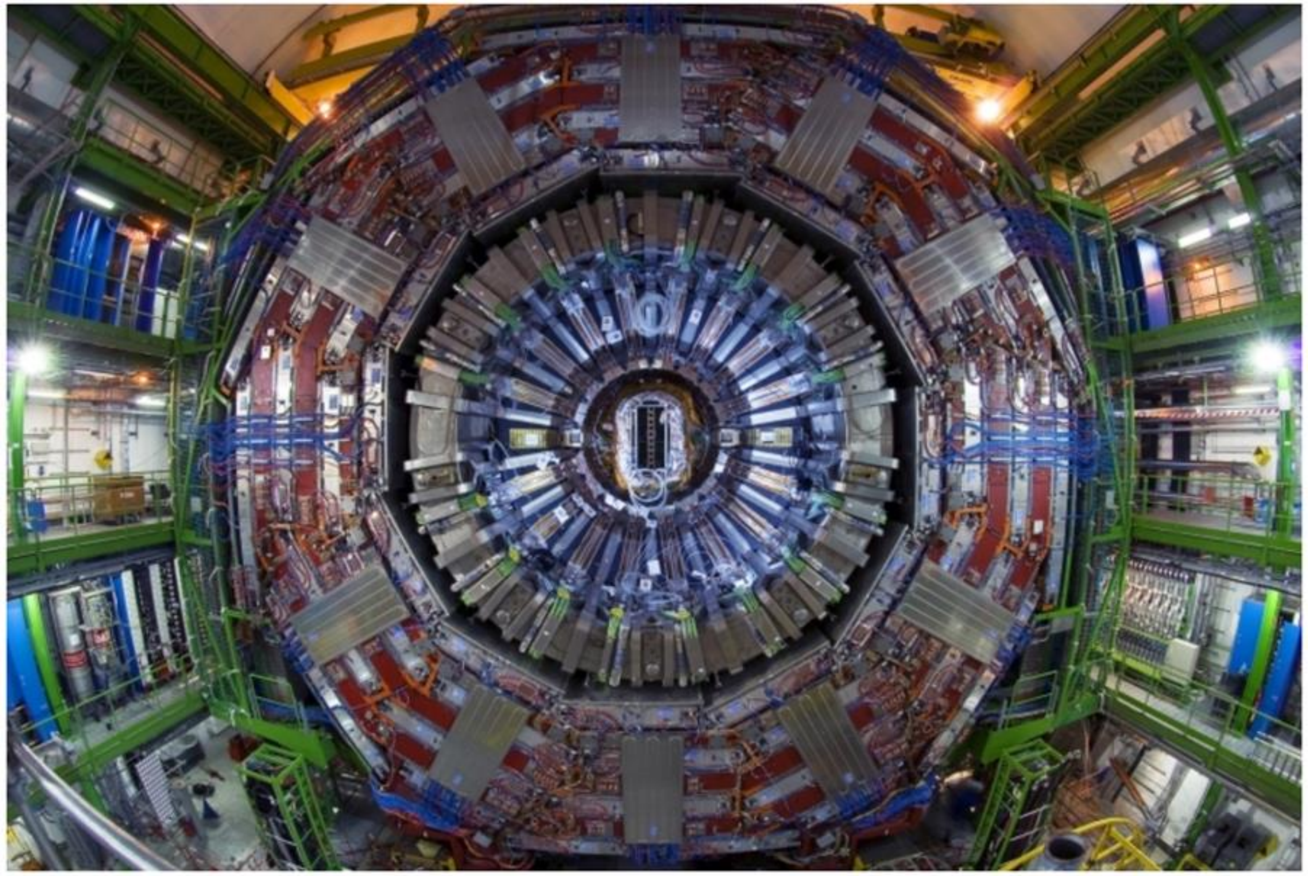
# Vector Boson Scattering
# Same Sign W Boson in CMS Experiment

Kourosh Mousavi
Supervisor: Prof. Michal Szleper

National Center for Nuclear Research
NCBJ, Warsaw, Poland
19th December 2024



# Outline

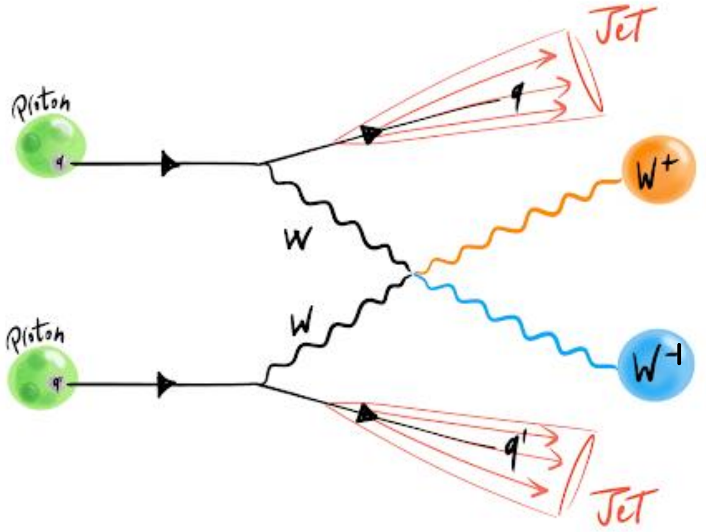
1. Vector Boson Scattering, its importance in SM and BSM
2. LHC, CMS detector, and Run 3
3. Signal and backgrounds
4. Identification, how to separate signal from background
5. Type of leptons
6. Definition of the signal region, identification of backgrounds
7. Definition of control regions and their role

# Vector Boson Scattering (VBS) in the SM

By vector boson, I mean W, Z boson, and photon here

What is VBS?

When two vector bosons scatter each other

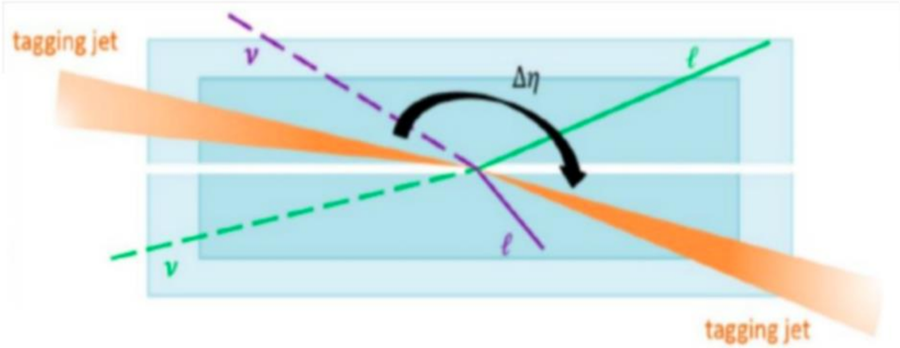



Signature of Vector Boson Scattering:

There are two important signature for VBS events:

1- Production of two energetic jets with large invariant mass
2- These two jets are separated by high rapidity from each other

