J/ψ,Υ and Z studies with CERN Open Data

Lekshmi Thulasidharan

Central University of Tamil Nadu, India

Supervised by: Achim Geiser

Deutsches Elektronen-Synchrotron (DESY), Hamburg, Germany

September 5, 2018

Abstract

The validation of 2010 Compact Muon Solenoid (CMS) open data is done for the full statistics, by successfully reconstructing the J/ψ, Υ and Z decay channels. Among the 14 primary datasets available in the CMS Open Data portal, the Muonia and Muon Datasets are used for the validation. The example to validate the usage of the MuOnia dataset is successfully finalized for full statistics. Analysis examples demonstrating the reconstruction of J/ψ,Υ and Z via their dimuon decay channel along with mass fits are finished. The successful conversion of A Toroidal LHC ApparatuS (ATLAS) data to a nano Analysis Object Data (nanoAOD) like format has also been done. The dimuon invariant mass plots were compared with the ones obtained from CMS.

Contents

1 Introduction

The original data used by the CMS collaboration of CERN can be accessed by the public via the CERN Open Data Portal (http://opendata.cern.ch/). The main purpose behind this is bi-fold. It will help the researchers outside the CMS collaboration to get an opportunity to do research with this data and to perform independent analysis, even without becoming a part of the collaboration. School and university students can utilize the simplified data for educational purposes and scientific outreach.

This project is mainly focused on the validation of the 2010 Muonia dataset and to set up analysis examples demonstrating the reconstruction of J/ψ,Υ and Z via their dimuon decay channel along with fits. The validation of Open Data is vital to ensure the data is clean, correct and useful, and can be safely used for research and educational purposes. In addition to this, an ATLAS open dataset is converted to nanoAOD-like format and the dimuon invariant mass plots are compared with those of CMS open data in the same format.

2 Exercise of 2010 Muon Dataset

An already available exercise in the CMS open data portal is done to reproduce the dimuon invariant mass distribution and is compared with the plot from the CMS collaboration paper [1].

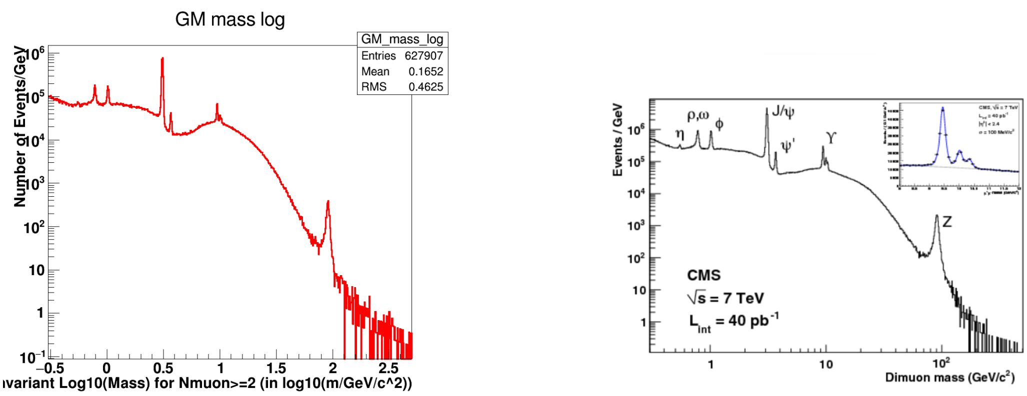

Figure 1: Di-muon invariant mass distribution using Mu sample

The figure on the left shows the number of entries as a function of the logarithm of invariant mass of the dimuons. One can identify J/ψ, Υ and Z peaks at log10(mass)=0.5, log10(mass)=0.98 and log10(mass)=1.95, respectively. This exercise was helpful in understanding how to deal with virtual machines and also helped to get familiar with the CMS environment. While doing this exercise, certain issues like ‘truncated files’ were found and were reported, and are currently getting fixed.

3 New Validation of 2010 MuOnia Dataset

Based on the previous work done by Bridget [5] and Irene [6] (on partial statistics), the 2010 MuOnia Dataset is validated for the full statistics and the invariant mass plots corresponding to J/ψ,Υ and Z are obtained.

