J/Ψ and Z Analysis with CMS and ATLAS Datasets - ATLAS Conversion to CMS

DESY Summer Student Programme, 2021

Leonardo Olivi
University of Turin, Italy

Supervisor
Achim Geiser
Co-Supervisor
Yewon Yang

September 8, 2021

Abstract

The aim of the project was to study the kinematic distributions of the J/Ψ and Z from different datasets: ATLAS 2012, CMS 2010, CMS 2011 and CMS 2012. In order to do that with the same code a common structure for datasets is needed, so ATLAS samples have been converted and validated. The results showed clear J/Ψ, Z and also $\Psi^{'}$ peaks with low background for all datasets. The comparisons between datasets showed different normalizations but similar distributions. Therefore, a performance comparison was done to obtain the relative resolution of CMS 2011 and ATLAS for the transverse (d0) and the longitudinal (z0) impact parameters obtaining $\sigma^{d0}_{CMS} = 0.6 \cdot \sigma^{d0}_{ATLAS}$, $\sigma^{z0}_{CMS} = 0.3 \cdot \sigma^{z0}_{ATLAS}$.

Contents

1 Introduction **1**

2 NanoAODplus Format **1**

3 ATLAS 2012 - NanoAODplus - Conversion **3**
3.1 Validation 5

4 ATLAS Open Data - 2012 **9**

5 CMS Double Muon - 2011 **14**

6 CMS Double Muon - 2012 **16**

7 Run 1 - Comparisons **18**

8 Conclusions **20**

9 Appendix **21**
9.1 CMS MuOnia - 2010 21
9.2 CMS MuOnia - 2011 22
9.3 CMS MuOnia - 2012 23
9.4 ATLAS 2012 - New vs Old 24
9.5 Next Steps 25
9.6 Code 27

1 Introduction

The core of the project is to plot kinematic distributions for J/Ψ and Z with a common analyis code for ATLAS 2012, CMS 2010, CMS 2011 and CMS 2012. In order to do that a similar format for the input is needeed.
CMS Run 1 data was taken in AOD (Analysis Object Data) format which needs a CMSSW environment to be read. Since CMS Run 2 the miniAOD and nanoAOD format were introduced, in which data is stored in flat ntuples and can be read without the CMSSW environment. The n-tuples are faster, smaller and can still cover up to 50 % of the analysis. There is an ongoing project [1] which is converting CMS Run 1 samples in nanoAODplus format which I used as input files. The "plus" stands for additional and useful variables which are not in the official nanoAODv8 list [2].
ATLAS and ZEUS conversion, on the other hand, was done directly in the summer student project. In particular I took care of the ATLAS 2012 Open Data [3] conversion, while Raphael Schwenzer (B12) took care of the ZEUS conversion and Aritra Bal (B11) took care of the ATLAS 2016 Open Data [4] conversion. Finally, a common code was written to generate kinematic distributions of J/Ψ and Z from all the different datasets.

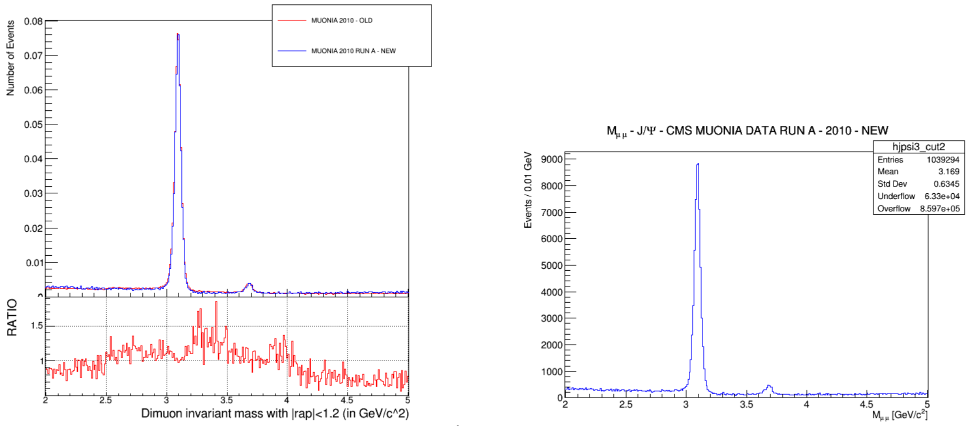
2 NanoAODplus Format

NanoAODplus format was introduced to translate CMS Run 1 data in flat n-tuples which can be read without a CMSSW environment. A new version of the project was completed lately (v0.8) [5], while the previous summer student project (Fabian Stäger) [6] used an older version. Therefore, new and old samples were compared in order to check whether the new ones are working properly.

CMS - MuOnia 2010 MuOnia dataset is a specific dataset used to obtain kinematic distributions for bounds state of charm quarks, called charmonium. In particular, I focused on the J/Ψ to compare the results from the old samples to the new ones.
The set of cuts in table 1 implemented for the plots are the CMS Open Data cuts already used by the summer student Fabian Stäger [6].

J/Ψ
$M_{\mu\mu} \geq 2.6GeV \wedge M_{\mu\mu} \leq 3.5GeV$
$Q_{\mu_1} + Q_{\mu_2} = 0$
$Rapidity_{\mu\mu} \leq 1.2$
$Muon\_isGlobal[m1/m2]$
$Muon\_gnValid[m1/m2] +$
$+ Muon\_nValidMu[m1/m2] \geq 12$
$Muon\_gnPix[m1/m2] \geq 2$
$Muon\_gChi2[m1/m2] \leq 4.0$

Table 1: J/Ψ Cuts

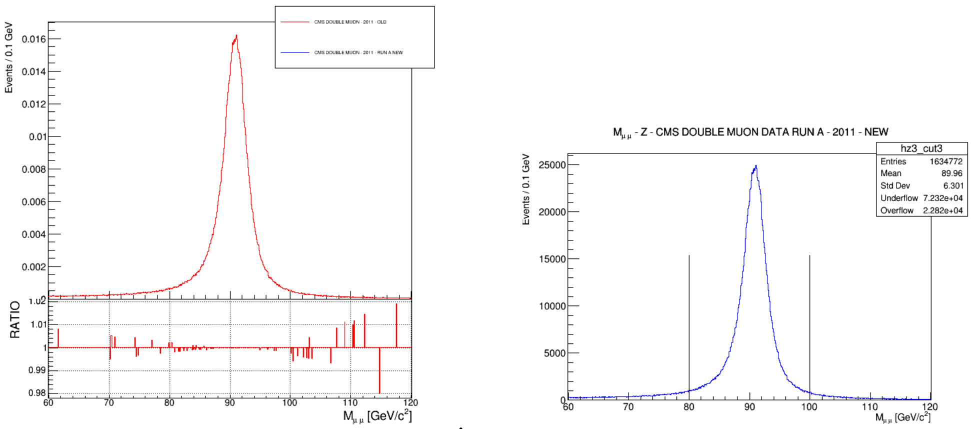


(a) MuOnia 2010 - New (Red) vs Old (Blue), (b) J/Ψ Mass Distribution - New

Figure 1: MuOnia 2010

The distributions have been scaled to 1 since the comparison was done only with Run A of the new MuOnia 2010 samples, which is only a subset of the total one. As can be seen by figure 1a the ratio is almost equal to 1 in the J/Ψ peak. The differences in the tails might come from the fact that we are not comparing the full new samples with the old ones, just Run A.

CMS - Double Muon 2011 The comparison was also done with CMS - Double Muon 2011 (Run A)



(a) CMS 2011 - New (Red) vs Old (Blue), (b) Z Mass Distribution - New (Run A)

Figure 2: CMS Double Muon 2011

The comparison in figure 2a shows that the ratio is almost exactly 1 (2 % deviations). This means that also CMS 2011 new samples have been correctly validated.

3 ATLAS 2012 - NanoAODplus - Conversion

As stated in the introduction, part of the project involved translating and validating the ATLAS 2012 Open Data [3] samples to the nanoAODplus format [1].
The translation was mostly done with official nanoAODv8 variables (not plus) [2], but this was not possible for all of them.

Common Variables - 2012/2016 The tables (2 and 3) which are presented here are for the common variables from both ATLAS 2012 [7] and 2016 [8], even if most of them have different inputs. ATLAS 2016 was done by summer student Aritra Bal, from project B11.

