The search for invisibly decaying Higgs bosons at the CMS experiment

Vukasin Milošević, on behalf of the VBF H->Inv. team

Imperial College London, 10.04.2019.

Overview

1

Analysis motivation

- VBF specific introduction

2

Analysis methods

- Background control
- Selection requirements

3

Creation of analysis specific trigger

- Structure and performance

4

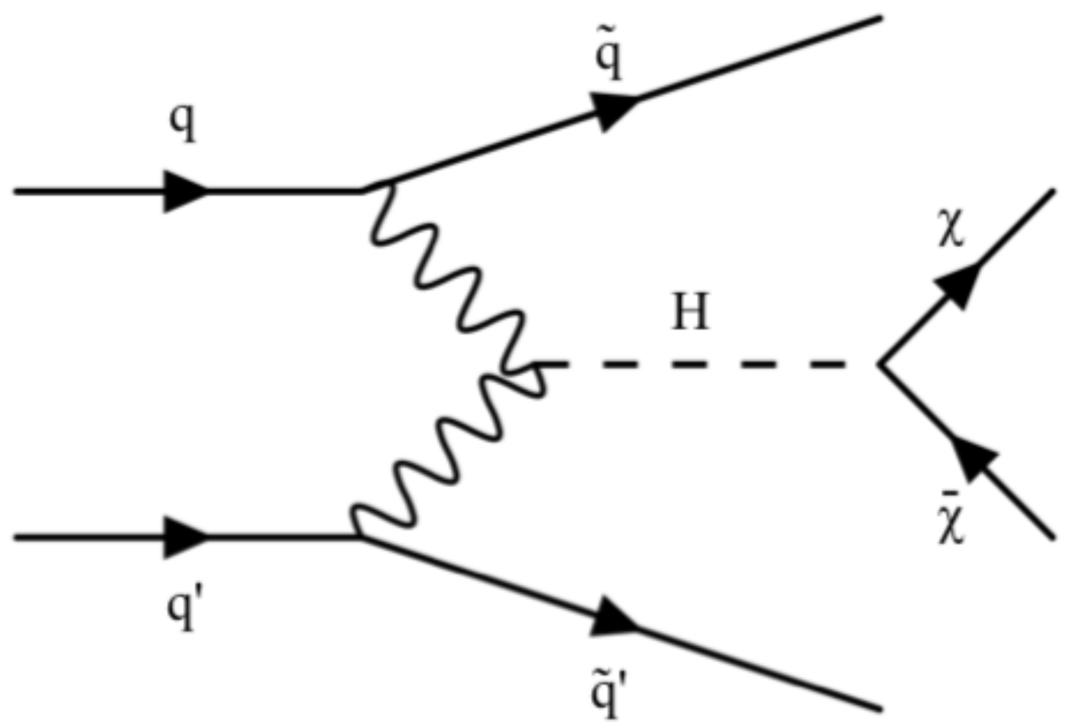
The road towards the first limit

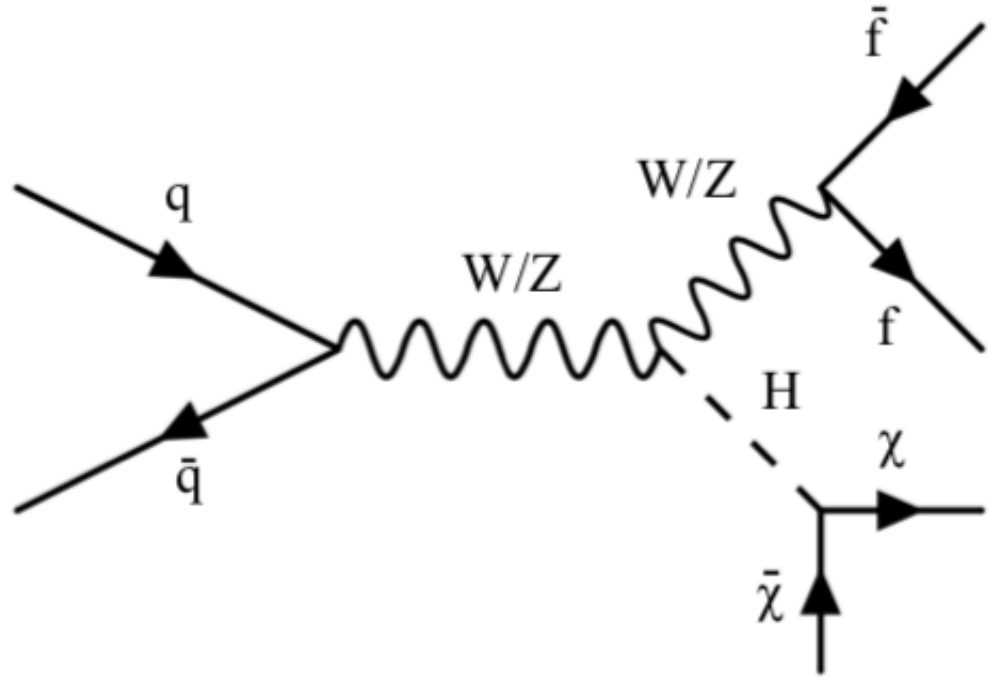
- First data/MC distributions for 2017 data

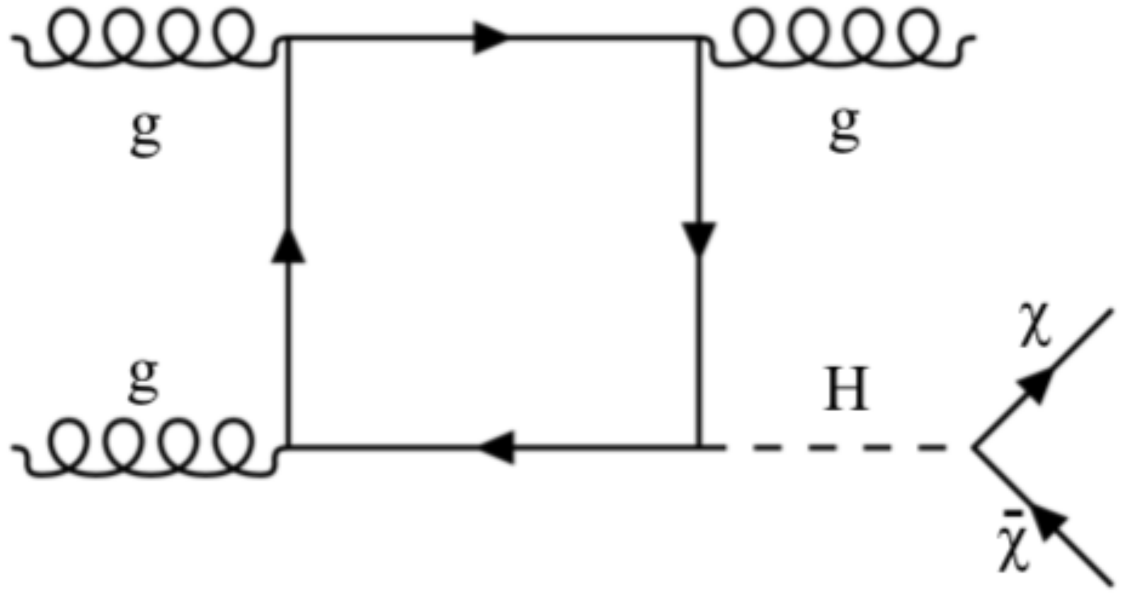
Conclusion

Introduction: Analysis motivation

- **Invisible decays of the Higgs boson**, as part of the “Higgs portal model” scenarios, are a good way of searching for new physics
  - **Higgs boson** a **mediator** between SM particles and ones that belong to the DM sector
- **Detection requires for the Higgs boson to recoil against a visible system:**
  - **qqH :** Higgs boson is produced in a vector boson fusion topology (VBF)
  - **VH:** Higgs boson production with a vector boson (V= Z or W)
  - **ggH:** Higgs boson produced via gluon fusion.

