Reconstruction and selection of $Z \rightarrow \tau \tau \rightarrow \mu + \tau$-jet + $\nu$'s decays at the CMS experiment

Letizia Lusito on behalf of the CMS Collaboration

Abstract

At the LHC, tau leptons are expected in final states of many important physics processes including Supersymmetry and the production of Higgs boson(s) and other exotic particles. An efficient and accurate $\tau$ reconstruction and identification are therefore an important part of the CMS physics programme. $Z^0 \rightarrow \tau^+\tau^-$ decays are often considered the “standard candle” of tau reconstruction as they validate tau lepton identification and provide a test bench for Higgs searches (for which they constitute the main irreducible background).

We describe techniques for selecting and reconstructing the $Z^0 \rightarrow \tau^\pm\tau^{\mp} \rightarrow \mu^\pm\nu\bar{\nu} + \tau$-jet+$\nu$($\bar{\nu}$) events that were developed for the measurement of the Z production cross-section by the CMS experiment using 200 pb$^{-1}$ of the LHC collision data at the center-of-mass energy $\sqrt{s} = 10$ TeV. We validate these techniques using simulated events and present a data-driven method for estimating background contributions to this measurement.

Presented at QCD@Work: International Workshop on QCD - Theory and Experiment, 20-23 June 2010, Martina Franca, Italy
Reconstruction and selection of $Z \rightarrow \tau \tau \rightarrow \mu + \tau$-jet + $\nu$’s decays at the CMS experiment

Letizia Lusito

via Orabona 4, I - 70126 Bari, Italy
CERN, CH-1211 Genève 23, Switzerland

Abstract.

At the LHC, tau leptons are expected in final states of many important physics processes including Supersymmetry and the production of Higgs boson(s) and other exotic particles. An efficient and accurate $\tau$ reconstruction and identification are therefore an important part of the CMS physics programme. $Z^0 \rightarrow \tau^+ \tau^-$ decays are often considered the “standard candle” of tau reconstruction as they validate tau lepton identification and provide a test bench for Higgs searches (for which they constitute the main irreducible background).

We describe techniques for selecting and reconstructing the $Z^0 \rightarrow \tau^+ \tau^- \rightarrow \mu^+ V_\mu (V_\mu \nu) + \tau$-jet $V_\tau (\bar{V}_\tau)$ events that were developed for the measurement of the Z production cross-section by the CMS experiment using 200 pb$^{-1}$ of the LHC collision data at the center-of-mass energy $\sqrt{s} = 10$ TeV. We validate these techniques using simulated events and present a data-driven method for estimating background contributions to this measurement.

Keywords: CMS, tau lepton decays
PACS: 13.35.Dx, 13.38.Dg, 14.60.Fg

THE COMPACT MUON SOLENOID EXPERIMENT

The Compact Muon Solenoid (CMS) [1] is a general purpose detector designed to record p-p collisions produced at the Large Hadron Collide. The CMS detector consists of the silicon-based inner tracking system, the scintillating-crystals-based electromagnetic calorimeter and the hadron calorimeter placed inside a high-field solenoid coil, which provides a magnetic field of 4 T (3.8 T at the start-up). The muon system (composed of Resistive Plate Chambers, Cathode Strip Chambers and Drift Tubes detectors) is placed in the return yoke. Besides the subdetector components, the trigger and data acquisition systems play a critical role in selecting and recording the data due to the very high interaction rates of p-p collisions.

Z BOSON PRODUCTION AND $\tau$ LEPTON PROPERTIES

In the first period of data-taking, events with the Z bosons decaying to leptons will serve as an important tool for detector calibration and alignment, tuning of the generators

---

1 On behalf of the CMS Collaboration
used for simulating data and for the Standard Model “re-discovery” at new center-of-
mass energy. At the leading order, Z bosons are produced via the process $q\bar{q} \rightarrow Z$; the
contribution from other processes, such as $qg$ or $gg$ scattering is subdominant.

