

DESY -HAMBURG-

---



Summer Student Programme 2010  
July 20<sup>th</sup>-September 9<sup>th</sup>

# Study of explicite $\gamma$ conversion reconstruction in hadronic $\tau$ lepton decays

Marilea Reale

Salento University - Physics Department (Italy)  
marilea.reale@alice.it

Supervisor: Michael Boehler

## Abstract

The ATLAS experiment is one of the experiments at the Large Hadron Collider (LHC) and its main aim is researching new elementary particles, such as supersymmetric particles or the Higgs boson. Very interesting decay channels in studying both SUSY and the Higgs boson are  $\tau$  lepton final states: this means that it is essential to have a successful reconstruction of  $\tau$  decays and to improve this reconstruction an explicite  $\gamma$  conversion reconstruction algorithm has been implemented.

The main purpose of this work is just to test the efficiency of this algorithm.

# Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Physical process overview</b>	<b>2</b>
2.1	$\tau$ lepton decay . . . . .	3
<b>3</b>	<b>Data analysis</b>	<b>4</b>
<b>4</b>	<b>MC data analysis results</b>	<b>5</b>
4.1	Tracks distribution of $\tau$ candidates . . . . .	5
4.2	Efficiencies of the $\gamma$ conversion reconstruction . . . . .	7
<b>5</b>	<b>Real data analysis results</b>	<b>8</b>
5.1	Investigating $\phi$ distributions of $\tau$ . . . . .	11
<b>6</b>	<b>Conclusion</b>	<b>13</b>
<b>7</b>	<b>Aknowledgements</b>	<b>13</b>

## 1 Introduction

ATLAS is one of the experiments installed along the LHC tunnel, whose main aim is to study Higgs boson or search for new beyond standard model scenarios. Very interesting decay channels in these studies are the  $\tau$  lepton final states: to improve the  $\tau$  decay reconstruction some tools that can efficiently identify photon conversions within the  $\tau$  decay cone with the high track density conditions have been implemented because the general purpose  $\gamma$  conversion reconstruction algorithm is not working in this environment. The aim of this work here in DESY is just to test the efficiency of an algorithm that try to identify the photon conversion in this environment.

## 2 Physical process overview

Before presenting the analysis made in order to test the  $\gamma$  conversion reconstruction algorithm efficiency, an overview of the physical processes involved and the problems to be overcome to improve the  $\tau$  identification should be given.

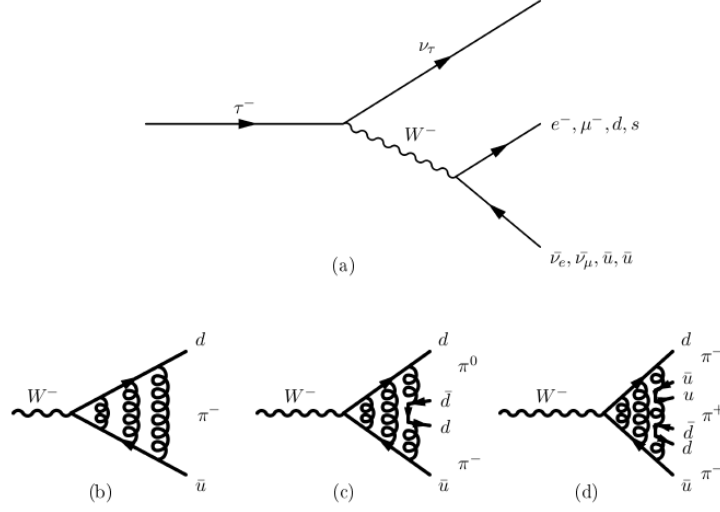


Figure 1:  $\tau$  decay Feynman diagrams.

## 2.1 $\tau$ lepton decay

Because of its large mass ( $m_\tau = 1.78$  GeV) the  $\tau$  lepton can decay into one of the lighter charged leptons ( $m_e = 0.51$  MeV and  $m_\mu = 105.6$  MeV) or hadronically into at least one charged pion ( $m_\pi = 139.6$  MeV).

The general Feynman diagrams of the possible  $\tau$  decays are represented in the Fig. 1(a): due to the flavour conservation the  $\tau^-$  decays into one  $\nu_\tau$  and one virtual  $W^-$  that can decay into  $e^- \bar{\nu}_e$  or  $\mu^- \bar{\nu}_\mu$  or into a quark pair ( $d\bar{u}$  or  $s\bar{u}$ ).

