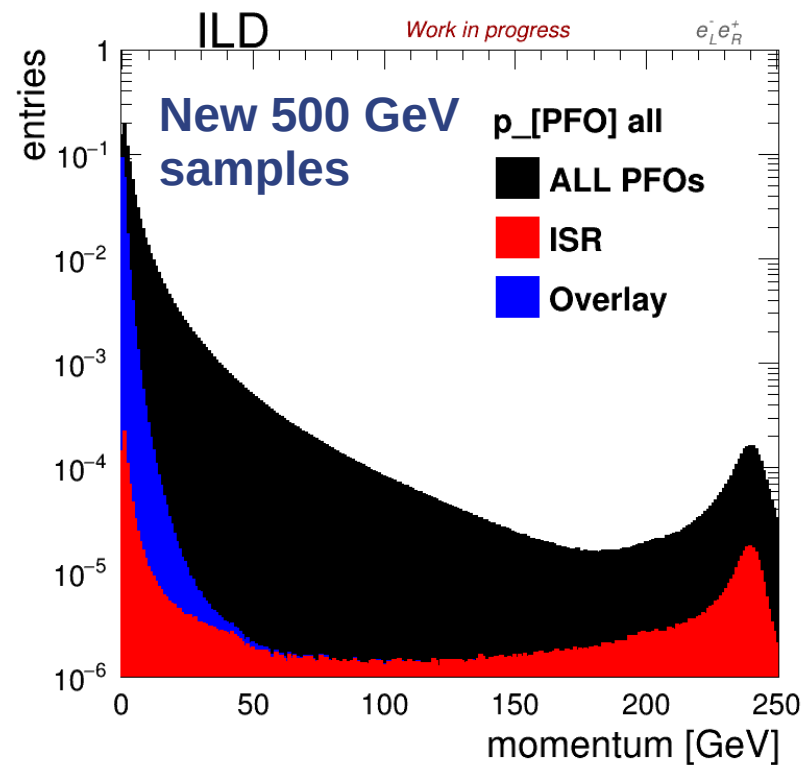
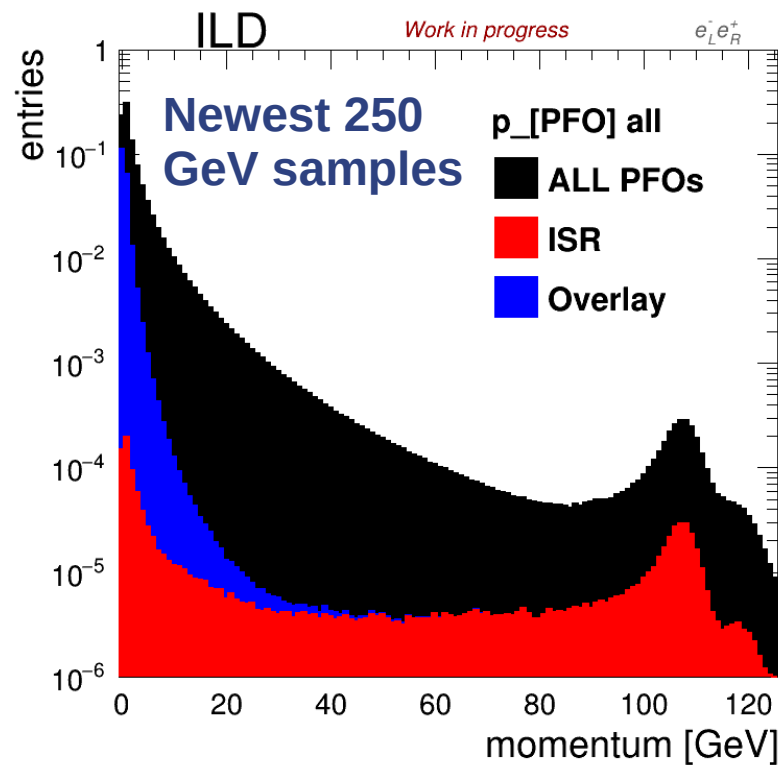




- We noticed an error in our code when counting the ISR's PFO in the old samples (left plot).
  - Better to compare this plot with the 250 GeV samples (next slide).

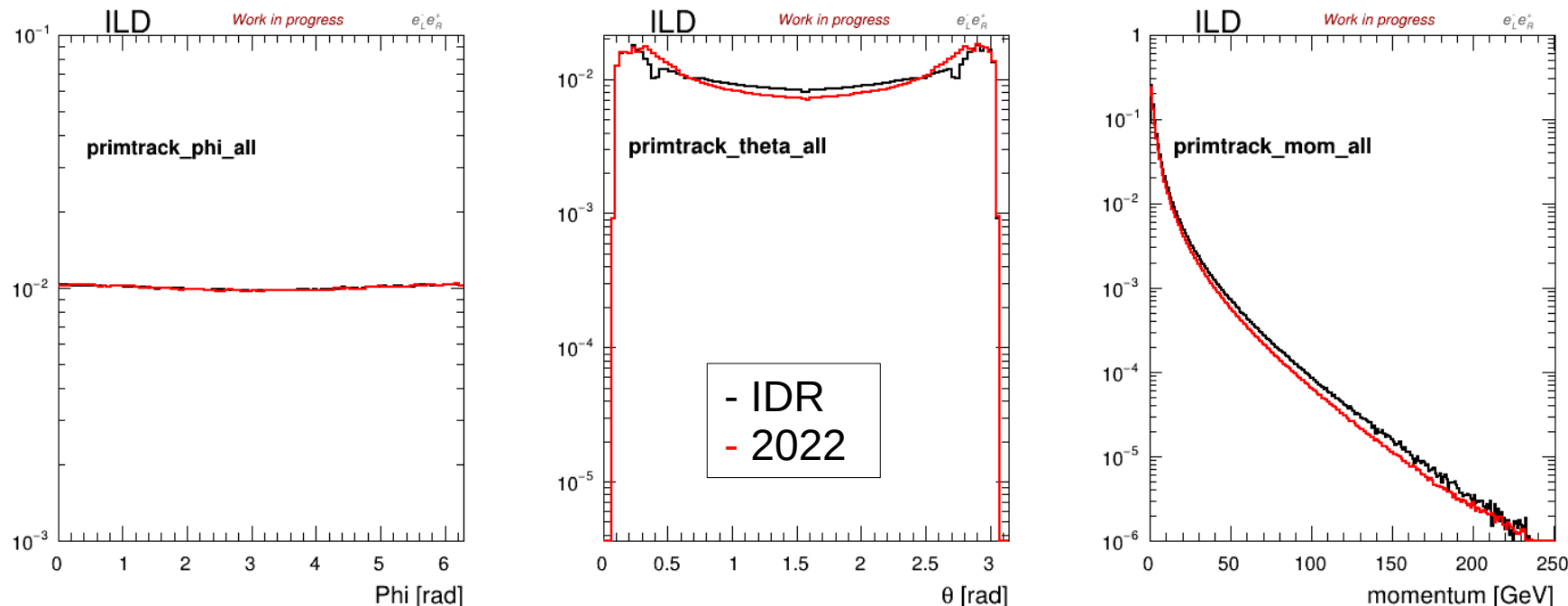


# Reco. level: Energy of PFOs

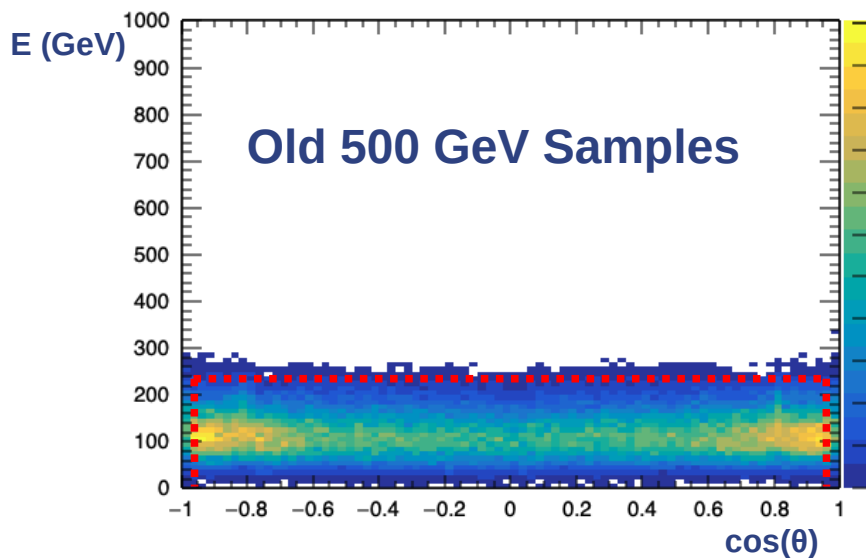


- PFO's content is as expected.

