Eur. Phys. J. C 46, s01, 1–21 (2006)
Digital Object Identifier (DOI) 10.1140/epjcd/s2006-02-001-y

EPJ C direct
electronic only

Scientific Note

Tau jet reconstruction and tagging with CMS

The CMS Collaboration

S. Gennai[1,a], F. Moortgat[2], L. Wendland[3], A. Nikitenko[4], S. Wakefield[4], G. Bagliesi[5], S. Dutta[5], A. Kalinowski[6], M. Konecki[6], D. Kotlinski[7]

[1] Centro Studi Enrico Fermi, Rome and INFN Pisa, Italy
[2] CERN, Geneva, Switzerland
[3] Helsinki Institute of Physics, Finland
[4] Imperiall College, London, UK
[5] INFN and Scuola Normale Superiore, Pisa, Italy
[6] Institute of Experimental Physics, Warsaw University, Warsaw, Poland
[7] Paul Scherrer Institut, 5232 Villigen, Switzerland

Received: 17 February 2006 / Revised version: 16 March 2006 /

Abstract. The τ indentification and reconstruction algorithms developed for the CMS experiment at LHC are presented. The algorithms are aimed to an efficient selection of hadronic decays of the τ lepton. Reconstruction methods suitable for use at the High Level Trigger and off-line are described in detail.

1 Introduction

The searches for the Higgs boson and SUSY signatures at the LHC relay very much on the identification of τ leptons in the final state. A number of methods to identify τ jets, i.e. hadronic decays of the τ lepton, used in CMS are described. The methods are based on τ properties such as its long lifetime, its mass and the small number of charged decay products. Usage of these methods in different combinations depends on the physics channel considered. The τ-jet identification requires an isolated and collimated jet made of charged particles reconstructed with the tracker; the τ identification can be improved combining this isolation criterium with other algorithms.

In the following, the basic properties of a τ jet are presented in Sect. 2. In Sect. 3 the off-line isolation and the other tagging methods are discussed. The impact parameter, flight-path and mass tagging are intended to be applied after isolation and so their performances have been computed over a preselected sample of isolated jets. In Sect. 4 the High Level Trigger chain is presented and discussed. In Sect. 5 the calibration of the energy of the τ jet is discussed and the basic ideas on how to estimate the tagging algorithms performances are presented.

2 Tau properties relevant to τ jet reconstruction and identification

The τ lepton decays hadronically 65% of the time, producing a τ jet, wghich is a jet-like cluster in the calorimeter containing a relatively small number of charged and neutral hadrons. When the p_T of the τ jet is large compared to the τ mass, these hadrons have relatively small momentum in the plane transverse to the τ jet axis. In 77% of hadronic τ decays, the τ jet consists of only one charged hadron and a number of π^0s (one-prong decays). Because of these features hadronic τ decays produce narrow jets in the calorimeter.

Figure 1 shows the ratio $r = E_T^{reco}/E_T^{MC}$ as a function of the reconstruction cone size for three bins of E_T^{MC}. The E_T^{reco} is the transverse energy reconstructed in the calorimeter with an iterative cone algorithm, while the E_T^{MC} is the Monte Carlo (MC) generated transverse energy. The thresholds on calorimeter towers, input to the jet finder, were set as $E_T = 0.5$ GeV and $E = 0.8$ GeV. The values of r in Fig. 1 were normalized to the value obtained with a cone size of 0.6. Figure 2 shows the transverse energy resolution of the τ jet as a function of the reconstruction cone size for the three different bins of E_T^{MC}. From Fig. 1 a cone size of 0.4 for τ jet reconstruction with the calorimeter was chosen since it contains a large fraction of the τ-jet energy (more that 98%) and the cone size smaller than 0.4 leads to a degradation of the τ-jet energy resolution as can be seen from Fig. 2. A larger cone size can lead to a con-

---

[a] e-mail: simone.gennai@cern.ch

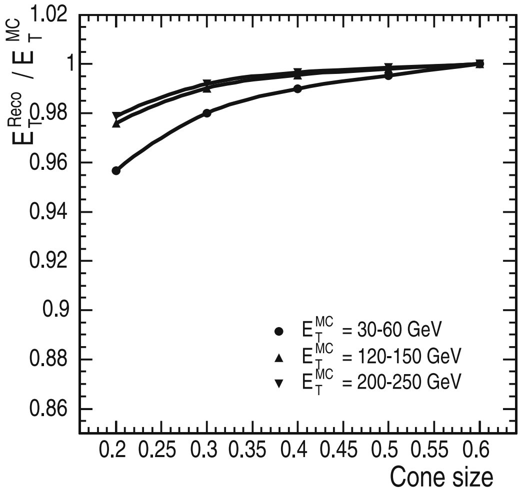

Fig. 1. Distribution of the ratio r = E_T^{reco}/E_T^{MC} as a function of the reconstruction cone size for the three bins of E_T^{MC}. The values of r were normalized to the r for a cone size of 0.6

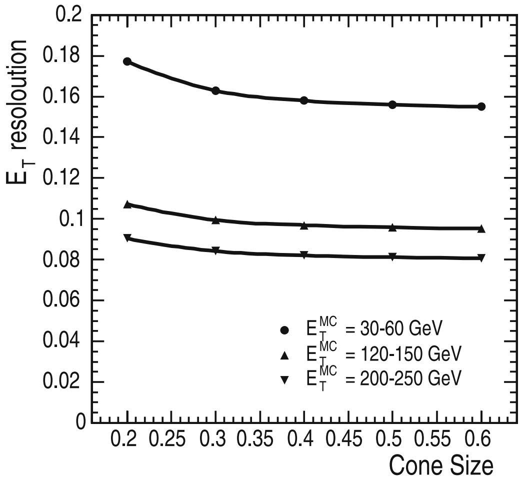

Fig. 2. Transverse energy resolution of τ jet as a function of the reconstruction cone size for the three bins of E_T^{MC}

tamination from other jets in multi-jet events. Figure 3 shows the difference in ϕ (left plot) and in η (right plot) between the jet-direction of the true τ jet and the τ jet reconstructed with the calorimeter, for the three intervals of the true τ-jet energy. The charge of the τ lepton is positive in these event samples. The 4 Tesla magnetic field leads to a systematic shift of $\simeq 0.02$ rad in the reconstructed τ jet direction in ϕ for τ jets with E_T between 40 and 60 GeV. The shift is reduced for the jets with larger E_T. The resolution in η is slightly worse than in ϕ and does not depend on E_T between 40 and 250 GeV.

The τ jet-identification requires a matching between the calorimeter jet axis and the charged particles from the hadronic τ decays measured with the tracker. Figure 4 shows the distance ΔR in η–ϕ space between the direction of the leading p_T track at the origin, reconstructed with the tracker, and the direction of the τ jet reconstructed with the calorimeter for three bins of the true τ-jet transverse energy E_T^{MC}. A cut of 3 GeV/c was applied on the p_T of the leading track. Both, the one and the three-prong (three charged particles in the decay product) τ decays are included. The value of ΔR does not exceed 0.1 for the range of E_T^{MC} considered.

In the case of hadronic τ decays with three charged particles in the final state, the three particles are produced within a narrow cone. Figure 5 shows the maximal distance ΔR in η–ϕ space between the leading p_T charged particle and other two charged particles in the three-prong τ decays for three bins of the true τ-jet transverse energy E_T^{MC}.

The τ-lepton lifetime ($c\tau = 87.11\,\mu$m) and the mass ($m_\tau = 1.78$ GeV/c^2) were used for the τ jet tagging with the track impact parameter measurement, vertex reconstruction (for three-prong decays) and constraining the effective mass of track(s) and calorimeter clusters.

3 Methods for τ tagging and performance

All the efficiencies shown in these sections are relative to events in which the calorimeter reconstructed jet matches

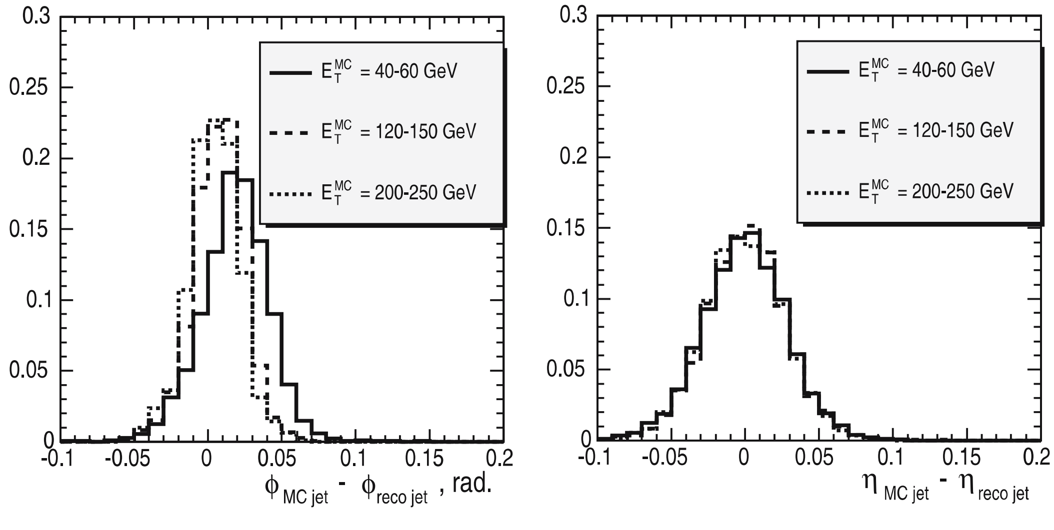

Fig. 3. Distribution of the difference in ϕ (*left plot*) and in η (*right plot*) between the true τ jet direction and the jet direction reconstructed with the calorimeter for the three different intervals of the true τ-jet energy. The τ lepton has a positive charge in these event samples

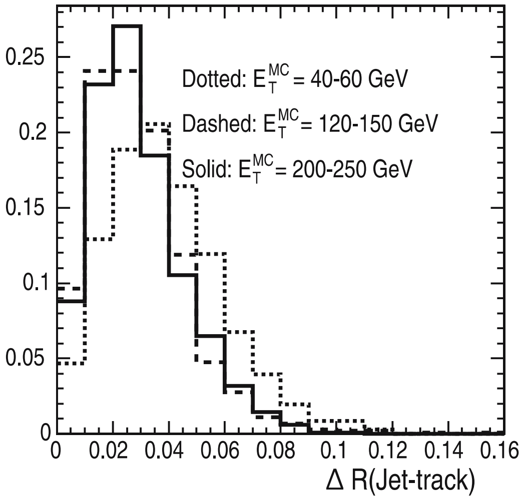

Fig. 4. Distribution of the distance, ΔR, in η–ϕ space between the leading p_T track reconstructed with the tracker and the direction of the τ jet reconstructed with the calorimeter for three bins of the true τ-jet transverse energy E_T^{MC}. The cut of 3 GeV/c was applied on the p_T of the leading track. Both, the one and the three-prong τ decays are taken into account

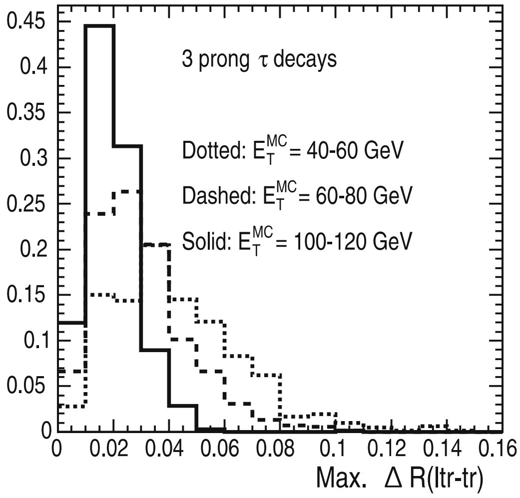

Fig. 5. Maximal distance ΔR in η–ϕ space between the leading p_T charged particle and other two charged particles in the three-prong τ decays for three bins of the true τ-jet transverse energy E_T^{MC}

with the jet reconstructed at the generator level (Monte Carlo jet). The matching criteria requires that the distance between the reconstructed and the Monte Carlo jet axis in the η–φ space should be less than 0.2. The performance of τ-tagging methods was evaluated for the sample of pure τ jets with transverse momentum $p_T^{\tau\ \mathrm{jet}} > 30$ GeV/c and distributed uniformly in pseudorapidity, up to $|\eta| < 2.2$, and φ. In order to study the characteristics of the τ jet in a simple environment, this sample has been simulated without the pile-up and the underlying event. The background rejection was computed on a sample of di-jet events generated with PYTHIA [1] (MSEL = 1) (QCD jets) with transverse energy between 30 and 150 GeV. For the background samples further MC selections have been applied: the two leading-E_T Monte Carlo jets in the di-jet sample were required to be separated in η–φ space by the distance of $\Delta R > 1.5$, and to be reconstructed inside $|\eta| < 2.1$. These two jets were propagated through the τ-identification criteria. The τ-tagging performance was evaluated as a function of the true transverse energy E_T^{MC} and the pseudorapidity of the jet. The true τ-jet transverse energy is defined as the energy of the τ lepton without neutrino energy in the decay $\tau \to \mathrm{hadrons} + \nu$. The true energy of the QCD jet is the energy of the Monte Carlo jet found using the cone algorithm with the cone size of 0.5. The cone algorithm uses the Monte Carlo stable particles, excluding neutrinos and muons, as an input. The efficiency for the QCD events, to pass the Monte Carlo preselection and the matching criteria was found to be of order of 12%. In the following sections (with the exception of the Sect. 4) the efficiencies are computed on these matched jets events, and they don't include the matching and preselection efficiency.

3.1 Isolation in the electromagnetic calorimeter

Hadronic τ decays produce localized energy deposit in the electromagnetic calorimeter (ECAL). Several variables were tried to quantify this feature and to use it for τ tagging and QCD-jet rejection [2]. The electromagnetic isolation parameter P_{isol} defined as

$$P_{isol} = \sum_{\Delta R < 0.40} E_T - \sum_{\Delta R < 0.13} E_T \qquad (1)$$

was found to provide the best efficiency for hadronic jet rejection. The sums run over transverse energy deposits in the electromagnetic calorimeter, and ΔR is the distance in η–φ space from the reconstructed τ jet axis. Jets with $P_{isol} < P_{isol}^{cut}$ are considered as τ candidates. More information about the choice of this variable and the parameters can be found in [3, 4].

Figure 6 shows the efficiency of the ECAL isolation for τ jets as a function of E_T^{MC} (left plot) and $|\eta^{MC}|$ (right plot) for $P_{isol}^{cut} = 5$ GeV. The efficiency is shown separately for four final states of hadronic decays of the τ lepton. Only a small ($\simeq 5\%$) variation with E_T^{MC} is observed over a large region of transverse energies from 30 to 300 GeV. The variation in pseudorapidity for τ decays with π^0's in the final state follows the η variation of the amount of the tracker material in front of the ECAL. This correlation is due to electrons and positrons from photon conversions in the tracker material contaminating the ECAL isolation region. Figure 7 shows the efficiency of the electromagnetic isolation for τ jets and QCD jets in several bins of the true transverse energy when the value of P_{isol}^{cut} is varied. The ECAL isolation can provide a rejection factor $\simeq 5$ against large E_T QCD jets (> 80 GeV) with the efficiency better than 80%. The efficiency for QCD jets decreases with increasing E_T of the jet. The explanation for this behaviour is that low energy charged particles ($p_T < 2$ GeV/c) are bend out of the 0.4 cone and so don't contribute to the energy sum in the P_{isol} formula [5].

