CERN Summer Student Report:
Analysis of Muon Trigger Efficiencies for the CMS Detector

Stefan Wunsch
CERN, Geneva, Switzerland / Karlsruhe Institute of Technology, Karlsruhe, Germany

Project supervisors: Carlo Battilana, Stefano Marcellini

Project Summary

The project took place within the Muon Physics Object Group (POG) of the CMS collaboration at CERN in 13 weeks from June to September 2016. In particular, I studied the performance of the trigger algorithms for high momentum muons. Such triggers are used to look for several beyond the standard model signatures, therefore their understanding is, for example, important for Z prime and W prime searches. Detailed studies on such front are in particular relevant in 2016 as the first level of the CMS trigger, implemented on custom electronics boards, was completely upgraded during the last shutdown. Therefore, the performance of the new system deserved thorough scrutiny.

The work can be separated in three parts. First, I spent roughly two weeks getting familiar with the needed tools, the workflow and the CMS trigger architecture. More precisely, this included the CMS software framework (CMSSW), the Tag-and-Probe efficiency measurement method in high detail, ROOT and the CERN computing infrastructure, e.g., using EOS for data access and accessing the local computing resources. For the next five to six weeks, I analysed continuously new incoming data from the CMS detector, which has been produced during my stay at CERN to measure the muon trigger efficiencies. Simultaneously, I placed emphasis on trying to make the Muon POG workflows smarter and more efficient. Out of this period resulted the most important activities, which I took care of. First of them was analysing the level 1 (L1) and high level muon trigger in the narrow eta region 1.4 to 1.6 due to significant inefficiencies in the high pt region. Second, I was filling a new software repository with convenient programs, which are more easily usable than the previously used script collection and therefore integrate well in the Muon POG workflow. Following sections describe the specific tasks in more precisely. The technical facts in highest detail can be found in the (CMS intern) slides presented at the Muon POG meeting (`indico.cern.ch/event/568217/`).

Trigger Efficiency Analysis of New Data

For this task, the standard workflow is as follows. The Muon POG produces centrally trees with $Z \to \mu\mu$ events out of the new data. Then, these have to be analysed regarding the most important variables, e.g., the geometric property eta or the transverse momentum pt and the different trigger paths using the data driven analysis method Tag-and-Probe. Typical efficiency measurements are presented in Fig. 1. Out of this studies, we discovered a so far unobserved efficiency drop in the 1.4 to 1.6 eta region. Furthermore, the repetitive character of this task led to the idea to use software designed for continuous integration for the Muon POG workflow.

Trigger Inefficiencies in 1.4 to 1.6 Eta Region

During analyses of new data, we discovered a trigger inefficiency of the high level trigger (HLT) Mu50 in the eta region 1.4 to 1.6 (see Fig. 2).

Fig. 2 shows how such inefficiency depends on the high pt muons. Fig. 3 shows instead the charge dependence of the effect. Two triggers based on independent reconstruction algorithms, Mu50 and TKMu50, are used by CMS to collect events with at least one high transverse momentum muon. As TkMu50 was probed to be immune to the problem (see Fig. 2), efficiency losses for many CMS analyses are largely mitigated. It is anyhow worth understand this issue in better detail and fix it, since, besides "general purpose" triggers, the full list of CMS HLT algorithms include some that do not profit of the backup of a second reconstruction logic.

Events were Mu50 trigger was failing were skimmed and inspected by means of detailed dumps. These allowed to identify the origin of the issue to be in the geometrical matching between L1 trigger candidates and muon tracks reconstructed at HLT using muon spectrometer information only (called L2 muons). For each inefficient event in the affected region, a low quality copy of the L1 muon candidate used to collect the event ("ghost") was found. The L2-L1 matching logic, at times, was found to prefer linking L2 tracks to ghost copies, hence the trigger algorithm stops at the stage when the correspondence between L2 tracks and L1 candidates is probed. A different matching logic is used by the TkMu trigger (were L2 tracks are not built at all), making it robust against such a flaw. Fig. 4 shows how a change in the matching logic can restore the efficiency.

