Search for non-resonant Higgs boson pairs production in the bb$\tau\tau$ final state in CMS

Jona Motta (LLR, École Polytechnique)
on behalf of the CMS Collaboration

Search for non-resonant Higgs boson pairs production in the bb$\tau\tau$ final state in CMS

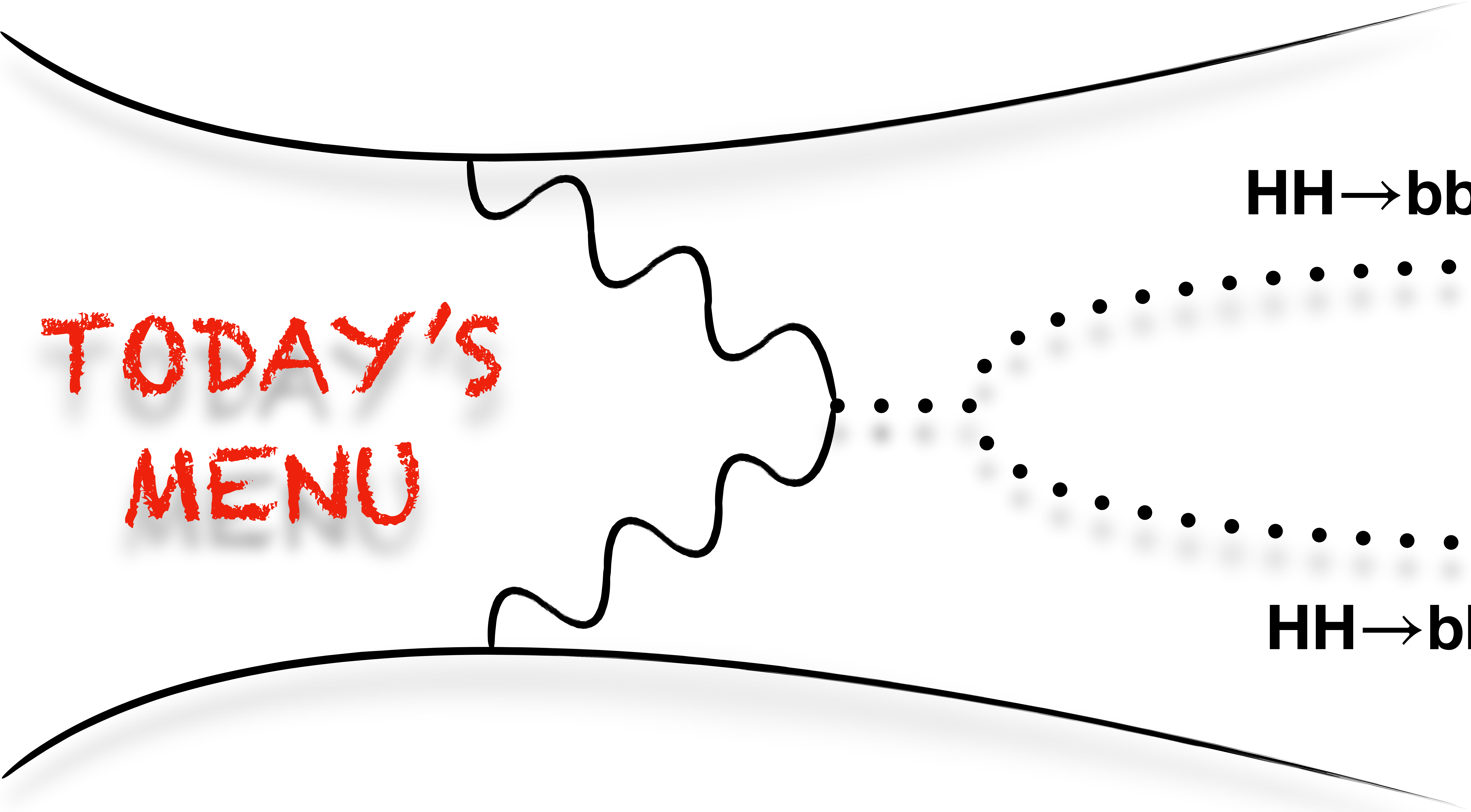
NEW RESULT!
ONE OF FIRST EVER
PUBLIC PRESENTATIONS

Jona Motta (LLR, École Polytechnique)
on behalf of the CMS Collaboration

ÉCOLE POLYTECHNIQUE
IP PARIS

Laboratoire Leprince-Ringuet

CMS



Introduction: what, why, where?

HH→bb$\tau\tau$ analysis strategy

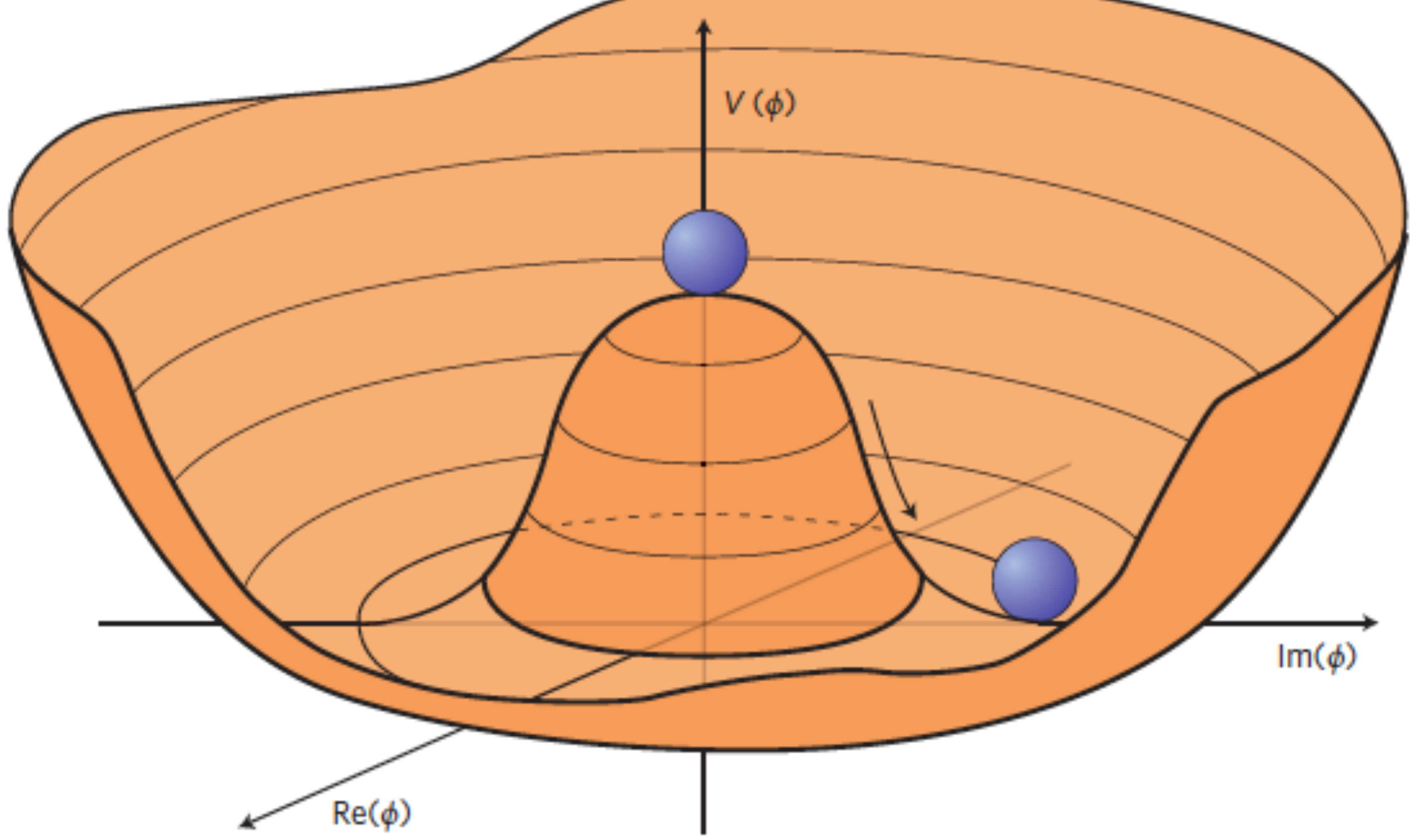
HH→bb$\tau\tau$ analysis results

Conclusions and outlook

Introduction: what, why, where?

What, why, and where

The Higgs mechanism is the most economical way to endow fundamental particles with mass while keeping the SM gauge invariant and predictive



The **Higgs field** is responsible for the spontaneous breaking of electroweak symmetry

$$V_H = \mu^2 + \frac{\mu^2}{v}H^3 + \frac{\mu^2}{4v}H^4 - \frac{1}{4}\mu^2 v^2$$
$$= \frac{1}{2}m_H^2 + \lambda_{HHH} v H^3 + \lambda_{HHHH} H^4 - \frac{1}{8}m_H^2 v^2$$

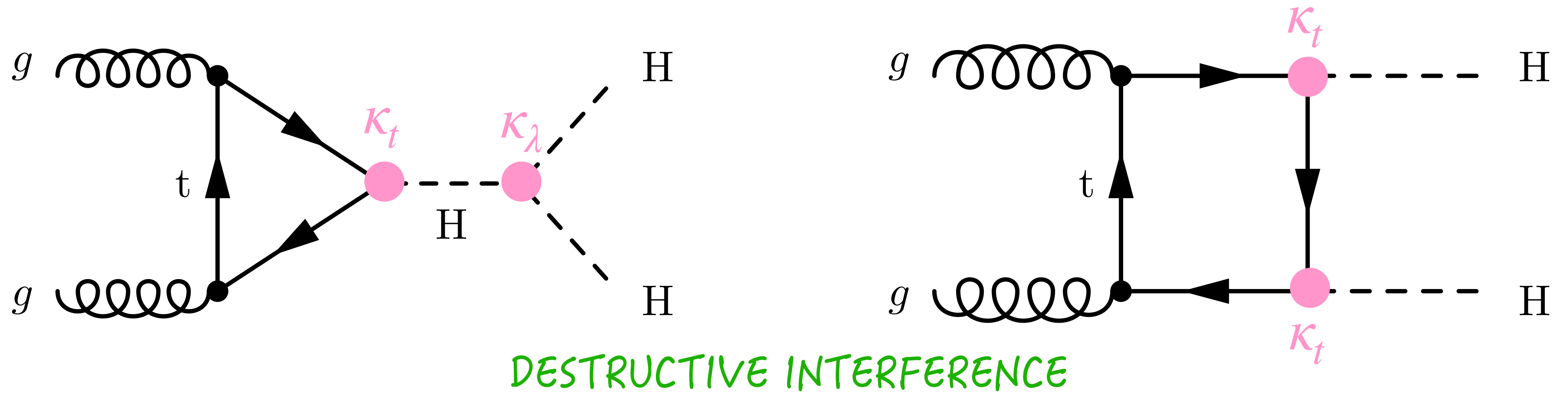
$$\lambda_{HHH} = 4\lambda_{HHHH} = \frac{m_H^2}{v^2}$$

only parameter regulating field's shape

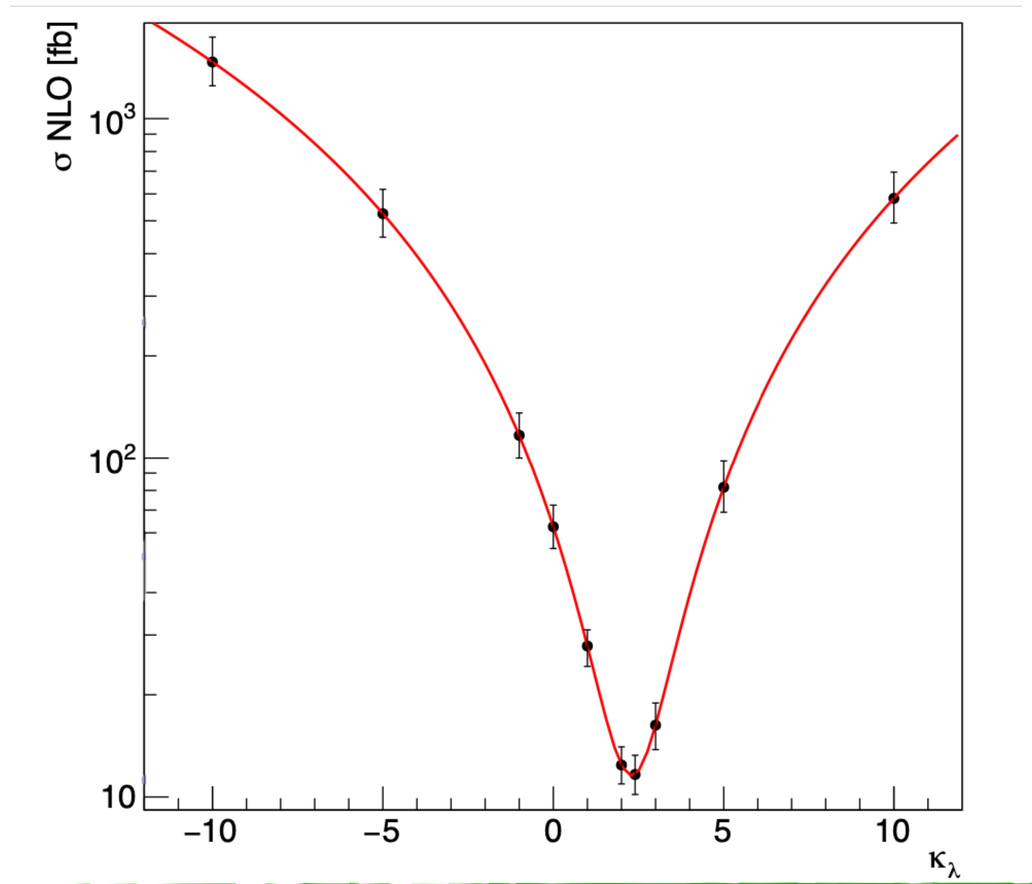
What, why, and where

1. Access the trilinear Higgs coupling λ_{HHH} **to study EWSB**
2. Set limits on the two main production mechanisms: **ggHH** and **qqHH**
3. Test deviation from the SM couplings (κ-framework): κ_λ , κ_t , κ_V , κ_{2V}

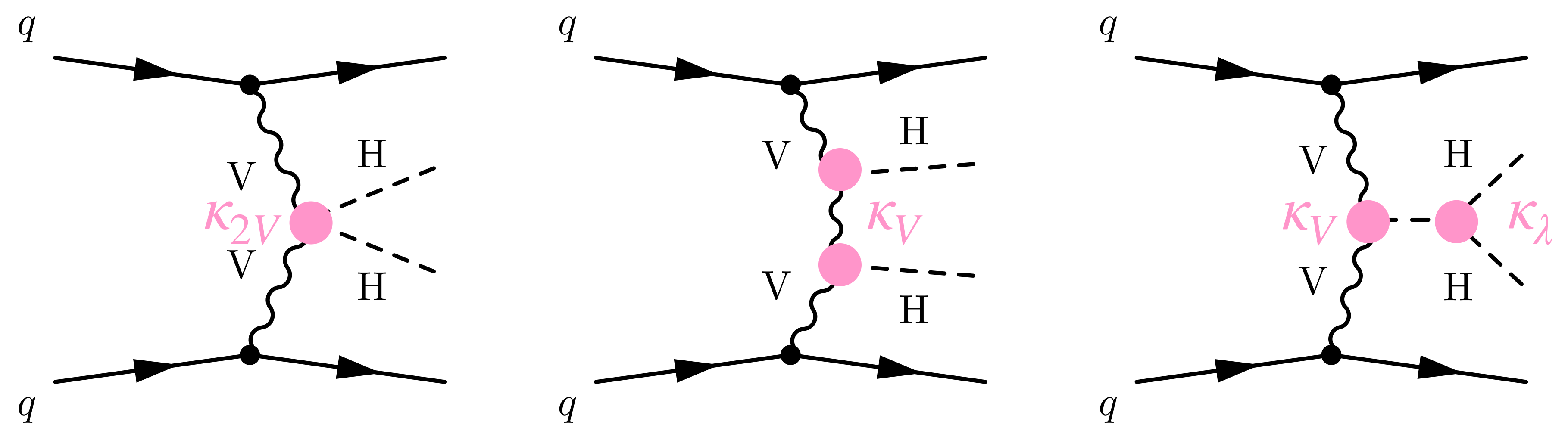
Gluon Fusion (ggHH)



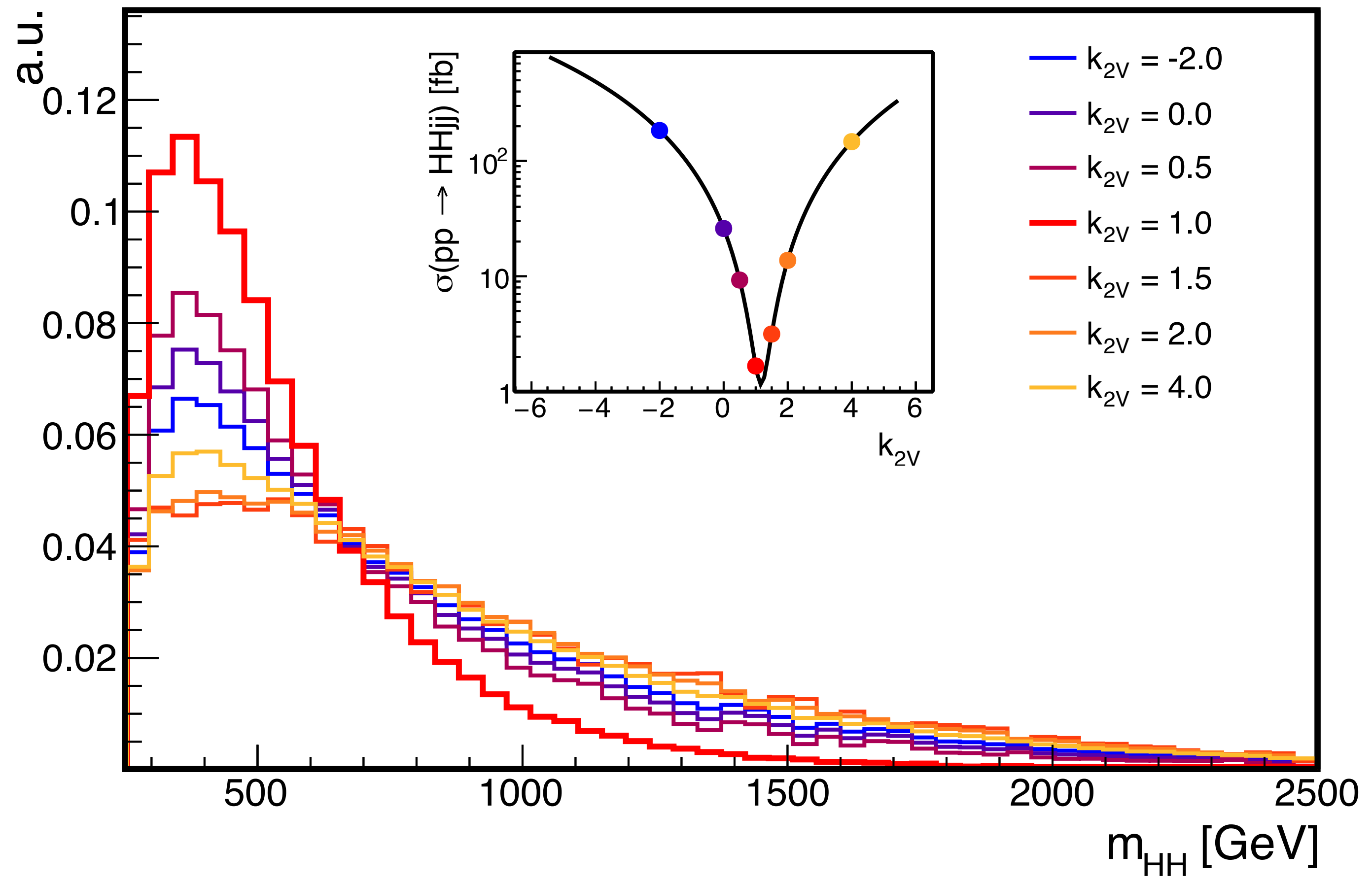
$\sigma^{ggHH}_{NNLO\text{-}FTapprox}(\sqrt{s} = 13\text{TeV}, m_H = 125\text{GeV}) \approx 31\text{fb}$