In case of a leptonic decay only the lepton is reconstructed and this can be done with an electron or muon reconstruction algorithm. In the hadronic decay the  $\tau$  reconstruction is more challenging: as shown in Fig. 1(b)(c)(d), if the virtual  $W^-$  decays into a quark pair, the colour connection leads to a meson production (a charged pion in the case of  $d\bar{u}$  or a charged kaon with  $s\bar{u}$ ) and, if there is enough energy, it is possible to have additional  $q\bar{q}$  pair production, and in these cases it is possible to have, in general, in order to conserve the total charge, 1 or 3 charged mesons and additional neutral mesons. These configurations are called *1prong* or *3prong* in reference to the charged tracks number. It is also possible to have 5 charged mesons, but with a very low branching ratio. The branching ratios of kaon production are lower than the pions production because of the different coupling of  $W$  boson to  $d$  ( $\propto \cos^2 \theta_C$ ) and to  $s$  ( $\propto \sin^2 \theta_C$ ). The different branching ratios are listed in Tab. 1: in 40 % of all  $\tau$  decays  $\pi^0$ s are produced. Each  $\pi^0$  decays with a

<b>leptonic</b>	35.2 %		
<b>1prong</b>	46.8%	$\tau \rightarrow \pi^\pm \nu_\tau$	23.4%
		$\tau \rightarrow n\pi^0 \pi^\pm \nu_\tau$	76.6%
<b>3prong</b>	13.9 %	$\tau \rightarrow 3\pi^\pm \nu_\tau$	64.6%
		$\tau \rightarrow n\pi^0 3\pi^\pm \nu_\tau$	35.4%
<b>decay in <math>K^\pm</math></b>	3.8 %		
<b>others</b>	3.8 %		

in general there is  $\pi^0$  production in 40% of all  $\tau$  decays

Table 1:  $\tau$  decay branching ratios.

branching ratio of 98.8% into two photons and if one of these interacts with detector materials it can convert into a  $e^+e^-$  pair and in this case it is difficult to distinguish a true 3prong decay from a 1prong with two additional charged tracks from  $e^+e^-$  pair. Since the photon momentum is not symmetrically shared to the  $e^+e^-$  pair, the momentum of one of the two produced leptons may be below the reconstructable threshold and so it is also possible to see a misreconstructed as *2prong* decay or maybe it is not possible to distinguish the remaining lepton from a prompt lepton. So to improve the  $\gamma$  conversion reconstruction in hadronic  $\tau$  lepton decay an algorithm has been implemented and now the aim of this work is to test the efficiency of this algorithm.

### 3 Data analysis

Before showing results about the photon conversion reconstruction algorithm, it is necessary to define what kind of cuts have been made on data.

Before beginning selecting events it is necessary to apply some cleaning cuts both on reconstructed data and on reconstructed MC data in order to select those  $\tau$ s which are in detector acceptance. These cuts are listed in the Tab. 2: with the first cleaning cut only  $\tau$ s reconstructed by both reconstruction algorithms, both calorimeter based and track based algorithm, are taken into account; the second cut is a geometrical cut and define the geometrical region in which  $\tau$ s can be reconstructed efficiently; the third cleaning cut is a  $p_t$  threshold below which it's not possible to reconstruct  $\tau$ s. After applying cleaning cuts, which are cuts applied on each  $\tau$ , it is possible to begin applying cuts to the whole event. Also these cuts are listed in Tab. 2 and they represent some good conditions for the whole event.

<b>objects cleaning cuts</b>	tau_author=3 eta_cut $ \eta  < 2.5$ each $\tau$ with at least $p_t=15\text{GeV}$
<b>events selection cuts</b> (PreSelection)	L1_J5 level trigger passed Primary vertex with at least 4 tracks JetIsGood selection passed

Table 2: Data cuts

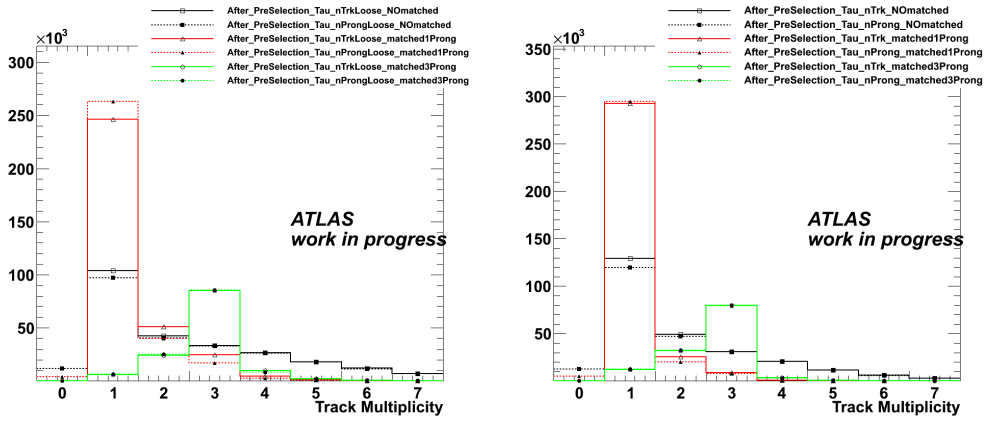


Figure 2:  $\tau$  tracks multiplicity distributions.