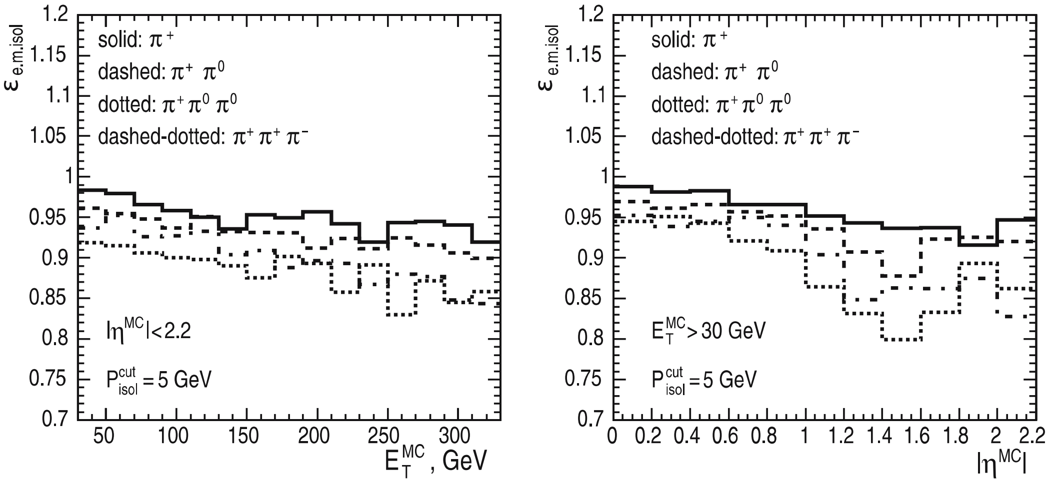

Fig. 6. Efficiency of the ECAL isolation for τ jets as a function of E_T^{MC} (*left plot*) and $|\eta^{MC}|$ (*right plot*) for $P_{isol}^{cut} = 5$ GeV. The efficiency is shown separately for several final states of hadronic decays of τ lepton

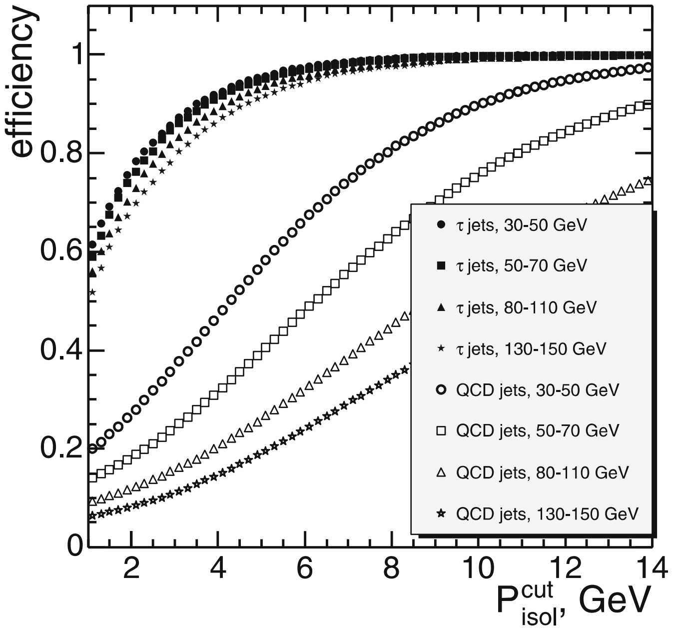

Fig. 7. Efficiency of the electromagnetic isolation for τ jets and QCD jets in the several bins of the true transverse energy when the value of P_{isol}^{cut} is varied

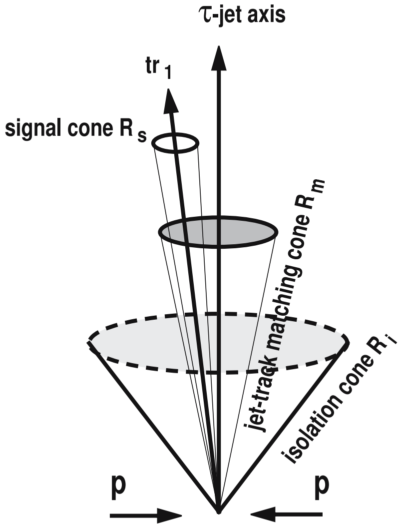

Fig. 8. Sketch of the basic principle of τ-jet identification using the tracker isolation

3.2 Tracker isolation

The principle of τ-jet identification using the tracker isolation is shown in Fig. 8. The direction of the τ jet is defined by the axis of the calorimeter jet. The tracks above a threshold of p_T^m and in a matching cone of radius R_m around the calorimeter jet direction are considered in the search for signal tracks. The leading track (tr_1 in Fig. 8) is defined as the track with the highest p_T. Any other track in the narrow signal cone R_s around tr_1 and with z-impact parameter z_{tr} close to the z-impact parameter of the leading track z_{tr}^{ltr} ($|z_{tr} - z_{tr}^{ltr}| < \Delta z_{tr}$) is assumed to come from the τ decay. Tracks with $|z_{tr} - z_{tr}^{ltr}|$ smaller than a given cut-off (Δz_{tr}) and transverse momentum above a threshold of p_T^i are then reconstructed inside a larger cone of the size R_i. If no tracks are found in the R_i cone, except for the ones which are already in the R_s cone, the isolation criteria is fulfilled.

Figure 9 shows the tracker isolation efficiency for the τ jets (left plot) and QCD jets (right plot) as a function of the isolation cone R_i for two values of the signal cone $R_s = 0.07$ and $R_s = 0.04$. The remaining tracker isolation parameters are: $R_m = 0.1$, $p_T^i = 1\,\mathrm{GeV}/c$, $\Delta z_{tr} = 2$ mm. The leading track p_T was required to be greater than $6\,\mathrm{GeV}/c$. Tracks were reconstructed with the combinatorial track finder algorithm [8] requiring at least 8 hits per track and the normalized $\chi^2 < 10$, with at least two reconstructed hits inside the pixel detector. Jets were reconstructed in the calorimeter with the iterative cone algorithm taking a cone size of 0.4. The reconstructed QCD jets should match the two leading E_T Monte Carlo jets. The tracker isolation can provide a rejection factor of more than 10 against QCD jets with an efficiency for τ jets above 70%. Inefficiencies for every step in the tracker isolation algorithm are presented in Table 1 for the τ jets in two bins of E_T^{MC}, 30–50 and 130–150 GeV (a closer view of the isolation features can be found in Sect. 4.2).

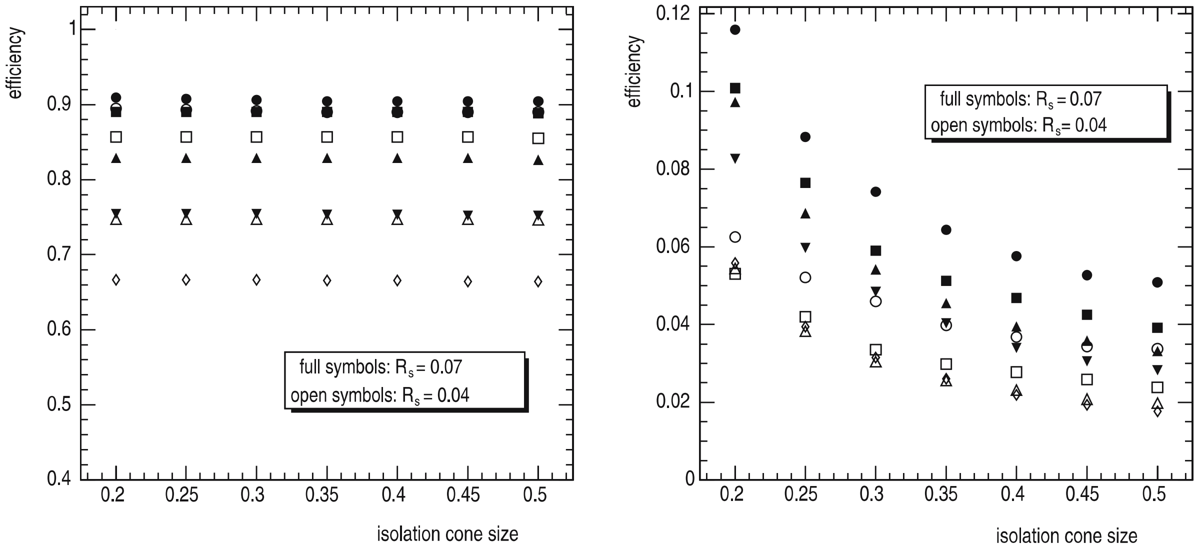

Fig. 9. Tracker isolation efficiency for the τ jets (*left plot*) and QCD jets (*right plot*) as a function of the isolation cone R_i for two values of the signal cone $R_s = 0.07$ (*full symbols*) and $R_s = 0.04$ (*open symbols*). In the order of the decreasing efficiency the symbols corresponds to E_T^{MC} bins of 130–150, 80–110, 50–70 and 30–50 GeV. The remaining tracker isolation parameters are: $R_m = 0.1$, $p_T^i = 1$ GeV/c, $\Delta z_{tr} = 2$ mm and the leading track $p_T > 6$ GeV/c. The efficiency has been computed after the MC preselection and matching

Table 1. Inefficiencies for every step of the tracker isolation algorithm for the τ jets in two bins of E_T^{MC}, 30–50 and 130–150 GeV. The isolation ring is defined as the ring between the signal and the isolation cone. The MC preselection and matching have been applied before the isolation

E_T^{MC}, GeV	≥ 1 track in the isolation ring	leading track with $p_T > 6$ GeV/c, $R_m = 0.1$	isolation $R_s = 0.07$	isolation $R_s = 0.04$	Total ineff. for $R_s = 0.04$
30–50	7.7%	10.2%	5.2%	14.2%	32.1%
130–150	4.8%	2.6%	1.0%	2.5%	9.9%

Table 2. Efficiency of the track counting requirement for τ and QCD jets in four bins of E_T^{MC}

QCD jets; E_T^{MC}, GeV	30–50	50–70	80–110	130–150
1 track	63%	72%	69%	60%
3 tracks	7%	9%	9%	13%
1 or 3 tracks	70%	81%	78%	73%
τ jets; E_T^{MC}, GeV	30–50	50–70	80–110	130–150
1 track	81%	77%	71%	70%
3 tracks	10%	16%	16%	20%
1 or 3 tracks	91%	93%	87%	90%

The hadronic τ decay products consist mainly of one or three charged particles (one and three-prong) in the final state. The one-prong decays represent $\simeq 77\%$ of all hadronic decay modes of the τ lepton. Therefore, the tracker isolation requirement can be naturally followed by the requirement to have only one or three reconstructed tracks in the signal cone. Table 2 shows the efficiency of the track counting requirement for τ and QCD jets in four bins of E_T^{MC}. The sizes of the isolation and the signal cones were fixed to $R_i = 0.4$ and $R_s = 0.07$. It can be concluded that the track counting criteria do not improve the suppression against QCD jets for events which pass the tracker isolation. To further suppress low E_T QCD jets, a strong cut on the p_T of the leading track can be used. In the analysis of the $H/A \to \tau\tau$ channel [9] a cut of 40 GeV/c was used to efficiently reduce the QCD background.

A number of tagging methods which can be applied after the isolation criteria are discussed in the following subsections. The events are preselected with the tracker isolation with the following parameters: $R_m = 0.1$, $R_s = 0.07$, $R_i = 0.4$, $p_T^i = 1$ GeV/c and $\Delta z_{tr} = 2$ mm. Only one or three tracks were required in the signal cone and the leading track p_T was required to be greater than 6 GeV/c.

3.3 Tagging with the impact parameter

The impact parameter (IP) tagging is based on the leading track of the jet. Figure 10 shows the unsigned transverse IP (IP_T) distributions for the τ and QCD jets with one or three reconstructed tracks. The upper plot shows a large tail for the background sample with one reconstructed track. Looking at the Monte Carlo (MC) information reveals, that most of the tracks in the tail are made up of hits

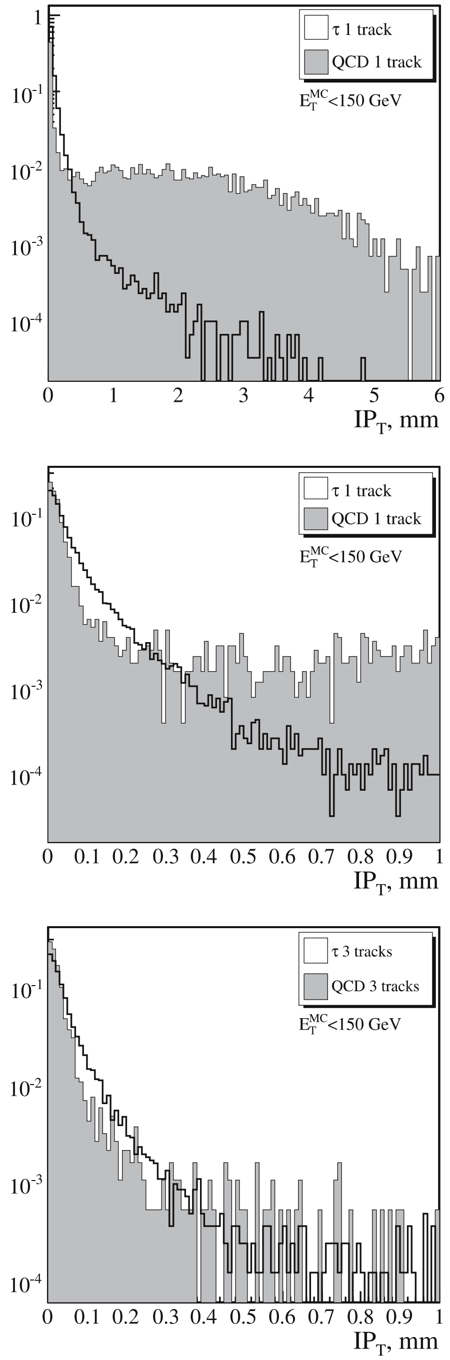

Fig. 10. Transverse unsigned IP distribution for τ and QCD jets with one or three reconstructed tracks. The *middle plot* is the same as the *upper plot*, but with enlarged scale in x-axis. The histograms have been normalized to unity

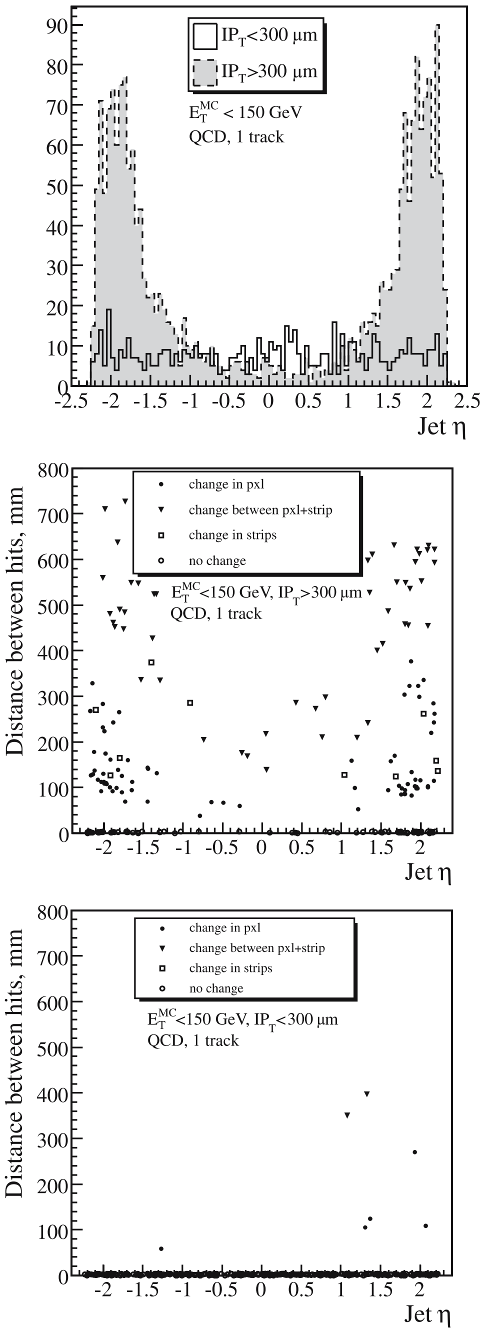

Fig. 11. Jet η distribution for the QCD jets with one reconstructed track (*upper*) and the propagation distance between two consecutive reconstructed hits, between which the simulated track association has changed, as a function of jet η for the tail region (*middle*) and peak region (*bottom*). The *middle* and *upper plots* are based on a smaller sub-sample of events than the *bottom* one

belonging to different MC tracks, of which only this mixed track survives the track reconstruction. The tracks in the tail have thus a higher fake hit rate than the tracks within the peak region.