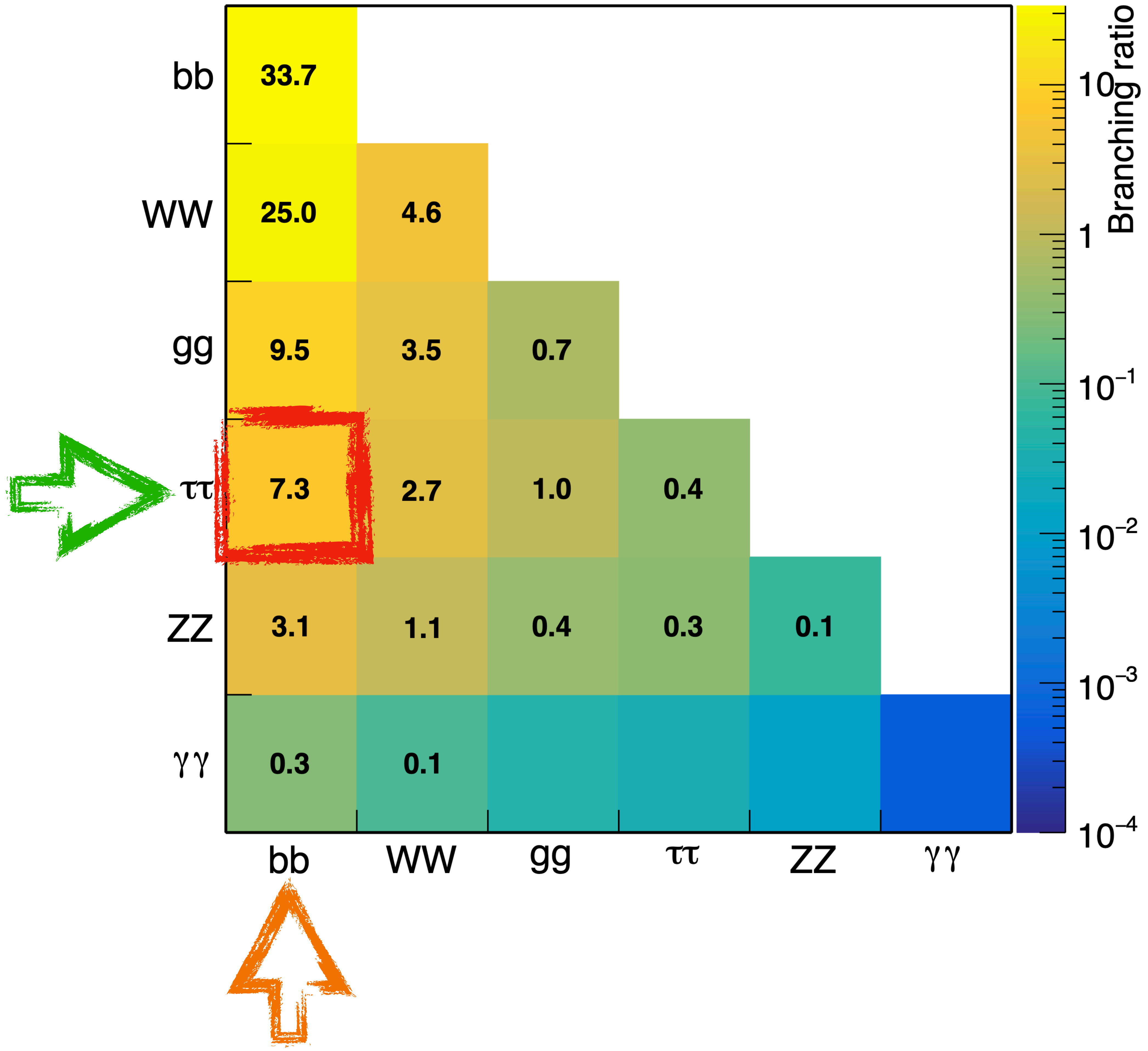
Vector Boson Fusion (qqHH)



$\sigma^{qqHH}_{N3LO}(\sqrt{s} = 13\text{TeV}, m_H = 125\text{GeV}) \approx 1.7\text{fb}$



What, why, and where

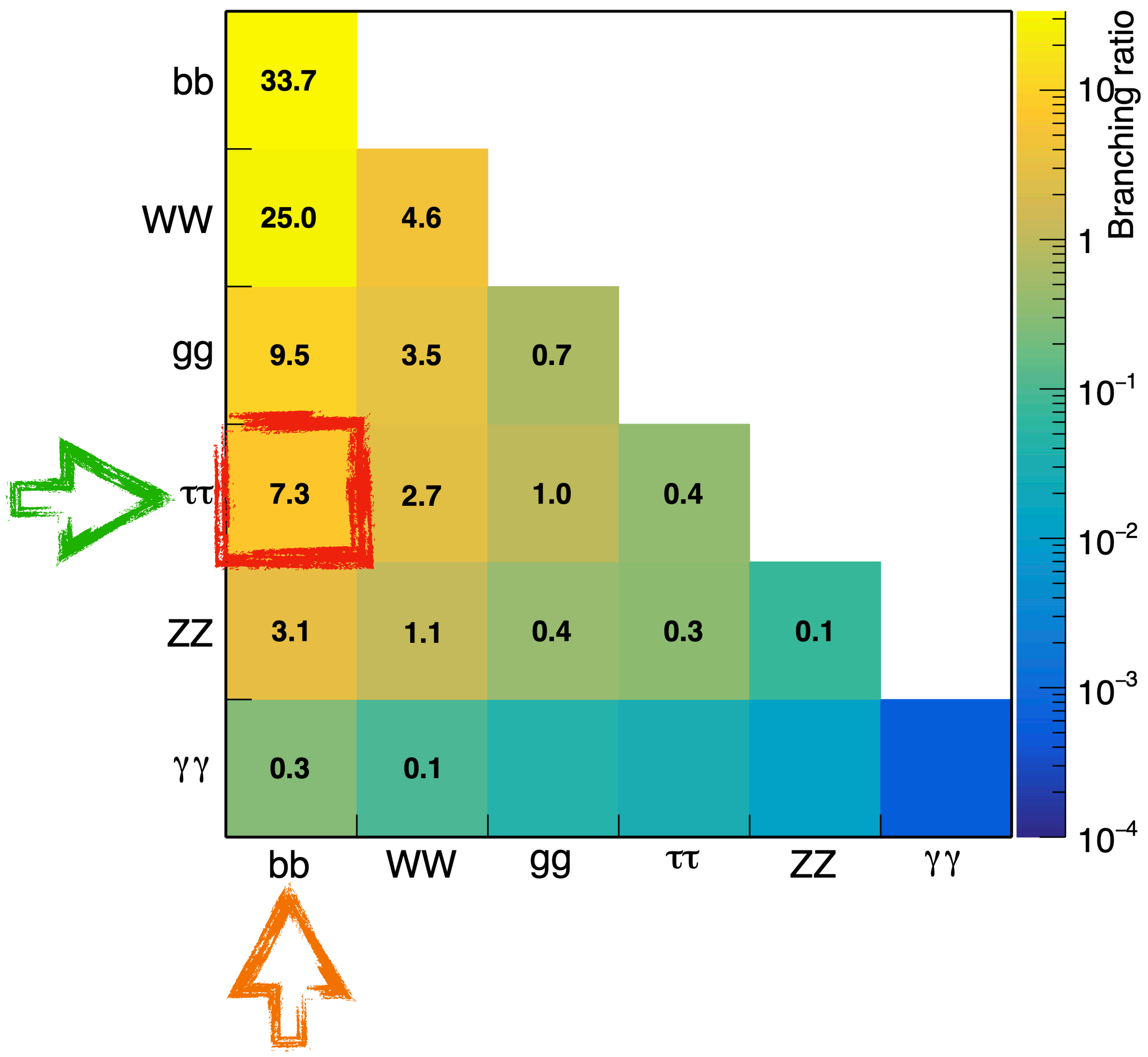


Why specifically this two legs?

1. **H→bb** : **large $\mathscr{B}$** (58%) thanks to big b-H coupling
2. **H→$\tau\tau$** : little $\mathscr{B}$ (6.3%) but **good purity**

⇛ bbττ good $\mathscr{B}$-purity compromise

What, why, and where



Why specifically this two legs?

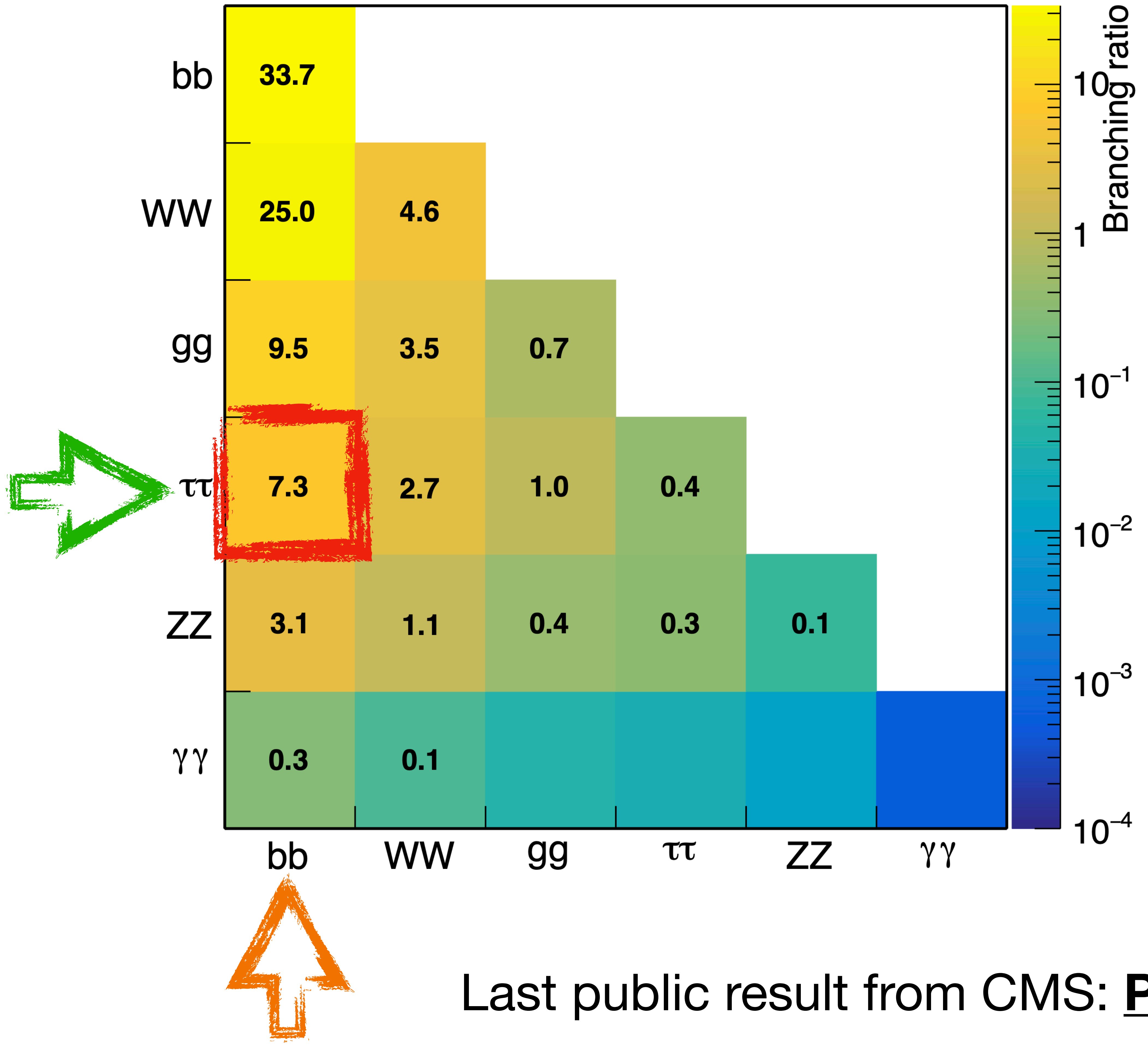
1. H→bb : **large $\mathscr{B}$** (58%) thanks to big b-H coupling
2. H→$\tau\tau$: little $\mathscr{B}$ (6.3%) but **good purity**

⇛ bbττ good $\mathscr{B}$-purity compromise

We focus on three specific channels of the ττ pair

⇒ $\tau_e\tau_h$, $\tau_\mu\tau_h$, and $\tau_h\tau_h$ which account for 88% of $\mathscr{B}_\tau$

What, why, and where



Why specifically this two legs?

1. **H→bb** : **large $\mathscr{B}$** (58%) thanks to big b-H coupling
2. **H→$\tau\tau$** : little $\mathscr{B}$ (6.3%) but **good purity**

⇒ bbττ good $\mathscr{B}$-purity compromise

We focus on three specific channels of the ττ pair

⇒ $\tau_e\tau_h$, $\tau_\mu\tau_h$, and $\tau_h\tau_h$ which account for 88% of $\mathscr{B}_\tau$

Last public result from CMS: **PLB 778 (2018) 101** used 2016 partial Run 2 dataset

Today the improved analysis performed with the full Run 2 dataset CMS-PAS-HIG-20-010 will be presented

HH→bb$\tau\tau$ analysis strategy

Analysis overview

H→$\tau\tau$ candidate selection

- Selection of two well identified leptons
- Veto additional light third lepton
- Selection of best leptons pair: $e\tau_h$ - $\mu\tau_h$ - $\tau_h\tau_h$

Analysis overview

H→$\tau\tau$ candidate selection

- Selection of two well identified leptons
- Veto additional light third lepton
- Selection of best leptons pair: $e\tau_h$ - $\mu\tau_h$ - $\tau_h\tau_h$

Analysis improvements w.r.t. PLB 778 (2018) 101

- New **VBF H → $\tau\tau$ trigger**
- New cross **$e\tau_h$ - $\mu\tau_h$ triggers**
- New τ_h identification with **DeepTau algorithm** (CNN)

Analysis overview

H→$\tau\tau$ candidate selection

- Selection of two well identified leptons
- Veto additional light third lepton
- Selection of best leptons pair: $e\tau_h$ - $\mu\tau_h$ - $\tau_h\tau_h$

H→bb candidate selection

- Selection of two well identified b-jets
- Selection of best AK4 b-jets pair
- Selection of best AK8 jet with AK4 b-subjets

Analysis improvements w.r.t. PLB 778 (2018) 101

- New **VBF H → $\tau\tau$ trigger**
- New cross **$e\tau_h$ - $\mu\tau_h$ triggers**
- New τ_h identification with **DeepTau algorithm** (CNN)
- New b-jets identification with **DeepJet algorithm** (RNN)
- New b-jets pair selection with **HH-btag algorithm** (RNN)

Analysis overview

H→$\tau\tau$ candidate selection

- Selection of two well identified leptons
- Veto additional light third lepton
- Selection of best leptons pair: $e\tau_h$ - $\mu\tau_h$ - $\tau_h\tau_h$

H→bb candidate selection

- Selection of two well identified b-jets
- Selection of best AK4 b-jets pair
- Selection of best AK8 jet with AK4 b-subjets

HH categorisation

- Definition of a signal enriched region
- Discrimination of signal against background
- Categorisation of the selected events

Analysis improvements w.r.t. PLB 778 (2018) 101

- New **VBF H → $\tau\tau$ trigger**
- New cross **$e\tau_h$ - $\mu\tau_h$ triggers**
- New τ_h identification with **DeepTau algorithm** (CNN)
- New b-jets identification with **DeepJet algorithm** (RNN)
- New b-jets pair selection with **HH-btag algorithm** (RNN)
- New **multiclass VBF categorisation** (DNN)
- **Improved background modelling**

Analysis overview

H→$\tau\tau$ candidate selection

- Selection of two well identified leptons
- Veto additional light third lepton
- Selection of best leptons pair: $e\tau_h$ - $\mu\tau_h$ - $\tau_h\tau_h$

H→bb candidate selection

- Selection of two well identified b-jets
- Selection of best AK4 b-jets pair
- Selection of best AK8 jet with AK4 b-subjets

HH categorisation

- Definition of a signal enriched region
- Discrimination of signal against background
- Categorisation of the selected events

