Reconstruction and classification of tau lepton decays at a future electron-positron Linear Collider

B. Xu S. Green, J. S. Marshall, M. A. Thomson

Cavendish Laboratory, University of Cambridge,
JJ Thomson Avenue, Cambridge, CB3 0HE, UK

E-mail: xu@hep.phy.cam.ac.uk

ABSTRACT: The ability to identify tau decay mode is important in a number of physics measurements, and provides a benchmark for detector performance. The reconstruction and the classification of major tau decay modes is studied, for a future electron-positron Linear Collider. This paper also presents an electromagnetic calorimeter optimisation study with varying square cell sizes, at different centre of mass energies. For the ILC 500 GeV and the CLIC 350 GeV scenarios, larger ECAL square cell sizes have a small impact on the tau hadronic decay classification probability. For the ILC 1 TeV and the CLIC 1.4 TeV and 3 TeV scenarios, a small ECAL cell size is needed to achieve a high classification probability,

1 Introduction

The tau lepton has been studied extensively in the past at the Large Electron Positron Collider (LEP)[1]. Tau lepton decays provide precision tests of the Standard Model and models beyond the Standard Model[2]. The spin state of the tau lepton, which can be inferred from kinematic properties of tau decay products, can be used to measure the CP (the product of charge conjugation and parity symmetries) of the Higgs, via H $\to \tau^+ \tau^-$ channel[3]. These measurements rely on the ability to identify and isolate different tau decay modes.

The ability to identify tau decay mode is important in a number of physics measurements, and therefore provides a benchmark for detector performance. Reconstructing tau decay products is crucial for tau decay mode measurements, as the lifetime of the tau is too short to be able to measure directly [4]. The performance of the calorimetric and track systems determines the ability to separate different tau decay modes.

This paper describes the reconstruction and classification of tau lepton decay modes. The classification is used for the ECal optimisation, with varying ECal square cell sizes of 3, 5, 7, 10, 15 and 20 mm, using different centre of mass energies ($\sqrt{s}$) of 100, 200, 500 and 1000 GeV for $e^+ e^- \to \tau^+ \tau^-$ interaction. The International Large Detector (ILD) concept [5] for the International Linear Collider (ILC)[6] is used for this study. Layout of the ILD detector is shown in figure 1. The detector consists of a vertex detector, a tracking system compromising a large time projection chamber (TPC) augmented with silicon tungsten layer, highly granular electromagnetic calorimeters (ECAL) and hadronic calorimeters (HCAL), muon chambers, forward calorimeters (FCAL), magnetic coils and iron yokes. The magnetic field strength is optimised for the particle flow approach[7] to calorimetry with high granularity in both the longitudinal and transverse directions.
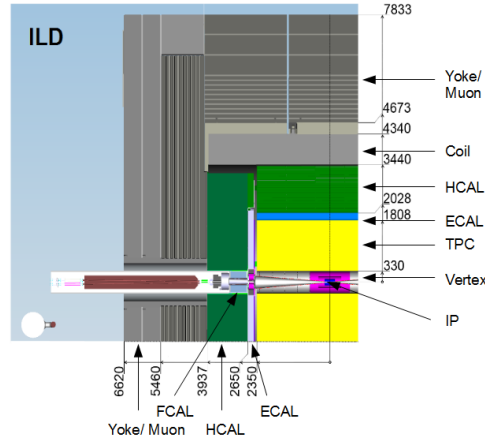
Figure 1: Longitudinal cross section of top quadrant of the ILD detector concept, taken from [8]. From interaction point (IP) outwards, there is a tracking system compromising a large time projection chamber (TPC) augmented with silicon tungsten layer, highly granular electromagnetic calorimeters (ECAL) and hadronic calorimeters (HCAL), muon chambers, forward calorimeters (FCAL), magnetic coils and iron yokes. Numbers are in units of mm.

The efficient tau lepton decay mode separation relies on the ability to reconstruct decay particles. Reconstruction of individual particles relies on highly granular detectors and sophisticated pattern recognition software to exploit the information. The PandoraPFA [9] particle flow software implements a sophisticated multi-algorithm particle reconstruction. It has been used in physics simulation and detector optimisation studies for future linear collider, such as the Compact Linear Collider[10] or the International Linear Collider[11]. Fitted helices of the hits in the tracking systems are projected to the front of the ECAL to associate calorimeter hits, and to form charged particles. Photons are reconstructed from calorimeter hits in the ECAL. A match is required between the photon clusters and the expected electromagnetic shower profiles. Neutral hadrons are reconstructed from calorimeter hits in the ECAL and the HCAL. The output of the PandoraPFA is a list of reconstructed particles, "Particle Flow Objects" (PFOs), which contains energy, momentum, position and particle id information.

