European Organization for Nuclear Research

CMS collaboration

Constraints on the Higgs self coupling via HH→ $b\bar{b}\mu\mu$

Cern Summer Student 2023 Report

Author

Eliacim Javier Velez

CERN Summer Student 2023

Supervisors

Brunella D'Anzi

University and INFN Bari

Nicola De Filippis

Politecnico and INFN Bari

Geneva, Switzerland

August 2023

Contents

List of Figures

List of Tables

Abbreviations

Abbreviations	Descriptions
SM	Standard Model of Particle Physics
BSM	Beyond Standard Model of Particle Physics
CMS	Compact Muon Solenoid
HH	Di-Higgs
ggF	gluon-gluon fusion
MC	MonteCarlo
VEV	Vacuum Expectation Value
SSB	Spontaneus Symmetry Breaking

Abstract

The discovery of the Higgs boson in 2012 toghether with the measurements of its properties have been the best achievement of the LHC experiments so far but shading light on the Higgs mechanism to give masses to the particles would require the inspection of the exact shape of the Higgs potential and in particular the measurement of the triple and quartic Higgs self-couplings. A direct constraint on the trilinear self-coupling can be extracted from the production of non-resonant Higgs boson pair (HH). For that purpose, our aim is to study the $b\bar{b}\mu\mu$ channel, motivated by the recent evidence of the $H \to \mu\mu$ final state of single Higgs production and the potentially lower background for double Higgs production.

CHAPTER 1

Di-Higgs Production and Data Analysis

1.1 The Standard Model of Particle Physics

The standard model currently stands as the most complete theory of matter, it classifies all known particles and comprises the fundamental interactions between these particles, namely electromagnetic, weak and strong interactions, this statement can be mathematically expressed in terms of symmetries such as $U(1)_Y \times SU(2)_L \times SU(3)_C$, where Y is the hypercharge, L reminds us that the weak interaction only works in left-handed fermions, and C is the color. Under the aforementioned symmetries, the Standard Model (SM) Lagrangian can be written as [1]:

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \bar{\psi}_i i \not{D} \psi_i + h.c. + \psi_i y_{ij} \psi_j \phi + h.c. + \left(D_\mu \phi\right)^\dagger D^\mu \phi - V(\phi) \tag{1.1}$$

The first line in (1.1) is the kinetic term for the gauge theory of electroweak interaction. The second term relates the matter fields ψ and the gauge fields. The third line involves the Yukawa interaction between the matter fields and the scalar field ϕ responsible for giving fermions their masses when the electroweak symmetry breaking happens. The final line is about the kinetic term for the scalar field ϕ and the final part of the last line about the Higgs potential [2].

1.1.1 Higgs Mechanism

The Higgs mechanism stands as a fundamental physical theory that illuminates the intricate process through which elementary particles attain their masses. This phenomenon is achieved by introducing a new field, named Higgs field [1], which is a scalar field that extends throughout the entirety of space. Remarkably, the Higgs field maintains a non-zero ground state value.

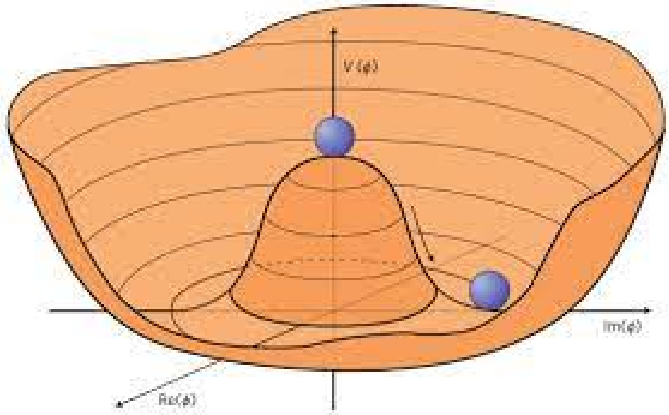


Figure 1.1: A prototypical effective 'Mexican hat' potential that leads to 'spontaneous' symmetry breaking.