Figure 11 shows that the tail are due to jets that are mostly emitted in the forward direction. By determining the MC track, to which most of the reconstructed hits are associated, it is possible to look for the region, in which the track reconstruction jumps from this most popular MC track to another track. The middle and bottom plots of Fig. 11 show, that for the tail events most of these jumps occur inside the pixel detector or between the pixel and strip detectors. Due to the large η, the distance between the two consecutive hits, where the jump occurs, is fairly large. The probability to assign wrong hits to the track is thus increased by the long propagation distance in the track building. The amount of events in the tail is increased with E_T^{MC}, because the track multiplicity increases with E_T^{MC} and thus also increases the probability to assign fake hits to the track.

The middle and bottom plots of Fig. 10 show that the tail can be cut away with minor losses on the signal events by placing an upper limit cut at $IP_T < 300\,\mu$m. Additionally, a lower limit cut on IP_T can be set to further increase background rejection. A cut on the IP sign was studied, but since almost half of the signal events would have been assigned a minus sign, it would have cut away too many signal events.

The error of the IP_T is somewhat larger for the background than for the signal events. The mean error is $15.0\,\mu$m and $16.7\,\mu$m for one-prong and three-prong signal events and $17.9\,\mu$m and $22.2\,\mu$m for one-prong and three-prong background events, respectively. The mean error on the full IP is $58.3\,\mu$m and $56.6\,\mu$m for one-prong and three-prong signal events and $75.7\,\mu$m and $60.4\,\mu$m for one-prong and three-prong background events, respectively.

Figure 12 shows efficiency curves, for jets with one reconstructed track, for several lower limit cuts on the IP_T and IP significance and by requiring upper limit cuts of $IP_T < 300\,\mu$m for the transverse and IP < 1 mm for the full significance plots. Since most of the background rejection is coming from the rejection of the tail events, the two curves are very similar. Figure 13 shows the same efficiency curves for the case, where the data samples have been additionally cleaned by requiring a $\chi^2_{\text{hit}} < 10$ for the estimates of the two first hits of a track prior to the efficiency calculation. For jets with three reconstructed tracks, the IP distributions are so similar for the τ and QCD samples, that the background efficiency is about 81%–88% and the signal efficiency is 95%.

3.4 Tagging with flight-path

The lifetime of the τ lepton ($c\tau = 87.11\,\mu$m) allows for the reconstruction of the decay vertex for the three and five-prong decay separated from the primary vertex. The lower cut-off on the flight-path as a distance between the primary vertex and the decay vertex of τ lepton is used for the τ tagging. The tracks in the signal cone are used as an input for the vertex fitter. Table 3 shows the fraction of τ jets with a certain number of reconstructed tracks in the signal cone for one, three and five-prong τ decays. The proportion of three-prong and five-prong τ decays of τ jets that have passed jet isolation is 23.9% and 0.3%, respectively. Due to this low statistics, the five-prong decays are discarded and three reconstructed tracks are required in the signal cone.

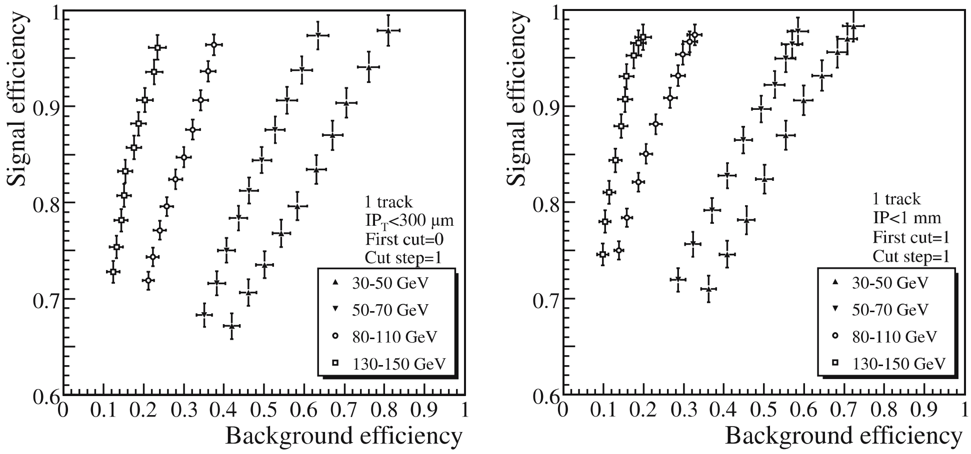

Fig. 12. Signal and background efficiency curves for four jet E_T^{MC} bins based on a lower limit cut on the unsigned IP_T significance (*upper*) and the unsigned full IP significance (*bottom*) with upper limit cuts of $IP_T < 300\,\mu$m and IP < 1 mm, respectively. The efficiency has been computed after having applied the MC preselection and matching, and the tracker isolation

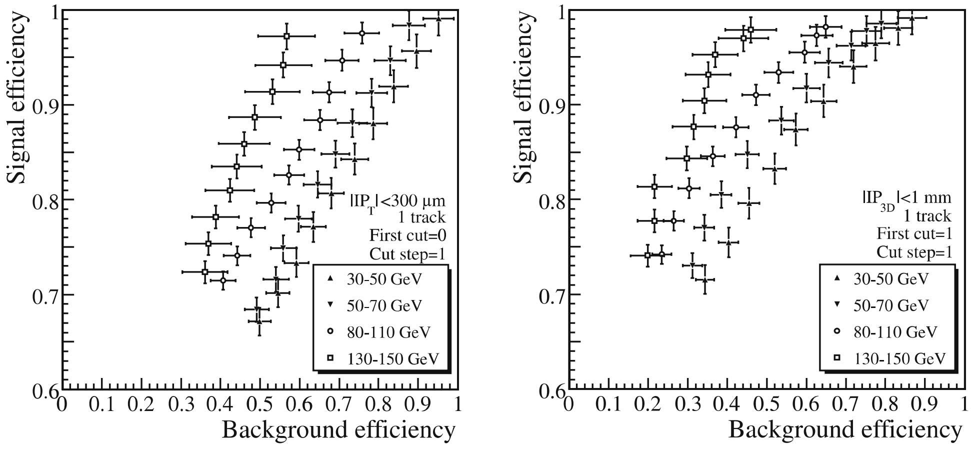

Fig. 13. Signal and background efficiency curves for different jet E_T^{MC} bins based on a lower limit cut on the unsigned IP_T significance (*upper*) and the unsigned full IP significance (*bottom*). In addition to the track quality cuts, $\chi^2_{\text{hit1}} < 10$ and $\chi^2_{\text{hit2}} < 10$ were required before the efficiency curves. The upper limit cut of $IP_T < 300\,\mu$m and IP < 1 mm, respectively, is included in the efficiency. The efficiency has been computed after having applied the MC preselection and matching, and the tracker isolation

Table 3. Fraction of τ leptons with a certain number of the reconstructed tracks in the signal cone for one, three and five-prong decay

	1 track	2 tracks	3 tracks	>3 tracks
1-prong τ	$88.4 \pm 0.3\%$	$6.1 \pm 0.1\%$	$4.1 \pm 0.1\%$	$1.4 \pm 0.1\%$
3-prong τ	$8.6 \pm 0.1\%$	$16.1 \pm 0.2\%$	$63.2 \pm 0.4\%$	$12.1 \pm 0.2\%$
5-prong τ	$13.1 \pm 1.7\%$	$4.4 \pm 1.0\%$	$11.7 \pm 1.7\%$	$70.9 \pm 4.1\%$

For three-prong τ decays the probability to reconstruct three tracks in the signal cone was found to be $\simeq 63\%$.

3.4.1 Secondary vertex resolution

The secondary vertex (SV) was fitted using the Kalman vertex fitter (KVF), adaptive vertex fitter (AVF) and principal vertex fitter (PVF). Figures 14 and 15 show the SV residuals and pulls for the x and z coordinates, respectively, for the signal sample. A double gaussian fit is used to estimate the central and tail parts of the distributions. The central part of the residual and pull is $180\,\mu$m and 1.1, respectively, independent of the vertex fitter used. The tails in the residual plots are coming predominantly from events, where two of the tracks are very close to each other.

In the plane transverse to the τ jet axis, the resolution $\sigma_{\mathrm{transverse}}$ of the secondary vertex is between 18 and 25 μm and does not depend too much on the τ energy for E_T^{MC} between 30 and 300 GeV as seen in Fig. 16. The resolution in the direction parallel to the τ jet axis ($\sigma_{\mathrm{longitudinal}}$) depends on the jet E_T due to kinematics as shown in the right plot of Fig. 16. The resolution is changed from 500 μm to $\simeq$ 1.5 mm when transverse τ jet energy is increased from 30 to 250 GeV. The three different vertex fitters deliver almost equal performance for both the transverse and longitudinal resolutions.

3.4.2 Background rejection based on flight-path

The reconstruction of the SV for the τ jets poses a challenge, because the tracks are very collimated in the pixel layers. The fraction of τ jets with $E_T^{MC} < 150$ GeV and with at least two shared hits in the reconstructed tracks are 22.8% and 9.6%, respectively, which can lead to a reconstruction of fake secondary vertices. The left plot on Fig. 17 shows the reconstructed transverse flight-path of the τ lepton for two intervals of the true transverse energy $E_T^{MC} < 150$ GeV and E_T^{MC} between 150 and 420 GeV. The fake secondary vertices reconstructed in the location of the first and the second barrel layers of the Pixel detector are visible as bumps at $\simeq 40$ mm and 70 mm. The contamination of the fake vertices is larger for more energetic τ jets since the charged particles are more collimated. The three charged particles from the high E_T τ jets can produce overlapping clusters in the pixel layer, thus leading to the reconstruction of one hit for all three tracks, which forces the vertex fitter to reconstruct the vertex position in the location of the pixel layer.

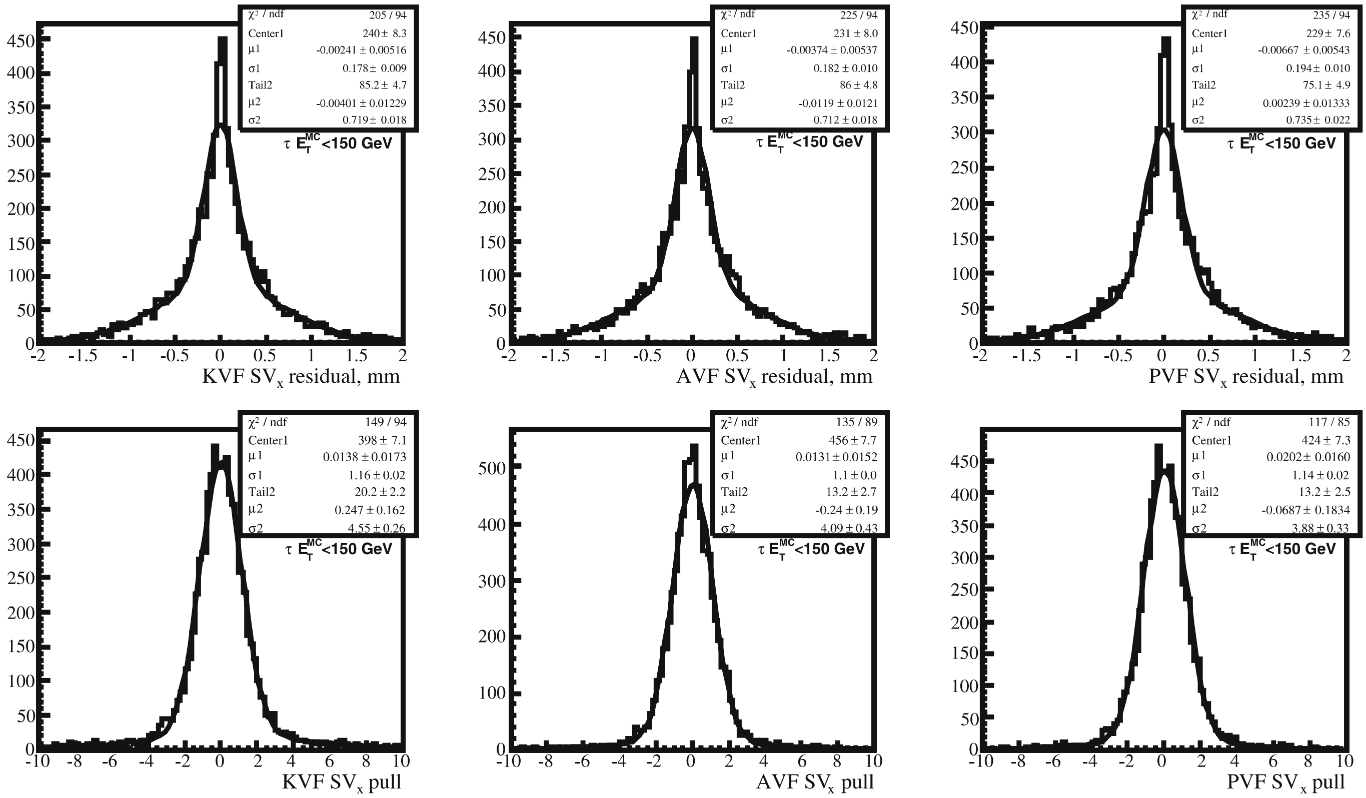

Fig. 14. Residual (*top*) and pull (*bottom*) of the secondary vertex x-coordinate for the KF, AV and PV Fitters