The ability to accurately separate tau decay modes relies strongly on the efficient photon reconstruction. Calorimeter hits in the ECal are clustered and carefully isolated from nearby charged particles. These photon candidates are compared with expected electromagnetic shower profiles. A likelihood based classifier is used to determine the photon id. Separate algorithms split merged photons and combine small photon fragments to main photons. Photon reconstruction in PandoraPFA is described in [12].

This paper is organised in the following way: event simulation and reconstruction, selected decay modes, pre-selection, discriminating varibles, $\rho\left(\pi^-\pi^0\right)$ and $a_1\left(\pi^-\pi^0\pi^0\right)$ resonance, multivariate analysis, results and the ECAL optimisation study.

2 Simulation and Reconstruction

Monte Carlo (MC) samples were generated with the generator software WHIZARD 1.95 [13]. PYTHIA 6.4 [14], tuned to the LEP results [], was used for the hadronisation. TAUOLA [15] was used to describe the tau lepton decays, including the effect of the spin correlation between decay products. Events were simulated with software MOKKA [16], based on the GEANT 4 package [17], with the ILD detector geometry description. Events reconstruction were performed using ilcsoft version v01-17-07 and PandoraPFA version v02-02-00, in the MARLIN software framework[18]. Two million $e^+ e^- \to \tau^+ \tau^-$ MC events were generated, for each ECal square cell size, and for each $\sqrt{s}$. In total, 48 million simulated events were generated. To maximise the impact of the ECAL design on the performance of the tau decay classification, the initial state radiation (ISR) and the beam induced background were not simulated, as these effects degrade the performance by the same amount irrespective of the ECAL design.

3 Selected decay modes

Tau lepton decays into a numerous number of final states. To study the major effect of the tau lepton decays, decay modes with branching ratio above 1% were chosen to study. This resulted in seven tau lepton decay modes considered, shown in table 1, which 92.58 % branching ratios of all tau lepton decays.

Decay Mode	Final State	Branching fraction / %
$\tau^- \to e^- \bar{\nu}_e \nu_\tau$	$e^- \bar{\nu}_e \nu_\tau$	17.83±0.04
$\tau^- \to \mu^- \bar{\nu}_\mu \nu_\tau$	$\mu^- \bar{\nu}_\mu \nu_\tau$	17.41±0.04
$\tau^- \to \pi^- \nu_\tau$	$\pi^- \nu_\tau$	10.83±0.06
$\tau^- \to \rho \left(\pi^- \pi^0 \right) \nu_\tau$	$\pi^- \pi^0 \nu_\tau$	25.52±0.09
$\tau^- \to a_1 \left(\pi^- \pi^0 \pi^0 \right) \nu_\tau$	$\pi^- \pi^0 \pi^0 \nu_\tau$	9.30±0.11
$\tau^- \to a_1 \left(\pi^- \pi^- \pi^+ \right) \nu_\tau$	$\pi^- \pi^- \pi^+ \nu_\tau$	8.99±0.06
$\tau^- \to \pi^- \pi^- \pi^+ \pi^0 \nu_\tau$	$\pi^- \pi^- \pi^+ \pi^0 \nu_\tau$	2.70±0.08

Table 1: Branching fractions[19] of the seven τ^- decay modes and final states studied. τ^+ decays similarly to τ^-.

4 Analysis

4.1 Event pre-selection

Some events are very difficult or almost impossible to reconstruct correctly. This can happen if particles are in the forward beam pipe direction, where the forward calorimeters are not simulated due to computational limitations, or if particles carry too little energies to be significant than noises. Photons that converted into electron pairs in the tracking systems are reconstructed as an electron pair, which will alter the event topology and decrease the classification efficiencies. Also for the ILD detector model concept, the gap region between barrel part of the calorimeter and the end cap part causes a significant drop in the particle reconstruction efficiency. These effects do not vary with change of the ECAL square cell sizes in the later optimisation study. Therefore, unreconstructable events are discarded from the subsequent analysis. The selection efficiency of each cut is presented in table 2, using nominal ILD detector model and $\sqrt{s} = 100$ GeV.

Selection Efficiency in %	No $\gamma \to e^+ e^-$	No $\gamma \to e^+ e^-$ and $E_{vis} > 5 GeV$	No $\gamma \to e^+ e^-$, $E_{vis} > 5 GeV$, and decay in barrel or end cap
$e^- \bar{\nu}_e$	100.0	84.7	66.2
$\mu^- \bar{\nu}_\mu$	100.0	85.2	66.7
π^-	100.0	88.3	60.9
$\rho \left(\pi^- \pi^0 \right)$	77.1	76.9	61.9
$a_1 \left(\pi^- \pi^0 \pi^0 \right)$	61.3	61.2	50.5
$a_1 \left(\pi^- \pi^- \pi^+ \right)$	100.0	100.0	78.0
$\pi^- \pi^- \pi^+ \pi^0$	77.0	77.0	61.8