A VBS event in the CMS detector

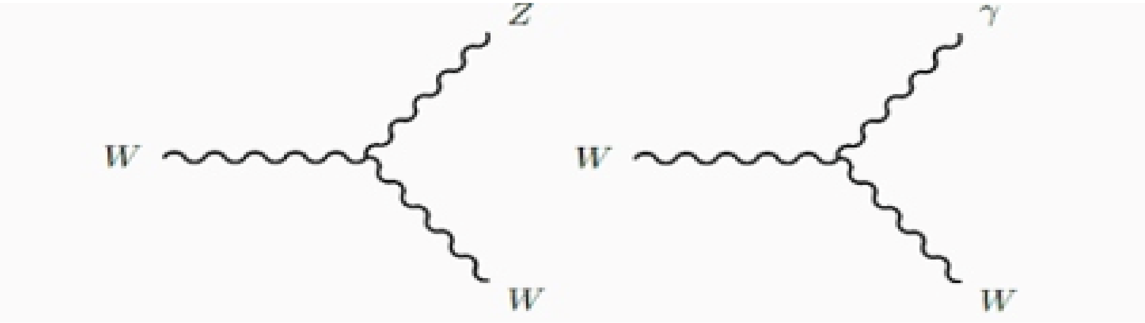


# Vector Boson Scattering (VBS) in the SM

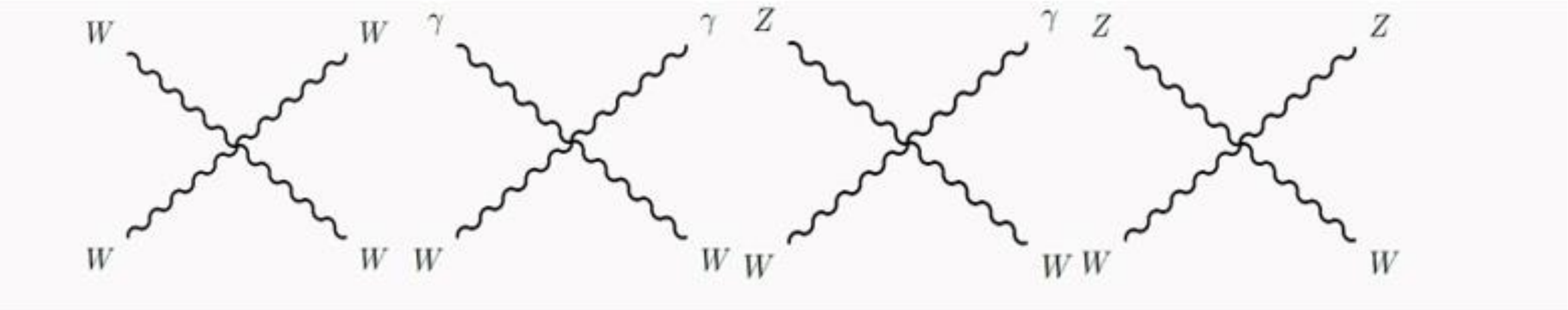
Why VBS is important?

- The VBS Lagrangian includes Triple and quartic gauge coupling constants

Triple




Quartic




Studying triple and quartic gauge couplings is important to test the SM

# Beyond the SM and VBS

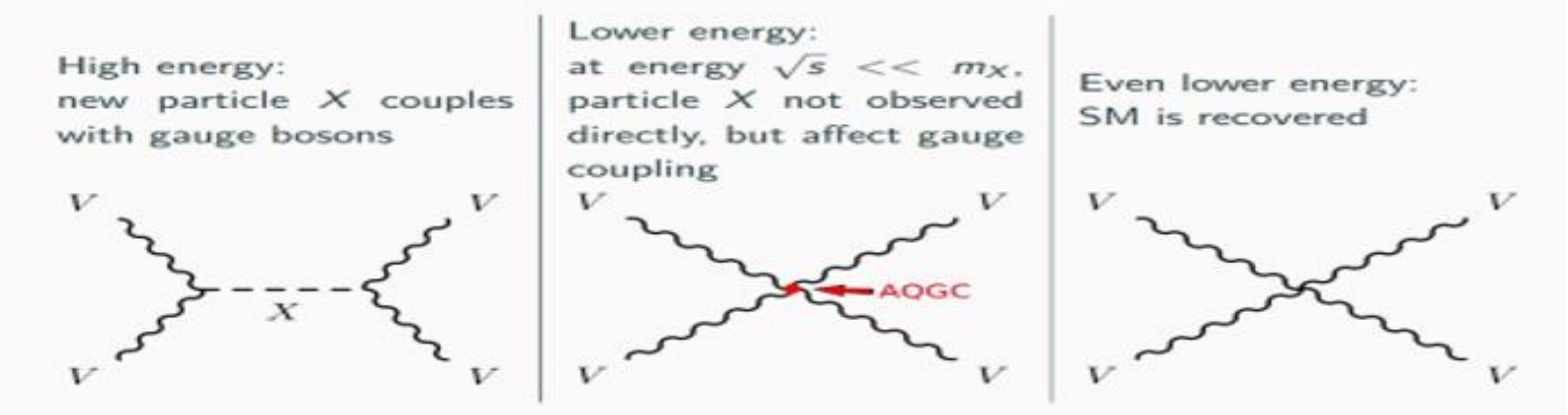
at low energy, one can extend the SM by adding higher dimension operators:

$$\mathcal{L} = \mathcal{L}_{SM} + \sum_{d>4} \sum_{i} \frac{C_i}{\Lambda^{d-4}} \mathcal{O}_i^d$$

where $\mathcal{O}_i^d$ are the new $d$-dimension operators, $C_i$ are dimensionless parameters and $\Lambda$ is the new physics scale.

By studying VBS, we can put limits on dimension-6 and dimension-8 parameters: vector boson fusion/scattering

Simple example to illustrate the idea:
Suppose there is a new particle X interacting with vector gauge bosons. Even if we do not have enough energy to produce it, its effect could be observed as an anomalous coupling.



## What is the Large Hadron Collider (LHC)?

- The LHC is the world's largest and most powerful particle accelerator.
- Protons (tiny building blocks of atoms) are sped up to nearly the speed of light and smashed together.
- This creates extremely high-energy collisions, recreating conditions similar to the Big Bang.

## Why Do We Smash Protons?

- To study the smallest particles in the universe.
- Understand fundamental forces and search for new physics (e.g., dark matter).
- Discover particles like the Higgs boson (the particle that gives others mass).

# What Does CMS Detect?

1. **Electrons:** Tiny particles with a negative charge, fundamental components of atoms.

2. **Muons:** Similar to electrons but about 200 times heavier.
3. **Photons:** Particles of light with no electric charge.
4. **Jets:** Sprays of particles produced when quarks or gluons fragment.
5. **Missing Transverse Energy (MET):** Represents invisible particles like neutrinos escaping the detector.

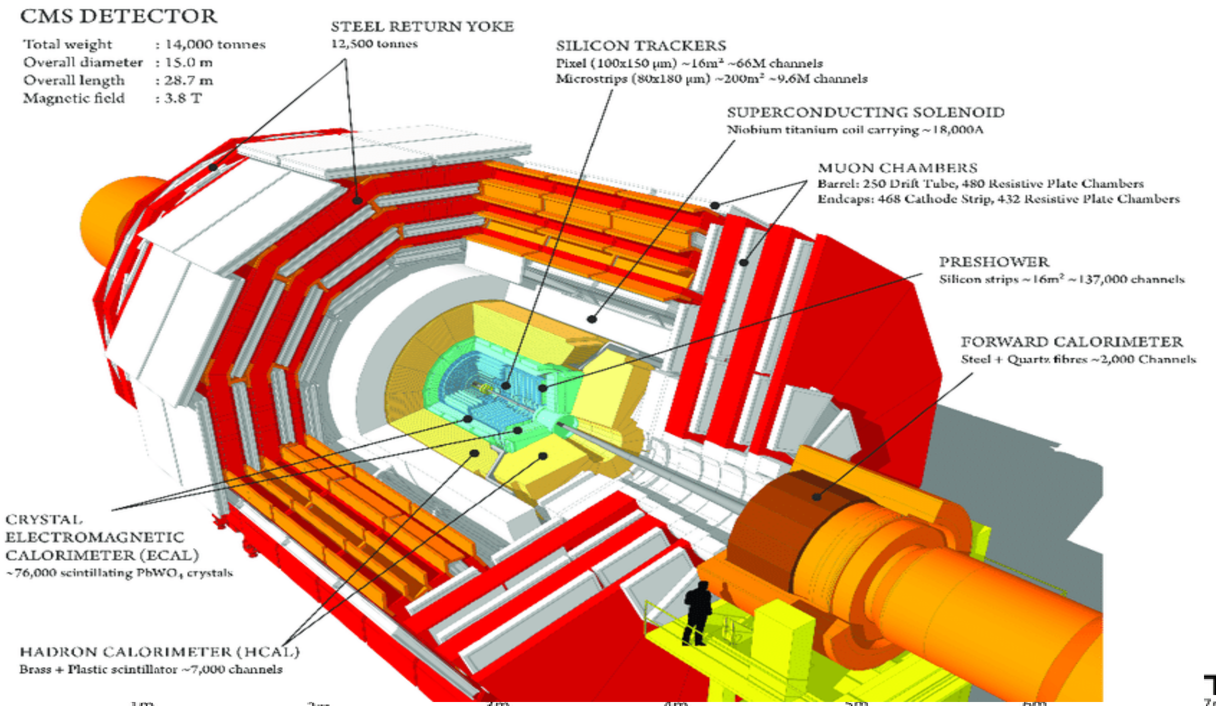
- **How CMS calculates it:**
- Measured by summing the energy and momentum of all visible particles.
- Any imbalance indicates missing energy carried away by neutrinos.