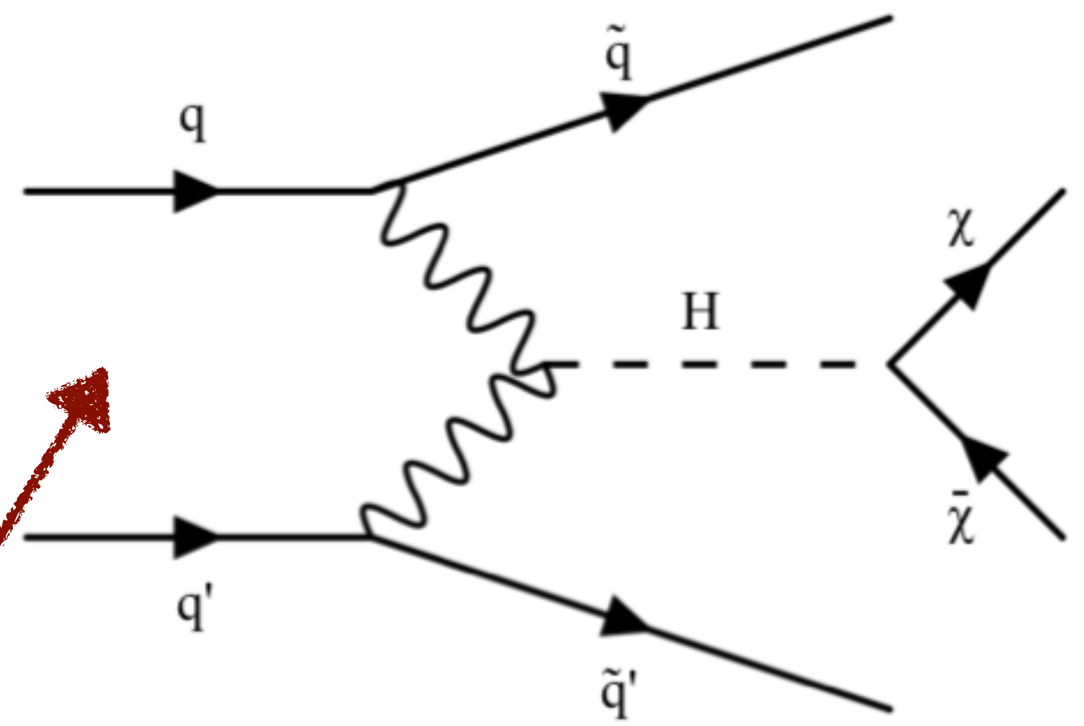


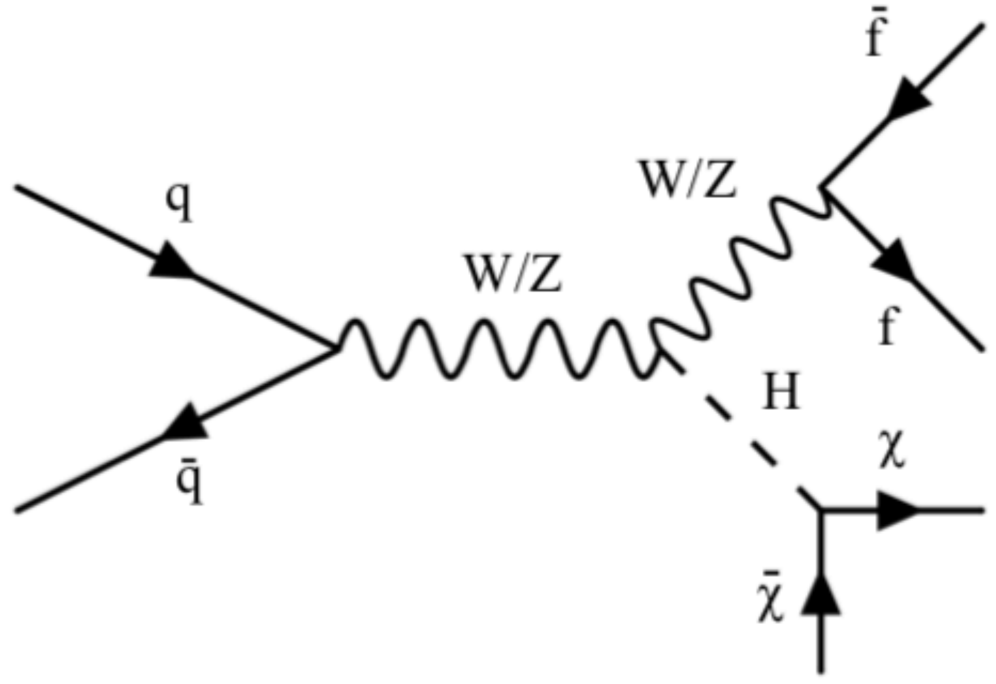


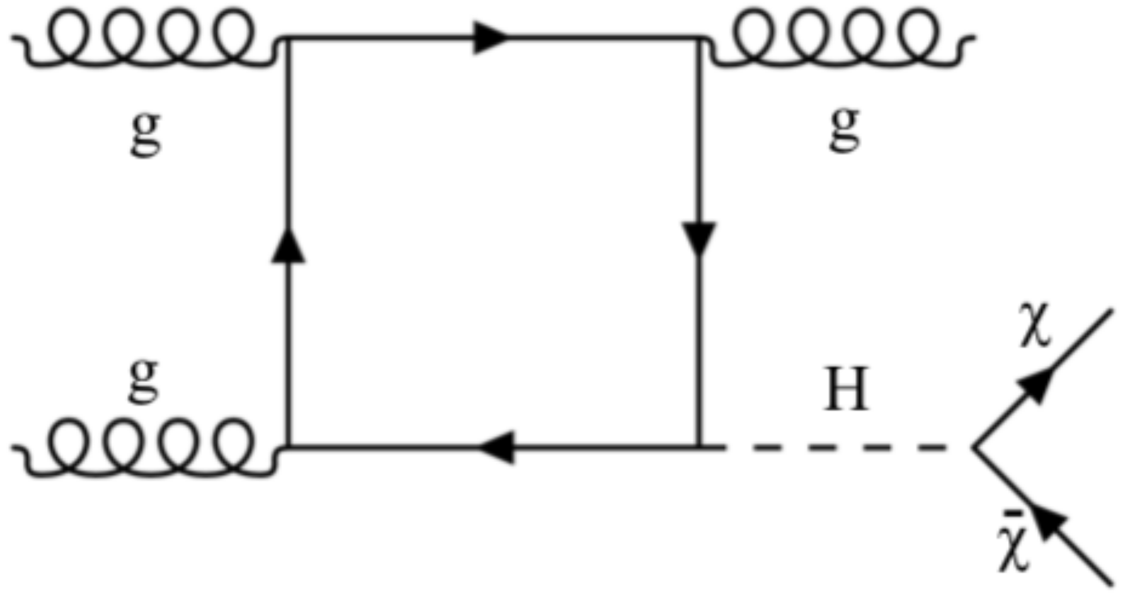


Introduction: Analysis motivation

- **Invisible decays of the Higgs boson**, as part of the “Higgs portal model” scenarios, are a good way of searching for new physics
  - **Higgs boson** a **mediator** between SM particles and ones that belong to the DM sector
- **Detection requires for the Higgs boson to recoil against a visible system:**
  - **qqH :** Higgs boson is produced in a vector boson fusion topology (VBF)
  - **VH:** Higgs boson production with a vector boson (V= Z or W)
  - **ggH:** Higgs boson produced via gluon fusion.

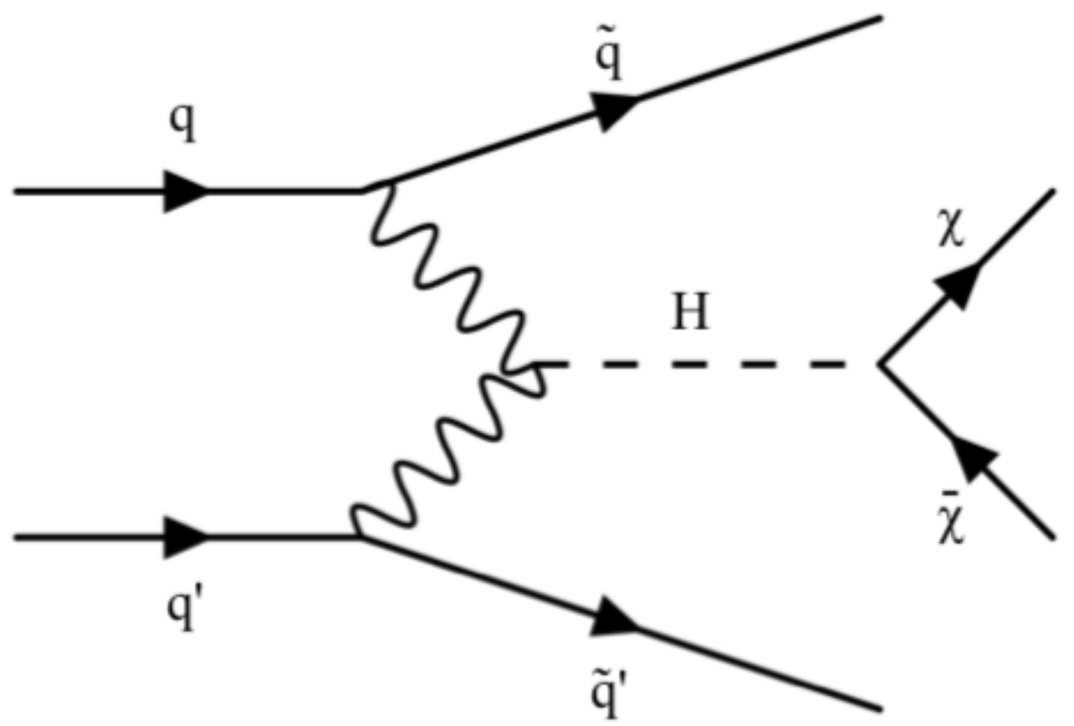


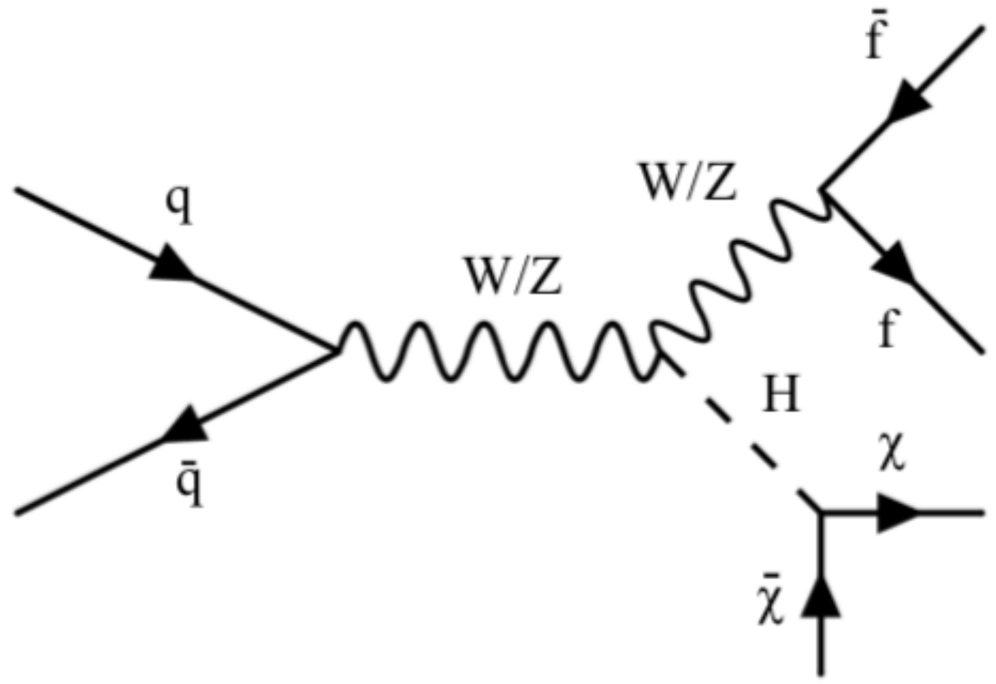


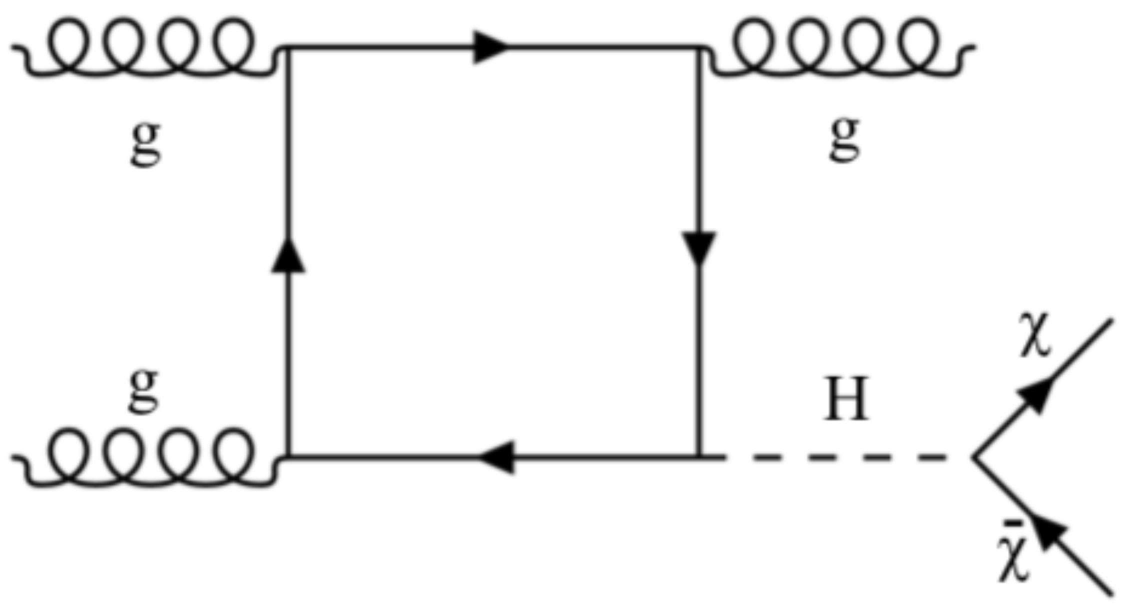


Introduction: Analysis motivation

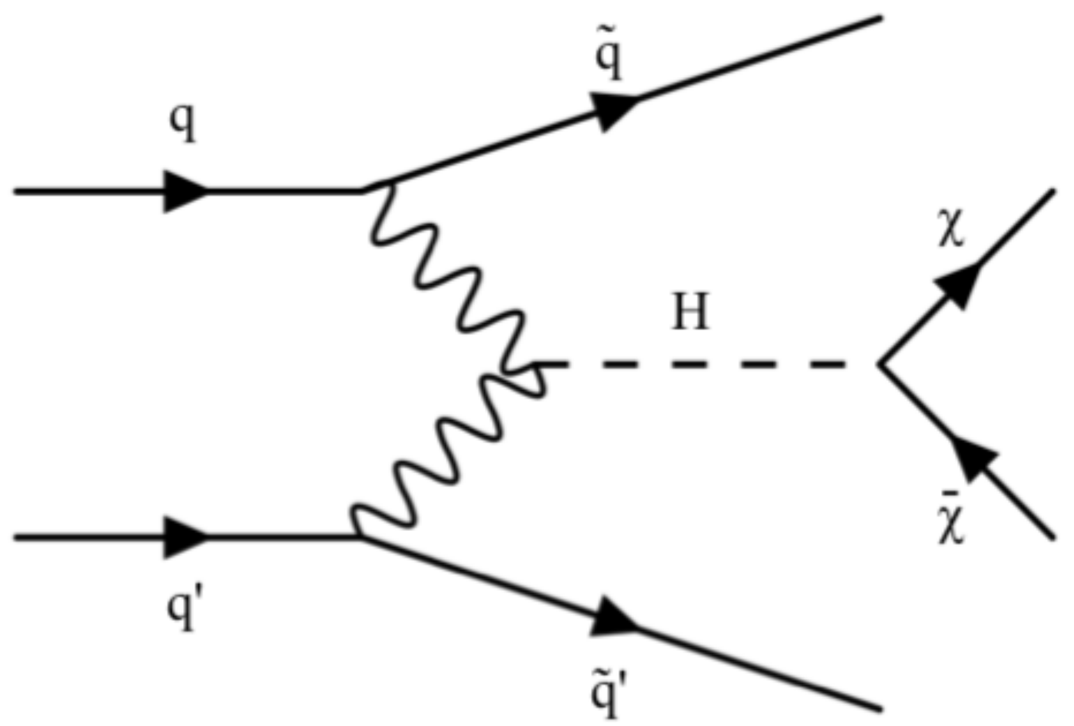
- **Invisible decays of the Higgs boson**, as part of the “Higgs portal model” scenarios, are a good way of searching for new physics
  - **Higgs boson** a **mediator** between SM particles and ones that belong to the DM sector
- **Detection requires for the Higgs boson to recoil against a visible system:**
  - **qqH :** Higgs boson is produced in a vector boson fusion topology (VBF)
  - **VH:** Higgs boson production with a vector boson (V= Z or W)
  - **ggH:** Higgs boson produced via gluon fusion.