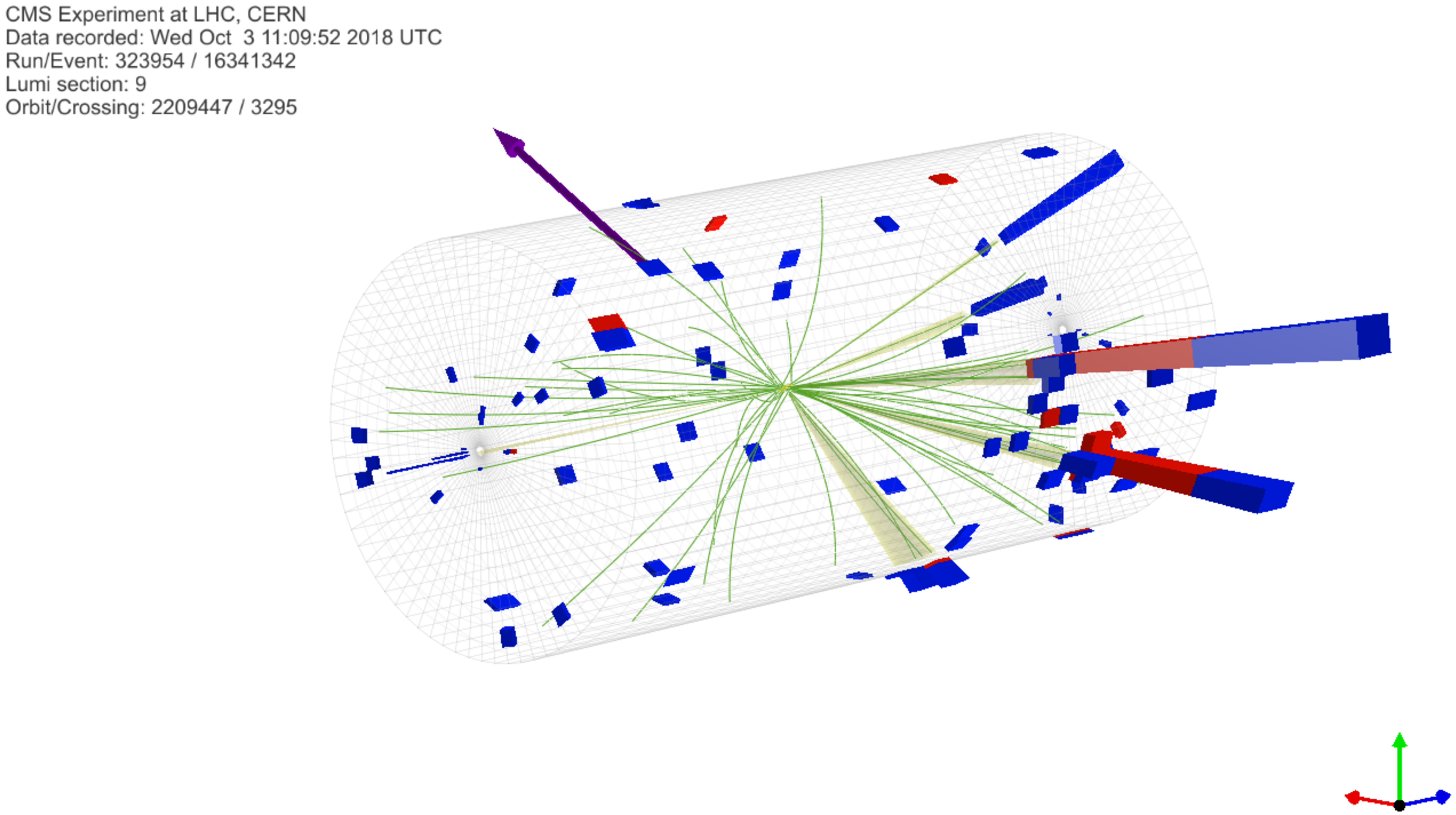
Signal extraction and limits

- Statistical model including all sources of uncertainties
- Signal extraction from shape of DNN discriminant

Analysis improvements w.r.t. PLB 778 (2018) 101

- New **VBF H → $\tau\tau$ trigger**
- New cross $e\tau_h$ - $\mu\tau_h$ **triggers**
- New τ_h identification with **DeepTau algorithm** (CNN)
- New b-jets identification with **DeepJet algorithm** (RNN)
- New b-jets pair selection with **HH-btag algorithm** (RNN)
- New **multiclass VBF categorisation** (DNN)
- **Improved background modelling**
- New **signal extraction algorithm** (DNN)
- Explicit **ggHH and qqHH limits setting**

H candidates selection

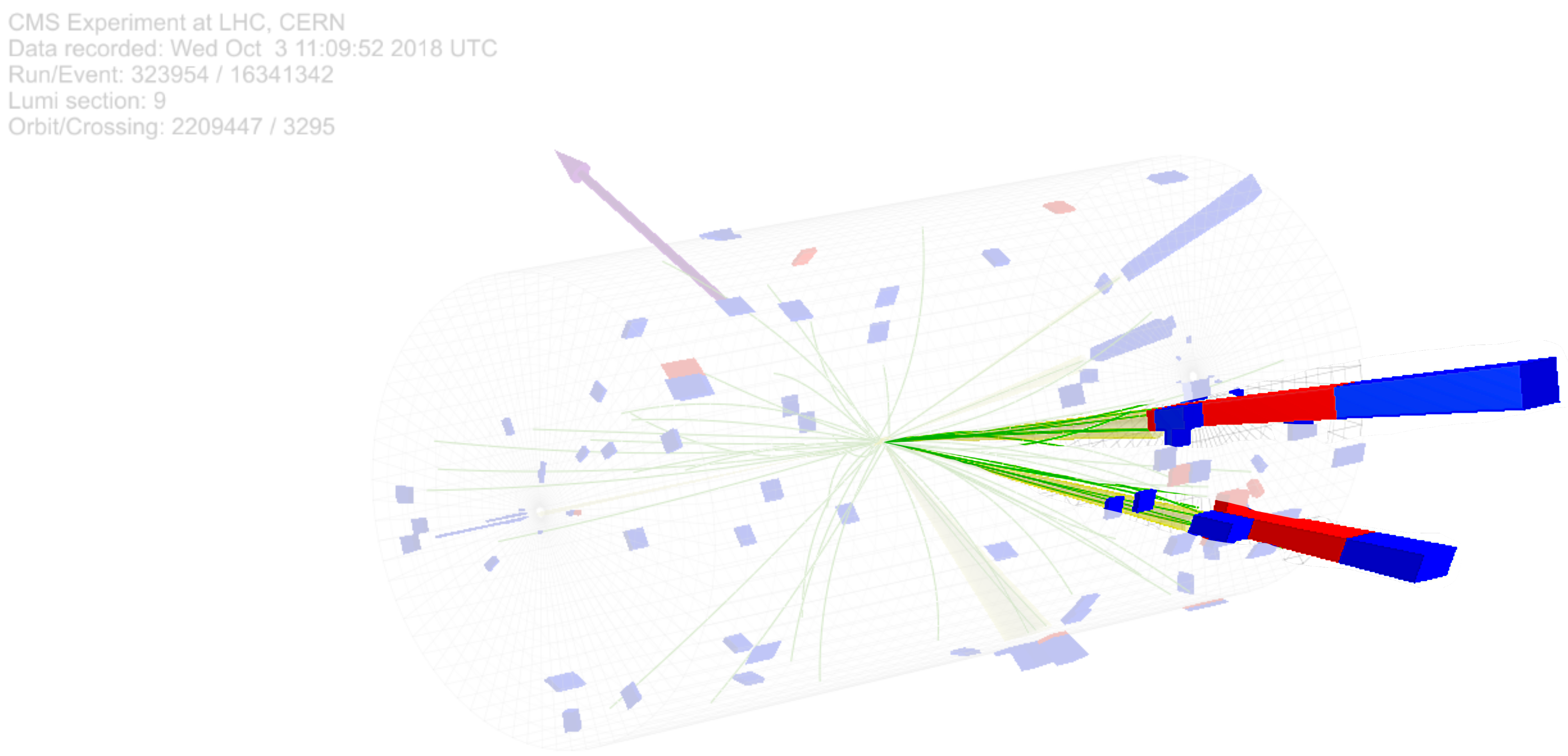


H candidates selection

$H \rightarrow \tau\tau$

⇛ reconstructed in 3 channels:

- ⇒ $e\tau_h$, $\mu\tau_h$, $\tau_h\tau_h$
- ⇒ $\boldsymbol{\tau_h}$ identified with **DeepTau ID** (CMS-TAU-20-001)

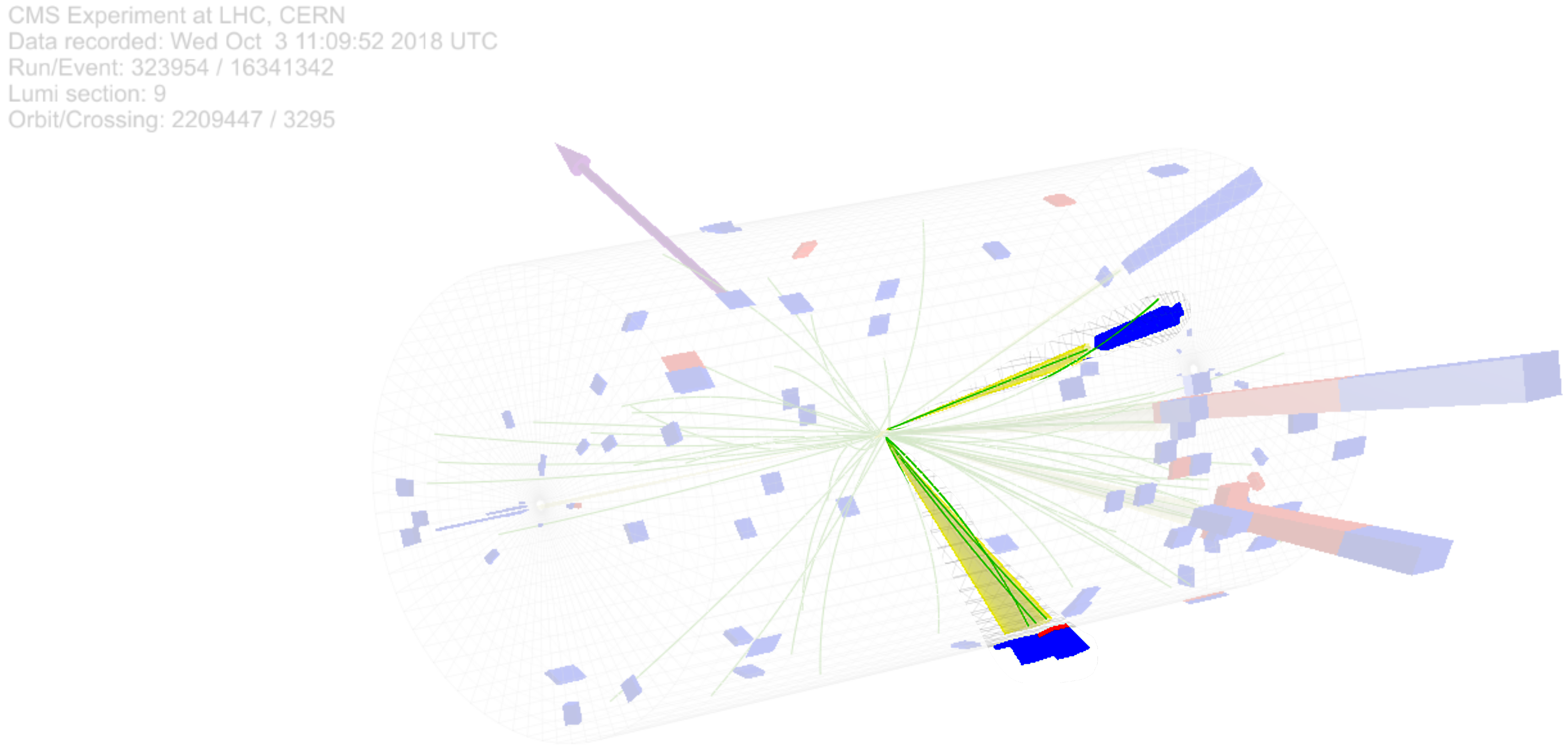


H candidates selection

$H \to \tau\tau$

⇛ reconstructed in 3 channels:

- ⇒ $e\tau_h$, $\mu\tau_h$, $\tau_h\tau_h$
- ⇒ $\boldsymbol{\tau_h}$ identified with **DeepTau ID** (CMS-TAU-20-001)



$H \to bb$

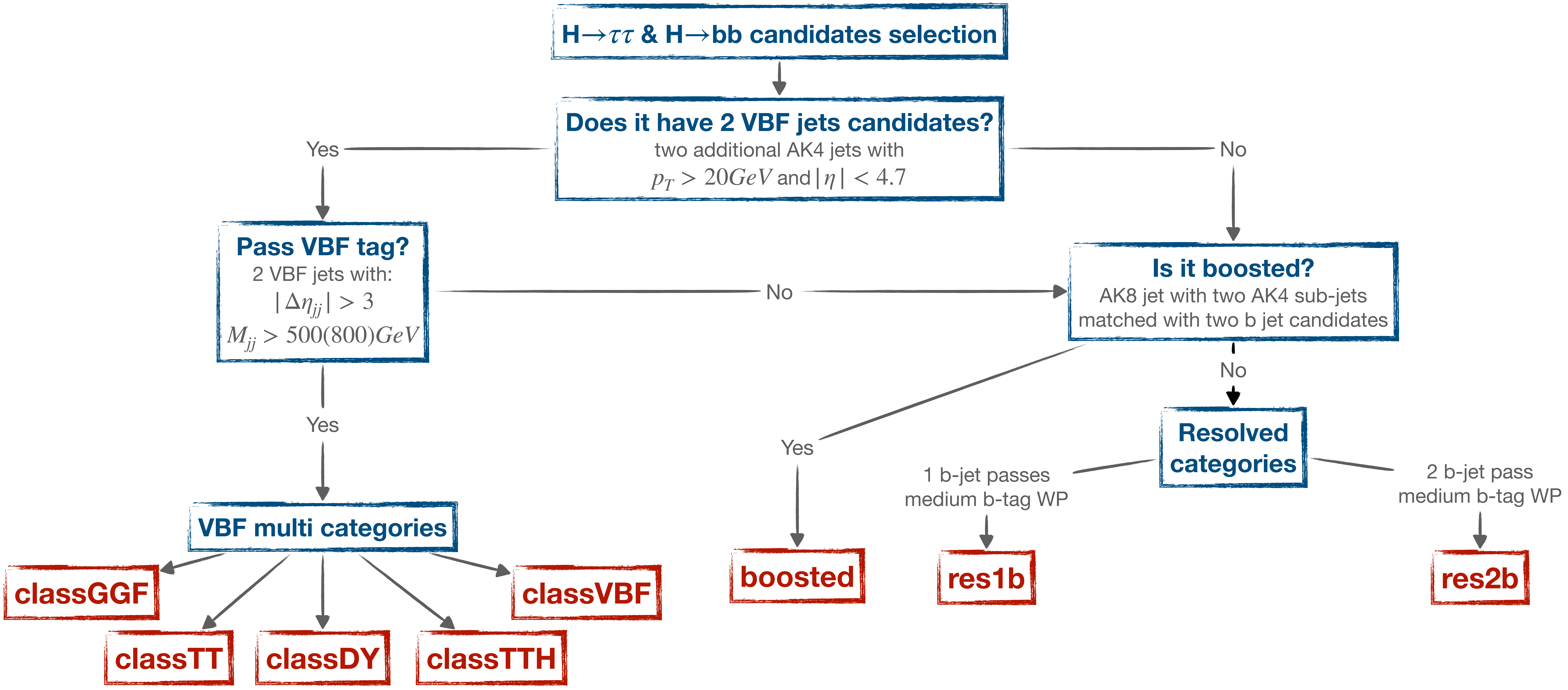
⇛ reconstructed as two AK4 jet or one AK8 jet with two AK4 sub-jets

- ⇒ **b-tagging** done with **DeepJet ID** (JINST 15 (2020) P12012)

⇛ best pair selection with **HH-btag tagger: RNN tagger** (tailored for this analysis)

