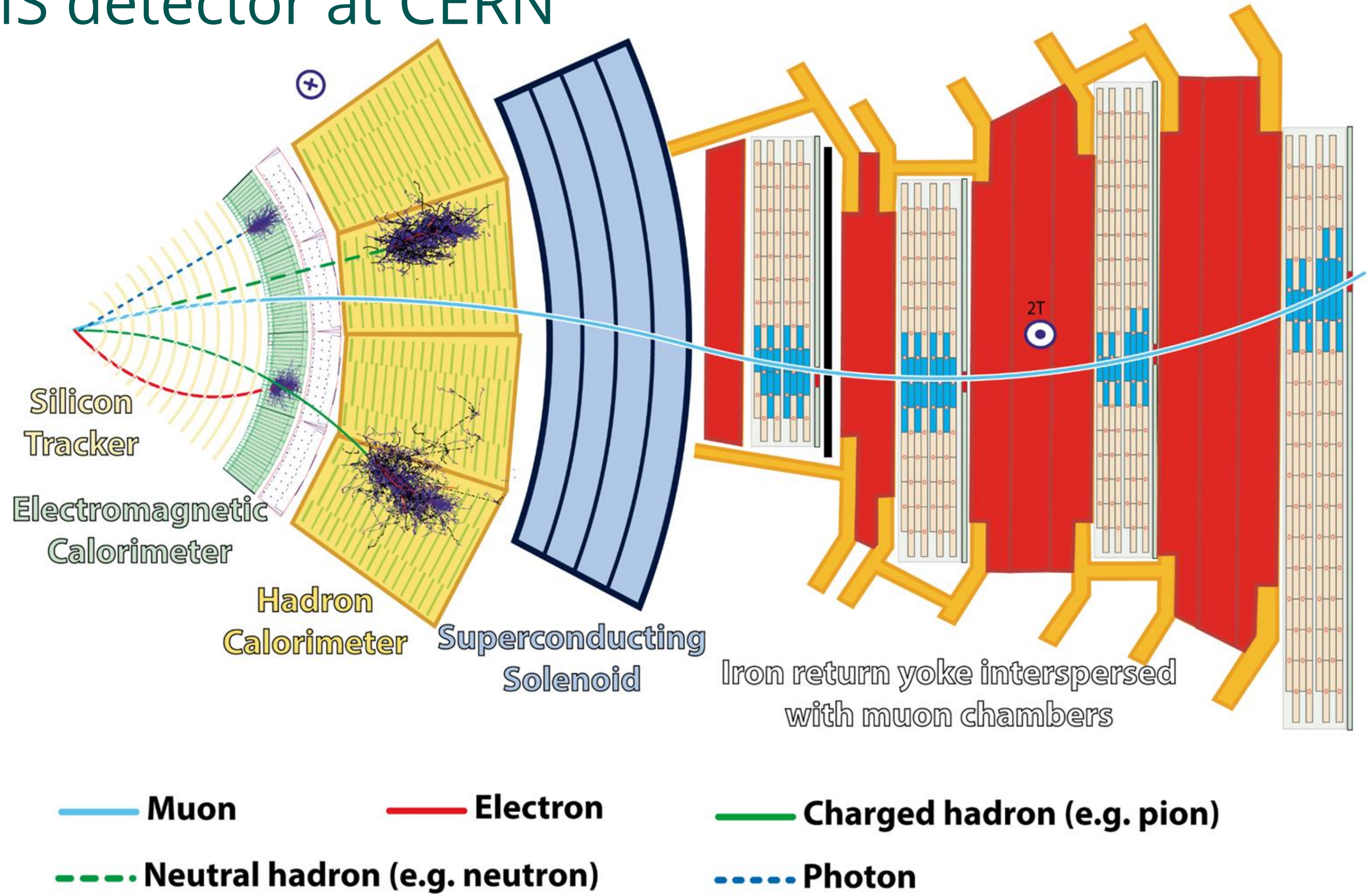
Measuring the low-p_T electron reconstruction efficiency for Higgs physics at CMS