There are two categories of triggers in the MuOnia dataset:

- $\sim$ 30% inclusive dimuon triggers with no explicit trigger transverse momentum p_T threshold or loosened requirements on 2^{nd} muon.
- $\sim$ 70% J/psi or quarkonium (dimuon) triggers with no or very loose p_T thresholds and cuts on the dimuon mass in the range 1.5-15 GeV.

3.1 Validation of MuOnia Datasets by the Reconstruction of Quarkonium States

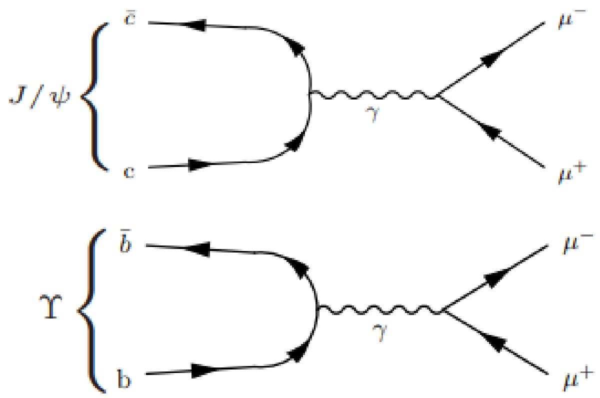

Figure 2: Decay of J/ψ and Υ to their dimuon final states

MuOnia datasets can be validated through the reconstruction of the quarkonium states J/ψ and Υ through their final decay to $\mu^+\mu^-$ final states. The figure shows the electromagnetic decay of J/ψ (bound state of c$\bar{c}$) and Υ (bound state of $b\,\bar{b}$) to their dimuon final states. Once the reconstruction is done successfully, one can compare the results with the plots in the published CMS collaboration papers to cross-check.

Separate cuts were applied to isolate J/ψ and Υ candidates. Quality criteria applied ensured the separation of muons from the decay of kaons or pions. For the reconstruction of the $J/\psi \rightarrow \mu^+\mu^-$ decay channel, the selection criteria used by the CMS collaboration in [1] is used. For the reconstruction of $\Upsilon \rightarrow \mu^+\mu^-$, the selection cuts are obtained from [2].

3.2 Reconstruction of the $J/\psi \rightarrow \mu^+\mu^-$ decay channel

The selection criteria obtained from [1], are implemented and the kinematic cuts were applied across three rapidity regions, such that muons should satisfy either:

$$\begin{array}{ll} p_T^\mu > 3.3GeV/c & \text{if } |\eta^\mu| < 1.3 \\ p_T^\mu > 2.9GeV/c & \text{if } 1.3 < |\eta^\mu| < 2.2 \\ p_T^\mu > 3.3GeV/c & \text{if } 2.2 < |\eta^\mu| < 2.4 \end{array}$$

Dimuon invariant mass plots are obtained after applying all the selection criteria from [1], in the region between 2.6 and 3.5 GeV/c^2, for three different rapidity regions as shown in the diagram. The plots are fitted using a Crystal Ball function, with a polynomial background. The plot obtained is compared with the corresponding plots in the published paper [1].The collaboration paper has partial statistics whereas the validation is done for full statistics.

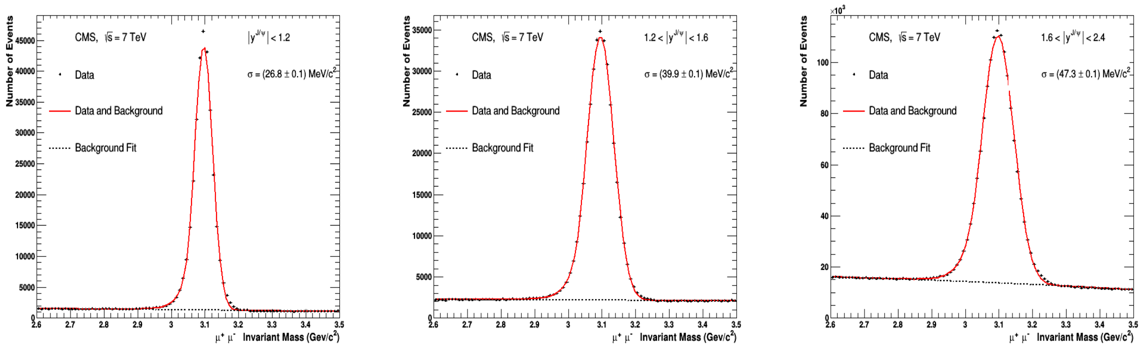

Figure 3: J/ψ Invariant Mass Peaks for three different rapidity regions (Integrated Luminosity $\sim 32\text{pb}^{-1}$), fitted with a Crystal Ball function

