Multi-Muon and Heavy-Flavour studies with CMS OpenData

Marc Quentin Führing, Technische Universität Dortmund, Germany

supervised by Achim Geiser (DESY)

September 6, 2017

Abstract

The CMS OpenData program gives external researcher and private persons access to CMS data. Within this project references, which can be used by external persons to validate analysis with CMS data from the year 2011, are created by reconstructing the di-muon mass. Furthermore, the 2012 data are validated for their upcoming release.
To demonstrate the potential of the CMS OpenData, the di-muon validation example is extended to tri-muon and tetra-muon studies. Within these studies there is no unexpected decay in tri-muon final states observable. The tetra-muon studies are not able to measure the predicted, but still unmeasured decay $J/\Psi \rightarrow \mu\mu\mu\mu$.
Besides the muon studies, a former heavy flavour analysis' capability to work with CMS OpenData is validated by reconstructing the B^+mass and prepared for upcoming analysis.

Contents

1 Introduction

As a part of the CERN OpenData program[1] the CMS collaboration releases its data to the public. This way a wider community is able to contribute research by analyzing data, the data can be used for education and cross-referencing between different experiments is possible. To guarantee correctness, it is essential to validate released subjects and to give the public an opportunity to validate their setups. While there are first analysis published by external researchers[2], these studies prepare further CMS OpenData releases by validating 2012 CMS data, which are in the pipeline to be released, and by partly validating a heavy flavour analysis.

2 The CMS-Experiment

The Compact Muon Solenoid (CMS) Experiment is one of the four big Large Hadron Collider (LHC) experiments at CERN in Geneva. The CMS-detector is a 4π-multipurpose detector. As it name says, the detector is a solenoid spectrometer. The superconducting solenoid magnet generates a magnetic field up to 3.8 T. Inside the solenoid's 6 m diameter are the silicon tracker, divided into pixel and strip tracker, the electromagnetic and hadronic calorimeter placed. The muon chambers are located outside of the solenoid.

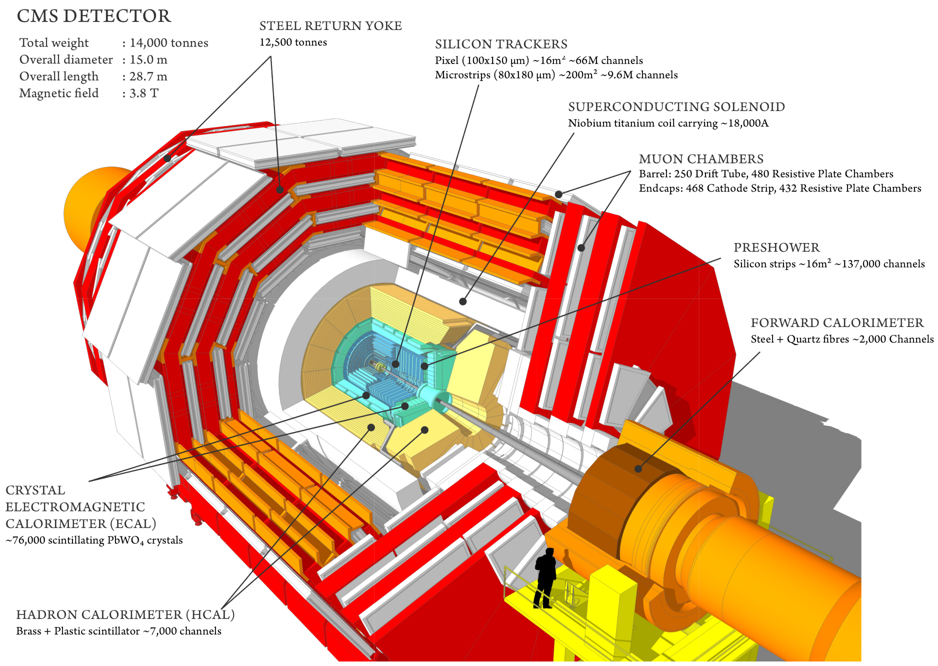


Figure 1: The CMS-detector[3].

