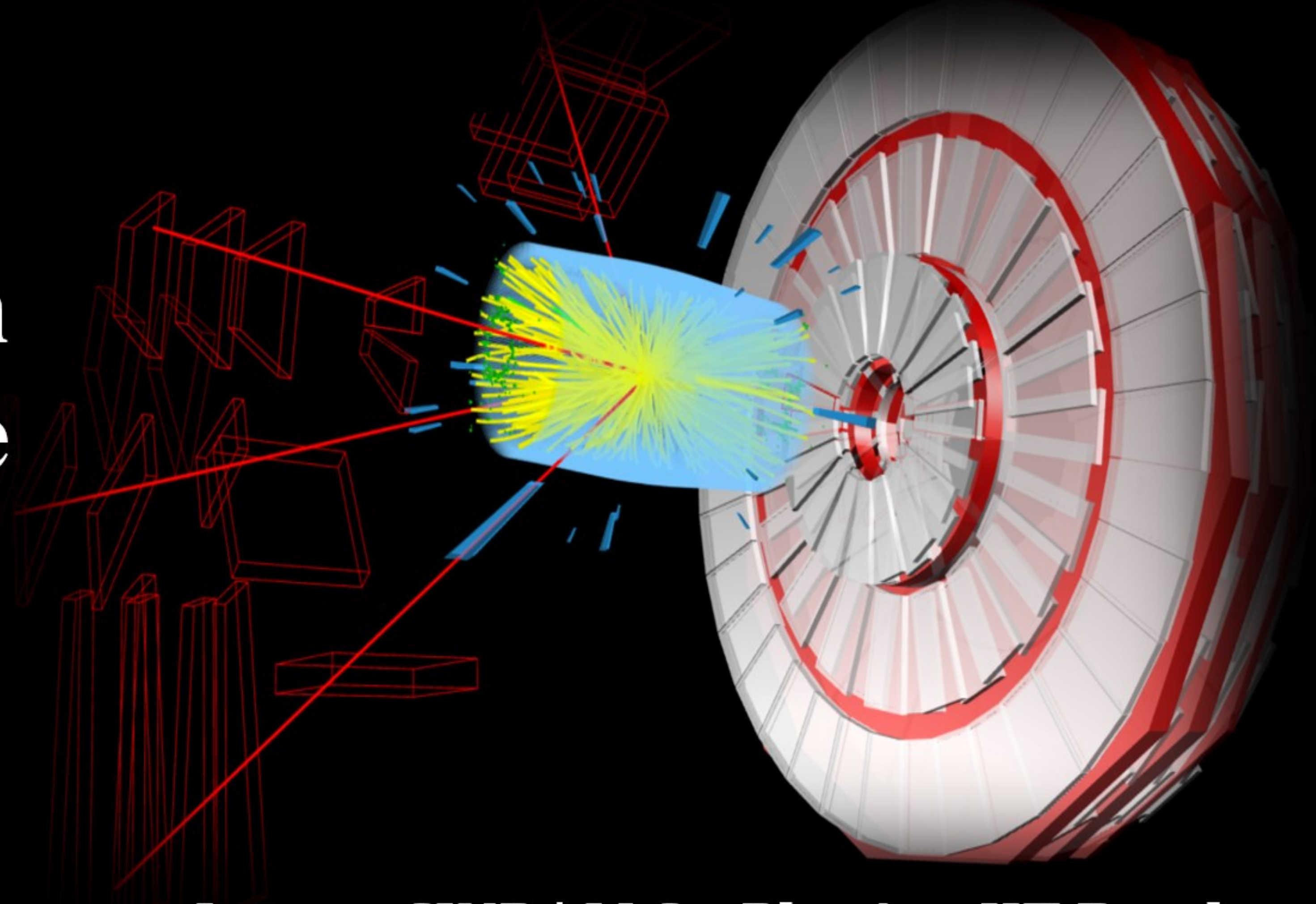
Efficiency of Muon Detection using the Tag and Probe method



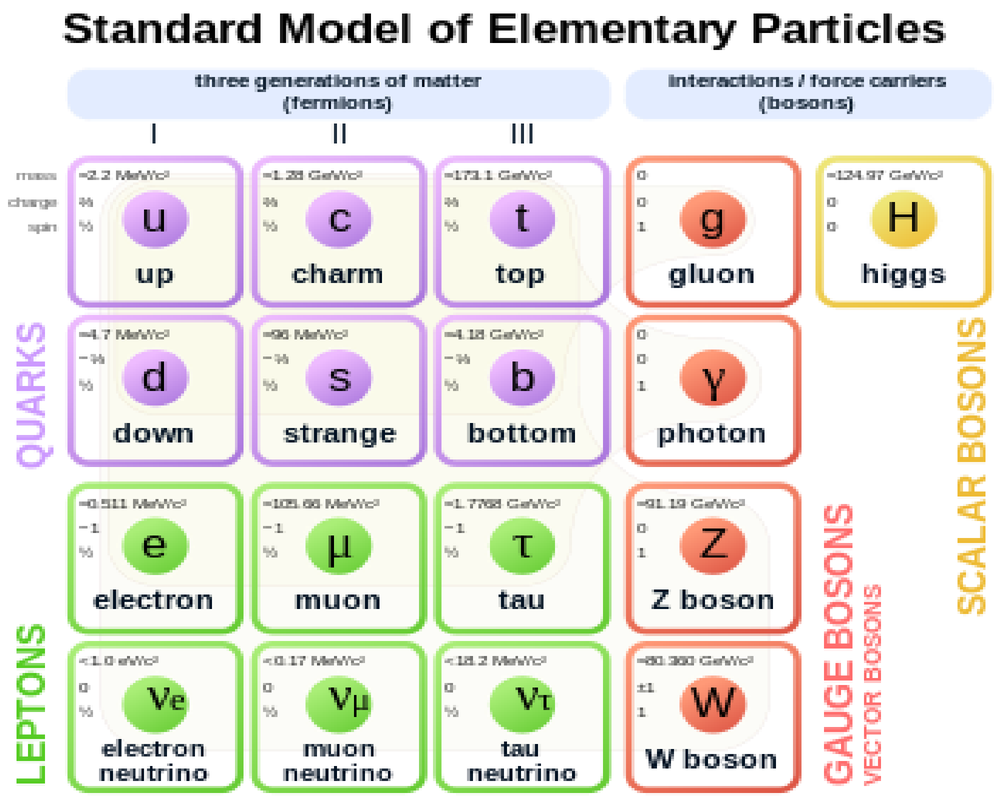
- Presented by : Sayan Dhani |Summer Intern, SINP | M.Sc. Physics, IIT Bombay
- Supervisor : Prof. Subir Sarkar , SINP
- Collaboration : Kuldeep Nishad(IIT Kanpur), Ratul Sarkar(IISER TVM)

Particle Physics

- Branch of physics in which we study the elementary particle and their interactions.
    - To understand the fundamental constituents of matter and the forces governing the universe.
- Standard model is one the most successful theories of Nature.
- To perform the experiments at high energies, we need large particle accelerators

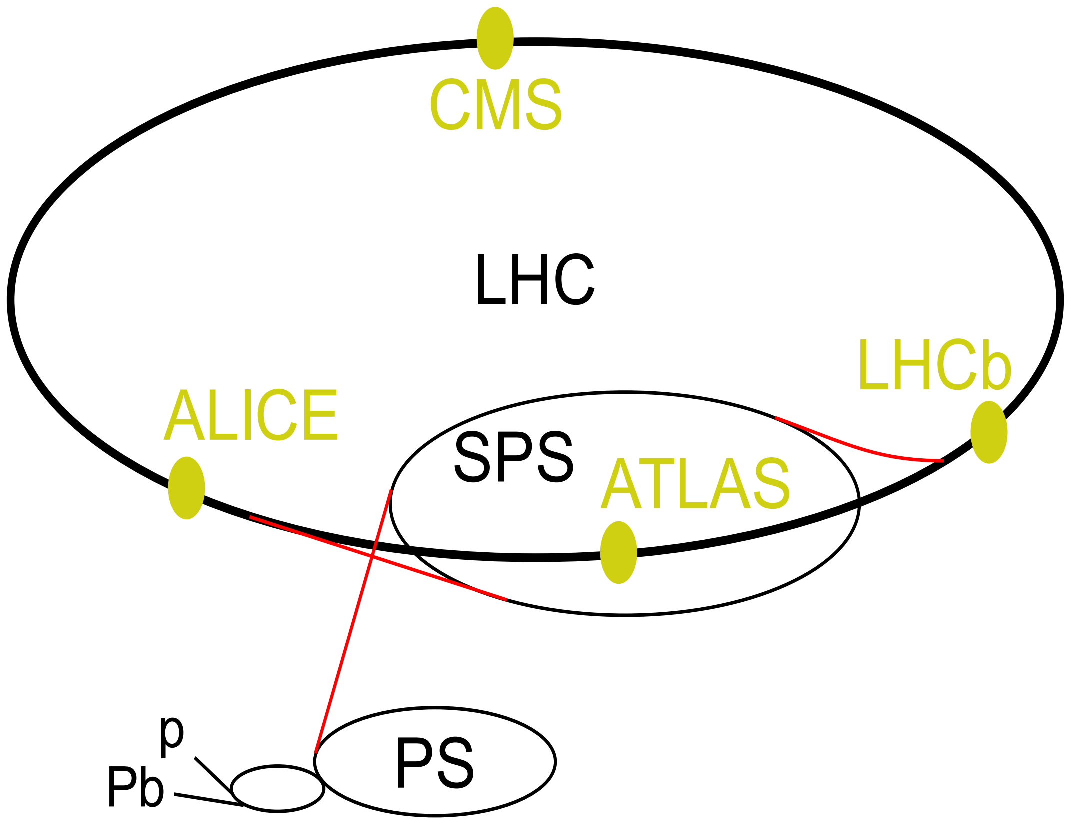
Standard Model

- Describes the elementary particles and their interactions.
- Encompasses particles such as quarks, leptons, gauge bosons and the Higgs boson.
- Describes three of the four fundamental forces of nature (electromagnetic, weak and strong Interactions excluding- Gravity).
- One of the most successful and precisely tested theory.
- Notable Colliders
  - ❑Large Electron Positron (LEP) collider
  - ❑Tevatron p-p collider (Fermilab)
  - ❑LHC pp / heavy ion collider (CERN)



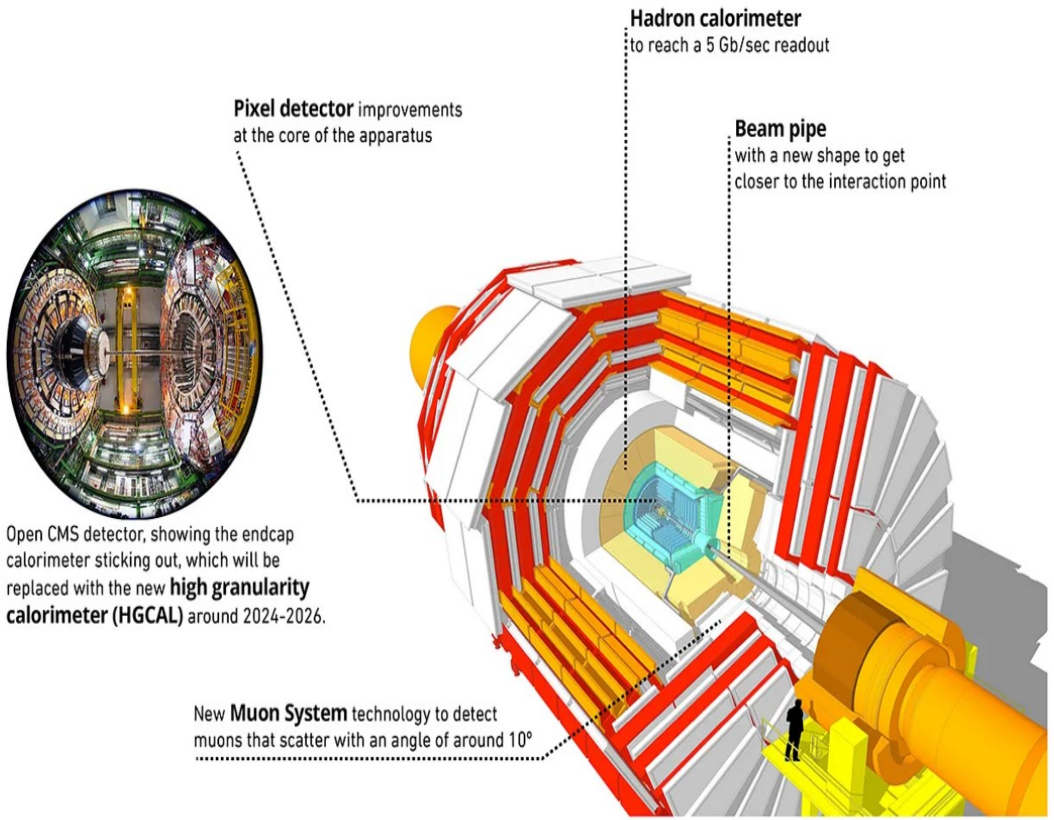
LHC (Large Hadron Collider)

- The LHC is the world's largest and most powerful particle accelerator, located at the European Organization for Nuclear Research (CERN) near Geneva, Switzerland.
- To investigate the fundamental nature of matter and the universe by colliding particles at extremely high energies.
- There are 4 major detector:
  - ATLAS, CMS, ALICE and LHCb
- Discovery of Higgs boson in 2012 (Mechanism of elementary particle mass generation).
- HL-LHC(High Luminosity LHC).

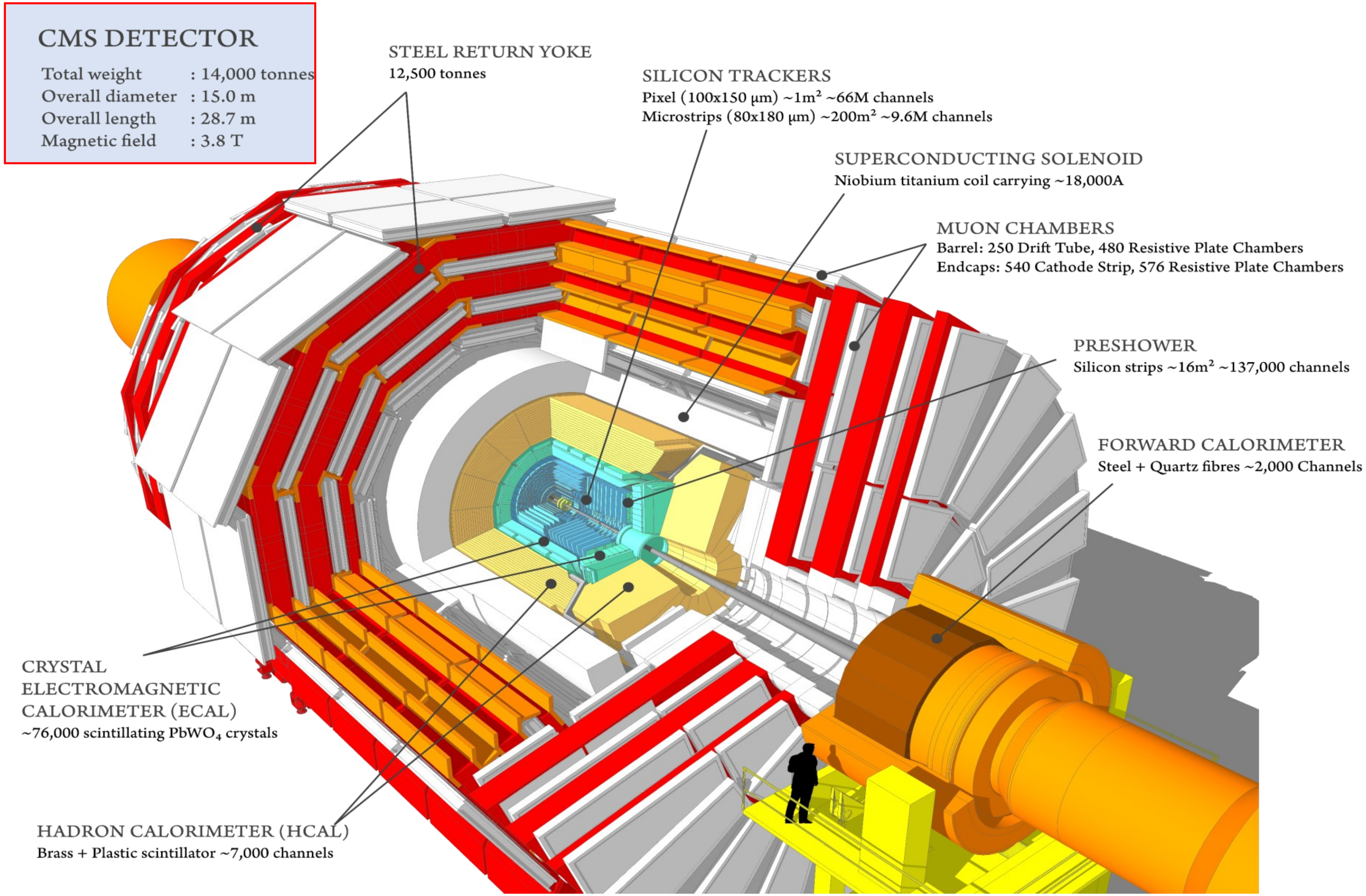


CMS Detector (Compact Muon Solenoid)

