Searches for rare Higgs boson production processes with the CMS detector

Tiziano Bevilacqua (UZH, PSI)
On behalf of the CMS Collaboration

PAUL SCHERRER INSTITUT PSI

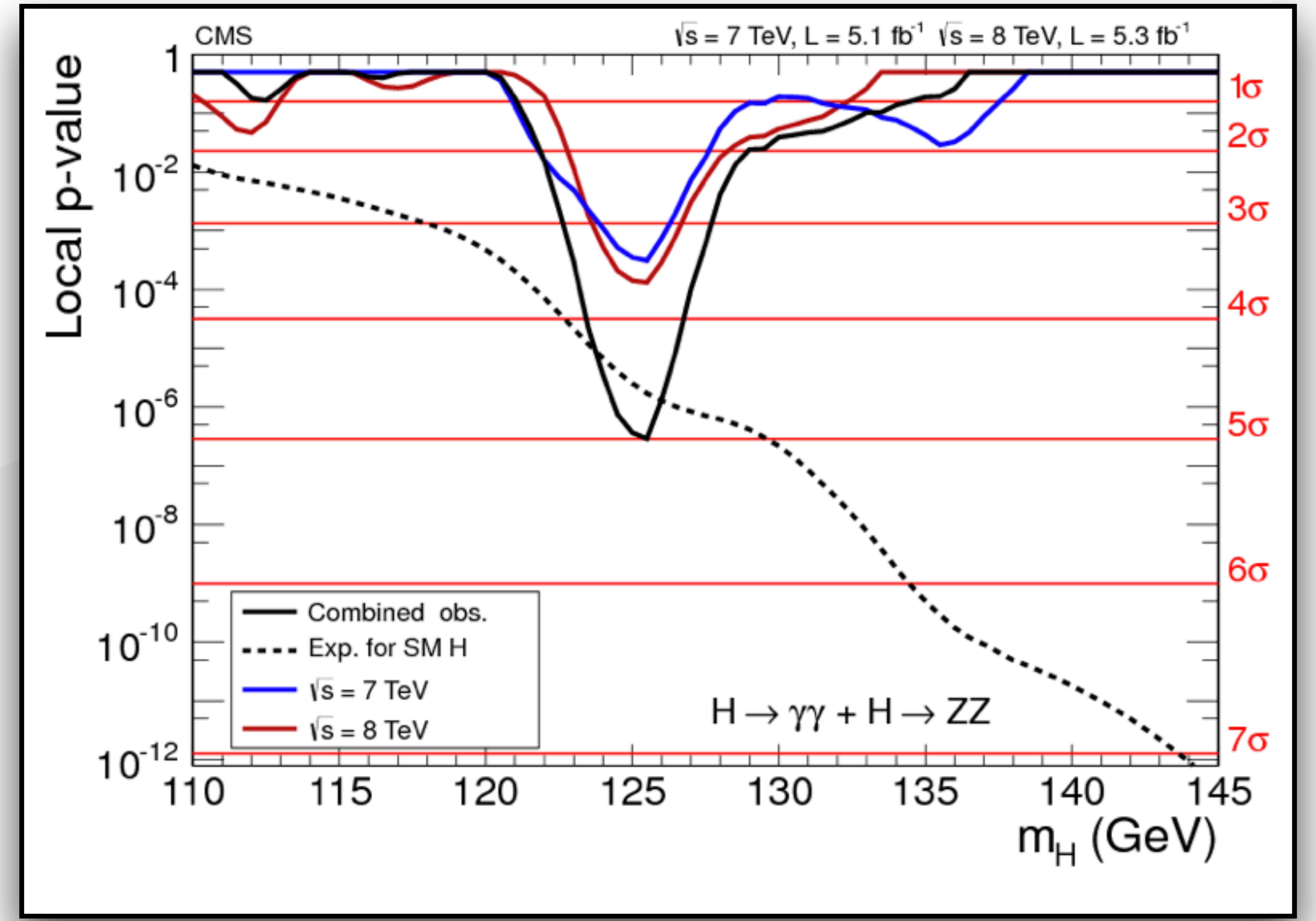
Outline

- **State of the art**:
  - Where we stand
- **Probing the Yukawa sector:**
  - H+c.
  - H+b.
- **Probing vector bosons couplings:**
  - WWHH production.
- **Conclusions.**

Where we stand 10+ years after

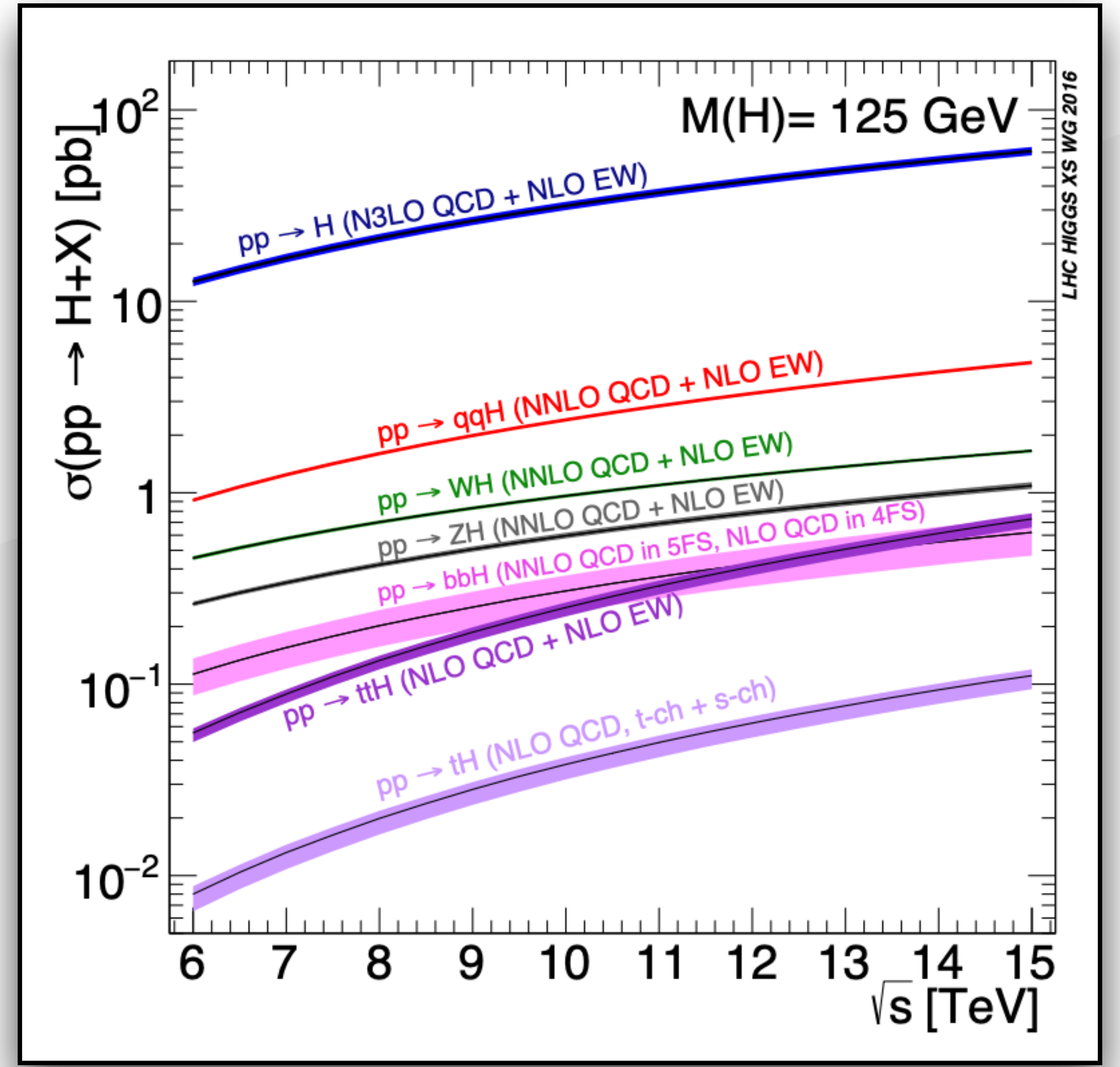
- **Run I:** Discovery of a boson in 2012 compatible with the scalar sector of the Standard Model.
  - Separate observation of the **bosonic** decay channels $O(15\%)$ and production in the most abundant channels.
  - Spin and parity compatible with $J^P = 0^+$.

[Phys. Lett. B 716 (2012) 30-61]



Where we stand 10+ years after

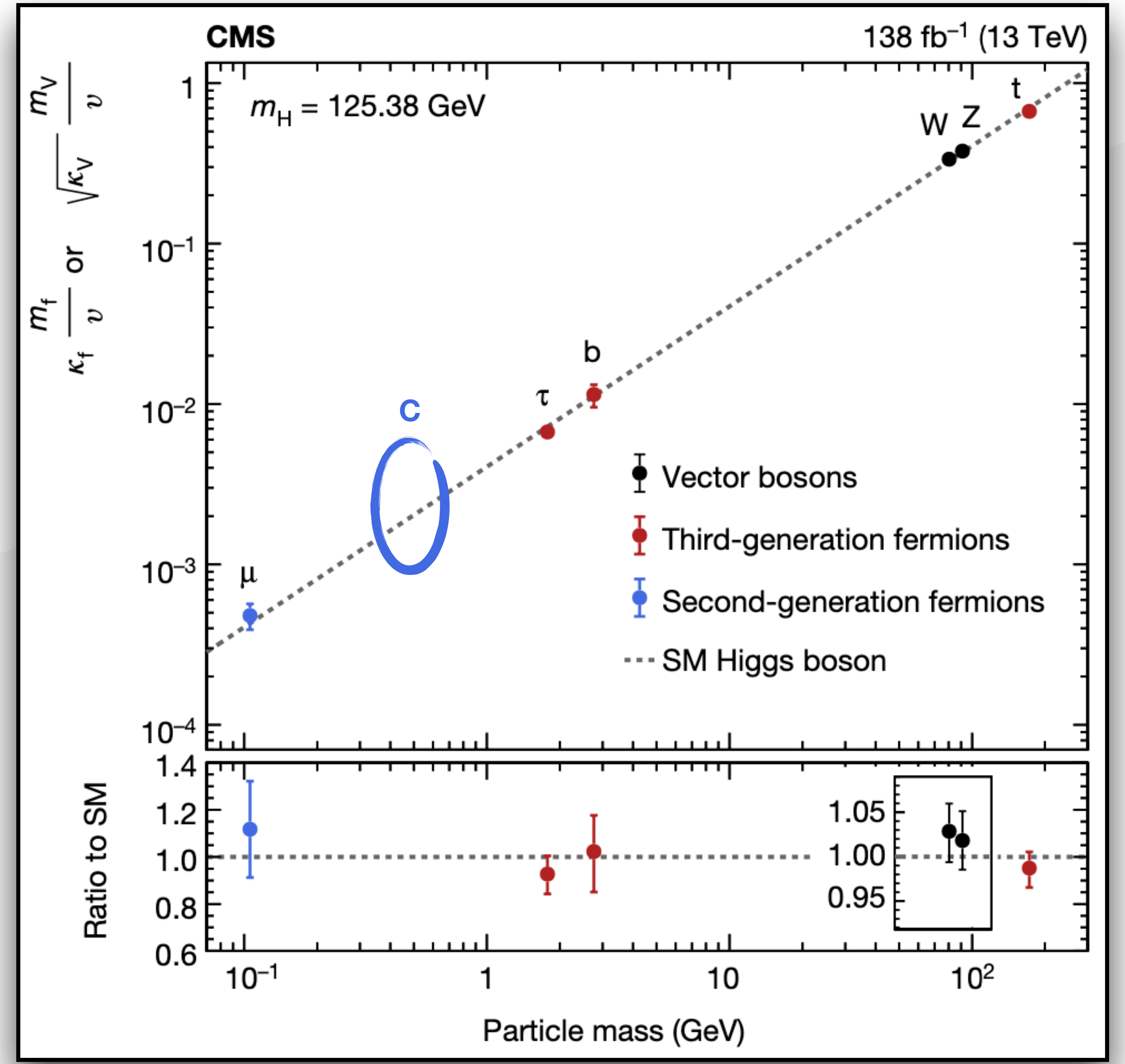
- **Run I:** Discovery of a boson in 2012 compatible with the scalar sector of the Standard Model.
  - Separate observation of the **bosonic** decay channels $O(15\%)$ and production in the most abundant channels.
  - Spin and parity compatible with $J^P = 0^+$.
- **Run II:** Firmly established the existence of Yukawa couplings.
  - Separate observation of **subdominant production modes:**
    $\Rightarrow$ **Vector Boson Fusion** (VBF),
    $\Rightarrow$ **Higgs-strahlung**, associated with W^{+-} and Z^0,
    $\Rightarrow$ t **and** $t\bar{t}$ **associated production**.



CERNYR-2017-002-M

Where we stand 10+ years after

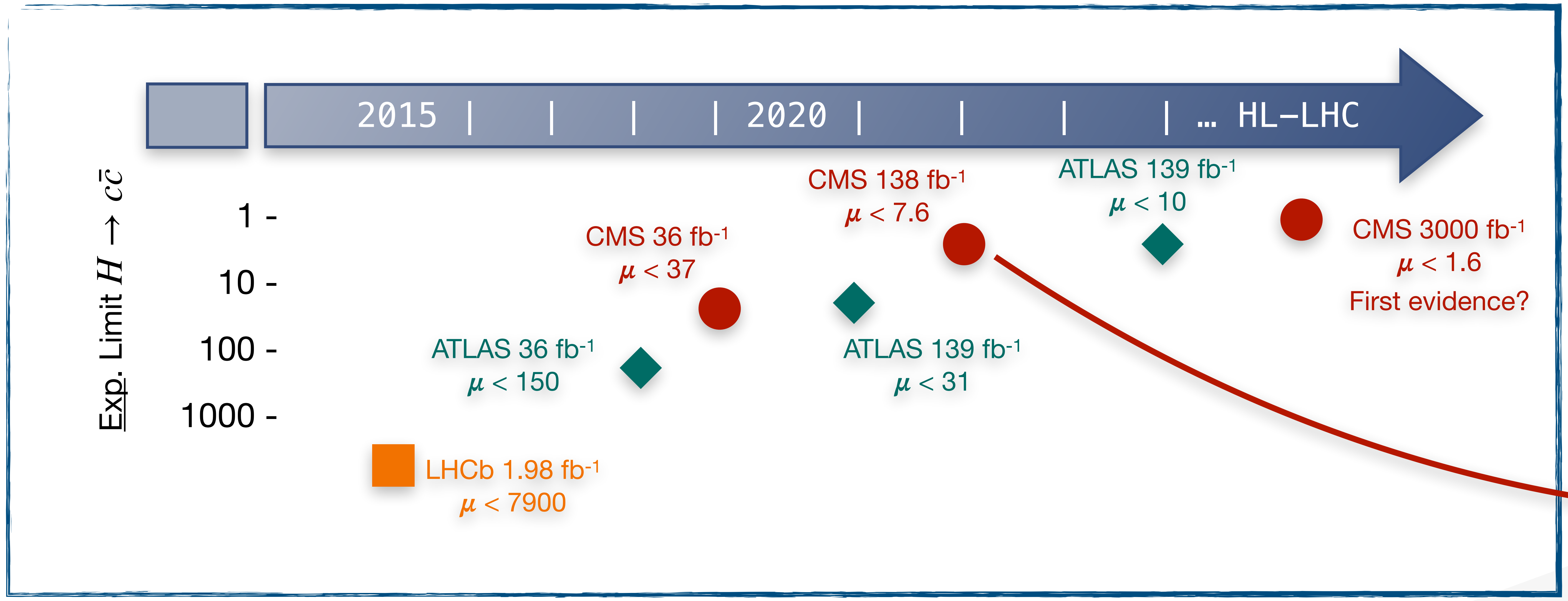
- **Run I:** Discovery of a boson in 2012 compatible with the scalar sector of the Standard Model.
  - Separate observation of the **bosonic** decay channels $O(15\%)$ and production in the most abundant channels.
  - Spin and parity compatible with $J^P = 0^+$.
- **Run II:** Firmly established the existence of Yukawa couplings.
  - Separate observation of **subdominant production modes:**
    $\Rightarrow$ **Vector Boson Fusion** (VBF),
    $\Rightarrow$ **Higgs-strahlung**, associated with W^{+-} and Z^0,
    $\Rightarrow$ t **and** $t\bar{t}$ **associated production.**
- **Run III and beyond:** Era of precision.
  - Tackle rare processes: **second generation** couplings and **Higgs self** and **quartic couplings** to complete the picture.



