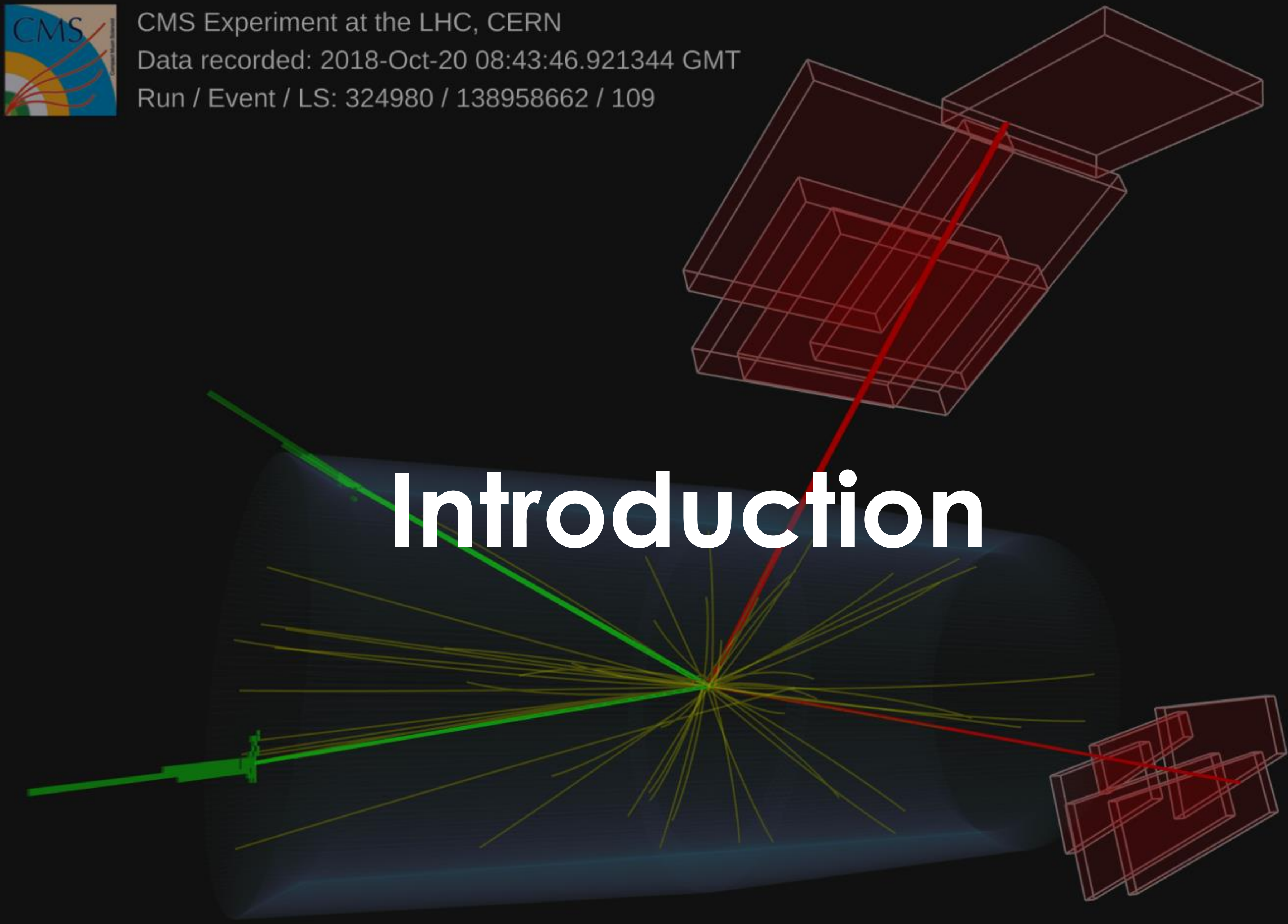
# Search for heavy resonances in the 4-lepton final state at CMS

## IRN Terascale

Nov. 13th-15th, 2024

**Geliang Liu on behalf of CMS Collaboration**

# Introduction



# Motivation

## Why new resonances?

**Despite the great success of the Standard Model (SM), many questions are not answered by it**

- Why 3 generations?
- Hierarchy problem?
- Dark matter and dark energy?
- Matter-anti-matter asymmetry?
- ……

**New resonances predicted by many theories beyond the SM (BSM)**

- **Additional Higgs bosons**:
  - extended SM / Higgs sector
  - supersymmetry
- **Radion**:
  - warped extra dimension

# Motivation

Why new resonances?

Despite the great success of the Standard Model (SM), many questions are not answered by it

- Why 3 generations?
- Hierarchy problem?
- Dark matter and dark energy?
- Matter-anti-matter asymmetry?
- ……

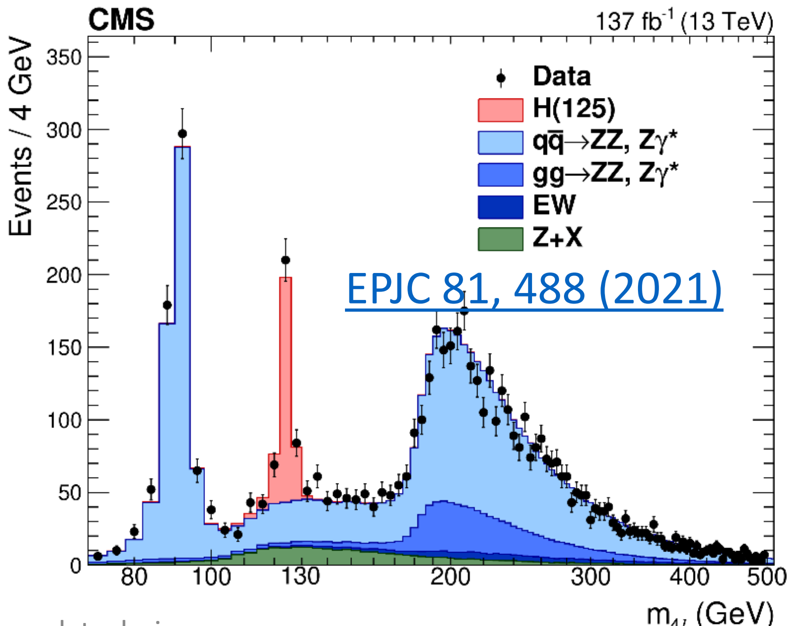
**New resonances** predicted by many theories beyond the SM (BSM)

- **Additional Higgs bosons**:
  - extended SM / Higgs sector
  - supersymmetry
- **Radion**:
  - warped extra dimension

## Why ZZ → 4 leptons?

**Golden channel:**

- Great S/B ratio for the SM Higgs boson
- High efficiency and good resolution of e, μ
- Well modeling of the background processes




EPJC 81, 488 (2021)

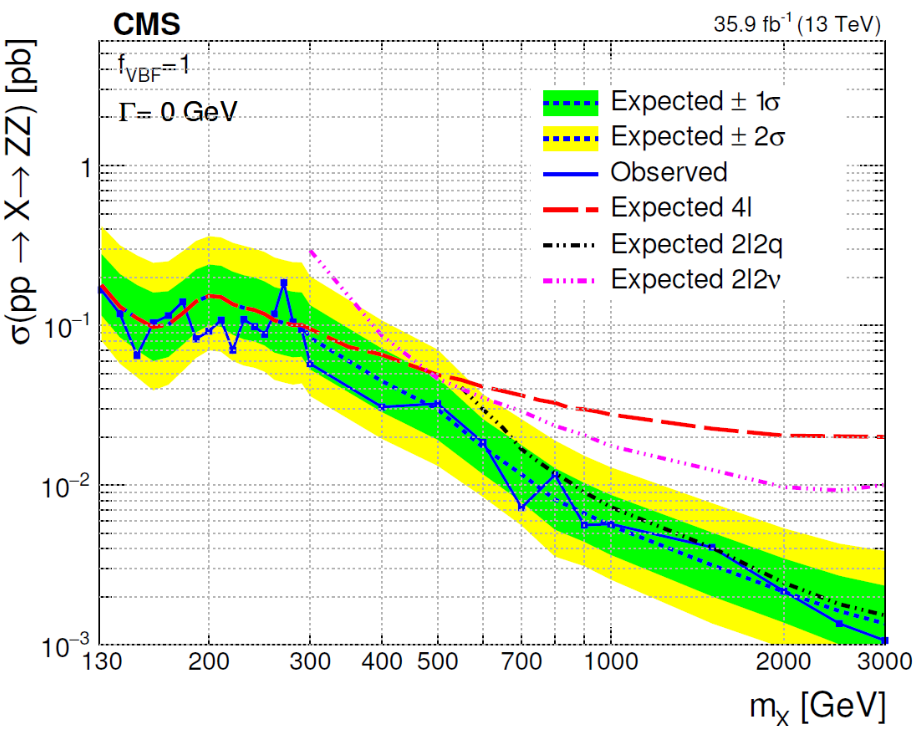
# Relevant public results

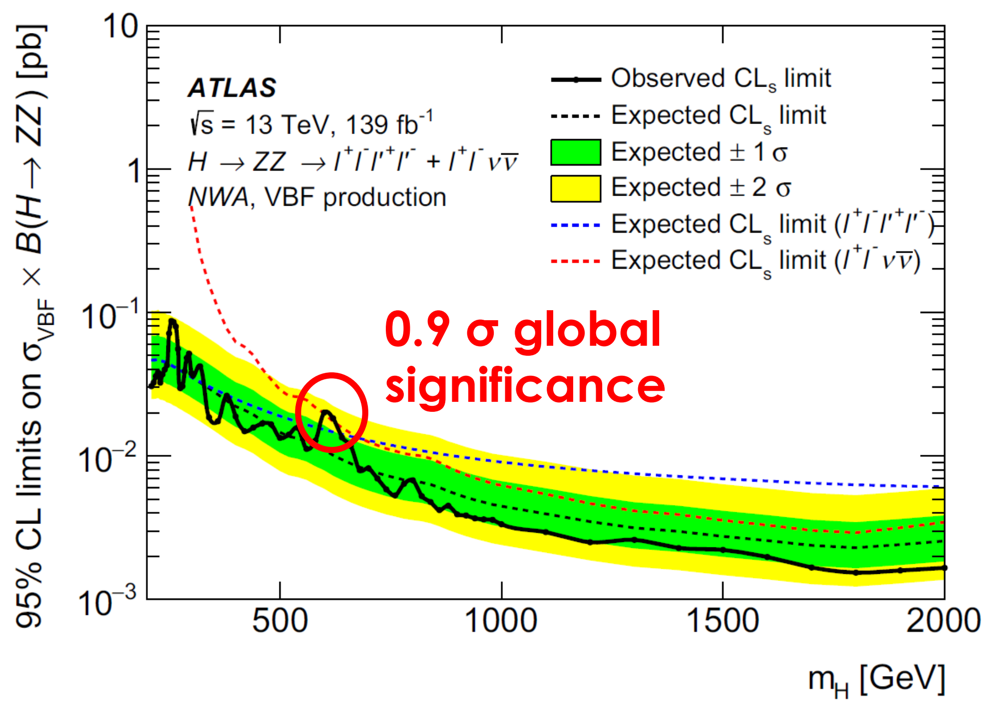
**From CMS experiments:**

- [JHEP 2018, 127 (2018)]
- **4l+2l2q+2l2ν** with the 2016 dataset
- Mass: 130 – 3000 GeV
- Width: 0 to 30% of mass
- **Model-independent searches**

**From ATLAS experiments:**

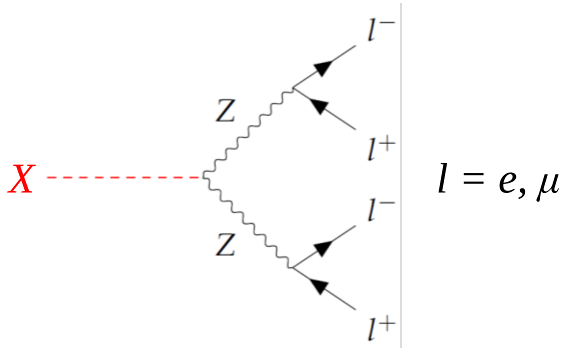
- [EPJC 81, 332 (2021)]
- **4l+2l2ν** with the Run 2 datasets
- Mass: 200 – 2000 GeV
- Width: 0 to 15% of mass
- **Model-independent searches + 2HDM + Kaluza-Klein graviton**



