CMS Internal Note

The content of this note is intended for CMS internal use and distribution only

December 21, 2007

Triggers for top physics

D. Author, E. Author,

Institute, Country

F. Author, G. Author[b)],

Institute, Country

K. Author, M. Author,

Institute, Country

L. Author, S. Author,

Institute, Country

Abstract

This note reviews the status of trigger studies, being performed in the Top Physics Analysis Group, in preparation for data analysis. Four main topics are investigated. First, as a reference, we report the expected trigger efficiencies for top events, depending on the trigger path and the selection channel, for a number of trigger paths, as defined in the HLT table for 10E32 luminosity. We also consider lower luminosity scenarios. Second, we propose new dedicated trigger paths for on-line selection of top-enriched events samples. Third, we propose a method for monitoring the trigger response in the kinematic regions of interest for top event selection. Finally, we address the question of the data-driven determination of trigger efficiency for the top signal and the background. We estimate the uncertainties on the trigger efficiency determination as a function of the sizes of the data samples used for the measurements.

DRAFT
Please contact Marta and Javier if you wish to contribute.

Time scale for completing this note:
Dec.22, 2007: Collected results
Jan.15, 2007: Finalized note

[b)] Also at

1 Introduction

Resp.; Marta

Discussion of trigger efficiency definitions: absolute efficiency, based on MC generator information, and "relative efficiency", relative to the off-line selection. The first definition is useful, for instance, to compare the efficiencies of a specific trigger on different signals, defined according to the MC information. The second definition is more useful for designing optimal triggers in the phase-space region of interest for a the signal. This second definition has also the advantage of providing a quantity which is measurable in the data, as the signal is defined on the basis of measured quantities, and not using MC generator information.

Explain why we need dedicated triggers.

Explain why and how to monitor triggers of interest on data samples selected by a fast (quasi-on line) top analysis: useful to give feedback to the trigger people on the trigger response to "top-like" selected events.

Illustrate with quantitative examples the methods envisaged to measure trigger efficiencies for top signal and background events.

Table 1: Trigger paths of interest for top events, with their pt or Et cuts, the $|\eta|$ coverage, the expected background rate and expected top event rate for a luminosity $10^{32}cm^{-2}s^{-1}$. These are HLT paths in CMSSW $131HLT6$ (also kept in further versions).

Trigger path name	pt or Et cut (GeV)	$\|\eta\|$ coverage	Rate	Top rate
HLT1MuonNonIso	16	2.5	22.7 ± 1.5	xxx
HLT1MuonIso	11	2.5	18.3 ± 2.2	xxx
HLT1ElectronRelaxed	18	2.5	9.6 ± 1.3	xxx
HLT1Electron	15	2.5	17.1 ± 2.3	xxx
HLT2MuonNonIso	3,3	2.5	12.3 ± 1.6	xxx
HLT2ElectronRelaxed	12,12	2.5	0.8 ± 0.1	xxx
HLTXElectronMuonRelaxed	10,10	2.5	0.1 ± 0.0	xxx
HLT1jet	200	5.0	9.3 ± 0.1	xxx

We define two trigger efficiencies $\epsilon_{MC}(trig)$ and $\epsilon_{RE}(trig)$.

$\epsilon_{MC,RE}(trig) = (Number\ of\ events\ preselected\ and\ passing\ the\ trigger)/(number\ of\ events\ preselected)$

In $\epsilon_{MC}(trig)$ events are preselected applying cut on the generated quantities, while in $\epsilon_{RE}(trig)$ cuts are applied to off-line reconstructed quantities..

The preselection used to calculate $\epsilon_{RE}(trig)$ requires:

- In **single lepton final states**, in ttbar events the $W \rightarrow \ell\nu$ decay; in off-line reconstructed events, the pT or ET of the lepton above 20 GeV with $|\eta_\ell| < 2.0$, no lepton isolation, plus two jets with $ET >$ 20 GeV and $|\eta_j| < 4$
- In **di-lepton final states**, in ttbar events both the W's decaying leptonically; in the reconstructed events, two leptons with opposite charge, one with pT or ET above 20 GeV and the other with pT or ET above 15 GeV, with $|\eta_\ell| < 2.0$, no isolation applied, and two jets with $ET >$ 20 GeV.and $|\eta_j| < 4$.

For the calculation of $\epsilon_{MC}(trig)$ the same cuts are applied on the generated quantities. The above cuts are chosen to ensure full trigger efficiency (plateau of the trigger efficiency turn-on curve in pT and eta) for the preselected events.