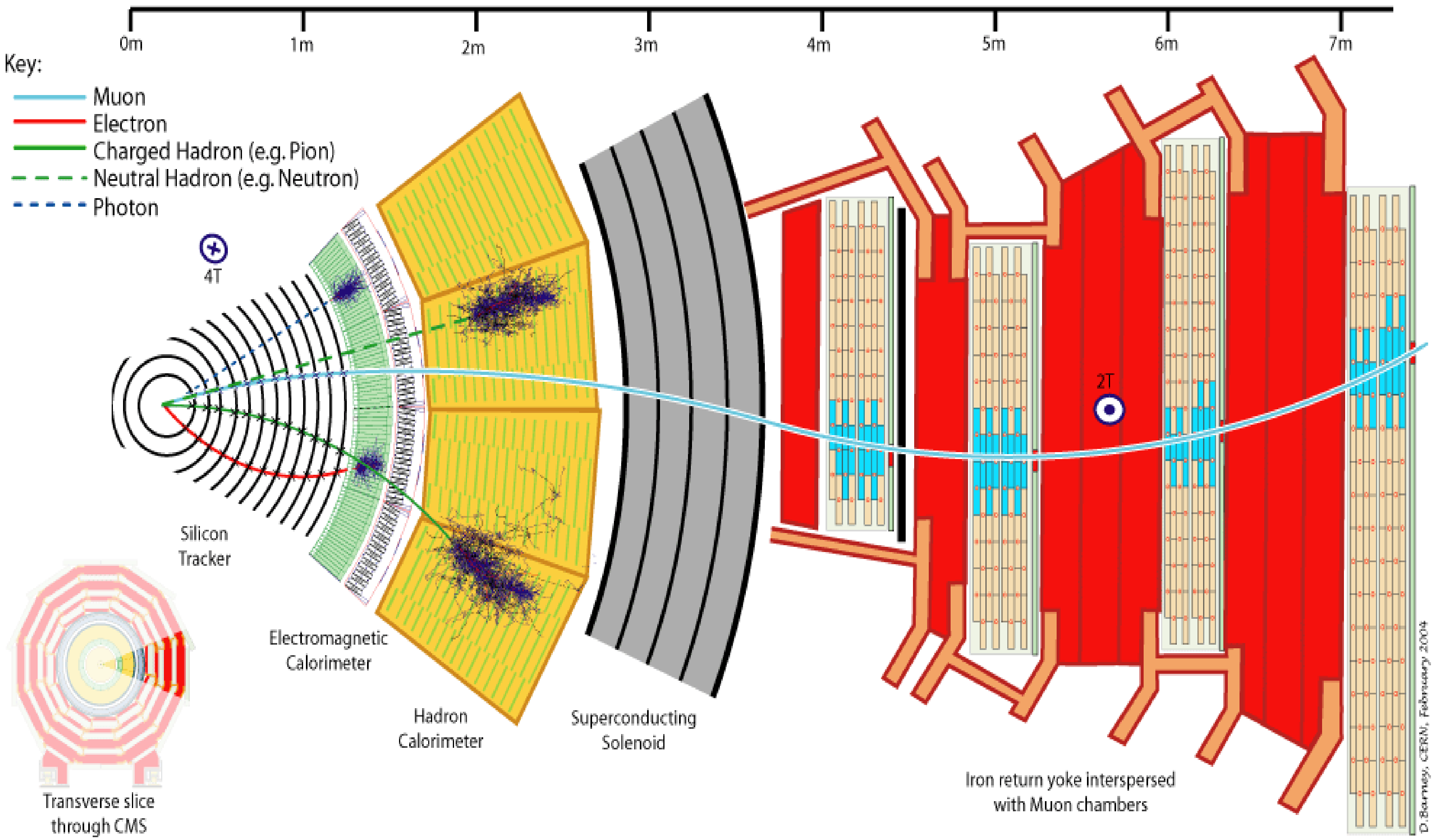
- General purpose detector.
- Played crucial role in the discovery of Higgs boson.
- The CMS detector features a powerful solenoid magnet, which produces a strong magnetic field (3.8 T) to bend the paths of charged particles.
  - enables the measurement of particle momenta and charge sign.
- Precision measurement of higgs boson properties.
- Search for Di-Higgs production (HHH coupling) to know more about the higgs potential.
- Search for new physics



CMS Detector



Particle Detection in CMS Detector



Motivation for Efficiency Estimation

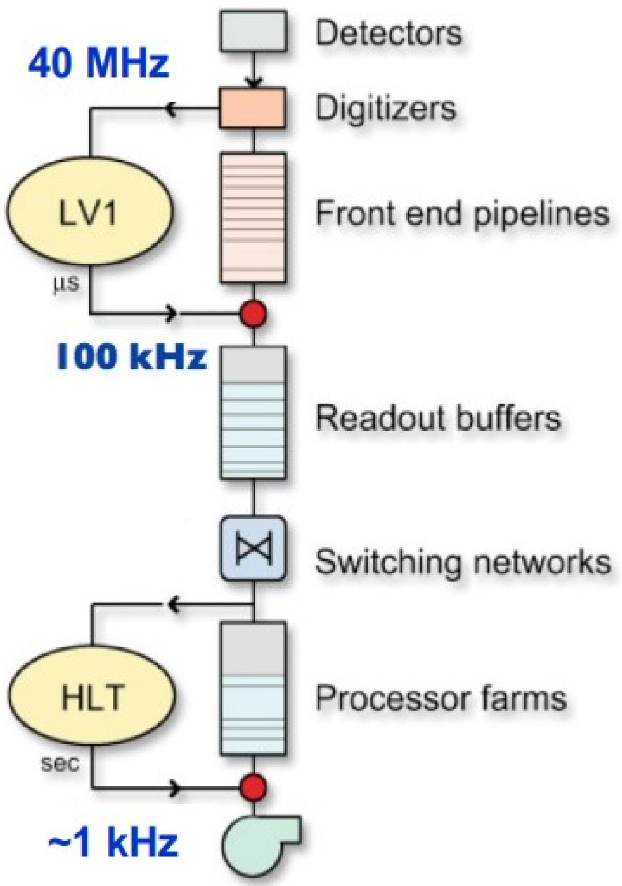
- Why Data and not MC
  - It accounts for detector-specific effects, such as inefficiencies, misalignments, or calibration uncertainties, which may not be fully captured by Monte Carlo simulations.
- Detector Performance Evaluation
  - By measuring the efficiency of various particle identification and reconstruction algorithms, we can evaluate the detector's capabilities and identify areas for improvement.
- Physics Analysis
  - Efficiency calculations are essential for physics analyses, such as cross-section measurements, searches for rare processes, or studies of particle decays.
- Calibration and Corrections
  - By comparing data-driven efficiencies with those obtained from Monte Carlo simulations, corrections and calibrations can be applied to the simulated data to align it with the observed data, improving the accuracy of simulation predictions.
- Accurate Signal Extraction
  - By quantifying the efficiency of selecting specific particles or processes of interest, we can distinguish signal events from background contributions more accurately, leading to precise measurements of physical properties or the discovery of new particles.

Signal and Background

- **Signal**
  - Desired physics processes or phenomena that the experiment aims to observe and study.
  - Here we are considering the physical process $Z \to \mu^+ \mu^-$
- **Backgrounds**
  - Events arising from known, established physics processes that can mimic signal
  - **QCD Backgrounds (jets of particles):** these events can mimic the Z boson signature by producing two high-energy muons.
  - **Cosmic Ray Muons**.
  - **Hadrons with semi-leptonic decays**, such as B mesons can produce a muon in their decay. These decays can mimic the Z boson signal if the accompanying particles are not detected.

The CMS Triggering System

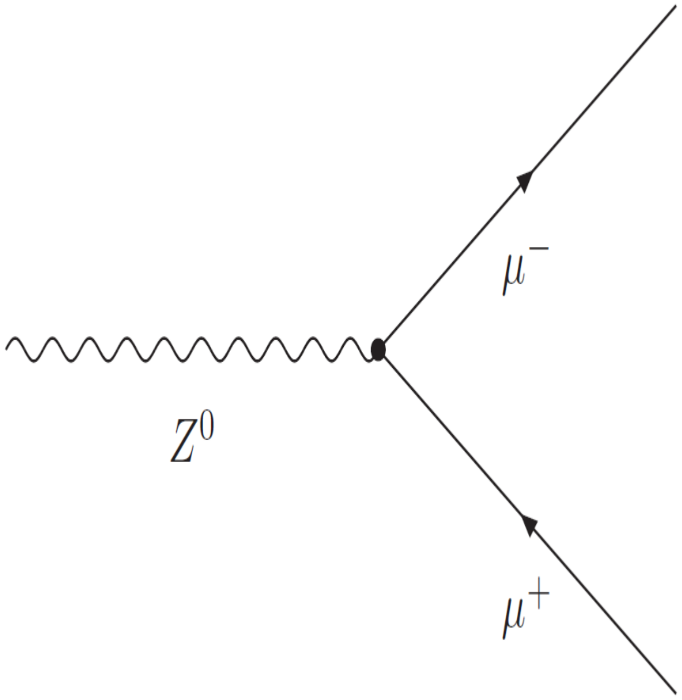
- Rate of collision in CMS around 40 MHz (time interval between two bunch crossing is around 25 ns).
- **Level-1 Trigger:**
  - Hardware based selection of interesting events
  - The L1 trigger operates at a very high speed
  - It reduces the event rate to ~ 100 kHz
- **HLT (High Level Trigger) :**
  - Software based Selection of events
  - Performs more sophisticated events selection and data processing than level-1 trigger
  - It reduces the event rate to an average of ~ 1 kHz



Decay of Z Boson

- Neutral, spin 1, $m_Z = 91.2\ \text{GeV}$ particle
- Decays to many particles

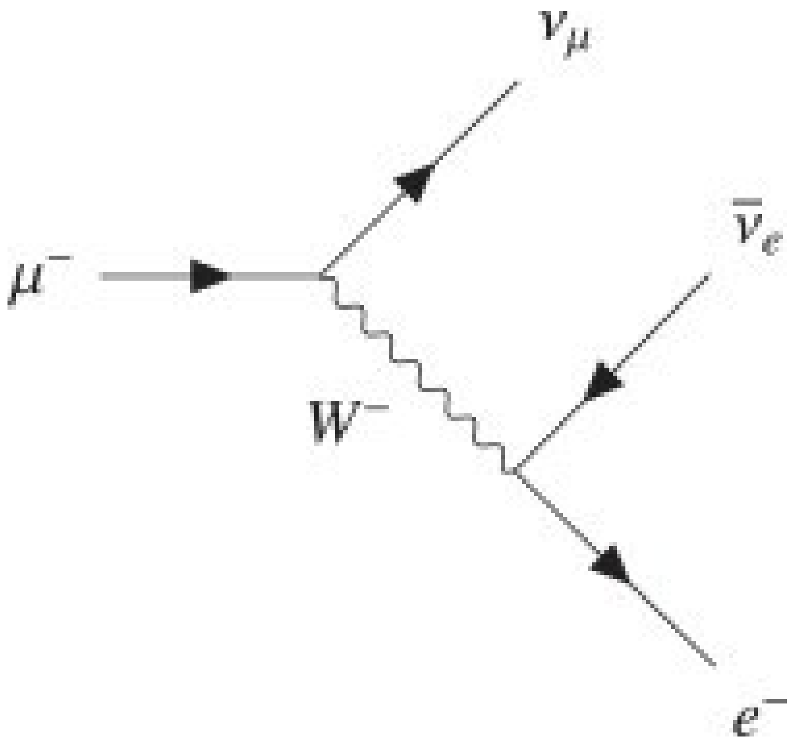
Daughter Particle	Possibilities	No. of Possible Decay
Charged leptons and anti-leptons pairs	10%	3
Neutrino and anti-neutrino pairs	20%	3
Quark and anti-quark (detected as jets)	70%	15 (Due to color charges of Quarks)



- Experimentally well understood and easy to reconstruct with low background.

Muon

- Second Generation, Elementary particle.
- Charge = - e, Spin 1/2 , ~ 207 times heavier than electron.
- has a mean life time of 2.2 μs .
- Muons decay to electron, a neutrino and an antineutrino.
- Muons reach up to Muon System before decay.
- Due to larger mass, muons accelerate slower than electrons in EM fields and emit less Bremsstrahlung radiation resulting far deeper penetration into matter.
- Muons are more efficiently measured than electrons in collider experiments due to their larger mass and better penetration through matter, resulting in less energy loss and stronger detection signals.



Key Variable

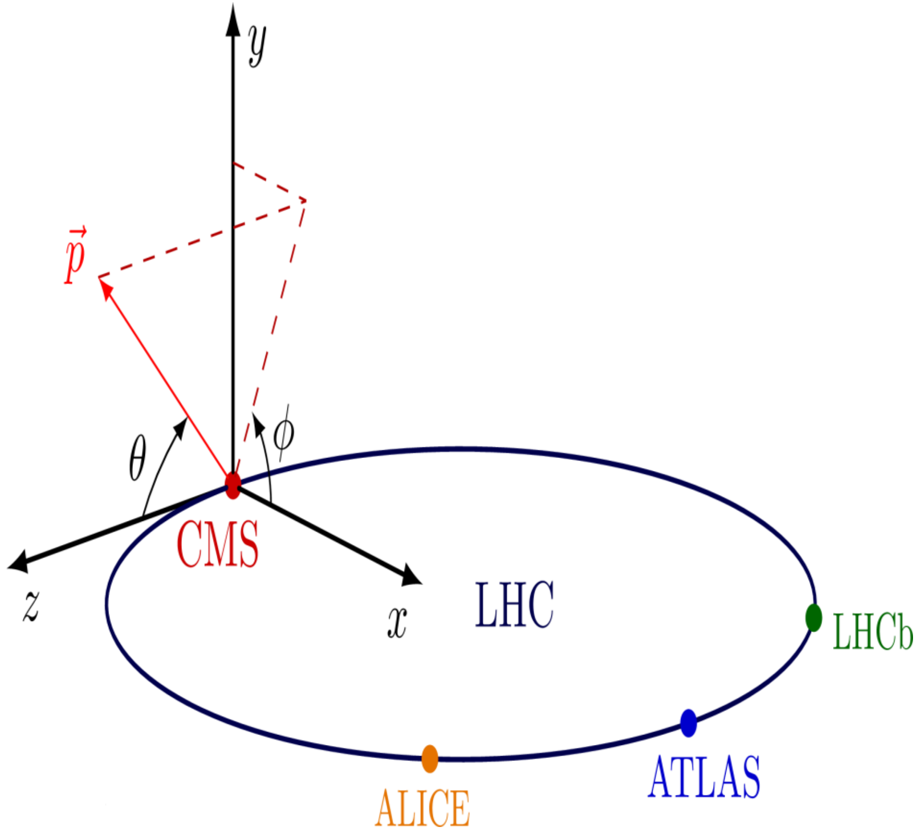
- $\eta = -\ln\left(\tan\left(\frac{\theta}{2}\right)\right)$ - Pseudo-rapidity: where θ *is the angle between momentum* $(\overrightarrow{p})$ *of the particle* and beam direction ($z - axis$).

- ϕ : Angle of the trajectory of the object in the plane transverse to the direction of the proton beams.

- pt : Transverse Momentum: projection of the particles momentum in the plane transverse of the proton beams .

- m: Invariant mass of muon.

- Using the above four quantity we prepare Lorentz vector for Tag and Probe muon and adding them we get the Lorentz Vector for corresponding Z particle.