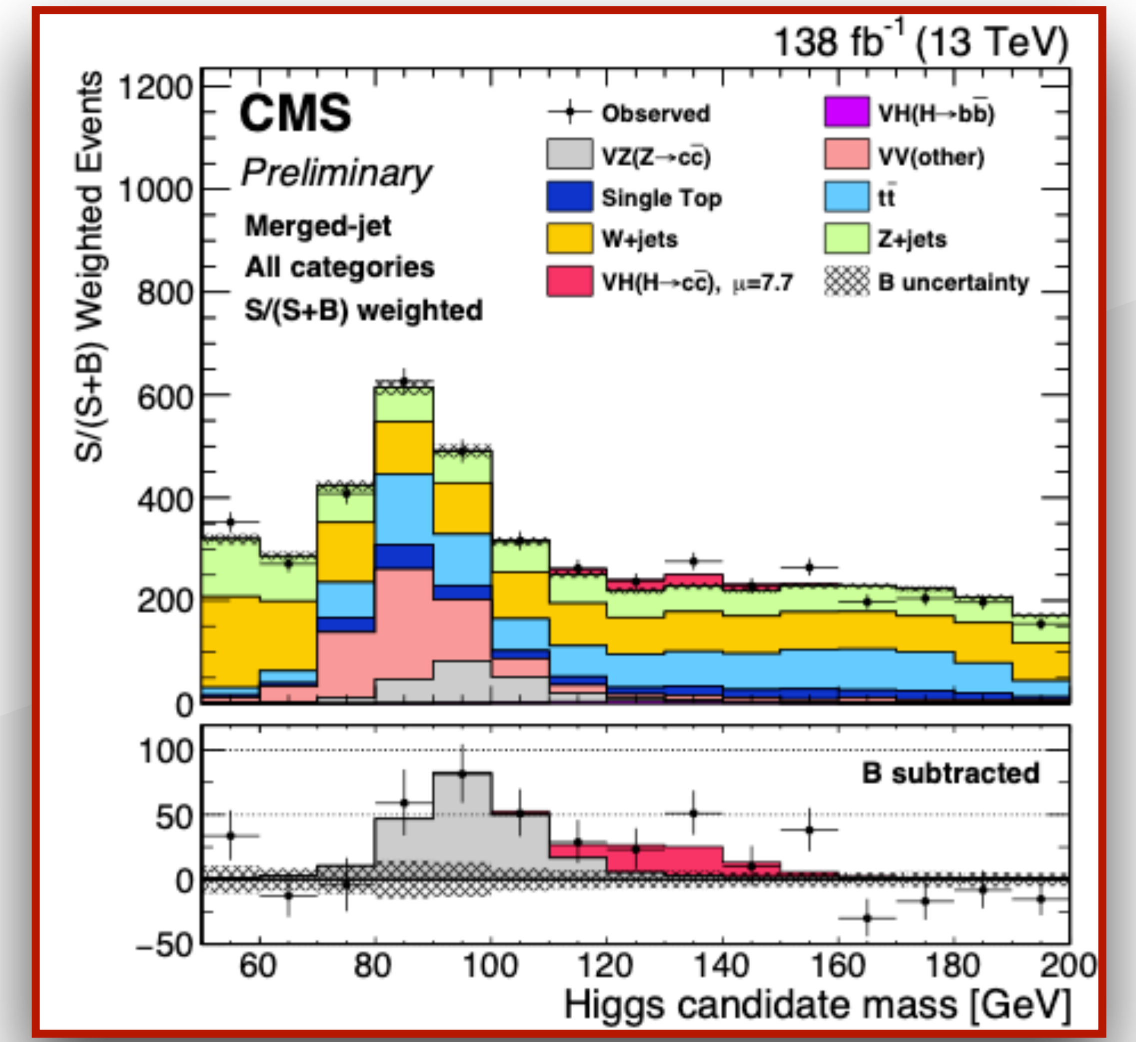
Nature 607, 60–68 (2022)

Probing the 2nd and 3rd generation Yukawas

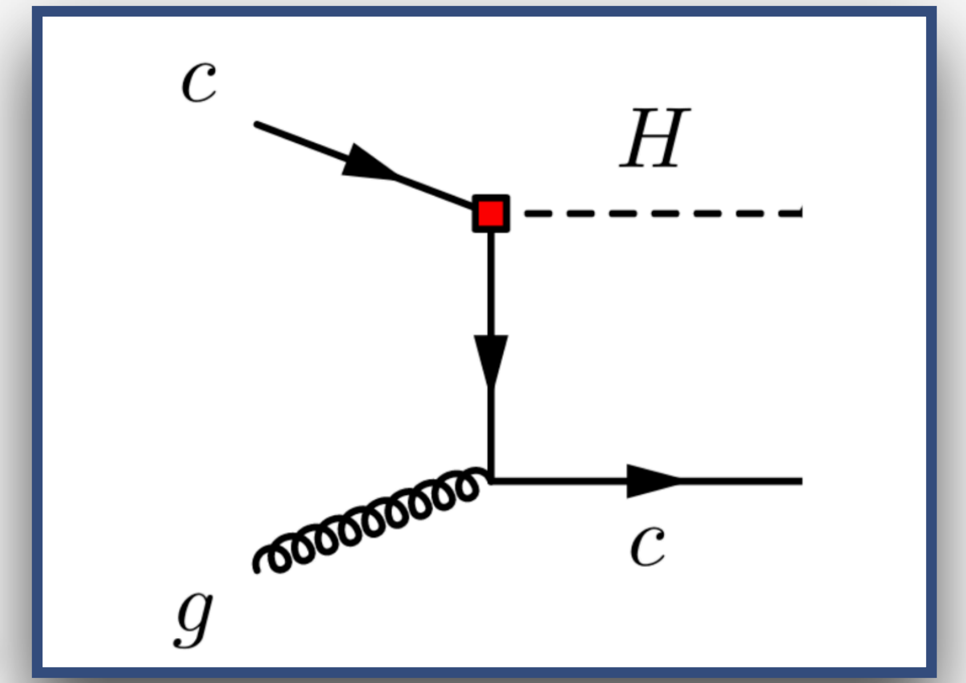
2nd generation Yukawas: Probing y_c



PRL 131 (2023) 061801



- Great improvements in the last few years.
- **Yukawa probed directly:** $VH(H \to c\bar{c})$ decay yields the most stringent (CMS) observed (exp.) limit of μ < 14 (7.6).
- Indirect approaches: Exclusive rare decays, $p_T(H)$ differential measurements.
- **New attempt:** Probe y_c in the production side with associated production.



$H(\gamma\gamma) + c$ analysis

- $H \to \gamma\gamma$ decay channel.

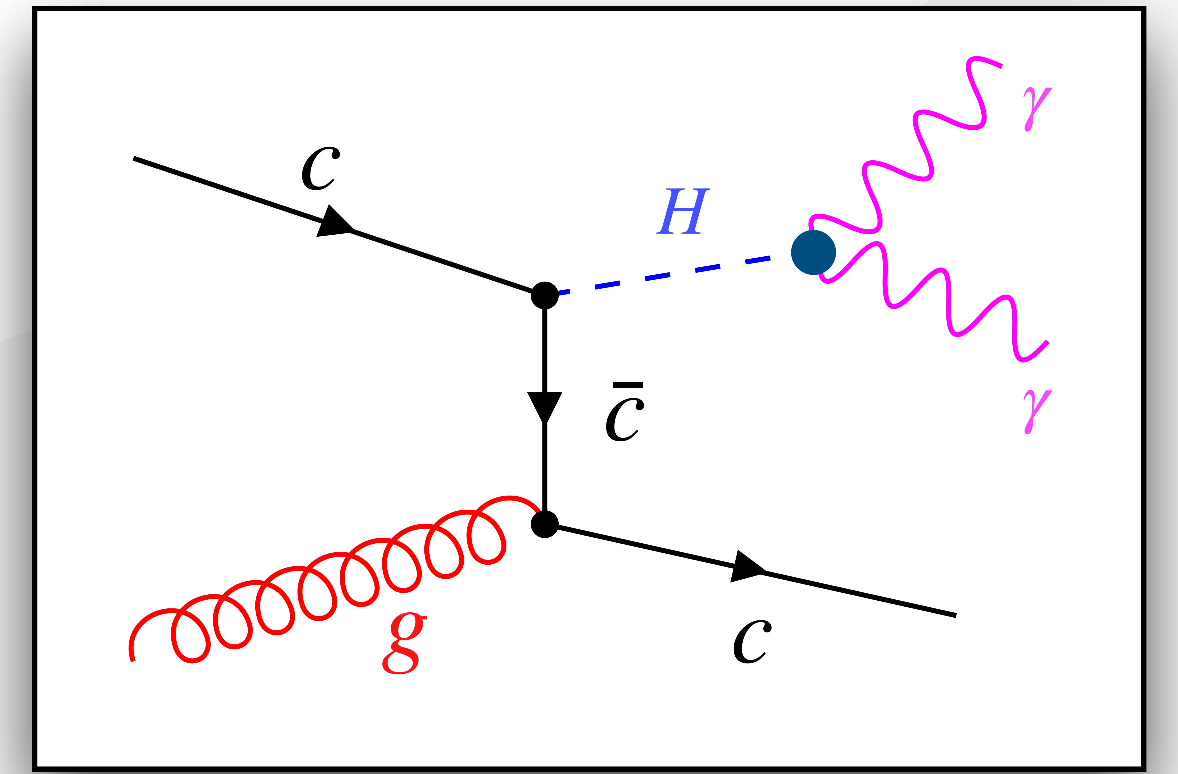
- **Main backgrounds**:
  - ⇒ Higgs production through gluon fusion (ggH),
  - ⇒ continuous diphoton background (CB) from $\gamma\gamma$ and $\gamma + jets$ events.

- **Full Run 2 dataset** of 138 fb^{-1}:
  - Uses dedicated NLO+PS simulation of the y_c dependent H production.

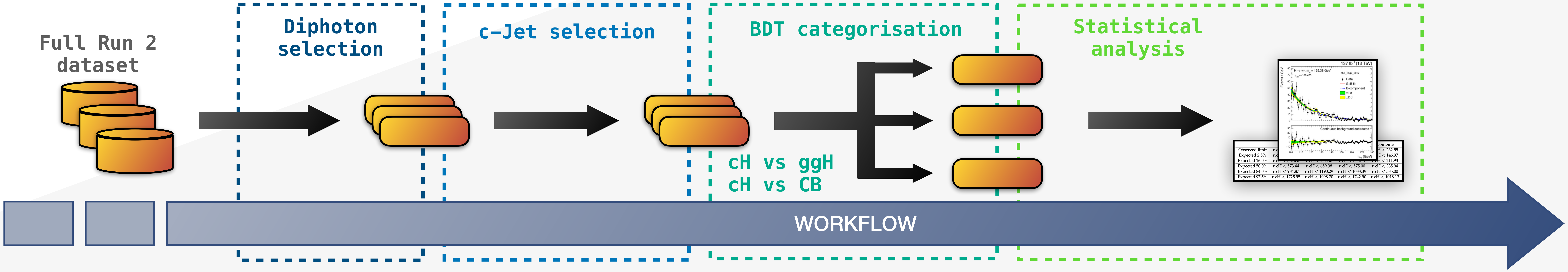
[GEN charm $p_T > 20$ GeV]

$[fb] \to A = 254.5,\ B = -3.5,\ C = 34.5$

$$\sigma(hc) = A + B \cdot y_c + C \cdot y_c^2$$



CMS-PAS-HIG-23-010



$H(\gamma\gamma)$ + c **analysis**

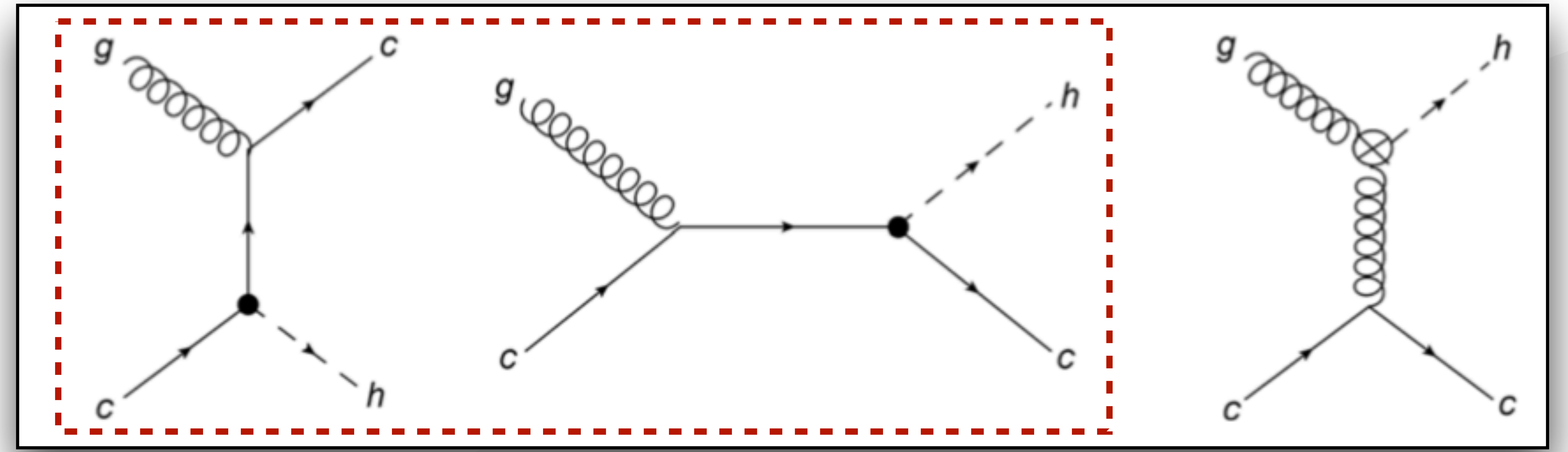
H+c signal:

- Focusing on the signal simulation for $H + c$ MC (not available in CMS up to now).

$$\sigma(hc) = A + B \cdot y_c + C \cdot y_c^2$$

	σ [fb]
A	254.5
B	-3.5
C	34.5

[GEN charm $p_T > 20$ GeV]



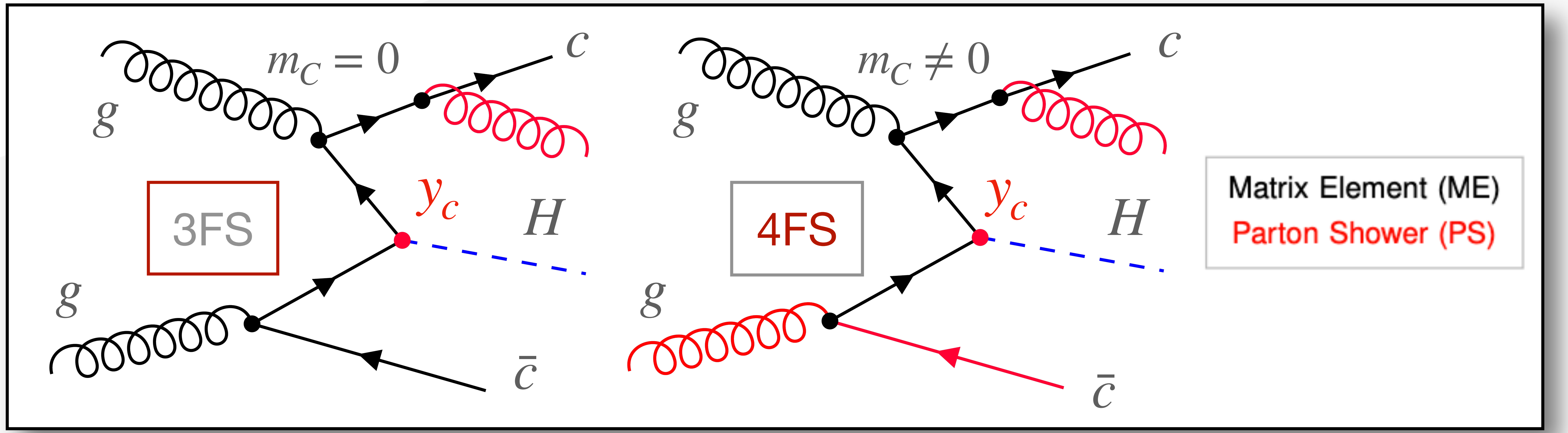
- Biggest contribution from the term that does not probe y_c.
- Small y_c proportional interference term (~10 times smaller than the y_c^2 dependent term), for sensitivity $O(10 \cdot SM)$ contribution of ~1%.
- As first approximation one can generate signal probing y_c^2 and bkgs/interference in separate MC, orthogonality with $H + jets$ MCs.

$H(\gamma\gamma)$ + c analysis

Focus on the y_c^2 term:

- Simulated with **MadGraph_aMC@NLO** ([QCD] NLO) + Pythia8 Parton Shower.
- Simulated using `loop_sm` model to have y_c in the $\overline{MS}$ renormalisation scheme and include **running of $y_c \to \overline{y}_c(\mu_R)$ and $m_c \to \overline{m}_c(\mu_R)$**.
- Simulated using **4 Flavour Scheme** (4FS), to have c-quarks in the initial state, and with FXFX-merging to **better describe the kinematics.**
  - To assess the 3FS vs 4FS theory uncertainty we **compare samples produced using both methods:**
    $\Rightarrow$ FS uncertainty $O(30\%)$ of the yields in analysis categories, dominant w.r.t. Scales, PDFs, PS.

