Adıyaman University Journal of Science
https://dergipark.org.tr/en/pub/adyujsci
DergiPark AKADEMİK
ISSN 2147-1630
e-ISSN 2146-586X

An Educational Higgs Study with CMS Open Data

Bora AKGUN[1,2,*]

[1]*Boğaziçi University, Department of Physics, Istanbul, Türkiye*

[2]*Feza Gürsey Center for Physics and Mathematics, Istanbul, Türkiye*

bora.akgun@boun.edu.tr, ORCID: 0000-0001-8888-3562

Received: 01.10.2021 **Accepted:** 03.12.2021 **Published:** 30.12.2021

Abstract

The Compact Muon Solenoid (CMS) Experiment is one of the particle detectors at CERN's Large Hadron Collider (LHC). The CMS Collaboration consists of more than 5000 scientists, engineers, technicians and students from more than 200 institutes and universities from more than 40 countries. The CMS collaboration has a wide physics program and published more than 1000 papers. The CMS has provided open access to 2 PB of its data recorded in proton-proton collisions for research and educational use. This paper gives a brief description on how to analyze the CMS open data and shows a simple Higgs Boson analysis with the data recorded in early 2011.

Keywords: CMS Open Data; Virtual Machine; Higgs.

CMS Açık Veri ile Eğitici Higgs Çalışması

Öz

Kompakt Müon Solenoid (CMS) Deneyi, CERN'in Büyük Hadron Çarpıştırıcısındaki (LHC) parçacık dedektörlerinden biridir. CMS İşbirliği, 40'tan fazla ülkeden 200'den fazla enstitü ve üniversiteden 5000'den fazla bilim insanı, mühendis, teknisyen ve öğrenciden oluşmaktadır. CMS İşbirliği geniş bir fizik programına sahiptir ve 1000'den fazla makale yayınlamıştır. CMS, araştırma ve eğitim amaçlı kullanım için proton-proton çarpışmalarında kaydedilen verilerinin 2 PB'sine açık erişim sağlamıştır. Bu makale, CMS açık verilerinin nasıl

* Corresponding Author DOI: 10.37094/adyujsci.1003333

analiz edileceğine dair kısa bir açıklama sunar ve 2011'in başlarında kaydedilen verilerle basit bir Higgs Boson analizini içerir.

Anahtar Kelimeler: CMS Açık Veri; Sanal Makine; Higgs.

1. Introduction

The Compact Muon Solenoid (CMS) [1] is a general-purpose detector at the Large Hadron Collider (LHC) at CERN. The CMS has a broad physics programme ranging from studying the Standard Model to searching for extra dimensions and particles that could make up dark matter. The CMS collaboration has published more than one thousand papers, one of the most important publications was on the observation of the Higgs Boson [2] in 2012. The CMS Collaboration continues its efforts on studying the properties of the Higgs Boson [3-38] ever since the first observation.

Since 2014 the CMS Collaboration started to release recorded data into open datasets. The CMS data are unique and are of interest to the scientific community as well as to those in education. Several papers [39-47] focusing on new methods and physics results have been published with the CMS open data.

This paper has educational intentions and focuses on CMS Open Data from 2011 (7 TeV) for a study on the Higgs Boson. Section 2 describes how to set up a virtual machine, the datasets and the event selection, Section 3 displays the event topology of the Higgs Boson candidates and Section 4 shows the results.

2. The Datasets and Event Selection

One of the decay channels of the Higgs Boson is $H \rightarrow ZZ \rightarrow 4l$, where the final decay products can be four muons, four electrons or two muons and two electrons. For this study, DoubleMuon [48] and DoubleElectron [49] datasets from RunA of 2011 were used since the events stored in these datasets have at least two high-energy muons or electrons.

A virtual machine[i] is installed and the CMS computing environment is set up[ii] to analyze the CMS Open Data. The Higgs Boson candidates were selected by requiring to have two Z boson candidates in the event. The Z boson candidates selection relied on the presence of a pair of muons or electrons. The muon candidates are selected from muons identified as both Global and Tracker Muon with the χ^2/d.o.f of the global-muon track fit is less than 10 and the number of inner-tracker hits are more than 10. The muon candidates are also required to have transverse momentum (p_T) higher than 5 GeV/c, absolute value of the pseudo-rapidity ($|\eta|$) less than 2.4 and the ratio of the

scalar sum of the P_T of tracks in the inner tracker (ΣP_T) and the transverse energies (ΣE_T) in calorimeter cells (both in the Electromagnetic Calorimeter (ECAL) and Hadronic Calorimeter (HCAL)) within a cone of radius $\Delta R = \sqrt{(\Delta\varphi)^2 + (\Delta\eta)^2} = 0.3$[iii] centered on the direction vector of the candidate particle ($dR_{0.3}$) and the P_T of the muon candidate less than 0.15. The electron candidates are selected from electrons passing loose identification requirements. The electron candidates are also required to have P_T higher than 7 GeV/c, $|\eta|$ less than 2.5 and additional isolation criteria: in the barrel region ($|\eta| < 1.44$) the ratio of the ΣP_T within $dR_{0.3}$ and the P_T of the electron candidate less than 0.09, the ratio of the ΣE_T in the ECAL within $dR_{0.3}$ and the P_T of the electron candidate less than 0.07 and the ratio of the ΣE_T in the HCAL within $dR_{0.3}$ and the P_T of the electron candidate less than 0.10, in the forward region ($|\eta| > 1.57$) the ratio of the ΣP_T within $dR_{0.3}$ and the P_T of the electron candidate less than 0.04, the ratio of the ΣE_T in the ECAL within $dR_{0.3}$ and the P_T of the electron candidate less than 0.05 and the ratio of the ΣE_T in the HCAL within $dR_{0.3}$ and the P_T of the electron candidate less than 0.0025. The region between 1.44 and 1.57 is excluded for the electrons due to low electrons reconstruction efficiency. Events with single Z boson candidate are also studied.

