Measurement of muon efficiencies in CMS open data

Tag and Probe method

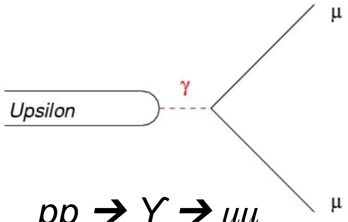
Sep. 11, 2020
LIP Internship 2020

Authors: Allan Jales and Thomas Gaehtgens
Advisors: Eliza Melo, Nuno Leonardo and Sandro Fonseca

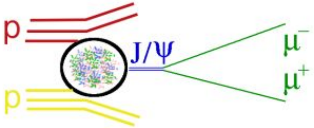
Outline

- A brief overview of the work;
- CMS Open Data;
- Tag & Probe workflow;
    - Tag & Probe method efficiency (ε);
    - Muon identification;
- Evaluating efficiencies;
- Signal extraction methods evaluation;
    - Sideband subtraction & results;
    - Fitting method & results;
- Comparison between sideband & fitting results;
- Summary and Next steps.

A brief overview of the work



$pp \rightarrow \Upsilon \rightarrow \mu\mu$



$pp \rightarrow J/\psi \rightarrow \mu\mu$

CMS Open Data

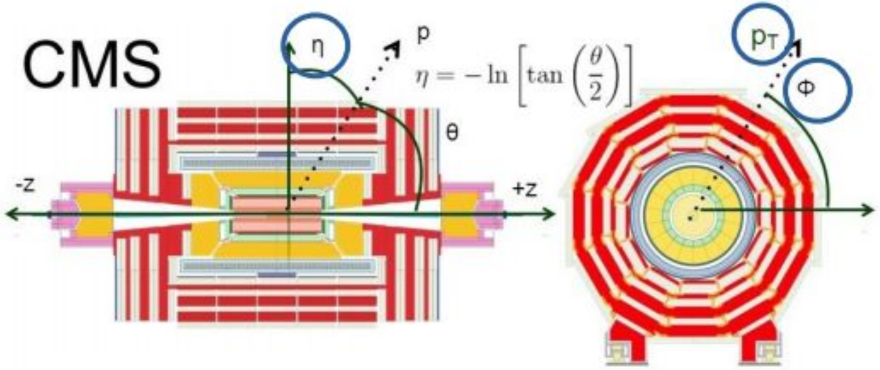
- Data (2011)
- Simulated Data

Tag & Probe →

efficiency (ε)

As function of:

- Transverse Momentum (p_T)
- Pseudorapidity (η)
- Azimuthal angle (ϕ)

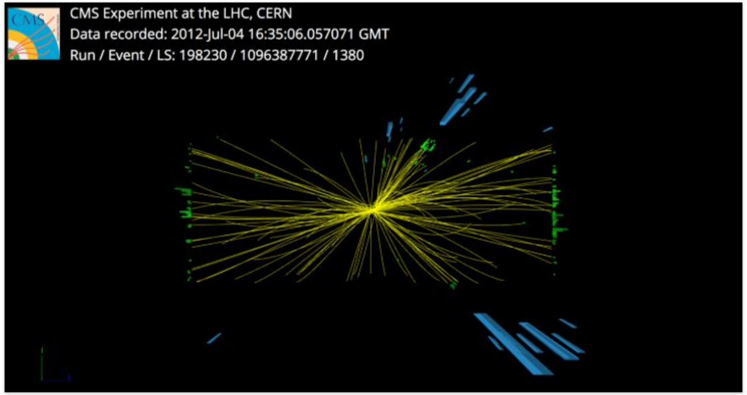


CMS Open Data

CMS releases more than one petabyte of open data

This release includes datasets that were used to discover the Higgs boson

20 DECEMBER, 2017 | By Achintya Rao

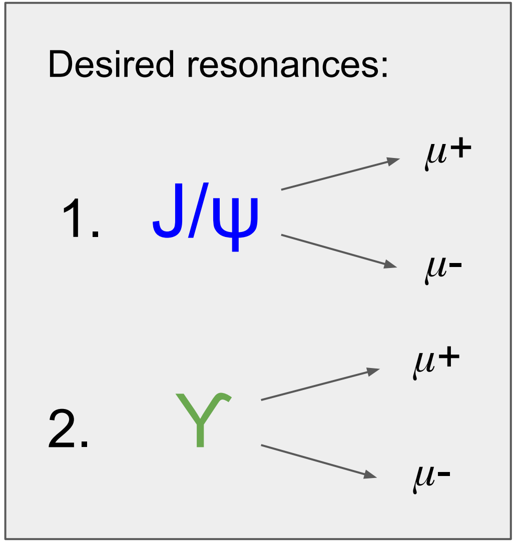


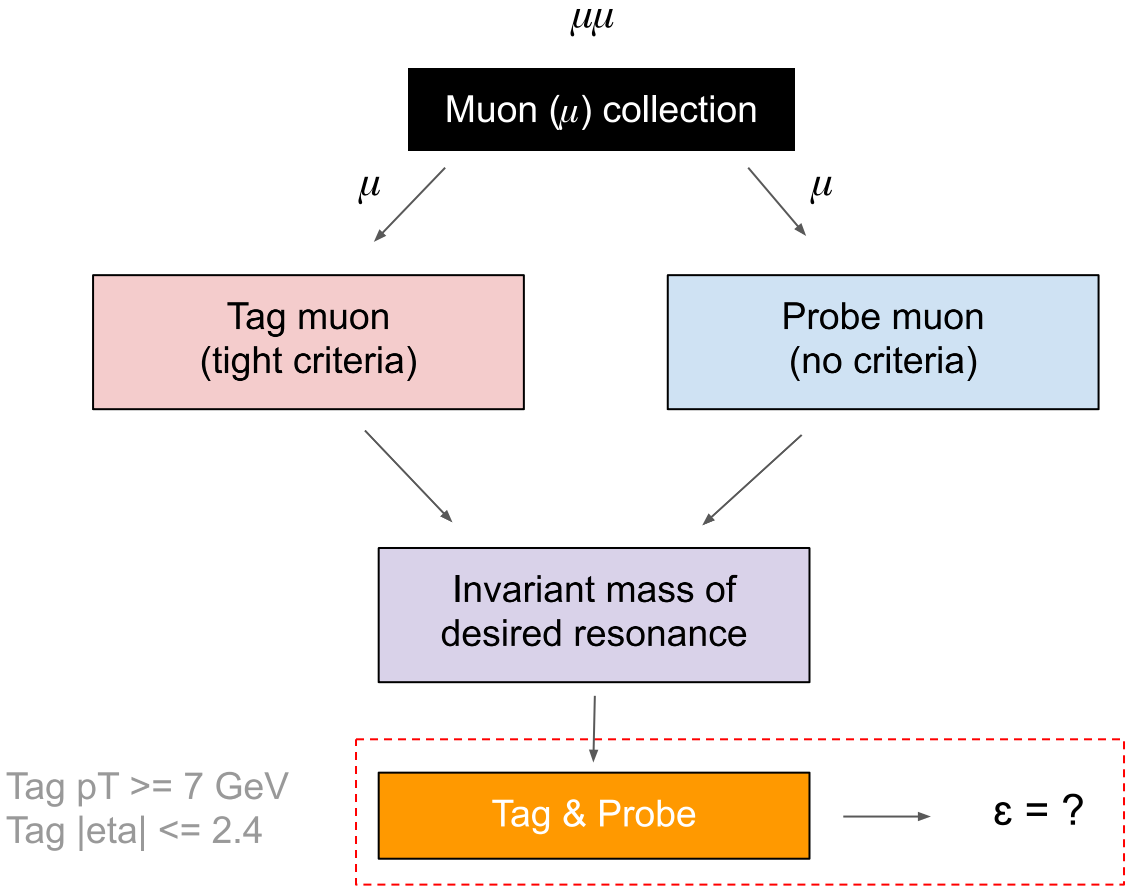
A collision event recorded by CMS in 2012 showing a "Higgs candidate", available on the CERN Open Data portal with the latest release of CMS Open Data. (Image: Tom McCauley/CMS/CERN)

The CMS Open Data is an initiative that makes scientific data from the CMS experiment available to the public.
With data transparency It tries to make science more inclusive and open.

https://home.cern/news/news/experiments/cms-releases-more-one-petabyte-open-data

Tag & Probe workflow





Tag & Probe method efficiency (ε)

The Tag & Probe method consists in calculating the efficiency of a detector using gathered data without the dependency on the simulations.
This method uses known resonances (e.g J/ψ,ϒ,Z)

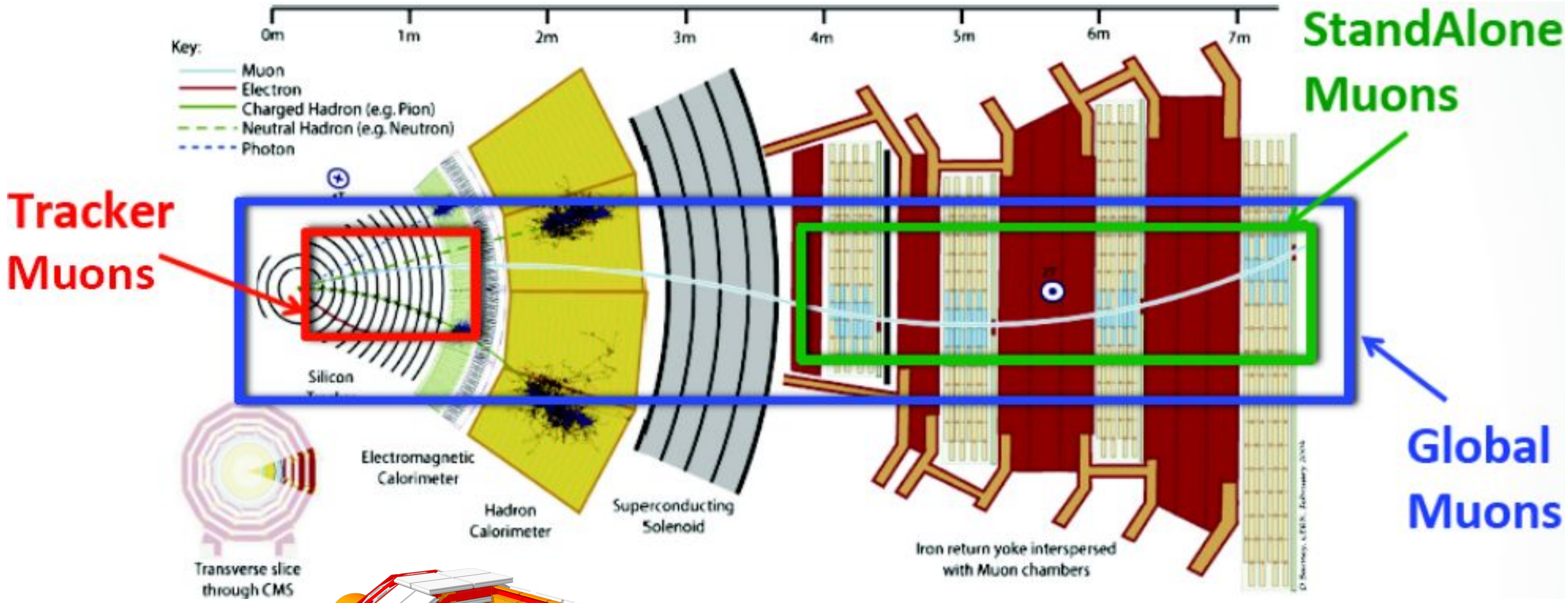
Tag muon = well identified, triggered muon (tight selection criteria)