## 4 MC data analysis results

First part of this work concerns the study of the MC data, on a Pythia generated MC sample of  $Z \rightarrow \tau\tau$  process, in order to check the  $\gamma$  conversion reconstruction efficiency and to try to reproduce results of previous studies on  $\gamma$  conversion reconstruction algorithm. In order to achieve these goals, first step is compare track multiplicity distributions of  $\tau$  candidates with and without  $\gamma$  conversion veto. Second step is to calculate efficiencies.

### 4.1 Tracks distribution of $\tau$ candidates

First results are shown in Fig. 2. First of all the histogram on the left shows  $\tau$ s reconstructed with calorimeter based algorithm and that one on the right is for  $\tau$ s reconstructed with track based algorithm. Then since this is an analysis made on MC data it is possible to know the true charged track number for each reconstructed  $\tau$ , so it is possible to split up the analysis for reconstructed  $\tau$ s matched to true 1prong  $\tau$ s, for those matched to true 3prong  $\tau$ s

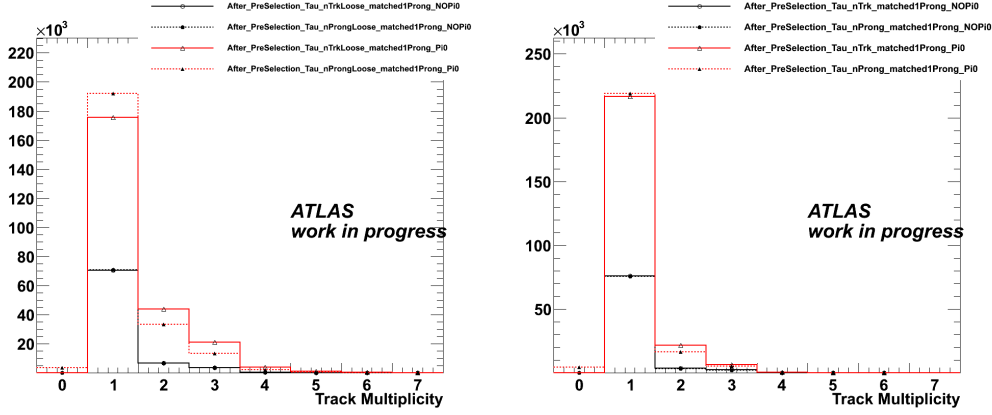


Figure 3: *Tracks multiplicity distributions for reconstructed  $\tau$ s matched to 1prong true  $\tau$ s.*

and also for those which are not matched and this means that they represent QCD background. For each  $\tau$  two type of variable are taken into account:  $nTrack$  and  $nProng$ . The first one represent the number of charged tracks in the reconstructed  $\tau$  cone and the second one represent the remaining number of charged tracks in reconstructed cone after  $\gamma$  conversion tracks are reconstructed explicitly and removed from the number of tracks.

The histograms show that after applying the  $\gamma$  conversion reconstruction algorithm there is an increase of the number of  $\tau$ s with 1 reconstructed charged track matched to true 1prong  $\tau$ s. This means that for  $\tau$ s matched to 1prong true  $\tau$ s the peak in the bin of 1 track is higher for the nProng variable (filled triangle) in comparison with the nTrk variable (open triangle). There is almost the same behaviour also for the  $\tau$ s matched to 3prong true  $\tau$ s and after  $\gamma$  conversion reconstruction algorithm application there is in general a decrease of the background. Finally it is important to observe that the improvement is higher for  $\tau$ s reconstructed with calorimeter based algorithm since they are reconstructed with looser conditions so usually there is an higher number of tracks in these  $\tau$ s and this means that it is possible to reconstruct an higher number of photon conversion.