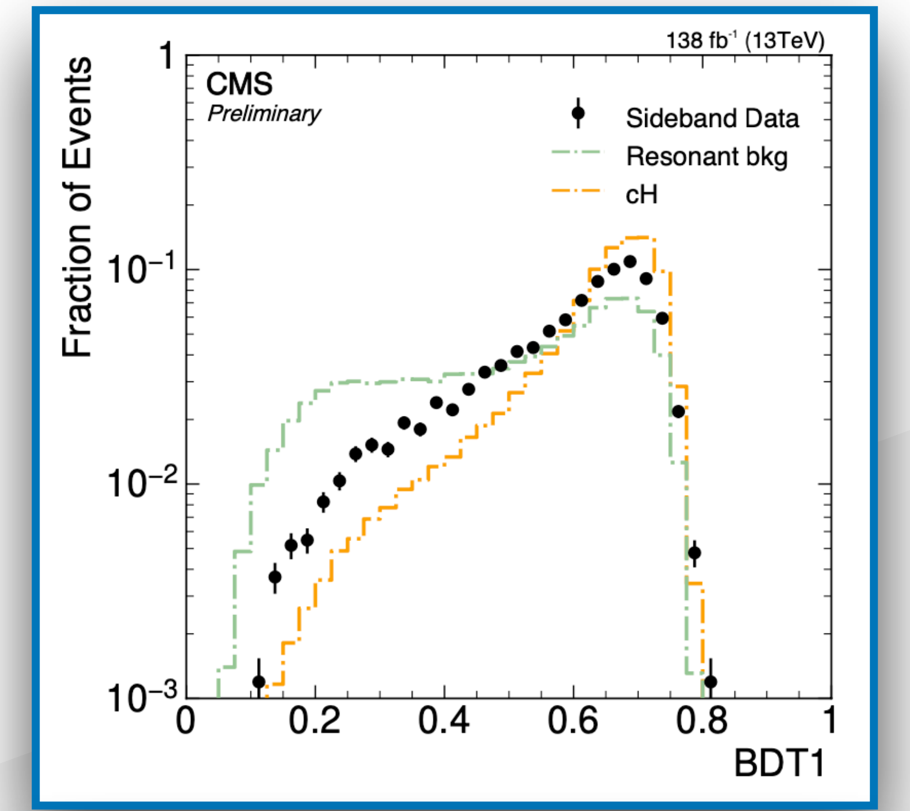
LO diagrams



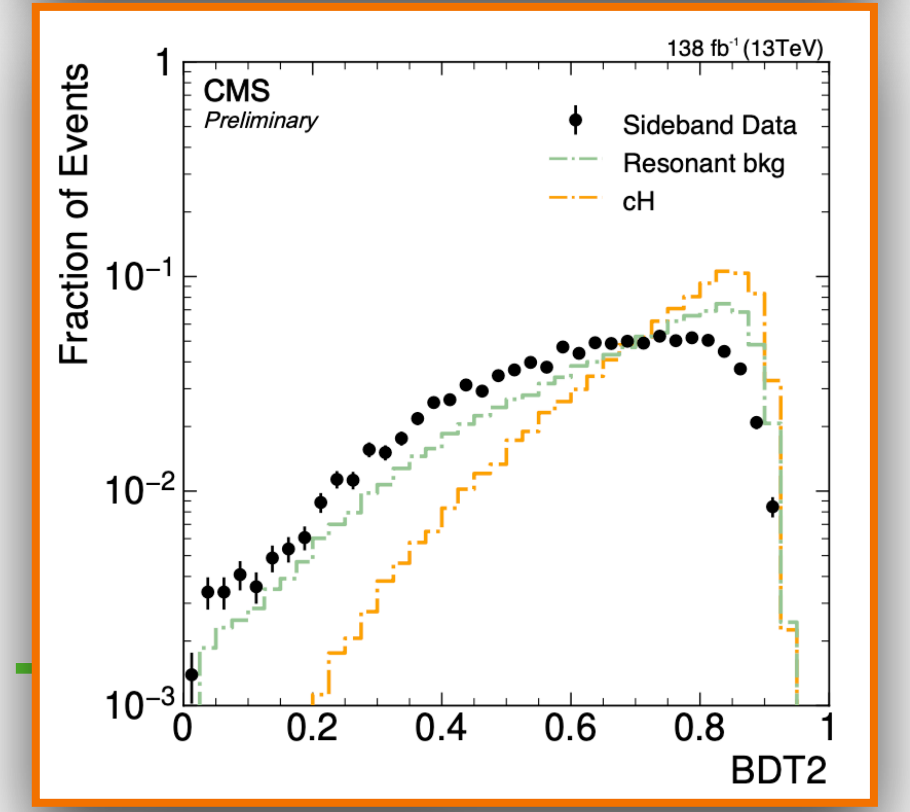
$H(\gamma\gamma) + c$ analysis

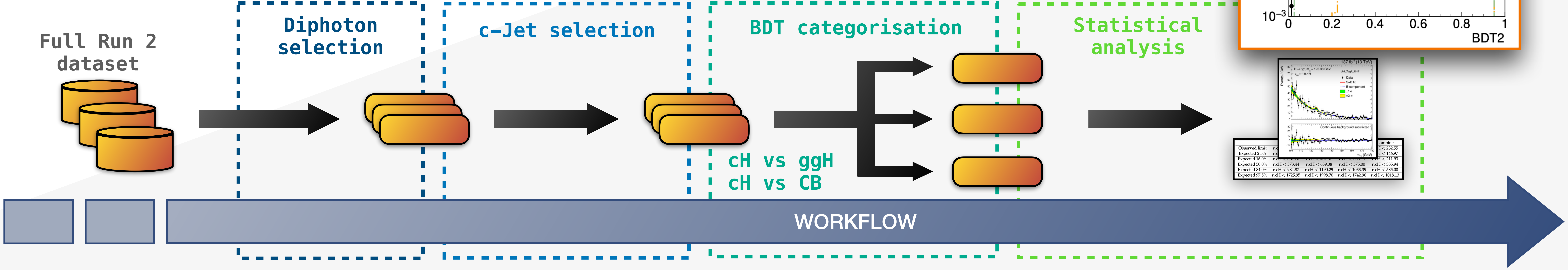
- $H \to \gamma\gamma$ decay channel.
- **Main backgrounds**:
  - ⇒ Higgs production through gluon fusion (ggH),
  - ⇒ continuous diphoton background (CB) from $\gamma\gamma$ and $\gamma + jets$ events.
- **Full Run 2 dataset** of 138 fb^{-1}:
  - Uses dedicated NLO+PS simulation of the y_c dependent H production.
  - Flavour Scheme studies to address signal generation theoretical uncertainty.
  - Kinematic based BDTs used to categorise the events.

cH vs ggH



cH vs CB



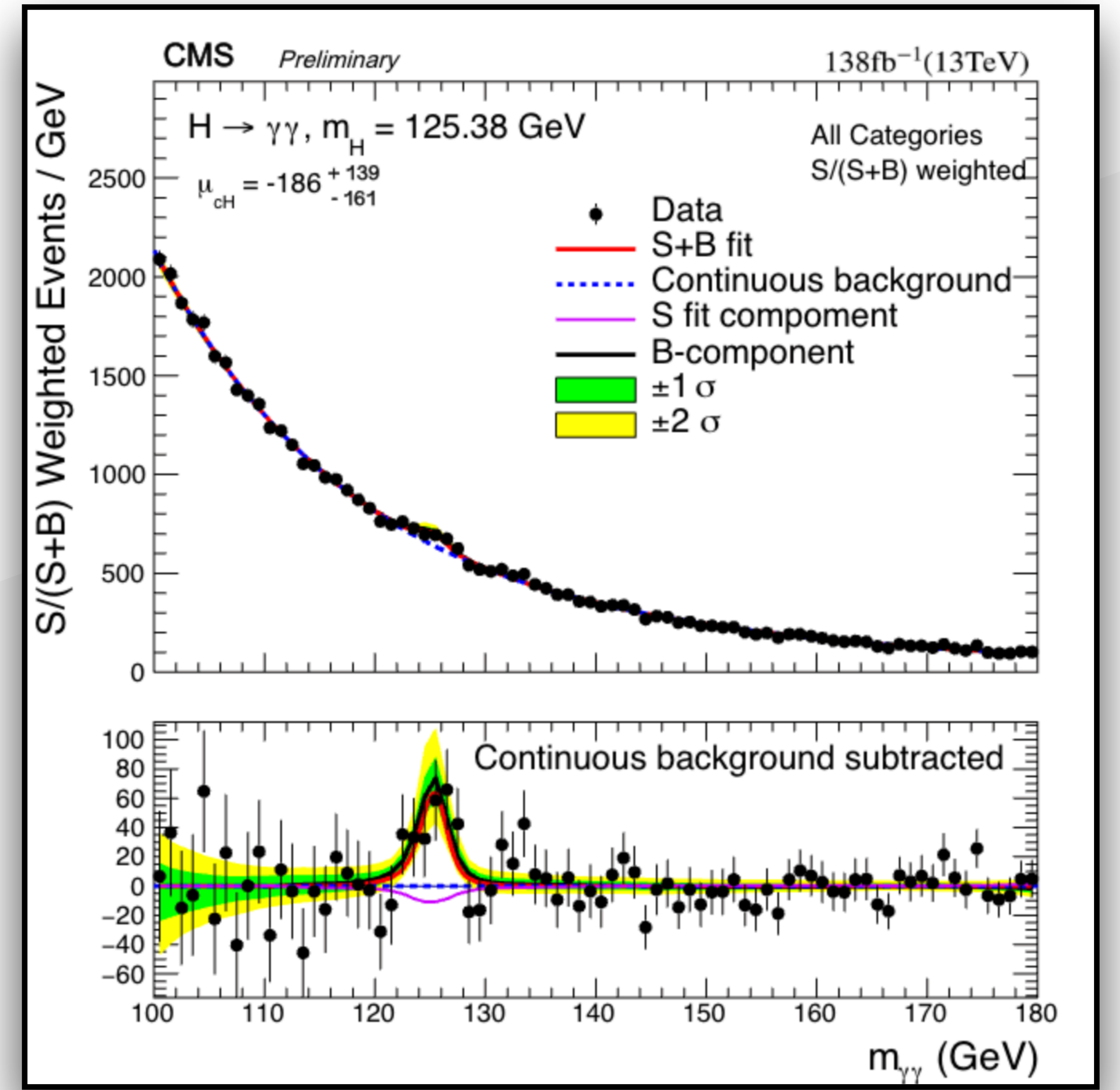


$H(\gamma\gamma)$ + c **analysis**

Results:

[CMS-PAS-HIG-23-010](CMS-PAS-HIG-23-010)

- μ_{cH} is extracted via a **simultaneous maximum likelihood fit** in the $m_{\gamma\gamma}$ distribution in the 27 event categories.
- Assuming the standard model (SM) cross sections times branching fractions for all other Higgs production processes.
- The **observed (expected)** upper limit at 95% confidence level on the cH signal strength is **243 (355)** times the SM prediction.
- Result interpreted considering the **"flat direction" approach** (PRD 100 (2019) 073013):
  - The **observed (expected)** allowed interval is $|\kappa_c|$ **< 38.1 (**$|\kappa_c|$ **< 72.5)** at 95% confidence level.



- Signal: cH
- Res. Bkg: ggH, ttH, VH, VBF, bH
- Cont. Bkg: $\gamma\gamma$, γ+jets

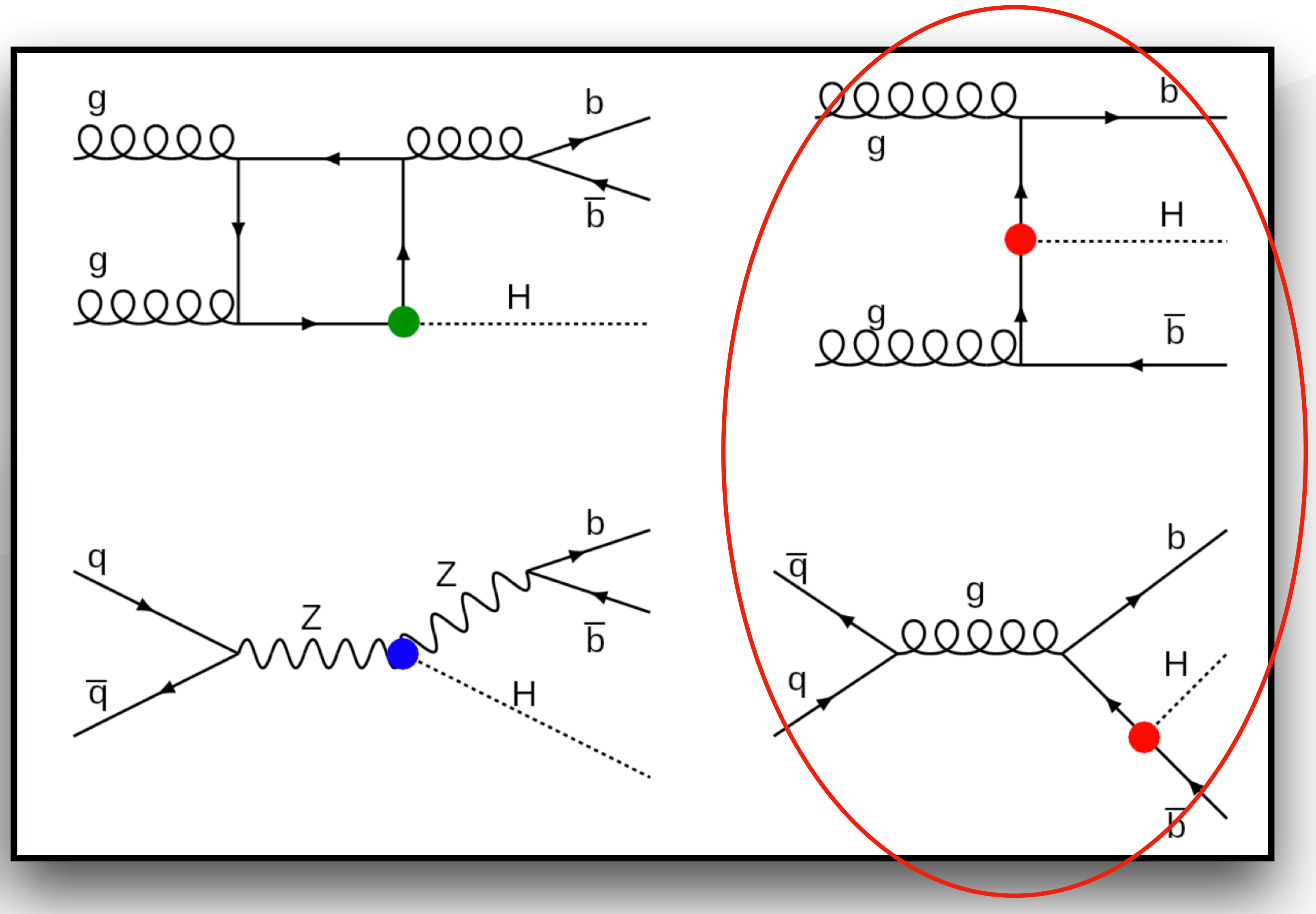
$$\mu_{cH} = \frac{\sigma_{cH}}{\sigma_{cH}^{SM}}$$

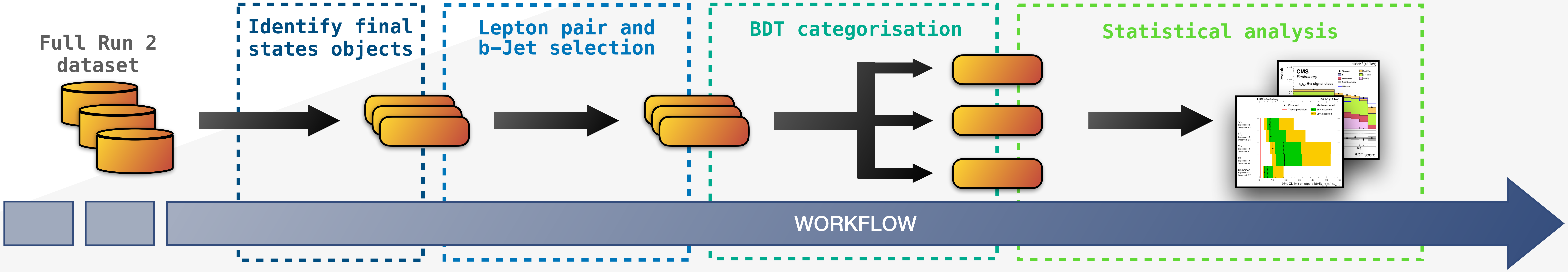
bbH associated production

- Search for b-quark associated Higgs boson production followed by decay to τ lepton pair or WW.
- Direct probe of Higgs couplings to the bottom quark (y_b) in production.
- Challenging analysis experimentally with 4 final states ($e\mu$, $e\tau_h$, $\mu\tau_h$, $\tau_h\tau_h$):
  - Larger background than in the *ttH* channel.
  - Complex contribution of y_t dependent diagrams and interference.