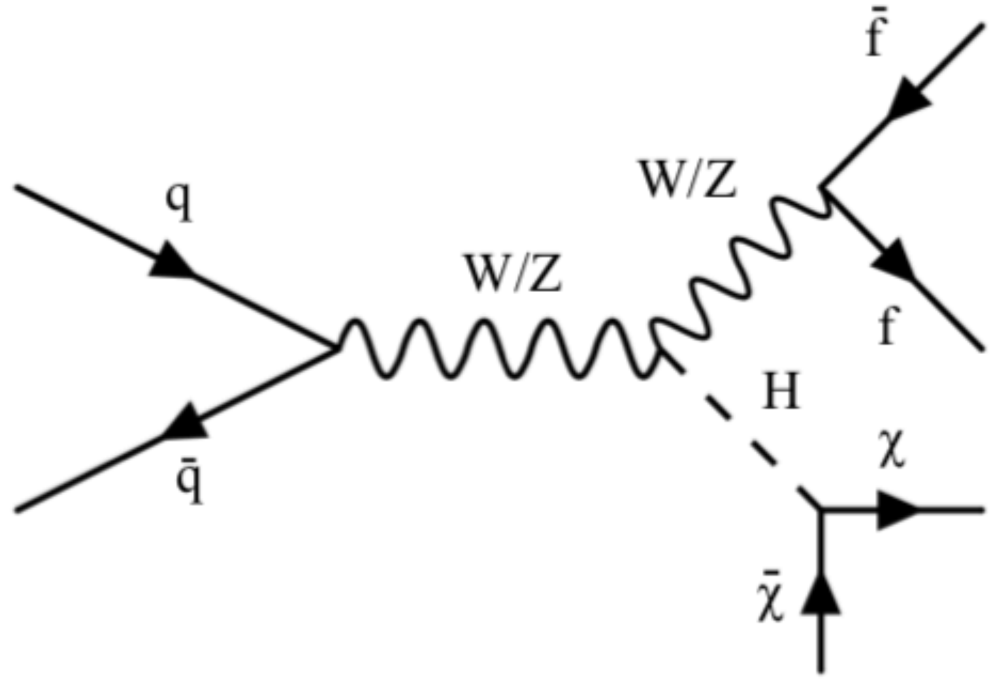




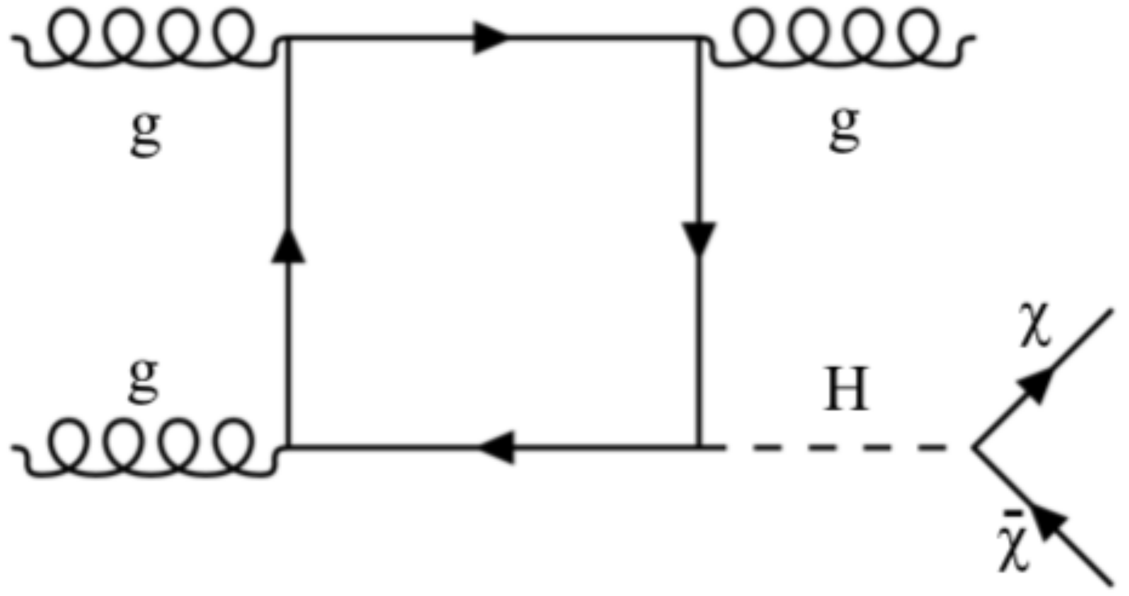
Introduction: Analysis motivation



- **Invisible decays of the Higgs boson**, as part of the "Higgs portal model" scenarios, are a good way of searching for new physics
  - **Higgs boson** a **mediator** between SM particles and ones that belong to the DM sector

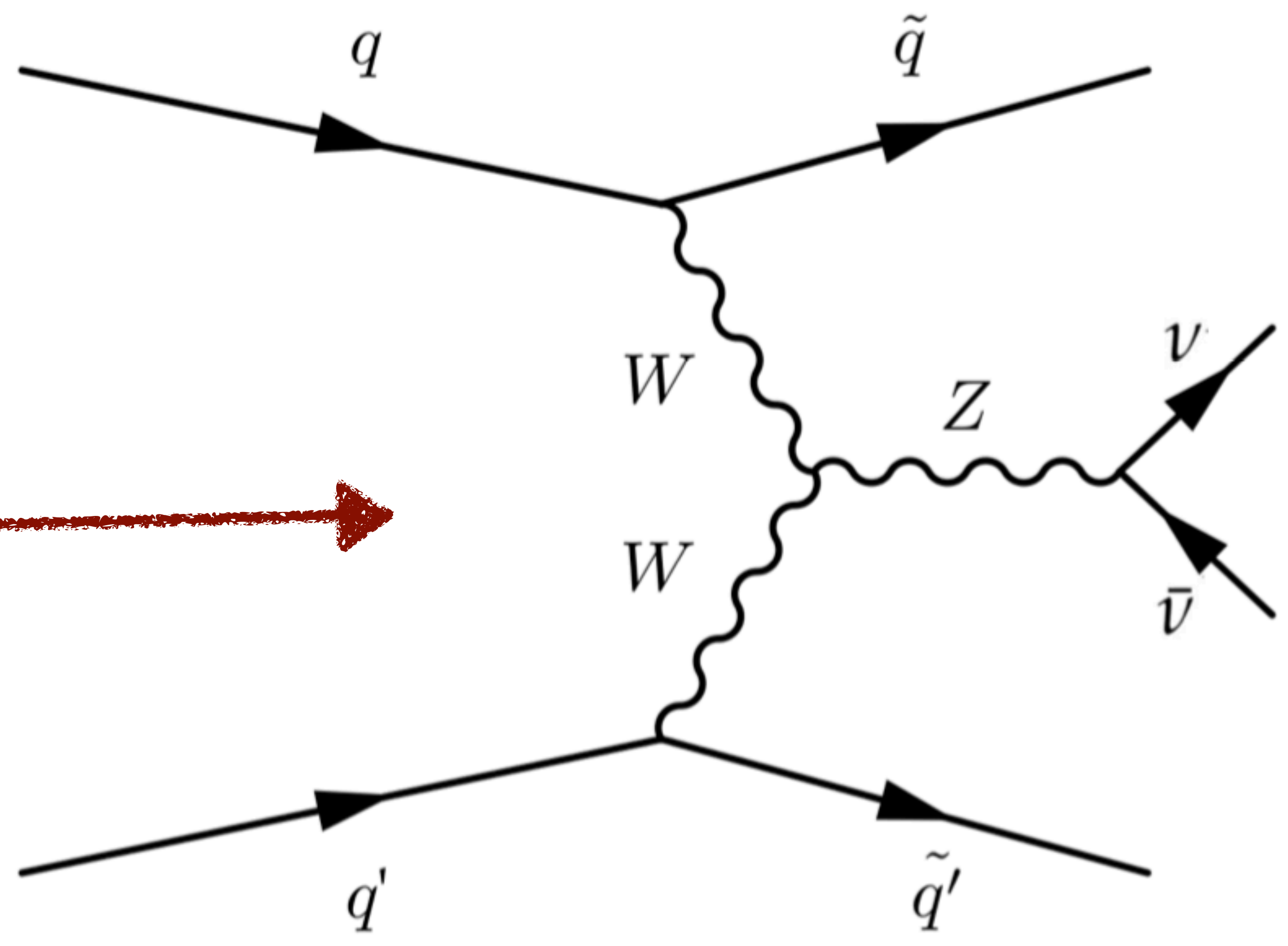


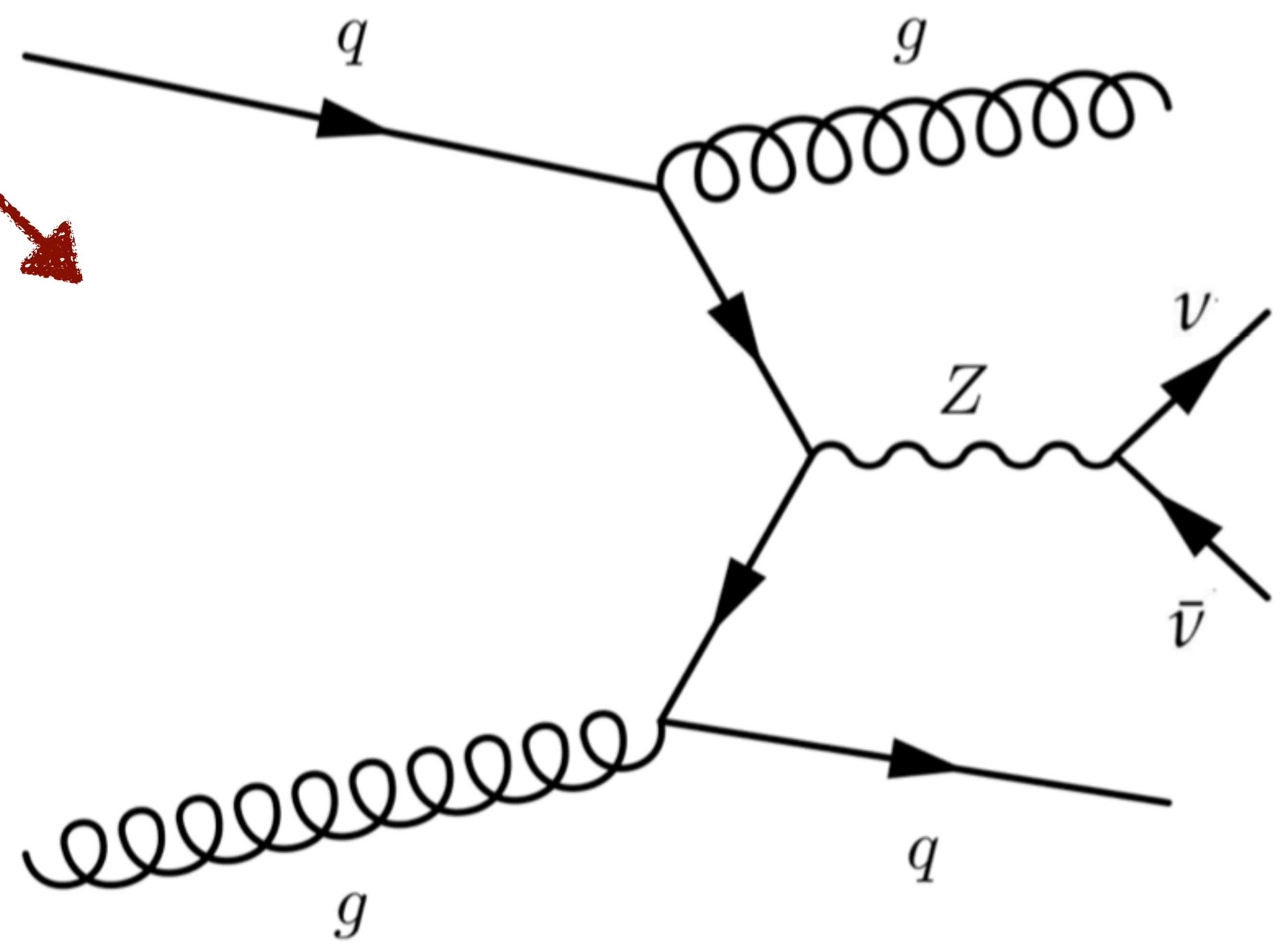
- **Detection requires for the Higgs boson to recoil against a visible system:**
  - **qqH :** Higgs boson is produced in a vector boson fusion topology (VBF)
  - **VH:** Higgs boson production with a vector boson (V= Z or W)
  - **ggH:** Higgs boson produced via gluon fusion.



