# Production of $q\bar{q}$ at 250 GeV: optimization of the cuts against ISR radiation

Adrián Irles, R. Pöschl, F. Richard
**IJCLab CNRS/IN2P3** (former LAL)

# Goal

- Measure EW observables at the **per mil level of experimental precision**
  - We want to measure differential distributions
  - Ratios are optimal to remove (or reduce) systematic uncertainties
  - These measurements will be done in the continuum *(cont.)* → i.e. Far from the Z-Pole

$$R_q^{cont.}(|cos\theta_q|) = \frac{\sigma^{cont.}_{e^-e^+\to q\bar{q}}(|\cos\theta|)}{\sigma^{cont.}_{had.}(|\cos\theta|)}.$$

$$A_{FB}^{q\bar{q}} = \frac{\sigma^{cont.,F}_{e^-e^+\to q\bar{q}} - \sigma^{cont.,B}_{e^-e^+\to q\bar{q}}}{\sigma^{cont.,F}_{e^-e^+\to q\bar{q}} + \sigma^{cont.,B}_{e^-e^+\to q\bar{q}}}$$

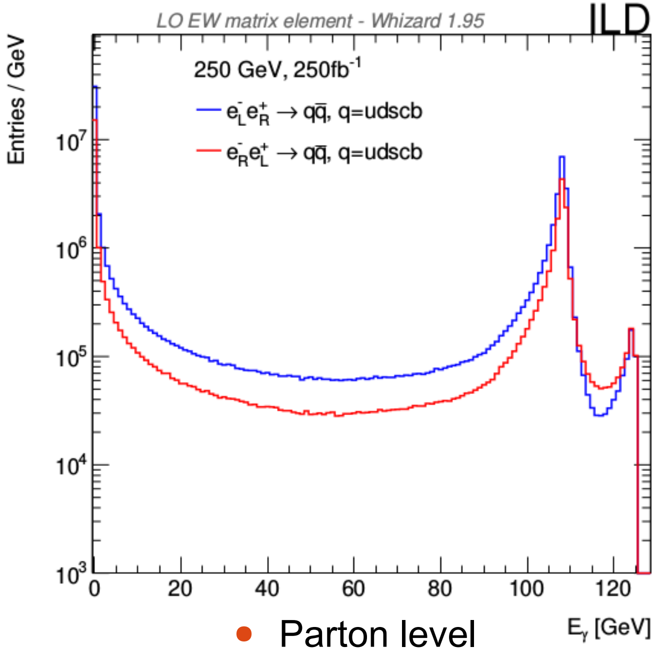
- For Rb main sources of uncertainty are the b-tagging efficiency and the normalisation (this talk)

# Predictions (as a function of the ISR)

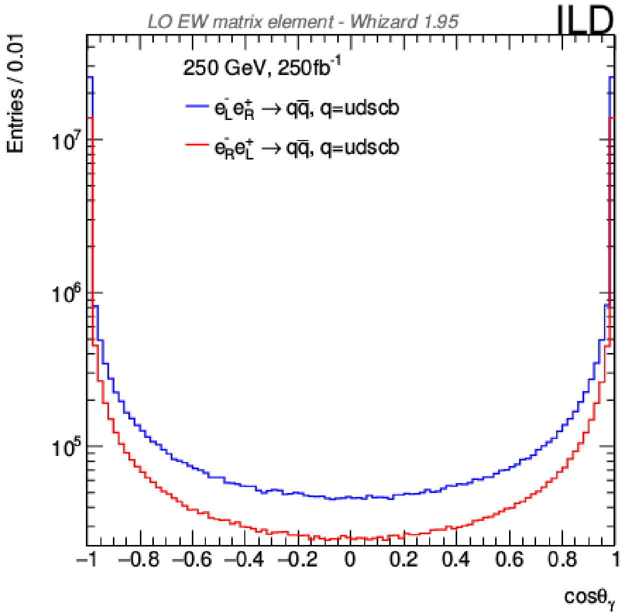
- The cross section depends on the "effective" center of mass energy
  - At which the Z/y couple to the quark-antiquark pair

$$\frac{d\sigma^{cont.}_{e^-e^+\to q\bar{q}}}{d\cos\theta_q}(s) \to$$

$$\to \frac{d\sigma^{cont.}_{q\bar{q}}}{d\cos\theta_q}(\hat{s} > s_{cut}) = \frac{d\sigma^{cont.}_{e^-e^+\to q\bar{q}}}{d\cos\theta_q}(E_\gamma < K_{cut})$$



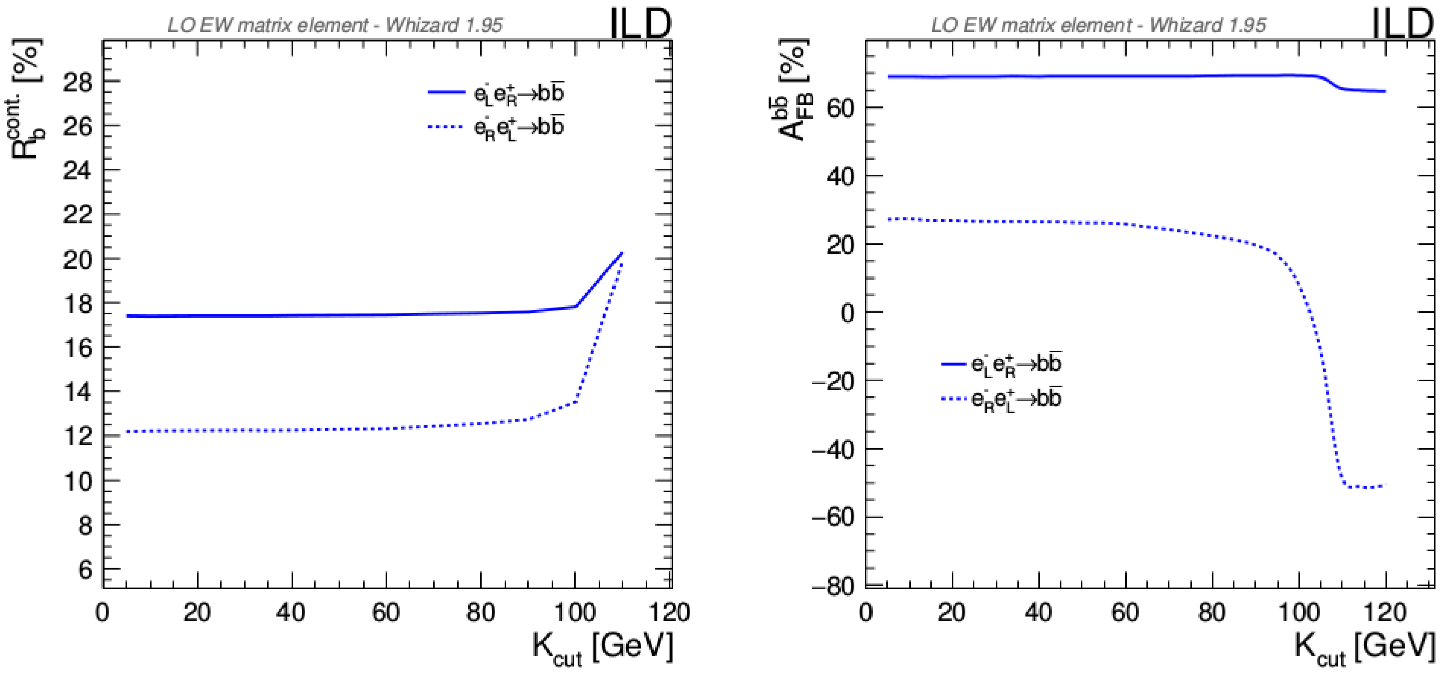

- Parton level




- Parton level

# Predictions (as a function of the ISR)

$$\frac{d\sigma^{cont.}_{e^-e^+\rightarrow q\bar{q}}}{d\cos\theta_q}(E_\gamma < K_{cut})$$




- The observables remain basically flat for a large range of the Kcut
- Drastic change when the photn ISR is large enough to produce a return to the Z-pole
  - We need to avoid that region of the phase space.

# Double Tag Method *(à la Z-pole, LEP/SLC)*

- The sample consisted on events made of two hadronic jets (qqbar)
  - The LEP/SLC preselection consisted on a "simple" veto of Z→ leptons events
- The method is based on the comparison of single vs double tagged samples → To remove Luminosity dependence. To remove modelling dependence on the efficiency of b-tagging
  - f1= ratio of jets that are tagged as b-jets
  - f2= ratio of events in which both jets are tagged as b-jets

$$f_1 = \epsilon_b R_b + \epsilon_c R_c + \epsilon_{uds} R_{uds}$$
$$f_2 = \epsilon_b^2 (1+\rho_b) R_b + \epsilon_c^2 R_c + \epsilon_{uds}^2 R_{uds}$$

- Rb and $\varepsilon_b$ are measured simultaneously.

- **$\varepsilon_b$ = b-tagging efficiency**
- $\rho_b$ = b-tagging correlation factor
- $\varepsilon_c$ = probability of tagging a c-quark jet as b-jet
- $\varepsilon_{uds}$ = probability of tagging an uds-quark jet as b-jet

These values must be as small as possible and with small uncertainties

to not spoil our accuracy (not covered in this talk)

# Double Tag Method *(in the continuum)*

- When we don't run at the Z-pole → **contamination from other processes**

$$f_1^{250} = \frac{\sum_{q=udscb} \epsilon_{q\bar{q}} \epsilon_q \left(\sigma_{q\bar{q}}^{cont.} + \sigma_{q\bar{q}}^{others}\right)}{\sum_{q=udscb} \epsilon_{q\bar{q}} \left(\sigma_{q\bar{q}}^{cont.} + \sigma_{q\bar{q}}^{others}\right)}$$

$\varepsilon_{q\bar{q}}$ = Preselection efficiency for qq events (q=udcsb)

$$f_2^{250} = \ldots$$

- **The first source of contamination** are the qq bar events produced together with photon **ISR (radiative return)**.
  - with **~3-10 times larger cross section than the signal**
  - Very different Rb and AFBb values
- **The second source of contamina**tion are the background events from completely different physical processes
  - **WW/HZ/ZZ → with ~0.5 times the signal cross section**

# Double Tag Method *(in the continuum)*

- When we don't run at the Z-pole → contamination from other processes

$$f_1^{250} = \frac{\sum_{q=udscb} \epsilon_{q\bar{q}}\, \epsilon_q \left(\sigma_{q\bar{q}}^{cont.} + \sigma_{q\bar{q}}^{others}\right)}{\sum_{q=udscb} \epsilon_{q\bar{q}} \left(\sigma_{q\bar{q}}^{cont.} + \sigma_{q\bar{q}}^{others}\right)}$$

$\varepsilon_{q\bar{q}}$ = Preselection efficiency for qq events (q=udcsb)

$$f_2^{250} = \ldots$$

- **The definition of the rations at 250 will match the on at the pole ( $f_{1/2}^{250} = f_{1/2}$ ) if:**
  - $\varepsilon_{q\bar{q}} = \varepsilon_{b\bar{b}} = \varepsilon_{c\bar{c}} = \varepsilon_{\overline{uds,uds}}$
  - **BKG** contribution is negligible
- If not, these factors will have to be modeled by MC and/or data driven methods
  - Challenges/spoils the goal of the per mile in the accuracy !