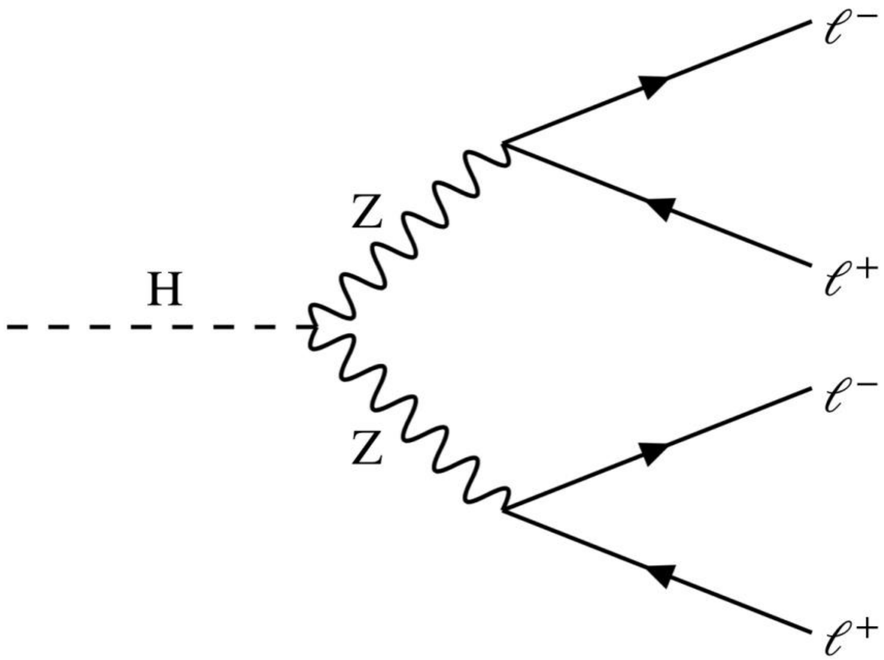
Antra Gaile[1], Jevgenijs Eugene Proskurins[1], Normunds Ralfs Strautnieks[2]

[1]Riga Technical University, Riga, Latvia
[2]University of Latvia, Riga, Latvia

CMS detector at CERN



We are interested in one of the main H boson decay channels:

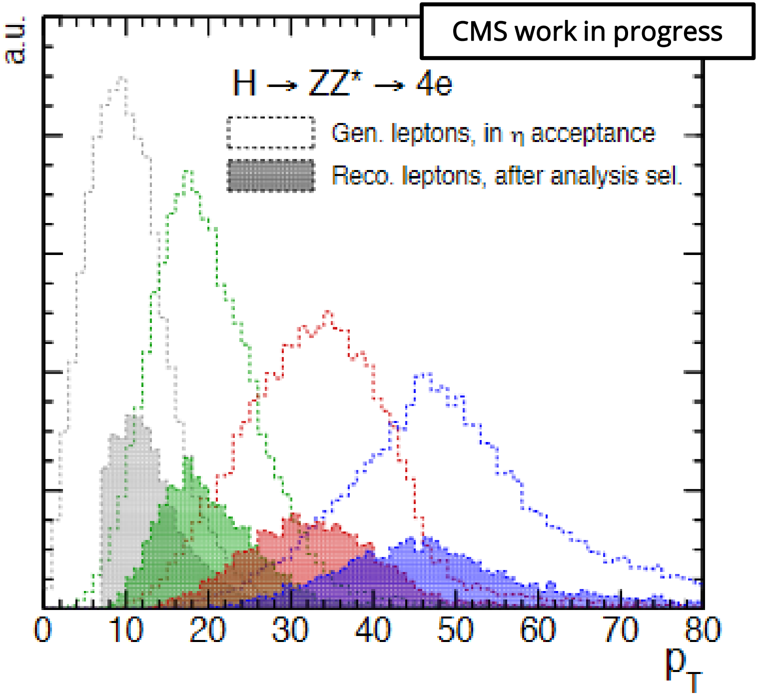


$H \rightarrow ZZ^* \rightarrow 4e$ decay channel

Z boson decays into two medium-p_T electrons ~ 45 GeV each on average.

Z* decays into two low-p_T electrons ~17 GeV[1] each on average.

To make precision measurements, all 4 leptons have to be reconstructed with high efficiency!



