CLICdp-Note-2017-007
21 December 2017

Separation of hadronic W and Z decays in the CLIC_ILD and the CLICdet detector models at 1.4 and 3 TeV

Sascha Dreyer [1)*†], Philipp Roloff[*], Rickard Ström[*], Matthias Weber[*]

[*] *CERN, Switzerland,* [†] *Heidelberg University, Germany*

Abstract

A study of the W and Z separation was performed for the CLIC_ILD and the CLICdet detector models for the proposed Compact Linear Collider (CLIC). Comparisons were done for fully-hadronic WW and ZZ events at the collision energies of 1.4 and 3 TeV. Particle flow objects are reconstructed using a full simulation of the events including relevant beam-induced background processes. Several different collections of particles, with varying level of background suppression, were compared for each of the detector models and optimal jet clustering parameters were found in each case, resulting in the best separation of the W and Z mass peaks. The CLICdet detector model performs similar to CLIC_ILD with an achieved jet mass separation of around 1.6 σ at 1.4 TeV and 1.3 σ at 3 TeV. For both detector models we achieve a better separation at 1.4 TeV when comparing dijet masses rather than large-R jet masses. At 3 TeV jets with a radius around $R = 0.5$ perform similarly well as dijets.

This work was carried out in the framework of the CLICdp Collaboration

[1] sascha.dreyer@cern.ch

1. Introduction

The Compact Linear Collider (CLIC) is a proposed electron-positron accelerator for the post-LHC era. It is planned to be built in several energy stages to maximise its physics potential, currently assumed at 380 GeV, 1.5 TeV and 3.0 TeV. The physics program covers the precise measurement of properties of both the Higgs boson and the top quark. This includes measurements of the Higgs self-coupling, Higgs couplings to fermions and gauge bosons, and the top quark mass, the latter measured in a dedicated $\mathrm{t\bar{t}}$ threshold scan. The program at high energies is targeted at direct and indirect searches for physics beyond the Standard Model.

The general detector design for CLIC includes a highly granular calorimeter in order to achieve a precise jet energy resolution of $3.5-5\%$ in the jet energy range of 50 GeV - 1 TeV. The Pandora Particle Flow Algorithm (Pandora PFA) [1] toolkit is used to identify particles in the detector and is based on a pattern recognition software also taking into account available tracking information. The output of the PFA are so called Particle Flow Objects (PFOs), including the particle type (ID) and its measured properties. For charged particles, the precise information of the tracker is used, while for photons we only use the information from the electromagnetic calorimeter (ECAL). Neutral hadrons are reconstructed using clusters in the ECAL and the less precise hadronic calorimeter (HCAL). The mis-assignment of clusters and tracks or wrong splitting/merging of calorimeter clusters contributes to the jet energy resolution. A good benchmark of the detector performance is the separation of W and Z jets by their mass. In fact, a reasonable separation was one of the key goals when designing the new unified CLIC detector model. The conclusion from the CDR studies was that a jet energy resolution of 5% or below is needed to achieve a separation of $2\,\sigma$ [2].

At CLIC, small beam sizes are required to achieve the desired luminosities of $\approx 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ [2]. This gives rise to beamstrahlung photons, produced when electrons or positrons interact with the electromagnetic field of the opposite beam. The products of these interactions lead to background particles similar to pile-up at hadron colliders. Contamination with larger transverse momenta occurs mainly due to $\gamma\gamma \rightarrow$ hadrons of which 90% affects the endcap regions of the detector. The impact of this background is reduced by applying p_{T} and additional timing cuts to the fully reconstructed PFOs. It is important to note that p_{T} cuts alone are not sufficient, and that the use of time information in this reconstruction step is crucial for the reduction of beam-induced background. On average, at most one hard $\mathrm{e^+\,e^-}$ interaction is expected per bunch train, consisting of 312 distinct bunch crossings separated by 0.5 ns [2]. The timing cuts on the PFOs reduce the background from interactions at a different time as the hard interaction. Three levels of timing cuts were applied, default, loose and tight. Typically considered are loose selected PFOs for the initial energy stage, selected PFOs for the intermediate energy stage and tight selected PFOs for the highest energy stage. Beamstrahlung and ISR photons reduce the centre-of-mass energy of the interaction, resulting in a luminosity spectrum with a tail towards lower energies.

In this note we present a comparison of the W candidate mass distributions in WW events using the different timing cut collections for the CLIC_ILD [2] and the new CLICdet [3] detector models at a collision energy of 1.4 [2] and 3 TeV using four quark final states, see Figure 1. The CLIC_ILD samples are used to optimise the jet clustering for a W/Z jet mass separation study. The results are used to compare the old and new detector models in terms of W/Z separation as benchmark for the performance of the new model.

1.1. Detector models

The CLIC_ILD detector model is discussed in detail in the CDR [2] while the new unified CLIC detector model CLICdet is described in [3]. The main difference between these models is the tracking sub-system.

[2] A slightly different staging scenario, with the intermediate stage at 1.5 TeV, has recently been adopted and will be used for future studies [4].

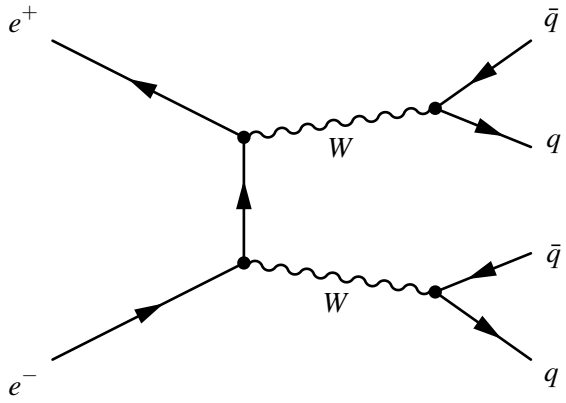


Figure 1: Feynman diagram for $e^+e^- \to WW \to qqqq$.