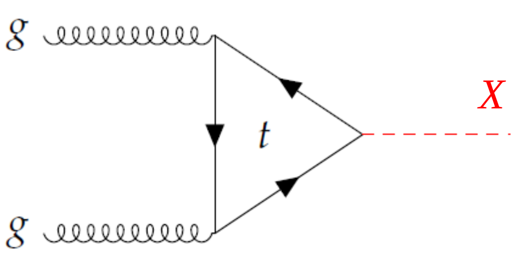



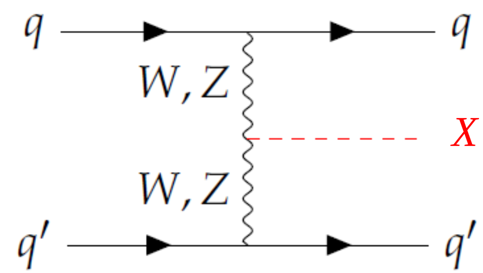
# Analysis strategy and status

- **Search for scalar resonances decaying to ZZ to 4 leptons (electron or muon)**




- **Production:**
  - gluon fusion (**ggF**)
  - vector boson fusion (**VBF**)



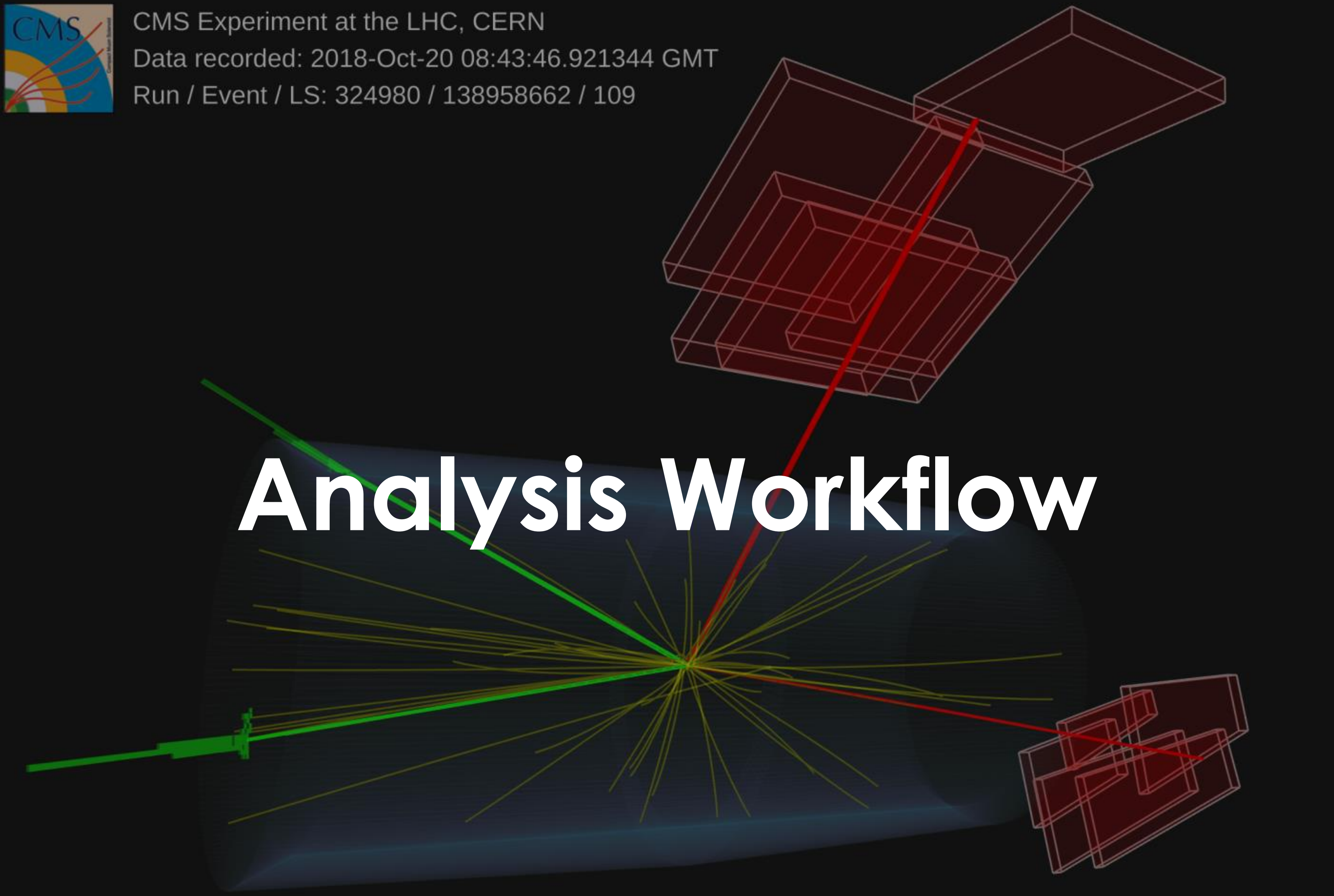




- **Resonance mass:**
  - 130 – 3000 GeV
- **Resonance width:**
  - 0 to 30% of mass

- **Model-independent searches: no specific physics model**

- **Full Run 2 datasets**
  - Integrated luminosity of 138 $\mathrm{fb}^{-1}$

- **Published as a CMS Physics Analysis Summary**
  - [CMS-PAS-HIG-24-002]
  - Will be submitted to a journal soon




# Analysis Workflow

# Event selection and categorization

## Trigger selection

- Trigger on leptons

| Datasets | SingleElectron/ EGamma | SingleMuon | DoubleEG | DoubleMuon | MuonEG |
|---|---|---|---|---|---|
| Triggers | single-e | single-μ | double-e<br>triple-e | double-μ<br>triple-μ | e+μ, e+2μ,<br>2e+μ |

## Event selection

**Lepton selection**

**Tight μ**: kinematic cuts, muon PF ID + tracker High-pT ID, iso
**Tight e**: kinematic cuts, dedicated ID and iso
**FSR** recovery

**Z selection**

Opposite sign, same flavor (ee or μμ)
12 GeV < $M_Z$ < 120 GeV

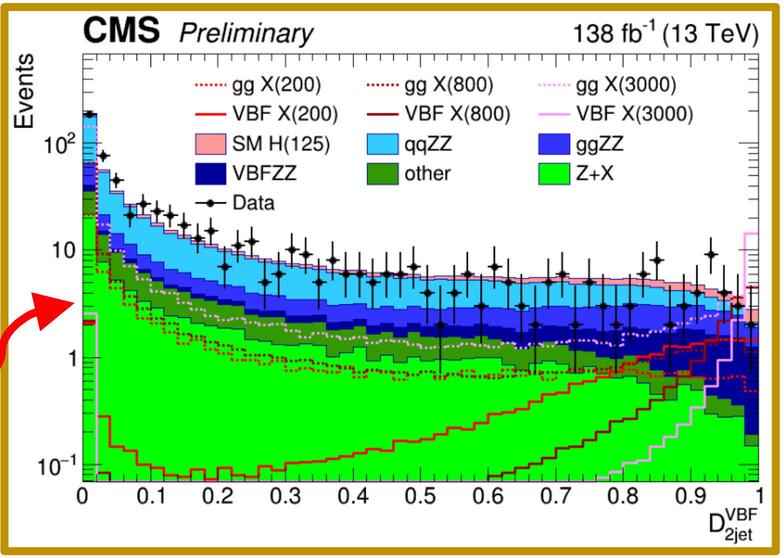
**ZZ selection**

$M_{ZZ}$ > 70 GeV, as well as other kinematic cuts
Three final states: **4μ, 4e, 2e2μ**

## Event categori-zation

Two categories: **VBF tagged and untagged**

- based on jet kinematics and $D_{2jet}^{VBF}$ computed from **Matrix Element Likelihood approach**

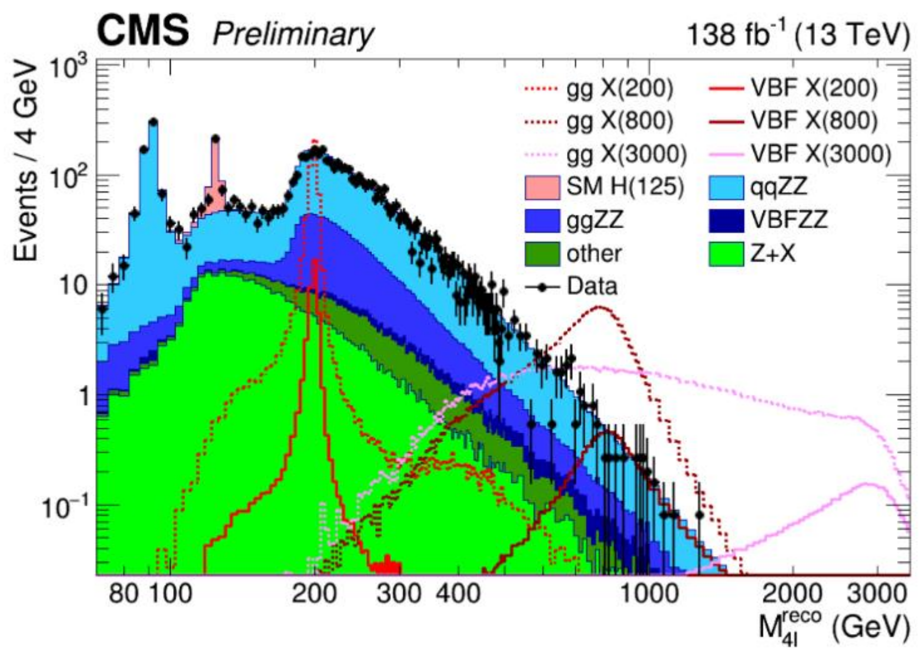


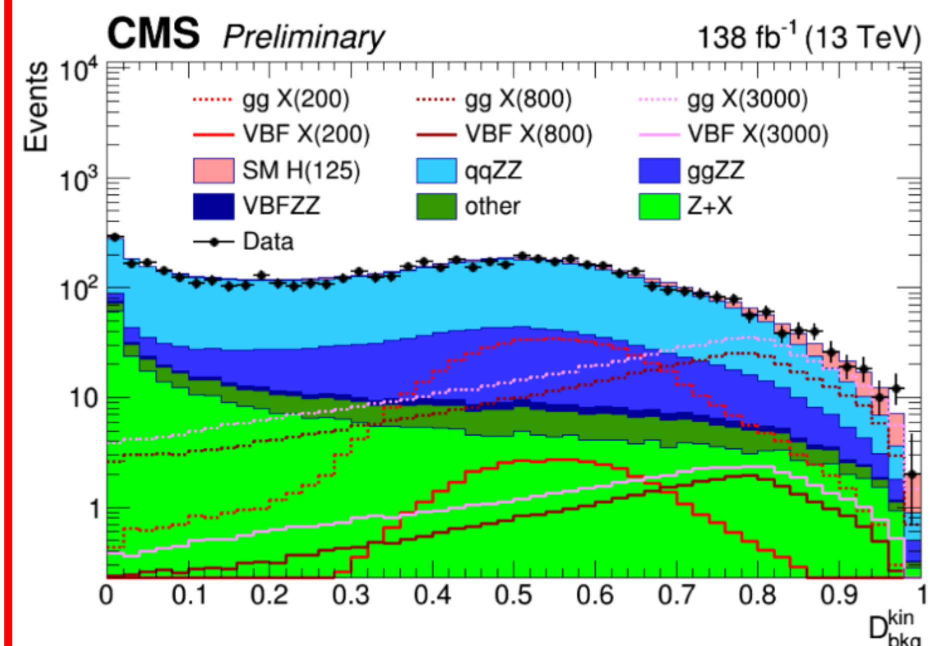
# Discriminating variables

Invariant mass of the 4-lepton final state: $M_{4l}^{reco}$

The discriminant built from MELA:

$$D_{bkg}^{kin} = \left[1 + \frac{P_{bkg}^{qqZZ}(\Omega^{4l}|m_{4l})}{P_{sig}^{ggH}(\Omega^{4l}|m_{4l})}\right]^{-1}$$




- All the processes are from **Monte Carlo (MC) simulation**, except **Z+X**.
- These distributions don't represent the statistical model we build.