More detailed results are presented in Fig. 3 and Fig. 4: for both  $\tau$ s matched to true 1prong and true 3 prong  $\tau$ s the analysis is splitted into matched to true  $\tau$ s with or without neutral pions. In Fig. 3 after  $\gamma$  conversion reconstruction there is a more evident increase of the peak in the 1 track bin for the distribution of  $\tau$ s in whose decay some  $\pi^0$ s have been produced since this is the situation in which photons are produced and conversion into  $e^+e^-$  is possible. The same behaviour is in Fig. 4 with the expected differences for

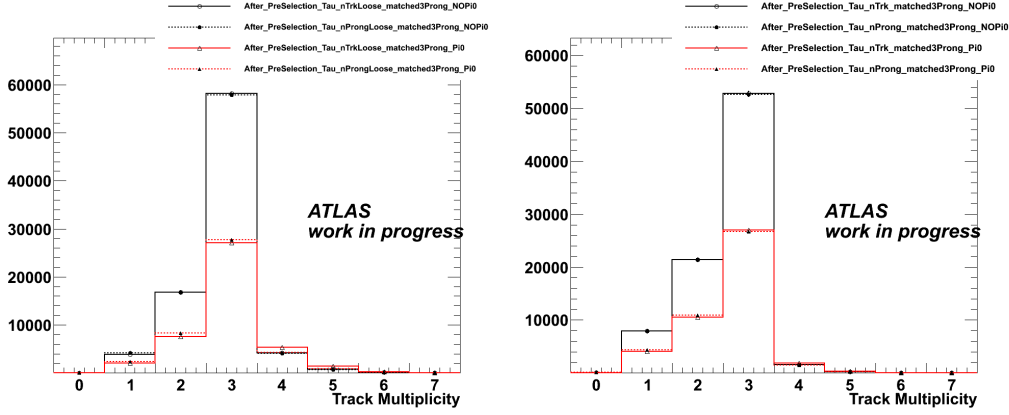


Figure 4: *Tracks multiplicity distributions for reconstructed  $\tau$ s matched to 3prong true  $\tau$ s.*

$\tau$ s matched to 3prong true  $\tau$ s: there is an increase of the peak in the 3 track bin. In this case, unlike the previous case, the peak of  $\tau$ s with  $\pi^0$ s production is lower than peak of  $\tau$ s without  $\pi^0$  production because there are less 3prong decays with  $\pi^0$ s due to branching ratios.

## 4.2 Efficiencies of the $\gamma$ conversion reconstruction

In order to have a quantitative idea of the  $\gamma$  conversion reconstruction algorithm efficiency it is calculated with the following definition:

$$eff = \frac{\text{reconstructed } \tau\text{s matched to true } \tau\text{s}}{\text{true } \tau\text{s}}.$$

The calculated efficiencies are shown in Tab. 3: in most cases efficiencies increase for the nProng variable in comparison with nTrk variable, this means that they increase after applying  $\gamma$  conversion reconstruction algorithm.

One of the goal of this work is to reproduce results from previous studies on this  $\gamma$  photon conversion algorithm and finally, in general, they have been very well reproduced even if there are some difference due to the fact that this study has been performed with a totally new software release in comparison with the previous study. Results in Tab. 3 are checking  $\tau$  candidates only ( $\tau$  candidates are  $\tau$ s after reconstruction step but before identification step). After seen a small improvement for  $\tau$  candidates it is also interesting to see if an improvement after identification can be achieved as well since this is a type of study that hasn't been done before. It is necessary to introduce different  $\tau$  identifications: this means that in selecting  $\tau$ s it is possible to add some cuts based on different sets of variables. These sets are combined as CutSafe and

	<b>nTrack</b>	<b>nProng</b>	<b>nTrackLoose</b>	<b>nProngLoose</b>
<b>1Prong</b>	57.05 $\pm$ 0.14	57.44 $\pm$ 0.14	48.02 $\pm$ 0.14	51.26 $\pm$ 0.14
<b>1Prong NO <math>\Pi^0</math></b>	63.57 $\pm$ 0.29	63.49 $\pm$ 0.29	59.19 $\pm$ 0.29	59.48 $\pm$ 0.29
<b>1Prong <math>\Pi^0</math></b>	55.08 $\pm$ 0.16	55.61 $\pm$ 0.16	44.64 $\pm$ 0.16	48.77 $\pm$ 0.16
<b>3Prong</b>	55.01 $\pm$ 0.27	54.72 $\pm$ 0.27	58.77 $\pm$ 0.26	59.00 $\pm$ 0.26
<b>3Prong NO <math>\Pi^0</math></b>	56.87 $\pm$ 0.33	56.73 $\pm$ 0.33	62.64 $\pm$ 0.33	62.35 $\pm$ 0.33
<b>3Prong <math>\Pi^0</math></b>	51.70 $\pm$ 0.45	51.14 $\pm$ 0.45	51.89 $\pm$ 0.45	53.05 $\pm$ 0.45

Table 3: After PreSelection efficiencies for total  $\tau$ s

CutSafeCalo and optimized for different reconstruction efficiencies (loose, medium, tight):

- CutSafeCaloLoose
- CutSafeCaloMedium
- CutSafeCaloTight
- CutSafeLoose
- CutSafeMedium
- CutSafeTight.