The two trigger efficiencies are used for monitoring the trigger performance, with respect to changes occurring in the trigger. Note that $\epsilon_{RE}(trig)$, a part for the requirement of the W decay, is a measurable quantity which can be used to monitor the trigger performance in data, as we will discuss in Section 4.

2 Expected trigger efficiencies for top events

Resp.: Javier

We report the Trigger efficiency matrix for a selection of trigger paths, listed in table, and report their efficiencies for top events selected according to the selection cuts listed in Table. The events are selected in the three lepton(electron, muon, or tau)+jets channels, the six dilepton channels and the fully hadronic channel (NdR this has to be changed depending on which final states we actually study).

Trigger efficiencies are calculated on the Spring07 samples with CMSSW-131HLT6, and using TopRex and Madgraph ttbar events. Efficiencies with Alpgen (and TopRex?) are reported for the CSA07 samples.

2.1 Electron and muon final states (W to electron or muon)

Resp. Roberto+Stephane

After MC-based and RECO-based preselection: give numbers of top events and background(W/Z+jets and QCD), for single electron, single muon and dilepton final states.

Show the pt and eta distributions for the leptons and the jets in the ttbar events events, superimposed to the background. For leptons show also the isolation distribution. These distribution are useful to understand the effect of the triggers.

Report the trigger efficiency tables.

In the following we report the trigger efficiencies $\epsilon_{MC}(trig)$ and $\epsilon_{RE}(trig)$ for the HLT paths listed in Table 1 and for the selections described in Section .

Table 2: Spring07 samples. Electron and muon final states. Trigger efficiency $\epsilon_{MC}(trig)$ for top events selected according to the five final state selections and for the HLT trigger paths of interested listed in the first column.

Final state	$\mu\nu bjjb$	$e\nu bjjb$	$\mu\nu be\mu\nu b$	$e\nu be\nu b$	$e\nu b\mu\nu b$
Trigger path name					
HLT1MuonNonIso					
HLT1MuonIso					
HLT1ElectronRelaxed					
HLT1Electron					
HLT2MuonNonIso					
HLT2ElectronRelaxed					
HLTXElectronMuonRelaxed					
HLT1jet					

Table 3: Spring07 samples. Electron and muon final states. Trigger efficiency $\epsilon_{RE}(trig)$ for top events selected according to the five final state selections and for the HLT trigger paths of interested listed in the first column.

Final state	$\mu\nu bjjb$	$e\nu bjjb$	$\mu\nu be\mu\nu b$	$e\nu be\nu b$	$e\nu b\mu\nu b$
Trigger path name					
HLT1MuonNonIso					
HLT1MuonIso					
HLT1ElectronRelaxed					
HLT1Electron					
HLT2MuonNonIso					
HLT2ElectronRelaxed					
HLTXElectronMuonRelaxed					
HLT1jet					

Mention anything else which you deem important.

2.2 Tau into electron and muon final states (W-to-tau-to-electron-or-muon)

Resp.: Michele, Michal

The "tau di-lepton" channel is the di-lepton $t\bar{t}$ decay with one electron or muon and one hadronically decaying tau lepton in the final state $\mathrm{t\bar{t}} \rightarrow (\ell\nu)(\tau\nu_\tau)\mathrm{b\bar{b}}$, $(\ell = \mathrm{e}, \mu, \tau \rightarrow \mathrm{e}, \mu)$.

The trigger paths used in the "tau di-lepton" analysis are the inclusive electron and muon triggers described in [2] and listed in Tab. 1. The trigger efficiency is calculated from a sample of Alpgen Monte Carlo $t\bar{t}$ events. The individual trigger paths (*HLT1MuonNonIso*, *HLT1MuonIso*, *HLT1ElectronRelaxed*, and *HLT1Electron*) and the logical OR of the trigger paths are considered.

Trigger efficiency, $\epsilon_0(trig)$, is defined as the ratio of the number of events after the HLT trigger paths and normalized to the expected event yield including the branching ratio of the given decay channel, $\sigma_{t\bar{t}}\times$BR (Tab. 4).

