Combined search for an invisibly decaying Higgs boson in hadronic channels at $\sqrt{s} = 13$ TeV with CMS

Eshwen Bhal (University of Bristol)

10th April, 2019

University of BRISTOL

CMS

Outline

- Background and motivation
- The CMS detector
- Higgs production modes
- Analysis strategy
- Signal region categorisation
- Background estimation

Background and motivation

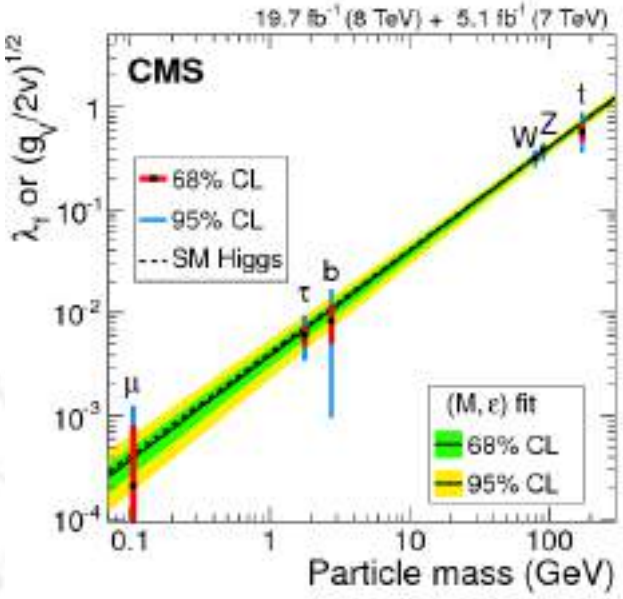
Why the Higgs?

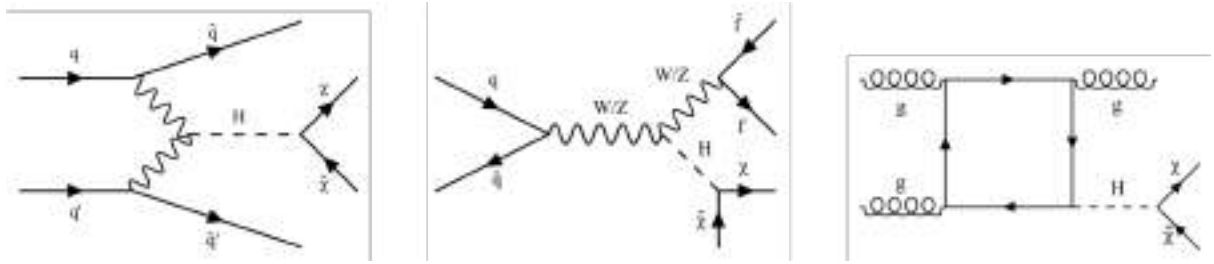
◎ The Higgs boson is the newest and least well-measured particle in the Standard Model (SM). Currently the **most interesting**

◎ It couples proportionally to mass, giving elementary particles theirs

◎ We know dark matter is massive. Does it also come from the Higgs?

◎ Our goal is to measure the $H \rightarrow$ inv. branching ratio (BR). A sufficiently constrained limit above the SM prediction could suggest **new physics**





University of BRISTOL

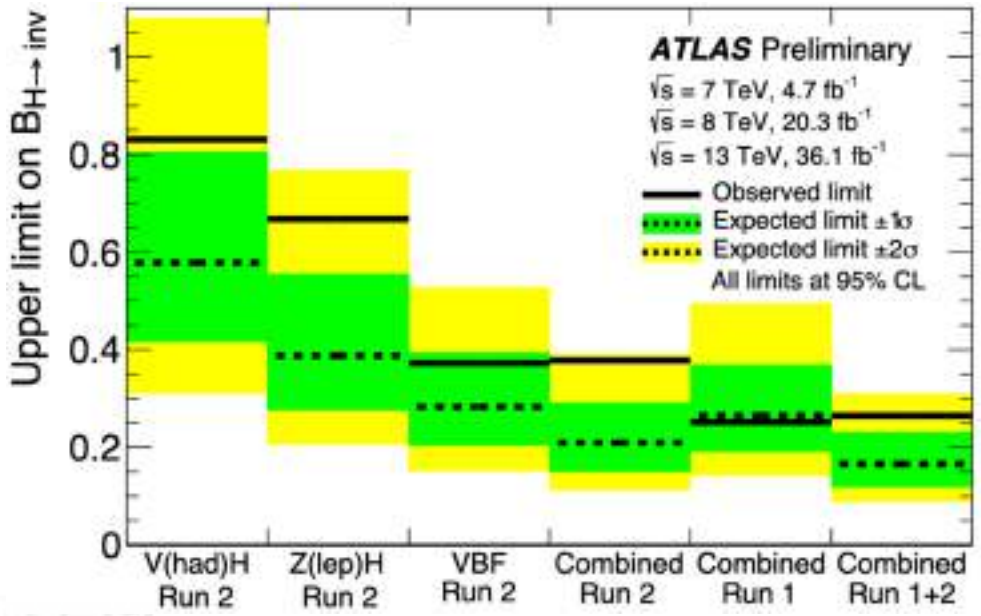
Current state of Higgs physics

◎ In SM, the Higgs can only decay invisibly via $H \rightarrow ZZ \rightarrow 4\upsilon$ with $\mathbf{BR} = \mathbf{0.1}\ \%$