Therefore this means that it possesses an inherent "default" value, persisting even in the absence of particles, as they interact with the Higgs field, particles acquire mass via a phenomenon called Spontaneous Symmetry Breaking (SBB). This means, when minimizing the potential illustrated in Figure 1.1, we find a non zero vaccum expectation value (VEV) of the Higgs field, whichever point lying around the bottom of the potential "breaks" the electroweak symmetry $SU_L \times U(1)_Y$. However the symmetry $U(1)_{EM}$ related to electromagnetism is kept intact, this is why the photon is kept massless after SSB. The mathematical form of the Higgs potential, in its conventional state, is expressed as follows [3]:

$$V(\phi) = -\mu^2 \phi^\dagger \phi + \lambda(\phi^\dagger \phi)^2 \tag{1.2}$$

In this context, the term λ assumes the identity of the quartic term, pivotal in bestowing the Higgs field its inherent self-interaction characteristics. The lowest point of the Higgs potential represents the moment when the Higgs field has a value other than zero in its empty state. This particular property is exactly what gives mass to all the basic particles in the world of particles.

Di-Higgs Production

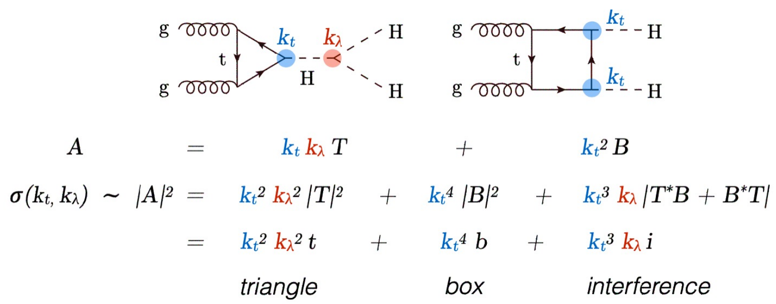
The expanded Higgs potential described by Eq. (1.3)

$$V = \frac{1}{2} m_H^2 H^2 + \lambda_{HHH} v H^3 + \frac{1}{4} \lambda_{HHHH} H^4 \tag{1.3}$$

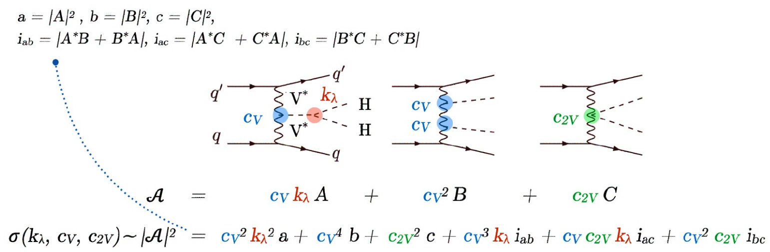
involves in the Standard Model theory $\lambda_{HHHH} = \lambda_{HHH} = \frac{m_H^2}{2v^2}$ where v is a measured parameter and m is determined experimentally. This means that the self- interaction of the Higgs boson even yields its own mass, the cubic and quartic self- interactions of the Higgs boson are directly related to the Higgs potential and are entirely determined by the Higgs mass and the VEV.

The coupling constant λ_{HHH} is related to the di-Higgs production mechanisms, at the CMS collaboration there are two specifically being researched.

- Gluon-gluon fusion (ggF)$gg \rightarrow$ is the dominant di-Higgs production mode, accounting for 90% of di-Higgs events, this is the production mode we focus on for the project. When two gluons fuse can form either a box or a triangle top quark loop, generating two on-shell Higgs bosons, as it can be seen in figure 1.2a. The box diagram and the triangle diagram are interfering destructively, resulting in the very small total cross section of Higgs boson pair production. The SM total cross section is $\sigma = 31.05fb$ at $\sqrt{s} = 13TeV$.
- Vector-boson fusion (VBF) $qq' \rightarrow HH$ this process involves two Higgs radiated from W or Z virtual bosons. A radiated off-shell Higgs can also split into two on-shell Higgs (figure 1.2b). The predicted total cross section is $\sigma = 1.7fb$.