Software Used During the Project

- Python & Jupyter Notebook
  - For general data manipulation, visualization, and analysis.
  - Provides a flexible and interactive environment.
- Root I/O (Reading & Writing Flat Root Tree)
  - Utilized for storing and accessing data in ROOT format.
  - ROOT is a popular data analysis framework in high-energy physics.
- PyRooT (ROOT Interface for Python)
  - Employed for data analysis within Python environment.
  - Facilitates access to ROOT's extensive features and data structures.
- RooFit (Fitting Mass Histogram)
  - Employed for advanced fitting of mass histograms.
  - Provides powerful tools for signal and background modeling.

Tag and Probe Method

- Data driven technique for measuring the efficiency.
- based on the resonances (e.g. J/ ψ, Υ and Z) to pairs of the particles being studied.
- We have used the Z resonance ($Z \rightarrow \mu^+ \mu^-$).
- Tag Muon: Well identified, muon that triggered the event (tight selection criteria).
- Probe Muon: unbiased set of muon candidates (within acceptance, no cut on pt), either passing or failing the criteria for which the efficiency is to be measured.
- In case of any discrepancies, simulation can be refined (scale factor).

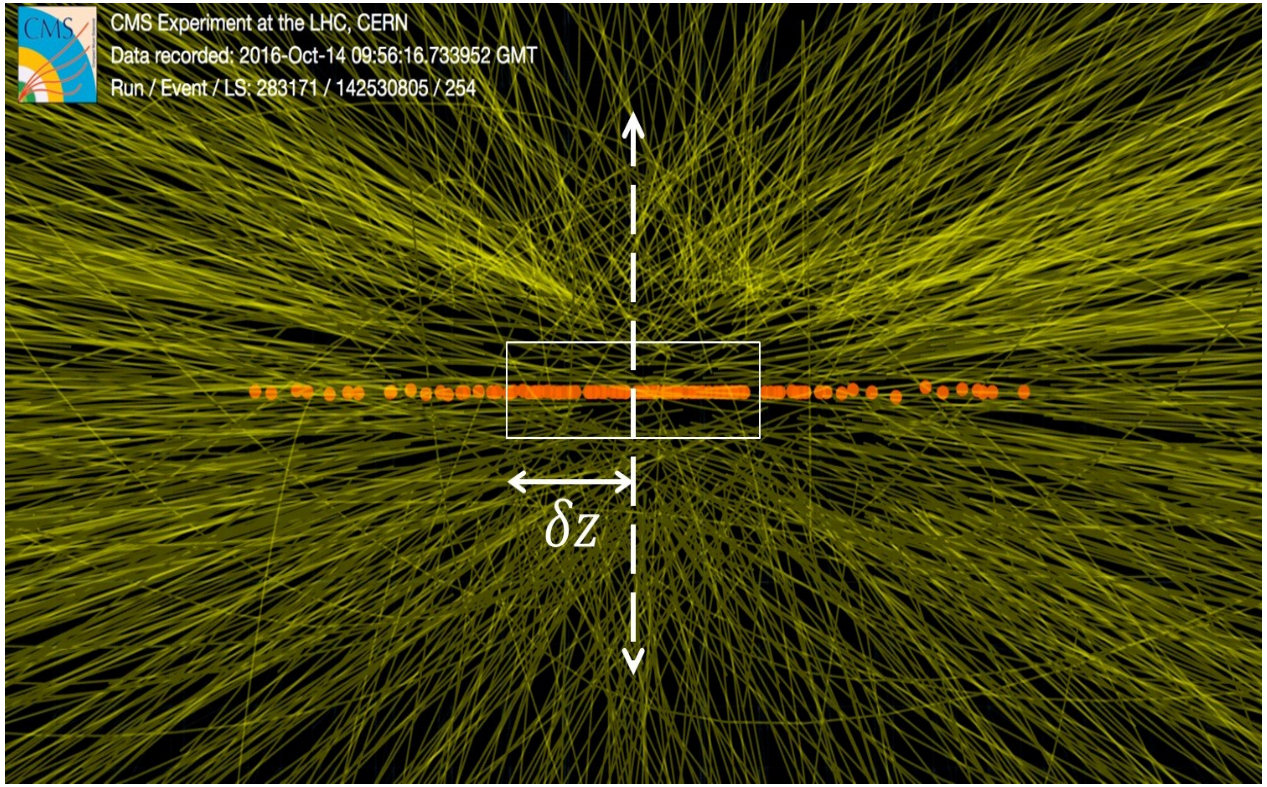
T&P: Object and Event Selection

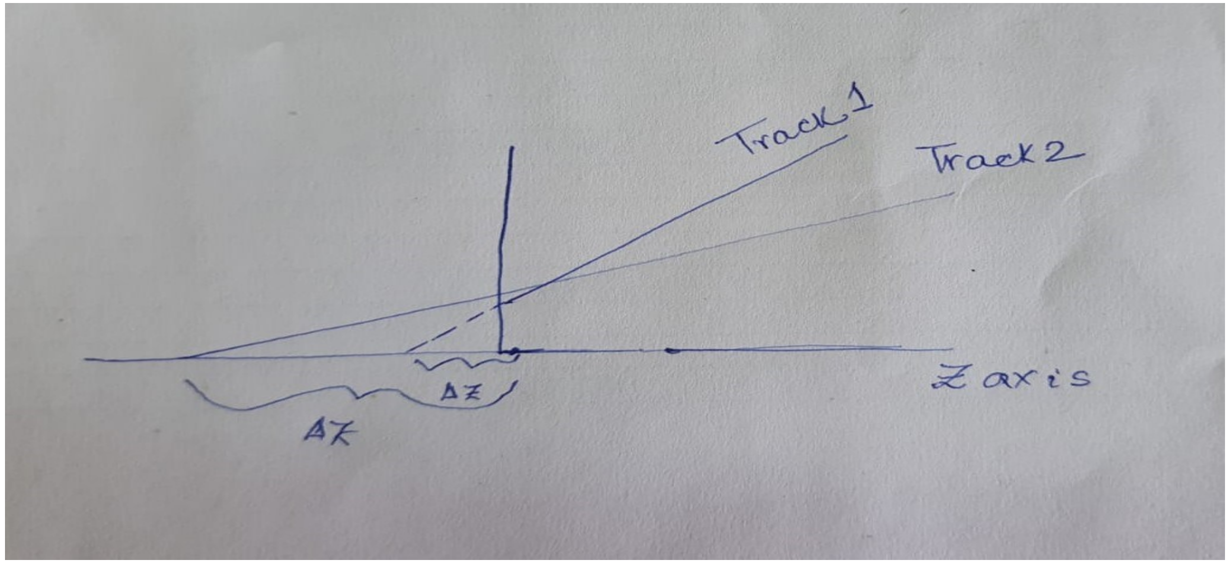
- **Event Selection:** select the interesting events having
  - a) at least 2 muon candidates
  - b) HLT Condition: at least one isolated muon with pt of at least 24 GeV and $|\eta| < 2.1$
  - c) MET < 40 GeV
- **Tag Muon Selection:** select those muons from each event which have
  - pt > 24 GeV, isolation < 0.3, $|\eta| < 2.1$
  - have the tight/best reconstruction.
  - put them in an array and shuffle; choose the 1st one
- **Probe Selection:** from the same event select muons having
  - Loop over all the muon except the tag muon,
  - opposite charge to the tag muon
  - $|\eta| < 2.1$,
  - and gives Z mass (adding tag and probe) between 20 – 160 GeV

- **Apply Veto**
- **Electrons**
  - pt > 10 GeV, $|\eta| < 2.5$, mvaId == 0 ,
  - DeltaR (ele, probe_u or tag_u) > 0.4
- **Jets**
  - pt > 30 GeV, $|\eta| < 4.7$
  - DeltaR (jet, probe_u or tag_u) > 0.4
- **bJets**
  - pt > 30 GeV, $|\eta| < 2.5$
  - DeltaR (jet, probe_u or tag_u) >0.4
  - jet_btagScore > 0.33

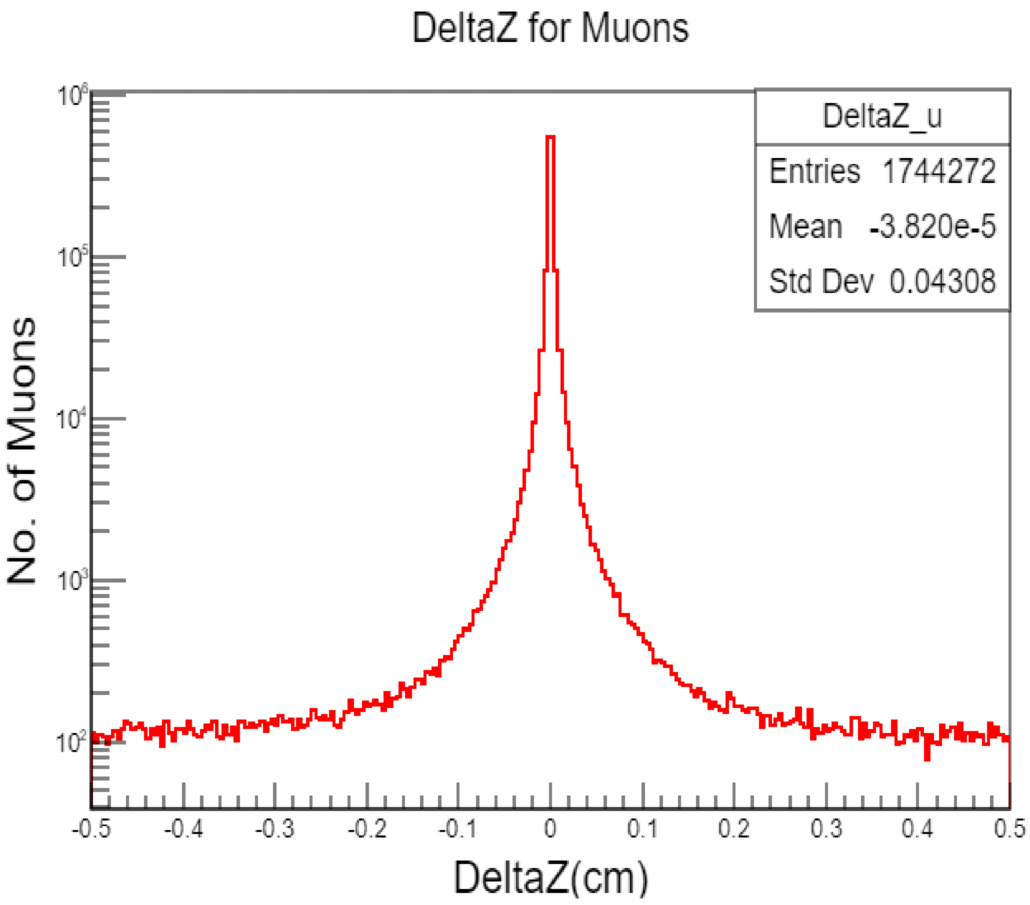
If the event has

- any selected electron, skip the event
- more than 3 Jets, skip the event
- any b-Jet, skip the event





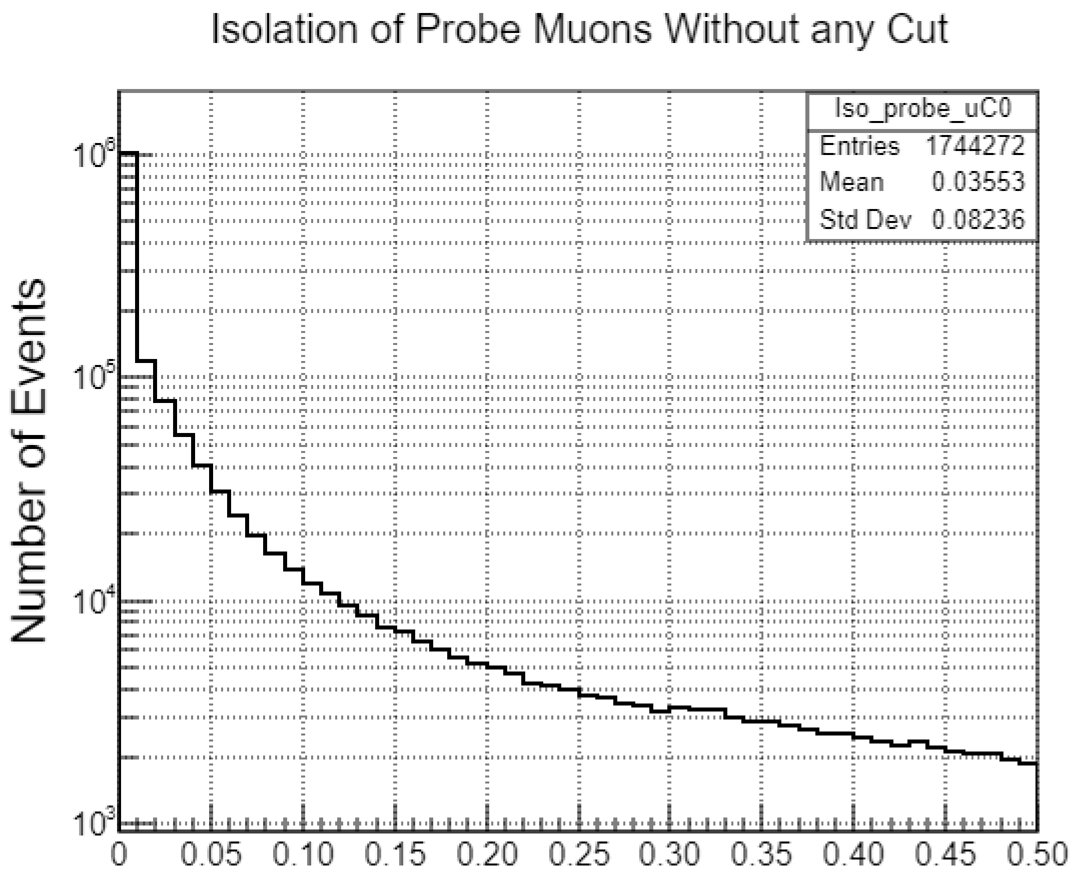
DeltaZ for a particle gives the information about the particle whether it is coming from (close to) primary vertex.

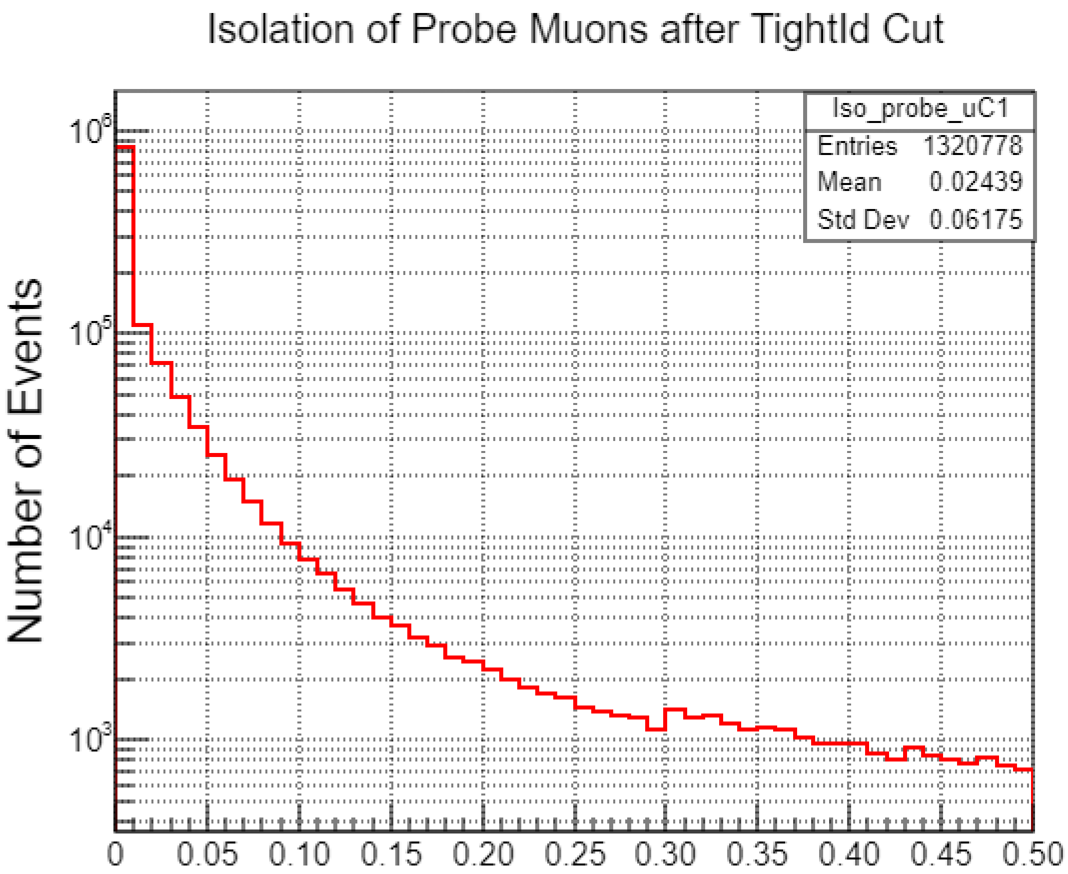


Isolation of probe muons

$$\textbf{Isolation} = \frac{\Sigma\, p_T^i}{p_T^{\mu}}$$

Where **i** is for other particles within a cone size around the muon.
and $\boldsymbol{\mu}$ muon for which we want to calculate isolation.