Table 2: Pre-selection efficiencies for seven tau decay modes, $e^- \bar{\nu}_e \nu_\tau$, $\mu^- \bar{\nu}_\mu \nu_\tau$, $\pi^- \nu_\tau$, $\rho \left(\pi^- \pi^0 \right) \nu_\tau$, $a_1 \left(\pi^- \pi^0 \pi^0 \right) \nu_\tau$, $a_1 \left(\pi^- \pi^- \pi^+ \right) \nu_\tau$, and $\pi^- \pi^- \pi^+ \pi^0 \nu_\tau$, using nominal CLIC_ILD detector model with $\sqrt{s} = 100$ GeV.

An low visible energy event is discarded if the total energy of the tau lepton visible decay

products is below $5\,\text{GeV}$. Requirement of the tau decay in the barrel or the end cap part of the calorimeter is defined as the polar angle of the tau is $17.2° < |\theta_Z| < 34.4°$ or $45.8° < |\theta_Z| < 90°$.

4.2 Select single tau decay

For convenience, $e^+ e^- \to \tau^+ \tau^-$ were used. Each of the taus was examined in turn. To select single tau decay, the scalar dot product between the thrust axis and the momentum vector is used to separate particles into one of the two collections. Thrust axis is the normal vector, $\hat{n}$, that maximises, $T = max_{\hat{n}} \frac{\Sigma_i |p_i \cdot \hat{n}|}{\Sigma_i |p_i|}$, where p_i is the momentum vector of the ith particle.
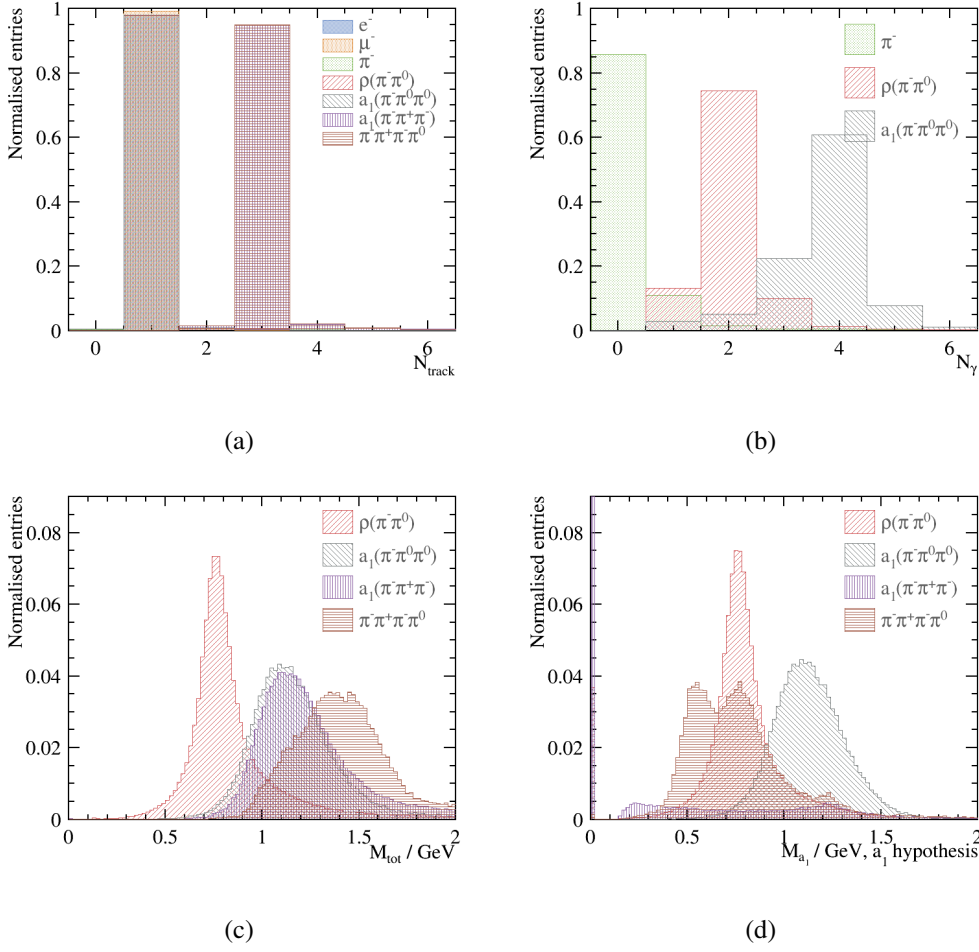
4.3 Discriminating variables

Having pre-selected samples, a set of discriminating variables was carefully developed to be used in the later multivariate analysis (MVA). These variables allow the multivariate classifier achieves a highly efficient tau decay mode classification. Different variables have discriminative powers against different final states. For convenience, variables are presented sequentially, although they participate the MVA simultaneously.