**Our goal is to define a preselection procedure that fulfills these conditions !**

# Preselection (Introduction):

- We need to understand the topology of the radiative return events to remove them as efficiently as possible
- Technical Issue:
  - The nominal sample (~250fb-1) has a cut at generator level $m_{qq}$ > 150 GeV

    *rv01-16-p10_250.sv01-14-01-p00.mILD_o1_v05.E250-TDR_ws.I110011.P2f-highM_z_h.eL.pR_dst_7637_XXXXX_DST.slcio*
    *rv01-16-p10_250.sv01-14-01-p00.mILD_o1_v05.E250-TDR_ws.I110012.P2f-highM_z_h.eR.pL_dst_7638_XXXXX_DST.slcio*

  - There is a second sample with lower statistics (~15fb-1) but includes the radiative return events

    *rv01-17-11-p02.sv01-14-01-p00.mILD_o1_v05.E250-TDR_ws.I106607.P2f_z_h.eL.pR_dst_00008992_XXXX-DST.slcio*
    *rv01-17-11-p02.sv01-14-01-p00.mILD_o1_v05.E250-TDR_ws.I106608.P2f_z_h.eR.pL_dst_00008992_XXXX-DST.slcio*

- For now on, we use the small sample for the design of the preselection procedure.
  - We will compare with the nominal one at the end.
  - The “chronological” approach was the opposite...

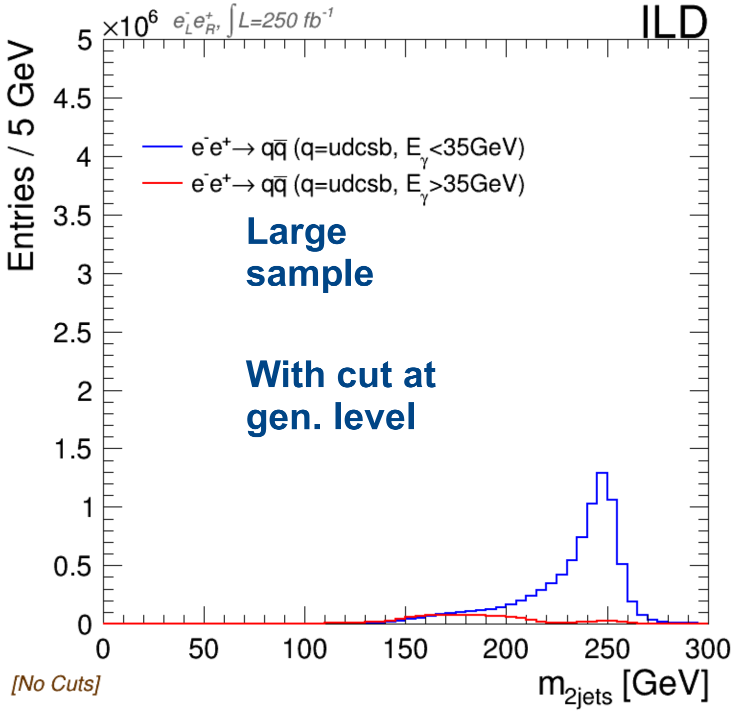
# Preselection (Introduction 2):

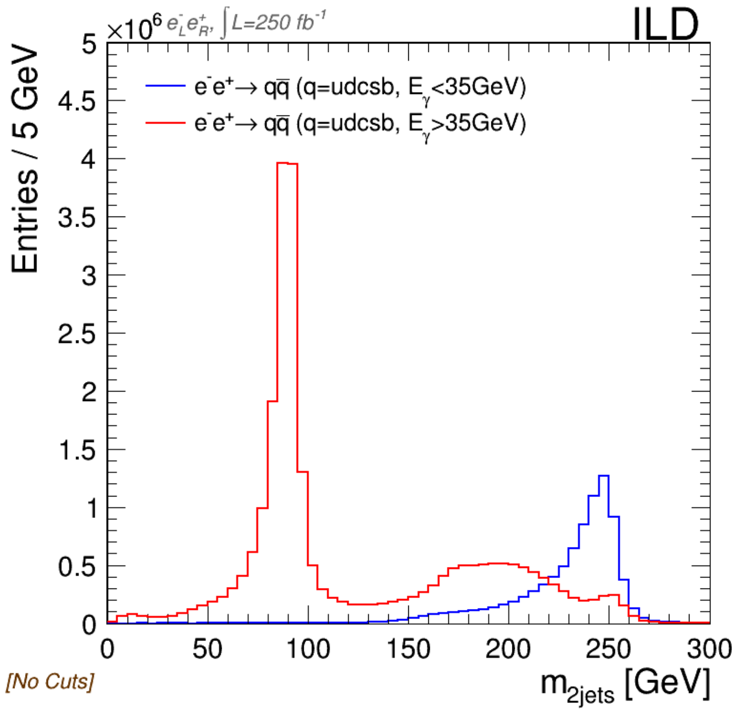
- We define our **signal with $K\gamma<35$ GeV** and the **radiative return with $K\gamma>35$ GeV**
  - parton level definition!
- Our first reconstruction step is to cluster the event in two exclusive jets
  - Generalized ee-kt algorithm

- The simplest variable for the removal of radiative return events would be the invariant mass.

# Preselection (Introduction 2):

- We define our **signal with Kγ<35 GeV** and the **radiative return with Kγ>35 GeV**




- According to the large sample, a simple cut on m(2jets)>180 GeV would be enough.

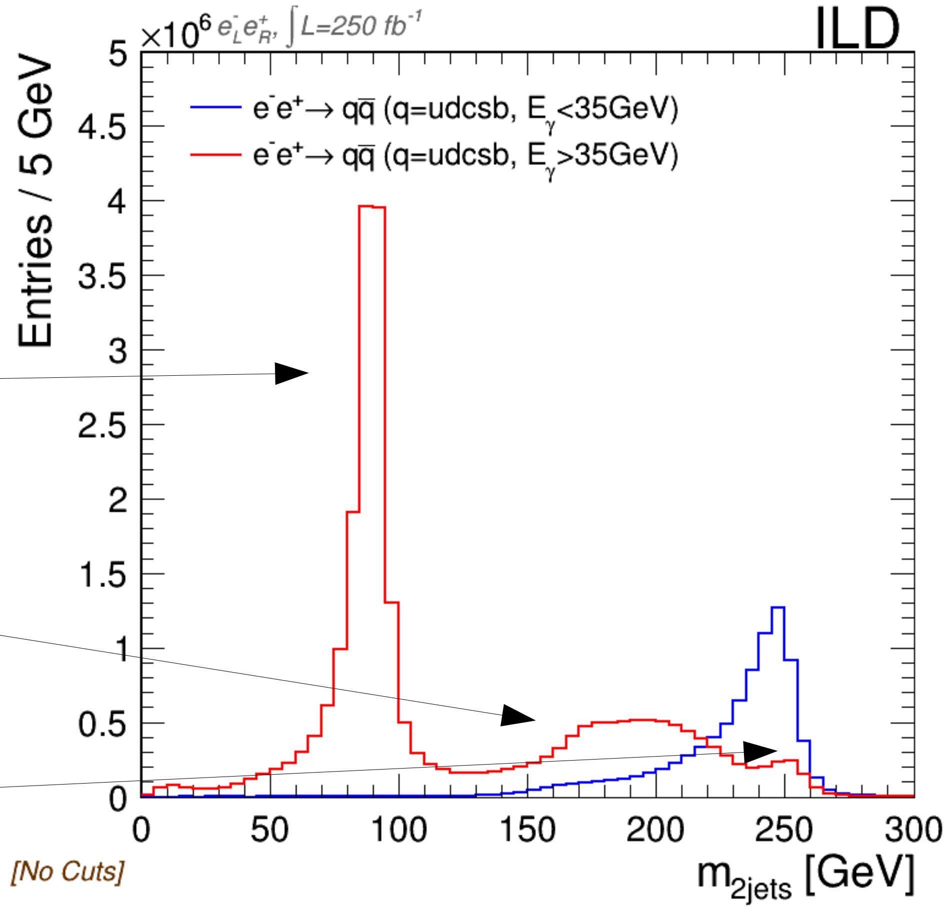
# Preselection (Introduction 2):

- We define our **signal with Kγ<35 GeV** and the **radiative return with Kγ>35 GeV**

highly energetic photon ISR escaping via the beam pipe

one quark is emitted in direction of the beam pipe and the ISR is partially reconstructed inside the detector.

the photon ISR is inside the detector and reconstructed together the two quark jets




- The cut on m(2jes)>180 GeV would need to be increased to ~220 GeV

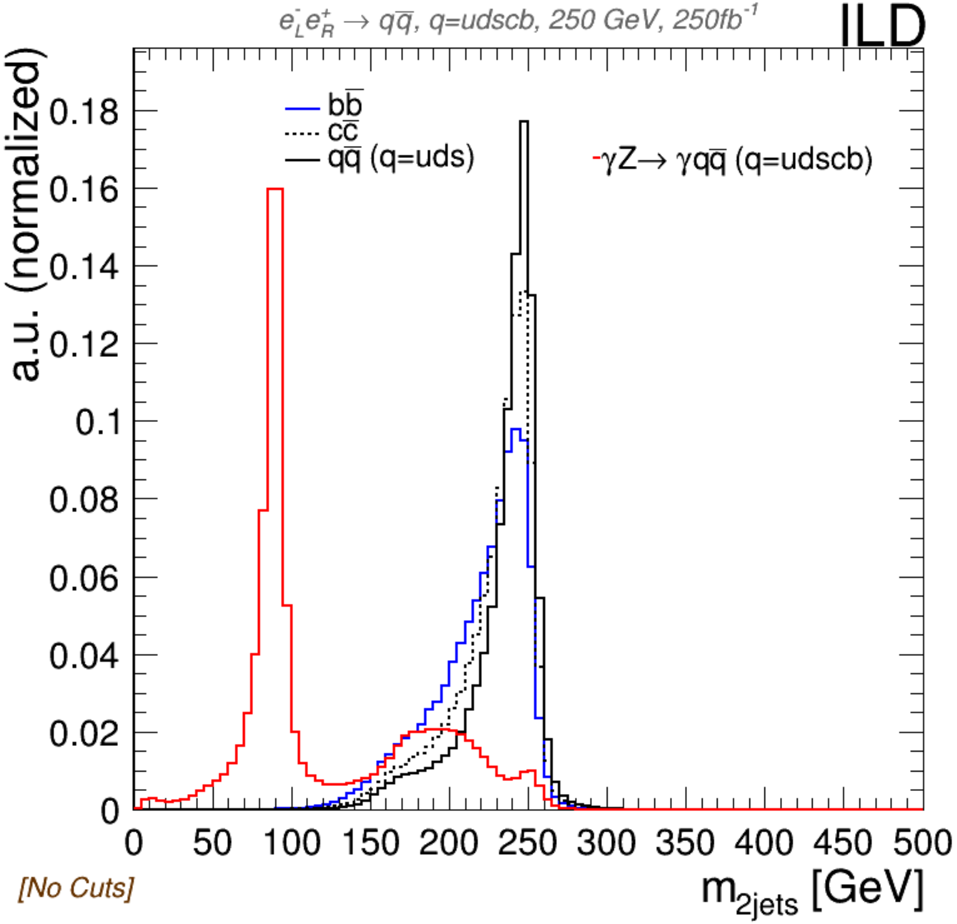
# Preselection (Introduction 2):

- Same histograms but normalized to 1

- To remove a large fraction of radiative return events we would need a cut on
  - I.e m(2jets)>220 GeV (or larger)
  - This cut would introduce large differences between flavors

    *Due to the presence of neutrinos in the hadronization/decay process.*

- *We will cut only at m(2jets)=130 GeV.*




[No Cuts]

# Preselection

- Alternatives to m(2jets) ?
- Estimator of the energy of the photon ISR using only the two reconstructed jets.
  - From momentum conservation (if the photon/s are emitted parallel to the beam pipe):

$$|\vec{k}| \approx K_{reco} = \frac{250\,\mathrm{GeV}}{\sin\Psi_{acol} + sin\theta_1 + sin\theta_2}.$$