The $\tau$ lepton is a point-like, spin 1/2 Dirac particle with a mass of 1.777 GeV/$c^2$
and a lifetime of $2.9 \times 10^{-17}$ s. Leptonic decays (containing an electron or a muon and
associated neutrinos) constitute 35% of the cases. The remaining 65% of the time $\tau$
leptons decay to hadrons, predominantly $\pi^\pm$’s and $\pi^0$’s, forming what is often referred
to as a $\tau$-jet. 77% of the time hadronic tau decays contain one charged particle (the
so-called 1 prong decays) and the majority of remaining the tau decays contain three
charged particles (3 prongs).

MUON AND TAU LEPTON RECONSTRUCTION

**Muons.** Muons are reconstructed by combining the information from the muon
detectors and the tracker (global muon reconstruction). To improve the rejection of pions
misreconstructed as muons, muon candidate selection relies on the calculation of the
probability that signals recorded in calorimeters and muon chambers are consistent with
being originated from a true muon (the so-called compatibility probabilities). Moreover,
uuons from Z decays can be discriminated from muons produced in soft jets from kaons
and pions using isolation defined as the sum of energy deposits and tracks transverse
momenta in cones defined around the muon track direction at vertex. Muon contribution
to this sum is excluded to improve discrimination against background events [2].

**Tau-jets.** Tau-jets are reconstructed using the Particle Flow (PF) algorithm designed
to provide a complete event description by reconstructing and identifying all stable par-
ticles in the event [3]. PF-jets are reconstructed by applying a standard jet clusterization
algorithm such as the Iterative Cone with a cone size of 0.5 in the ($\eta$, $\phi$) space to the
list of particles in the event. For tau-jet selection, PF-jets are required to have at least
one charged hadron with $p_T > 5$ GeV/$c$. Around the PF-jet direction, a matching cone is
built in ($\eta$, $\phi$) space (typical size is 0.1). The highest-momentum track reconstructed in
the matching cone becomes the leading track of the tau-jet. Around the leading track di-
rection, a signal cone (which is expected to contain all the tau decay products due to the
narrowness of the $\tau$-jet) and a larger isolation cone (in which little activity is expected if
the tau-jet is isolated) are built. A $\tau$-jet is considered isolated if no charged hadrons or
photons are reconstructed in the isolation annulus, where the isolation annulus is defined
as the space between the signal and the isolation cones. The ”shrinking” signal cone def-
inition (where the signal cone size varies as $S/E_T$, with $E_T$ being the transverse energy of
the tau-jet) improves the tau reconstruction performance since it allows a better recov-
ery of three prongs decays. Next, discriminators against muons or electrons are applied
to avoid misidentifications of isolated electrons and muons as taus, thus suppressing
$Z \rightarrow e e$ and $Z \rightarrow \mu \mu$ background contributions.
Data samples. Simulated $Z \rightarrow \tau \tau$ events are produced using Pythia 6 [4] Monte Carlo generator complemented with the TAUOLA [5] package designed for high accuracy tau decay generation. Detector response is emulated using standard CMS detector simulation. To account for background contributions, two categories of processes are considered: electroweak-like backgrounds ($Z \rightarrow \mu \mu$, $Z \rightarrow e e + \text{jets}$, $W + \text{jets}$ and $t \bar{t} + \text{jets}$) with final isolated leptons similarly to the signal, and the QCD multijet background, i.e. the process $pp \rightarrow \mu + X$, where at least one muon from heavy flavor decays or decays in flight is produced in the final state. These background processes would completely dominate the signal if not properly rejected. In order to increase the equivalent integrated luminosity of the QCD sample, given finite CPU resources, a set of preselection at the MC-truth level are applied at the generation time. Table 1 describes the MC samples used in this study including the product of cross-section and the preselection filter efficiency, the number of analyzed events and the equivalent integrated luminosity. Pythia and MadGraph event generators are used for data simulation [6].