term	σ(pb)
y_t^2	1.040 (+0.468 -0.489)
y_b^2	0.482 (+0.048 -0.070)
$y_b y_t$	– 0.033 (+0.007 -0.008)

Subdominant y_b dependent contribution





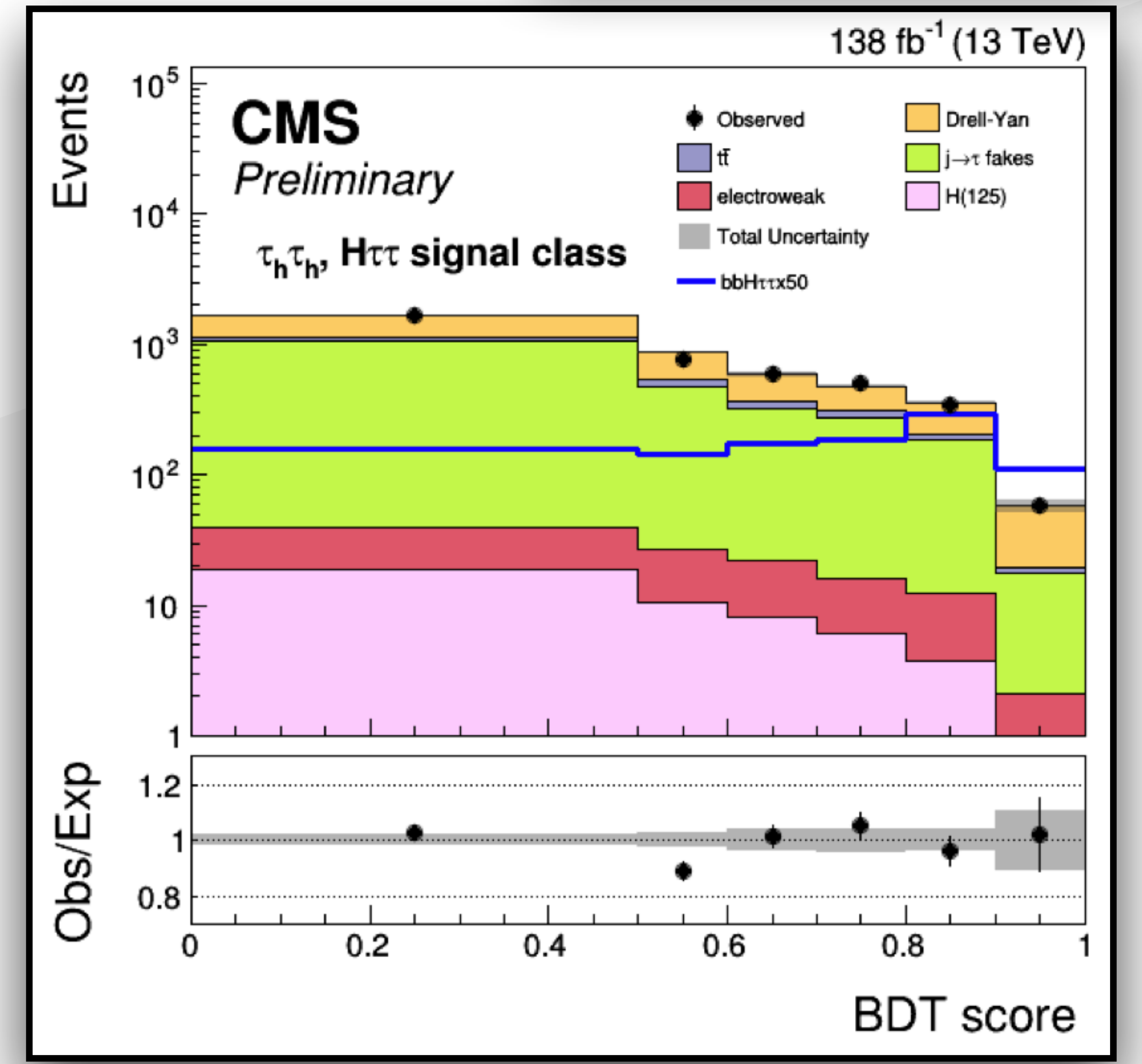
bbH associated production

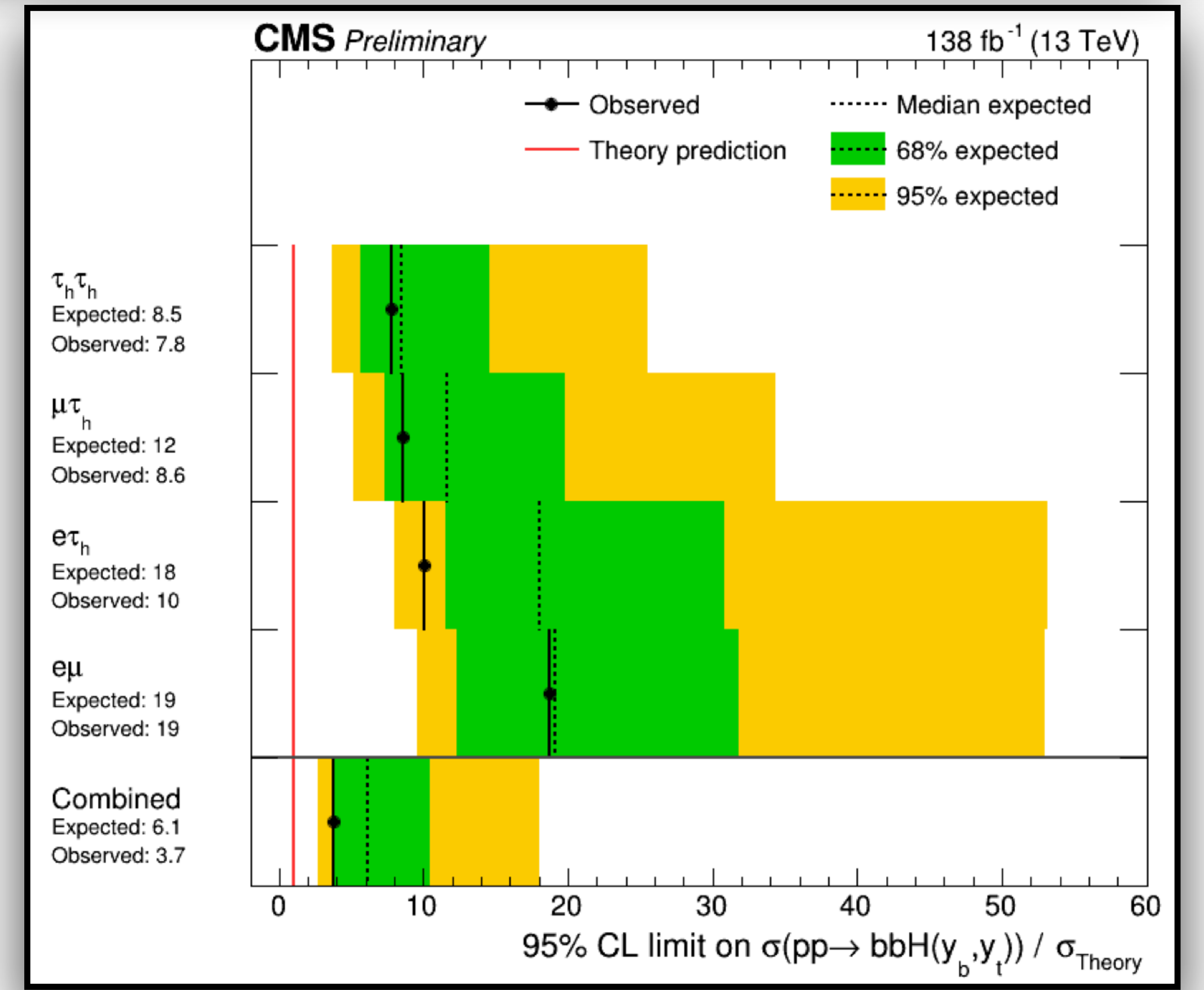
Results:

- Dominant backgrounds: $t\bar{t}$, DY, and $j \rightarrow \tau_h$ misidentification ⇒ require dedicated classes.
- Fit to BDT score, inclusive measurement:
  - The different contributions to the signal are scaled by varying proportionally the y_b^2, y_t^2 and $y_b y_t$ terms.

Channel	$e\mu$	$e\tau_h$	$\mu\tau_h$	$\tau_h\tau_h$
BDT Categories	DY, TT, bbH(→WW), bbH(→ $\tau\tau$)	DY, TT, bbH(→ $\tau\tau$)	DY, TT, bbH(→ $\tau\tau$)	DY+Higgs, TT, j→τ_h fakes, bbH(→ $\tau\tau$)

- Observed **obs (exp)** upper limits at **3.7 (6.1)** times the SM expectation.



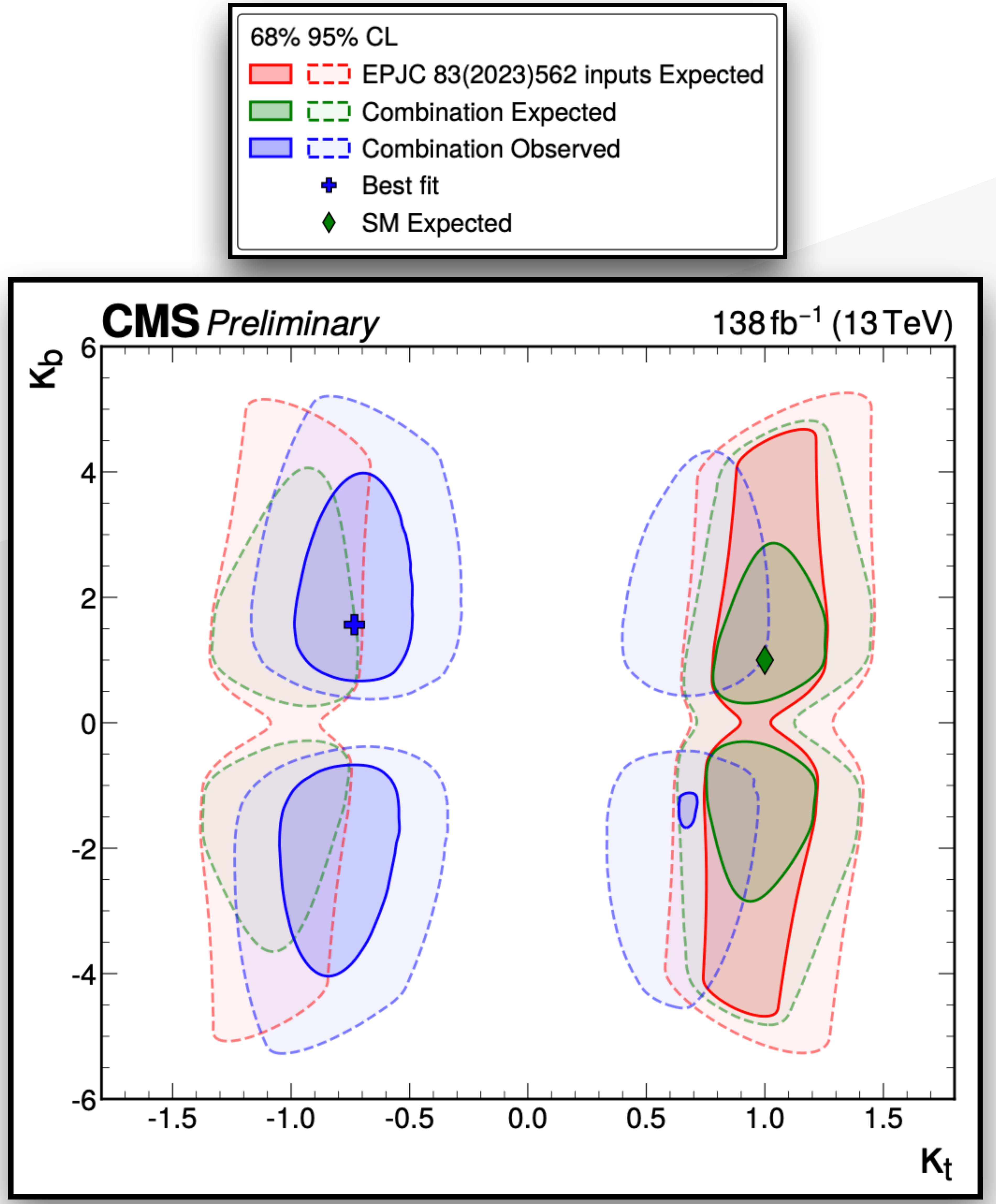


CMS PAS HIG-23-003

bbH associated production

K-framework interpretation:

- Scan performed on coupling modifiers k_t and k_b, with k_τ freely floating.
- Combined with the results from STXS $H \rightarrow \tau\tau$ cross-section measurement (with veto on b-jets) to better constrain k_t.
- The best fit point is (k_t,k_b) = (-0.73,1.58)
- Limits on the couplings are **compatible with the SM** at 95% CL.



CMS PAS HIG-23-003

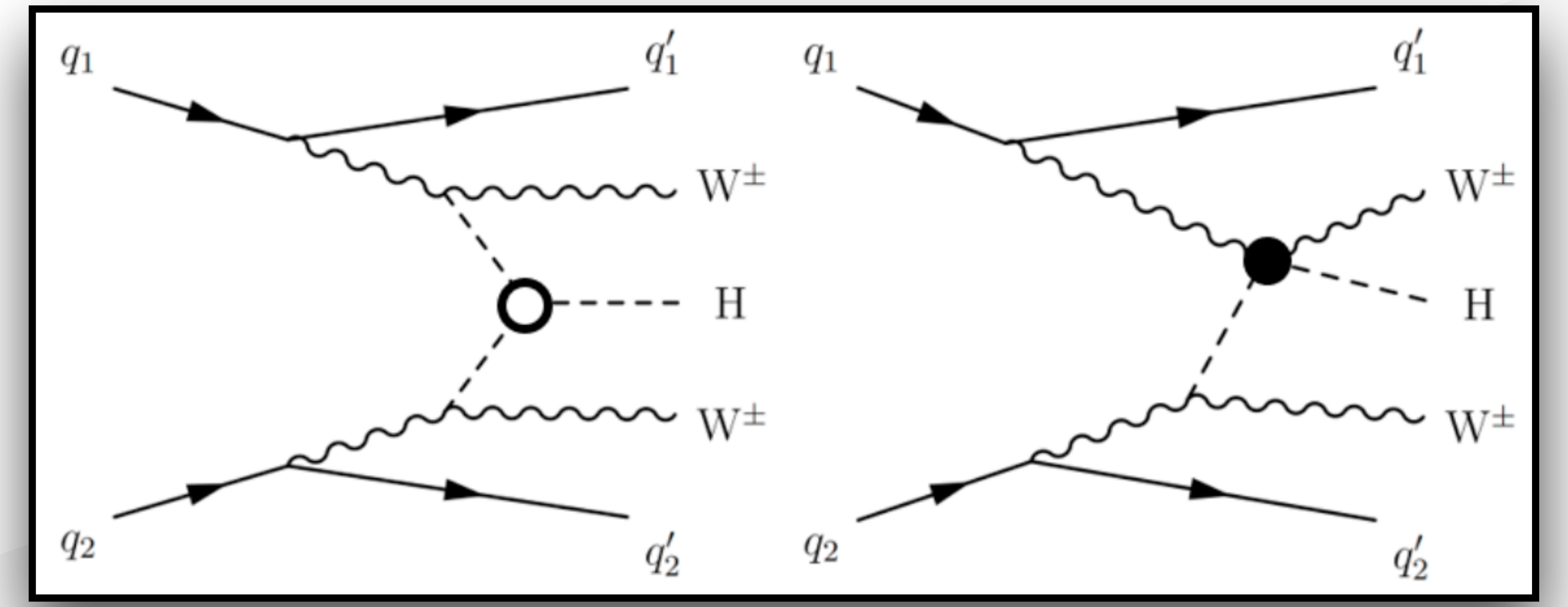
Vector bosons couplings

$H + W^{\pm}W^{\pm}$ boson production through VBS

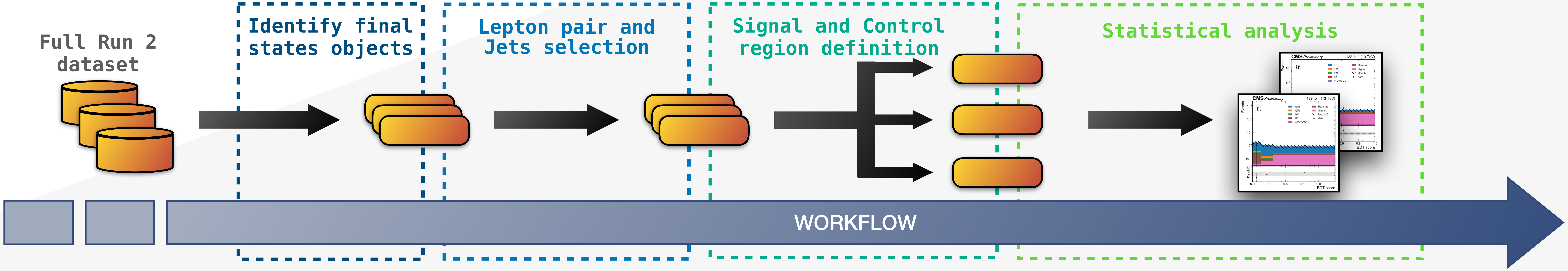
- H production through Vector Boson Scattering (VBS) is sensitive to $HHWW$ (k_{VV}) and HHH (k_{λ}) couplings.