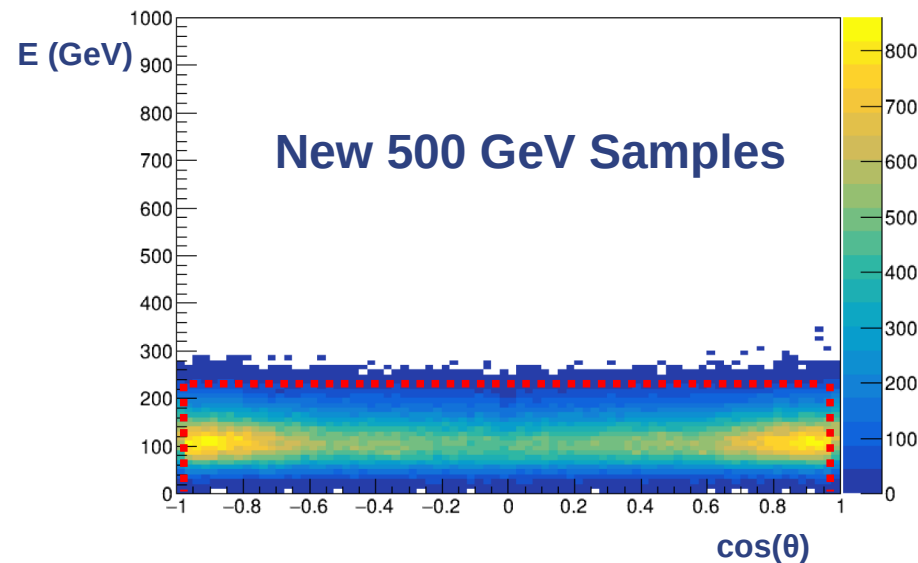




- These two “inverted horns” in the forward/backward distribution disappear.
  - Mismatch between tracks and calorimeter objects in the ECAL barrel-endcap transition).



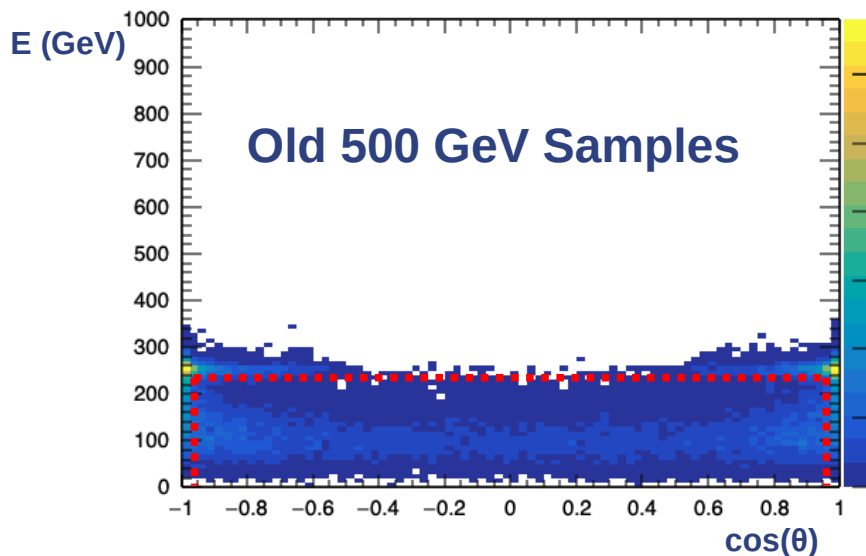
$q\bar{q}$  events (uds)



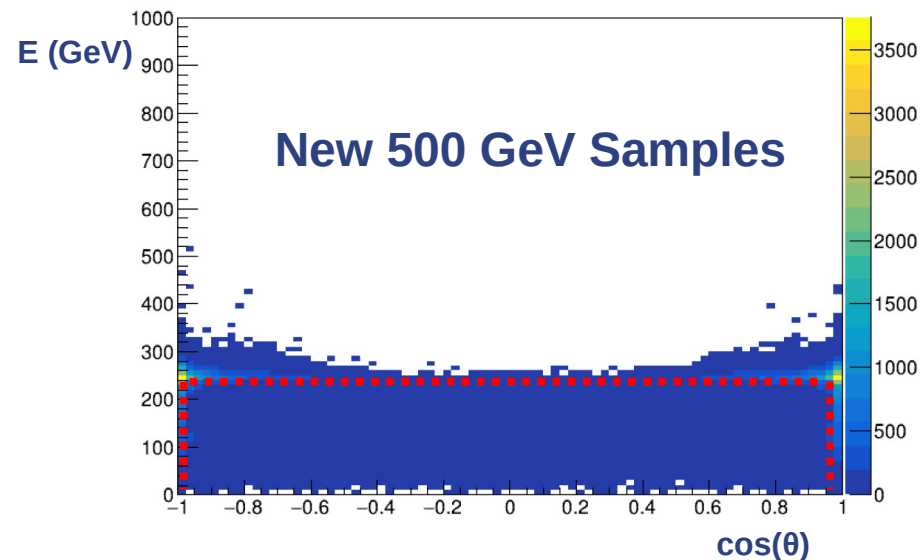
$E_{\text{ISR}} < 50$  GeV

- Energy vs  $\cos(\theta)$  distribution of neutral PFOs (photon ISR candidates), identified with PandoraPFO. Useful for ISR removal (1<sup>st</sup> cut to preselect the  $q\bar{q}$  signals).