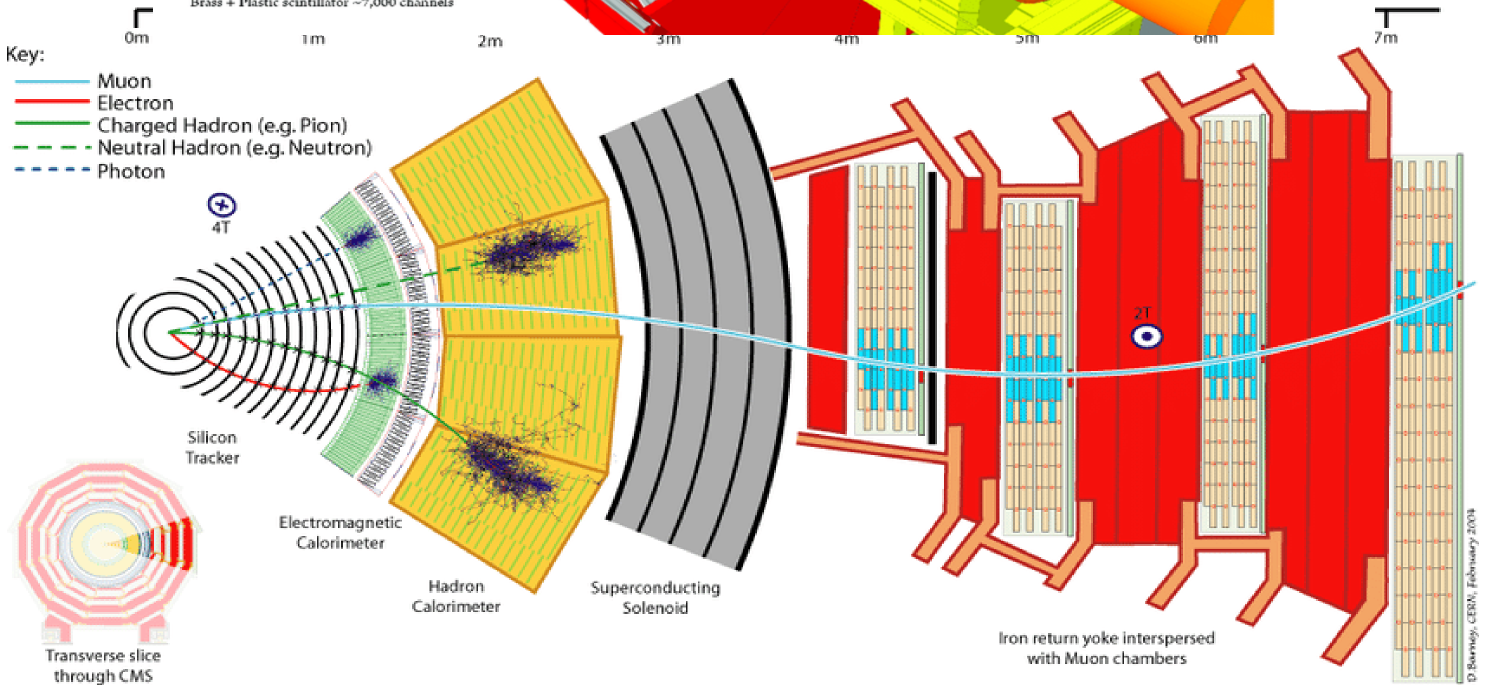
# CMS Detector

- **What is the CMS Detector?**
- CMS stands for Compact Muon Solenoid.
- It's a giant, layered machine that surrounds the collision point and acts like a "3D camera."

It is a **5-story-tall cylindrical detector** made of multiple layers:

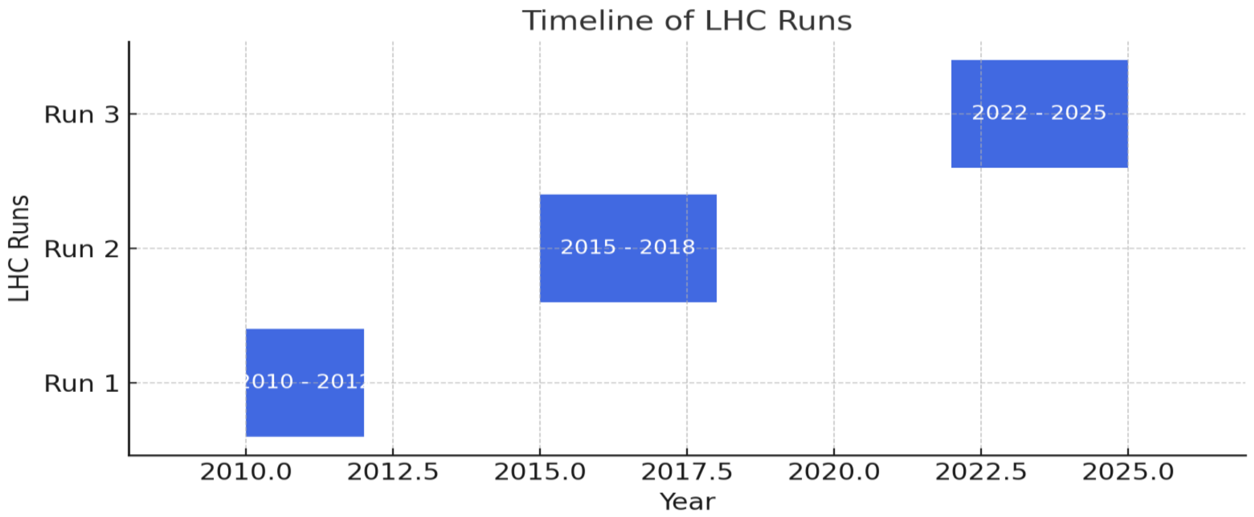
1. **Tracker:** Measures the paths of charged particles.
2. **Electromagnetic Calorimeter (ECAL):** Measures energy of photons and electrons.
3. **Hadronic Calorimeter (HCAL):** Measures energy from jets (quarks/gluons).
4. **Solenoid Magnet:** Produces a strong magnetic field to bend charged particles.
5. **Muon Chambers:** Detects muons, which pass through other layers.



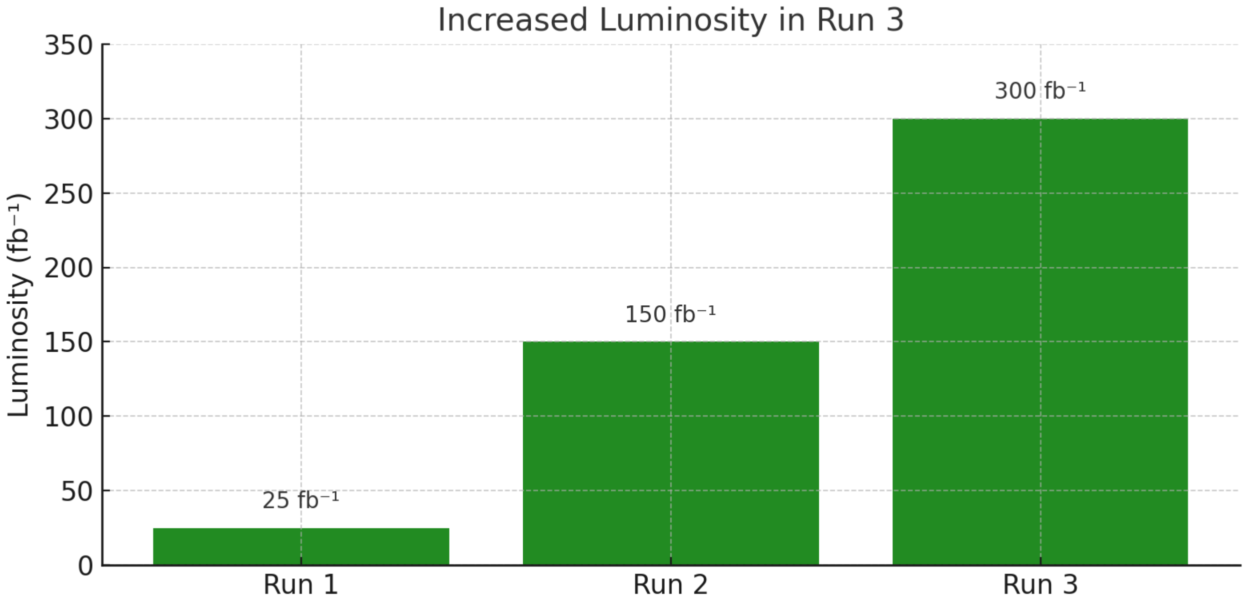



# LHC and Run 3

**What is Run 3?**
•**Run 3** is the third operational phase of the LHC, which began in **2022** and runs until 2025.
•The LHC is operating at its **highest energy (13.6 TeV)** to date and with improved detectors.