Sculac, T. (2019, January 9). *Measurements of Higgs boson properties in the four-lepton channel in pp collisions at centre-of-mass energy of 13 TeV with the CMS detector*. CERN Document Server. https://cds.cern.ch/record/2653094?ln=en

[1] In CMS, anything below 20 GeV is considered low-p_T

To estimate electron reconstruction efficiency a scale factor is needed

A scale factor between real data and simulation corrects any differences identified as resulting from **detector effects** or other effects**.**

An **ideal** set of selection cuts would have **identical** efficiency in data like in the simulation, and it would **efficiently eliminate the background** events while keeping the **signal rates untouched**.

In $H \to ZZ^*$ analysis, scale factors and their corresponding uncertainties enter the final event selection **with the power of four**.

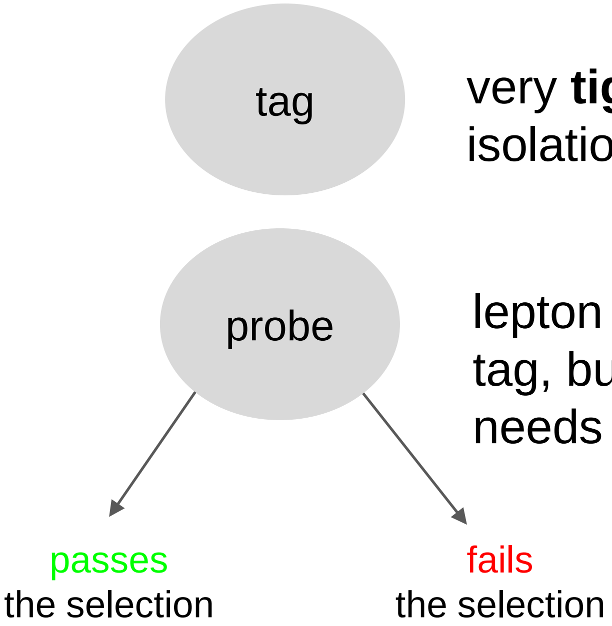
It is, therefore, crucial to measure them precisely with a good understanding of any deviations from unity.

The CMS collaboration *et al* 2021 *JINST* 16 P05014, doi: 10.1088/1748-0221/16/05/P05014

Sculac, T. (2019, January 9). *Measurements of Higgs boson properties in the four-lepton channel in pp collisions at centre-of-mass energy of 13 TeV with the CMS detector*. CERN Document Server. https://cds.cern.ch/record/2653094?ln=en

The Tag and Probe method is used to find the reconstruction and selection efficiencies, and scale factors

It chooses a tag and a probe from, for example, a Z →e+e- decay



very **tight selection** requirements (identification + isolation), trigger

lepton of the same flavor and opposite charge to the tag, but with as **loose selection** as possible; it needs to be unbiased!

Z peak is often used as the standard candle because these events are abundant, it is “very clean” and well studied, its invariant mass is close to Higgs boson mass!

The CMS collaboration *et al* 2021 *JINST* 16 P05014,
doi: 10.1088/1748-0221/16/05/P05014

The efficiency of the selection criteria is defined as the ratio of the number of **passing** probes NP to the **total** number of probes:

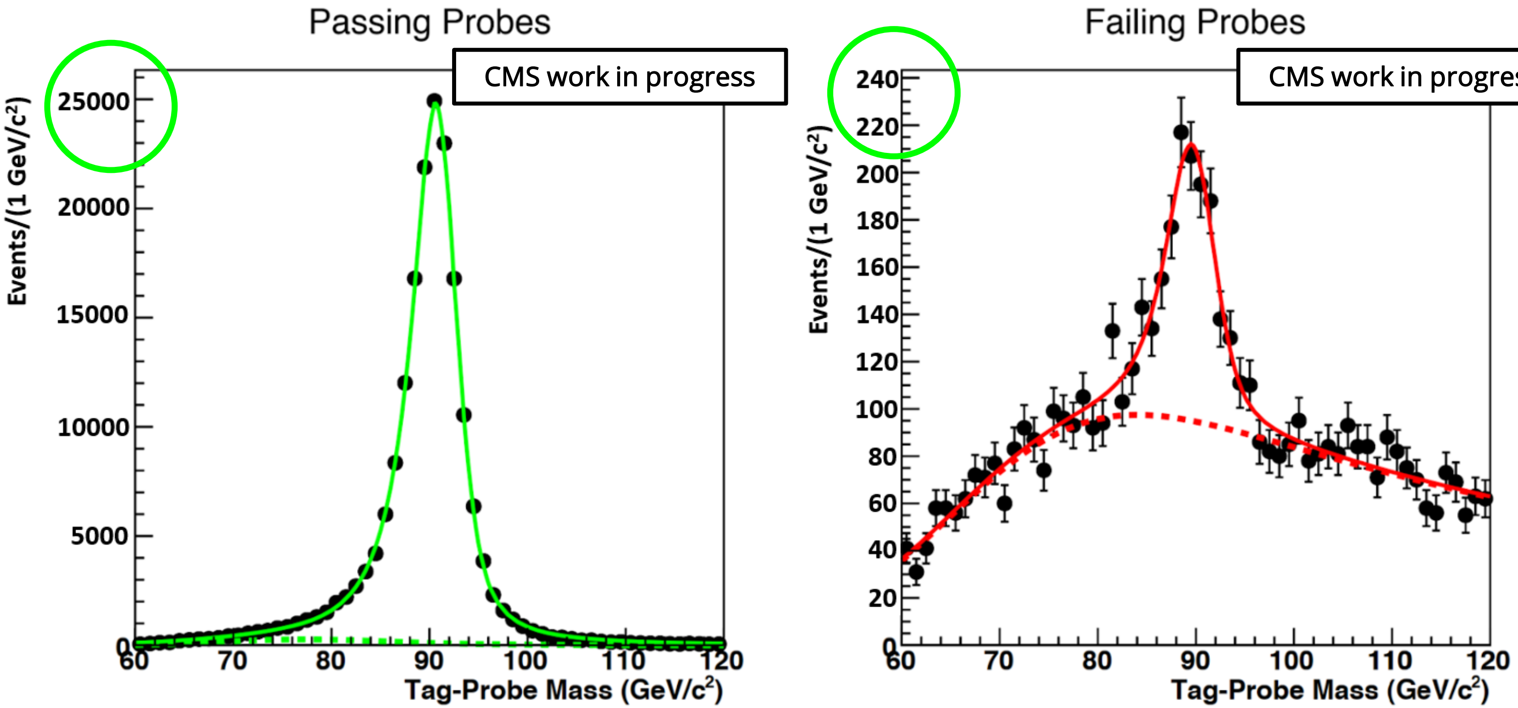
$$\epsilon_{sel.} = \frac{N_P}{N_P + N_F}$$

Scale factors are defined as a ratio of the efficiency of a given selection cut in **data** and the **simulation**:

$$SF_l(p_{\mathrm{T}}^l, \eta^l) \equiv \frac{\epsilon_{data}(p_{\mathrm{T}}^l, \eta^l)}{\epsilon_{MC}(p_{\mathrm{T}}^l, \eta^l)}$$

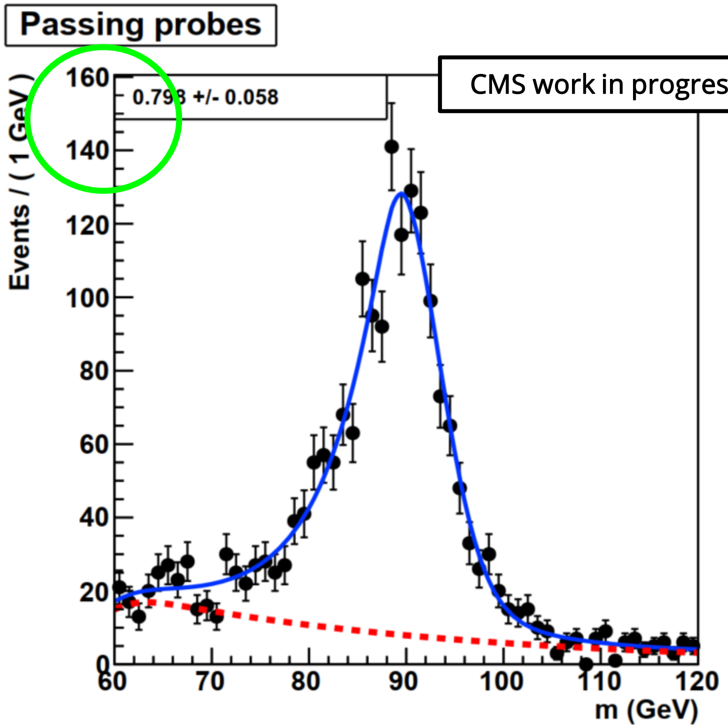
This procedure is usually performed in bins of pt and η of the probe electron, to measure efficiencies as a function of those variables.