Background control

- The SM backgrounds:
  - ***V+jets:*** *Z*(*νν*)+jets and *W*(*lν*)+jets
    - where the charged lepton is unidentified
  - originating from **QCD multijet** production processes.
  - **Diboson** and **top quark** processes
    - estimated from simulation.





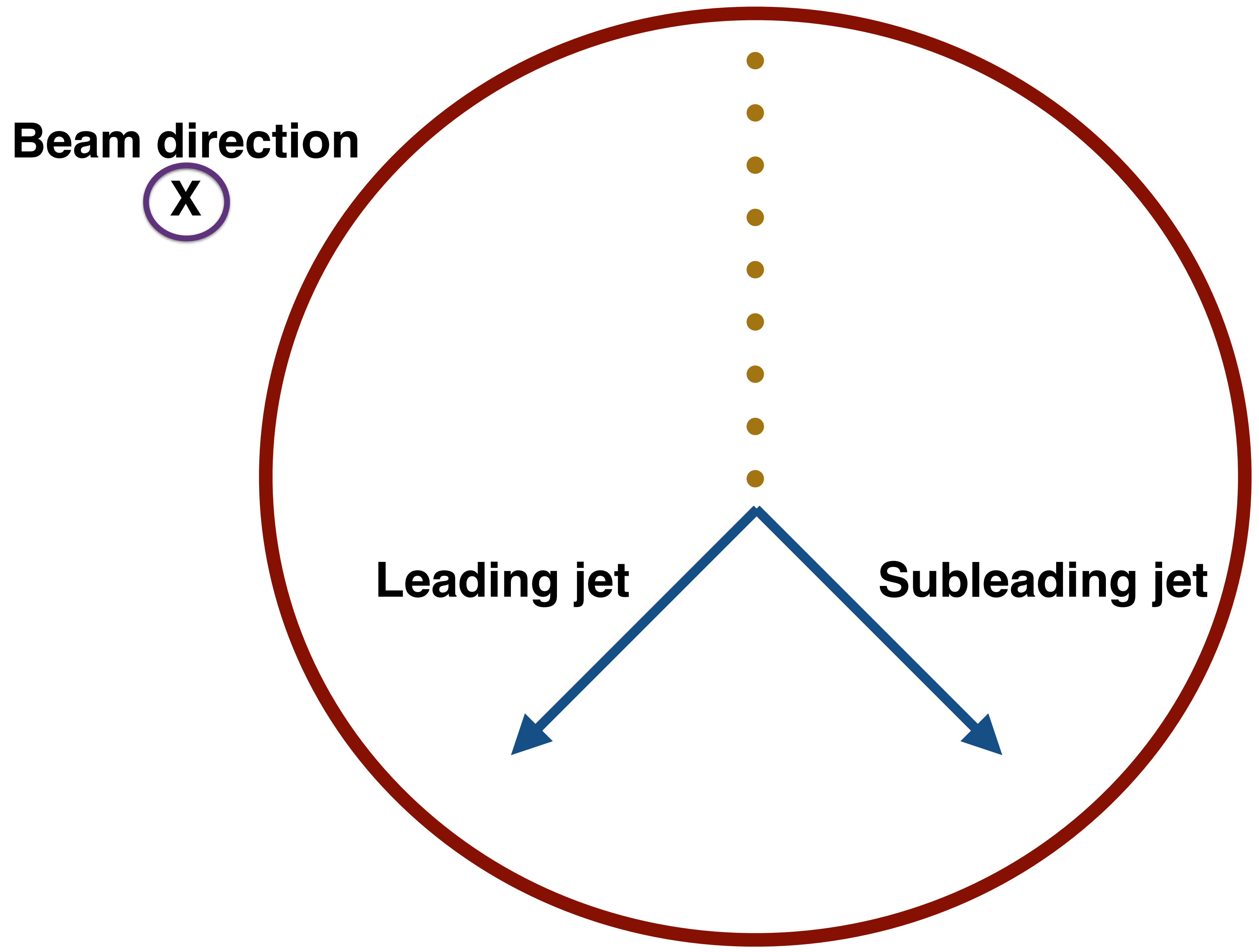
- V+Jets: Dedicated control regions in data
  - *Z* or *W* boson associated with the same dijet topology
- This means that we can have the following scenarios:
  - $Z \to e^+e^-$, $Z \to \mu^+\mu^-$, $W \to e\nu$, $W \to \mu\nu$

Selection requirements

- Missing transversal energy (MET):

$$E_{T,miss} = -|\sum_i \vec{p}_{T,i}|$$

  - Existence of MET can imply presence of "invisible" objects



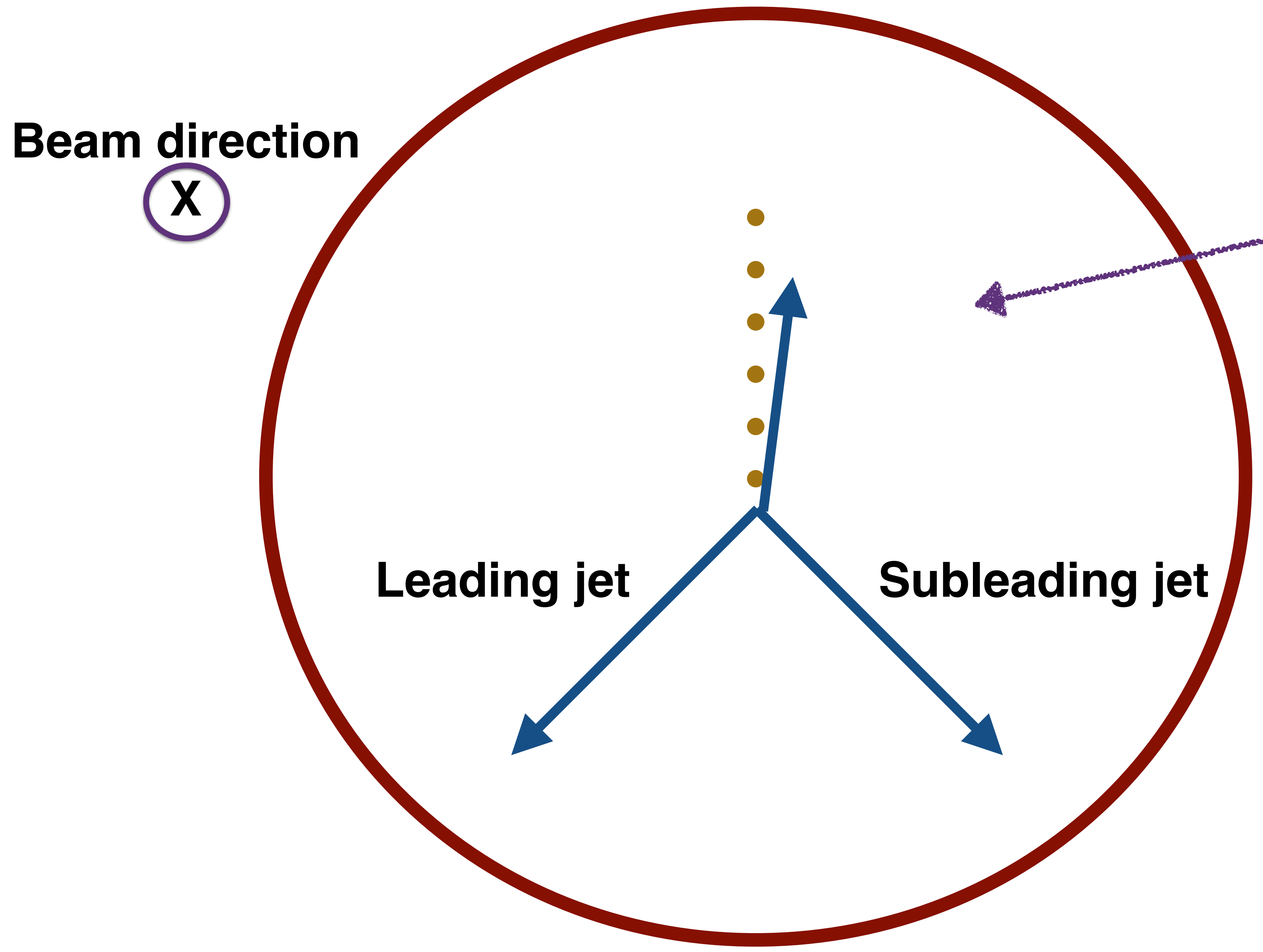
Variable	$\sqrt{s} = 13$ TeV
$p_T^{j_1}$	> 80 GeV
$p_T^{j_2}$	> 40 GeV
E_T^{miss}	> 250 GeV
$min\Delta\phi(\vec{p}_T^{miss}, j)$	> 0.5
$\Delta\eta(j_1, j_2)$	> 4.0
$\Delta\phi(j_1, j_2)$	< 1.5
m_{jj}	> 1300 GeV

