CMS Calibration

Mehmet Ehliz

M2 Internship at LLR

Adviser: Dr. Christophe Ochando

Measurements of electron reconstruction efficiency with the CMS detector and data recorded in 2024

30.06.2025

@cms


How can we calibrate
the CMS Detector?
...

MOTIVATION

$$H \rightarrow 2Z \rightarrow 4l$$

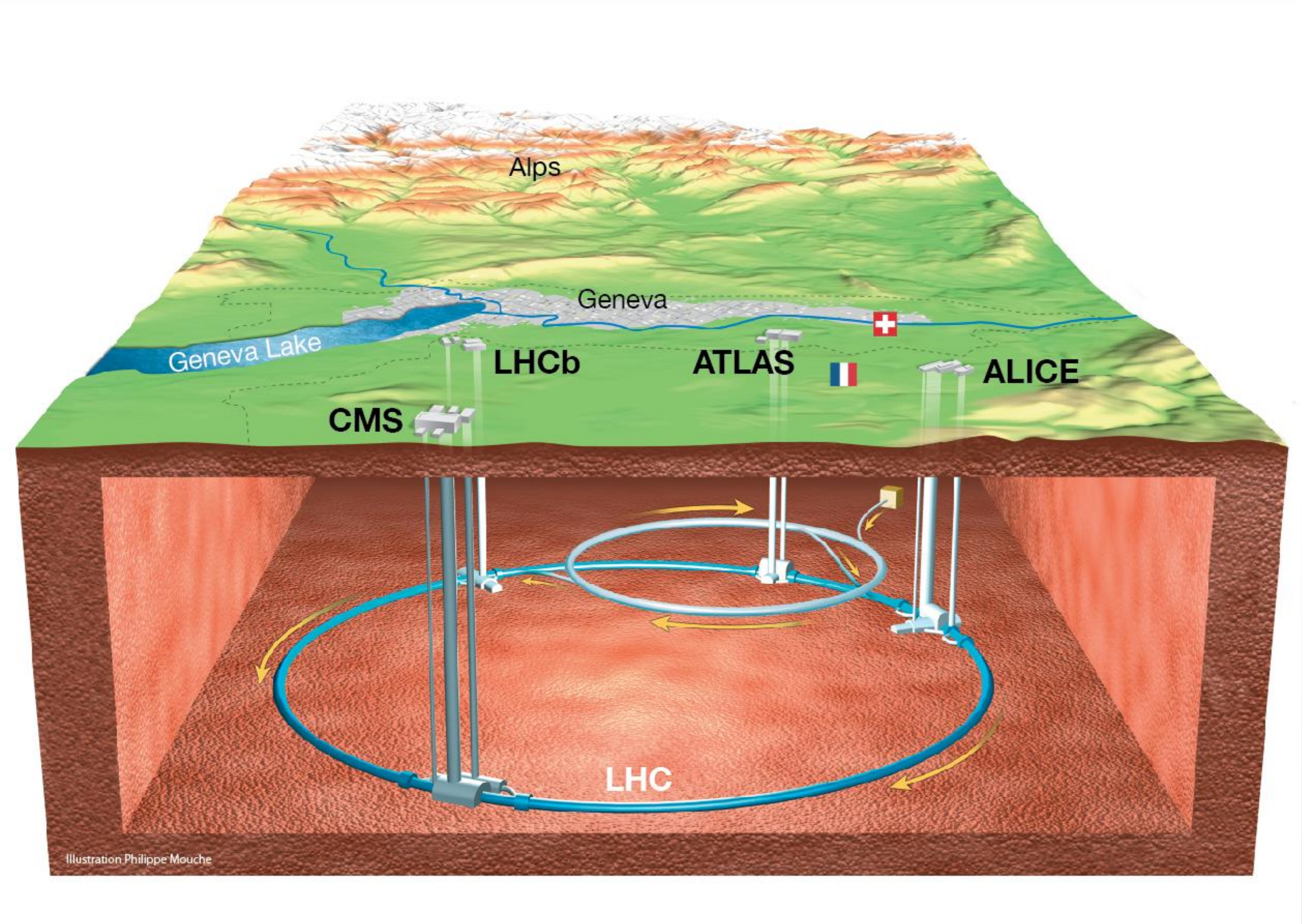
Mainly at low energy, but why ?

In this channel, electrons can have p_T as low as 7 GeV, and the uncertainty on their scale factors (SFs) at low p_t is one of largest experimental systematics