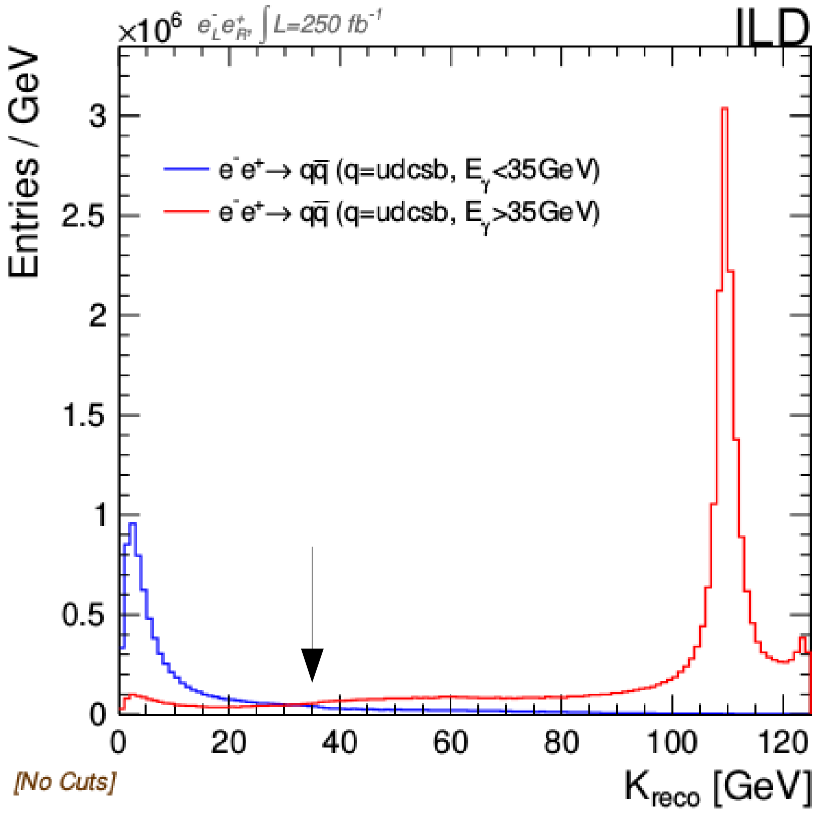
Two jet acolinearity

Jet angular variables (w.r.t. detector frame)

$$\sin\Psi_{acol} = \frac{\vec{p_{j_1}} \times \vec{p_{j_2}}}{|\vec{p_{j_1}}| \cdot |\vec{p_{j_1}}|}$$

# Preselection : Kreco

- Estimator of the energy of the photon ISR
- We apply a cut of Kreco<35 GeV
- Some signal events have larger Kreco (~15%)
  - Because of detector resolution and double photon ISR
- Some radiative return events have Kreco<35GeV (~7%)
  - Because the photon(s) has not escaped through the beam pipe
- Can we identify the photon clustered in one or both jets and veto these events?

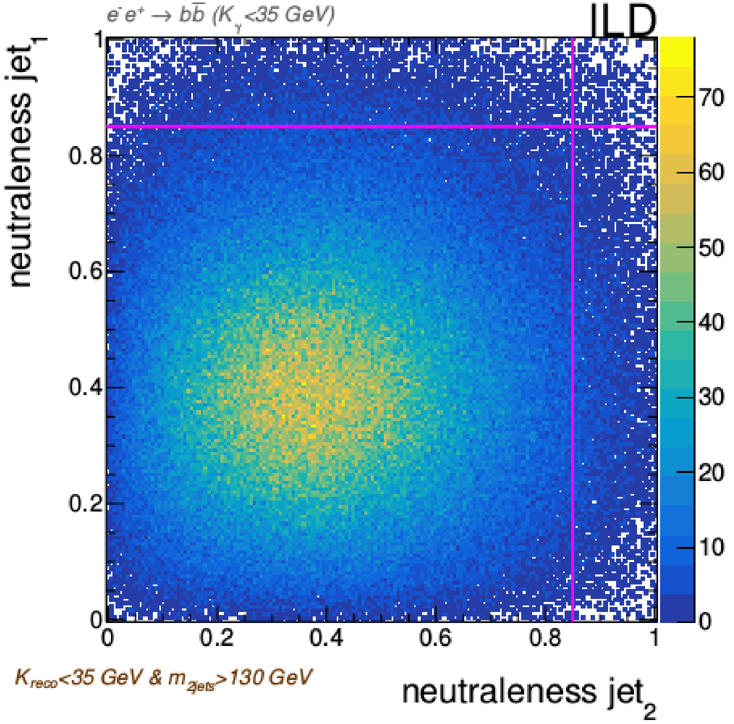


# Preselection : Photon Veto

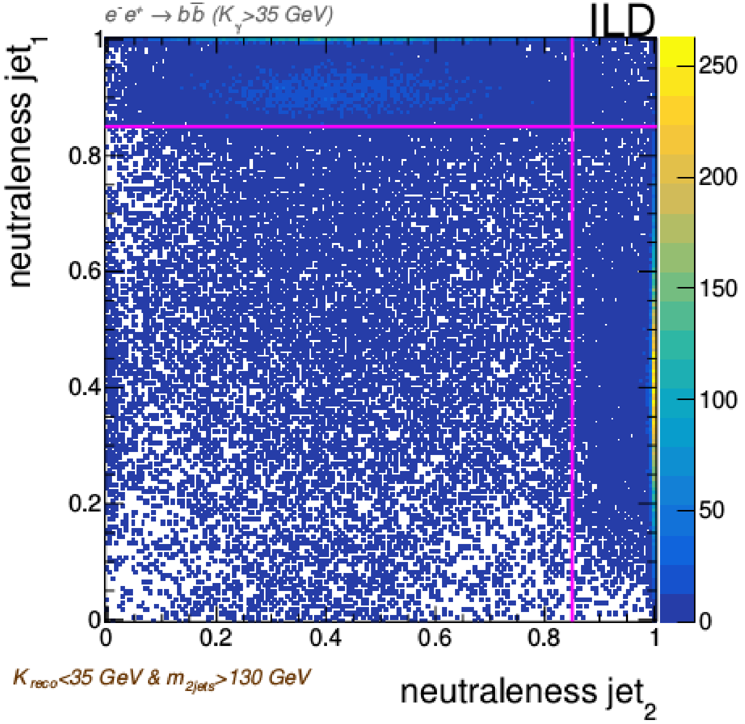
- We look at the neutraleness of the jets

$$\text{neutraleness}_j = \frac{\sum_{i=nPFO} E_i}{E_j}.$$

**signal**




**radiative return**



# Preselection : Photon Veto

- Vetoing events with at least one jet with neutraleness>0.85,
  - The signal efficiency remains almost unchanged
  - The efficiency of mis-selection radiative return events is reduced from the 7% to the 1.5%

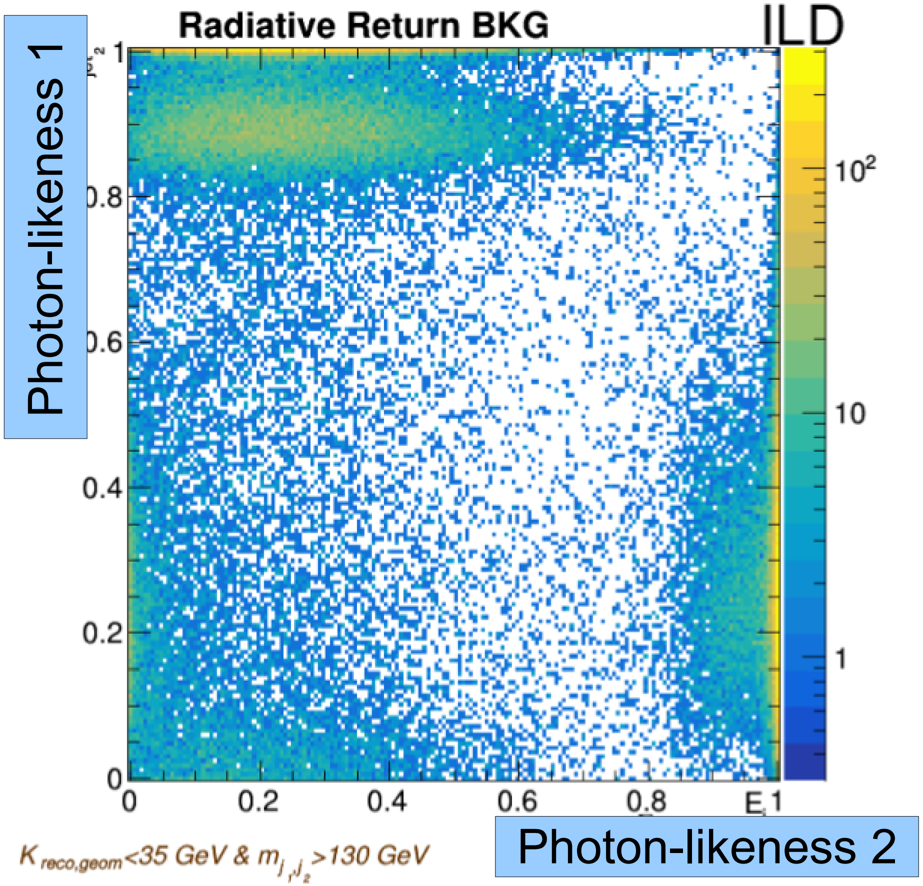
- In addition, we veto:
  - Events with jets with less than 5 PFOs (to veto events with photon conversion)
  - Events with energetic neutral PFOs at very large angles
  - The mis-selection efficiency is reduced to the ~0.5% (Details on the backup slides)

# why not looking at the "photon-likeness" ?

- Similar definition but using only PFOs tagged as photons instead of neutral PFOs

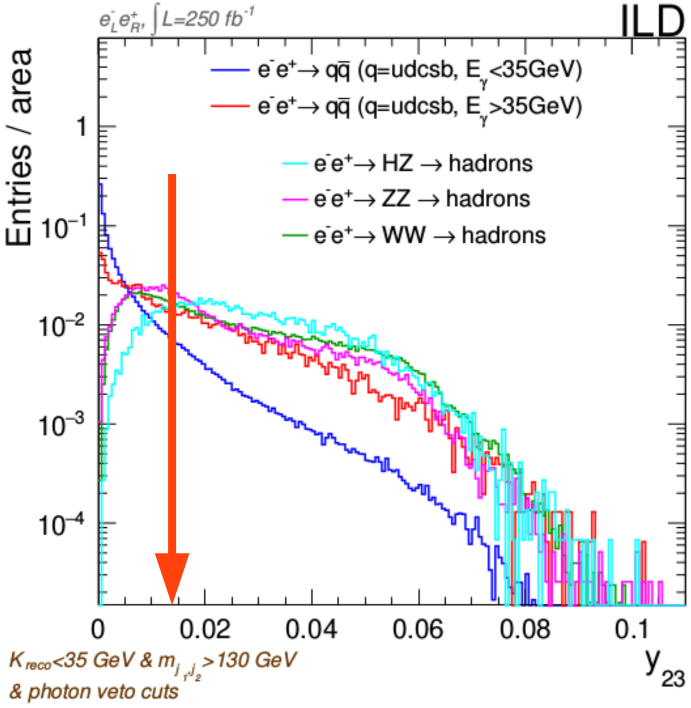
- It shows large inefficiencies
  - ~50% of the events that we remove now would not be removed by vetoing identified photons

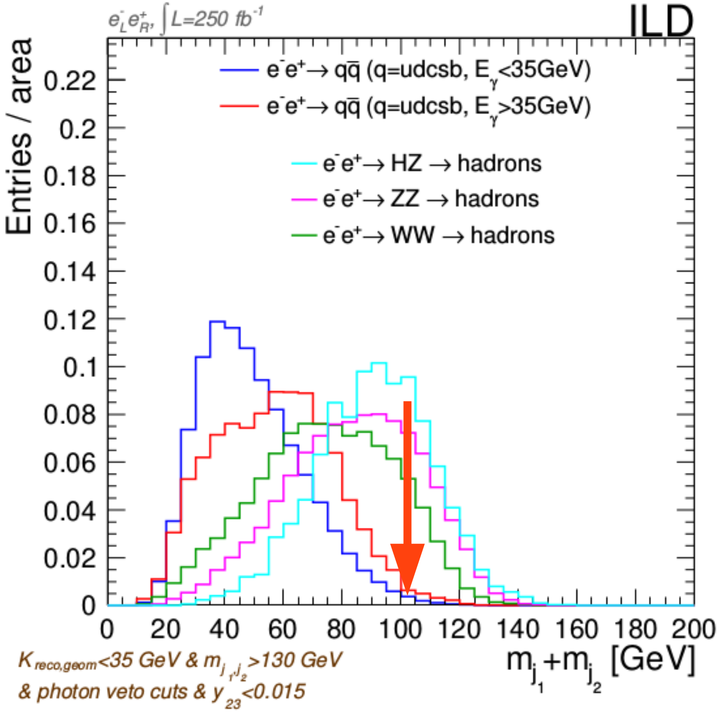
- To be investigated in detail with the new software and samples.