Solutions to overcome the issue were proposed to the Muon HLT group as a consequence of this study.

Evaluation of Continuous Integration for the Muon POG Workflow

I tried to integrate the Muon POG workflow of analysing new data from CMS in a continuous integration framework. The study was done using Jenkins because the CERN computing resources offer a Jenkins service and it is highly not encouraged to host server software locally on your own. Although, a first prototype of the developed machinery was fully working, we decided not to employ it in the Muon POG daily workflow. The advantages such as triggering the analysis task automatically on new data and publishing the results as a webpage do not outweigh the problem of a fairly complicated Jenkins setup, which is most likely unmaintainable for users with no continuous integration experience. As a result of this, we decided to make the collection of currently used scripts easier to use and add newly developed ones in a new Muon POG repository.

Muon POG Utility Repository

The Muon POG utility repository resulted out of the idea to make the daily work easier. The programs are publicly available at `www.github.com/cms-MuonPOG/TnPUtils`. The already available scripts, developed for very specific use cases, were ported to Python so that a program and argument description is inline available and callable directly from the terminal without looking at the source itself. Furthermore, the usage of PyROOT reduces the code base substantially and makes argument handover much easier, which results in a better maintainability and usability. As well, the so far used code was partially only working within specific workflows, which was generalised during the code porting. Besides, I added a set of new tools. First, I introduced programs for generic editing of Tag-and-Probe trees. Especially skimming trees and reducing the subset of stored variables is getting increasingly important because the trees produced by the Muon POG serve a wide range of use cases (hence are rather large in terms of used memory per event). Moreover, this year the amount of data collected by CMS has got large enough to make the analysis programs fail due to excessive memory consumption if no reduction of the input data is performed before the actual analysis of the trees. Analysis programs are being optimised, but at present the afore mentioned tools are essential to make analysis workflows leaner and faster. Second, plotting and publishing tools were developed, which enables the user to produce publishable plots from Tag-and-Probe analyses rapidly and put them together on a webpage for a more productive collaborative workflow. Last, simple and out of the box working examples were given in the repository itself to make the use of the programs easily understandable for new users. Note, that all efficiency plots in this report are directly generated from Tag-and-Probe result files using programs from this repository.

Systematical Uncertainties Measurement

So far unmentioned, I have set up a machinery to measure the systematical uncertainty on a Tag-and-Probe efficiency measurement. This task was mainly done out of educational purposes and did not contribute directly to the Muon POG activities. Nevertheless, it is worthwhile to mention in this report because I learned all details of providing systematical uncertainties, which is a highly sophisticated task. E.g., the differences between simple variance and maximum error approach or the usage of bias correction were evaluated with respect to the final goal to present a fully justifiable systematical uncertainty estimation. The performed case studies on new data fully agreed with the expectations and measurements previously performed by the Muon POG.