Selection requirements

- Missing transversal energy (MET):

$$E_{T,miss} = -|\sum_i \vec{p}_{T,i}|$$

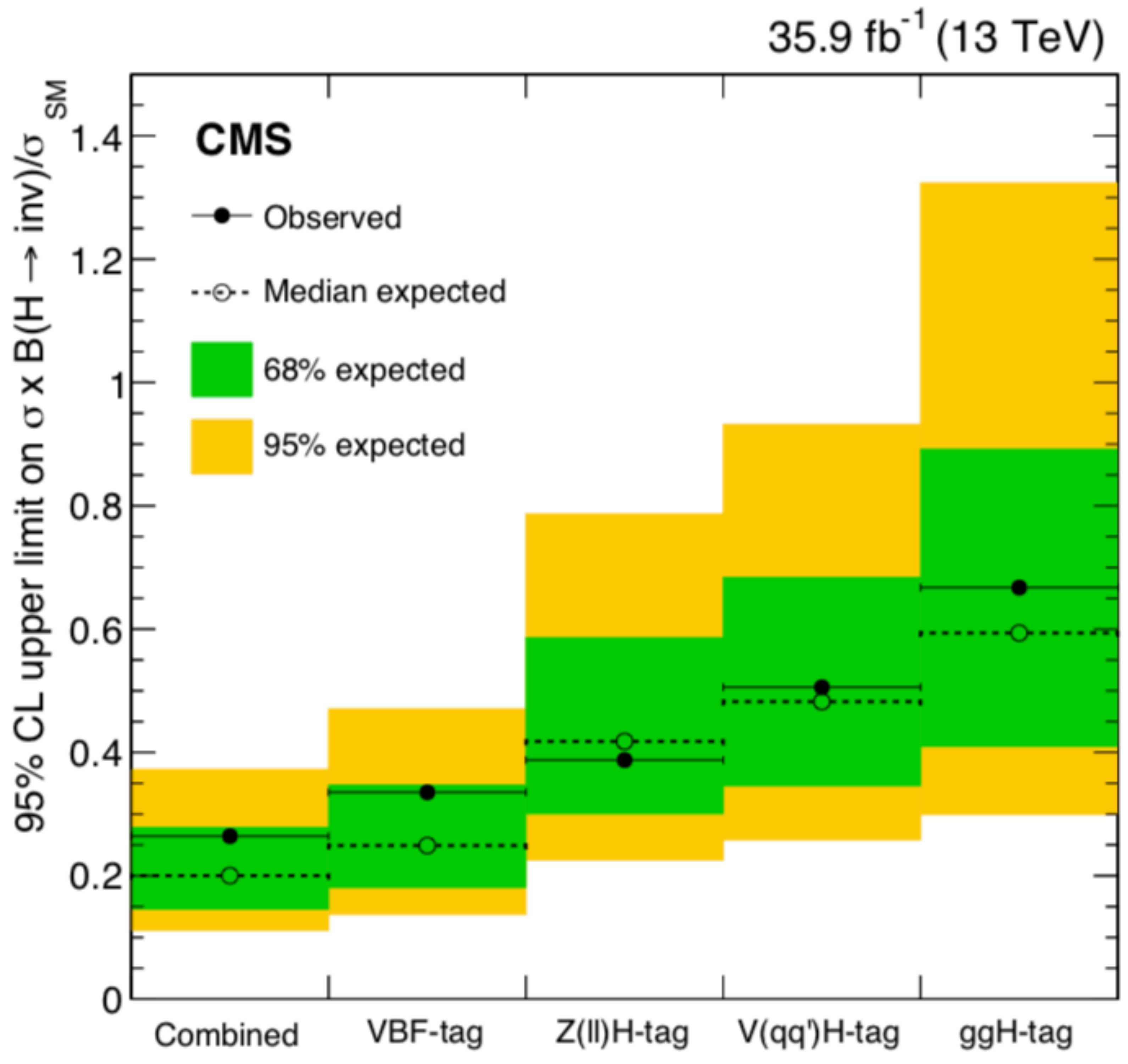
  - Existence of MET can imply presence of “invisible” objects



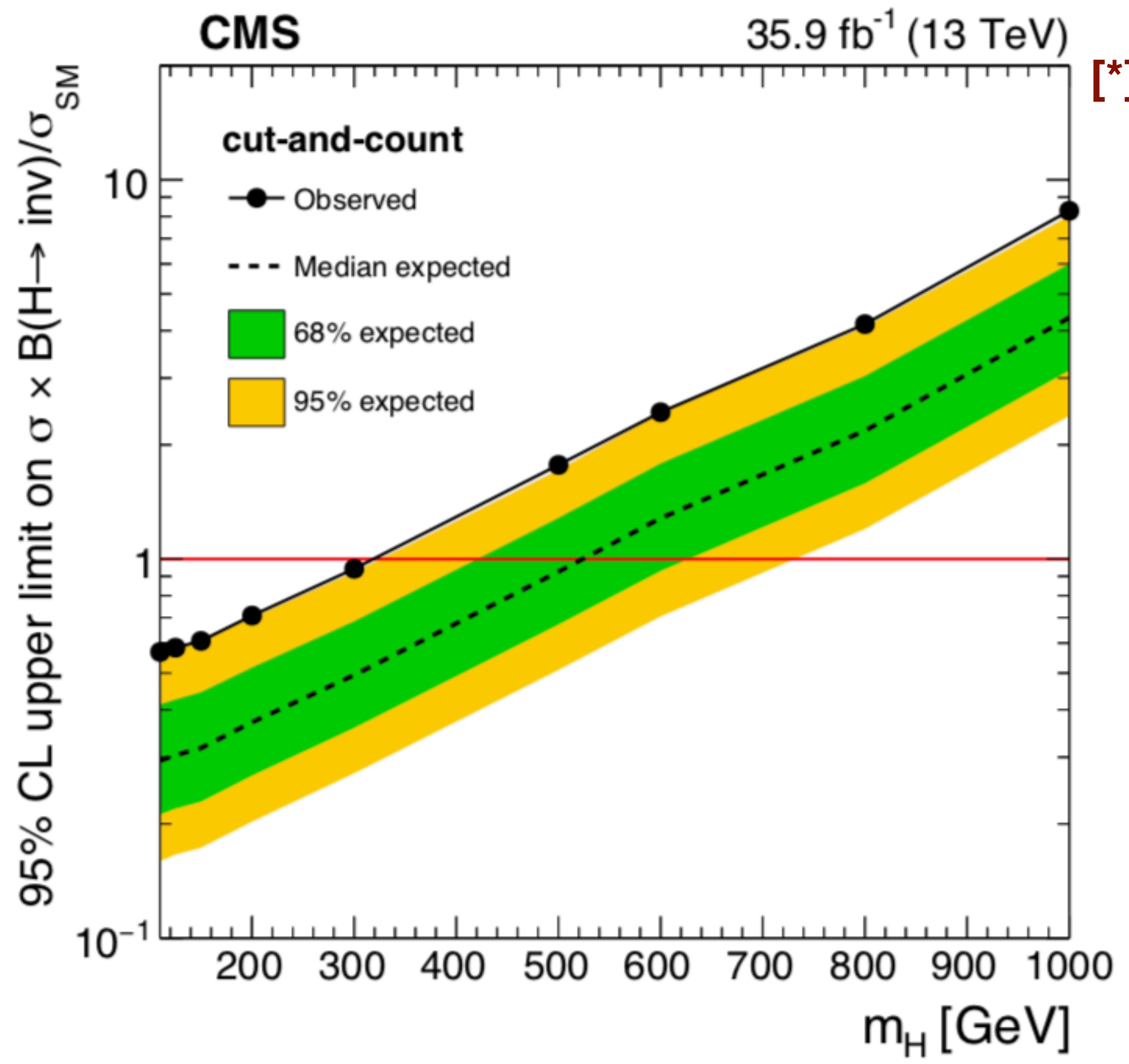
Variable	$\sqrt{s} = 13$ TeV
$p_T^{j_1}$	> 80 GeV
$p_T^{j_2}$	> 40 GeV
E_T^{miss}	> 250 GeV
$min\Delta\phi(\vec{p}_T^{miss}, j)$	> 0.5
$\Delta\eta(j_1, j_2)$	> 4.0
$\Delta\phi(j_1, j_2)$	< 1.5
m_{jj}	> 1300 GeV

Introduction: Current results

- **The latest analysis, containing data collected during 2016, has been submitted to "Physics Letters B" [*]**



[*]



[*]

[*] link to the paper on arxiv

High-Level Trigger: Introduction

- Up until now relied solely on generic SUSY trigger based around missing energy
- Upgraded CMS Level-1 (L1) trigger system:
  - Detailed probing of topologies of interest
  - Allows separation of events by:
    - Jet transverse momenta (pt)
    - Dijet mass (mjj)
- Building upon the L1 strategy, new VBF based trigger:
  - Loosens the missing energy requirement
  - Covers additional population of events

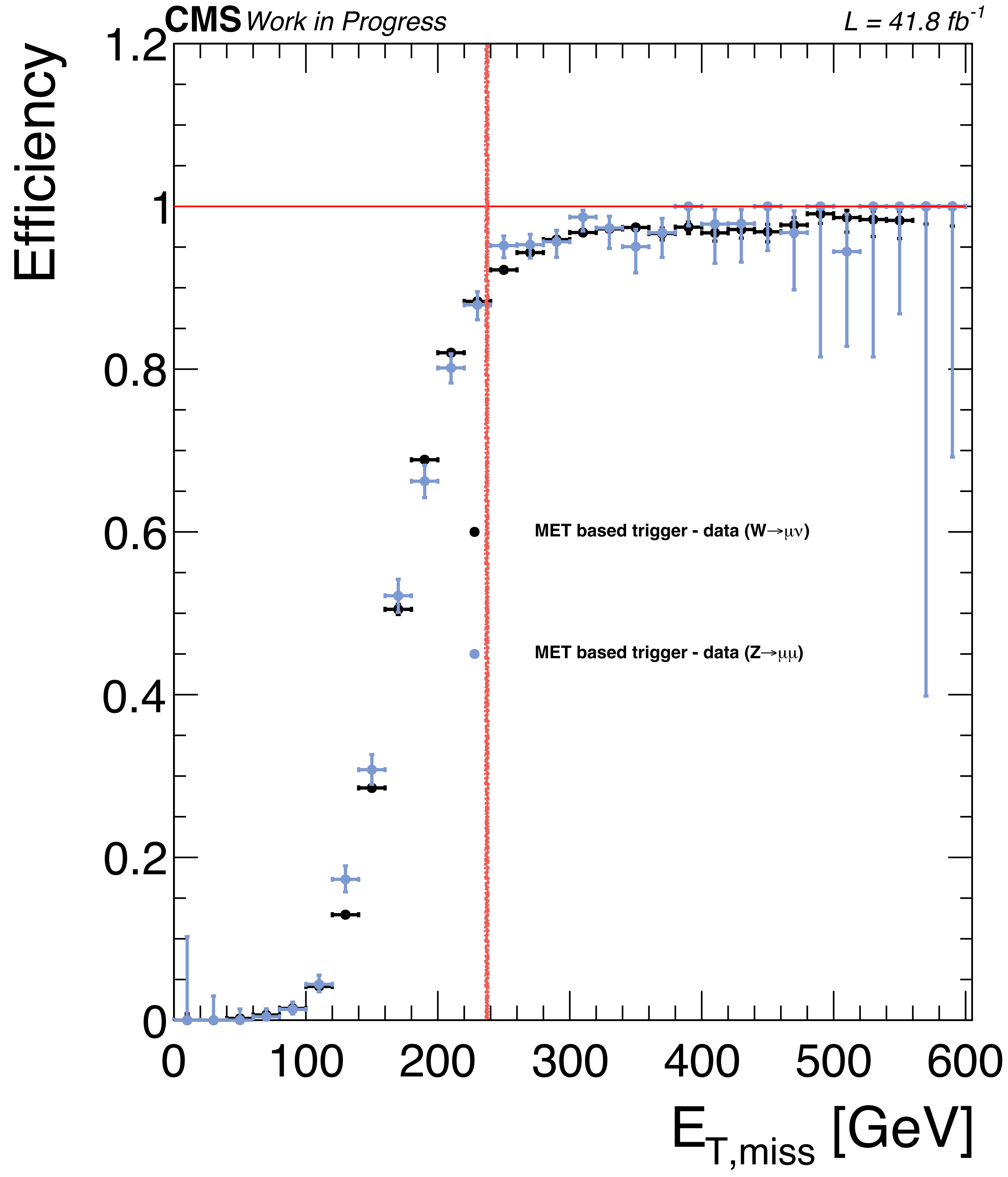
L1 requirement
$P_{T_{j,1}} > 110\ GeV$
$P_{T_{j,2}} > 35\ GeV$
$m_{jj} > 650\ GeV$