**Goals of Run 3:**
- Collect **more data** than ever before to improve precision.
- Look for **new physics** beyond the Standard Model (e.g., dark matter, exotic particles).

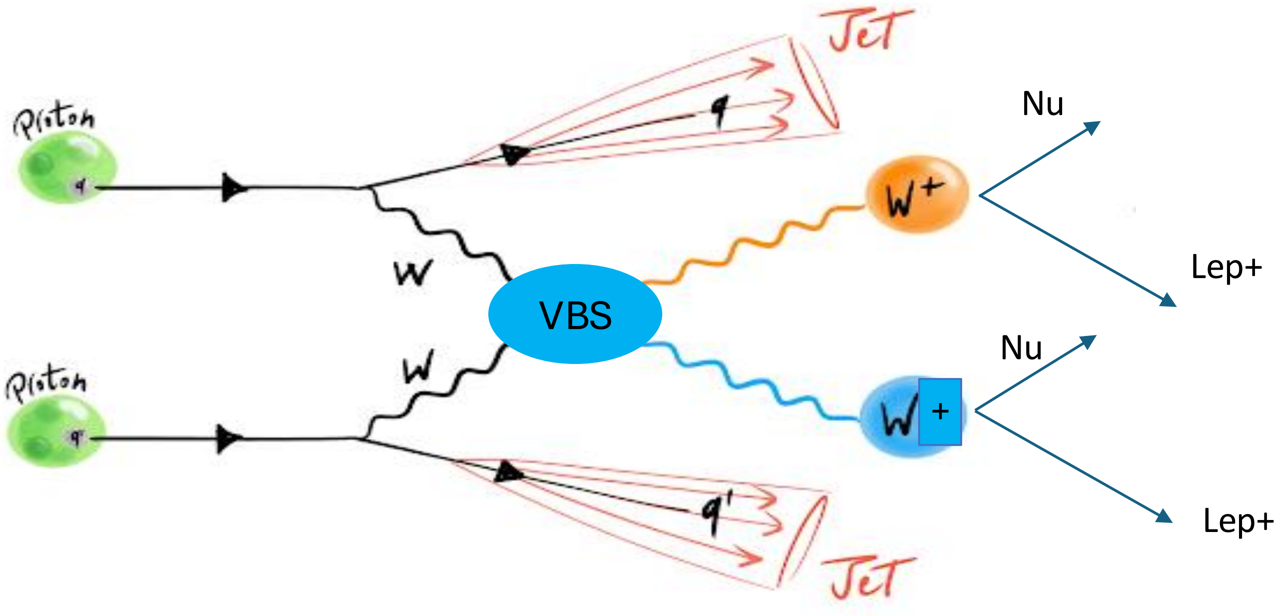


# My Analysis:
# Same sign W Boson Electroweak Scattering

Signal:

The **electroweak** production of same-sign W boson pairs in the two jets and two same-sign leptons final state

$$q\,q \to W^+W^+q\,q \to \ell^+\nu\;\ell^+\nu\;q\,q$$



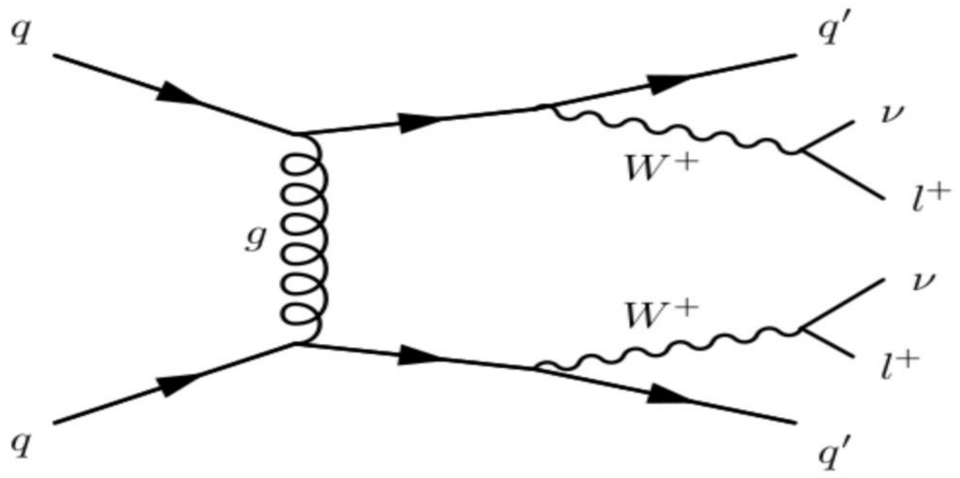
# Same sign W Boson Scattering

Background:

The background events can be split in two categories:

1- **Irreducible backgrounds:**

- have the **same final state particles** as the signal process and cannot be distinguished from VBS on an event-by-event basis.