- ⇒ H→bb candidate tagging efficiency ≈95%
- ⇒ m_{bb} resolution improved by 25%

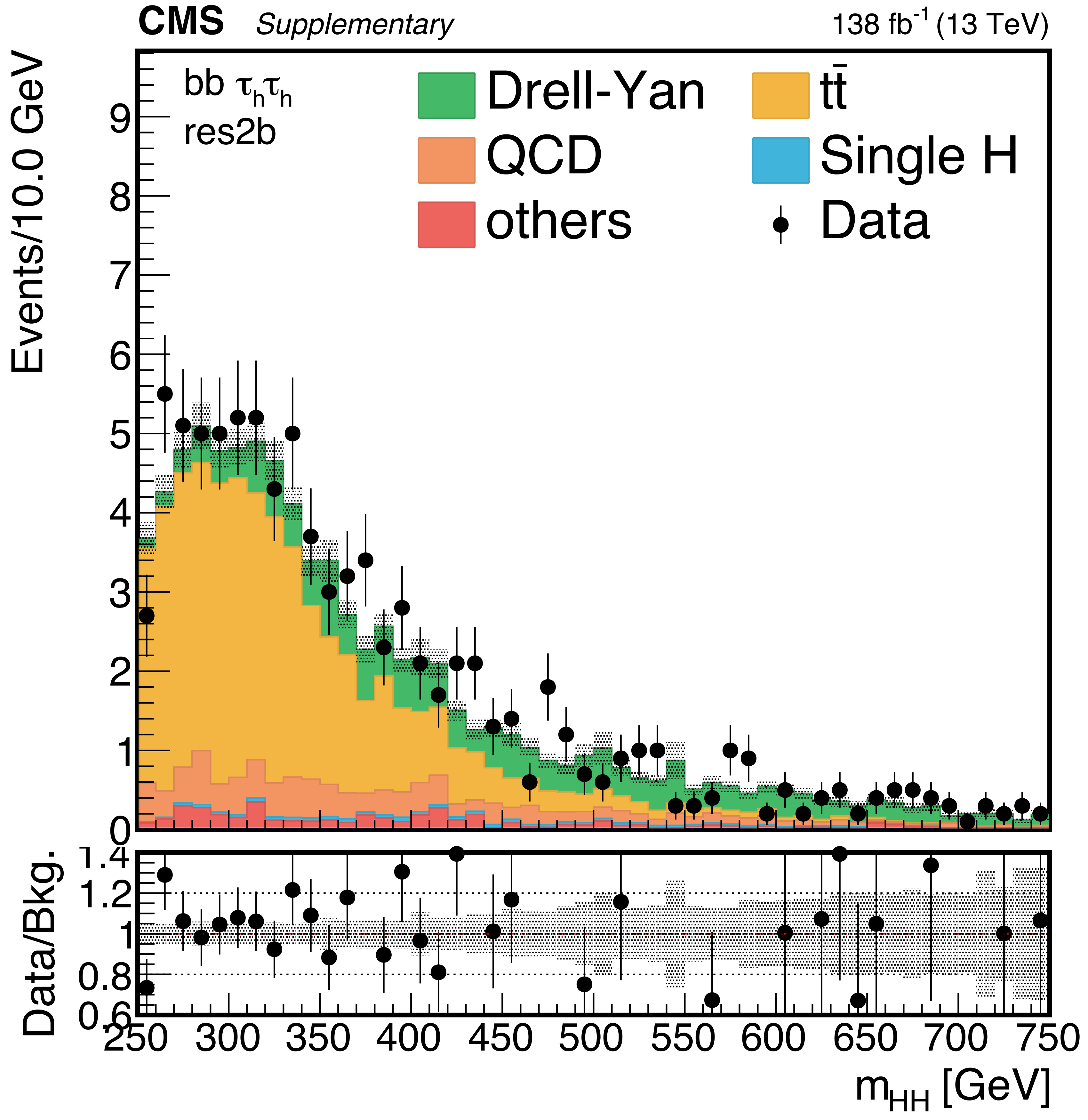
HH selection and categorisation



HH selection and categorisation

⇒ 8 categories × 3 $\tau\tau$ channels × 3 years = 72 categories

Background modelling



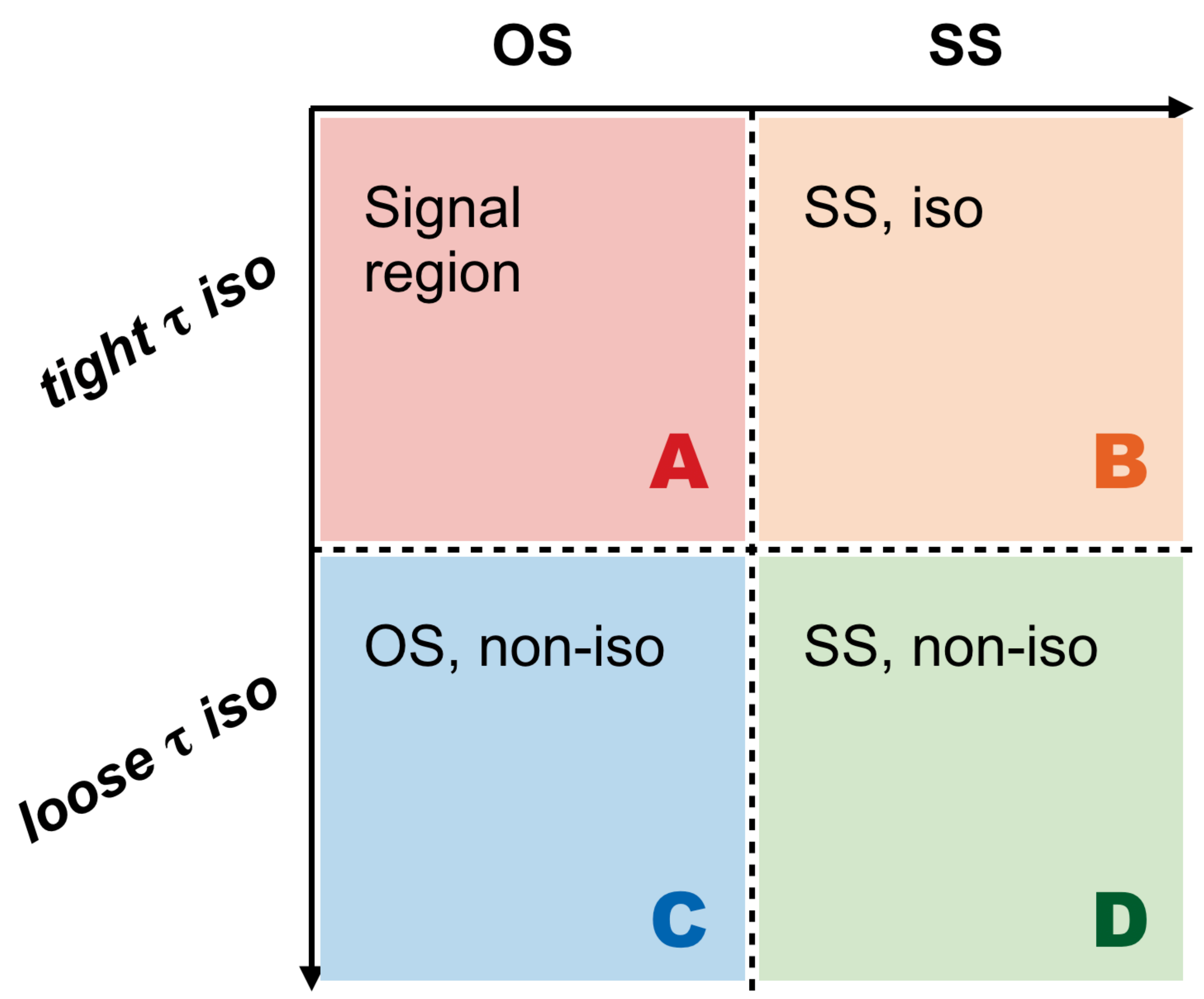
DY
Shape from MC
Normalisation from ZZ CR

$t\bar{t}$
Shape from MC
Normalisation from $t\bar{t}$ CR

Single Higgs
Shape & normalisation from MC

Others
Shape & normalisation from MC

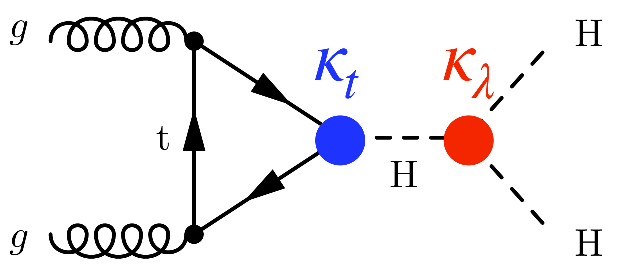
QCD
Shape & normalisation data-driven

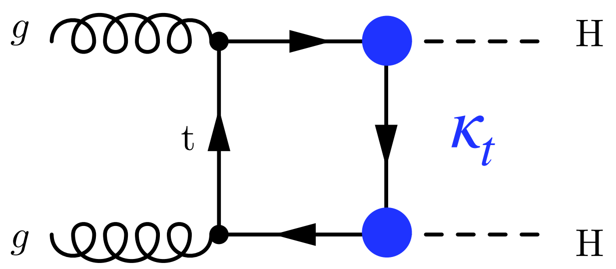


$$\left.\begin{array}{l} A' = C \cdot B/D \\ A'' = B \cdot C/D \end{array}\right\} A = \frac{A' + A''}{2}$$

Signal modelling and statistical interpretation

SIGNAL MODELLING





Signal modelling at inference stage

⇒ **3 ggHH samples solve the system**

$$A = \kappa_t \kappa_\lambda T + \kappa_t^2 B$$

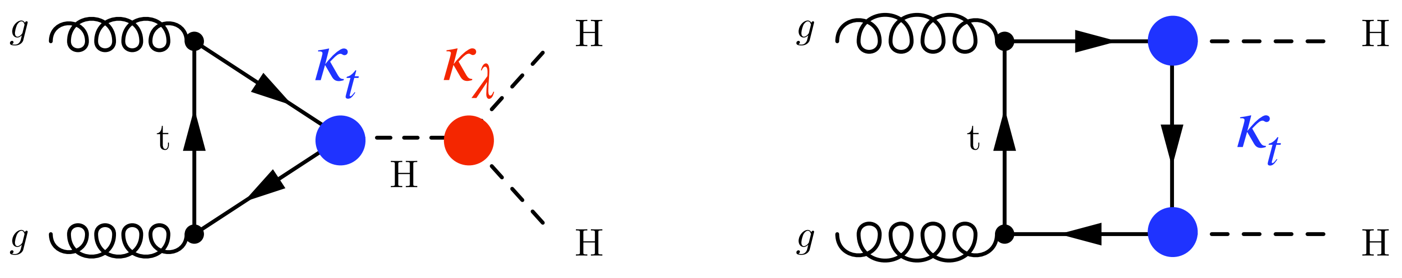
$$\sigma(\kappa_t, \kappa_\lambda) \propto |A|^2 = \kappa_t^2 \kappa_\lambda^2 |T|^2 + \kappa_t^4 |B|^2 + \kappa_t^3 \kappa_\lambda |T^*B + B^*T|$$

$$= \kappa_t^2 \kappa_\lambda^2 t + \kappa_t^4 b + \kappa_t^3 \kappa_\lambda i$$

= triangle +box +interference

Signal modelling and statistical interpretation

SIGNAL MODELLING



Signal modelling at inference stage

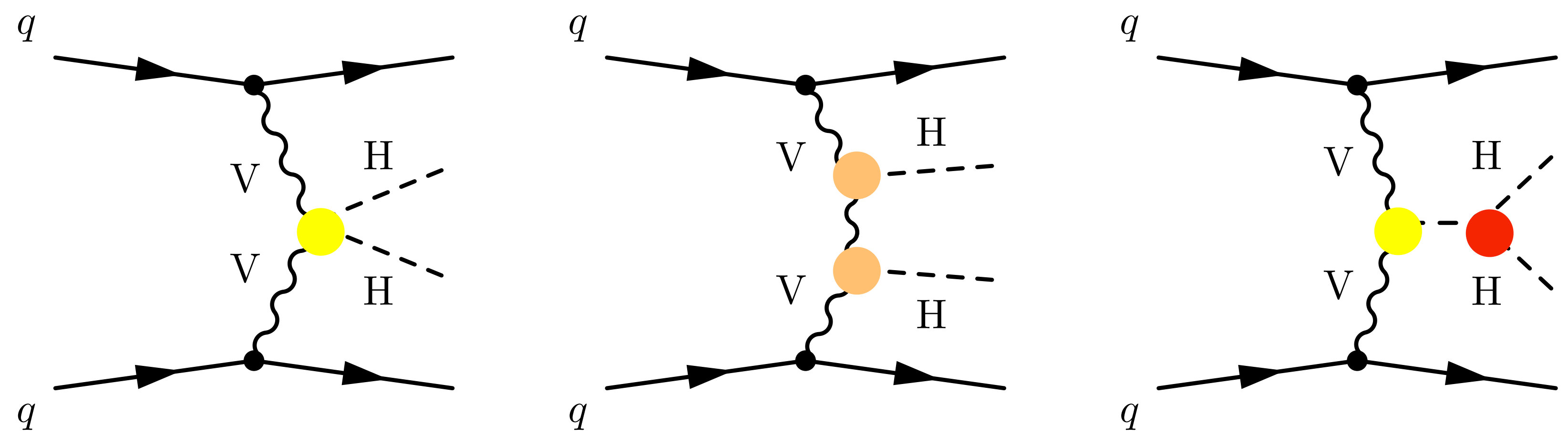
⇒ **3 ggHH samples solve the system**

$$A = \kappa_t \kappa_\lambda T + \kappa_t^2 B$$

$$\sigma(\kappa_t, \kappa_\lambda) \propto |A|^2 = \kappa_t^2 \kappa_\lambda^2 |T|^2 + \kappa_t^4 |B|^2 + \kappa_t^3 \kappa_\lambda |T^*B + B^*T|$$

$$= \kappa_t^2 \kappa_\lambda^2 t + \kappa_t^4 b + \kappa_t^3 \kappa_\lambda i$$

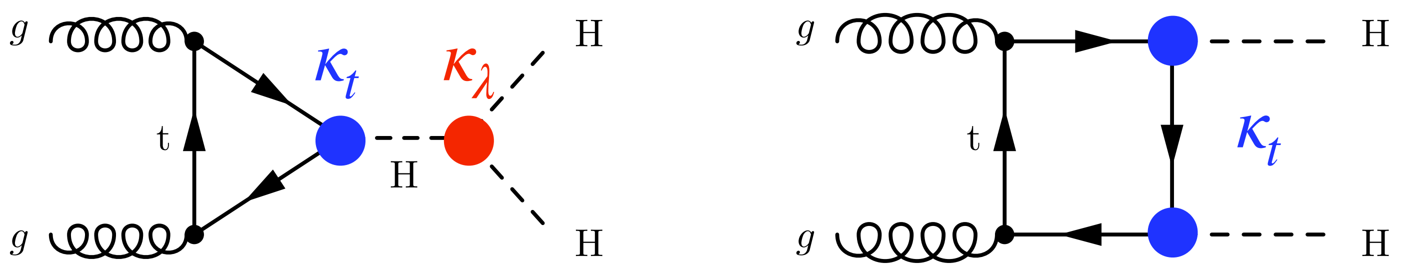
= triangle + box + interference



⇒ **6 qqHH samples solve the system**

Signal modelling and statistical interpretation

SIGNAL MODELLING



Signal modelling at inference stage

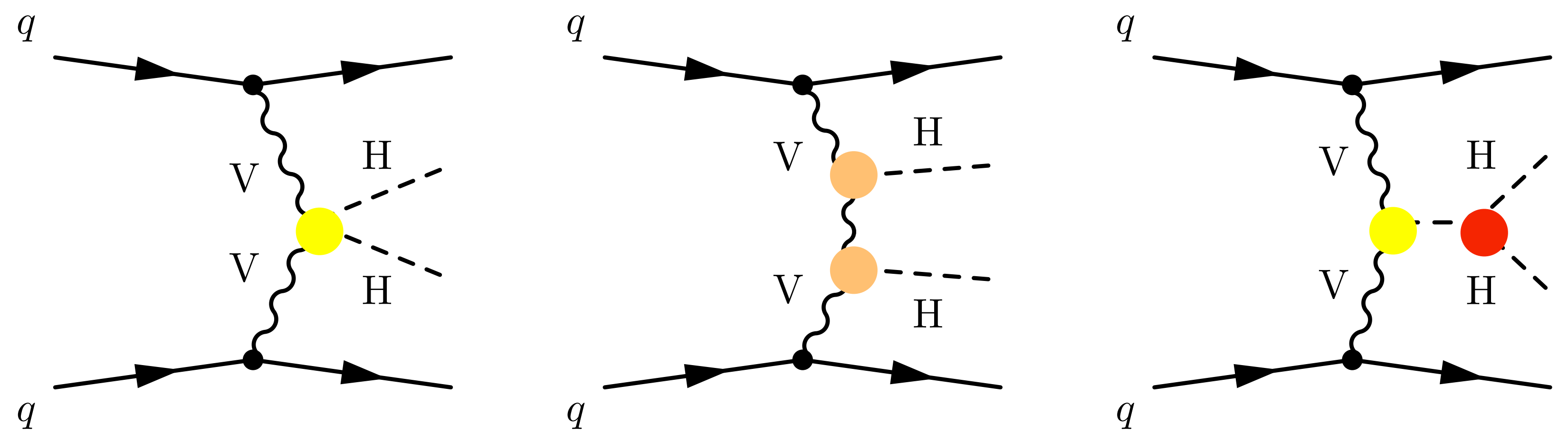
⇒ **3 ggHH samples solve the system**

$$A = \kappa_t \kappa_\lambda T + \kappa_t^2 B$$

$$\sigma(\kappa_t, \kappa_\lambda) \propto |A|^2 = \kappa_t^2 \kappa_\lambda^2 |T|^2 + \kappa_t^4 |B|^2 + \kappa_t^3 \kappa_\lambda |T^*B + B^*T|$$