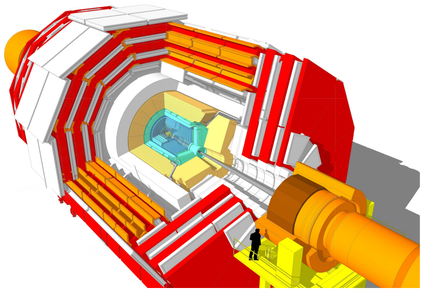
Probe muon = unbiased set of the desired particle type with a very loose selection criteria

$$\varepsilon = \frac{Passing\ probe\ muon\ criteria}{All\ probe\ muon}$$

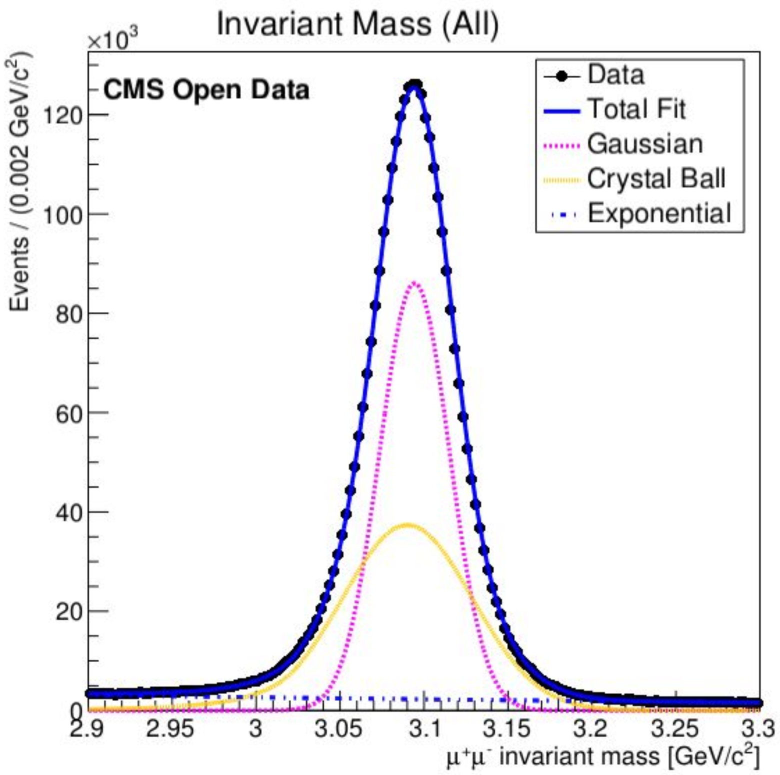
→ Tracker muon

Muon identification





Evaluating efficiencies

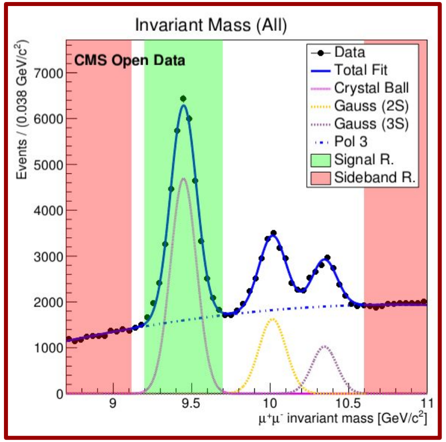


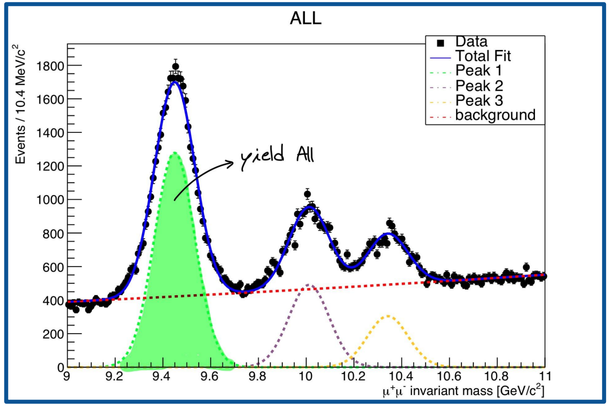
The efficiency is calculated using only **signal**
For example using J/ψ invariant mass distribution we have:

Signal = Gaussian + Crystal Ball

How can we extract signal quantities probe muon (p_T, η, ϕ)?

Signal extraction methods evaluation



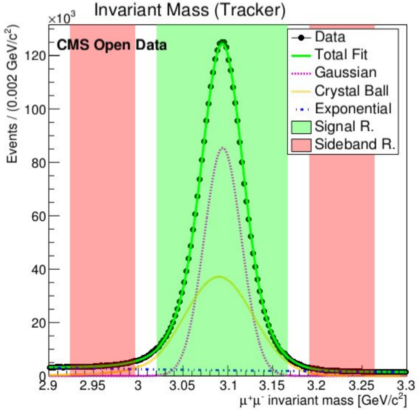


We would like to evaluate two signal extraction methods:

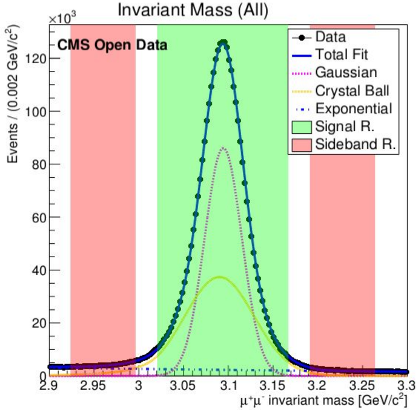
- Sideband subtraction
- Fitting

Extracting signal: sideband subtraction using J/ψ candidates

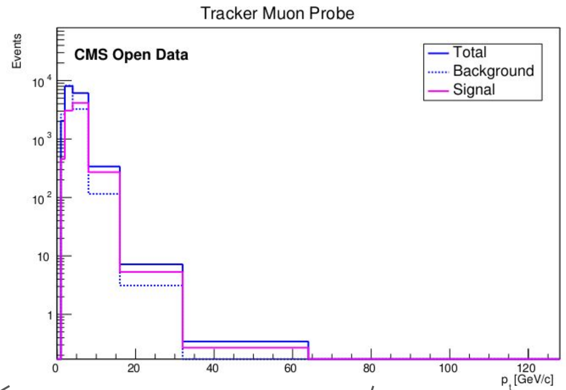
(tag muon + **passing** probe)



(tag muon + **any** probe)



example for tracker probes p_T



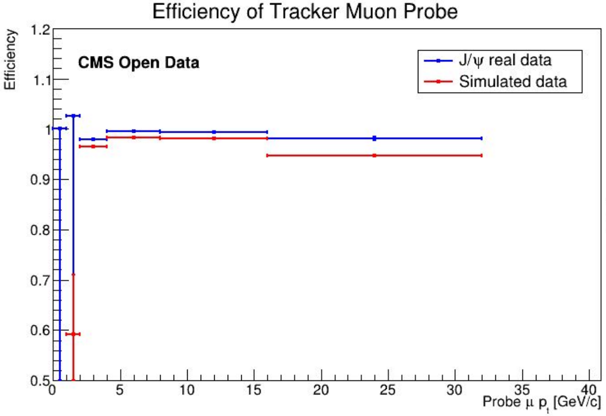
signal region → Total (signal + background)

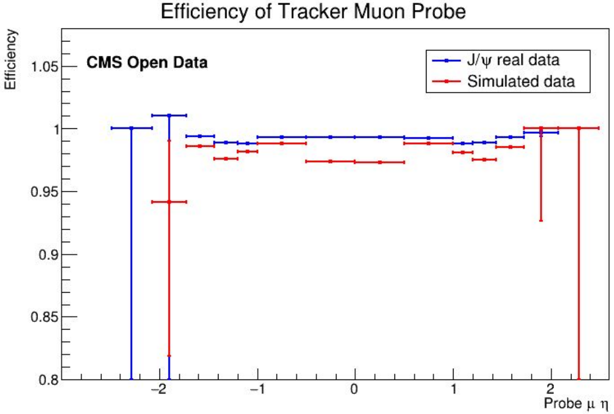
sideband regions → - - - Background (background)