To differentiate 1-prong final states from 3-prong final states, number of tracks, shown in figure 2a, provides excellent separation power. Selection efficiencies are over 95% for the 1-prong final states, and over 90% for the 3-prong final states. For the 1-prong final states, they are further separated based on the number of photons reconstructed. Clear distinction between $\pi^- \nu_\tau$, $\rho\left(\pi^-\pi^0\right) \nu_\tau$ and $a_1\left(\pi^-\pi^0\pi^0\right) \nu_\tau$ decays mode can be seen in figure 2b, where the overlap between any two decay modes is below 15%. The two leptonic decay modes, $\mu^- \bar{\nu}_\mu \nu_\tau$ and $e^- \bar{\nu}_e \nu_\tau$, have very different event topologies to other decay modes. For the $\mu^- \bar{\nu}_\mu \nu_\tau$ decay mode, only muons deposit energies in the muon chamber. For the $e^- \bar{\nu}_e \nu_\tau$ decay mode, electrons shower in the ECAL can be modelled by electromagnetic shower profiles. This information is used to distinguish a electron from a charged pion, including comparison of the observed and the expected electromagnetic shower profile, the matching between inner detector tracks and calorimeter clusters, and the fraction of calorimeter hits registered as minimum ionised particles.

4.4 Reconstruction of $a_1\left(1260\right)$ and $\rho\left(770\right)$ resonances

Varibles described above separate 1-prong final states from 3-prong final states, and separate different 1-prong final states. Figure 2c shows the differences in the total invariant mass. However, there is an overlap between the mass distribution for different decay modes. To further enhance the separation of the $a_1\left(\pi^-\pi^0\pi^0\right) \nu_\tau$ and $\rho\left(\pi^-\pi^0\right) \nu_\tau$, the $a_1\left(1260\right)$ and $\rho\left(770\right)$ resonances are reconstructed by considering all possible combination of reconstructed photon candidates. For example, the $a_1\left(1260\right) \to \pi^- \pi^0 \pi^0$ decay produces four photons. If two photons are very close together, they can be merged in the reconstruction. Alternatively, if some of the energy deposition from the π^- appears as fragments, additional photon photons can be reconstructed. To enhance the resonance reconstruction, $a_1\left(\pi^-\pi^0\pi^0\right) \nu_\tau$ decay is reconstructed by minimising χ_a^2, defined as

$$\chi_a^2 = \left(\frac{m_a^* - m_a}{\sigma_a}\right)^2 + \left(\frac{m_{\pi^0,1}^* - m_{\pi^0}}{\sigma_{\pi^0,1}}\right)^2 + \left(\frac{m_{\pi^0,2}^* - m_{\pi^0}}{\sigma_{\pi^0,2}}\right)^2, \tag{4.1}$$

Figure 2: Normalised distribution for discriminative variables for seven tau decay modes, $e^- \bar{\nu}_e \nu_\tau$, $\mu^- \bar{\nu}_\mu \nu_\tau$, $\pi^- \nu_\tau$, $\rho\left(\pi^-\pi^0\right)\nu_\tau$, $a_1\left(\pi^-\pi^0\pi^0\right)\nu_\tau$, $a_1\left(\pi^-\pi^-\pi^+\right)\nu_\tau$, and $\pi^-\pi^-\pi^+\pi^0\nu_\tau$, with $\sqrt{s}$ = 100 GeV for nominal ILD detector model. Figure 2a is the number of tracks. Figure 2c is the invariant mass of visible PFOs. Figure 2b is the number of photons. Figure 2d is the invariant mass of $a_1\left(\pi^-\pi^0\pi^0\right)$, calculated with $a_1(1260) \rightarrow \pi^- \pi^0 \pi^0$ resonance reconstruction.

where $m^*_{\pi^0,1}$ and $m^*_{\pi^0,2}$ are the invariant masses of two photons combinations, σ_a and σ_{π^0} are the half width of the invariant mass distribution of reconstructed a and π^0 using the truth information, and m_a and m_π are the masses of a and π^0, taken from [19]. This improves the invariant mass of the $a_1(1260)$ resonance. Shown in figure 2d, only the $a_1\left(\pi^-\pi^0\pi^0\right)\nu_\tau$ decay mode has a peak structure at the $a_1(1260)$ resonance, and there is a little overlap between $a_1\left(\pi^-\pi^0\pi^0\right)\nu_\tau$ decay mode and other decay modes comparing to the distribution in the figure 2c.

The $\rho(770) \rightarrow \pi^-\pi^0$ decay has the similar reconstruction issue, and is enhanced with a similar χ^2 minimisation. It results in similar improvement in $\rho(770)$ resonance reconstruction and better separation between the $\rho\left(\pi^-\pi^0\right)\nu_\tau$ decay mode and other decay modes.

