Overview of the HH $\rightarrow$ bb $\tau\tau$ searches with the CMS experiment

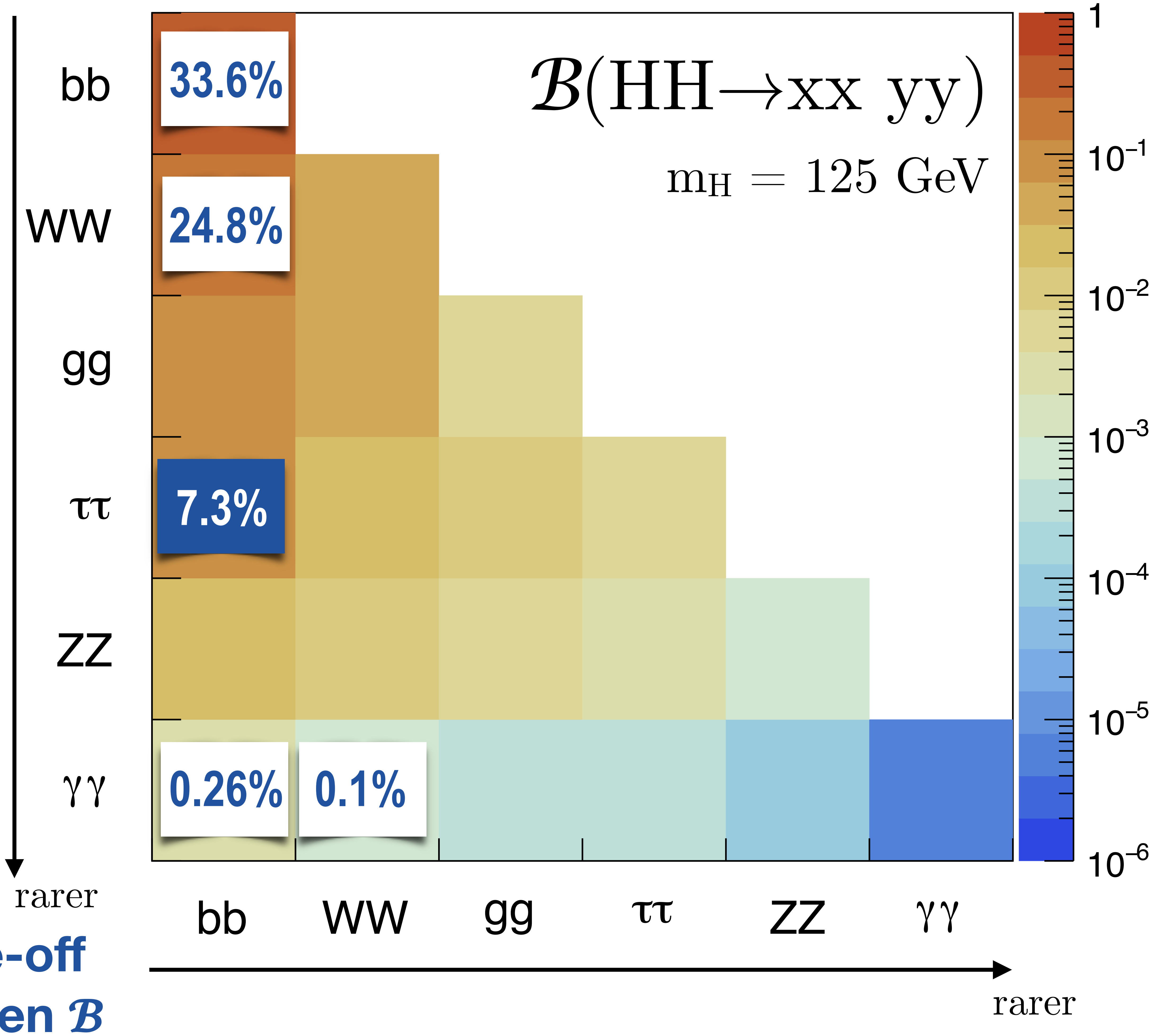
Luca Cadamuro

on behalf of the CMS Collaboration

Double Higgs Production at Colliders workshop

Fermilab, September 6th, 2018

Why bb $\tau\tau$?



Trade-off between $\mathcal{B}$ and purity

Keep $\mathcal{B}$ high enough

Opportunities

One of the most sensitive channels at the LHC

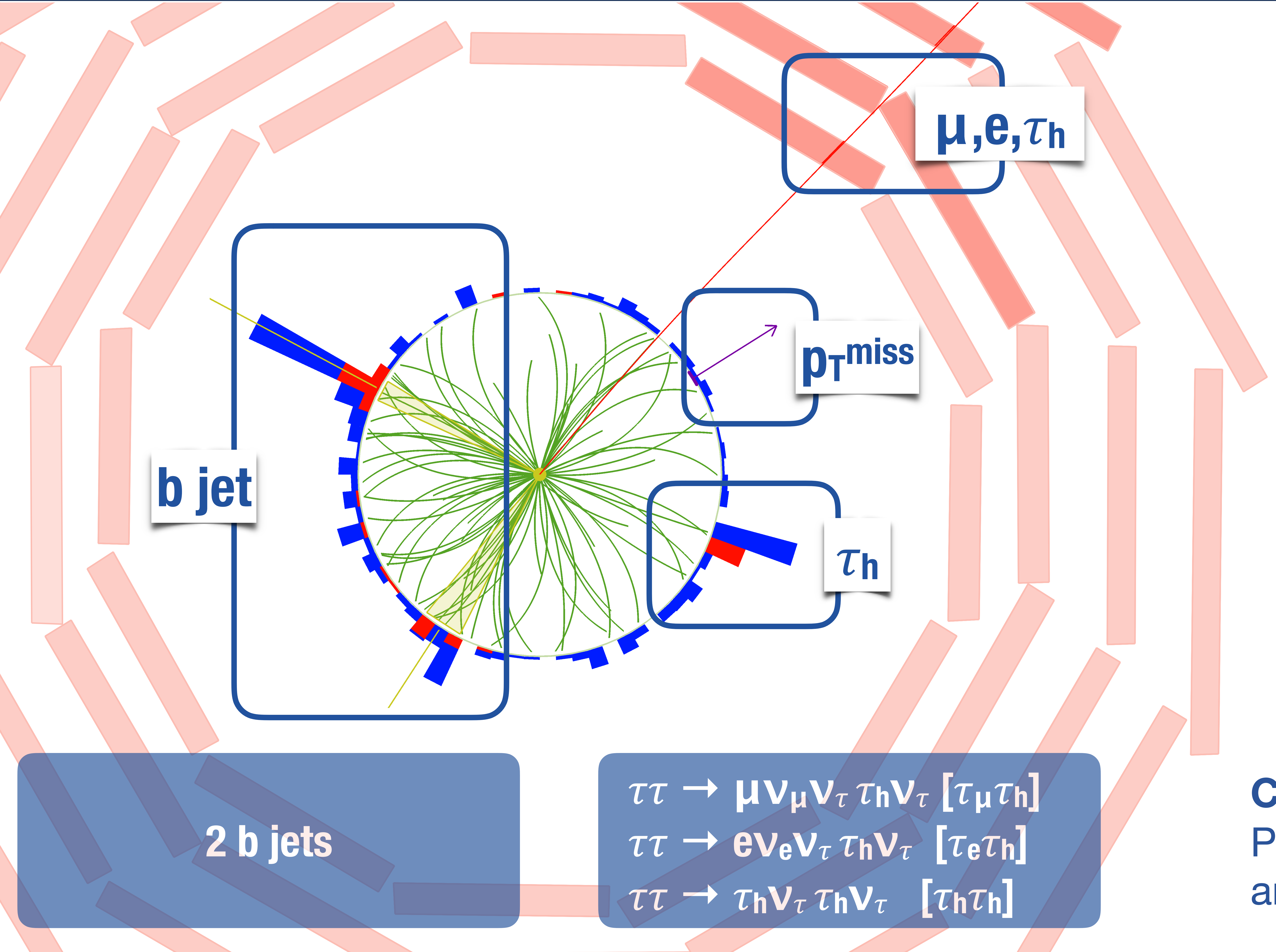
- Sizeable branching fraction
- Moderate contamination from background processes

Challenges

Sophisticated analysis techniques are required

- τ decays result in several final states
- Incomplete final state reconstruction
- Offline reconstruction and q/g jet rejection
- Difficult signature for trigger
  - hadronic final state in 42% of $\tau\tau$ decays

The bb $\tau\tau$ decay channel



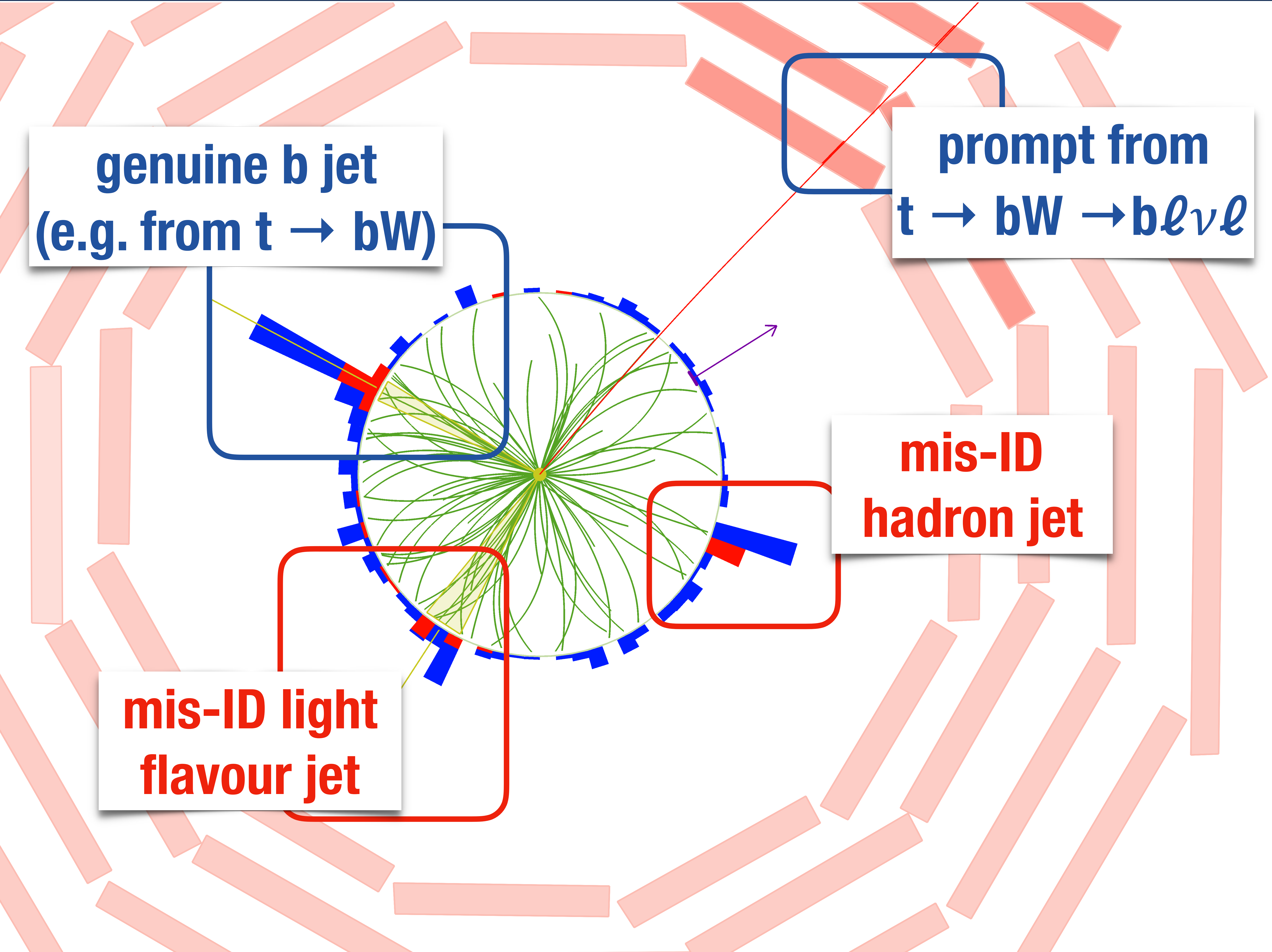
- Three $\tau\tau$ final states considered
  - $\mu\tau_h$, $e\tau_h$, $\tau_h\tau_h$: 88% of $\tau\tau$ decays
  - highest sensitivity
  - many different final state objects
- Different trigger strategy depending on the channel
- Different background contamination for $\ell\tau_h$ and $\tau_h\tau_h$ $\Longrightarrow$ specific analysis techniques
- Two analyses targeting **resolved** and **boosted** $\tau\tau$ topologies

CMS analyses references:

PLB 778 (2018) 101 nonresonant and low mass resonant

arXiv:1808.01365 high mass resonant

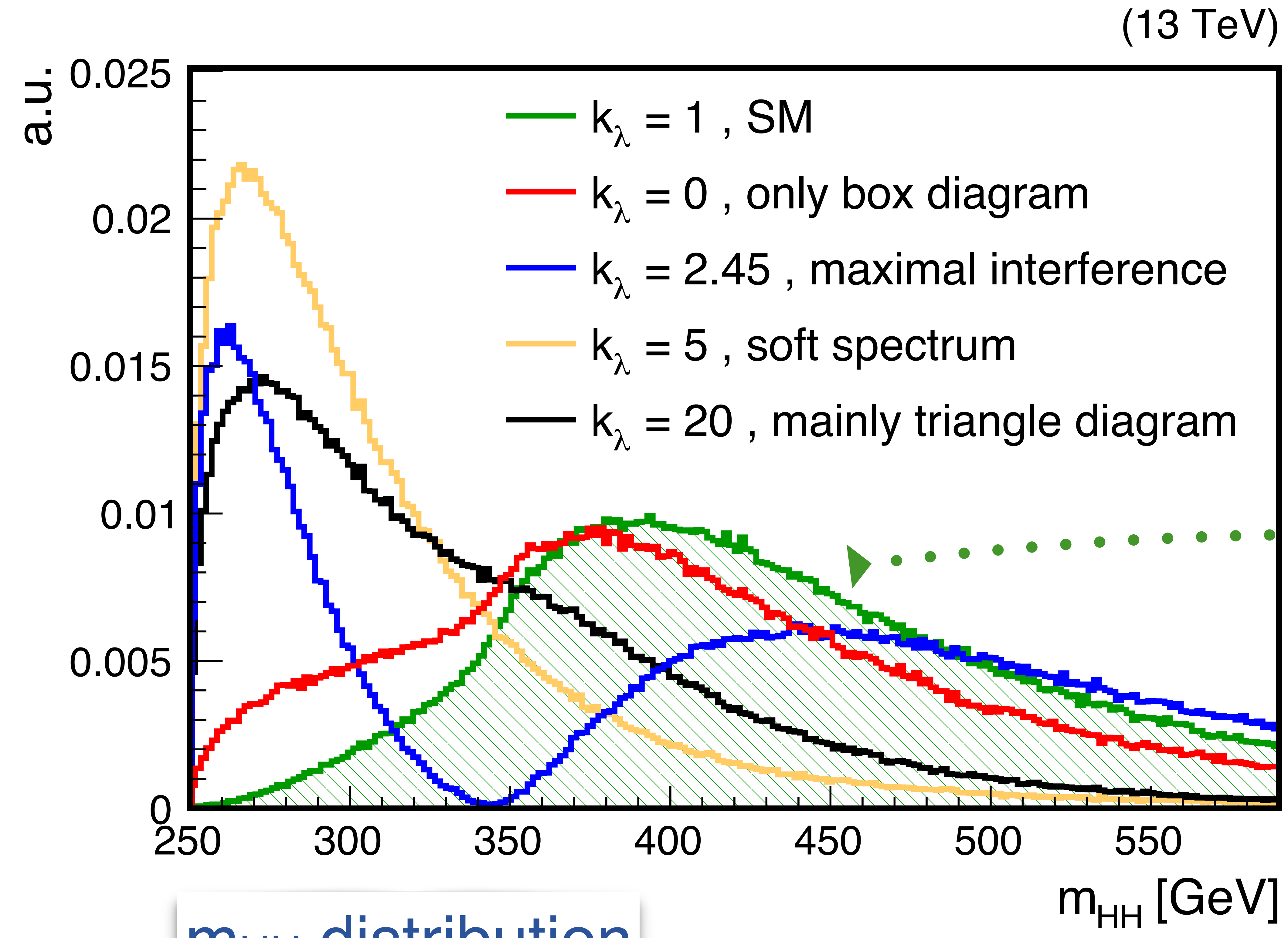
The bb $\tau\tau$ decay channel



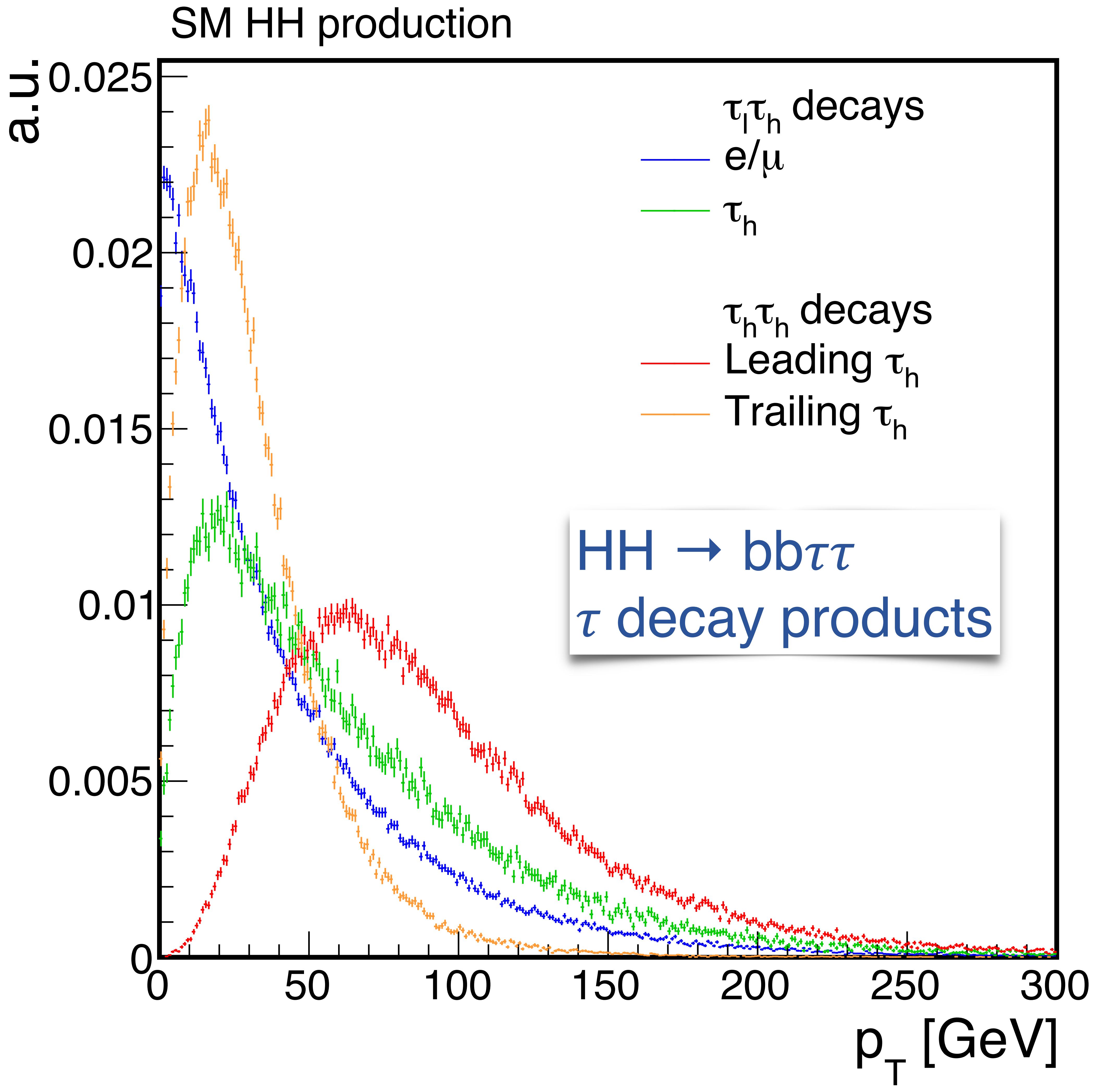
- **Irreducible** backgrounds
  - tt → bbWW → bb $\tau\tau$
  - Z/γ^* → $\tau\tau$ + 2 b jets
  - di-boson, ZH (minor)

- **Instrumental (reducible)** backgrounds
  - tt, Z/γ^*, multijet with misidentified jets as τ_h or b jet
  - single top, W+jets (minor)

The HH → bbττ signal

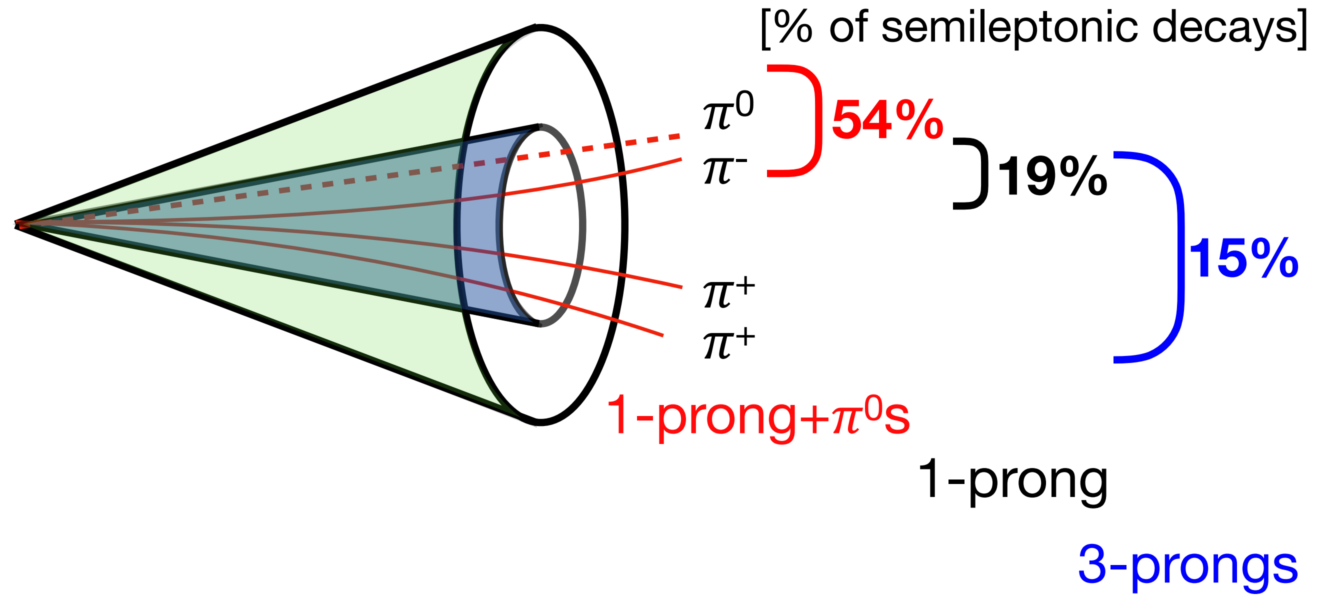


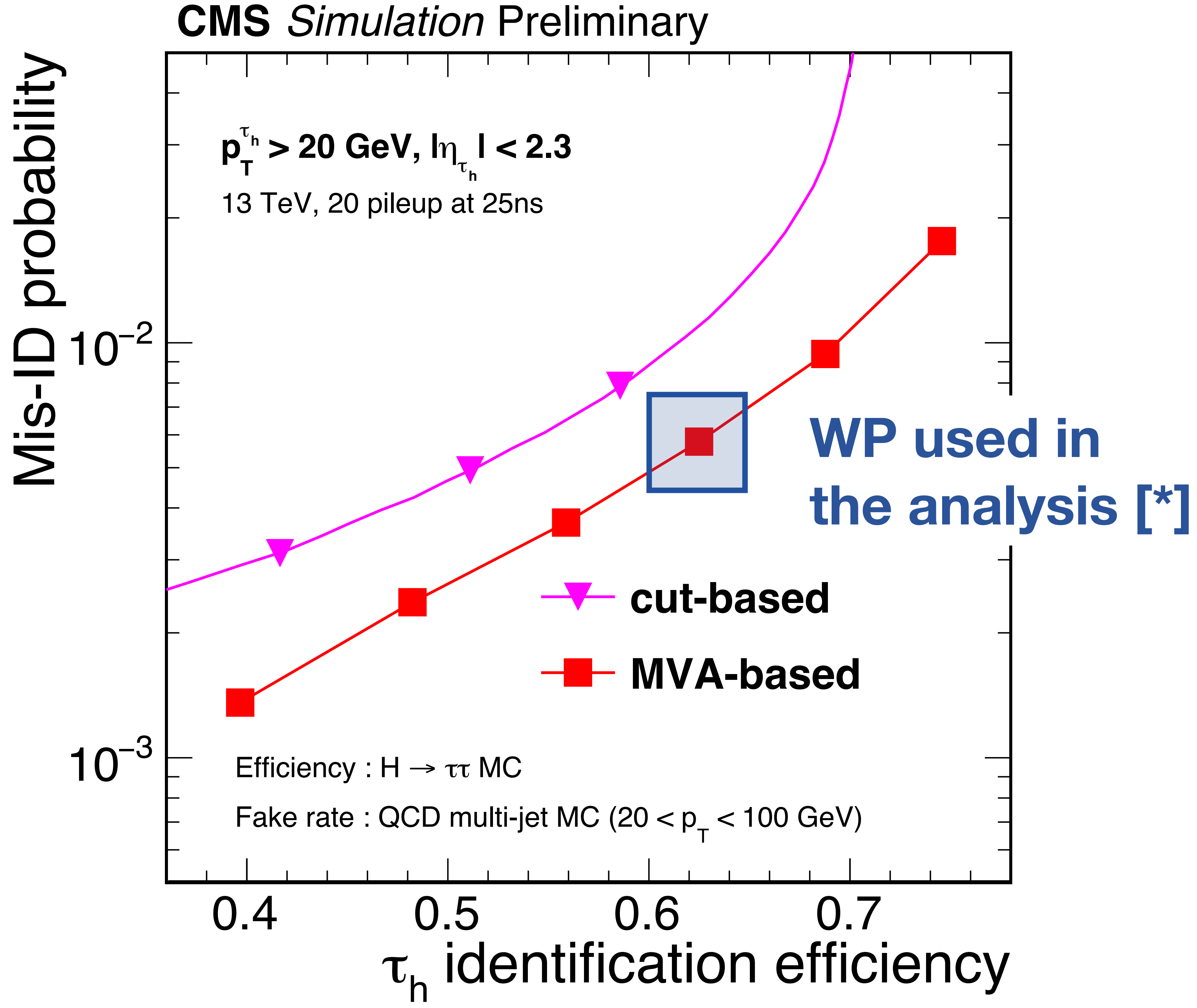
m_{HH} distribution



- Low p_T objects involved in the analyses
- High p_T / high boost (∝ m_{HH}) events relevant for resonant and BSM nonresonant searches

τ leptons in CMS





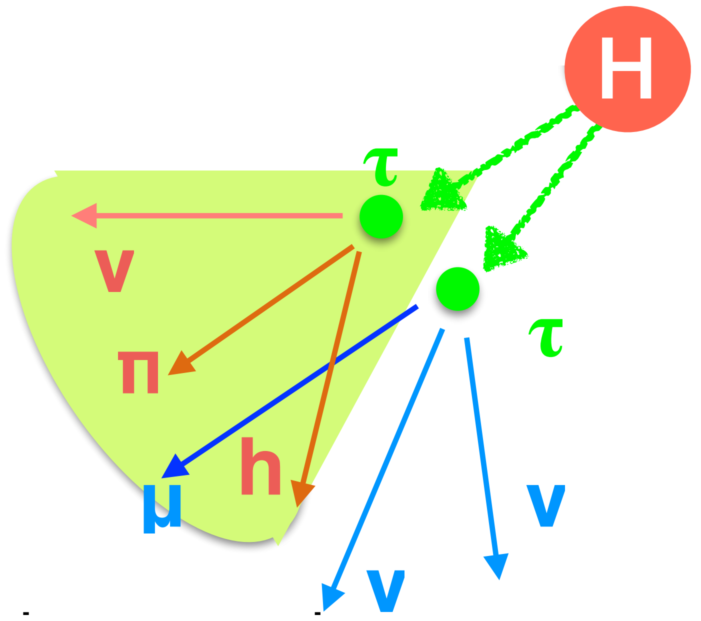
- $\tau \rightarrow e/\mu + \nu$s : clean signature from electron and muon
- $\tau \rightarrow$ hadrons (τ_h) + νs
  - dominant decay mode
  - 1 or 3 charged particles w/ or w/o $\pi^0 \Longrightarrow$ many signatures
  - **particle flow reconstruction** using the hadron-plus-strips (HPS) algorithm to combine charged and $\pi^0 \rightarrow \gamma\gamma$ candidates in jets using mass reconstruction
  - multivariate isolation discriminant to separate signal from q/g jets

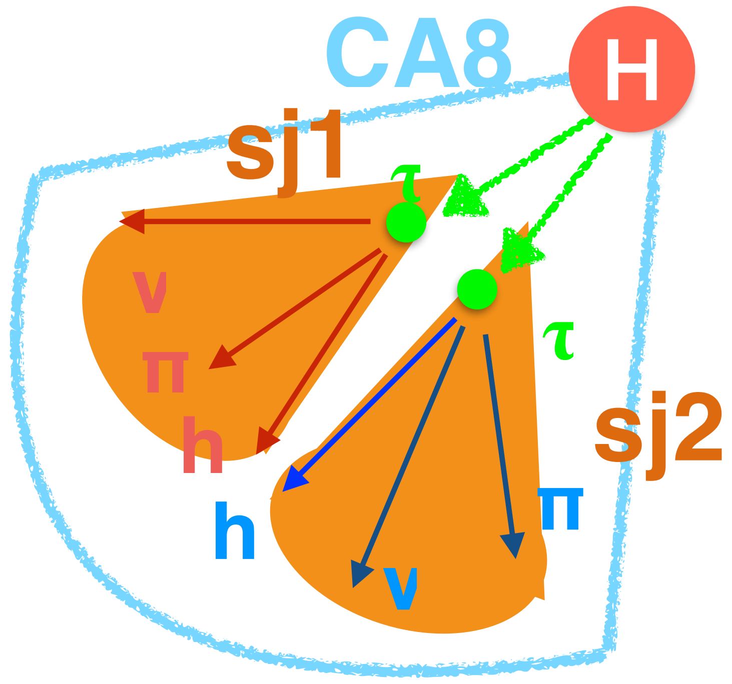
Efficient τ reconstruction and isolation
Important to suppress instrumental backgrounds