$$= \kappa_t^2 \kappa_\lambda^2 t + \kappa_t^4 b + \kappa_t^3 \kappa_\lambda i$$

= triangle + box + interference



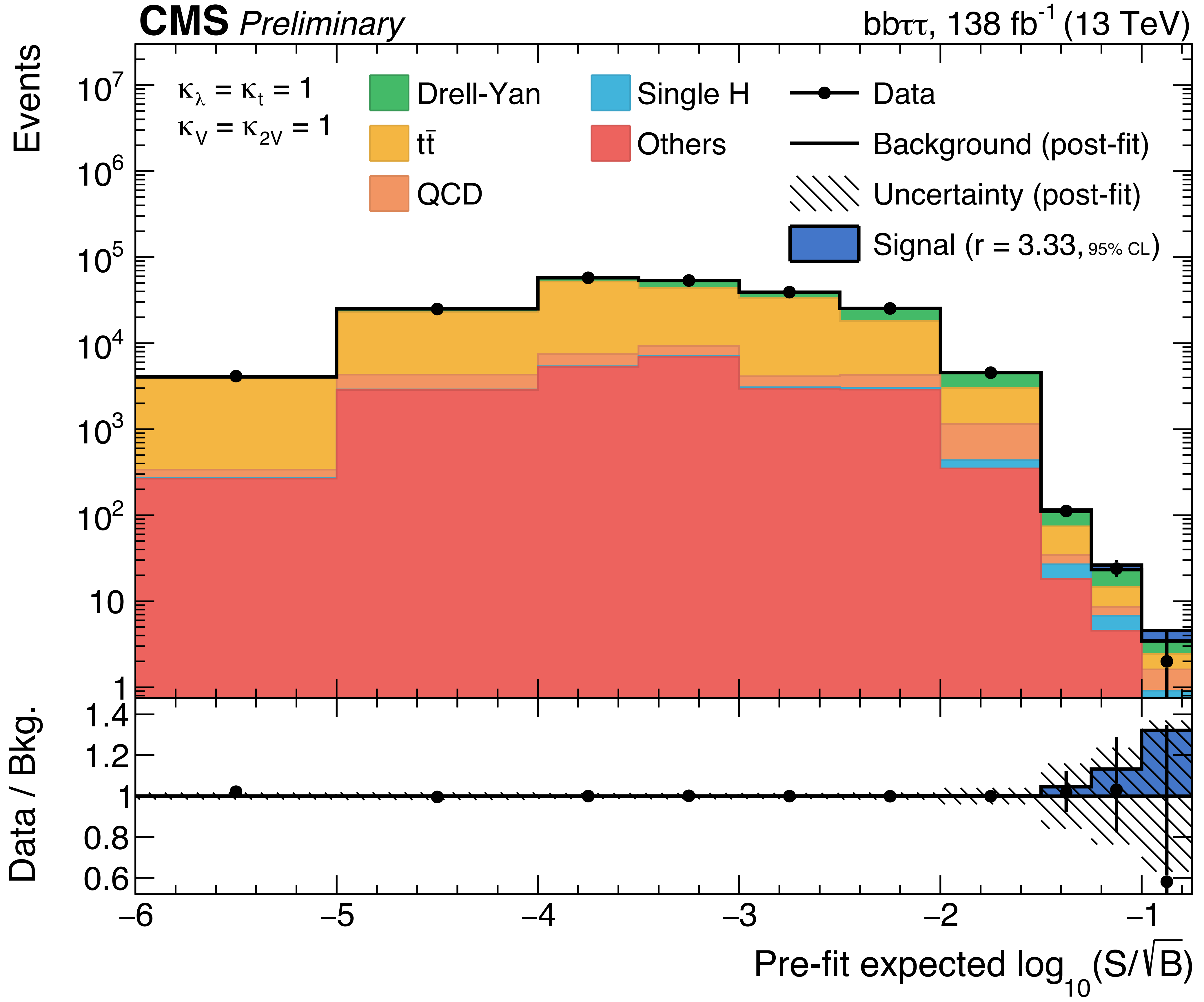
⇒ **6 qqHH samples solve the system**

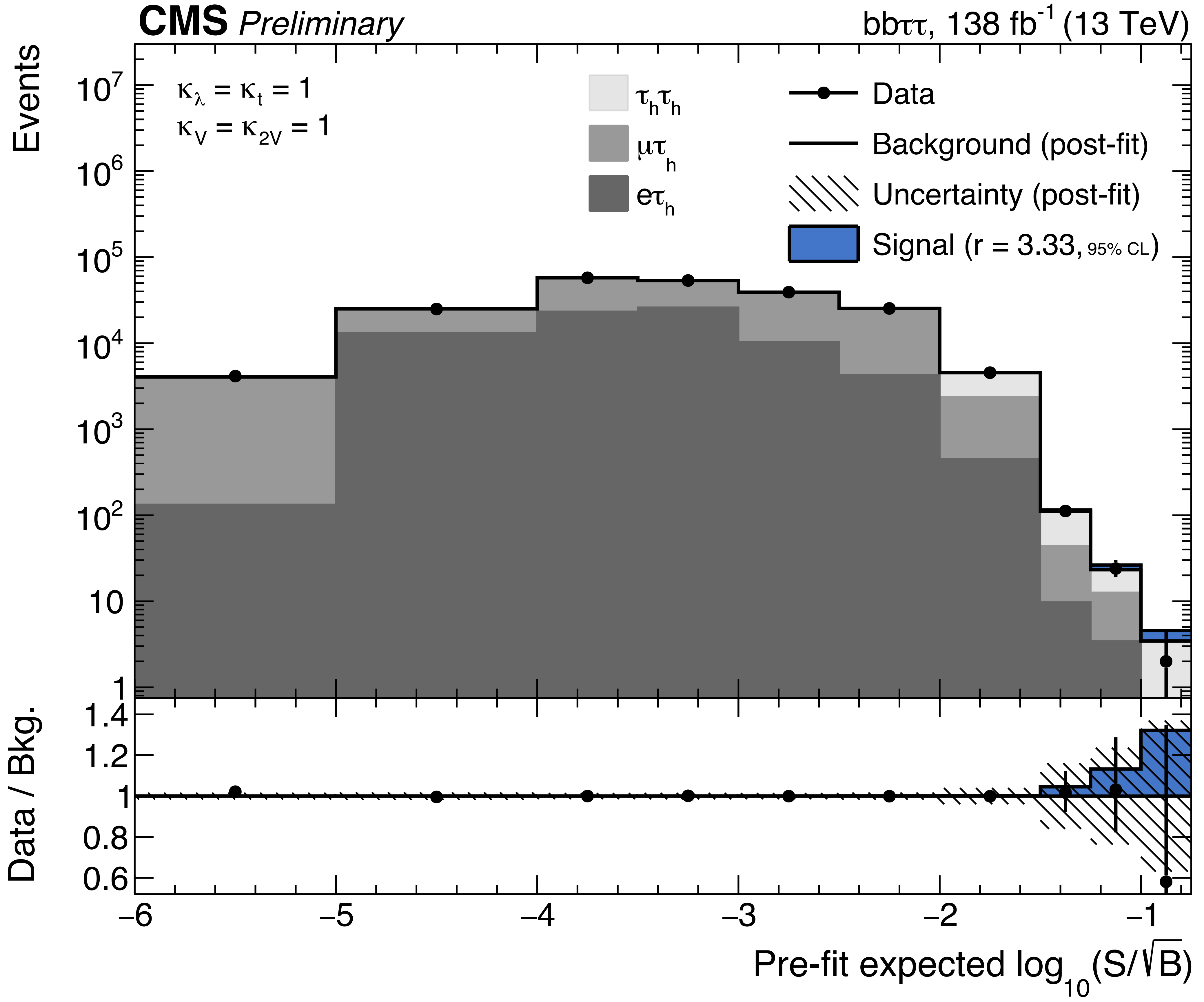
STATISTICAL INTERPRETATION

- A **single DNN** trained and used for all the categories of the analysis to extract the HH signal
- **Binned DNN** output used to **set upper limits** on the signal normalisations at **95% CL using profile-likelihood test statistic** and modified **frequentist CLs technique**, under **asymptotic approximation**

HH→bb$\tau\tau$ results

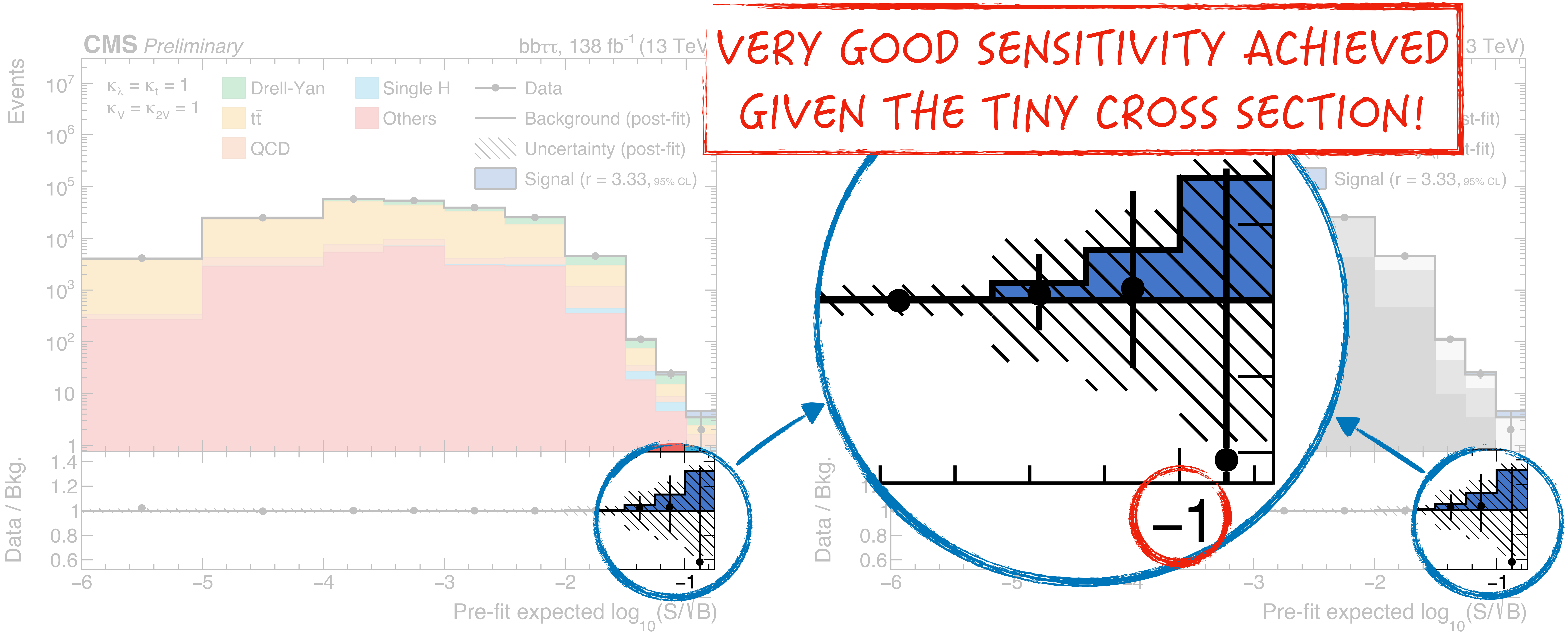
Pre-fit expected sensitivity





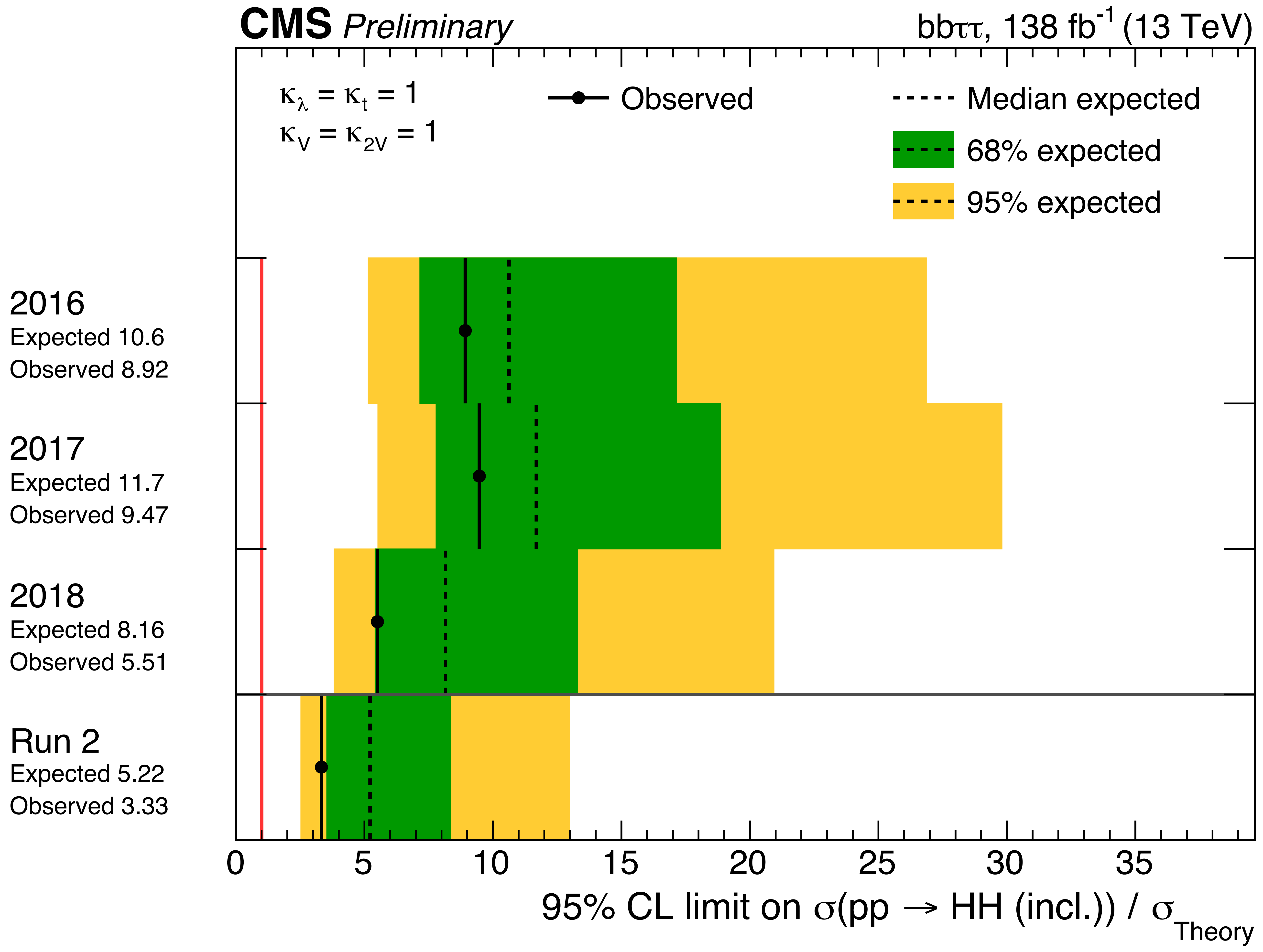
- Smooth background behaviour over sensitivity range
- Sensitivity smoothly goes as $\tau_e\tau_h < \tau_\mu\tau_h < \tau_h\tau_h$

Pre-fit expected sensitivity



- Smooth background behaviour over sensitivity range
- Sensitivity smoothly goes as $\tau_e\tau_h < \tau_\mu\tau_h < \tau_h\tau_h$

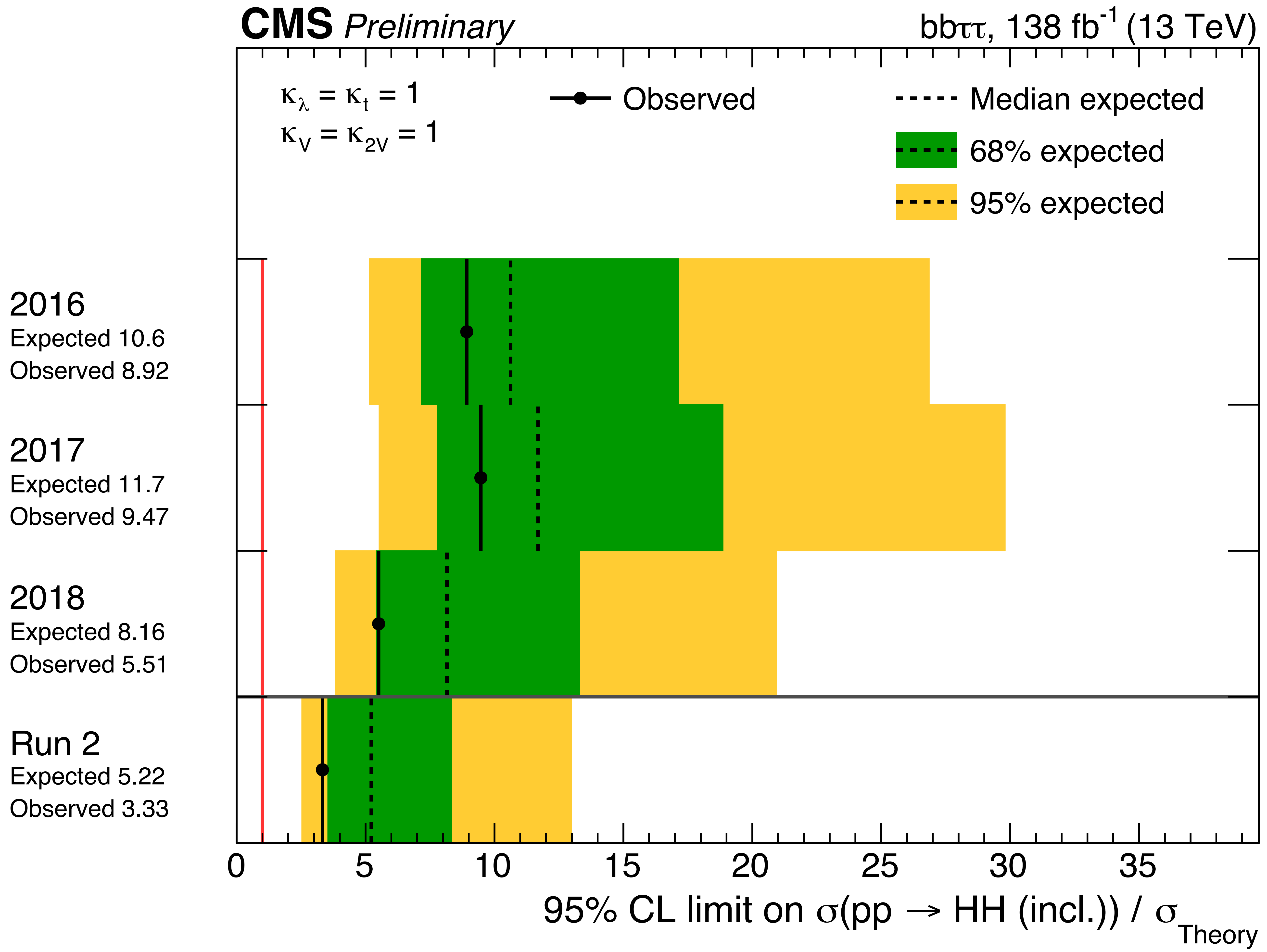
Standard model HH cross section



Observed (expected) 95% CL upper limit for the SM point

$$\sigma_{ggHH+qqHH} = 3.33(5.2) \cdot \sigma^{SM}_{ggHH+qqHH}$$

Standard model HH cross section



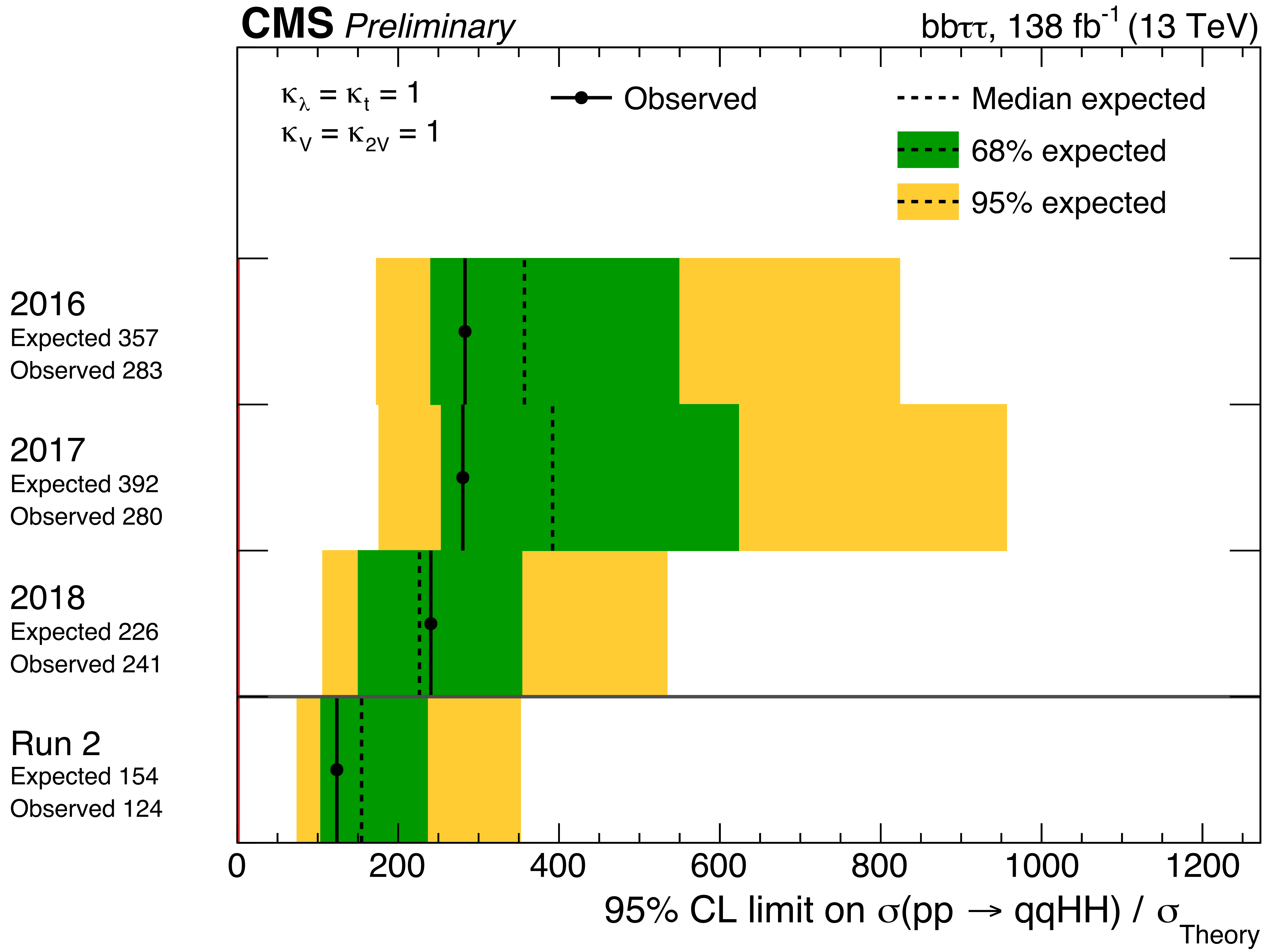
Observed (expected) 95% CL upper limit for the SM point