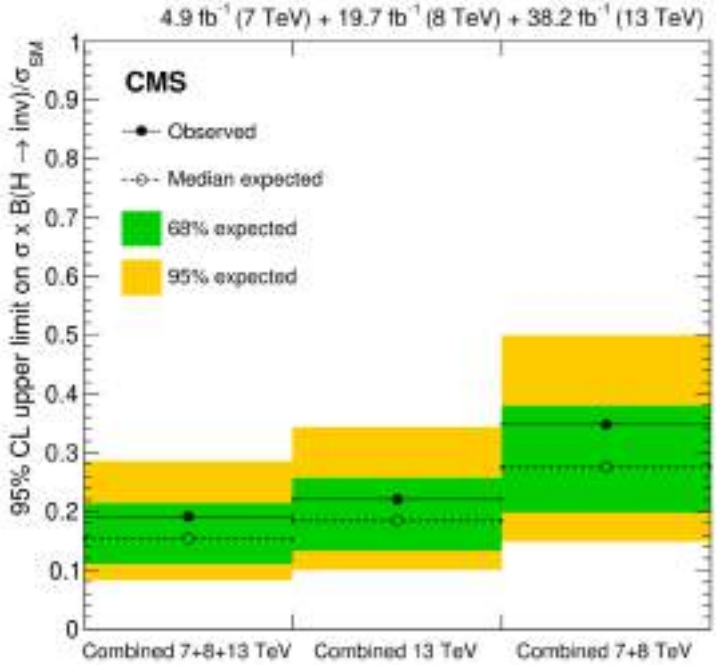
◎ Best experimental limit comes from combination over all Higgs production modes

◎ Historically, each channel analysed in separate publication then combined posthoc

ATLAS-CONF-2018-054



arXiv:1809.05937



The CMS detector

The CMS detector

Just kidding, we've all seen it before

Higgs production modes

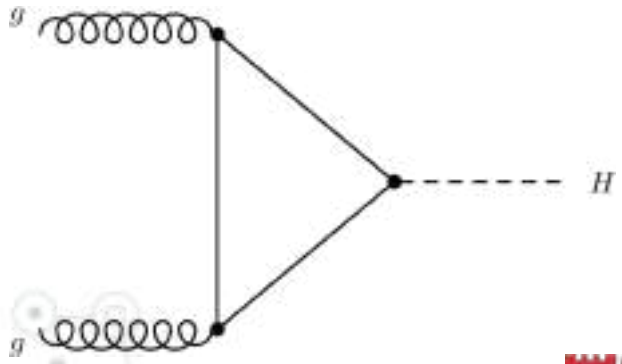
Production modes

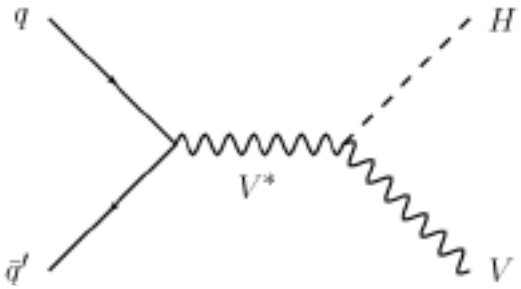
◎ Higgs production split into two main mechanisms: VBF (single most-sensitive channel, see Vukasin Milosevic's talk) and non-VBF

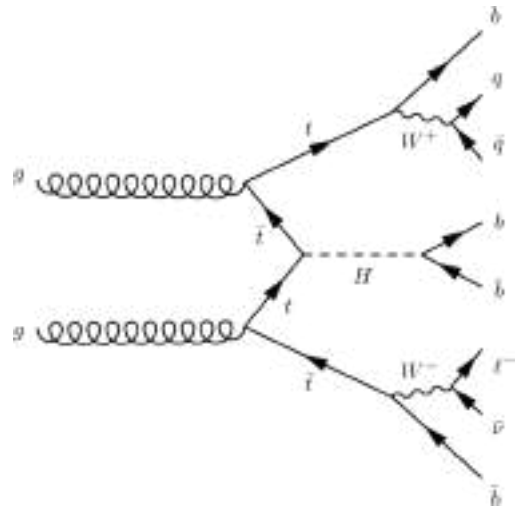
◎ non-VBF channels:

- Gluon-gluon fusion ($ggF/\ ggH$)
- Top quark pair production in association with a Higgs (ttH)
- Vector boson production in association with a Higgs (VH: W^+H, W^-H, ZH)

◎ Each have different characteristics, kinematics, cross sections. Possible to study in separate, orthogonal categories to maximise sensitivity