(a) ggF



(b) VBS

1.2 Analysis Tools

The analysis has been carried out by using a framework called Bamboo [4] which is a HEP library for ROOT: RDataFrame by installing it in our lxplus machine and using Microsoft's Visual Studio Code in tandem with the windows Powershell to manipulate Bamboo and see the results. The datasets analyzed are in NanoAOD format [5].

Our reasons of choosing Bamboo as our analysis facility is treefold:

- First of all Bamboo has a lot of documentation which is of extremely utility whenever you have to troubleshoot for issues whenever it is being used.
- Several examples of Di-Higgs analysis have been made using Bamboo, which for our case in particular was of great help because it provided a starting point for our analysis and served us as examples to study so we could understand better our own analysis.
- It's fairly easy to use in comparison to other existing ROOT based analysis facilities.

All the MC samples used in this work have been taken from the CMS Data Agregation System (DAS) with dbs instance prod/global as depicted in fig 1.3.

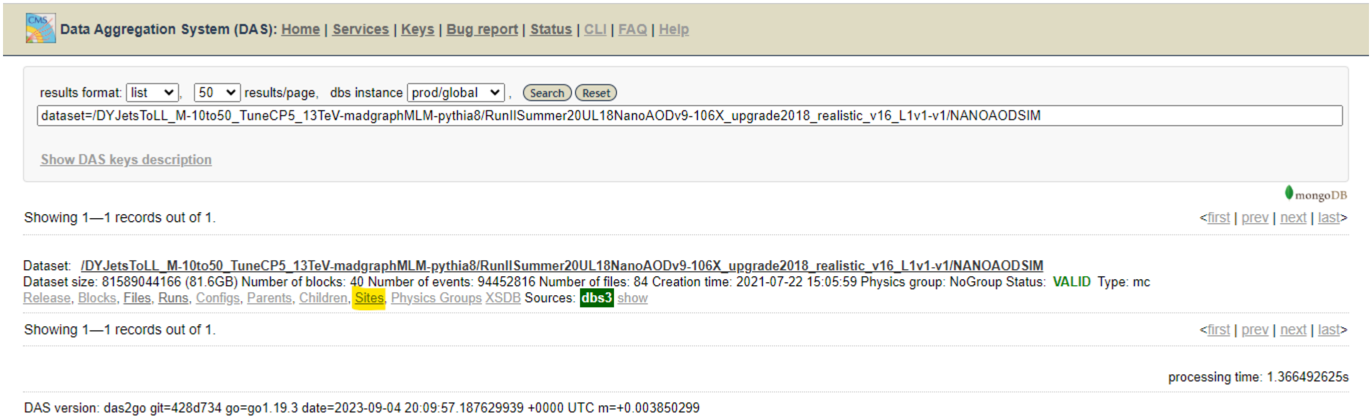


Figure 1.3: Snippet of a sample of type Drell-Yan (DY) in CMS DAS

CMS DAS is only accesible with a security certificate in format p.12 installed into the browser. It is also possible to verify the block completion of diferent datasets accross all the tiers 1,2 and 3 everywhere there is one. For example taking again as an example our dataset of type DY in DAS, if we click into the "sites" (fig 1.4) we see our DY sample is completely present in basically all CMS datacenters worldwide.