3. Event Displays

The CMS collaboration has developed a tool for visualizing particle collisions. Figure 1 shows event displays of Higgs Boson candidates from 2011 and 2012 decaying to four muons, four electrons or two muons and two electrons. The muon tracks and the muon stations with a muon hit are shown in red, the electron tracks and the energy deposited in ECAL cells are shown in green, and the energy deposited in the HCAL cells are shown in blue in the event displays.

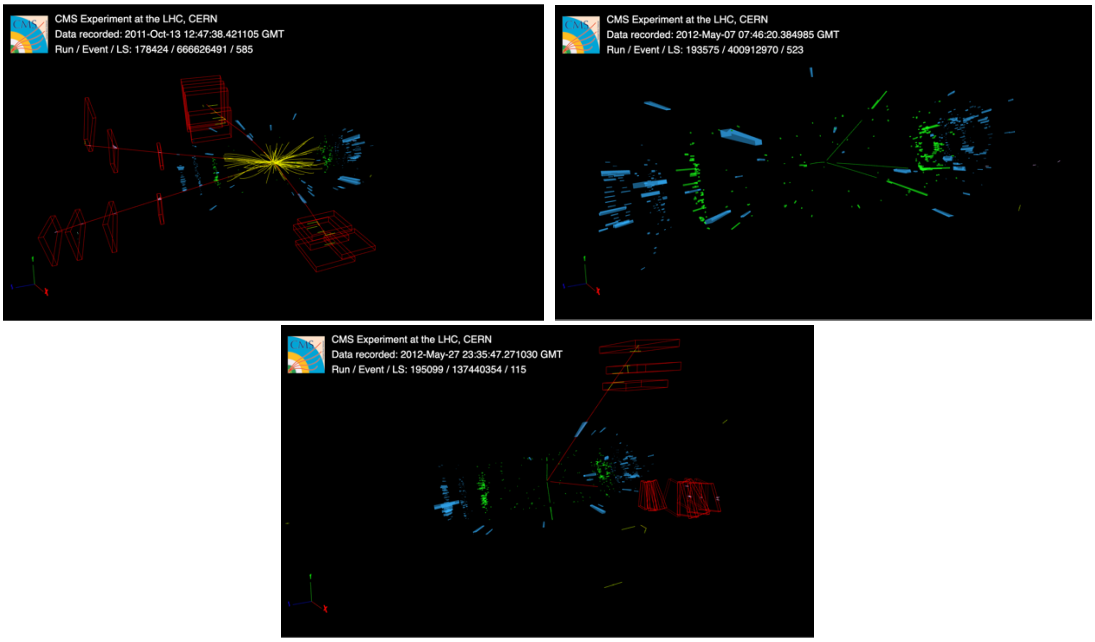

Figure 1: The event displays of the Higgs Boson candidates. The up-left display shows four muons, the up-right display shows four electrons, and the low-middle display shows two muons and two electrons. The particle tracks in the CMS tracker are shown in yellow in the upper-left display, for the other two displays visualization of the particle tracks are disabled to make the electron tracks (in green) visible to the reader

4. Results

One million events per dataset (DoubleMu and DoubleElectron) were analyzed for Z candidates. Figure 2 shows the kinematic properties of the muons used for Z candidate reconstruction. Figure 3 shows the kinematic properties and the mass of the Z candidates reconstructed by the muon pairs. Figure 4 shows the kinematic properties of the electrons used for Z candidate reconstruction. Figure 5 shows the kinematic properties and the mass of the Z candidates reconstructed by the electron pairs. Both the muons and electrons are mostly in the barrel region and their transverse momenta peaks at around a bit higher than 40 GeV/c. The Z candidates are mostly in the forward region. The reconstructed mass of the Z candidates centered at 91 GeV/c^2. The mass of the Z candidates have a narrower mass peak for the muon case than the electron case. This suggests that the muon resolution of the CMS detector is better than its electron resolution. The details of the muon reconstruction performance can be found at [50].