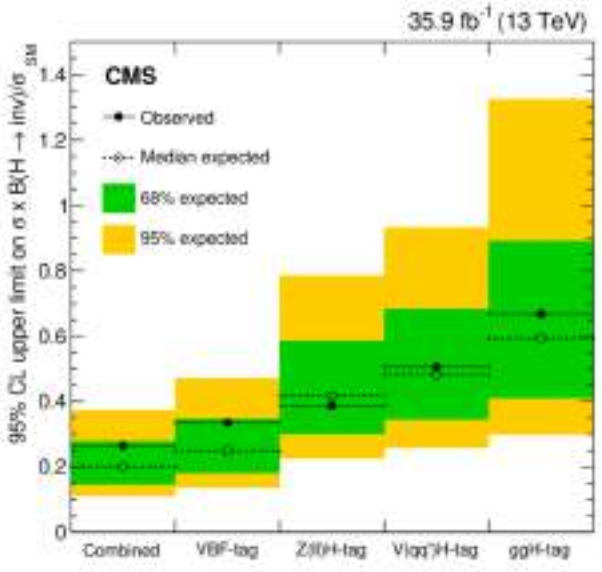
Sensitivity of each production mode

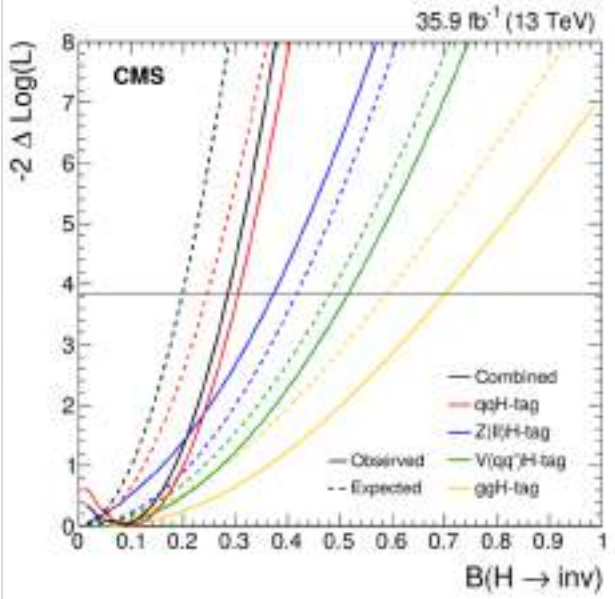
VBF, VH, ggF:

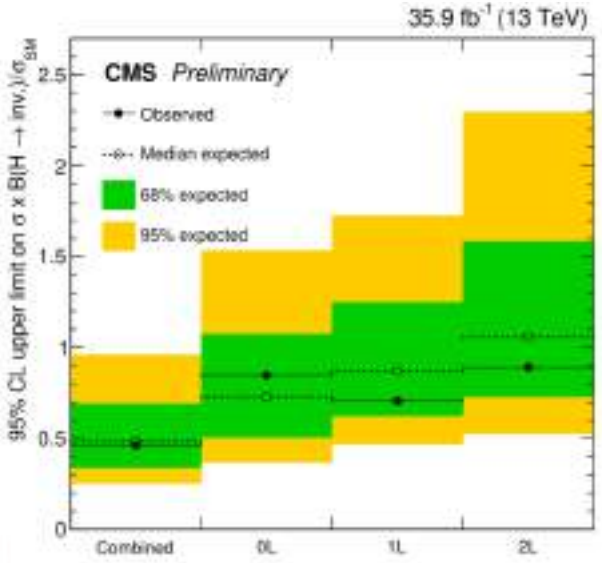
arXiv:1809.05937

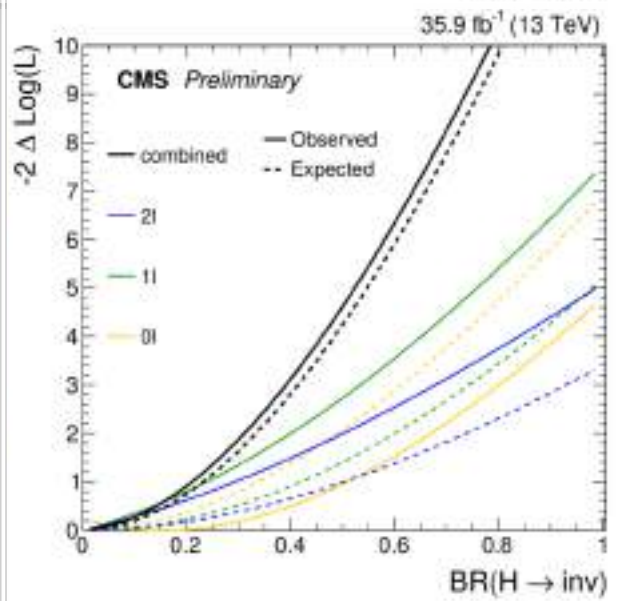
ttH:

CMS-PAS-HIG-18-008









Analysis strategy

Analysis strategy

◎ Our strategy is to perform combination of production modes from the start

- Taking inspiration from SUSY analyses; searching for many topologies simultaneously
- Using common workflow/tools for each mode
- Final combined result much more straightforward, sharing systematics and background methods

◎ Analysing each channel in an orthogonal search region should give high degree of sensitivity and phase space coverage

Signal region categorisation

Signal region categorisation

◎ Events categorised according to hadronic variables: H_{T}, $H_{\mathrm{T}}^{\mathrm{miss}}$, n_{jet}, $n_{\mathrm{b-jet}}$

- Shapes and characteristics quite different depending on production mode

◎ Baseline selection:

- $e,\ \mu,\ \gamma$ vetoes
- $H_{\mathrm{T}},\ H_{\mathrm{T}}^{\mathrm{miss}},\ n_{\mathrm{jet}}$ cuts
- $\Delta\phi_{\mathrm{min}}(\mathrm{jet},\ p_{\mathrm{T}}^{\mathrm{miss}})$ for QCD rejection (also considering others – arXiv:1803.07942)
- "Not" of the VBF preselection