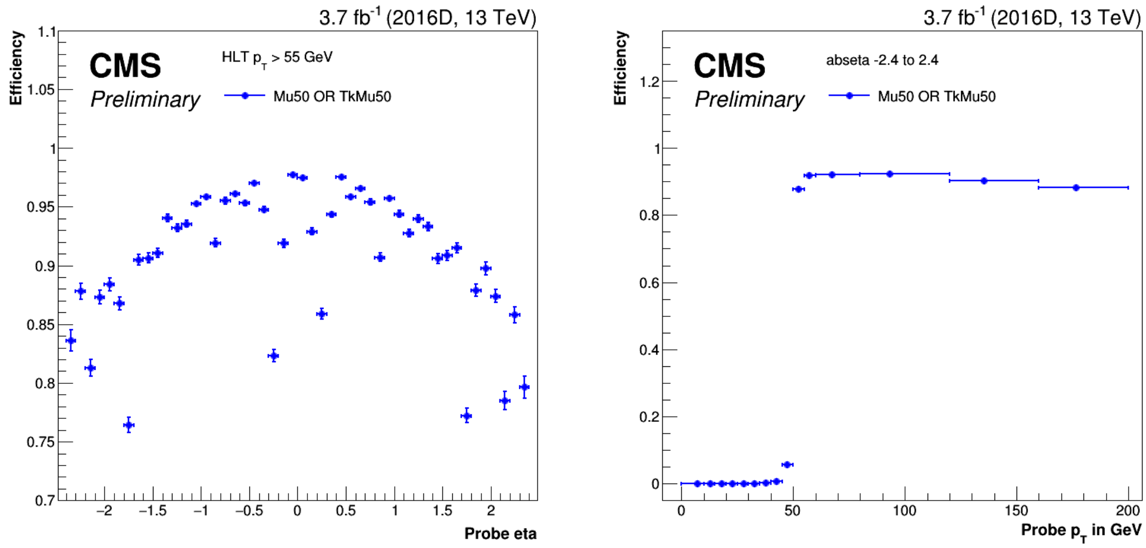

Fig. 1: Efficiency measurements using Tag-and-Probe respective to eta and pt

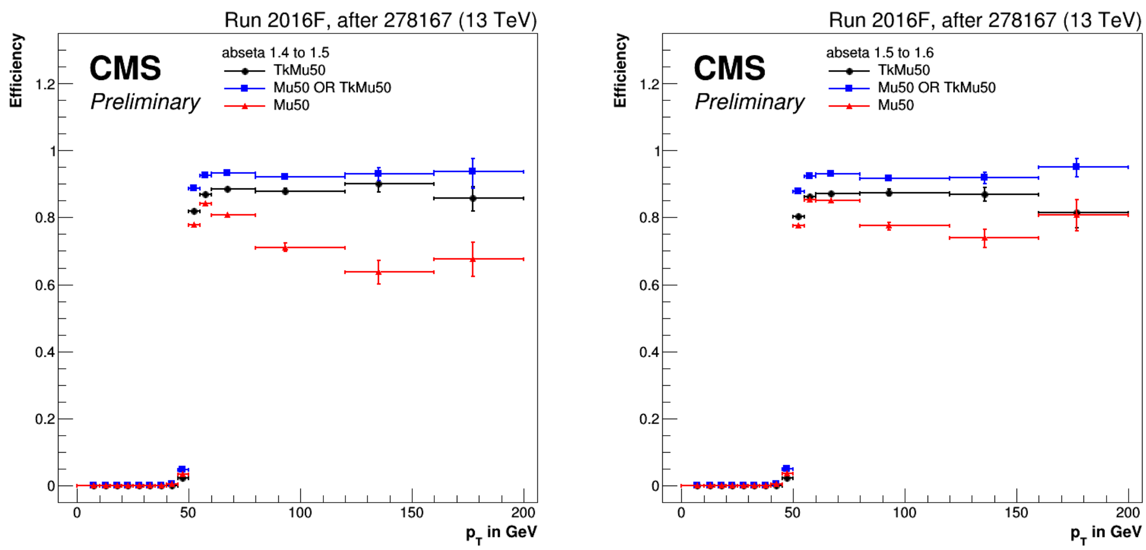

Fig. 2: Comparison of Mu50 and TkMu50 high level trigger