2.1 CMS OpenData

The CMS OpenData program is the preceding program within the CERN OpenData program. The goal of this program is to give the public access to CMS data. This way the data can be used for education, can be analyzed by a wider community of researchers and can be used for cross-referencing between different experiments. Some people also see a contribution to data preservation in the OpenData program by spreading the data. Within this program CMS releases both data and analysis tools. The released datasets are primary datasets, which are prepared after the trigger selections, but there no further selection cuts applied. There is also a collection of simulated datasets avaible. At the moment the CMS data from 2010 and 2011 are released and the 2012 data are in the pipeline to be released soon. The analysis tools contain the CMSSW software to run CMS code and virtual machines (VM) as environment for the software.

2.1.1 EOS-Issues

EOS is a disk-based storage service at CERN. Next to several different projects EOS hosts the CERN OpenData program. During this project, there are two issues observed:

- A large amount of unpermitted writing attempts on EOS server, while reading the data files.
- Temporal not accessible data files.

After consultation with both EOS and CMSSW, the writing attempts are declared to be harmless. They are caused by the interplay of CMSSW code and updated EOS code. Meanwhile this issue is fixed.
According to EOS, the not accessible data files are caused by double disk failures. It is point out by EOS, that sporadic double disk failures are still possible in the future. In case, it is recommended to run another data file first and to rerun the affected data file a few hours later. This should be documented and released to the OpenData user, who could face this issue. Furthermore, this doubts the OpenData program's suitability of additional data preservation, which some people are likely to see, in its actual implementation.

3 Multi-Muon studies

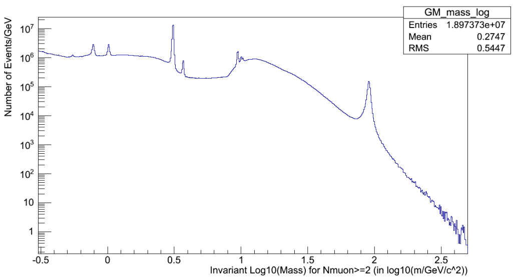
The multi-muon studies aim to create multi-muon mass spectra to observe resonances. The used datasets are the SingleMu and the DoubleMu datasets with events, which are selected because of the presence of at least one (SingleMu) or two (DoubleMu) high-energy muons. For 2011 data the run A data are used, which contain all 2011 data, for 2012 the run B data only are used, because the run C data are not avaible yet.

3.1 Di-Muon Validation Example

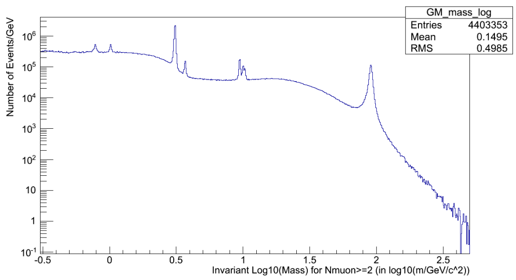
The di-muon validation example[4] has already been used at the earlier CMS OpenData Releases of 2010 and 2011 data. It is a simple analysis, which is reconstructing the di-muon mass from data. Surveying the di-muon mass spectrum turned out to be a good way to validate the published data and analysis tools. In this project the di-muon validation example shall be extended by reference results for the 2011 data. This gives the public an opportunity to validate the setup and the data. Furthermore, the 2012 data shall be validated to prepare their upcoming release.

Both SingleMu and DoubleMu datasets are used for the validation. Only a few loose selection cuts are applied to the muon candidates to keep the example as simple as possible. So a muon candidate has to have at least 12 hits in the silicon tracker, with at least one hit in the pixel detector, to separate muon from decays of long lived pions and kaons. Only muon pairs with a total charge of $Q(\mu\mu) = 0$ are considered, because the standard model predicts decays in pure di-muon final states only for neutral particles.

In the reconstructed di-muon spectrum, shown in figures 2 and 3, the expected resonances of several neutral mesons and of the Z^0-boson are observable. So the 2012 run B DoubleMu and SingleMu datasets are validated. Besides, the 2011 software environment's predicted capability to work with 2012 data is validated.