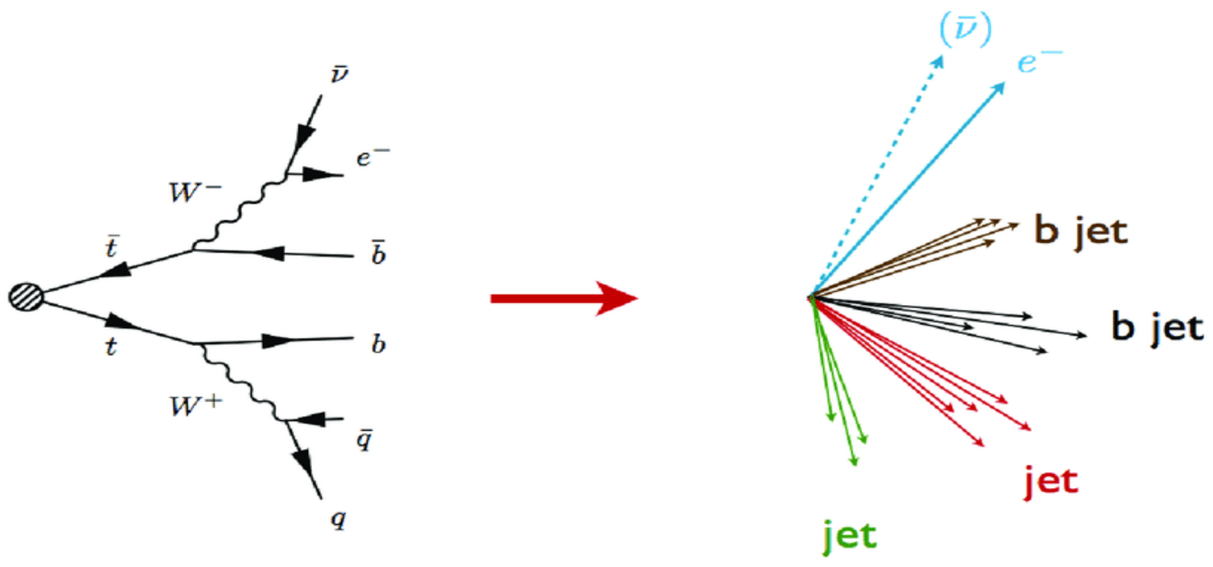

QCS same-sign WW -> 2l + 2nu

# Same sign W Boson Scattering

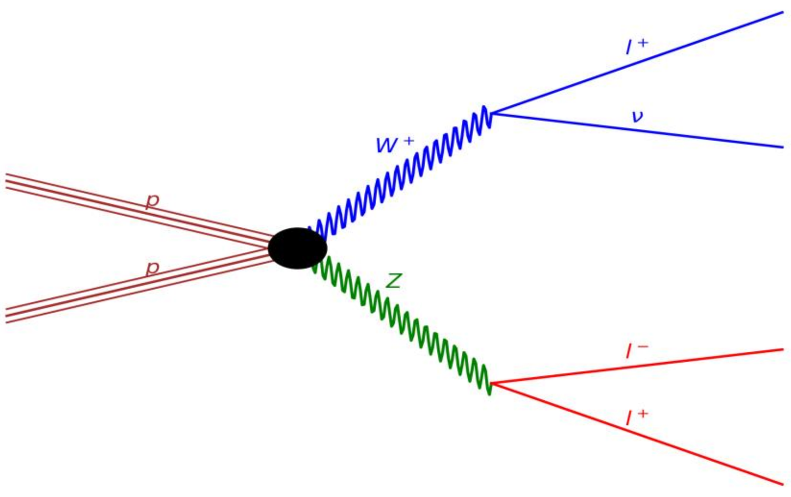
## Background:

2- **Reducible backgrounds:** are processes with **different final state from VBS**, but can enter the signal region.

One example 1:
when objects are not correctly reconstructed (e. g. a jet that is misreconstructed as a lepton)
(Like ttbar->L+2Q and W+Jet -> L+Nu+Jet)




One example 2:
when objects r are not within the detector acceptance (abs(Lepton_Eta) < 2.5).




W+ Z -> 3 leptons + nu

# Identification

How to separate signal from backgrounds

1- Define suitable kinematic for leptons to make them close to real

2- Signal region (with specified cuts and conditions) to allow the signal events separate from backgrounds

3- Identifying all background processes and simulate them by MC samples

4- Use data-driven method to count non-prompt background contribution

*** In following slides, I go step by step through each stage

# Type of Leptons

Generally, we categorize leptons in two types:

**1- Prompt Lepton:**

A **prompt lepton** is a lepton (electron, muon, or tau) that originates directly from the **primary vertex** of a particle collision.

It is typically produced in **hard processes** such as:

- Decays of bosons (e.g., $W \rightarrow e + nu$, $Z \rightarrow e+ \ e-$).
- Decays of heavy quarks (e.g., top quarks).

**2- Fake Lepton:**

A **fake lepton** is a reconstructed object in the detector that is **misidentified as a lepton**, but it does not originate from a genuine lepton.

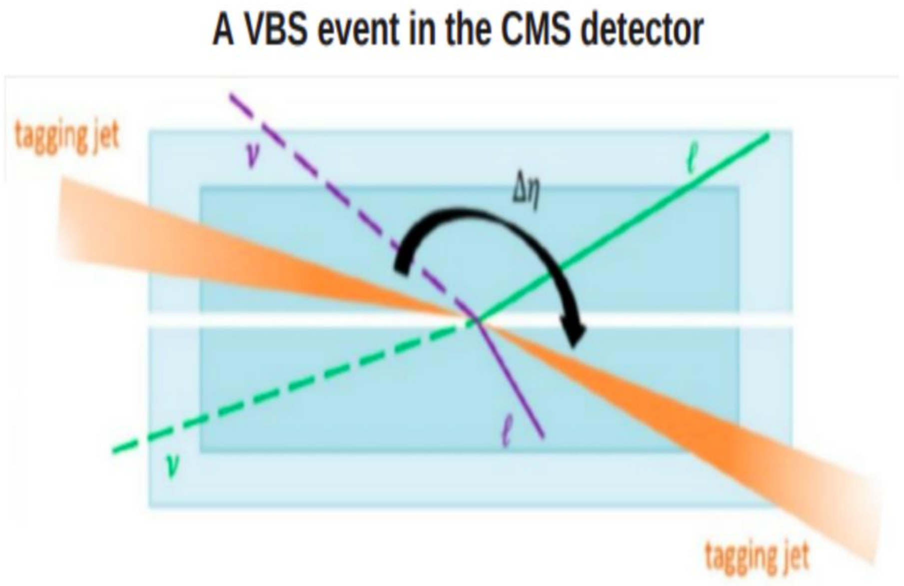
Common sources of fake leptons include:

- **Hadrons misidentified as leptons**: Charged pions or kaons can mimic leptonic signals in certain detector layers.
- **Non-prompt leptons**: Leptons produced in secondary decays, such as those from hadron decays (e.g., $b \rightarrow e+X$, $c \rightarrow \mu+X$)

## Definition of Signal Region (SR)

We first need to define signal region which is enriched for signal events leading separating them from other backgrounds

| | Signal Region Cuts |
|---|---|
| Pt_Leptons | Pt1 > 25 GeV, Pt2 > 20 GeV |
| Pt_Jets | >50 GeV |
| Pt_Miss (Pt of Neutrino) | >30 GeV |
| Anti Btagging | applied |
| Abs(Delta_Eta(Jet1, Jet2)) | >2.5 |
| Invariant Mass of Diljet (M_jj) | >500 GeV |



## Identifying All Background Processes

Now we need to identify all background processes contribute in SR and simulate them by MC generators (like MadGraph)

| **Backgrounds Processes** |
|---|
| WZ -> 3l + nu |
| QCD same-sign WW -> 2l + 2 nu |
| Gamma + Vector Boson (VG) -> 3l |
| ZZ -> 4l |
| Misscharge Identification |
| Three Vector Bosons Scattering (VVV) |

***** There is another important background (non-prompt background) which cannot be obtained by simulation and need to be calculated by data-driven method directly from data**

# Data-Driven Method to Compute Non-Prompt Background

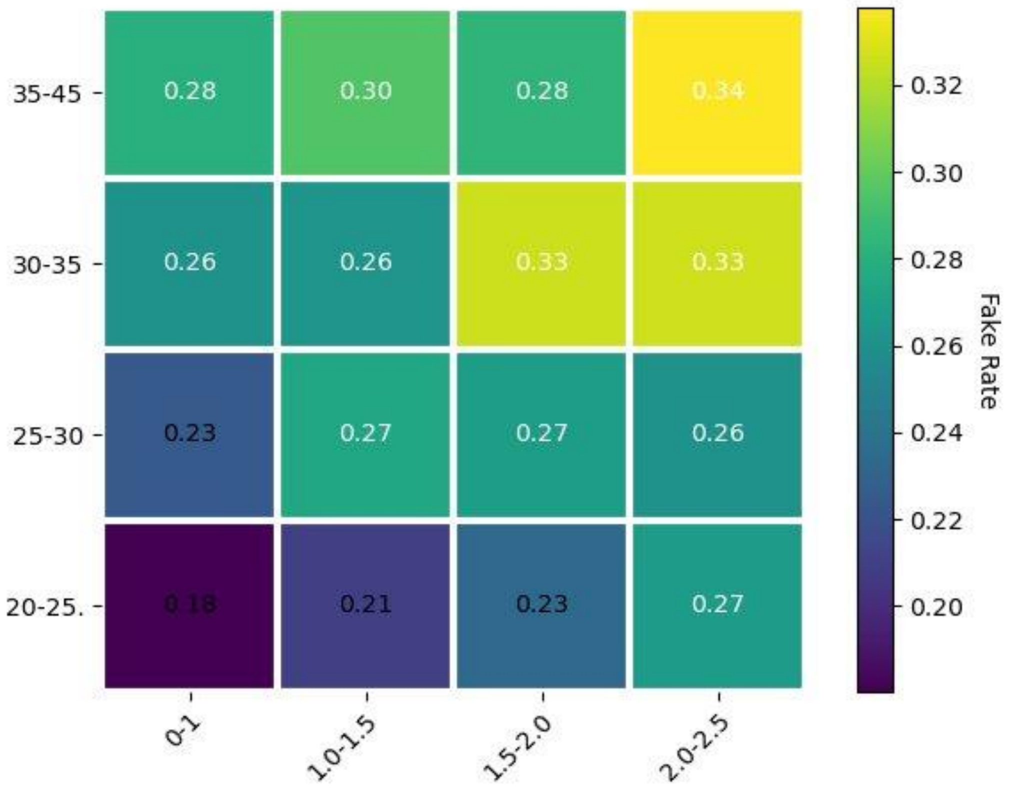
The data-driven (dijet) method that is used to estimate the non-prompt background