Improving those SFs gives you a more precise Higgs measurement

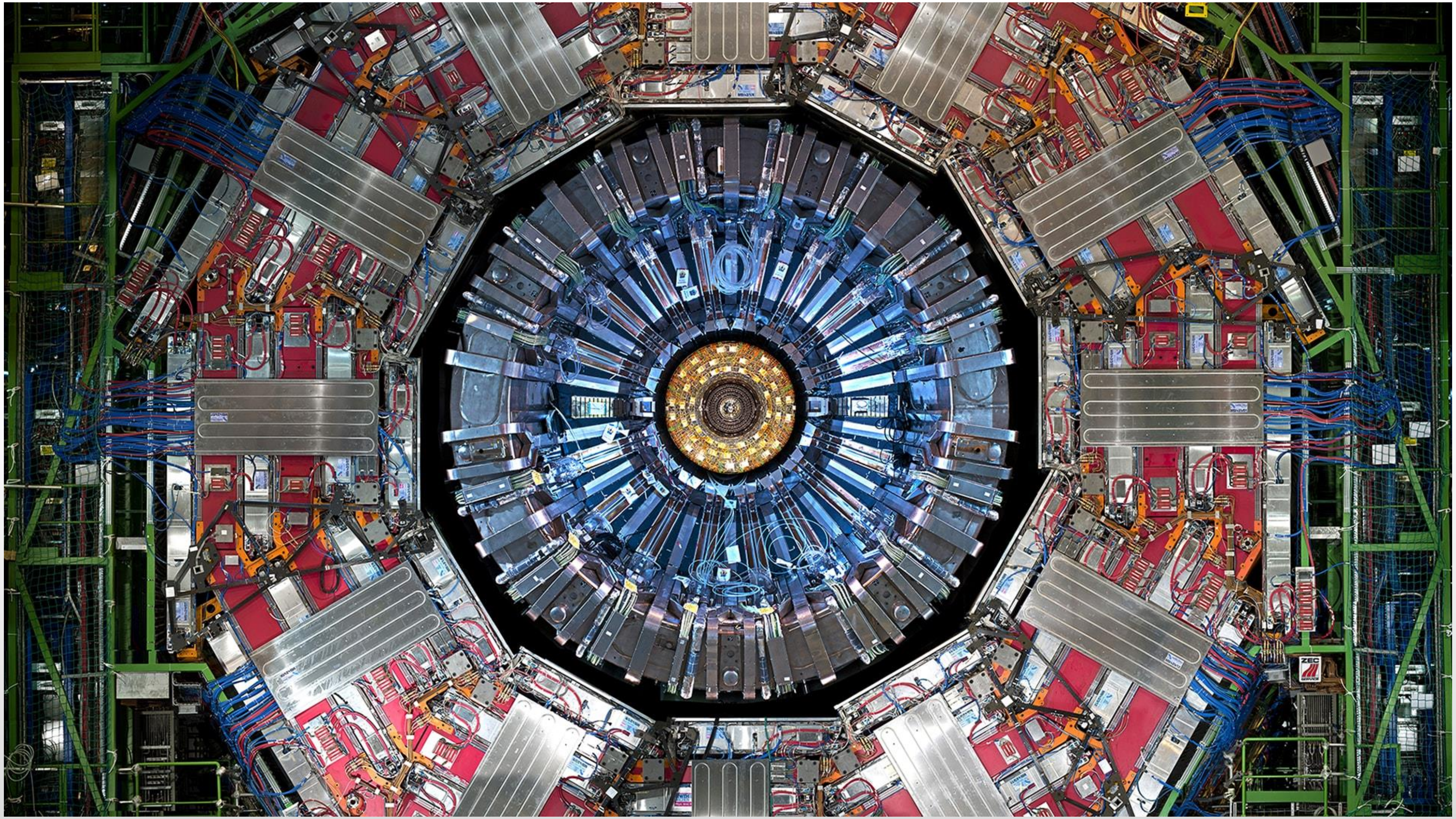
CONTENTS

INTRODUCTION



@cern

INTRODUCTION

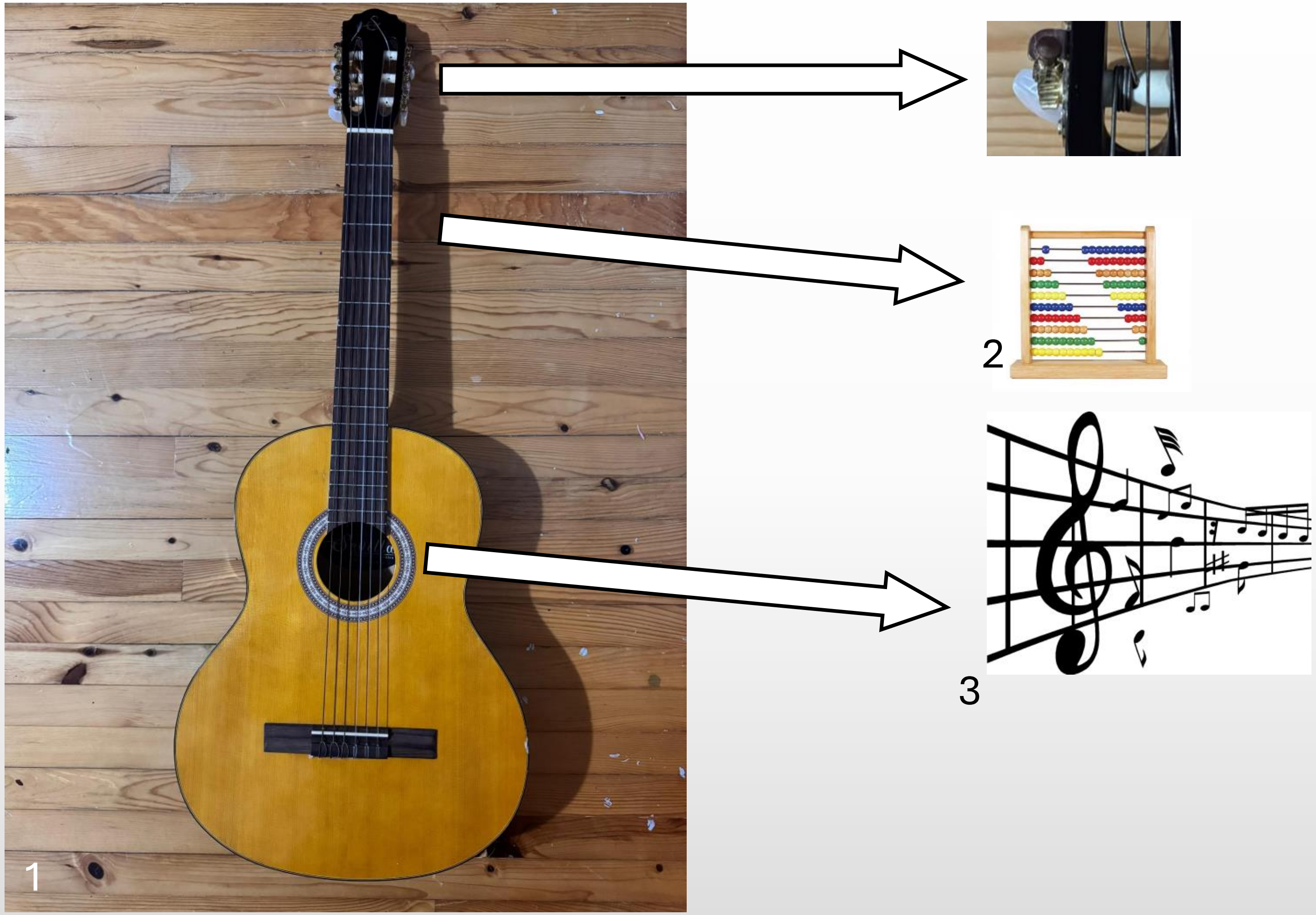


@cms

INTRODUCTION

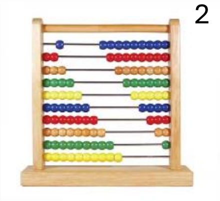
@M.Hakkı Ehliz

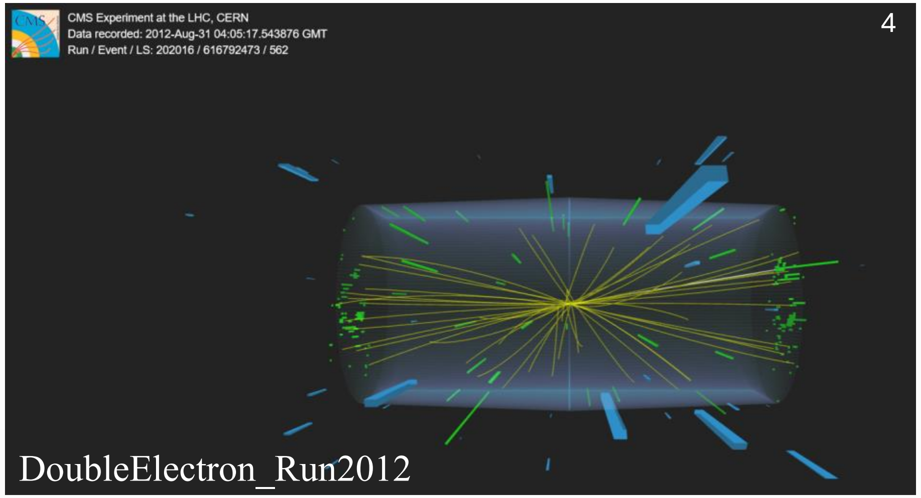
INTRODUCTION



1 @M.Hakkı Ehliz
2 @blogs.ubc.ca
3 @elestuchemusical.blogspot.com

INTRODUCTION

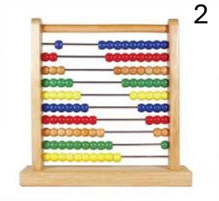




1 @M.Hakkı Ehliz
2 @blogs.ubc.ca
3 @elestuchemusical.blogspot.com
4 @cms

INTRODUCTION

Scale Factor



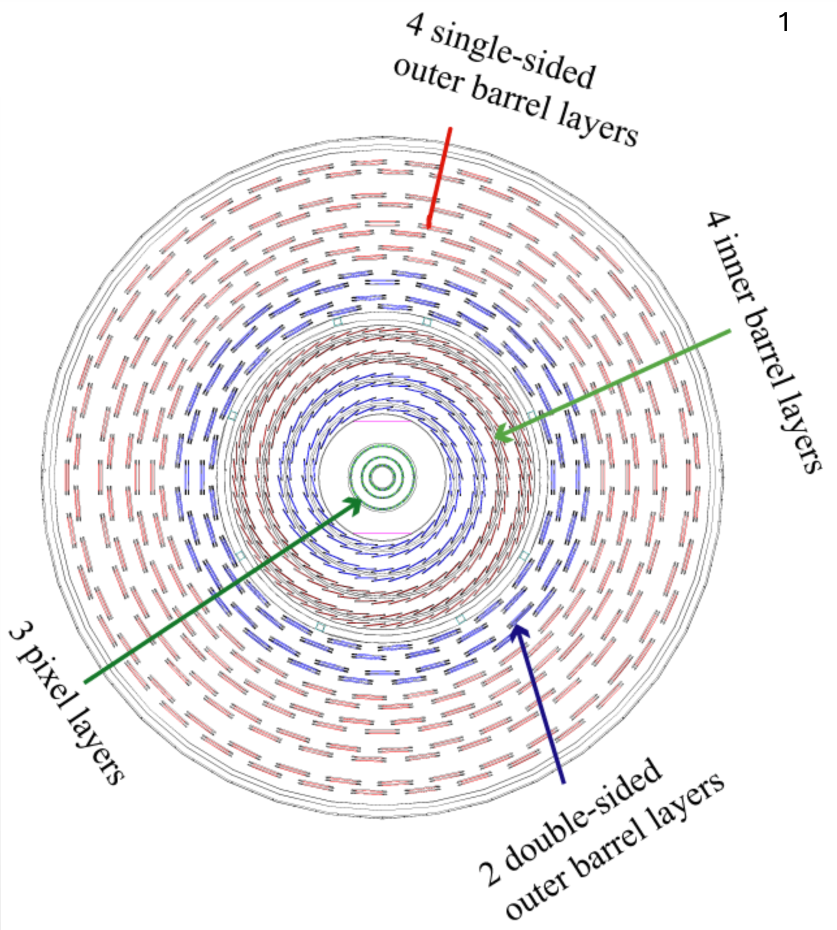
Methods

DoubleElectron_Run2012

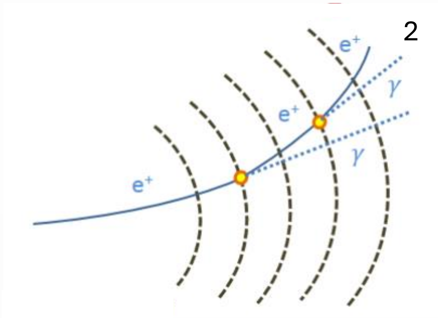
Data

1 @M.Hakkı Ehliz
2 @blogs.ubc.ca
3 @elestuchemusical.blogspot.com

Tracking and Calorimeter Systems

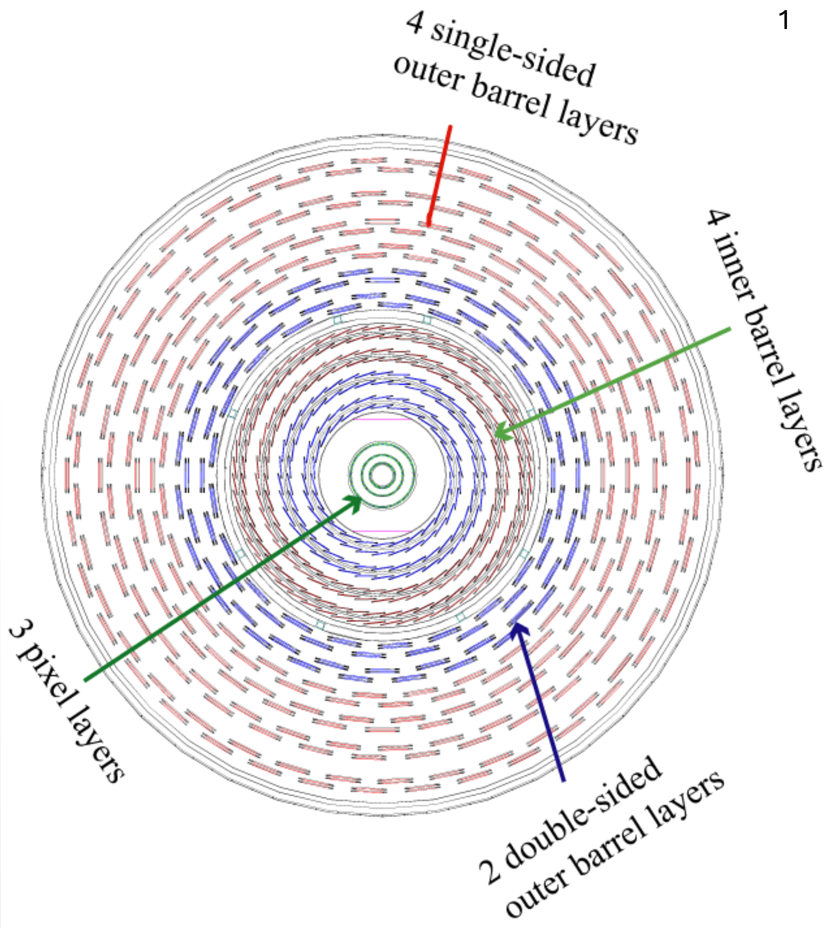


Kalman Filter → Gaussian Sum Filter

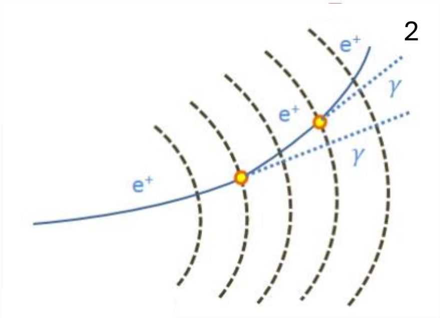


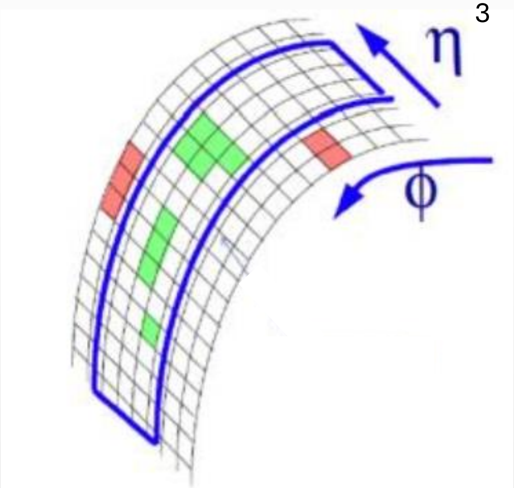
1 @cms
2 @S. Harper

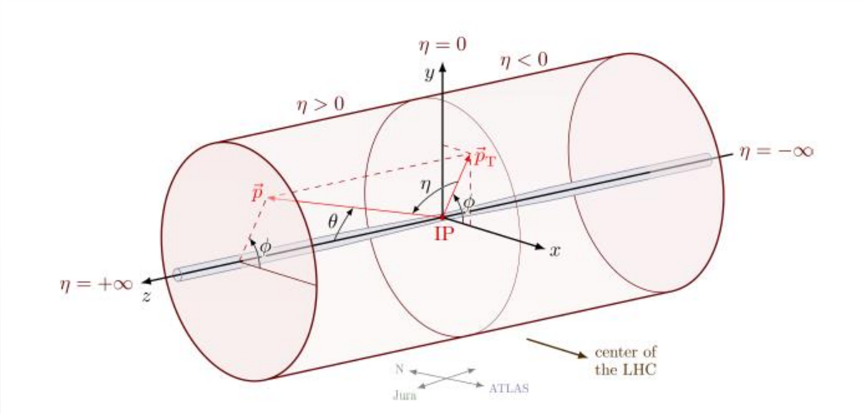
Tracking and Calorimeter Systems



Kalman Filter → Gaussian Sum Filter





