# Analysis of the Higgsstrahlung process ($e^+e^- \rightarrow Z(qq)H$) in the hadronic decay mode for the context of the ILC with the SDHCAL.

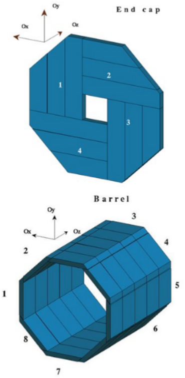
Héctor García Cabrera

CIEMAT

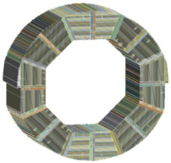
# Setup

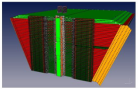
- Analysis of the Higgsstrahlung process in the context of the ILC.
- The detector model used is the ILD $l5\_o2\_v02$. Large model with the SDHCAL in the Tesla geometry and the SiWECal. List of all models.

## SiWECAL




## SDHCAL





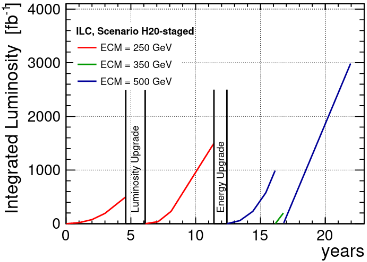
# Running scenarios



- Original scenarios in the proposal of 2015: ILC Operating Scenarios

| | $\int \mathscr{L} dt$ [fb$^{-1}$] | | | |
|---|---|---|---|---|
| $\sqrt{s}$ | G-20 | H-20 | I-20 | Snow |
| 250 GeV | 500 | 2000 | 500 | 1150 |
| 350 GeV | 200 | 200 | 1700 | 200 |
| 500 GeV | 5000 | 4000 | 4000 | 1600 |

- H-20 scenario consolidated in the report to Snowmass 2021: ILC report to Snowmass 2021



# MC Production



- At the moment of the request the request the Snowmass 2015 scenario was used as reference. L = 1150 fb -> 10.5 years of operation, which equals to $\sim$ 8.5 years in the H-20 scenario due to the early luminosity upgrade.
- The beam polarization values of operation in all scenarios are $P_{e^-}(80\%)$ and $P_{e^+}(30\%)$ creating the following proportions in the case of the $(-,+)$ as example:

$$eL = 90\% \quad eR = 10\%$$
$$pL = 35\% \quad pR = 65\%$$

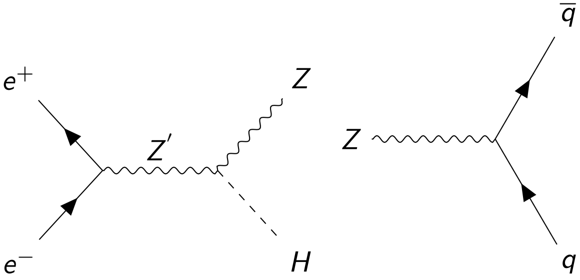
- The Snowmass 2015 scenario has the polarization mixture of: $(-,+) = 67.5\%$, $(+,-) = 22.5\%$, $(-,-) = 5\%$ and $(+,+) = 5\%$. Producing the following polarization factors for the total samples:

$$eLpR = 42.175\% \quad eLpL = 25.825\%$$
$$eRpL = 17.425\% \quad eRpR = 14.575\%$$

- H-20 polarization mixture: $(-,+) = 45\%$, $(+,-) = 45\%$, $(-,-) = 5\%$ and $(+,+) = 5\%$. Easily adapted from the Snowmass 2021 by weighting the histograms.

# Signal

The signal studied in this analysis is the Higgsstrahlung process with the Z decaying hadronically:



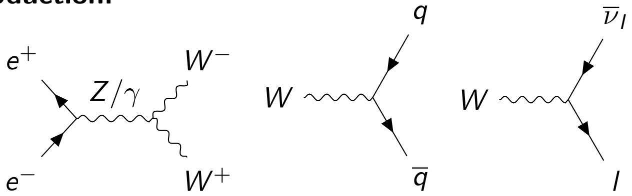

Also the MC request includes all the Higgs decay modes to make per channel tests: $H \longrightarrow b\overline{b}$, $H \longrightarrow c\overline{c}$, $H \longrightarrow gg$, $H \longrightarrow WW^*$, $H \longrightarrow ZZ^*$, $H \longrightarrow \mu\overline{\mu}$, $H \longrightarrow \tau\overline{\tau}$, $H \longrightarrow \gamma\gamma$ and $H \longrightarrow Z\gamma$

# Backgrounds I

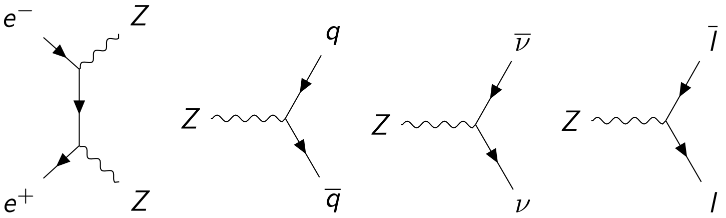


The background includes all processes with two or more jets in the final state.