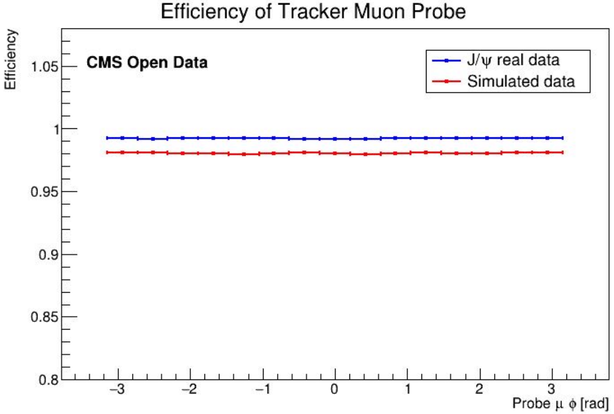
Signal = Total - α x Background

$$\alpha = \frac{yield\ of\ background\ in\ fit\ region}{yield\ of\ background\ in\ sideband\ region}$$

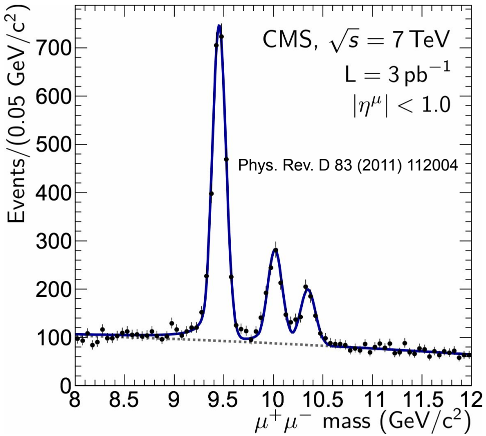
Results: efficiencies for J/ψ using sideband subtraction



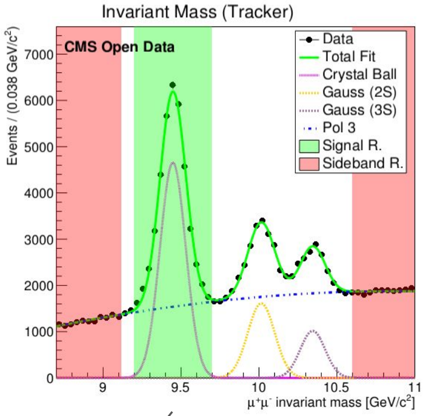




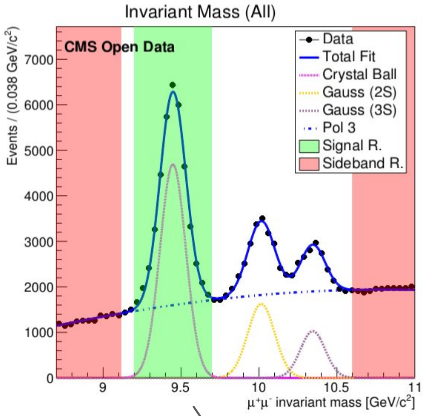
Extracting signal: sideband subtraction for ϒ (1S)

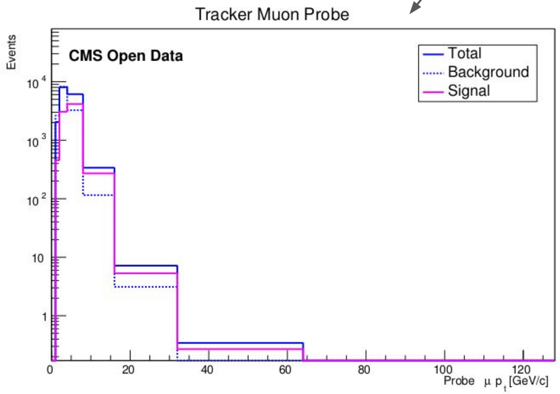


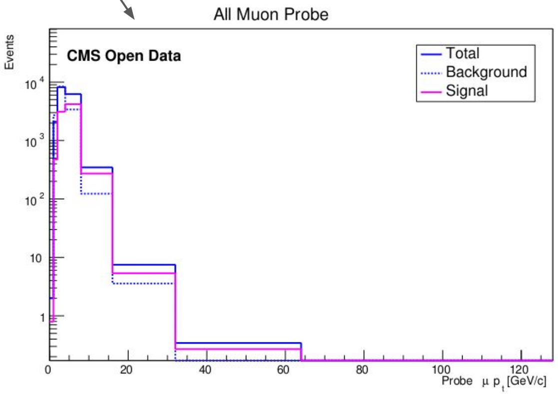
(tag muon + **passing** probe)



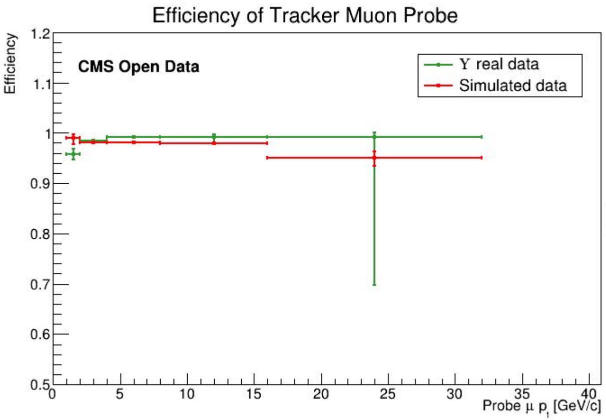
(tag muon + **any** probe)

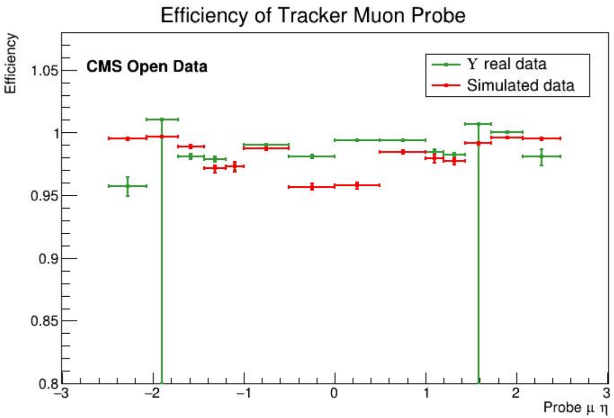


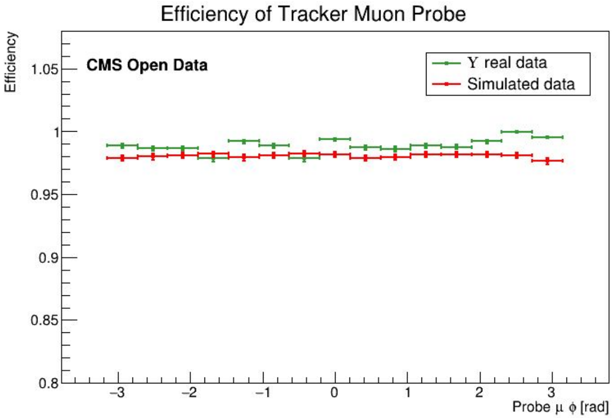




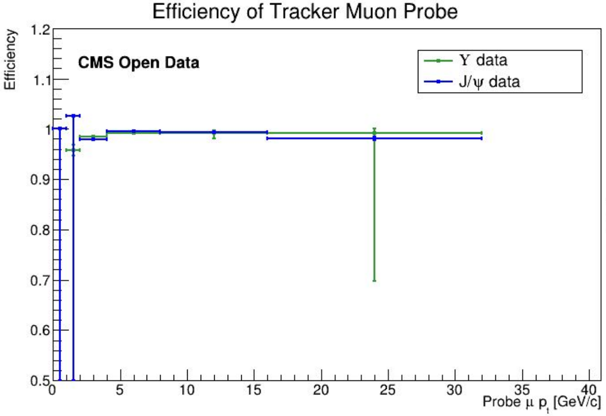
Results: efficiencies for Υ (1S) using sideband subtraction