A Time Projection Chamber (TPC) was proposed for CLIC_ILD, whereas a full silicon tracker was chosen for CLICdet, partially motivated by the challenging beam-induced background conditions. Both models feature a tungsten-based ECAL with interlaid silicon sensors. Also the hadronic calorimeters differ between the models with CLIC_ILD featuring tungsten absorbers in the barrel and steel in the endcaps, while the CLICdet design features an all steel HCAL. For both detector models, the tracker and calorimeter systems are enclosed by a large-radius solenoid magnet of 4 T. An iron return yoke with muon detection layers, Resistive Plate Chambers (RPC) or crossed scintillator bars, is placed outside the magnet.

1.2. Event samples and reconstruction

To study the impact of the different timing cuts on the W candidate mass distributions, we use the event samples described in Table 1, which also include the expected integrated luminosity at the corresponding energy stages. The samples include WW and ZZ events where the bosons decay hadronically. The expected Standard Model (SM) cross-section for ZZ is around ten times less than for WW. Monte Carlo information was used to tag events as either WW or ZZ. The selection is based on a mass window as well as charge and quark flavour. Cuts were applied on the direction of all quarks to exclude the very forward region, $|\cos\theta| < 0.95$, where θ is the polar angle measured from the beam direction. A cut on the invariant mass of 1.2 TeV for the 1.4 TeV samples and 2.8 TeV for the 3 TeV samples was chosen for later extractions of the WW cross-section at the nominal energies. The study was performed with the ILCSOFT v2017-07-27 used by CLICdp.

The VLC jet clustering algorithm [5] was used in most cases because of its three adjustable parameters, jet radius, β and γ. The jet radius R defines the maximal angular distance to which PFOs are considered for clustering. The β parameter controls the clustering order. The default choice 1.0 results in an order similar to the k_T-algorithm used at hadron colliders. The γ parameter controls the rate of shrinking in jet size in the forward region. The resulting distance parameter to which PFOs are clustered into a jet is intrinsically well suited for the e^+ e^- environment [5] and is given by

$$d_{ij} = 2\min(E_i^{2\beta}, E_i^{2\beta})(1-\cos\theta_{ij})/R^2, \tag{1}$$

with the beam distance:

$$d_{iB} = E_i^{2\beta}\sin\theta_{iB}{}^{2\gamma}. \tag{2}$$

When the smallest distance in a clustering step is the beam distance of an object, this object is clustered to the beam-jet and rejected. This behaviour helps to reject isolated forward particles that are likely to

Table 1: Four quark samples for CLIC_ILD and CLICdet, N describes the expected number of events from $L \times \sigma$.

Model	$\sqrt{s}$ [TeV]	ProdID	N	L [ab^{-1}]	σ [fb]	overlay	version	generator
CLIC_ILD	1.4	4034	1 867 650	1.5	1245.1	yes	CDR	WHIZARD V55
CLIC_ILD	3	6776	1 639 500	3	546.5	yes	CDR	WHIZARD V57
CLICdet	1.4	8307	1 867 650	1.5	1245.1	yes	o3_v12	WHIZARD V55
CLICdet	3	8299	1 639 500	3	546.5	yes	o3_v12	WHIZARD V57
CLICdet	1.4	8581	1 867 650	1.5	1245.1	no	o3_v12	WHIZARD V55
CLICdet	3	8618	1 639 500	3	546.5	no	o3_v12	WHIZARD V57

originated from background. Two exclusive jets are used in the comparison of timing cut collections. For the W/Z separation, we consider both, two and four exclusive jets. Four exclusive jets are paired into W and Z candidates by minimising $\chi^2 = (m_{ij} - m_{\mathrm{W/Z}})^2 + (m_{kl} - m_{\mathrm{W/Z}})^2$.

2. Timing cut collections

The W candidate mass distribution is studied for three different PFO collections: loose selected, selected and tight selected. The VLC clustering algorithm with a radius of R = 0.8 and $\gamma = \beta = 1.0$ is used throughout this study. For CLICdet, a comparison with the reference case without the $\gamma\gamma \to$ hadrons overlay is done.

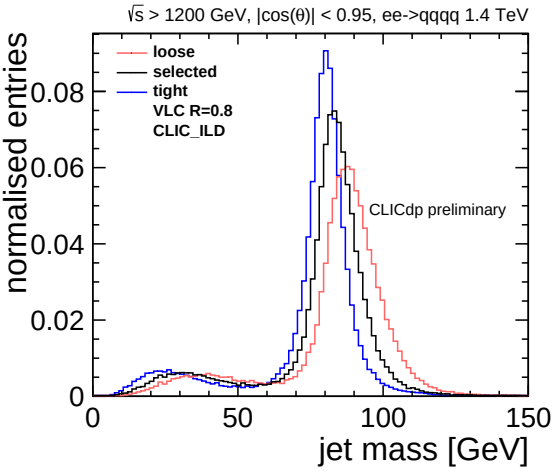
2.1. Comparison at 1.4 TeV

Figure 2 (top) shows the jet mass distributions for WW events for the CLIC_ILD and the CLICdet model at 1.4 TeV using two exclusive jets. The largest differences between the models can be observed in the loose selected PFOs, where for CLICdet the distribution is narrower and peaked at smaller masses. A possible cause could be the timing window used in the HCAL barrel which was changed from 100 to 10 ns [3] when changing from tungsten to steel. It might be interesting to revise the study on the initial energy stage in the light of the differences in the loose selected PFO distributions. The selected distribution is slightly narrower and peaks closer to the W mass for CLICdet while there are no distinct differences in the W jet mass distributions for the tight selected collection observed. Comparing the different timing cuts for CLICdet with the line representing all PFOs without overlay, the tight selected distribution is closest to the reference. For studies involving hadronically decaying W bosons, the tight selected PFO collection should therefore be already considered at the intermediate energy stage.