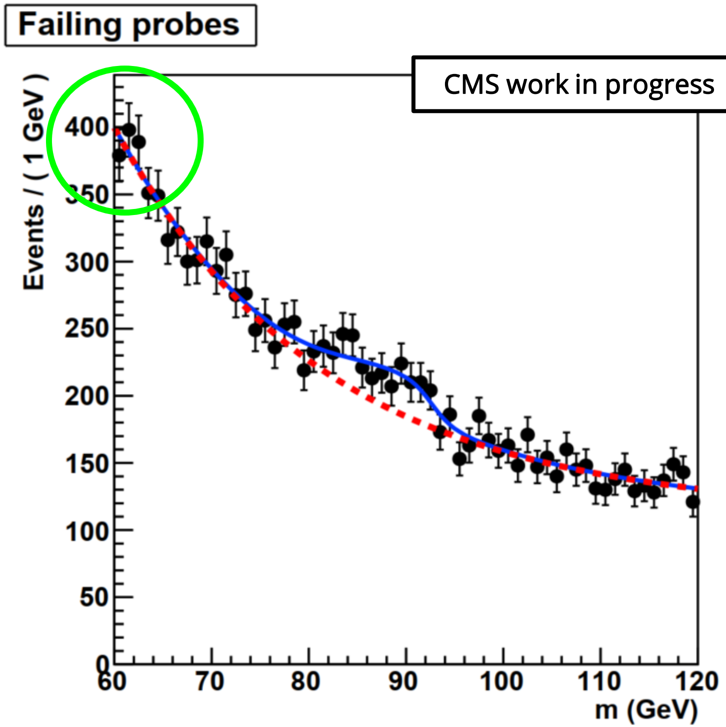
In the **high-p_T** electron case, only a couple of percent of events ends up as failing probes



Sculac, T. (2019, January 9). *Measurements of Higgs boson properties in the four-lepton channel in pp collisions at centre-of-mass energy of 13 TeV with the CMS detector*. CERN Document Server. https://cds.cern.ch/record/2653094?ln=en

In the **low-p_T** electron case, the number of failing probes are of same order as passing probes



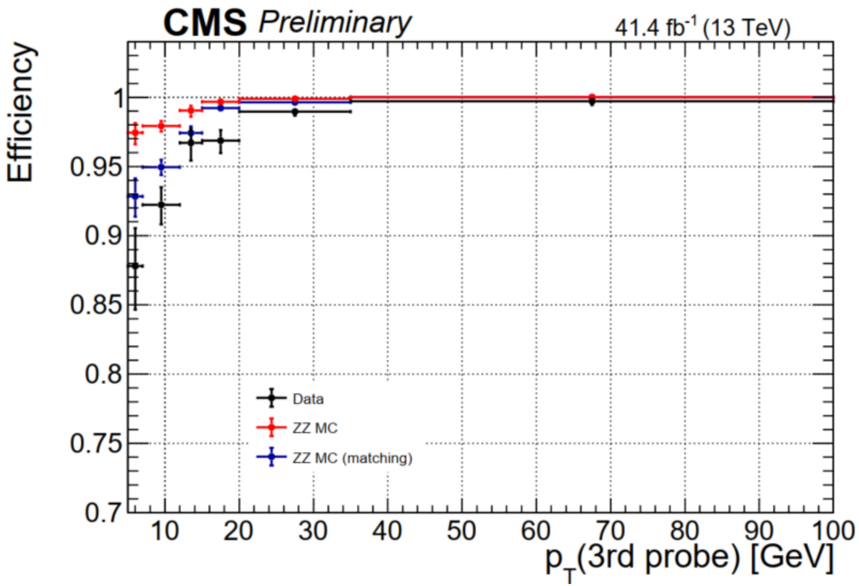


Sculac, T. (2019, January 9). *Measurements of Higgs boson properties in the four-lepton channel in pp collisions at centre-of-mass energy of 13 TeV with the CMS detector*. CERN Document Server. https://cds.cern.ch/record/2653094?ln=en