# Signal modeling

Parametric approach to model all processes, because:

- Model independent: signal models parameterized on $M_X, \Gamma_X$.
- the statistics at high mass of backgrounds are low.

# Signal modeling

Parametric approach to model all processes, because:

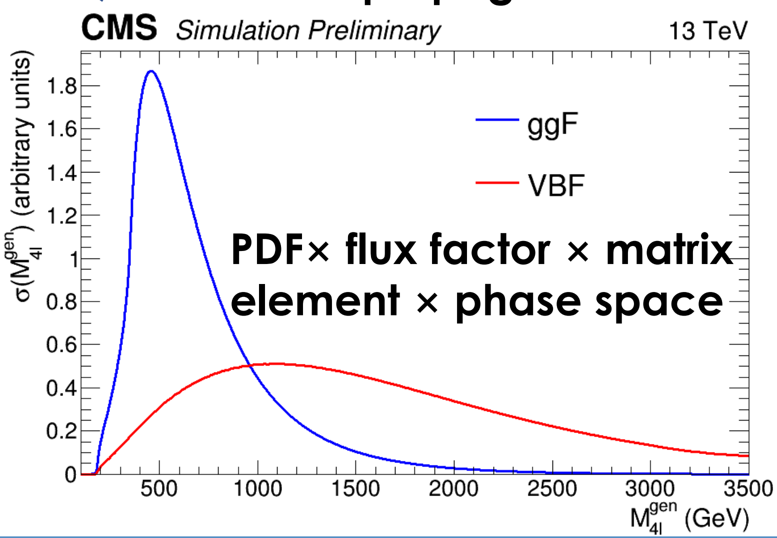
- Model independent: signal models parameterized on $M_X, \Gamma_X$.
- the statistics at high mass of backgrounds are low.

$$P(M_{4l}^{reco}, D_{bkg}^{kin}) = \{[P(M_{4l}^{gen}|m_X,\Gamma_X) \times eff(M_{4l}^{gen})] \otimes R(M_{4l}^{reco}|M_{4l}^{gen})\} \cdot P(D_{bkg}^{kin}|M_{4l}^{reco})$$

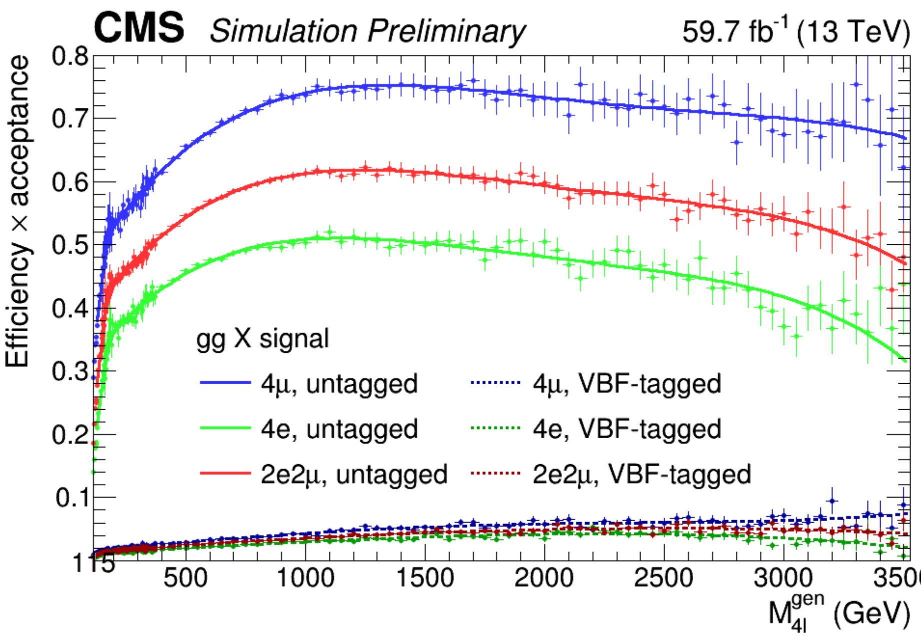
Lineshape:

$$P(m_H^{gen}|m_H^{pole},\Gamma_H) = \sigma(m_H^{gen}) \cdot \frac{2 m_{4l}^{gen} m_H^{pole}}{\left[(m_{4l}^{gen})^2 - (m_H^{pole})^2\right]^2 + (m_H^{pole}\Gamma_H)^2}$$

**propagator**




Signal efficiency: $eff(m_H^{gen})$ **from MC**



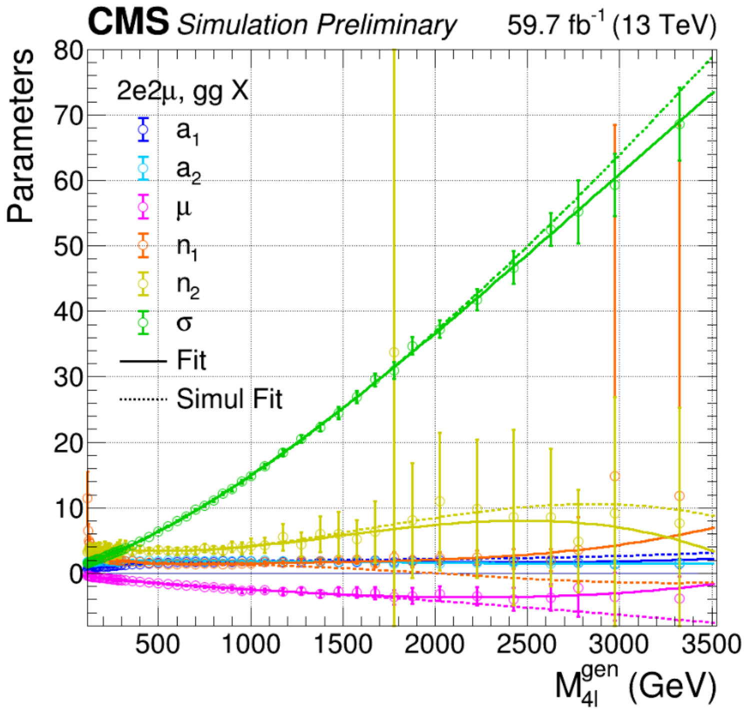
# Signal modeling

Parametric approach to model all processes, because:

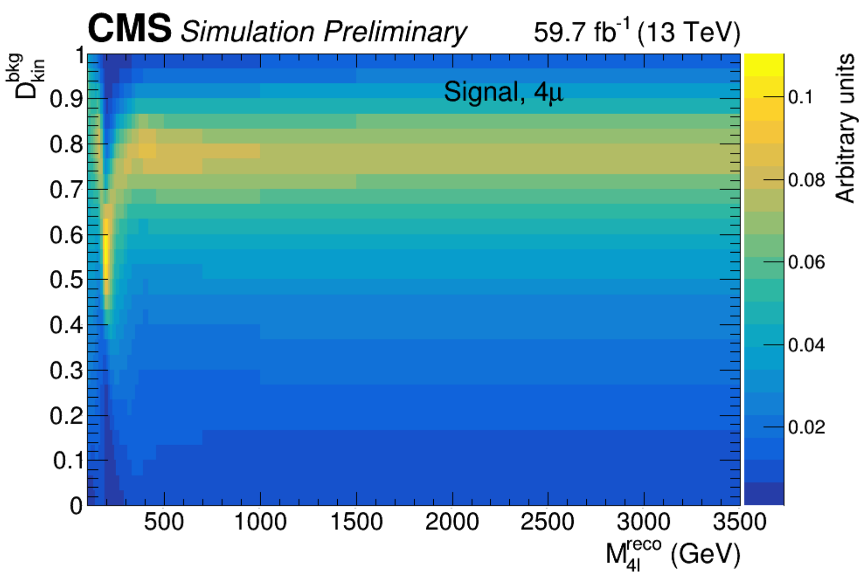
- Model independent: signal models parameterized on $M_X, \Gamma_X$.
- the statistics at high mass of backgrounds are low.

$$P(M_{4l}^{reco}, D_{bkg}^{kin}) = \{[P(M_{4l}^{gen}|m_X,\Gamma_X) \times eff(M_{4l}^{gen})] \otimes R(M_{4l}^{reco}|M_{4l}^{gen})\} \cdot P(D_{bkg}^{kin}|M_{4l}^{reco})$$

Signal resolution **from MC** : $R(m_H^{reco}|m_H^{gen}) = \mathrm{DCB}(m_H^{reco} - m_H^{gen}|\, \mu, \sigma, \alpha_1, n_1, \alpha_2, n_2)$

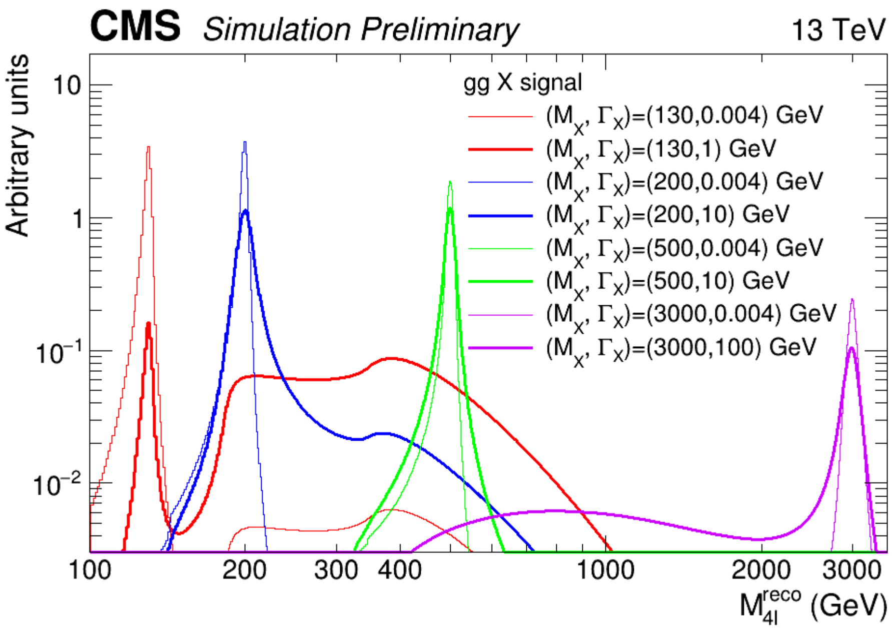


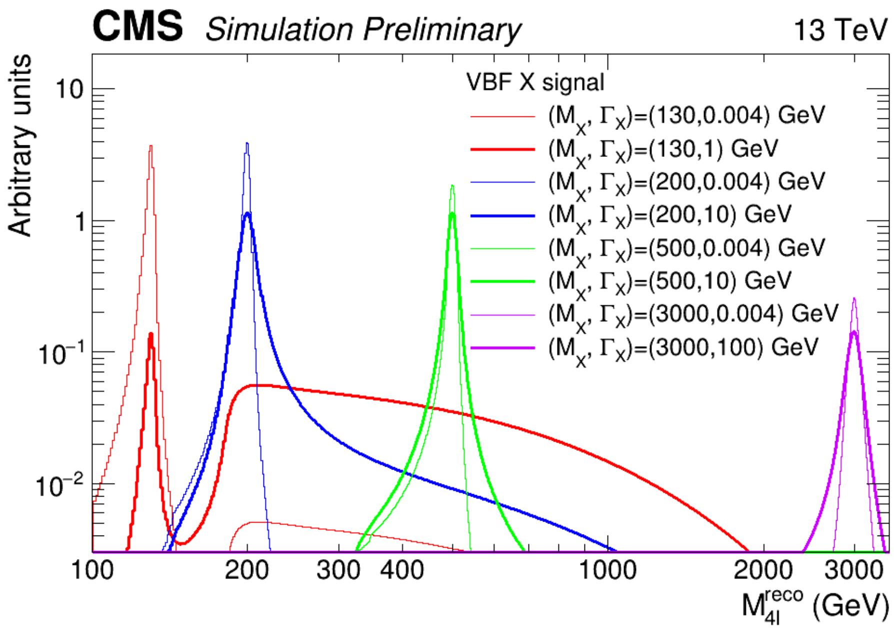

2D templates: $P(D_{bkg}^{kin}|m_{4l}^{reco})$ **from MC** as the 2D histograms



# Signal modeling

$$P(M_{4l}^{reco}, D_{bkg}^{kin}) = \{[P(M_{4l}^{gen}|m_X,\Gamma_X) \times eff(M_{4l}^{gen})] \otimes R(M_{4l}^{reco}|M_{4l}^{gen})\} \cdot P(D_{bkg}^{kin}|M_{4l}^{reco})$$