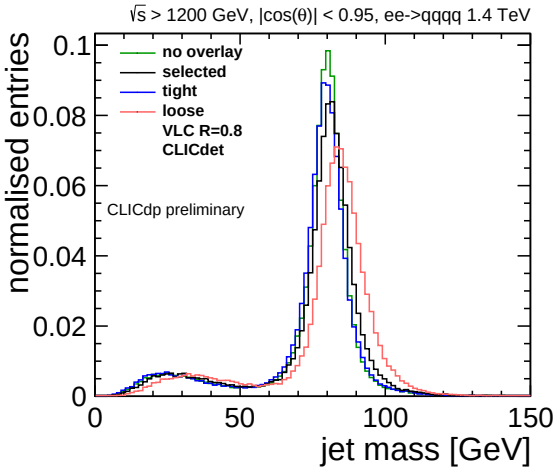
2.2. Comparison at 3 TeV

Figure 2 (bottom) shows the jet mass distributions for WW events for the CLIC_ILD and the CLICdet model at 3 TeV. Similarly to the case at 1.4 TeV, the distributions for loose selected PFOs differ the most. Smaller differences can be observed for the selected and tight selected collections. The distributions for CLICdet are slightly narrower in general and peaked at lower masses, closer to the W mass. The tight selected PFO collection is closest to the reference without overlay and is therefore considered to be the best choice for CLICdet at 3 TeV, as expected.

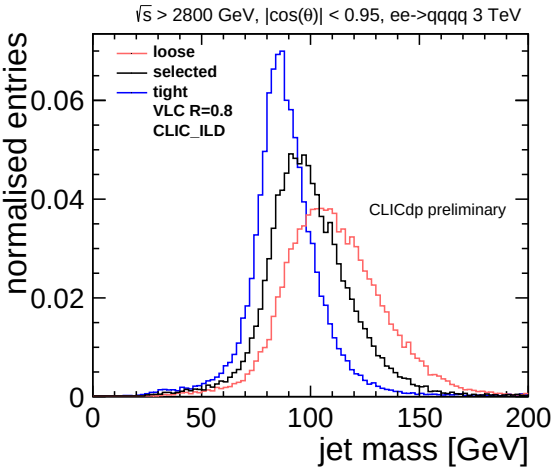
Overall, the timing cuts work without re-optimisation for the new CLICdet detector model.



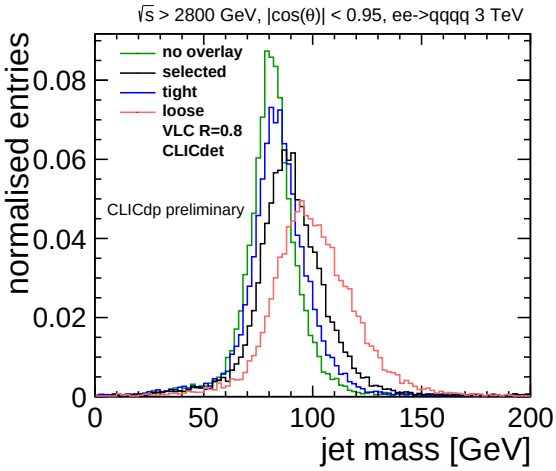
(a) CLIC_ILD at 1.4 TeV



(b) CLICdet at 1.4 TeV



(c) CLIC_ILD at 3 TeV



(d) CLICdet at 3 TeV

Figure 2: Jet mass distributions in WW events for the CLIC_ILD (left) and the new CLICdet detector model (right) at 1.4 TeV (top) and 3 TeV (bottom) using different timing cuts.

3. Separation of WW and ZZ

The separation of W and Z jet masses is a good benchmark to evaluate the CLIC detector performance. We first optimise the jet clustering for a good separation with the CLIC_ILD model and then compare the achieved separation for both detector models. The tight selected PFO collections were used for this study. Similar studies have been performed in [1, 2].

3.1. Jet clustering optimisation

The generalised e^+e^- k_T-algorithm (ee-genkt) is compared to the k_T-algorithm used at hadron colliders and the VLC algorithm. The resulting jet mass distribution in WW events for a radius of R = 1.0 and standard parameters is shown in Figure 3. While the distributions for VLC and k_T appear quite similar, the ee-genkt is missing the low-mass bump and instead features a long tail to higher masses. The bump at low masses appears when the jet does not catch all W decay products, e.g. because the radius is too small. For the separation of W and Z jets, a long tail to high masses is less desired. The ee-genkt

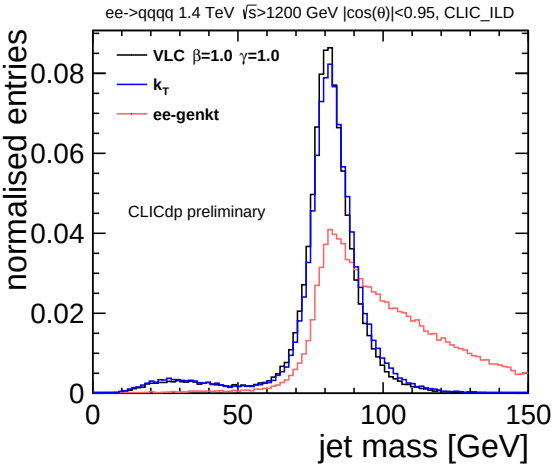


Figure 3: Jet mass distributions in WW events for CLIC_ILD using different jet algorithms.