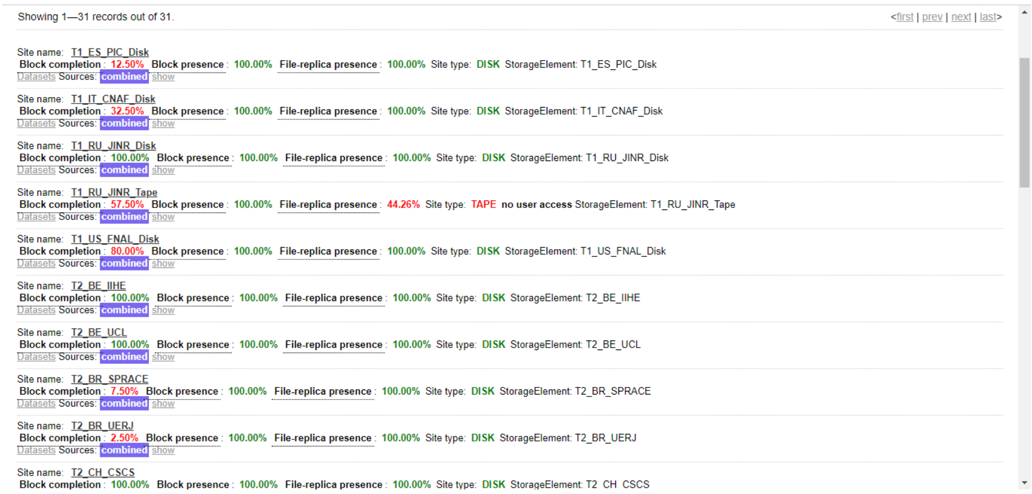


Figure 1.4: State of dataset completion accross all CERN datacenters worldwide

In table 1 we show a list of the signals and backgrounds used in the project, all of them taken from CMS DAS.

Group	Process CMS DAS Path	$\sigma \cdot BR(pb)$
Signal	/HIG-RunIISummer20UL18wm LHEGEN_HH_bbmumu_13TeV-powheg_500k/ bdanzi-NANOAOD- 0000000000000000000000000000 /USER instance=prod/phys03	0.0000078
DY	/DYJetsToLL_0J_TuneCP5_13TeV-amcatnloFXFX-pythia8 /RunIISummer20UL18 NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-v1 /NANOAODSIM	18610
DY	/DYJetsToLL_1J_TuneCP5_13TeV-amcatnloFXFX-pythia8 /RunIISummer20UL18 NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-v1 /NANOAODSIM	4757.0
SingleT	/ST_tW_top_5f_NoFullyHadronicDecays_TuneCP5_13TeV-powheg-pythia8 /RunIISummer20UL18 NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-v1 /NANOAODSIM	35.85
SingleT	/ST_s-channel_4f_hadronicDecays_TuneCP5_13TeV-amcatnlo-pythia8 /RunIISummer20UL18 NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-v1 /NANOAODSIM	35.85
SingleT	/ST_s-channel_4f_leptonDecays_TuneCP5_13TeV-amcatnlo-pythia8 /RunIISummer20UL18 NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-v1 /NANOAODSIM	35.85
DY	/DYJetsToLL_2J_TuneCP5_13TeV-amcatnloFXFX-pythia8 /RunIISummer20UL18 NanoAODv9-106X_upgrade2018_realistic_v16_L1v1-v1 /NANOAODSIM	859.589
SMHiggs	/GluGluHToZZTo4L_M125_TuneCP5_13TeV_ powheg2_minloHJJ_JHUGenV7011_pythia8 /RunIISummer20UL18 NanoAODv9-106X_ upgrade2018_realistic_v16_L1v1-v2 /NANOAODSIM	0.0121
GluGluTo ContinTozz	/GluGluToContinToZZTo4mu_TuneCP5_13TeV-mcfm701-pythia8 /RunIISummer20UL18 NanoAODv9-106X_ upgrade2018_realistic_v16_L1v1-v2 /NANOAODSIM	0.0015

Table 1.1: $HH \to b\bar{b}\mu\mu$ signal (highlithed in red) and the backgrounds used, toghether with their cross sections σ in picobarns.

1.3 Data Analysis Strategy and Results

In the analysis, our first focus lies on the reconstruction of the Higgs and Z bosons through the decay channel involving two muons, the cleanest one. To elaborate further, we will delve into the intricacies of reconstructing muons and electrons, key particles in this process.

- Muons: they are very well reconstructed in the CMS detector thanks to the muon chambers (RPC, DT, CSC) and thanks to the fact that they leave a clear signature passing the full dector at MIP. Thanks to this their tracks are also clearly identified by the tracker.
- Electrons: we reconstruct electrons mainly thanks to the traker and the electromagnetic calorimeter (ECAL). The resolution of the calorimeter, the segmentation of the tracker and the fact that electrons are always produced within jets, makes the reconstruction more complicated with respect to the muons one. By appling some offline tools we can anyway achieve a good performance.