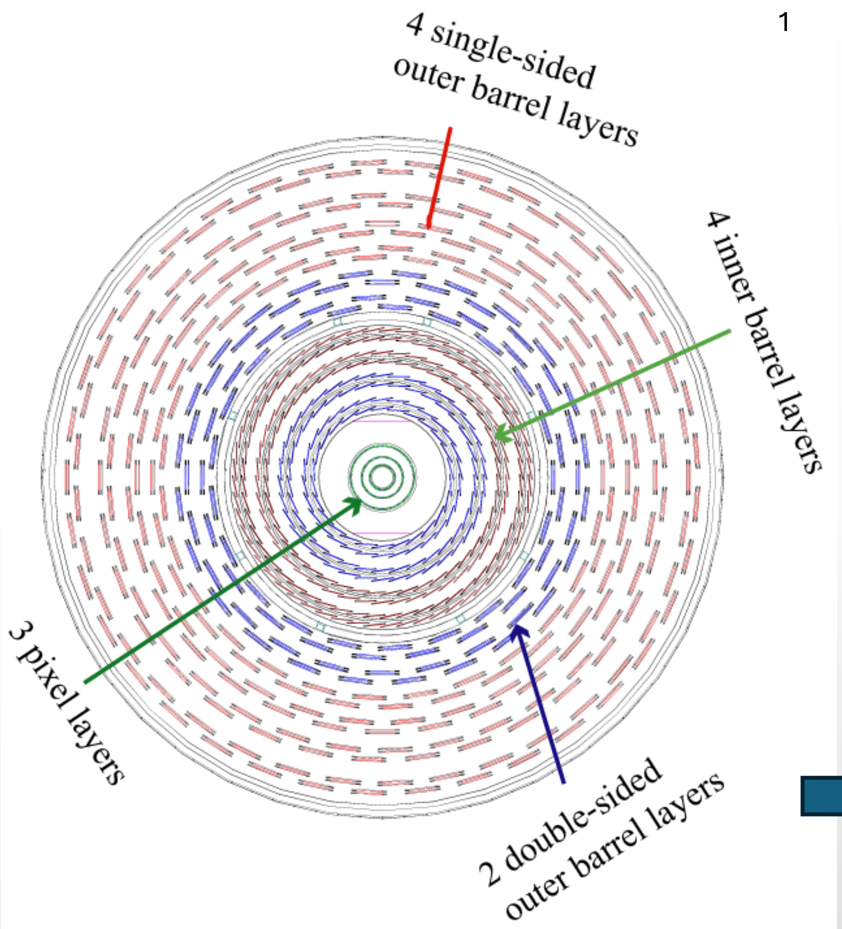


1 @cms

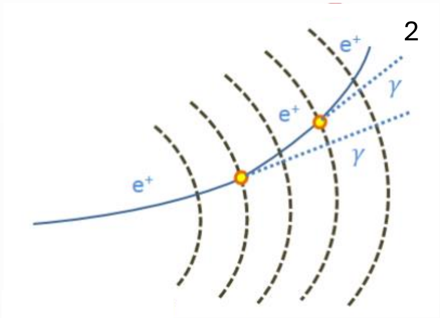
2 @S. Harper

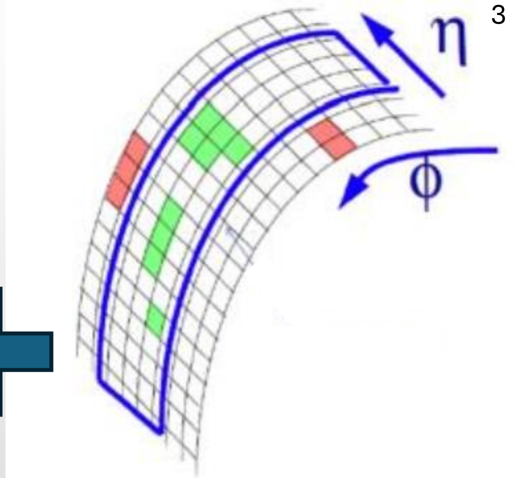
3 @ Yong Yang

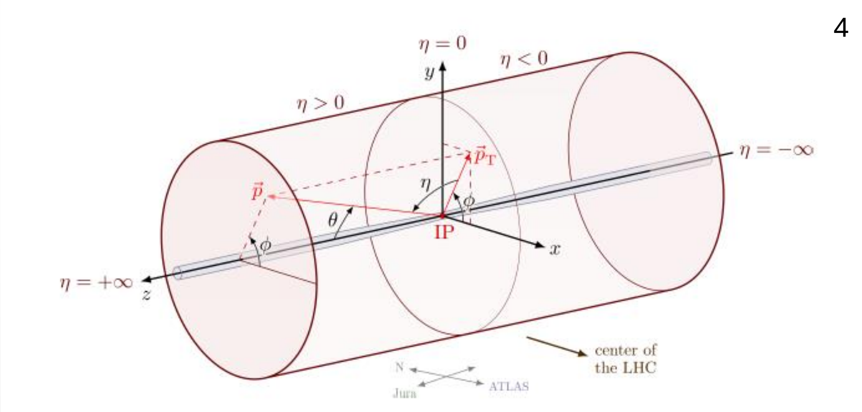
Tracking and Calorimeter Systems



Kalman Filter → Gaussian Sum Filter







+ … = Matching or not

1 @cms

2 @S. Harper

3 @ Yong Yang

4 @ tikz

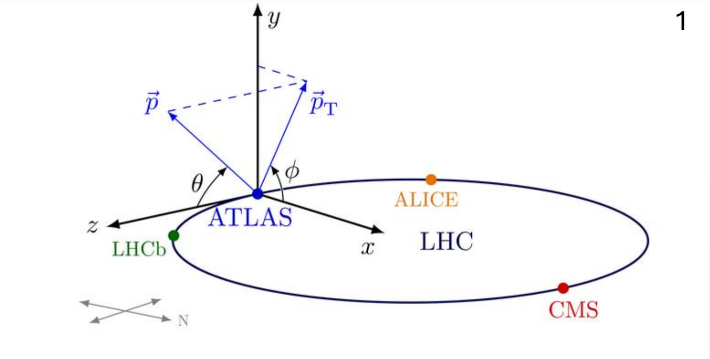
Datasets

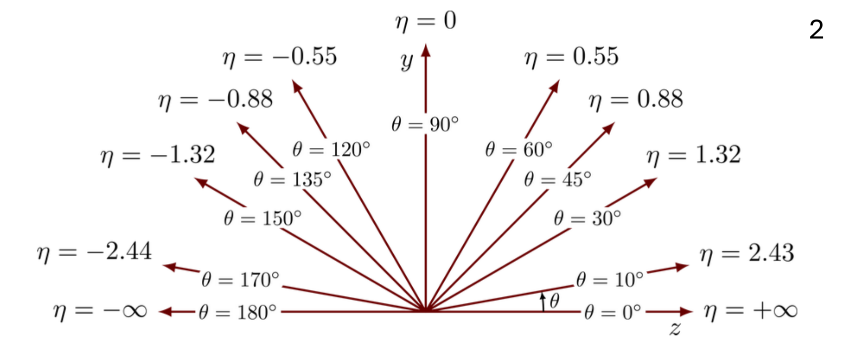
Data

Run-3 data at $\sqrt{s} = 13.6$ TeV from 2024 with an integrated luminosity of about 109.8 fb^{-1}

MC

Monte Carlo samples with the aMC@NLO generator at next-to-leading order in QCD, then ran them through a full GEANT4 simulation of the CMS detector





$$\eta \equiv -\ln[\tan(\frac{\theta}{2})]$$

1,2 @ tikz

Tag and Probe Method

Data-driven approach

$$Z \rightarrow e^{+}e^{-}$$

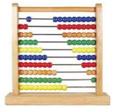
Tag

- Strict selection criteria (isolation, trigger…)
- Introduces bias

Probe

- Loose selection criteria (enough selection to form Z candidate with the Tag)
- It is unbiased