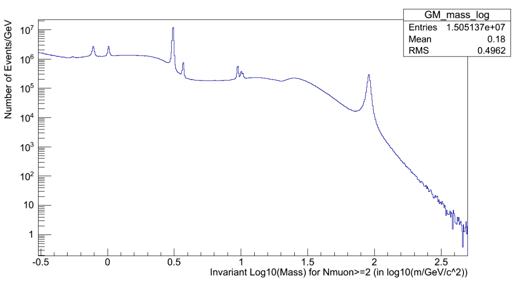


(a) DoubleMu dataset.

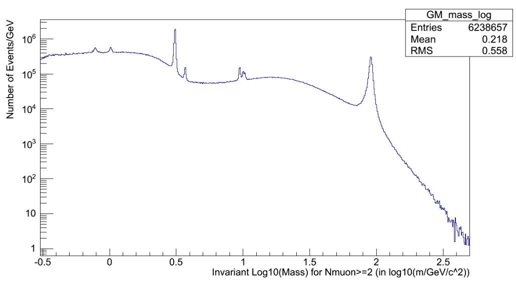


(b) SingleMu dataset.

Figure 2: 2011 di-muon mass spectra.



(a) DoubleMu dataset.



(b) SingleMu dataset.

Figure 3: 2012 run B di-muon mass spectra.

3.2 Tri-Muon Studies

One of the OpenData program's goals is cross-referencing. So LHCb studies[5] about the flavour-violating decay $\tau^- \to \mu^-\mu^+\mu^-$ inspired to have a look at the tri-muon spectrum. To keep these studies close to the original di-muon Example 3.1, there are only a few loose selection cuts applied to the muon candidates. Again every muon candidate has to have at least 12 hits in the silicon tracker of which at least one hit has to be in the pixel detector.

The reconstructed tri-muon spectrum, shown in figure 4, is divided into two charge configuration: a total tri-muon charge of $Q(\mu^\pm\mu^\mp\mu^\pm) = \pm 1$ and of $Q(\mu^\pm\mu^\pm\mu^\pm) = \pm 3$. It shows no unexpected features. Though it is to mention, that this does not exclude the possibility to observe tri-muon final states within these data with deeper analysis.

The mass peak at $\log_{10}[M(\mu^\pm\mu^\mp\mu^\pm)/\mathrm{GeV/c^2}] = 2$ is caused by reconstructing a reflected Z^0-boson $M(Z^0(\mu^+\mu^-) + \mu)$. Similar to this the Bump at $\log_{10}[M(\mu^\pm\mu^\mp\mu^\pm)/\mathrm{GeV/c^2}] = 0.6$ is also caused by reconstructing a reflected J/Ψ-meson $M(J/\Psi(\mu^+\mu^-) + \mu)$.