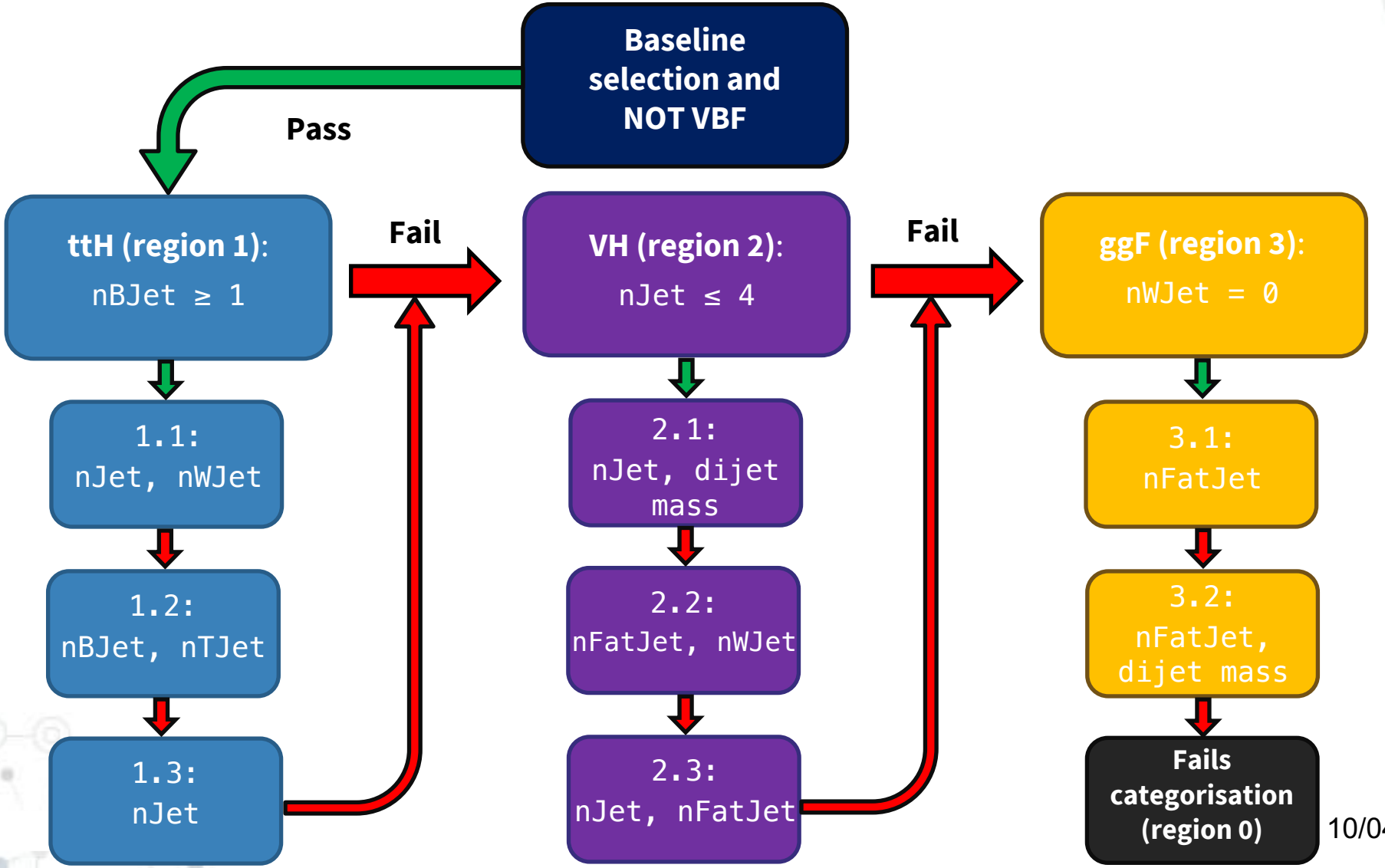
◎ Categories not final yet, still evolving

$$H_{\mathrm{T}} = \sum_{i} p_{\mathrm{T}}^{\mathrm{jet},\, i}$$

$$H_{\mathrm{T}}^{\mathrm{miss}} = -\sum_{i} \vec{p}_{\mathrm{T}}^{\mathrm{jet},\, i}$$

Signal region categorisation

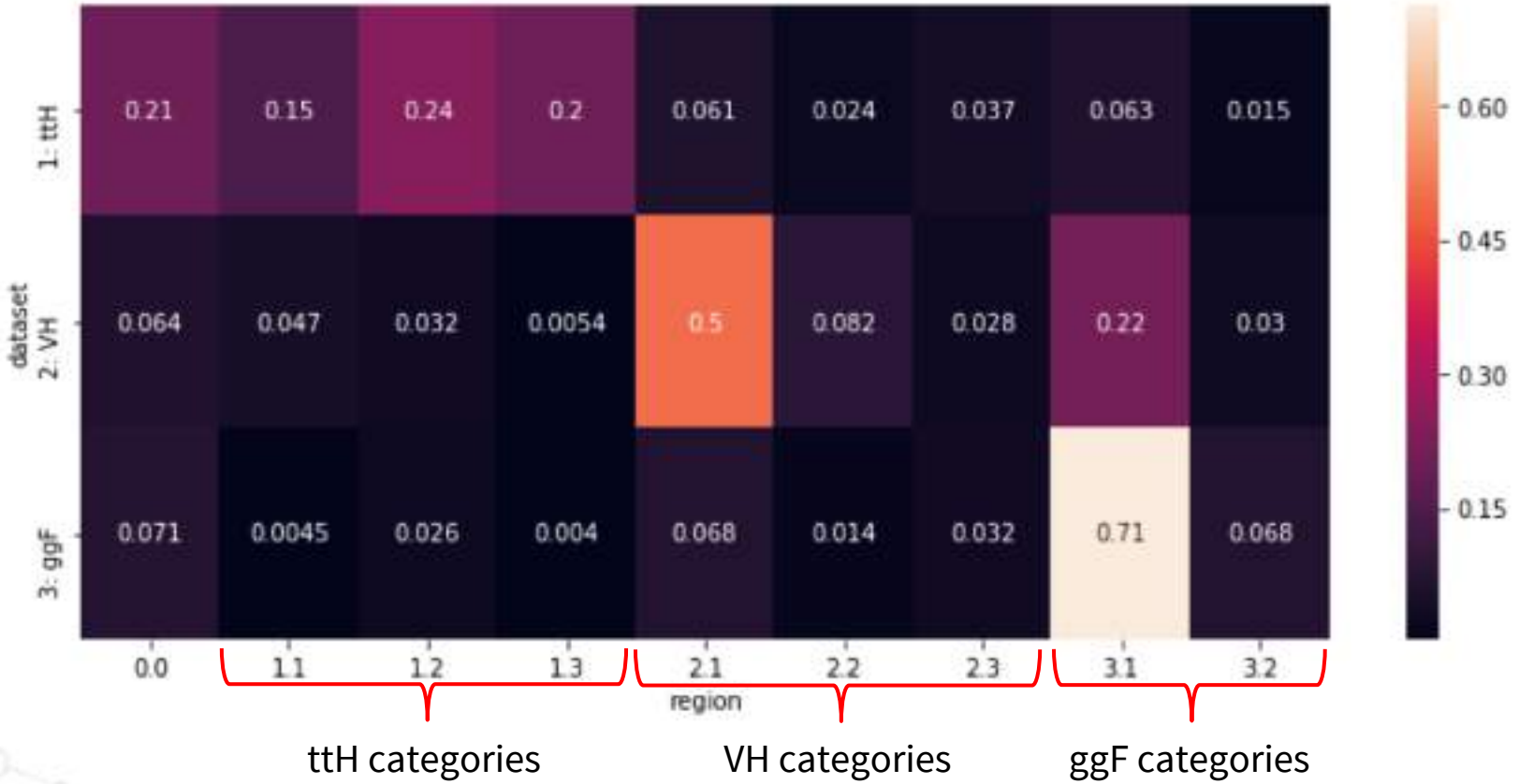
◎ Top and W jet tagging used to aid categorisation: arXiv:1011.2268, arXiv:1402.2657