Tag and Probe Method



Simple Example

- **Baseline (loose) selection:**
  $E_T > 20\ GeV$
  Has a matching tracker track

- **Study (tight) selection to be measured:**
  H/E < 0.05, where H/E is the ratio of hadronic to electromagnetic energy

Tag and Probe Method

Simple Example

- **Baseline (loose) selection:**
  $E_T > 20\ GeV$
  Has a matching tracker track

- **Study (tight) selection to be measured:**
  $H/E < 0.05$, where H/E is the ratio of hadronic to electromagnetic energy

Probe ID	E_T [GeV]	Track Matched?	H/E	Baseline?	Study Cut?	Category
1	25	Yes	0.03	Yes	Yes	Passing
2	22	Yes	0.08	Yes	No	Failing
3	18	Yes	0.02	No	—	Not a probe
4	30	No	0.01	No	—	Not a probe

Tag and Probe Method

Simple Example

- **Baseline (loose) selection:**
  $E_T > 20\ GeV$
  Has a matching tracker track

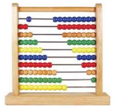
$$\epsilon = \frac{N_{pass}}{N_{pass} + N_{fail}} = \frac{1}{2}$$

- **Study (tight) selection to be measured:**
  $H/E < 0.05$, where H/E is the ratio of hadronic to electromagnetic energy

Probe ID	E_T [GeV]	Track Matched?	H/E	Baseline?	Study Cut?	Category
1	25	Yes	0.03	Yes	Yes	Passing
2	22	Yes	0.08	Yes	No	Failing
3	18	Yes	0.02	No	—	Not a probe
4	30	No	0.01	No	—	Not a probe

without background

Tag and Probe Method



Simple Example

- **Baseline (loose) selection:**

  $E_T > 20\ GeV$

  Has a matching tracker track

$$\epsilon = \frac{N_{pass-sig}}{N_{pass-sig} + N_{fail-sig}} \longrightarrow SF_s(E_T,\eta) = \frac{\epsilon_{\ Data}(E_T,\eta)}{\epsilon_{MC}(E_T,\eta)}$$

- **Study (tight) selection to be measured:**

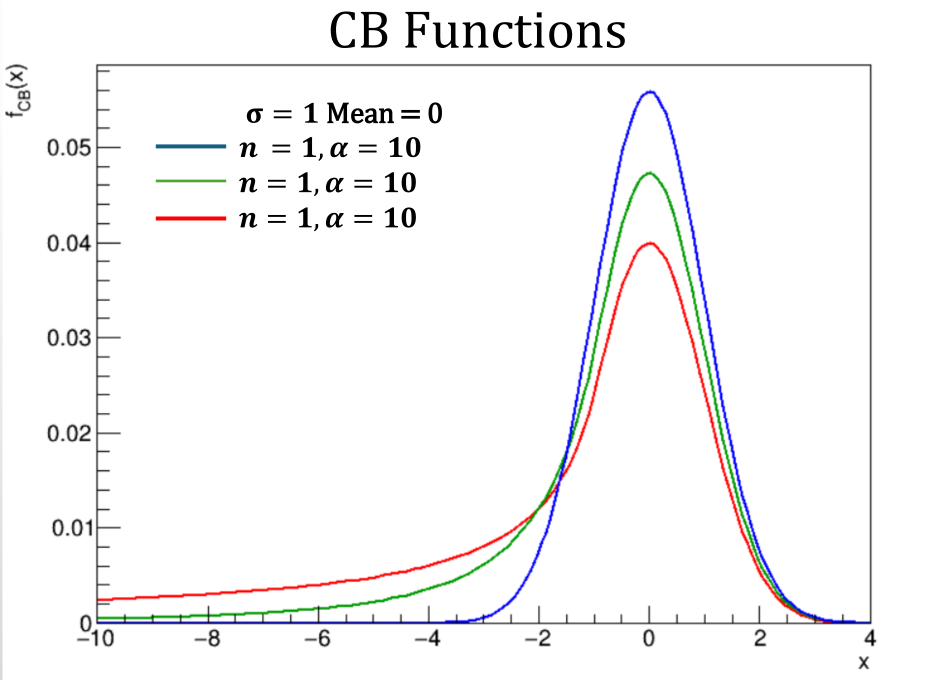
  $H/E < 0.05$, where H/E is the ratio of hadronic to electromagnetic energy

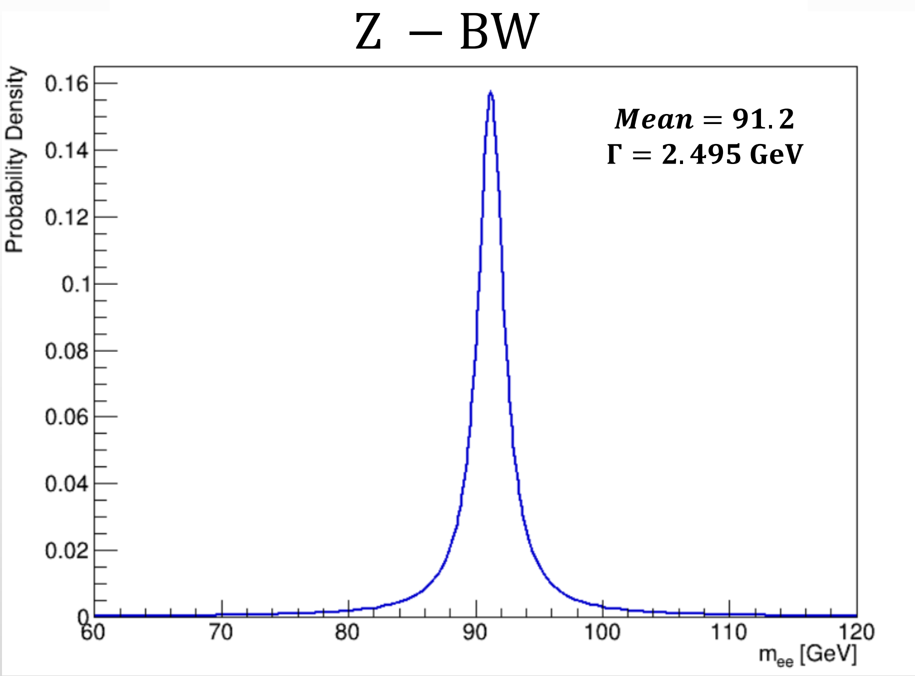
Probe ID	E_T [GeV]	Track Matched?	H/E	Baseline?	Study Cut?	Category
1	25	Yes	0.03	Yes	Yes	Passing
2	22	Yes	0.08	Yes	No	Failing
3	18	Yes	0.02	No	—	Not a probe
4	30	No	0.01	No	—	Not a probe

Models of signal & background

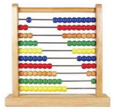
Signal
- Breit–Wigner
- Gaussian
- Crystal Ball
- Double Crystal Ball

Background
- Exponential
- Bernstein Polynomial
- Chebychev Polynomial
- RooCMSShape





Models of signal & background



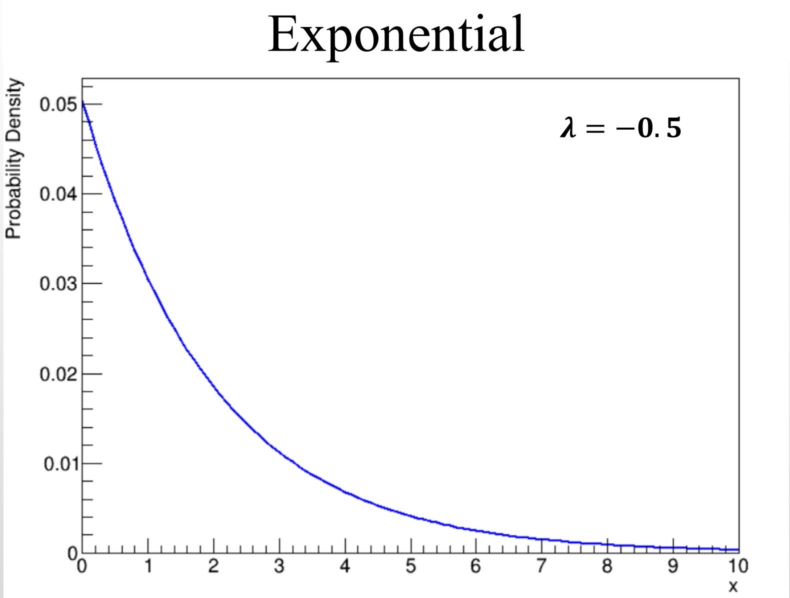
Signal
- Breit–Wigner
- Gaussian
- Crystal Ball
- Double Crystal Ball

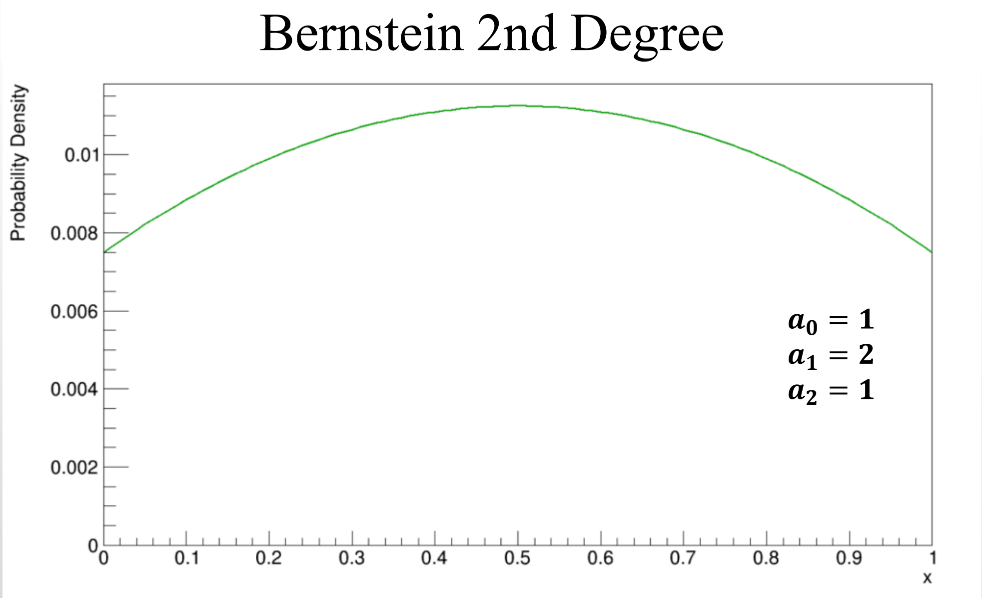
Background
- Exponential
- Bernstein Polynomial
- Chebychev Polynomial
- RooCMSShape