4.5 Multivariate analysis

A set of 29 discriminating variables was used to classify tau decays into one the following decay modes: $e^- \bar{\nu}_e \, \nu_\tau$, $\mu^- \bar{\nu}_\mu \, \nu_\tau$, $\pi^- \, \nu_\tau$, $\rho \left(\pi^- \pi^0 \right) \nu_\tau$, $a_1 \left(\pi^- \pi^0 \pi^0 \right) \nu_\tau$, $a_1 \left(\pi^- \pi^- \pi^+ \right) \nu_\tau$, and $\pi^- \, \pi^- \, \pi^+ \, \pi^0$ ν_τ. The variables fall into 5 distinctive categories. Energy variables include:

- Total energy deposited in ECal and HCal, divided by the total energy of charged PFOs,

- Total energy deposited in ECal and HCal, divided by the total energy of all PFOs,

- Invariant mass of total PFOs in GeV,

- Total energy of all PFOs, divided by the energy of the τ^-,

- Total energy of charged PFOs, divided by the energy of τ^-,

- Total energy of muons, divided by the energy of τ^-,

- Total energy of electrons, divided by the energy of τ^-,

- Total energy of photons, divided by the energy of τ^-,

- Total energy of charged pions, divided by the energy of τ^-

where the energy of the τ is taken to be the same as the energy of $e^\pm$ beam, which is half of the $\sqrt{s}$ energy.

Invariant mass variables include:

- Invariant mass of charged PFOs in GeV,

- Invariant mass of neutral PFOs in GeV,

- Invariant mass of photons in GeV,

- Invariant mass of charged pions in GeV.

Number of PFOs variables include:

- Number of charged PFOs,

- Number of muons,

- Number of electrons,

- Number of photons,

- Number of charged pions.

Variables to distinguish electrons from charged pions include:

- Average energy deposited in a calorimeter cell in GeV,

- Average transverse shower width, for all clusters in the ECal,

- Average longitudinal shower start layer for all clusters in the ECal,

- Average discrepancy in longitudinal electromagnetic shower between the observed and the expected profile, for all clusters in the ECal,

- Average fraction of calorimeter hits registered as minimum ionised particles, all clusters in the ECal,

- Average ratio between energy and momentum, for all clusters in the ECal,

The $a_1(1260)$ and $\rho(770)$ resonances reconstruction variables include:

- Best reconstructed invariant mass of $a_1(1260)$, for $a_1\left(\pi^-\pi^0\pi^0\right)$ resonance reconstruction,

- Best reconstructed invariant mass of π^0, for $a_1\left(\pi^-\pi^0\pi^0\right)$ resonance reconstruction,

- Second best reconstructed invariant mass of π^0, for $a_1\left(\pi^-\pi^0\pi^0\right)$ resonance reconstruction,

- Best reconstructed invariant mass of $\rho(770)$, for $\rho\left(\pi^-\pi^0\right)$ resonance reconstruction,

- Best reconstructed invariant mass of π^0, for $\rho\left(\pi^-\pi^0\right)$ resonance reconstruction.

For the multivariate analysis, the multiclass of the TMVA package [20] was used to perform a multiclass classification, which trains the seven final states simultaneously. The multiclass class is an extension of the standard two-class signal-background classifier. For each final state, the multiclass classifier will train the final state as the signal against all other final states as the background. This process is repeated for each final state. The classifier output for a single event is a normalised response for each final state, where the sum is one. The response of each final state of an event can be treated as the likelihood. The event is classified into a particular final state if the final state has the highest classifier output response. The advantage of using the multiclass is that the correlation between different final states are accounted for and the classifier output are correctly adjusted for multiple final states, hence one event can only be classified into one final state. The issue with the multiclass is that discriminative variables for each decay mode need enter the training stage, resulting in a large number of variables.

Half of the events, randomly selected, were used in the training process and the other half were used for testing. The TMVA multiclass classifier is a boosted decision tree with gradient boosting (BDTG), as it was found to give for the best performance. The MVA classifier is trained and optimised to give the best product of correction classification probabilities of all final states.

5 Results

Using the discriminating variables and trained the MVA classifier, the classification results of the MVA is shown in table 3, for seven tau decay modes of using the nominal ILD detector model, for $\sqrt{s} = 100\,\text{GeV}$. A perfect classification would only have diagonal terms in the table.