Signal region categorisation

◎ Signal efficiencies in each region after applying the baseline event selection and categorisation

◎ Contamination from other production modes can also be seen



Background estimation

Background estimation

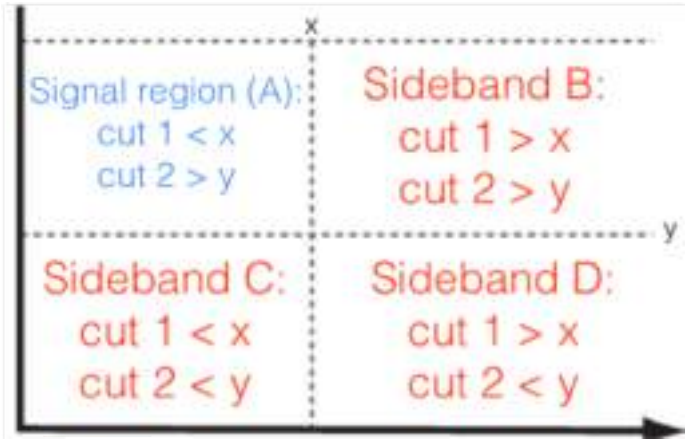
◎ Four dedicated control regions enriched in background processes that also occur in signal regions:

- Single muon and single electron to predict $W \to \ell\nu$ (the lepton being "lost")
- Double electron and double muon to predict $Z \to \nu\nu$

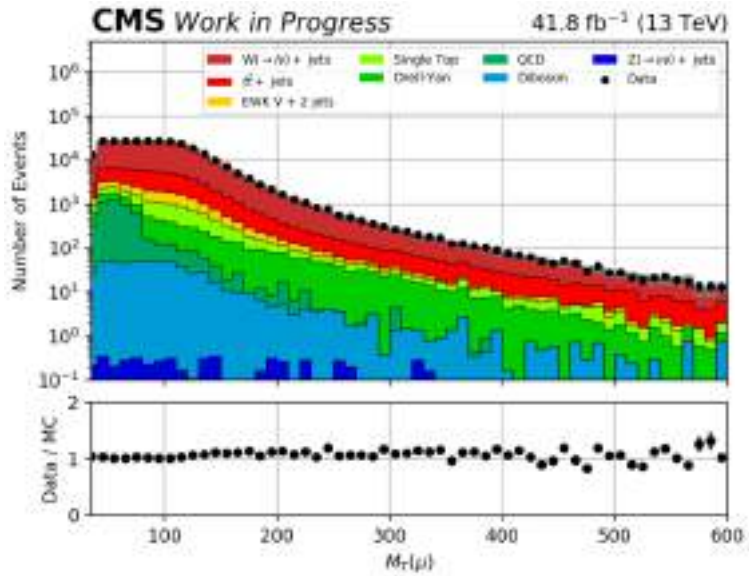
◎ Double lepton regions can **model $Z \to \nu\nu$ kinematics** by excluding leptons from MET calculation

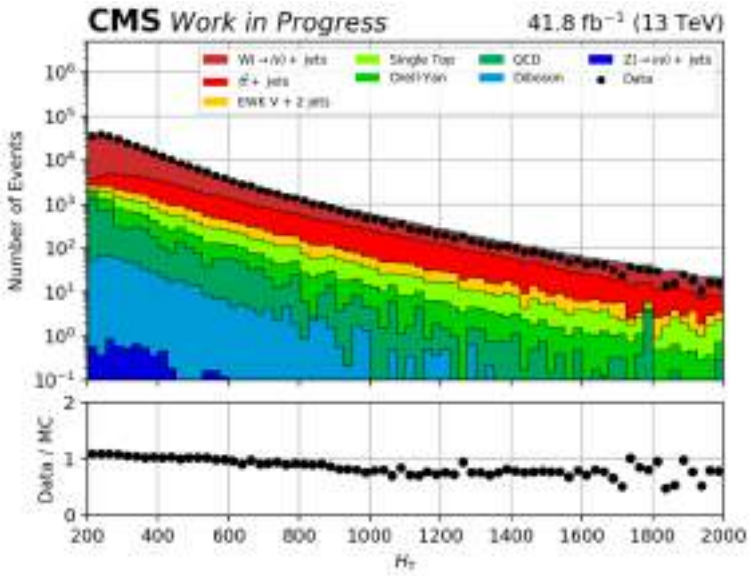
◎ Can extrapolate yields to estimate backgrounds in signal region