$$b_{v,n}(x) = \binom{n}{v} x^{v}(1-x)^{n-v} \, for \; v = 0, \ldots, n$$

$$T_n(x) = \cos(n \arccos(x)), x \, \epsilon [-1,1]$$

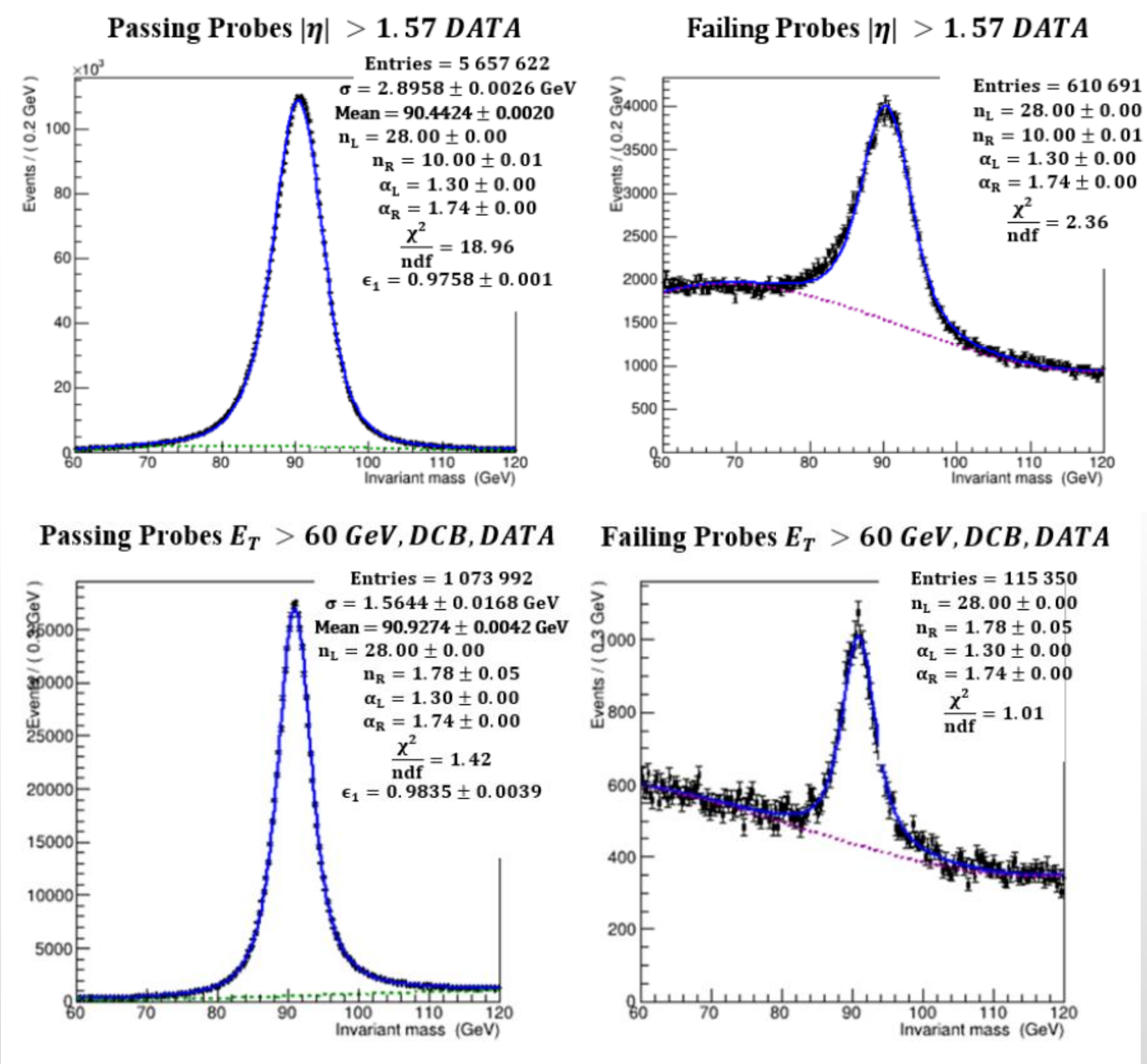
$$f(x) = e^{\lambda}$$

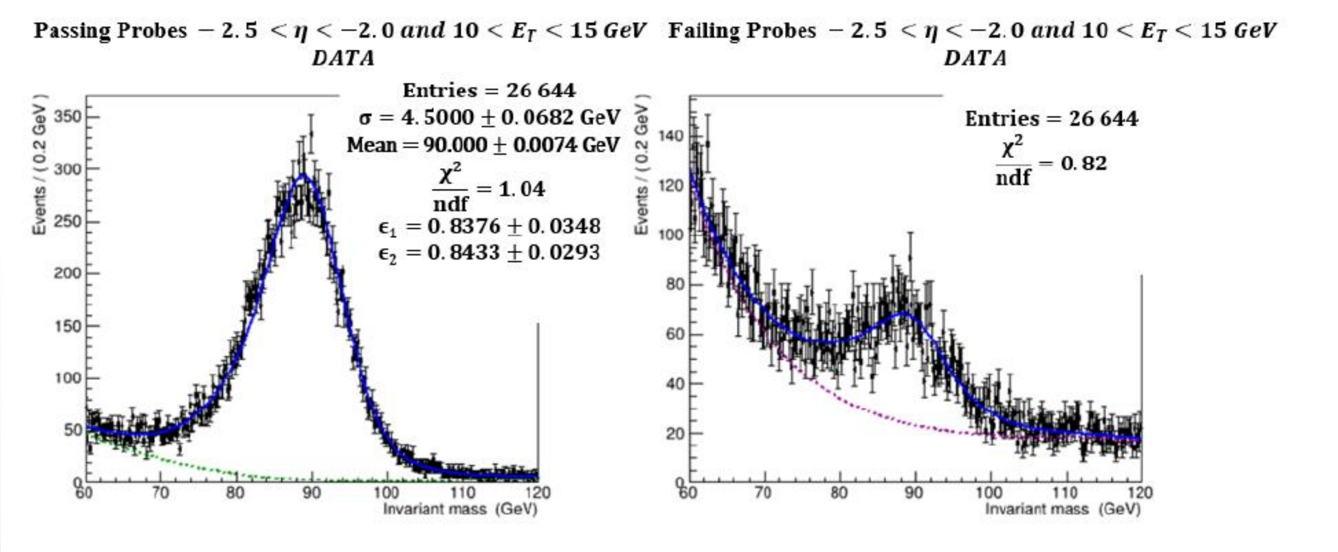




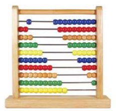
$$B_2(x) = a_0(1-x)^2 + a_1 2x(1-x) + a_2 x^2$$