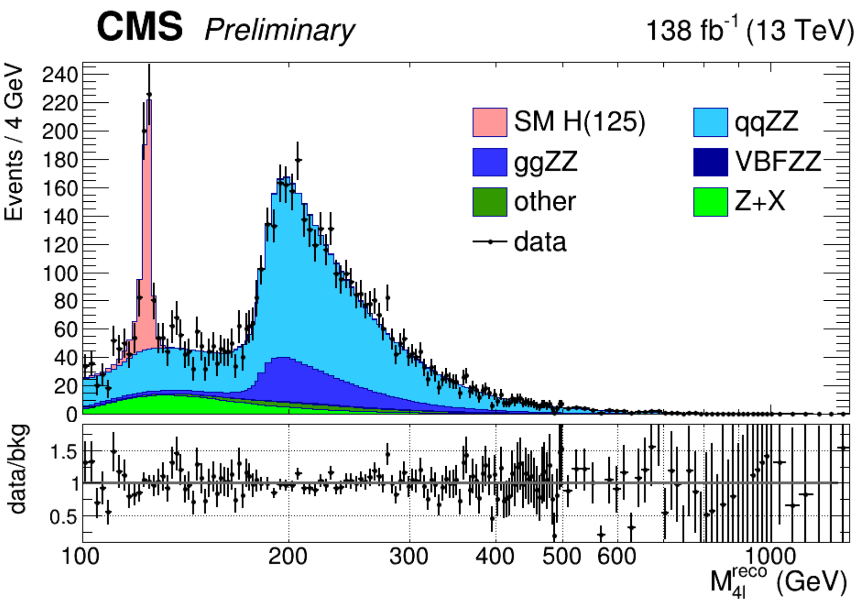

**Validated by comparing with shapes from MC simulation**

# Backgrounds and interferences

## Background modeling

- SM Higgs boson: the same as signals, with $M_H$ = 125 GeV and $\Gamma_H$ = 4.1 MeV
- qqZZ (dominating): by MC (NLO in QCD, with NNLO QCD k-factor and NLO EW k-factor)
- ggZZ: by MC (LO, with NNLO k-factor)
- Triboson and tops+bosons, VBFZZ: by MC
- Z+X: estimated from control region data using a data-driven method

**$M_{4l}^{reco}$ is parameterized with empirical functions**



# Backgrounds and interferences

## Background modeling

SM Higgs boson: the same as signals, with $M_H$ = 125 GeV and $\Gamma_H$ = 4.1 MeV

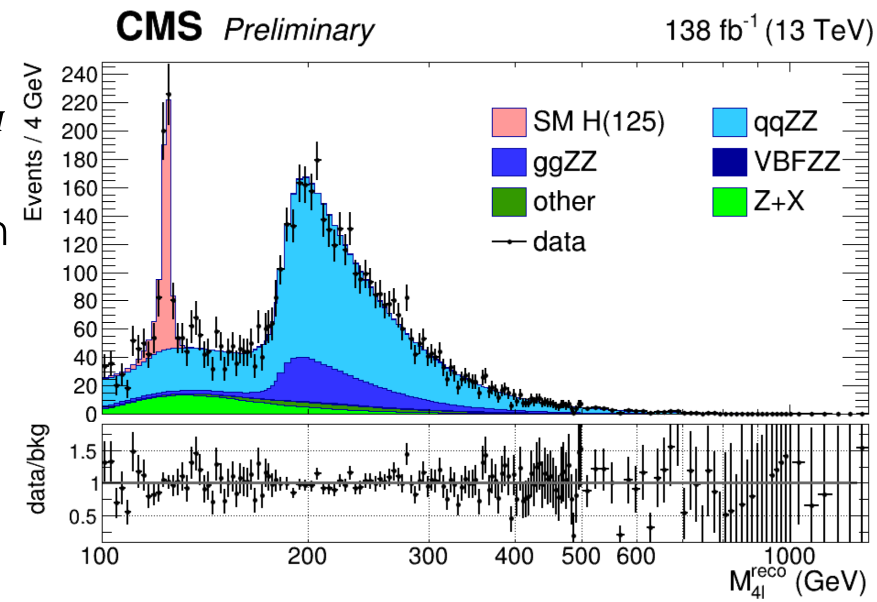
qqZZ (dominating): by MC (NLO in QCD, with NNLO QCD k-factor and NLO EW k-factor)

ggZZ: by MC (LO, with NNLO k-factor)

Triboson and tops+bosons, VBFZZ: by MC

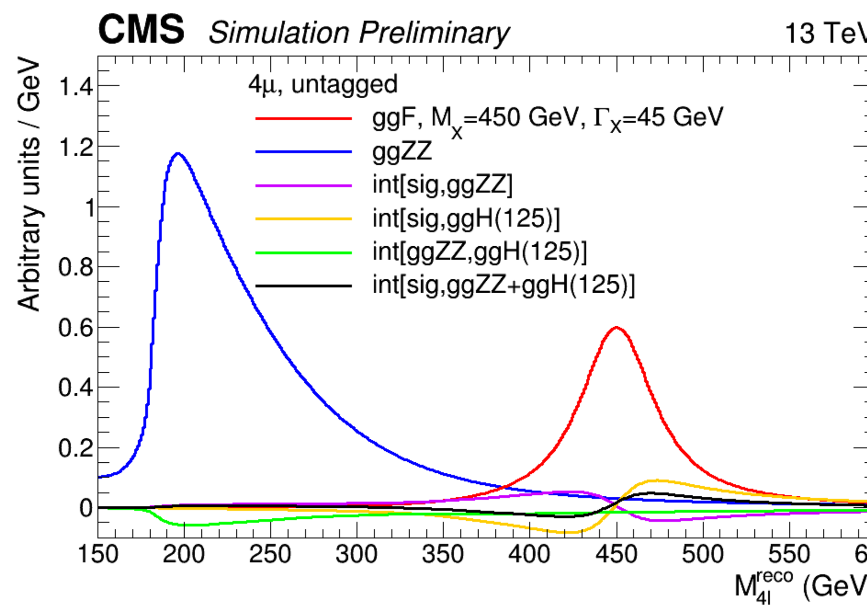
Z+X: estimated from control region data using a data-driven method

**$M_{4l}^{reco}$ is parameterized with empirical functions**

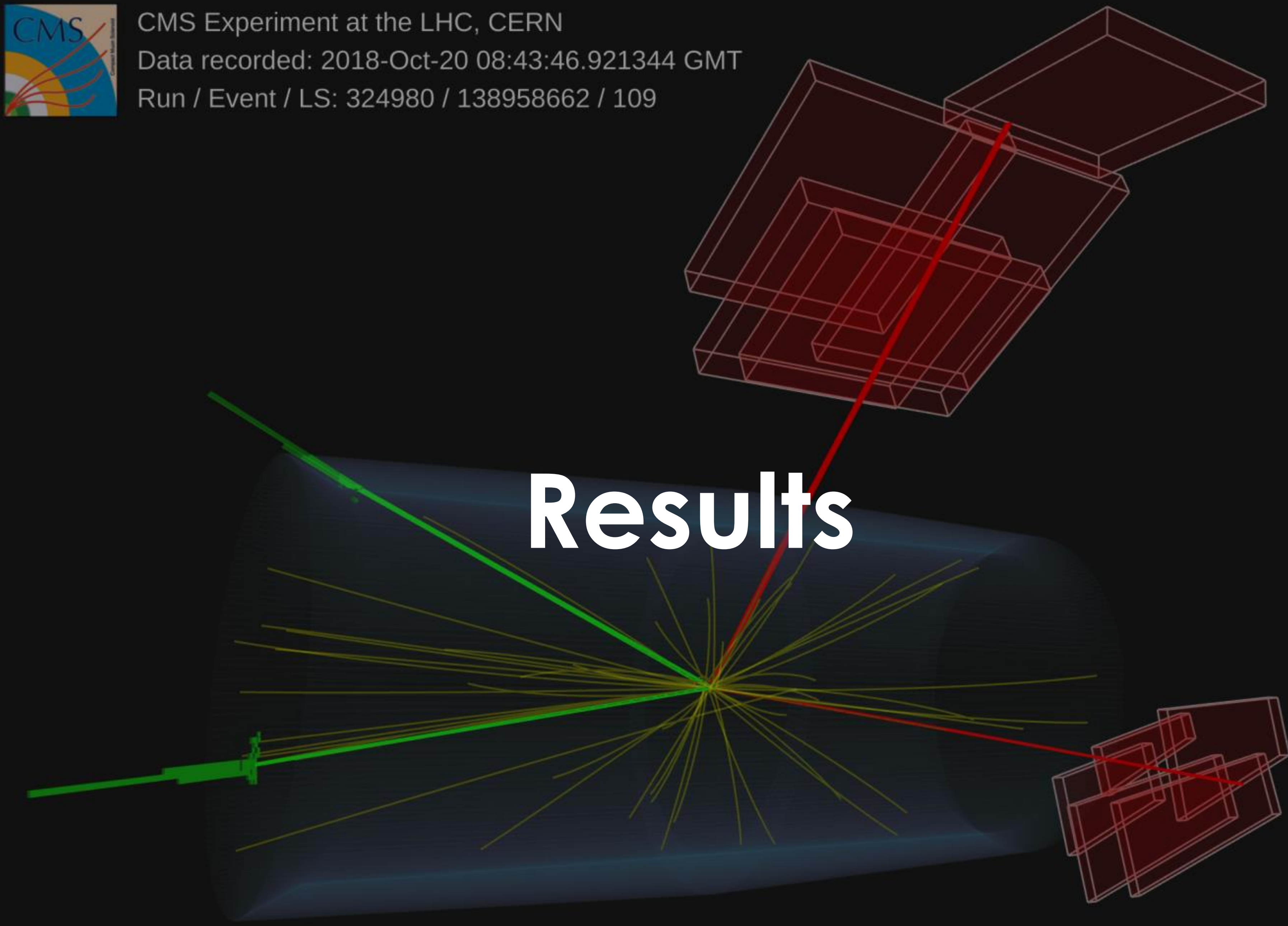



## Interference modeling

- Three components for each production: **signal v.s. SM Higgs boson v.s. non-resonant bkg**
- **Amplitudes** from signal and background models
- **Phases** from generators and kinematics