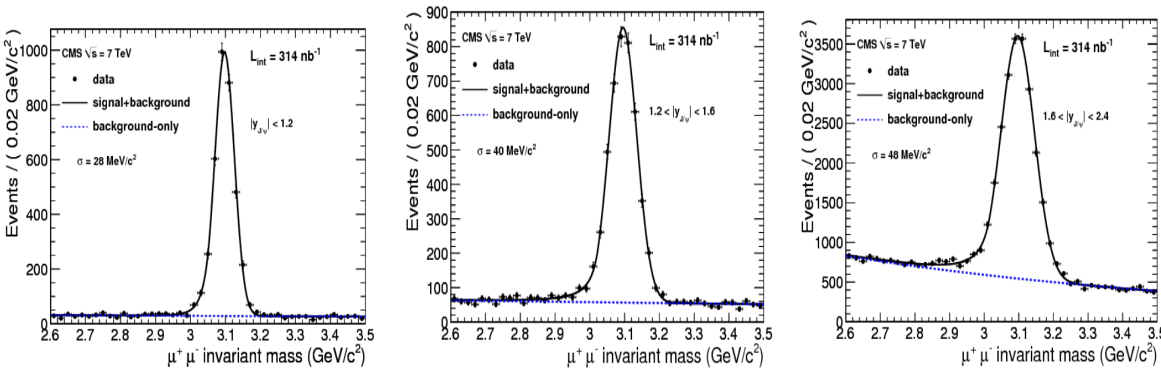

Figure 4: Dimuon invariant mass distribution between 2.6 and 3.5 GeV/c^2 obtained from [1]

4 Reconstruction of the $\Upsilon \rightarrow \mu^+\mu^-$ decay channel

For the reconstruction of the $\Upsilon \rightarrow \mu^+\mu^-$ decay channel, the selection criteria obtained from [2] are used. Kinematic cuts were applied across two pseudorapidity ranges, in which the muons were required to satisfy either:

$$
\begin{array}{ll}
p_T^\mu > 3.5 GeV/c & \text{if } |\eta^\mu| < 1.6| \\
p_T^\mu > 2.5 GeV/c & \text{if } 1.6 < |\eta^\mu| < 2.4
\end{array}
$$

The dimuon invariant mass is obtained after applying these cuts, in the region between 8 and 14 GeV/c^2. The plots are fitted using a triple Crystal Ball function with an exponential or polynomial background. The plot obtained is compared with the corresponding plots in the published paper[2]. The collaboration paper has partial statistics whereas the validation is done for full statistics. The fit for the $|\eta^\mu| < 1.6$ produced a Υ(1s) mass of 9.4562 ± 0.0008 GeV/c^2, which is consistent with PDG world average value of 9.46030 ± 0.00026 GeV/c^2.

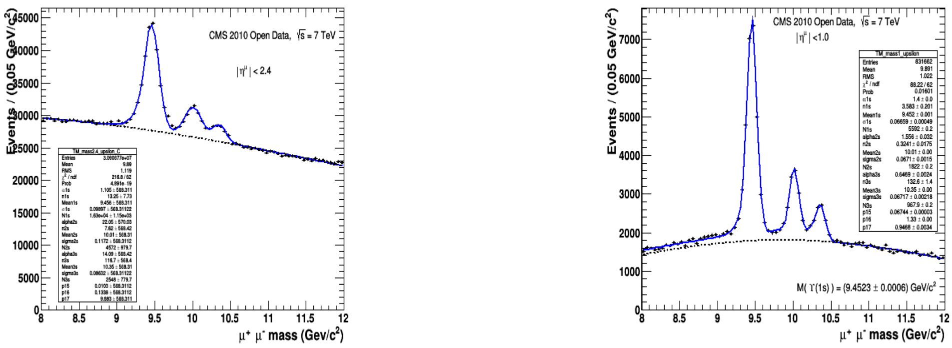

Figure 5: Υ invariant mass peaks for different pseudorapidity ranges $|\eta^\mu| < 2.4$ and $|\eta^\mu| < 1.0$ (Luminosity $\sim 32\text{pb}^{-1}$), fitted with Crystal Ball function