$$\sigma_{ggHH+qqHH} = 3.33(5.2) \cdot \sigma^{SM}_{ggHH+qqHH}$$

⇒ x5 improvement w.r.t. the previous HH→bb$\tau\tau$ analysis PLB 778 (2018) 101

⇒ improvement much larger than simple luminosity scaling

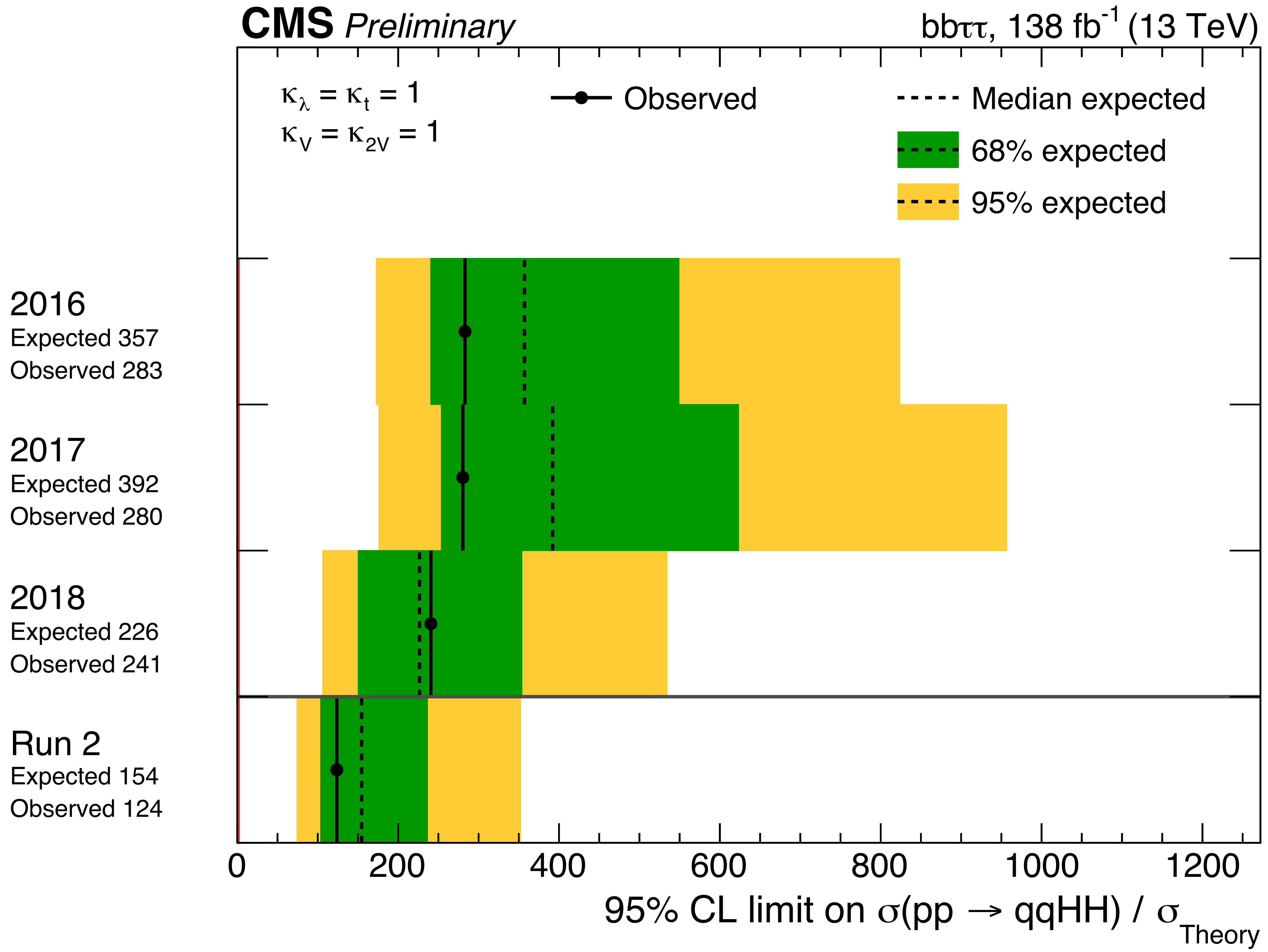
Standard model VBF HH cross section



Observed (expected) 95% CL upper limit for the SM point

$$\sigma_{qqHH} = 124(154) \cdot \sigma_{qqHH}^{SM}$$

Standard model VBF HH cross section



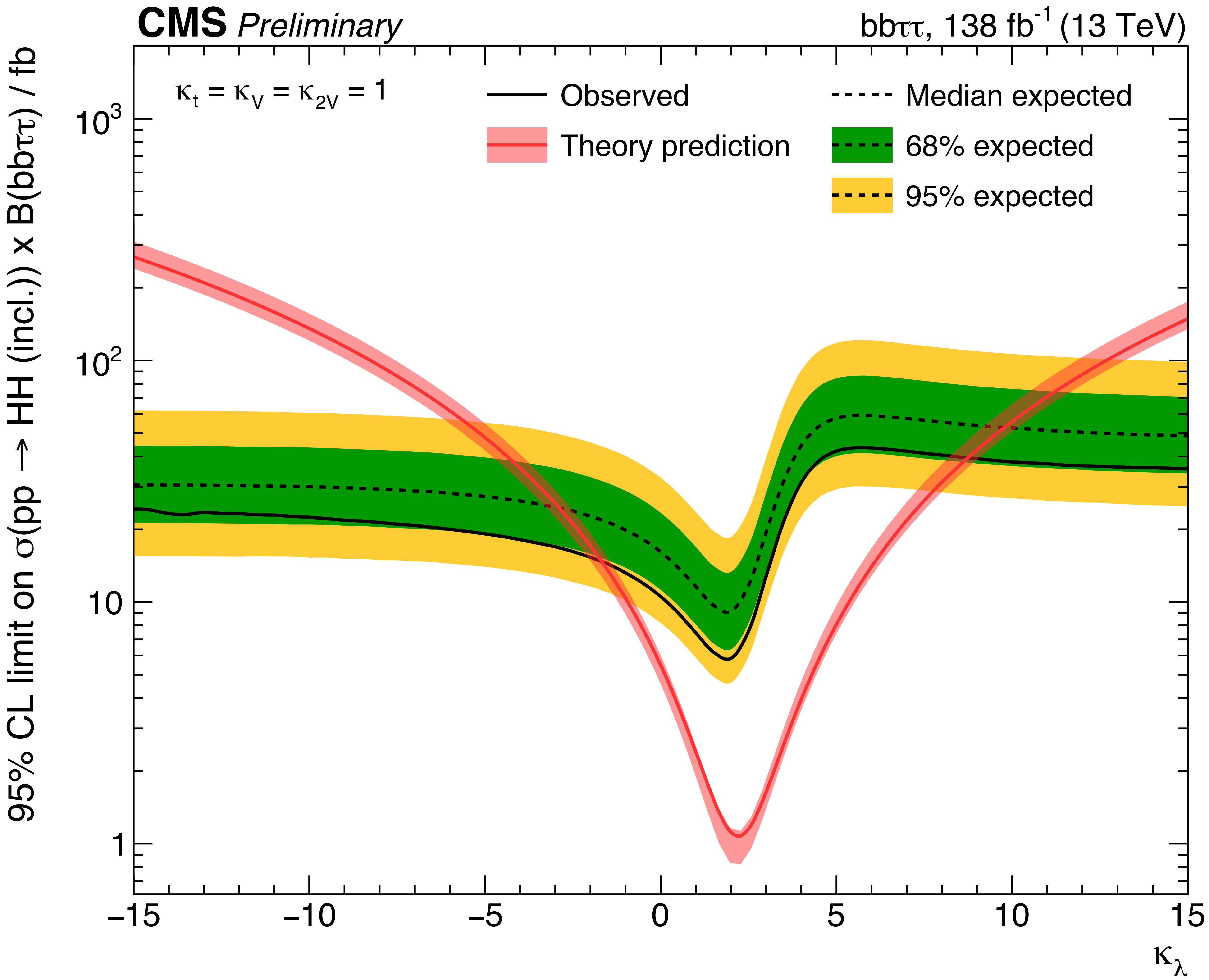
Observed (expected) 95% CL upper limit for the SM point

$$\sigma_{qqHH} = 124(154) \cdot \sigma_{qqHH}^{SM}$$

⇛ new result not present in the previous HH→bb$\tau\tau$ analysis PLB 778 (2018) 101

⇛ current most stringent limit on σ_{qqHH}

Constraints on κ modifiers: κ_λ

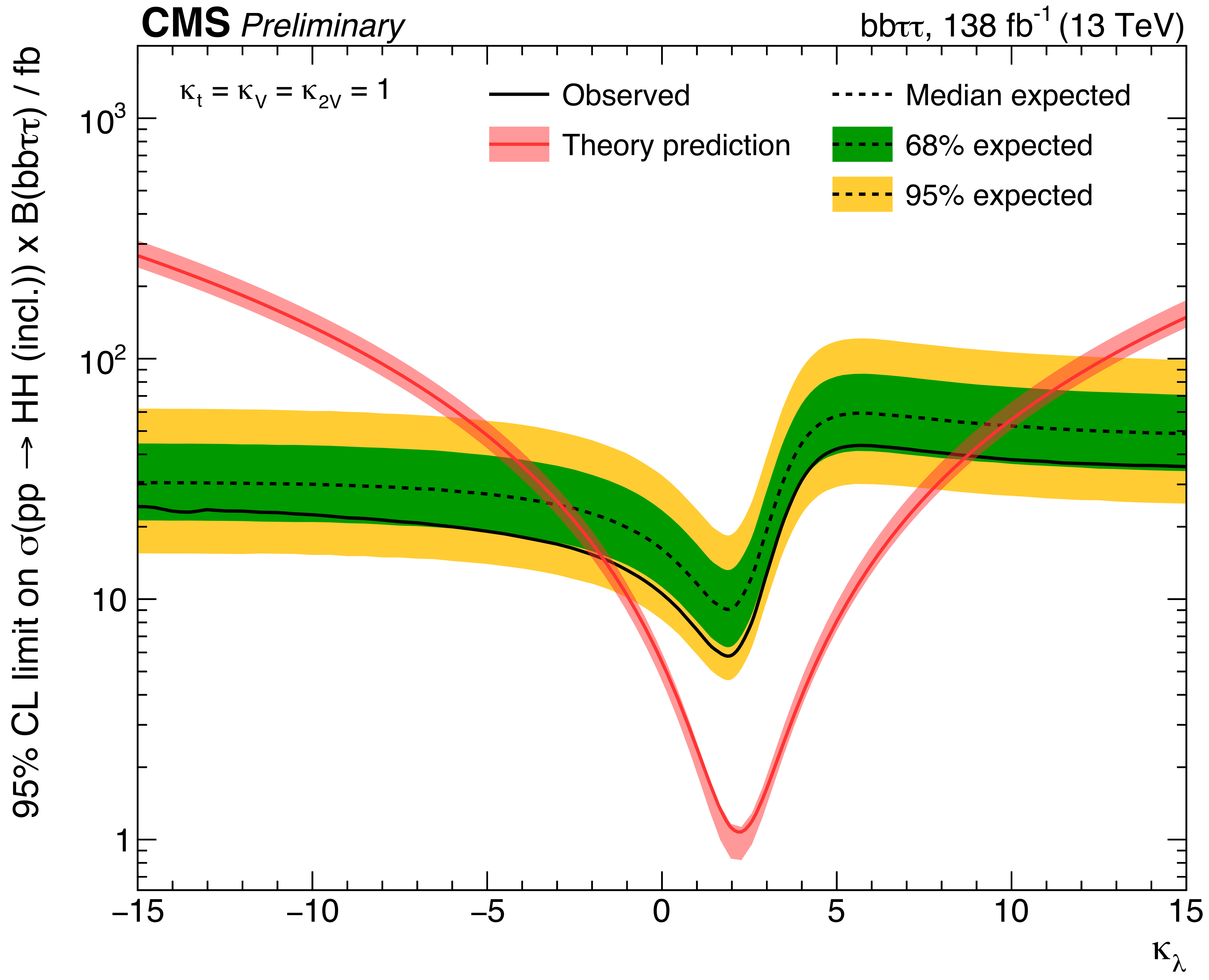


95% CL constraints on κ_λ from $\sigma_{ggHH+qqHH} \cdot \mathscr{B}(HH \rightarrow bb\tau\tau)$

Expected: $-3 < \kappa_\lambda < 9.9$

Observed: $-1.8 < \kappa_\lambda < 8.8$

Constraints on κ modifiers: κ_λ



95% CL constraints on κ_λ from $\sigma_{ggHH+qqHH} \cdot \mathscr{B}(HH \to bb\tau\tau)$

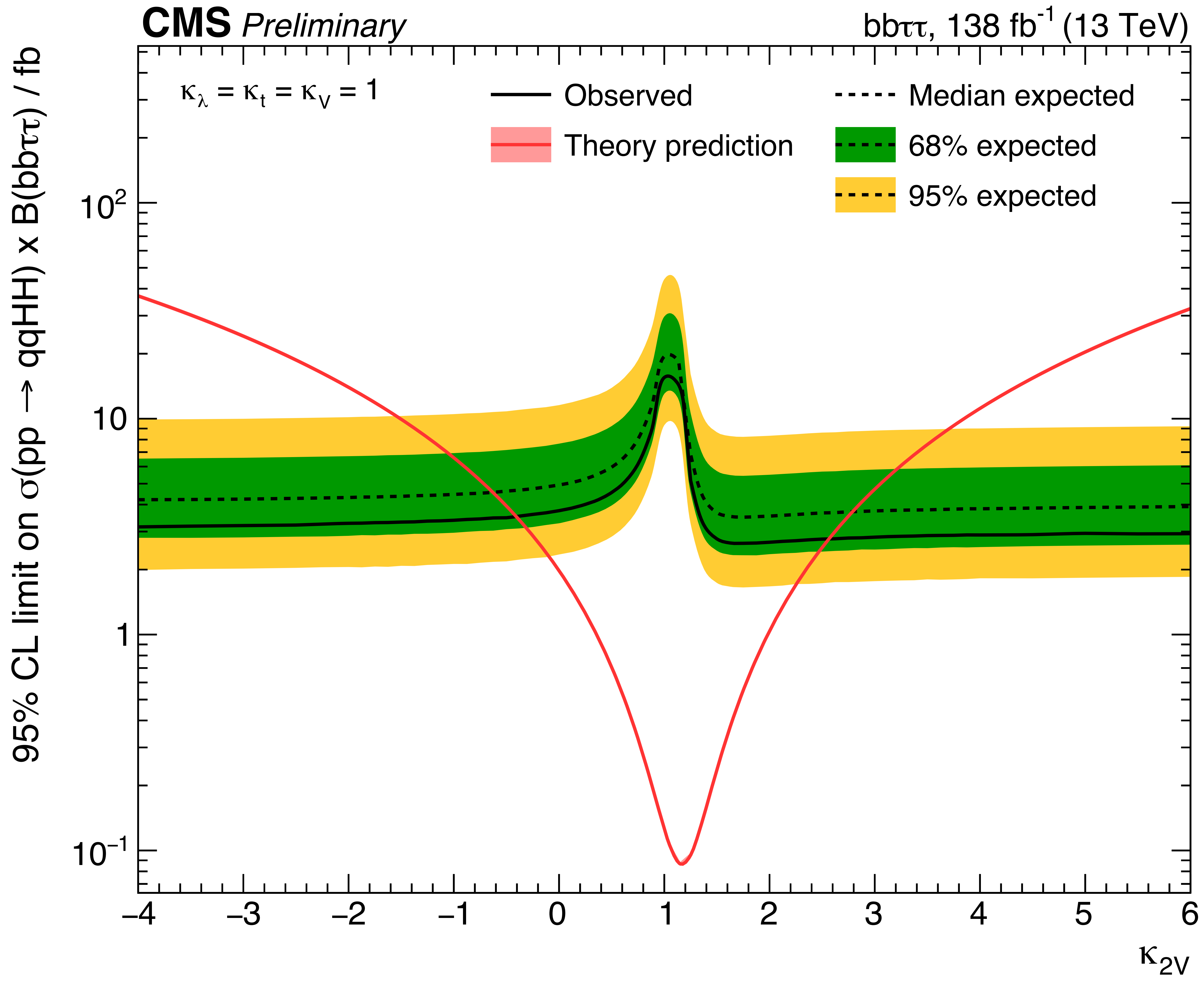
Expected: $-3 < \kappa_\lambda < 9.9$

Observed: $-1.8 < \kappa_\lambda < 8.8$

⇛ new result not present in the previous HH→bb$\tau\tau$ analysis PLB 778 (2018) 101