**W pair production.**




**Z pair production.**

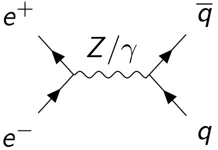


# Backgrounds II

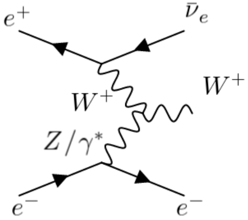


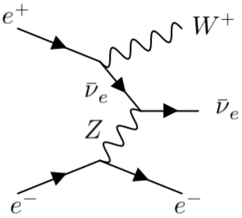
The background includes all processes with two or more jets in the final state.

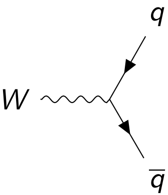
**Z hadronic decay.**




**Single W production (+ charge conjugation).**



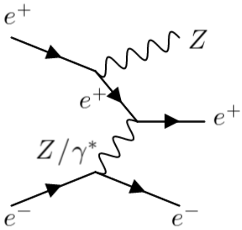



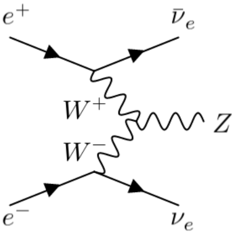
# Backgrounds III

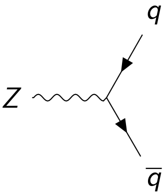


The background includes all processes with two or more jets in the final state.

**Single Z production (+ charge conjugation).**



# Event selection

- The samples include the simulation, reconstruction, digitizacion and overlay.
- The next step is to reconstruct the jets using FastJet.
- An event is selected if it has a dijet compatible with the Z mass.
- The parameters from FastJet are optimized by maximizing the percentage of well reconstructed jets. ($Z_{Tag} > 0.9$ and $Z_{Purity} > 0.9$).

$$\{j_1, j_2\} = argmin|m_{j1,j2} - m_Z| \;\; ; \;\; j_1, j_2 \in N_{Jets} \; ; j_1 \neq j_2$$

$$Z_{Tag} = \frac{E_Z^{j1,j2}}{E_Z^{Total}} \;\; ; \;\; Z_{Purity} = \frac{E_Z^{j1,j2}}{E^{j1,j2}}$$

# $M_Z$ and $M_{Recoil}$

Taking into account the incident angle of the two beams of 14mrad and $\sqrt{s} = 250 GeV$ the computations of the recoil mass are the following:

$$E_{q\overline{q}} = E_q + E_{\overline{q}} \; ; \; \overrightarrow{p_{q\overline{q}}} = \overrightarrow{p_q} + \overrightarrow{p_{\overline{q}}} \; ; \; M_Z = \sqrt{E_{q\overline{q}}^2 - \overrightarrow{p_{q\overline{q}}}^2}$$

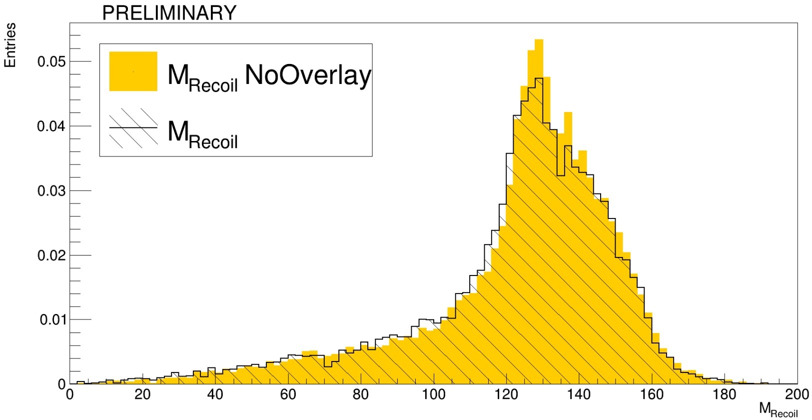
$$M_{Recoil}^2 = (2E + E_{q\overline{q}})^2 - ((2p_x + p_{q\overline{q}_x})^2 + p_{q\overline{q}_y}^2 + p_{q\overline{q}_z}^2)$$

$$2E = 250.0061252 GeV \; ; \; 2p_x = 1.7500286 GeV$$

# Recoil mass ee_kt_genkt



With this algorithm the optimal parameters have been found to be YCut = 0.002 and an impact of less than 0.3 % from overlay in the Z reconstruction in the study of the signal:



# Cross section

CIEMAT

$$\sigma_{ZH} = \frac{N_{Total} - N_B}{BR(Z \to q\bar{q}) \cdot \epsilon \cdot L}$$

The error of the cross section is computed as:

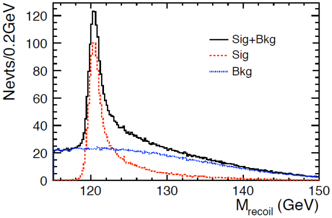
$$\Delta\sigma = \left(\frac{\Delta N_{Total}}{\epsilon L}\right)_{stat} \oplus \left(\frac{\Delta N_B}{\epsilon L} \oplus \sigma \frac{\Delta \epsilon}{\epsilon} \oplus \sigma \frac{\Delta L}{L}\right)_{syst} \; ; \; \frac{\Delta g_{HZZ}}{g_{HZZ}} = \frac{\Delta \sigma_{ZH}}{2\sigma_{ZH}}$$

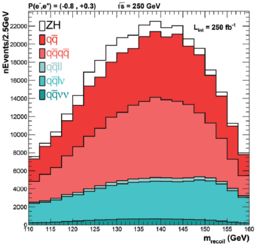
- $\left(\frac{\Delta\sigma}{\sigma}\right)_{stat} = S^{-1}$ ; $S = \frac{N_S}{\sqrt{N_S + N_B}}$
- $\Delta\epsilon$ and $\Delta N_B$ are the greatest challange since it involves the error propagation from the background rejection process.
- $\Delta L/L \sim 0.1\%$ from studies of the ILC and it is negligible compared to the others.

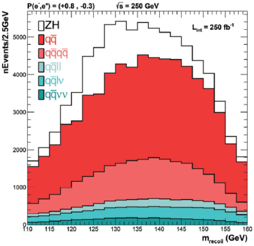
# Results from other studies

CIEMAT

$$F_M(x;\ \mathrm{M_H}, N_S) = N_S \cdot F_S(x;\ \mathrm{M_H}) + N_B \cdot F_B(x)$$



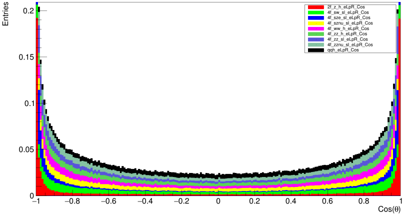



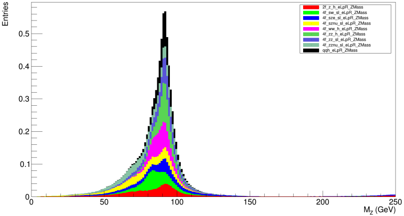



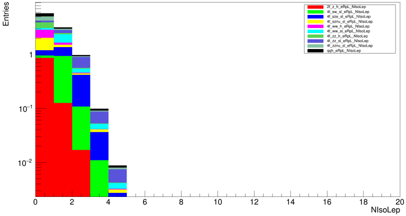
# Background rejection

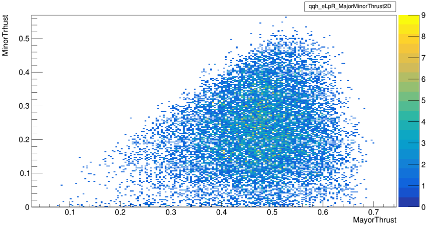


Proposal of background rejection variables: $Cos(\theta)$ and $M_Z$ and the Number of Isolated Leptons