ATLAS	2012 Type	2016 Type	CMS	CMS Type
runNumber	Int_t	Int_t	run	UInt_t
eventNumber	Int_t	Int_t	event	ULong64_t
mcWeight	Float_t	Int_t	Generator_weight	Int_t
lep_n	UInt_t	UInt_t	nMuon, nElectron	UInt_t
lep_charge	Float_t[]	Vector<Float_t>	Muon_charge, Electron_charge	Int_t[]
lep_pt/1000.0	Float_t[]	Vector<Float_t>	Muon_pt, Electron_pt	Float_t[]
lep_eta	Float_t[]	Vector<Float_t>	Muon_eta, Electron_eta	Float_t[]
lep_phi	Float_t[]	Vector<Float_t>	Muon_phi, Electron_phi	Float_t[]
lep_E	Float_t[]	Vector<Float_t>	/	/
lep_z0/10.0	Float_t[]	Vector<Float_t>	Muon_dz, Electron_dz	Float_t[]
lep_d0/10.0	Float_t[]	Vector<Float_t>	Muon_dxy, Electron_dxy	Float_t[]
lep_sd0/10.0	Float_t[]	Vector<Float_t>	Muon_dxyErr, Electron_dxyErr	Float_t[]
lep_charge	Float_t[]	Vector<Float_t>	Muon_charge, Electron_charge	Float_t[]
lep_type	UInt_t[]	Vector<Int_t>	11 = e, 13 = μ, 15 = τ	-
lep_ptcone30/lep_pt	Float_t[]	Vector<Float_t>	Muon_pfRelIso03_chg, Electron_pfRelIso03_chg	Float_t[]
(lep_etcone20/lep_pt) ×9.0/4.0	Float_t[]	Vector<Float_t>	Muon_pfRelIso03_all	Float_t[]
(lep_etcone20/lep_pt) ×9.0/4.0	Float_t[]	Vector<Float_t>	Electron_pfRelIso03_all	Float_t[]
(lep_etcone20/lep_pt) ×16.0/4.0	Float_t[]	Vector<Float_t>	Muon_pfRelIso04_all	Float_t[]
lep_isTightID	/	Vector<Bool_t>	Muon_tightID	Bool_t[]
$\sqrt{\text{lep\_z0}^2 + \text{lep\_d0}^2/10.0}$	Float_t[]	Vector<Float_t>	Muon_ip3d, Electron_ip3d	Float_t[]
lep_n	UInt_t	/	nTau	UInt_t
lep_charge	Float_t[]	/	Tau_charge	Float_t[]
lep_pt/1000.0	Float_t[]	/	Tau_pt	Float_t[]
lep_eta	Float_t[]	/	Tau_eta	Float_t[]
lep_phi	Float_t[]	/	Tau_phi	Float_t[]
lep_z0/10.0	Float_t[]	Vector<Float_t>	Tau_dz	Float_t[]
lep_d0/10.0	Float_t[]	/	Tau_dxy	Float_t[]
tau_n	/	Vector<Float_t>	nTau	Float_t
tau_pt/1000.0	/	Vector<Float_t>	Tau_pt	Float_t
tau_eta	/	Vector<Float_t>	Tau_eta	Float_t
tau_phi	/	Vector<Float_t>	Tau_phi	Float_t
tau_charge	/	Vector<Int_t>	Tau_charge	Int_t

Table 2: ATLAS - Translation

ATLAS	2012 Type	2016 Type	CMS	CMS Type
met_et/1000.0	Float_t	Float_t	Met_sumEt	Float_t
met_phi	Float_t	Float_t	Met_phi	Float_t
jet_n	/	Int_t	nJet	UInt_t[]
alljet_n	Int_t	/	nJet	UInt_t[]
jet_pt/1000	Float_t[]	Vector<Float_t>	Jet_pt	Float_t[]
jet_eta	Float_t[]	Vector<Float_t>	Jet_eta	Float_t[]
jet_phi	Float_t[]	Vector<Float_t>	Jet_phi	Float_t[]
jet_m/1000	Float_t[]	/	Jet_mass	Float_t[]
scaleFactor_PILEUP	Float_t[]	Float_t[]	sf_pileup	Float_t[]
scaleFactor_ELE	Float_t[]	Float_t[]	sf_ele	Float_t[]
scaleFactor_MUON	Float_t[]	Float_t[]	sf_muon	Float_t[]
scaleFactor_BTAG	Float_t[]	Float_t[]	sf_btag	Float_t[]
scaleFactor_TRIGGER	Float_t[]	Float_t[]	sf_trigger	Float_t[]
scaleFactor_JVFSF	Float_t[]	Float_t[]	sf_jvfsf	Float_t[]
scaleFactor_ZVERTEX	Float_t[]	Float_t[]	sf_zvertex	Float_t[]
trigE	Bool_t[]	Bool_t[]	Trig_goodMuTrigger	Bool_t[]
trigM	Bool_t[]	Bool_t[]	Trig_goodETrigger	Bool_t[]
passGRL	Bool_t[]	/	GoodLumisection	Bool_t[]
hasGoodVertex		/	PVtx_isGood = *true*	
hasGoodVertex	Bool_t[]	/	PVtx_isMain = *true*	Bool_t[]
hasGoodVertex		/	PVtx_isValid = *true*	
-	-	-	Muon_isPFcand = *true*	Bool_t[]
-	-	-	Muon_isGlobal = *true*	Bool_t[]
-	-	-	Muon_softID = *true*	Bool_t[]
-	-	-	Muon_isTracker = *true*	Bool_t[]
-	-	-	Electron_isPFcand = *true*	Bool_t[]
-	-	-	Muon_softID = *true*	Bool_t[]
if |lep_eta|<1.37	-	-	Electron_isEB = *true*	Bool_t[]
if |lep_eta|>1.52	-	-	Electron_isEE = *true*	Bool_t[]
-	-	-	Muon_sip3d = 0.0	Float_t[]
-	-	-	Electron_sip3d = 0.0	Float_t[]
-	-	-	Electron_lostHits = 0	UChar_t[]
-	-	-	Muon_mass = 0.105658 GeV/c^2	Float_t[]
-	-	-	Electron_mass = 0.0005109989 GeV/c^2	Float_t[]
-	-	-	Tau_mass = 0.177686 GeV/c^2	Float_t[]

Table 3: ATLAS - Translation

Remark: lep_d0 is actually lep_trackd0pvunbiased, lep_sd0 is actually lep_tracksigd0pvunbiased.

Unique Variables - 2012 ATLAS 2012 Open Data [3] samples have also some unique variables (4), that are not present in 2016.

ATLAS 2012	TYPE	CMS	TYPE
pvxp_n	Int_t	PV_npvs	Int_t
vxp_z	Float_t	PV_z	Float_t
lep_truthMatched	Bool_t []	Muon_simIdx	Int_t []
lep_trigMatched	Bool_t []	Muon_trigIdx	Int_t []

Table 4: ATLAS 2012 - Translation

Still Missing The following variables (5)are the ones not yet implemented during my project and remained for next summer student project.

ATLAS	TYPE	CMS	TYPE
channelNumber	*Int_t*	?	?
lep_flag	*Int_t*[ ]	?	?
jet_jvf	*Float_t*[ ]	?	?
jet_flag	*Int_t*[ ]	?	?
jet_trueflav	*Int_t*[ ]	?	?
jet_truthMatched	*Int_t*[ ]	?	?
jet_SV0	*Float_t*[ ]	?	?
jet_MV1	*Float_t*[ ]	?	?

Table 5: ATLAS 2012 - Still Missing

3.1 Validation

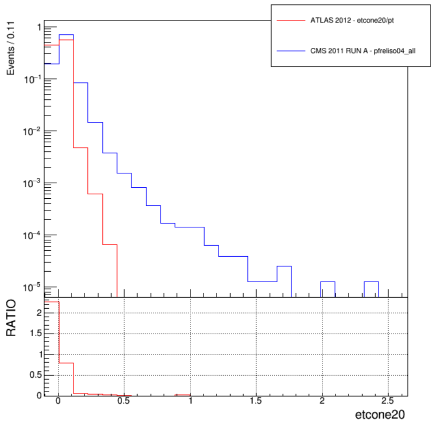
For some variables the definition was enough to understand that the conversion was correct (e.g. lep_pT [1], lep_eta, etc.).
On the other hand, some variables needed to be validated.

Isolation Variables The definitions of the isolation variables for ATLAS [7] are:

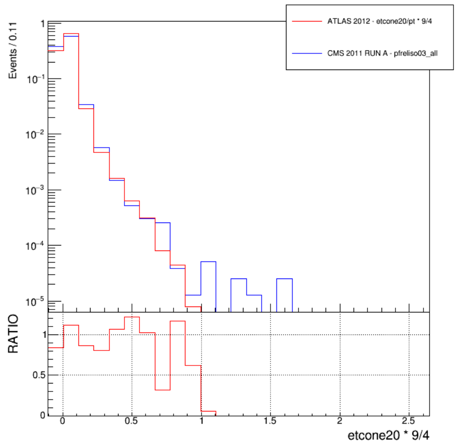
- **ptcone30**: the scalar sum of pT's not including pT of lepton itself in a cone of radius 0.3 cm around the lepton
- **etcone20**: the scalar sum of eT's not including eT of lepton itself in a cone of radius 0.2 cm around the lepton

The corresponding CMS variables are **Muon_pfRelIso03_chg** for **ptcone30** and **Muon_pfRelIso03_all**, **Muon_pfRelIso04_all** for **etcone20**.
The first difference is the fact that the ATLAS variable needs to be divided by pT [9]. Then, since the radius of the cone isn't always the same, a scaling factor equal to the ratio of the areas of the cone was applied. These are the results:

[1] The 1000 factor comes from the fact that ATLAS is in MeV, and CMS is in GeV

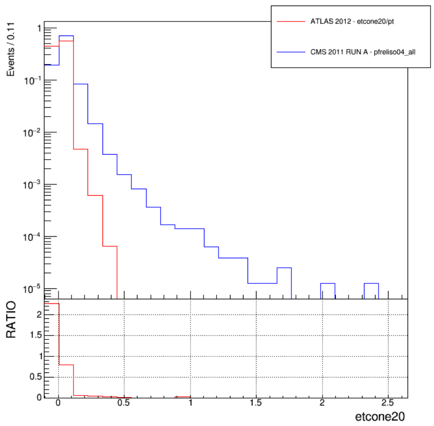


(a) Without scaling factor

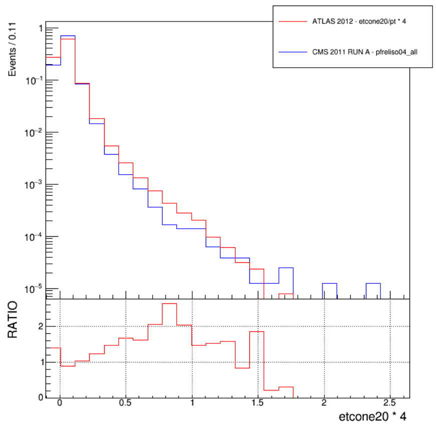


(b) With scaling factor of 9/4

Figure 3: ATLAS 2012 (red), CMS 2011 pfRelIso03_all (blue)



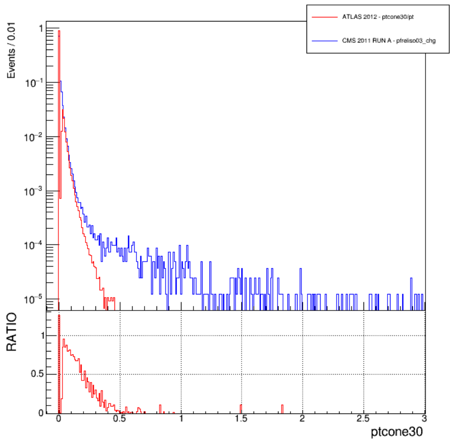
(a) Without scaling factor



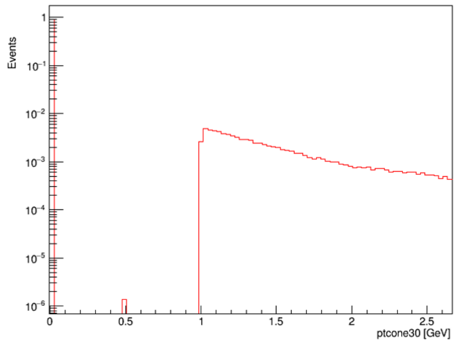
(b) With scaling factor of 16/4

Figure 4: ATLAS 2012 (red), CMS 2011 pfRelIso04_all (blue)

These results show that the translation with the scaling factors has been validated for pfRelIso03_all and pfRelIso04_all. The results are a bit different for pfRelIso03_chg.



(a) ATLAS 2012 (red), CMS 2011 pfRelIso04_all (blue)

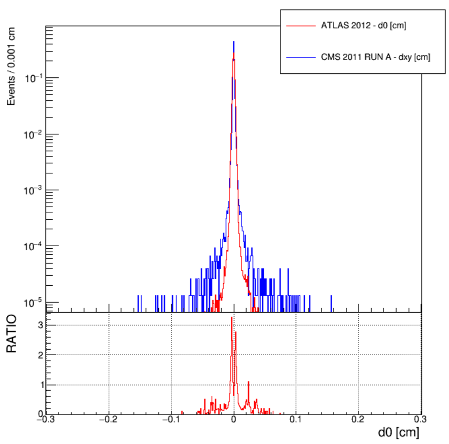


(b) ptcone30 (not divided) - Cut @ 1 GeV

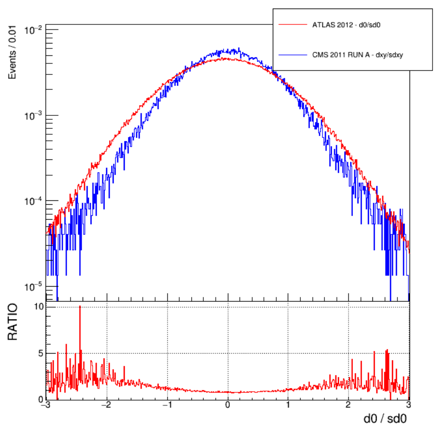
Figure 5: ptcone30 - Validation

Comparison shows that for this variable the distributions are a bit different, this is probably due to fact that ptcone30 has a cut at 1 GeV (5b) which is causing the difference, but the conversion is still correct.

Impact Parameter ATLAS variables are stored in mm while the CMS ones are in cm ([9]). This means that there is a factor of 10 for the impact parameter variables.
For ATLAS the transverse impact parameter is called **lep_trackd0pvunbiased**, its uncertainty is **lep_tracksigd0pvunbiased** and the longitudinal one is **lep_z0**.
The CMS variables are **Muon_ip3d** for the 3D impact parameter (it doesn't exist for ATLAS), **Muon_dxy** for the transverse one, **Muon_dxyErr** is the respective uncertainty and finally **Muon_dz** for the longitudinal one.

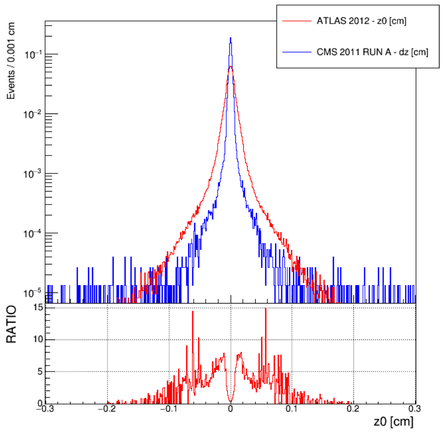


(a) Transverse Impact Parameter (cm)

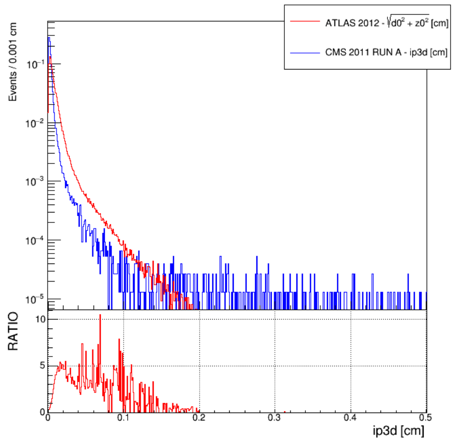


(b) Sigma of Transverse I.P., plots refer to $d0/sd0$

Figure 6: ATLAS 2012 (red), CMS 2011 (blue)



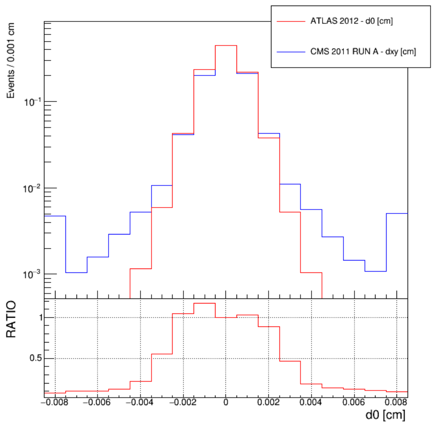
(a) Longitudinal Impact Parameter (cm)



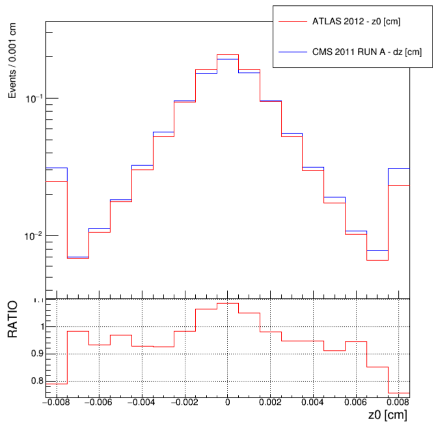
(b) Impact Parameter 3D, ATLAS $= \sqrt{d0^2 + z0^2}/10$.

Figure 7: ATLAS 2012 (red), CMS 2011 (blue)

Comparisons show that results are not exactly the same for ATLAS and CMS. This might not be caused by a fault in the conversion but rather in a different resolution between ATLAS and CMS.
Therefore, a suitable scaling factor was applied to the core of the d0 and z0 distribution (ATLAS), such that the ratio could be almost 1 and flat in the core of the distribution. This is an approximate measurement of the relative resolution between CMS and ATLAS. The results are:



(a) Transverse I.P. (cm) - Rescaled (0.6)



(b) Longitudinal I.P. (cm) - Rescaled (0.3)

Figure 8: ATLAS 2012 (red), CMS 2011 (blue)

$$\sigma^{d0}_{CMS} = 0.6 \cdot \sigma^{d0}_{ATLAS} \tag{1}$$

$$\sigma^{z0}_{CMS} = 0.3 \cdot \sigma^{z0}_{ATLAS} \tag{2}$$

This is just an approximation and it should be investigated with gaussian fits (see 9.5)

4 ATLAS Open Data - 2012

After translating the samples, a code was written to plot the kinematic distributions for both the Z and the J/Ψ, reconstructed from double muons.