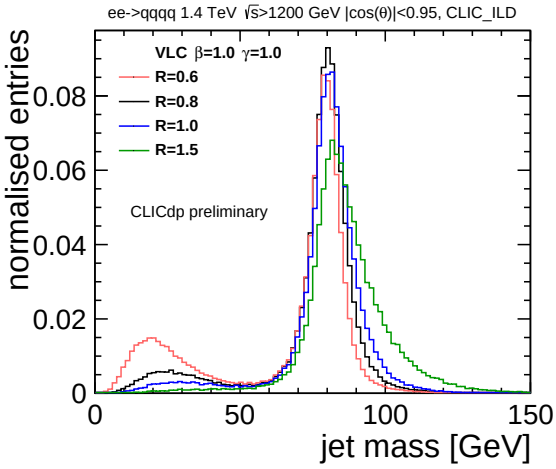


Figure 4: Jet mass distributions in WW events for CLIC_ILD using VLC with different jet radii.

algorithm is therefore discarded in the following. Figure 4 shows the performance of VLC with different radius parameters. The jet mass distribution for a low jet radius, here R = 0.6, has a sharp peak at the W mass but features a pronounced bump at low masses because the active area is too small. A large jet radius catches the whole W but also picks up undesired energy depositions resulting in a longer tail to high masses. These depositions are mostly remnants from the $\gamma\gamma \to$ hadrons background.

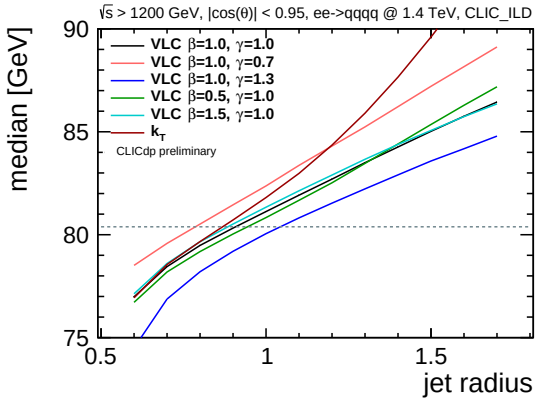
Due to the asymmetric, non-gaussian shape of the jet mass distributions, the median and the IQR_{34} were chosen as figures of merit for the optimisation of the jet clustering parameters, following the procedure in [5]. The IQR_{34} is defined as the interval around the median that contains 68% of the distribution. Figure 5 shows the results for the median and the IQR_{34}/median for different jet clustering options. The optimal settings are found to be VLC with R = 0.8, $\beta = 1.0$ and $\gamma = 0.7$, yielding a low IQR_{34}/median and a median near the W mass. The differences between positive values of β are observed to be negligible. This is consistent with the expectation from the definition of the distance parameter Equation (1). A value lower than 1.0 for γ increases the active area for forward jets relative to higher values, reducing the low mass tail. This effectively reduces the width of the distribution and thereby the IQR_{34}. Even lower values of γ lead to an increased tail to high masses, see further examples in Appendix A. The median shows the expected rising nature when increasing the radius.

3.2. Results at 1.4 TeV

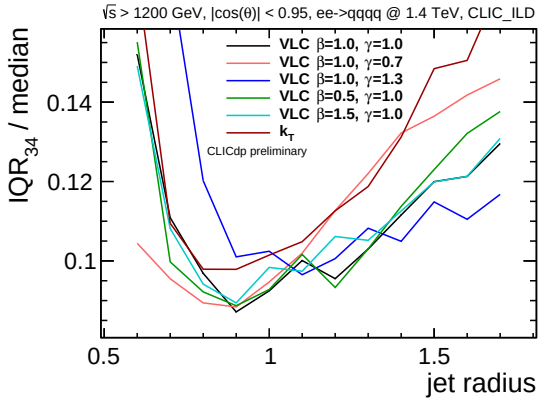
The resulting invariant mass distributions for WW and ZZ events with CLIC_ILD and the optimised clustering settings are shown in Figure 6(a). The distributions are nicely separated when scaled to the same number of entries, with a distance between both medians is around 9 GeV. This shows a good performance of the detector design, PFA and timing cuts.

Figure 7 shows the jet mass distributions in WW and ZZ events for CLIC_ILD at 1.4 TeV and at the expected SM cross-section ratio with optimised jet clustering parameters. This case is important for a study of the $\mathrm{e^+e^-} \to \mathrm{W^+W^-}$ cross-section. The jet mass distribution for ZZ events lies in the tail of the WW distribution, it is fairly difficult to separate both jet mass peaks.

Figure 6(b) shows the performance when running the VLC algorithm with four exclusive jets. The same clustering parameters were found to perform best. The jets were paired to W and Z candidates via minimising $\chi^2 = (m_{ij} - m_{\mathrm{W/Z}})^2 + (m_{kl} - m_{\mathrm{W/Z}})^2$. Due to the combination of two jets, it is less likely to loose decay products of the W or Z even for large angles between the quarks. This effectively reduces



(a) Median of the jet mass distribution in WW events.



(b) IQR_{34}/median of the jet mass distribution in WW events.

Figure 5: Results of the optimisation for CLIC_ILD with the k_T-algorithm and VLC with different values for β and γ.