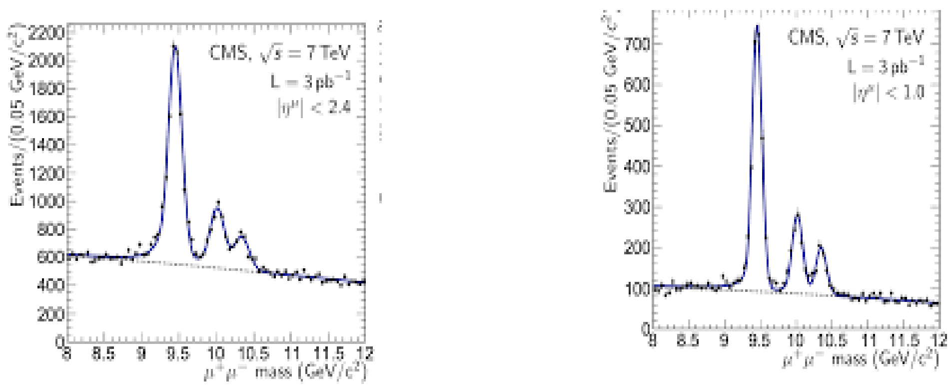

Figure 6: Reference plots from [2]

5 Analysis using 2010 Muon dataset- Reconstruction of $Z \rightarrow \mu^+\mu^-$ decay channel

In order to probe into the high mass region one can't use the MuOnia dataset since there are cuts on the dimuon mass in the range 1.5-15 GeV. Therefore, the analysis is also done using the Muon dataset for full statistics to obtain the invariant mass peak of Z. The same kinematic cuts were applied across the whole pseudorapidity region, in which the muon is required to satisfy :

$$p_T^\mu > 20 GeV/c \qquad \text{if } |\eta^\mu| < 2.1$$

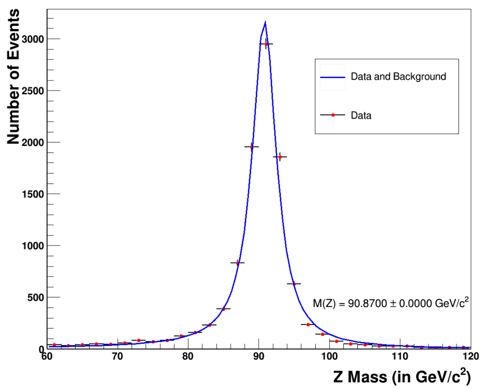

Figure 7: Z Invariant mass peak fitted using Breit- Wigner function

The Dimuon invariant mass plot between the region 60 GeV/c^2 and 120 GeV/c^2 is obtained, after applying these cuts. The plots are fitted using a Breit-Wigner function. The Z mass peak is fitted reasonably and produced a mass of 90.87 GeV/c^2 with a negligible statistical uncertainity and is reasonably agreeing with PDG average of 91.1876 $\pm$0.0021GeV/c^2. We still can improve the fit of the Z mass peak.

6 Summary of Validation project

The results are consistent with the published ones. We have successfully finalized the example to validate the usage of the MuOnia dataset for full statistics. Analysis examples demonstrating the reconstruction of J/ψ, Υ and Z via their dimuon decay channel along with the fits are now finished. The Validation and Analysis examples, including the fits and reference plots are ready to go to the CMS Open data portal. The validation exercises require a CMS environment via the virtual machines, and may not be easy for a student or a researcher who is not familiar with this, to reproduce these results. Therefore, it makes sense to have a data format which can be accessed through ROOT and to perform analysis outside the CMS environment. This gives us a motivation for the nanoAOD-like format.