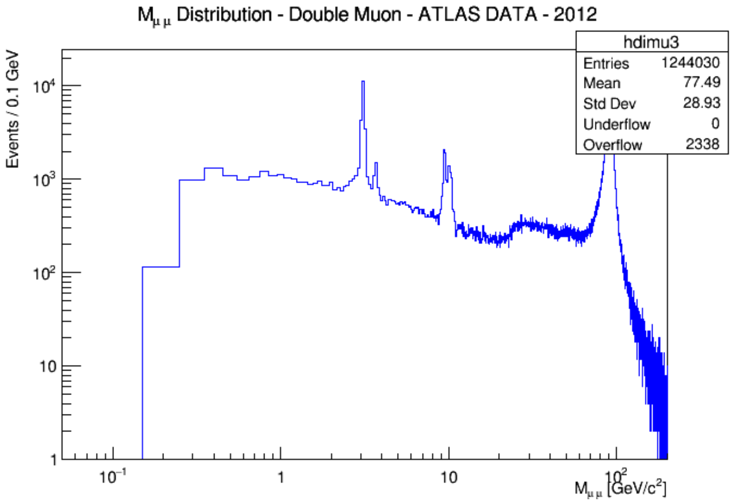


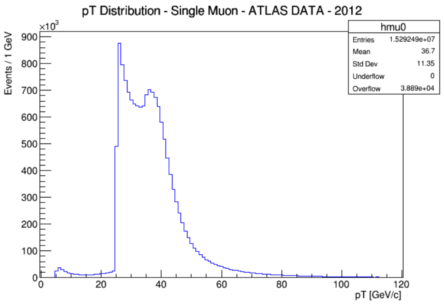
Figure 9: Full Double Muon Spectrum - ATLAS 2012

The full dimuon spectrum is shown in figure 9. Peaks from J/Ψ, Z and also Υ are easily visible without any cuts. This is due to the fact that good quality muons are implemented in Open Data.
There is also a set of preselection cuts ([9]) listed in table 6:

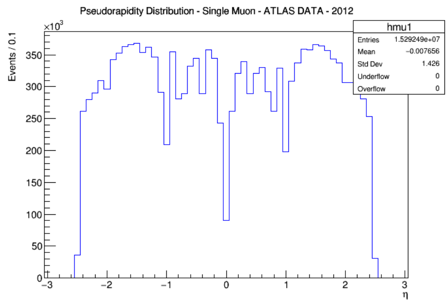
Electron	Muon
pT $\geq$ 5 GeV	pT $\geq$ 5 GeV
$\lvert\eta\rvert \leq 2.5$	$\lvert\eta\rvert \leq 2.5$
$\lvert z_0\rvert \leq 2$ mm	$\lvert z_0\rvert \leq 2$ mm

Table 6: Preselection cuts

This is also verified by looking at the single muon distributions of η and pT



(a) pT - Single Muon Distribution



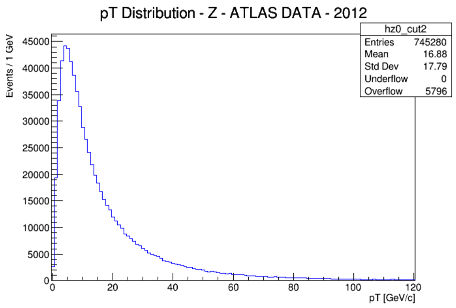
(b) η - Single Muon Distribution

Figure 10: ATLAS 2012 - Single Muon

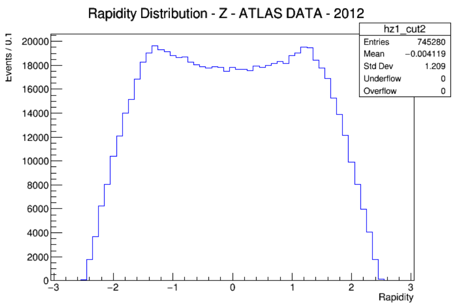
Z Analysis - Data The set of cuts shown in the table 7 was used for the Z ([7])

Z
$Q_{\mu_1} + Q_{\mu_2} = 0$
$Muon\_pt[m1] \geq 25GeV$
$Muon\_pt[m2] \geq 25GeV$
$Muon\_pfRelIso03\_all[m1] \leq 0.15$
$Muon\_pfRelIso03\_all[m2] \leq 0.15$
$Muon\_pfRelIso03\_chg[m1] \leq 0.15$
$Muon\_pfRelIso03\_chg[m2] \leq 0.15$

Table 7: Z - Cuts

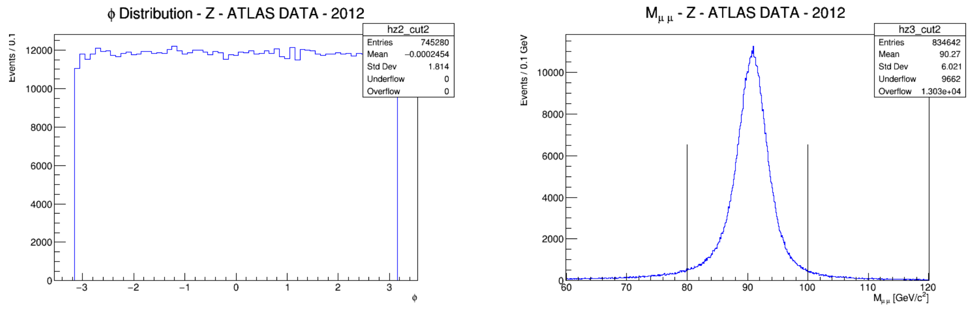


(a) Z - pT Distribution



(b) Z - Rapidity Distribution

Figure 11: ATLAS 2012

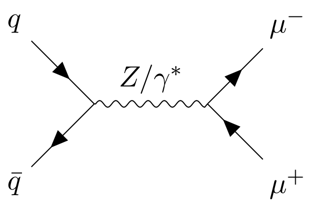


(a) Z - ϕ Distribution (b) Z - Mass Distribution

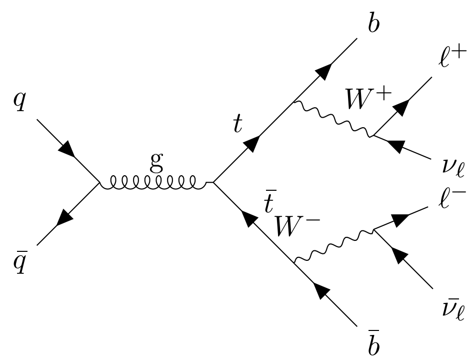
Figure 12: ATLAS 2012

These plots show that the conversion is working fine, at least for data and for the Z. The double muon mass distribution, in particular, shows a peak with very low background.

Z Analysis - Monte Carlo Monte Carlo samples have also been translated with the same mapping table. Some of the Feynman diagrams of the most significant processes for both signal and background are:



(a) Drell - Yan



(b) $t - \bar{t}$

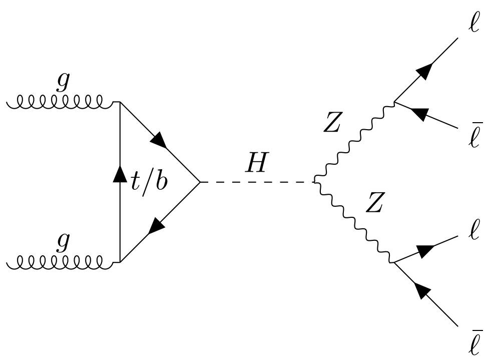


Figure 13: ggHZZ

Once the samples have been generated, they have been scaled to the luminosity times cross sections, multiplied by some scale factors. The results are:

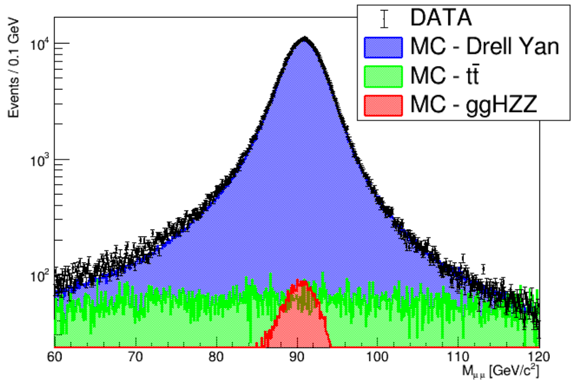


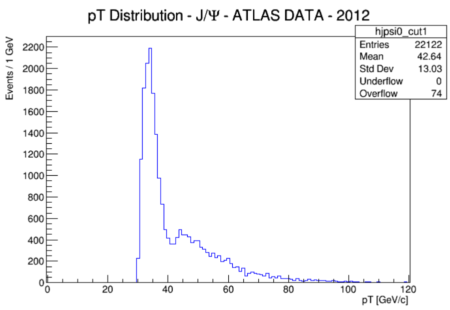
Figure 14: Drell Yan (Blue), $t\bar{t}$ (Green), ggHZZ (Red), Data (Black Dots)

The figure 14 shows that the Drell Yan contribution is the most significant for the signal and that the $t\bar{t}$ is the most significant for the background. The Higgs contribution, on the other hand, is very low, as expected.
Besides, this plot also shows that the conversion is working on simulated datasets.