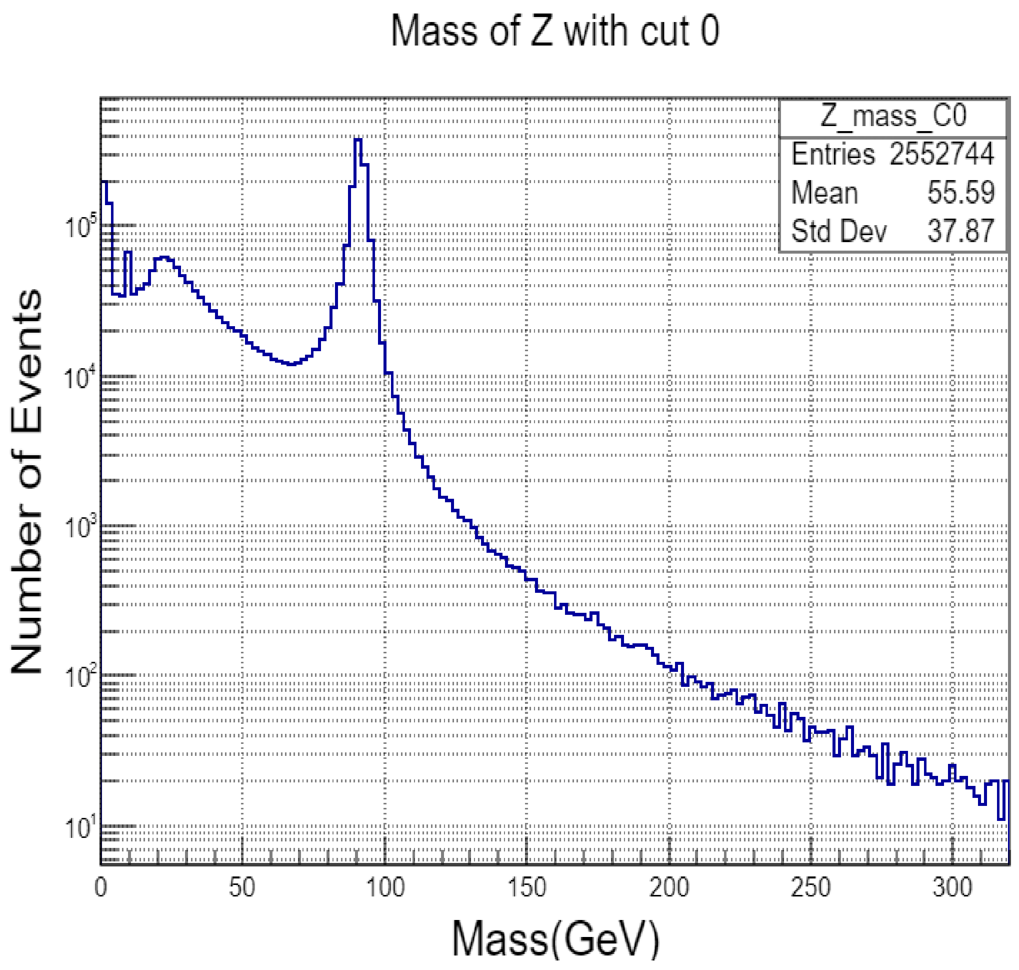
Efficiency Calculation

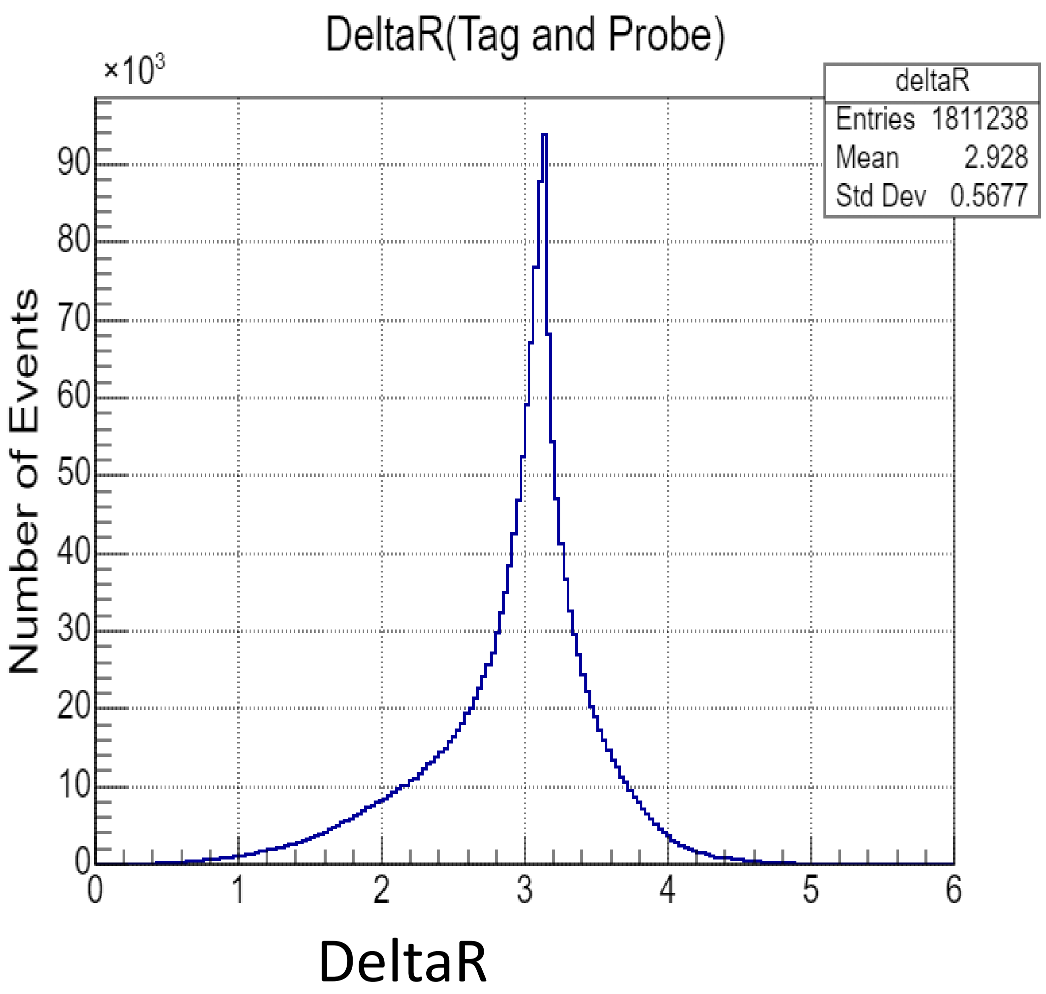
- Fraction of Probe muons that pass the particular criteria.
- The aim is to measure a certain ID and/or ISO efficiency over the "probe" muons as.

$$\epsilon = \frac{N_{pass}}{N_{all}}$$

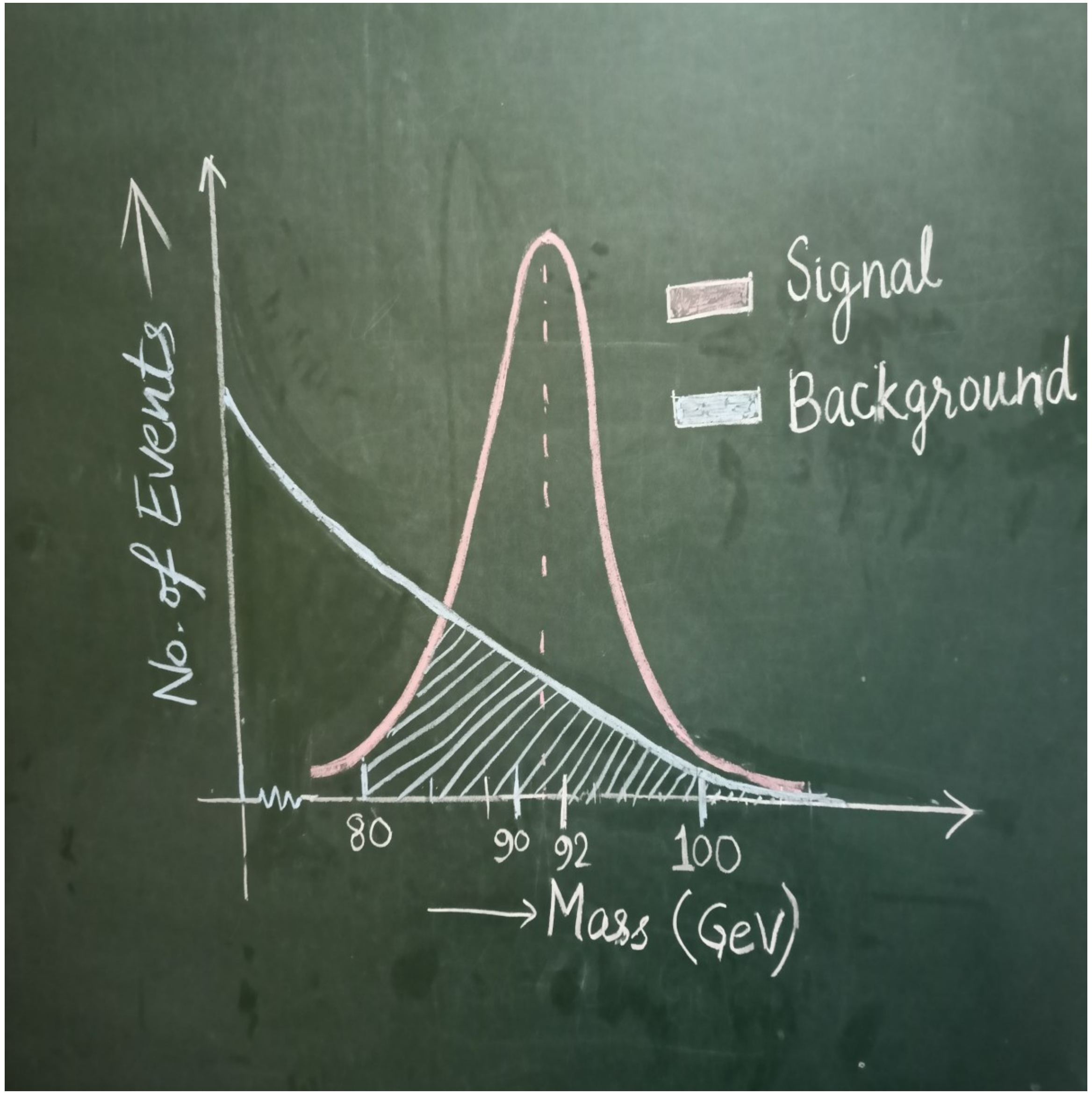
N_{pass} = The number of muon passing certain ID (Reconstruction) cut e.g. : tight Id, medium Id, loose Id and/or Isolation cut.

N_{all} = The number of probable candidate for $Z\ to\ \mu^{+} and\ \mu^{-}$ in a mass window of (20 – 160) GeV.

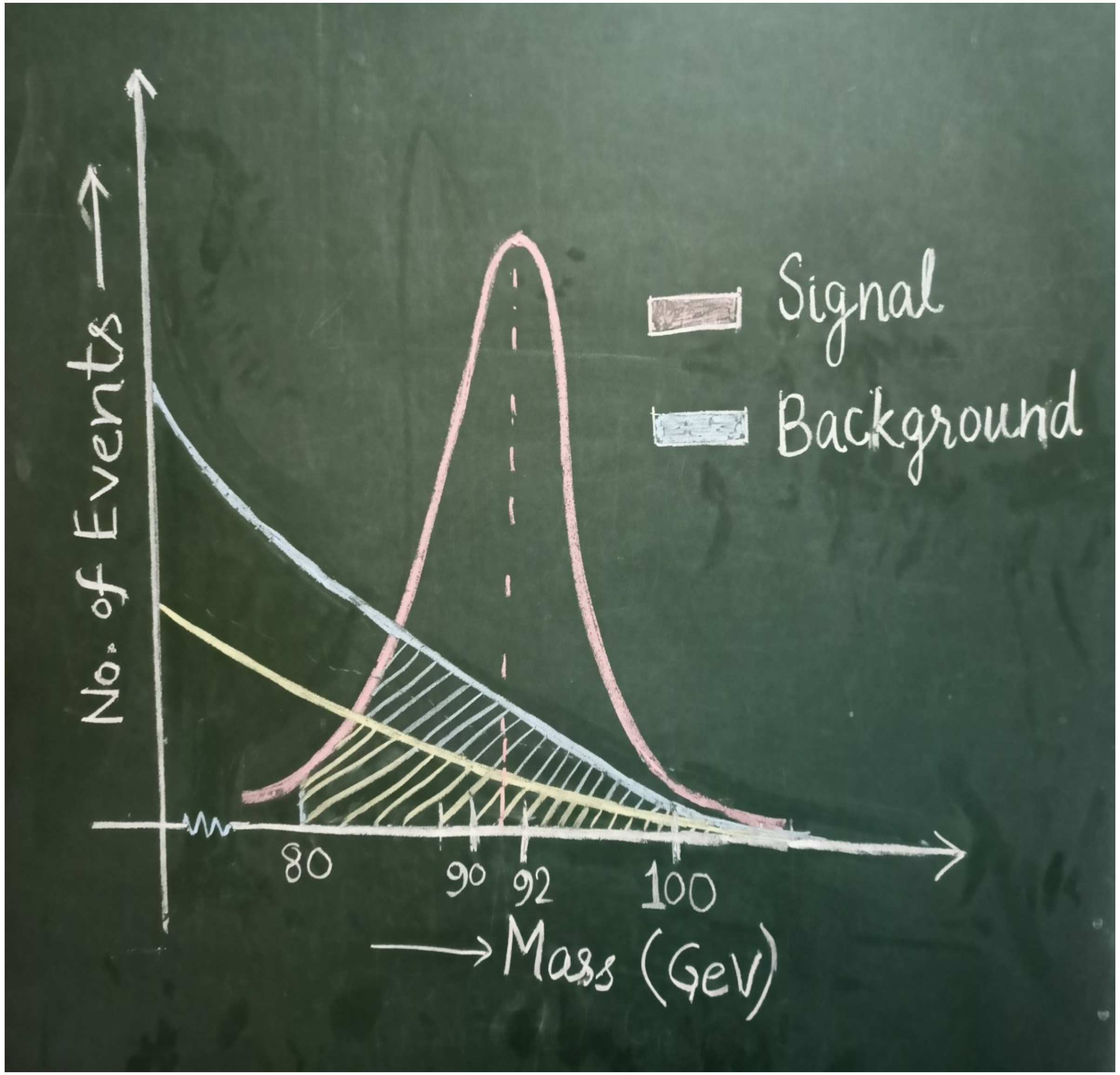

Mass of Z with cut 0
Z_mass_C0
Entries 2552744
Mean 55.59
Std Dev 37.87
Number of Events
Mass(GeV)
0
50
100
150
200
250
300


DeltaR(Tag and Probe)
×10³
deltaR
Entries 1811238
Mean 2.928
Std Dev 0.5677
Number of Events
DeltaR
0
10
20
30
40
50
60
70
80
90
1
2
3
4
5
6