- **Focus on** $k_{VV} \rightarrow$ other channels are more performing in constraining k_{λ}.

- $H + W^{\pm}W^{\pm}$ **events are selected requiring**:
  - $n = 2$ same sign leptons,
  - 2 well separated Jets with a radial dimension $\Delta R < 0.4$,
  - A merged b-Jet pair in a large radius bb-tagged Jet with $\Delta R < 0.8$.



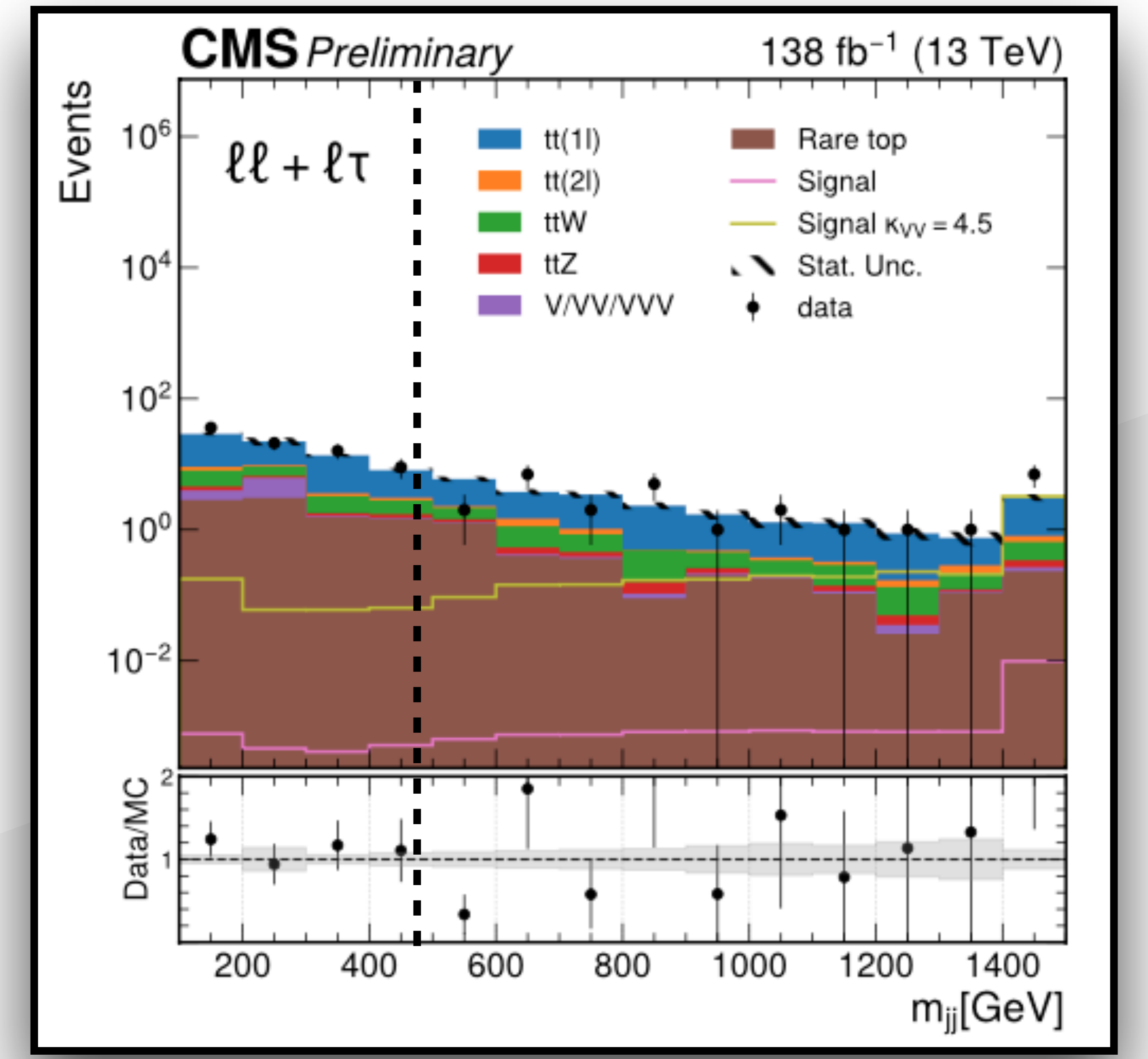
CMS PAS HIG-24-001

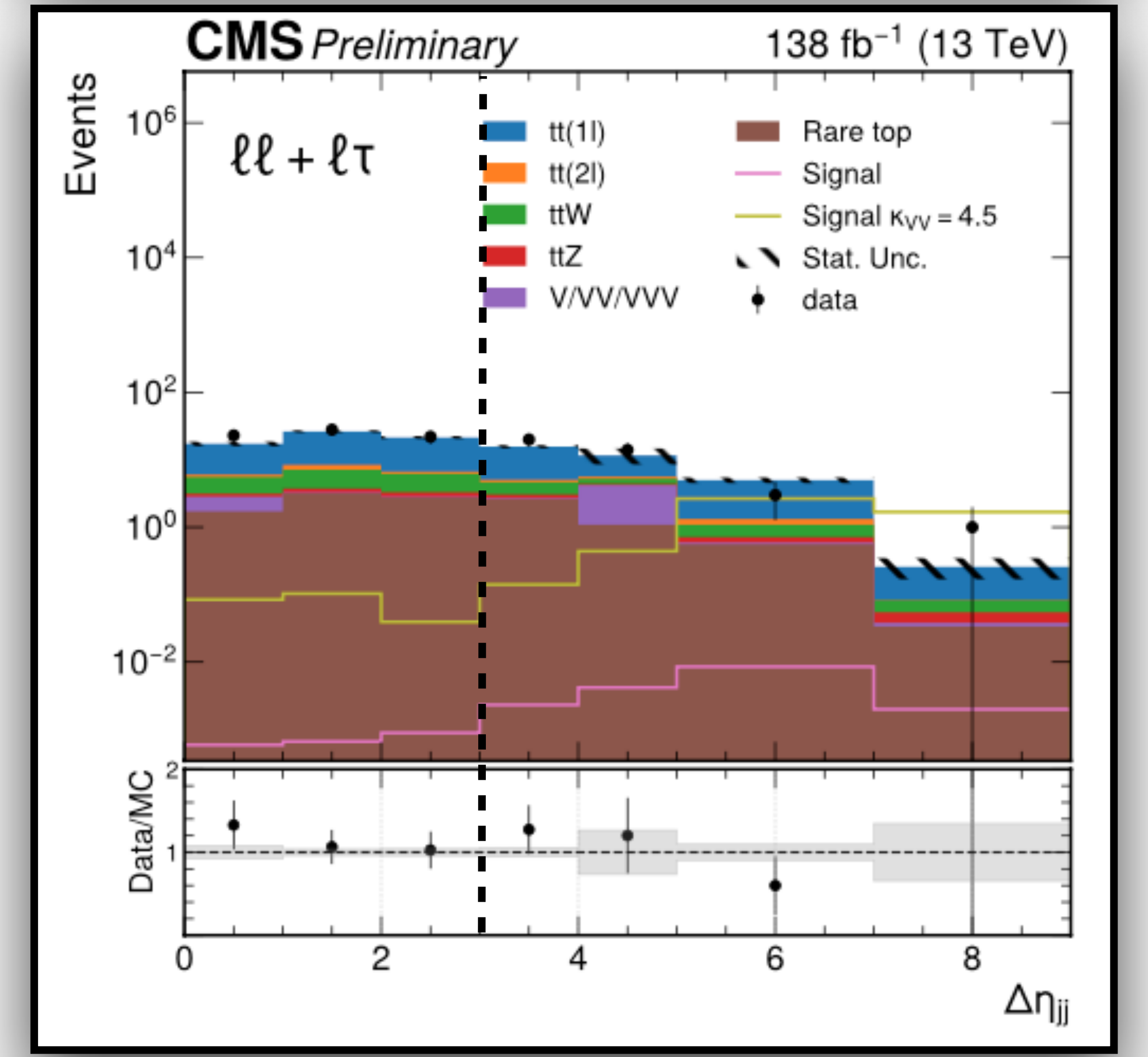


$H + W^{\pm}W^{\pm}$ boson production through VBS

Categorisation and background rejection:

- Only the $H \rightarrow b\bar{b}$ decay in the **boosted topology** is used.
- **Main backgrounds** for the analysis are various top and vector boson production modes:
  - Events with additional leptons are vetoed to suppress contribution from $HHZZ$ couplings.
  - The VBS AK4 Jets must fail the medium b-tagging WP.
  - $m_{jj} > 100$ GeV requirement to remove the non-VBS-like events.
- **Signal Region (SR)**: $m_{jj} > 500$ GeV, $|\Delta\eta_{jj}| > 3$ and $Xbb > 0.9$.
- **Control Region (CR)**: events failing at least one requirement (m_{jj} or $|\Delta\eta_{jj}|$).
- Events are split in categories according to the number of leptons: $ll, l\tau$.



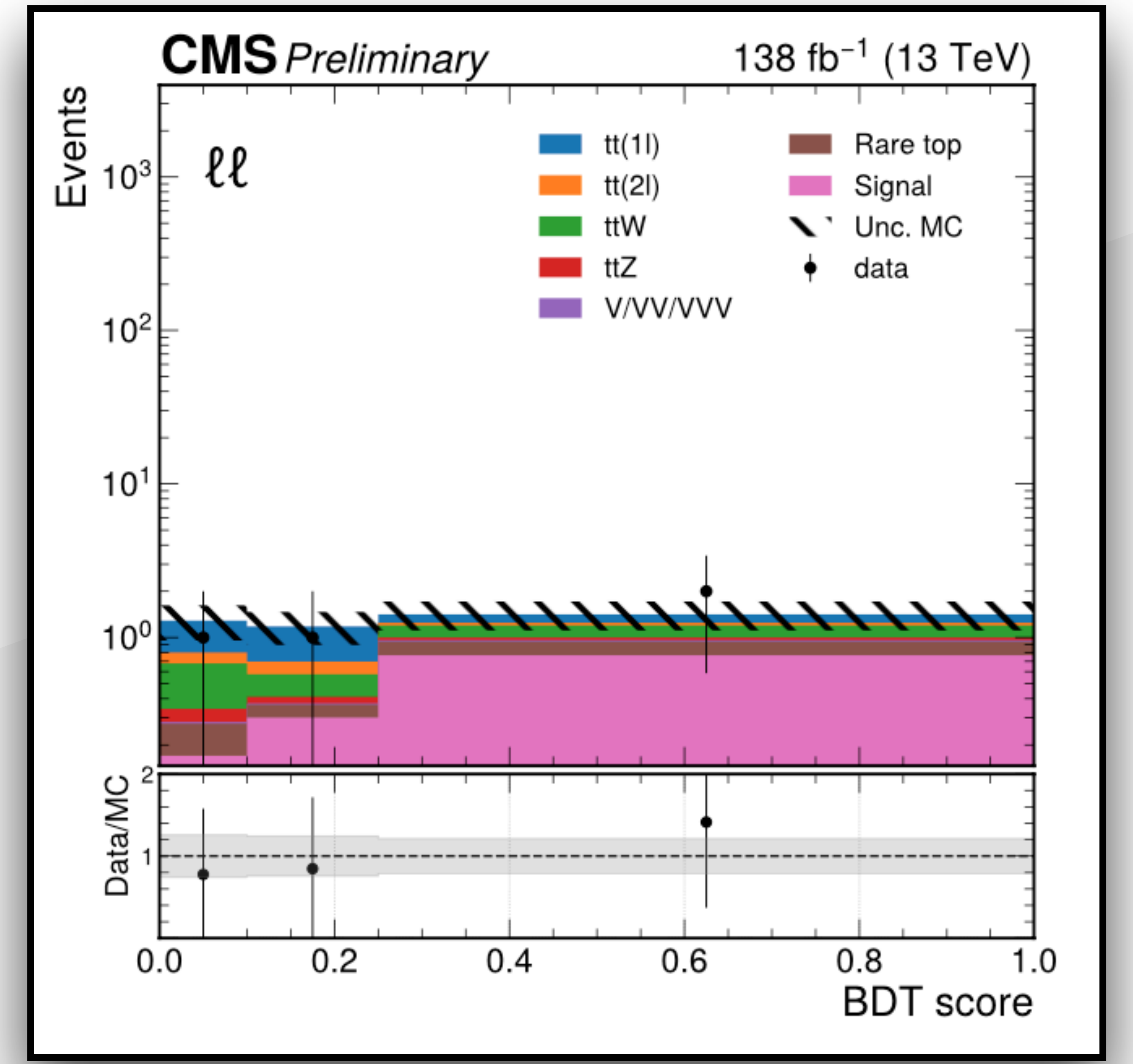


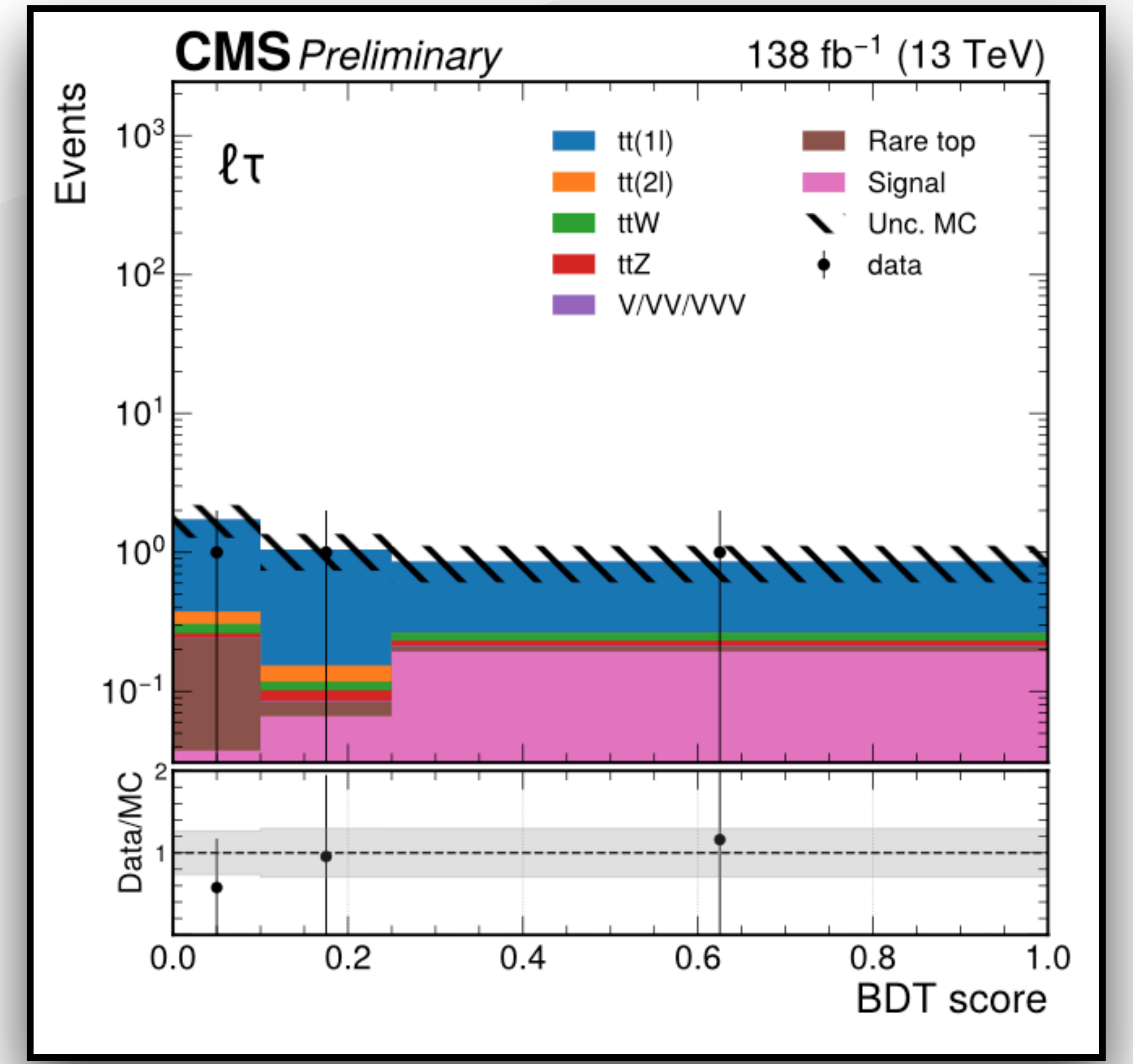
CMS PAS HIG-24-001

$H + W^{\pm}W^{\pm}$ boson production through VBS

Statistical analysis:

- A BDT is trained to discriminate signal and background using kinematics of Jets, leptons and MET.
- **Simultaneous fit of the BDT score distribution** in the two lepton flavour categories:
  $\Rightarrow$ Best fit yields $\mu = 169^{+199}_{-137}$.
- The backgrounds normalisation is constrained from fit in the CRs.