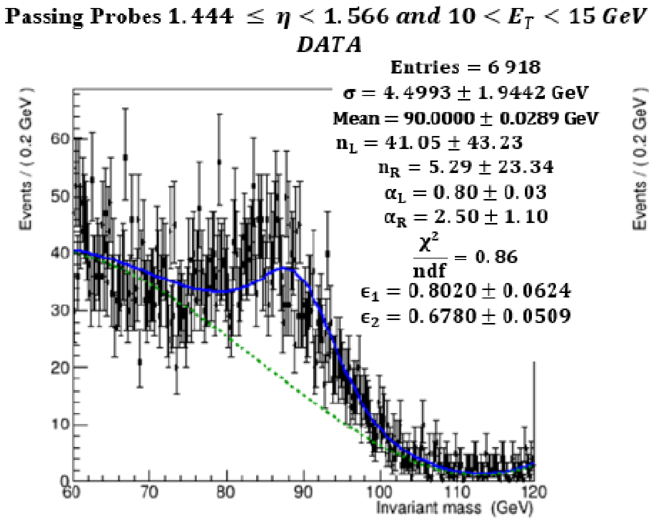
Measurements for several E_T and η bins

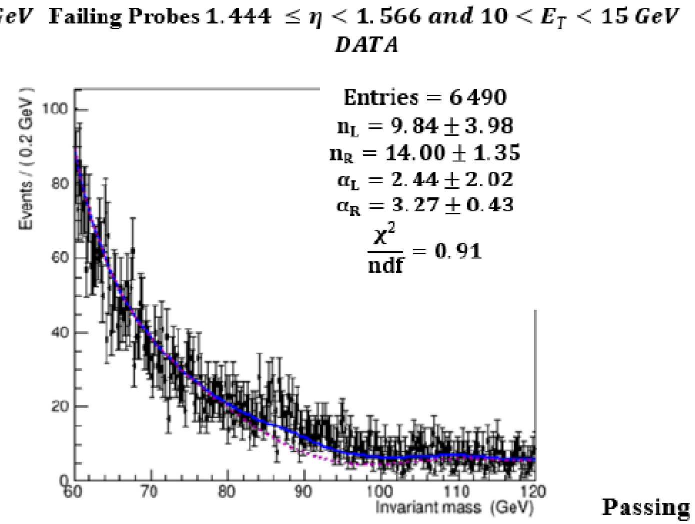


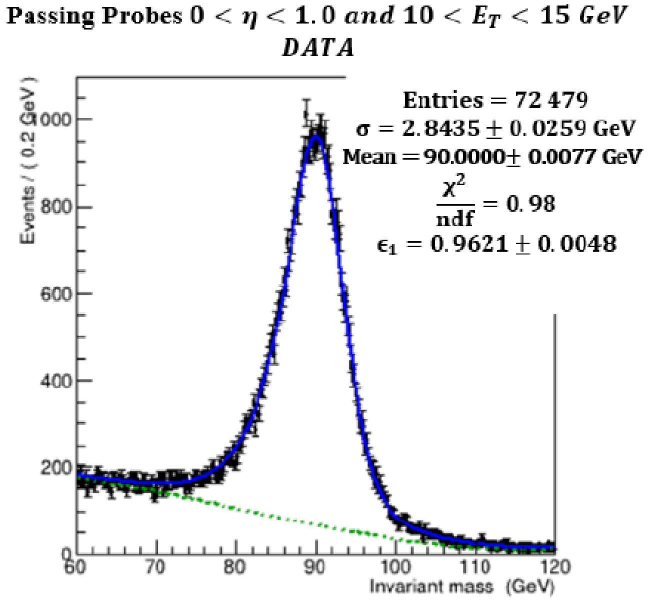


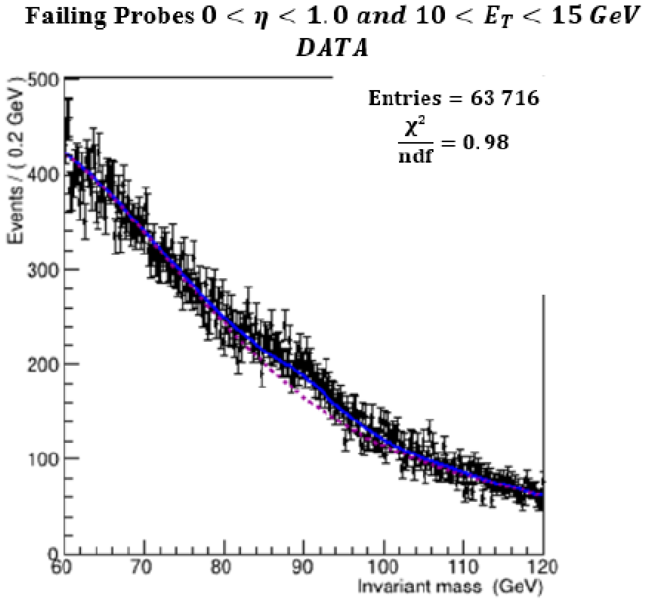
Measurements for several E_T and η bins



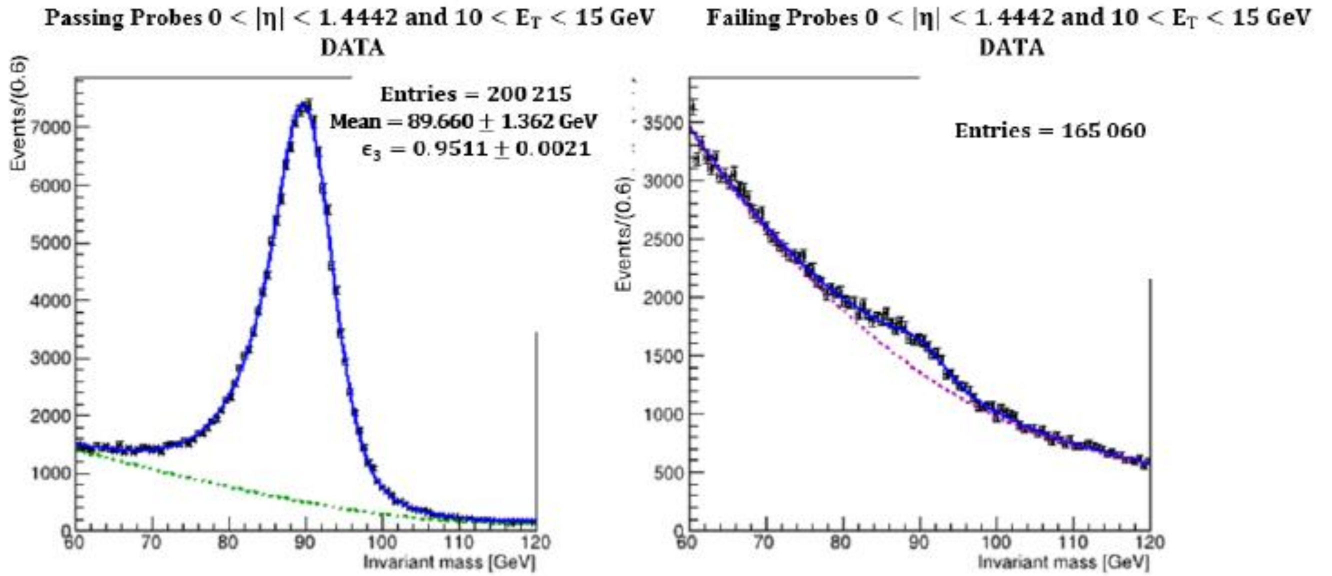








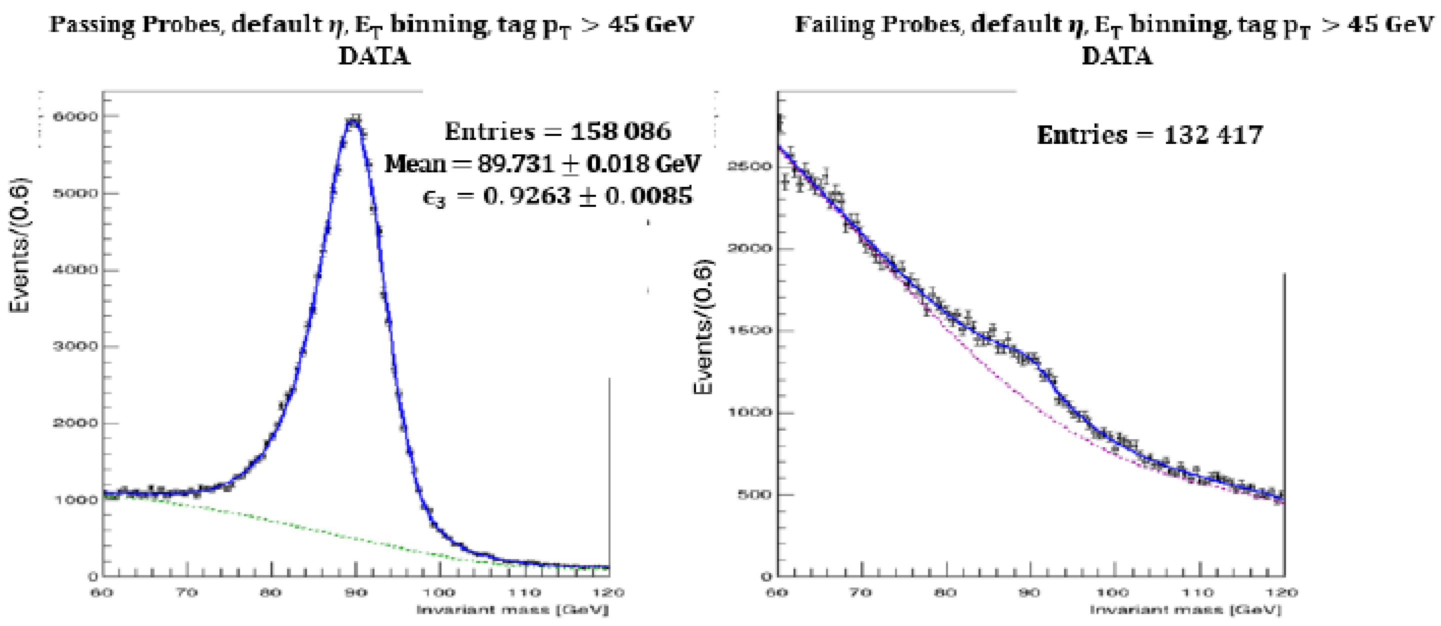
Measurements for several E_T and η bins



Data: unbinned fit in the barrel ($|\eta| < 1.4442$), $10 < E_T < 15$ GeV, tag $p_T > 35$ GeV.