# Results

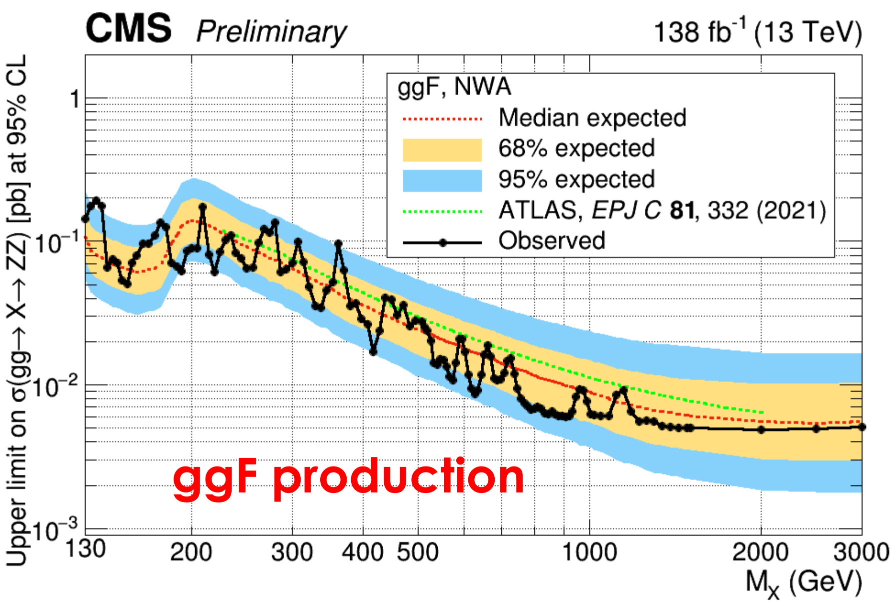

CMS Experiment at the LHC, CERN
Data recorded: 2018-Oct-20 08:43:46.921344 GMT
Run / Event / LS: 324980 / 138958662 / 109

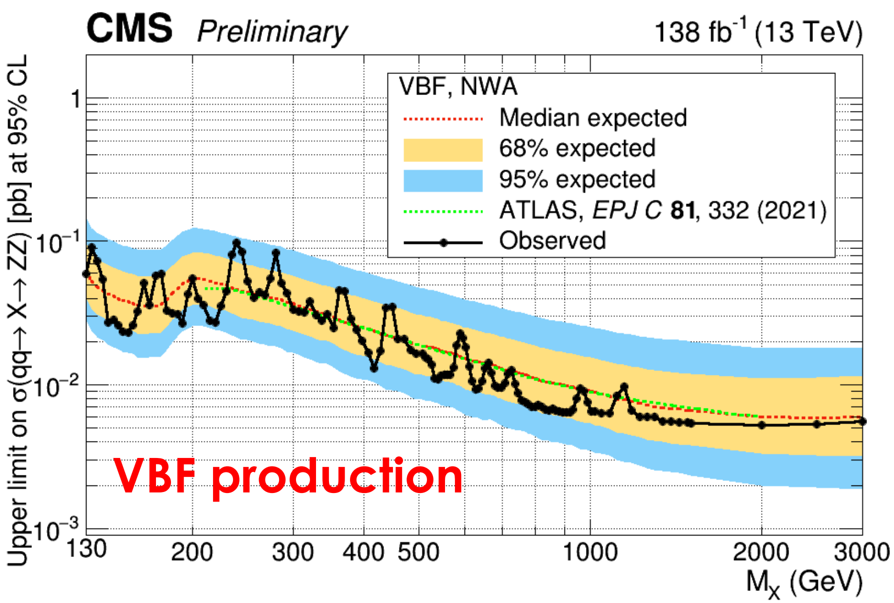
# Statistical inferences

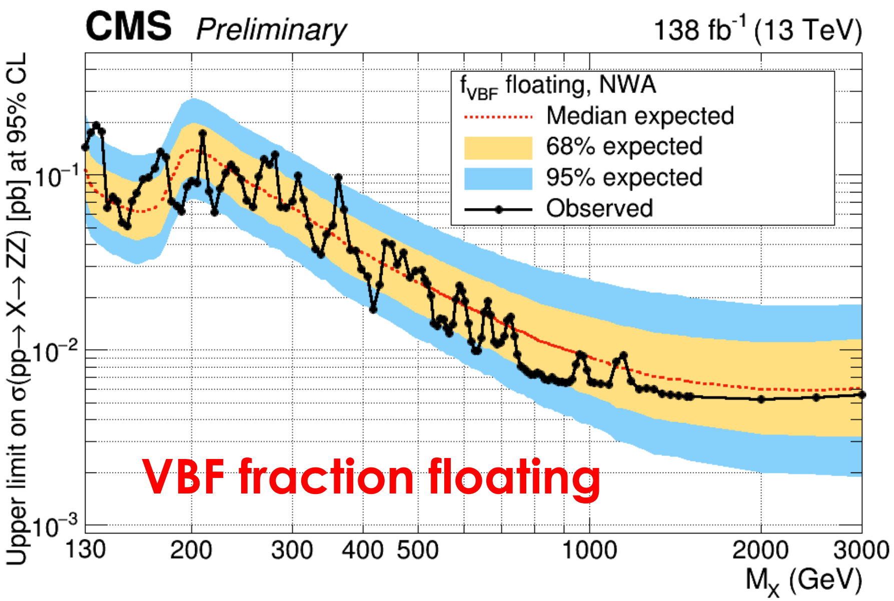
- Parameter of interest: signal strength $\boldsymbol{r} = \frac{\boldsymbol{\sigma(pp \to X \to ZZ)}}{\boldsymbol{\sigma_{\text{model}}(pp \to X \to ZZ)}}$
  - $r_{ggF} = r \cdot (1 - f_{VBF})$
  - $r_{VBF} = r \cdot f_{VBF}$
  - $P_{tot} = r_{ggF/VBF} \cdot P_{sig} + \sqrt{r_{ggF/VBF}} \cdot P_{int} + P_{bkg}$

- Statistical method:
  - Unbinned likelihood fits on $(m_{4l}^{reco}, D_{bkg}^{kin})$
  - Since no significant excess is observed, we use the **CLs method** to compute **upper limits at 95% confidence level** on signal strength $r$ with asymptotic formulae, as a function of $M_X$ and $\Gamma_X$, for different production mechanisms.