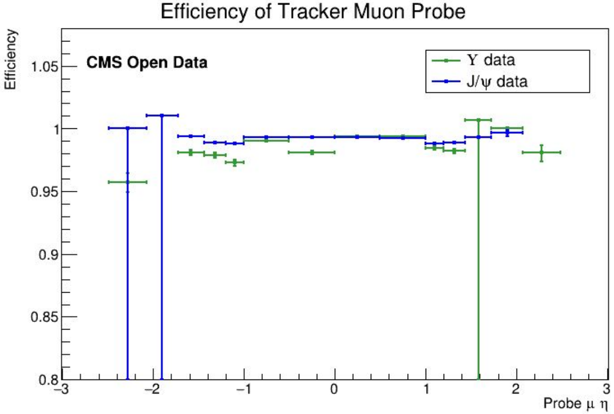


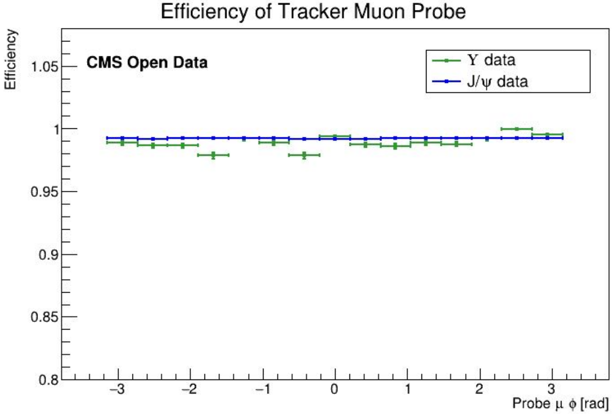




Results: comparison between J/ψ and ϒ (1S) using sideband subtraction

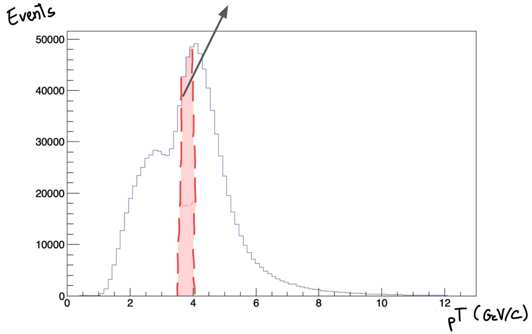






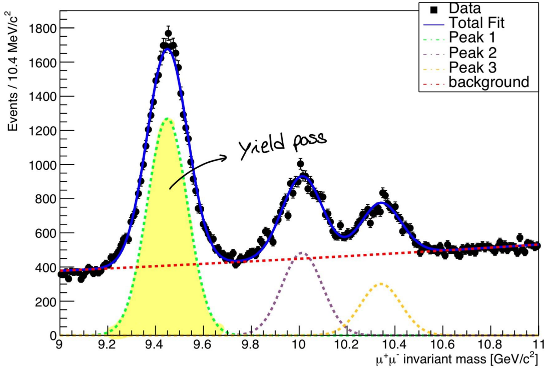
Extracting signal: Fitting method

Select region

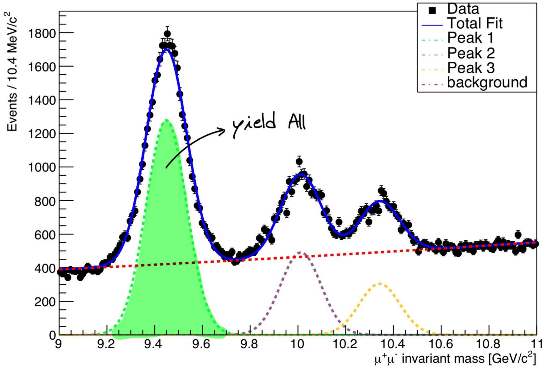


Fit Invariant Mass on Selected Region

PASS



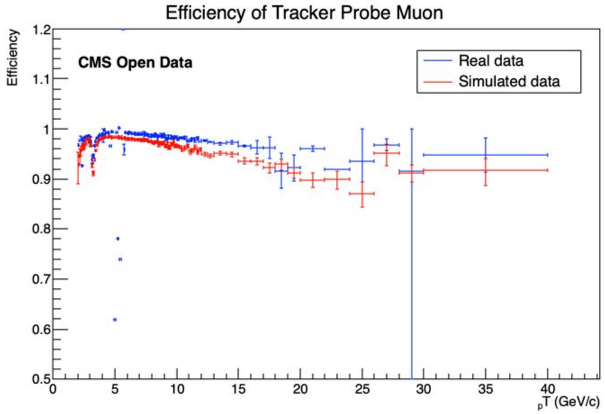
ALL

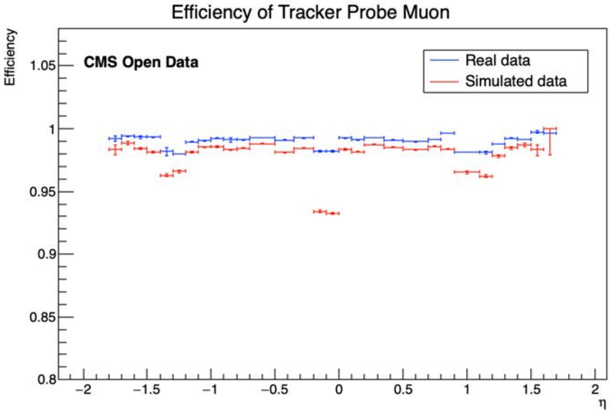


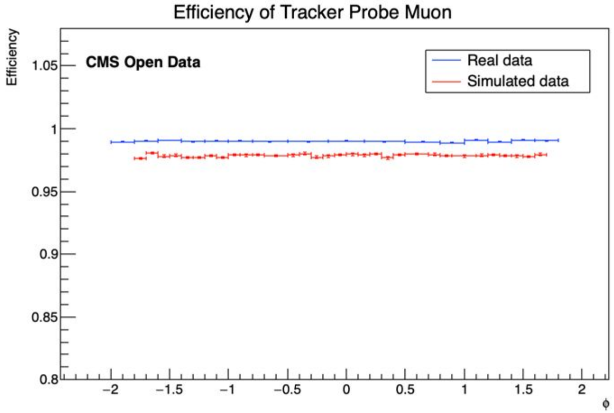
Compute Efficiency

$$\varepsilon = \frac{\text{Yield Pass}}{\text{Yield All}}$$

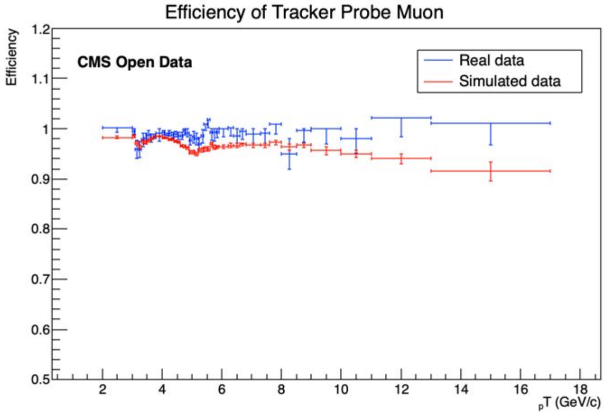
Results: efficiencies for J/ψ using fitting method

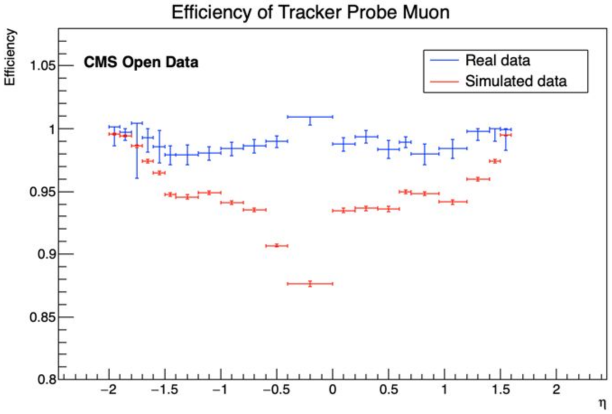


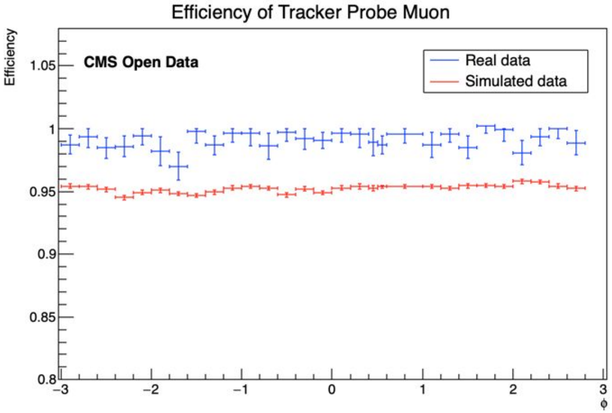




Results: efficiencies for Υ (1S) using fitting method

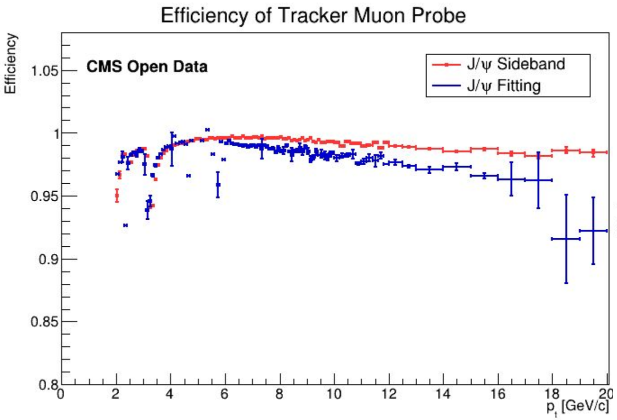


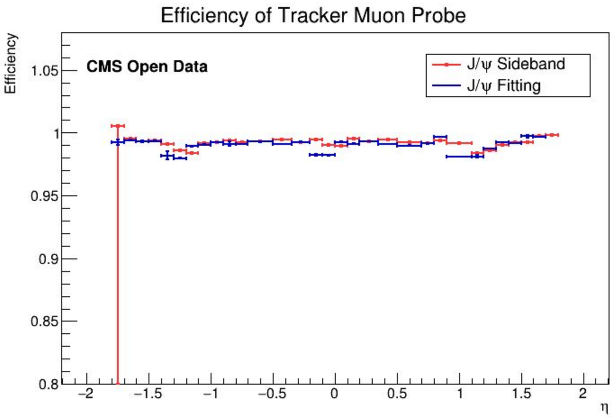


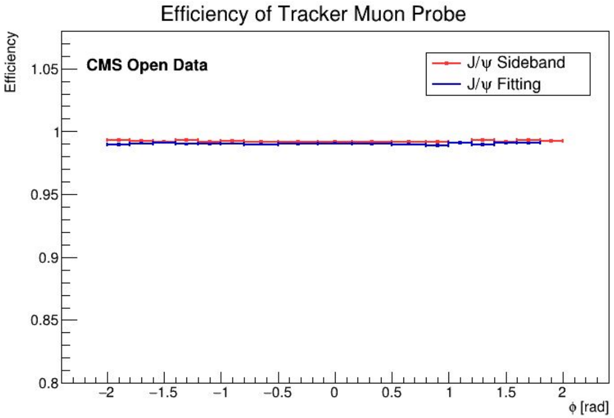


Comparison between both methods

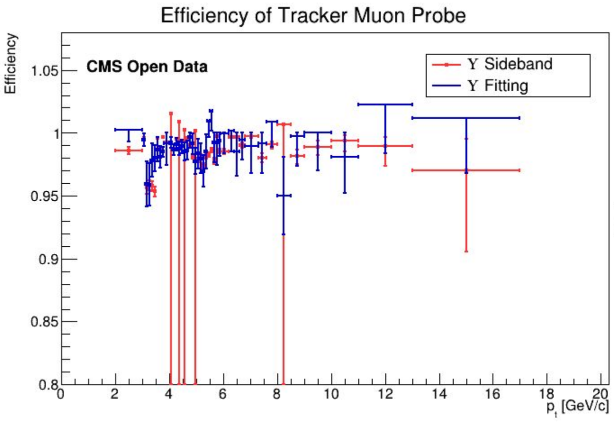
Results: comparison between sideband subtraction and fitting for J/ψ