Signal and Background for Z Resonance



After Tight Event Selection



Estimation of Efficiency: Method 1

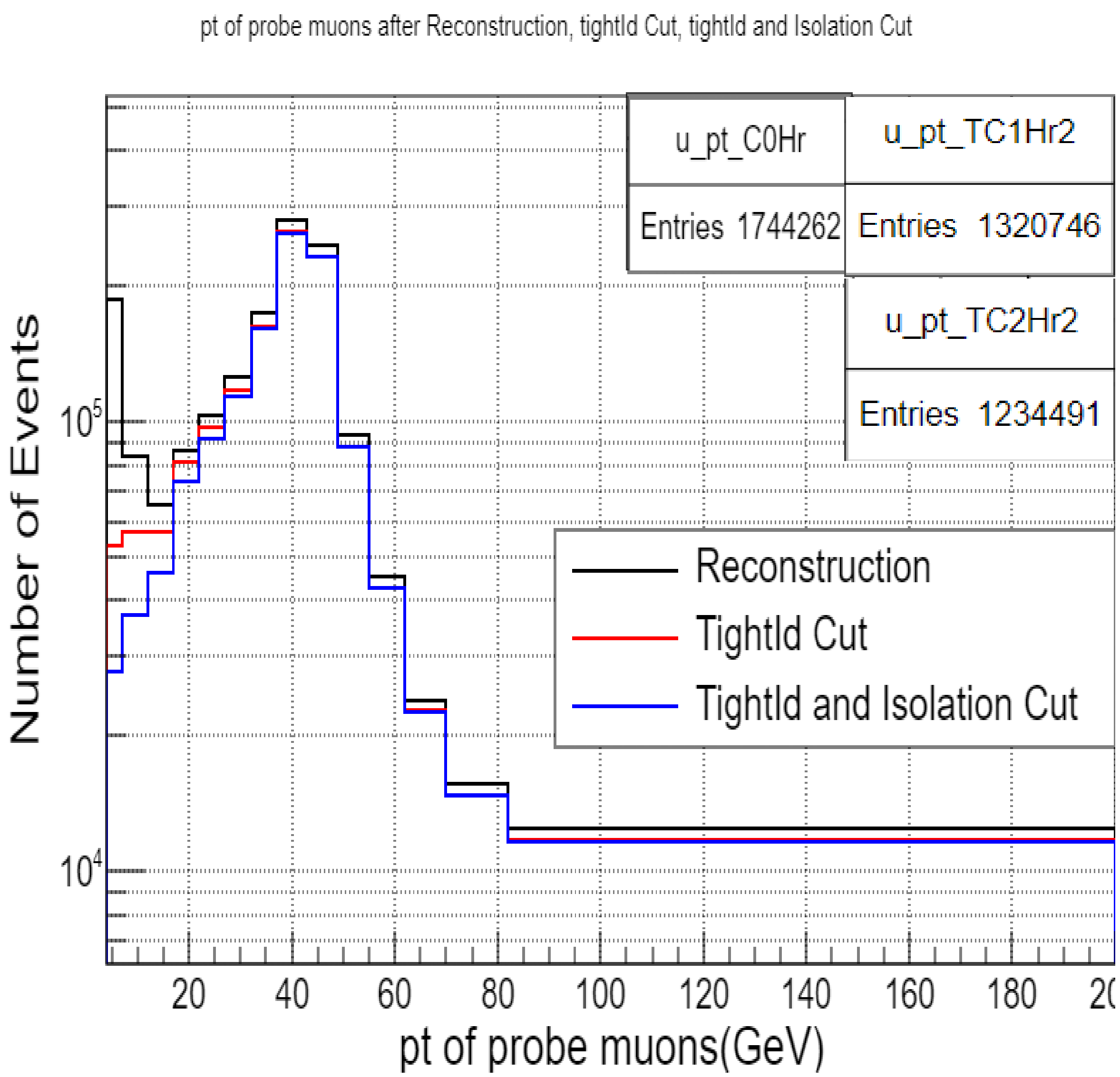
- After tight event selection: No. of probe muons = N_0
- Now apply tight Id cut: No. of probe muons = N_1
- Again apply Isolation cut: No. of probe muons = N_2

Efficiency of tight Id Cut = $\frac{N_1}{N_0}$

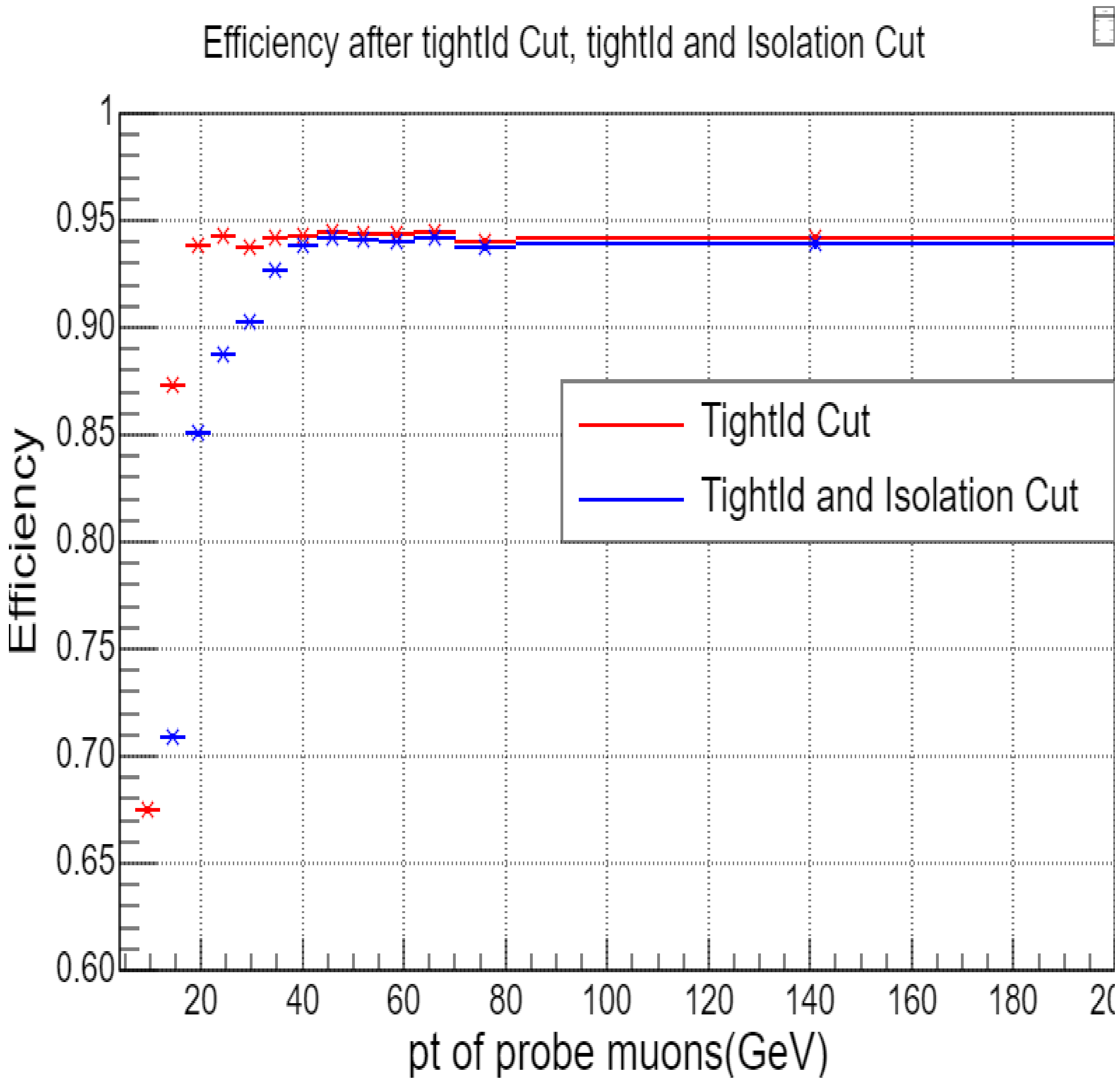
Efficiency of tight Id and Isolation Cut = $\frac{N_2}{N_0}$

But Efficiency depends on pt of the muons. High pt muons can be detected more efficiently.

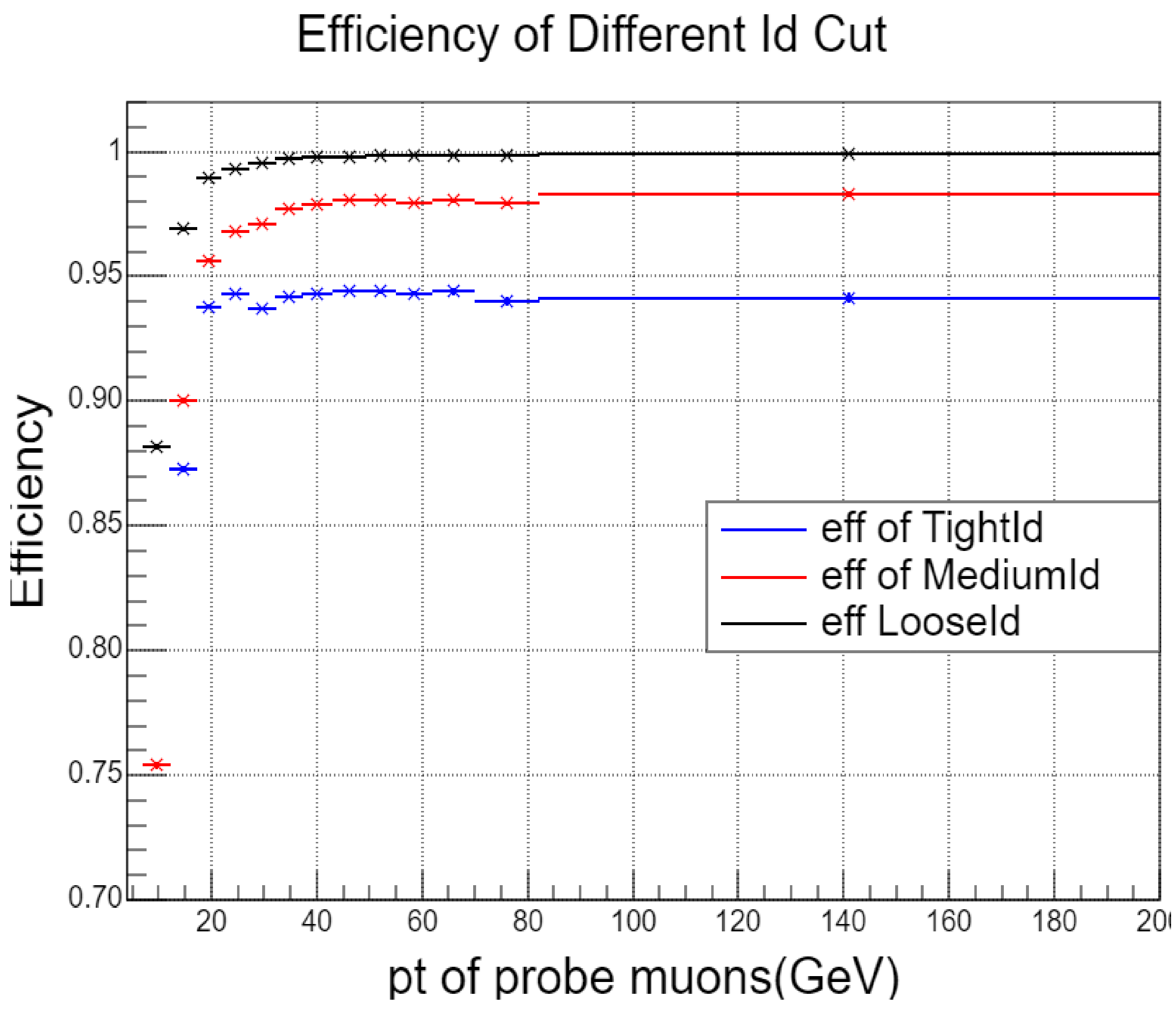
Hence we estimate the efficiency by plotting the muon pt histograms(after different cuts) and dividing the histograms.



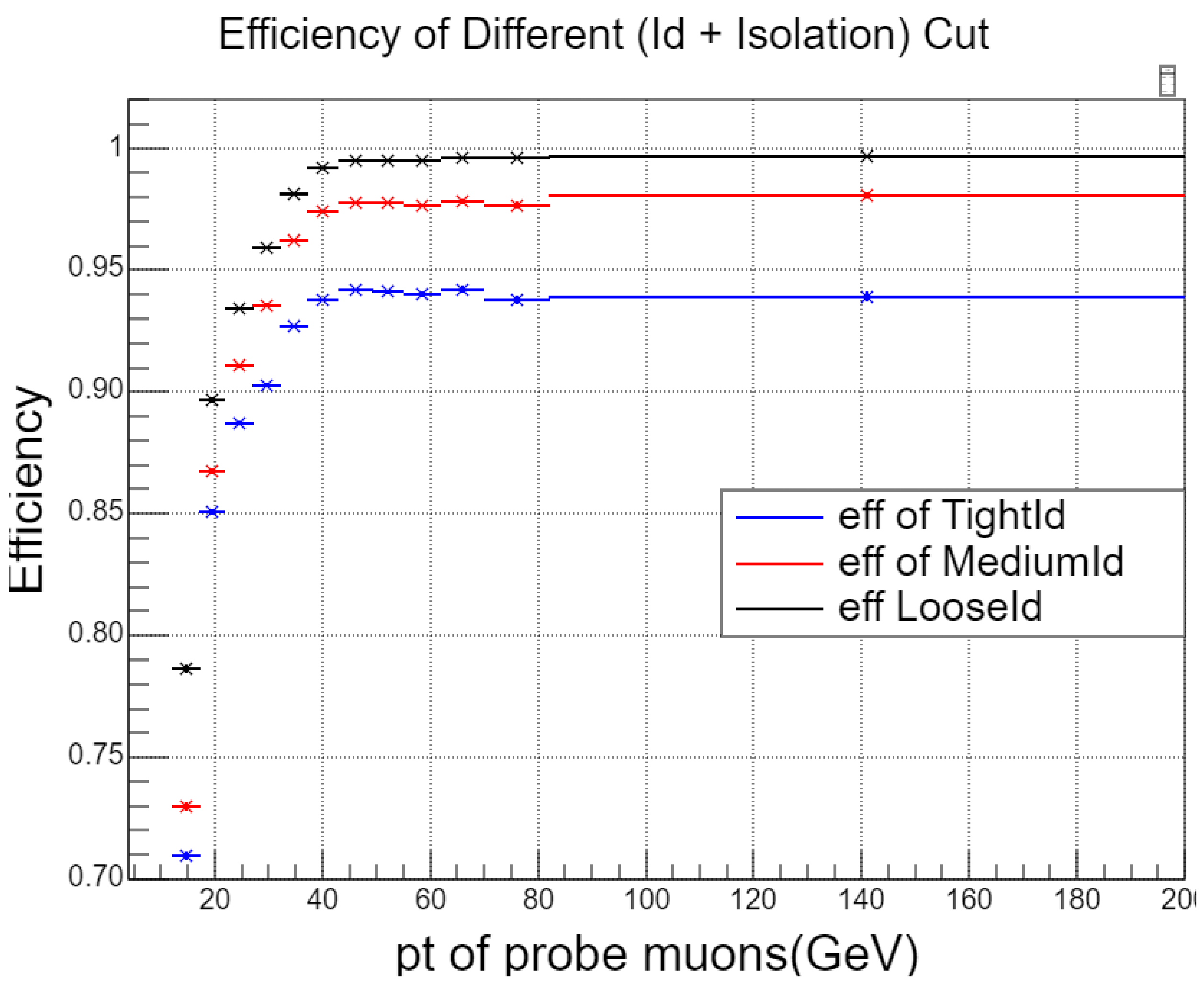
pt of probe muons after Reconstruction, tightId Cut, tightId and Isolation Cut
u_pt_C0Hr
Entries 1744262
u_pt_TC1Hr2
Entries 1320746
u_pt_TC2Hr2
Entries 1234491
Reconstruction
TightId Cut
TightId and Isolation Cut
Number of Events
pt of probe muons(GeV)