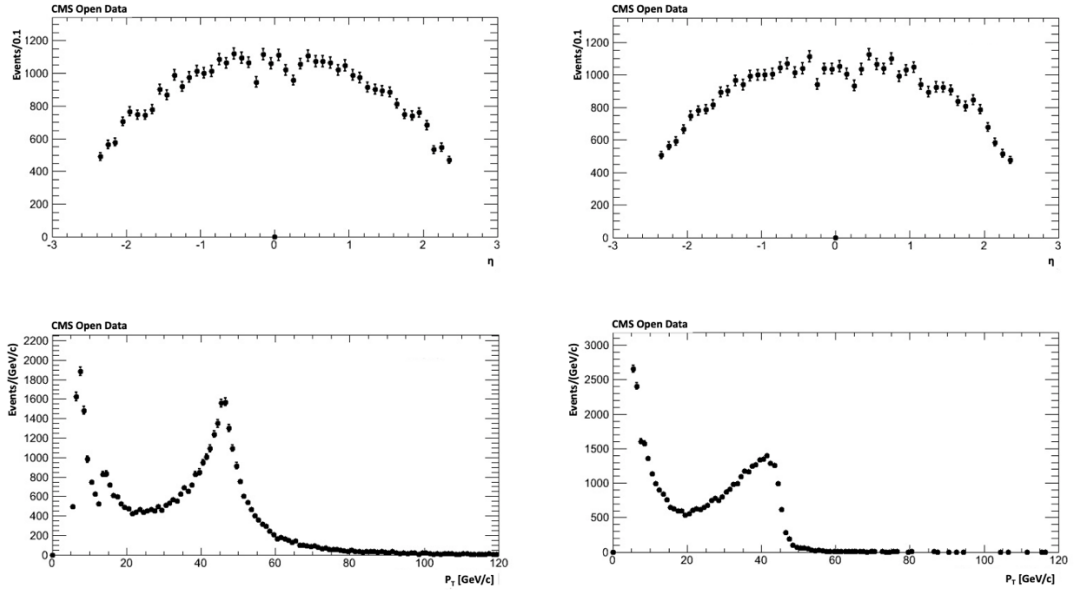

Figure 2: The left column shows the pseudo-rapidity (up-left) and transverse momentum (low-left) distributions of the more energetic muon. The right column shows the pseudo-rapidity (up-right) and transverse momentum (low-right) distributions of the less energetic muon

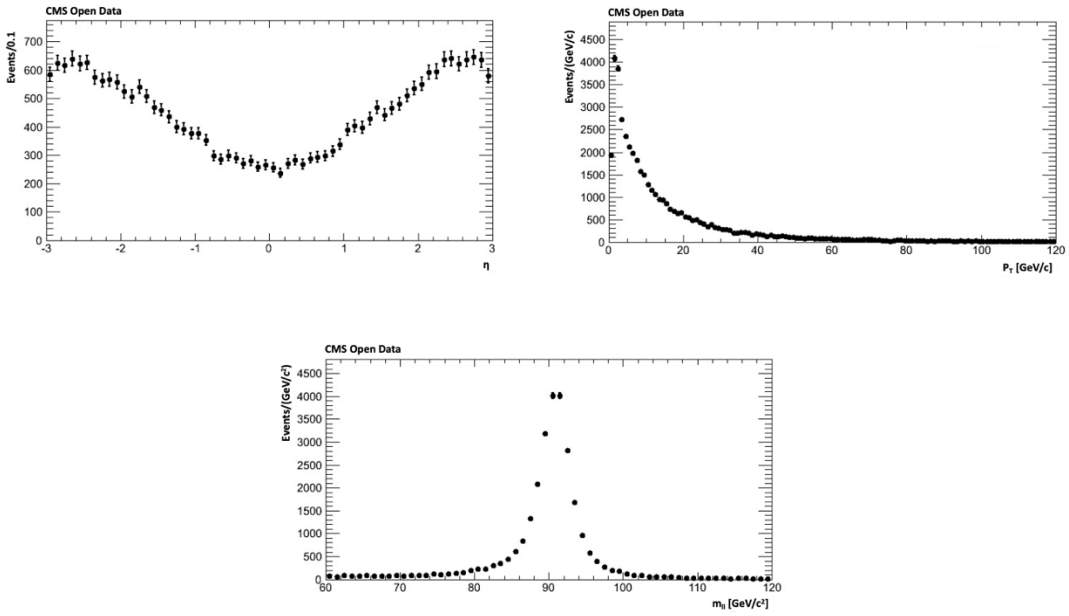

Figure 3: The pseudo-rapidity (up-left) and transverse momentum (up-right) distributions of the Z candidates reconstructed by the muon pairs. The mass distribution (low-middle) of the Z candidates reconstructed by the muon pairs

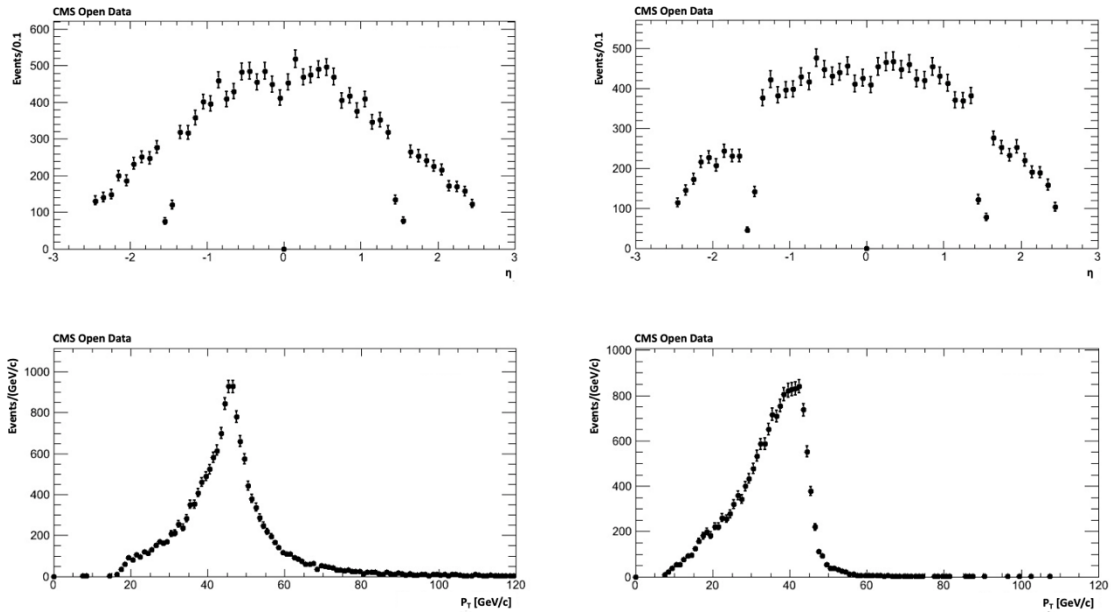

Figure 4: The left column shows the pseudo-rapidity (up-left) and transverse momentum (low-left) distributions of the more energetic electron. The right column shows the pseudo-rapidity (up-right) and transverse momentum (low-right) distributions of the less energetic electron