Optimizations at Low Probe E_T Cut

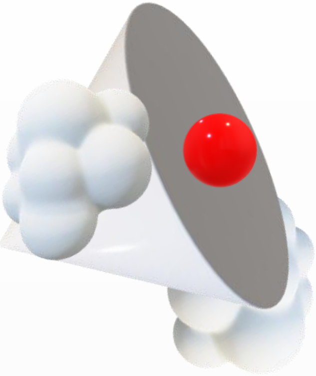
Tag p_T Threshold

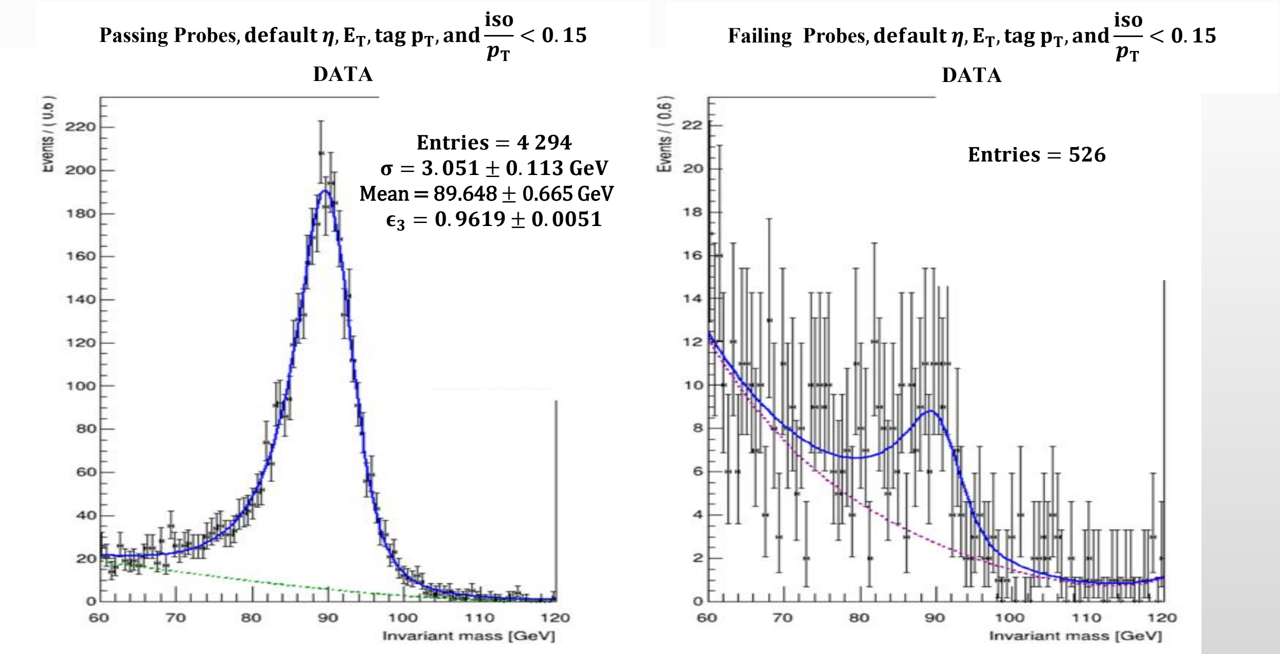


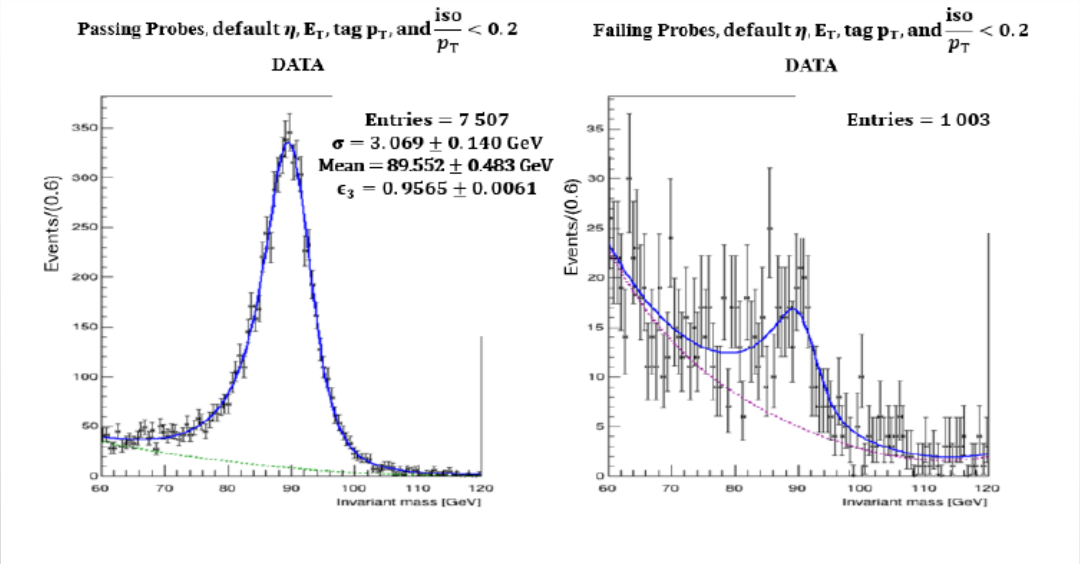
Data (default η, E_T binning, tag $p_T > 45\,\text{GeV}$)

Optimizations at Low Probe E_T Cut

Track Isolation

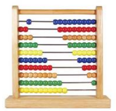






(a) Data sample with $iso/p_T < 0.2$

Optimizations at Low Probe E_T Cut



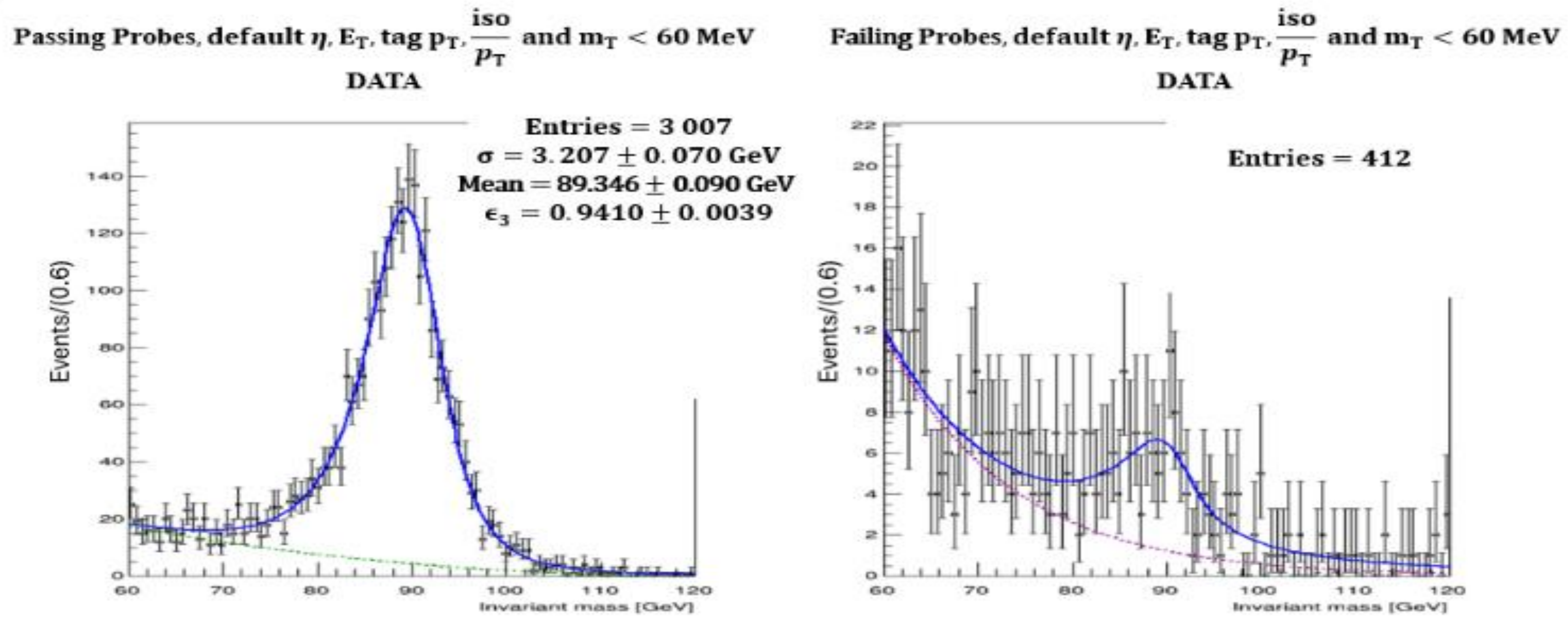
Transverse-Mass Veto

$W \rightarrow e\nu$

$$m_T = \sqrt{2\, p_{T,e}\, E_T^{miss}(1 - \cos \Delta\phi\,)}$$

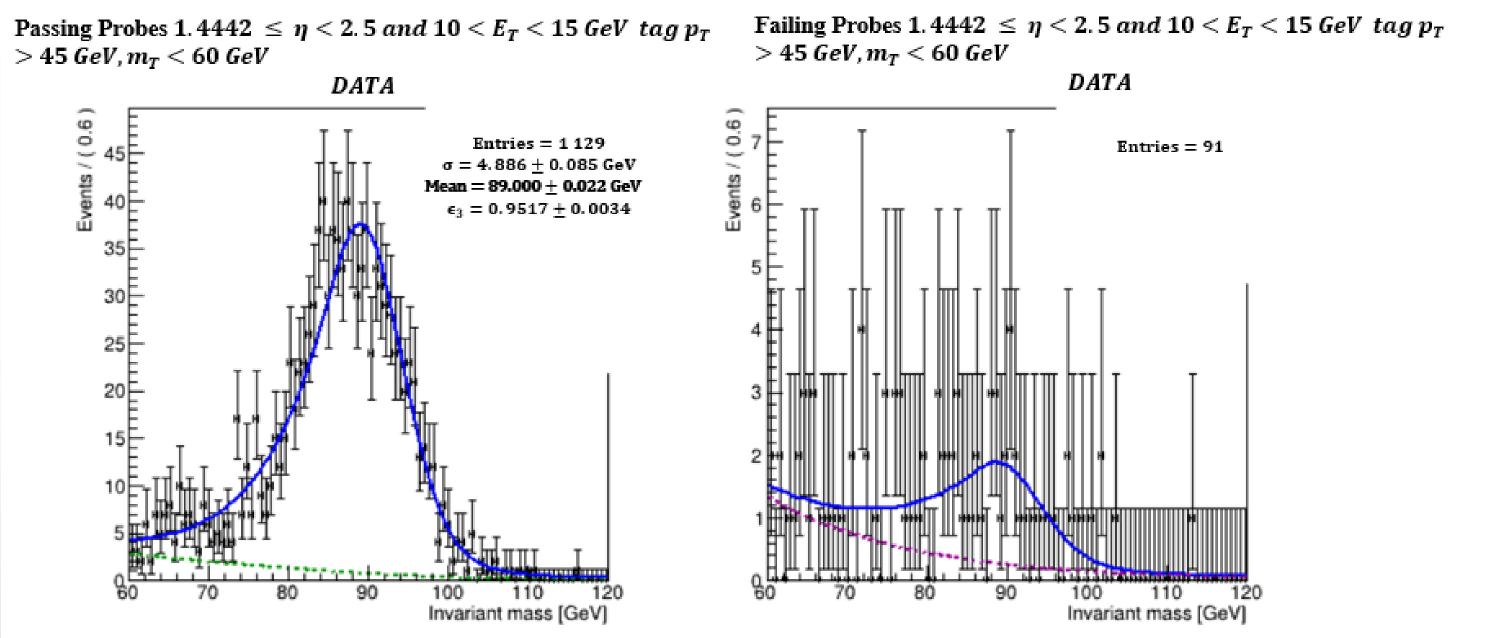
$\Delta\phi$ is the angle between the electron and the missing ergy vector

E_T^{miss} is missing transverse energy



Data sample with previous cuts and adding $(m_T < 60)$ GeV

Optimizations at Low Probe E_T Cut



Results

Tag-and-Probe efficiencies and scale factors (SF$\equiv \epsilon_{\text{data}}/\epsilon_{\text{MC}}$) for $10 < E_T < 15$ GeV

$\lvert\eta\rvert$ range	Tag p_T	ISO/p_T	m_T	ϵ_{data}	SF
< 1.4442	> 35 GeV	—	—	0.9511 ± 0.0021	1.0293 ± 0.0023
1.5556–2.5	> 35 GeV	—	—	0.8428 ± 0.0145	0.9467 ± 0.0163
< 1.4442	> 45 GeV	—	—	0.9263 ± 0.0082	1.0024 ± 0.0089
1.5556–2.5	> 45 GeV	—	—	0.8861 ± 0.0067	0.9955 ± 0.0075
< 1.4442	> 45 GeV	< 0.2	—	0.9565 ± 0.0061	1.0352 ± 0.0066
1.5556–2.5	> 45 GeV	< 0.2	—	0.9334 ± 0.0096	1.0485 ± 0.0108
1.5556–2.5	> 45 GeV	< 0.15	—	0.9258 ± 0.0828	1.0400 ± 0.0931
< 1.4442	> 45 GeV	< 0.15	—	0.9619 ± 0.0051	1.0413 ± 0.0055
< 1.4442	> 45 GeV	< 0.15	< 60 GeV	0.9723 ± 0.0457	1.0525 ± 0.0495
1.5556–2.5	> 45 GeV	< 0.15	< 60 GeV	0.9276 ± 0.1500	1.0422 ± 0.1685
1.5556–2.5	> 45 GeV	< 0.25	< 60 GeV	0.9517 ± 0.0034	1.0692 ± 0.0038
< 1.4442	> 45 GeV	< 0.25	< 60 GeV	0.9465 ± 0.0032	1.0248 ± 0.0035

Conclusion and the Future

- *Electron Reconstruction Efficiency*: Measured using the Tag-and-Probe method on CMS Run-3 data, comparing real data and Monte Carlo in bins of E_T and η, calculated scale factors

- *Background Suppression*: Applied optimized cuts to reduce background under the Z peak, improving the peak in the falling probes in low E_T bins.