# Final steps of the preselection

- Cut on y23<0.015 (jet distance at which the 2 jet event would be clustered in 3 jets)
- Cut on mj1+mj2<100 GeV



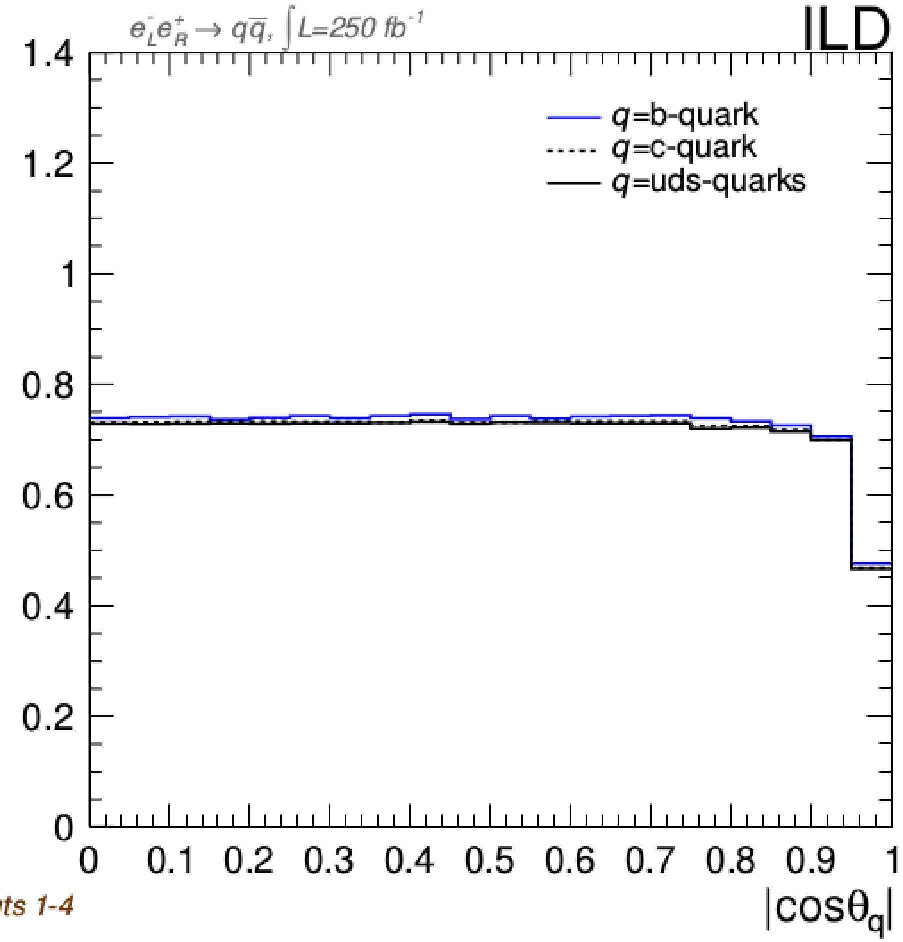



# Preselection summary

- Cut 1:
  - Kreco<GeV & m(2jets)>130 GeV
- Cut 2:
  - Photon veto cuts
- Cut 3:
  - y23<0.015
- Cut 4:
  - mj1+mj2<100 GeV

- What is the preselection efficiency $\varepsilon_{q\bar{q}}$ for each flavour?
  - It is flat in almost all the detector
  - Almost equal for all flavours

$\varepsilon_{q\bar{q}}$




Cuts 1-4

# Final values of the preselection

- Using the "small sample with radiative return included"
  - Efficiency of signal selection of the 71%
  - Radiative return contamination (**B/S) of the 2.8%**
- Using the "nominal large sample with radiative return removed at generator level"
  - Efficiency of signal selection of the 71%
  - Radiative return contamination **(B/S) of the 1.6%** (almost half than expected!)
- B/S for the other backgrounds are of ~0.5%
- **Preselection cuts have to be carefully adjusted**
  - To avoid biases due to cuts at generator level

# Conclusions & prospects

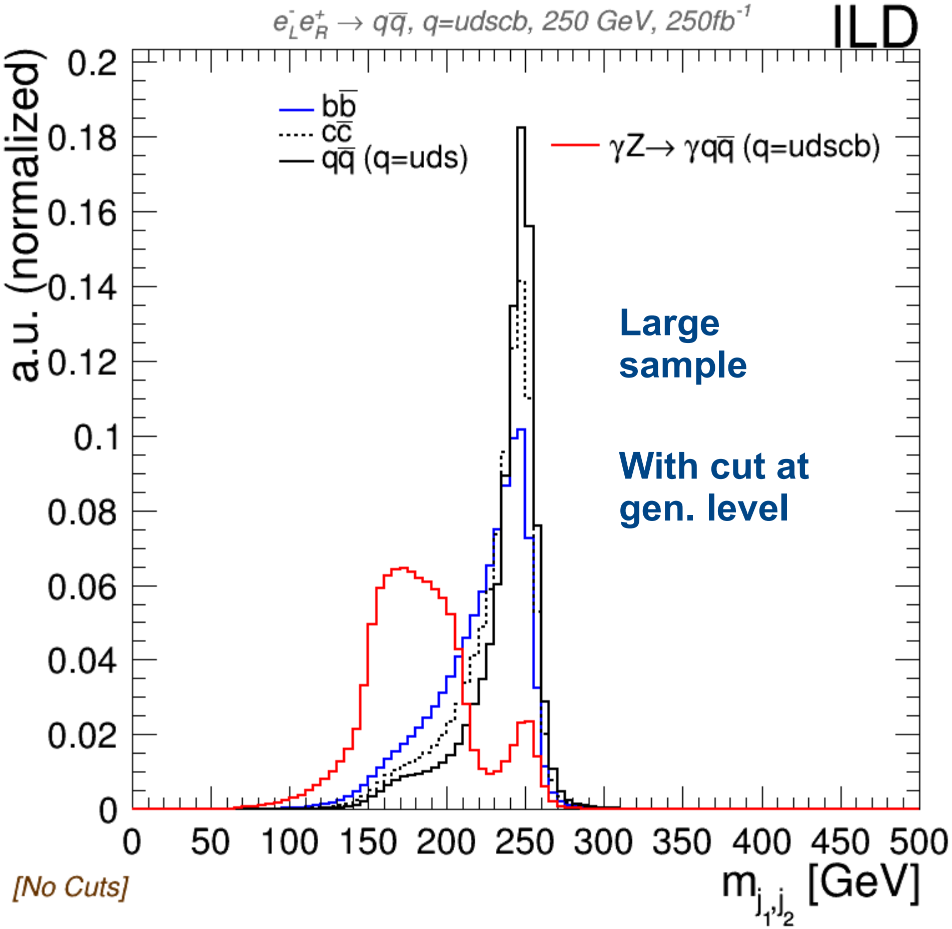
- We have defined a new and more robust preselection procedure and a more robust signal & radiative return definition
  - Have observed significant differences on background between full ISR sample and pre-selected sample
  - The procedure relies on the veto of photons → further investigation on photon ID capabilities with the new releases is planned.

- Reduced B/S for radiative return and other backgrounds
- The $\varepsilon_{q\bar{q}}$ are the same for all flavors and had no angular dependence
  - **The per mile level of precision is not compromised !!**

- Several other improvements on the method have been carried out (not discussed here)
  - We are now in the process of upgrading the ILD note draft for circulation.