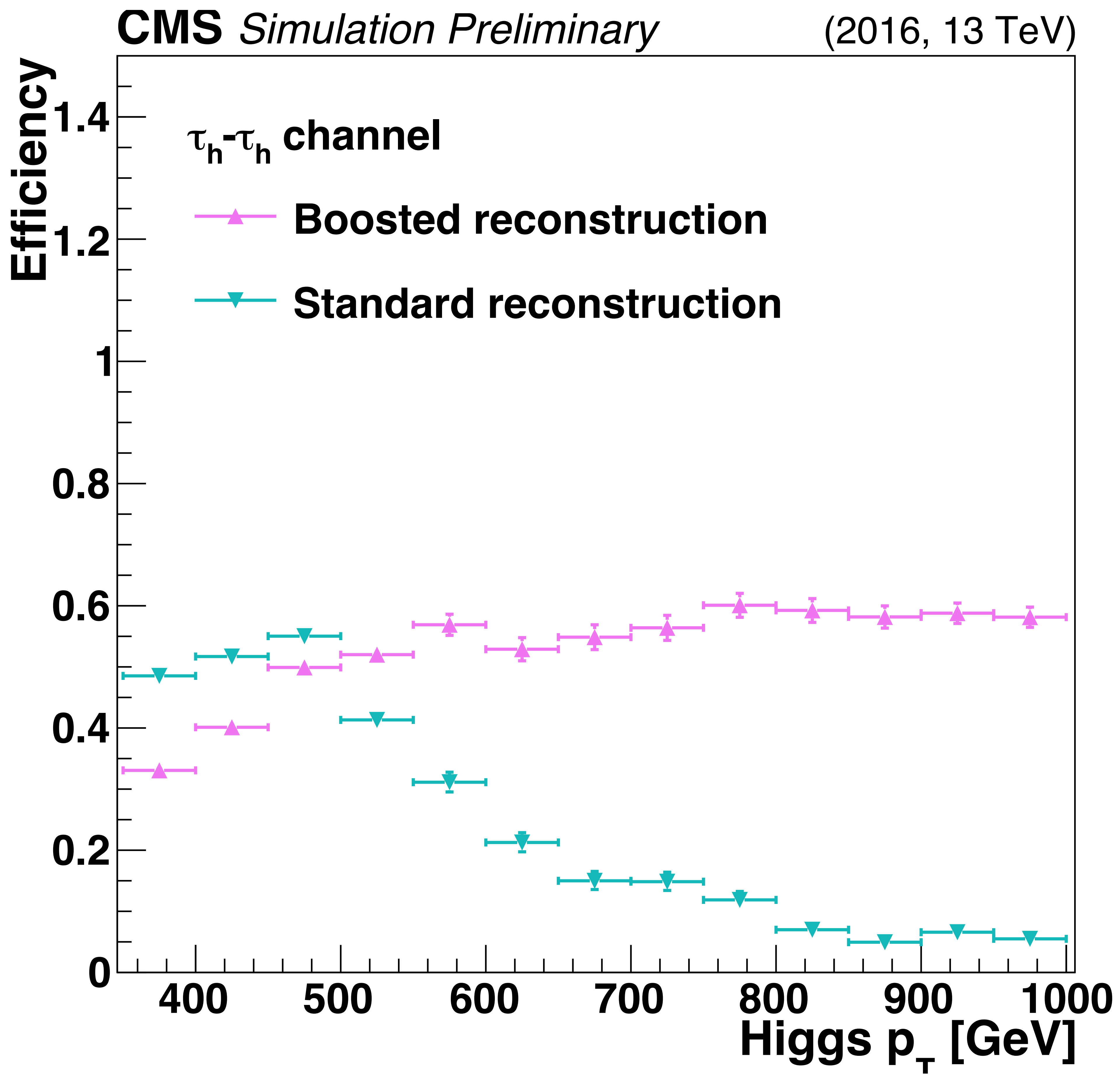
[*] Larger efficiency for $\tau_h\tau_h$ because of correlation with trigger isolation

Boosted tau leptons

- Reconstruction of boosted $\tau\tau$ decays starts from large radius Cambridge-Aachen R = 0.8 jets to identify subjets used as "seeds" to the HPS algorithm
  - $\tau\tau \rightarrow \ell\tau_h + \nu s$: remove the identified lepton from the jet
  - $\tau\tau \rightarrow \tau_h\tau_h + \nu s$: find the two subjets (with quality criteria)
  - relaxed decay mode criteria for the τ_h candidates



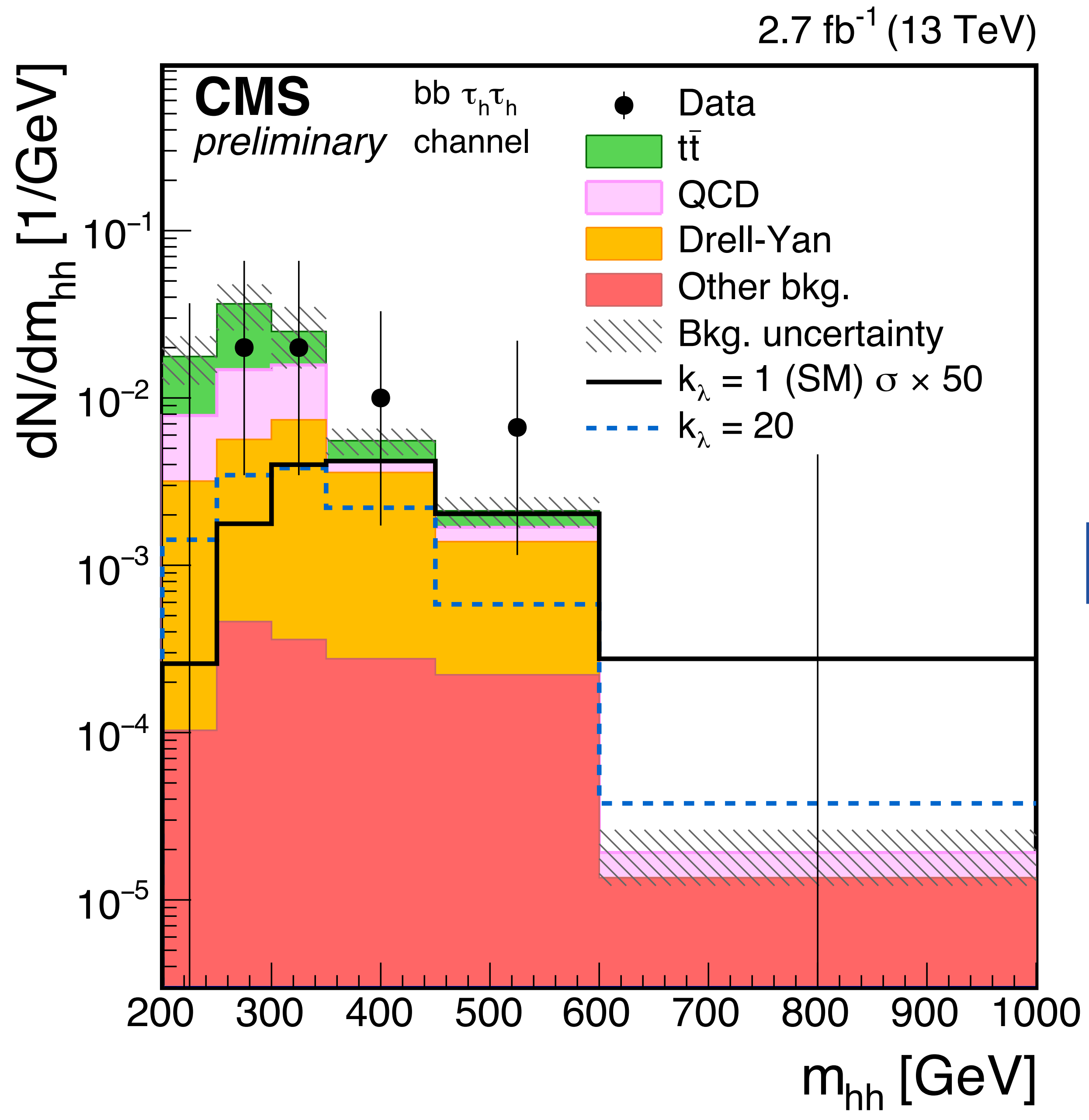




Key element for the high mass analysis
Improves the reconstruction efficiency of high p_T H

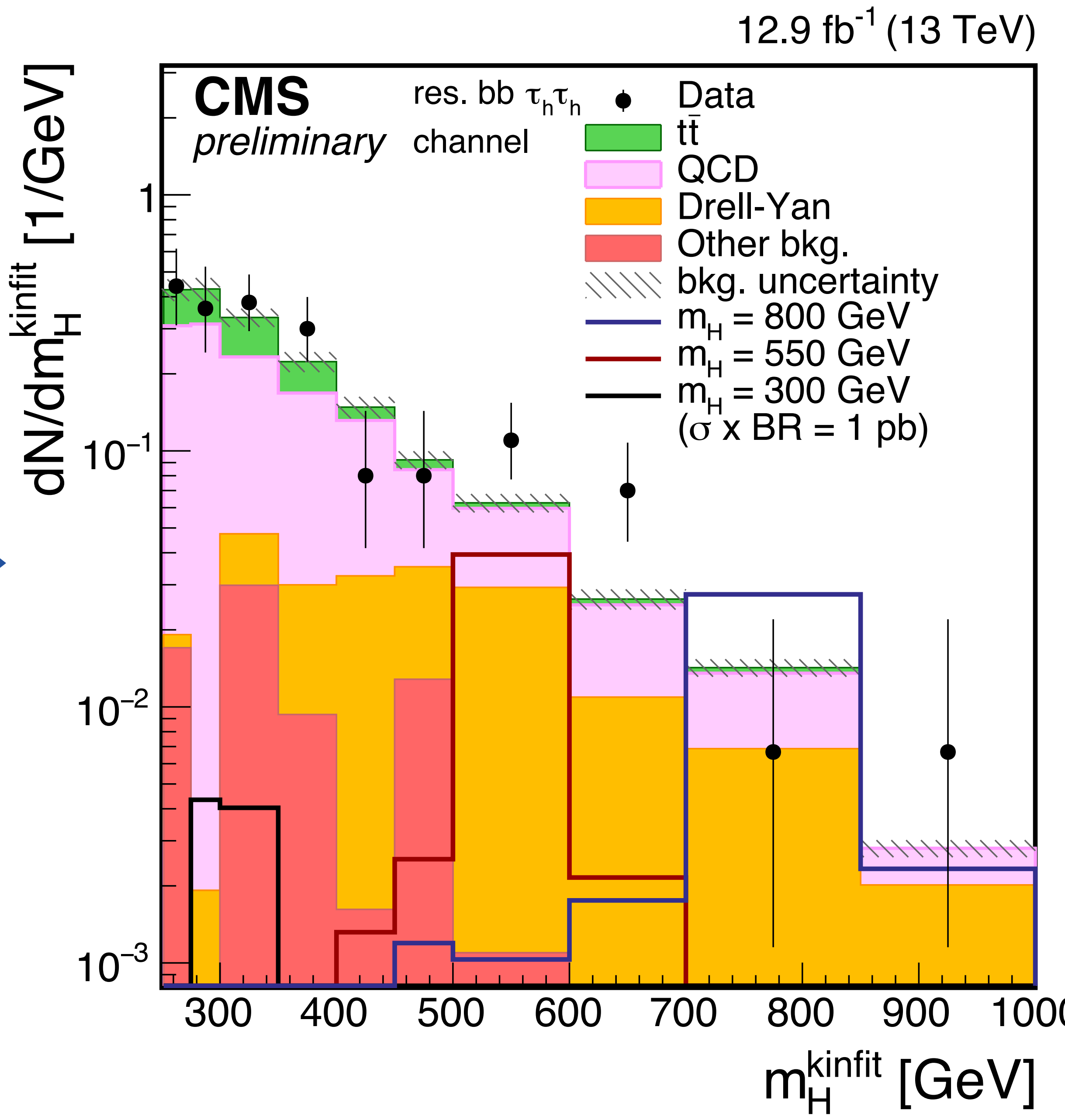
Updated Run II plots soon available

bb $\tau\tau$: evolving results



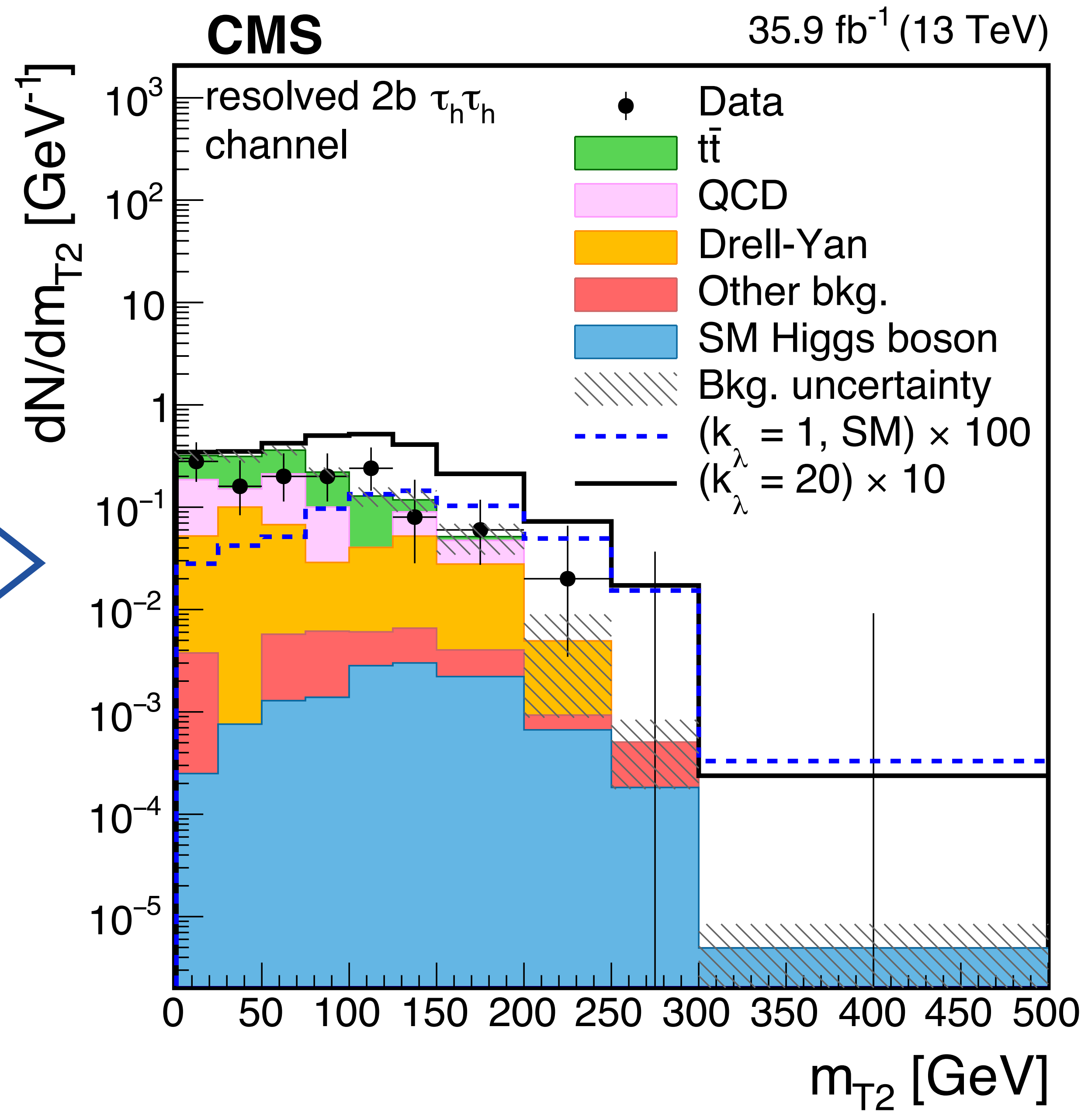
PAS-HIG-16-012/013 — 2.7 fb^{-1}

First LHC 13 Tev HH result
Assessment of techniques and sensitivity



PAS-HIG-16-028/029 — 12.9 fb^{-1}

Intermediate result for ICHEP 2016
Introduction of boosted categories to explore the high mass range