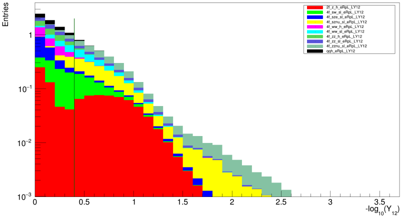



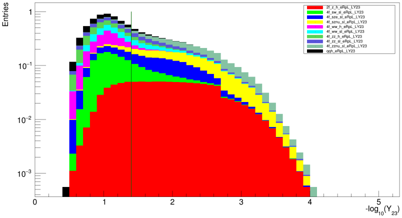



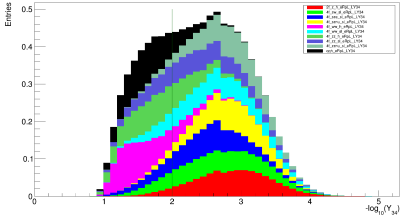
# Background rejection. YCuts

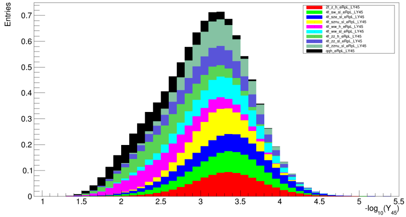
CIEMAT

Other background rejection variables: $Y_{12}$, $Y_{23}$, $Y_{34}$ and $Y_{45}$



# Background rejection. Cuts

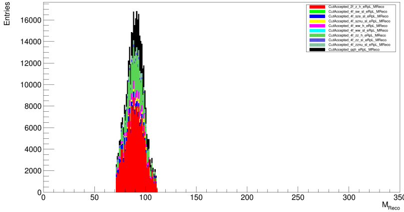


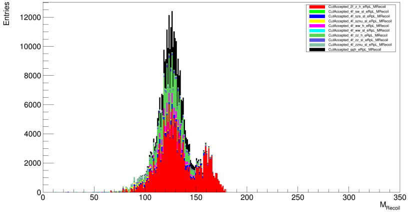
| | |
|---|---|
| Cut 1 | $M_{Reco} \in\ Mass_Z \pm 20GeV$ |
| Cut 2 | $\lvert Cos(\theta_{DiJet})\rvert < 0.9$ |
| Cut 3 | $-Log_10(Y12) < 0.4$ |
| Cut 4 | IF $-Log_10(Y23) > 1.4$ & $NJets + NIsoLep >= 2$<br>$\lvert\lvert \rightarrow TwoJetTopologyAnalysis = true$ |
| Cut 5 | ELSE IF $-Log_10(Y34) > 2.0$ & $NJets + NIsoLep >= 3$<br>$\lvert\lvert \rightarrow ThreeJetTopologyAnalysis = true$ |
| Cut 6 | ELSE IF $-Log_10(Y45) > 1.4$ & $NJets + NIsoLep >= 4$<br>$\lvert\lvert \rightarrow FourJetTopologyAnalysis = true$ |
| Cut 7 | ELSE<br>$\lvert\lvert \rightarrow OtherJetTopologyAnalysis = true$ |
| 2Jets | IF $E_{DiJet}/E_{Visible} < 0.9$<br>$\lvert\lvert \rightarrow$ Continue to next topology<br>ELSE IF $\lvert M_{DiJet} - M_Z\rvert < 10$ & NIsoLep $= 0$<br>$\lvert\lvert \rightarrow$ True |
| 3Jets | IF $\chi^2_Z < \chi^2_W$ & NIsoLep $= 0$<br>$\lvert\lvert \rightarrow$ Continue to next topology |
| 4Jets | Must pass all the following cuts:<br>$Thrust_{Major} < 0.4$ & $Thrust_{Minor} < 0.15$<br>$\lvert Cos(\theta_{DiJet_A})\rvert < 0.9$ & $\lvert Cos(\theta_{DiJet_B})\rvert < 0.9$<br>$\lvert M_{DiJet} - M_Z\rvert < 20$<br>$\chi^2_{ZH} < \chi^2_{WW}$ & $\chi^2_{ZH} < \chi^2_{ZZ}$<br>$\lvert\lvert \rightarrow$ True |
| Other | $Thrust_{Major} < 0.4$ & $Thrust_{Minor} < 0.15$<br>$\lvert\lvert \rightarrow$ True |

| Polarizations:<br>Final State | $P(eLpR)$<br>$\sigma(fb)$ | $P(eRpL)$<br>$\sigma(fb)$ | $P(eLpL)$<br>$\sigma(fb)$ | $P(eRpR)$<br>$\sigma(fb)$ |
|---|---|---|---|---|
| $Z(q(\bar{q})) + H$ | 343.03 | 219.49 | | |
| $q\bar{q}$ | 127965.53 | 70416.74 | | |
| $q\bar{q}\nu\nu$ | 1063.75 | 392.79 | | |
| $q\bar{q}l\nu$ | 29043.16 | 260.16 | 190.53 | 190.64 |
| $q\bar{q}ll$ | 2261.39 | 1686.21 | 1155.83 | 1157.20 |
| $q\bar{q}q\bar{q}$ | 16271.48 | 743.53 | | |