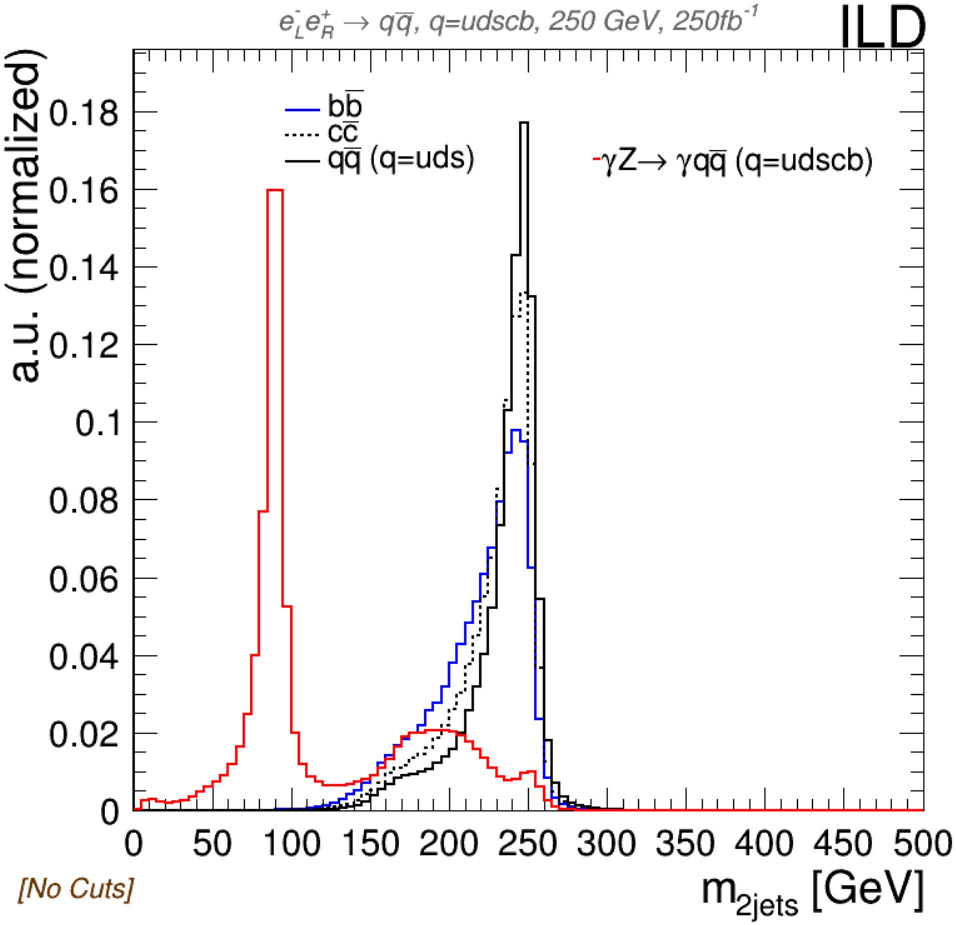
# Back-up slides

# m2jets (large vs small sample)




**Large sample**

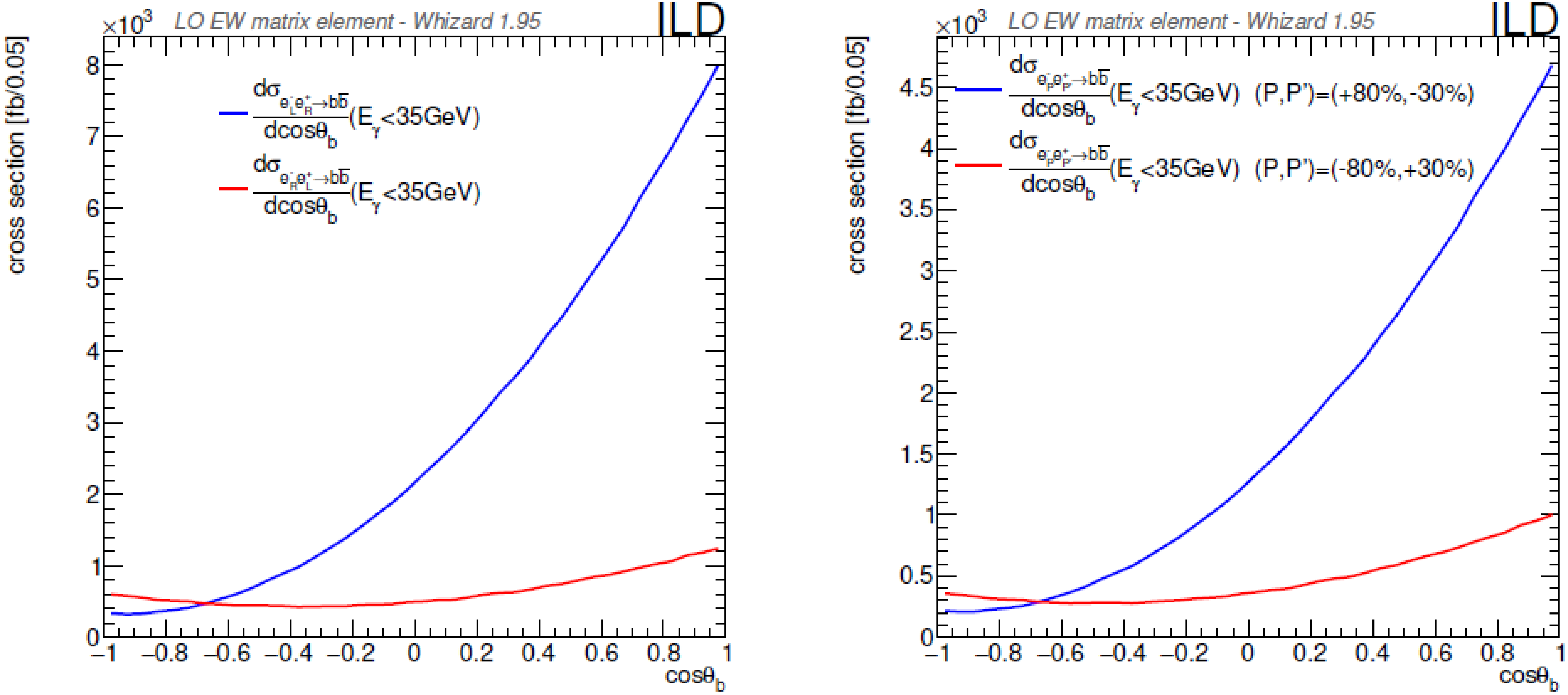
**With cut at gen. level**

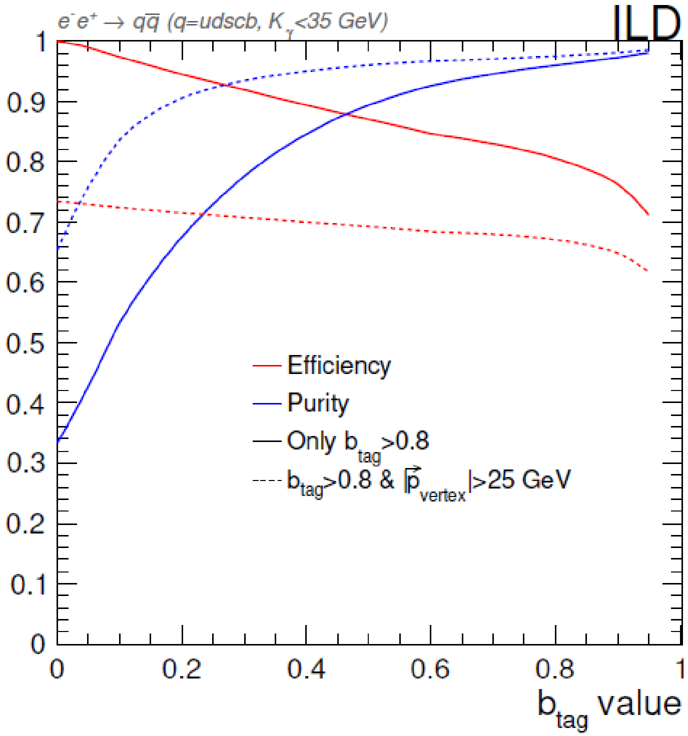


| Polarization | $\sigma_{e^-e^+\to q\bar{q}}(E_\gamma < 35\,GeV)$[fb] | | | $\sigma_{e^-e^+\to q\bar{q}}(E_\gamma > 35\,GeV)$[fb] | | |
|---|---|---|---|---|---|---|
| | $b\bar{b}$ | $c\bar{c}$ | $q\bar{q}$ $(q = uds)$ | $b\bar{b}$ | $c\bar{c}$ | $q\bar{q}$ $(q = uds)$ |
| $e^-_L e^+_R$ | 5677.2 | 8518.1 | 18407.3 | 20531.4 | 18363.8 | 57651.3 |
| $e^-_R e^+_L$ | 1283.2 | 3565.0 | 5643.5 | 12790.8 | 11810.8 | 36179.5 |
| | $R_q^{cont.}(E_\gamma < 35\,GeV)$ | | | $R_q^{cont.}(E_\gamma > 35\,GeV)$ | | |
| | $q = b$ | $q = c$ | $q = uds$ | $q = b$ | $q = c$ | $q = uds$ |
| $e^-_L e^+_R$ | 0.17457 | 0.26053 | 0.56480 | 0.17875 | 0.24826 | 0.57300 |
| $e^-_R e^+_L$ | 0.12183 | 0.33913 | 0.53905 | 0.13658 | 0.31175 | 0.55167 |
| | $A_{FB}^{q\bar{q}}(E_\gamma < 35\,GeV)$ | | | $A_{FB}^{q\bar{q}}(E_\gamma > 35\,GeV)$ | | |
| | $q = b$ | $q = c$ | $q = uds$ | $q = b$ | $q = c$ | $q = uds$ |
| $e^-_L e^+_R$ | 0.69057 | -0.59611 | -0.09905 | 0.62376 | -0.42540 | 0.10900 |
| $e^-_R e^+_L$ | 0.26489 | -0.68455 | -0.45878 | -0.58853 | 0.30890 | -0.14354 |

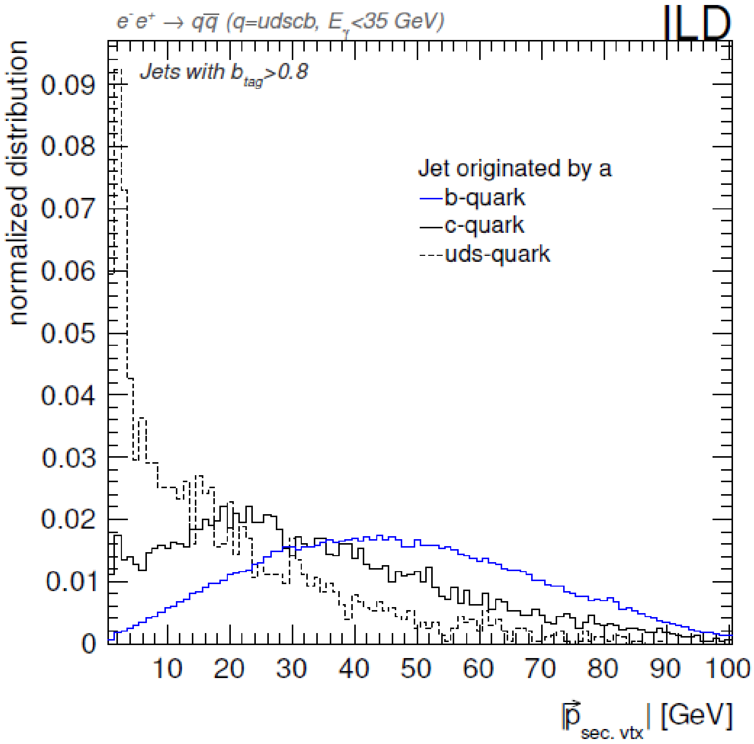
Table 1: Production cross section of quark pairs at 250 GeV of center of mass using polarized beams.

| Channel | $\sigma_{e^-_L e^+_R \to X}$ [fb] | $\sigma_{e^-_R e^+_L \to X}$ [fb] |
|---|---|---|
| $X = WW \to \nu q \nu q$ | 14874.4 | 136.4 |
| $X = ZZ \to q\bar{q}q\bar{q}$ | 1402.1 | 605.0 |
| $X = HZ \to q\bar{q}q\bar{q}$ | 346.0 | 222.0 |

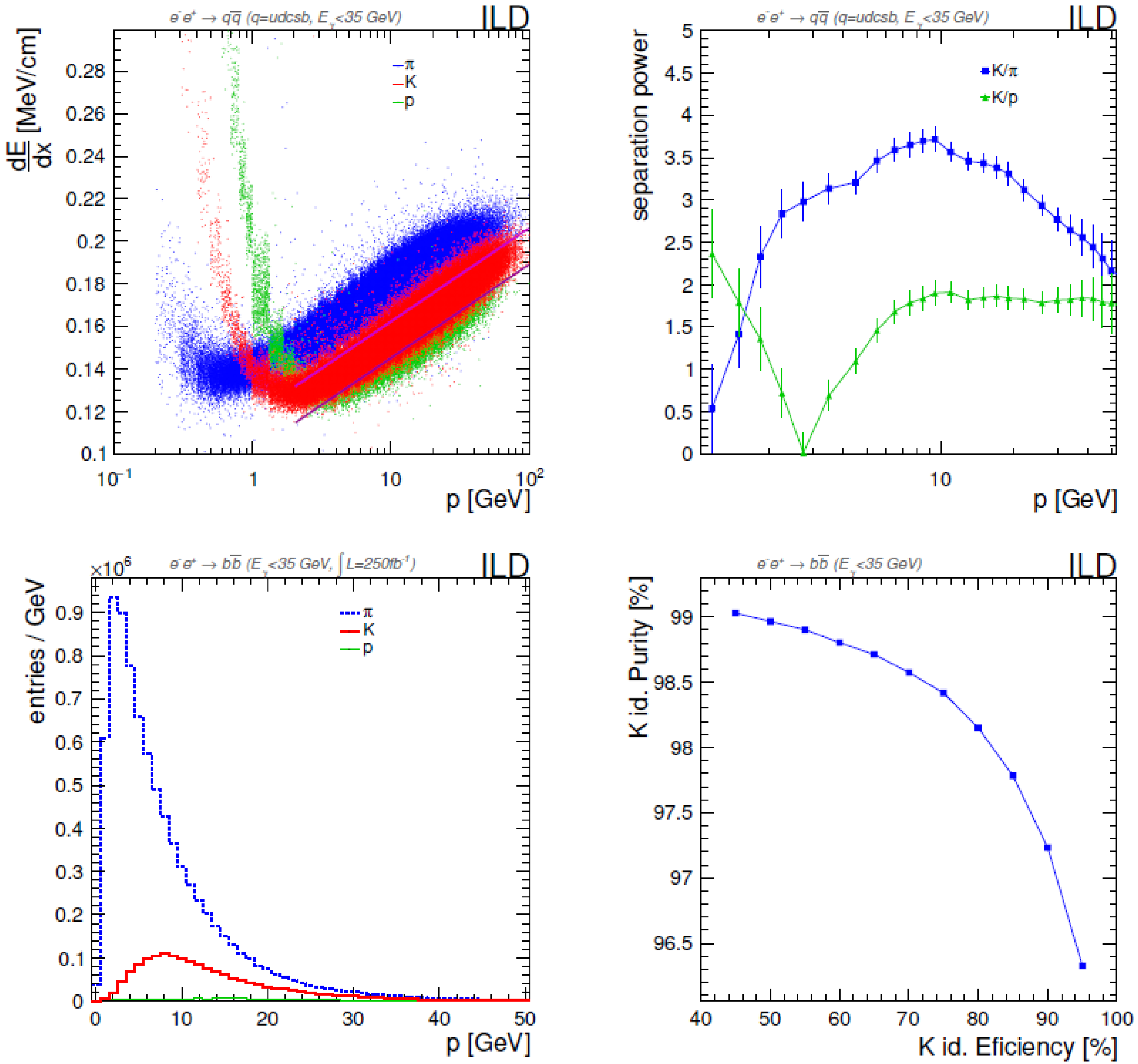

LO EW matrix element - Whizard 1.95
ILD
cross section [fb/0.05]
$\times 10^3$
$\frac{d\sigma_{e^-_L e^+_R \to b\bar{b}}}{d\cos\theta_b}(E_\gamma<35\text{GeV})$
$\frac{d\sigma_{e^-_R e^+_L \to b\bar{b}}}{d\cos\theta_b}(E_\gamma<35\text{GeV})$
$\cos\theta_b$
LO EW matrix element - Whizard 1.95
ILD
cross section [fb/0.05]
$\times 10^3$
$\frac{d\sigma_{e^-_P e^+_{P'} \to b\bar{b}}}{d\cos\theta_b}(E_\gamma<35\text{GeV})$ (P,P')=(+80%,-30%)
$\frac{d\sigma_{e^-_P e^+_{P'} \to b\bar{b}}}{d\cos\theta_b}(E_\gamma<35\text{GeV})$ (P,P')=(-80%,+30%)
$\cos\theta_b$



e⁻e⁺ → qq̄ (q=udscb, K_γ<35 GeV)
ILD
Efficiency
Purity
Only b_tag>0.8
b_tag>0.8 & |p_vertex|>25 GeV
b_tag value