CutSafe sets are based on calorimeter information only and CutSafeLoose sets are based on both calorimeter and track information. It is interesting to calculate efficiencies for different  $\tau$  identifications to check their behaviour in relation to the different cuts. The calculated efficiencies are shown in Tab. 4 and Tab. 5: going from loose to medium to tight cuts efficiencies decrease as expected, but for each different  $\tau$  identification efficiencies have still the same behaviour of increasing after applying  $\gamma$  conversion reconstruction algorithm and this means that the identification cuts and the  $\gamma$  conversion track veto are uncorrelated.

## 5 Real data analysis results

In addition to MC studies, an analysis on data has been also performed on samples of data taken from the end of March up to the mid of July, in different data taking periods (A, B, C, and D) corresponding to a total integrated luminosity of 304.7 nb $^{-1}$ . During studying  $\gamma$  photon conversion



After PreSelection efficiencies for CutSafeCaloLoose  $\tau$ s

	nTrack	nProng	nTrackLoose	nProngLoose
1Prong	49.95 $\pm$ 0.14	50.42 $\pm$ 0.14	42.53 $\pm$ 0.14	45.46 $\pm$ 0.14
1Prong NO $\Pi^0$	52.60 $\pm$ 0.30	52.58 $\pm$ 0.30	49.81 $\pm$ 0.30	50.03 $\pm$ 0.30
1Prong $\Pi^0$	49.16 $\pm$ 0.16	49.78 $\pm$ 0.16	40.34 $\pm$ 0.16	44.09 $\pm$ 0.16
3Prong	53.65 $\pm$ 0.27	53.36 $\pm$ 0.27	56.91 $\pm$ 0.27	57.13 $\pm$ 0.27
3Prong NO $\Pi^0$	55.19 $\pm$ 0.33	55.04 $\pm$ 0.33	60.32 $\pm$ 0.33	60.04 $\pm$ 0.33
3Prong $\Pi^0$	50.92 $\pm$ 0.45	50.36 $\pm$ 0.45	50.85 $\pm$ 0.45	51.96 $\pm$ 0.45

After PreSelection efficiencies for CutSafeCaloMedium  $\tau$ s

	nTrack	nProng	nTrackLoose	nProngLoose
1Prong	33.43 $\pm$ 0.14	34.15 $\pm$ 0.14	29.53 $\pm$ 0.13	31.44 $\pm$ 0.13
1Prong NO $\Pi^0$	40.01 $\pm$ 0.29	40.05 $\pm$ 0.29	38.58 $\pm$ 0.29	38.72 $\pm$ 0.29
1Prong $\Pi^0$	31.43 $\pm$ 0.15	32.36 $\pm$ 0.15	26.79 $\pm$ 0.14	29.24 $\pm$ 0.15
3Prong	42.91 $\pm$ 0.27	42.71 $\pm$ 0.27	45.05 $\pm$ 0.27	45.20 $\pm$ 0.27
3Prong NO $\Pi^0$	43.44 $\pm$ 0.33	43.34 $\pm$ 0.33	46.93 $\pm$ 0.34	46.72 $\pm$ 0.34
3Prong $\Pi^0$	41.98 $\pm$ 0.44	41.59 $\pm$ 0.44	41.71 $\pm$ 0.44	42.51 $\pm$ 0.44

After PreSelection efficiencies for CutSafeCaloTight  $\tau$ s

	nTrack	nProng	nTrackLoose	nProngLoose
1Prong	19.64 $\pm$ 0.11	20.34 $\pm$ 0.12	17.84 $\pm$ 0.11	19.05 $\pm$ 0.11
1Prong NO $\Pi^0$	27.11 $\pm$ 0.26	27.13 $\pm$ 0.26	26.44 $\pm$ 0.26	26.54 $\pm$ 0.26
1Prong $\Pi^0$	17.38 $\pm$ 0.12	18.28 $\pm$ 0.13	15.24 $\pm$ 0.12	16.79 $\pm$ 0.12
3Prong	25.14 $\pm$ 0.23	25.05 $\pm$ 0.23	26.49 $\pm$ 0.24	26.51 $\pm$ 0.24
3Prong NO $\Pi^0$	25.27 $\pm$ 0.29	25.23 $\pm$ 0.29	27.34 $\pm$ 0.30	27.24 $\pm$ 0.30
3Prong $\Pi^0$ $\Pi^0$	24.92 $\pm$ 0.39	24.71 $\pm$ 0.39	24.97 $\pm$ 0.39	25.23 $\pm$ 0.39

Table 4: After PreSelection efficiencies for different  $\tau$  identifications based on calorimeter informations only. Going from the top to the bottom there are results of loose, medium and tight cut.