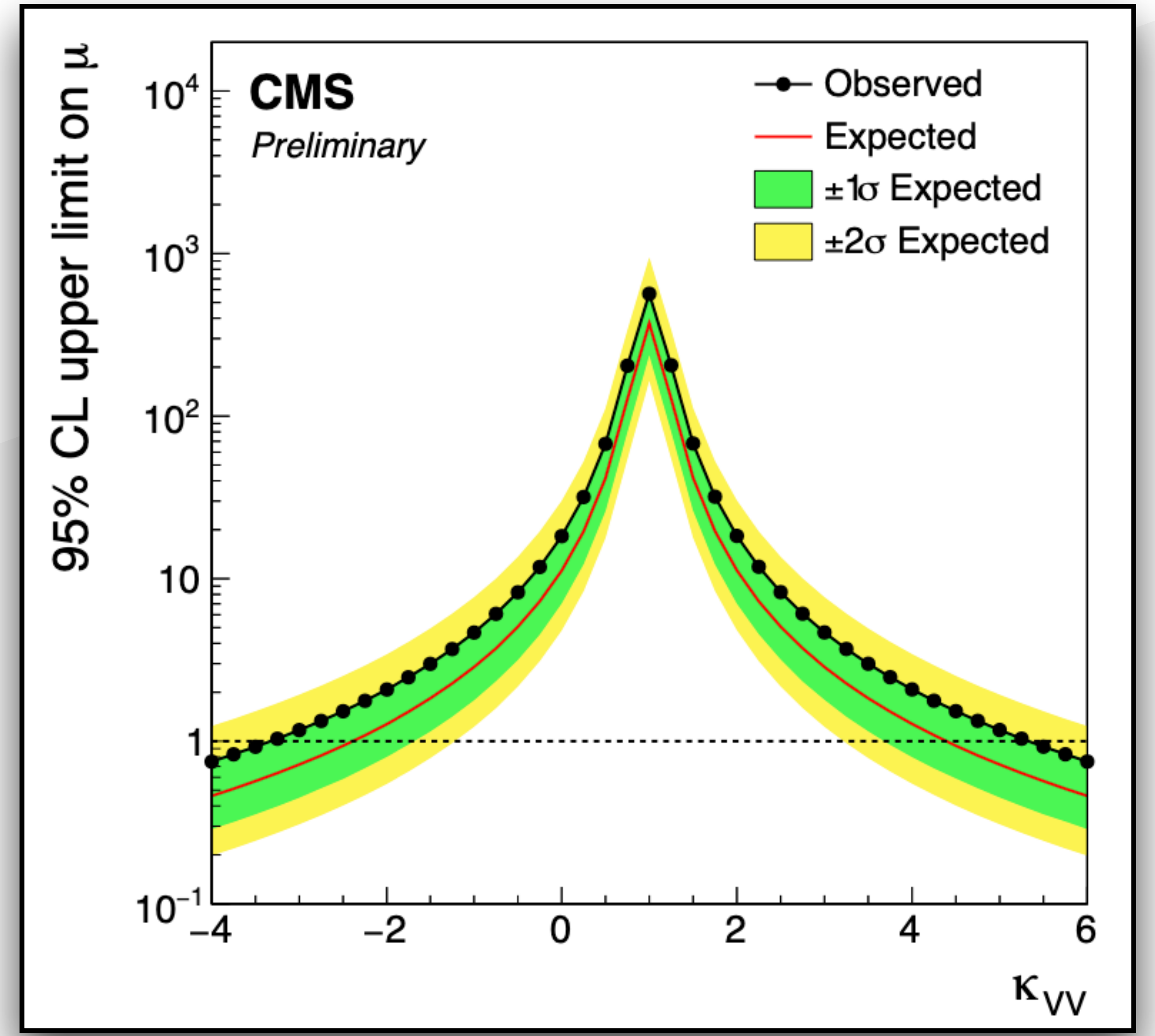


CMS PAS HIG-24-001

$H + W^{\pm}W^{\pm}$ boson production through VBS

Limits:

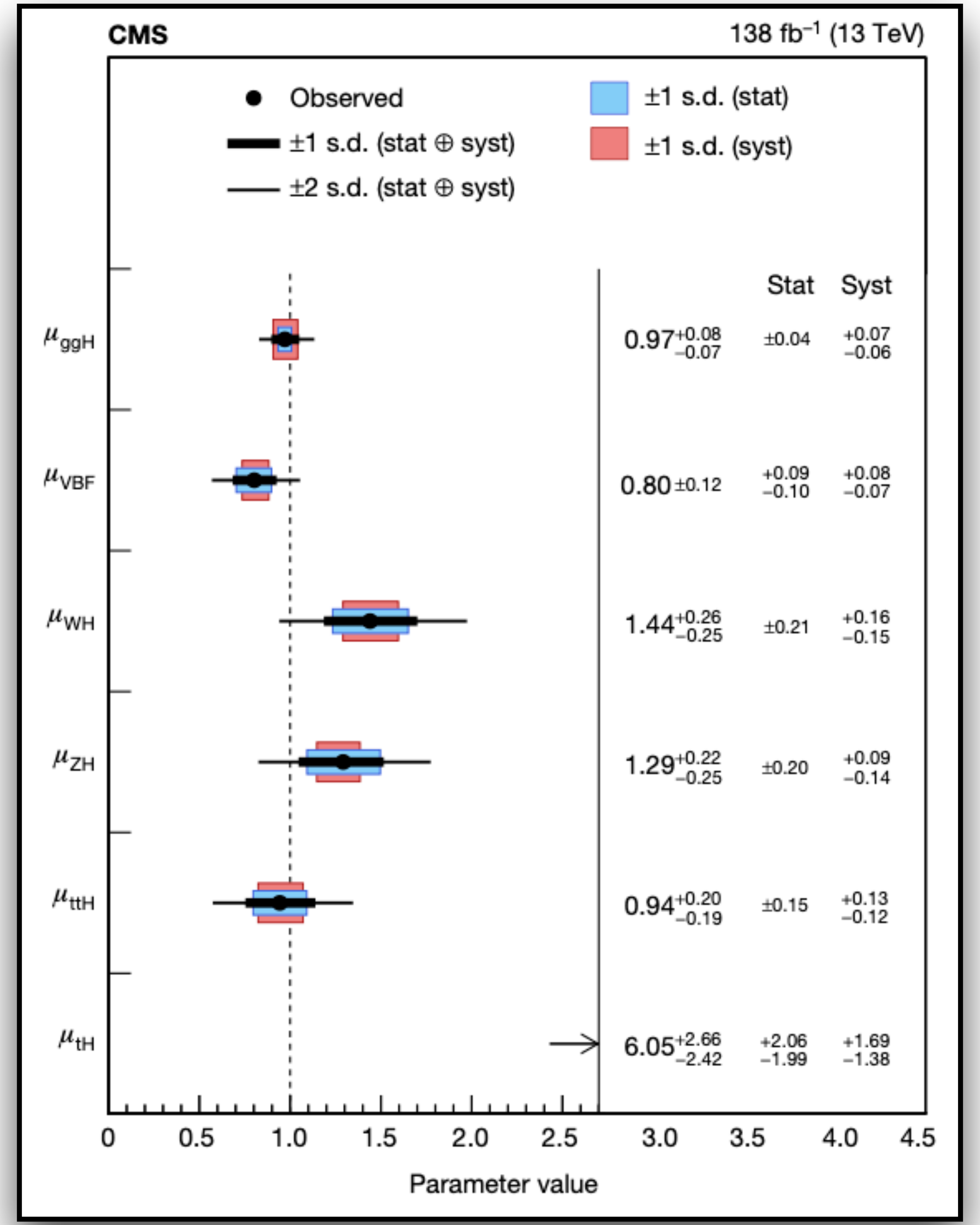
- Fit with different k_{WW} hypotheses:
  - Upper limit at 95% CL on the signal strength as a function of k_{WW}.
- Allowed range k_{WW} found by the intersections of this with the line $\mu = 1$.
- The observed (expected) 95% CL constrained interval is $k_{ww} \in [-3.3,\ 5.3]$ ($[-2.4, 4.4]$).



CMS PAS HIG-24-001

Summary

- After more than 10 years of studying the properties of the Higgs Boson we're now targeting precision measurements of its properties.
- **We've taken the most out of the Run 2 dataset:**
  - Subdominant production modes are known with $O(20\%)$ uncertainty.
  - 2nd generation couplings and rare production modes are becoming accessible.
- Today I presented new results focusing on the measurements of k_c, k_b and k_{VV} exploiting rare processes.
- Much more to come! Looking forward to Run 3 and beyond.

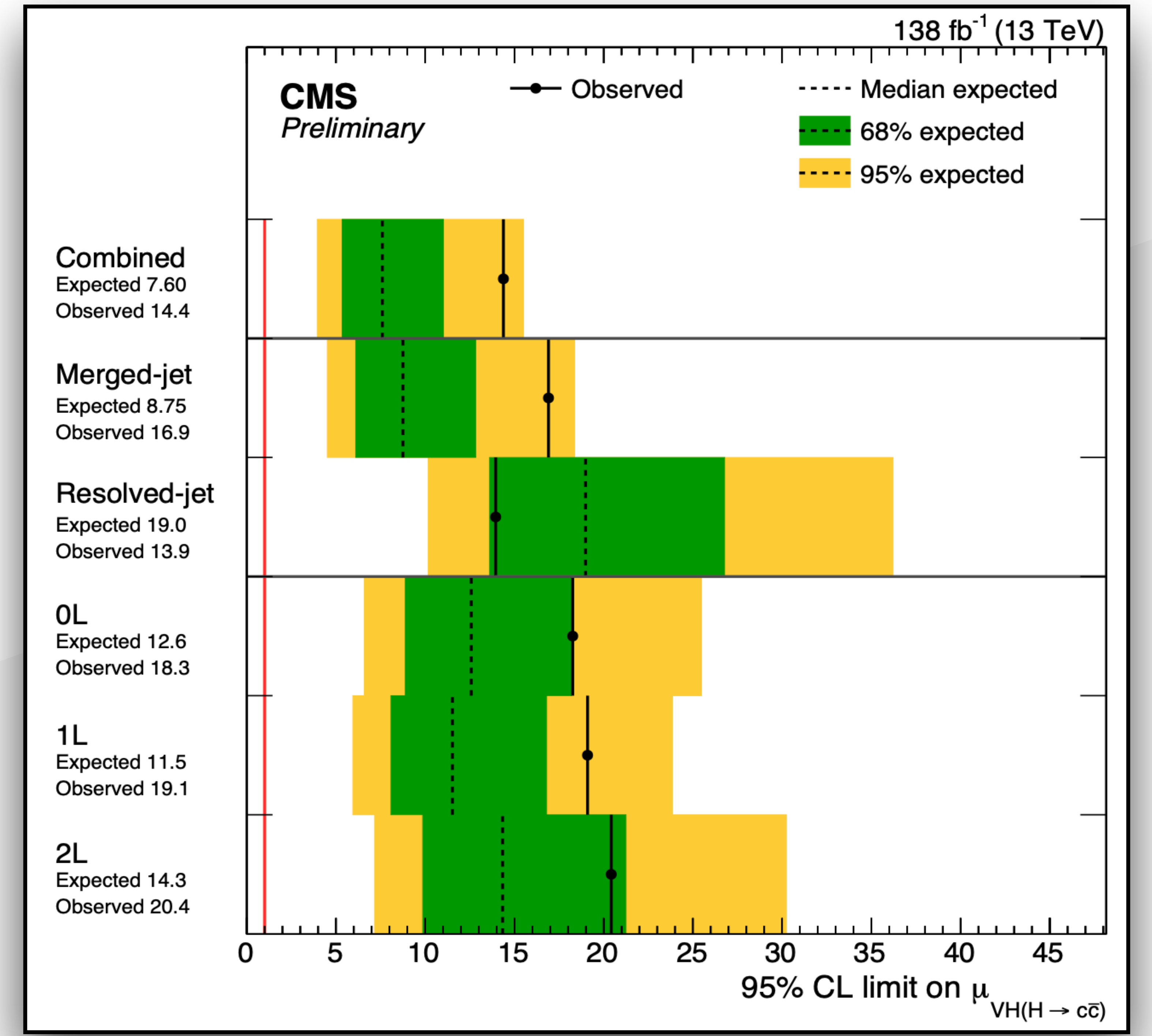


Nature 607, 60–68 (2022)

Back Up

Motivation

- Direct search for $VH(H \to c\bar{c})$ arXiv:2205.05550: recent improvements, most stringent limit on $H \to c\bar{c}$.
  - Upper limit $\mu_{VH(H\to c\bar{c})} < 14\ (7.6)$ observed (expected).
  - $1.1 < |k_c^{[*]}| < 5.5\ (|k_c| < 3.4)$ observed (expected) at 95% C.L.
    [ATLAS : $|k_c| < 8.5 (12.4)$ obs (exp) at 95% C.L.]
  - First observation of $Z \to c\bar{c}$ at a hadron collider (5.7σ)



- Boosted $ggH(H \to c\bar{c})$ HIG-21-012:
  - $\mu < 38\ (45)$ observed (expected) at 95% C.L.

- Exclusive $H \to J/\Psi + \gamma$ decays, clean signature, $J/\Psi \to \mu\mu$ but very rare process:
  - $BR/BR_{SM} < 220\ (170)$ observed (expected) at 95% C.L.
    [ATLAS : proj. for $3\ ab^{-1}$ $\mu < \mu_{SM}$ at 95% C.L.]

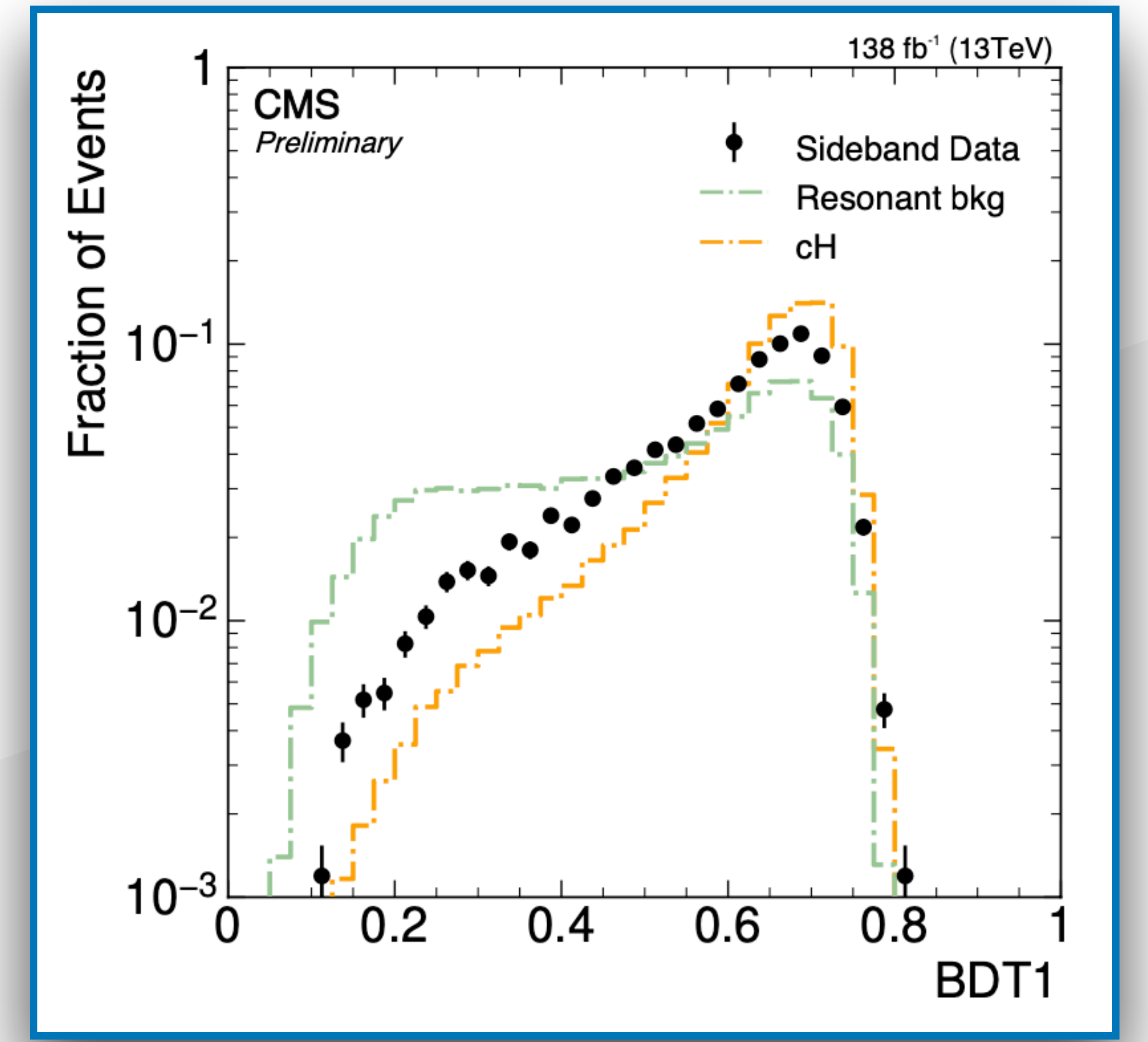
- H differential measurements, variation of $p_T(H)$ as a function of k_c:
  - $-4.9 < k_c < 4.8\ (-6.1 < k_c < 6.0)$ observed (expected) at 95% C.L.