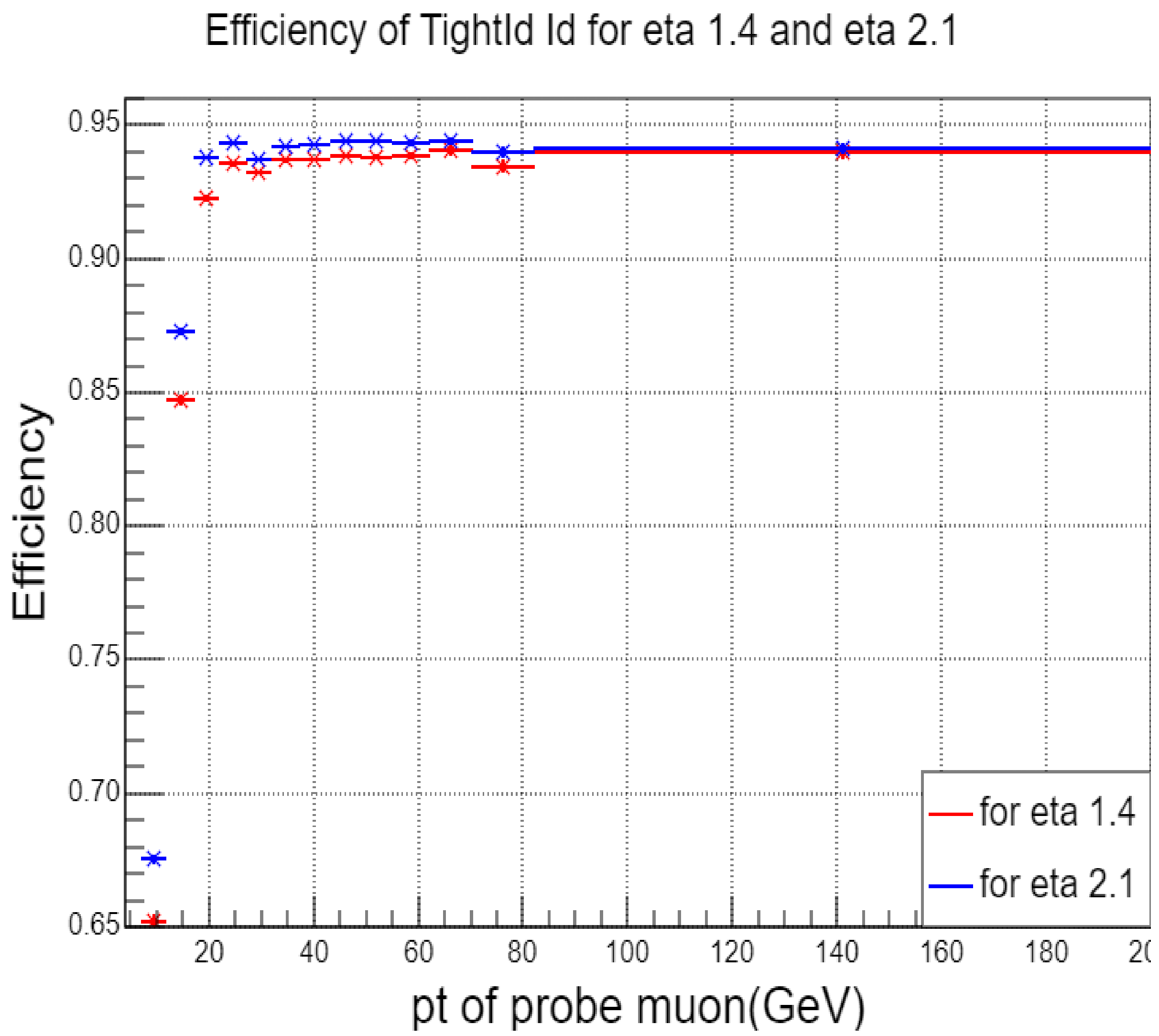
Efficiency after tightId Cut, tightId and Isolation Cut
TightId Cut
TightId and Isolation Cut
Efficiency
pt of probe muons(GeV)

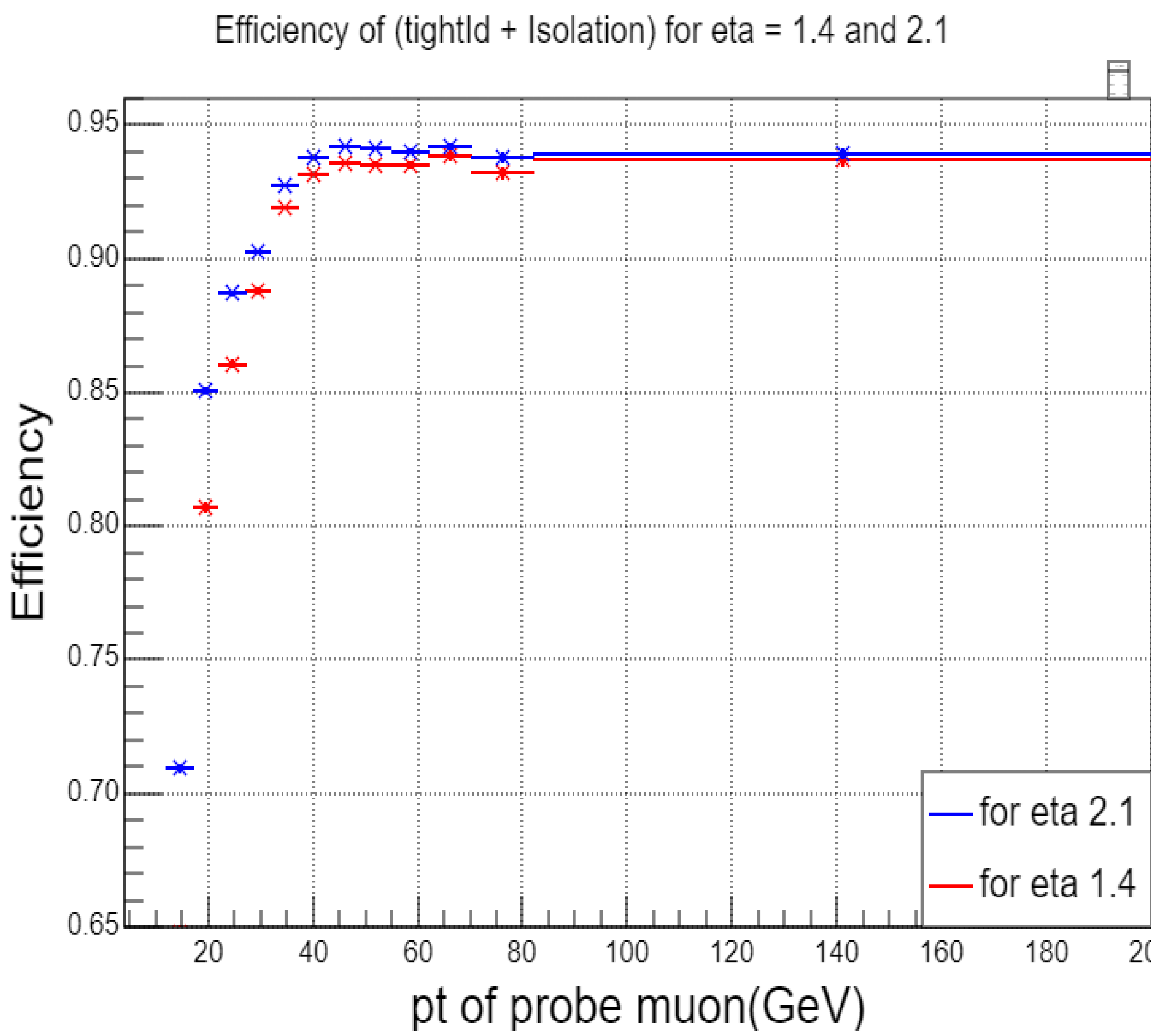


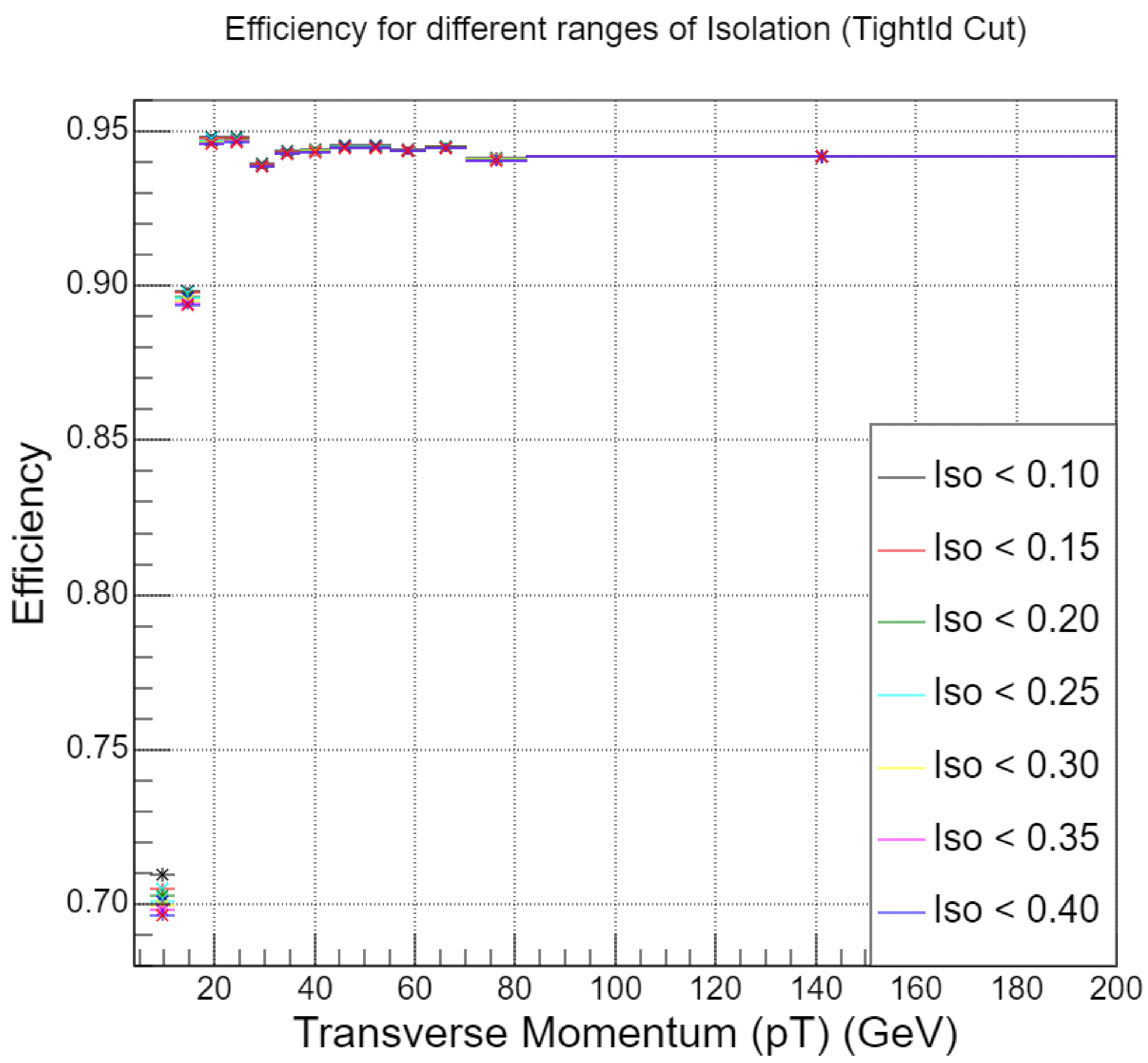
Efficiency of Different Id Cut
Efficiency
pt of probe muons(GeV)
eff of TightId
eff of MediumId
eff LooseId



Efficiency of Different (Id + Isolation) Cut
Efficiency
pt of probe muons(GeV)
eff of TightId
eff of MediumId
eff LooseId


Efficiency of TightId Id for eta 1.4 and eta 2.1
Efficiency
0.95
0.90
0.85
0.80
0.75
0.70
0.65
20
40
60
80
100
120
140
160
180
pt of probe muon(GeV)
for eta 1.4
for eta 2.1

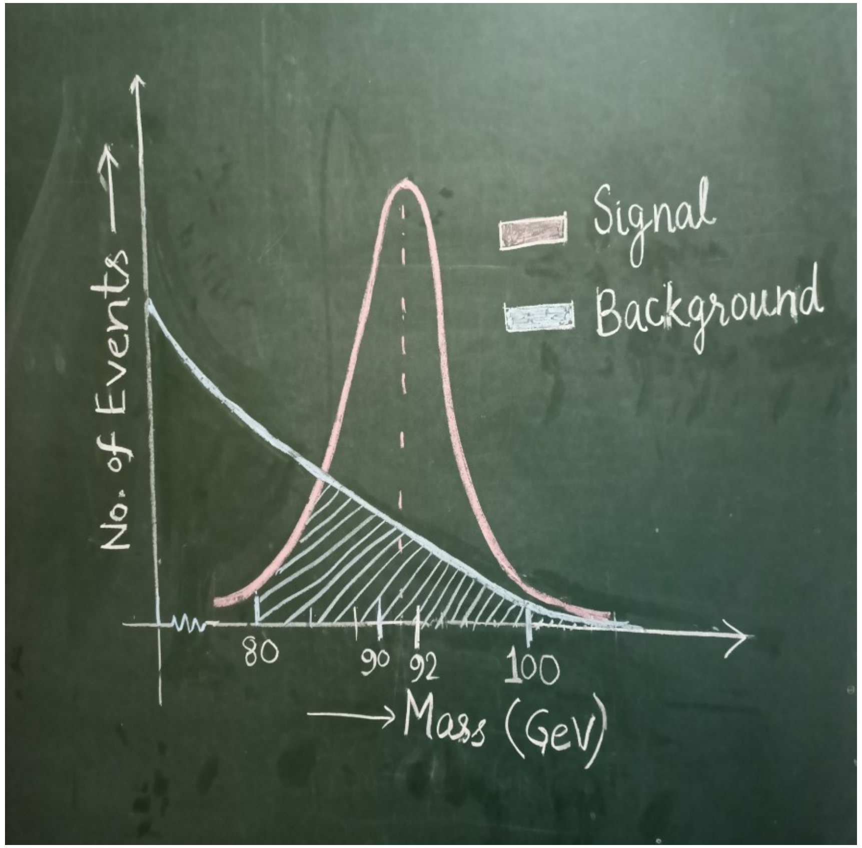

Efficiency of (tightId + Isolation) for eta = 1.4 and 2.1
Efficiency
0.95
0.90
0.85
0.80
0.75
0.70
0.65
20
40
60
80
100
120
140
160
180
pt of probe muon(GeV)
for eta 2.1
for eta 1.4



Efficiency for different ranges of Isolation (TightId Cut)
Efficiency
0.95
0.90
0.85
0.80
0.75
0.70
20
40
60
80
100
120
140
160
180
200
Transverse Momentum (pT) (GeV)
Iso < 0.10
Iso < 0.15
Iso < 0.20
Iso < 0.25
Iso < 0.30
Iso < 0.35
Iso < 0.40

Estimation of Efficiency: Method 2

- Although after the tag and probe method and event veto we manage to get a more pure sample of signal, we can't remove the background completely.
- It will be good if we model the background and signal and try to fit (using RooFit) the data.
- In order to do this we will have to use the invariant mass histograms (for different pt ranges) which will look like the given figure.
- The shape of the mass distribution is dominated by the background contribution.
- After the fit, we get no of signal events after each cut and calculate the efficiency which should be more reliable.



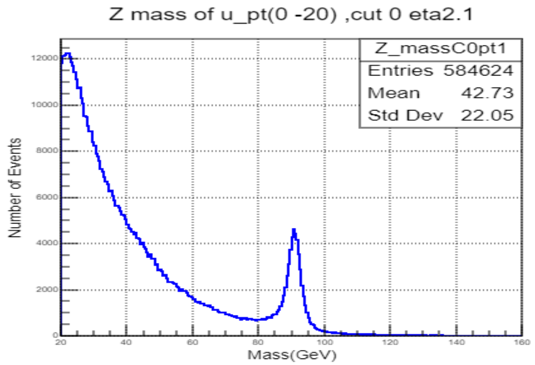
Signal and Background Modeling

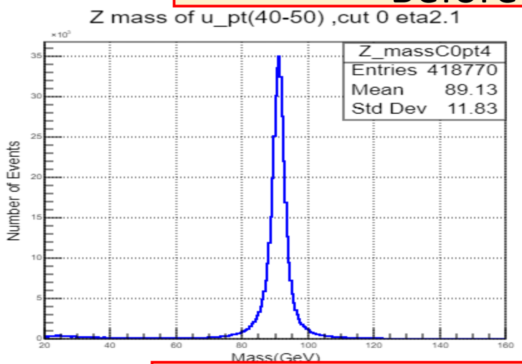
- RooFit
  - RooFit is a powerful toolkit in ROOT for advanced data modeling and fitting.
  - Enables precise signal and background modeling in high-energy physics analyses.
  - Offers various probability density functions (PDFs) for accurate fitting of data.
- Double-Sided Crystal Ball (DSCB)
  - DSCB is a specialized PDF used to model signal and background distributions.
  - Combines Gaussian and power-law functions, suitable for asymmetric peaks.
  - Captures resonant signals and smoothly describes non-Gaussian backgrounds.
- Signal Modeling
  - In signal modeling, DSCB represents the characteristic peak of the physics process.
  - DSCB parameters adjust to the observed mass peak, width, and tail behavior.
- Background Modeling
  - In background modeling, DSCB describes non-Gaussian or tails of distributions.
  - Fits complex background shapes with smooth transitions to the signal region.