In this method, we compute the probability for a fake lepton passing a prompt lepton conditions which is called fake rate

The fake rate is computed from this formula:

Ɛ(fake rate) = NP/(NP+NF)

NP = number of prompt lepton,  NF = number of fake lepton




Muon fake rate as a function of pt and eta of lepton

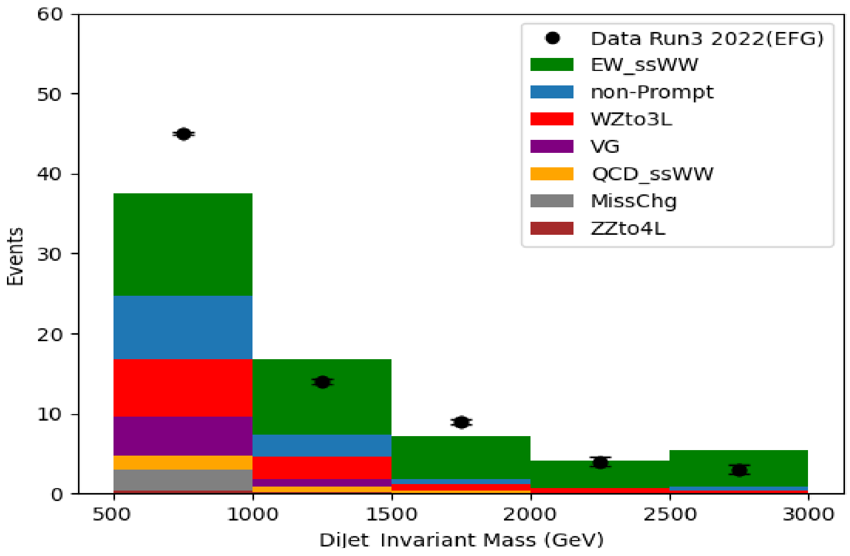
# Statistics for SR

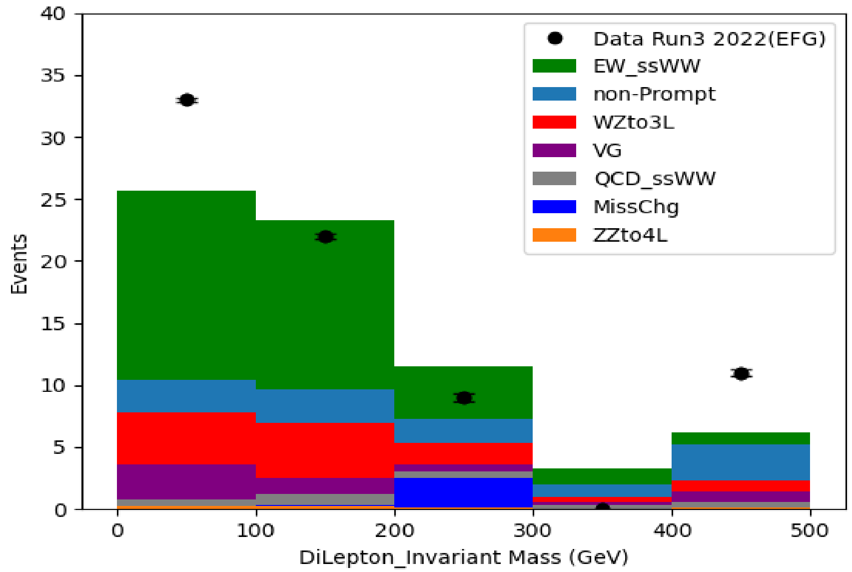
Now that we defined our signal region and its kinematics, and identified backgrounds, now we do calculations. The results are as below:

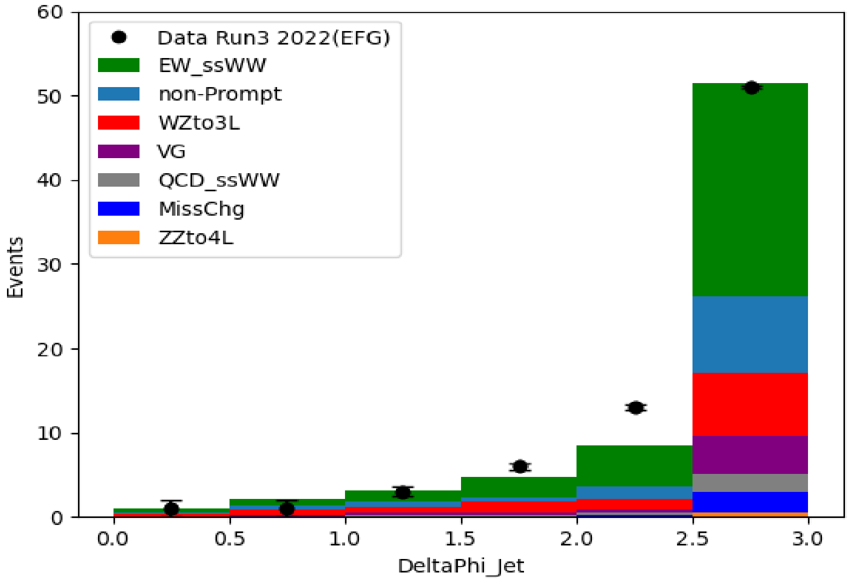
| | Signal Region |
|---|---|
| Data | 75 |
| **Non-Prompt** | **12.3** |
| **EW_ssWW** | **35.4** |
| WZto3L | 11.55 |
| QCD_ssWW | 2.76 |
| VG | 5.8 |
| Miss_Charge | 2.53 |
| ZZto4L | 0.73 |
| Sum_All_BK(including non_prompt) | 35.65 |
| Predicted = Sum(Signal + All BK) | 71.05 |
| **Ratio (Sig/All_BK)** | **0.99** |
| **%Ratio (Sig/Data)** | **%47.2** |
| **%Ratio (Sig/Predicted)** | **%50** |
| **%Ratio (Non-Prompt/Data)** | **%16.4** |
| **%Ratio(Predicted/Data)** | **%94.7** |

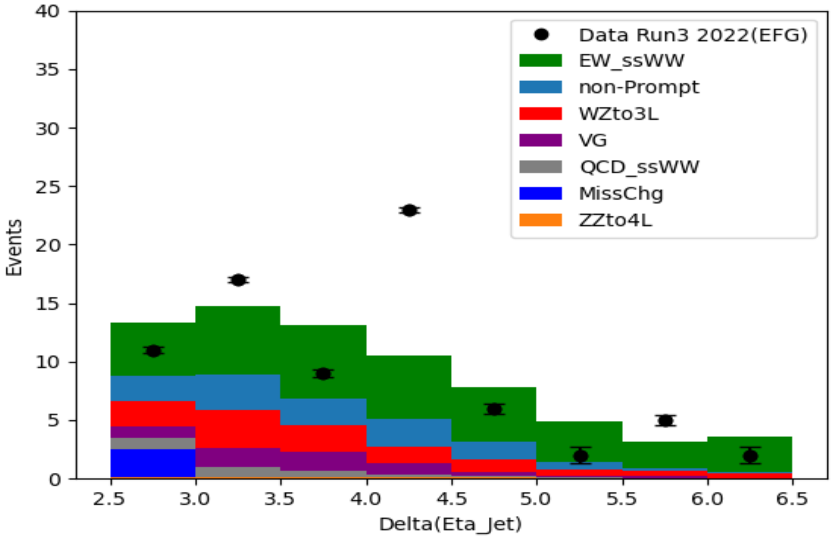
*****The results are for Run 2022 EFG and Summer22EE**
***** In Plots, errors are relative errors. I am working on them and they are developing.**

# Signal Region for leplep



# Control Region

To test the accuracy of our calculations in signal region, we define some control regions (CR). CRs are defined as to be rich for a background process

Two results are expected from a control region:

1- Contribution and number of signal events must decreases

2- Contribution and number of the specific background events (which CR is defined for that) must increased