The correct classification probability reflects the difficulty to classify a final state. The $\mu^- \bar{\nu}_\mu$ and $e^- \bar{\nu}_e$ decay modes have 99.5% and 99.8% correct classification probabilities, as these two

Reco ↓ True →	$e^- \bar{\nu}_e$	$\mu^- \bar{\nu}_\mu$	π^-	$\rho\left(\pi^-\pi^0\right)$	$a_1\left(\pi^-\pi^0\pi^0\right)$	$a_1\left(\pi^-\pi^-\pi^+\right)$	$\pi^-\,\pi^-\,\pi^+\,\pi^0$
$e^- \bar{\nu}_e$	**99.7**	-	0.9	0.6	0.4	-	-
$\mu^- \bar{\nu}_\mu$	-	**99.5**	0.6	-	-	-	-
π^-	-	0.3	**94.0**	0.8	-	0.4	-
$\rho\left(\pi^-\pi^0\right)$	-	-	3.4	**93.6**	9.5	0.6	2.3
$a_1\left(\pi^-\pi^0\pi^0\right)$	-	-	-	4.5	**89.7**	-	0.6
$a_1\left(\pi^-\pi^-\pi^+\right)$	-	-	0.9	-	-	**96.8**	6.4
$\pi^-\,\pi^-\,\pi^+\,\pi^0$	-	-	-	0.3	-	2.0	**90.6**

Table 3: Classification probability in percentage for tau decay modes using the nominal ILD detector model , for $\sqrt{s} = 100$ GeV . Bold numbers show the correct classification probability. Numbers below 0.25%, and ν_τ are not shown in decay modes, for display purposes. Statistical uncertainties are less than 0.25%.

decay modes have very different topologies. For 1-prong final states, wrong classification mainly occur between the π^-, $\rho\left(\pi^-\pi^0\right)$, and $a_1\left(\pi^-\pi^0\pi^0\right)$ decay modes, because if two photons are very close together, they can be merged and the information about the two photons are lost. For 3-prong final states, wrong classifications mainly occur between $a_1\left(\pi^-\pi^-\pi^+\right)$, and $\pi^-\,\pi^-\,\pi^+\,\pi^0$, for the similar reason that photons close together can be merged and the information is lost. A correct classification probability over 88.2% for each tau decay mode is achieved, using the nominal ILD detector with $\sqrt{s} = 100\,\mathrm{GeV}$.

6 Electromagnetic calorimeter optimisation

To study of the impact of the ECal square cell sizes on the classification of tau decays, the above procedure was repeated for ECal square cell sizes at 3, 5, 7, 10, 15 and 20 mm, at four $\sqrt{s}$ energy of 100, 200, 500, 1000 GeV. Other ECAL dimensions were kept constant. Because the lepton reconstruction mostly rely on the tracking system, which was not varied in this study, only the hadronic tau decays were investigated and compared between different ECAL square cell sizes.

6.1 Impact of ECal square cell sizes

Figure 3 shows correct classification probabilities for hadronic tau decay final states as a function of the ECal square cell sizes, using the nominal CLIC_ILD detector model with $\sqrt{s} = 100$, 200, 500 and 1000 GeV. Overall, the correct classification probabilities for hadronic tau decays decrease with increasing $\sqrt{s}$ and increasing ECAL cell sizes. At a higher $\sqrt{s}$, decay products are closer together in the detector, and therefore it is more difficult to reconstruct individual decay product. A larger ECAL cell sizes offers a lower spatial resolutions. The reconstruction of two photons close together is more difficult. If photons are merged in the reconstruction, classification between different decay modes is more challenging and the correct classification probability is lower.

For the $\rho\left(\pi^-\pi^0\right)$ ν_τ decay mode, shown in figure 3b, the correct classification probabilities for $\sqrt{s} = 500\,\mathrm{GeV}$ rises for ECal square cell sizes from 15 to 20 mm, and for $\sqrt{s} = 1000\,\mathrm{GeV}$ from 7, to 20 mm. This is due to the optimisation of the MVA favours the overall correct classification

(a)

(b)

(c)

(d)