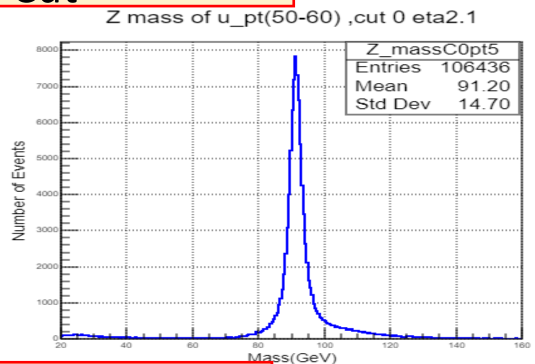
Invariant mass histogram of Z particle (Tag and Probe pair) for different pt ranges

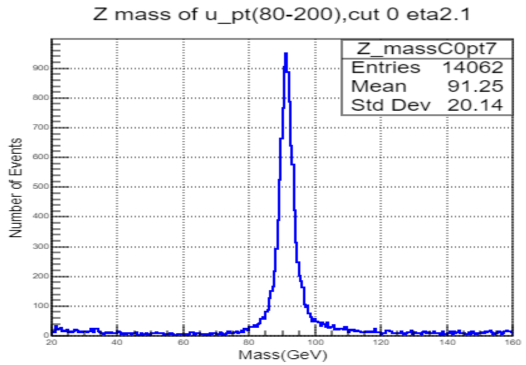
pt range (0-20) GeV | pt range (40-50) GeV | pt range (50-60) GeV | pt range (80-200) GeV

Before any Cut

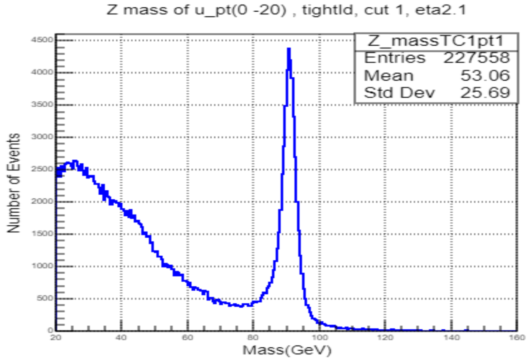


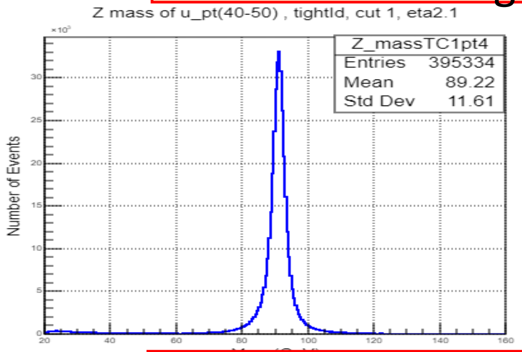


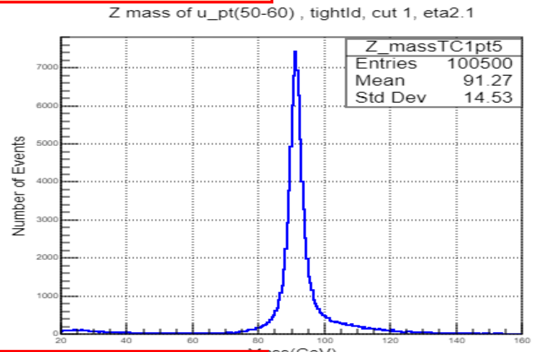


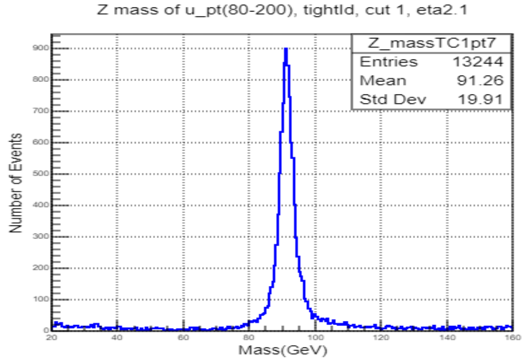


After Tight Id Cut

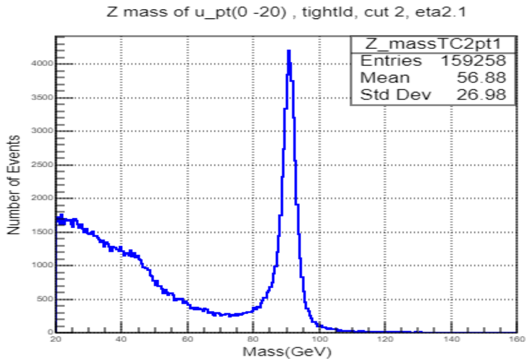


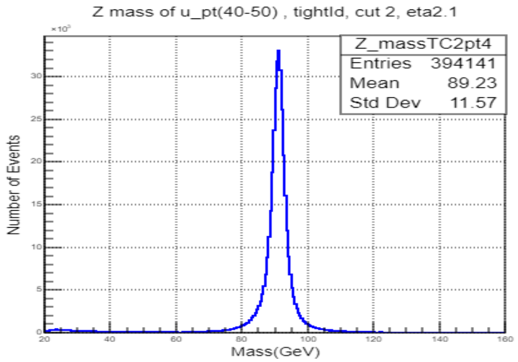


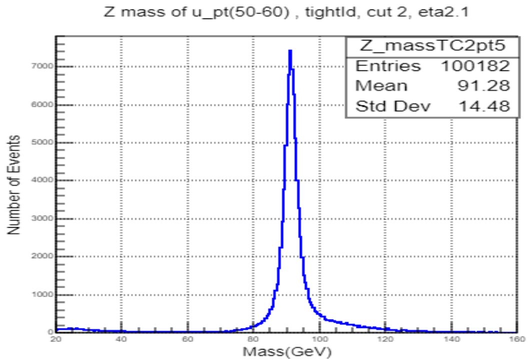


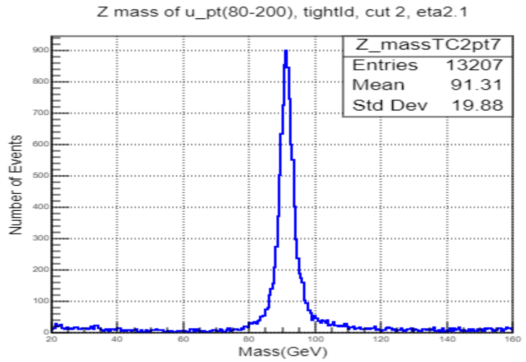


After (Tight Id + Isolation) Cut





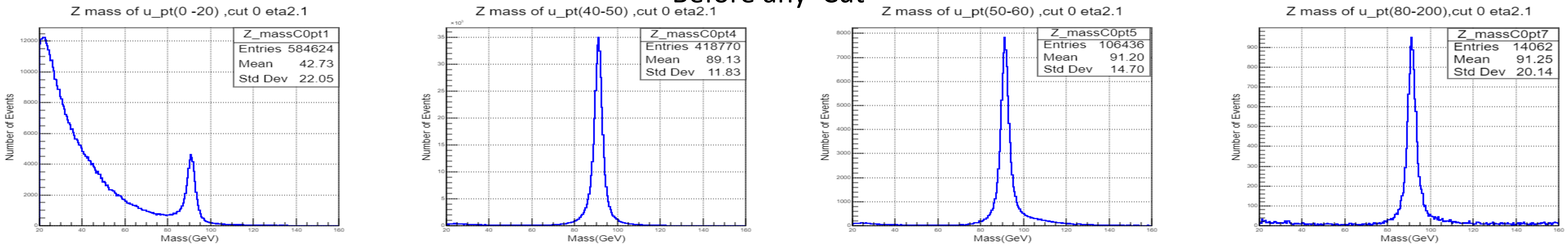




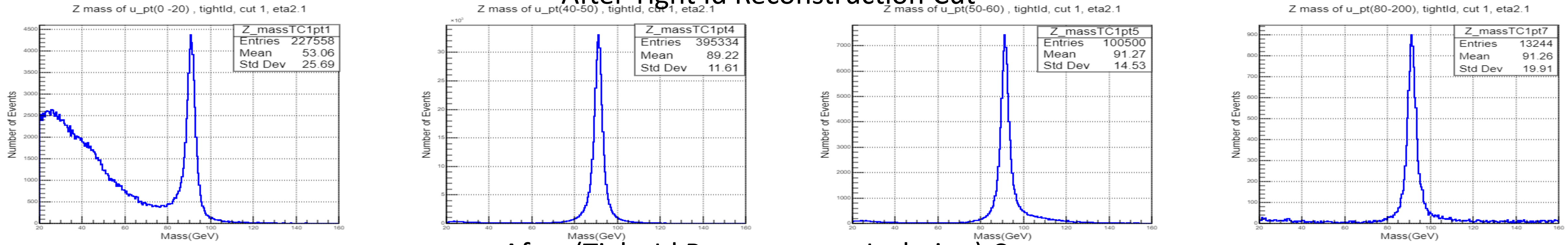
Invariant mass histogram of Z particle (Tag and Probe pair) for different pt ranges

pt range (0-20) GeV | pt range (40-50) GeV | pt range (50-60) GeV | pt range (80-200) GeV

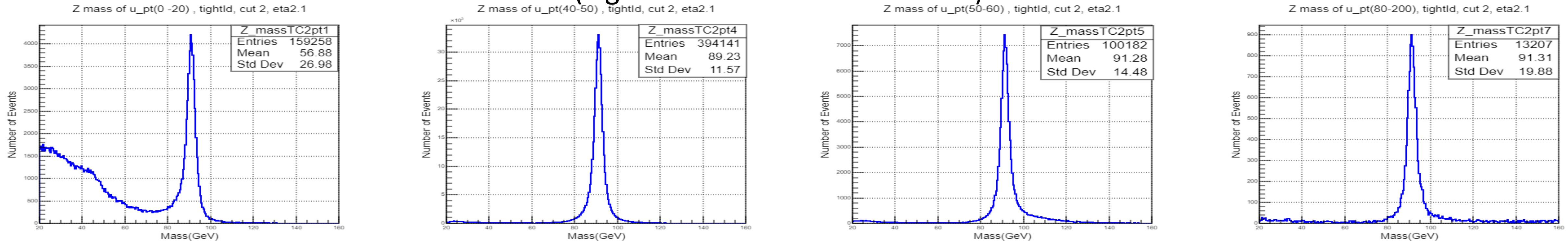
Before any Cut



After Tight Id Reconstruction Cut

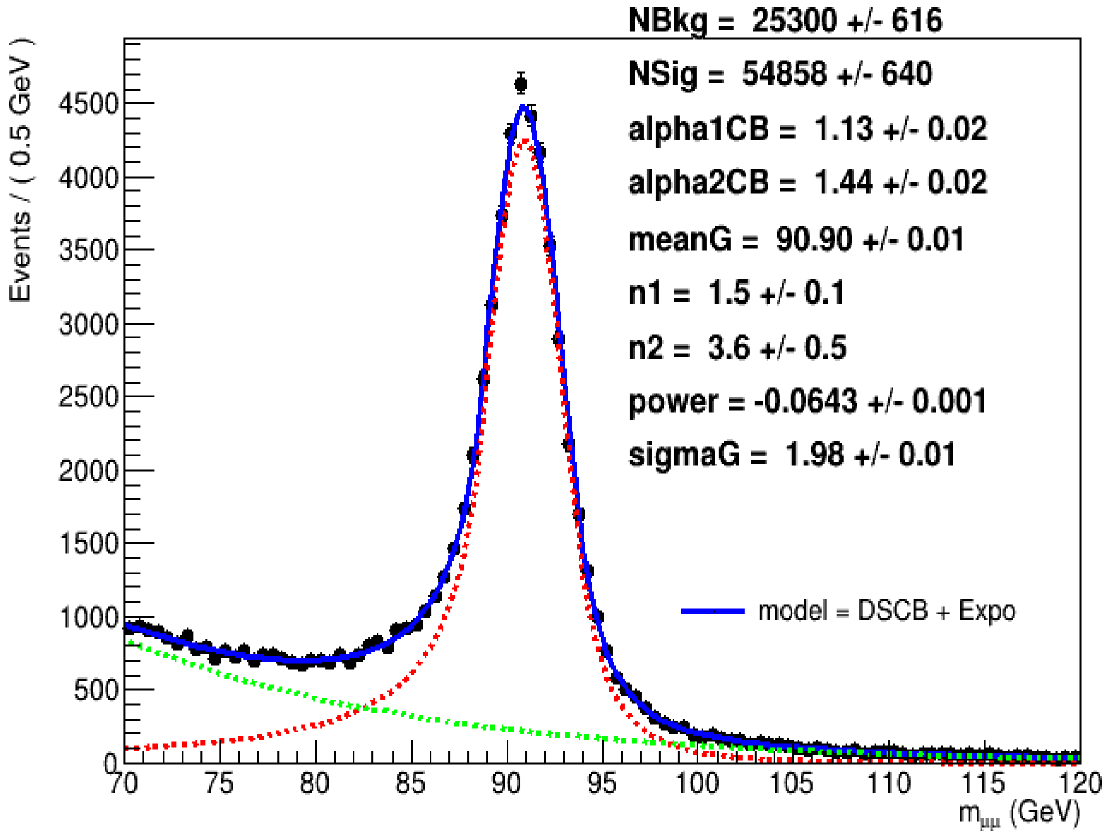


After (Tight Id Reconstruct + Isolation) Cut

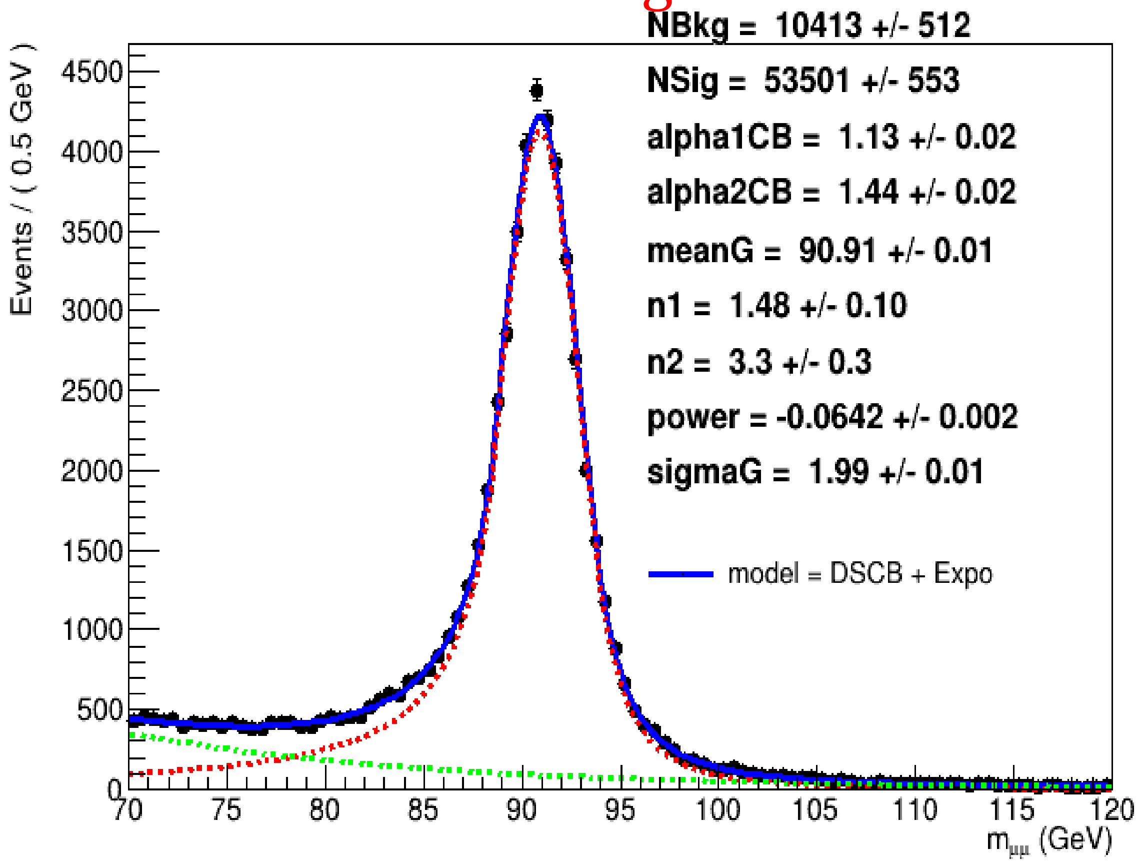


Signal and Background: probe pt range(0-20)