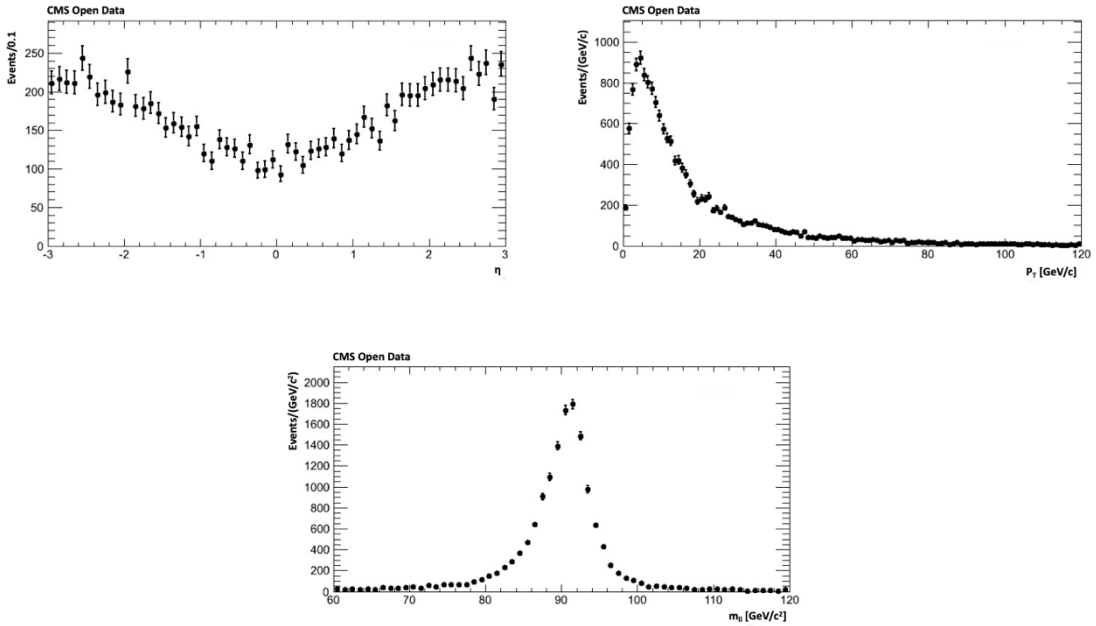

Figure 5: The pseudo-rapidity (up-left) and transverse momentum (up-right) distributions of the Z candidates reconstructed by the electron pairs. The mass distribution (low-middle) of the Z candidates reconstructed by the electron pairs.

All events were analyzed for Higgs Boson candidates; 35,329,695 events in DoubleMu dataset and 3,835,254 events in DoubleElectron dataset. Figure 6 shows the kinematic properties of the Z candidates used for Higgs Boson reconstruction by using the DoubleMu dataset. Figure

7 shows the kinematic properties and the mass of the Higgs Boson candidates by using the DoubleMu dataset. Figure 8 shows the kinematic properties of the Z candidates used for Higgs Boson reconstruction by using the DoubleElectron dataset. Figure 9 shows the kinematic properties and the mass of the Higgs Boson candidates by using the DoubleElectron dataset. The Z candidates are evenly distributed in pseudo-rapidity for both datasets. The Higgs candidates are mostly in the forward region for both datasets. Even though it is not possible to make a conclusion, the mass of the Higgs Boson candidates with the DoubleMu dataset shows a slight excess in the 120-130 GeV/c2 range. A similar excess doesn't exist in the Higgs Boson candidate mass plot with the DoubleElectron dataset.