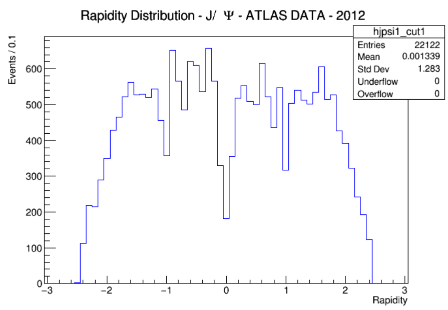
J/Ψ Analysis The cuts shown in the table 8 were used for the J/Ψ:

J/Ψ
pT(μ_1) $\geq$ 25 GeV $\wedge$ pT(μ_2) $\geq$ 5 GeV
$Q_{\mu_1} + Q_{\mu_2} = 0$
$M_{\mu\mu} \leq$ 3.4 GeV $\wedge$ $M_{\mu\mu} \geq$ 2.8 GeV
$Muon\_isGlobal$ = true
$Muon\_softId$ = true

Table 8: J/Ψ - Cuts

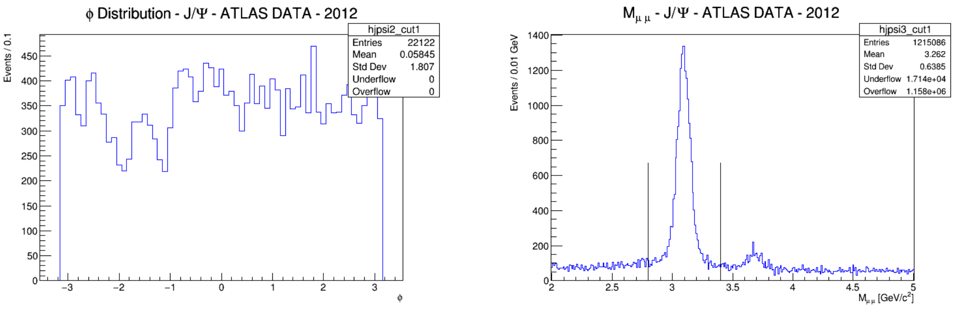


(a) J/Ψ - pT Distribution



(b) J/Ψ - Rapidity Distribution

Figure 15: ATLAS 2012



(a) J/Ψ - ϕ Distribution (b) J/Ψ - Mass Distribution

Figure 16: ATLAS 2012

The mass distribution shows that there is a clear peak for the J/Ψ and also a small but very recognizable peak for the $\Psi^{'}$.

5 CMS Double Muon - 2011

The CMS Run 1 samples used in the analysis are 2011 Run A/B and 2012 Run B/C. In this section the analysis was performed on CMS 2011 Run A. Results for Run B are not shown but they are similar.

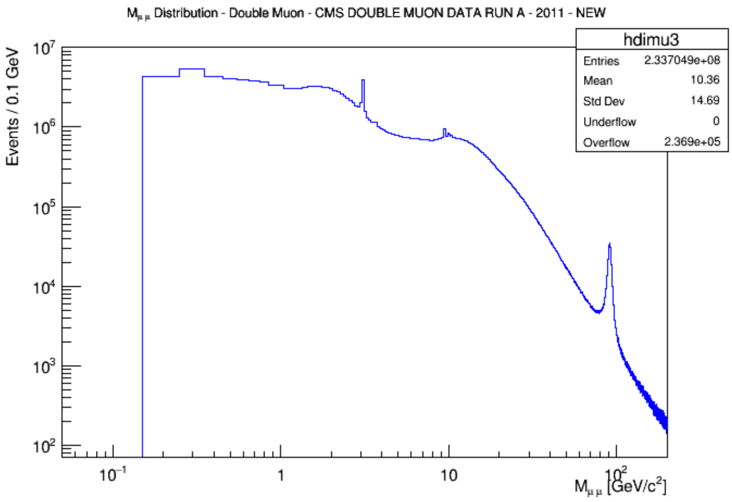


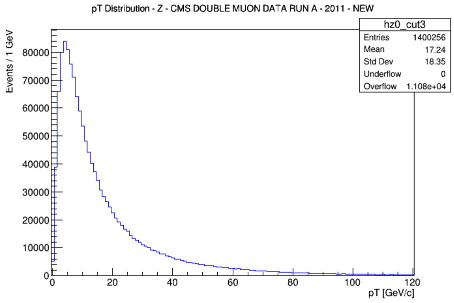
Figure 17: Full Double Muon Spectrum - CMS 2011

The full dimuon spectrum still shows peaks from J/Ψ, Z and also Υ but they are less clear than the ATLAS ones, because CMS samples have also bad quality muons. On the other hand the number of entries is much higher.

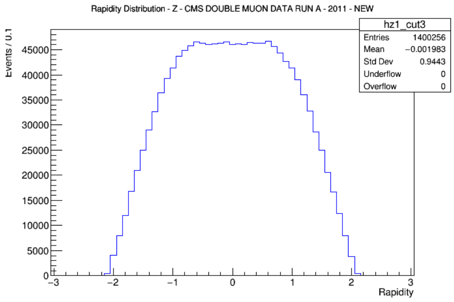
Z Analysis - Data The CMS Open Data set of cuts for the Z is listed in table 9:

Z
$Q_{\mu_1} + Q_{\mu_2} = 0$
$M_{\mu\mu} \leq 3.4$ GeV $\wedge$ $M_{\mu\mu} \geq 2.8$ GeV
$Muon\_isGlobal[m1/m2]$
$Muon\_pfRelIso03\_all[m1/m2] \leq 0.15$
$Muon\_pt[m1/m2] \geq 20GeV/c$
$\|Muon\_eta[m1/m2]\| \leq 2.1$
$\|Muon\_dxy[m1/m2]\| \leq 0.2$

Table 9: Z - Cuts

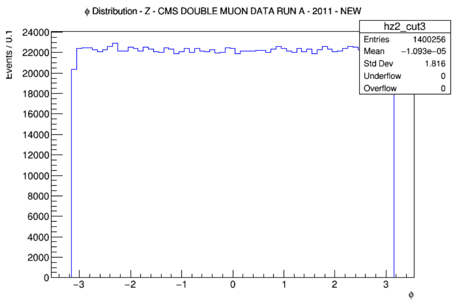


(a) Z - pT Distribution

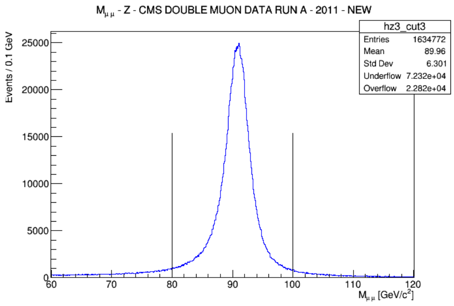


(b) Z - Rapidity Distribution

Figure 18: CMS 2011 - Run A



(a) Z - ϕ Distribution



(b) Z - Mass Distribution

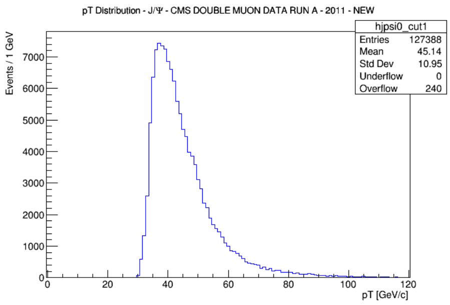
Figure 19: CMS 2011 - Run A

The Z peak is very clear also with CMS 2011. This means that the code is working also on the nanoAODplus samples.

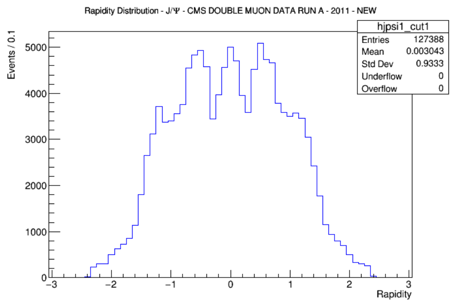
J/Ψ Analysis The CMS Open Data set of cuts for the J/Ψ is in table 10:

J/Ψ
$M_{\mu\mu} \geq 2.6GeV \wedge M_{\mu\mu} \leq 3.5GeV$
$Q_{\mu_1} + Q_{\mu_2} = 0$
$Rapidity_{\mu\mu} \leq 1.2$
$Muon\_isGlobal[m1]$
$Muon\_isGlobal[m2]$

Table 10: J/Ψ - Cuts

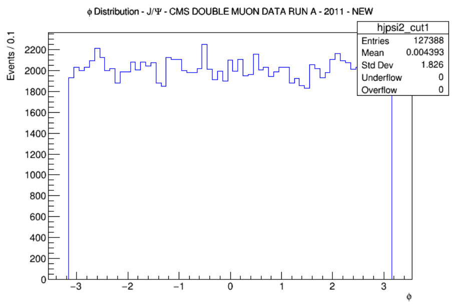


(a) J/Ψ - pT Distribution

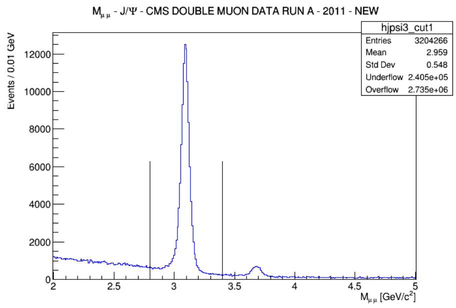


(b) J/Ψ - Rapidity Distribution

Figure 20: CMS 2011 - Run A



(a) J/Ψ - ϕ Distribution



,

(b) J/Ψ - Mass Distribution

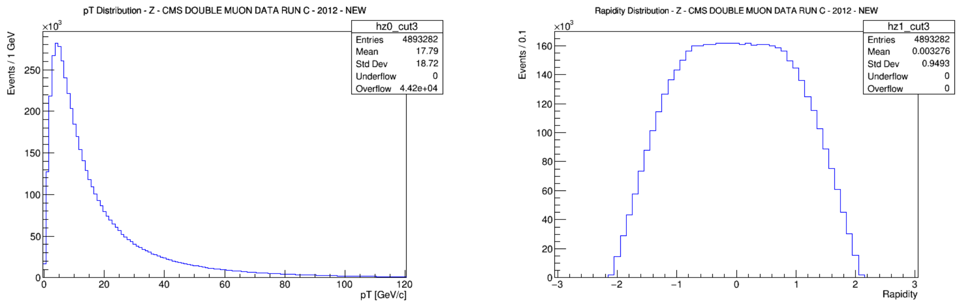
Figure 21: CMS 2011 - Run A

The J/Ψ peak in the mass distribution is very clear, but the background is higher than the ATLAS 2012 one. This is a further proof of the fact that ATLAS only implements good quality muons.