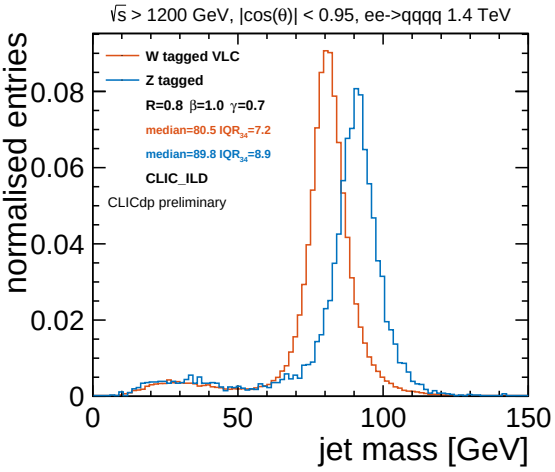
the lower mass bump and thereby reducing the IQR_{34} and improving the separation power by about 15%.

The median of the jet mass distributions with CLICdet is slightly lower compared to CLIC_ILD. This was also observed when comparing the timing cut collections in Section 2. A slightly larger radius parameters of R = 0.9 was chosen to compensate for this. The results are shown in Figure 6(c) and Figure 6(d). The same β and γ values used for CLIC_ILD yield optimal results. The overall result is similar to the CLIC_ILD case with a separation of about 9 GeV between the medians. As previously observed, using four exclusive jets help to reduce the low-mass bump (especially for the Z). But in this case the bump is already less pronounced due to the larger radius used. A visualisation of the achieved separation is given in Figure 8 showing the masses of both jet combinations.

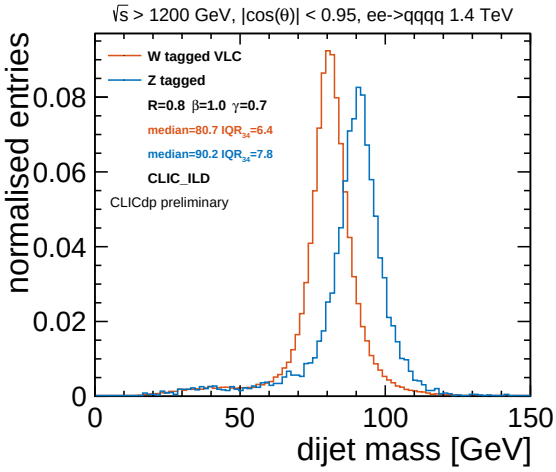
We further evaluated the W/Z separation performance in terms of gaussian standard deviations. Figure 9 shows the jet mass distributions for W and Z candidates with a gaussian fitted to the core of the distributions. The separation is then given by the distance between both mean values normalised to the standard deviation of the W candidate distribution. CLIC_ILD reaches 1.7 σ and CLICdet 1.6 σ separation. This is in good agreement with the particle flow performance study in [1]. The cut is placed at the intersection of the gaussian fits. The total misidentification, for the same amount of W and Z events, is then defined as the part of the Z candidate distribution left of the cut plus the part of the W candidate distribution right of the cut, divided by the sum of both distributions. Achieved are rates close to 30%, the theoretical minimum using the natural W and Z widths is given by 5.6% [1].

3.3. Results at 3 TeV

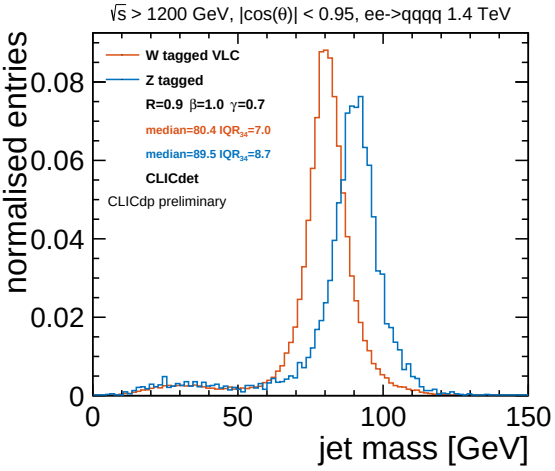
Events at 3 TeV are expected to be more boosted, resulting in a smaller opening angle between the decay products of the W and Z bosons. This can be estimated via the rule of thumb $\Delta R \approx 2m/p_T$. The optimal jet radius is therefore expected to be smaller compared to the 1.4 TeV case. The optimal settings found are R = 0.5, $\beta = 1.0$ and $\gamma = 0.7$ for CLICdet and R = 0.4 for CLIC_ILD (same β and γ), yielding the best compromise of median position, width and low mass tail. The separation for CLIC_ILD and CLICdet with these settings are shown in Figure 10. The medians are separated by around 9 GeV, similar to the 1.4 TeV case. The overall separation of W and Z candidates however is worse due to the wider distributions. This translates to a separation of around 1.1–1.3 σ and a misidentification rate slightly above 30% as can be seen in Figure 12. Clustering four exclusive jets does not have the same impact as for the 1.4 TeV event samples due to the small angle between the quarks from the W or Z decay, respectively. The resulting distributions can be found in Appendix C.



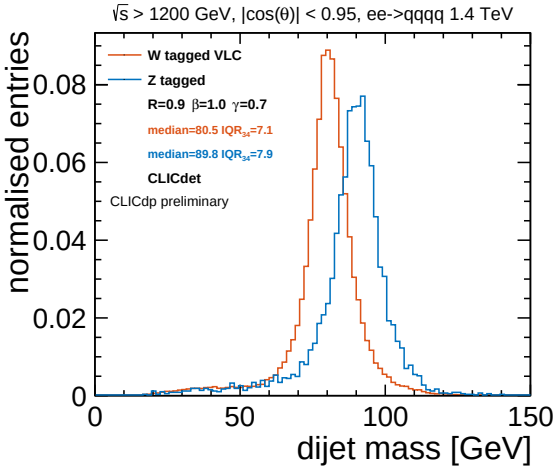
(a) CLIC_ILD, two exclusive jets.