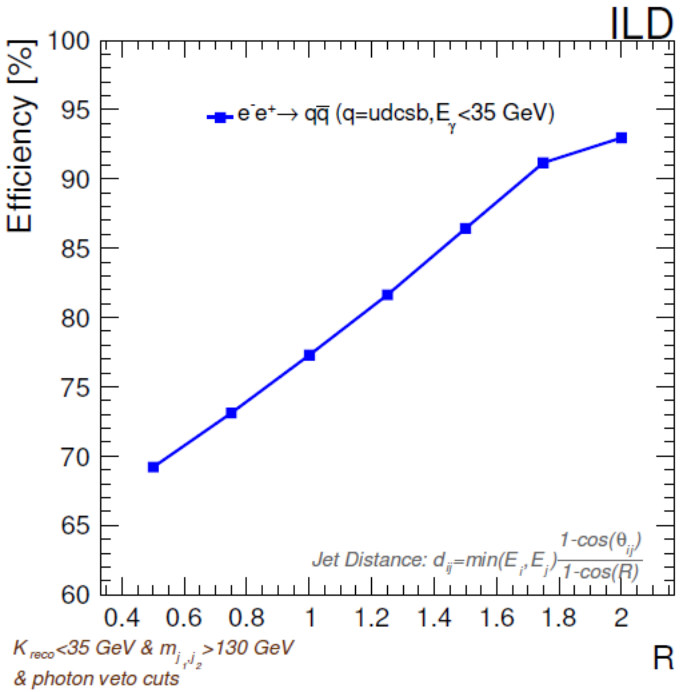
e⁻e⁺ → qq̄ (q=udscb, E_γ<35 GeV)
ILD
Jets with b_tag>0.8
Jet originated by a
b-quark
c-quark
uds-quark
normalized distribution
|p_sec. vtx| [GeV]

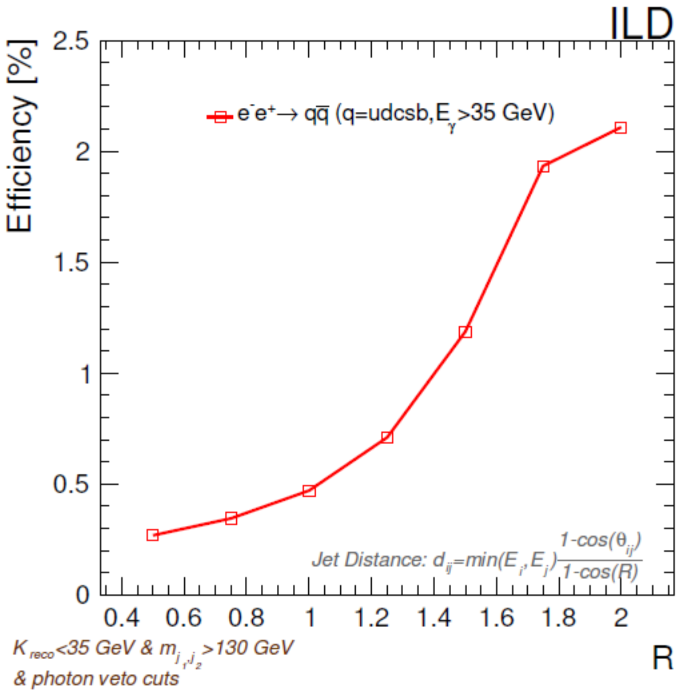


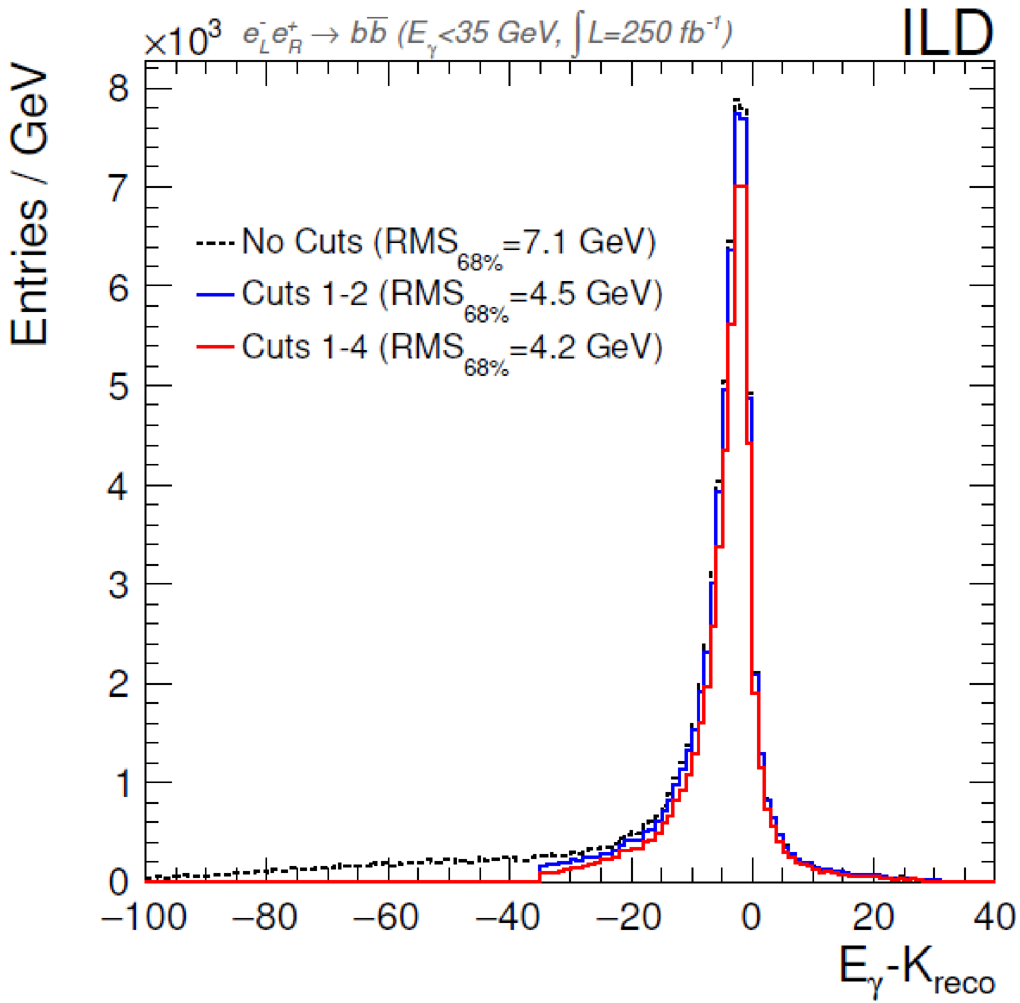
e⁻e⁺ → qq̄ (q=udcsb, Eγ<35 GeV)
ILD
dE/dx [MeV/cm]
p [GeV]
π
K
p
separation power
K/π
K/p
e⁻e⁺ → bb̄ (Eγ<35 GeV, ∫L=250fb⁻¹)
×10⁶
entries / GeV
e⁻e⁺ → bb̄ (Eγ<35 GeV)
K id. Purity [%]
K id. Eficiency [%]