After PreSelection efficiencies for CutSafeLoose $\tau$ s				
	nTrack	nProng	nTrackLoose	nProngLoose
1Prong	56.19 $\pm$ 0.14	56.48 $\pm$ 0.14	47.37 $\pm$ 0.14	50.50 $\pm$ 0.14
1Prong NO $\Pi^0$	61.46 $\pm$ 0.29	61.39 $\pm$ 0.29	57.48 $\pm$ 0.29	57.74 $\pm$ 0.29
1Prong $\Pi^0$	54.60 $\pm$ 0.16	55.00 $\pm$ 0.16	44.31 $\pm$ 0.16	48.31 $\pm$ 0.16
3Prong	54.04 $\pm$ 0.27	53.75 $\pm$ 0.27	57.59 $\pm$ 0.27	57.82 $\pm$ 0.27
3Prong NO $\Pi^0$	55.86 $\pm$ 0.33	55.72 $\pm$ 0.33	61.37 $\pm$ 0.33	61.09 $\pm$ 0.33
3Prong $\Pi^0$	50.80 $\pm$ 0.45	50.24 $\pm$ 0.45	50.88 $\pm$ 0.45	52.01 $\pm$ 0.45

After PreSelection efficiencies for CutSafeMedium $\tau$ s				
	nTrack	nProng	nTrackLoose	nProngLoose
1Prong	41.28 $\pm$ 0.14	41.31 $\pm$ 0.14	36.02 $\pm$ 0.14	37.73 $\pm$ 0.14
1Prong NO $\Pi^0$	45.40 $\pm$ 0.30	45.38 $\pm$ 0.30	43.41 $\pm$ 0.29	43.55 $\pm$ 0.29
1Prong $\Pi^0$	40.04 $\pm$ 0.16	40.08 $\pm$ 0.16	33.79 $\pm$ 0.15	35.97 $\pm$ 0.16
3Prong	45.73 $\pm$ 0.27	45.54 $\pm$ 0.27	47.76 $\pm$ 0.27	47.88 $\pm$ 0.27
3Prong NO $\Pi^0$	49.84 $\pm$ 0.34	49.73 $\pm$ 0.34	53.45 $\pm$ 0.34	53.24 $\pm$ 0.34
3Prong $\Pi^0$	38.42 $\pm$ 0.44	38.11 $\pm$ 0.44	37.65 $\pm$ 0.43	38.36 $\pm$ 0.44

After PreSelection efficiencies for CutSafeTight $\tau$ s				
	nTrack	nProng	nTrackLoose	nProngLoose
1Prong	24.26 $\pm$ 0.12	24.04 $\pm$ 0.12	21.88 $\pm$ 0.12	22.41 $\pm$ 0.12
1Prong NO $\Pi^0$	31.18 $\pm$ 0.27	31.14 $\pm$ 0.27	30.23 $\pm$ 0.27	30.28 $\pm$ 0.27
1Prong $\Pi^0$	22.17 $\pm$ 0.14	21.89 $\pm$ 0.14	19.35 $\pm$ 0.13	20.02 $\pm$ 0.13
3Prong	30.86 $\pm$ 0.25	30.81 $\pm$ 0.25	31.25 $\pm$ 0.25	31.30 $\pm$ 0.25
3Prong NO $\Pi^0$	39.25 $\pm$ 0.33	39.18 $\pm$ 0.33	40.31 $\pm$ 0.33	40.22 $\pm$ 0.33
3Prong $\Pi^0$	15.96 $\pm$ 0.33	15.94 $\pm$ 0.33	15.16 $\pm$ 0.32	15.47 $\pm$ 0.32

Table 5: After PreSelection efficiencies for different  $\tau$  identifications based on both calorimeter and tracking informations. Going from the top to the bottom there are results of loose, medium and tight cut.