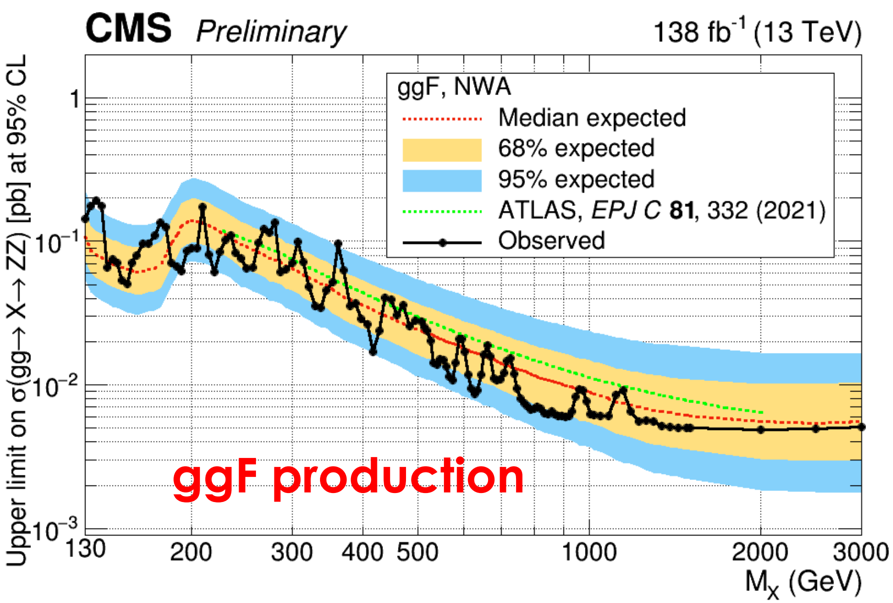
# Results: narrow width assumption

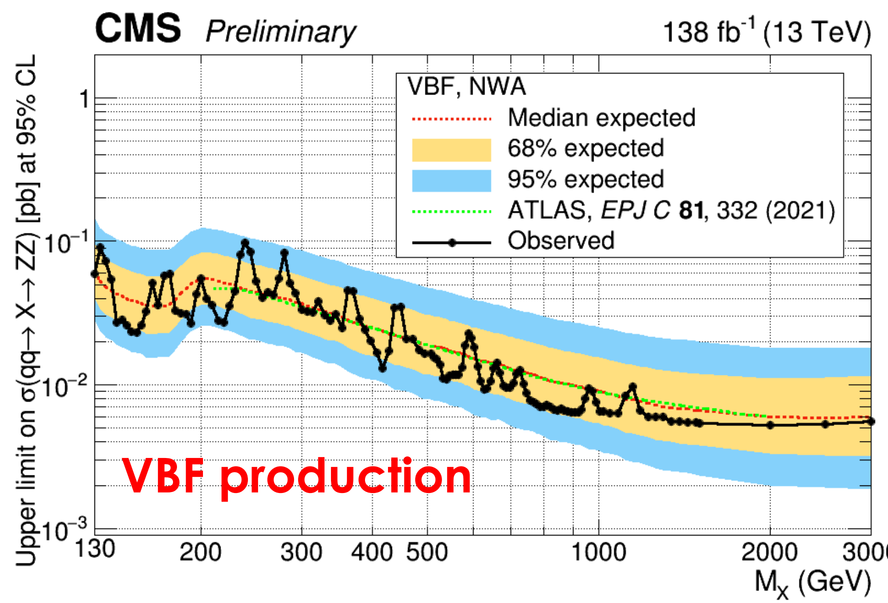


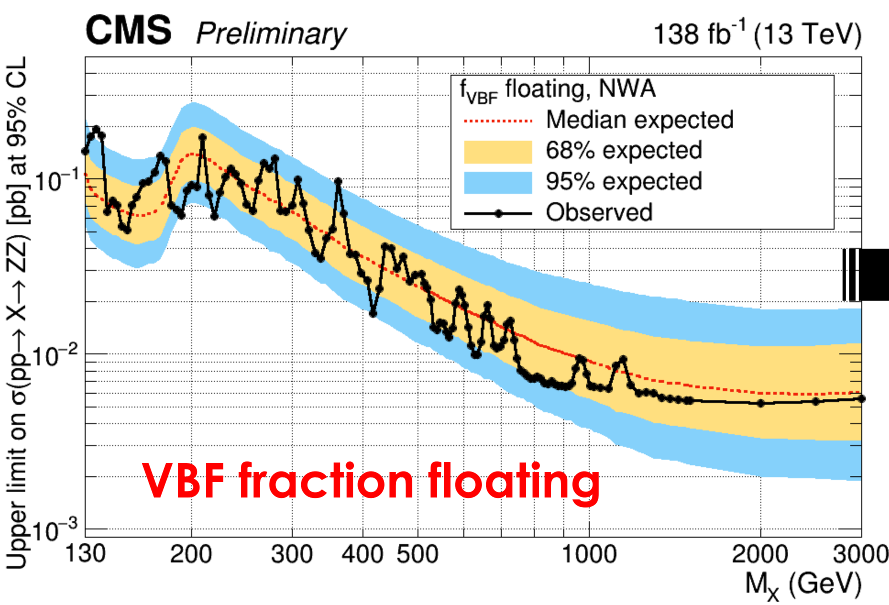



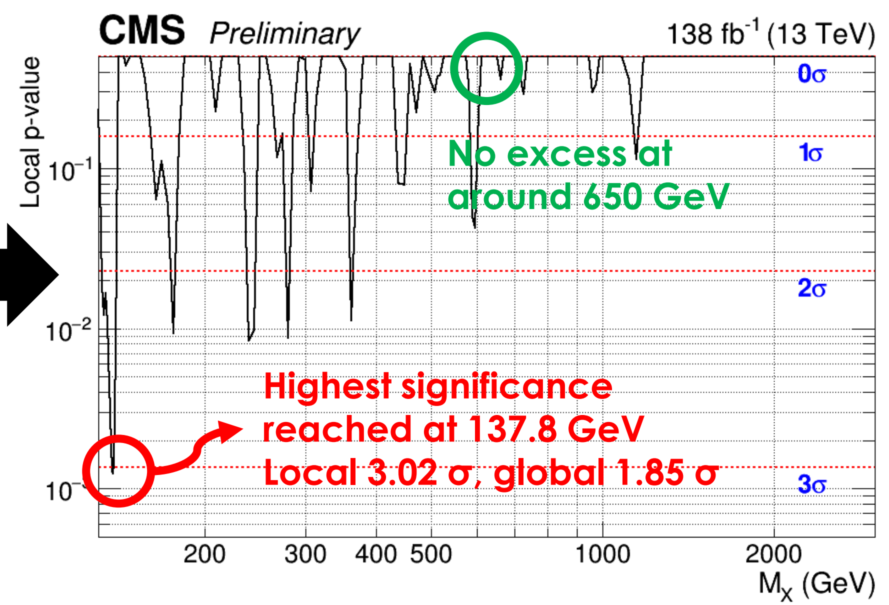



# Results: narrow width assumption



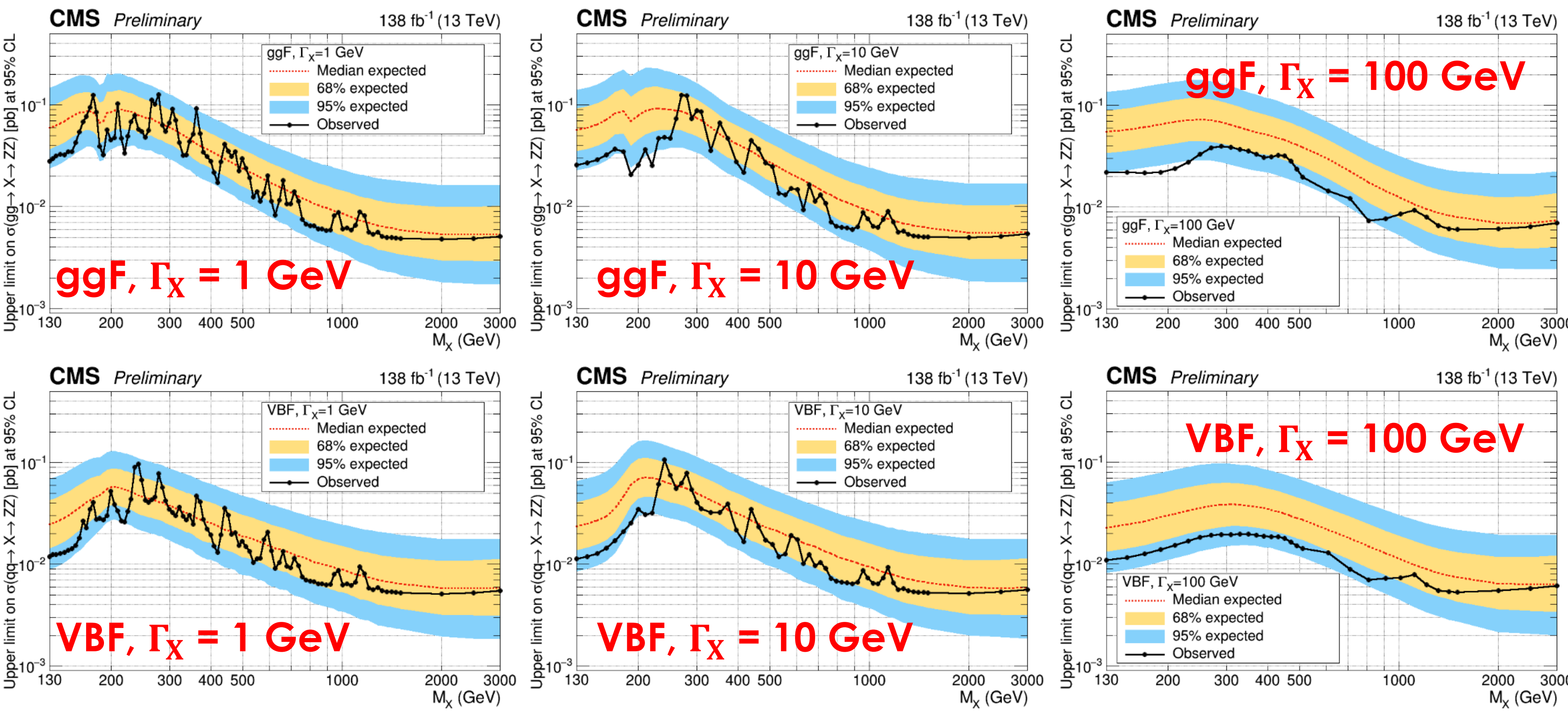



# Results: various width assumptions

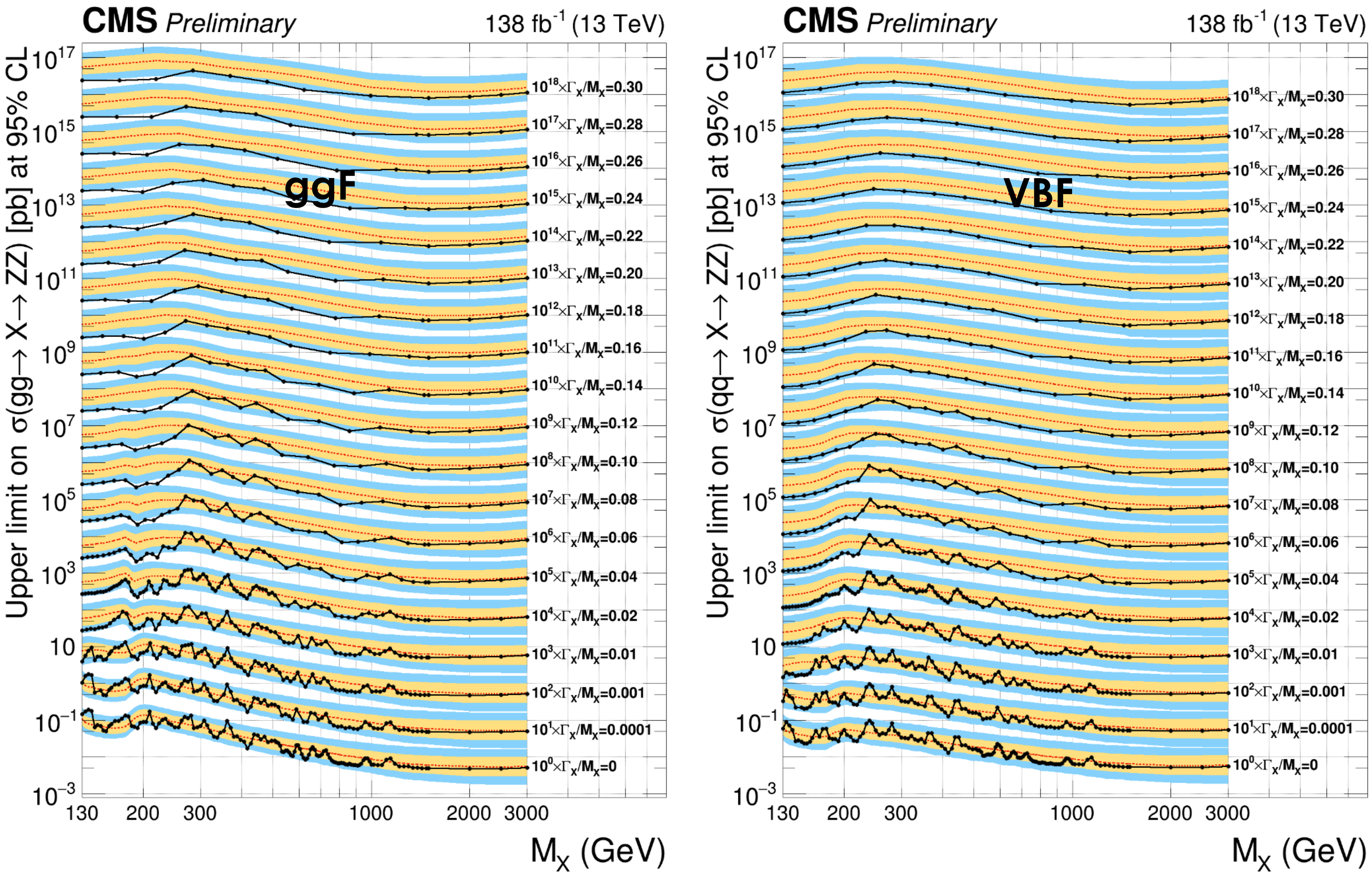
- $\Gamma_X$ = 1, 10, 100 GeV



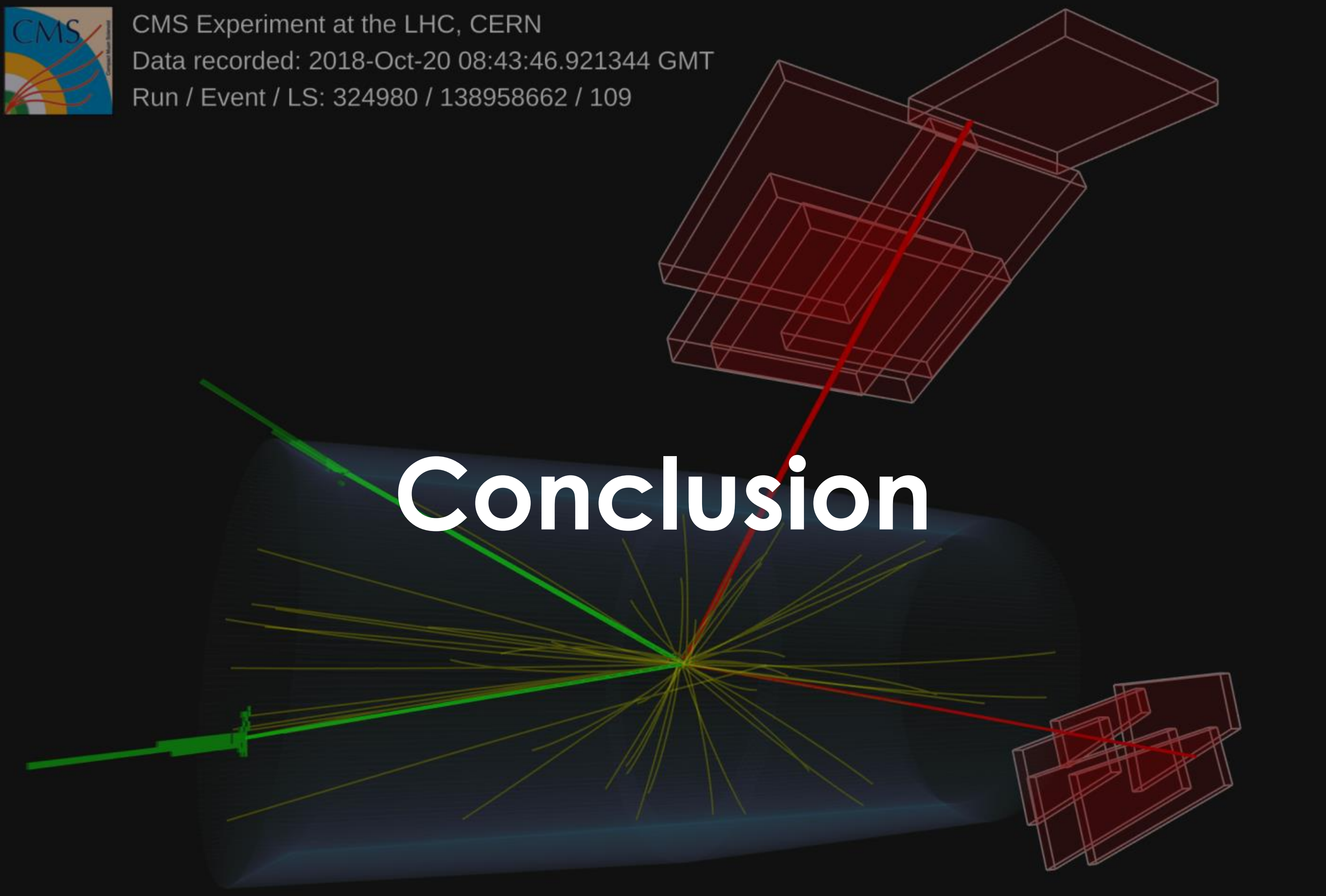

- No additional excess is observed.
- When the signal is widely spread, the observed results are below the expected ones, because the expected background yield is higher than data.