$$d_{ij} = min(E_i^{2p}, E_j^{2p}) \frac{1 - cos(\theta_{ij})}{1 - cos(R)}$$

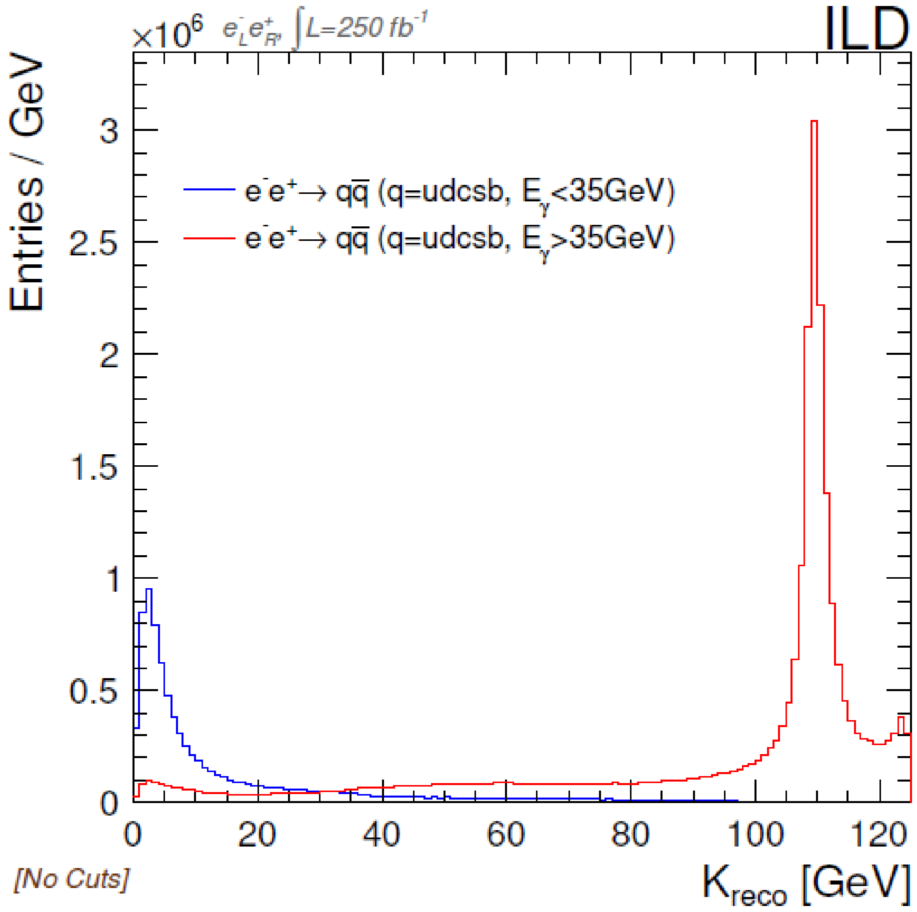
$$d_{iB} = E_i^{2p}$$

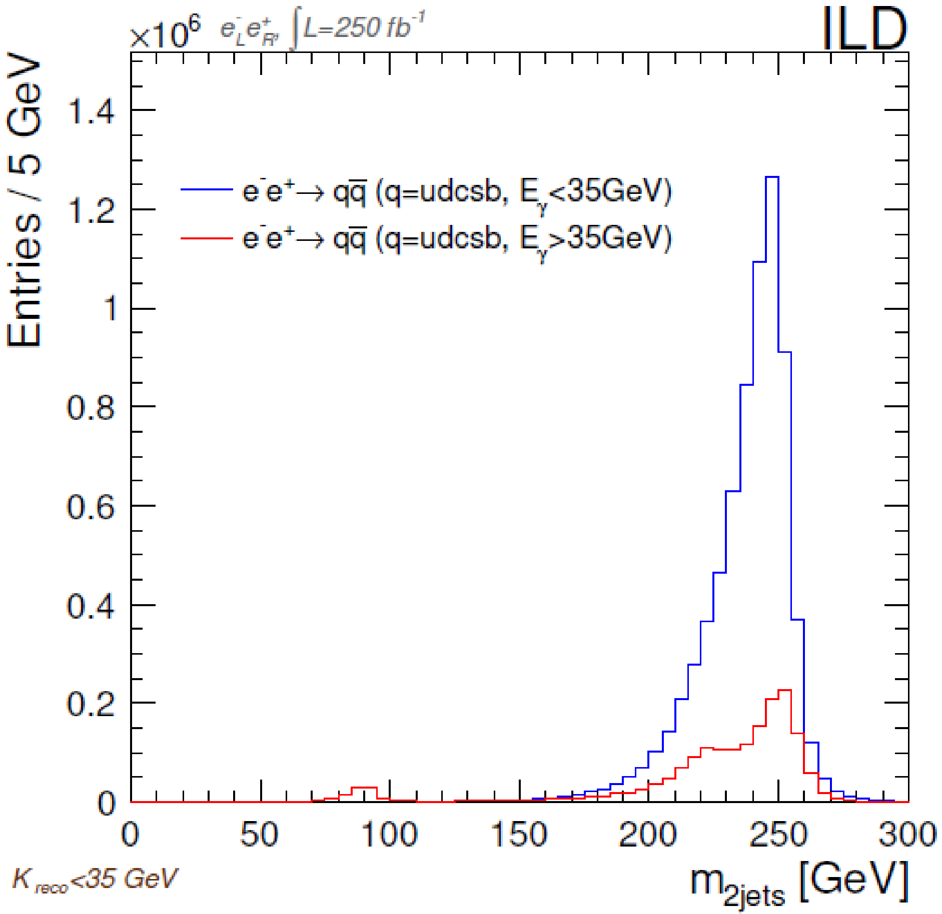


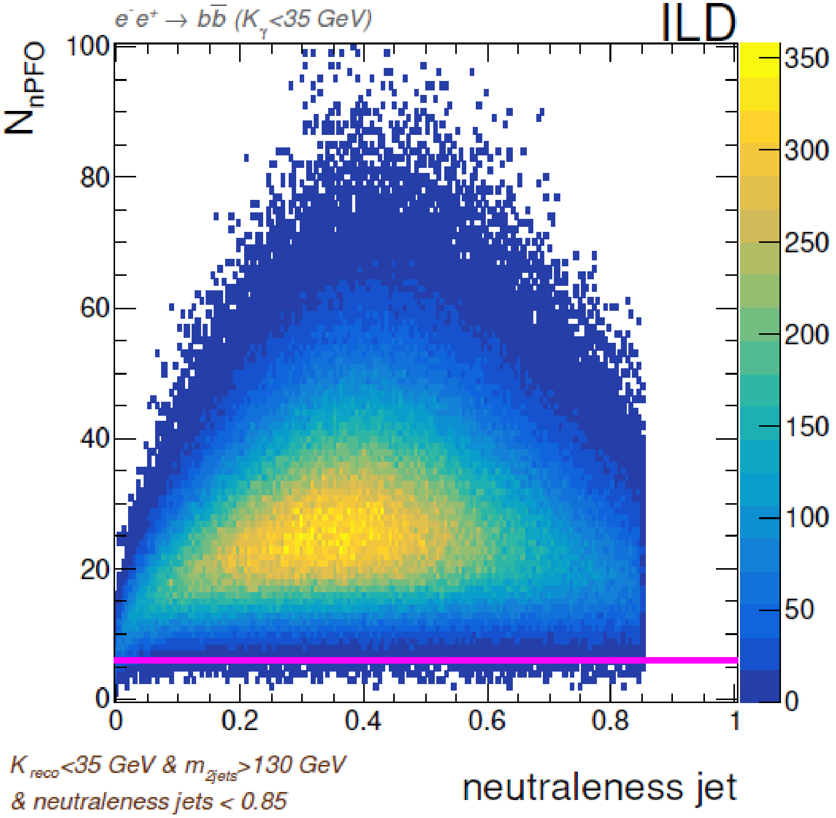





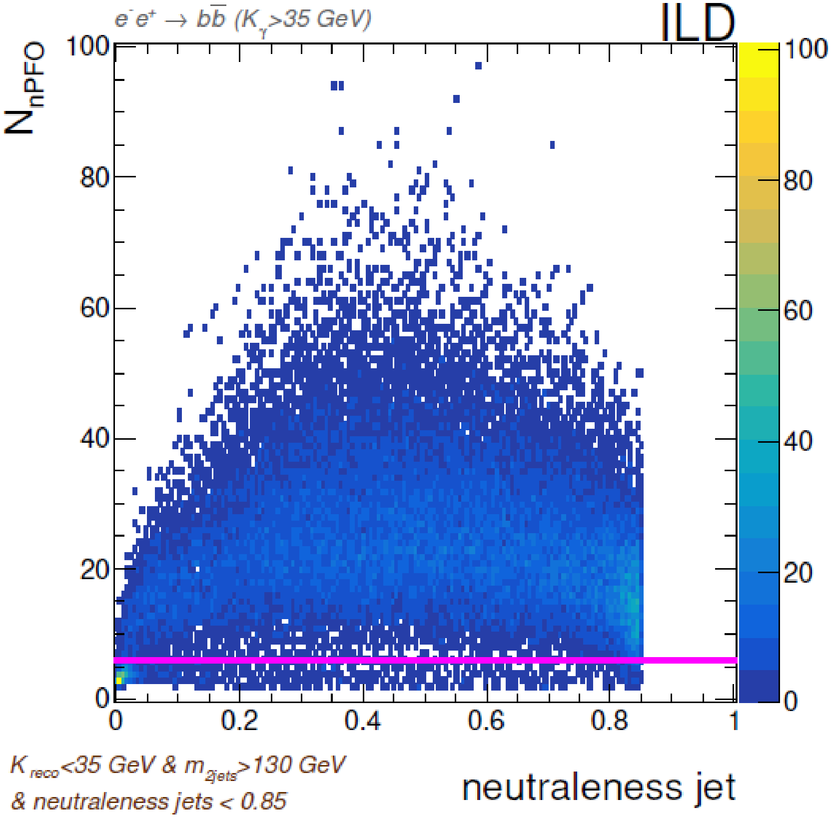
$\times 10^3$ $e^-_L e^+_R \to b\bar{b}$ ($E_\gamma$<35 GeV, $\int L$=250 fb$^{-1}$)
ILD
Entries / GeV
No Cuts ($RMS_{68\%}$=7.1 GeV)
Cuts 1-2 ($RMS_{68\%}$=4.5 GeV)
Cuts 1-4 ($RMS_{68\%}$=4.2 GeV)
$E_\gamma$-$K_{reco}$


×10^6 e-L e+R ∫L=250 fb^-1
ILD
Entries / GeV
e-e+→ qq̄ (q=udcsb, Eγ<35GeV)
e-e+→ qq̄ (q=udcsb, Eγ>35GeV)
Kreco [GeV]
[No Cuts]

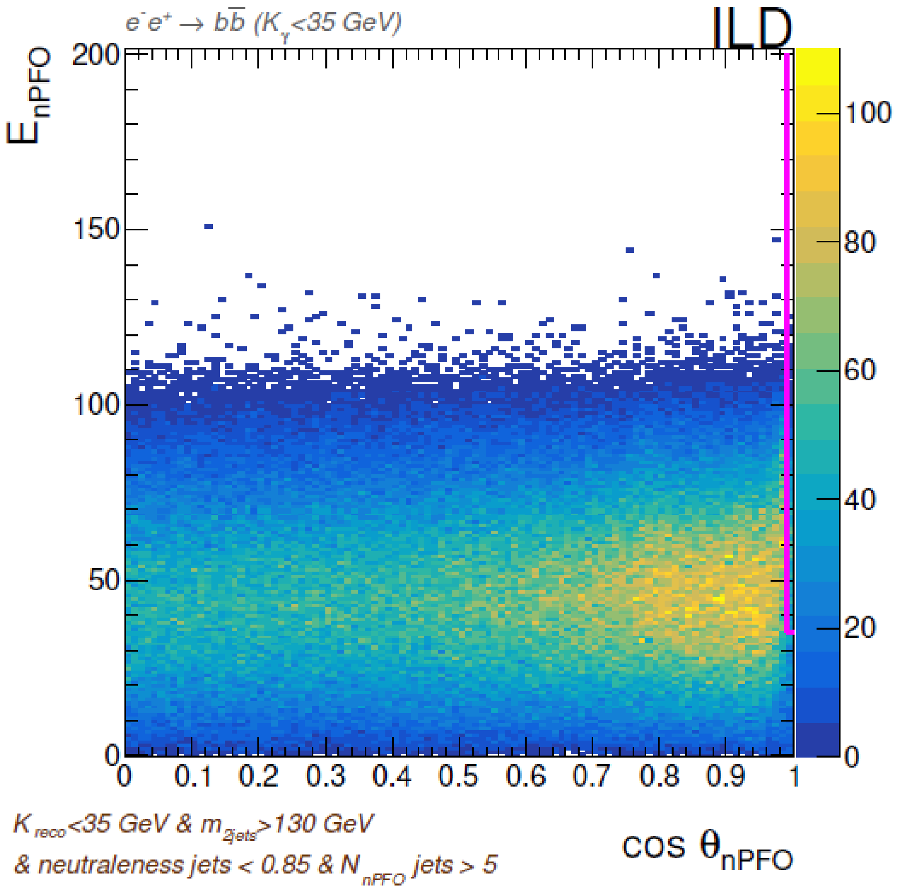

×10^6 e-L e+R ∫L=250 fb^-1
ILD
Entries / 5 GeV
e-e+→ qq̄ (q=udcsb, Eγ<35GeV)
e-e+→ qq̄ (q=udcsb, Eγ>35GeV)
m2jets [GeV]
K reco<35 GeV