6 CMS Double Muon - 2012

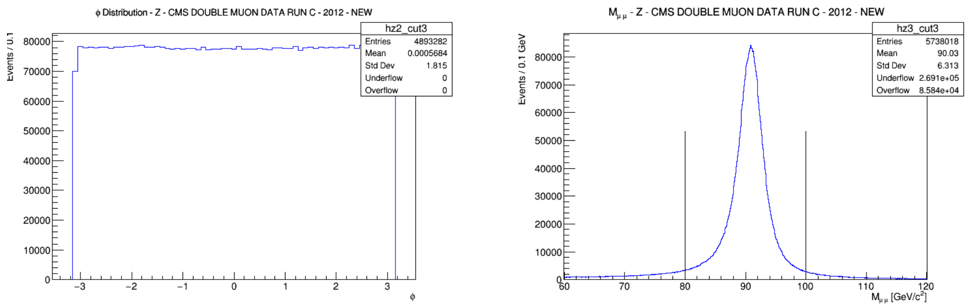
Results for CMS 2012 are for Run C (Similar for Run B)

Z Analysis - Data The same set of cuts of CMS 2011 has been implemented for 2012.



(a) Z - pT Distribution (b) Z - Rapidity Distribution

Figure 22: CMS 2012 - Run C

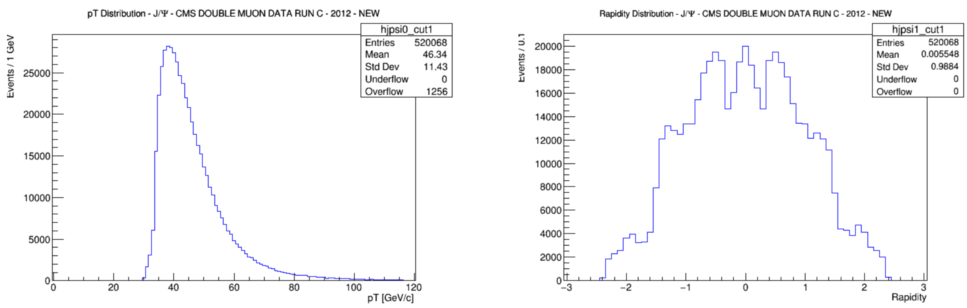


(a) Z - ϕ Distribution (b) Z - Mass Distribution

Figure 23: CMS 2012 - Run C

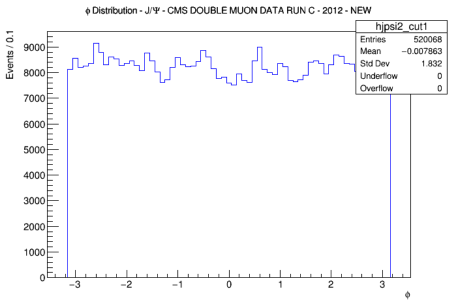
Results are very similar to 2011, as expected. There is a clear Z peak with very low background.

J/Ψ Analysis The J/Ψ was selected with the same cuts as CMS 2011.

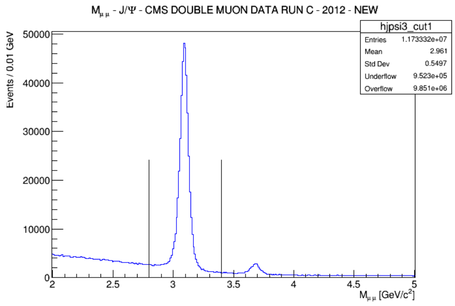


(a) J/Ψ - pT Distribution (b) J/Ψ - Rapidity Distribution

Figure 24: CMS 2012 - Run C



(a) J/Ψ - ϕ Distribution



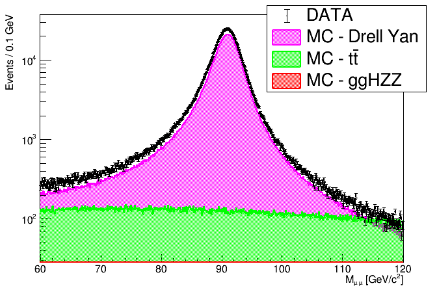
(b) J/Ψ - Mass Distribution

,

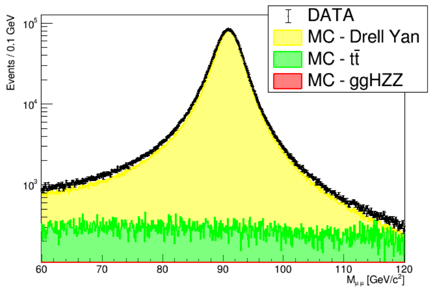
Figure 25: CMS 2012 - Run C

The results show a clear J/Ψ peak, and a small Ψ' peak. The background is higher than the ATLAS 2012 results, showing that also bad quality muons have been implemented in CMS 2012 samples.

Z Analysis - Monte Carlo The Monte Carlo samples have been scaled to the luminosity times the cross sections and multiplied by some scale factors like the ATLAS ones.



(a) CMS 2011 - Run A



(b) CMS 2012 - Run C

Figure 26: Drell Yan (Violet/Yellow), $t\bar{t}$ (Green), ggHZZ (Red), Data (Black Dots)

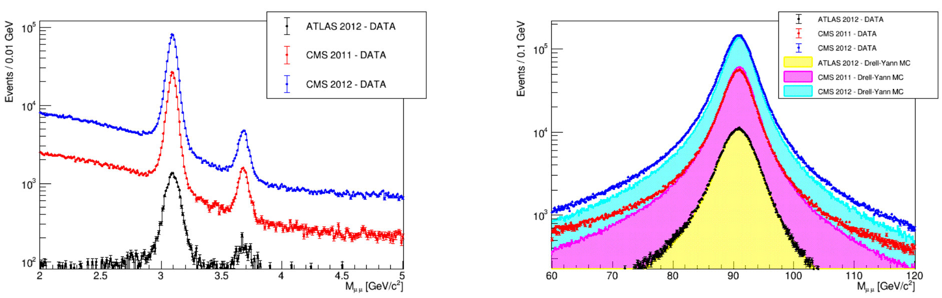
The comparison with simulated datasets shows again that the Drell-Yan process is the most significant for the signal, while the $t\bar{t}$ is the most significant for the background. The Higgs contribution is not even visible, because it has very few entries.

7 Run 1 - Comparisons

Finally, once that it is checked that the code is working on all of the Run 1 samples comparisons between datasets were performed both for the Z and the J/Ψ. The set of cuts used to compare them was the "tighter" one, therefore the one used for ATLAS results listed in table 11:

Z	J/Ψ
$Q_{\mu_1} + Q_{\mu_2} = 0$	$Q_{\mu_1} + Q_{\mu_2} = 0$
$Muon\_pt[m1] \geq 25GeV$	$Muon\_pt[m1] \geq 25GeV$
$Muon\_pt[m2] \geq 25GeV$	$Muon\_pt[m2] \geq 5GeV$
$M_{\mu\mu} \leq 100$ GeV $\wedge$ $M_{\mu\mu} \geq 80$ GeV	$M_{\mu\mu} \leq 3.4$ GeV $\wedge$ $M_{\mu\mu} \geq 2.8$ GeV
$Muon\_pfRelIso03\_all[m1] \leq 0.15$	/
$Muon\_pfRelIso03\_all[m2] \leq 0.15$	/
$Muon\_pfRelIso03\_chg[m1] \leq 0.15$	/
$Muon\_pfRelIso03\_chg[m2] \leq 0.15$	/
/	$Muon\_isGlobal =$ true
/	$Muon\_softId =$ true

Table 11: J/Ψ and Z - Cuts



(a) J/Ψ (b) Z

Figure 27: ATLAS 2012 Data (Black Dots), CMS 2011 Data (Red Dots), CMS 2012 Data (Blue Dots)

	ATLAS 2012	CMS 2011	CMS 2012
$\sqrt{s}$	8 TeV	7 TeV	8 TeV
$\mathcal{L}$	1.00 fb^{-1}	5.1 fb^{-1}	11.6 fb^{-1}

Table 12: Luminosity and $\sqrt{s}$

Both comparisons show that the normalizations are different. This might be caused both from the fact that ATLAS 2012 luminosity ($\mathcal{L} = 1.00\ fb^{-1}$) ([3]) is much lower than the CMS 2012 one (8.6 %, see table 12) and both from the fact that ATLAS 2012 samples have been preselected.
However, distributions for all datasets are similar for both Z and J/Ψ. This is why it is allowed to do performance comparison like the one for the d0 and z0 resolution in section 3.1.