PLB 778 (2018) 101 — 25.9 fb^{-1}

First LHC full Run 2 result (Moriond 2017)
Optimisation of techniques to exploit the large 2016 CMS dataset

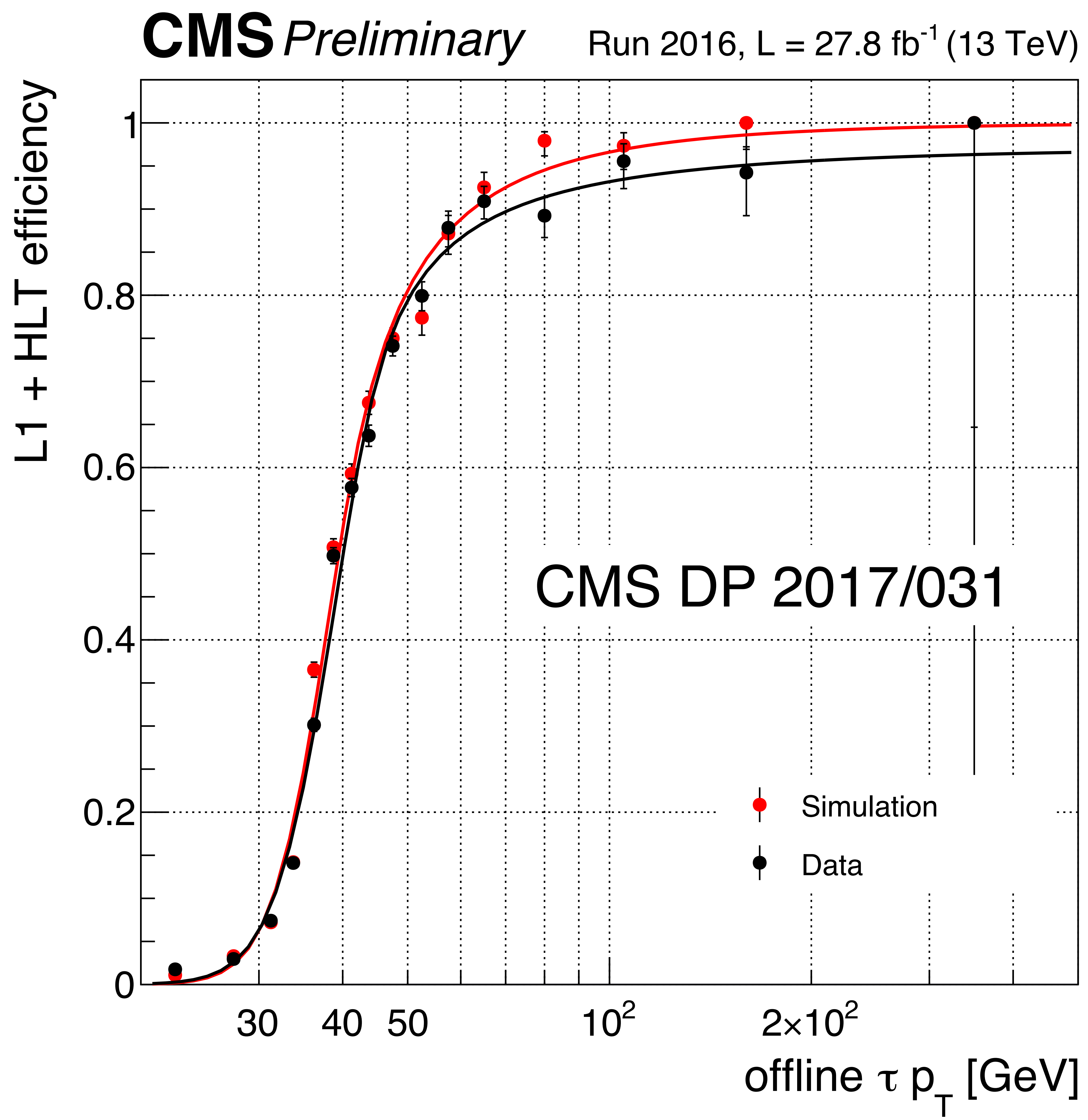
Constant improvements of methods ⟹ sensitivity increasing faster than luminosity
Good prospects for full Run II results and beyond

Resolved search strategy

Trigger requirements	Record data for offline analysis
Object preselections	Ensure reconstruction and quality of the objects Reject instrumental backgrounds
$H \to \tau\tau$ and $H \to bb$ candidates	Choice of the best object pair
Event categorisation	Assessment of $\tau\tau$ decay mode Categorisation on bb topology and b-tag
Signal regions	Selection on $\tau\tau$ and bb invariant mass BDT to reject tt contamination in $\tau_{\mu}\tau_{h}$ / $\tau_{e}\tau_{h}$

Trigger requirements

- $\ell\tau_h$: **single lepton triggers** (1μ or 1e)
  - moving to cross-triggers (ℓ + τ_h) for full Run 2 to increase acceptance and reduce rate
- $\tau_h\tau_h$: **double τ_h trigger**
  - low thresholds (35 GeV) thanks to L1 trigger upgrade in 2016
  - tuning of L1 and HLT algorithms to sustain the increase of luminosity in 2017 and 2018 without increasing the thresholds
- Boosted analysis: **missing E_T / H_T triggers**

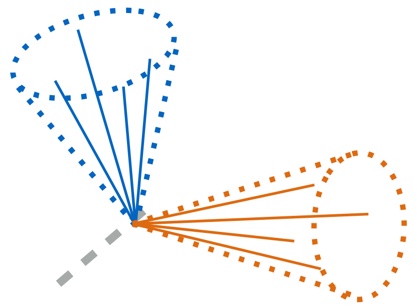


More in Agni's and Chiara's talks

Event categorisation

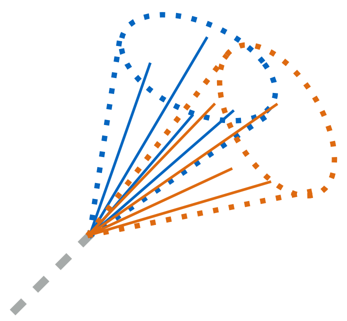
1 AK8 jet, p_T > 170 GeV, m_{AK8} > 30 GeV
2 subjets matched to preselected AK4 jets

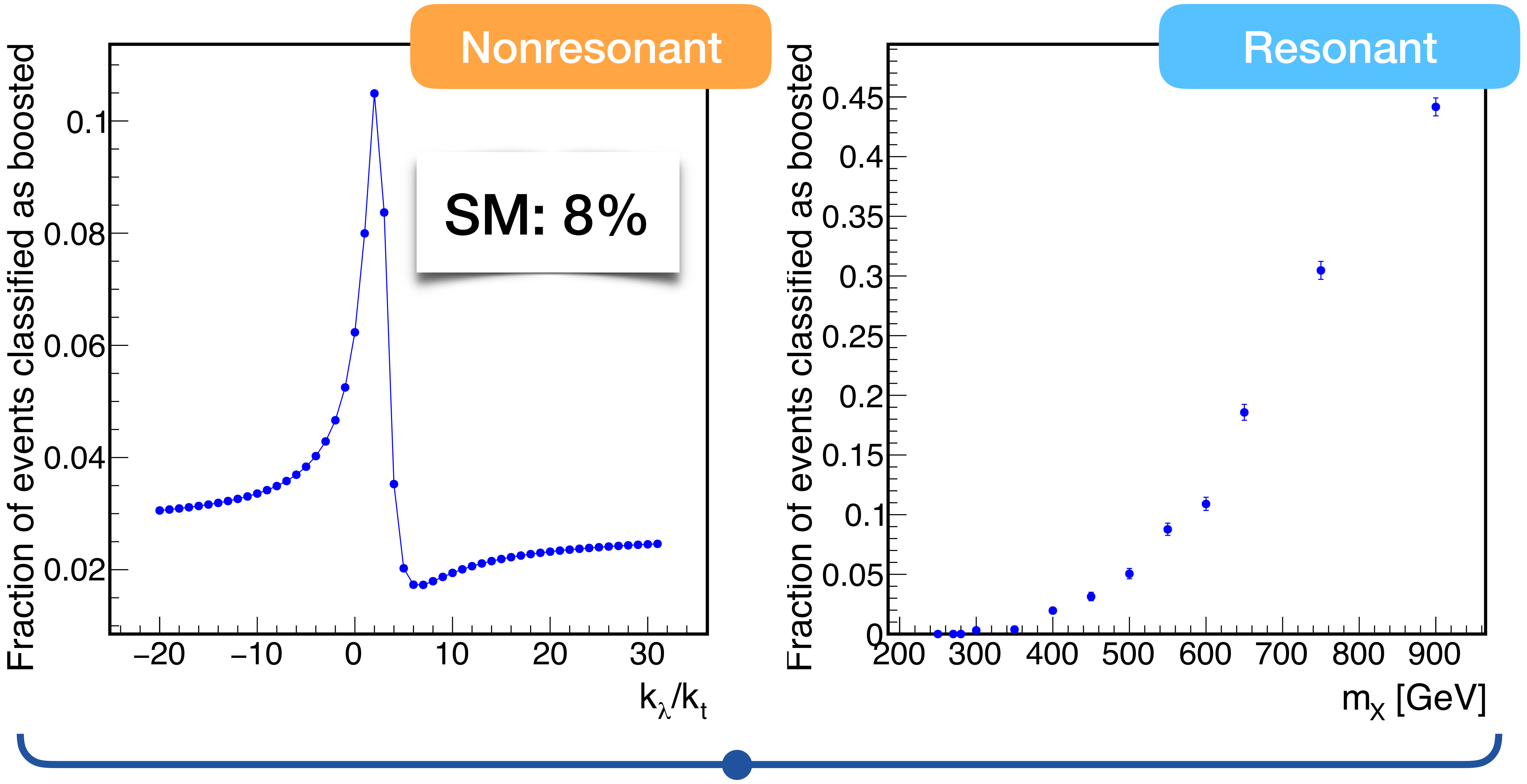
✗ → Number of b-tagged jets?



- 1 jet → **1b1j resolved**
- 2 jets → **2b resolved**

✔ → Both AK4 jets b-tagged (loose WP) → **boosted**



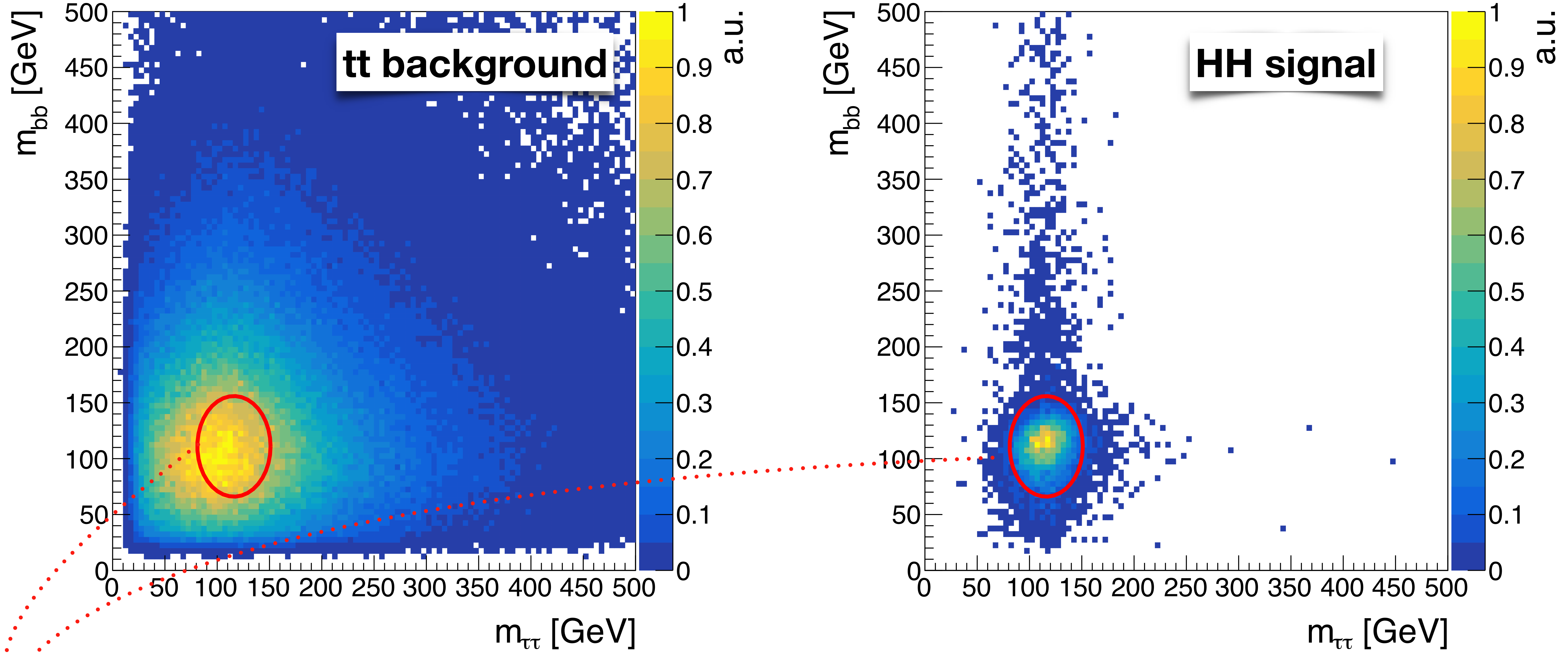


1b1j and 2b resolved categories to increase signal acceptance

Boosted topologies improve sensitivity for high-mass resonances

Signal regions

- H→$\tau\tau$: **likelihood method** to reconstruct m_H (SVFit)
- H→bb: AK4 jet pair invariant mass (resolved) or AK8 jet mass (boosted)



Resolved:

$$\frac{\left(m_{\tau\tau} - 116\,\mathrm{GeV}\right)^2}{(35\,\mathrm{GeV})^2} + \frac{\left(m_{\mathrm{bb}} - 111\,\mathrm{GeV}\right)^2}{(45\,\mathrm{GeV})^2} < 1$$

Boosted:

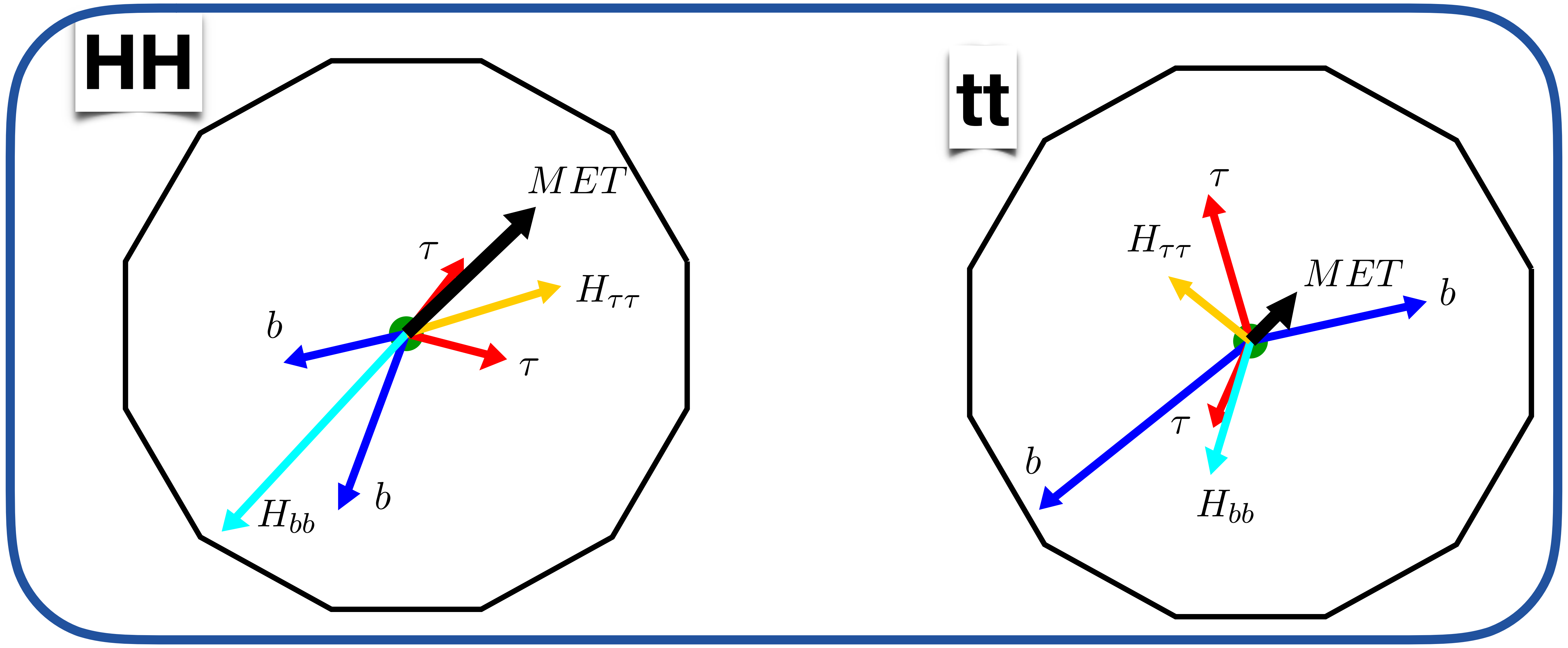
$$80 < m_{\tau\tau} < 152\,\mathrm{GeV}$$
$$90 < m_{\mathrm{bb}} < 160\,\mathrm{GeV}$$

- Selections based on the mean and resolution of $\tau\tau$ and bb mass distributions
  - signal efficiency: ~ 80%
  - background rejection: ~85%

Multivariate methods

tt is the dominant background in $\ell\tau_h$ final states

Dedicated MVA methods based on kinematics



- **8** input **geometrical** variables (minimise NLO effects, more stable vs m_X)

- Two separate trainings:
  - low-mass (LM, $m_X \leq 350$ GeV) + nonresonant
  - high-mass (HM, $m_X > 350$ GeV)

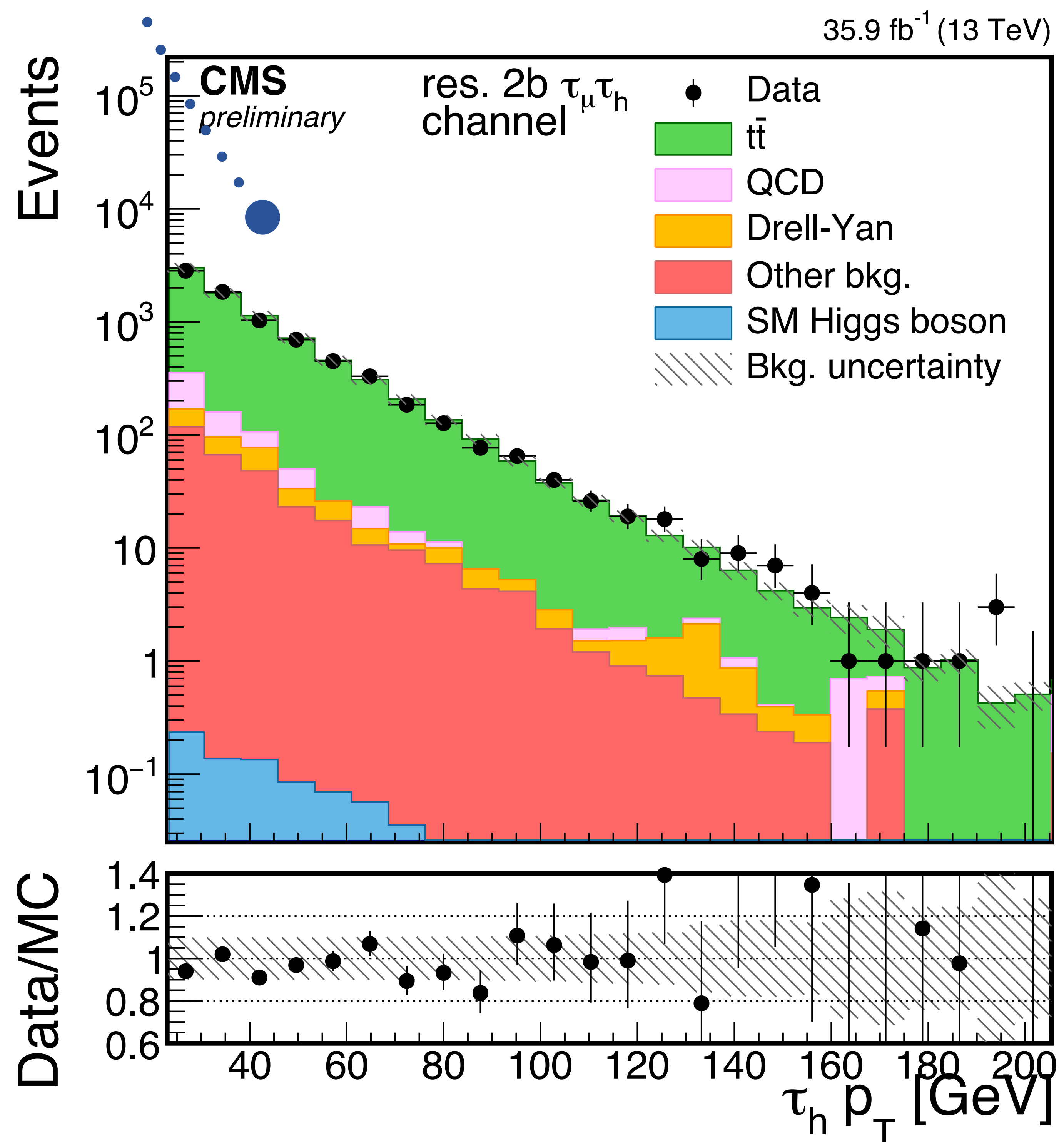
- 70-90% of tt rejection for 65-95% of signal efficiency
- **Sensitivity improved by 20-80% in $\ell\tau_h$**

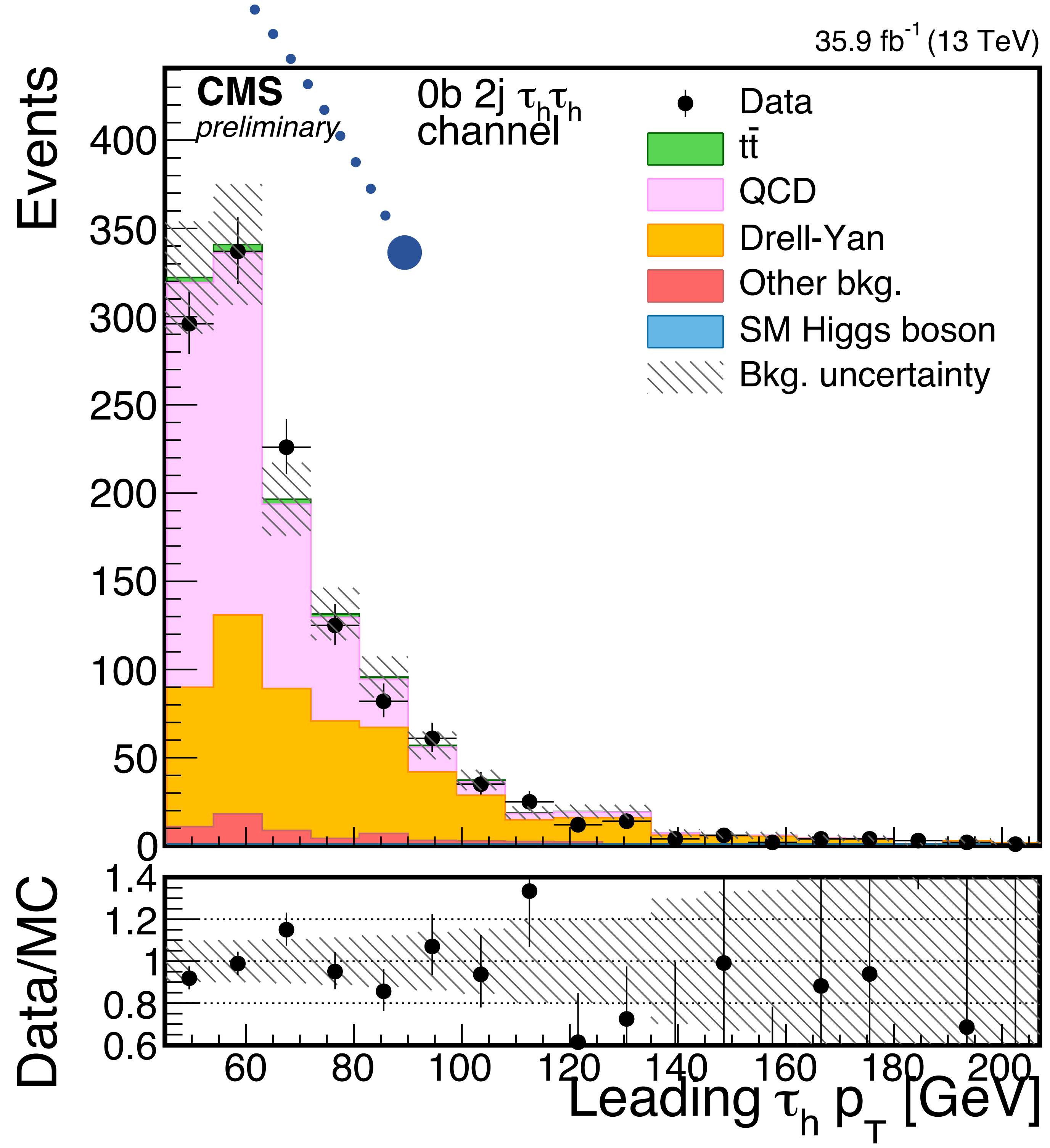
More in Arnaud's talk

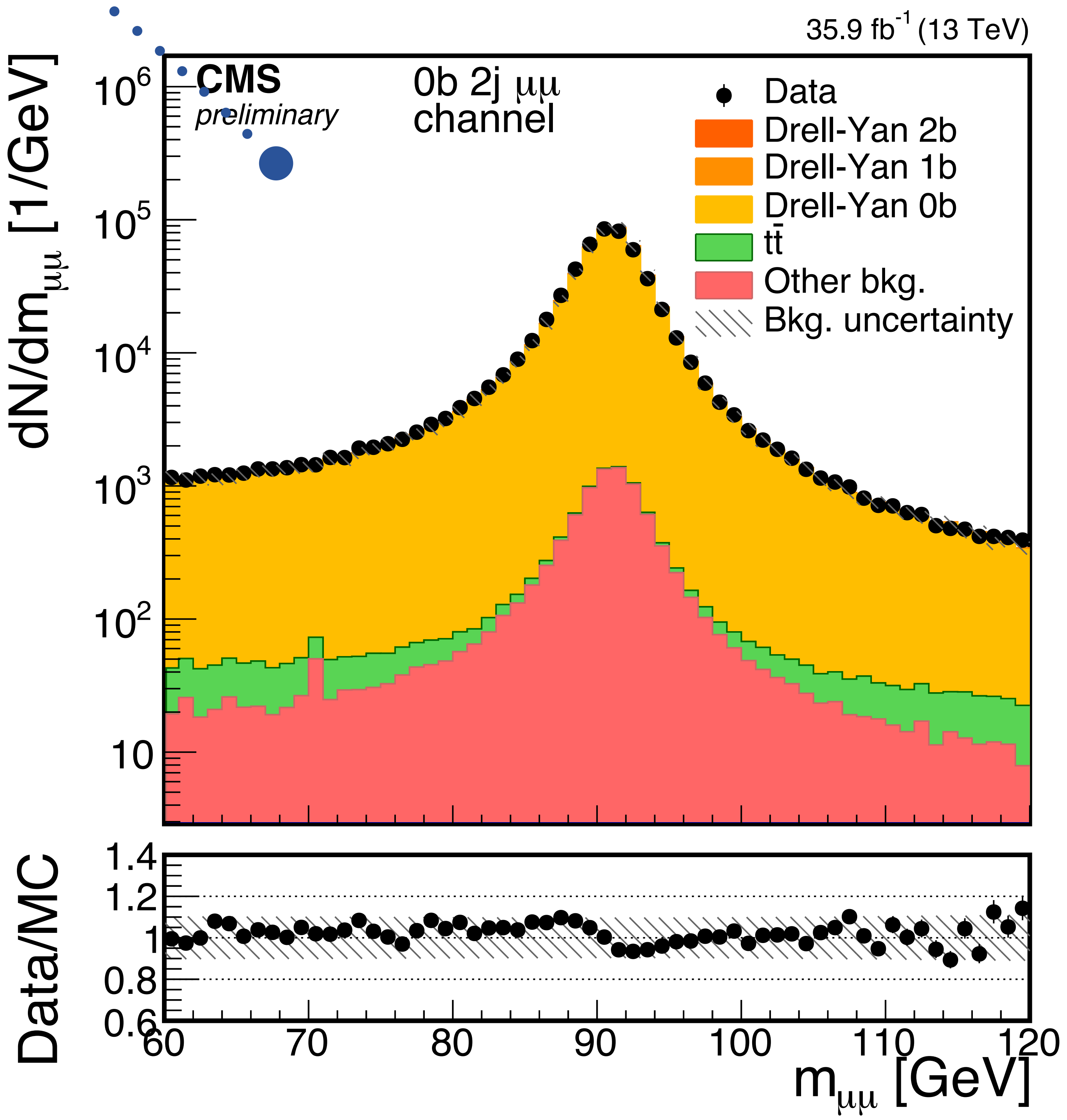
Background estimation

- tt: MC simulation + dedicated top p_T shape uncertainty
- Z/γ^*: MC + **data-driven correction**
- multijet : **data-driven**
- W+jets, single top, single Higgs, di-bosons, EWK W/Z production: MC simulation