# Results






L $= 250\ fb^{-1}$

| Polarizations: | | $P(eLpR)$ | | $P(eRpL)$ | | $P(eLpL)$ | | $P(eRpR)$ |
|---|---|---|---|---|---|---|---|---|
| Final State | $\varepsilon$ | Events | $\varepsilon$ | Events | $\varepsilon$ | Events | $\varepsilon$ | Events |
| $Z(q(\bar{q})) + H$ | 24.28 % | 20822 | 23.3967 % | 12838 | | | | |
| $q\bar{q}$ | 0.40 % | 129204 | 0.22 % | 38592 | | | | |
| $q\bar{q}\nu\nu$ | 4.54 % | 12073 | 3.35 % | 3293 | | | | |
| $q\bar{q}l\nu$ | 0.62 % | 44870 | 0.19 % | 124 | 1.09 % | 521 | 1.10 % | 528 |
| $q\bar{q}ll$ | 1.43 % | 8113 | 0.93 % | 3945 | 0.67 % | 1928 | 0.68 % | 1963 |
| $q\bar{q}q\bar{q}$ | 7.25 % | 294906 | 9.80 % | 18210 | | | | |
| $S = \frac{N_S}{\sqrt{N_S+N_B}}$ | | 29.157 | | 46.26 | | | | |
| $\frac{\Delta\sigma_{ZH}}{\sigma_{ZH}} = S^{-1}$ | | 3.42 % | | 2.16 % | | | | |
| $S_F$ | 22.99 | | | | | | | |
| $\frac{\Delta\sigma_{ZH}}{\sigma_{ZH}}_F = S^{-1}$ | 4.34 % | | | | | | | |

# Conclusion



- Results similar to previous MVA studies.
- Still fit to $F_B + F_S$ seems not possible.
- NEXT STEP: Test model Independence.
- LAST STEP: Compute systematic errors.

# BACKUP

# Open questions



- Processes uncertain if they apply as background: *aa_2f* ; *aa_4f* ; *ae_3f* ; *ae_5f*

# Setup



- The latest MC production *sv02-02-01*, set 250 GeV, reconstructed with the SDHCAL *rv02-02-01*.mILD_l5_o2_v02 has now included the overlay events.
- Number of available events 50k in the DST-Merged files analyzed.
- Polarization eLpR (analysis of other polarizations is planned in the future).
- Beam crossing angle correction applied.

The objective of this preliminary analysis is to understand the effect of the newly included overlay into the signal events.

# Z identification (previous results)



- We run over all possible jet combinatorial possibilities.
- Then we choose the pair that give us the closest mass to the Z known mass.
- Two quality variables are defined:

$$Z_{Tag} = \frac{E_Z^{j_1,j_2}}{E_Z^{Total}} \quad Z_{Purity} = \frac{E_Z^{j_1,j_2}}{E^{j_1,j_2}}$$

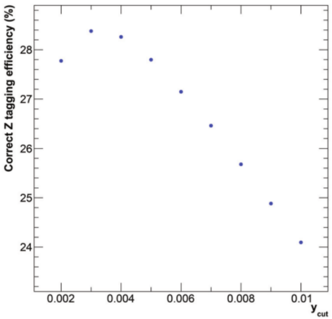
where $E_Z^{j_1,j_2}$: MC-truth di-jet energy from the Z, $E_Z^{Total}$: MC-truth Z energy and $E^{j_1,j_2}$: MC-truth total di-jet energy.

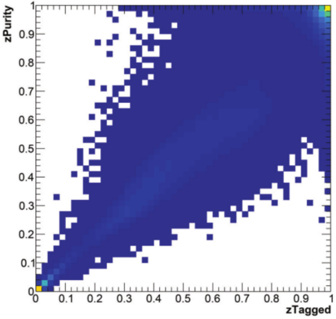
- The $E_Z^{j_1,j_2}$ is computed through a loop over all the PFOs. The associated MC particle energy is weighted using the RecoMCTrughLink excluding the particle with a Higgs as a parent (pdgId = 25).

# Z identification (previous results)

CIEMAT

The algorithm used in this study was *ee_kt_algorithm* within fast-jet with ExclusiveYCut strategy. An optimization of $y_{cut}$ has to be done in order to maximize the di-jet reconstruction selection efficiency. A di-jet is considered efficient if $Z_{Tag} > 0.9$ and $Z_{Purity} > 0.9$.




The optimal value in this previous study was $y_{cut} = 0.003$.

# FastJet algorithms



We want to find the set of parameters from the jet clustering algorithms from which the impact of the overlay is minimal. The two algorithms studied are:

- *Generalized $k_t$ algorithm for $e^+e^-$ collisions (ee_genkt) in the YCut exclusive mode.*

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

- *$k_t$ algorithm for $e^+e^-$ collisions (ee_kt) in the YCut exclusive mode.*

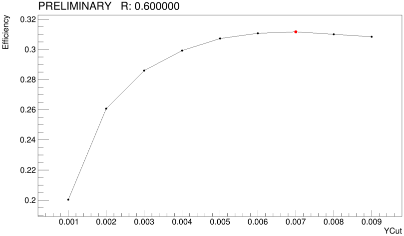
$$d_{ij} = 2min(E_i^2, E_j^2)(1 - cos(\theta_{ij}))$$