8 Conclusions

In conclusion, the project first involved a validation of some of the new nanoAODplus samples for CMS Run 1 which have been produced. For MuOnia 2010 the results were similar but the ratio was quite different from a flat 1 ideal distribution. On the other hand, for Double Muon 2011 ratio was exactly 1 with some 2 % deviations.
The core of the project was performing the analysis with the same code for CMS and ATLAS 2012 and in order to do that ATLAS converted samples were produced and validated. The results showed clear J/Ψ, Z and also Ψ' peaks with low background for all datasets (ATLAS 2012, CMS 2011, CMS 2012).
The comparison between all datasets also showed that normalizations are different, which might be caused by differences in luminosity and preselection, and that the distributions for both Z and J/Ψ are similar.
Therefore, the setup was used to compare resolutions from CMS 2011 and ATLAS 2012 for the transverse (d0) and the longitudinal (z0) impact parameters. The approximate results are:

$$\sigma^{d0}_{CMS} = 0.6 \cdot \sigma^{d0}_{ATLAS} \tag{3}$$

$$\sigma^{z0}_{CMS} = 0.3 \cdot \sigma^{z0}_{ATLAS} \tag{4}$$

9 Appendix

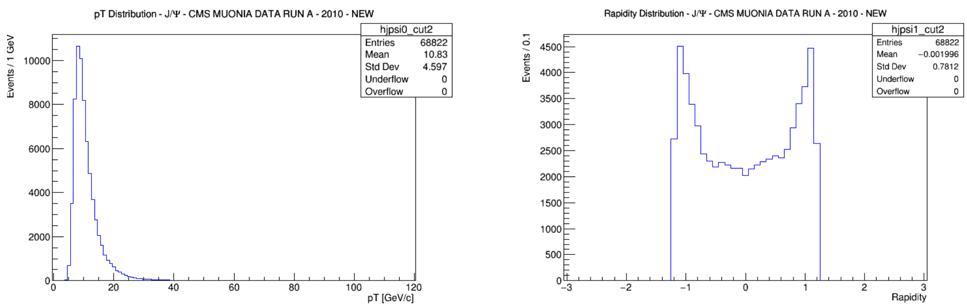
9.1 CMS MuOnia - 2010

MuOnia samples are special datasets used to obtain kinematic distributions for bounds state of charm quarks, called charmonium. In this case I will show results only for Run A (similar results were obtained for Run B).

J/Ψ
$M_{\mu\mu} \geq 2.6GeV \wedge M_{\mu\mu} \leq 3.5GeV$
$Q_{\mu_1} + Q_{\mu_2} = 0$
$Rapidity_{\mu\mu} \leq 1.2$
$Muon\_isGlobal[m1]$
$Muon\_isGlobal[m2]$
$Muon\_gnValid[m1] +$
$+ Muon\_nValidMu[m1] \geq 12$
$Muon\_gnValid[m2] +$
$+ Muon\_nValidMu[m2] \geq 12$
$Muon\_gnPix[m1] \geq 2$
$Muon\_gnPix[m2] \geq 2$
$Muon\_gChi2[m1] \leq 4.0$
$Muon\_gChi2[m2] \leq 4.0$

Table 13: J/Ψ - Cuts

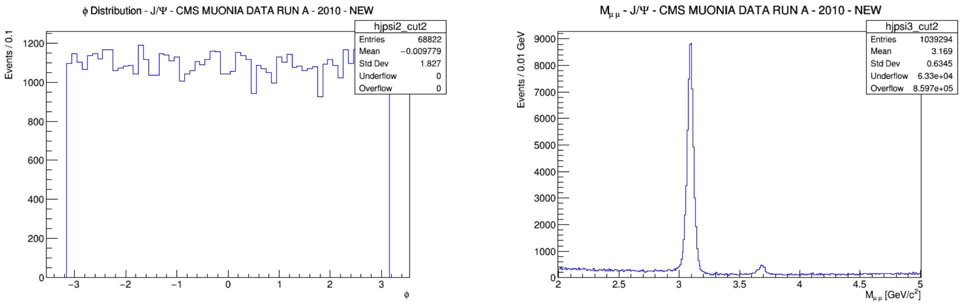
The set of cuts listed in table 13 implements also "plus" variables (gnValid, nValidMu, gnPix, gChi2) ([5])



(a) J/Ψ - pT Distribution

(b) J/Ψ - Rapidity Distribution

Figure 28: MuOnia 2010

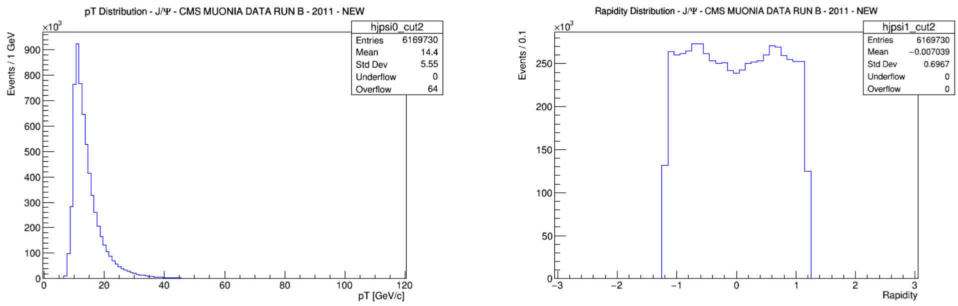


(a) J/Ψ - ϕ Distribution (b) J/Ψ - Mass Distribution

Figure 29: MuOnia 2010

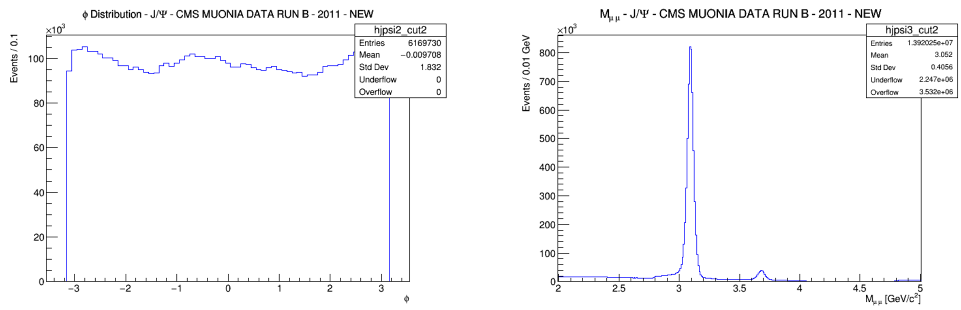
The mass distribution shows a clear peak for the J/Ψ and Ψ', with very low background.

9.2 CMS MuOnia - 2011



(a) J/Ψ - pT Distribution (b) J/Ψ - Rapidity Distribution

Figure 30: MuOnia 2011

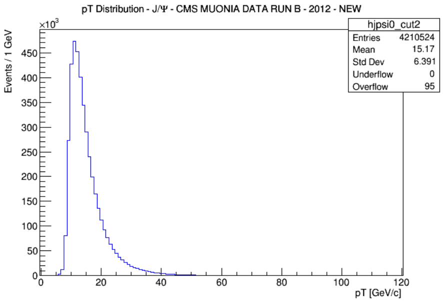


(a) J/Ψ - ϕ Distribution (b) J/Ψ - Mass Distribution

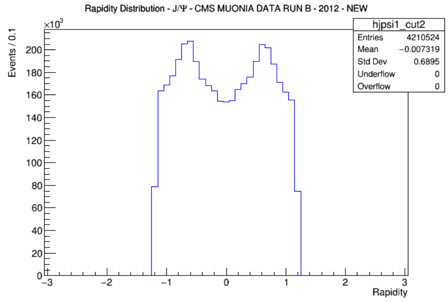
Figure 31: MuOnia 2011

Plots refer to Run B. The mass distribution shows that the background is lower than the CMS Double Muon 2011 results (see 21b). This proves that MuOnia datasets are actually more suited to do J/Ψ analysis, the signal is much more enriched.

9.3 CMS MuOnia - 2012

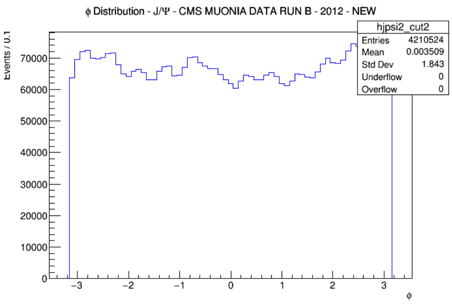


(a) J/Ψ - pT Distribution

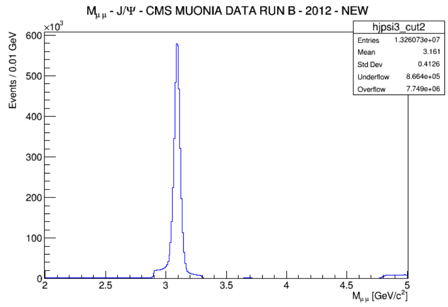


(b) J/Ψ - Rapidity Distribution

Figure 32: MuOnia 2012



(a) J/Ψ - ϕ Distribution



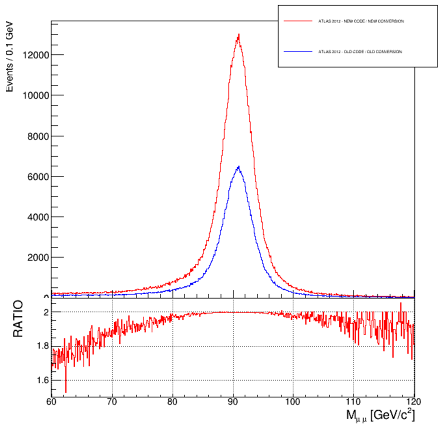
(b) J/Ψ - Mass Distribution

Figure 33: MuOnia 2012

Plots refer to Run B. The mass distribution again clearly shows that the background is lower than the CMS Double Muon 2012 results (see 25b). This is even more noticeable than the CMS 2011 one. On the other hand, there are not entries for the Ψ'.