More in Francesco's talk



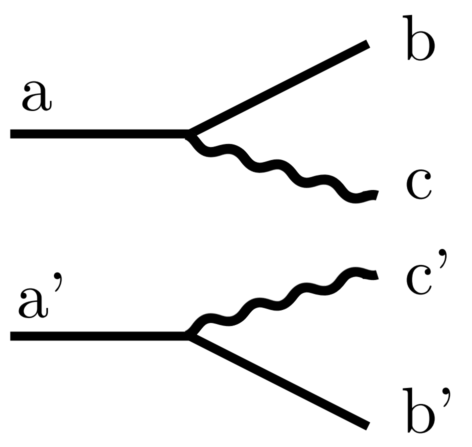




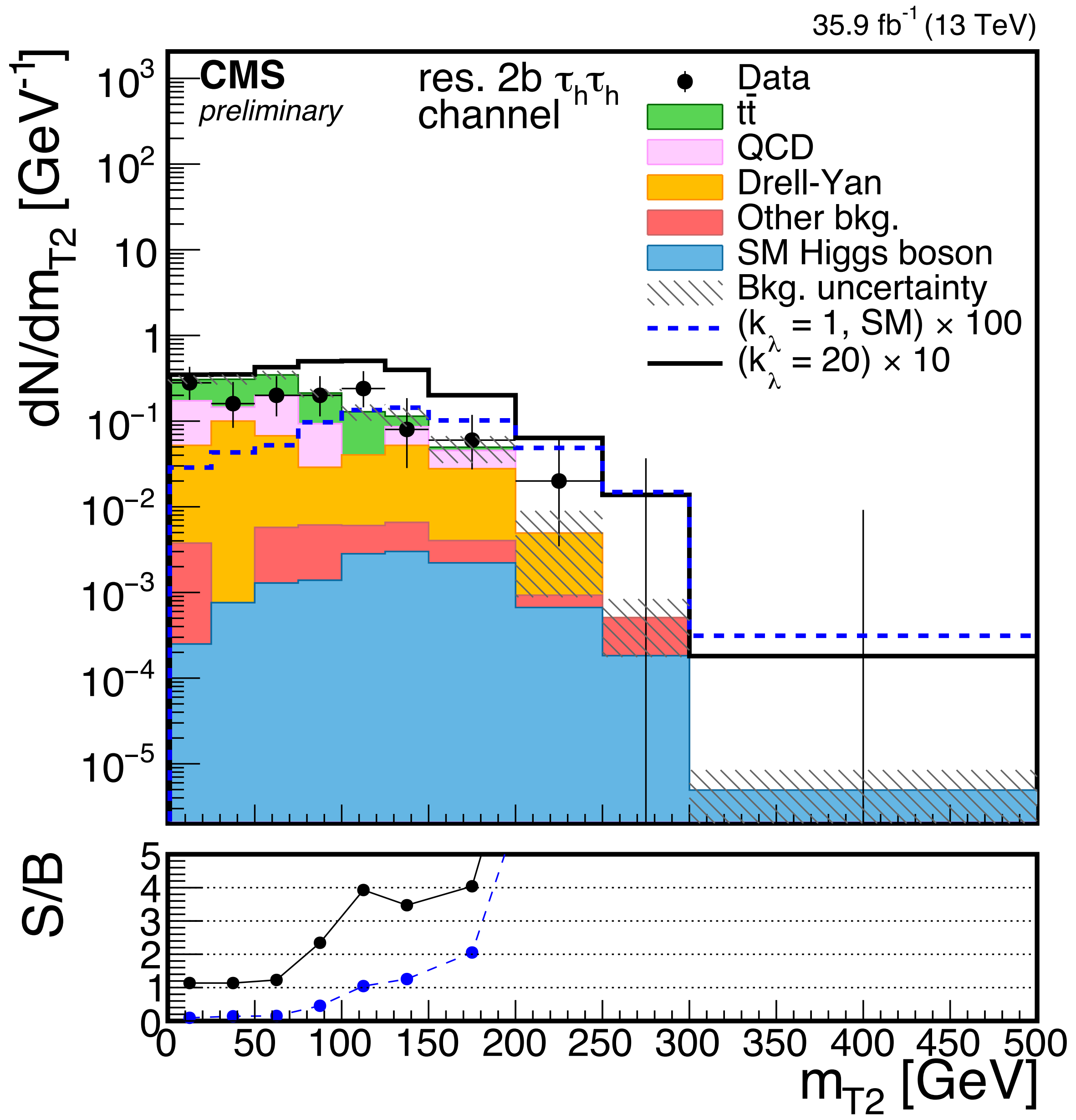
Discriminant observables

Nonresonant search | m_{T2}

- Generalisation of the transverse mass for two equal-mass particle decays
  - optimally separates tt and other backgrounds (low m_{T2}) from signal (high m_{T2})
  - improve the sensitivity to SM by ~40% over visible $bb\tau\tau$ invariant mass

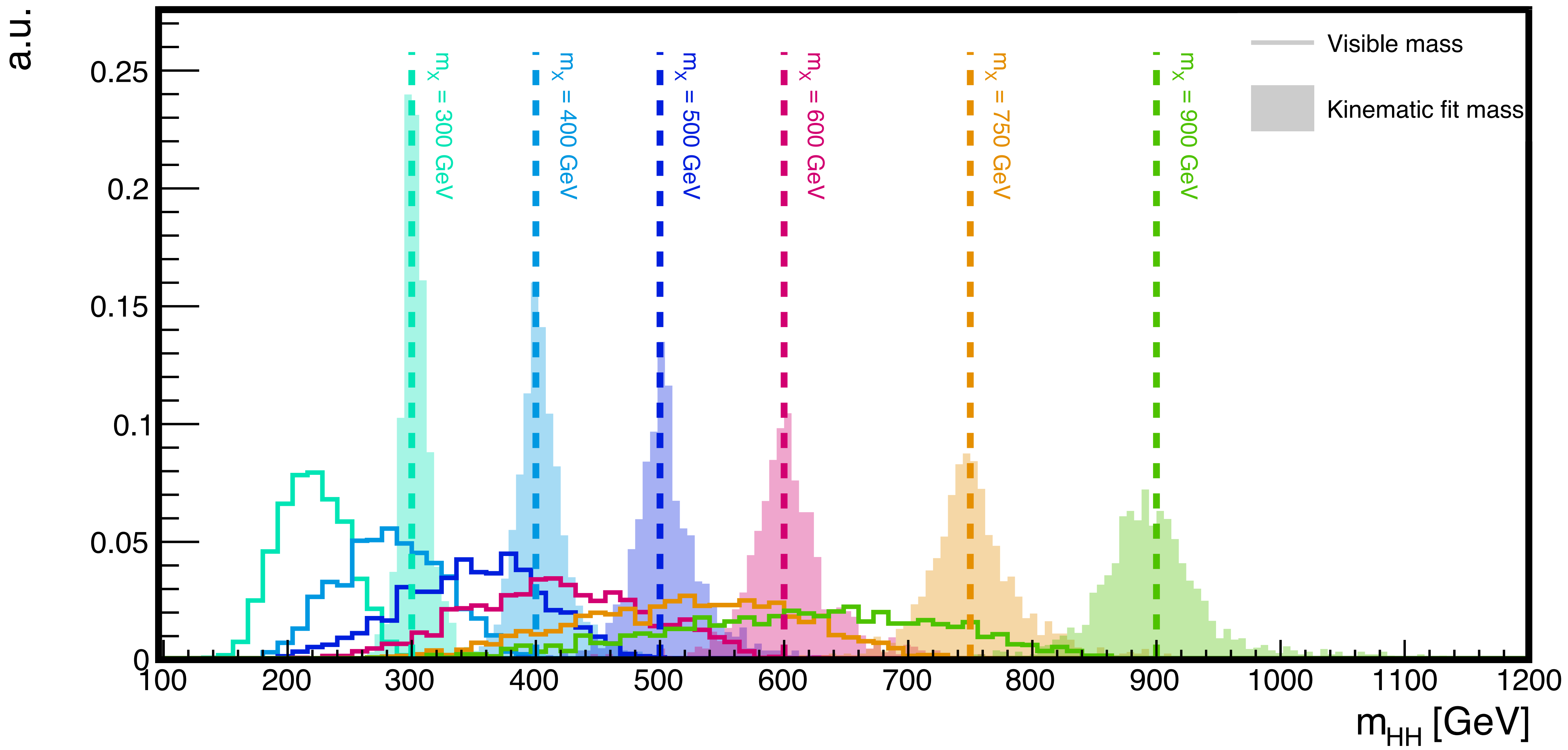


$$m_{T2}\left(m_b, m_{b'}, \vec{b_T}, \vec{b'_T}, \vec{\Sigma}_T, m_c, m_{c'}\right) = \min_{\vec{c_T}+\vec{c_T}'=\vec{\Sigma}_T}\{\max(m_T, m'_T)\}$$



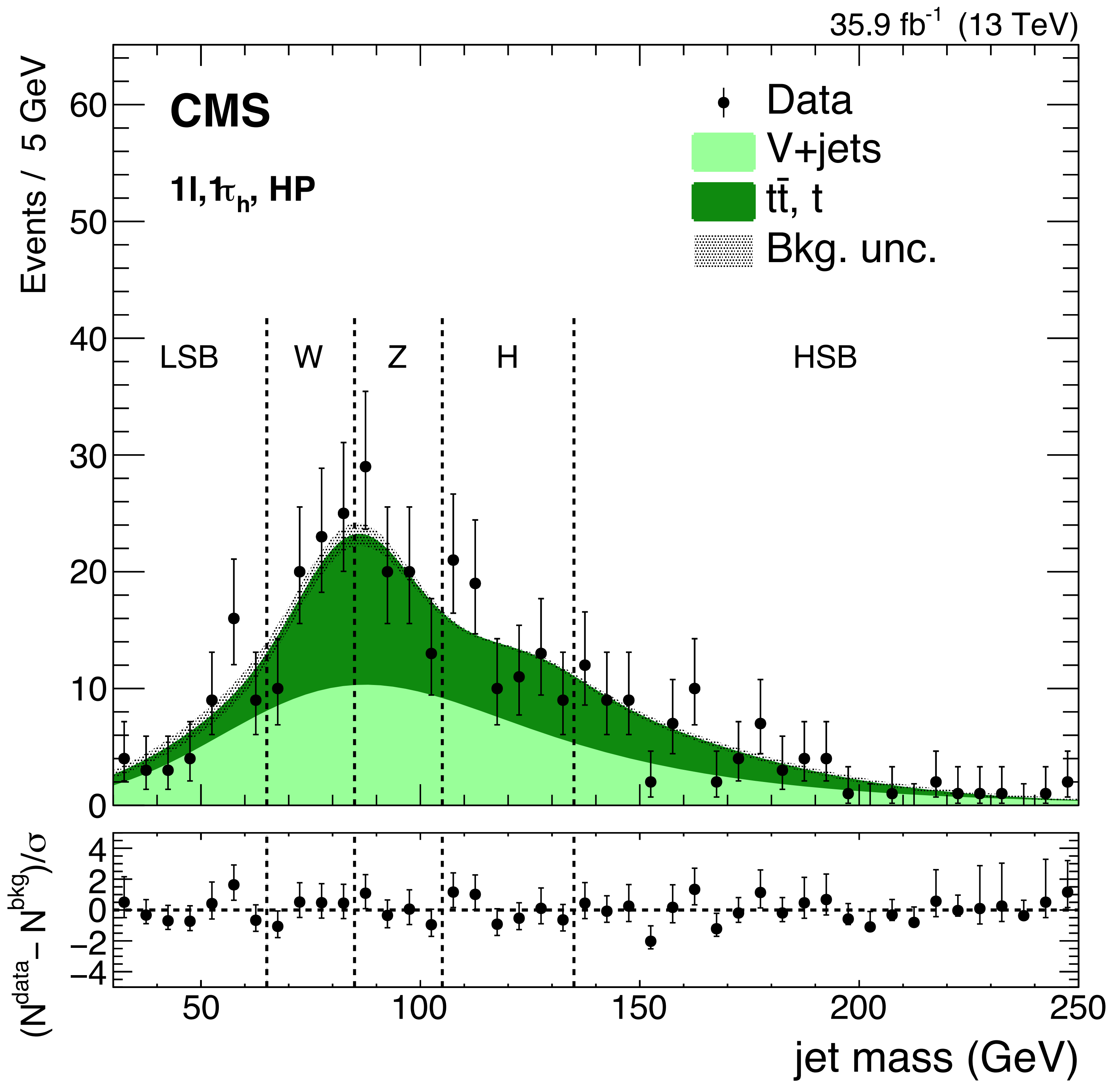
Resonant search | m_{HH}^{KinFit}

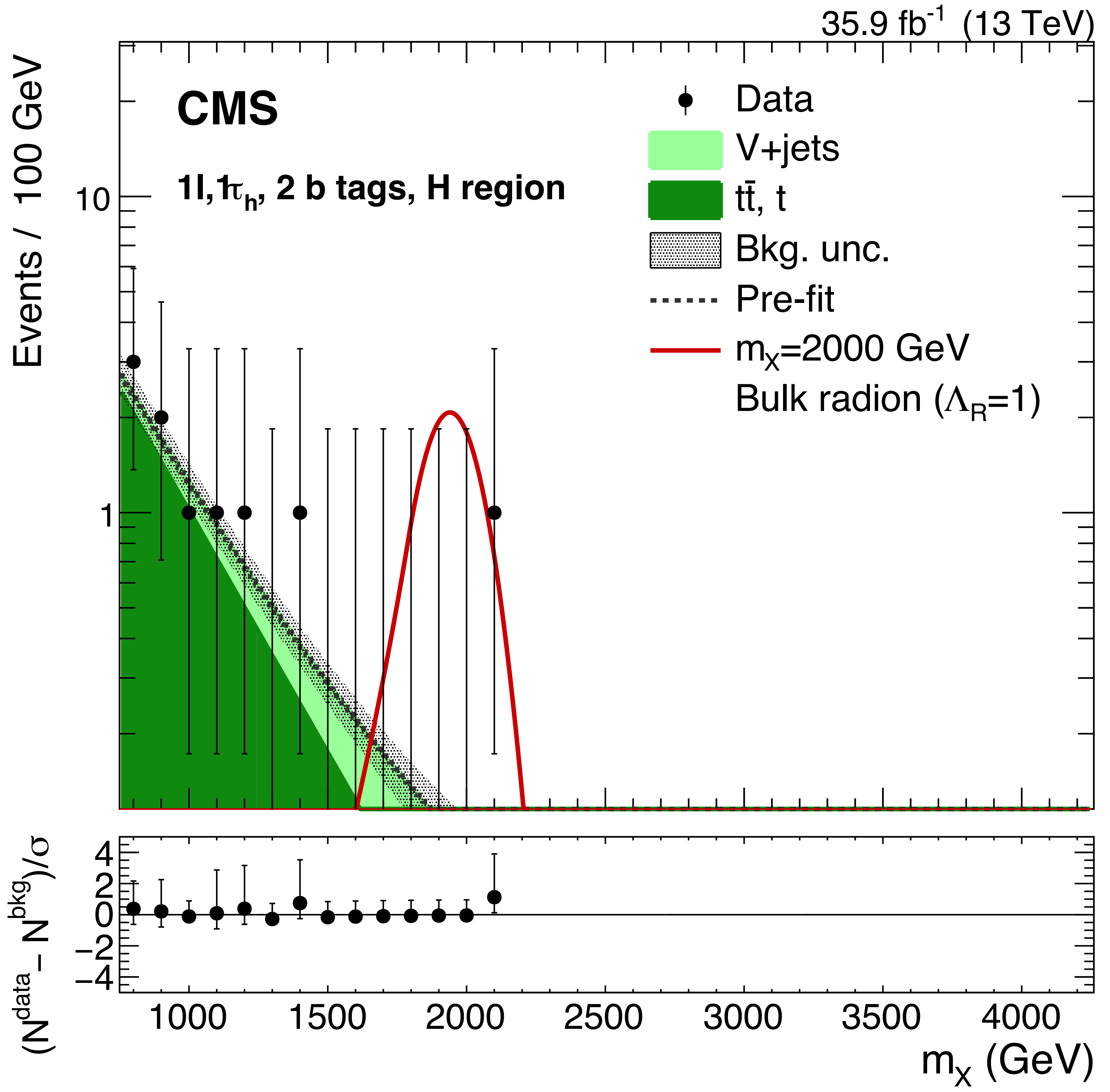
- Kinematic reconstruction of the $bb\tau\tau$ system
  - resolution improved by a factor of 4 over visible $bb\tau\tau$ invariant mass
  - reconstruction of the true m_X value