ISR events

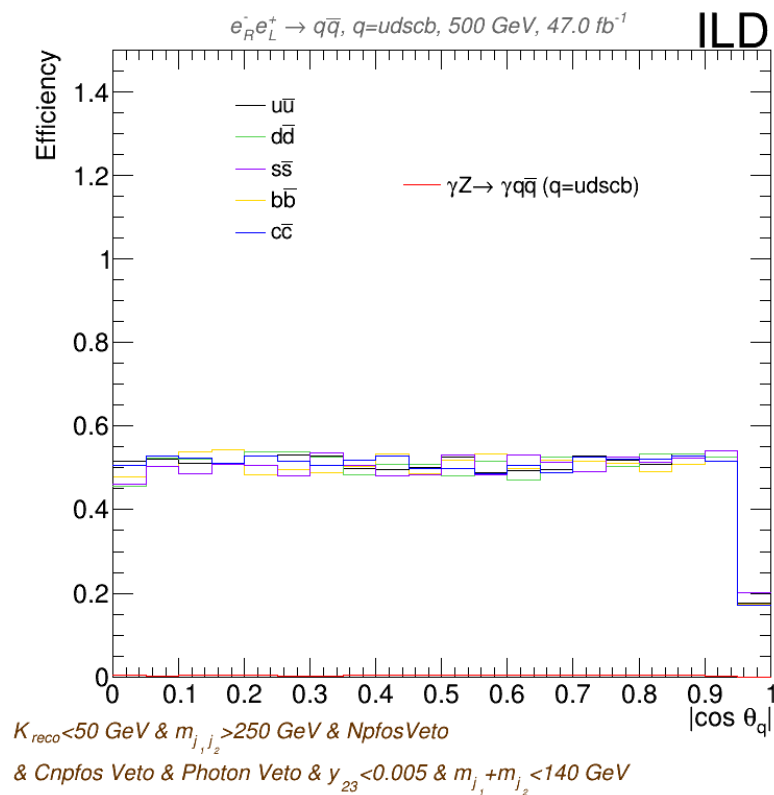
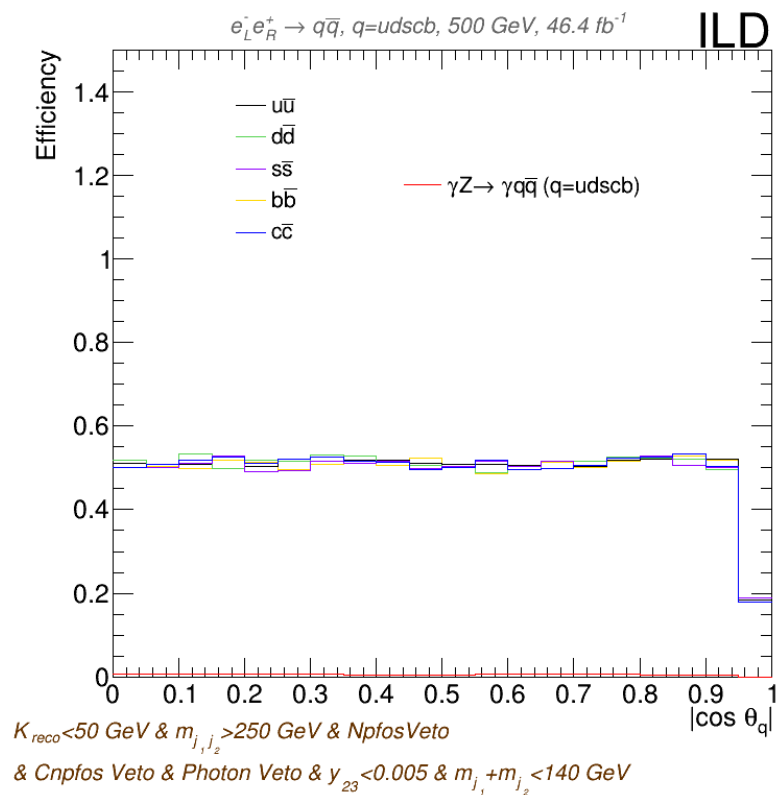


$E_{\text{ISR}} > 50 \text{ GeV}$

- Energy vs  $\cos(\theta)$  distribution of neutral PFOs (photon ISR candidates), identified with PandoraPFO. Useful for ISR removal (1<sup>st</sup> cut to preselect the  $q\bar{q}$  signals).

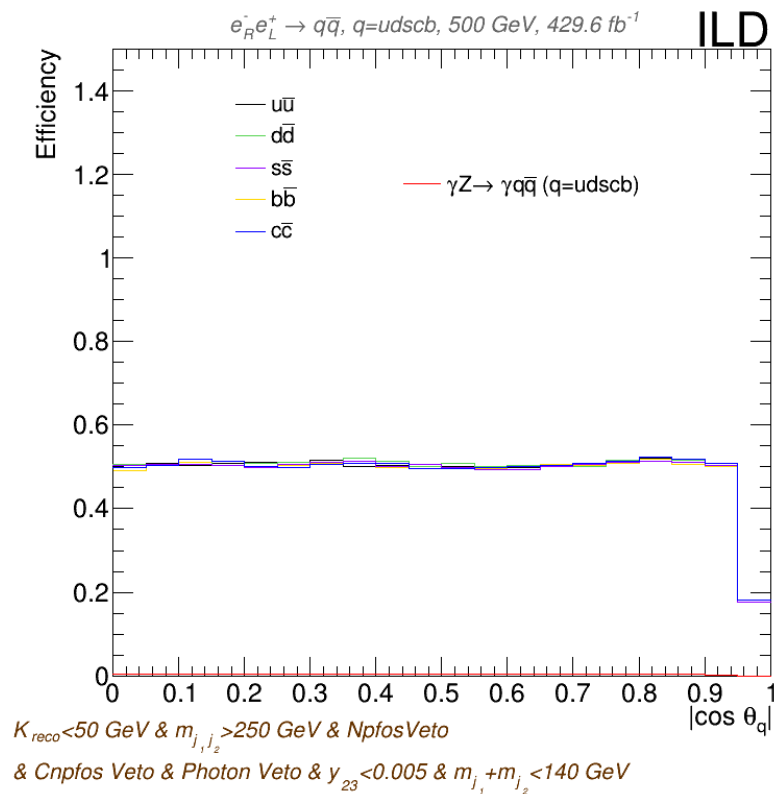
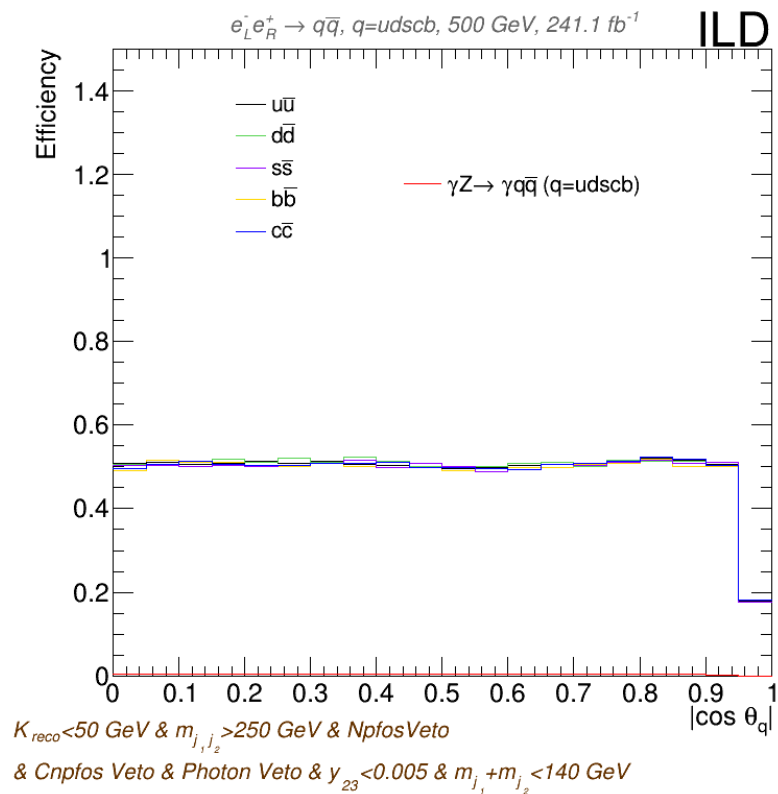