The "relative" trigger efficiency, $\epsilon_{RE}(trig) = N_{trg}/N_{pre}$, is calculated with respect to the number of preselected events for the different decay channels. Events are preselected with at least one high-p_T lepton (electron or muon, with $p_T > 20$ GeV and $|\eta| < 2.4$) (Tab. 5). The efficiency of the preselection $\epsilon(pre)$ is a ratio of the number of events passing preselection requirments to the expected event yield including the branching ratio of the given decay channel, $\sigma_{t\bar{t}}\times$BR (Tab. 6)

Table 4: Spring07 samples. Tau final states in $t\bar{t}$ decays with $\tau \rightarrow$ had.: $\mathrm{t\bar{t}} \rightarrow (\ell\nu)(\tau\nu_\tau)\mathrm{b\bar{b}}$, ($\ell = \mathrm{e}, \mu, \tau \rightarrow \mathrm{e}, \mu$). Trigger efficiency $\epsilon_0(trig)$ (in %) for signal events ($\ell\tau$) and for the other $t\bar{t}$ decay channels (di-lepton, single-lepton and all-hadronic), for the HLT trigger paths listed in the first column.

	signal		other $t\bar{t}$			
Trigger path name	$e\tau$	$\mu\tau$	$\ell\ell$	$e(q\bar{q})$	$\mu(q\bar{q})$	full-had.
HLT1MuonNonIso	3.6 ± 1.0	64.5 ± 4.3	56.8 ± 2.1	3.7 ± 0.3	64.4 ± 1.4	3.4 ± 0.2
HLT1MuonIso	0.0 ± 0.0	57.9 ± 4.1	51.6 ± 2.0	0.1 ± 0.1	56.8 ± 1.3	0.1 ± 0.0
HLT1ElectronRelaxed	48.5 ± 3.8	0.0 ± 0.0	40.1 ± 1.8	46.0 ± 1.2	0.3 ± 0.1	0.4 ± 0.1
HLT1Electron	47.6 ± 3.8	0.6 ± 0.4	38.3 ± 1.7	42.9 ± 1.2	0.4 ± 0.1	0.4 ± 0.1
All HLT1Lepton	54.5 ± 4.0	68.5 ± 4.4	85.2 ± 2.6	50.3 ± 1.3	68.0 ± 1.4	4.0 ± 0.2

Table 5: Spring07 samples. Tau final states in $t\bar{t}$ decays with $\tau \rightarrow$ had.: $\mathrm{t\bar{t}} \rightarrow (\ell\nu)(\tau\nu_\tau)\mathrm{b\bar{b}}$, ($\ell = \mathrm{e}, \mu, \tau \rightarrow \mathrm{e}, \mu$). Trigger efficiency $\epsilon_{RE}(trig)$ (in %) for signal events ($\ell\tau$) and for the other $t\bar{t}$ decay channels (di-lepton, single-lepton and all-hadronic), for the HLT trigger paths listed in the first column.

	signal		other $t\bar{t}$			
Trigger path name	$e\tau$	$\mu\tau$	$\ell\ell$	$e(q\bar{q})$	$\mu(q\bar{q})$	full-had.
HLT1MuonIso	0.0 ± 0.0	75.2 ± 5.4	56.1 ± 2.1	0.3 ± 0.1	77.6 ± 1.7	4.0 ± 2.3
HLT1MuonNonIso	4.7 ± 1.6	85.4 ± 5.8	62.4 ± 2.2	2.5 ± 0.4	87.7 ± 1.9	21.3 ± 5.3
HLT1Electron	67.0 ± 5.9	0.4 ± 0.4	46.4 ± 1.9	71.1 ± 2.0	0.5 ± 0.1	12.0 ± 4.0
HLT1ElectronRelaxed	74.9 ± 6.3	0.8 ± 0.6	49.4 ± 2.0	78.0 ± 2.1	0.5 ± 0.1	14.7 ± 4.4
All HLT1Lepton	75.9 ± 6.3	85.4 ± 5.8	92.8 ± 2.7	79.1 ± 2.1	88.1 ± 1.9	36.0 ± 6.9

Table 6: Spring07 samples. Tau final states in $t\bar{t}$ decays with $\tau \rightarrow$ had.: $\mathrm{t\bar{t}} \rightarrow (\ell\nu)(\tau\nu_\tau)\mathrm{b\bar{b}}$, ($\ell = \mathrm{e}, \mu, \tau \rightarrow \mathrm{e}, \mu$). Efficiency of preselection $\epsilon(pre)$ (in %) for signal events ($\ell\tau$) and for the other $t\bar{t}$ decay channels (di-lepton, single-lepton and all-hadronic). Number of events for $\mathcal{L} = 100\mathrm{pb}^{-1}$.

	signal		other $t\bar{t}$			
	$e\tau$	$\mu\tau$	$\ell\ell$	$e(q\bar{q})$	$\mu(q\bar{q})$	full-had.
# Expected	1872.2 ± 23.9	1872.2 ± 23.9	6903.6 ± 88.1	17551.4 ± 223.9	17551.4 ± 223.9	54649.3 ± 697.0
# Preselected	820.1 ± 59.3	1090.6 ± 68.4	5453.0 ± 153.0	7737.2 ± 182.3	10970.4 ± 217.0	322.0 ± 37.2
Efficiency (%)	43.8 ± 3.2	58.3 ± 3.7	79.0 ± 2.4	44.1 ± 1.2	62.5 ± 1.5	0.6 ± 0.1

2.3 First results with the CSA07 samples (CMSSW-160

Resp: Javier

Here there should be trigger efficiencies calculated as usual but for the CSA07 samples and CMSSW-160.