- *Impact and Outlook*: Results show per-cent level agreement between data and simulation

- *Methods extensions:* to other resonances (J/ψ)

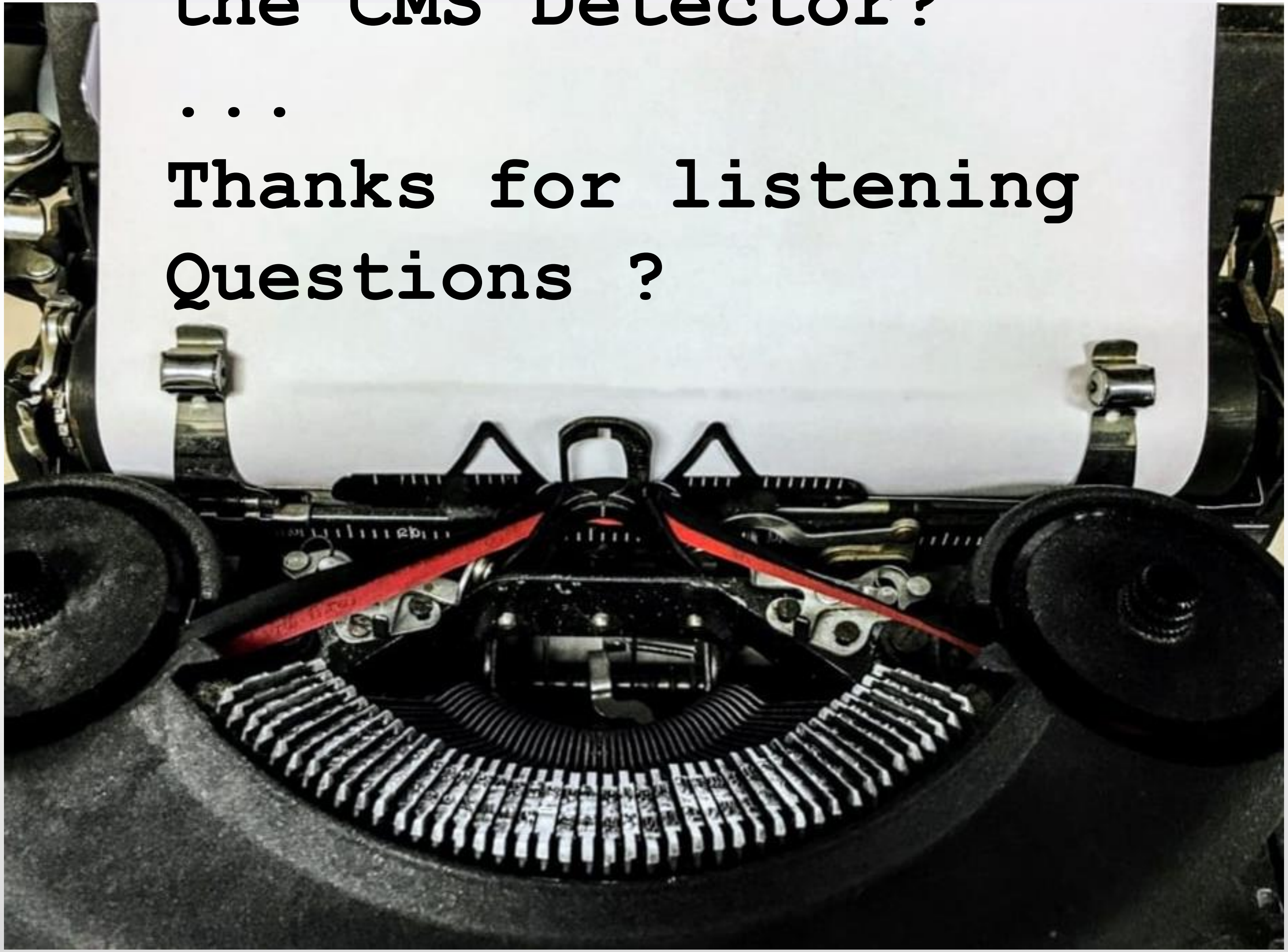
- *2nd Uncertainty*: Calculation of systematic uncertainty

the CMS Detector?

...

Thanks for listening

Questions ?



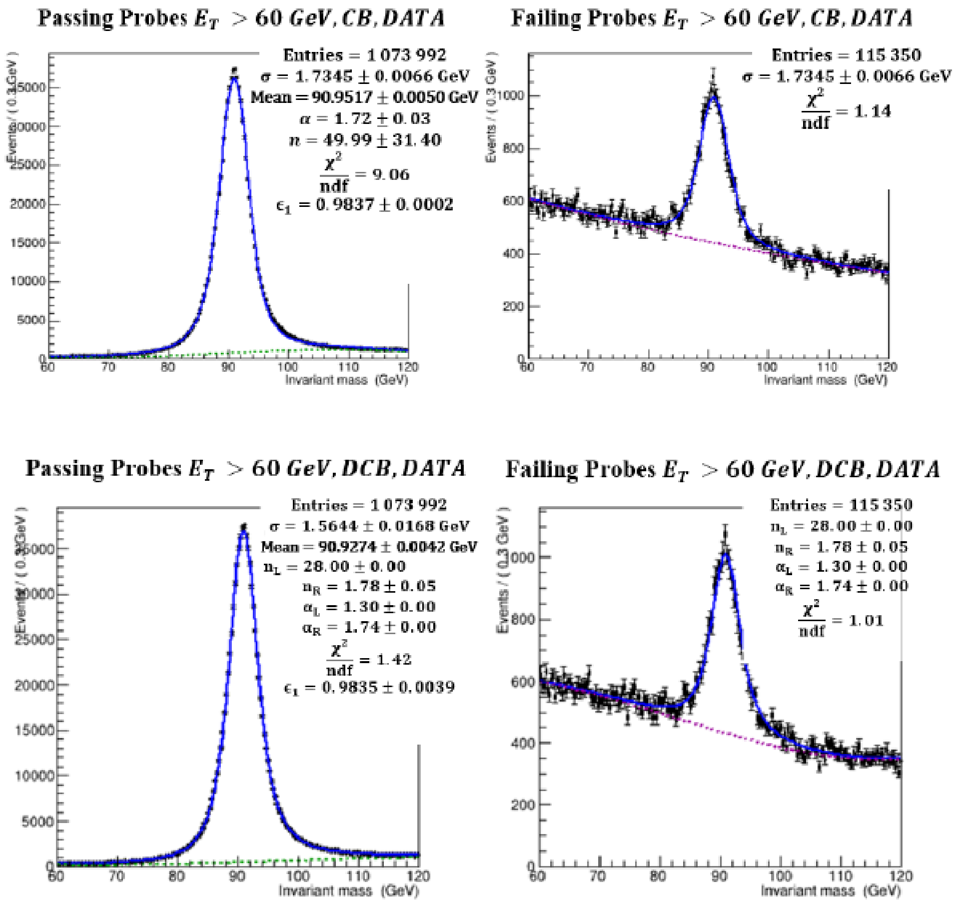
Backup…

Efficiency Measurements

Comparison of Efficiency Extraction Methods

Aspect	Histogram-Ratio	Simultaneous Likelihood Fit	Unbinned Extended Likelihood Fit
Data format	Binned histograms	Binned histograms	Raw (unbinned) events
Fit strategy	Simultanous fits for "pass" and "fail" samples, then calculate ratio manually.	Simultanous fits for "pass" and "fail" samples, then calculate ratio automatically	Same model as Simultaneous Fit, applied directly on unbinned events.
PDFs used	$f(m) = S \cdot P_{\text{sig}}(m) + B \cdot P_{\text{bkg}}(m)$	pass: $\varepsilon N P_{\text{sig}}(m) + (1-\varepsilon) N P_{\text{bkg}}(m)$ fail: $(1-\varepsilon) N P_{\text{sig}}(m) + \varepsilon N P_{\text{bkg}}(m)$	Same model, applied to individual events.
Efficiency	Computed after fit: $\varepsilon_1 = \frac{S_{\text{pass}}}{S_{\text{pass}}+S_{\text{fail}}}$	ε_2 is fitted directly.	ε_3 is fitted directly, with better precision.
Error propagation	Manual, based on yield errors.	Automatic, via RooFit covariance.	Automatic, via unbinned likelihood.
Correlations handled?	Slightly yes – ϵ isn't a parameter.	Yes – shared parameters are linked.	Yes – full parameter dependencies preserved.
Binning effects	May bias results.	Less sensitive.	No binning bias.
Best for...	Simple setups, high stats.	Moderate stats, automatic extraction.	Low stats, highest precision.

Extra Plots



Top: Fits using a single Crystal Ball (CB) for $E_T > 60$ GeV.
Bottom: Fits using a Double Crystal Ball (DCB) for $E_T > 60$ GeV.