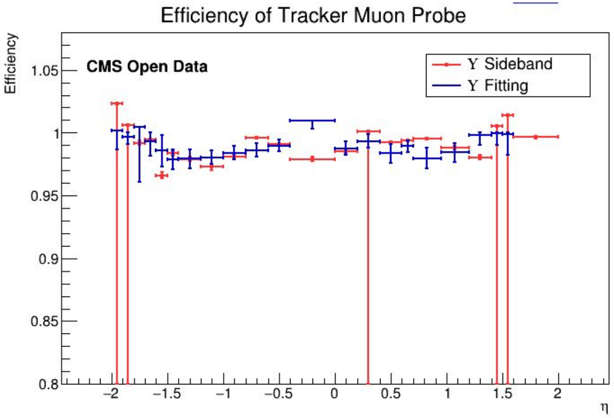


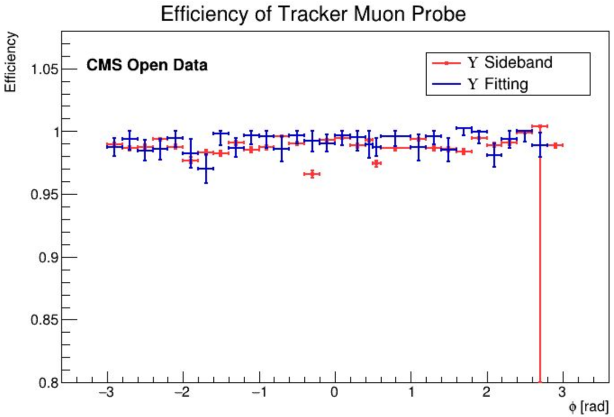




Results: comparison between sideband subtraction and fitting for Υ (1S)







Summary

- We developed a tool to evaluate tag & probe method using CMS Open Data for J/ψ and ϒ decaying in dimuons;
- We measured efficiency for tracker muon ID making a comparison between data and MC;
- Tested sideband subtraction and fitting method to extract signal;
  - At first look, fitting seems to be a better method;

Next steps

- Use standalone and global muon IDs;
- Optimize the code for CMS Open Data users;
- Implement s-Plot method;
- Evaluate efficiency for Z boson
- Improvement propagating uncertainties in sideband method

References

Our Workspace:

https://cern.ch/cms-lip/internship/tnp

Githubs:

https://github.com/allanjales/tag-probe

https://github.com/AthomsG/LIP_INTERNSHIP

Used Ntupples:

shorturl.at/dqv45

Tag and probe Wiki:

shorturl.at/imILM

References

Datasets:

- Jpsi MC: http://opendata.cern.ch/record/1335
- Upsilon MC: http://opendata.cern.ch/record/1522
- Data (2011 Legacy): http://opendata.cern.ch/record/27

CMS Open data:
http://opendata.cern.ch/search?experiment=CMS

HLT Triggers Path (J/ψ - Data/MC):
➢ HLT_Dimuon10_Jpsi_Barrel_v*

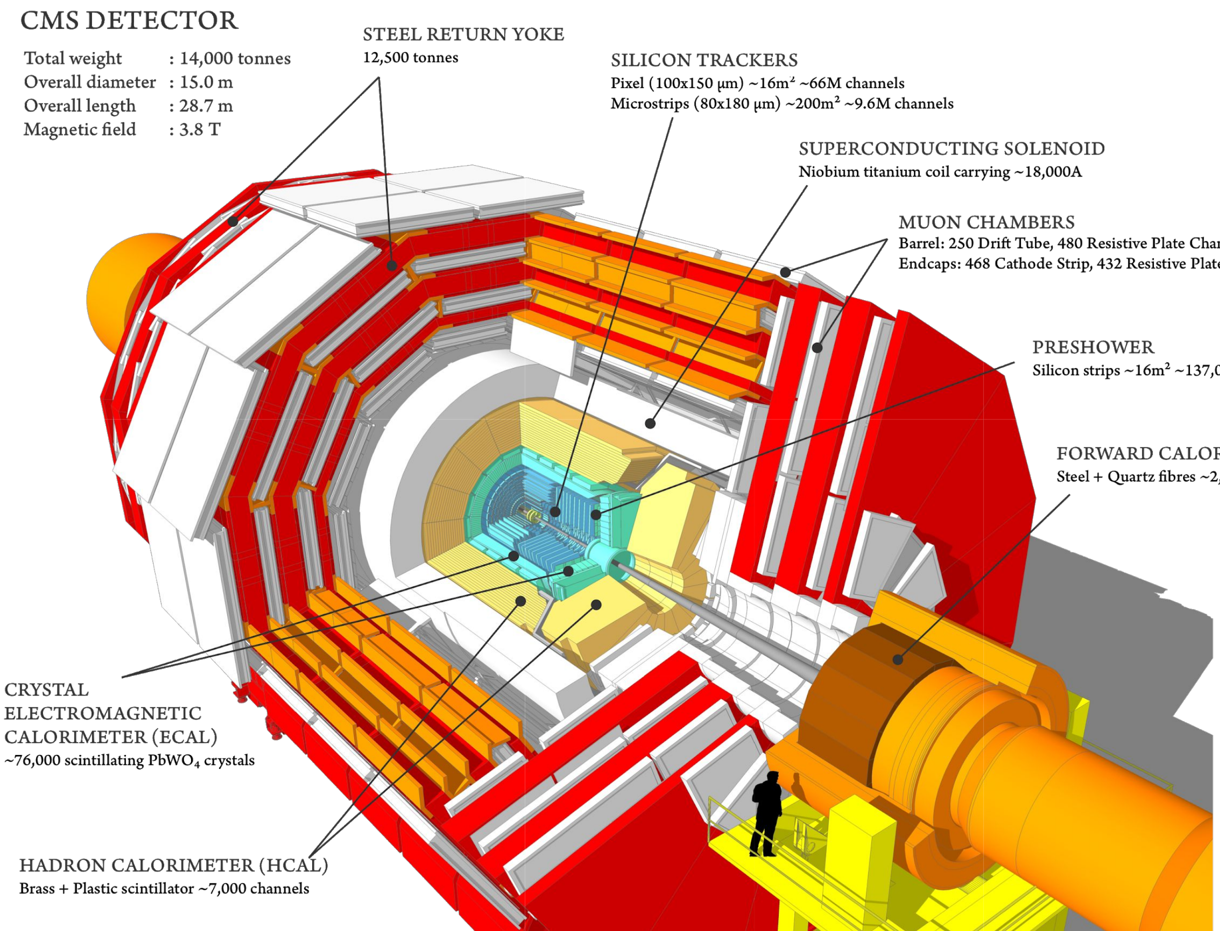
HLT Triggers Path (ϒ - Data):
➢ HLT_Dimuon0_Barrel_Upsilon