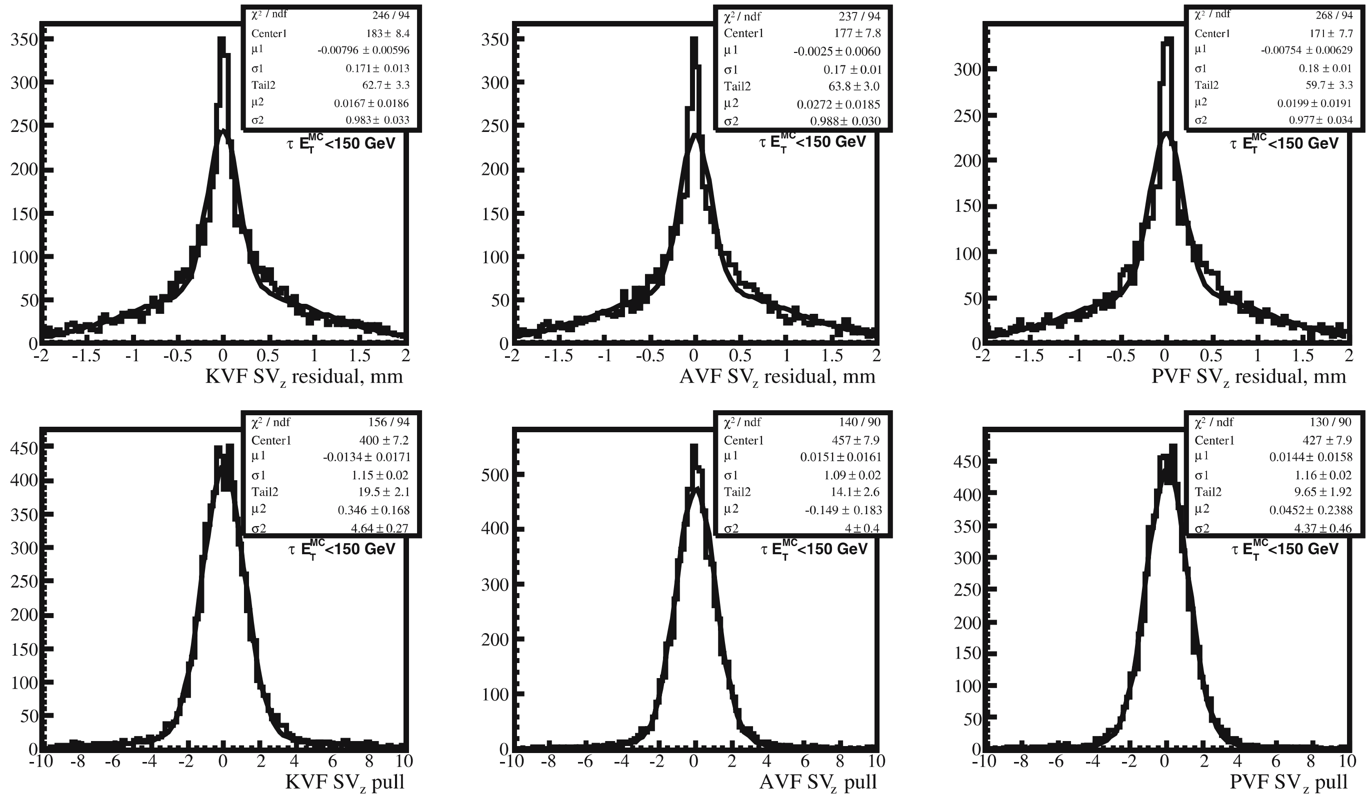

Fig. 15. Residual (*top*) and pull (*bottom*) of the secondary vertex z-coordinate for the KF, AV and PV Fitters

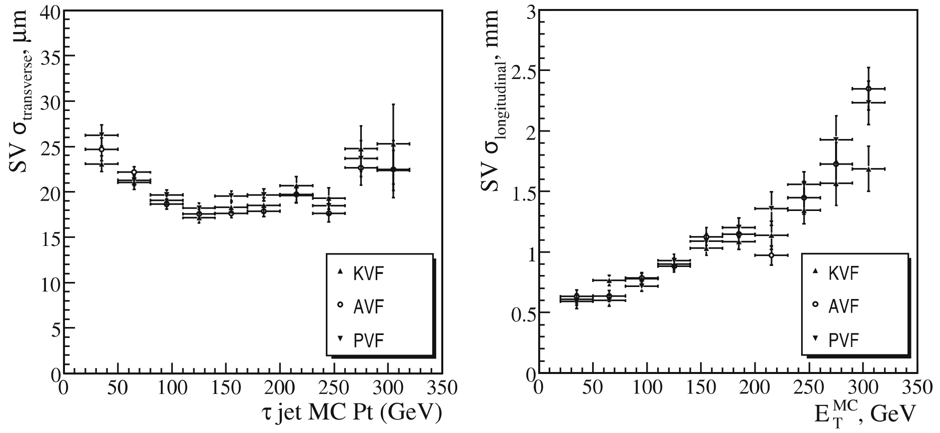

Fig. 16. Resolution of the central gaussian of the secondary vertex projected transverse to τ jet axis (*left*) and parallel to τ jet axis (*right*) for the different vertex fitters

The right plot on Fig. 17 shows the reconstructed transverse flight-path for the QCD jets generated with $\hat{p}_T$ between 80 and 120 GeV. The jets were required to pass the tracker isolation and to have three reconstructed tracks in the signal cone. The fake vertices produce a bump at $\simeq$ 40 mm in the location of the first barrel pixel layer and the tail. Even for the most dangerous QCD p_T bin, the jet isolation criteria and the request for three reconstructed tracks kill most of the background coming from QCD jets containing b and c quarks, which can produce B and D mesons with a non-zero flight-path length. A contribution from c- and b-quark jets in the sample is also shown in Fig. 17. Almost no difference is visible between the distributions for the the light and heavy quarks. The content of c and b jets is presented in the Table 4 for four intervals of E_T^{MC}.

Table 4. Fraction of the c and b jet in QCD jets in four bins of E_T^{MC} after the tracker isolation and the requirement to have three reconstructed tracks in the signal cone

E_T bin, GeV	fraction of c jets, %	fraction of b jets, %
30–50	13.1 ± 2.1	3.1 ± 1.0
50–70	12.3 ± 1.6	4.4 ± 1.0
80–110	13.4 ± 1.4	2.7 ± 0.6
130–150	12.4 ± 1.9	3.0 ± 1.0

The primary vertex position in the z coordinate was found with the pixel vertex finder for the QCD multi-jet events. For single τ jets the Monte Carlo primary vertex

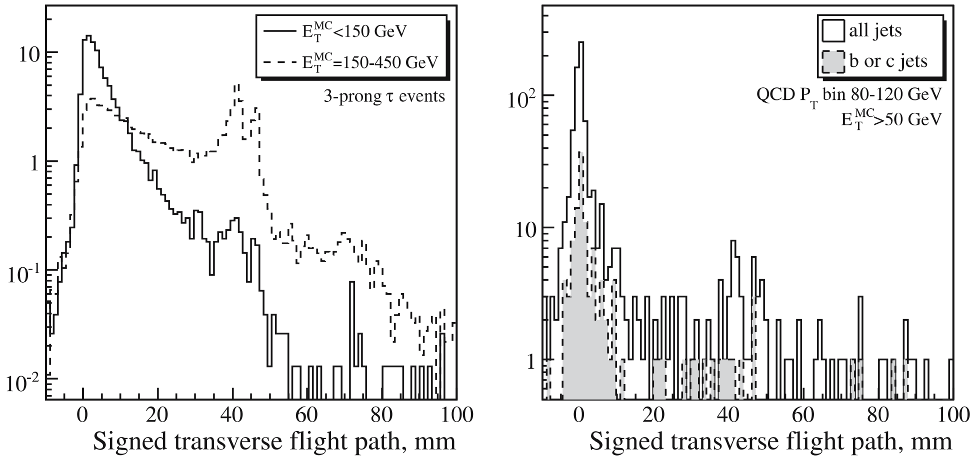

Fig. 17. Signed transverse flight-path for τ (*left*) and QCD (*right*) jets passed the tracker isolation. Three reconstructed tracks was required to be in the signal cone. The detailed explanation can be found in the text

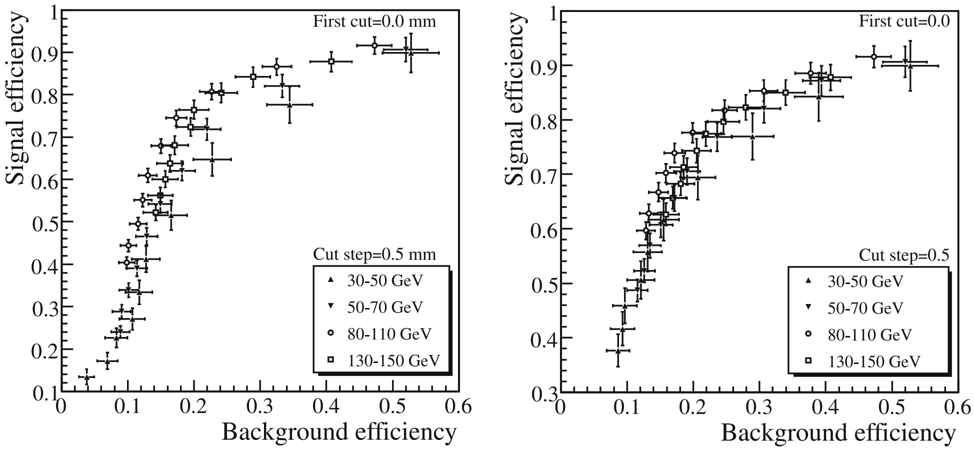

Fig. 18. Background rejection efficiency based on a cut on the signed transverse flight-path (*left*) and on the signed transverse flight-path significance (*right*) for four E_T^{MC} bins. The efficiency has been computed after applying MC preselection and matching and after the tracker isolation

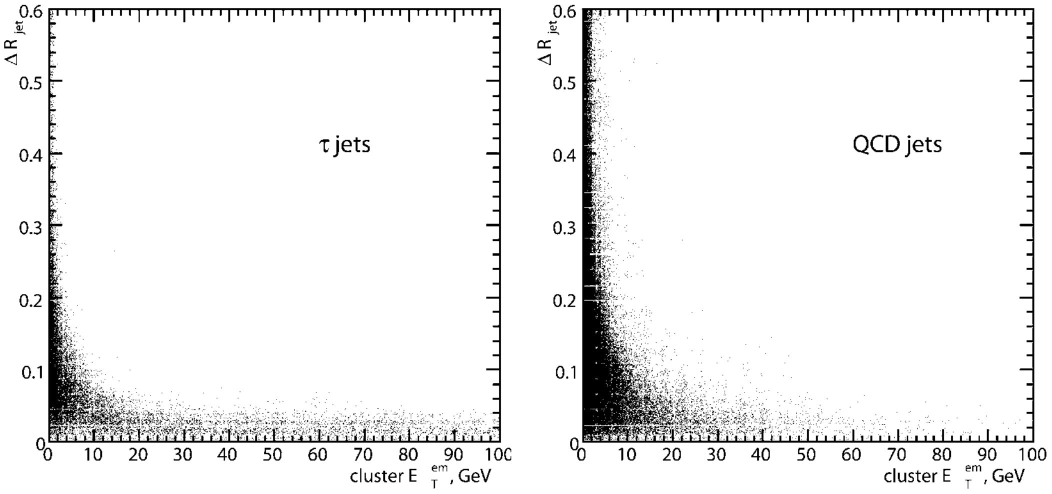

Fig. 19. Scatter plots of the transverse energy of the electromagnetic clusters, E_T^{em}, and the distance ΔR_{jet} in the η–φ space between the calorimeter jet axis and the clusters for the τ jets (*left plot*) and for the QCD jets E_T^{MC} between 30 and 150 GeV (*right plot*)

position in z coordinate was smeared with a resolution of 60 μm. It was found that the error in the flight-path is completely dominated by the secondary vertex resolution.

The performance of the τ-jet tagging with the flight-path was evaluated using Kalman vertex fitter. Since the background rejection is based on a cut on the flight-path length or on the flight-path significance, the large tail in the background distribution reduces the rejection efficiency. The upper cut of 35 mm on the transverse flight path was used to remove a part of the fake vertices. The efficiency of this cut is included in the performance plots. Additionally, a minus sign is assigned to the flight-path length and flight-path significance, if the SV is reconstructed behind the primary vertex compared to the reconstructed jet axis. Figure 18 shows the rejection efficiency based on the signed transverse fligh-path length and signed transverse flight-path significance for several bins of E_T^{MC}. It can be concluded that a rejection factor of 5 can be achieved with the efficiency of 70%–80% for jets E_T between 30 and 150 GeV.

3.5 Tagging with the τ mass reconstruction

In this section the τ-tagging method using the constraint on the reconstructed mass of the τ jet, $M_{\tau\ \mathrm{jet}}$, is discussed. After the tracker isolation the τ-jet mass is reconstructed from the momentum of the tracks in the signal cone and the energy of the clusters in the electromagnetic calorimeter within a certain cone in the η–φ space around the calorimeter jet axis.

Figure 19 shows the scatter plots of the transverse energy of the electromagnetic clusters, E_T^{em}, and the distance ΔR_{jet} in the η–φ space between the calorimeter jet axis and the clusters for the τ jets (left plot) and the QCD jets with E_T^{MC} between 30 and 150 GeV (right plot). The ΔR_{jet} and E_T^{em} are strongly correlated, a constraint on one variable would naturally restrict the other one. The clusters are closer to the jet axis for the τ jets than for the QCD jets and for both τ and the QCD jets the energetic clusters with $E_T^{em} > 10$ GeV are located mostly within the cone of the size 0.1 around the calorimeter jet axis. The cone size of 0.4 was found to be the optimal for the τ-tagging performance. A smaller cone size reduces the capability to distinguish the τ and the QCD jets with the constraint on the $M_{\tau\ \mathrm{jet}}$ value.

The $M_{\tau\ \mathrm{jet}}$ calculated from the tracks in the signal cone and from the clusters with $\Delta R_{jet} < 0.4$ shows a very broad distribution with the long tail as shown in Fig. 20. This tail is due to double counting, when the clusters in the ECAL produced by the charged particles are taken in the $M_{\tau\ \mathrm{jet}}$ calculation. These clusters are rejected with a requirement on a track-cluster matching. The cluster, taken for the mass calculation, must be separated from the track impact point on the ECAL surface by the distance $\Delta R_{track} > 0.08$. The improvement in the $M_{\tau\ \mathrm{jet}}$ the distribution is shown in Fig. 20. There is a large peak at zero value of $M_{\tau\ \mathrm{jet}}$ due to the single track events with no ECAL clusters satisfying the constraints on ΔR_{jet} and ΔR_{track}.