The steps taken in for the analysis are summarized in the following

- Events passing SingleMuon trigger: HLTIsoMu24
- Selection of the muons with $|\eta_\mu| < 2.4$ and initial transverse momentum cut $p_T > 5$
- Selection of the longitudinal and transverse components of the muon impact parameters at $|d|_z < 1$ and $|d|_{xy} < 0.5$
- Selection of the electrons with $|\eta_{e^-}| < 2.5$ and $p_T > 7$
- Selection of the transverse and longitudinal components of the impact parameter at $|d|_{xy} < 0.5$ and $|d|_z < 1$
- Check agreement between data and Monte Carlo in Z boson $\rightarrow \mu\mu$

We require two final muons with opposite charges and $p_{T_{l1}} > 20$ and $p_{T_{l2}} > 10$, the code will then compare the mass of all the possible pairs of generated muons to the Z and Higgs candidate masses, and the muon pair mass approximates the best to the nominal masses of the

Z/ Higgs bosons. With the methodology and tools elucidated earlier now at our disposal, we stand prepared to unveil our findings and draw forth our final conclusions.

1.3.1 Results

We highlight what is happening in the first plot (Fig 1.5), where we see a considerable discrepancy between the data and MC generated data, this is due low statistics and with MC events the discrepancies can be fixed.

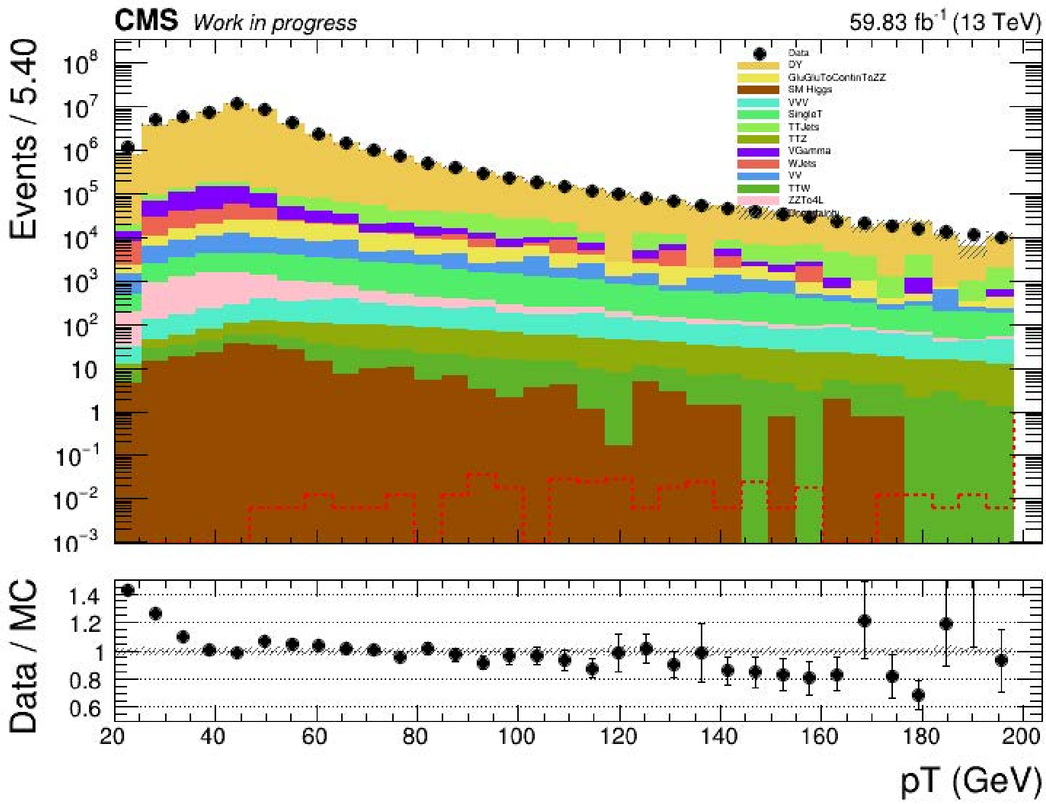


Figure 1.5: Data (black dots) and MC for different samples as a function of transverse momentum p_T