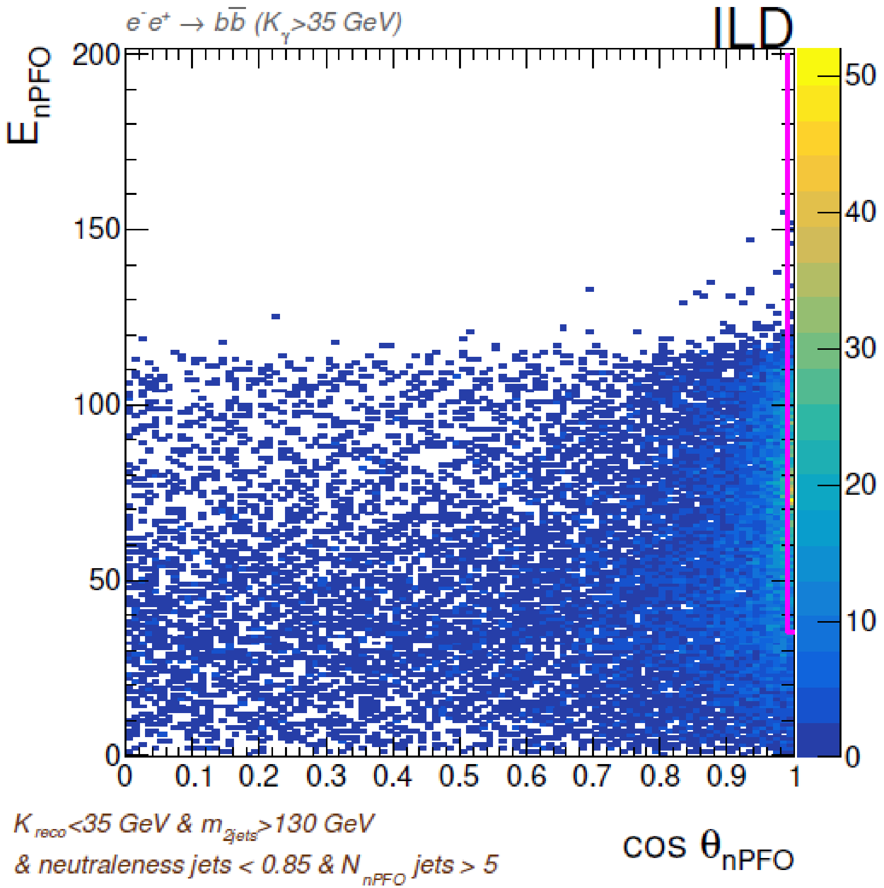
e⁻e⁺ → bb̄ (Kγ<35 GeV)
ILD
NnPFO
neutraleness jet
K reco<35 GeV & m2jets>130 GeV
& neutraleness jets < 0.85

e⁻e⁺ → bb̄ (Kγ>35 GeV)
ILD
NnPFO
neutraleness jet
K reco<35 GeV & m2jets>130 GeV
& neutraleness jets < 0.85



$e^-e^+ \to b\bar{b}$ ($K_\gamma$<35 GeV)
ILD
$E_{nPFO}$
200
150
100
50
0
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
100
80
60
40
20
0
$K_{reco}$<35 GeV & $m_{2jets}$>130 GeV
& neutraleness jets < 0.85 & $N_{nPFO}$ jets > 5
cos $\theta_{nPFO}$

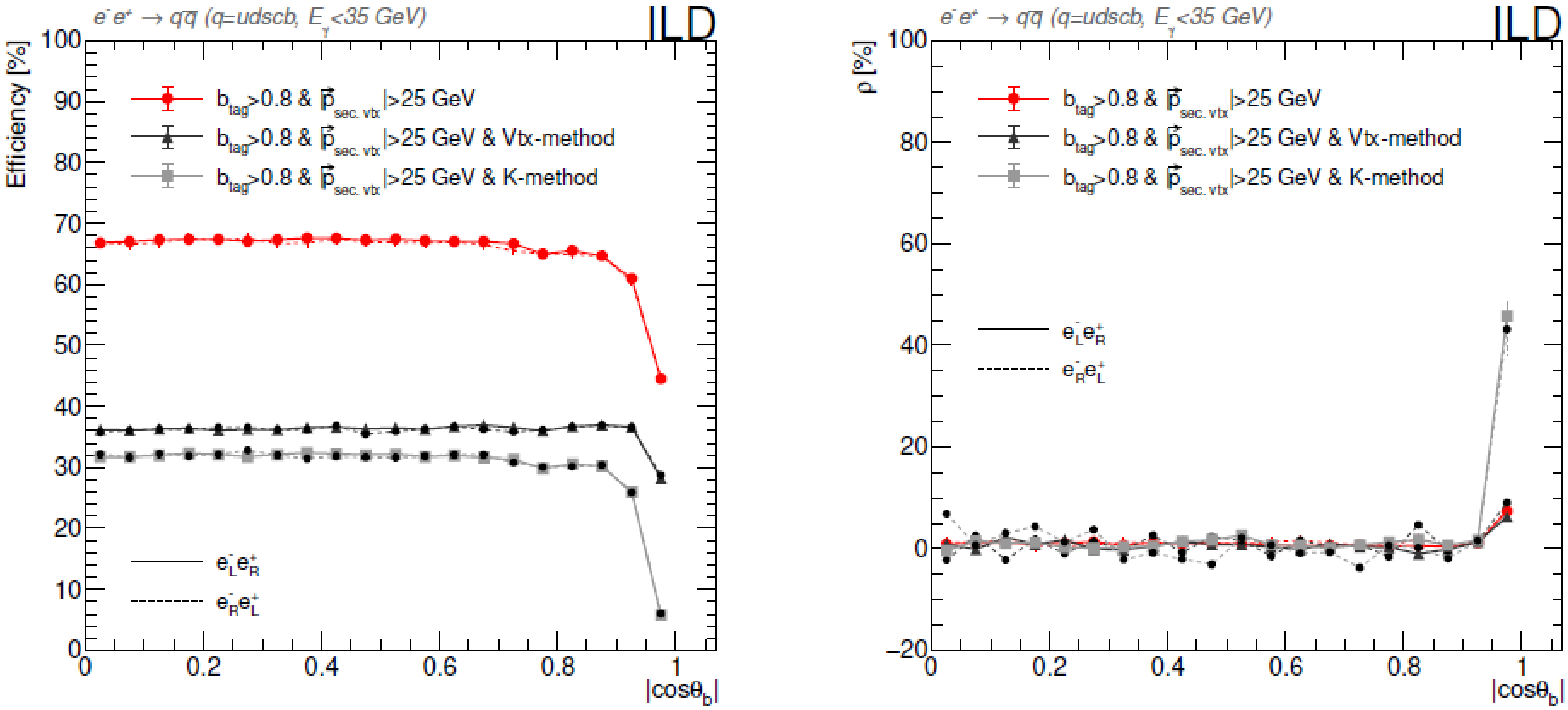
$e^-e^+ \to b\bar{b}$ ($K_\gamma$>35 GeV)
ILD
$E_{nPFO}$
200
150
100
50
0
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
50
40
30
20
10
0
$K_{reco}$<35 GeV & $m_{2jets}$>130 GeV
& neutraleness jets < 0.85 & $N_{nPFO}$ jets > 5
cos $\theta_{nPFO}$

## Efficiency of selection for $e^-_L e^+_R \to X$ [%]

| | $X = q\bar{q}$ ($E_\gamma < 35\,GeV$) | | | $X = q\bar{q}$ ($E_\gamma > 35\,GeV$) | | | |
|---|---|---|---|---|---|---|---|
| | $b\bar{b}$ | $c\bar{c}$ | $q\bar{q}$ (uds) | $q\bar{q}$ (udscb) | $X = ZZ$ | $X = WW$ | $X = HZ$ |
| No cuts | 100% | 100% | 100% | 100% | 100% | 100% | 100 |
| Cut 1 | 84.5% | 84.9% | 86.4% | 6.7% | 12.3% | 11.7% | 12.6 |
| + Cut 2 | 82.8% | 82.0% | 80.3% | 1.2% | 12.1% | 11.1% | 11.8 |
| + Cut 3 | 72.1% | 71.7% | 71.3% | 0.7% | 2.5% | 5.0% | 4.5 |
| + Cut 4 | 71.5% | 71.1% | 70.7% | 0.7% | 1.6% | 3.6% | 3.8 |

## Efficiency of selection for $e^-_R e^+_L \to X$ [%]

| | $X = q\bar{q}$ ($E_\gamma < 35\,GeV$) | | | $X = q\bar{q}$ ($E_\gamma > 35\,GeV$) | | | |
|---|---|---|---|---|---|---|---|
| | $b\bar{b}$ | $c\bar{c}$ | $q\bar{q}$ (uds) | $q\bar{q}$ (udscb) | $X = ZZ$ | $X = WW$ | $X = HZ$ |
| No cuts | 100% | 100% | 100% | 100% | 100% | 100% | 100 |
| + Cut 1 | 84.1% | 85.2% | 86.5% | 7.0% | 12.5% | 12.6% | 12.4 |
| + Cut 2 | 82.6% | 82.2% | 81.1% | 0.7% | 12.3% | 11.8% | 11.8 |
| + Cut 3 | 71.6% | 72.3% | 72.2% | 0.4% | 2.5% | 5.6% | 1.8 |
| + Cut 4 | 71.1% | 71.6% | 71.6% | 0.4% | 1.7% | 4.3% | 1.6 |



$e^-e^+ \to q\bar{q}$ (q=udscb, $E_\gamma$<35 GeV)
ILD
Efficiency [%]
$b_{tag}$>0.8 & $|\vec{p}_{sec.\ vtx}|$>25 GeV
$b_{tag}$>0.8 & $|\vec{p}_{sec.\ vtx}|$>25 GeV & Vtx-method
$b_{tag}$>0.8 & $|\vec{p}_{sec.\ vtx}|$>25 GeV & K-method
$e^-_L e^+_R$
$e^-_R e^+_L$
$|\cos\theta_b|$
$e^-e^+ \to q\bar{q}$ (q=udscb, $E_\gamma$<35 GeV)
ILD
ρ [%]
$b_{tag}$>0.8 & $|\vec{p}_{sec.\ vtx}|$>25 GeV
$b_{tag}$>0.8 & $|\vec{p}_{sec.\ vtx}|$>25 GeV & Vtx-method
$b_{tag}$>0.8 & $|\vec{p}_{sec.\ vtx}|$>25 GeV & K-method
$e^-_L e^+_R$
$e^-_R e^+_L$
$|\cos\theta_b|$

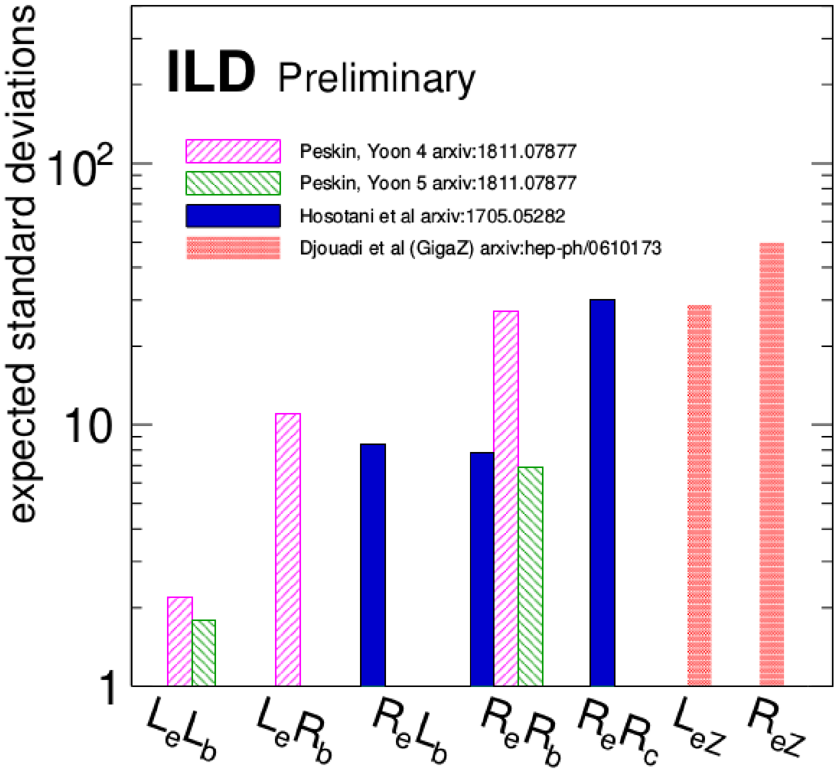



Figure 7: Expected number of standard deviations for different BSM scenarios when determining the different EW couplings to $c$- and $b$-quark at ILC250. The $x$-axis shows the different couplings for different chiralities, following the notation from [20]. The expectations for different BSM scenarios are shown: 1) for Djouadi [24] one assumes $m_{Z'} = 3$ TeV; 2) for the Peskin *et al.* model [26], two versions are given, labelled as Peskin 4 and Peskin 5; 3) for Hosotani et al.[25] one assumes $m_{Z'} \sim 8$ TeV for the 3 resonances. These prospects assume the input from the ILC GigaZ programme running at the $Z$-Pole [22] in order to improve by a factor $\sim$5 the current precision on the SM $Z$-boson couplings to the different quarks measured at the $Z$-pole.