# Results: various width assumptions

- Scan both $M_X$ and $\Gamma_X/M_X$



# Conclusion


CMS Experiment at the LHC, CERN
Data recorded: 2018-Oct-20 08:43:46.921344 GMT
Run / Event / LS: 324980 / 138958662 / 109

# Conclusion

- Searches for **heavy scalar resonances decaying into 4 leptons** are performed with the Run 2 dataset collected from the CMS detector.
- The **model-independent approach** is used.
- Upper limits at 95% confidence level on $\sigma(pp \rightarrow X \rightarrow ZZ)$ are computed, as a function of $M_X$, $\Gamma_X$, for different production mechanisms.
- No excess is observed, except at 137.8 GeV with narrow width assumption, corresponding to a global significance of 1.85 σ.

**Future plans**

- Will be submitted to a journal soon.
- Combination with the **2l2q and 2l2v** final states with the full Run 2 dataset:
  - 2l2q ongoing (also with contributions from LLR).
  - 2l2v at a very young stage.
- Perform the searches with **Run 3 data**, trying to confirm or exclude the existing excesses.

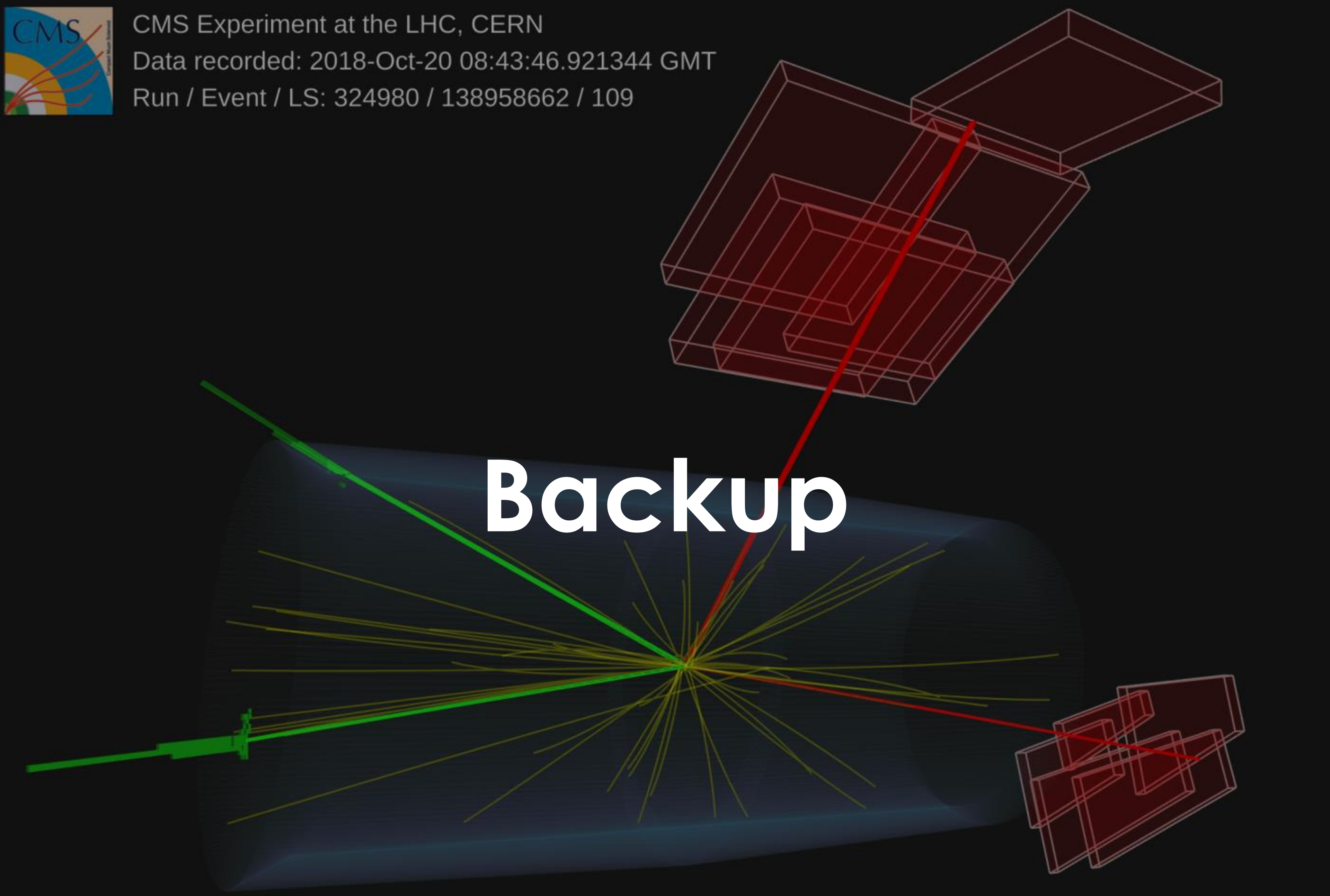
# Conclusion

- Searches for **heavy scalar resonances decaying into 4 leptons** are performed with the Run 2 dataset collected from the CMS detector.
- The **model-independent approach** is used.
- Upper limits at 95% confidence level on $\sigma(pp \to X \to ZZ)$ are computed, as a function of $M_X$, $\Gamma_X$, for different production mechanisms.
- No excess is observed, except at 137.8 GeV with narrow width assumption, corresponding to a global significance of 1.85 σ.

**Future plans**

- Will be submitted to a journal soon.
- Combination with the **2l2q and 2l2v** final state with the full Run 2 dataset:
  - 2l2q ongoing (also with contributions from LLR).
  - 2l2v at a very young stage.
- Perform the searches with **Run 3 data**, trying to confirm or exclude the existing excesses

**Thanks for your attention!**


CMS Experiment at the LHC, CERN
Data recorded: 2018-Oct-20 08:43:46.921344 GMT
Run / Event / LS: 324980 / 138958662 / 109

# Backup

# Object selection

| Particle type | Selection requirements |
|---|---|
| Electrons | $p_T^e > 7$ GeV<br>$\|\eta_e\| < 2.5$<br>$d_{xy} < 0.5$ cm<br>$d_z < 1.0$ cm<br>$\|SIP_{3D}\| < 4$<br>ID from BDT score<br>Isolation from BDT score |
| Muons | Global or Tracker Muon<br>Discard Standalone Muon tracks<br>$p_T^\mu > 5$ GeV<br>$\|\eta_\mu\| < 2.4$<br>$d_{xy} < 0.5$ cm<br>$d_z < 1.0$ cm<br>$\|SIP_{3D}\| < 4$<br>PF muon ID if $p_T^\mu < 200$ GeV<br>Tracker High-$p_T$ muon ID if $p_T^\mu > 200$ GeV<br>$\mathcal{I}^\mu < 0.35$ |
| Photons | $p_T^\gamma > 2$ GeV<br>$\|\eta_\gamma\| < 2.4$<br>$\mathcal{I}^\gamma < 1.8$<br>$\Delta R(l,\gamma) < 0.5$<br>$\Delta R(l,\gamma)/(p_T^\gamma)^2 < 0.012$ GeV$^{-2}$ |
| Jets | $p_T^{jet} > 30$ GeV<br>$\|\eta_{jet}\| < 4.7$<br>$\Delta R(l/\gamma, jet) > 0.4$<br>Cut-based ID<br>Jet pileup ID |

# ZZ candidate selection

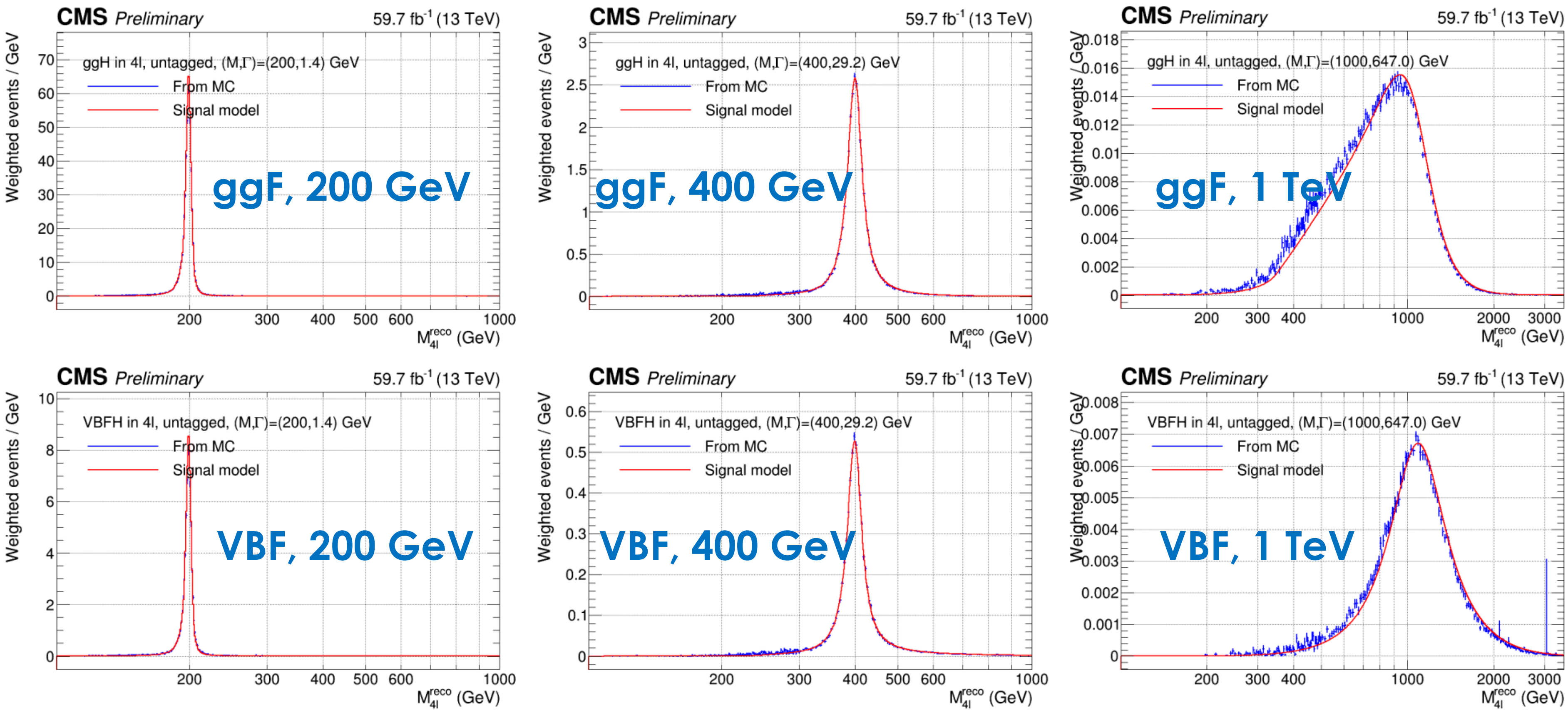
- Any two leptons have $\Delta R(l_1, l_2) > 0.02$.
- The Z candidate with mass closer to Z mass is defined as $Z_1$, the other $Z_2$. $Z_1$ has $M_{Z_1} > 40$ GeV.
- The highest lepton pT is > 20 GeV, second highest > 10 GeV.
- Any two leptons with opposite sign have $M_{l^+l^-} > 4$ GeV.
- In the 4μ and 4e final state, swap two positive leptons to form another two Z candidates, and the one with mass closer to Z mass is is defined as $Z_a$, the other $Z_b$. The ZZ candidate cannot satisfy both $|M_{Z_a} - M_Z| < |M_{Z_1} - M_Z|$ and $M_{Z_b} < 12$ GeV.
- $M_{ZZ} > 70$ GeV.

# Signal model validation

$$P(M_{4l}^{reco}, D_{bkg}^{kin}) = \{[P(M_{4l}^{gen}|m_X, \Gamma_X) \times eff(M_{4l}^{gen})] \otimes R(M_{4l}^{reco}|M_{4l}^{gen})\} \cdot P(D_{bkg}^{kin}|M_{4l}^{reco})$$

**Validation: check if the signal model describes the MC simulation**

- MC produced with CPS, reweighted to the Breit-Wigner lineshape based on MELA.
- The mass, width and cross section are set to be the same as the MC.