# High Level Reco: Signal preselection

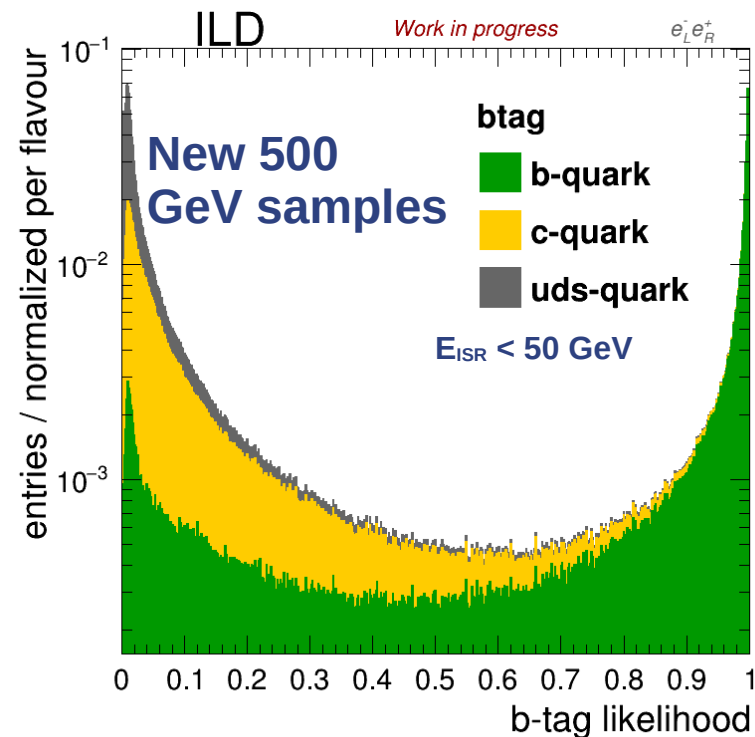
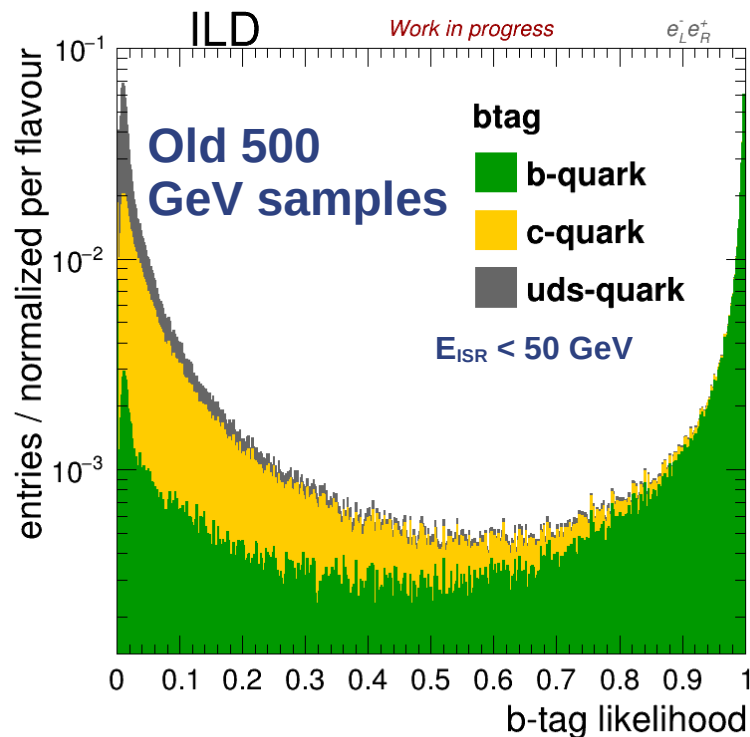


Old 500 GeV Samples

# High Level Reco: Signal preselection

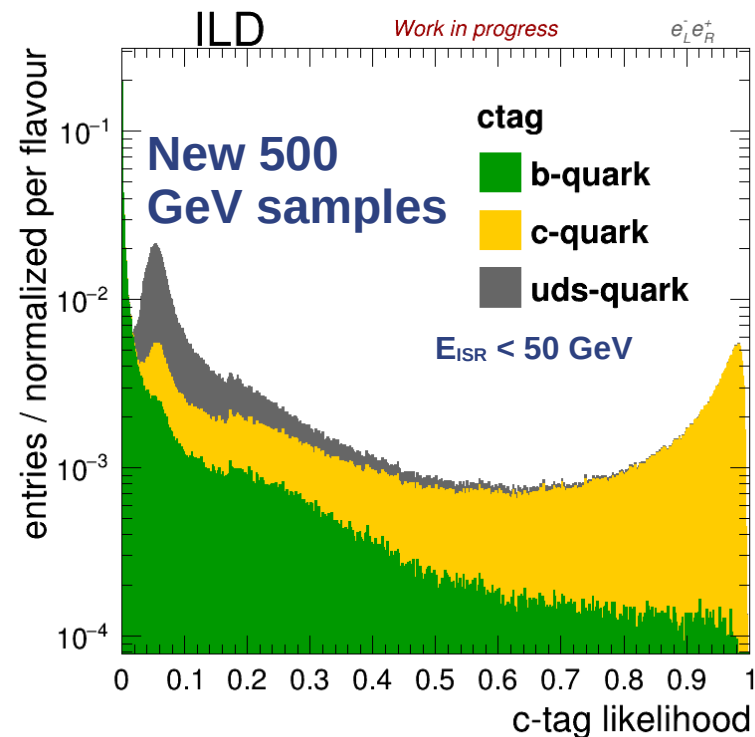
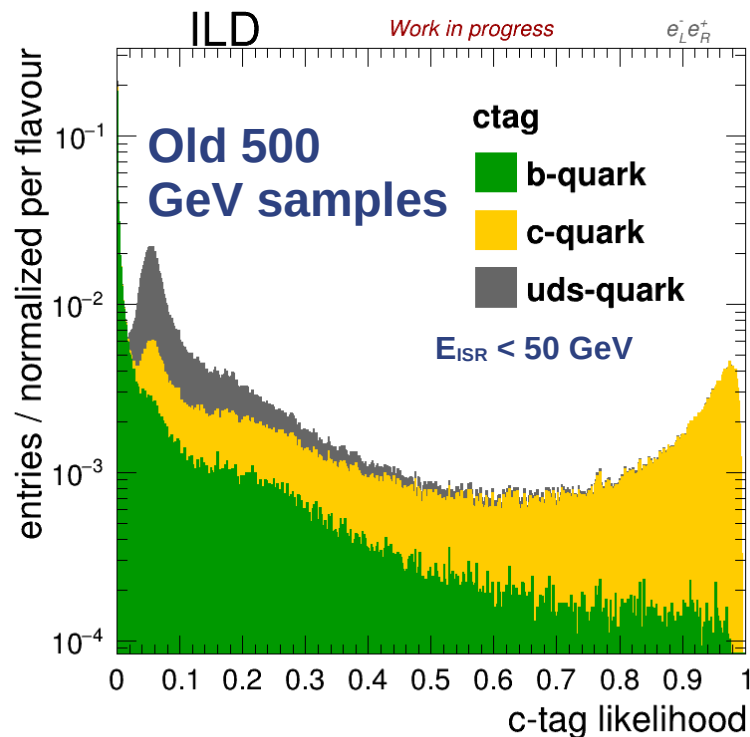


**New 500 GeV Samples**



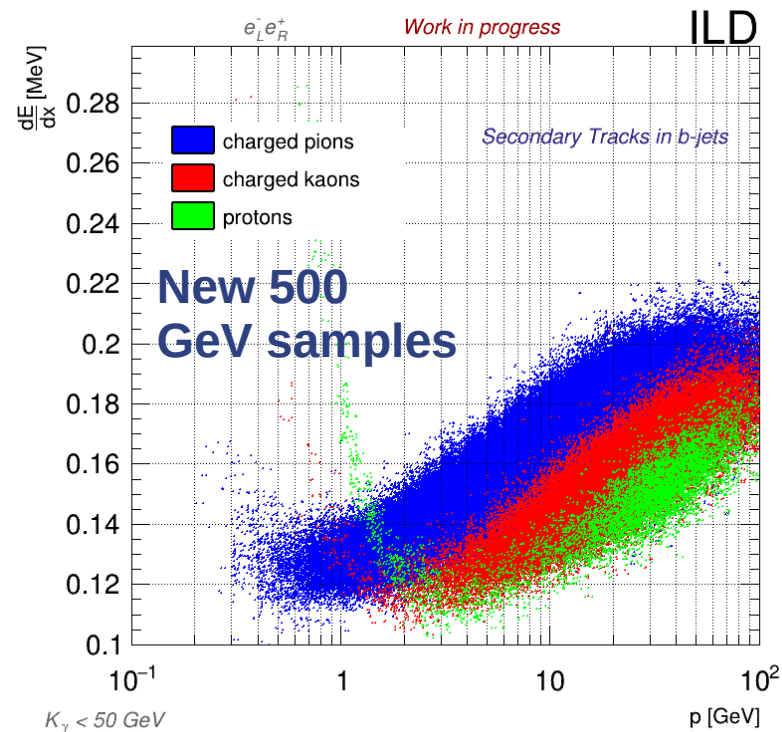
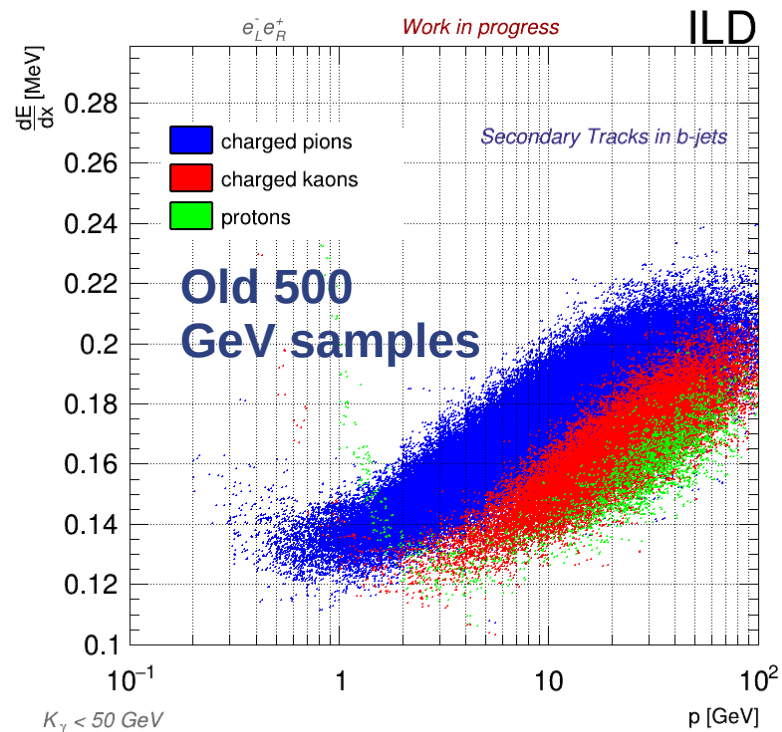
- Weight files: 2q250\_v04\_p00\_ildl5 - files
- VTX files: d0probv2\_ildl5\_2q250.root (and z0)