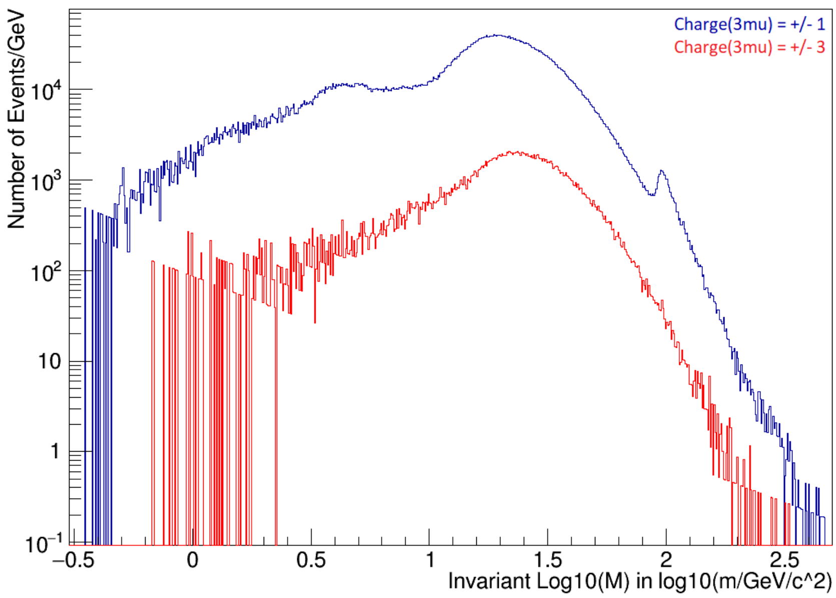


Figure 4: Tri-muon mass spectrum (2011 DoubleMu dataset). The blue line shows the configuration with a total Charge of ± 1, the red line a charge of ± 3.

A Dalitz-like plot 5 is used to have a look at substructures. To create this plot the muon with unique charge within the muon-trio is used to estimate the di-muon masses with both of the remaining muons. Finally, the di-muon mass of the higher p_T muon-pair is plotted against the lower p_T muon-pair's di-muon mass. Similiar to the di-muon mass spectrum, the resonances of several neutral mesons and of the Z^0-boson are observable in the Dalitz-like plot. With this illustration the features of the reflected masses in figure 4 is more obvious.

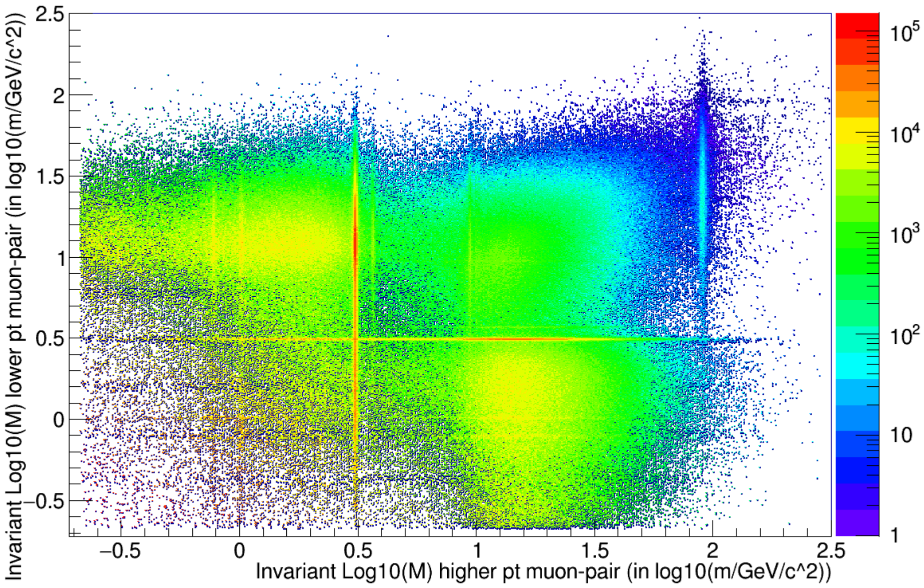


Figure 5: Di-muon mass $M(\mu_0\mu_1)$ vs $M(\mu_0\mu_2)$.With $Q(\mu_0) \neq Q(\mu_1) = Q(\mu_2)$ and $p_T(\mu_1) \geq p_T(\mu_2)$ (2011 DoubleMu dataset).

3.3 Tetra-Muon Studies

The decay $J/\Psi \to \mu\mu\mu\mu$ is predicted but not measured yet. This decay is caused by final state radiation in the $J/\Psi \to \mu\mu$ decay and suppressed by a factor α^2 relative to it. To keep these studies consistent to the di-muon 3.1 and tri-muon 3.2 studies, there are only the same few loose cuts applied to muon candidates. So only reconstructed muons with at least one pixel tracker hit and at least 12 silicon tracker hits are assumed. The resulting tetra-muon mass spectrum, shown in figure 6, is divided by the total charge of the four muons into three spectra: $Q(\mu^\pm\mu^\mp\mu^\pm\mu^\mp) = 0$, $Q(\mu^\pm\mu^\mp\mu^\pm\mu^\pm) = \pm 2$ and $Q(\mu^\pm\mu^\pm\mu^\pm\mu^\pm) = \pm 4$.