⇛ constraints very competitive with analyses in other decay channels

Constraints on κ modifiers: κ_{2V}



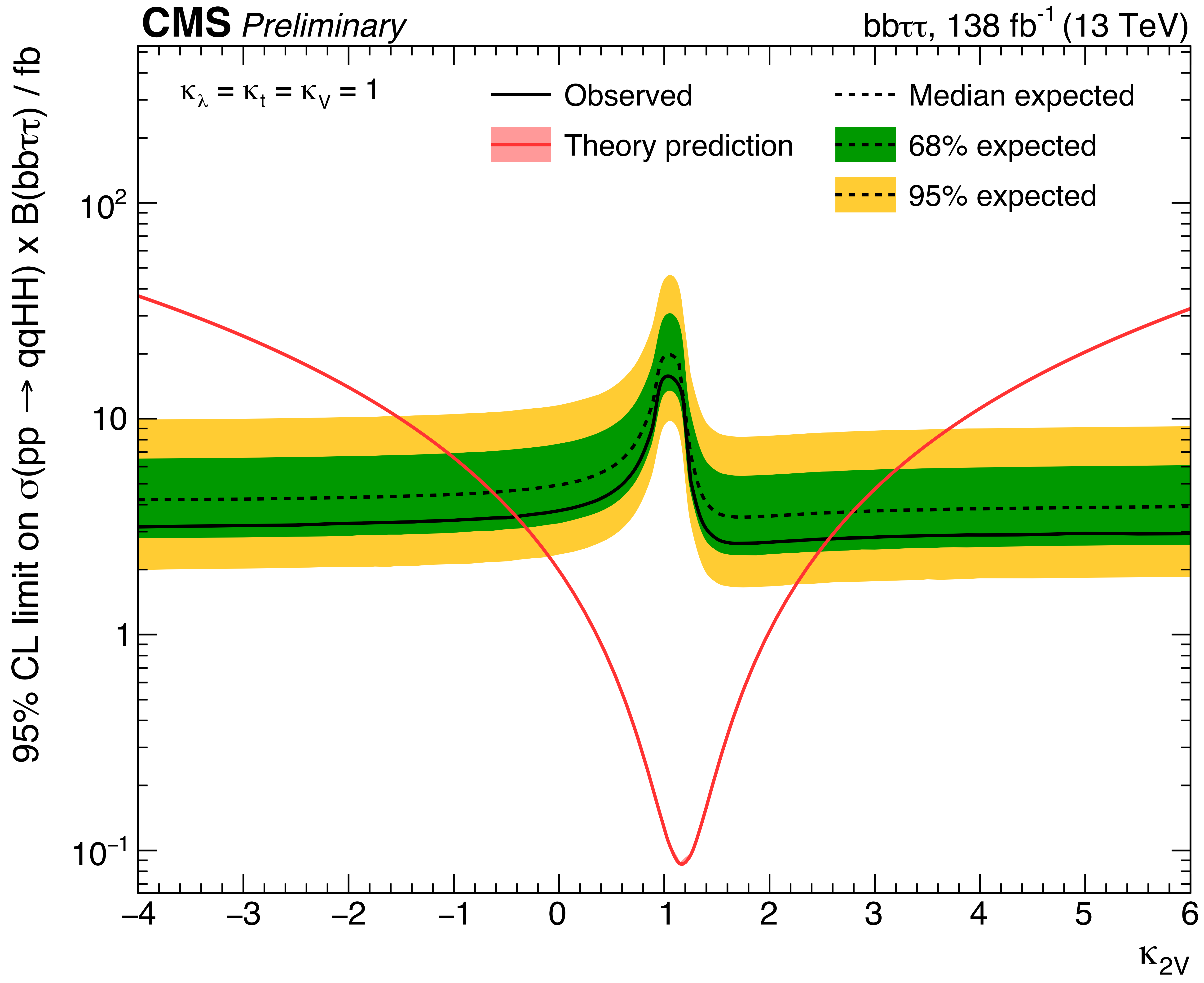
95% CL constraints on κ_{2V} from

$$\sigma_{qqHH} \cdot \mathcal{B}(HH \rightarrow bb\tau\tau)$$

Expected: $-0.6 < \kappa_{2V} < 2.8$

Observed: $-0.4 < \kappa_{2V} < 2.6$

Constraints on κ modifiers: κ_{2V}



95% CL constraints on κ_{2V} from $\sigma_{qqHH} \cdot \mathcal{B}(HH \to bb\tau\tau)$

Expected: $-0.6 < \kappa_{2V} < 2.8$
Observed: $-0.4 < \kappa_{2V} < 2.6$

⇛ new result not present in the previous HH→bb$\tau\tau$ analysis PLB 778 (2018) 101

⇛ constraints very competitive with analyses in other decay channels

Conclusions and outlook

Conclusions and outlook

- The **non-resonant HH→bb$\tau\tau$** analysis performed with data collected by CMS during Run 2 was presented
- **Considerable improvement** over the previous public results
  → much **more than simple luminosity scaling** thanks to **advance ML techniques**
- **Wide range of results** produced: σ_{HH} and σ_{qqHH} limits; κ_λ and κ_{2V} constraint; 2D exclusion contours; 1D likelihood scans

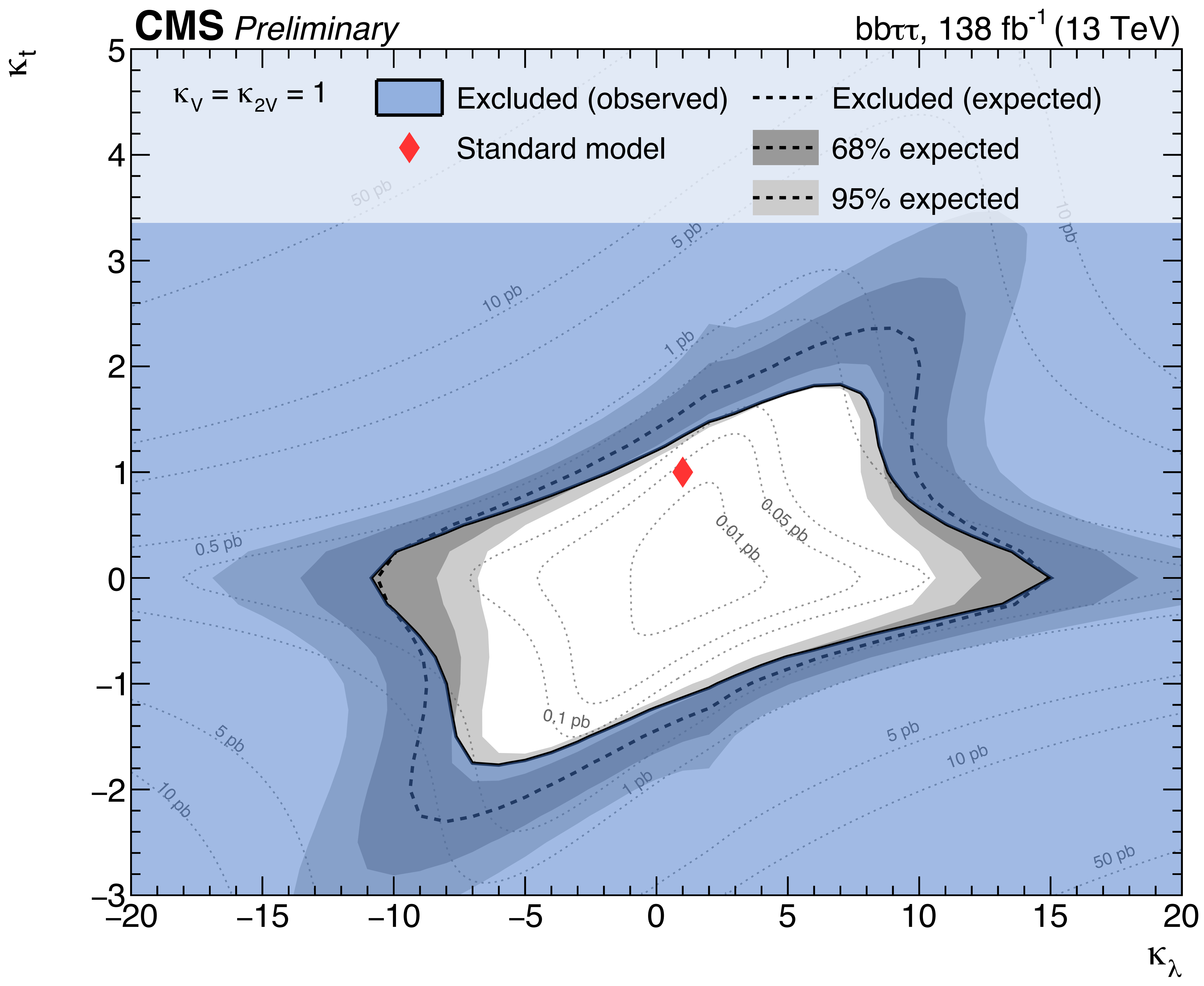
95% CL limit on σ_{HH}	**obs(exp) $3.3(5.2) \times \sigma^{SM}_{HH}$**
95% CL limit on σ_{qqHH}	**obs(exp) $124(154) \times \sigma^{SM}_{qqHH}$**

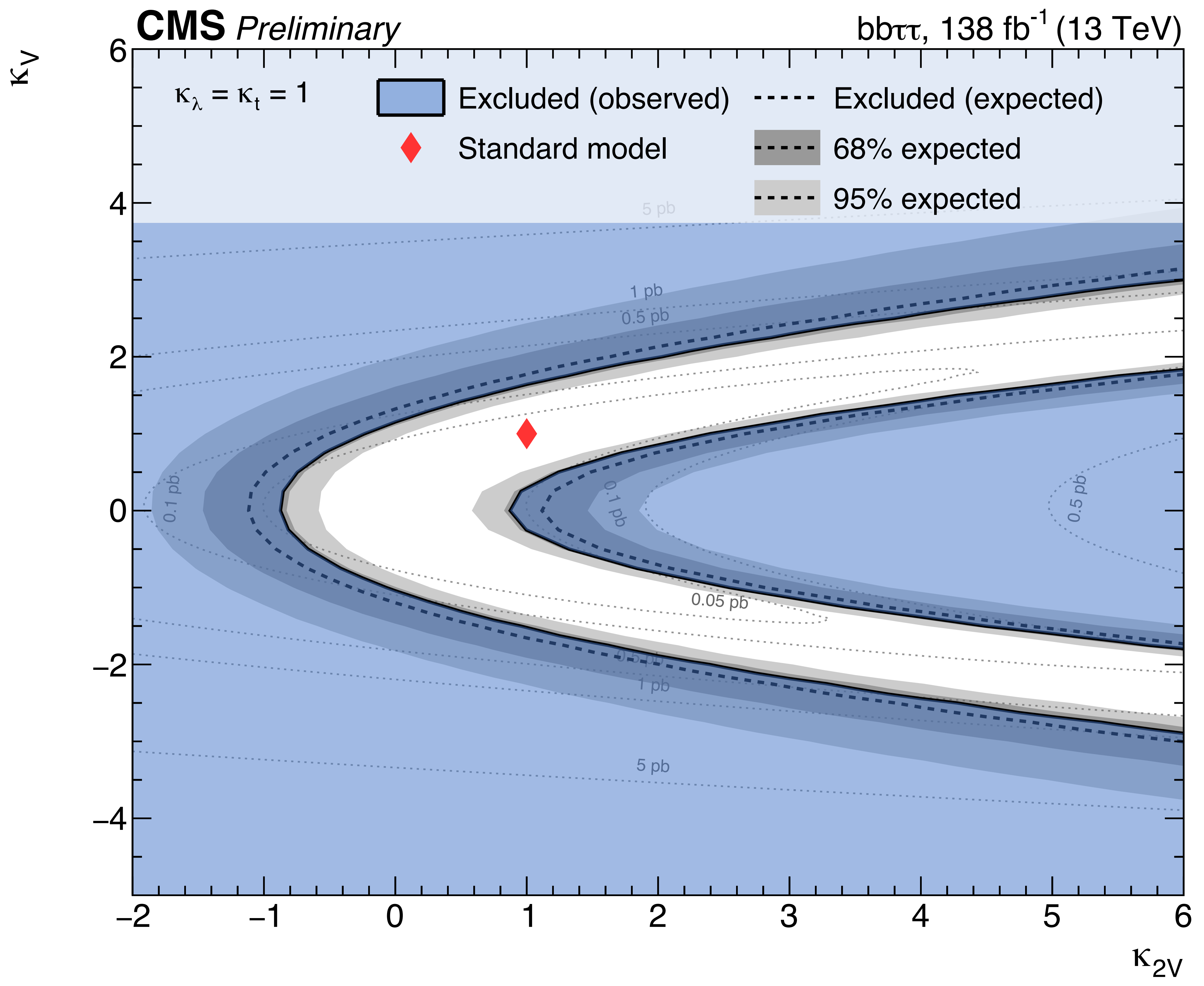
← BEST LIMIT TO DATE

- The resonant analysis is currently underway and will enlarge this already wide set of results
- We also look forward to analyse Run 3 data in order to better this even more