(b) CLIC_ILD, four exclusive jets.



(c) CLICdet, two exclusive jets.



(d) CLICdet, four exclusive jets.

Figure 6: Jet mass distributions in WW and ZZ events for CLIC_ILD (top) and CLICdet(bottom) at 1.4 TeV using optimised jet clustering parameters with two exclusive jets (left) and four exclusive jets (right).

4. Conclusions and outlook

This note discussed the performance for the reconstruction of hadronic W decays of the new CLICdet detector model compared to the CLIC_ILD model from the CDR. We studied the timing cut collections used to minimise the impact of $\gamma\gamma \to$ hadrons for both detector models. Furthermore, the W/Z separation in the full hadronic channel was used as benchmark to test the CLICdet design with overlay and to compare to CLIC_ILD. The jet clustering was optimised in terms of median, IQR_{34} and separation.

The timing cuts are found to work for CLICdet out of the box. The W candidate mass distribution with the loose PFO collection for CLICdet is much narrower, presumably due to a changed timing window for the calorimeter. For some applications this can already be too tight, an even looser collection might be worth studying. The tight PFO collection was found to yield results closest to the reference without overlay at both considered energies for CLICdet. Using the tight PFO collection already at the intermediate energy stage should be considered for the measurement of $\mathrm{e}^+\mathrm{e}^- \to \mathrm{W}^+\mathrm{W}^-$.

Both models achieve a similar separation of the median of the jet mass distributions in WW and ZZ

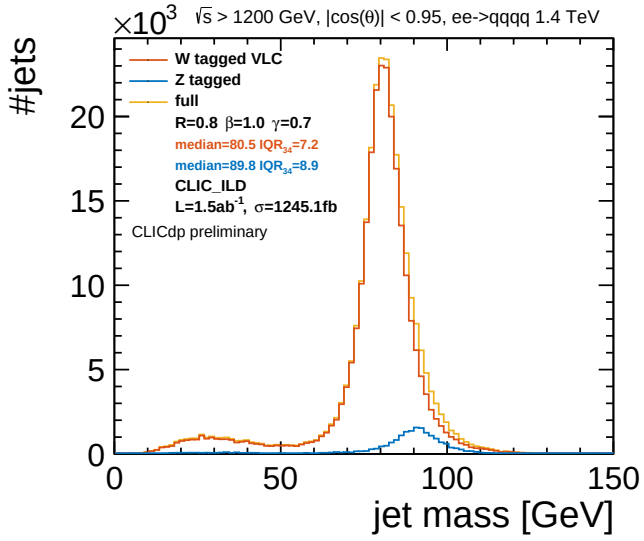


Figure 7: Jet mass distributions in WW and ZZ events for CLIC_ILD at the expected SM cross-section ratio using two exclusive jets.

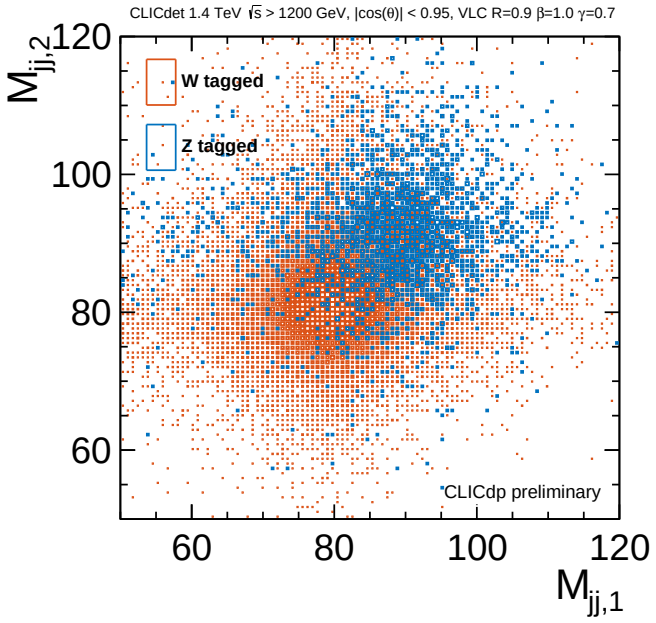
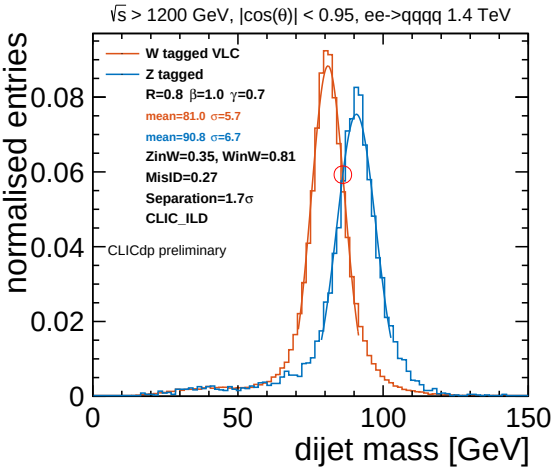


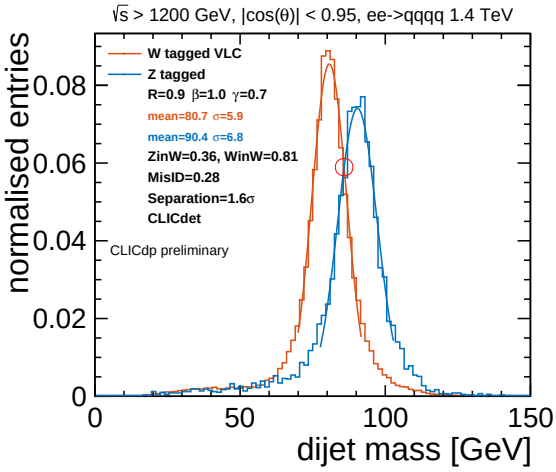
Figure 8: Masses of both resulting combinations of four exclusive jets in WW and ZZ events for CLICdet at 1.4 TeV, normalised entries.

events of the order of 8–9 GeV. Assuming a gaussian distribution, the separation is found to be around 1.6 σ at 1.4 TeV and around 1.3 σ at 3 TeV. Clustering four exclusive jets was found to be helpful at 1.4 TeV to reduce the low mass tail. At 3 TeV clustering in a two jet exclusive mode performs equally well.