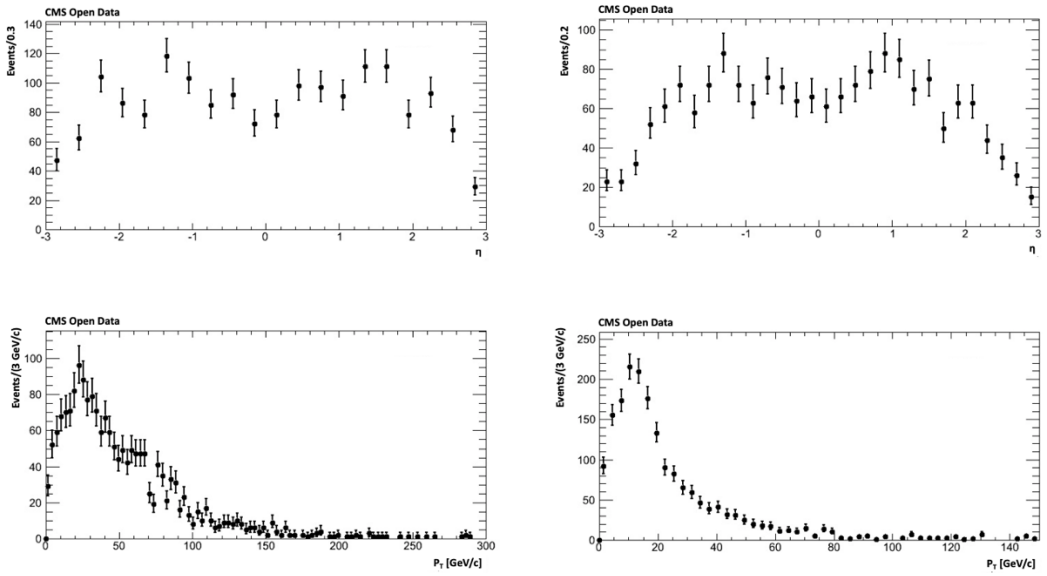

Figure 6: The left column shows the pseudo-rapidity (up-left) and transverse momentum (low-left) distributions of the real Z candidates. The right column shows the pseudo-rapidity (up-right) and transverse momentum (low-right) distributions of the virtual Z candidates. The plots are produced by using the DoubleMu dataset

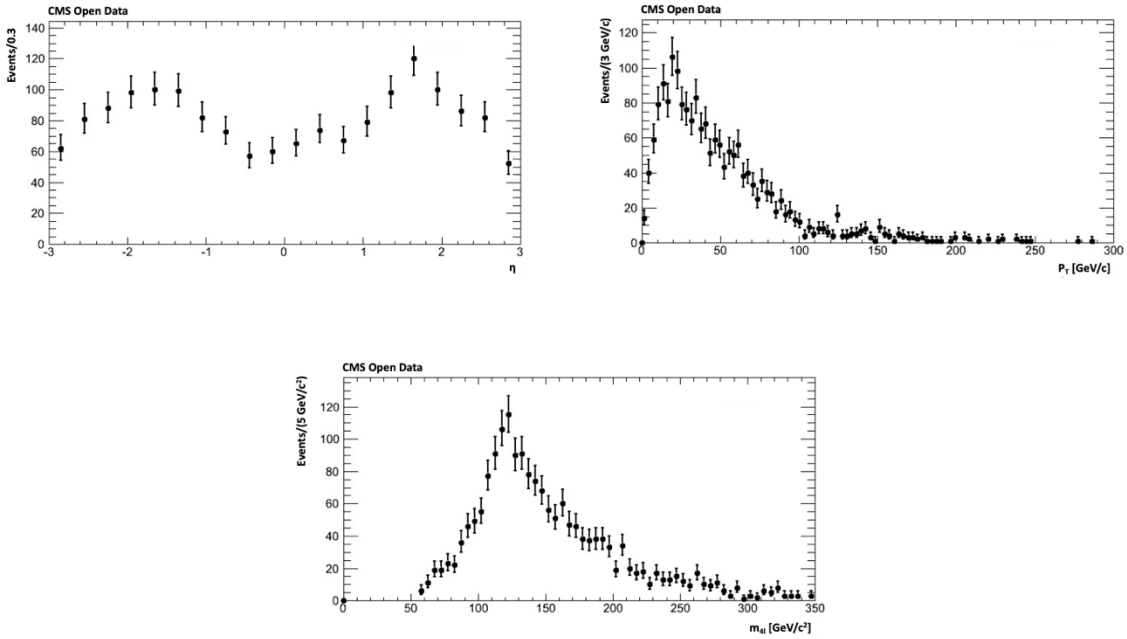

Figure 7: The pseudo-rapidity (up-left), transverse momentum (up-right) and the mass distribution (low-middle) of the Higgs candidates reconstructed by using the DoubleMu dataset

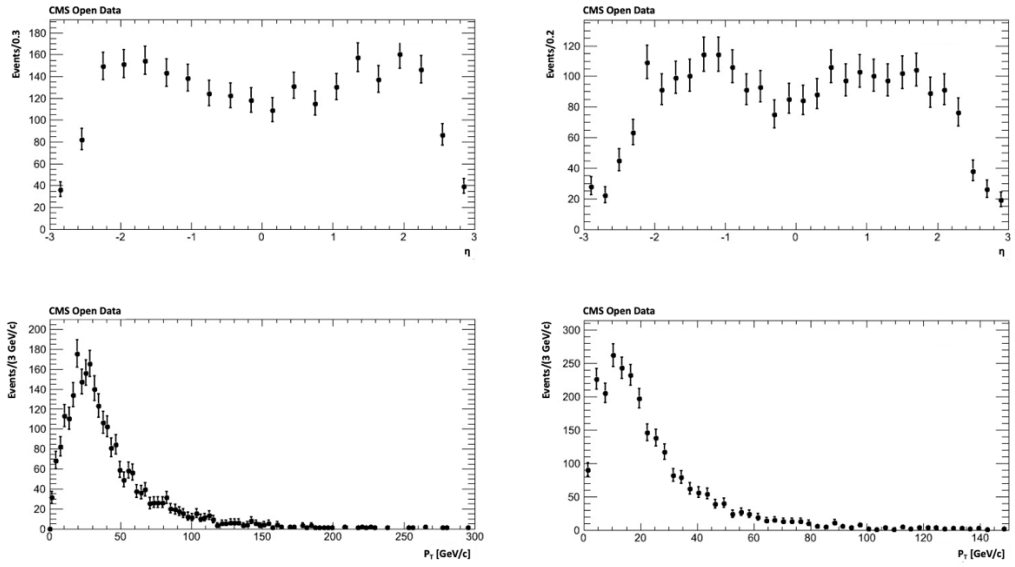

Figure 8: The left column shows the pseudo-rapidity (up-left) and transverse momentum (low-left) distributions of the real Z candidates. The right column shows the pseudo-rapidity (up-right) and transverse momentum (low-right) distributions of the virtual Z candidates. The plots are produced by using the DoubleElectron dataset