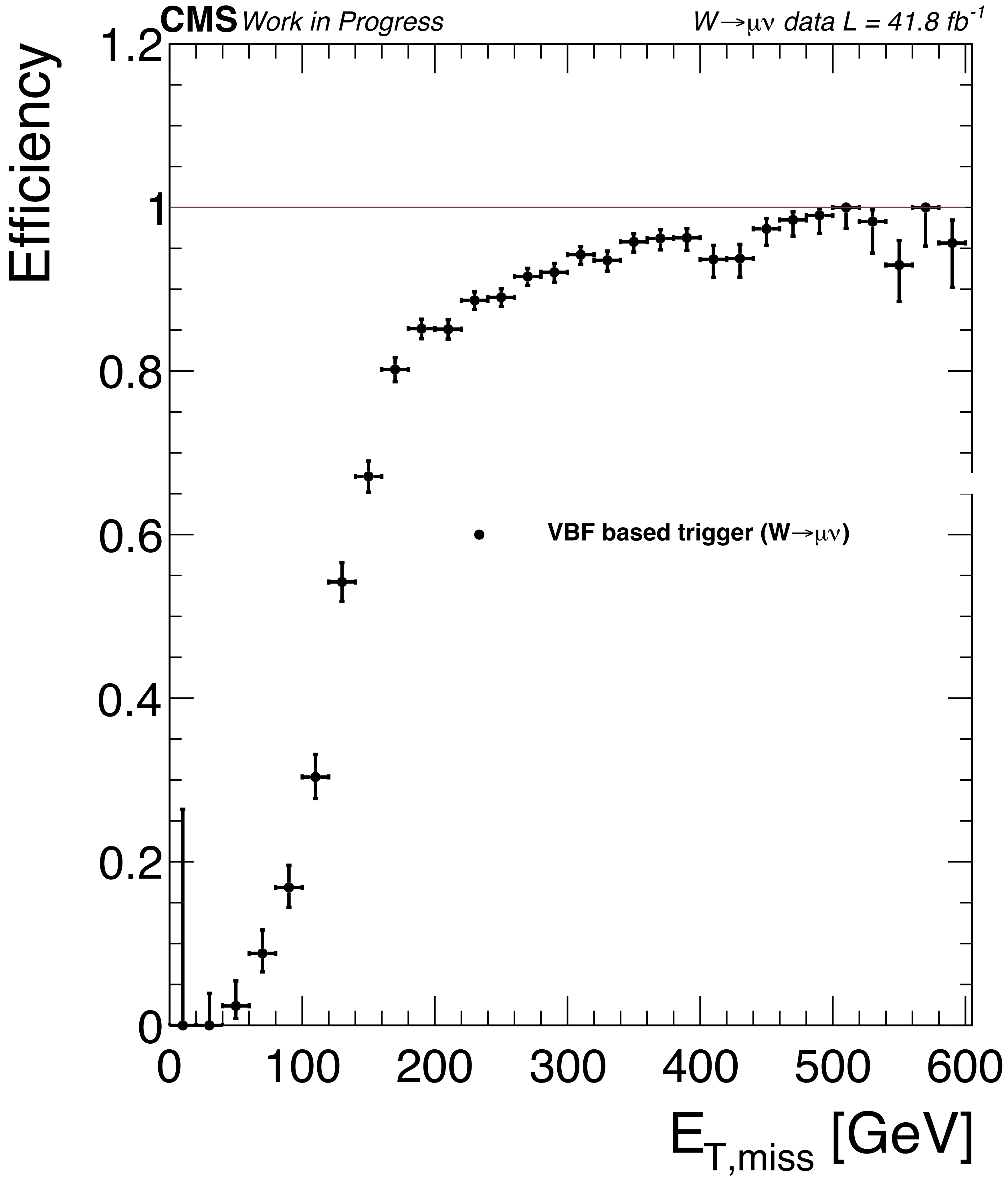
For signal efficiency and removal of QCD

HLT requirement
$E_{T,miss_{calo}} > 66\ GeV$
$E_{T,miss_{PF}} > 110\ GeV$
$P_{T_{j,1}} > 110\ GeV$
$P_{T_{j,2}} > 35\ GeV$
$m_{jj} > 650\ GeV$

Control of rates and timing

High-Level Trigger: Performance





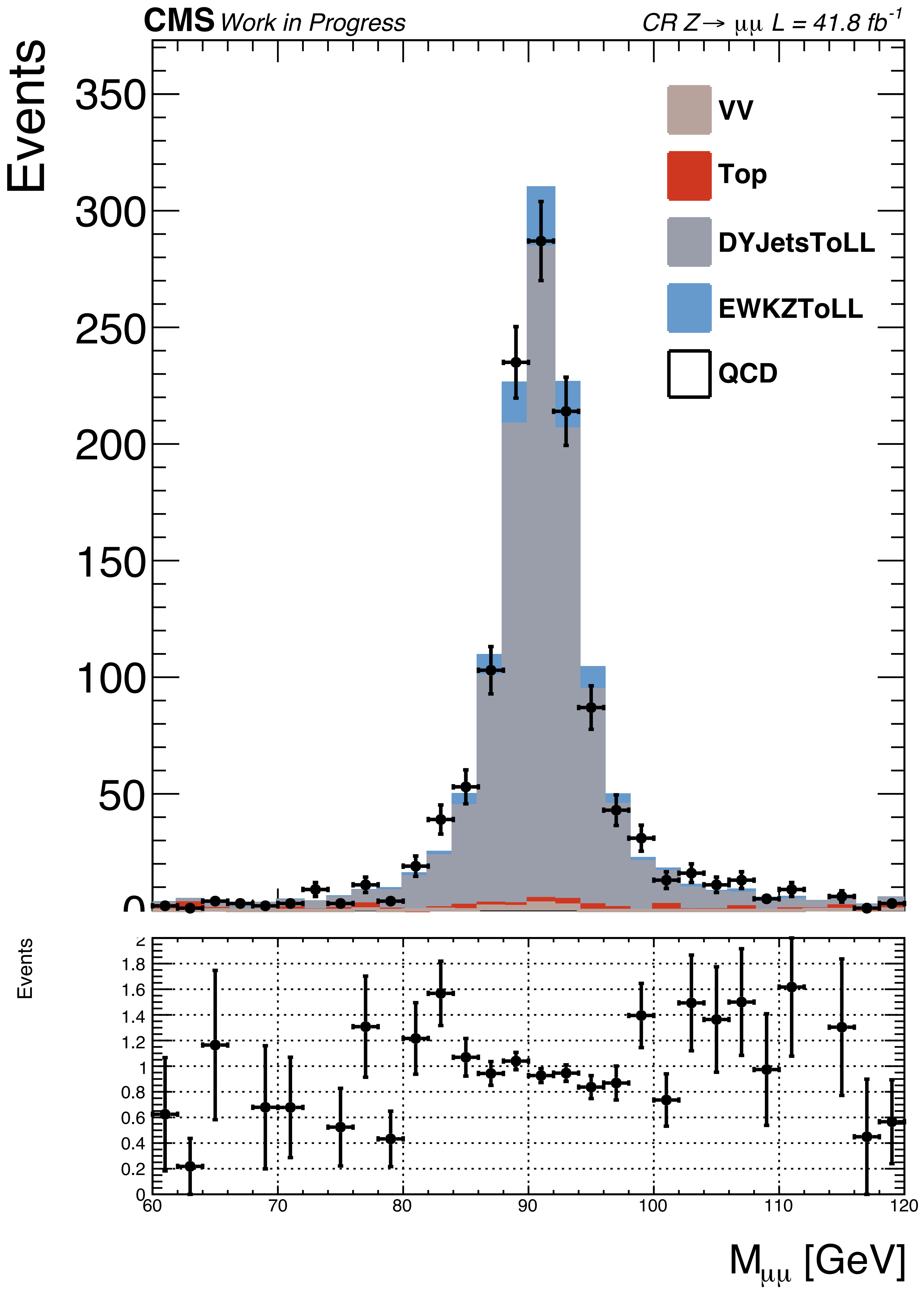
- Better performance of standard MET triggers
- Lower MET requirement

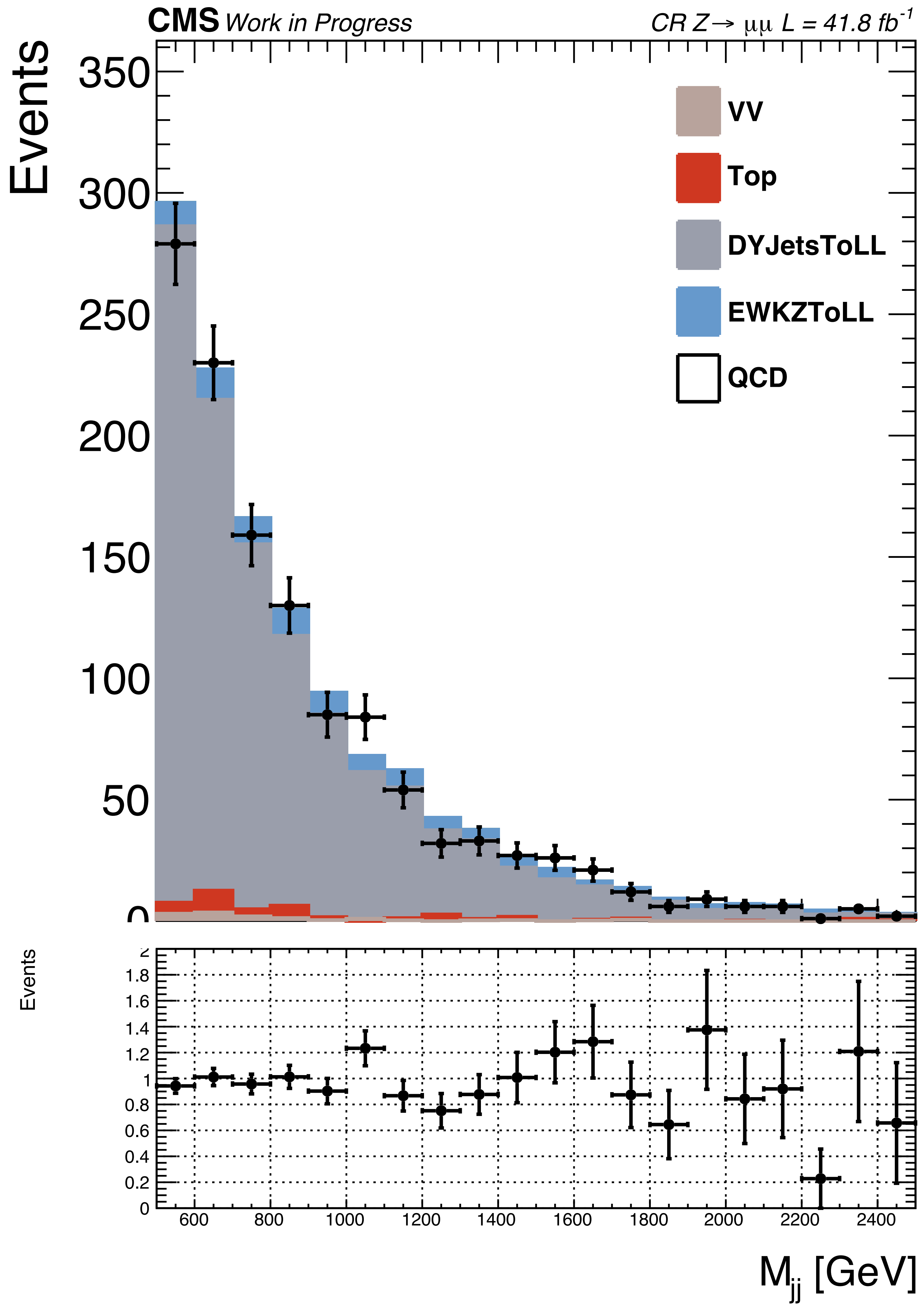
- Preliminary gain in signal acceptance ~10%
- Allows us to probe a new part of the phase space

Increasing the sensitivity: A first look at 2017 data

- **Zμμ control region:**
  - **Preliminary results**
  - **Full 2017 dataset: 41.8 fb-1**

Variable	Requirement
$p_T^{j_1}$	> 80 GeV
$p_T^{j_1}$	> 40 GeV
E_T^{miss}	> 250 GeV
$min\Delta\phi(E_T^{miss},\ j)$	> 0.5
$\Delta\eta(j_1,\ j_2)$	> 2
$\Delta\phi(j_1,\ j_2)$	< 1.5
m_{jj}	> 500 GeV