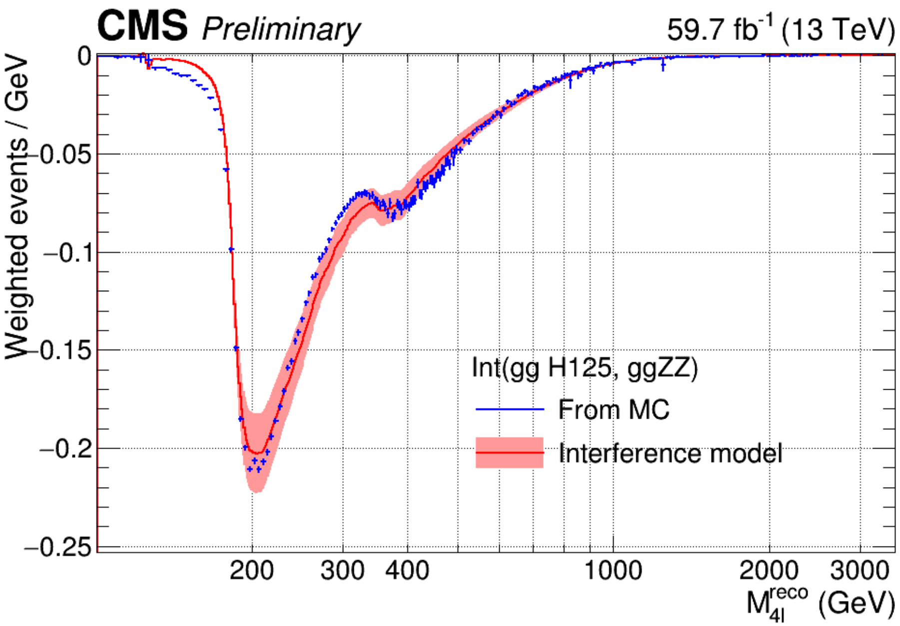


# Interference model validation

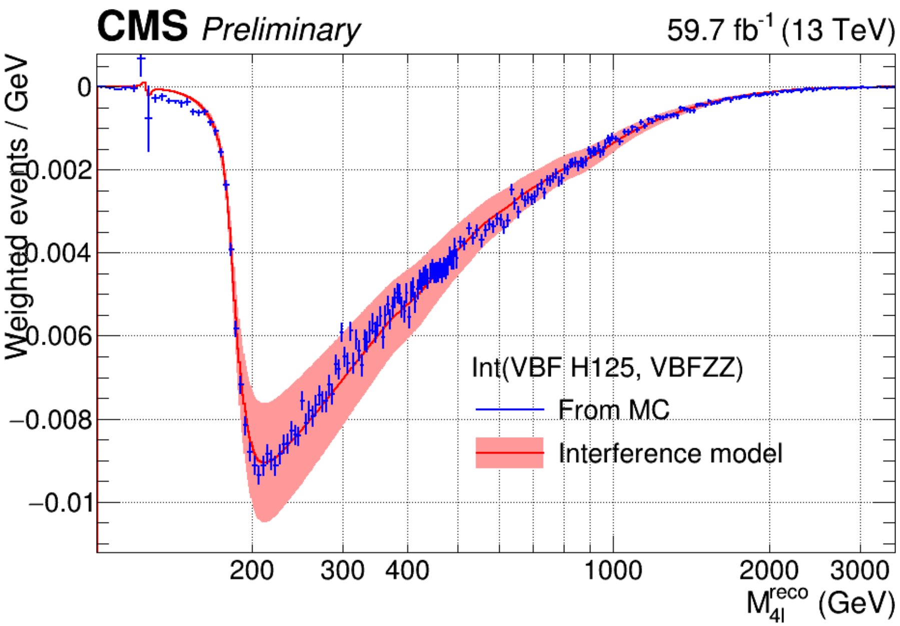
**Validation: SM Higgs boson v.s. non-resonant backgrounds (ggZZ or VBFZZ)**

- Compare to MC simulation: off-shell production samples with weights from MELA to model the H125-background interference

**ggH125 v.s. ggZZ**




**VBFH125 v.s. VBFZZ**



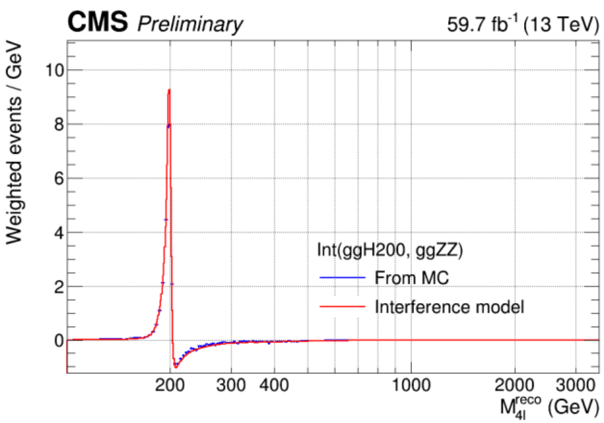

- Uncertainties to take into account the potential difference (total yield and category migration) between MC and the interference model

# Interference model validation

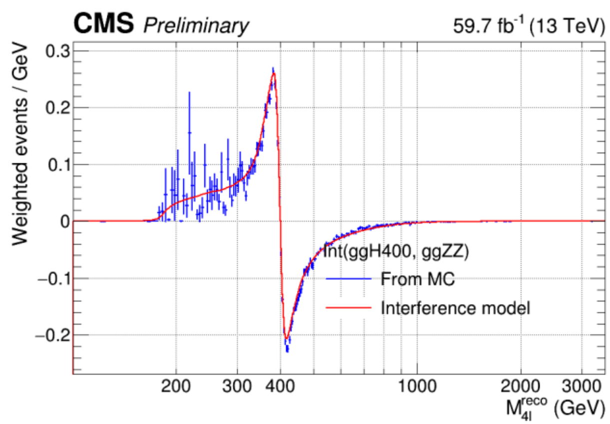
## Validation: signals v.s. non-resonant backgrounds (ggZZ or VBFZZ)

- Compare to MC simulation: signal samples with weights from MELA to model the interferences

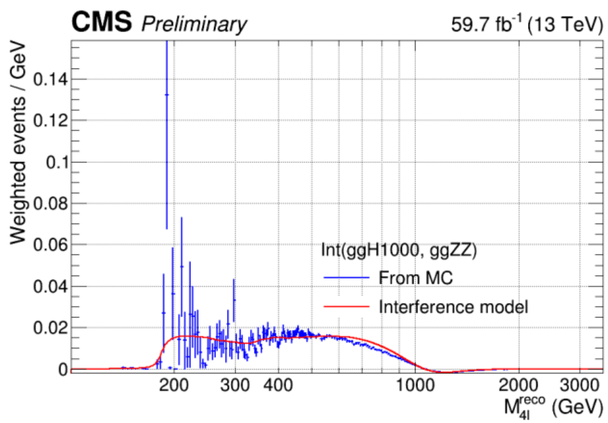
**ggH200 v.s. ggZZ**




**ggH400 v.s. ggZZ**

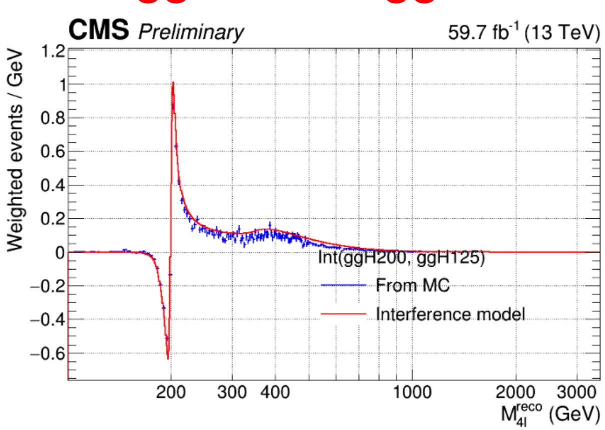



**ggH1000 v.s. ggZZ**

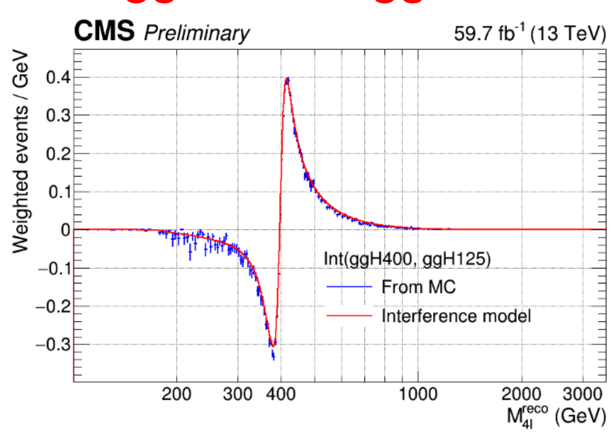



## Validation: signals v.s. SM Higgs boson

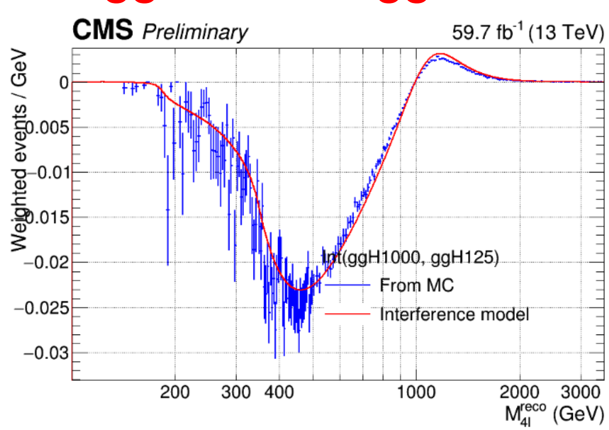
**ggH200 v.s. ggH125**




**ggH400 v.s. ggH125**




**ggH1000 v.s. ggH125**






VBF in backup

# Systematic uncertainties

| Source | 2016 | 2017 | 2018 | Effects | Affected processes |
|---|---|---|---|---|---|
| Experimental uncertainties | | | | | |
| Luminosity | 1.2% | 2.3% | 2.5% | norm | all except Z +X |
| e efficiency | 3-10% | 3-9% | 3-9% | norm | all except Z +X |
| $\mu$ efficiency | 1-2% | 1-2% | 1-2% | norm | all except Z +X |
| e($\mu$) energy scale | 0.15(0.03)% | | | shape | $\mathcal{R}(M_{4\ell}^{\text{reco}} \mid M_{4\ell}^{\text{gen}})$ |
| e($\mu$) energy resolution | 10(3)% | | | shape | $\mathcal{R}(M_{4\ell}^{\text{reco}} \mid M_{4\ell}^{\text{gen}})$ |
| jet energy scale | $\approx$ 1% | | | norm | all except Z +X |
| jet energy resolution | $\approx$ 1% | | | norm | all except Z +X |
| jet b-tagging efficiency | 0.1% | | | norm | all except Z +X |
| pileup reweighting | 0-1% | | | norm | All except Z +X |
| L1 prefiring | 0-1% | | | norm | All except Z +X |
| ggF interference | 9-11% | | | norm | interferences |
| VBF interference | 13-18% | | | norm | interferences |
| Z + X (4$\mu$) | 30% | 30% | 30% | norm | Z +X |
| Z + X (4e) | 31% | 31% | 30% | norm | Z +X |
| Z + X (2e2$\mu$) | 31% | 30% | 30% | norm | Z +X |
| Theoretical uncertainties | | | | | |
| BR(X $\to$ ZZ) | 2% | | | norm | signals |
| QCD scale | 1-15% | | | shape: qqZZ<br>norm: others | all except Z +X |
| PDF | 0.1-7% | | | shape: qqZZ<br>norm: others | all except Z +X |
| $\alpha_S$ | 0.1-7% | | | norm | all except Z +X |
| underlying events | 0.4-10% | | | norm | all except Z +X |
| qqZZ $K$ factor | 0.1-30% | | | shape | qqZZ |
| ggZZ $K$ factor | 10% | | | norm | ggZZ |

# Background-only fits