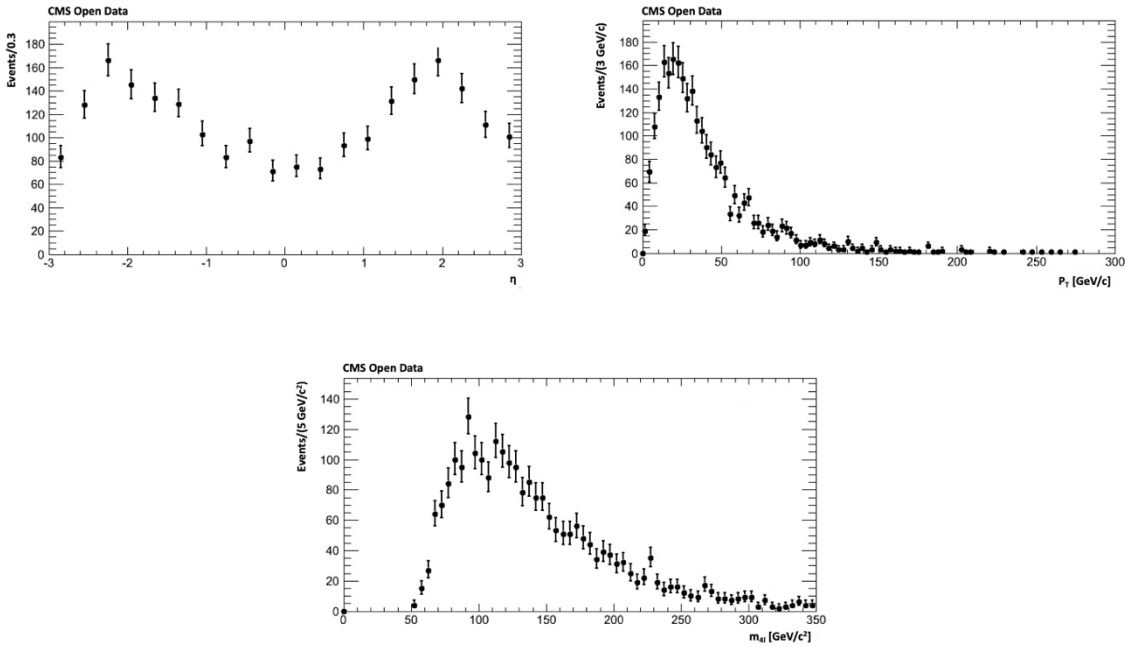

Figure 9: The pseudo-rapidity (up-left), transverse momentum (up-right) and the mass distribution (low-middle) of the Higgs candidates reconstructed by using the DoubleElectron dataset

5. Conclusion

The CMS Open Data collections are released for research and educational use. The data can be analyzed by installing a virtual machine and setting up the CMS computing environment. It is possible to visualize the collision data and make histograms. This paper aims to give pointers to guidelines on how to analyze the data and presents a simple Higgs Boson analysis with the data recorded by CMS experiment in early 2011.

References

[1] The CMS Collaboration, *The CMS experiment at the CERN LHC*, JINST 3 S08004, 2008.

[2] The CMS Collaboration, *Observation of a new boson with mass near 125 GeV in pp collisions at √s = 7 and 8 TeV*, JHEP 06, 081 (2013), doi:10.1007/JHEP06(2013)081.

[3] The CMS Collaboration, *Measurement of higgs boson production and properties in the ww decay channel with leptonic final state*s, JHEP 01, 096 (2014), doi:10.1007/JHEP01(2014)096.

[4] The CMS Collaboration, *Measurement of the properties of a Higgs boson in the four-lepton final state*, Phys. Rev. D 89, no.9, 092007 (2014), doi:10.1103/PhysRevD.89.092007.

[5] The CMS Collaboration, *Evidence for the 125 GeV Higgs boson decaying to a pair of τ leptons*, JHEP 05, 104 (2014), doi:10.1007/JHEP05(2014)104.

[6] The CMS Collaboration, *Evidence for the direct decay of the 125 GeV Higgs boson to fermions*, Nature Phys. 10, 557-560 (2014), doi:10.1038/nphys3005.

[7] The CMS Collaboration, *Observation of the diphoton decay of the Higgs boson and measurement of its properties*, Eur. Phys. J. C 74, no.10, 3076 (2014), doi:10.1140/epjc/s10052-014-3076-z.

[8] The CMS Collaboration, *Combined measurement of the Higgs boson mass in pp collisions at √s = 7 and 8 TeV with the atlas and cms experiments*, Phys. Rev. Lett. 114, 191803 (2015), doi:10.1103/PhysRevLett.114.191803.

[9] The CMS Collaboration, *Limits on the Higgs boson lifetime and width from its decay to four charged leptons*, Phys. Rev. D 92, no.7, 072010 (2015), doi:10.1103/PhysRevD.92.072010.

[10] The CMS Collaboration, *Measurement of differential cross sections for Higgs boson production in the diphoton decay channel in pp collisions at √s = 8 Te*V, Eur. Phys. J. C 76, no.1, 13 (2016), doi:10.1140/epjc/s10052-015-3853-3.