◎ QCD can be estimated separately using sidebands: signal region with two cuts inverted (ABCD method)

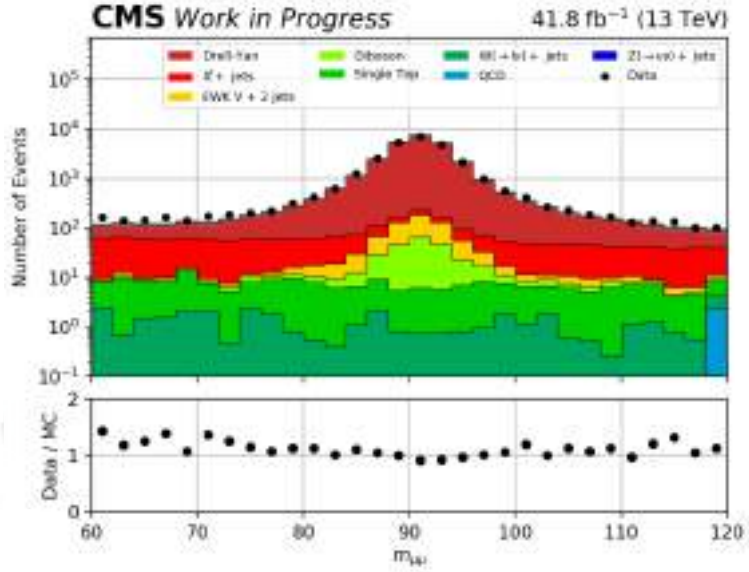


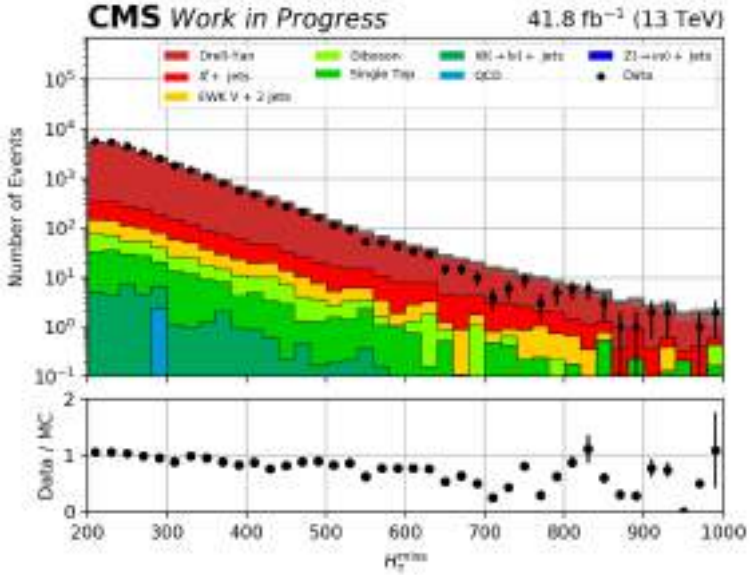
Muon control regions





Single muon





Double muon

Summary and plans

◎ With the Higgs discovery, many analyses underway to understand it and its influence on SM. Possible entry points to the dark sector

◎ Presented **control region studies** and first look at **signal region categorisation** for non-VBF production

◎ A short to-do list:

- Refine control regions (particularly electron ones) for signal region backgrounds
- Finalise signal categorisation
- Take a first look at sensitivity/limit with systematic uncertainties included

◎ Final result to include full 2016-18 datasets; over 130 fb^{-1} at $\sqrt{s} = 13$ TeV

Thanks for listening!