Again there is a bump caused by the reflection of the J/Ψ mass $M(J/\Psi(\mu^+\mu^-) + \mu\mu)$ visible in the range of $0.5 \leq \log_{10}[M(\mu\mu\mu\mu)/\mathrm{GeV/c^2}] \leq 1$. This becomes obvious by looking at the substructure of the tetra-muon final state shown in the Dalitz-like plot 7. The substructure shows the expected J/Ψ resonances and a small Z^0 resonance decaying in a high p_T muon pair. The peak, which is visible at the crossing point of both J/Ψ resonances, indicates, that there are several events, which contain at least two J/Ψ mesons decaying to two muons each.

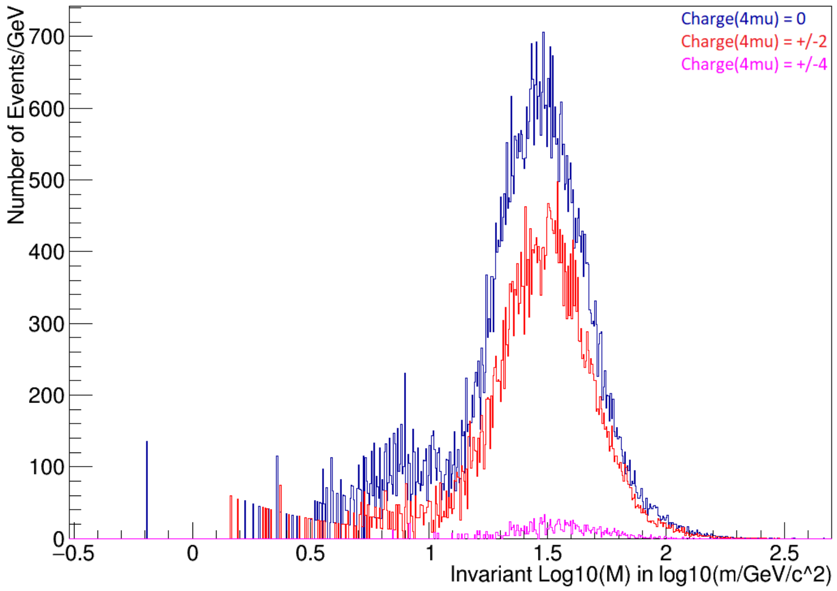


Figure 6: Tetra-muon mass spectrum (2011 DoubleMu dataset). The blue line shows the configuration with a total Charge of 0, the red line a charge of ±2 and the pink of ±4.