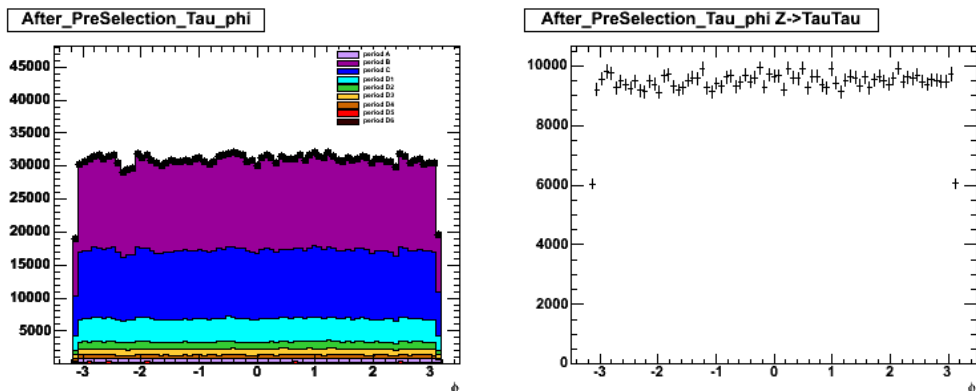


Figure 5:  $\phi$  distribution for reconstructed real  $\tau$ s and reconstructed MC  $\tau$ s.

algorithm efficiencies a lot of histograms have been produced in order to monitor data and to understand the different variables behaviour. In doing this, a  $\phi$  variable distribution not totally flat has been observed for  $\tau$ s, so a more detailed analysis has been implemented on this variable.

## 5.1 Investigating $\phi$ distributions of $\tau$

The  $\phi$  distribution for reconstructed  $\tau$ s is shown in Fig. 5: contrary to the expected totally flat distribution, it is possible to identify some slight dips. The "unflat" shape of the  $\phi$  distribution has been discussed in the ATLAS  $\tau$  working group. In order to understand the distribution shape different distributions for different number of charged tracks in the  $\tau$  cone have been analyzed. Additional distributions are shown in Fig. 6, Fig. 7 and Fig. 8: even if data distributions fit with MC distributions, the interesting shape of these distributions makes it necessary to have a more detailed analysis to try to understand the reason of this strange behavior. To obtain more informations some two dimensional histograms are produced and are shown in Fig. 9. They represent the distribution of  $\tau$ s with 1, 2 or 3 charged tracks on plane  $\eta$  over  $\phi$ : for  $\tau$ s with 1 charged tracks there are some peak regions in the plane and on the contrary for  $\tau$ s with 2 or 3 charged tracks there are some dip regions. These results imply that in the highlighted regions there are some pixel modules not well working, so it is possible to loose some tracks and to have an increase (peaks) of 1 track reconstructed  $\tau$ s and a decrease (dips) of 2 tracks and 3 tracks reconstructed  $\tau$ s. After discussing with  $\tau$  working group it turns out that the not working modules are already simulated and this explains the agreement between data and MC distribution.

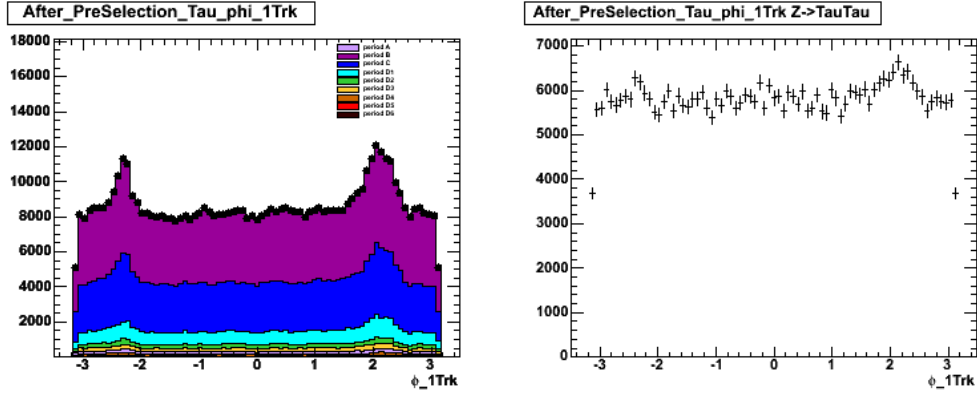


Figure 6:  $\phi$  distribution for reconstructed real  $\tau$ s and reconstructed MC  $\tau$ s with 1 charged track.