Event selection. Events are first preselected using the following criteria:

- at least one global muon candidate with $p_T > 8 \text{ GeV/c}$ and $|\eta| < 2.5$
- at least one ParticleFlow $\tau$-jet (PFTau) candidate with a leading pion (charged or neutral) having $p_T > 5 \text{ GeV/c}$ OR at least one $\tau$-jet candidate reconstructed using only calorimeter informations (CaloTau) with a leading track having $p_T > 5 \text{ GeV/c}$.
- at least one pair muon+tau-jet (where the tau-jet can be “PF” or “Calo”) with the two lepton candidates separated by a distance $\Delta R = \sqrt{\Delta \phi^2 + \Delta \eta^2} > 0.7$
The trigger selection is based on the two trigger paths “HLT_IsoMu11” and “HLT_Mu15”, which require respectively an isolated muon with \( p_T > 11 \text{ GeV/c} \) and a not necessarily isolated muon with \( p_T > 15 \text{ GeV/c} \). Events are then selected by requiring the presence of a well reconstructed vertex and lepton kinematics cuts: the global muon is required to have \( p_T > 15 \text{ GeV/c} \) and \( |\eta| < 2.1 \) while the tau-jet is required to have \( p_T > 20 \text{ GeV/c} \), \( |\eta| < 2.1 \) and to not overlap with the muon. Muon isolation is applied by selecting only muons with an absolute isolation variable value (sum of transverse momenta in tracker and energy deposits in ECAL and HCAL associated to particles reconstructed in a cone of size \( \Delta R = 0.6 \) around the muon direction) < 1 GeV. Background events with charged pions misidentified as muons are suppressed using the compatibility probabilities for segments in the muon chambers and energy deposited in calorimeters. Events are further selected by applying a cut of 500 \( \mu \text{m} \) on muon transverse impact parameter \( d_0 \). This cut is effective in removing backgrounds in which muons originate from decays of long-lived mesons. The \( \tau \)-jet candidates are further required to have only one or three charged tracks reconstructed inside the signal cone with a total charge of \( \pm 1 \). Then a discriminator against muons misidentified as 1-prong tau jets is applied to reject \( Z \rightarrow \mu \mu \) events. Events are then required to contain at least one lepton pair where \( \Delta R(\mu, \tau\text{-jet}) > 0.7 \). To reject \( W(\rightarrow \mu\nu) + \text{jets} \) events, only events for which the transverse mass between the muon and the \( E_T \) is < 50 GeV and \( p_\zeta - 1.5 p_\zeta^{\text{vis},\tau} > -20 \) [7] are finally considered. The variables \( p_\zeta \) and \( p_\zeta^{\text{vis},\tau} \) are defined to be the projections of \( \vec{p}_Z^\zeta \) and \( \vec{p}_T^{\text{vis},\tau} \) onto the axis \( \vec{\zeta} \), defined along the bisection direction of the angle formed by the \( \tau \) visible decay products:

\[
\begin{align*}
  p_\zeta^Z &= \vec{p}_Z^\zeta \cdot \vec{\zeta} \\
  p_\zeta^{\text{vis},\tau} &= \vec{p}_T^{\text{vis},\tau} \cdot \vec{\zeta}
\end{align*}
\]

where \( \vec{p}_Z^\zeta \) is the transverse momentum of \( Z \) decaying to tau pairs and \( \vec{p}_T^{\text{vis},\tau} \) is the sum of the transverse momenta of the visible products of the two tau decays.

**SIGNAL EXTRACTION**

For \( \sqrt{s} = 10 \text{ GeV} \), with a dataset corresponding to \( \mathcal{L} = 200 \text{ pb}^{-1} \), the selections described above yield 1330\( \pm 36 \) (statistical error calculated considering a binomial statistics) \( Z \rightarrow \tau\tau \rightarrow \mu + \tau\text{-jet} + \nu \)'s events with the expected background of 374 events. The corresponding significance parameter \( S/\sqrt{(S+B)} \) is 32. The background is dominated by \( W + \text{jets} \) (275 events). Figure 1 shows the visible mass distribution \( M_{\mu,\tau}^{\text{vis}} \) of the selected signal and background events. During the first period of data-taking, the visible mass distribution is used instead of the full mass distribution to reduce the dependence of the analysis results on the performance of the \( E_T^{\text{miss}} \) measurement, which may require more time to fully understand and calibrate [6].
**THE DATA-DRIVEN BACKGROUND ESTIMATION**

The template fit method allows to estimate contributions of background processes to $Z \rightarrow \tau \tau \rightarrow \mu + \tau$-jet + $\nu$’s in a data-driven way.