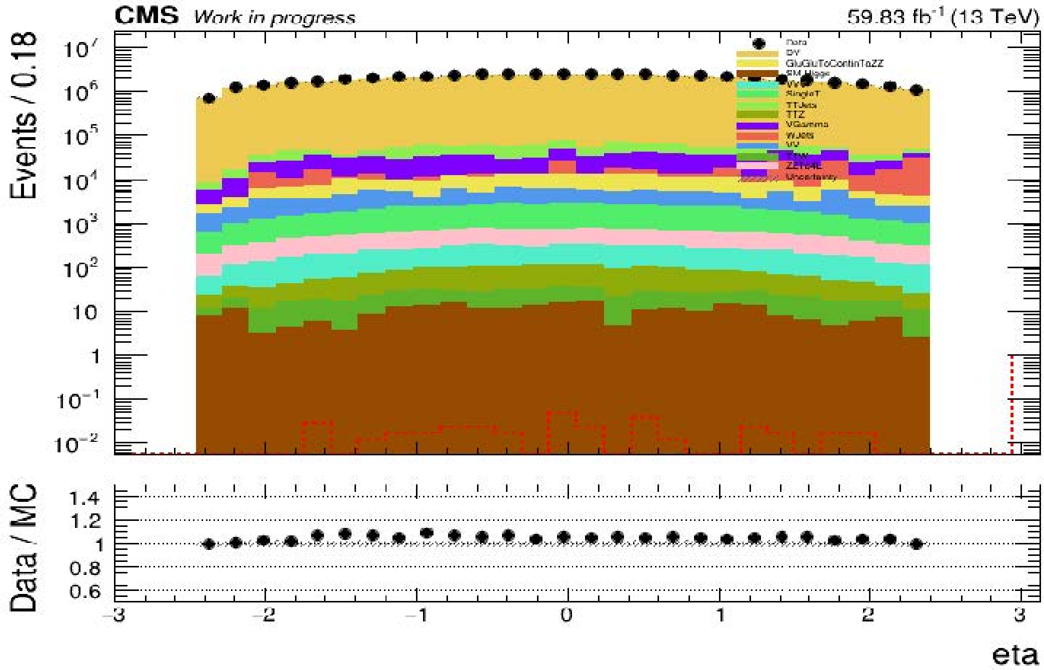


Figure 1.6: Data (black dots) and MC for different samples as a function of the Pseudorapidity η