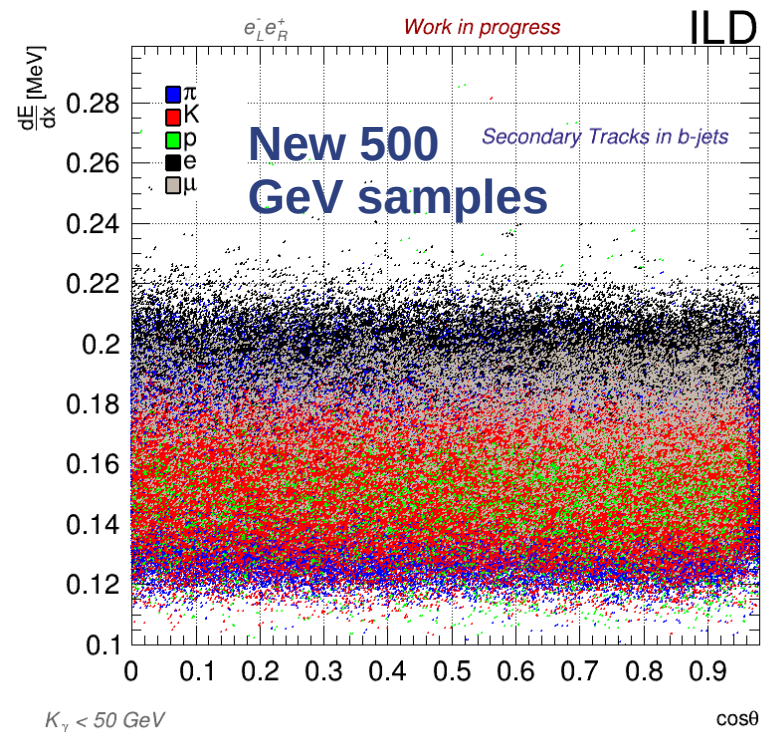
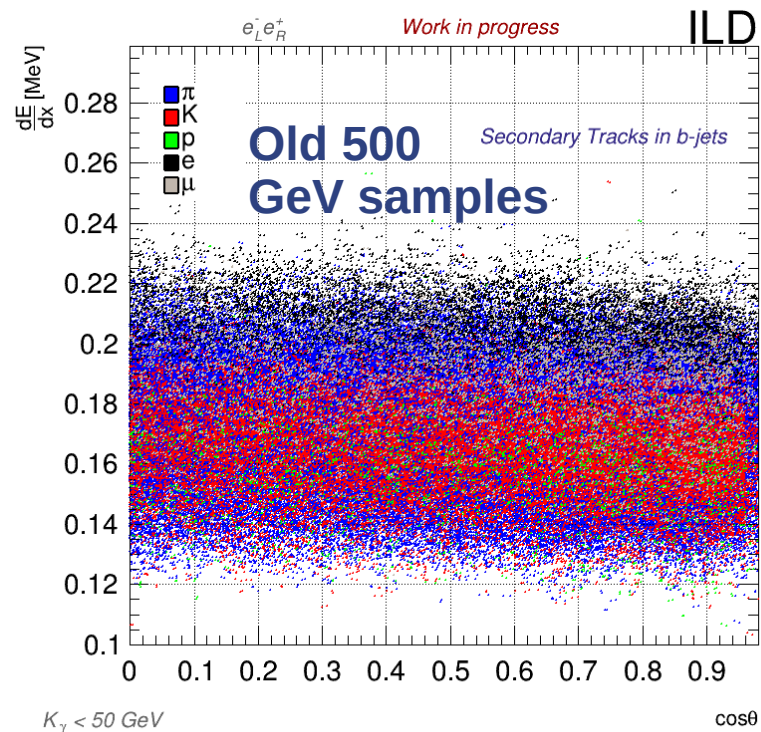


- Weight files: 2q250\_v04\_p00\_ildl5 - files
- VTX files: d0probv2\_ildl5\_2q250.root (and z0)

# High Level Reco: dEdx



# High Level Reco: dEdx



- Samples study and testing:
  - More high level reconstruction studies:
    - Particle ID for Kaons, Pions and protons.
      - Including dE/dx & TOF
  - Look for the full performance of the Preselection process (and refine it).
- Flavour tagging:
  - Use physical samples from the new simulation to get new FT weight files.
  - Working on a high-level optimization for the LCFI+ weight files:
    - Test to avoid overtraining when running ROOT's TMVA for FT.
    - Optimize for the different LCFI+ categories separately.
    - Improve the performance of the BDTs.

**PSO (next slide)**



- Particle Swarm Optimization is a Gradient-free, bio-inspired, stochastic, population-based algorithm to optimize any kind of process towards a certain goal:
  - No maths involved in the optimization (no gradients or loss function used).
  - It just keeps trying configurations and saves the best-performing one.
- How it works:
  - We have N “particles”, in our case: configurations of the BDT. Then:
    - 1) The BDT runs with the configuration of the particle.
    - 2) When finished, each particle gets a performance score.
      - We define a Function Of Merit (FOM) for this scoring.
      - And we test for overfitting.
    - 3) We track each particle’s best configuration and the best global one.
    - 4) The particles moves to a new configuration, approaching a better one using the previous results.