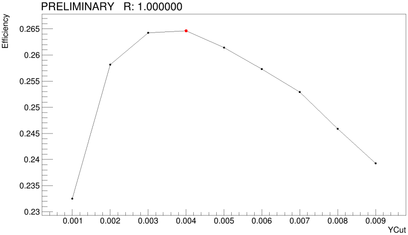
Parameters: ee_genkt -> **YCut**, **R** and **P** ; ee_kt -> **YCut**

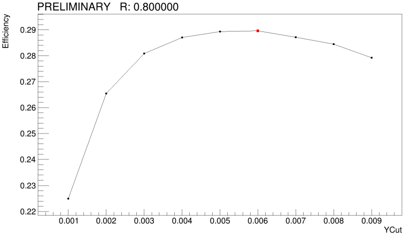
# YCut scans ee_genkt

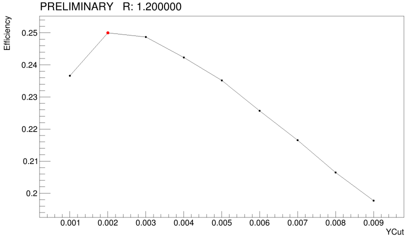
CIEMAT

**YCut** scans like the ones in Slide 4 are performed for different values of **R** and **P** for the ee_genkt algorithm and another one for the ee_kt algorithm.



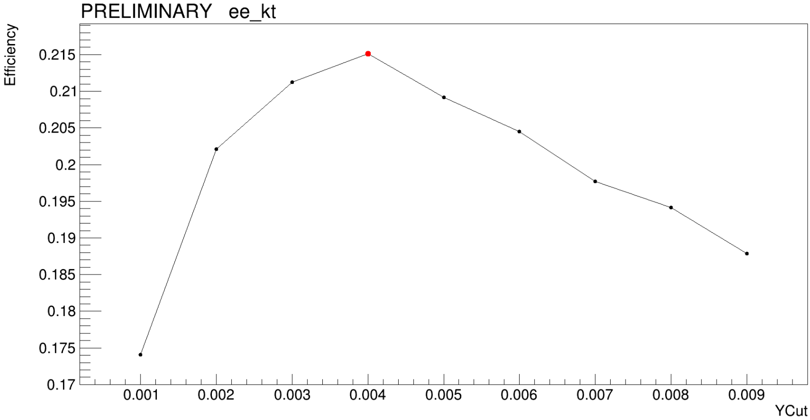



# YCut scans ee_kt



**YCut** scans like the ones in Slide 4 are performed for different values of **R** and **P** for the ee_genkt algorithm and another one for the ee_kt algorithm.



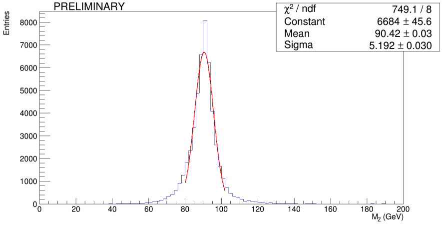
# Quality variables

CIEMAT

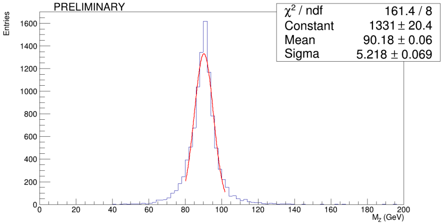
From the studied parameters R = 0.6 gives the highest efficiency. However this does not mean that gives the best jet reconstruction. The quality variables computed are:

- $Z_{Mass}$ differences (%). The $Z$ mass is reconstructed, following the previous procedure, with and without overlay. This distribution is fitted to a gaus function and then two differences are computed:
  $ZDiff = |M_Z - M_Z^{NoOverlay}|/M_Z$
  $OverlayDiff = |M_Z^{Overlay} - M_Z^{NoOverlay}|/M_Z^{NoOverlay}$