[11] The CMS Collaboration, *Measurement of differential and integrated fiducial cross sections for Higgs boson production in the four-lepton decay channel in pp collisions at √s = 7 and 8 TeV*, JHEP 04, 005 (2016), doi:10.1007/JHEP04(2016)005.

[12] The CMS Collaboration, *Measurement of the transverse momentum spectrum of the Higgs boson produced in pp collisions at s = √8 TeV using H → WW decays*, JHEP 03, 032 (2017), doi:10.1007/JHEP03(2017)032.

[13] The CMS Collaboration, *Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at √s = 7 and 8 TeV*, JHEP 08, 045 (2016), doi:10.1007/JHEP08(2016)045.

[14] The CMS Collaboration, *Measurements of properties of the Higgs boson decaying into the four-lepton final state in pp collisions at √s = 13 TeV*, JHEP 11, 047 (2017), doi:10.1007/JHEP11(2017)047.

[15] The CMS Collaboration, *Observation of the Higgs boson decay to a pair of τ leptons with the CMS detector, Phys. Lett. B 779, 283-316 (2018)*, doi:10.1016/j.physletb.2018.02.004.

[16] The CMS Collaboration, *Inclusive search for a highly boosted Higgs boson decaying to a bottom quark-antiquark pair*, Phys. Rev. Lett. 120, no.7, 071802 (2018), doi:10.1103/PhysRevLett.120.071802.

[17] The CMS Collaboration, *Evidence for the Higgs boson decay to a bottom quark–antiquark pair*, Phys. Lett. B 780, 501-532 (2018), doi:10.1016/j.physletb.2018.02.050.

[18] The CMS Collaboration, *Evidence for associated production of a Higgs boson with a top quark pair in final states with electrons, muons, and hadronically decaying τ leptons at √s = 13 TeV*, JHEP 08, 066 (2018), doi:10.1007/JHEP08(2018)066.

[19] The CMS Collaboration, *Measurements of Higgs boson properties in the diphoton decay channel in proton-proton collisions at √s = 13 TeV*, JHEP 11, 185 (2018), doi:10.1007/JHEP11(2018)185.

[20] The CMS Collaboration, *Measurements of properties of the Higgs boson decaying to a W boson pair in pp collisions at √s = 13 TeV*, Phys. Lett. B 791, 96 (2019), doi:10.1016/j.physletb.2018.12.073.

[21] The CMS Collaboration, *Measurement of inclusive and differential Higgs boson production cross sections in the diphoton decay channel in proton-proton collisions at √s = 13 TeV*, JHEP 01, 183 (2019), doi:10.1007/JHEP01(2019)183.

[22] The CMS Collaboration, *Observation of Higgs boson decay to bottom quarks*, Phys. Rev. Lett. 121, no.12, 121801 (2018), doi:10.1103/PhysRevLett.121.121801.

[23] The CMS Collaboration, *Combined measurements of Higgs boson couplings in proton–proton collisions at √s = 13 TeV*, Eur. Phys. J. C 79, no.5, 421 (2019), doi:10.1140/epjc/s10052-019-6909-y.

[24] The CMS Collaboration, *Combination of searches for Higgs boson pair production in proton-proton collisions at √s = 13 TeV*, Phys. Rev. Lett. 122, no.12, 121803 (2019), doi:10.1103/PhysRevLett.122.121803.

[25] The CMS Collaboration, *Measurement and interpretation of differential cross sections for Higgs boson production at √s = 13 TeV*, Phys. Lett. B 792, 369-396 (2019), doi:10.1016/j.physletb.2019.03.059.

[26] The CMS Collaboration, *Measurements of the Higgs boson width and anomalous HVV couplings from on-shell and off-shell production in the four-lepton final state*, Phys. Rev. D 99, no.11, 112003 (2019), doi:10.1103/PhysRevD.99.112003.

[27] The CMS Collaboration, *Constraints on anomalous HVV couplings from the production of Higgs bosons decaying to τ lepton pairs*, Phys. Rev. D 100, no.11, 112002 (2019), doi:10.1103/PhysRevD.100.112002.

[28] The CMS Collaboration, *Search for Higgs and Z boson decays to J/ψ or Y pairs in the four-muon final state in proton-proton collisions at s=13TeV*, Phys. Lett. B 797, 134811 (2019), doi:10.1016/j.physletb.2019.134811.

[29] The CMS Collaboration, *A measurement of the Higgs boson mass in the diphoton decay channe*l, Phys. Lett. B 805, 135425 (2020), doi:10.1016/j.physletb.2020.135425.

[30] The CMS Collaboration, *Measurements of $t\bar{t}H$ production and the CP structure of the Yukawa interaction between the Higgs boson and top quark in the diphoton decay channel*, Phys. Rev. Lett. 125, no.6, 061801 (2020), doi:10.1103/PhysRevLett.125.061801.

[31] The CMS Collaboration, *Search for resonant pair production of Higgs bosons in the bbZZ channel in proton-proton collisions at √s = 13 TeV*, Phys. Rev. D 102, no.3, 032003 (2020), doi:10.1103/PhysRevD.102.032003.

[32] The CMS Collaboration, *Inclusive search for highly boosted Higgs bosons decaying to bot- tom quark-antiquark pairs in proton-proton collisions at √s = 13 TeV*, JHEP 12, 085 (2020), doi:10.1007/JHEP12(2020)085.

[33] The CMS Collaboration, *Measurement of the inclusive and differential Higgs boson production cross sections in the leptonic WW decay mode at √s = 13 TeV*, JHEP 03, 003 (2021), doi:10.1007/JHEP03(2021)003.