Backup

The CMS detector

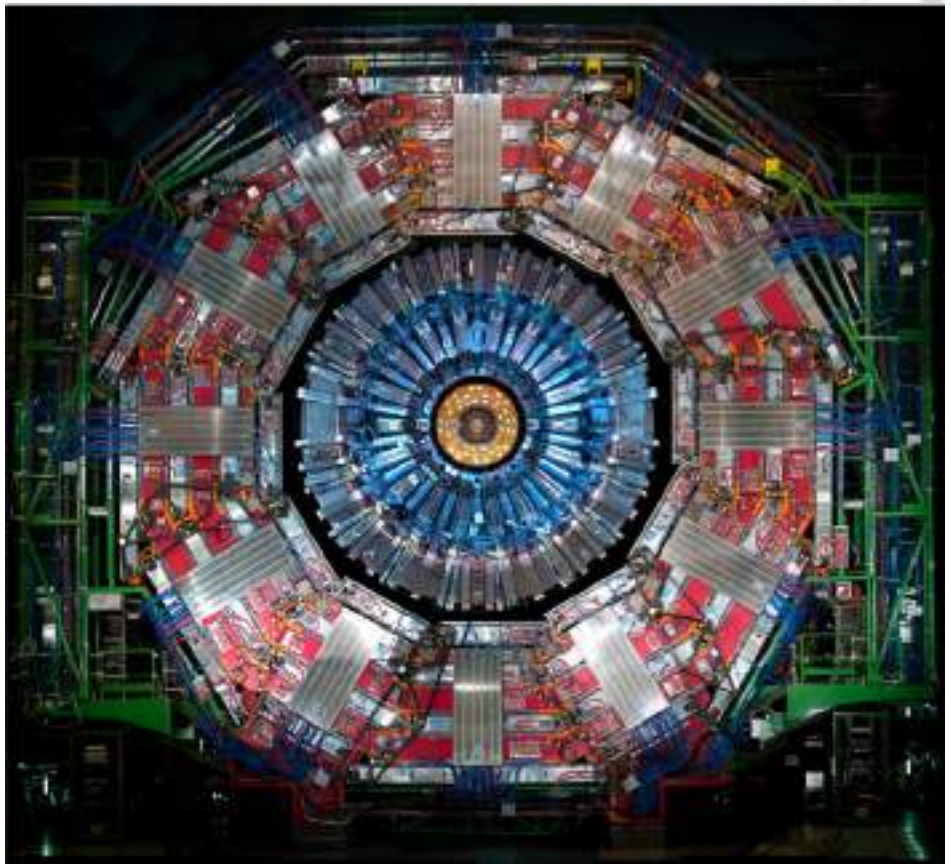
◎ General purpose detector

◎ Weighs 14,000 tons

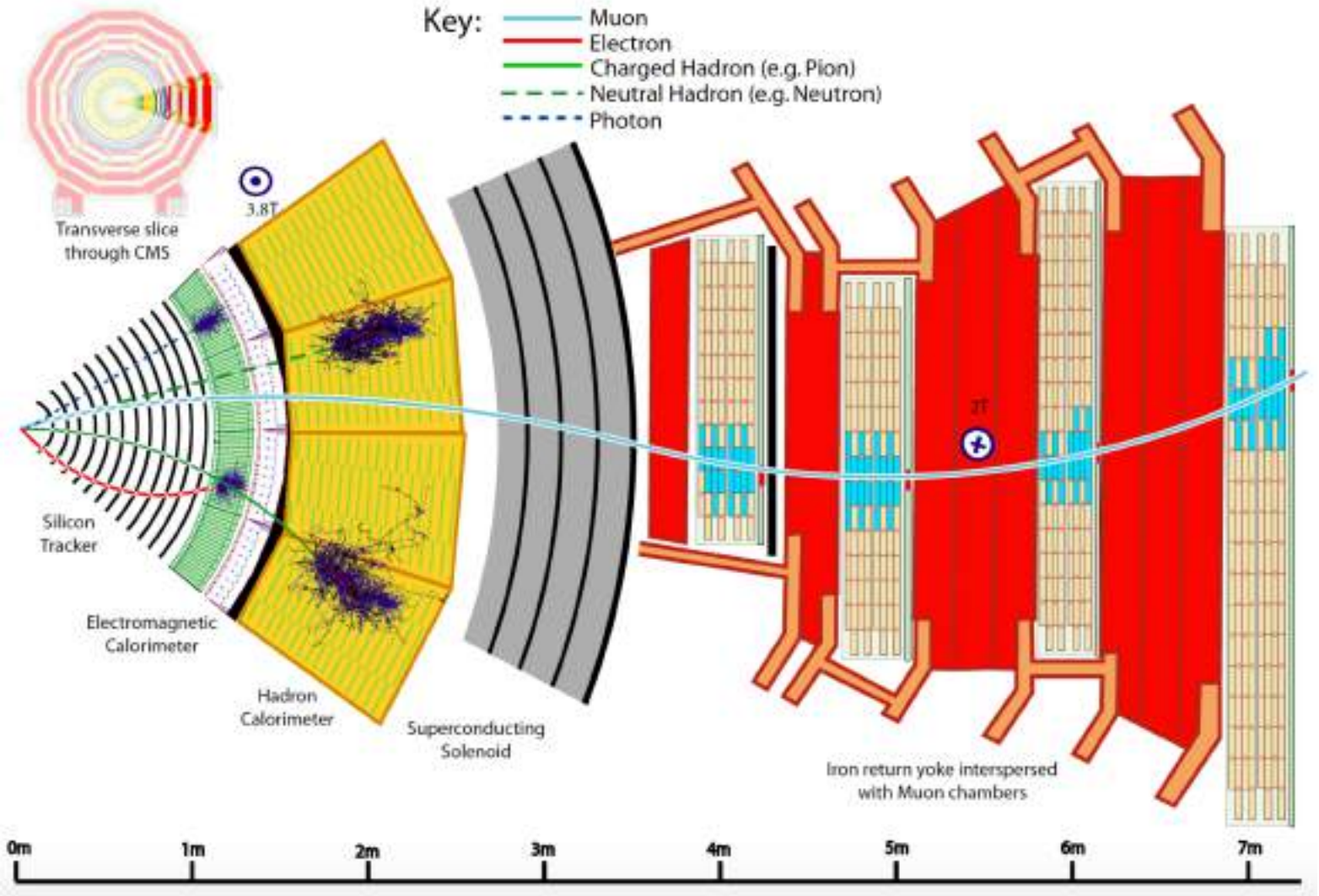
◎ 4 T magnet storing 2.7 GJ of energy

◎ Level-1 and High Level Triggers reduce event rate from 40 MHz to ~1 khZ: O(GB/s) written to disk

◎ Analyses on supersymmetry & exotic searches, Higgs, top, *b*-physics, heavy ions, SM precision measurements



The CMS detector



Software and tools

◎ Preceding the analysis, a lot of time has been spent developing tools to optimise future work

◎ Using CMS' nanoAOD format: flat tree with small event size (~1 kB). Allows for quicker I/O

◎ Analysis makes use of FAST tools: fast-carpenter, fast-curator, fast-plotter

◎ Read trees using uproot, converting output into pandas dataframes or awkward jagged arrays

◎ Vectorised formats allow minimal loops and **fast operations**

◎ Parallelisation locally and on batch mean *all* samples for analysis can be run in O(hours)