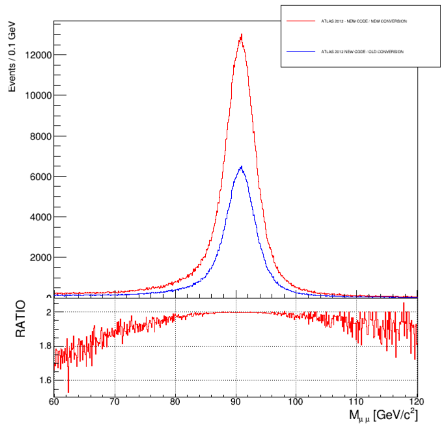
9.4 ATLAS 2012 - New vs Old

ATLAS 2012 conversion was partially already performed by internship student Matthew Snape [10]. To check that the new conversion was compatible with his old results comparisons were made.



,

(a) New Plotting Code running on New Conversion (red) vs Old Plotting Code running on Old Conversion (blue)

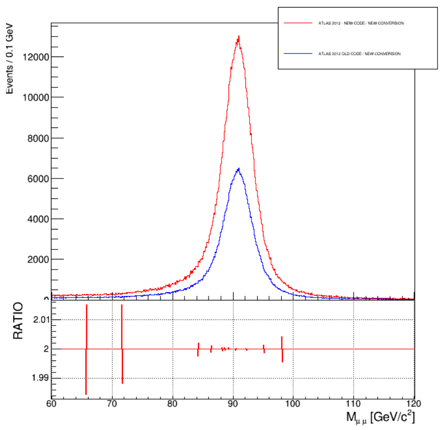


,

(b) New Plotting Code running on New Conversion (red) vs New Plotting Code running on Old Conversion (blue)

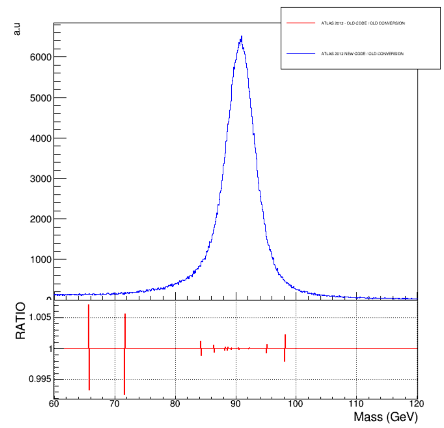
Figure 34: ATLAS Comparison

The plot in figure 34a shows that there is a factor 2 of difference between new results and old results. The one in figure 34b shows that running the same plotting code (the new one) doesn't remove the factor 2. Therefore, the factor 2 might come from the conversion.



,

(a) New Plotting Code running on New Conversion (red) vs Old Plotting Code (*) running on New Conversion (blue)



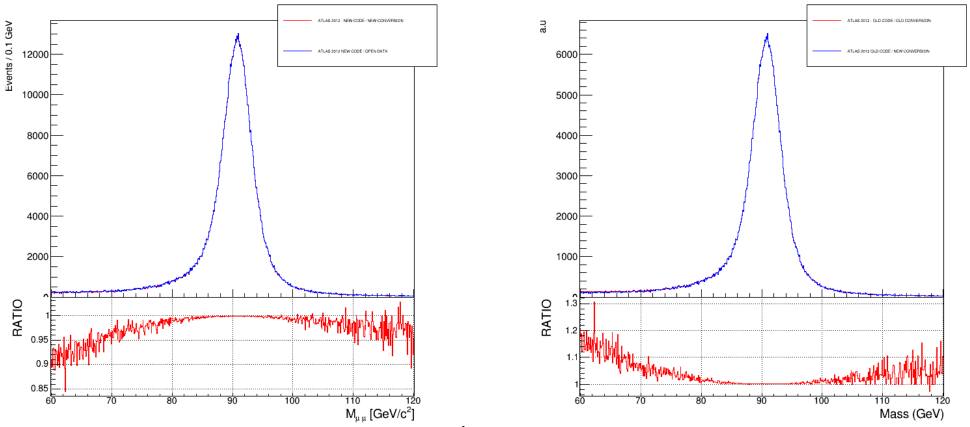
,

(b) Old Plotting Code running on Old Conversion (red) vs New Plotting Code running on Old Conversion (blue)

Figure 35: ATLAS Comparison

The plot in figure 35a shows that running with different code in the same converted

dataset (the new one) doesn't remove the factor 2, which is something unexpected, since figure 34b suggested that the problem might come from the conversion. However, the old plotting code had to be modified with unsigned integers instead of integers to read the number of muons in the new converted samples (which implement unsigned integers).
Besides, plot on figure 35b shows that running different plotting code without any modification on the same input (the old converted samples) gives a factor 1, because the two distributions both have half of the entries (~ 6000). This means that the problem is indeed in the old converted samples.



(a) New Plotting Code running on New Conversion (red) vs New Plotting Code running on Open Data (blue)

(b) Old Plotting Code running on Old Conversion (red) vs Old (*) Plotting Code running on New Conversion (blue)

Figure 36: ATLAS Comparison

The plot in figure 36a shows that running on Open Data and comparing it to the new results gives a factor 1. This means that the new conversion has been validated. However, there is a factor ~ 0.9 in the tails which still needs to be investigated.
Finally, plot in figure 36b shows that both samples have half of the entries (~ 6000). The problem on the red distribution might come from the old conversion, as discussed before, while the one in red might come from the fact that the old plotting has been modified with unsigned integers.

9.5 Next Steps

The purpose of this section is to give the next summer student who will work on this project general guidelines to continue the project.

ATLAS Translation Regarding the translation, few variables (table 14) still need to be mapped:

ATLAS	TYPE	CMS	TYPE
$channelNumber$	Int_t	?	?
lep_flag	$Int\_t[\,]$	?	?
jet_jvf	$Float\_t[\,]$	?	?
jet_flag	$Int\_t[\,]$	?	?
$jet\_trueflav$	$Int\_t[\,]$	?	?
$jet\_truthMatched$	$Int\_t[\,]$	?	?
jet_SV0	$Float\_t[\,]$	?	?
jet_MV1	$Float\_t[\,]$	?	?

Table 14: ATLAS Translation - Still Missing

Besides, jets, electrons and taus have not been yet used in the analysis, even if their variables have been (mostly) translated to CMS. Therefore, a suggestion would be to use them carefully at the beginning.

ATLAS Validation Regarding the validation of the conversion, a suggestion would be to find out what might cause the factor 2 in the old and new comparison. A starting point would be to look at differences with unsigned integers and integers.
Besides, comparison with Open Data (36a) shows that the new translation gives a factor 0.9 in the tails of the Z distribution. This still needs investigation.

Performance Comparison Resolutions were obtained only approximately for longitudinal and transverse impact parameters. This needs to be extended to even more variables, using gaussian fits when it's possible.

9.6 Code

- Code to produce plots: `/afs/desy.de/user/l/lolivi/public/ATLAS_CMS_MAP/cmsanalysisv8.C`
- Code to validate translation: `/afs/desy.de/user/l/lolivi/public/ATLAS_CMS_MAP/ValidationATLAS.C`
- Code to compare histograms: `/afs/desy.de/user/l/lolivi/public/ATLAS_CMS_MAP/CompareHistos.C`
- Code to convert ATLAS 2012 to CMS: `/afs/desy.de/user/l/lolivi/public/ATLAS_CMS_MAP/atlascmsv8_2012`
- Code to produce Run 1 plots: `/afs/desy.de/user/l/lolivi/public/ATLAS_CMS_MAP/Run1.C`
- Code to compare data and MC: `/afs/desy.de/user/l/lolivi/public/ATLAS_CMS_MAP/DataMC.C`

References

[1] CMS NanoAOD Work Book.
https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookNanoAOD

[2] CMS NanoAODv8 Documentation.
https://www.desy.de/~geiser/nano/nanoAODv8.html

[3] Review of ATLAS Open Data 8 TeV datasets, tools and activities.
https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-OREACH-PUB-2018-001

[4] Review of the ATLAS 13 TeV Open Data dataset.
https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-OREACH-PUB-2020-001

[5] CMS NanoAODplus Documentation.
https://www.desy.de/~geiser/nano/nanoAODplusv0.8.html

[6] 2019 Summer Project Report - Fabian Stager.
https://www.desy.de/f/students/2019/reports/Fabian.St%c3%a4ger.pdf

[7] Full list of Branches and Variables: ATLAS 8 TeV Open Data dataset.
http://opendata.atlas.cern/books/current/openatlasdatatools/_book/

[8] Full list of Branches and Variables: ATLAS 13 TeV Open Data dataset.
http://opendata.atlas.cern/release/2020/documentation/datasets/dataset13.html

[9] Review Studies for the ATLAS Open Data Dataset.
https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-OREACH-PUB-2016-001

[10] 2019 Summer Project Report - Matthew Snape.
/afs/desy.de/user/g/geiser/public/summerstudents/Matthew/Final_presentation.pdf