In the following we report the trigger efficiencies $\epsilon_{MC}(trig)$ and $\epsilon_{RE}(trig)$ for the HLT paths listed in Table 1 and for the selections described in Section .

Table 7: CSA07 samples. Electron and muon final states. Trigger efficiency $\epsilon_{MC}(trig)$ for top events selected according to the five final state selections and for the HLT trigger paths of interested listed in the first column.

Final state	$\mu\nu bjjb$	$e\nu bjjb$	$\mu\nu be\mu\nu b$	$e\nu be\nu b$	$e\nu b\mu\nu b$
Trigger path name					
HLT1MuonNonIso					
HLT1MuonIso					
HLT1ElectronRelaxed					
HLT1Electron					
HLT2MuonNonIso					
HLT2ElectronRelaxed					
HLTXElectronMuonRelaxed					
HLT1jet					

Table 8: CSA07 samples. Electron and muon final states. Trigger efficiency $\epsilon_{RE}(trig)$ for top events selected according to the five final state selections and for the HLT trigger paths of interested listed in the first column.

Final state	$\mu\nu bjjb$	$e\nu bjjb$	$\mu\nu be\mu\nu b$	$e\nu be\nu b$	$e\nu b\mu\nu b$
Trigger path name					
HLT1MuonNonIso					
HLT1MuonIso					
HLT1ElectronRelaxed					
HLT1Electron					
HLT2MuonNonIso					
HLT2ElectronRelaxed					
HLTXElectronMuonRelaxed					
HLT1jet					

3 New dedicated triggers for on-line top-like event selection

3.1 Trigger with muon+Njets

Resp.: Silvia

Show ttbar rates in a lepton pt vs jet Et (same Et cut for all jets) plane. Lepton and jet variable can be the offline reconstructed ones, or those measured by the HLT, if one knows how to use the HLT software and analyzer. Discuss triggers with an isolated and relaxed muon.

3.2 Trigger with electron+Njets

Resp.: Marta Show lepton pt vs jet Et (same Et cut for all jets). Discuss isolated and relaxed lepton.

3.3 Trigger with lepton(electron/muon)+b-jet+Njets

Resp.: Muriel(?) Show lepton pt vs jet Et (same Et cut for all jets). Study b-jet related variables for the optimization?

3.4 Triggers with two leptons + N jets

Resp.: Javier, Oviedo people Show lepton pt (same pt cut for both leptons) vs jet Et (same Et cut for all jets) Lepton and jet variable can be the offline reconstructed ones, or those measured by the HLT, if one knows how to use the HLT software and analyzer.

4 Monitoring triggers for top events

Resp.; Marta, Greg (L1 monitoring), Javier (HLT monitoring)

In this Section we address the issue of monitoring the performance of the trigger system for top-like events of interest for top physics. The reason why we need to select top-like events and monitor the trigger performance in this selected sample are given below. Trigger performance will be monitored by many groups. The on-line group will monitor the total data stream. PO groups will monitor inclusive lepton and jet samples. PA groups should analyze more exclusive samples, with events having characteristics close to the signal of interest. In the case of the top PAG, we should monitor the trigger performance for top-like events. By top-like events we mean events with similar characteristics as for top events, namely with at least a lepton of relatively high transverse momentum and at least two jets. One of the reasons why we need to monitor performance within an exclusive top-like sample is the following. Let us imagine that inclusive single lepton and jet rates are as expected and that the POG monitoring analyses do not detect any anomaly. It may happen that the rate of top-like lepton+jets events is eg a half of what expected from MC predictions. The reason could be that the isolation of the lepton in top-like events (mainly W+jets and genuine top events) is worse than the one predicted by the MC calculation, because eg more energy is deposited around the lepton than the MC calculation predicts. This effect may be detectable only on an exclusive lepton+jet event sample, because eg the calculation may underestimate the number of tracks or their momenta in the production of this type of events.. For this and other reasons we need to define a top-like sample which we use to monitor the trigger performance. Depending on the luminosity, we should define the selection so that we have a relatively good purity of genuine top-like events, while having a reasonable statistics to make our checks. We have to decide how often we can monitor (make plots). We could decide to make plots at the end of each fill. If the luminosity is reasonable the collected statistics of top-like events could be sufficient to study the effects we mentioned above.