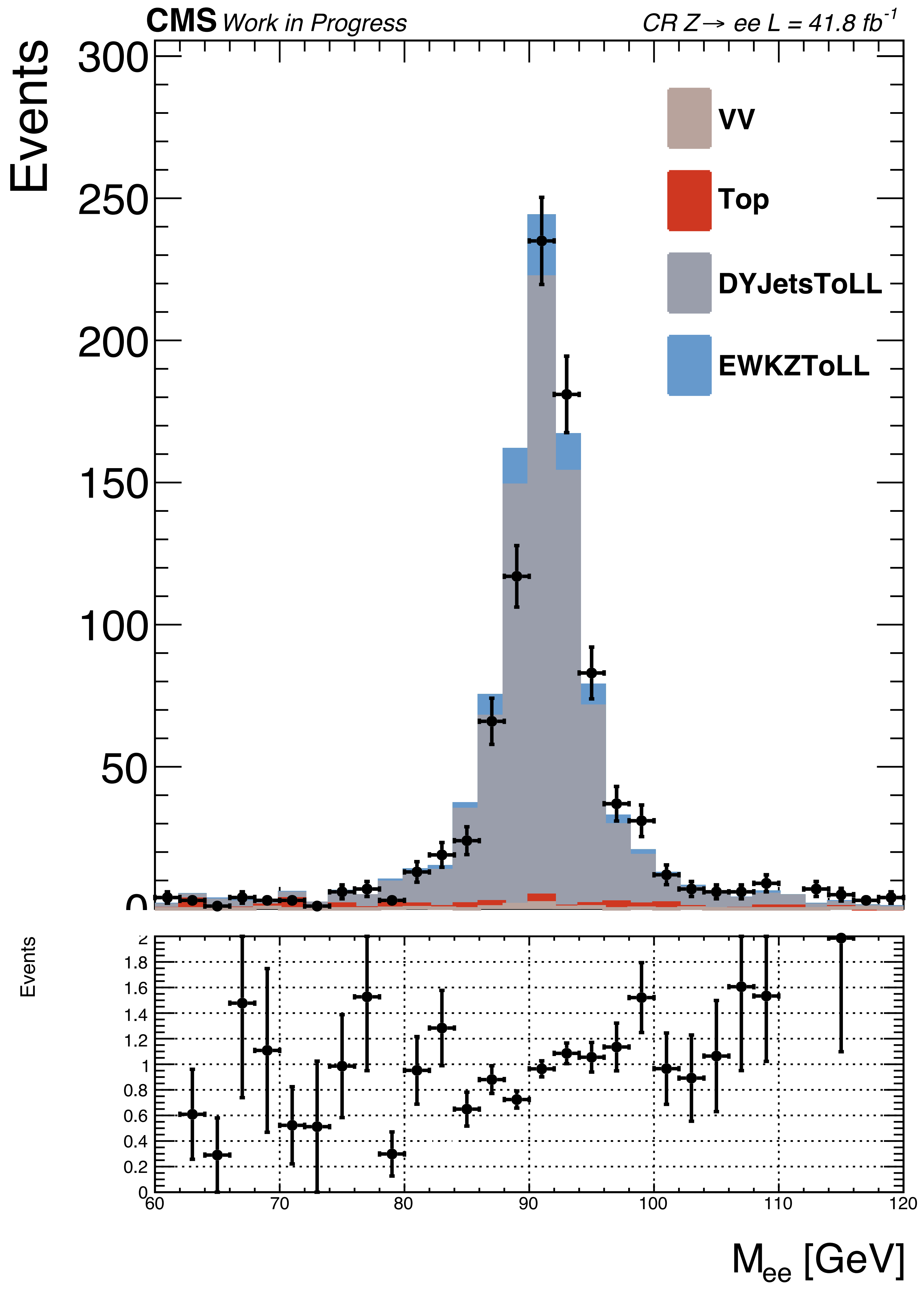


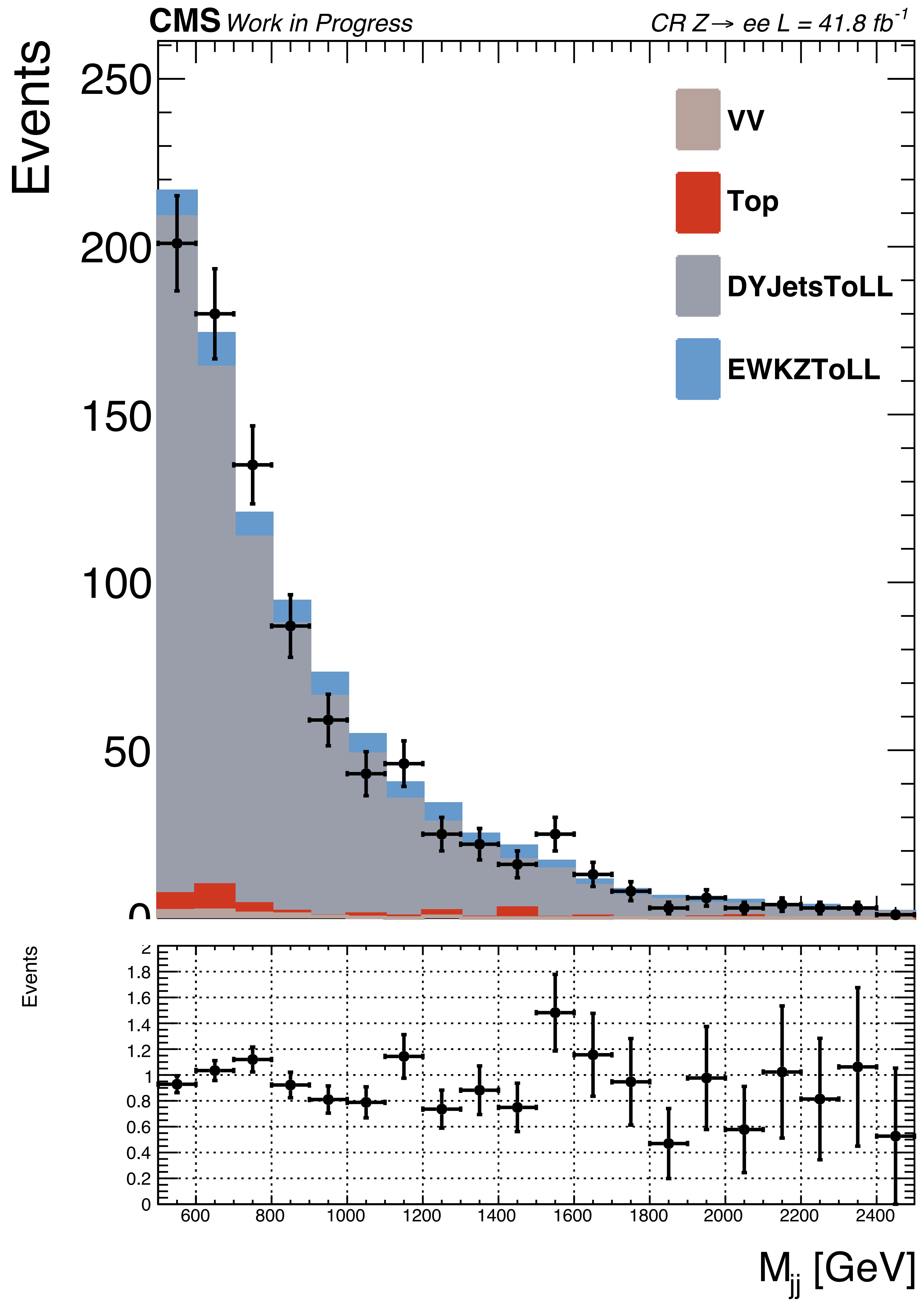


Increasing the sensitivity: A first look at 2017 data

- **Zee control region:**
  - **Preliminary results**
  - **Full 2017 dataset: 41.8 fb-1**

Variable	Requirement
$p_T^{j_1}$	> 80 GeV
$p_T^{j_1}$	> 40 GeV
E_T^{miss}	> 250 GeV
$min\Delta\phi(E_T^{miss},\ j)$	> 0.5 GeV
$\Delta\eta(j_1,\ j_2)$	> 2
$\Delta\phi(j_1,\ j_2)$	< 1.5
m_{jj}	> 500 GeV





Conclusion: Plans for the legacy paper

- Plans for the VBF analysis:
  - Usage of new VBF based triggers
  - Better trigger performance: looser selection requirements on MET
  - Moving from a single bin to a multi bin fit for obtaining the final result

- Legacy publication containing the **entire Run 2** information
  - A collaboration between several UK CMS teams (more details can be seen in a given by Esh -Link-)

    - Parallel study of **all modes** (VBF, ttH, VH and ggH):
      - Built-in orthogonality
      - Sharing background estimation methods
      - Using the same systematics wherever possible

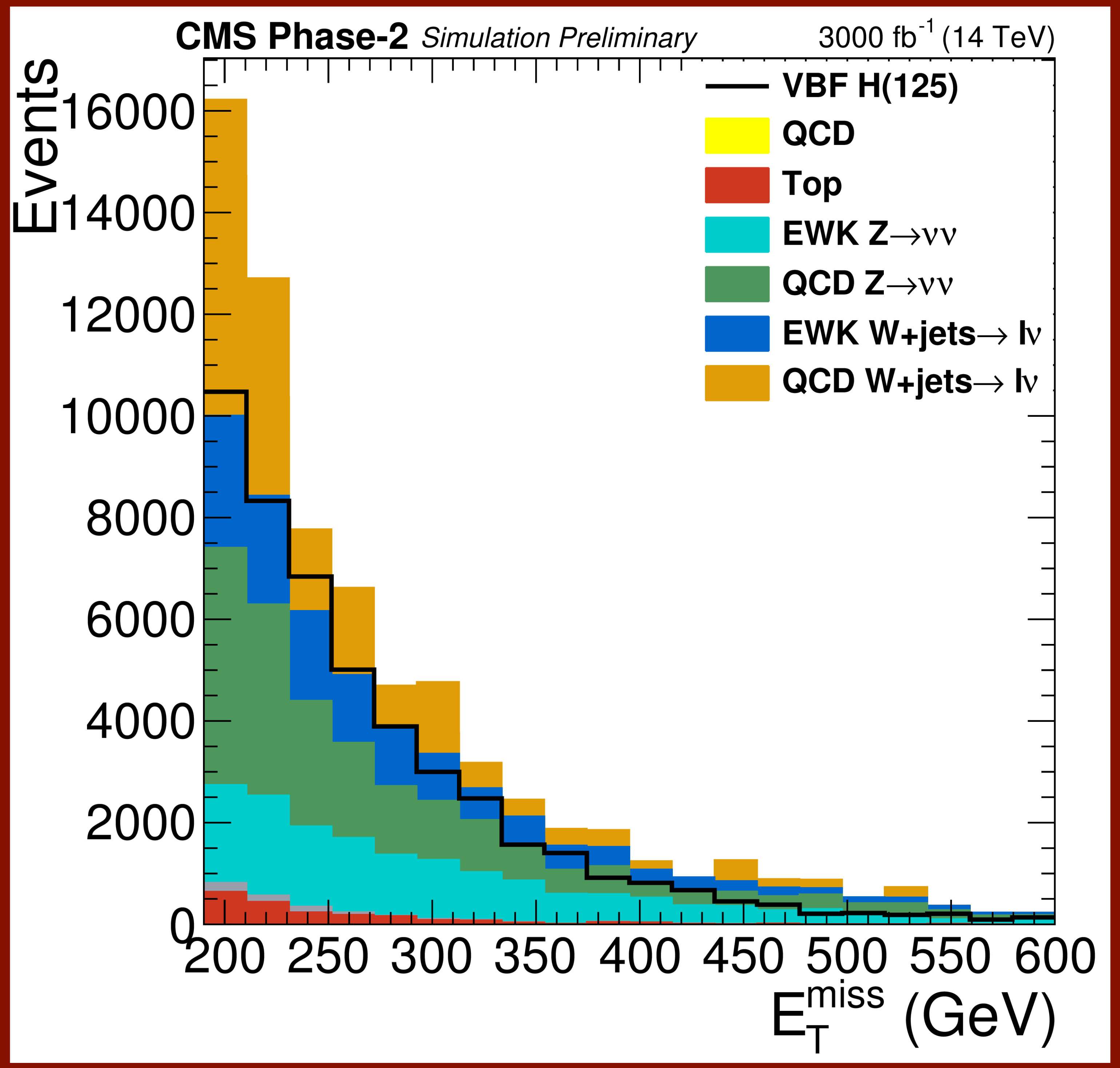
  - Analysis is currently being prepared with the end of summer 2019 as its goal

Thank you for your time!