Problems arise from the background in low-p_T region

Background mostly comes from QCD multi-jet production and usually is at a few percent levels, but it **increases exponentially in the low-p_T region.**

Therefore it is extremely challenging to distinguish any signal amongst the background.



Sculac, T. (2019, January 9). *Measurements of Higgs boson properties in the four-lepton channel in pp collisions at centre-of-mass energy of 13 TeV with the CMS detector*. CERN Document Server. https://cds.cern.ch/record/2653094?ln=en

It is not a problem only for $H \rightarrow ZZ^* \rightarrow 4e$ analysis

Due to low-pt electron reconstruction efficiency, the **systematic uncertainty is high.**

It is the **highest contributor** to the overall systematic uncertainty in the $H \rightarrow ZZ^*$ decay channel

By improving the reconstruction efficiency, we could reduce this systematic error, and additionally, it would benefit any group that uses low-p_T electrons in their analysis!

Near future goals

By using the TnP method on **J/ψ resonances**, we aim to increase electron reconstruction efficiency in the low-p_T region.

After the analysis of the J/ψ meson is complete, it is possible to conclude whether for future low-p_T electron studies it is more suitable to:

1) use the same triggers as previously, but with different data selection criteria;
2) create new HLT paths to collect more precise data and signal efficiency in the low-p_T region;

This could greatly decrease the largest uncertainty and would benefit other analyses in the H$\rightarrow$ZZ* group.

Many thanks to our mentors - Ana, Toni, and Christophe - for guiding us through the vastness of HZZ and electron analysis!

Backup slides

The Compact Muon Solenoid (CMS) experiment is one of the four large experiments based on the Large Hadron Collider (LHC) at CERN. CMS specializes in the ultra-precise reconstructions of muons and uses its onion-like sub-detector structure and its 4T solenoid for the reconstruction and identification of various particle species.

CMS has a high charged-particle reconstruction efficiency in the high transverse momentum (pT) range, however, increasingly, further improvement of a plethora of physics results relies on an improved charged particle, especially, lepton reconstruction at the lower pT range. One such analysis is the H->ZZ analysis, where the inability to determine the lepton, especially electron, reconstruction efficiency is one of the leading systematic uncertainties.

A team of early-career researchers from Riga Technical University and the University of Latvia has joined the CMS Higgs Physics Analysis Group (Higgs PAG) to work on the improvement of the low-pT electron reconstruction efficiency determination. In this talk, we will broadly cover the reasons for needing and key obstacles to achieving an improved electron reconstruction efficiency at CMS as well as the work-in-progress status of the current electron efficiency studies within the Higgs PAG.