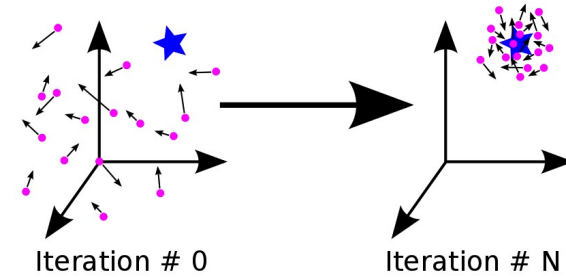


Image taken from a [website](#)

**For each iteration**

**After N iterations we will have optimal classification while avoiding overfitting**

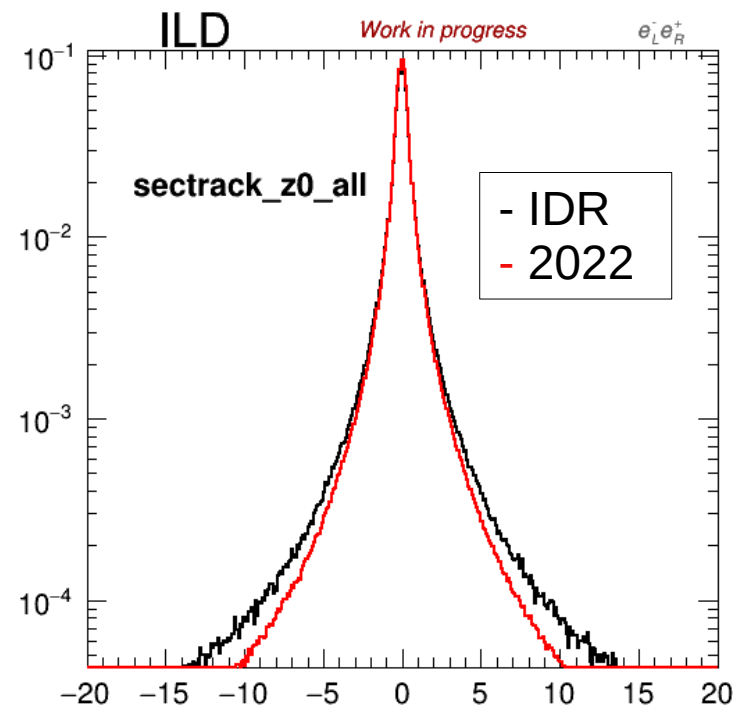
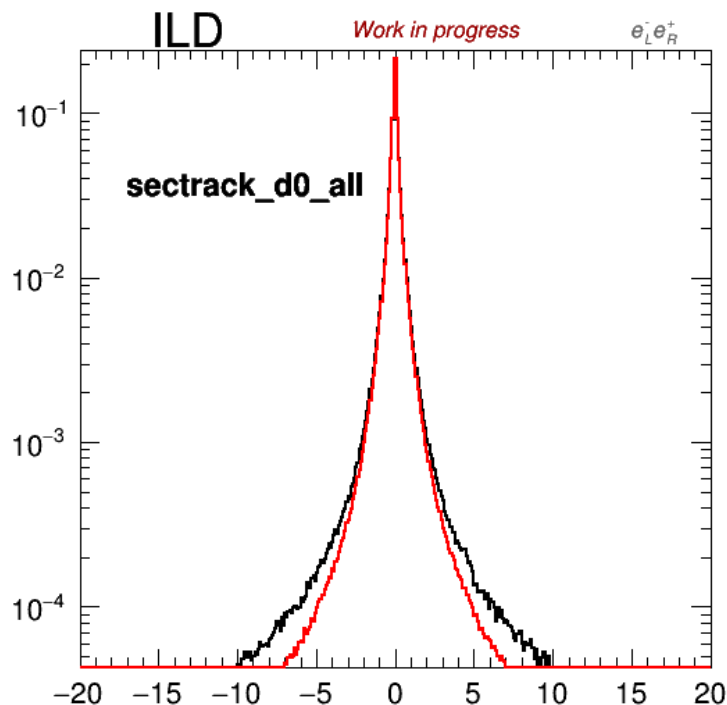


---

# Back-up

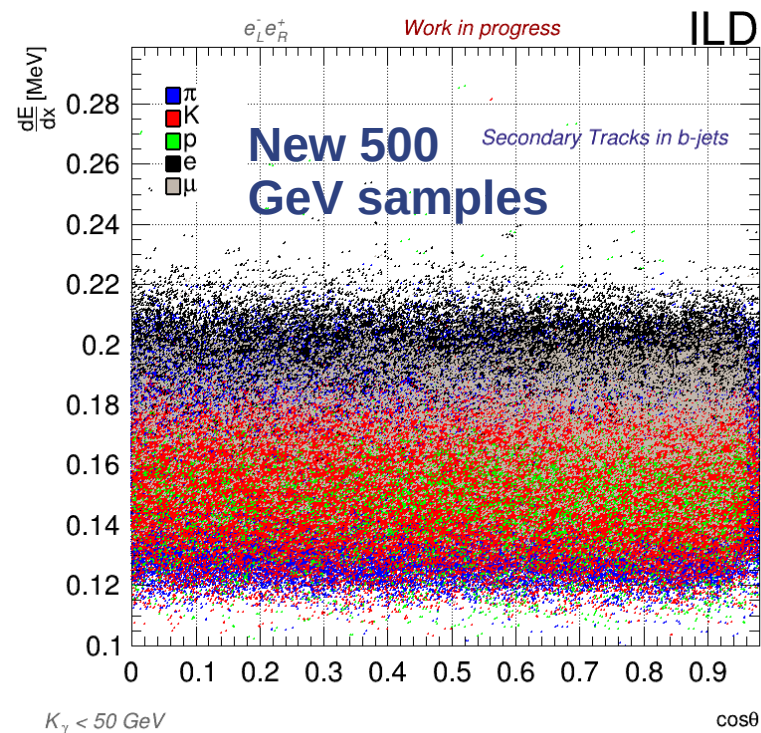
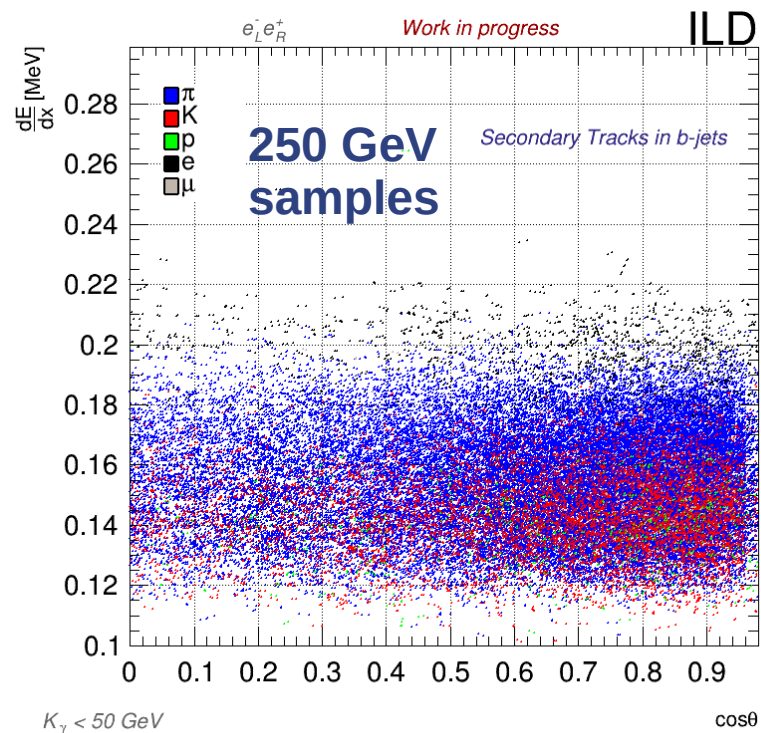
---





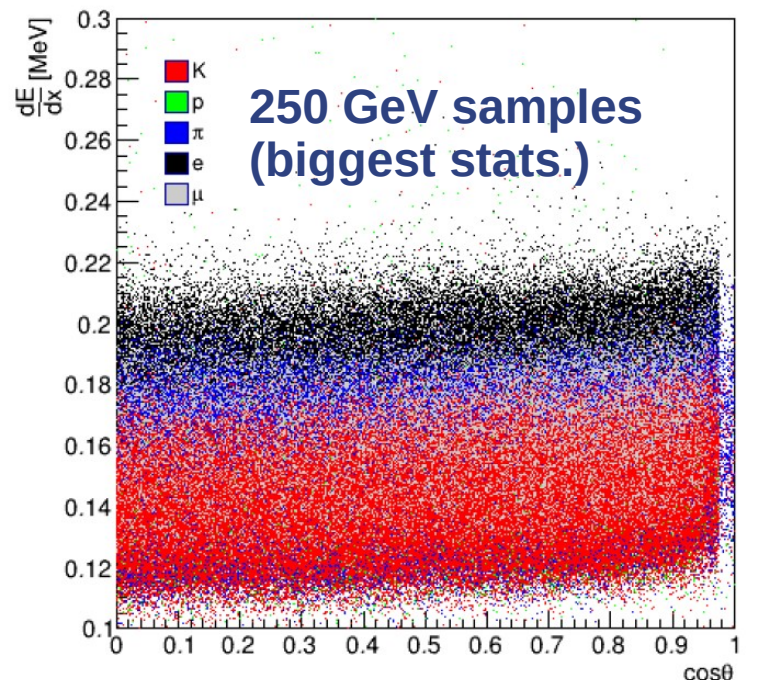
- Secondary tracks slightly more collimated in the new samples.