We define two control regions:

1- Non-Prompt Control Region: Which is expected to enhance the contribution of non-prompt background

2- QCD Control Region: Which is expected to enhance the contribution of QCD same-sign WW background

## Definition of Control Region

| | non-prompt CR | QCD CR |
|---|---|---|
| Pt_Leptons | Pt1 > 25 GeV, Pt2 > 20 GeV | Pt1 > 25 GeV, Pt2 > 20 GeV |
| Pt_Jets | >50 GeV | >50 GeV |
| Pt_Miss | >30 GeV | >30 GeV |
| Anti Btagging | inverted | applied |
| Abs(Delta_Eta(Jet1, Jet2)) | >2.5 | <2.5 |
| M_jj | >500 GeV | >200 GeV<br>< 500 GeV |

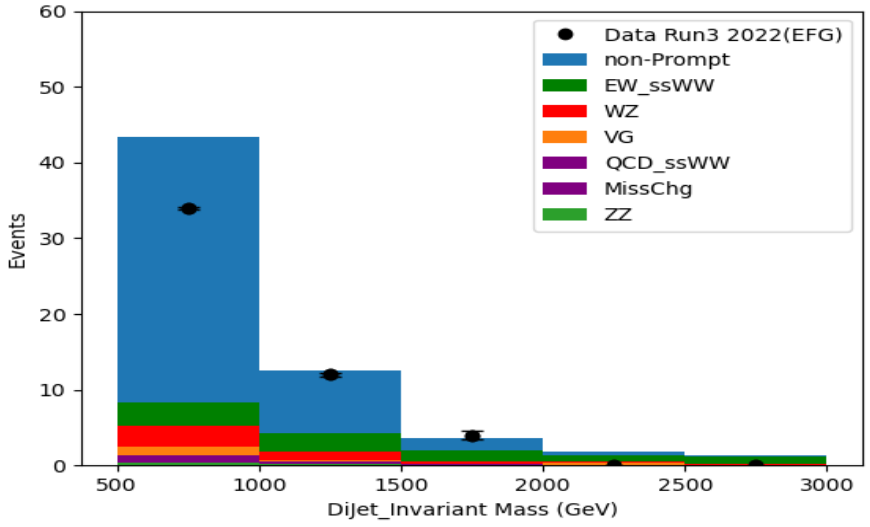
## Statistics for Control Regions

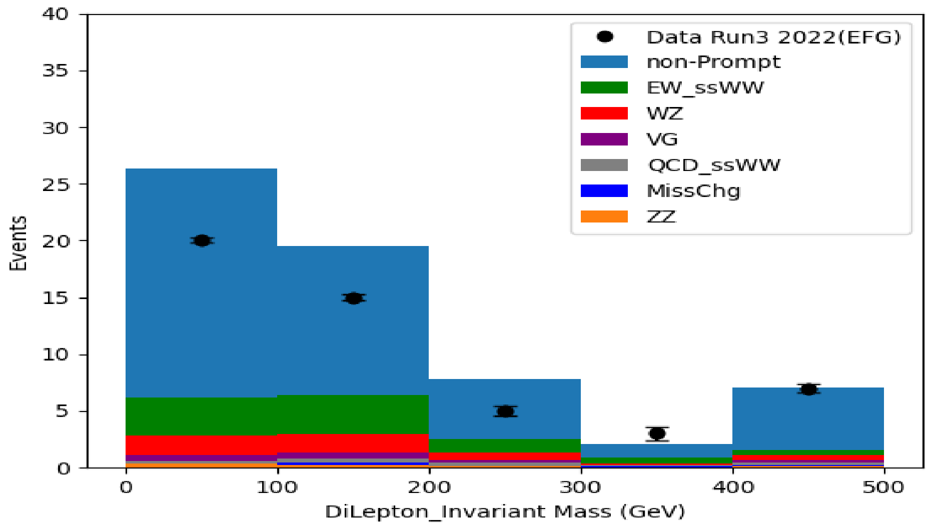
| | SR | Non-PR CR | QCD CR |
|---|---|---|---|
| Data | 75 | 50 | 83 |
| **Non-Prompt** | **12.3** | **45.3** ↑ | **20.15** ↑ |
| **EW_ssWW** | **35.4** | **8.92** ↓ | **6.96** ↓ |
| QCD_ssWW | 2.76 | 1.18 | 9.7 ↑ |

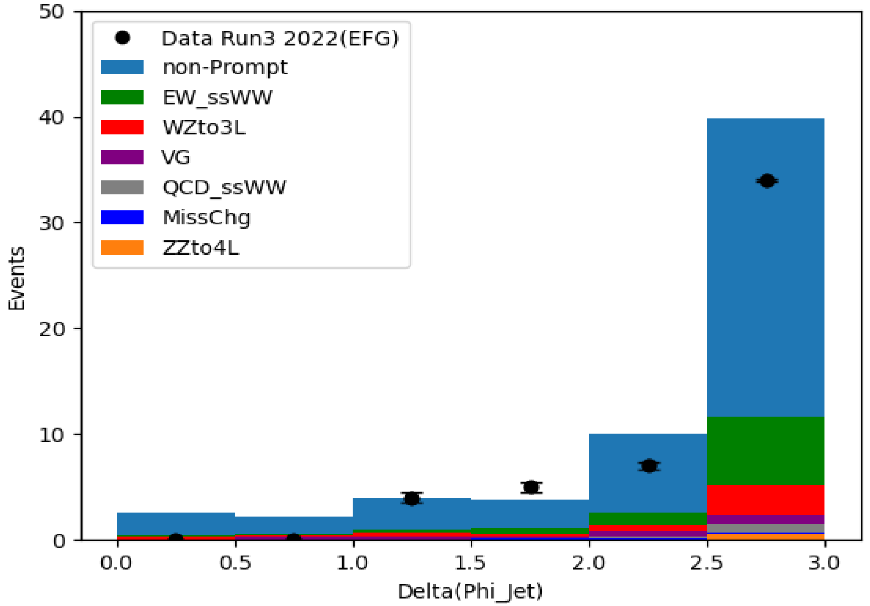
## The percentage of signal and each background contribution from the predicted calculation by MC samples

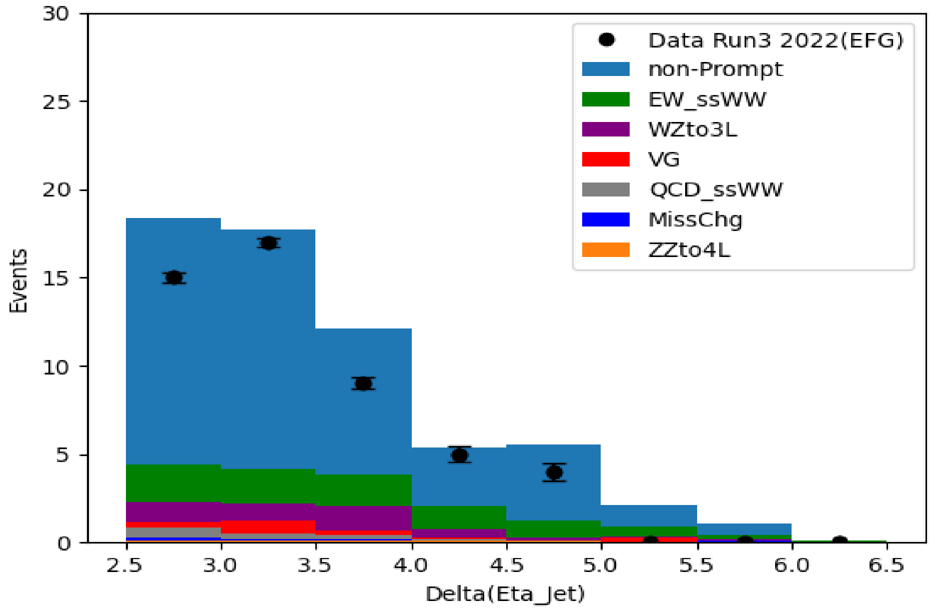
| | SR | Non-PR CR | QCD CR |
|---|---|---|---|
| Predicted Events | 71.05 | 62.46 | 76.53 |
| **EW_ssWW** | **%49.8** | **%14.3** ↓ | **%9.09** ↓ |
| **Non-Prompt** | **%17.3** | **%72.5** ↑ | **%26.32** |
| WZto3L | %16.26 | %7.2 | %30.22 |
| VG | %8.16 | %2.49 | %18.9 |
| QCD_ssWW | %3.9 | %1.18 | %12.67 ↑ |
| Miss_Charge | %3.6 | %1.9 | %2.4 |
| ZZto4L | %1 | %1.1 | %0.3 |