HLT Triggers Path (ϒ - MC):
➢ HLT_Dimuon0_Upsilon

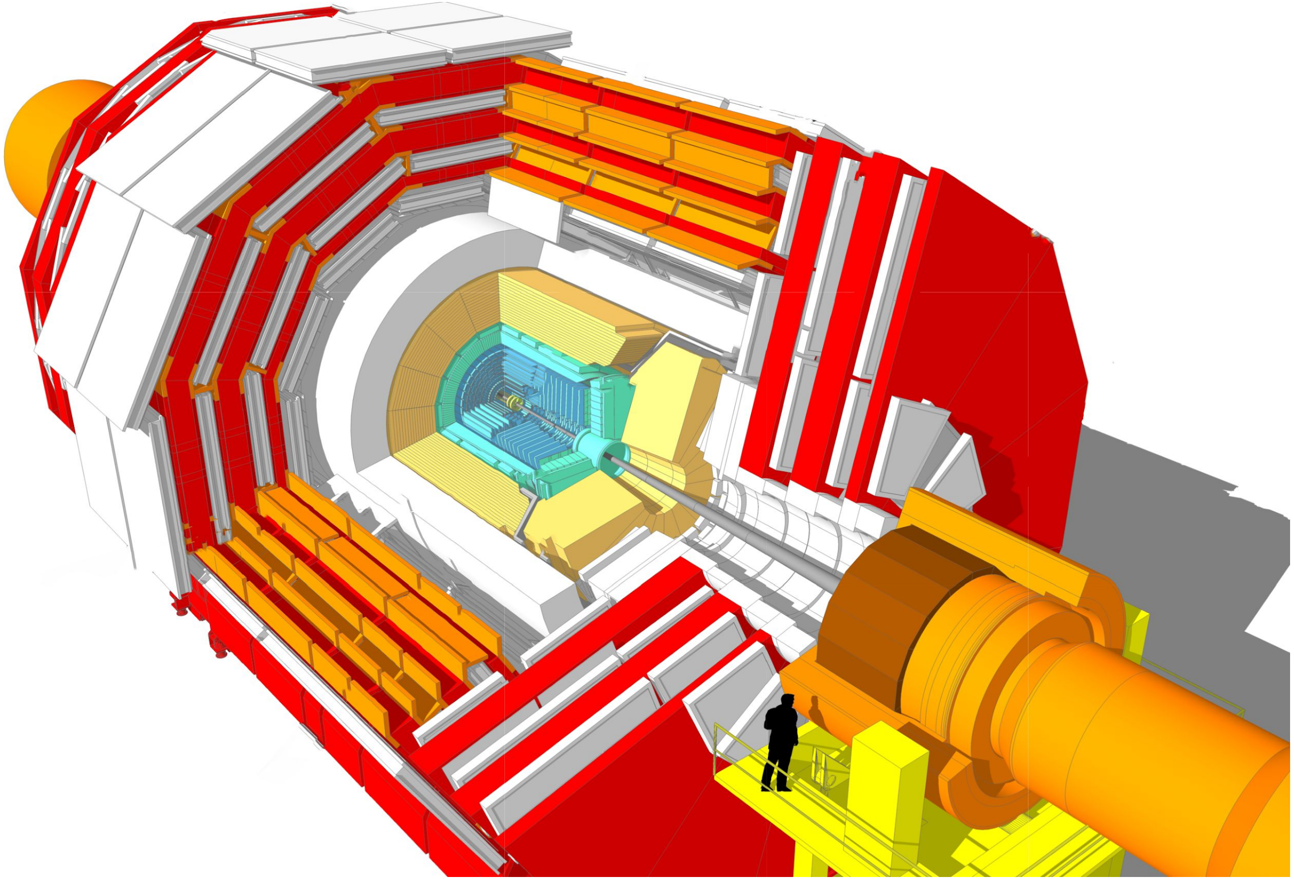
END

Backup

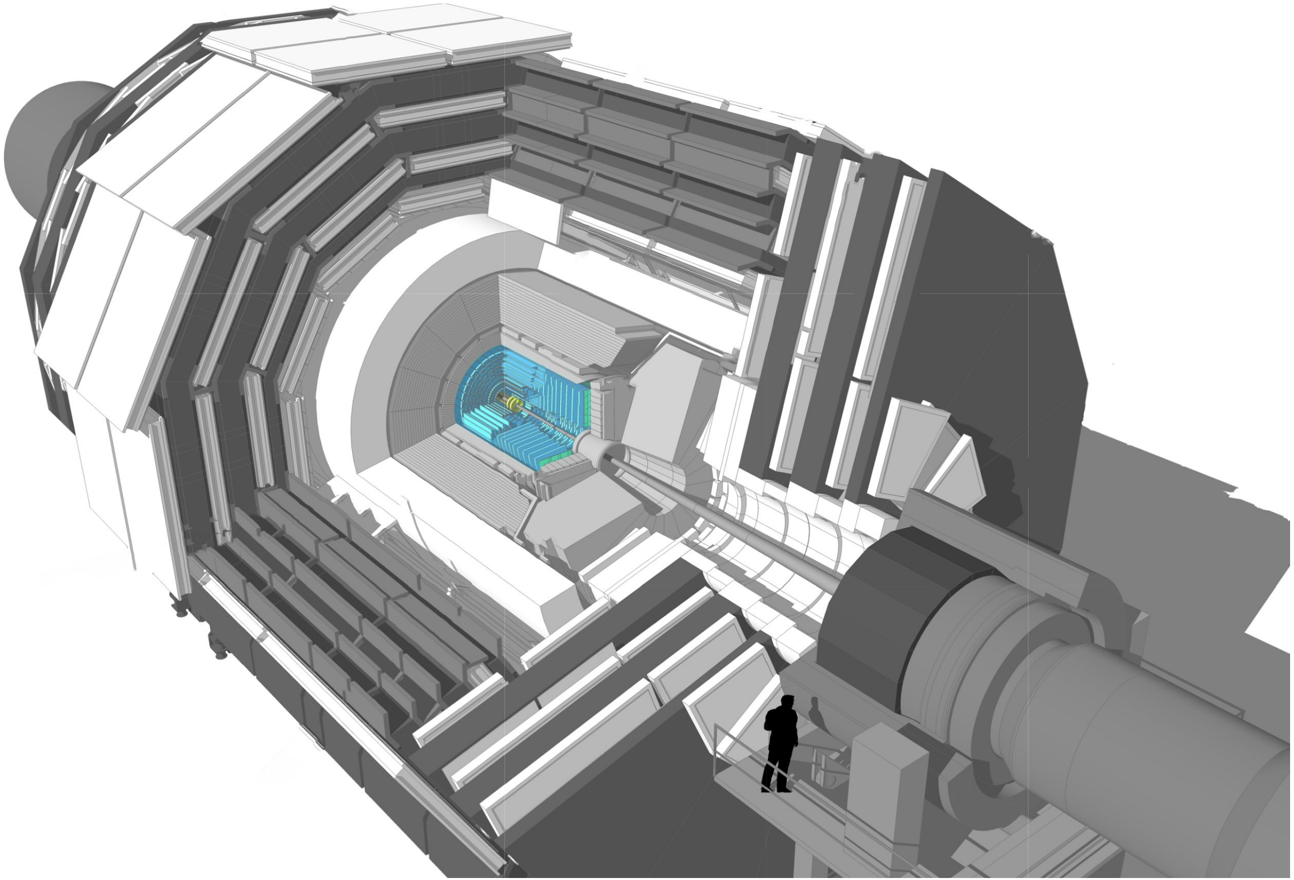
CMS Detector



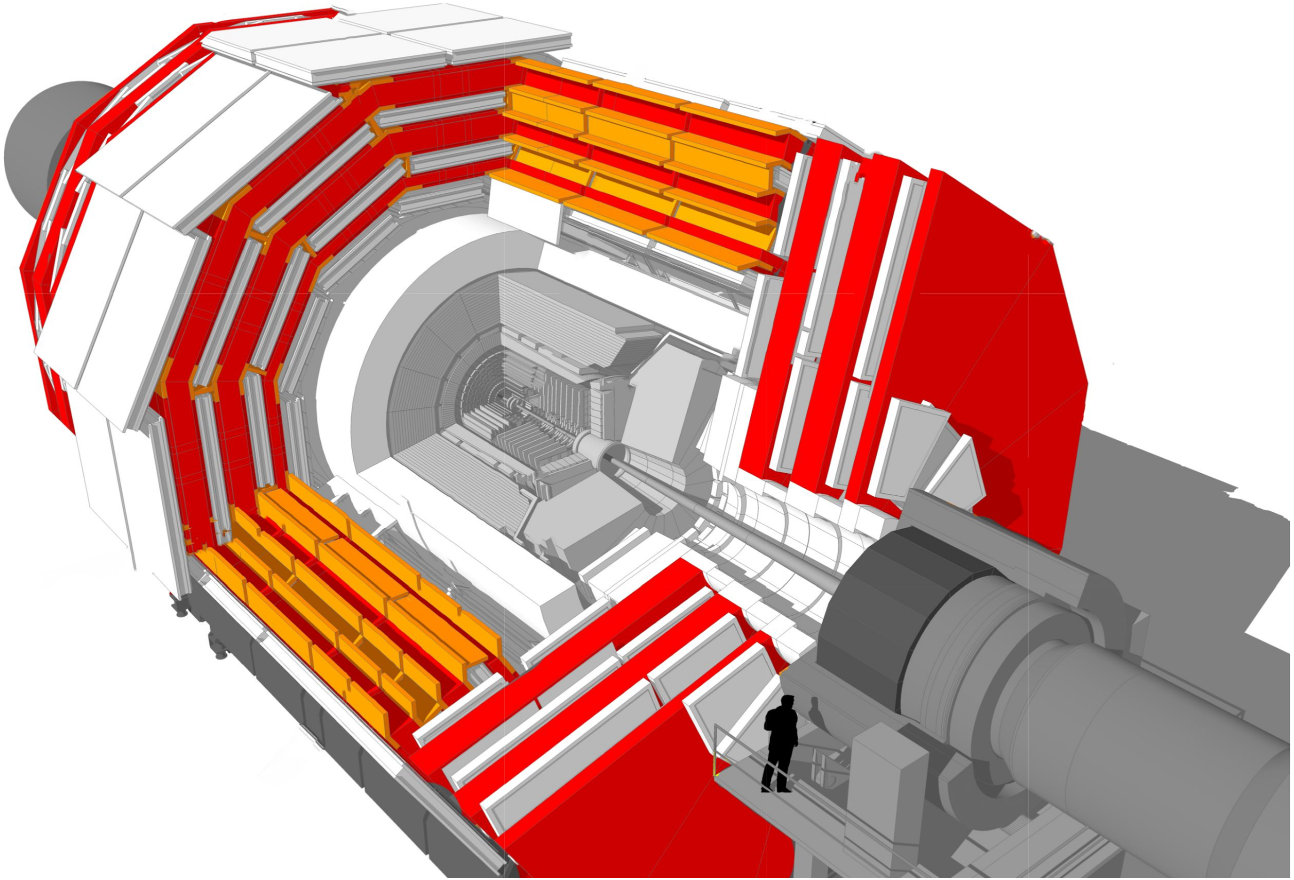
CMS Detector



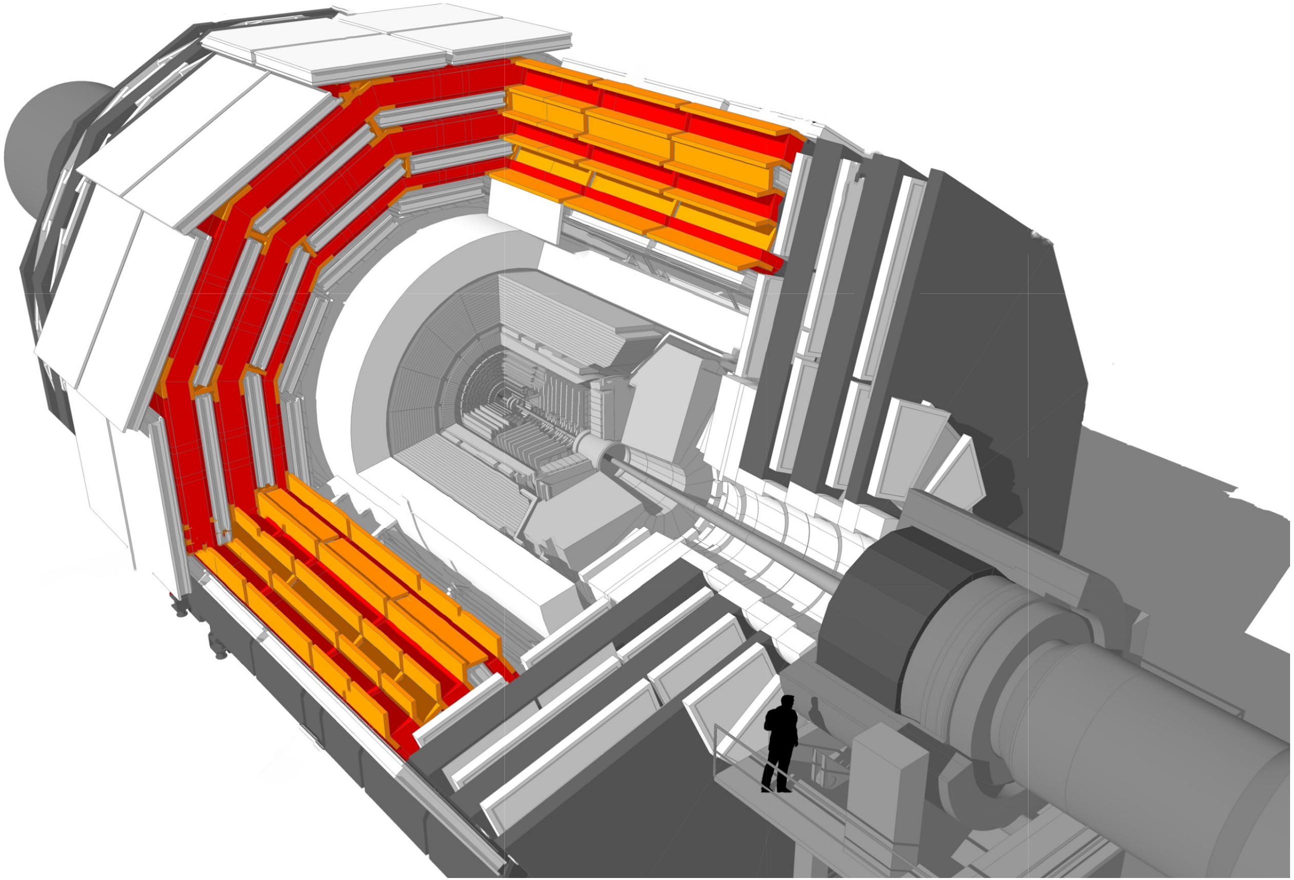
Inner Tracker



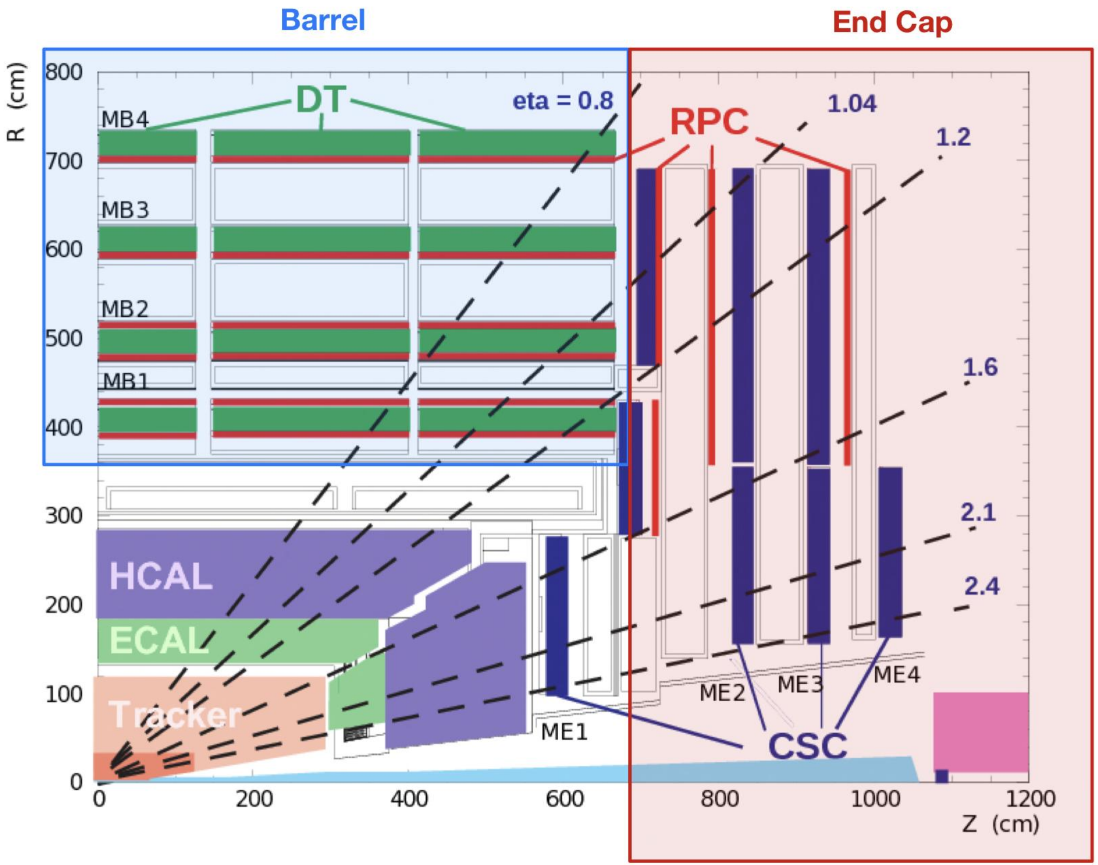
Muon Chamber



Barrel

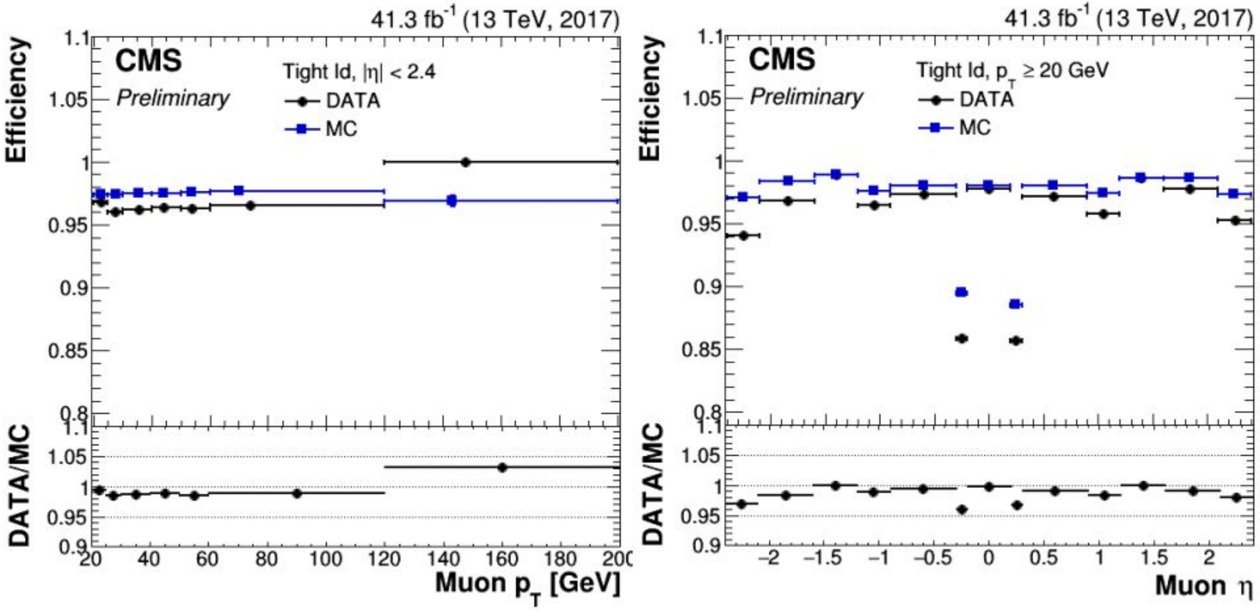


Barrel



Barrel
End Cap
R (cm)
800
700
600
500
400
300
200
100
0
MB4
MB3
MB2
MB1
DT
eta = 0.8
RPC
1.04
1.2
1.6
2.1
2.4
HCAL
ECAL
Tracker
ME1
ME2
ME3
ME4
CSC
0
200
400
600
800
1000
1200
Z (cm)

Tight Id efficiency vs p_T, η



* Error bars in the plot include only statistical uncertainty

Tight muon Id efficiency as a function of p_T and η for 2017 data and MC. The denominator is all tracker tracks with $p_T > 20$ GeV. The drops at around $|\eta| = 0.2$ are due to the cracks between wheels in the muon detectors. There is no significant dependency with the number of primary vertex. The drops in the forward region are due to inactive chambers not modelled in the MC.