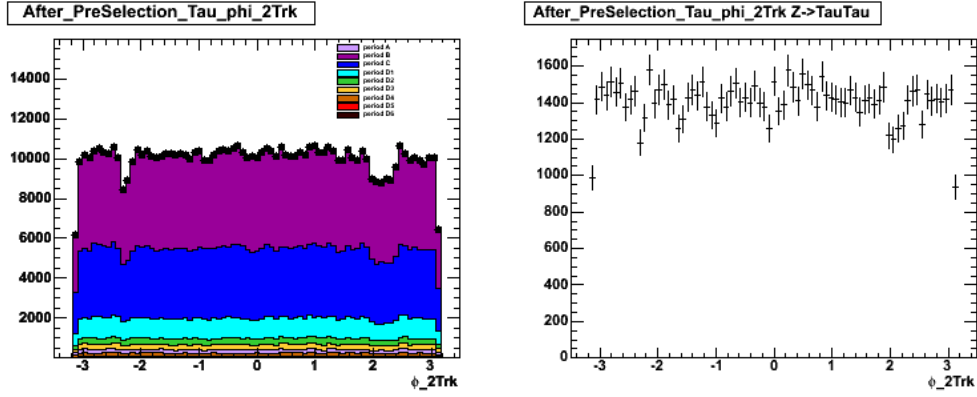


Figure 7:  $\phi$  distribution for reconstructed real  $\tau$ s and reconstructed MC  $\tau$ s with 2 charged tracks.

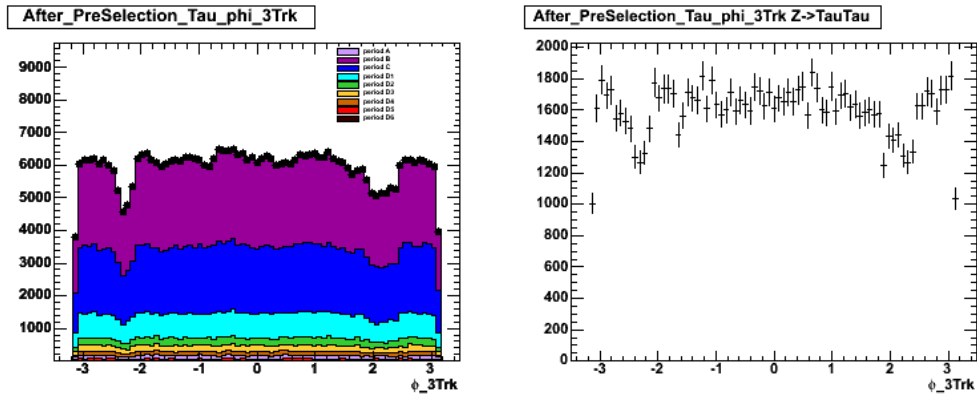


Figure 8:  $\phi$  distribution for reconstructed real  $\tau$ s and reconstructed MC  $\tau$ s with 3 charged tracks.

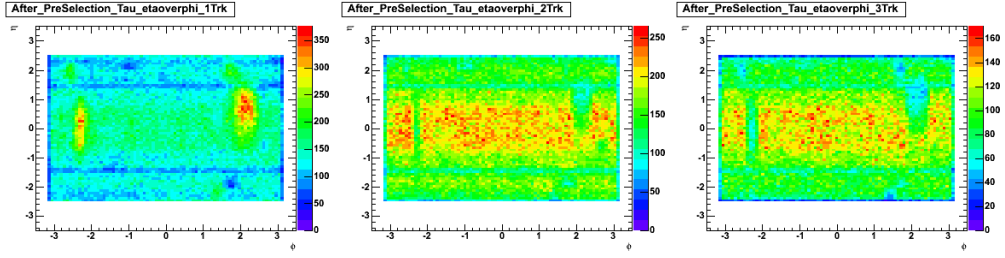


Figure 9: *distribution for reconstructed real  $\tau$ s for different number of charged tracks in a  $\eta$  over  $\phi$  plane.*

## 6 Conclusion

Summarizing the results obtained in MC data analysis, it is important to remember that results from previous studies on the  $\gamma$  conversion reconstruction algorithm efficiencies have been very well reproduced. It was also shown that  $\gamma$  conversion reconstruction algorithm is uncorrelated with different  $\tau$ s identifications. Concerning real data analysis an interesting result was reached by studying  $\phi$  distribution for reconstructed  $\tau$ s: the  $\phi$  could be explained by some pixel modules which are not well working: this leads to a loss of tracks in a specific  $\eta - \phi$  region and to an increase of 1 track reconstructed  $\tau$ s and a decrease of 2 tracks and 3 tracks reconstructed  $\tau$ s.

## 7 Acknowledgements

First of all I would like to thank DESY for arranging the Summer Student Programme and all the people involved in the organization: I had a great experience during the programme. My special thanks go to my supervisor, Michael Boehler, for his continuous support and for always having time and patience for my questions and finally I would like to thank everyone in the ATLAS group for support in every kind of problems.