To determine the contributions of the signal process and of all background processes to the event sample selected in the final analysis, a set of template histograms is used to fit the selected data sample. The visible mass distribution (invariant mass between muon and tau decay products) $M_{\text{vis}}^{\mu+\tau-\text{jet}}$ is used as the template histogram because its shape is uncorrelated with the selection criteria used in the determination of phase space regions.

The shapes of these template histograms represent the distributions of events expected for the signal process under study and for individual background processes and are determined from control samples selected in data. The templates are normalized to unit area so that the fit relies only on the shape of these distributions, while normalization factors are determined by the fit. The normalization factors obtained for the signal and for the individual background processes then provide an estimate of the signal and background contributions to the selected data sample.

The shape templates for individual background processes are obtained from the dedicated control samples. The control samples are selected in a region of phase-space similar to that of the $Z \rightarrow \tau \tau$ candidate events, but with one of the event selection criteria inverted in each sample in order to enhance the contribution of a single background process and suppress as much as possible the contributions from all other background processes with a good compromise between purity and statistics.

The shape template for the $Z \rightarrow \tau \tau$ signal is obtained from a data sample of $Z \rightarrow \mu \mu$ events, selected by applying an ad-hoc event selection criteria for $\mu$’s pair final states, substituting reconstructed muons with simulated tau decay products.

The biases in the template shapes due to the contamination of control samples by other background processes and the signal itself are negligible.

The distributions of the template histograms scaled by the normalization factors obtained in the fit (colored lines, where the black line represents their sum) and of the
visible mass $M_{\mu+\tau-jet}^{{\text{vis}}}$ (invariant mass of the muon and tau visible decay products) of the $\mu+\tau$-jet events selected with the analysis selection (points) are reported in Figure 2.

![Distribution of the visible mass of the muon plus the tau jet, $M_{\mu+\tau-jet}^{{\text{vis}}}$, in $Z\rightarrow\mu+\tau$-jet + $\nu$'s candidate events (points) compared to signal and background templates scaled by the normalization factors determined in the fit (colored lines).](image)

**FIGURE 2.** Distribution of the visible mass of the muon plus the $\tau$-jet, $M_{\mu+\tau-jet}^{{\text{vis}}}$, in $Z\rightarrow\mu+\tau$-jet + $\nu$'s candidate events (points) compared to signal and background templates scaled by the normalization factors determined in the fit (colored lines).

We conservately estimate the accuracy of the present method in determining the normalization of the signal be of the order of $O(15\%)$.

**CONCLUSIONS**

The techniques planned to be used in the first measurement of $Z$ cross section in the di-tau channel are presented and their sensitivity is evaluated in simulated data. A good efficiency and good background rejection (significance $S/\sqrt{(S+B)} = 32$) are obtained. Results and techniques presented here can serve as a basis for other analyses with tau in final states and currently are being validated and optimized in real p-p collisions. The data-driven estimation of background via the template fitting method proposed in this analysis shows good agreement with the expectation and will be utilized in this and other CMS measurements.

**REFERENCES**

1. CMS Collaboration, “The CMS experiment at the CERN LHC”, Journal of Instrumentation 3, 08, S08004 (2008).
2. CMS Collaboration, “Towards a measurement of $W$ and $Z$ cross sections into muons in pp collisions at $\sqrt{s}=10$ TeV”, CMS-PAS-EWK-09-001 (2009).
3. CMS Collaboration, “Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET”, CMS-PAS-PFT-09-001 (2009).
4. T. Sjostrand, S. Mrenna, P. Skands, “PYTHIA 6.4 Physics and Manual”, arXiv:hep-ph/0603175 (2006).
5. Z. Was et al., “TAUOLA the library for $\tau$ lepton decay”, Nucl. Phys. Proc. Suppl. 98, 96 (2001).
6. L. Lusito, Reconstruction and selection of $Z\rightarrow\tau\tau\rightarrow\mu+\tau$-jet + $\nu$'s decays at the CMS experiment, CMS TS 2010/006, (2010).
7. The CDF Collaboration, “Search for neutral MSSM Higgs bosons decaying to tau pairs in p-pbar collisions at $\sqrt{s}=1.96$ TeV”, Phys. Rev. Lett. 96, 011802 (2006).