In this note we studied the exclusive clustering to two and four jets. Additionally, one could think about intermediate solutions e.g. using different number of jets on an event by event basis depending on merging scales. Another interesting field to study in this context is the use of jet grooming techniques such as trimming used at the LHC. These could help to reduce the $\gamma\gamma \rightarrow$hadrons background, replacing

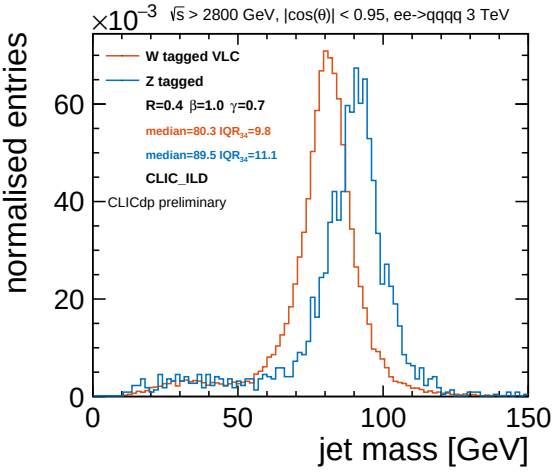


(a) CLIC_ILD.

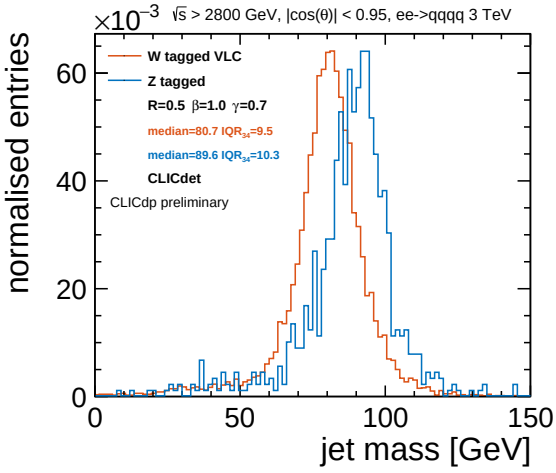


(b) CLICdet.

Figure 9: Separation between W and Z candidates in terms of gaussian widths for CLIC_ILD and CLICdet at 1.4 TeV using four exclusive jets.



(a) CLIC_ILD, R = 0.4, $\beta = 1.0$ and $\gamma = 0.7$.



(b) CLICdet, R = 0.5, $\beta = 1.0$ and $\gamma = 0.7$.

Figure 10: Separation of W and Z jets for CLIC_ILD (left) and CLICdet (right) at 3 TeV using optimised jet clustering parameters with two exclusive jets.

or in addition to the timing cuts.

Further studies with truth particles would be beneficial to isolate the different contributions from clustering effects, detector resolution and smearing due to $\gamma\gamma$ background. In addition, we see potential in further studying the correlation between the boson candidate masses, as illustrated in figure 8. This could give us a better handle to improve separation between fully-hadronic W/W and Z/Z decays.

References

[1] J. S. Marshall, A. Munnich, M. A. Thomson, *Performance of Particle Flow Calorimetry at CLIC*, Nucl. Instrum. Meth. CU-HEP-12-12 **A700** (2013) 153, DOI: 10.1016/j.nima.2012.10.038.

[2] L. Linssen et al., eds., *Physics and Detectors at CLIC: CLIC Conceptual Design Report, vol.* CERN-2012-003, CERN, 2012, URL: http://cds.cern.ch/record/2254048.

[3] N. Alipour Tehrani et al., *CLICdet: The post-CDR CLIC detector model*, CLICdp-Note-2017-001 (2017), URL: http://cds.cern.ch/record/2254048.

[4] M. J. Boland et al., CLICdp, CLIC, *Updated baseline for a staged Compact Linear Collider*, CERN-2016-004 (2016), ed. by P. Lebrun et al., DOI: 10.5170/CERN-2016-004.

[5] M. Boronat et al., *Jet reconstruction at high-energy lepton colliders*, CLICDP-PUB-2017-002 (2016), arXiv: 1607.05039 [hep-ex].