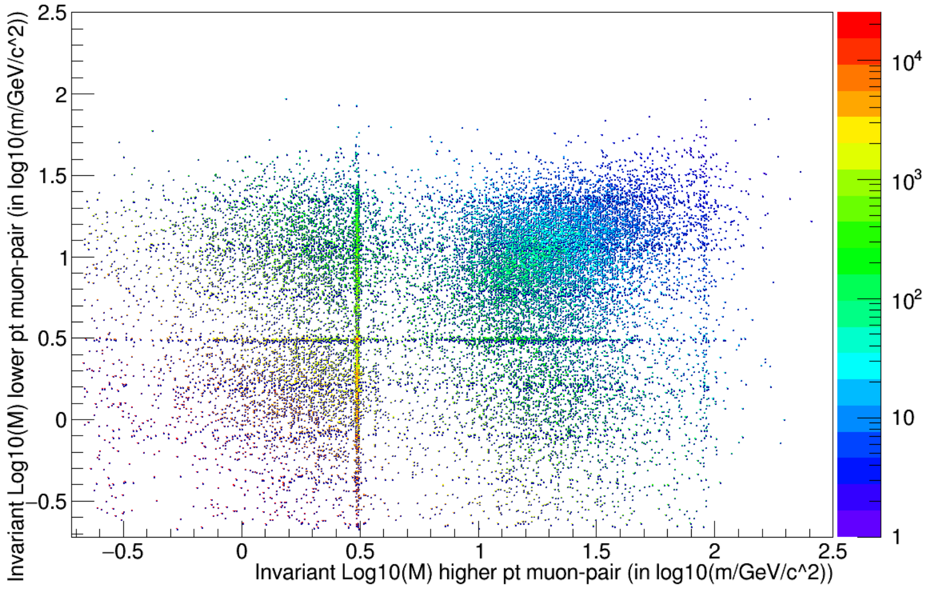


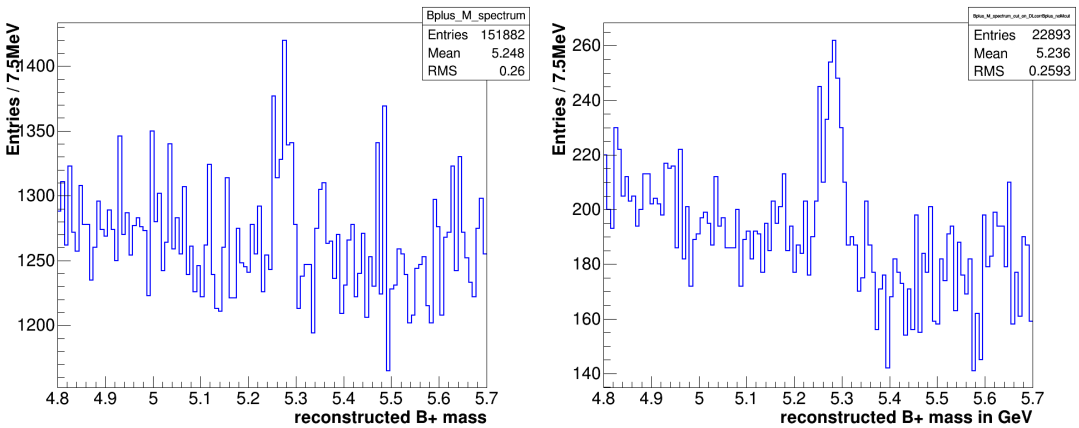
Figure 7: Di-muon mass of uncharged muon pairs: the uncharged muon pair combination with highest p_T against the remaining muon pair (2011 DoubleMu dataset).

4 Heavy Flavour Studies

Upcoming studies, which are basing on a former heavy flavour analysis[6], require to recreate and validate the analysis on CMS OpenData. In this context the CMS 2010 OpenData are used. During recreating and running the analysis on 2010 data, the used 2010 VM shows performance and stability issues. These stability issues occur by running the heavy flavour analysis, with improved vertex fitting and decay length calculation, on higher statistics and have been seen at the 2010 VM only, yet.
Due to these circumstances, the study is limited to the validation of the B^+ meson reconstruction on small statistics.

4.1 Beauty Production Validation

In this study B^+ mesons are reconstructed by the the decay $B^+ \to J/\Psi(\to \mu\mu)K^+$, where the J/Ψ is reconstructed by its di-muon decay. The cuts applied to the candidates, used to reconstruct the B^+, are listed in table 1. A improved vertex fitter[6] is used to associate candidates to matching vertices and to calculate the decay length. The decay can be used as additional selection criteria for the short lived B meson. The resulting B^+ mass spectrum, shown in figure 8, shows the expected resonance at 5.279 GeV.



(a) Mass spectrum, without decay length cut.

(b) Mass spectrum, with cut on decay length > 0.1 mm.

Figure 8: Reconstructed B^+ mass spectrum.

The relations between analyzed events, reconstructed B^+ candidates before and after the decay length cut, which are shown in table 2 are well fitting in there order of magnitude to the original analysis' observation. So the analysis is considered to be validated with CMS OpenData from 2010.