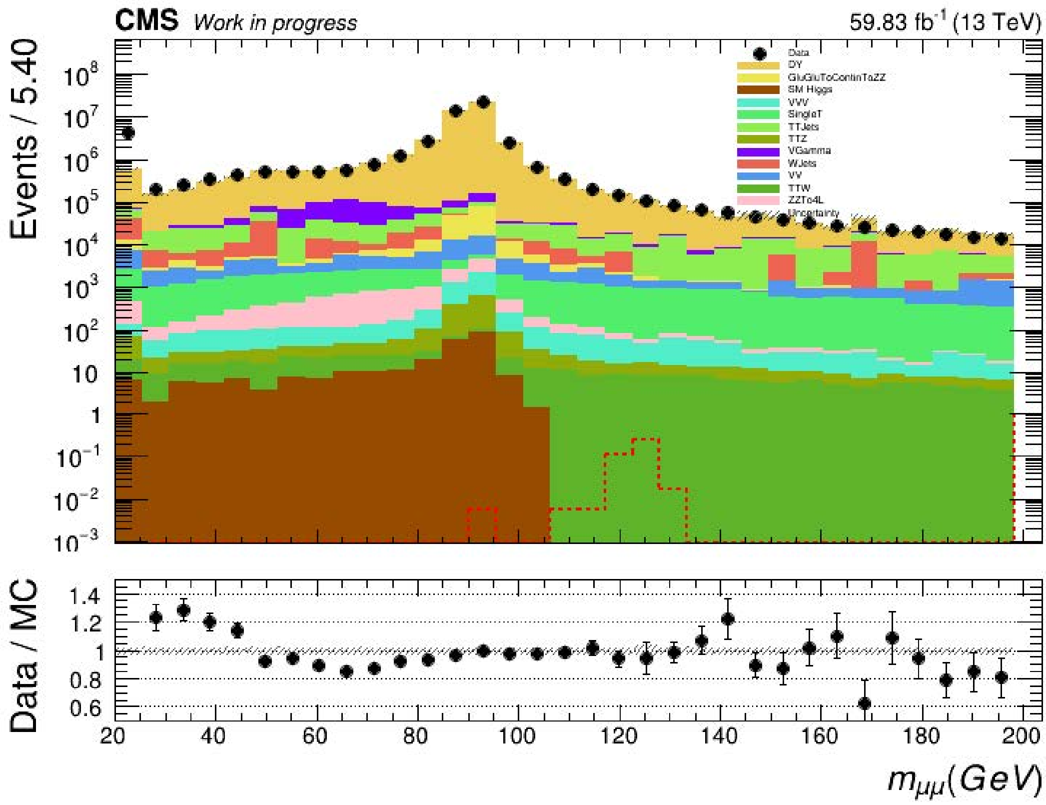


Figure 1.7: Invariant mass of the muon pairs recreating the mass of the Z boson.The mass of the Higgs boson has also been recreated (red hashed peak)

Conclusions and Future Works

Conclusions

The analysis of the obtained results leads to several key conclusions. Firstly. It has been observed a considerable level of agreement between the data and Monte Carlo simulations when examining lepton kinematic variables. However, it is essential to acknowledge that some discrepancies have been found, primarily stemming from the limitations of our Monte Carlo simulations due to low statistics. Nonetheless, despite these challenges, we successfully reconstructed the invariant masses of both the Z boson and the Higgs boson from a pair of muons . These findings shed light on the intricacies of our study and underscore the importance of accounting for statistical limitations in our Monte Carlo simulations.

Future Works

Our future goals involve. First and foremost, the aim to seamlessly integrate HTCondor into Bamboo, enhancing our computational capabilities and efficiency. Additionally, our objectives include the reconstruction of the Higgs boson into b$\bar{b}$.Furthermore, we are dedicated to the meticulous selection of variables to optimize the training of a Deep Neural Network, thus enhancing its discrimination power between signal and background. This will allow us to extract crucial upper limits on the cross section and the Higgs self-coupling. Lastly, we plan to expand the usage of the real datasets from CMS using the Double Muon and the MuonEG ones.

Acknowledgements

I wish to extend my sincere and profound gratitude to my supervisors, Brunella D'Anzi and Nicola De Filippis, for their invaluable support, assistance, and boundless patience. Without their guidance, this project would not have reached its fruition. Additionally, my appreciation goes to the CMS collaboration for granting me access to their facilities, tools, and data, enabling the successful completion of this endeavor.

Bibliography

[1] F. Englert and R. Brout. “Broken Symmetry and the Mass of Gauge Vector Mesons”. In: *Phys. Rev. Lett.* 13 (9 1964), pp. 321–323. DOI: 10.1103/PhysRevLett.13.321. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.321.

[2] John Ellis. *Higgs Physics*. 2013. arXiv: 1312.5672 [hep-ph].

[3] Peter W. Higgs. “Broken Symmetries and the Masses of Gauge Bosons”. In: *Phys. Rev. Lett.* 13 (16 1964), pp. 508–509. DOI: 10.1103/PhysRevLett.13.508. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.508.

[4] Pieter David. “Readable and efficient HEP data analysis with bamboo”. In: *EPJ Web Conf.* 251 (2021), p. 03052. DOI: 10.1051/epjconf/202125103052. arXiv: 2103.01889 [physics.data-an].

[5] Marco Peruzzi, Giovanni Petrucciani, Andrea Rizzi, and for the CMS Collaboration. “The NanoAOD event data format in CMS”. In: *Journal of Physics: Conference Series* 1525.1 (2020), p. 012038. DOI: 10.1088/1742-6596/1525/1/012038. URL: https://dx.doi.org/10.1088/1742-6596/1525/1/012038.