7 Introduction to nanoAOD-like format

For Run 1, the data format used in the CMS experiment is AOD or Analysis Object Data. Since this is huge, the CMS collaboration introduced a new data format which is the miniAOD for Run 2. This is smaller in size as compared to AOD. Both AOD and miniAOD are Event Data Model (EDM) files. This makes it difficult to access it directly via ROOT. Therefore, CMS introduced a new format- the nanoAOD format, which is nothing but a flat ROOT ntuple. This makes sense to convert the Run 1 and Run 2 datasets into similar format.The nanoAOD like format is a flat ROOT ntuple having the same variables as the nanoAOD, which is produced by taking AOD as an input file. The motivation behind the nAOD-like format is predominantly because of three reasons:

- The Simplified format will make it easier to handle the data.
- It is independent of CMSSW versions - enabling the analysis in a non-CMS environment.
- It allows a comparison with the results from other experiments like ATLAS.

8 Conversion of ATLAS Open Data to nanoAOD like format

ATLAS Open Data [4] is available to the public for educational purposes and can be accessed via the CERN Open data portal. We made use of the 1 fb^{-1} of 2012 pp μ data (Cuts - $p_{T\mu_1} > 25GeV$ and $p_{T\mu_2} > 5GeV$) for converting it into nanoAOD-like format. Successful conversion of ATLAS Open data to the nAOD-like is done and the dimuon invariant mass plots are compared with the former to latter. The plots looks identical as shown in Figure 8.

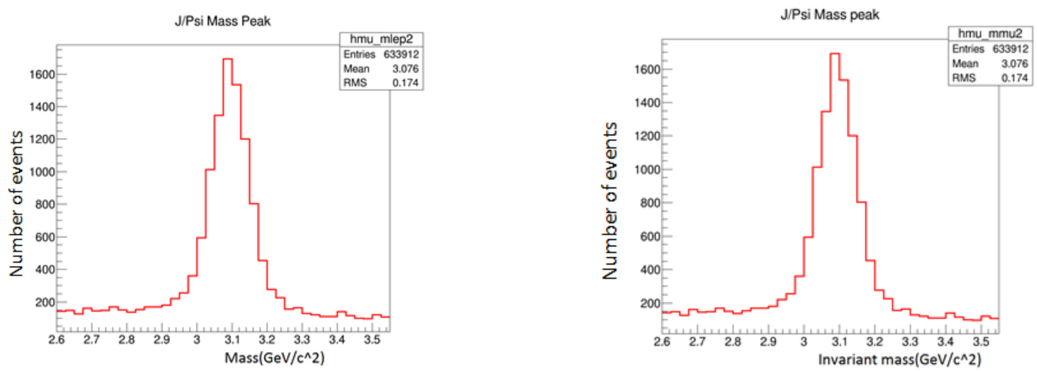

Figure 8: (From Left)Comparison of J/ψ invariant mass peak obtained using ATLAS Open data and nanoAOD-like format (after conversion)

9 Comparison of ATLAS Open Data (nanoAOD-like) with CMS Open Data (nanoAOD-like)

For the CMS 2011 Muonia dataset is used, applying similar cuts as ATLAS (mentioned in section 8). In order to compare the dimuon mass plots obtained from CMS and ATLAS, it is useful to have it in similar data format. Therefore, we compare the invariant mass plots of the CMS nanoAOD-like dataset with the ATLAS nanoAOD like dataset, which will also give an insight on the respective mass resolutions of ATLAS and CMS.

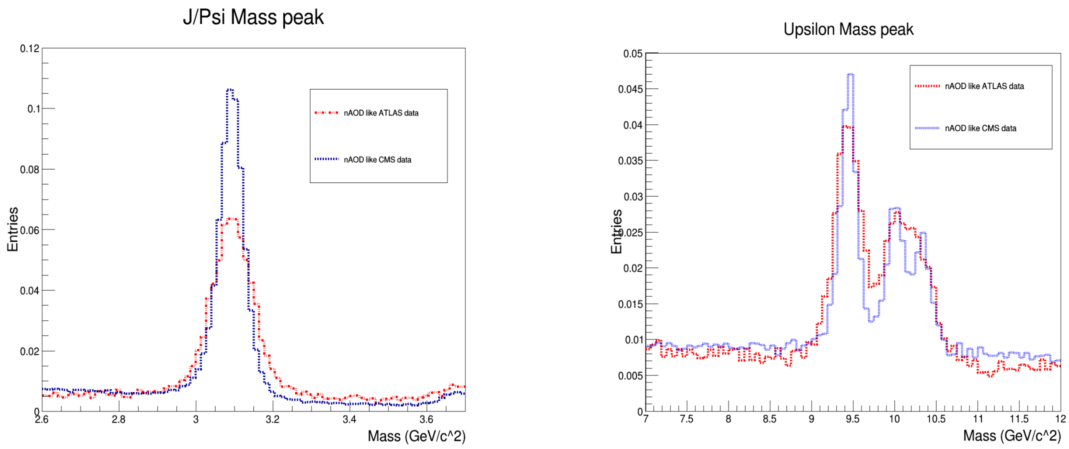

Figure 9: Resolution comparison of the J/ψ and Υ invariant mass peaks from ATLAS and CMS Open data