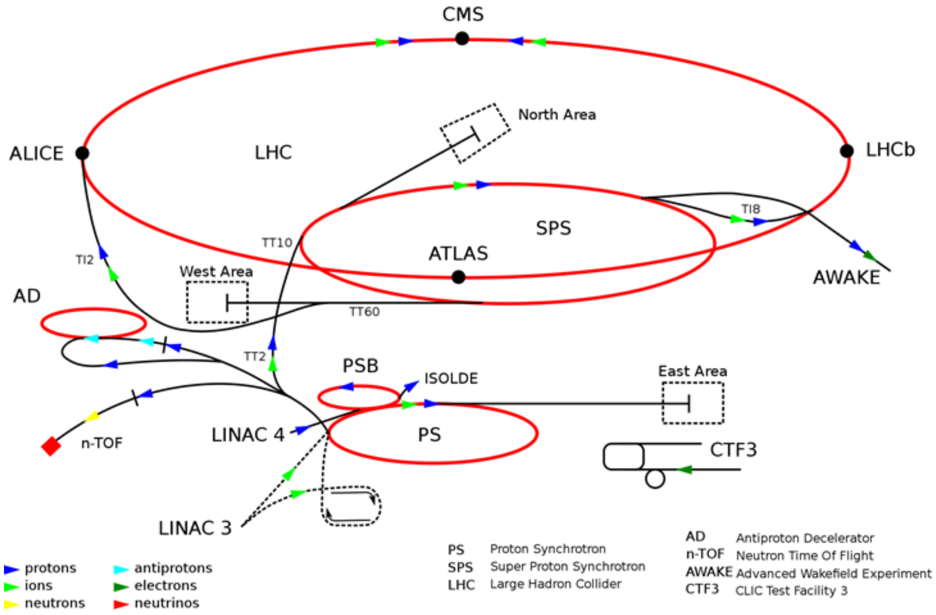
CMS detector at CERN

CMS - one of the 4 large detector experiments at CERN

CMS and ATLAS are the two largest detectors at CERN, built at the same time and designed to complement each other

CMS - is a superconducting solenoid of 6m internal diameter, a length of 21.6 m, and a total weight of 12500 tons, providing a magnetic field of 3.8 T.

CMS detector specializes in ultra-precise muon reconstruction but is used in very broad research topics of particle physics

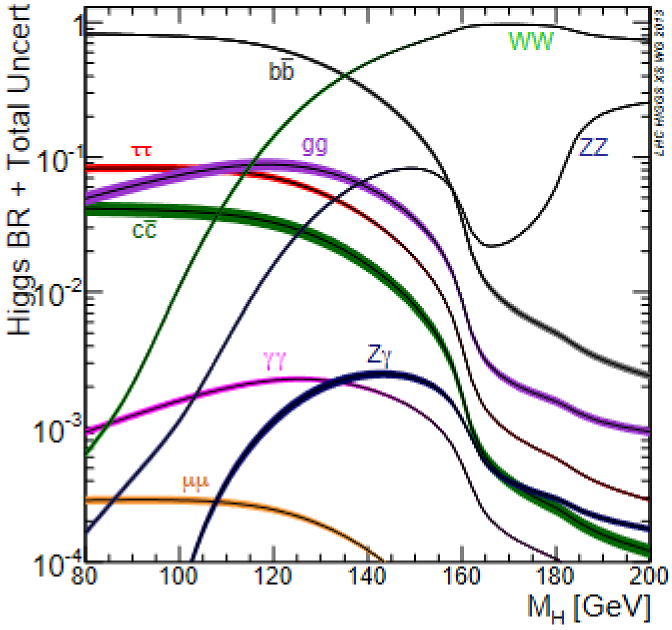


Measurements of the H boson properties

One of the main decay channel for the H boson is H->ZZ*

It gives us the possibility to fully reconstruct the final state and estimate the precise properties of the H boson

Has generally good signal to background ratio



H→ZZ* decay channel

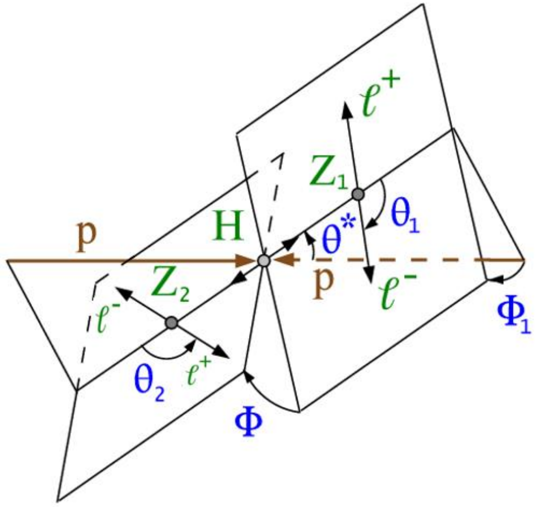
As H boson (125 GeV) decays to two Z bosons, it is clear that they both cannot be regular bosons