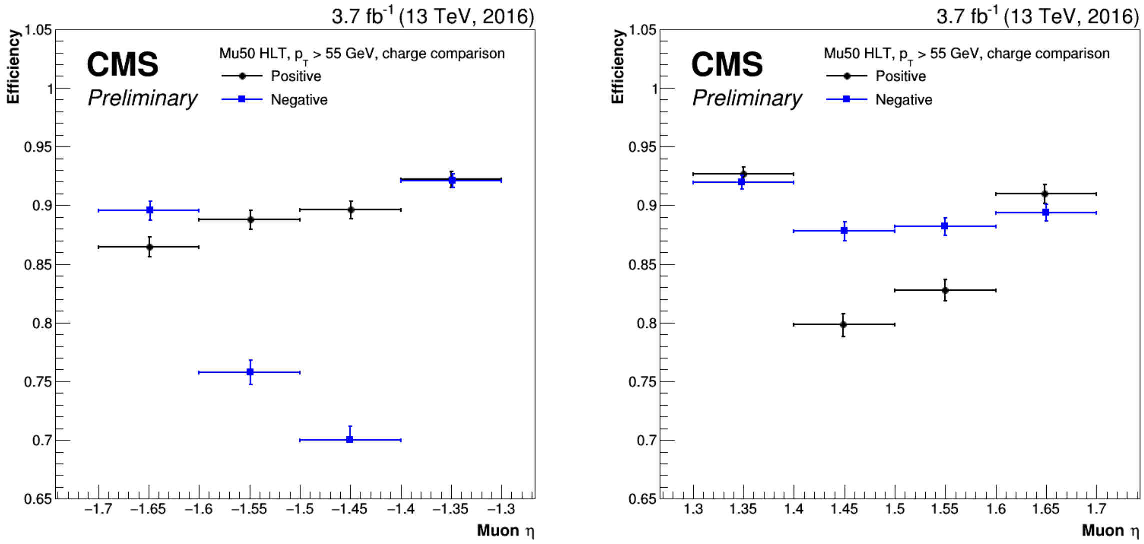

Fig. 3: Charge dependency of Mu50 high level trigger inefficiency

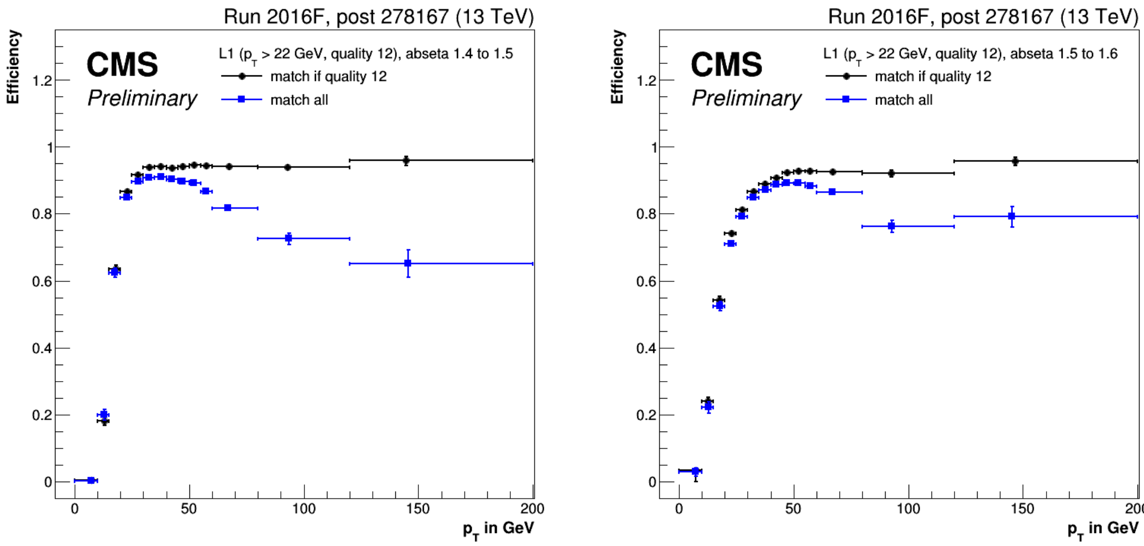

Fig. 4: L1 trigger efficiency recovery with changed matching logic