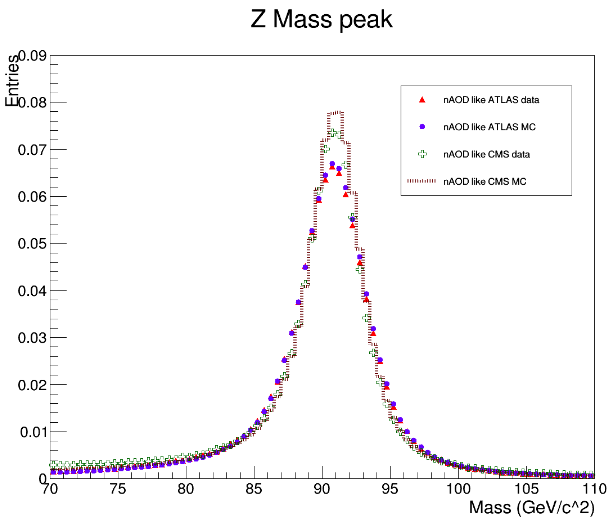

Figure 10: Comparison of the Z invariant mass peaks from ATLAS (nanoAOD-like) and CMS (nanoAOD-like) Data and Monte-Carlo

From Figure 9, it is evident that the J/ψ peak is narrower for CMS compared to ATLAS. Similarly, for Υ, the peaks are not well resolved in the case of ATLAS. Hence, we can conclude that, CMS has better resolution than ATLAS in the low mass region, at least for these cuts ($p_{T\mu_1} > 25GeV$ and $p_{T\mu_1} > 5GeV$). From Figure 10, by looking at the dimuon invariant mass distribution in the region 70 GeV/c^2 to 110 GeV/c^2 , it can be inferred that the resolution for CMS and ATLAS are almost the same in the high mass region.

10 Conclusion

The validation of the 2010 Muonia dataset, which is one among the thirteen primary datasets available in the CMS Open data portal, is successfully finalized and is ready to go to the CMS open data portal. Results obtained from Validation and analysis are in agreement with the ones published by the CMS collaboration.

A 2012 ATLAS open dataset is successfully converted to nanoAOD-like format and compared with dimuon invariant mass spectra obtained from CMS. We have also compared the resolutions of CMS and ATLAS in the high and low mass regions and observed that, though in the high mass realms, the resolutions of CMS and ATLAS are almost the same, ATLAS resolution in the low mass region is not in pace with that of CMS, at least for these cuts.

11 Acknowledgement

I express my sincere thanks to my supervisor Achim Geiser, for his continuous support and invaluable guidance throughout the project. I owe a special thanks to Nur Zulaiha Jomhari for her great support and encouragement. I also thank the summer student organizers, for giving me an opportunity to do summer internship as well as for providing a conducive and enjoyable environment for research.

References

[1] **CMS Collaboration. Performance of CMS muon reconstruction in pp collision events at** $\sqrt{s} = 7TeV$ **(DOI: 10.1088/1748-0221/7/10/P10002)**.

[2] **CMS Collaboration. Upsilon production cross section in pp collisions at** $\sqrt{s} = 7TeV$ **(arXiv:1012.5545.)**

[3] **CERN. CMS Open Data (http://opendata.cern.ch/research/CMS)**

[4] **CERN. ATLAS Open Data (http://opendata.atlas.cern/)**

[5] **B. Sheeran. Validation of CMS open data in dimuon systems and first results from** $D^{*\pm}$ **reconstruction. Summer Student Report 2015,DESY, Hamburg**

[6] **I. Dutta. Validation of CMS 2010 Open Data .Summer Report 2015. DESY, Hamburg July 2015.**