A. Different values of γ and β

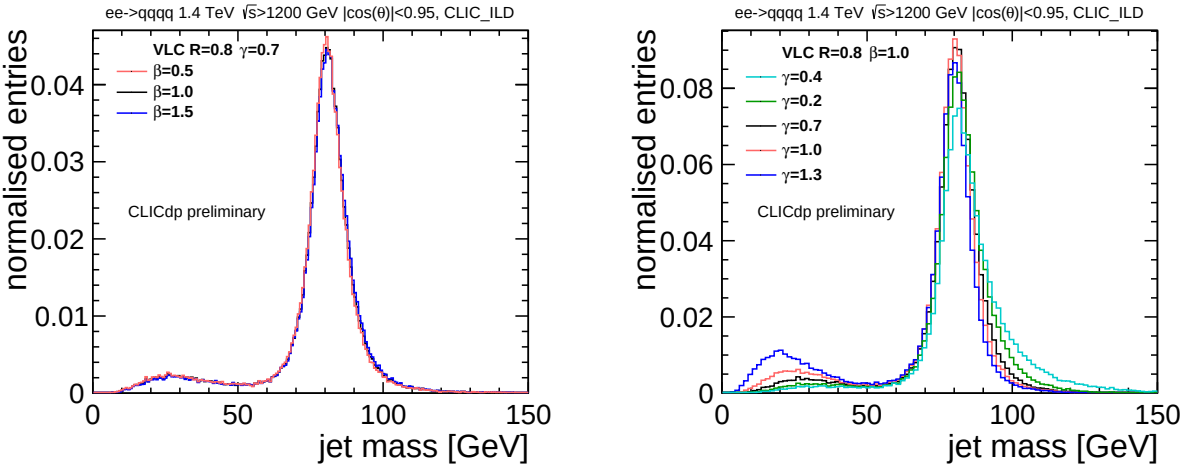
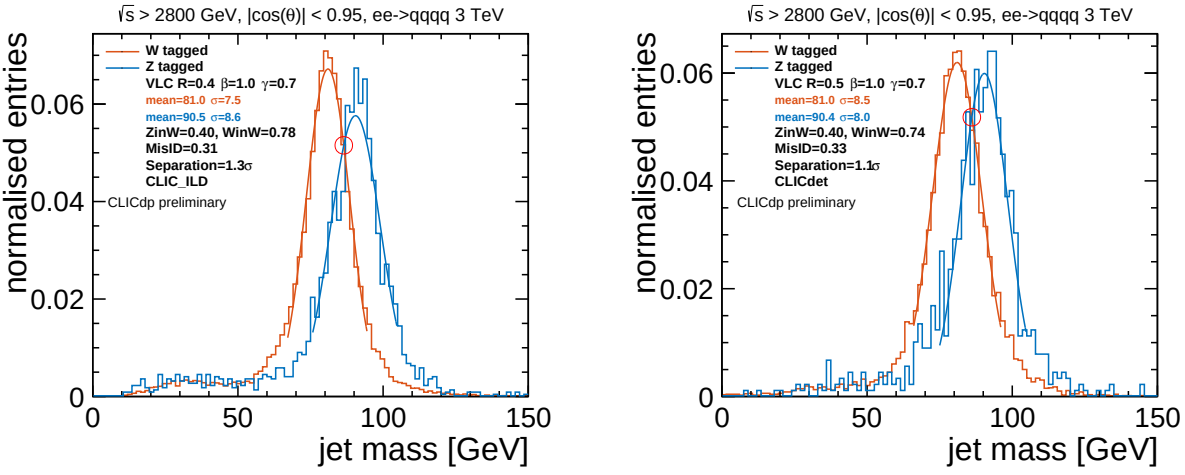


Figure 11: Jet mass distributions in WW events for CLIC_ILD at 1.4 TeV using VLC with different values of β (left) and γ (right). While low values of γ lead to undesired tails to higher masses, high values lead to an increase of mis-reconstructed events manifested as a small peak at low mass. Varying β has a very small impact on the jet mass distribution.

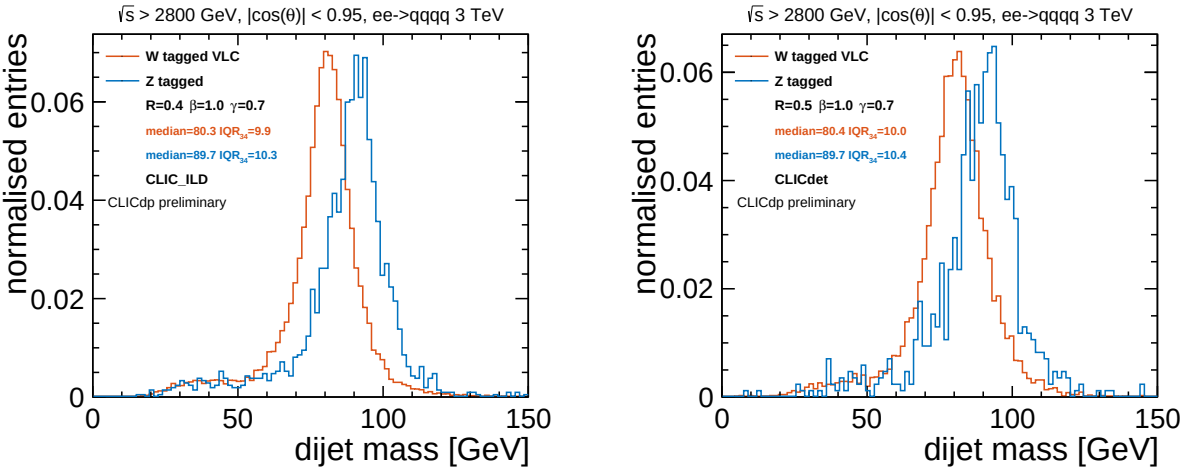
B. Gaussian fits to the W and Z candidate distributions at 3 TeV



(a) CLIC_ILD (b) CLICdet

Figure 12: Separation between W and Z candidates in terms of gaussian widths for CLIC_ILD and CLICdet at 3 TeV.

C. Four exclusive jets at $\sqrt{s} = 3\,\text{TeV}$



(a) CLIC_ILD

(b) CLICdet

Figure 13: Separation of W and Z jets for CLIC_ILD (left) and CLICdet (right) at 3 TeV using optimised jet clustering parameters with four exclusive jets.