One of the decay products is a regular Z boson ~91 GeV

The other one is Z* off-shell boson ~34 GeV

Both of these bosons decay to 2 leptons

In our study of interest - each of them decays to 2 electrons



electron selection

1) Kinematic and impact parameters:
   aims to reduce backgrounds from displaced electron candidates (photon conversions and the B-meson decay) - differentiate from signal electrons by their distance from the primary vertex;
2) electron identification (ID):
   to remove sizable contamination of fake electrons mostly from hadronic jets and photon conversions (e.g. $\pi^0 \rightarrow \gamma\gamma$);
3) electron isolation:
   a tool to suppress the dominant fake electron background from hadronic jets, a signal electron will not be surrounded by other particles coming from the same hard interactions (other PF candidates in a cone around the electron is a sign of fake electrons).

electron selection

$H \to ZZ^{\star} \to 4\ell$ → Signal electrons

the Drell-Yan sample → BG electrons

the per-lepton efficiency propagates to event selection with the power of four:

1) one has to study carefully the dependence of the signal and background efficiencies on the candidate kinematics;
2) signal and background efficiency are expected to have no pileup dependence;
3) the agreement between the efficiency in data and simulation should be cross-checked and validated in control regions with sufficient statistics.

a selection is performed on signal electrons and background or fake electrons in three sub-selection steps:

1) kinematic and impact parameter requirements;
2) identification;
3) isolation.

https://cds.cern.ch/record/2653094?ln=en (T.Šćulac, PhD Thesis)

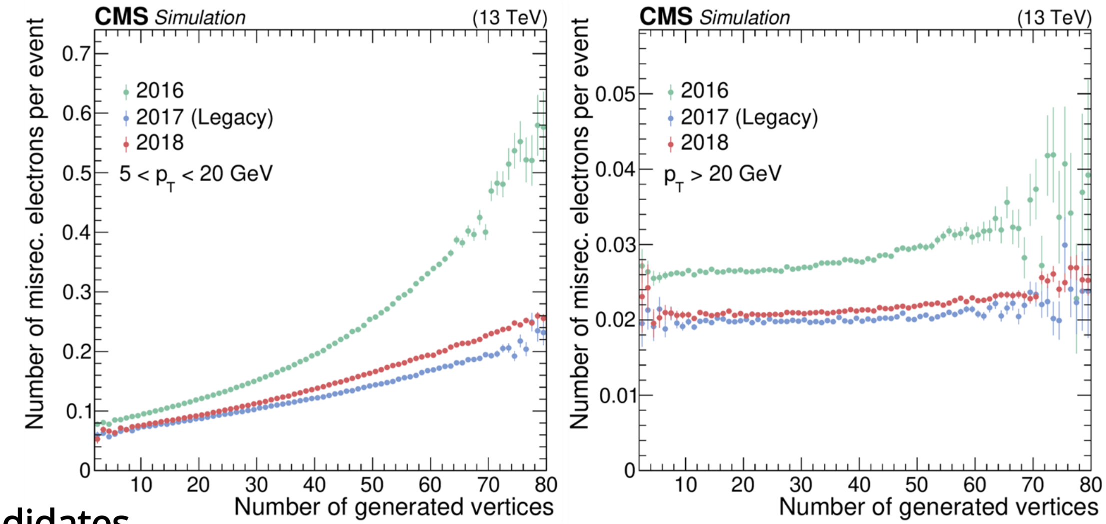
Uncertainties

A total uncertainty for the measurement of the scale factors is a quadratic sum of statistical and systematic uncertainties.

1) Statistical uncertainties are obtained directly from the fits.
2) Systematic uncertainties require a detailed study:
   - uncertainty in the accuracy of the signal model
   - uncertainty in the background modeling
   - uncertainty in the background coming from the tag selection
   - uncertainty in the overall event description
   - uncertainty in the pileup estimation

The total systematic uncertainty **is of the order of 10%** for the low-pT bins and around 1-2% otherwise.

Number of misidentified electron candidates per event



Number of **misidentified electron candidates** per event as a function of the number of generated vertices in DY + jets MC events simulated with the different detector conditions of the Run 2 data taking period. Results are shown for electrons with pT in the range 5-20 GeV (left) and electrons with pT> 20 GeV (right) without further selection.

The CMS collaboration *et al* 2021
JINST 16 P05014, doi: 10.1088/1748-0221/16/05/P05014