Table 1: Selection cuts for B^+ reconstruction

Candidate	cut
Muon ($0.0 < \|\eta\| < 1.3$):	$p_T > 3.3\,\mathrm{GeV/c^2}$
Muon ($1.3 < \|\eta\| < 2.2$):	$p_T > 2.9\,\mathrm{GeV/c^2}$
Muon ($2.2 < \|\eta\| < 2.4$):	$p_T > 0.8\,\mathrm{GeV/c^2}$
Muon:	$\chi^2_{\mathrm{track}} < 4$
J/Ψ:	$2.95\,\mathrm{GeV/c^2} < M < 3.25\,\mathrm{GeV/c^2}$
J/Ψ:	$\|y\| < 2.4$
B^+:	$4.8GeV/c^2 < M < 5.7\,\mathrm{GeV/c^2}$
B^+:	$\|y\| < 2.4$
B^+:	$p_T > 5\,\mathrm{GeV}$
B^+:	$DL < 0.1\,\mathrm{mm}$

Table 2: Comparison of used statistics.

	Original Analysis	Validation
Datasets:	minBias, MuOnia	MuOnia
Events:	approx. 16×10^6	3×10^6
B^+ candidates before DL cut:	992 168	151 882
Candidates per event:	0.0620	0.056
B^+ candidates after DL cut:	173 581	22 893
Candidates per event:	0.010	0.008

5 Summary

Working on this project issues in the interplay with EOS, the for the OpenData storage responsible service, are observed. The more harmless issue of writing attempts from CMS software at EOS server is solved now. The rare, but more harmful issue of double disk failures at EOS is documented. It is highly recommended to inform the OpenData user about this issue. Furthermore, the EOS double disk failures doubt the OpenData program's suitability of additional data preservation, which some people are likely to see. Using the existing di-muon example reference plots are created, which gives the opportunity to validate the CMS OpenData setup to the public. Besides to this, 2012 run B data are validated for their umpcoming OpenData release. The predicted feature of the 2011 software environment to work with 2012 data is validated, too.
Further the di-muon example is extended to tri- and tetra-muon studies. Within the tri-muon studies no unexpected decay in a tri-muon final state is observed. The predicted decay $J/\Psi \to \mu\mu\mu\mu$ is not observable in the tetra-muon studies. Though, upcoming, more complex studies than this one, which is kept simple to keep it as close as possible to the di-muon example, could see more.
A former heavy flavour analysis is validated by reconstructing the B^+ mass and prepared for further studies. In this context stability issues of the used CMS OpenData 2010 VM are observed, which are not seen at the newer 2011 VM, yet.

Acknowledgements

I like to thank the DESY for the opportunity to be part of this years Summer Student Program. I enjoyed the prgram a lot: the lectures have been really instructive, working at DESY has been a great and unique expirience and I have got a nice view in various cultures.
A special thanks goes to Achim Geiser, not only for supervising my project, but also for teaching me a lot and for being always avaible for giving advices. I also want to thank Nazar Stefaniuk, who contributed with many advices, although he is not employed at DESY anymore, and Nur Zulaiha Jomhari, who has always been willing to help, even in her stressful acclimation period here at DESY.
Furthermore, I thank the organisation team for the smoothnes of all procedures here at DESY.
At last I do not want to forget to mention Olaf Behnke and all the other Summies, who have always been great company here in Hamburg and who caused a lot of fun.

References

[1] URL: http://opendata.cern.ch.

[2] "Jet Substructure Studies with CMS Open Data". In: MIT-CTP 4890 (May 2017). URL: https://arxiv.org/abs/1704.05842.

[3] URL: http://cms.cern/detector.

[4] URL: http://opendata.cern.ch/getting-started/CMS.

[5] "Search for the lepton flavour violating decay $\tau^- \to \mu^-\mu^+\mu^-$". In: *JHEP* 02.LHCB-PAPER-2014-052. LHCB-PAPER-2014-052. CERN-PH-EP-2014-236 (Oct. 2014). Comments: 19 pages, 6 figures, submitted to JHEP, 121. 19 p. URL: http://cds.cern.ch/record/1951383.

[6] Nazar Stafaniuk. "Measurement of the total and differential b cross sections at HERA and CMS tracker alignment at LHC". Hamburg: DESY, 2017. URL: https://csd.cern.ch/record/2275332/files/TS2017_016_2.pdf.