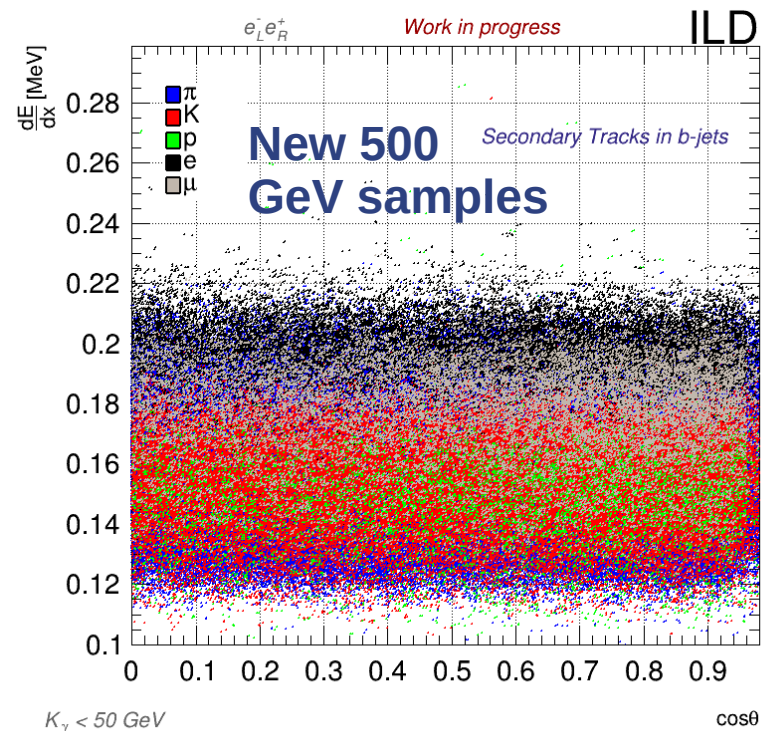


- Angular dependence observed in 250 GeV not present with 500 GeV samples





Plots from Y. Okugawa



- Angular dependence observed in 250 GeV not present with 500 GeV samples



# Old preselection for $K_{reco} < 50$ GeV ( $e_L p_R$ )

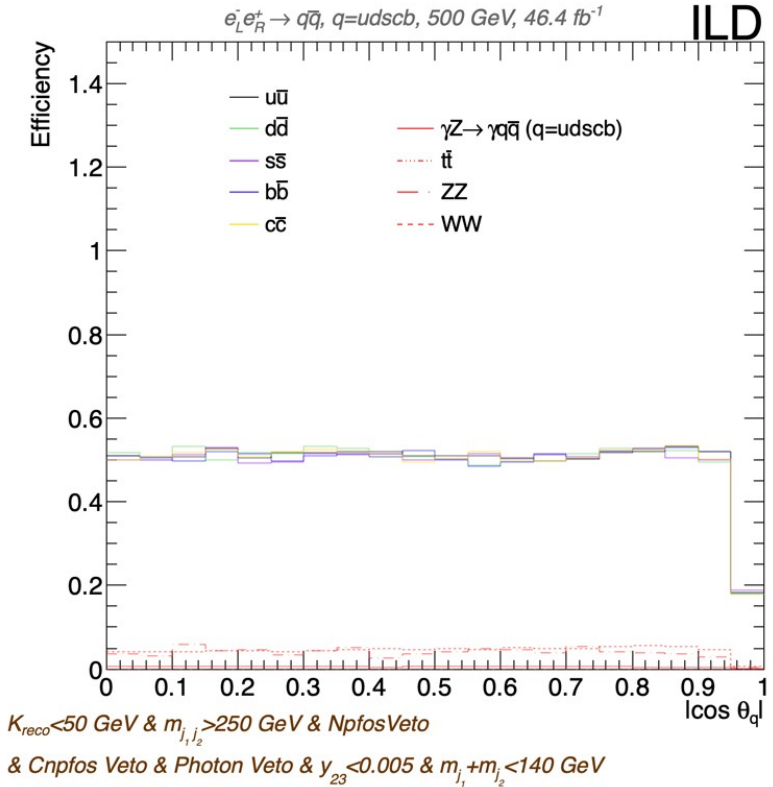
## Cuts:

- $K_{reco} < 50$  GeV
- $m_{2jets} > 250$  GeV
- Charged N pfos  $> 0.5$
- Neutral N pfos  $> 3.5$
- Photon veto
- $y_{23} < 0.005$
- $m_{j1} + m_{j2} < 140$  GeV

## VLC Algorithm parameters:

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

	Efficiency (%)			Background/Signal			
	$b\bar{b}$	$c\bar{c}$	$q\bar{q}$	ISR	WW	ZZ	$t\bar{t}$
No cut	100	100	100	3.50	1.06	0.09	0.10
+ Cut 1	74.9	74.7	74.7	0.76	0.77	0.06	0.01
+ Cut 2	74.8	74.6	74.7	0.74	0.77	0.06	0.01
+ Cut 3	74.8	74.5	74.3	0.16	0.77	0.06	0.01
+ Cut 4	74.7	74.5	74.1	0.11	0.77	0.06	0.01
+ Cut 5	72.1	71.7	71.1	0.05	0.58	0.05	0.01
+ Cut 6	49.6	49.7	49.6	0.03	0.09	0.01	1e-04
+ Cut 7	48.6	48.7	48.7	0.02	0.06	5e-03	5e-06



Old 500 GeV samples

# 1<sup>st</sup> test preselection for $K_{reco} < 50$ GeV ( $e_L p_R$ )

Cuts:

- $K_{reco} < 50$  GeV
- $m_{2jets} > 250$  GeV
- Charged N pfos  $> 0.5$
- Neutral N pfos  $> 3.5$
- Photon veto
- $y_{23} < 0.005$
- $m_{j_1} + m_{j_2} < 140$  GeV

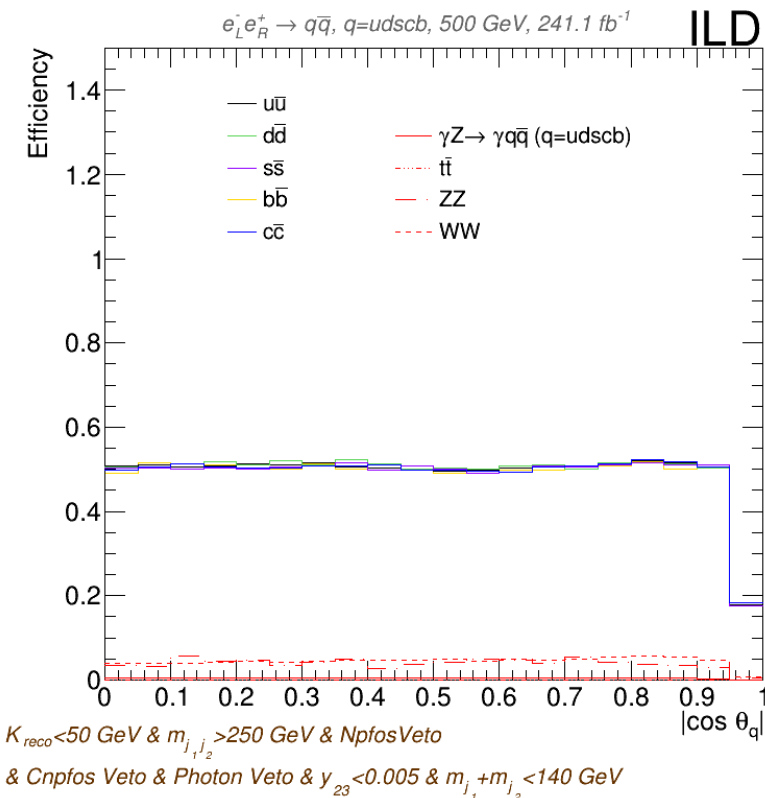
VLC Algorithm parameters:

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

	Efficiency (%)			Background/Signal			
	$b\bar{b}$	$c\bar{c}$	$q\bar{q}$	ISR	WW	ZZ	$t\bar{t}$
No cut	100	100	100	5.43	1.56	0.14	0.14
+ Cut 1	75.0	75.2	75.1	1.23	1.12	0.08	0.01
+ Cut 2	75.0	75.2	75.1	1.21	1.12	0.08	0.01
+ Cut 3	74.9	75.1	74.9	0.22	1.12	0.08	0.01
+ Cut 4	74.7	74.8	74.7	0.15	1.12	0.08	0.01
+ Cut 5	72.3	72.4	71.8	0.07	0.84	0.07	0.01
+ Cut 6	49.0	49.5	48.9	0.03	0.13	0.01	2e-04
+ Cut 7	47.9	48.2	48.3	0.03	0.09	7e-03	8e-06

Need to rebalance

Jesús P. Márquez Hernández - ILD SW&ANA 26/10/22



New 500 GeV samples

# Old preselection for $K_{reco} < 50$ GeV ( $e_R p_L$ )

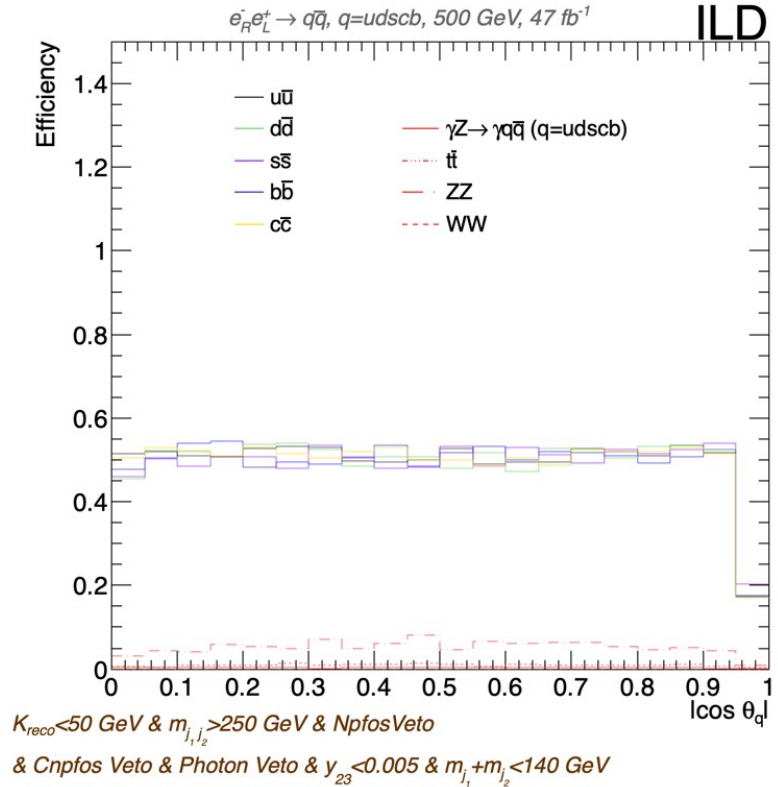
Cuts:

- $K_{reco} < 50$  GeV
- $m_{2jets} > 250$  GeV
- Charged N pfos  $> 0.5$
- Neutral N pfos  $> 3.5$
- Photon veto
- $y_{23} < 0.005$
- $m_{j1} + m_{j2} < 140$  GeV (optional)

VLC Algorithm parameters:

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

	Efficiency (%)			Background/Signal			
	$b\bar{b}$	$c\bar{c}$	$q\bar{q}$	ISR	WW	ZZ	$t\bar{t}$
No cut	100	100	100	6.51	0.01	0.11	0.10
+ Cut 1	74.6	74.6	75.0	1.45	0.01	0.07	0.01
+ Cut 2	74.5	74.5	75.0	1.43	0.01	0.07	0.01
+ Cut 3	74.5	74.4	74.7	0.26	0.01	0.07	0.01
+ Cut 4	74.5	74.4	74.5	0.18	0.01	0.07	0.01
+ Cut 5	71.9	71.7	71.5	0.07	0.01	0.06	0.01
+ Cut 6	49.5	49.6	49.6	0.03	5e-04	0.01	9e-05
+ Cut 7	48.5	48.8	48.7	0.03	3e-04	8e-03	3e-06



Old 500 GeV samples

# 1<sup>st</sup> test preselection for $K_{reco} < 50$ GeV ( $e_R p_L$ )

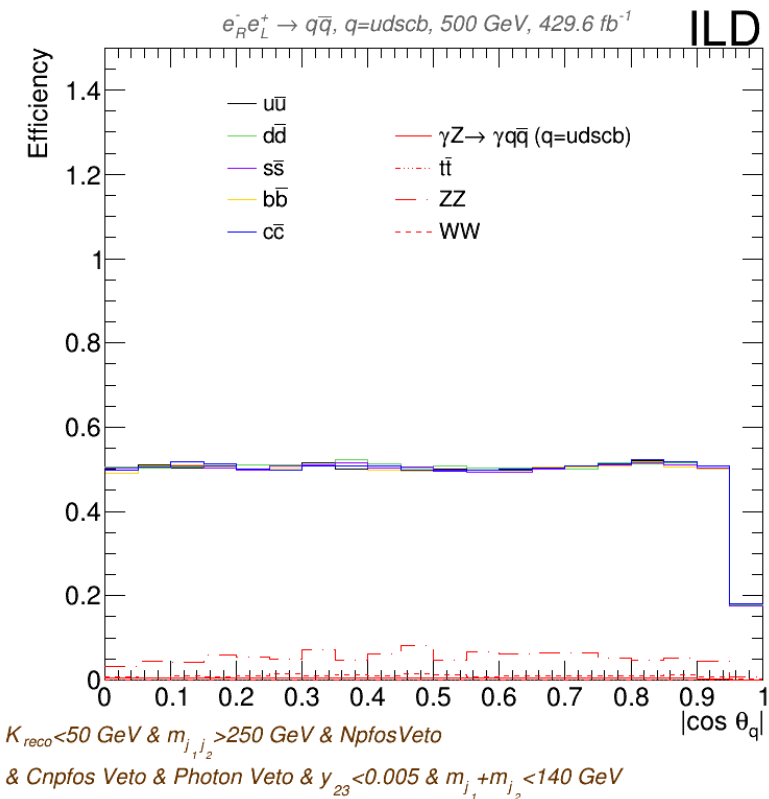
Cuts:

- $K_{reco} < 50$  GeV
- $m_{2jets} > 250$  GeV
- Charged N pfos  $> 0.5$
- Neutral N pfos  $> 3.5$
- Photon veto
- $y_{23} < 0.005$
- $m_{j_1} + m_{j_2} < 140$  GeV (optional)

VLC Algorithm parameters:

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

	Efficiency (%)			Background/Signal			
	$b\bar{b}$	$c\bar{c}$	$q\bar{q}$	ISR	WW	ZZ	$t\bar{t}$
No cut	100	100	100	4.78	0.01	0.09	0.08
+ Cut 1	75.2	75.2	75.1	1.08	0.01	0.06	0.01
+ Cut 2	75.1	75.2	75.1	1.06	0.01	0.06	0.01
+ Cut 3	75.1	75.1	74.8	0.20	0.01	0.06	0.01
+ Cut 4	75.1	75.1	74.6	0.14	0.01	0.06	0.01
+ Cut 5	72.5	72.3	71.8	0.06	0.01	0.05	0.01
+ Cut 6	49.1	49.4	49.2	0.03	4e-04	0.01	7e-05
+ Cut 7	48.0	48.3	48.2	0.03	2e-04	6e-03	3e-06



**New 500 GeV samples**