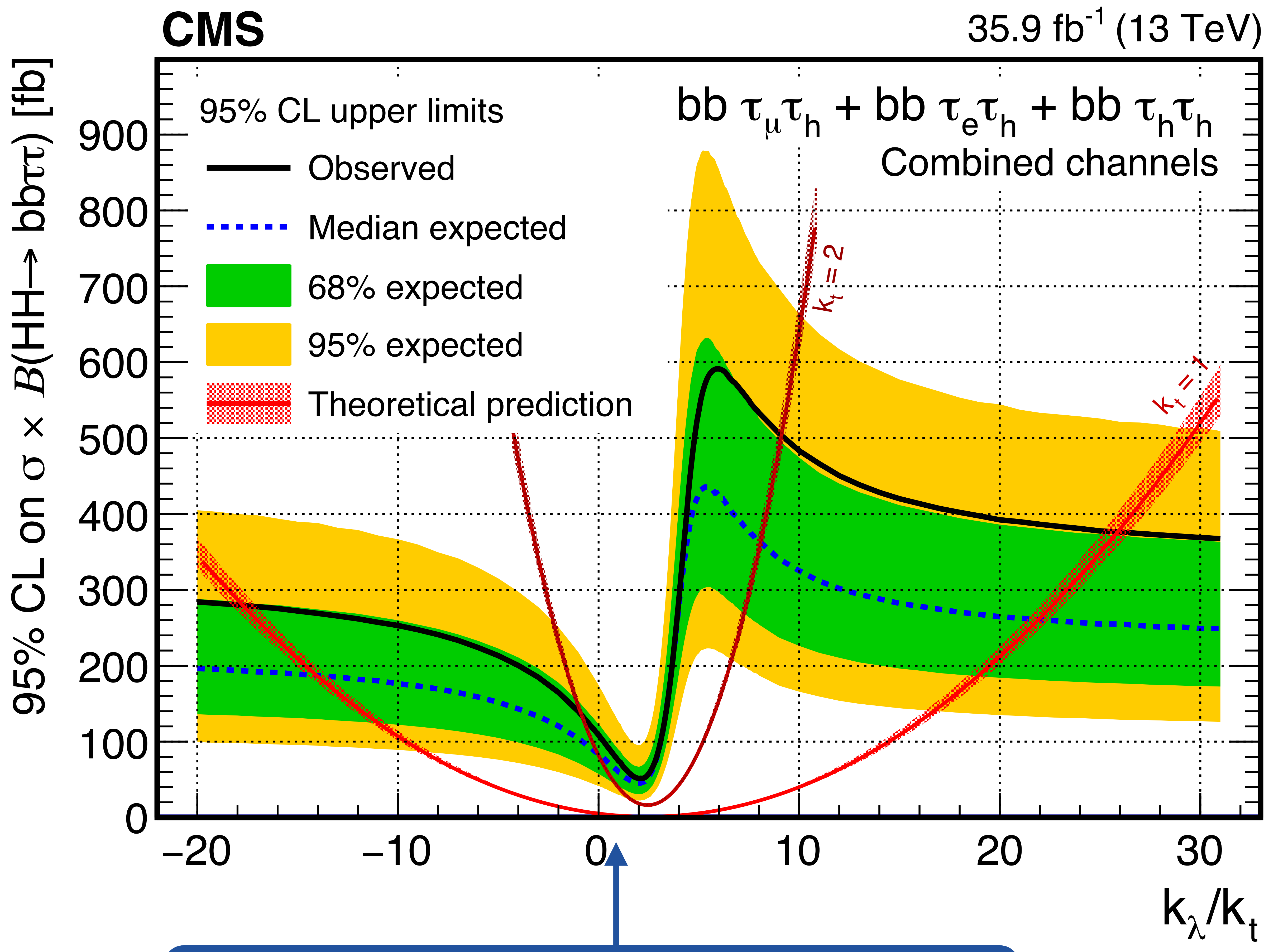
High mass resonant search

- Substructure, mass and b tag criteria applied to the H→bb jet candidate
- H→$\tau\tau$ mass from SVFit, then combined with bb into m_X
- Use data to estimate backgrounds
  - top : MC + control region corrections
  - V+jets : data driven estimation with "alpha ratio" method (based on shape parametrisation and fit in control regions + transfer in signal region)
- Look for a signal in the m_X spectrum





Nonresonant production



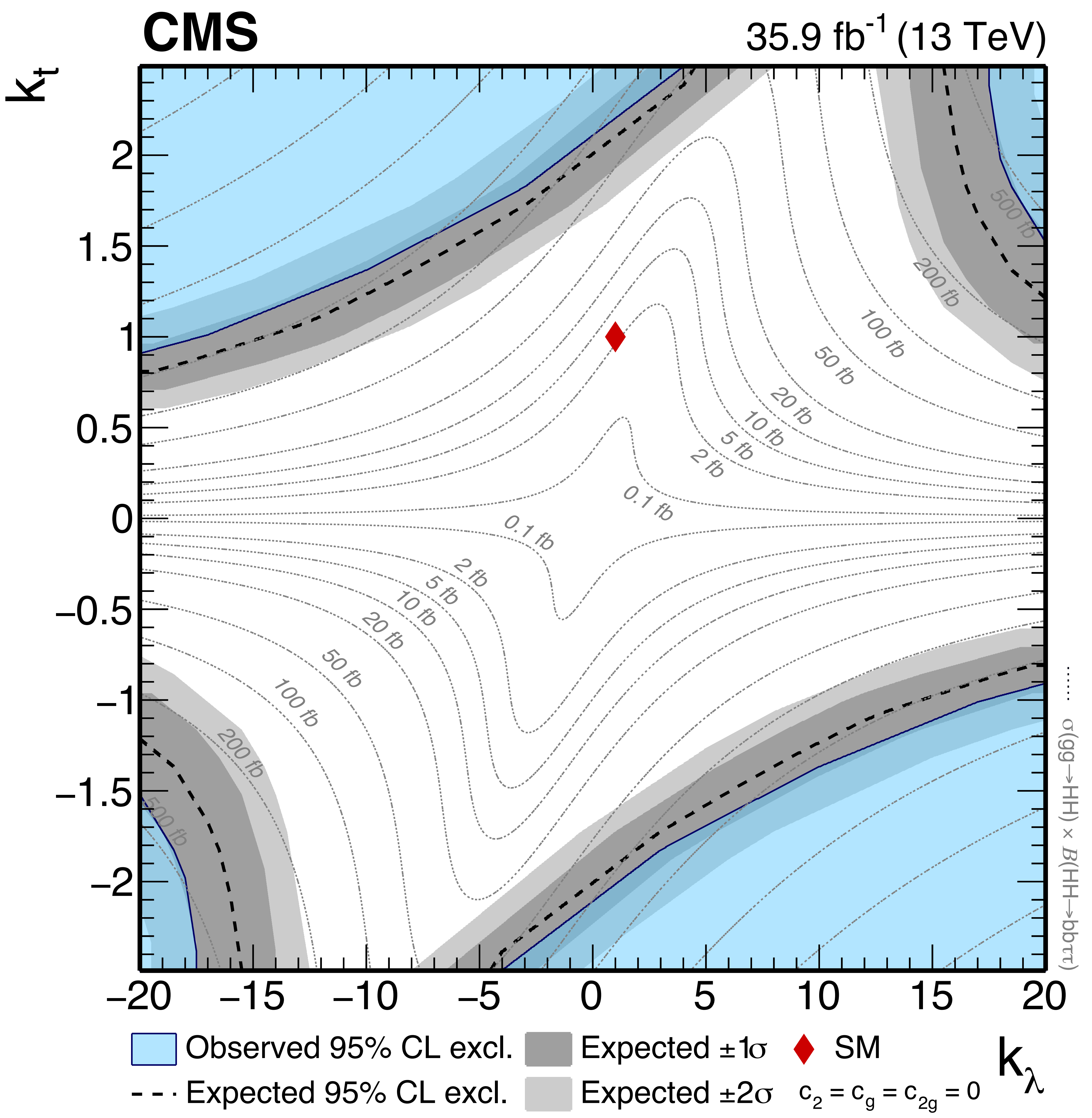
SM: obs. (exp) 75.4 (61.0) fb = 31 (25) x σ_{HH}^{SM}

including theoretical signal uncertainties

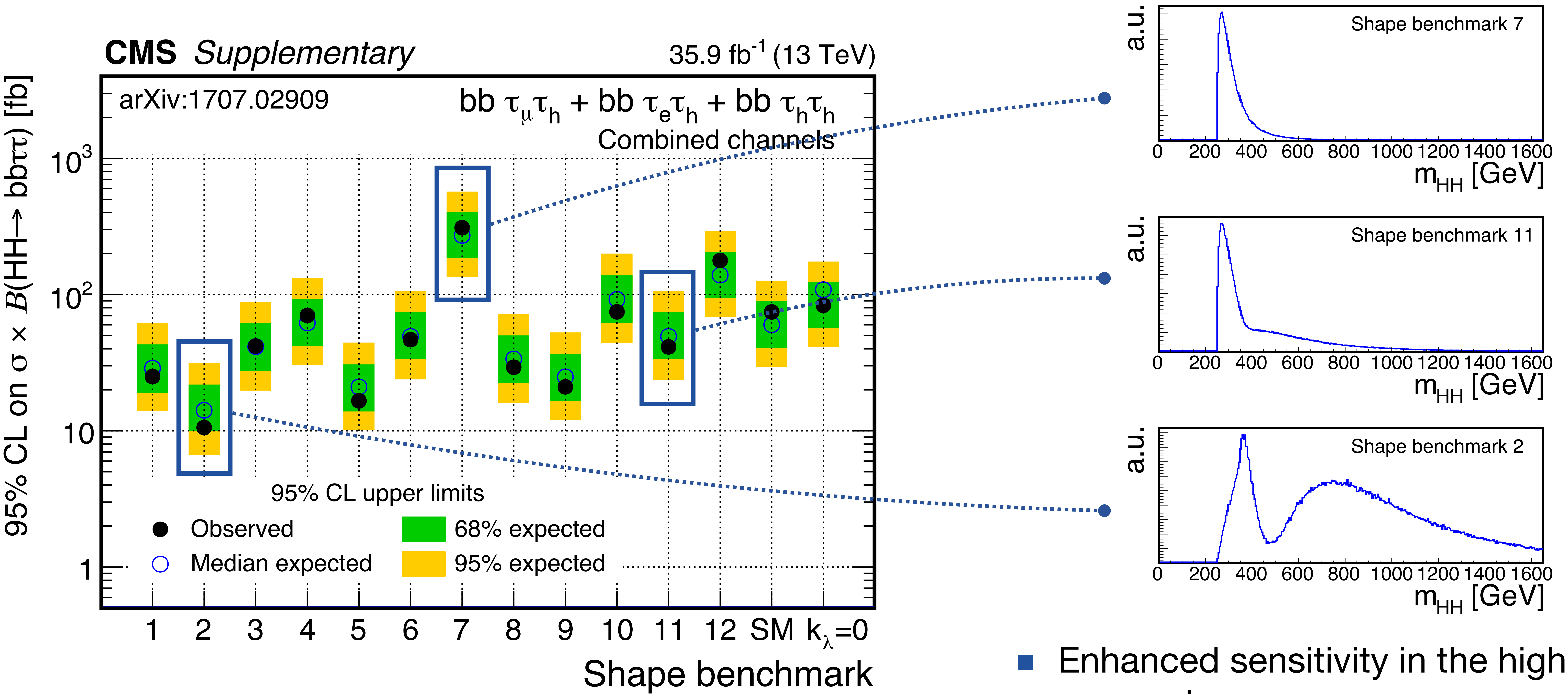
- Limit set on σ_{HH} as a function of the k_λ/k_t ratio
- Exclusion depends on HH kinematics, that affects acceptance and m_{T2} shape

Nonresonant production

- 2D exclusions at 95% CL of anomalous λ_{HHH} and y_t couplings
- Cross section and signal modelling symmetric under $(k_\lambda, k_t) \leftrightarrow (-k_\lambda, -k_t)$ exchange
- If SM k_λ assumed, sensitive to the sign of the top-Yukawa coupling