# Non-Prompt Control Region for leplep

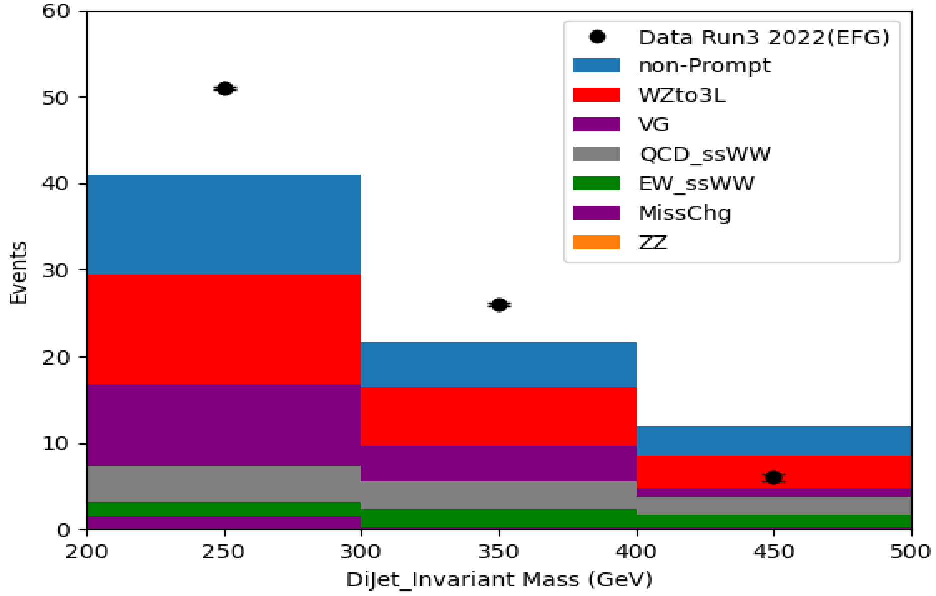


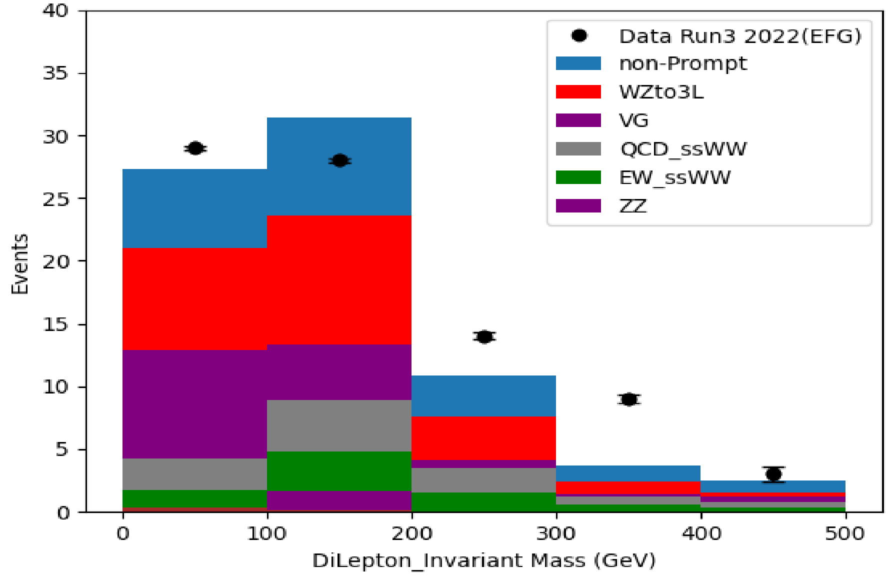



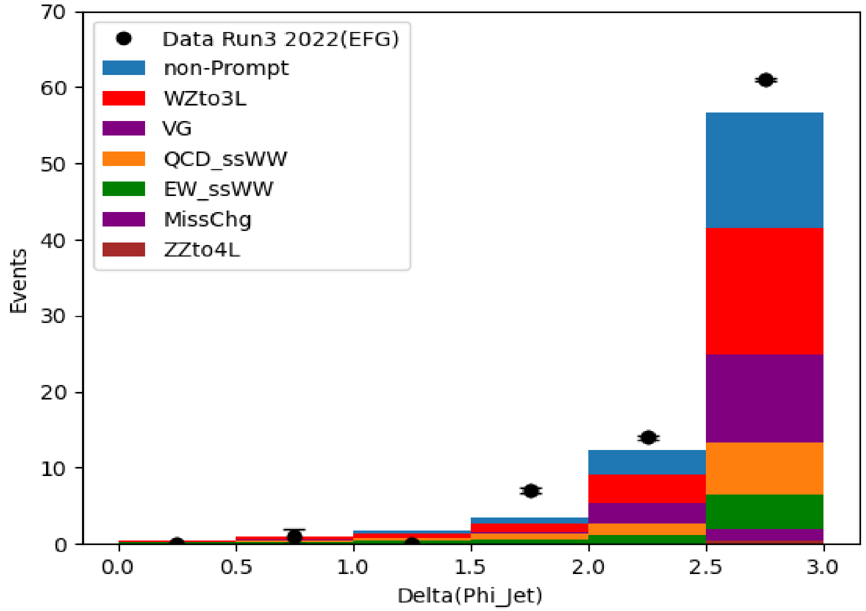



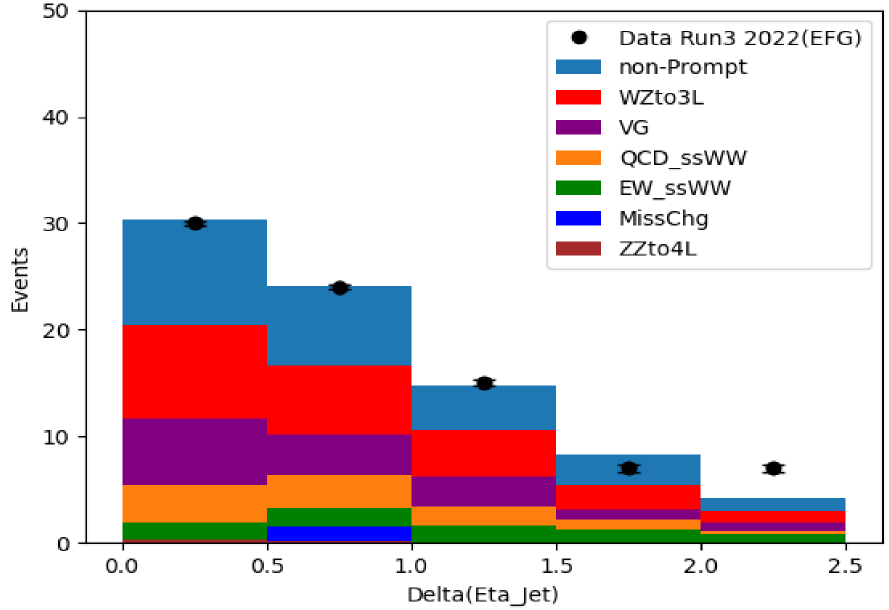



# QCD Control Region for leplep



# Conclusion and Future Plans

Conclusions:

By checking the measurements in signal and control regions, we could understand:

1- Our definitions, cuts, and calculation of signal and control regions are well performed

2- Very good enhancement in contribution of signal events w.r.t other backgrounds (we could increase signal and decrease background contribution as much as possible)

Future Plans:

1- Read newer data and MC samples for year 2023 and 2024

2- Measure the cross section of same-sign WW -> 2l + 2nu

3- Put some upper limits to the Lagrangian of process using Wilson's coefficient

*** The systematic uncertainties has been done but because integrating all results by them takes long time, so at the momentof the presentation, they were not ready