Software and tools

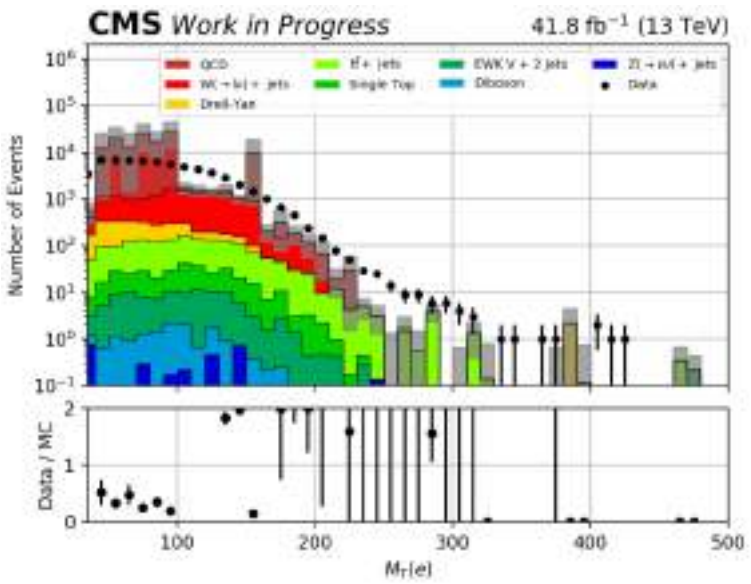
◎ Entire analysis can be specified in a YAML config file. Can include

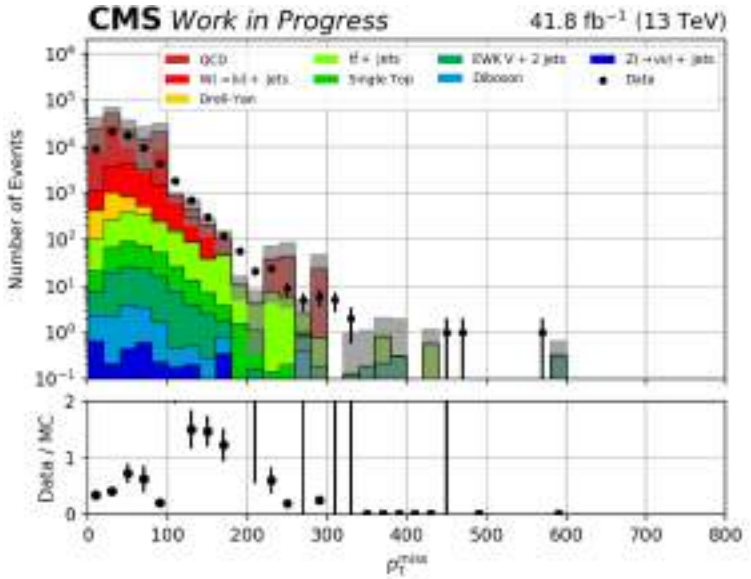
- Event selections
- Custom Python classes to create new variables, operate on existing ones
- Object and event weights
- Variables to plot and respective binning schemes

◎ Output from fast-carpenter in binned dataframe format. Can be easily fed into matplotlib for visualisation

◎ Better introduction at fast-carpenter: turning trees into tables

Electron control regions



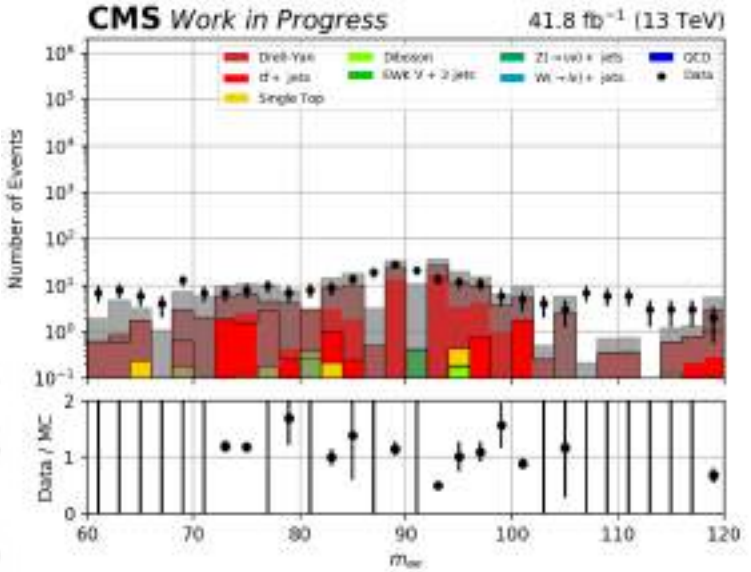


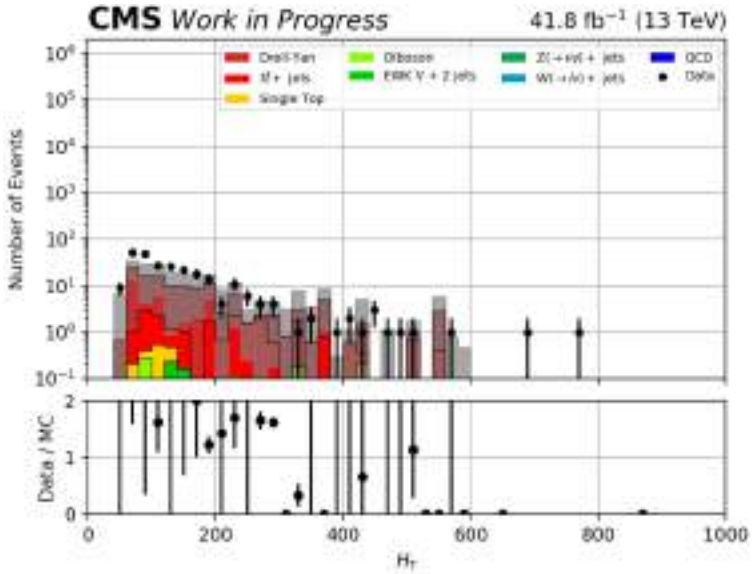
Single electron

◎ Not very good data-MC agreement currently

◎ Low stats in double electron region

◎ Have iterated extensively, still working on them





Double electron