**R = 0.6**




**R = 0.6 No Overlay**



# Quality variables

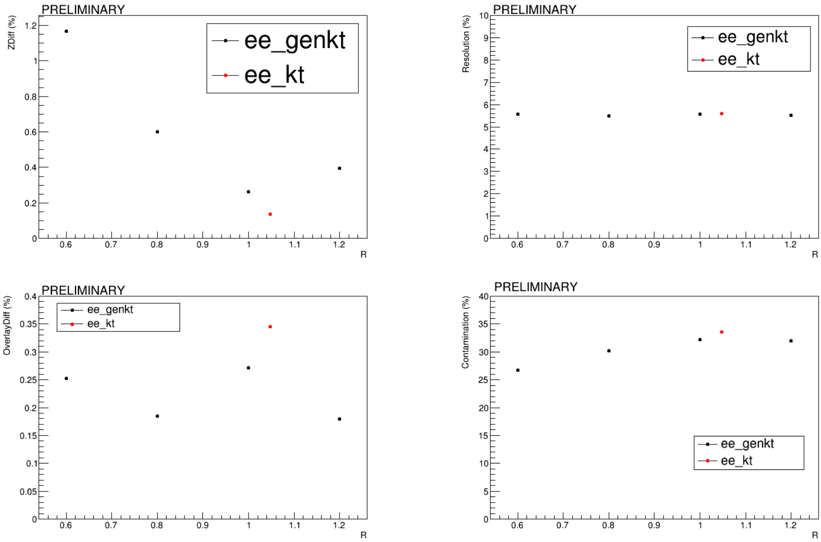


From the studied parameters R = 0.6 gives the highest efficiency. However this does not mean that gives the best jet reconstruction. The quality variables computed are:

- $Z_{Mass}$ differences (%). The $Z$ mass is reconstructed, following the previous procedure, with and without overlay. This distribution is fitted to a gaus function and then two differences are computed:
  $ZDiff = |M_Z - M_Z^{NoOverlay}|/M_Z$
  $OverlayDiff = |M_Z^{Overlay} - M_Z^{NoOverlay}|/M_Z^{NoOverlay}$
- $Z_{Mass}$ "resolution". $(\sigma/M_Z)$ from the previous fit.
- Jet Contamination (%). The mean percentage of jets that have some overlay contribution.

# Quality variables

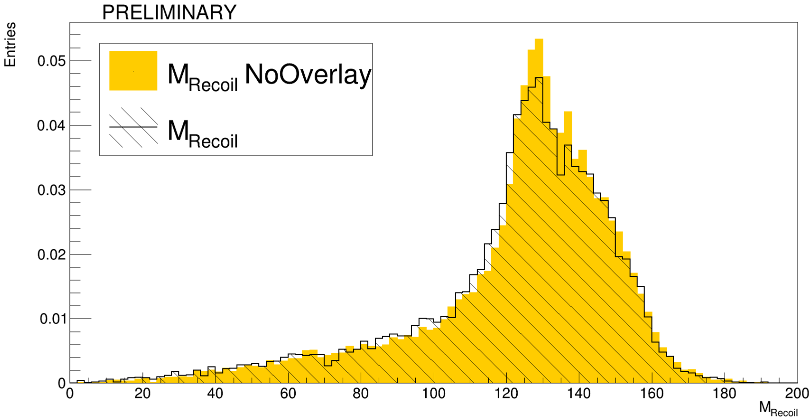




# Recoil mass ee_genkt



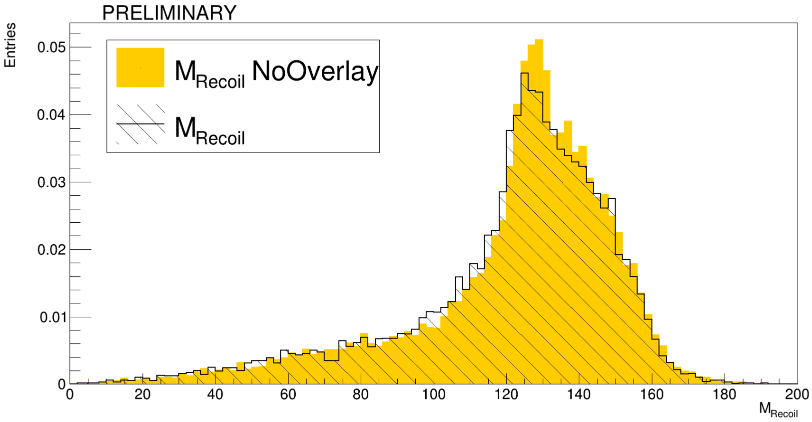
With this algorithm the optimal parameters have been found to be R = 1.0 and YCut = 0.004 and an impact of less than 0.3 % from overlay in the Z reconstruction.



# Recoil mass ee_kt



With this algorithm the optimal parameter is YCut = 0.004 and an impact of 0.35 % from overlay in the Z reconstruction.



# Conclusion



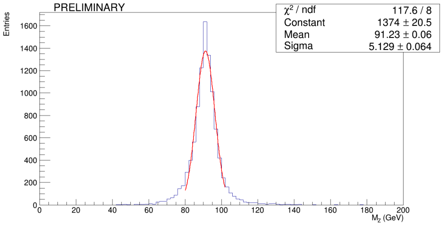
It have been shown the set of parameters that minimizes the overlay effect while efficiently reconstructing the Z mass and showing healthy recoil mass distributions.

A slightly better shape in the recoil mass is obtained from the ee_kt method that indicates that it is the algorithm that should be used.