From these results a data-to MC scale factor of ~98% is derived with a systematic uncertainty < 0.5%.

CMS DP-2018/042
https://cds.cern.ch/record/2629364/files/DP2018_042.pdf

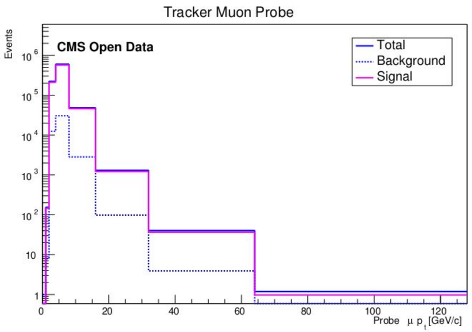
Method and samples

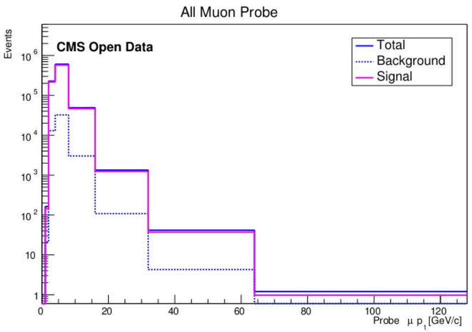
- Efficiencies are computed by means of a **tag-and-probe method** exploiting the $Z \to \mu^+\mu^-$ resonance.
- **Data:**
  - Proton-proton collision data at 13 TeV corresponding to an integrated luminosity of 16.3 fb^{-1} (2016), 41.3 fb^{-1} (2017) and 11.8 fb^{-1} (2018)
  - Events collected using single muon triggers.
  - 2017 and 2016: reconstruction using realistic calibrations and alignment.
  - 2018: prompt reconstruction (using startup calibration and alignment)
- **Simulation**
  - Drell-Yan + jets LO sample (madgraph)
  - Events are re-weighted to match the pileup distribution in data.