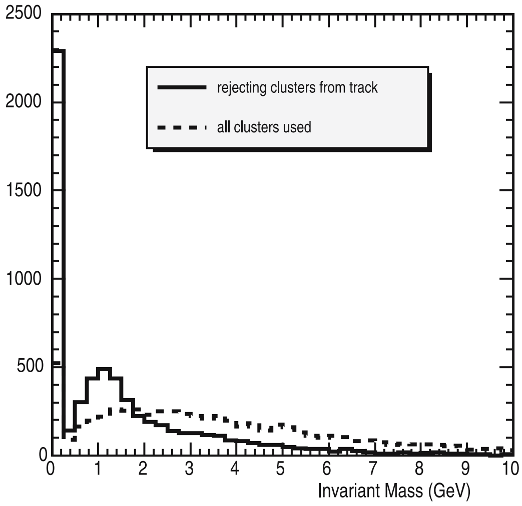

Fig. 20. $M_{\tau\ \mathrm{jet}}$ distribution for τ jets when all clusters with $\Delta R_{jet} < 0.4$ are taken (*dashed line*) and when clusters not matched with tracks ($\Delta R_{track} > 0.08$) are used (*solid line*)

The distributions for different E_T^{MC} bins of τ and QCD jets are shown in Figs. 21 and 22. The selection efficiency of the cut $M_{\tau\ \mathrm{jet}} < 2.5$ GeV is shown in Table 5 for τ and QCD jets in three intervals of E_T^{MC}. The efficiency for the τ jets depends only slightly on the jet E_T, while there is a strong dependence for QCD jets leading to a better rejection factor for the larger E_T jets.

3.6 Rejection of electrons

A genuine electron passes throw all the τ tagging criteria described above, except the impact parameter tagging when the electron originates from the primary signal vertex. Two methods providing similar performance were tested to suppress the electron-τ jet miss-identification rate in off-line analysis. In the first method the electron

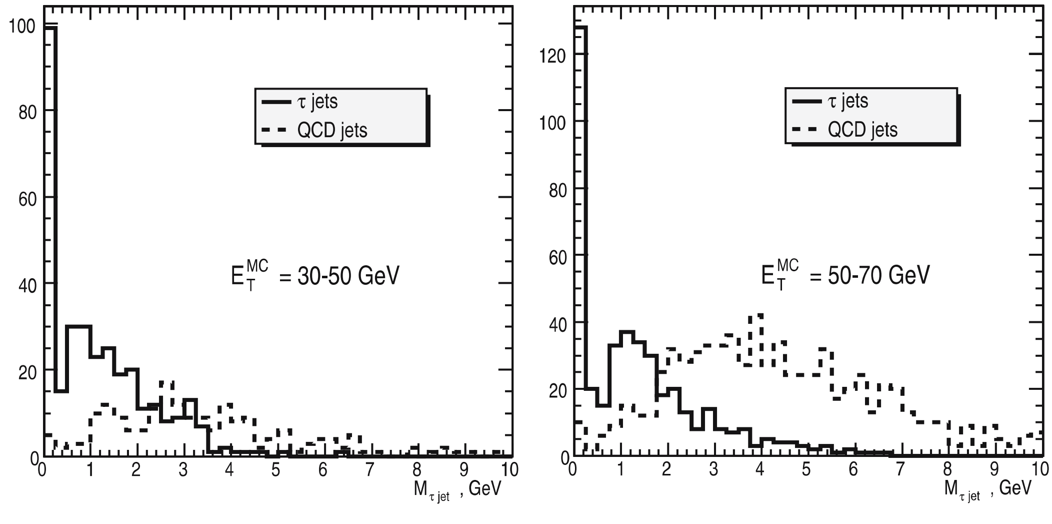

Fig. 21. Distribution for $M_{\tau\ \mathrm{jet}}$ of τ jets (*solid line*) and QCD jets (*dashed line*) in the E_T^{MC} bins of 30–50 GeV (*left plot*) and 50–70 GeV (*right plot*)

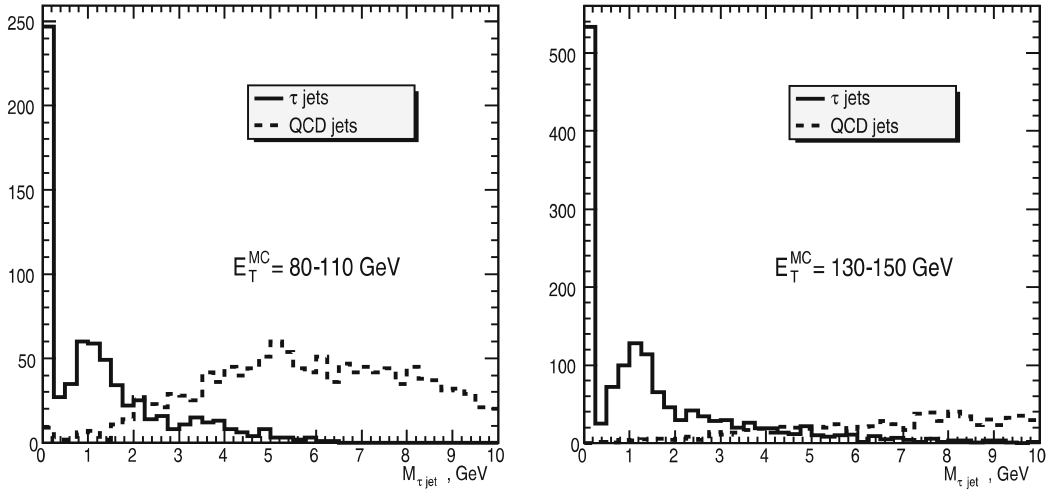

Fig. 22. Distribution for $\mathrm{M}_{\tau\ \mathrm{jet}}$ of τ jets (*solid line*) and QCD jets (*dashed line*) in the $\mathrm{E}_\mathrm{T}^\mathrm{MC}$ bins of 80–110 GeV (*left plot*) and 130–150 GeV (*right plot*)

Table 5. Selection efficiency of the cut $\mathrm{M}_{\tau\ \mathrm{jet}} < 2.5\ \mathrm{GeV}/c^2$ for the τ and the QCD jets for different intervals of $\mathrm{E}_\mathrm{T}^\mathrm{MC}$. The efficiency has been computed after having applied the MC preselection and matching and after the tracker isolation

$\mathrm{E}_\mathrm{T}^\mathrm{MC}$ bins, GeV	30–50	50–70	80–110	130–150
Eff. for τ jets, %	86.32	82.27	83.02	80.76
Eff. for QCD jets, %	33.67	19.16	6.05	2.47

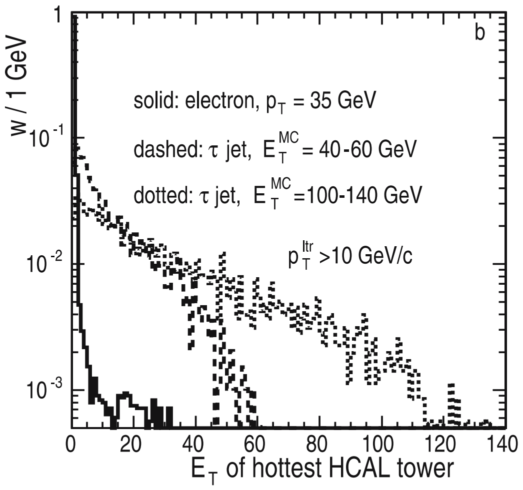

Fig. 23. Transverse energy of the maximal E_T HCAL tower belonging to the reconstructed jet. *Solid line* – electron of $\mathrm{p}_\mathrm{T} = 35\ \mathrm{GeV}/c$ reconstructed as a jet. *Dashed* (*dotted*) *line* – τ jet in the range true τ-jet transverse energy of 40–60 GeV (100–140 GeV). The cut on the measured transverse momentum of the leading track in τ jet, $\mathrm{p}_\mathrm{T}^\mathrm{ltr} > 10\ \mathrm{GeV}/c$, was applied. All histograms are normalized on one

rejection is done using a lower cut-off on the transverse energy of the maximal E_T Hadron Calorimeter (HCAL) tower belonging to the reconstructed jet. In the second method a lower cut-off was applied on the value of $\mathrm{E}_\mathrm{T}^\mathrm{hadr}/\mathrm{p}_\mathrm{T}^\mathrm{ltr}$, where $\mathrm{E}_\mathrm{T}^\mathrm{hadr}$ is the transverse energy of the τ jet measured in the HCAL only and $\mathrm{p}_\mathrm{T}^\mathrm{ltr}$ is the transverse momentum of the leading track measured in the tracker. In the following the first method is presented. Figure 23 shows the transverse energy of the maximal E_T HCAL tower for an electron with $\mathrm{p}_\mathrm{T} = 35\ \mathrm{GeV}/c$ reconstructed as a jet and for τ jets in two ranges of the true τ-jet transverse energy, 40–60 GeV and 100–140 GeV. A cut on the measured transverse momentum of the leading track in the τ jet, $\mathrm{p}_\mathrm{T}^\mathrm{ltr} > 10\ \mathrm{GeV}/c$, was applied. Table 6 shows the efficiency for two cuts on the transverse energy of the maximal HCAL tower belonging to the jet for an electron of $\mathrm{p}_\mathrm{T} = 35\ \mathrm{GeV}/c$ reconstructed as a jet and for a τ jet in two ranges of the true transverse energy of the τ jet and with two cuts on the transverse momentum of the leading track in the τ jet. The high E_T tail for the electron in Fig. 23 corresponds to electrons going through the η/ϕ gaps of the ECAL and barrel/endcap cracks.

Misidentification between a muon and a τ jet was not considered, since the average energy losses of the muon in the calorimeter are an order of a few GeV, therefore well

Table 6. Efficiency of the cut on the transverse energy of the maximal E_T HCAL tower for a jet reconstructed from an electron with $\mathrm{p}_\mathrm{T} = 35\ \mathrm{GeV}/c$ and for a τ jet in two ranges of the true transverse energy and for two cuts on transverse momentum of the leading track ($\mathrm{p}_\mathrm{T}^\mathrm{ltr}$) in τ jet

cut	electron	τ jet E_T 40–60 GeV		τ jet E_T 100–140 GeV	
		$\mathrm{p}_\mathrm{T}^\mathrm{ltr} > 10\ \mathrm{GeV}/c$	$\mathrm{p}_\mathrm{T}^\mathrm{ltr} > 25\ \mathrm{GeV}/c$	$\mathrm{p}_\mathrm{T}^\mathrm{ltr} > 10\ \mathrm{GeV}/c$	$\mathrm{p}_\mathrm{T}^\mathrm{ltr} > 25\ \mathrm{GeV}/c$
> 1 GeV	0.08	0.936	0.971	0.977	0.991
> 2 GeV	0.03	0.854	0.917	0.942	0.969

below the lowest E_T threshold for τ jet used in the physics analyses (15–30 GeV).

4 High level trigger

4.1 Introduction

The ECAL and tracker isolation are used at the High Level trigger (HLT) [6] for the τ-jet identification. The performance of the τ tagging at the HLT was evaluated for the most difficult case of triggering on the decay of the MSSM neutral Higgs boson into two τ leptons when both τ's decay hadronically, thus producing two τ jets in the final state. At the HLT a double τ-jet identification is needed to suppress the rate from the single or double Level 1 τ trigger [6]. Table 7 shows the QCD multi-jet background rate in kHz at the luminosity of $2 \times 10^{33}\,\mathrm{s^{-1}cm^{-2}}$ for the Level 1 single, double, and single or double τ triggers with the single (double) trigger threshold of 93 (66) GeV optimized in [6, 9]. The rate is shown for six $\hat{p}_T$ bins of the background between 30 and 300 GeV/c. The three $\hat{p}_T$ bins in the interval between 50 and 170 GeV/c give the dominant ($> 90\%$) contribution to the rate. Therefore, these three bins are used to evaluate the rejection factor at the HLT with the double τ-jet tagging. The signal efficiency of the HLT selections was evaluated for two masses of the MSSM neutral Higgs boson: 200 and 500 GeV/c^2 produced in the association with a $b\bar{b}$ quark pair. The presence of b jets in the final state can lead to the presence of tracks inside the τ jet coming from the b quark decays, which can affect the tagging performances. All the efficiencies presented in this section are given with respect to events which pass the Level 1 single or double τ trigger.

At the HLT the two jets are reconstructed with the calorimeter in the regions given by the first and the second Level 1 τ jets. If the second Level 1 τ jet does not exist in the Global Level 1 calorimeter trigger output, the jet is reconstructed in the region of the first Level 1 Central jet (the jets are ordered in transverse energy, so the first jet means the largest E_T jet). For the signal events the two jets selected in this way have a good purity, 97% for the first and 82% for the second jet, moreover the purity does not depend on the Higgs boson mass between 200 and 800 GeV/c^2. The other tagging methods previously discussed (track counting, impact parameter, flight path, invariant mass) are not used at the High Level trigger since the isolation alone and the cut on the p_T of the leading track were proved to be sufficient to reject the QCD background down to the acceptable level.

Table 7. Rate for the QCD multi-jet background in kHz at a luminosity of $2 \times 10^{33}\,\mathrm{s^{-1}cm^{-2}}$ for the Level 1 single, double, and single or double τ triggers with single (double) trigger threshold of 93 (66) GeV. The rate is shown for six $\hat{p}_T$ bins of the background generation

$\hat{p}_T$, GeV/c	cross section, fb	Rate, kHz		
		single τ	double τ	single or double τ
30–50	1.56×10^{11}	0.04	0.08	0.12
50–80	2.09×10^{10}	0.59	0.70	1.19
80–120	2.94×10^{9}	1.32	0.75	1.65
120–170	5.00×10^{8}	0.46	0.16	0.48
170–230	1.01×10^{8}	0.10	0.03	0.10
230–300	2.39×10^{8}	0.02	0.007	0.021
total rate		2.53	1.73	3.56

The next step in the HLT selection is the τ tagging of the two τ jets coming from the $H \to \tau\tau$. Two different approaches are investigated:

- ECAL isolation, followed by the tracker isolation with the tracks reconstructed using only the pixel detector.
- Tracker isolation with the regional track reconstruction using both the pixel and the silicon tracker layers.

The first approach is fast and gives a good performance as far as the isolation algorithm is concerned. It is therefore the preferred approach for decays with two taus in the final state (like $A/H \to \tau\tau$) where the isolation is sufficient to reach the required background rejection factor. The second approach is slower but gives a more accurate estimation of the track momenta. It is therefore useful in the channels like charged Higgs boson decay into τ lepton and neutrino. The HLT selection for these events requires a large missing E_T and the tracker τ isolation must complete the selection with a tight cut on the momentum of the leading p_T track in the signal cone [6]. More details on the logic of the trigger system can be found in [2, 6, 7, 10].

4.2 The τ selection based on ECAL and pixel isolation

The ECAL plus the pixel-track isolation at the HLT is referred to as the Calo+Pxl τ trigger. In this approach, the Level 1 rate is first suppressed with a factor of $\simeq 3$ with the ECAL isolation applied to the first jet. Figure 24 shows the efficiency of the ECAL isolation for the signal

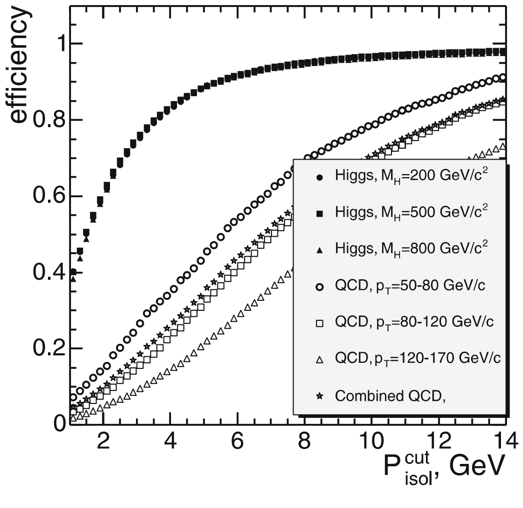

Fig. 24. Efficiency of the ECAL isolation at the High Level trigger for the signal and for the QCD multi-jet background as a function of the cut $P_{\mathrm{isol}}^{\mathrm{cut}}$ on the ECAL isolation parameter

and for the QCD multi-jet background as a function of the ECAL isolation parameter cut $P_{\rm isol}^{\rm cut}$. The efficiency is shown for events which pass the Level 1 single or double τ trigger. A rejection factor of three can be achieved with $P_{\rm isol}^{\rm cut} = 5$ GeV. The remaining background rate is suppressed with the tracker isolation using the information only from the Pixel detector. The isolation requirement is applied to both jets.