(e)
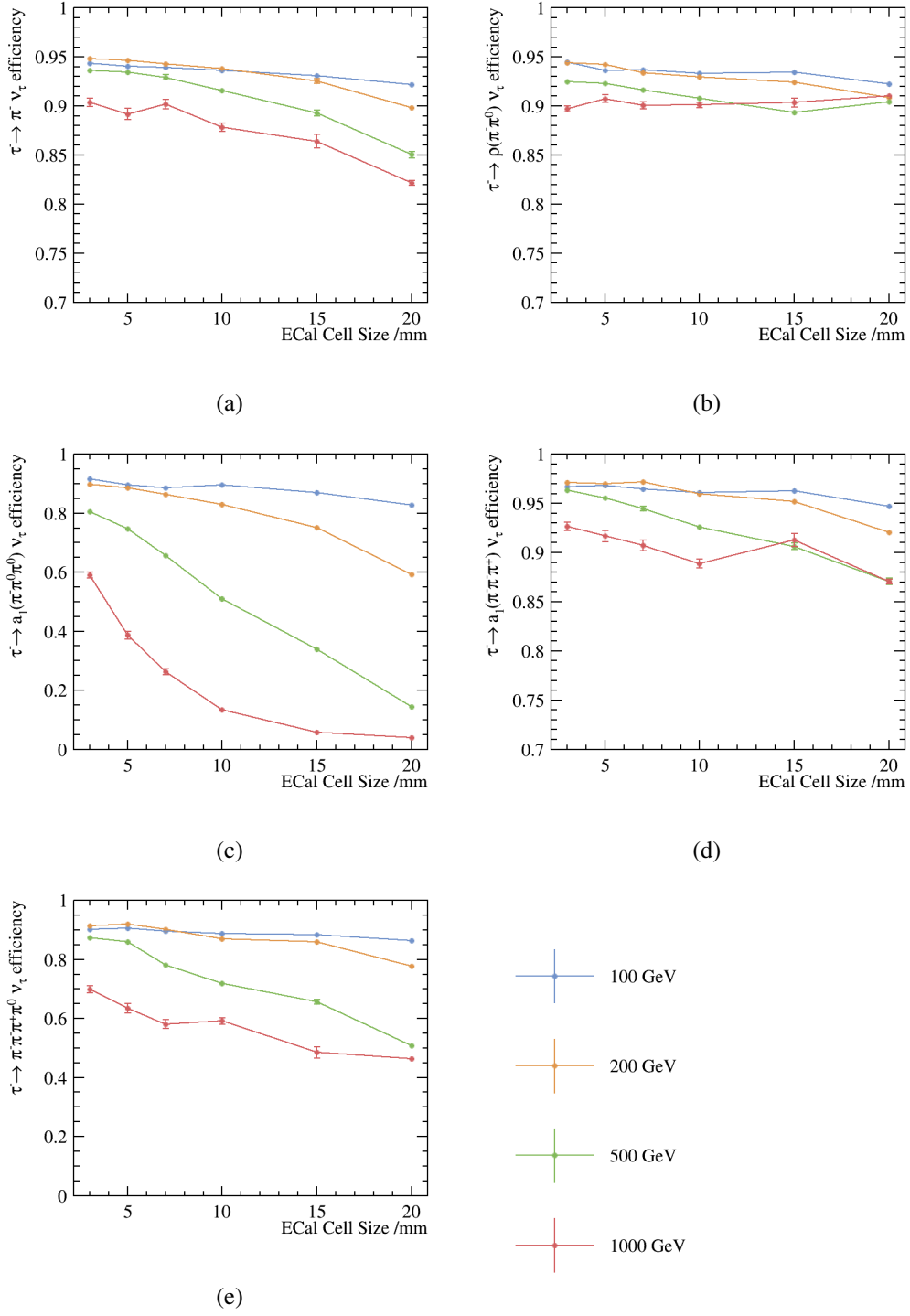
Figure 3: The correct classification probabilities for hadronic tau decay final states as a function of the ECal square cell sizes, using the nominal CLIC_ILD detector model with $\sqrt{s} = 100$, 200, 500 and 1000 GeV. Figure 3a, 3b, 3b, 3c, 3d show the $\pi^- \nu_\tau$, $\rho \left(\pi^- \pi^0 \right) \nu_\tau$, $a_1 \left(\pi^- \pi^0 \pi^0 \right) \nu_\tau$, $a_1 \left(\pi^- \pi^- \pi^+ \right) \nu_\tau$ and $\pi^- \pi^- \pi^+ \pi^0 \nu_\tau$ decay modes, respectively. Blue, orange, green and red lines represents $\sqrt{s} = 100$, 200, 500 and 1000 GeV respectively. Note that the the scale on the y axis are not the same, for display purposes.

rate, which can be seen in the sharp drop in correct classification probabilities for the $a_1\left(\pi^-\pi^0\pi^0\right)$ ν_τ decay mode for the same $\sqrt{s}$ and the same energy, shown in figure 3c.

TODO: For $\sqrt{s} = 100$ and $200\,\text{GeV}$, the correct classification probabilities for the 5 mm ECal cell size is better than that of the 3 mm. This is most noticeably shown in figure 3c. One possible explanation is that PandoraPFA has been optimised for the nominal ILD detector with the 5 mm ECal cell size, which shares the same ECal structure with the nominal CLIC_ILD detector.