$H(\gamma\gamma)$ + c analysis

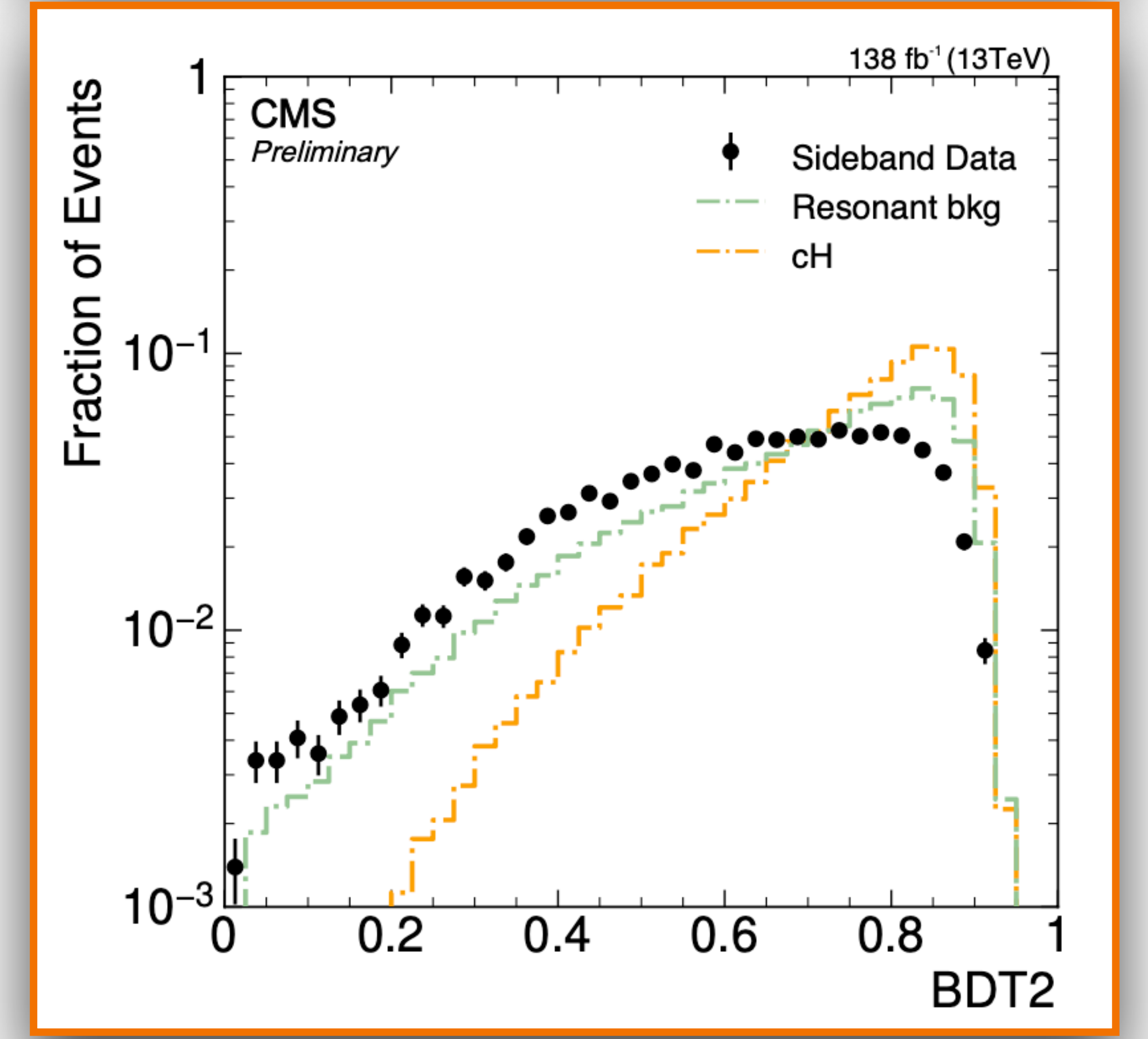
BDT training:

- We use two Boosted Decision Trees (BDT) to separate the two main backgrounds:
  - **cH vs ggH** (BDT1): cH (signal), ggH (background),
  - **cH vs CB** (BDT2): cH (signal), $\gamma\gamma$ and $\gamma + jets$ (backgrounds).
- Separation is achieved exploiting the **kinematics of the Photons and Jets** in the event.
- c-tag scores are NOT used in the training, so that $ggH + c$ fraction is stable w.r.t BDT outputs. This limits the impacts of $ggH + HF$ mismodelling.
- Training performed with the XGBoost package.

cH vs ggH



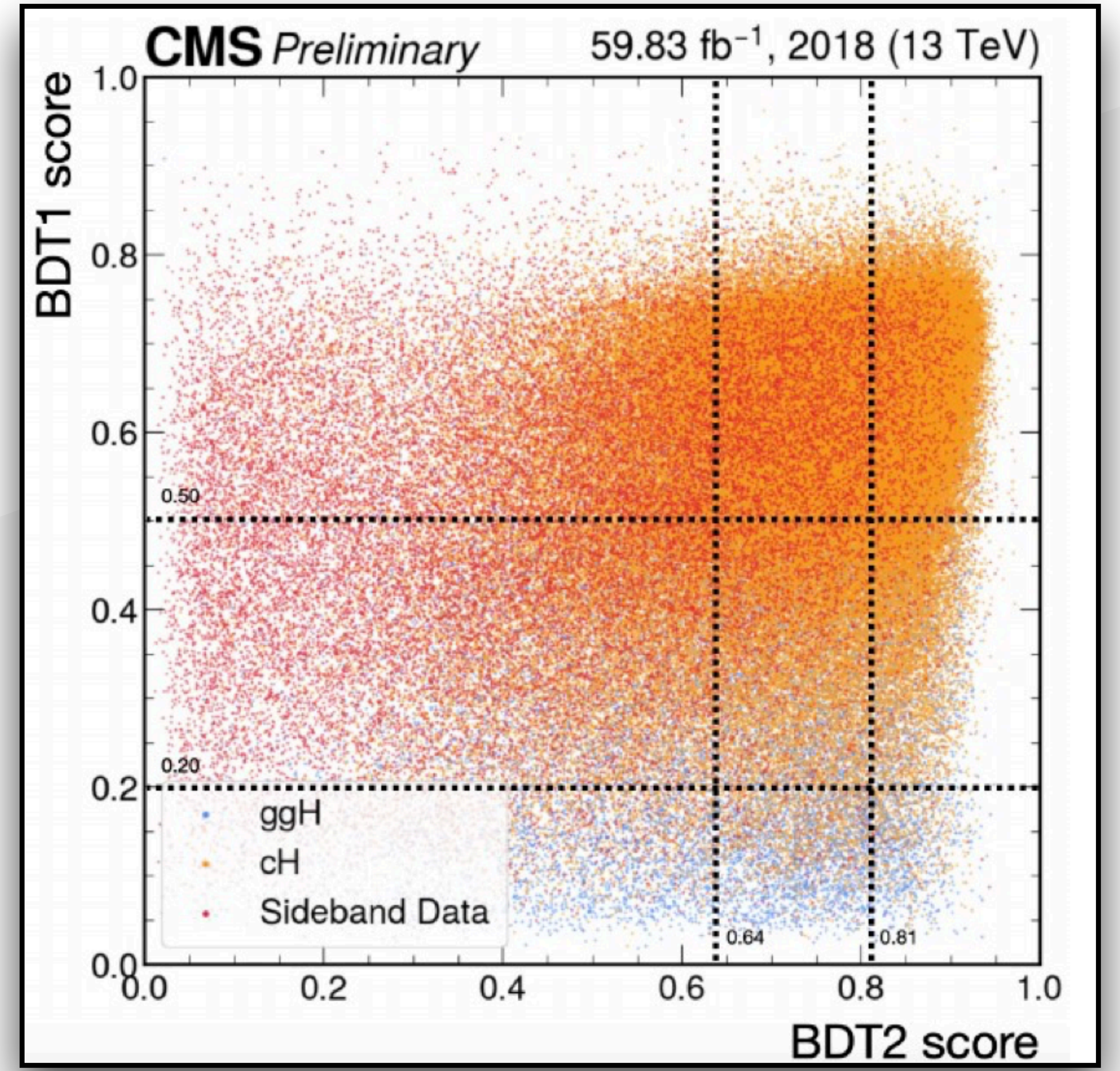
cH vs CB



$H(\gamma\gamma) + c$ analysis

Categorisation:

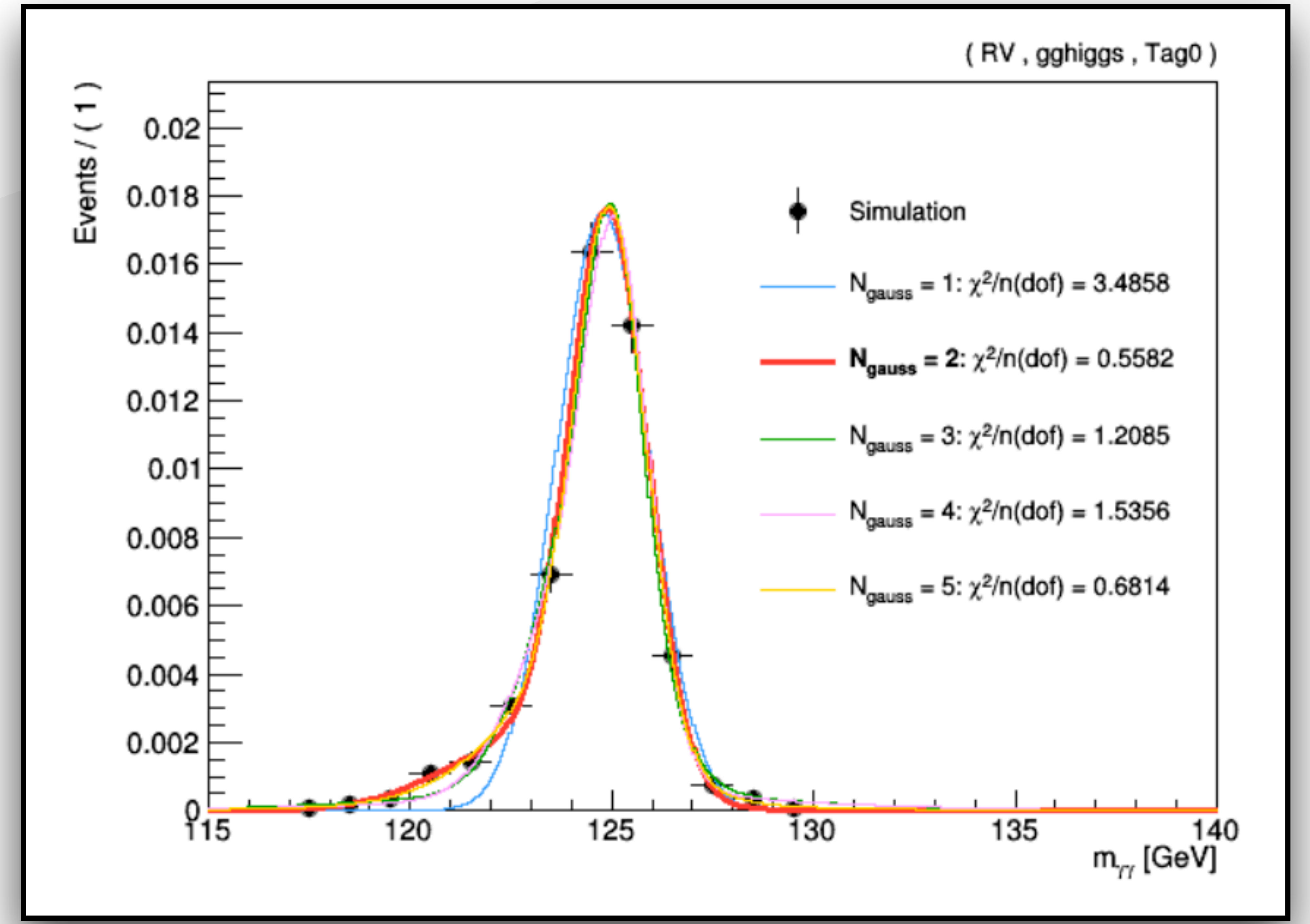
- The events are divided into **9 categories for each year**, according to the scores of BDT1 and BDT2.
- The category boundaries are simultaneously optimised using MCs:
  - To reduce the correlation between the cH and ggH processes,
  - To maximise the sensitivity.
- Boundaries are optimised separately for each year.
- Migration uncertainties and data/MC agreements are extracted from $Z \to e^+e^-$ events.



$H(\gamma\gamma) + c$ **analysis**

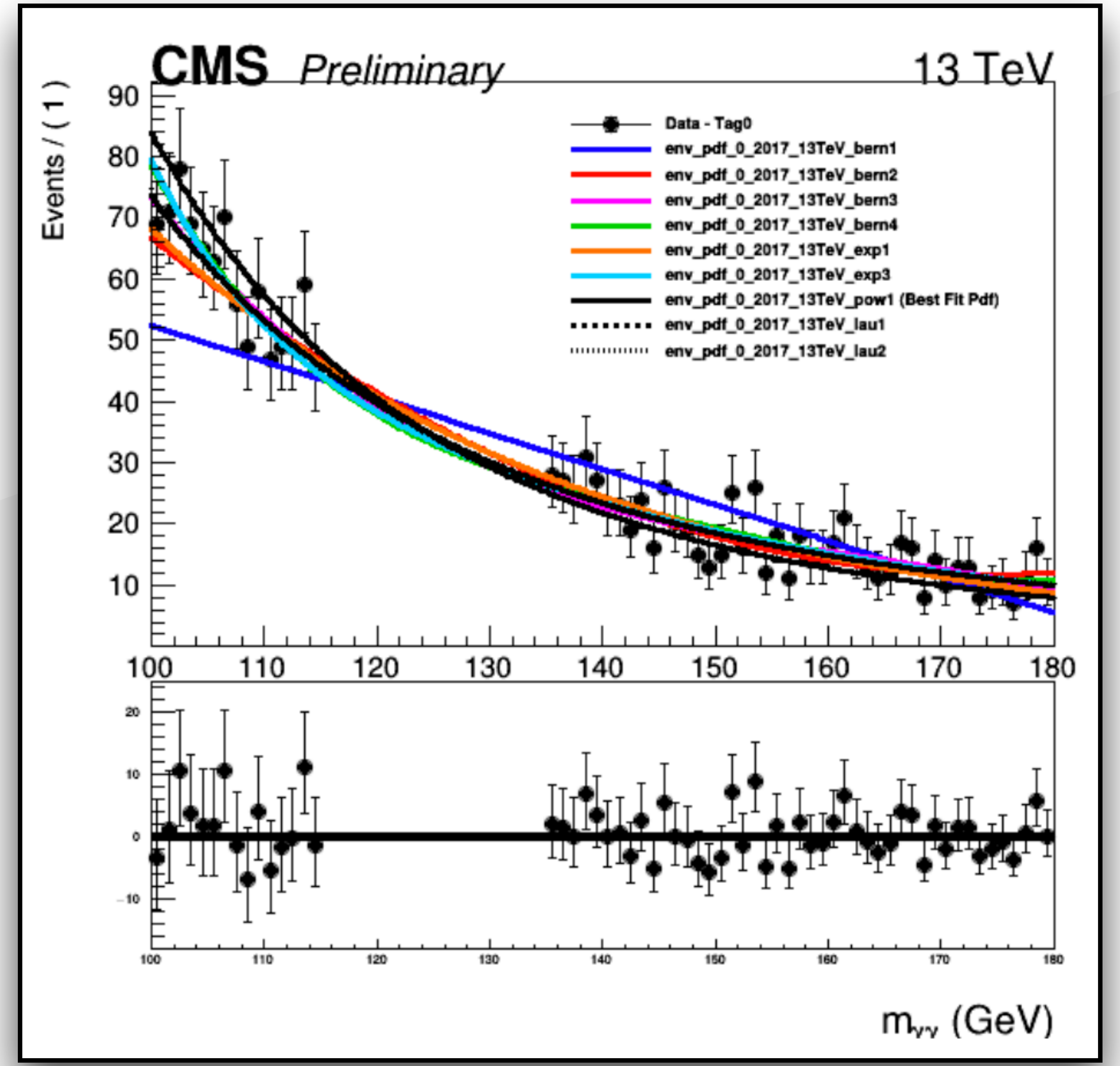
Higgs processes modelling:

- The statistical analysis is performed with the $H \to \gamma\gamma$ FlashggFinalFit framework.
- The signal and the Higgs backgrounds are modelled with MC simulations.
- **Mass shapes** are:
  - Parametrised with **sum of multiple gaussian functions.**
  - The parametrisation is derived **independently for each process X category X vtx scenario.**
  - Yields are extracted from simulation.