With three pixel hits one can reconstruct tracks using the pixel detector only, such tracks are called pixel-tracks. The algorithm used to find tracks is explained in detail in [11], here only the most important details are given. Pixel hit pairs from the first two layers (barrel+barrel or barrel+endcap) are matched in r–ϕ and z–r planes to establish track candidates. The cuts are optimized for a minimum track p_T of 1 GeV/c. Valid pixel pairs are matched with a third pixel hit forming pixel-tracks. The momentum of the pixel-tracks is then reconstructed from the three pixel hits without the primary vertex constraint. The number of fake pixel-tracks in the isolation cone was found to be very low (3%–4%).

Using pixel-tracks a list of primary vertices is formed at z values where several tracks cross the z axis. Only primary vertices with at least 3 valid tracks are kept and the position of each vertex is estimated as the mean value of the z impact parameters of all tracks assigned to it [11].

Pixel-tracks are then used by the isolation algorithm which was described in Sect. 3.2. The leading p_T track found in the cone $R_m = 0.1$ around the calorimeter jet axis must have a transverse momentum $p_T^{\rm ltr}$ larger than 3 GeV/c. The leading track from the first jet defines the primary vertex. This selection is very pure, in 99% of the cases it corresponds to the position of the true Monte Carlo vertex. All tracks used for the τ tagging must be associated with this vertex. In addition, all tracks considered in the selection of the second jet also have to be associated with the primary vertex used for the first jet. Therefore tracks from other vertices are ignored.

The signal cone R_s was set to 0.07 and the isolation cone R_i was varied as a free parameter to adjust the trigger rate. Figure 25 shows the performance of the Calo+Pxl trigger. The selection efficiency plotted on both axes is defined relative to the events passing the Level 1 single or double τ trigger. The upper plot in Fig. 25 shows the performance for the trigger for the first jet and the bottom plot shows the performance of the double jet trigger. The nine points correspond to a step of 0.5 of the isolation cone R_i between 0.2 and 0.6. The required suppression of the QCD multi-jet background of about 10^{-3} can be achieved with R_i between 0.45 and 0.50, with the signal efficiency of 0.29–0.32.

In order to test the Primary Vertex (PV) requirement the trigger efficiency versus the background rejection without the common PV constraint is shown in Fig. 26. These results should be compared to those of Fig. 25. Without a PV constraint the efficiency for the signal increases by few percent due to the primary vertex reconstruction inefficiency. However, the PV constraint becomes very important at high luminosities where multi-jet background originating from several pile-up interactions becomes significant.

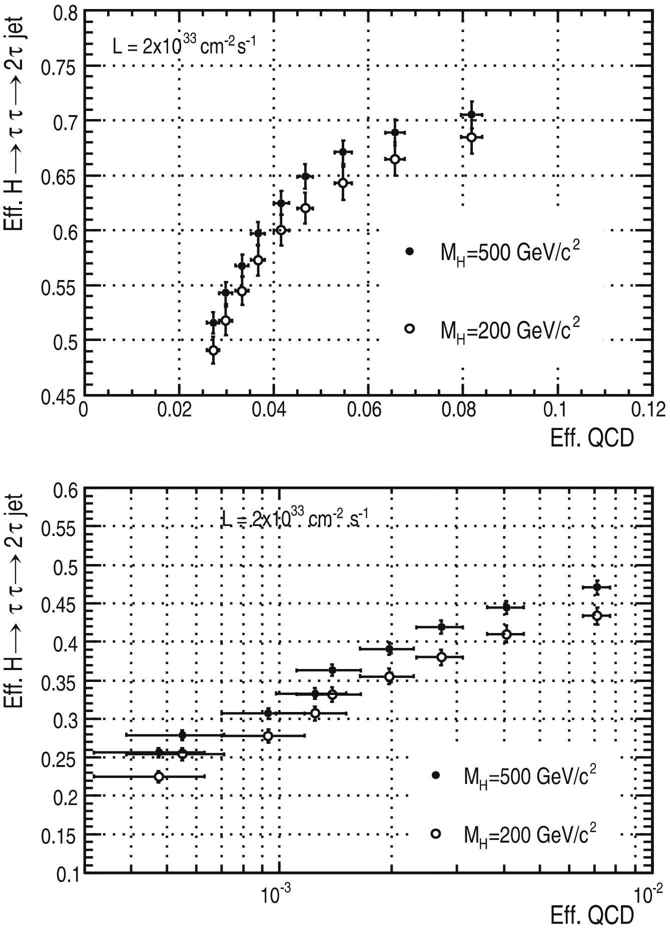

Fig. 25. Efficiency of the Calo+Pxl trigger applied to the first jet (*upper plot*) and to both jets (*bottom plot*) for signal events versus QCD multi-jet events. The efficiencies are shown for two Higgs boson masses of $M_H = 200$ and 500 GeV/c^2.The isolation cone is varied from 0.2 to 0.6 in steps of 0.05, the signal cone is fixed to 0.07, the matching cone to 0.1 and the p_T of the leading tracks is required to exceed 3 GeV/c

A similar study was performed in [6]. The results presented there were obtained under somewhat different conditions: simpler multiple interaction model in PYTHIA, different detector simulation model, different strategy to search for the second τ-jet candidate after the Level 1 trigger and in particular, smaller electronic noise in the HCAL. In addition, for the signal events preselection cuts at the generator level were used. Within these limitations the agreement is satisfactory. In Fig. 27 all events were preselected (at the generator level) with the selection cuts used in [6]. For the Higgs boson with $M_H = 500$ GeV/c^2 the efficiency for the first τ jet was found to be 0.68 in [6] as compared to 0.67 obtained presently at the same value of $R_i = 0.35$.

In the following paragraphs a closer examination of the Calo+Pxl isolation is presented. To better understand the features of the algorithms two special “signal” Monte Carlo samples were used. One is the “pure-tau” sample where only the tracks from the decay of the two τ’s were included in the simulation (Sect. 1). The second one is the “pure-

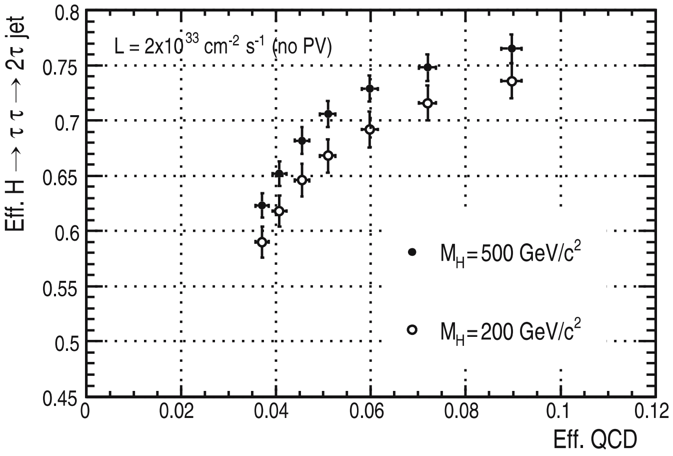

Fig. 26. Calo+Pxl τ efficiency for the 1st τ jet without using the primary vertex constraint. The variable R_i is varied from 0.2 to 0.5

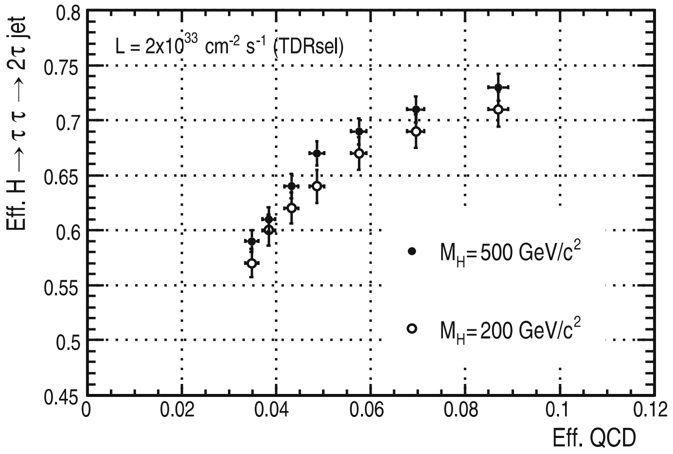

Fig. 27. Calo+Pxl τ efficiency for the 1st τ jet with the DAQ-TDR MC event preselection. The variable R_i is varied from 0.2 to 0.5

tau-PU” sample where in addition the low luminosity pile-up was included. These very “clean” events are compared to the standard gg → bbH/A → bb$\tau\tau$ events used for all the other HLT studies. These genuine signal events present more complex final states where tracks from the b jets and from the underlying event can be mismatched with the tracks coming from the hadronic decays of τ.

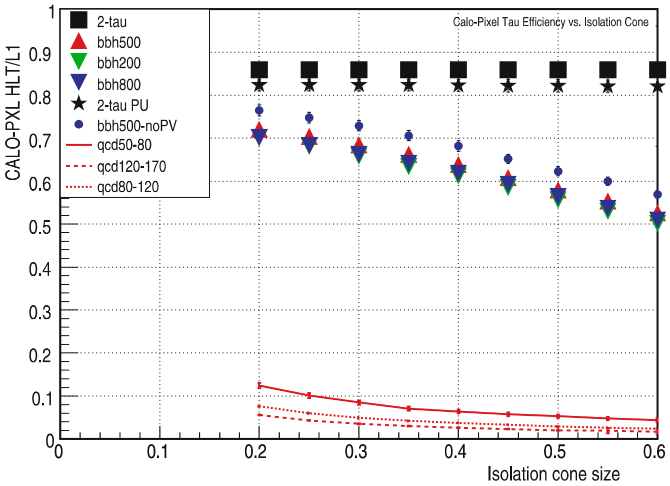

Fig. 28. Efficiency of the Calo+Pxl τ HLT for the 1st τ jet versus the size of the isolation cone

In Fig. 28 the efficiency of the Calo+Pxl HLT is shown versus the isolation cone (R_i) for several samples of MC events. All efficiencies are calculated with respect to the events which pass the Level 1 trigger.

For the pure-tau events the efficiency is constant and equal to 86%. The sources of inefficencies are presented in Table 8. The numbers are for $R_i = 0.35$. The label “no tracks in jet cone” means that no pixel tracks were found in the jet cone (isolation cone + signal cone), which can happen if the τ jet has no charged tracks above the 1 GeV/c cut-off or the tracks were lost due to the pixel detector inefficiencies. The label “no leading track” means that no track candidate was found above the 3 GeV/c cut-off. The label “not isolated” means that tracks were found in the isolation cone. The samples bbH200, bbH500 and bbH800, are the signal samples with respective Higgs boson masses of 200, 500 and 800 GeV/c^2. Jets reconstructed at the second trigger level are labelled as L2 Calo jets.

For the pure-tau events with pile-up the efficiency is lower (82%) than for the pure-tau events without pile-up. The loss of effciency is due to a lower (by about 2%) effi-

Table 8. Sources of an HLT negative response for different event types with $R_i = 0.35$ for the 1st τ jet

Event type	No L2 Calo jets	No tracks in jet cone	No leading track	Not Isolated Not Isolated	No PV No PV
pure-tau	2.9%	9.5%	1.2%	0.4%	–
pure-tau-PU	5.0%	10.0%	2.2%	0.8%	–
bbH500	9.0%	7.9%	3.2%	9.3%	4.4%
bbH200	9.4%	7.5%	5.1%	11.0%	3.4%
bbH800	9.3%	8.7%	3.2%	9.9%	4.8%
qcd50-80	49%	1.7%	6.1%	36.0%	0.3%
qcd80-120	64%	0.8%	3.1%	27.0%	0.2%
qcd120-170	75%	0.5%	1.6%	20.0%	0.2%

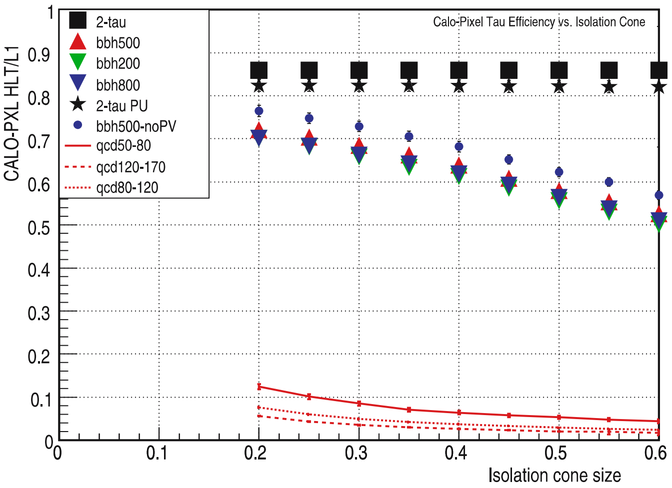

Fig. 29. Efficiency of the Calo+Pxl HLT for both τ jets versus $\mathrm{R_i}$ (linear scale)

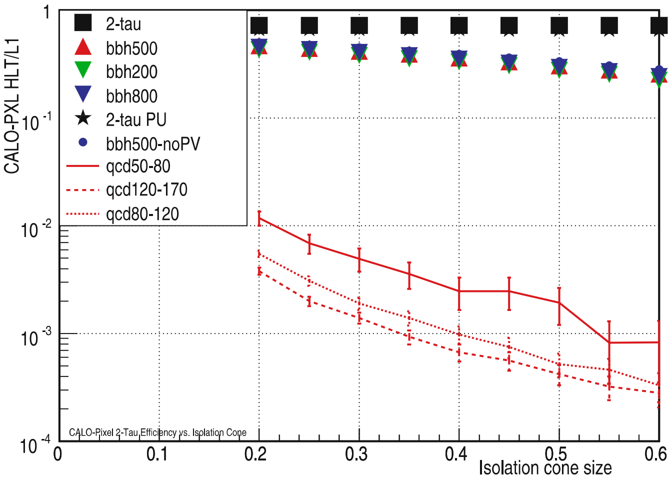

Fig. 30. Efficiency of the Calo+Pxl HLT for both τ jets versus $\mathrm{R_i}$ (logarithmic scale)

ciency of the calorimeter reconstruction at the Level 2 trigger and a decreasing pixel inefficiency. The presence of pile-up does not affect significantly the isolation performance. This fact is visible also in the behaviour of the efficiency as a function $\mathrm{R_i}$ in Fig. 28. All the other MC event samples of Fig. 28 have been simulated with a low luminosity pile-up. All the three signal event samples (the three Higgs boson masses) show a similar behavior. Without using the PV constraint one gains few percent in efficiency which reflects the finite PV finding efficiency. The QCD backgrounds fall steeply until $\mathrm{R_i}$ of 0.35–0.40 after which the rejection gain slows down.