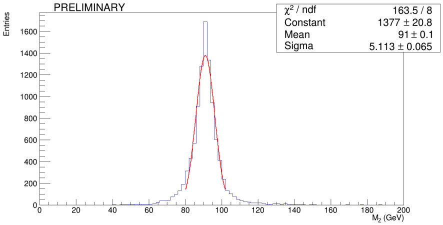
# $M_Z$ R = 1.0



## Overlay



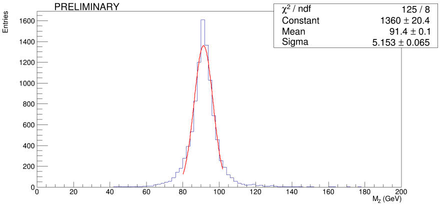

## No Overlay



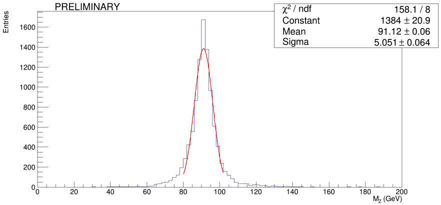
# $M_Z$ ee_kt

CIEMAT

## Overlay



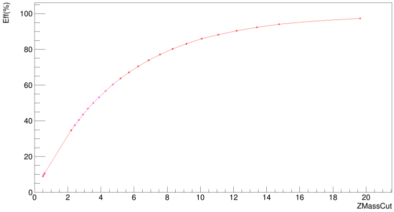

## No Overlay

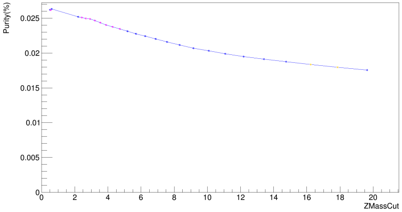


# Cut selection

$$\epsilon = \frac{N^{sel}_{signal}}{N^{tot}_{signal}} \; ; \; Purity = \frac{N^{sel}_{signal}}{N^{tot}_{events}} \implies \frac{\Delta\sigma}{\sigma} = \frac{\Delta Purity}{Purity} \oplus \frac{\Delta\epsilon}{\epsilon} \oplus \frac{\Delta L}{L}$$

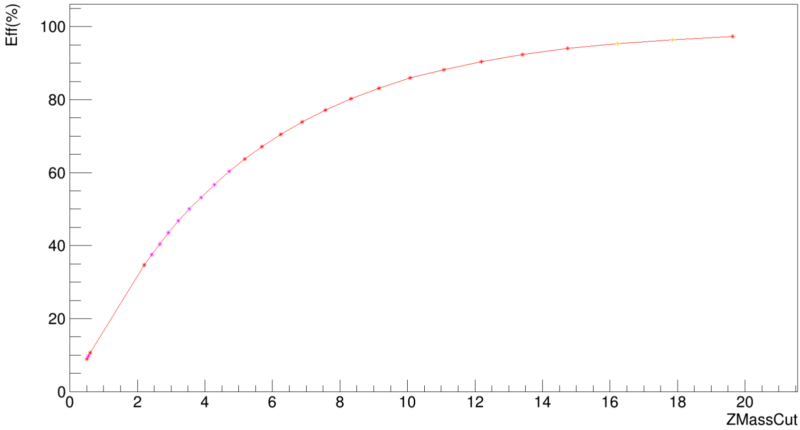
And using as target a resolution of 3% from previous analysis this creates partial and optimal ranges of selection for the cuts. Cut in the Z Mass as an example:



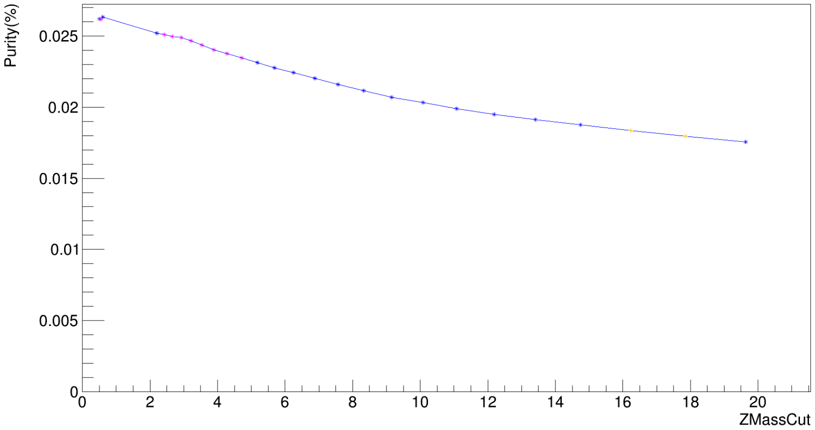
# Efficiency





# Purity




Purity(%)
0.025
0.02
0.015
0.01
0.005
0
0
2
4
6
8
10
12
14
16
18
20
ZMassCut