For this reason we have proposed in Section 2, new exclusive triggers for top selection. These triggers are optimized to select on-line top-enriched samples in the different channels investigated (lepton+jets, dilepton+jets, Multijets) For each of these top-exclusive triggers, we propose to monitor the absolute rate and the overlap with inclusive triggers (single lepton, di-jet, dilepton). We also propose to monitor the differential rates as a function of L1 and HLT reconstructed quantities (pT/ET, eta, phi of the lepton(s) and the jets) The trigger-turn on curves for these exclusive triggers, as a function of the off-line reconstructed quantities (ndr: define trigger efficiency), should also be measured (offline). In Fig. we show the expected rate distributions for the SM soup (what should be seen in the data) and for the top signal alone.(should see that the top signal is significant in these distributions).

5 Data-driven trigger effi ciency determination for signal and background

Resp.: Marta, who else interested? The importance of the trigger efficiency measurements is easily seen in the cross-section measurement. In the simplest approach, the number of off-line selected events is:

$$N_{sel} = N_{sig} - N_{bkg}; N_{sig} = L(\sigma_{sig}\epsilon^{trg}_{sig}\epsilon^{sel}_{sig}) = N_{sel} - N_{bkg}$$

Let us assume that we have independently measured the number of background events N_{bkg}. Then the measurement of the cross-section σ_{sig}, implies the measurement of the integrated luminosity L, of the trigger efficiency ϵ^{trg}_{sig} and of the off-line selection efficiency ϵ^{sel}_{sig} for signal events. The trigger efficiency ϵ^{trg}_{sig} is the signal efficiency for a specific trigger. It assumes that the off-line selected events have all fired that specific trigger. For the sake of simplicity, let us discuss the case of single-object triggers. The trigger efficiency should be measured as a function of the off-line reconstructed pT, eta and phi of the triggering object. From this measurement, a parametrization can be derived which provides for any off-line selected event, the probability to have fired the specific trigger path (see D0 method or CDF method) as a function of the off-line reconstructed pT, eta and phi of the object identified as the triggering object (geometrical matching in DeltaR between the offline reco object and the online triggering object).

Here we briefly discuss the methods to measure the trigger efficiency (tag-and-probe, bootstrap methods, if existing work in CMS, refer to them). Single-object trigger efficiencies as a function of pT/ET, eta and phi should be centrally provided for the use of all analysis groups. Eventually they should be provided in a data-base as average values, per run, or block of runs.

It is useful to investigate what is the uncertainty on the trigger efficiency measurement as a function of the luminosity and its impact on the measurement uncertainties.

Here we also address the issue of the background measurement. It is also important to evaluate the statistical size of needed background samples. The QCD background can be measured using *e.g.* a 4-jet data sample, selected with a 4-jet trigger. The number of selected 4-jet events is given by

$$N^{4j} = N^{4j}_{QCD} + N^{4j}_{W+Nj} + N^{4j}_{top}$$

Out of this sample, it is then possible to select events with an additional lepton, and plot the lepton distribution, see Fig. (coming). The high-pT part is dominated by top and W+jets events. The low-PT is dominated by QCD. Cutting pT¡XY GeV selects a pure sample of 4jet QCD events from which we can measure the QCD 4jet cross-section and extrapolate the contribution in the top (lepton high-pT region). To measure the 4jet QCD cross section, we also need to measure the 4jet trigger efficiency. We propose to do this using the so-called bootstrap method. It implies that we select on-line, in addition to the 4jet sample which we need for the measurement, also 3-jet, 2-jet, single-jet and min-bias samples. We measure the relative trigger efficiency for Njet triggered events with respect to the N-1-jet triggered sample. The single jet trigger efficiency is measured with respect to the min-bias triggered events (this is needed for several other measurements).

We need to calculate the needed sample statistics as a function of the desired precision on the 4jet QCD cross section measurement and extrapolation.

6 Summary and outlook

Resp.: Javier, Marta

Summarize what is done and what needs further investigation.

7 References

References should be placed at the end of the note (see example [1]).

References

[1] **CMS Note 2005/000**, X.Somebody et al., *"CMS Note Template"*.

[2] **CMS Note 2007/021**, X.Somebody et al., *"CMS High Level Trigger"*.