The efficiency of the pure-tau trigger is shown in Fig. 29 for the same MC event samples. The pure-tau events again do not show much dependence on the isolation cone size and the three signal event samples show similar behavior. The rejection factor for the QCD events is high and therefore the same distributions are shown again in a logarithmic scale in Fig. 30. The fluctuations at large $\mathrm{R_i}$ are due to the low MC statistics which is also indicated with the large error bars. The QCD background with $\mathrm{E_T}$ between 50 and 80 GeV (qcd50-80 sample) is the most difficult one to reduce and requires $\mathrm{R_i}$ of about 0.5 for the 10^{-3} suppression.

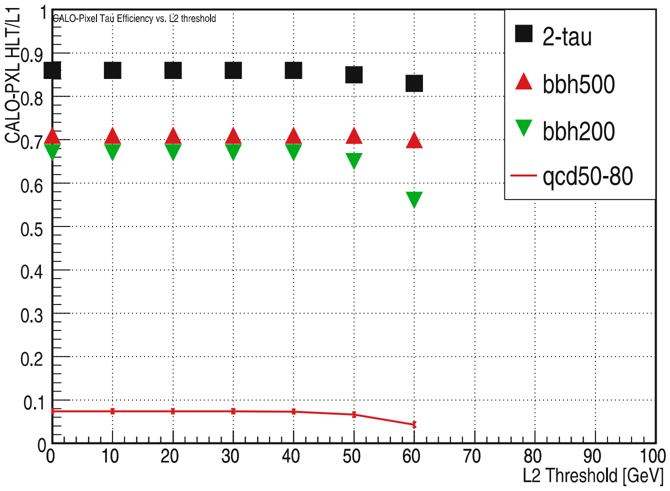

Fig. 31. Efficiency of the Calo+Pxl τ HLT for the 1st τ jet versus the L2 Calo jet $\mathrm{E_T}$ cut

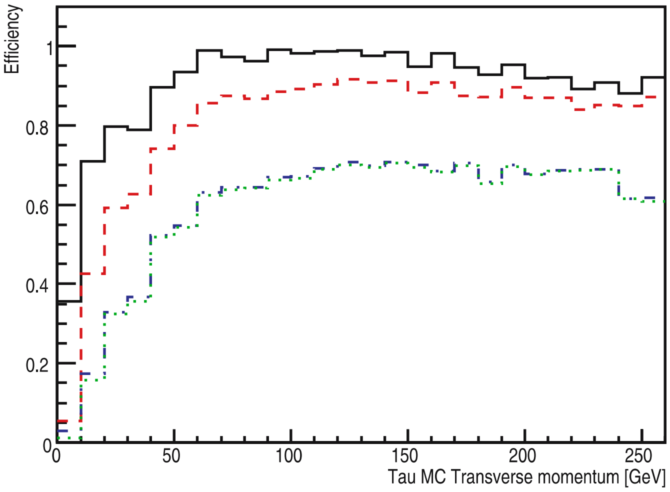

Fig. 32. Calo+Pxl τ efficiency for τ jets from bbh500 event sample versus the MC τ transverse momentum. Efficiency is with respect to the events passing the Level 1. The meaning of the lines is the following: *Solid* – L2 Calo jet reconstruction efficiency, *dashed* – L2 Calo jets matched to MC, *dashed-dotted* – pixel-tau HLT efficiency, *dotted* – pixel-tau matched to MC

An attempt was made to vary the cut on the $\mathrm{E_T}$ jet threshold at the second level trigger (L2 Calo $\mathrm{E_T}$) for an additional rejection. Figure 31 shows that a visible background reduction requires a threshold of 50–60 GeV for which the signal (especially for the light Higgs) is also reduced.

A more detailed study of the efficiencies was done for the bbH500 sample. The HLT efficiency versus the $\mathrm{p_T}$ of the MC τ is plotted in Fig. 32. The L2 Calo jet reconstruction efficiency increases rapidly above 10 GeV$/c$ and

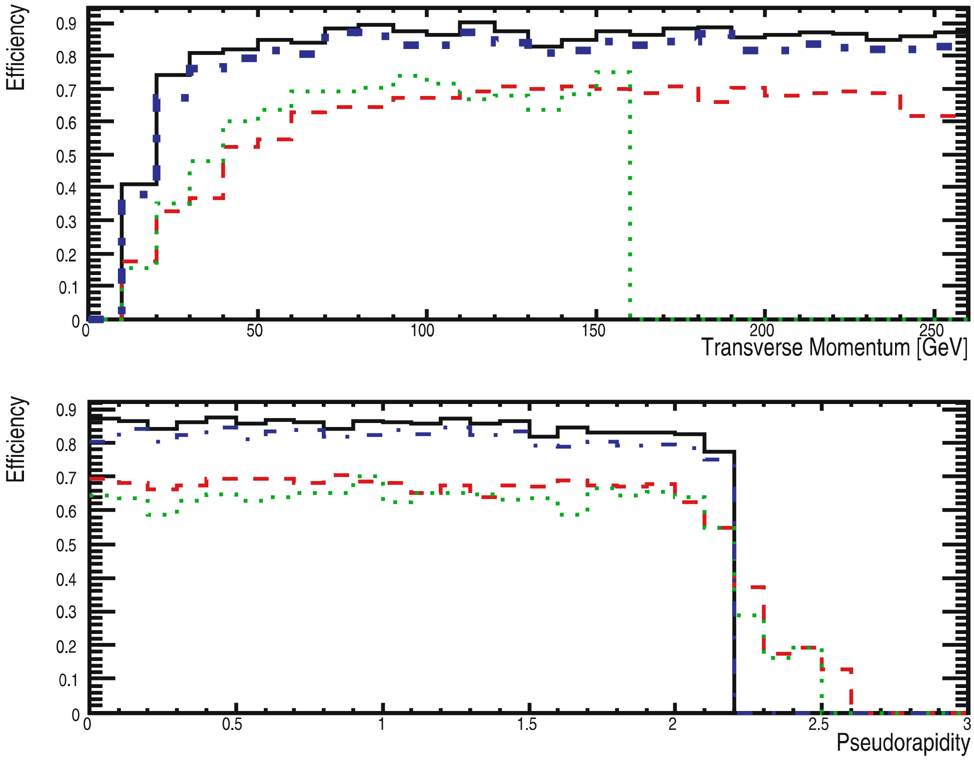

Fig. 33. Comparison of Calo+Pxl τ efficiencies for various event types versus p_T (*upper figure*) and η (*lower figure*) of the MC τ. The events are: pure-tau (*solid line*), pure-tau-PU (*dashed-dotted line*), bbH500 (*dashed line*) and bbH200 (*dotted line*)

reaches an almost 100% plateau for $\mathrm{p_T} > 60$ GeV. Figure 32 shows also that the jet reconstructed at the second level of the trigger (L2 jets) is well matched ($\Delta \mathrm{R} \leq 0.5$) to the direction of the MC τ. About 10% of the jets are not matching with the true τ. The efficiency of the pixel HLT selection rises lower with the $\mathrm{p_T}$ of the τ and reaches a plateau of about 65% at 60 GeV. All the τ jets passing the pixel trigger match well the MC τ's, which shows that the pixel-tau HLT has a high purity.

The samples for the pure-tau, pure-tau-PU, bbH500 and bbH200 are compared in Fig. 33 as a function of $\mathrm{p_T}$ and η of the MC τ. As a function of $\mathrm{p_T}$ a plateau is reached around 60 GeV/c for all the samples. As a function of η the τ efficiency is flat in rapidity up to $\eta = 2.1$ after which the pixel detector loss of coverage is visible. The efficiency is similar for the three Higgs boson masses.

4.3 The τ selection based on silicon tracker isolation

The algorithm described in this paragraph is referred to as the Trk-Tau trigger. Due to the time limitation it is not possible to perform a full tracker reconstruction of the whole event after the Level 1 Tau trigger. It is possible however, to read and reconstruct a selected part of the tracker data. The Trk-Tau trigger reconstructs only of tracks confined in the restricted regions of interest ("regional tracking"), defined with a cone around the calorimeter jet direction. The primary signal vertex needed in the Trk-Tau trigger is obtained using only the pixel detector in order to ensure the fast reconstruction. Once the signal vertex is found the regional track reconstruction starts. The tracker isolation algorithm is then applied using the tracks with the z-impact parameter close to the z position of the signal vertex.

4.3.1 Trk-Tau trigger performance

Figure 34 shows the Trk-Tau trigger performance in terms of the signal versus background efficiency for the events that have passed the Level 1 single or double Tau trigger. The upper plot shows the efficiency for the first jet, while the bottom plot shows the double jet tagging efficiency. The size of the isolation cone $\mathrm{R_i}$ is varied between 0.2 and 0.45 with a step of 0.05. The matching cone is set to 0.1 and the signal cone to 0.07. The leading track momentum $\mathrm{p_T^{ltr}}$ must exceed 6 GeV/c. The rejection factor of $\simeq 10^3$ against the QCD multi-jet background can be achieved with $\mathrm{R_i}$ around 0.40.

The main difference between the performances of the Trk and Calo+pxl algorithm comes from the better momentum resolution of the tracks reconstructed using also the silicon tracker layers. The Trk-Tau trigger allows a stronger cut on the $\mathrm{p_T}$ of the leading track. The higher resolution and sensitivity on the $\mathrm{p_T}$ leads to a different efficiency for the two Higgs boson masses (200 and 500 GeV/c^2).

4.3.2 Comparison with older studies

The main differences between the studies in [6] and the studies of this work were mentioned in the previous chapter. Another source of difference is the new implementation of a regional seeding [12] for the track reconstruc-

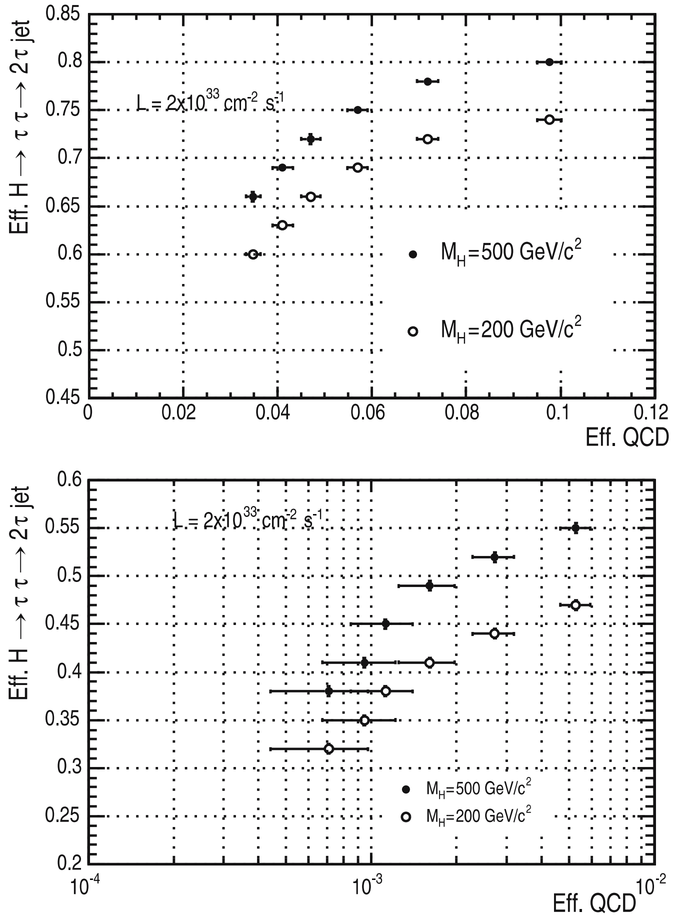

Fig. 34. Efficiency of the Trk-Tau trigger applied to the first jet (*upper plot*) and to the two jets (*bottom plot*) for signal events versus efficiency for QCD multi-jet background. Two masses of the Higgs boson, $\mathrm{M_H} = 200$ and 500 GeV/c^2, are considered. Isolation cone is varied from 0.2 to 0.45 with steps of 0.05, signal cone is fixed to 0.07, matching cone to 0.1 and the $\mathrm{p_T}$ of the leading tracks is required to exceed 6 GeV/c

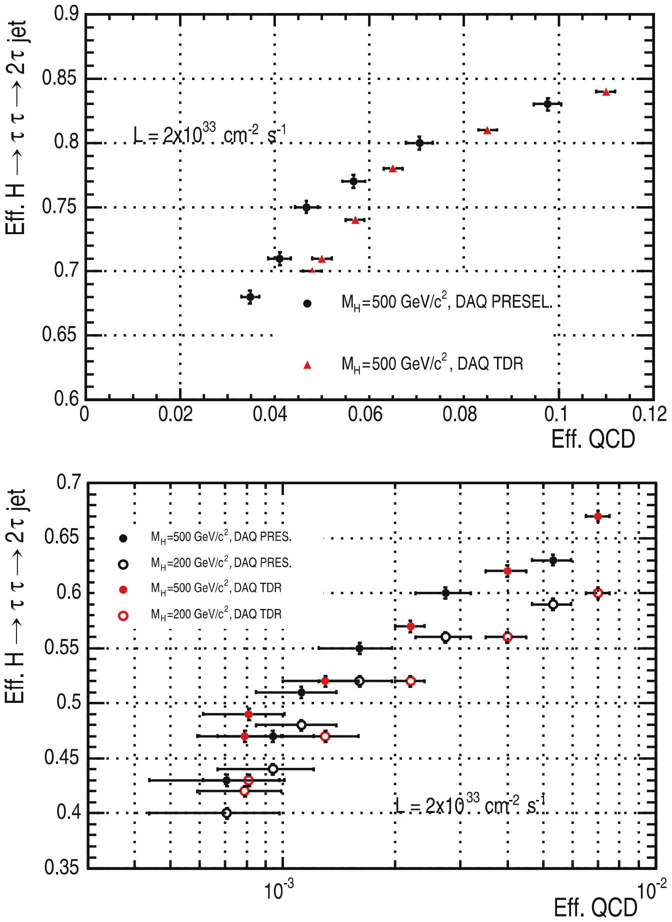

Fig. 35. Efficiency of the Trk-Tau trigger applied to the first jet (*upper plot*) and to the two jets (*bottom plot*) for signal events versus the efficiency for the QCD multi-jet background, using the same preselections for the signal as in [6]. Two masses of the Higgs boson, $\mathrm{M_H} = 200$ and 500 GeV/c^2, are considered. The isolation cone is varied from 0.2 to 0.45, the signal cone is fixed to 0.07, the matching cone to 0.1 and the $\mathrm{p_T}$ of the leading tracks must exceed 6 GeV/c. The results of [6] are also shown in the plots

tion and the primary vertex reconstruction algorithm. In order to compare with the results found in [6], the same preselections at the generation level for the signal events $\mathrm{H/A} \to \tau\tau \to \mathrm{jet}\ jet$ were applied: $\mathrm{p_T^{\tau\ jet}} > 45$ GeV/c and $|\eta^{\tau\ \mathrm{jet}}| < 2.4$.