[34] The CMS Collaboration, *Evidence for Higgs boson decay to a pair of muons,* JHEP 01, 148 (2021), doi:10.1007/JHEP01(2021)148.

[35] The CMS Collaboration, *Measurement of the Higgs boson production rate in association with top quarks in final states with electrons, muons, and hadronically decaying tau leptons at √s = 13 TeV*, Eur. Phys. J. C 81, no.4, 378 (2021), doi:10.1140/epjc/s10052-021-09014-x.

[36] The CMS Collaboration, *Search for nonresonant Higgs boson pair production in final states with two bottom quarks and two photons in proton-proton collisions at √s = 13 TeV*, JHEP 03, 257 (2021), doi:10.1007/JHEP03(2021)257.

[37] The CMS Collaboration, *Measurements of production cross sections of the Higgs boson in the four-lepton final state in proton–proton collisions at √s = 13TeV*, Eur. Phys. J. C 81, no.6, 488 (2021), doi:10.1140/epjc/s10052-021-09200-x.

[38] The CMS Collaboration, *Measurements of Higgs boson production cross sections and couplings in the diphoton decay channel at √s = 13 TeV*, JHEP 07, 027 (2021), doi:10.1007/JHEP07(2021)027.

[39] Larkoski, A., Marzani, S., Thaler, J., Tripathee, A. and Xue, W., *Exposing the QCD Splitting Function with CMS Open Data*, Phys. Rev. Lett. 119 (2017) no.13, 132003, doi:10.1103/PhysRevLett.119.132003.

[40] Tripathee, A., Xue, W. , Larkoski, A., Marzani, S. and Thaler, J., *Jet substructure studies with CMS Open Data*, Phys. Rev. D 96 (2017) no.7, 074003, doi:10.1103/PhysRevD.96.074003.

[41] Andrews, M., Paulini, M., Gleyzer, S. and Poczos, B., *End-to-end physics event classification with CMS Open Data: applying image-based deep learning to detector data for the direct classification of collision events at the LHC*, Comput. Softw. Big Sci. 4 (2020) no.1, 6, doi:10.1007/s41781-020-00038-8.

[42] Cesarotti, C., Soreq, Y. , Strassler, M.J., Thaler, J., and Xue, W., *Searching in CMS open data for dimuon resonances with substantial transverse momentum*, Phys. Rev. D 100 (2019) no.1, 015021, doi:10.1103/PhysRevD.100.015021.

[43] Andrews, M., Alison, J., An, S., Bryant, P., Burkle, B., Gleyzer, S., Narain, M. , Paulini, M., Poczos, B., and Usai, E., *End-to-end jet classification of quarks and gluons with the CMS Open Dat*a, Nucl. Instrum. Meth. A 977 (2020), 164304, doi:10.1016/j.nima.2020.164304.

[44] Apyan, A., Cuozzo, W., Klute, M., Saito, Y., Schott, M., and Sintayehu, B., *Opportunities and challenges of Standard Model production cross section measurements in proton-proton collisions at √s =8 TeV using CMS Open Data*, JINST 15 (2020) no.01, P01009, doi:10.1088/1748-0221/15/01/P01009.

[45] Paktinat Mehdiabadi, S., and Fahim, A., *Explicit jet veto as a tool to purify the underlying event in the Drell–Yan process using CMS Open Data*, J. Phys. G 46 (2019) no.9, 095003, doi:10.1088/1361-6471/ab33a9.

[46] Komiske, P.T., Mastandrea, R., Metodiev, E. M., Naik, P. and Thaler, J., *Exploring the Space of Jets with CMS Open Data,* Phys. Rev. D 101 (2020) no.3, 034009, doi:10.1103/PhysRevD.101.034009.

[47] Knapp, O., Cerri, O., Dissertori, G., Nguyen, T.Q., Pierini, M. and Vlimant, J.R., *Adversarially learned anomaly detection on CMS Open Data: re-discovering the top quark,* Eur. Phys. J. Plus 136 (2021) no.2, 236, doi:10.1140/epjp/s13360-021-01109-4.

[48] The CMS Collaboration*, DoubleMu primary dataset in AOD format from RunA of 2011 (/DoubleMu/Run2011A-12Oct2013-v1/AOD*). CERN Open Data Portal, (2016). doi:10.7483/OPENDATA.CMS.RZ34.QR6N.

[49] The CMS Collaboration, *DoubleElectron primary dataset in AOD format from RunA of 2011 (/DoubleElectron/Run2011A-12Oct2013-v1/AOD)*, CERN Open Data Portal, (2016). doi:10.7483/OPENDATA.CMS.MQXP.QNMB.

[50] The CMS Collaboration, *Performance of CMS muon reconstruction in pp collision events at* $\sqrt{s} = 7$ *TeV*, JINST 7 (2012) P10002, doi:10.1088/1748-0221/7/10/P10002.

[i] VirtualBox version 6.1.18 and virtual machine image CMS-OpenData-1.5.3.ova are used. The links for the VirtualBox and the image can be found at: https://www.virtualbox.org/wiki/Download_Old_Builds_6_1 and https://opendata.cern.ch/record/252. The recipe can be found at: https://github.com/cms-opendata-analyses/OutreachExercise2011.

[ii] The recipe can be found at: https://github.com/cms-opendata-analyses/OutreachExercise2011.

[iii] http://opendata.cern.ch/visualise/events/cms.