CMS DP-2018/042
https://cds.cern.ch/record/2629364/files/DP2018_042.pdf

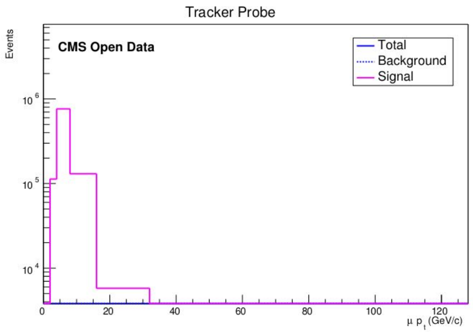
Results for J/ψ in sideband subtraction

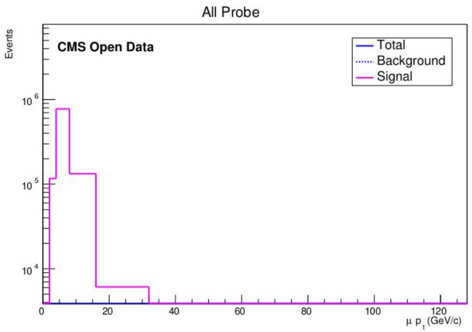
Data (2011)





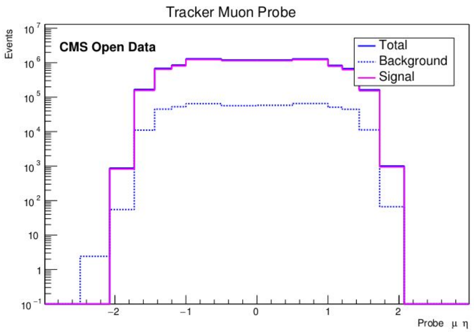
Simulated Data

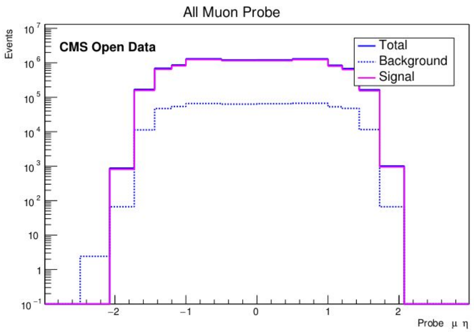




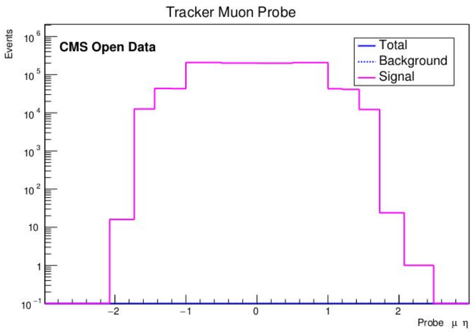
Results for J/ψ in sideband subtraction

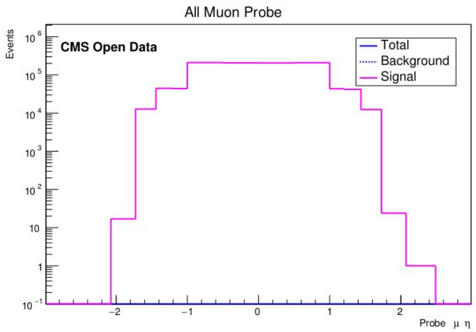
Data (2011)





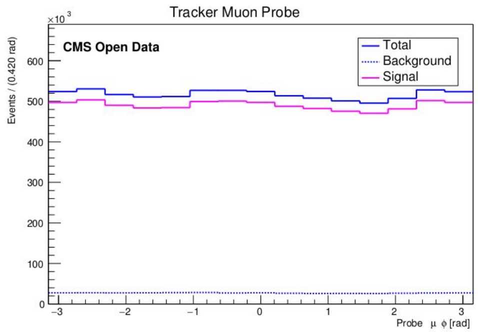
Simulated Data

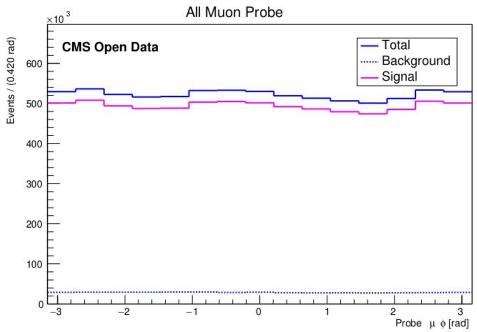




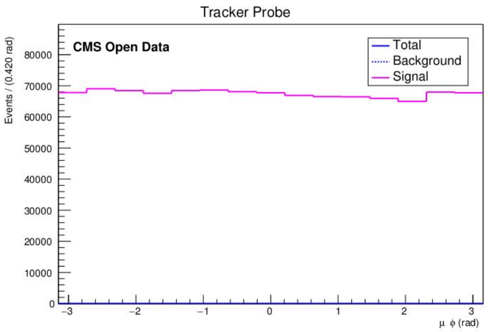
Results for J/ψ in sideband subtraction

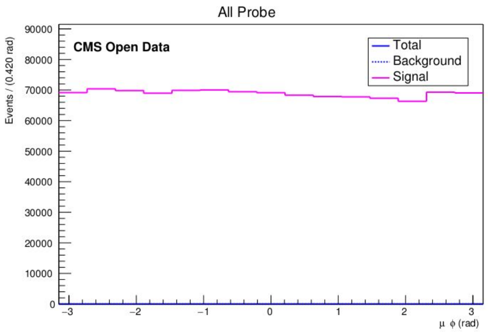
Data (2011)





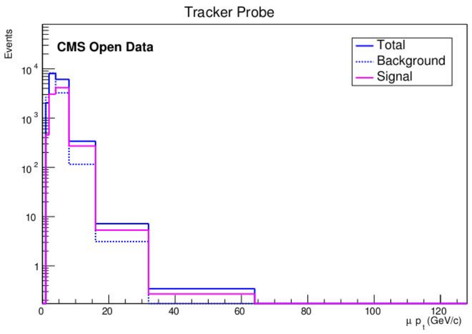
Simulated Data

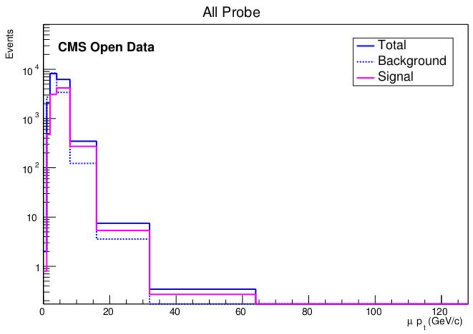




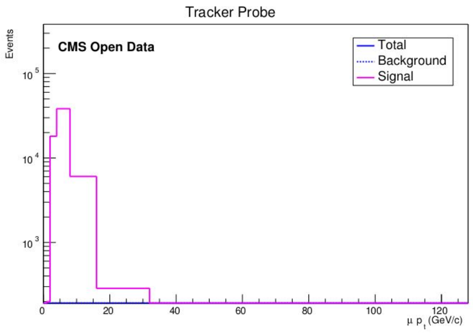
Results for Υ in sideband subtraction

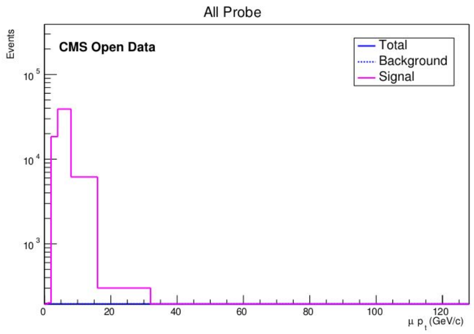
Data (2011)





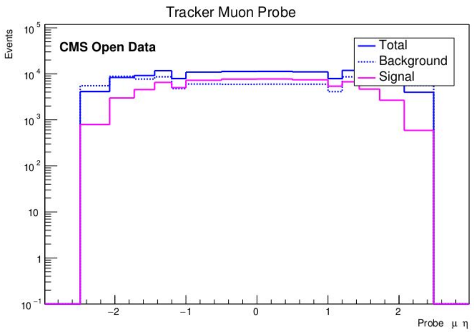
Simulated Data

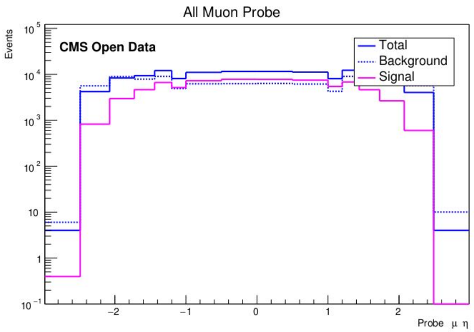




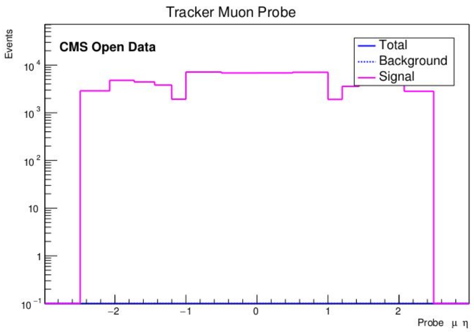
Results for Υ in sideband subtraction

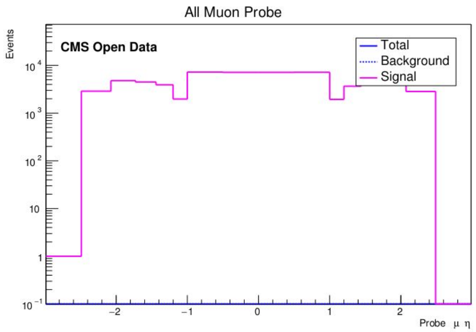
Data (2011)





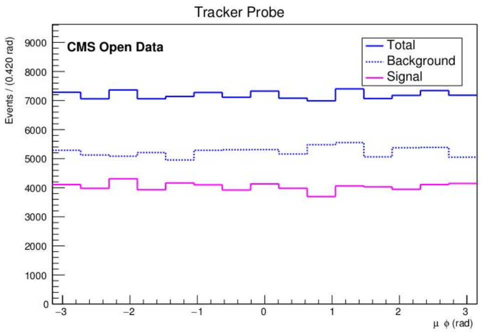
Simulated Data

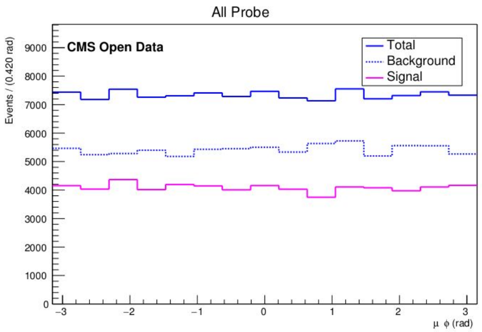




Results for Υ in sideband subtraction

Data (2011)





Simulated Data