Figure 35 compares the Trk-Tau trigger performance obtained in the present study and the studies in the [6]. The upper plot shows the efficiency for the first jet, while the bottom plot shows the double jet tagging efficiency. The points correspond to the six values of the isolation cone $\mathrm{R_i}$ between 0.2 and 0.45 with the step of 0.05. The efficiency for the Higgs boson with a mass of 200 GeV/c^2 is in good agreement with the results shown in [6], while that for the mass of 500 GeV/c^2 is $\simeq 5\%$ lower. The lower efficiency of the double τ-jet tagging can be explained, in particular, by $\simeq 8\%$ lower purity of the second jet in the present study. The rejection factor for the QCD multi-jet events is compatible, within 1.5 sigma, with the results in [6]. The disagreement with the last left point ($\mathrm{R_i} = 0.45$) is due to a too small track reconstruction cone used in [6].

5 Calibration and tagging efficiency

5.1 Tau jet energy scale and calibration with calorimeter

The τ-jet energy measurement with the calorimeter requires smaller energy corrections than "normal" QCD jets. The reasons are that, first, the average transverse momentum of the charged hadrons is larger, and second, the fraction of the electromagnetic energy in the τ jet due to π^0s is larger. Figure 36 shows the τ jet energy scale, the ratio $\mathrm{E_T^{reco}/E_T^{MC}}$, as a function of the $\mathrm{E_T^{MC}}$ and τ jet pseudorapidity for four final states of hadronic decays of τ. The jets were reconstructed with the iterative cone algorithm with a cone size of 0.4. The thresholds on the calorimeter towers were set to $\mathrm{E_T} = 0.5$ GeV and $E = 0.8$ GeV. One can see that the τ jet formed from the three charged pions has the lowest response in the calorimeter in comparison to the jet

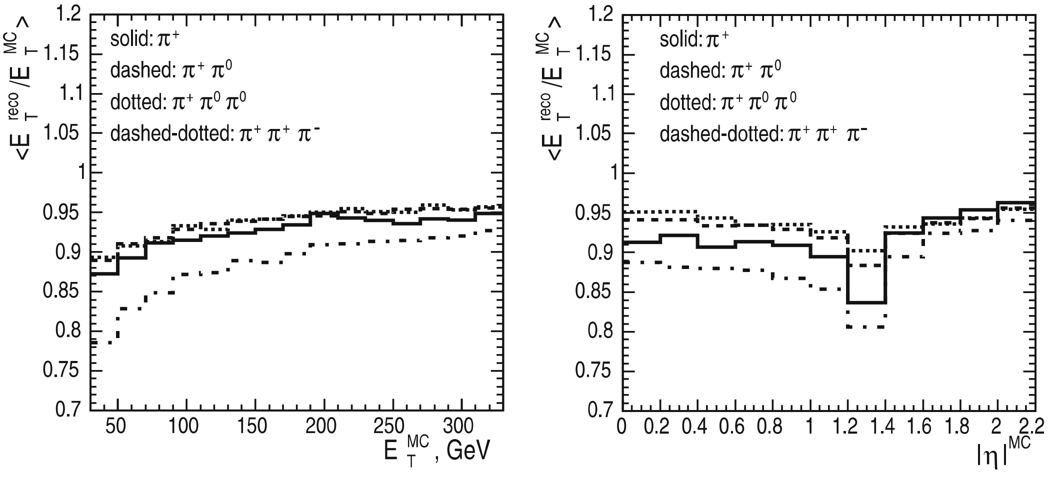

Fig. 36. Mean value of the ratio $\mathrm{E_T^{reco}}/\mathrm{E_T^{MC}}$ as a function of the $\mathrm{E_T^{MC}}$ (*left*) and τ jet pseudorapidity (*right*) for different final states of hadronic decays of τ lepton

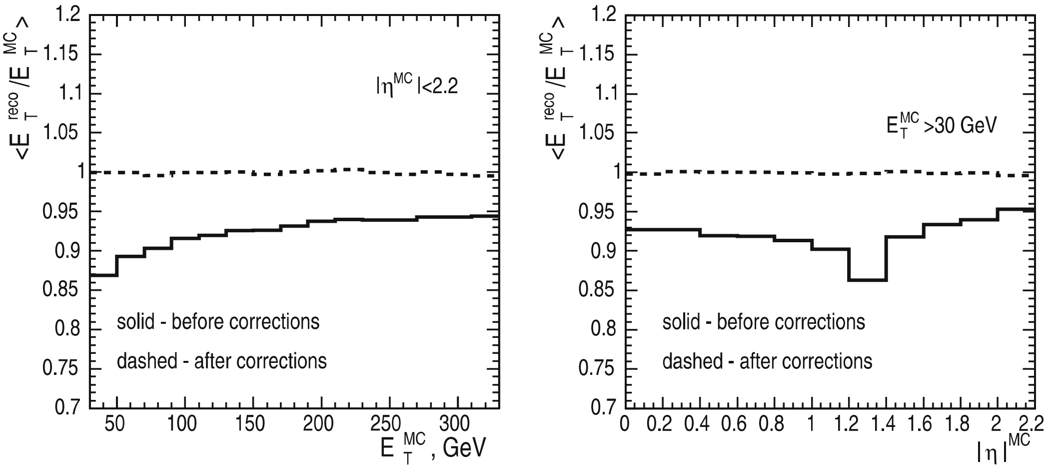

Fig. 37. Mean value of the ratio $\mathrm{E_T^{reco}}/\mathrm{E_T^{MC}}$ as a function of the $\mathrm{E_T^{MC}}$ (*left*) and τ jet pseudorapidity (*right*) before and after the energy corrections were applied

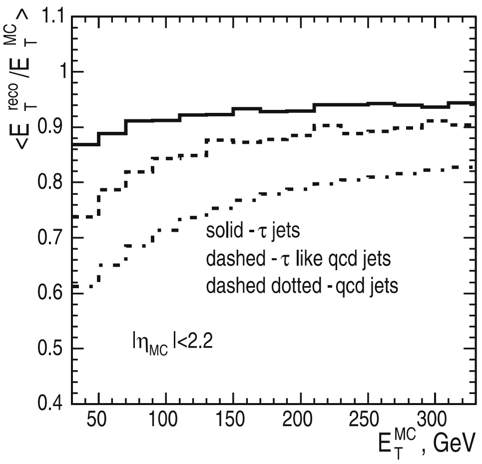

Fig. 38. Mean value of the ratio $\mathrm{E_T^{reco}}/\mathrm{E_T^{MC}}$ for the QCD jets without preselection (*dashed-dotted line*), τ-like QCD jets (*dashed line*) and the real τ jets (*solid line*) as a function of $\mathrm{E_T^{MC}}$

containing only one charged and one or two neutral pions for the same $\mathrm{E_T^{MC}}$. The drop in the response at the pseudorapidity $\simeq 1.4$ is due to the instrumentation gap between the barrel and the end-cap calorimeters.

The energy correction function was obtained from the parameterization of the $\mathrm{E_T^{MC}}$ and $|\eta^{\mathrm{MC}}|$ dependence of the ratio $\mathrm{E_T^{reco}}/\mathrm{E_T^{MC}}$. Figure 37 shows the $\mathrm{E_T^{reco}}/\mathrm{E_T^{MC}}$ ratio before and after the energy corrections were applied. The resolution of the transverse energy of the τ jet after energy corrections can be parameterized with the equation: $\sigma(\mathrm{E_T})/\mathrm{E_T} \simeq \mathrm{a}/\sqrt{\mathrm{E_T}} \oplus \mathrm{b}$, where $\mathrm{a} = 0.883\,\mathrm{GeV}$ and $\mathrm{b} = 0.058$ for τ jets with $\mathrm{E_T}$ between 30 and 300 GeV and pseudorapidity less than 2.2.

The γ-plus-jet events where the jet passes the τ-identification criteria and thus becomes a τ-like jet can be used to setup the initial τ-jet energy scale from the real data. In the following the preliminary results are presented. Figure 38 shows the mean value of the distribution of the ratio $\mathrm{E_T^{reco}}/\mathrm{E_T^{MC}}$ for the unpreselected QCD jets, τ-like QCD jets and the real τ jets. Both the QCD and the τ jets were reconstructed in the calorimeter with a cone size of 0.4. The same cone size was used to evaluate the true transverse energy $\mathrm{E_T^{MC}}$ of the Monte Carlo QCD jets. The τ jet identification includes the ECAL and the tracker isolation with the parameters $\mathrm{P_{isol}^{cut}} = 5\,\mathrm{GeV}$, $\mathrm{R_i} = 0.4$, and $\mathrm{R_s} = 0.07$. The one or three tracks were required to be in the signal cone and a cut $\mathrm{p_T} > 10\,\mathrm{GeV}/c$ on the transverse momentum of the leading track was applied. One can see that the τ-like QCD jets produce a higher calorimeter response than the unpreselected QCD jets, which is only 5%–10% smaller than the response of the real τ jets. More studies are needed to understand the sources of the remaining difference and the calibration uncertainties.

Another method to evaluate the τ-jet energy scale with the data is to use $Z \to \tau\tau \to \ell +$jet events and to reconstruct the Z mass peak. This method, however has two disadvantages: the background contamination and the uncertainty of the missing E_{T} measurement.

5.2 Measurement of jet $\to \tau$ misidentification from the data

The measurement of the jet $\to \tau$ misidentification rate can be done again with the γ-plus-jet events used for the calorimeter calibration. About 10^5 such events are expected in a 10 fb^{-1} data sample for each $\mathrm{E}_{\mathrm{T}}^{\gamma}$ bin of a size of $0.1 \times \mathrm{E}_{\mathrm{T}}^{\gamma}$, with $|\eta^{\mathrm{jet}}| < 3$ and $\mathrm{E}_{\mathrm{T}}^{\gamma}$ in the interval between 30 and 300 GeV. The mistagging rate can be evaluated then as a fraction of events where the jet passed the τ-jet identification criteria. Taking into account the jet rejection factor, for example, with the tracker isolation and the mass tagging (evaluated from the right plot of Fig. 9 for $\mathrm{R_i} = 0.4$ and $\mathrm{R_s} = 0.04$ and from the results of Table 5) one could expect a 4%–10% uncertainty in the estimated mistag rate per energy bin in the jet $\mathrm{E_T}$ interval of 30–150 GeV with a 10 fb^{-1} data sample.

5.3 Measurement of τ-tagging efficiency from the data

The τ-tagging efficiency can be evaluated (and compared with the Monte Carlo) from the ratio of $Z \to \tau\tau \to \mu +$jet and $Z \to \mu\mu$ events selected with the single muon trigger stream. The reconstruction efficiency of the second muon in the $Z \to \mu\mu$ events is assumed to be known. The preliminary estimates were obtained based on the search for MSSM H/A $\to \tau\tau \to \mu +$jet channel described in [14]. The $Z \to \tau\tau \to \mu +$jet event selections are the same as used in [14], but without the b tagging and the jet veto. The systematical uncertainty in the selection cuts on $\mathrm{m_T}(\ell, \mathrm{E}_{\mathrm{T}}^{\mathrm{miss}})$, $\mathrm{E}_{\mathrm{T}}^{\tau\ \mathrm{jet}}$, and $\mathrm{E_T}$ of reconstructed neutrinos, which contain calorimeter informations, were taken into account, as well as the uncertainty of the background evaluation.

With an integrated luminosity of 30 fb^{-1} the total uncertainty of the τ-tagging efficiency is expected to be between 4 and 5%.

6 Conclusions

All the available methods to identify the τ lepton hadronic decays were discussed. The primary requirement for the τ-jet identification is the isolation of af collimated jet made of charged particles reconstructed with the tracker. This method can be completed with a cut on the $\mathrm{p_T}$ of the leading track, impact parameter and vertex tagging and mass tagging. The usage of these methods in different combinations depends on the physics channel considered. A brief recipe to complement the tracker isolation with the other methods can be the following: for the one-prong decays, the tagging with impact parameter is suggested. The mass tag can be used for both one and three-prong decays. The cut on the $\mathrm{p_T}$ of the leading track in one or three-prong decays was found to be very effective to suppress the QCD background in the analysis of the A/H $\to \tau\tau$ and $\mathrm{H}^{\pm} \to \tau\nu$ decay channels. For the three-prong decays the vertex tag can be used. For the HLT, the Calo+Pxl approach is faster and gives a good performance as far as the isolation algorithm is concerned. It is therefore the preferred approach for the decays with two τ's in the final state (like A/H $\to \tau\tau$) where the isolation is sufficient to reach the required background rejection factor. The Trk approach is slower but gives a more accurate estimation of the track momenta. It is therefore useful in channels like the charged Higgs boson decay into a τ lepton and a neutrino where a stronger cut on the $\mathrm{p_T}$ of the leading track is required to achieve the desired trigger rate.

Acknowledgements. The authors would like to thank Ritva Kinnunen, Claire Shepherd-Themistocleous and Vincent Lemaitre for having read the manuscript and their useful comments and suggestions.

References

1. T. Sjostrand, P. Eden, C. Friberg, L. Lonnblad, G. Miu, S. Mrenna, E. Norrbin, Comput. Phys. Commun. **135**, 238 (2001) [hep-ph/0010017]
2. S. Eno, S. Dasu, W. Smith, R. Kinnunen, A. Nikitenko, A Study of a First and Second Level Tau Trigger, CMS NOTE 2000/055
3. D. Denegri, R. Kinnunen, A. Nikitenko, Study of Calorimeter Calibration with Tau's in CMS, CMS NOTE 1997/039
4. S. Abdullin et al., Summary of the CMS Potential for the Higgs Boson Discovery, CMS NOTE 2003/033
5. A. Nikitenko, S. Kunori, R. Kinnunen, Missing Transverse Energy Measurement with Jet Energy Corrections, CMS NOTE 2001/040
6. CMS Collaboration, G.L. Bayatian, CMS: The Tridas Project. Technical Design Report, Vol. 2: Data Acquisition And High-Level Trigger, CERN/LHCC/2002-26, CMS TDR 6.2, December 2002; CMS Trigger and Data Acquisition Group, The CMS High Level Trigger, to be published in The European Physical Journal C
7. D. Kotlinski, A. Nikitenko, R. Kinnunen, Study of a Level-3 Tau Trigger with the Pixel Detector, CMS NOTE 2001/017
8. CMS Collaboration, G.L. Bayatian, The Tracker Project Technical Design Report, CERN/LHCC 98-6; CMS Collaboration, G.L. Bayatian, Addendum to the Tracker TDR, CERN/LHCC 2000-016
9. R. Kinnunen, A. Nikitenko, Study of H → Tau Tau with Hadronic Tau Decays in CMS, CMS NOTE 2003/006
10. G. Bagliesi, S. Gennai, G. Sguazzoni, A L2 Trigger for Tau Hadronic Decays with Tracker Isolation in the Low Luminosity Scenario, CMS NOTE 2002/018

11. S. Cucciarelli, M. Konecki, D. Kotlinski, T. Todorov, Track-Parameter Evaluation and Primary-Vertex Finding with the Pixel Detector, CMS NOTE 2003/026
12. CMS Collaboration, G.L. Bayatian, CMS OO Reconstruction, http://cmsdoc.cern.ch/orca/
13. S. Cucciarelli, D. Kotlinski, T. Todorov, Position Determination of Pixel Hits, CMS NOTE 2002/049
14. R. Kinnunen, S. Lehti, F. Moortgat, A. Nikitenko, M. Spira, Measurement of the H/A → tau tau Cross Section and Possible Constraints on Tan(beta), CMS NOTE 2004/027