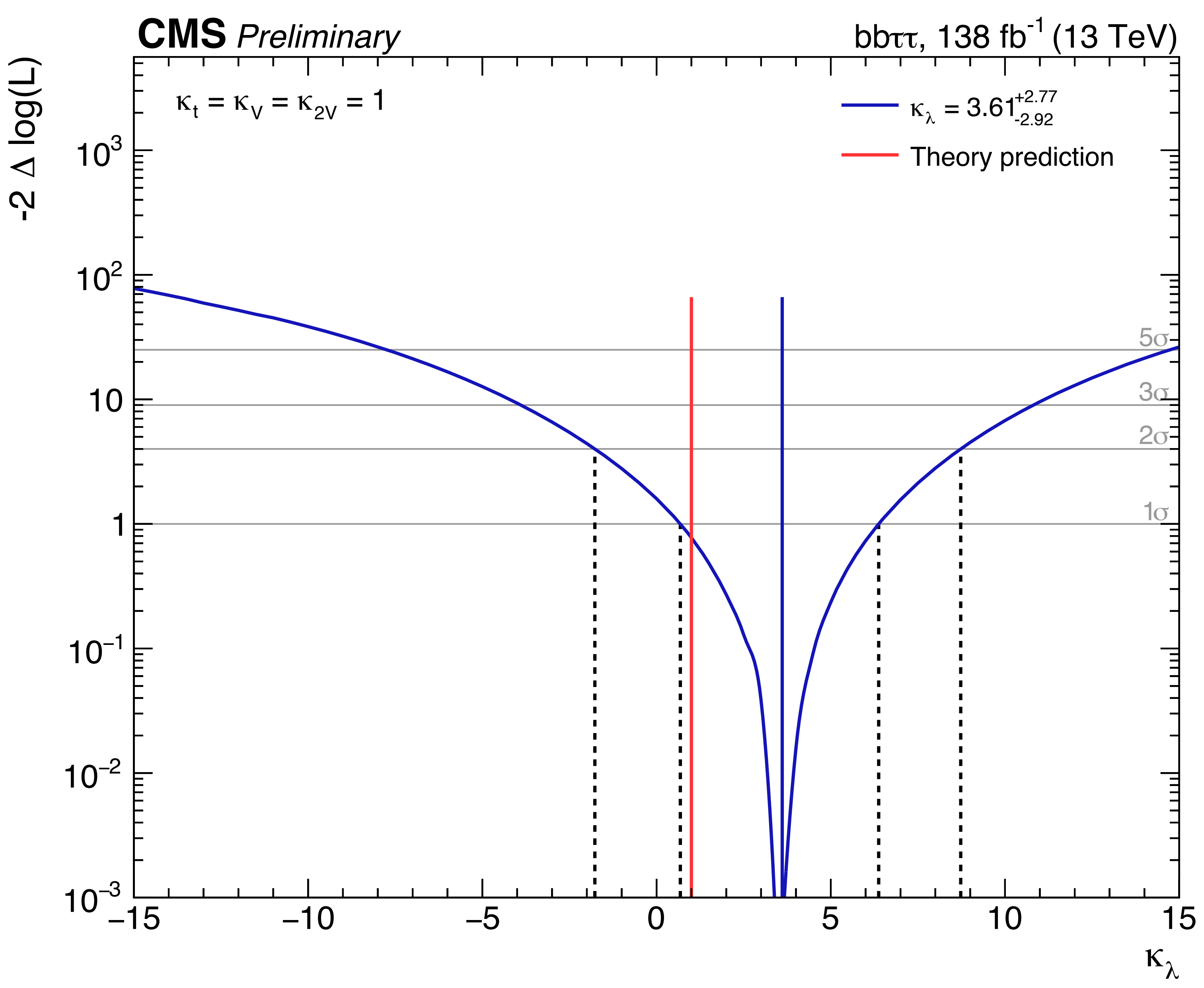
BACKUP

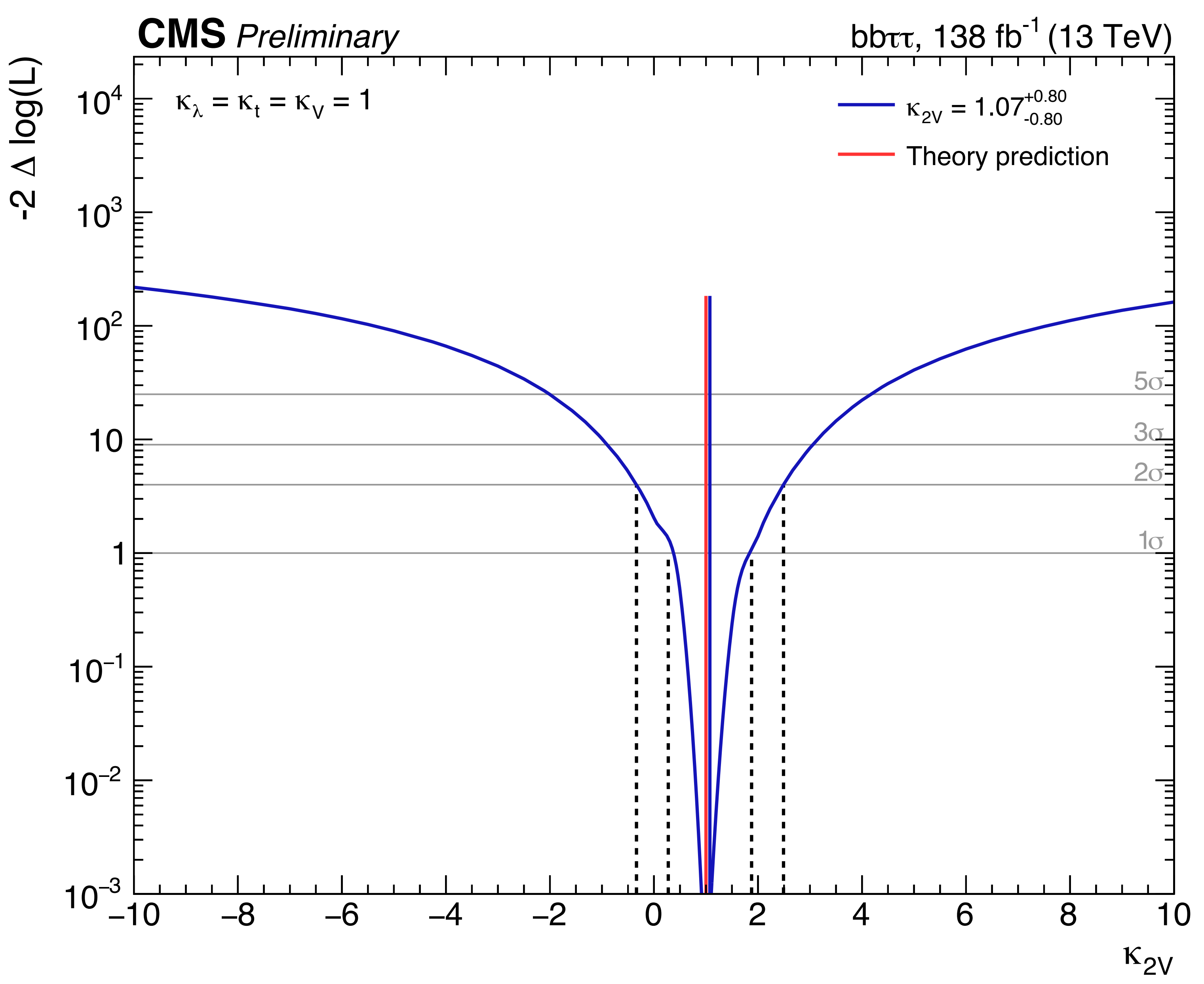
BACKUP: 2D exclusion limits

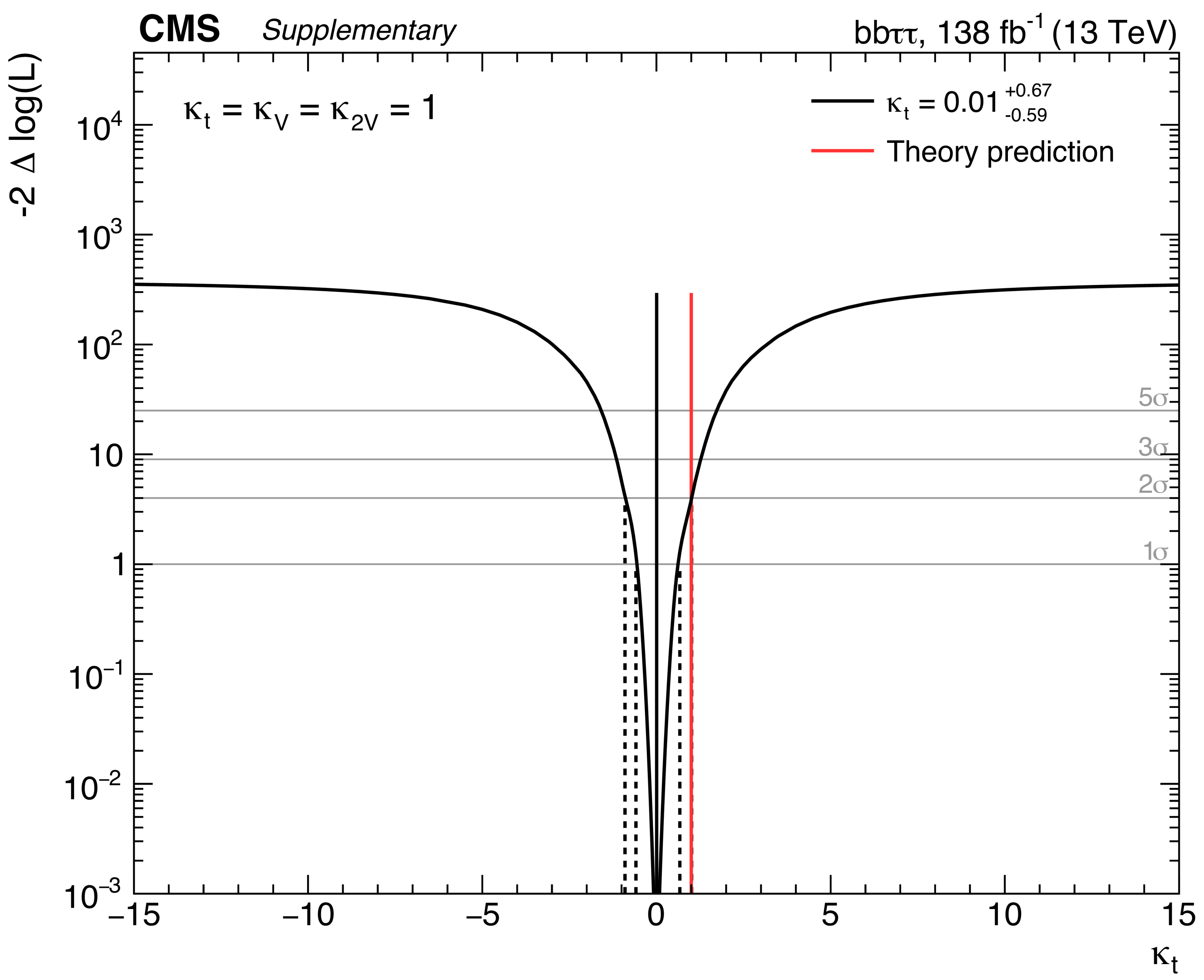




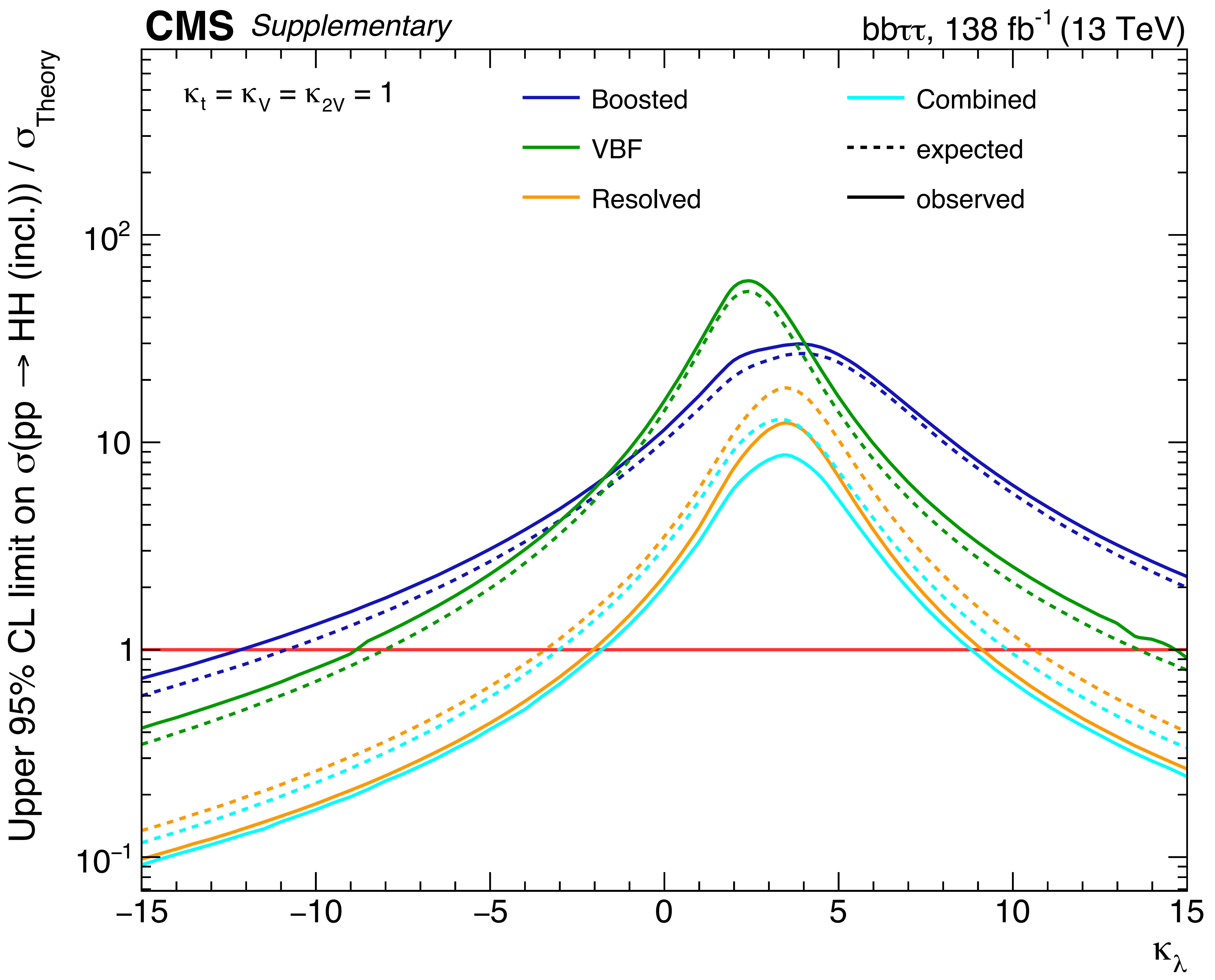
BACKUP: likelihood scan vs. κ_λ / κ_{2V} / κ_t

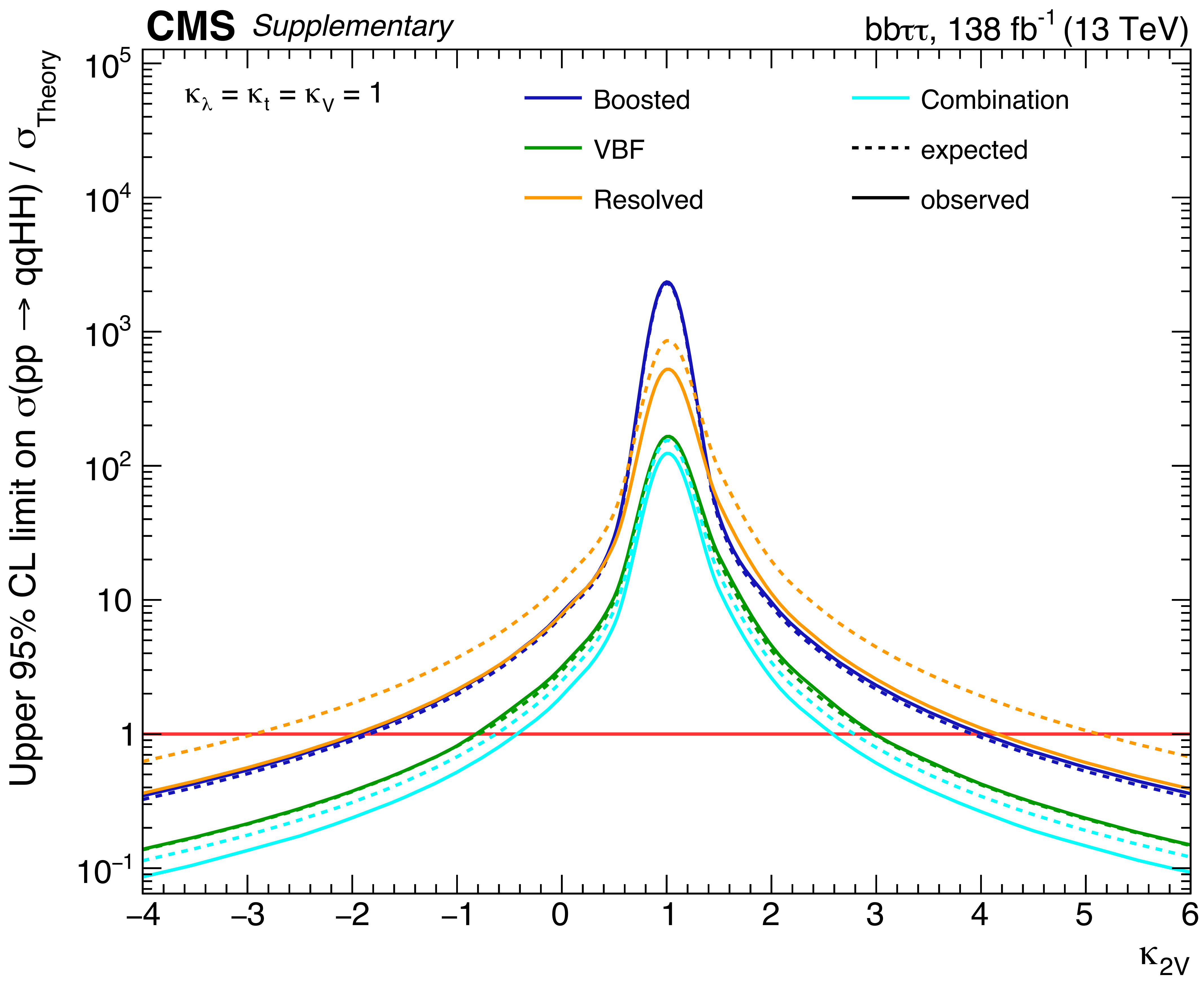




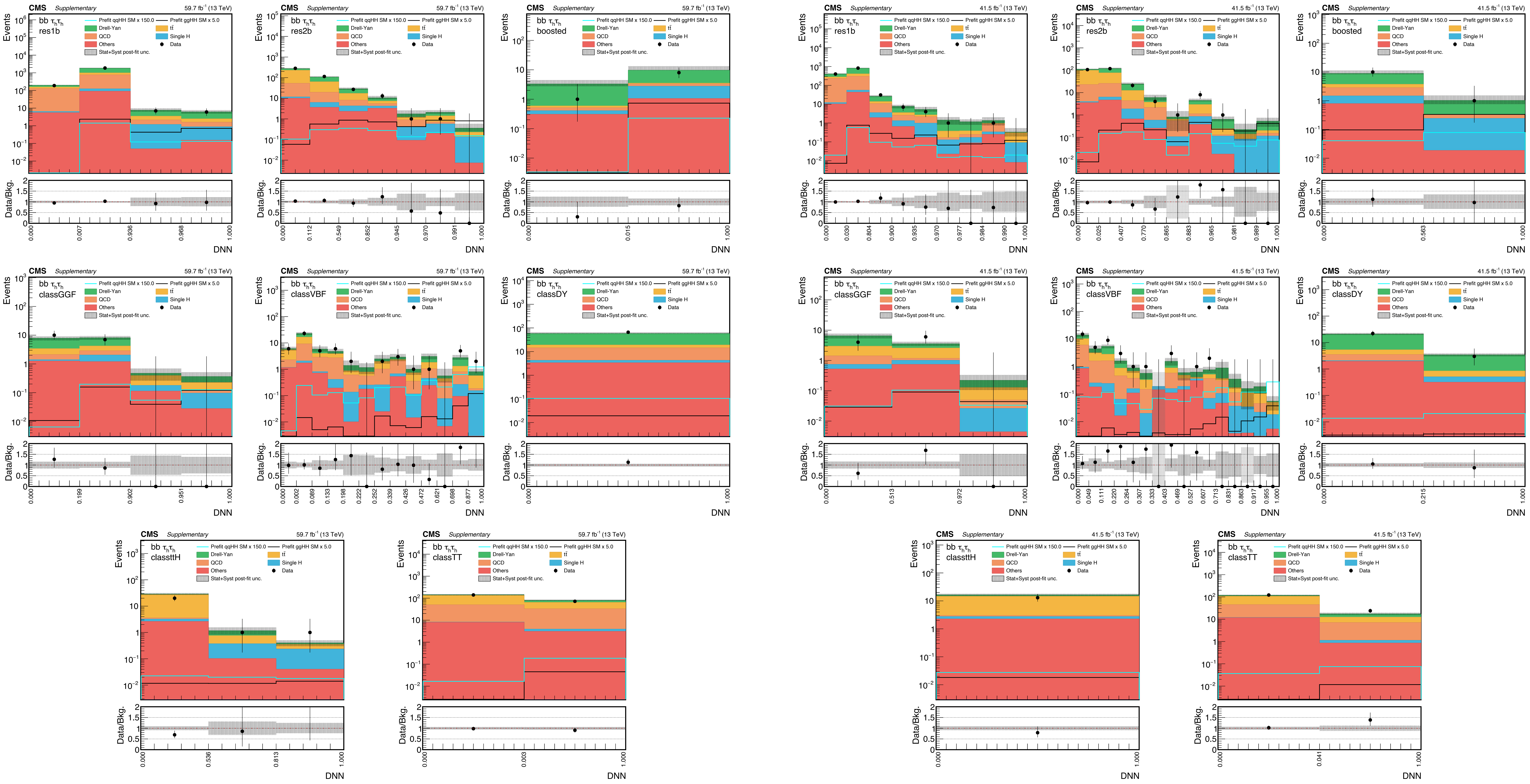


BACKUP: likelihood scan by category





BACKUP: example DNN distributions



BACKUP: objects selections

Online p_T trigger thresholds	single-e: $p_T > 25(32)$ GeV, cross-e: $p_T > 24$ GeV single-μ: $p_T > 22(24)$ GeV, cross-μ: $p_T > 19(20)$ GeV di-tau: $p_T > 35$ GeV, di-tau VBF: $p_T > 20$ GeV
Offline p_T thresholds	1 GeV (electrons and muons), 5 GeV (taus)
η thresholds	electrons and muons: $\lvert\eta\rvert < 2.1$ tau: $\lvert\eta\rvert < 2.1$ (2.3) for di-tau and cross (single) triggers
Lepton ID and Isolation	Tight electron MVA ID+Iso, Tight muon ID and Iso
τ_h isolation ($\tau_e\tau_h$, $\tau_\mu\tau_h$ channels)	Medium DeepTauVsJet Tight DeepTauVsMu Very-loose DeepTauVsEle
τ_h isolation ($\tau_h\tau_h$ channel)	Medium DeepTauVsJet Very-loose DeepTauVsMu Very-very-loose DeepTauVsEle
Distance to PV	$\lvert d_{xy}\rvert < 0.045$ cm (electrons and muons only) $\lvert d_z\rvert < 0.2$ cm
Pair selections	opposite sign, $\Delta R > 0.5$

BACKUP: uncertainty sources

NORMALISATION UNCERTAINTIES

- Luminosity
- e, μ ID & efficiency
- L1 pre-firing (2016 & 2017)
- PU reweighing
- DY normalisation corrections
- $t\bar{t}$ normalisation corrections
- QCD iso/non-iso extrapolation (stat.)
- QCD iso/non-iso extrapolation (syst.)
- HH cross-section (theoretical)
- Higgs branching fractions
- SM process cross-sections
- VBF dipole recoil

SHAPE UNCERTAINTIES

- τ ID and energy scale (ES)
- Custom τ ID SF in 2017
- e, μ faking τ ID & ES
- Jet ES – 11 split sources
- Jet energy resolution
- QCD shape
- Trigger corrections
- b-tag corrections – 7 split sources
- Pileup jet ID corrections
- Finite MC statistics