Continuous background modelling:

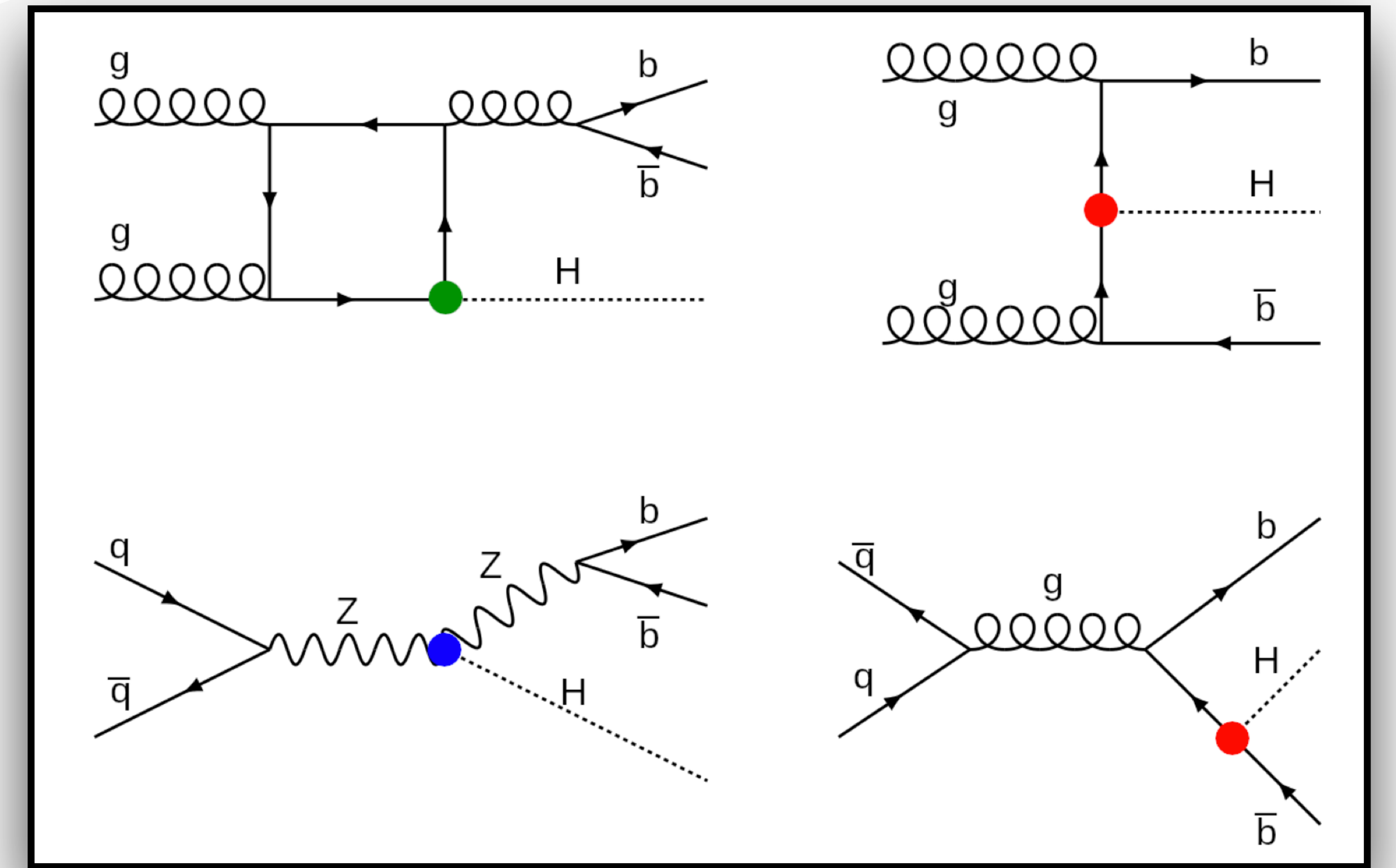
- The continuous background is modelled with a **data driven approach**.
- The functional form and normalisation are extracted by fitting the data.
- An F-test is performed to chose from different orders and families of analytical functions.
- A **discrete nuisance parameter** is used to extract an uncertainty due to the choice of one functional form over the others.
- Normalisation is extracted from data.



bbH associated production

- Cross section components coming from the different processes.
- Inclusive measurement: The different contributions to the signal are scaled by varying proportionally the y_b^2, y_t^2 and $y_b y_t$ terms.
  - Infer limits on the Higgs coupling structure → done by introducing the coupling scaling parameters k_t and k_b, and performing a likelihood ratio scan over the $k_t - k_b$ parameter space.
  - b-quark contribution to the quark loop in the y_t^2 process, are accounted for by scaling it by $1.04k_t^2 - 0.04k_b k_t + 0.002k_b^2$, while the y_b^2 contribution and the interference term are scaled by k_b^2 and $k_b k_t$ respectively.

term	σ(pb)
y_t^2	1.040 (+0.468 -0.489)
y_b^2	0.482 (+0.048 -0.070)
$y_b y_t$	– 0.033 (+0.007 -0.008)



bbH associated production

- The fit is performed on the BDT score distributions for both signal and background categories.
- Lower-score regions are still dominated by background processes,
- Higher BDT score regions show an increasing contribution from bbH process in final states with τ leptons or W bosons.
- There must be a $e\mu$, $e\tau_h$, $\mu\tau_h$, or $\tau_h\tau_h$ pair with opposite electric charge.
- No additional electrons or muons may be present in the event.
- The leptons and τ_h candidates must be separated by $\Delta R >$ 0.5(0.3) in the $e\tau_h$, $\mu\tau_h$, $\tau_h\tau_h$ ($e\mu$) channels.
- There must be at least one, and no more than two, b-tagged jet.

Variable	$e\mu$	$e\tau_h$	$\mu\tau_h$	$\tau_h\tau_h$
$m_{\tau\tau}$	×	✓	✓	✓
m_{vis}	✓	✓	✓	✓
Collinear mass	×	✓	✓	×
D_ζ	✓	✓	✓	×
$\Delta\eta$ between lepton and τ_h	×	✓	✓	×
Total transverse mass	✓	×	×	×
Di-τ p_T	✓	✓	✓	✓
Electron p_T	✓	×	×	×
Muon p_T	✓	×	×	×
p_T of leading τ_h	×	×	×	✓
p_T of trailing τ_h	×	×	×	✓
Transverse mass	×	✓	✓	×
Number of b-jets	✓	×	×	✓
p_T of leading b-jet	✓	✓	✓	✓
p_T of trailing b-jet	×	✓	✓	×
B-tag score for leading b-jet	×	✓	✓	✓
$\Delta\eta$ between di-τ p_T and leading b-jet	×	✓	✓	×
B-tag score for trailing b-jet	×	✓	✓	✓
Number of jets	✓	×	×	✓
p_T of leading jet	✓	×	×	✓
p_T of trailing jet	✓	×	×	✓
Di-jet invariant mass	×	×	×	✓
Di-jet $\Delta\eta$	✓	×	×	✓
p_T^{miss}	×	×	×	✓

$H + W^{\pm}W^{\pm}$ boson production through VBS

- All the MC events passing the signal region and control region selection criteria are used for BDT training and testing.
- The allowed range of the Xbb score is extended to include (0.3,0.9], leaving all other signal region requirements unchanged. This relies on the lack of correlation between the Xbb score and the BDT input variables.
- The signal events are weighted with the cross-section weight associated to the κVV = 4.5 point.

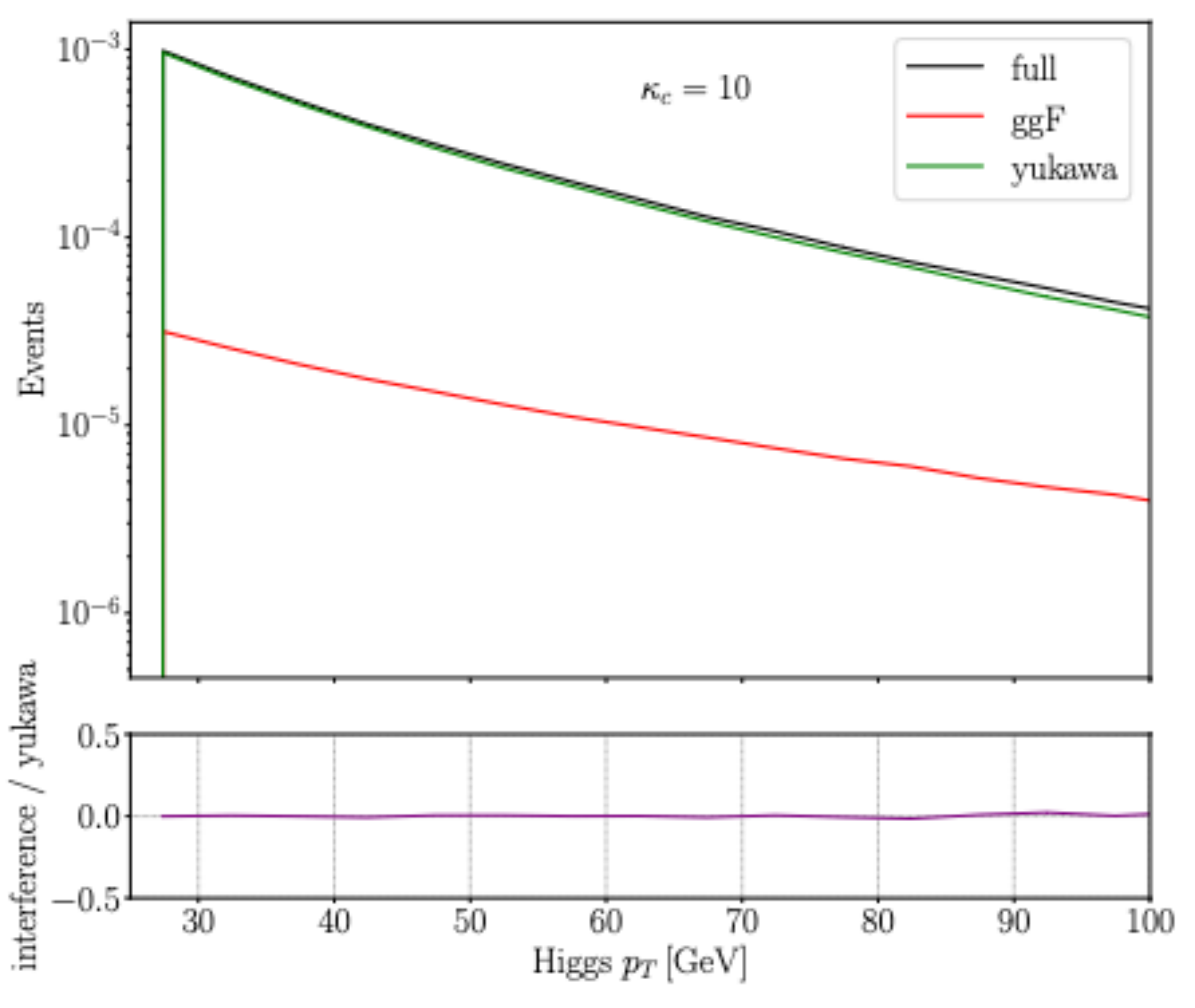
Shorthand	Description
η_J	η of the leading merged jet
$p_{\mathrm{T},J}$	p_{T} of the leading merged jet
$p_{\mathrm{T},jj}$	p_{T} of the VBS-jet system
P_{j_0}	magnitude of the three-momentum of the leading VBS jet
P_{j_1}	magnitude of the three-momentum of the subleading VBS jet
$M_{\ell\ell}$	invariant mass of the SS dilepton system
p_{T,ℓ_0}	p_{T} of the leading lepton
p_{T,ℓ_1}	p_{T} of the subleading lepton
$E_{\mathrm{T}}^{\mathrm{miss}}$	missing transverse energy
L_T	scalar sum of $p_{\mathrm{T},\ell_0}, p_{\mathrm{T},\ell_1}$, and $E_{\mathrm{T}}^{\mathrm{miss}}$
S_T	scalar sum of $p_{\mathrm{T},J}$ and L_T

MC samples: H+c simulation

Interference term:

- The interference term negligibility has been further studied following a request from the ARC.
- Simulation at LO using the `heft` UFO model both in 3FS and 4FS:
  - $full \rightarrow$ all the diagrams, $ggF \rightarrow$ only the non$-y_c$ dependant, $yukawa \rightarrow$ only the y_c^2 dependent, $int. \rightarrow full - ggF - yuk$.
- The **interference component**:
  - **Is small**, especially for high k_c values (table).
  - **Is constant** also looking **differentially** w.r.t. $p_T(H)$ (plot).
- The relative interference contribution scales $\sim 1/k_c$.

cross-section (fb)	$\kappa_c = 1$	$\kappa_c = 3$	$\kappa_c = 10$
full	0.267	0.631	4.348
ggF	0.228	0.228	0.229
yukawa	0.042	0.412	4.15
interference	-0.003	-0.01	-0.031
interference/yukawa	-0.075	-0.023	-0.007



MC samples: H+c simulation

Loop induced part:

- Contributions from loop induced diagrams that are not included in the signal simulation have also been studied.
- These processes don't feature any charm quark in the initial or final state, thus they don't interfere with the rest of the H+c simulation.
- In the 4FS due to the fact that c-quarks are massless these contributions are effectively zero (helicity conservation).
- In 3FS they yield non-zero contribution $O(5\ fb)$ further reduced by the $n_{jet} \geq 1$ requirement.

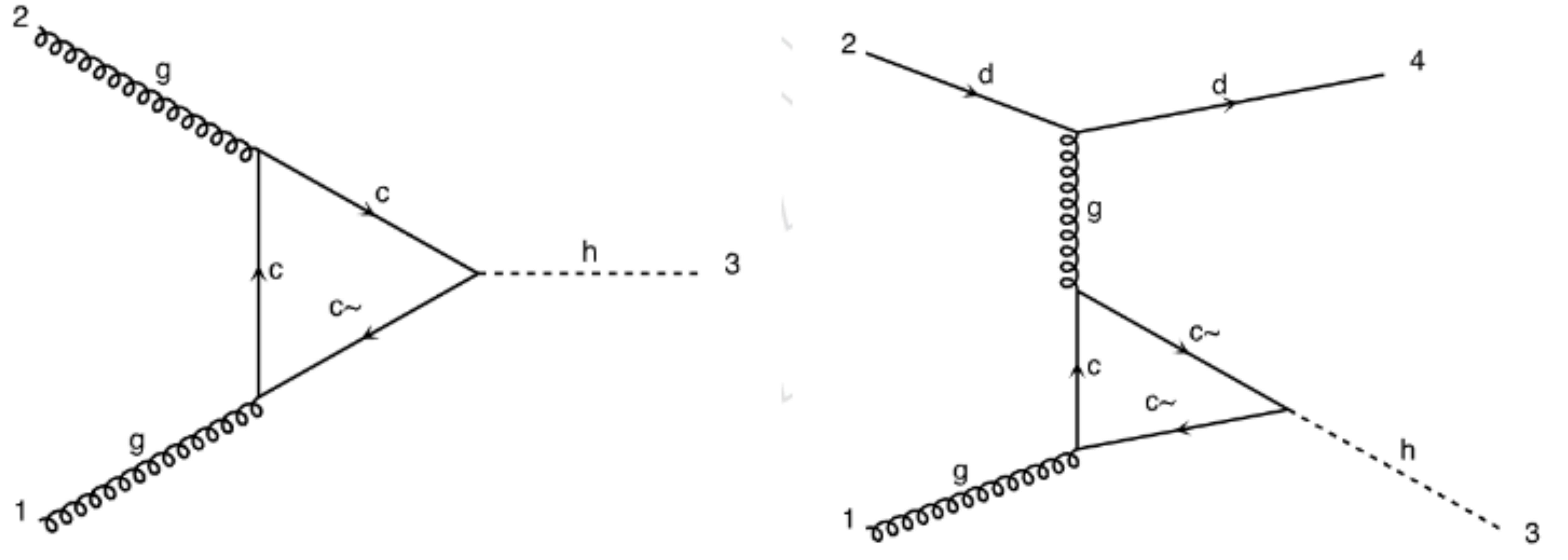


Figure 6: Examples of loop induced diagrams with ch coupling