Backup

Conclusion: Future of the analysis

- The main idea behind High Luminosity LHC (HL-LHC) is to significantly increase the number of proton-proton collisions per second
- Effects on the CMS experiment:
  - Increased radiation
  - Large in-time event pileup
- Proposed plan: the High Granularity Calorimeter (HGCal):
  - Replacing the present endcap calorimeters
  - Uses a combination of transverse and longitudinal segmentation for all calorimetry components
- The VBF H → inv. is interesting due to its MET dependance
  - Full performance study published as a part of the HL-LHC YR
- Ongoing studies: Trigger requirements for Run 3



B

2017 data analysis

- Idea to use the new “FAST” framework, a software used by a collaboration between several UK CMS teams for the “Combined Higgs to Invisible Project - CHIP”
  - Plan is to combine all hadronic analyses (within UK)
  - Modular approach - allows us to build analysis specific computations by specifying which parts of the code we need/ adding new custom-made packages
  - nanoAOD friendly approach: Complete inclusion of nanoAODtools and the new Ntuple format
  - Removal of ROOT dependency: Binning the data into data frames instead of creating a new “mediator” tree after the selection
  - Configuration files: Summarising all the variables (binning, ranges, selections) needed for studies in one YAML configuration file

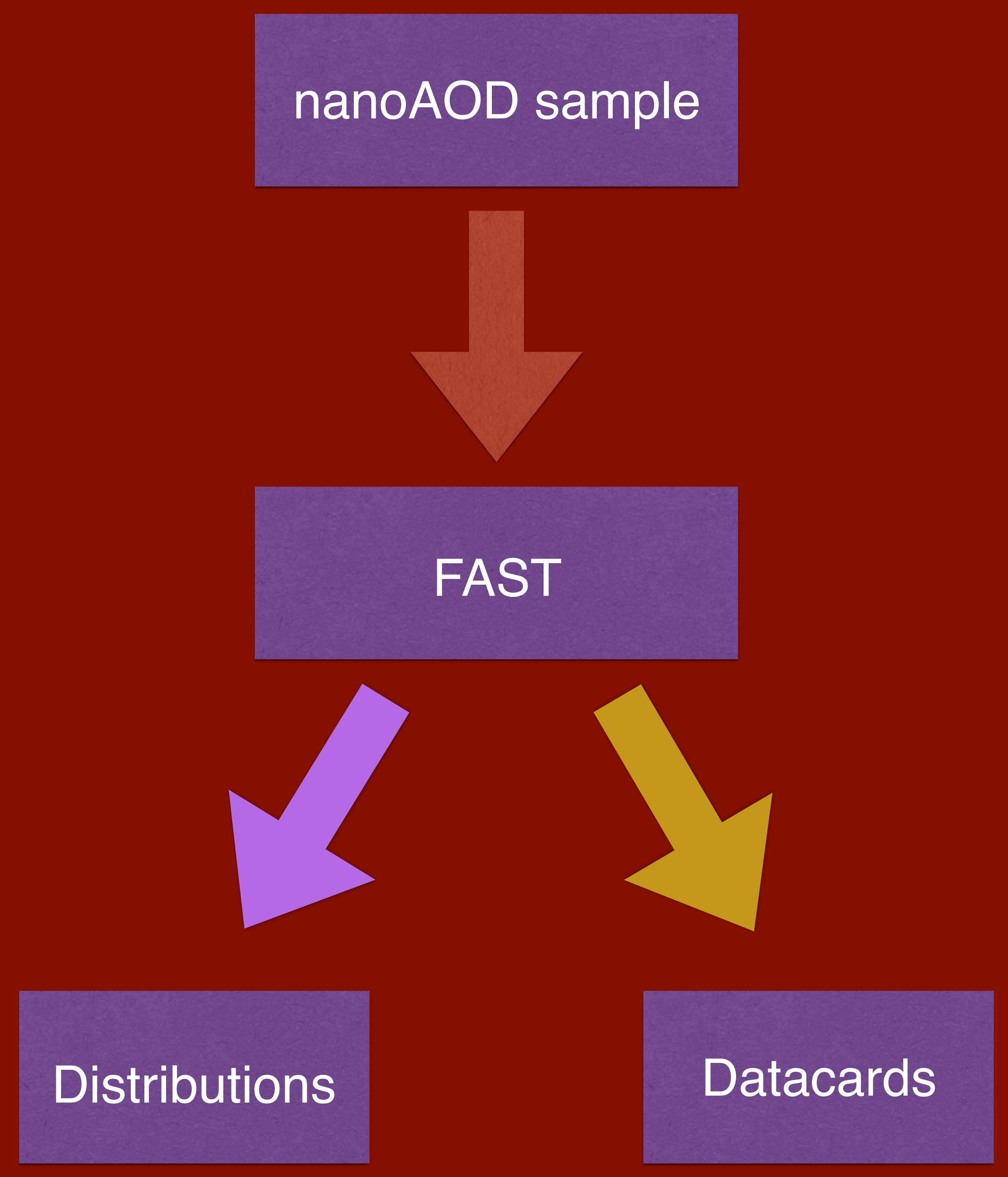


Figure 19: Diagram of the workflow of the FAST software

$$N^{SR}_{expected} = \frac{\sigma(\mathrm{Z} \to \nu\nu)}{\sigma(\mathrm{Z} \to ll)} \cdot \frac{\epsilon^{SR}}{\epsilon^{CR}} \cdot (N^{CR}_{data} - N^{CR}_{bkg})$$

$$\epsilon^{SR} = \frac{\sigma(\mathrm{Z} \to \nu\nu, \mathrm{EWK}) \cdot \frac{N^{SR}_{MC}(\mathrm{EWK})}{N_{gen}(\mathrm{M_Z,EWK})} + \sigma(\mathrm{Z} \to \nu\nu, \mathrm{QCD}) \cdot \frac{N^{SR}_{MC}(\mathrm{QCD})}{N_{gen}(\mathrm{M_Z,QCD})}}{\sigma(\mathrm{Z} \to \nu\nu, \mathrm{EWK}) + \sigma(\mathrm{Z} \to \nu\nu, \mathrm{QCD})}$$

$$\epsilon^{CR} = \frac{\sigma(\mathrm{Z} \to ll, \mathrm{EWK}) \cdot \frac{N^{CR}_{MC}(\mathrm{EWK})}{N_{gen}(\mathrm{EWK})} + \sigma(\mathrm{Z} \to ll, \mathrm{QCD}) \cdot \frac{N^{CR}_{MC}(\mathrm{QCD})}{N_{gen}(\mathrm{QCD})}}{\sigma(\mathrm{Z} \to ll, \mathrm{EWK}) + \sigma(\mathrm{Z} \to ll, \mathrm{QCD})}$$

$$N^{SR}_{expected} = \frac{N^{SR}_{MC}}{N^{CR}_{MC}} \cdot (N^{CR}_{data} - N^{CR}_{bkg})$$

- **Region-A**: min-$\Delta\phi(j, E_{\mathrm{T}}^{\mathrm{miss}}) < 0.5$ and $100 < E_{\mathrm{T}}^{\mathrm{miss}} < 160$ GeV.
- **Region-B**: min-$\Delta\phi(j, E_{\mathrm{T}}^{\mathrm{miss}}) > 0.5$ and $100 < E_{\mathrm{T}}^{\mathrm{miss}} < 160$ GeV
- **QCD-CR**: min-$\Delta\phi(j, E_{\mathrm{T}}^{\mathrm{miss}}) < 0.5$ and $E_{\mathrm{T}}^{\mathrm{miss}} > 250$ GeV.
- **Signal region**: min-$\Delta\phi(j, E_{\mathrm{T}}^{\mathrm{miss}}) > 0.5$ and $E_{\mathrm{T}}^{\mathrm{miss}} > 250$ GeV.

$$N_{QCD}^{SR}(m_{jj}) = \left(N_{Data}^{CR}(m_{jj}) - \sum_{i}^{bkg} N_i^{CR}(m_{jj}) \right) \cdot r(m_{jj})$$

$$r = \frac{\min \Delta\phi(\mathrm{jet}, \mathrm{E}_{\mathrm{T}}^{\mathrm{miss}}) > 0.5}{\min \Delta\phi(\mathrm{jet}, \mathrm{E}_{\mathrm{T}}^{\mathrm{miss}}) < 0.5}$$

The V+jets background yields are determined using a maximum-likelihood fit, performed simultaneously across all CRs and the SR. The likelihood function is defined as:

$$\begin{aligned}
\mathcal{L}(\mu,\kappa^{\nu\overline{\nu}},\boldsymbol{\theta}) = & \prod_i \mathrm{P}\left(d_i \middle| B_i(\boldsymbol{\theta}) + (1+f_i(\boldsymbol{\theta})_{\mathrm{Q}})\kappa_i^{\nu\overline{\nu}} + R_i^{\mathrm{Z}}(1+f_i(\boldsymbol{\theta})_{\mathrm{E}})\kappa_i^{\nu\overline{\nu}} + \mu S_i(\boldsymbol{\theta})\right) \\
& \prod_i \mathrm{P}\left(d_i^{\mu\mu} \middle| B_i^{\mu\mu}(\boldsymbol{\theta}) + \frac{\kappa_i^{\nu\overline{\nu}}}{R_i^{\mu\mu}(\boldsymbol{\theta})_{\mathrm{Q}}} + \frac{R_i^{\mathrm{Z}}\kappa_i^{\nu\overline{\nu}}}{R_i^{\mu\mu}(\boldsymbol{\theta})_{\mathrm{E}}}\right) \\
& \prod_i \mathrm{P}\left(d_i^{\mathrm{ee}} \middle| B_i^{\mathrm{ee}}(\boldsymbol{\theta}) + \frac{\kappa_i^{\nu\overline{\nu}}}{R_i^{\mathrm{ee}}(\boldsymbol{\theta})_{\mathrm{Q}}} + \frac{R_i^{\mathrm{Z}}\kappa_i^{\nu\overline{\nu}}}{R_i^{\mathrm{ee}}(\boldsymbol{\theta})_{\mathrm{E}}}\right) \\
& \prod_i \mathrm{P}\left(d_i^{\mu} \middle| B_i^{\mu}(\boldsymbol{\theta}) + \frac{f_i(\boldsymbol{\theta})_{\mathrm{Q}}\,\kappa_i^{\nu\overline{\nu}}}{R_i^{\mu}(\boldsymbol{\theta})_{\mathrm{Q}}} + \frac{R_i^{\mathrm{Z}} f_i(\boldsymbol{\theta})_{\mathrm{E}}\,\kappa_i^{\nu\overline{\nu}}}{R_i^{\mu}(\boldsymbol{\theta})_{\mathrm{E}}}\right) \\
& \prod_i \mathrm{P}\left(d_i^{\mathrm{e}} \middle| B_i^{\mathrm{e}}(\boldsymbol{\theta}) + \frac{f_i(\boldsymbol{\theta})_{\mathrm{Q}}\,\kappa_i^{\nu\overline{\nu}}}{R_i^{\mathrm{e}}(\boldsymbol{\theta})_{\mathrm{Q}}} + \frac{R_i^{\mathrm{Z}} f_i(\boldsymbol{\theta})_{\mathrm{E}}\,\kappa_i^{\nu\overline{\nu}}}{R_i^{\mathrm{e}}(\boldsymbol{\theta})_{\mathrm{E}}}\right) \prod_j \mathrm{P}(\theta_j)
\end{aligned} \tag{1}$$