6.2 Tau hadronic decay correct classification probability
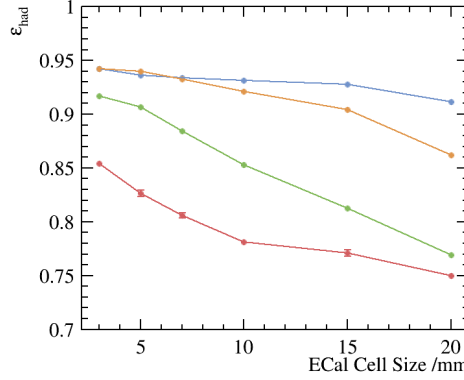
Figure 4: The tau lepton hadronic decay correct classification probability, ε_{had}, as a function of the ECal square cell size for different $\sqrt{s}$. Blue, orange, green and red lines represent $\sqrt{s} = 100$, 200, 500 and 1000 GeV, respectively.

Overall classification performance across all decay modes can be compared by constructing a single parameter function, which is a weighted average of the correct classification probability for each decay modes. This single parameter allows the direct comparison between different ECal cell sizes, for different $\sqrt{s}$. The chosen single parameter function is the tau lepton hadronic decay correct classification probability, ε_{had}, defined as

$$\varepsilon_{had} = \frac{\left(\Sigma_i Br_i \varepsilon_i\right)}{\Sigma_i Br_i}, \tag{6.1}$$

where Br_i is the branching ratio of a hadronic decay mode after the generator level cut in section 4.1. ε_i is the correct classification probability for decay mode i. Figure 4 shows ε_{had} as a function of the ECal square cell size for different $\sqrt{s}$. As expected ε_{had} decreases with increasing ECal square cell sizes, and increasing $\sqrt{s}$.

TODO: //For $\sqrt{s} = 100$ and $200\,\text{GeV}$, ε_{had} for the 5 mm ECal square cell size is better than that of the 3 mm. This is possibly due the optimisation of the software for the nominal ILD 5 mm square cell size.

CONCLUSION MAY CHANGE For the ILC 500 GeV and the CLIC 350 GeV scenarios, larger ECAL square cell sizes up to 15 mm have a small impact on ε_{had}. The decrease in ε_{had} from 3 mm to 15 mm ECAL cell for $\sqrt{s} = 100$ and $200\,\text{GeV}$ is below 4%. For the ILC 1 TeV and

the CLIC 1.4 TeV and 3 TeV scenarios, an increase of 1 mm in the ECAL square cell size leads a decrease of approximately 0.8% in ε_{had}. To achieve a high ε_{had}, a small ECAL cell size is needed.

Acknowledgments

References

[1] ALEPH collaboration, S. Schael *et al.*, Phys. Rept. **421**, 191 (2005).

[2] K. Riles, Int. J. Mod. Phys. **A7**, 7647 (1992).

[3] S. Berge, W. Bernreuther, and S. Kirchner, Phys. Rev. **D92**, 096012 (2015).

[4] DELPHI collaboration, P. Abreu *et al.*, Phys. Lett. **B267**, 422 (1991).

[5] ILD Concept Group, T. Abe *et al.*, (2010), 1006.3396.

[6] ILC collaboration, H. Baer *et al.*, (2013), 1306.6352.

[7] M. A. Thomson, Nucl. Instrum. Meth. **A611**, 25 (2009).

[8] ILC collaboration, T. Behnke *et al.*, (2007), 0712.2356.

[9] J. S. Marshall and M. A. Thomson, Eur. Phys. J. **C75**, 439 (2015).

[10] CLIC collaboration, L. Linssen, A. Miyamoto, M. Stanitzki, and H. Weerts, (2012), 1202.5940.

[11] ILC collaboration, J. Brau *et al.*, (2007).

[12] B. Xu, Improvement of photon reconstruction in PandoraPFA, in *Proceedings, International Workshop on Future Linear Colliders (LCWS15): Whistler, B.C., Canada, November 02-06, 2015*, 2016, 1603.00013.

[13] W. Kilian, T. Ohl, and J. Reuter, Eur. Phys. J. **C71**, 1742 (2011).

[14] T. Sjostrand, (1995), hep-ph/9508391.

[15] S. Jadach, Z. Wąs, R. Decker, and J. H. Kuhn, Comput. Phys. Commun. **76**, 361 (1993).

[16] P. Mora de Freitas and H. Videau, Detector simulation with MOKKA / GEANT4: Present and future, in *Linear colliders. Proceedings, International Workshop on physics and experiments with future electron-positron linear colliders, LCWS 2002, Seogwipo, Jeju Island, Korea, August 26-30, 2002*, pp. 623–627, 2002.

[17] GEANT4 collaboration, S. Agostinelli *et al.*, Nucl. Instrum. Meth. **A506**, 250 (2003).

[18] F. Gaede, Nucl. Instrum. Meth. **A559**, 177 (2006).

[19] Particle Data Group, K. A. Olive *et al.*, Chin. Phys. **C38**, 090001 (2014).

[20] TMVA Core Developer Team, J. Therhaag, AIP Conf.Proc. **1504**, 1013 (2009).