Before any cut on probe muon

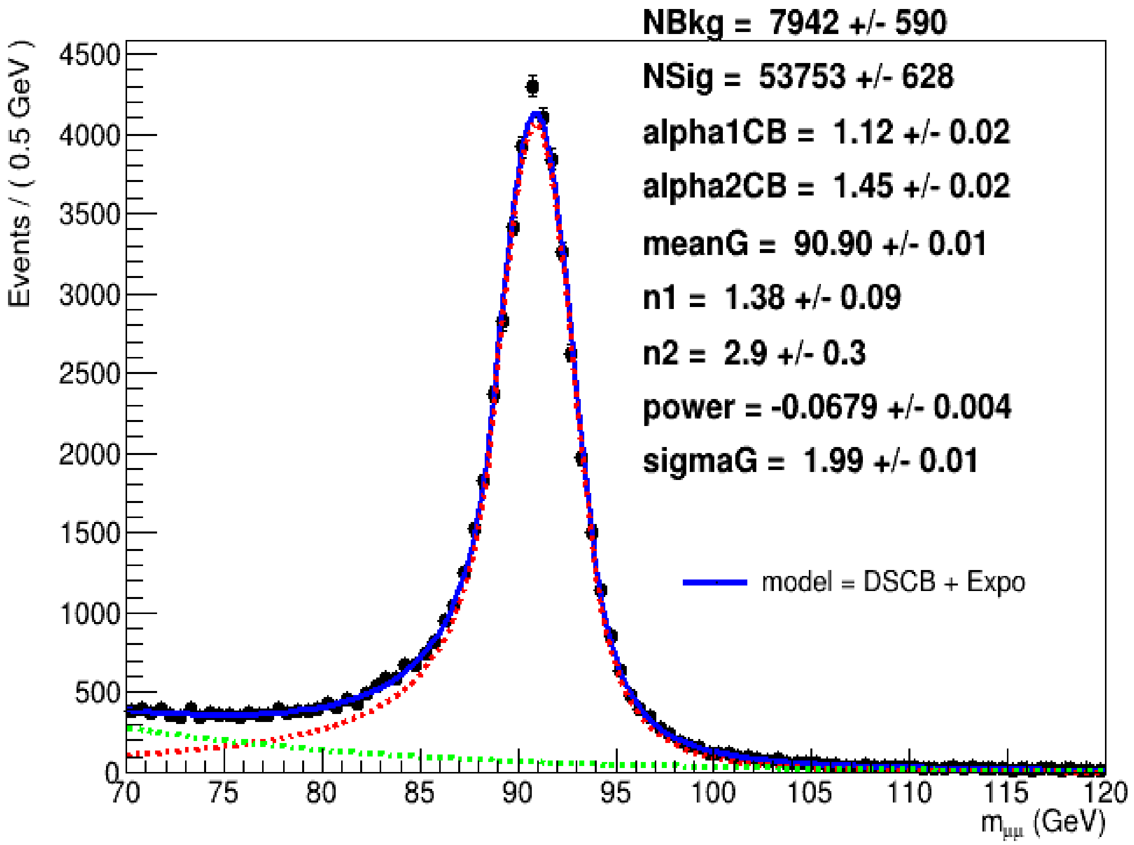


After TightId Cut



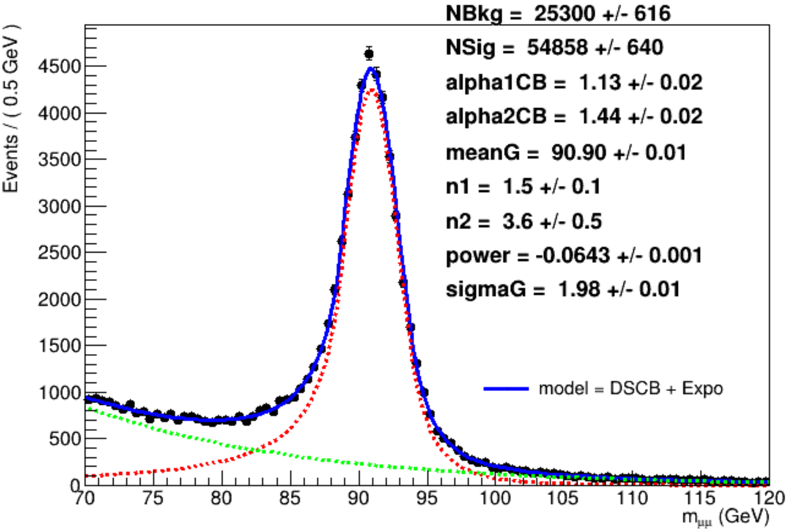
Signal and Background: probe pt range(0-20)

After (tightId + Isolation) cut

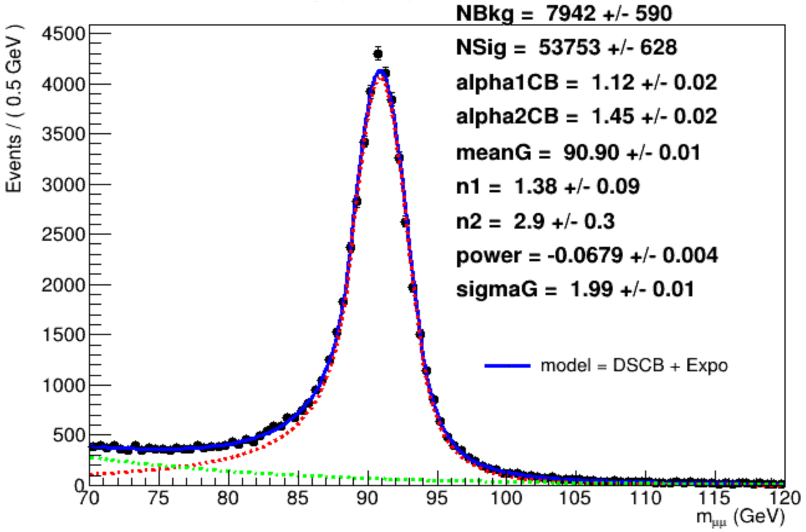


Signal and Background: probe pt range(0-20)

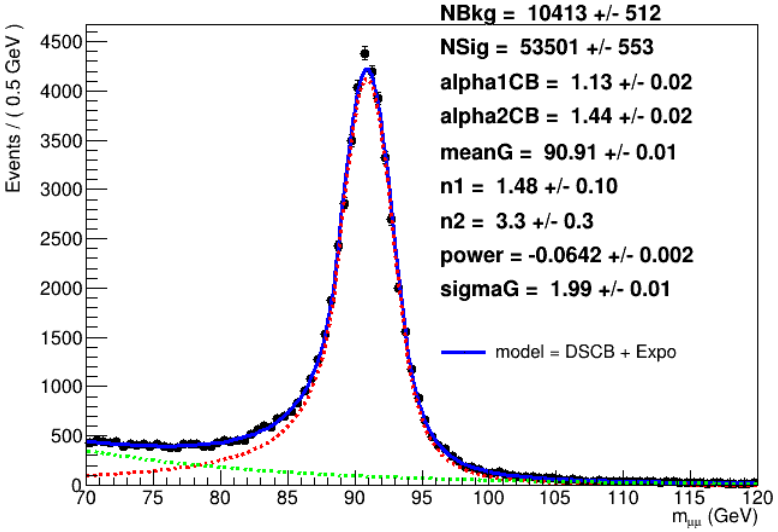
Before any cut on probe muon



After Tight ID Reconstruction Cut

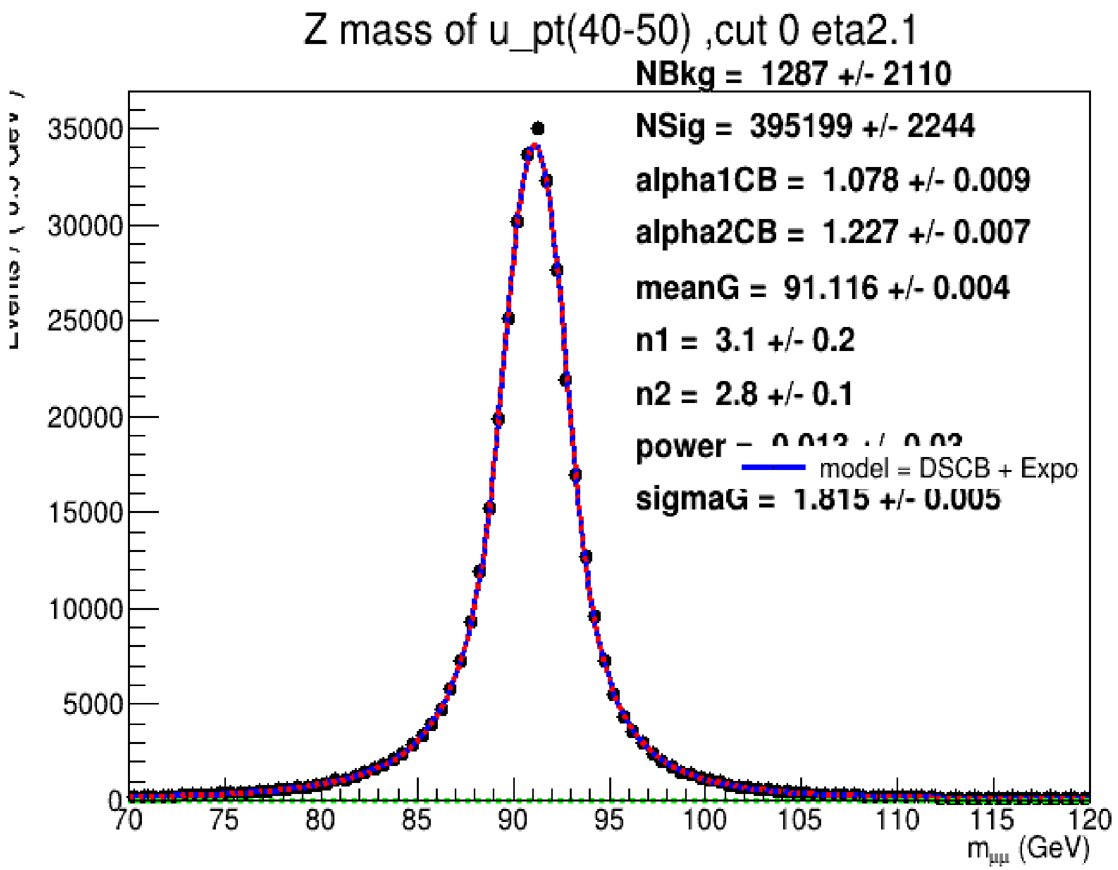


After(Recons + Isolation) Cut

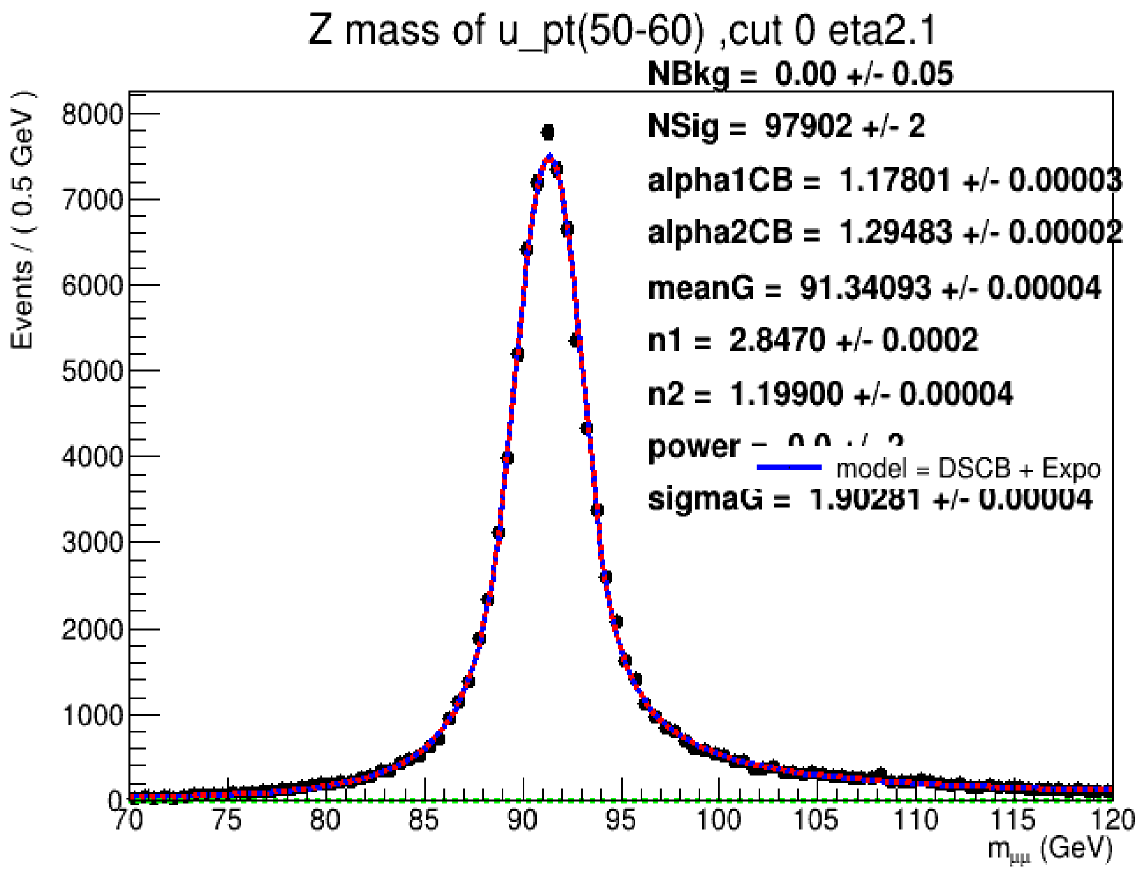


Signal and Background: probe pt ranges (40-50)&(50-60)

Before any cut on probe muon



Before any cut on probe muon



Efficiency for different pt ranges calculated by fitting the mass histogram

pt Range(GeV)	Cut 0	Ns	TightId Cut 1	Ns	TightId Cut 2	Ns	T_eff10	T_eff20
0-20	Z_massC0pt1	54858	Z_massTC1pt1	53501	Z_massTC2pt1	52968	0.975	0.965
20-30	Z_massC0pt2	148307	Z_massTC1pt2	140487	Z_massTC2pt2	142784	0.947	0.962
30-40	Z_massC0pt3	329268	Z_massTC1pt3	311223	Z_massTC2pt3	307709	0.945	0.934
40-50	Z_massC0pt4	395199	Z_massTC1pt4	369996	Z_massTC2pt4	374229	0.936	0.946
50-60	Z_massC0pt5	97902	Z_massTC1pt5	92630	Z_massTC2pt5	92402	0.946	0.943
60-80	Z_massC0pt6	40352	Z_massTC1pt6	36978	Z_massTC2pt6	38069	0.916	0.943
80-200	Z_massC0pt7	11498	Z_massTC1pt7	9216	Z_massTC2pt7	11287	0.801	0.981
							0.924	0.954

Conclusion

- The objective of this project was to measure the efficiency of muon detection using the Tag and Probe method using CMS data.
- By measuring efficiency we can evaluate detectors' capabilities and identify areas of improvement, cross section measurement, look for new physics.
- Observations
  - Efficiency of loose Id cut was highest and that of tight Id cut was lowest
  - Efficiency after isolation cut was not affected too much (slight change)
  - Efficiency was slightly higher for higher value of eta
  - Background contribution in the low pt range was significant but in higher pt range it was low.
- I have learnt basic Particle Physics, Python, ROOT software package, PyRooT (Root interface for python), RooFit etc. during the project

Acknowledgement

"I would like to express my gratitude to the mentor Prof. Subir Sarkar , the colleagues Kuldeep Nishad and Ratul Sarkar and lab mates Dr. Pritam Palit and Suman Dasgupta for their support and guidance throughout this project."

Thank you !!

Back-Up

DSCB

- The Double-Sided Crystal Ball (DSCB) is a probability density function (PDF) used in data analysis, particularly in high-energy physics, to model signal and background distributions. It is a specialized PDF that combines a Gaussian core with power-law tails on both sides, making it suitable for modeling asymmetric peaks often observed in particle physics experiments.

- **x** is the variable being modeled (e.g., invariant mass in a physics analysis).
- **α** is the threshold parameter defining where the Gaussian core transitions to the power-law tails.
- **n** is the power-law index, controlling the tail behavior.
- **μ** is the mean (peak) of the Gaussian core.
- **σ** is the standard deviation (width) of the Gaussian core.
- **N** is a normalization constant ensuring that the entire function integrates to unity.
- **A** and **B** are constants defined in terms of **n** and **α** to ensure a smooth transition between the Gaussian core and the power-law tails.

- By fitting the DSCB to experimental data, physicists can extract important parameters such as the signal peak position, width, and the background's tail behavior, contributing to precise measurements and searches for new particles or phenomena.