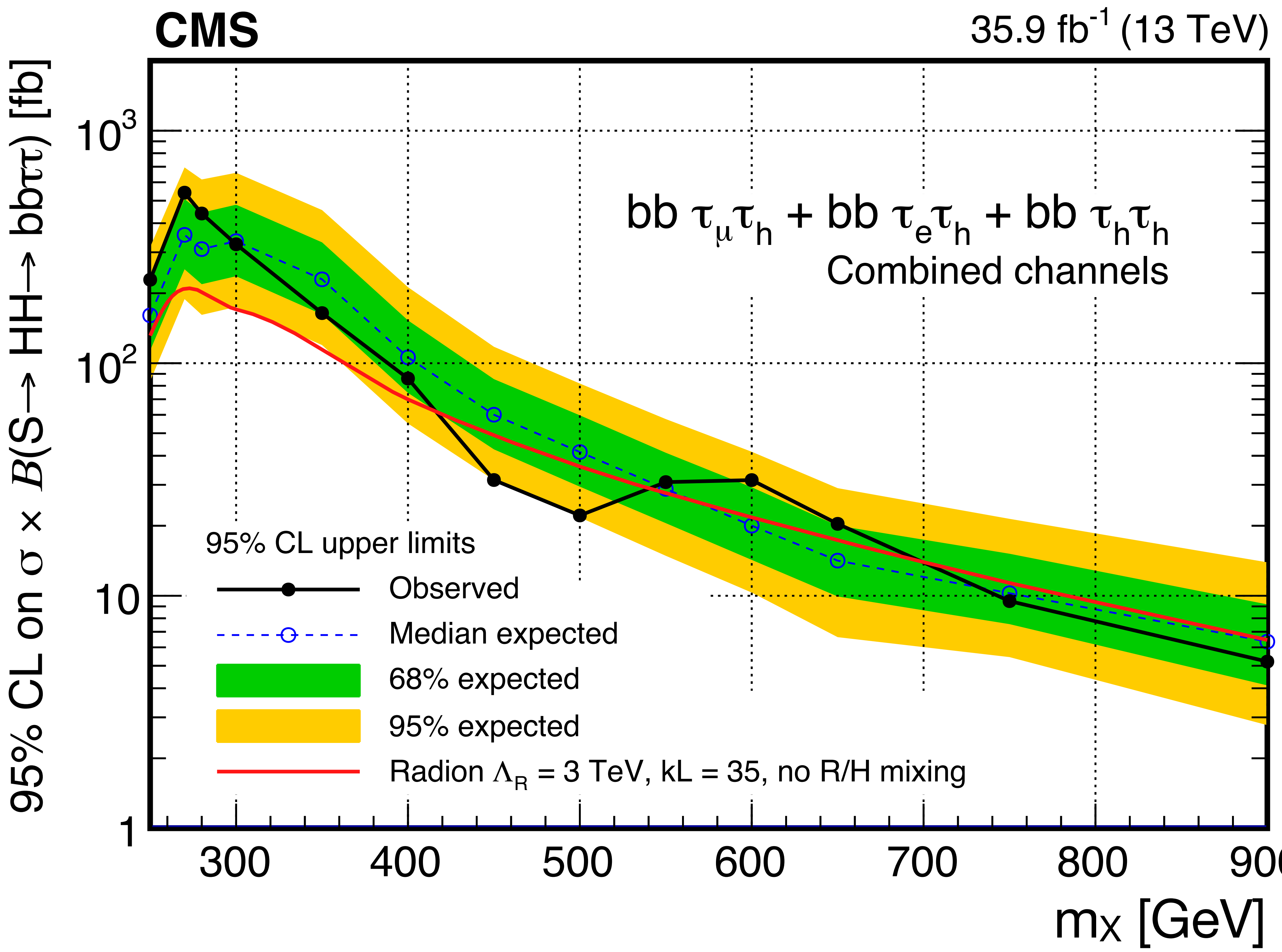


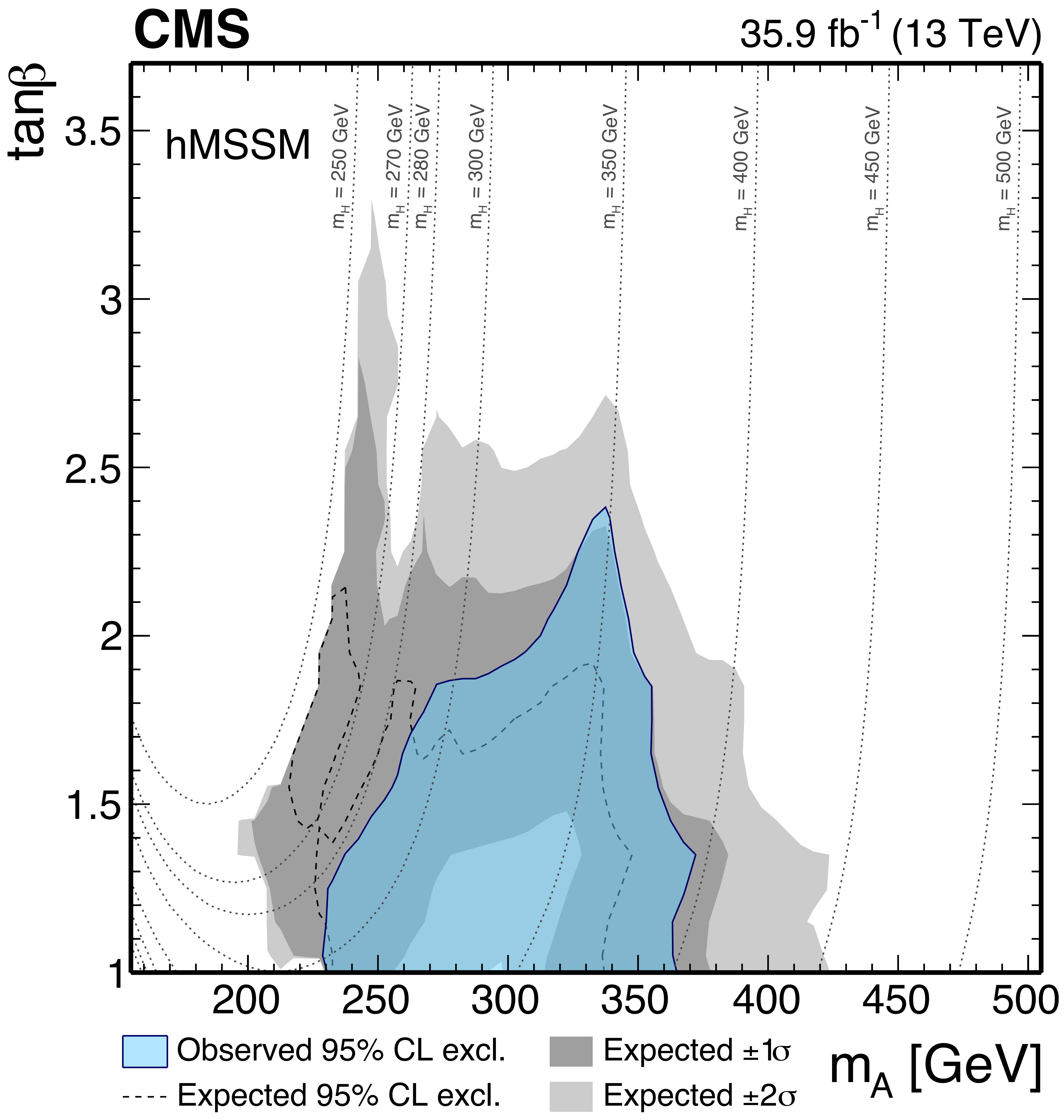
Nonresonant production



- Enhanced sensitivity in the high m_{HH} region

Resonant production - resolved

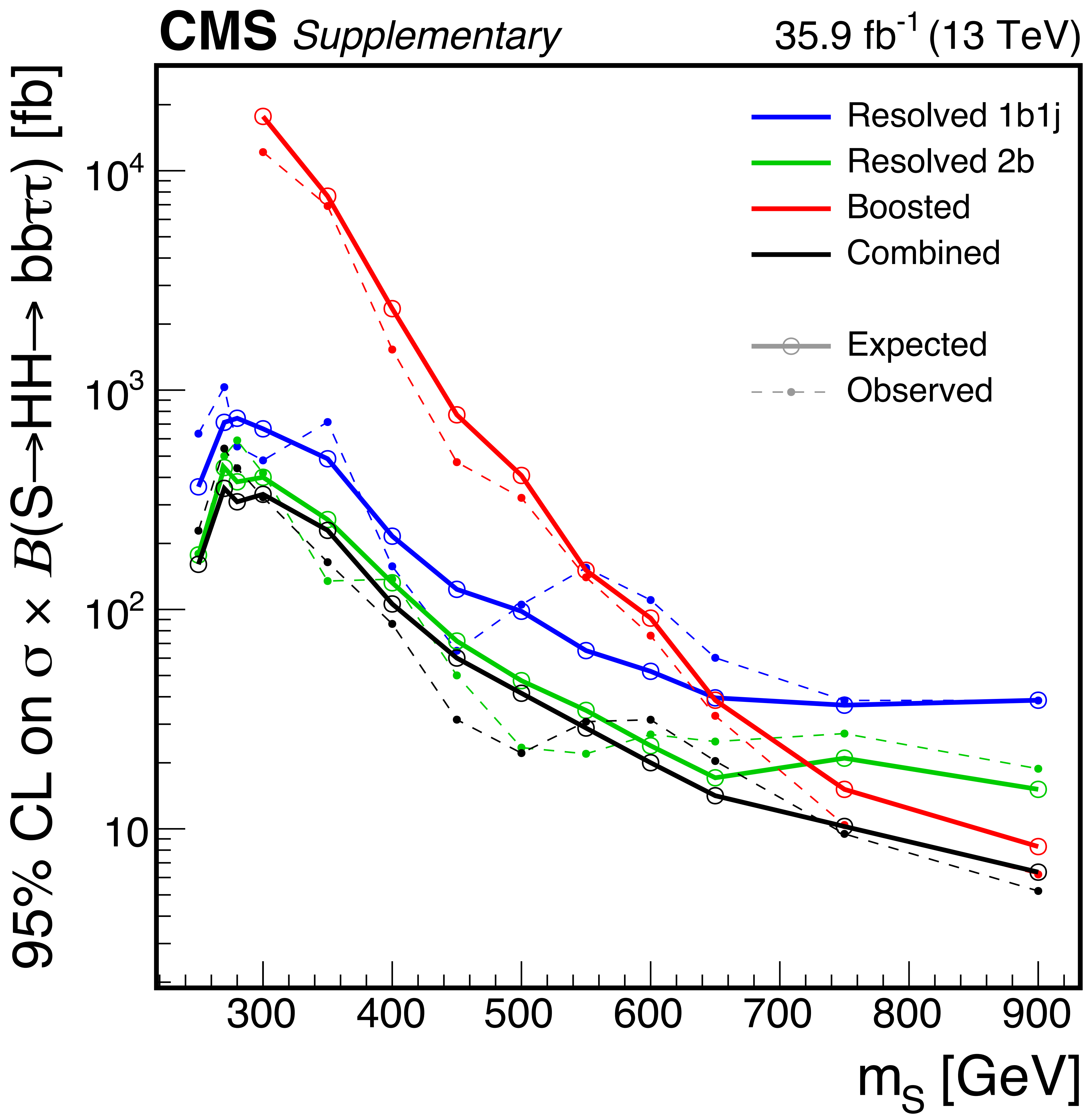




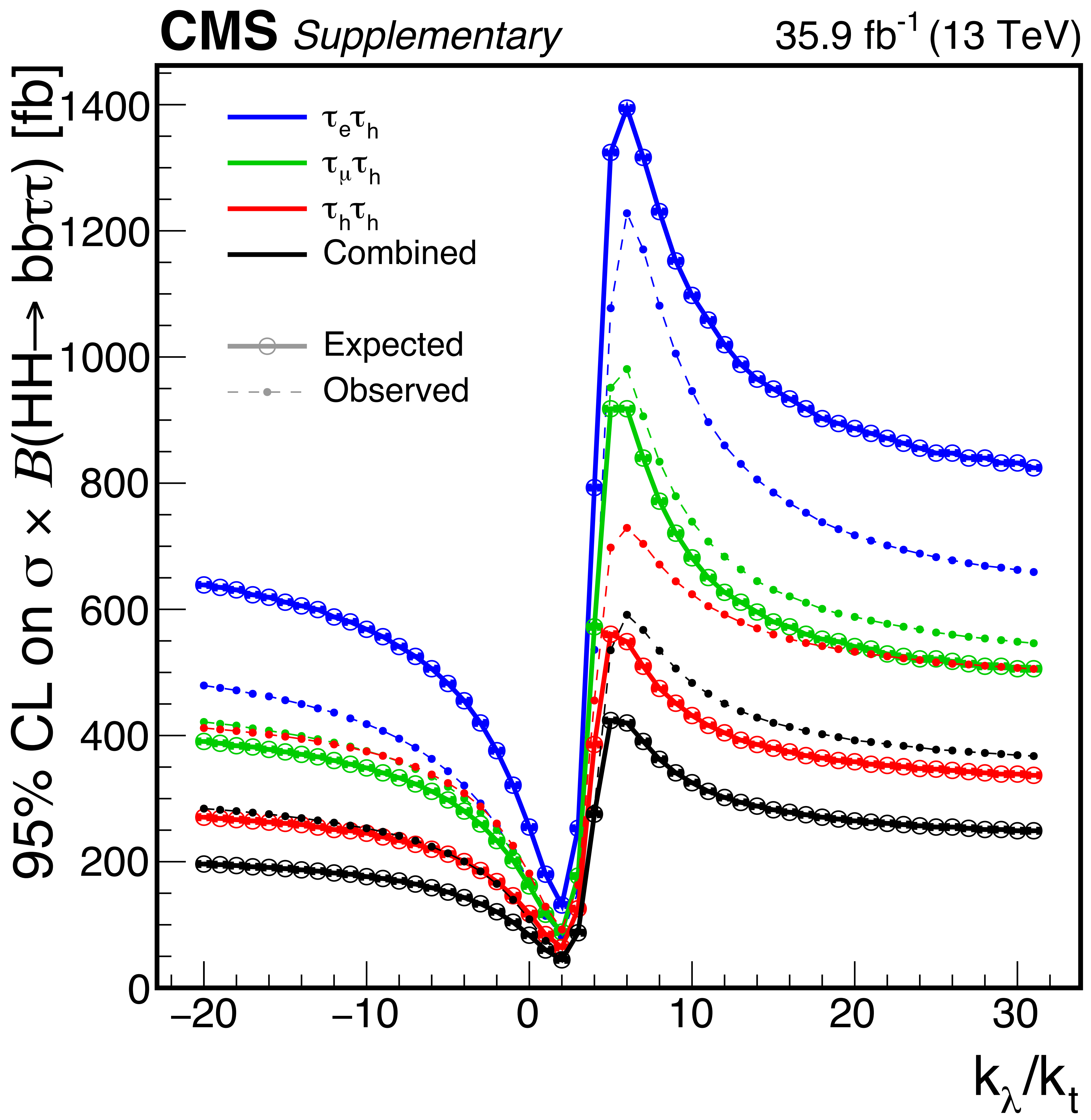
- 95% CL upper limits set on σ(gg→HH) x B(HH→bb$\tau\tau$)
- Limits set for both spin-0 and spin-2 hypotheses
- Reinterpreted in MSSM scenarios

How do categories and channels compare?

Categories

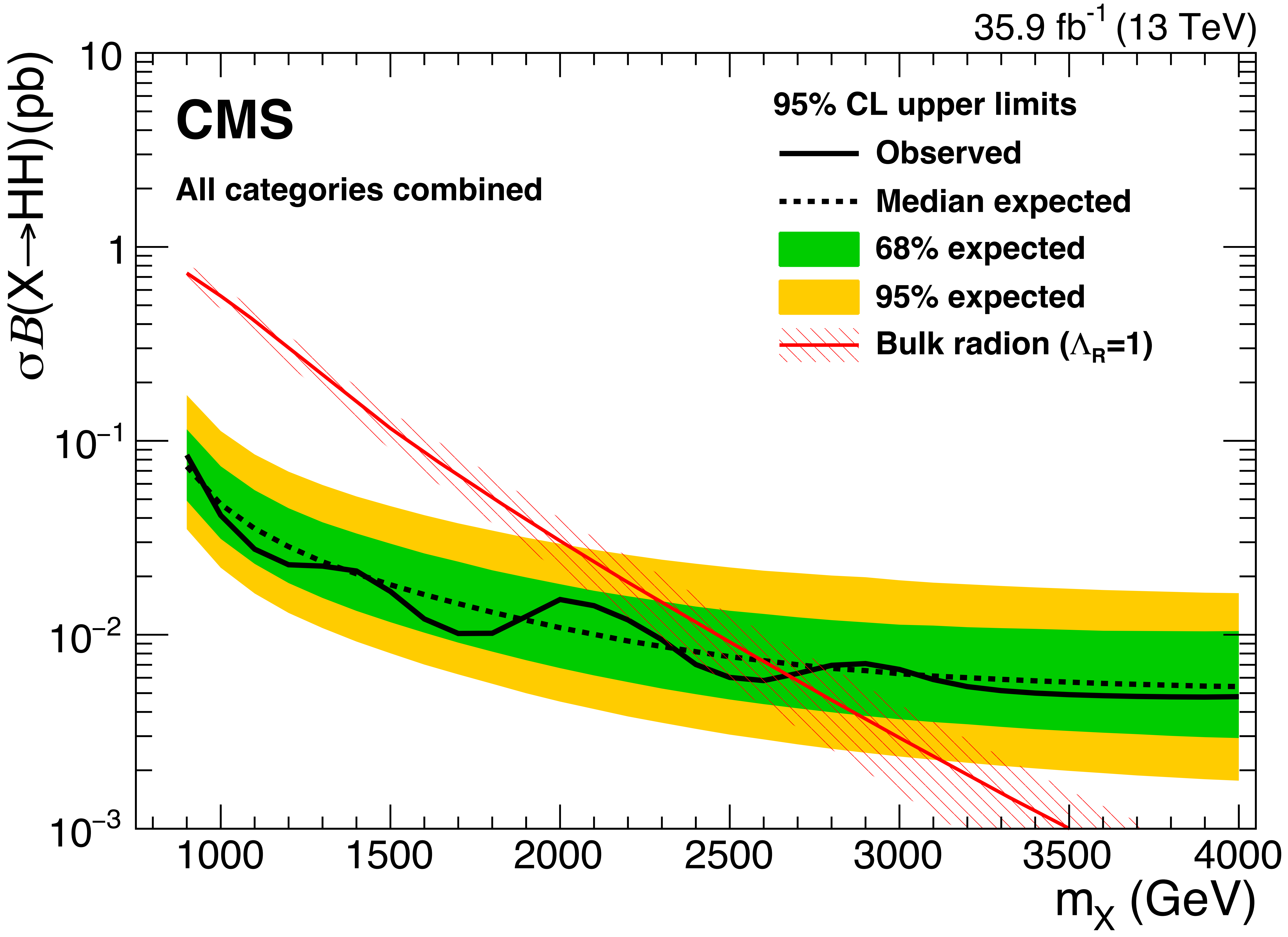


Channels



- $\tau_h\tau_h$ is the most sensitive channel
  - $\ell\tau_h$ is important in the low mass region because of better acceptance
- 2b resolved category is the most sensitive
  - boosted category sensitive beyond ~700 GeV

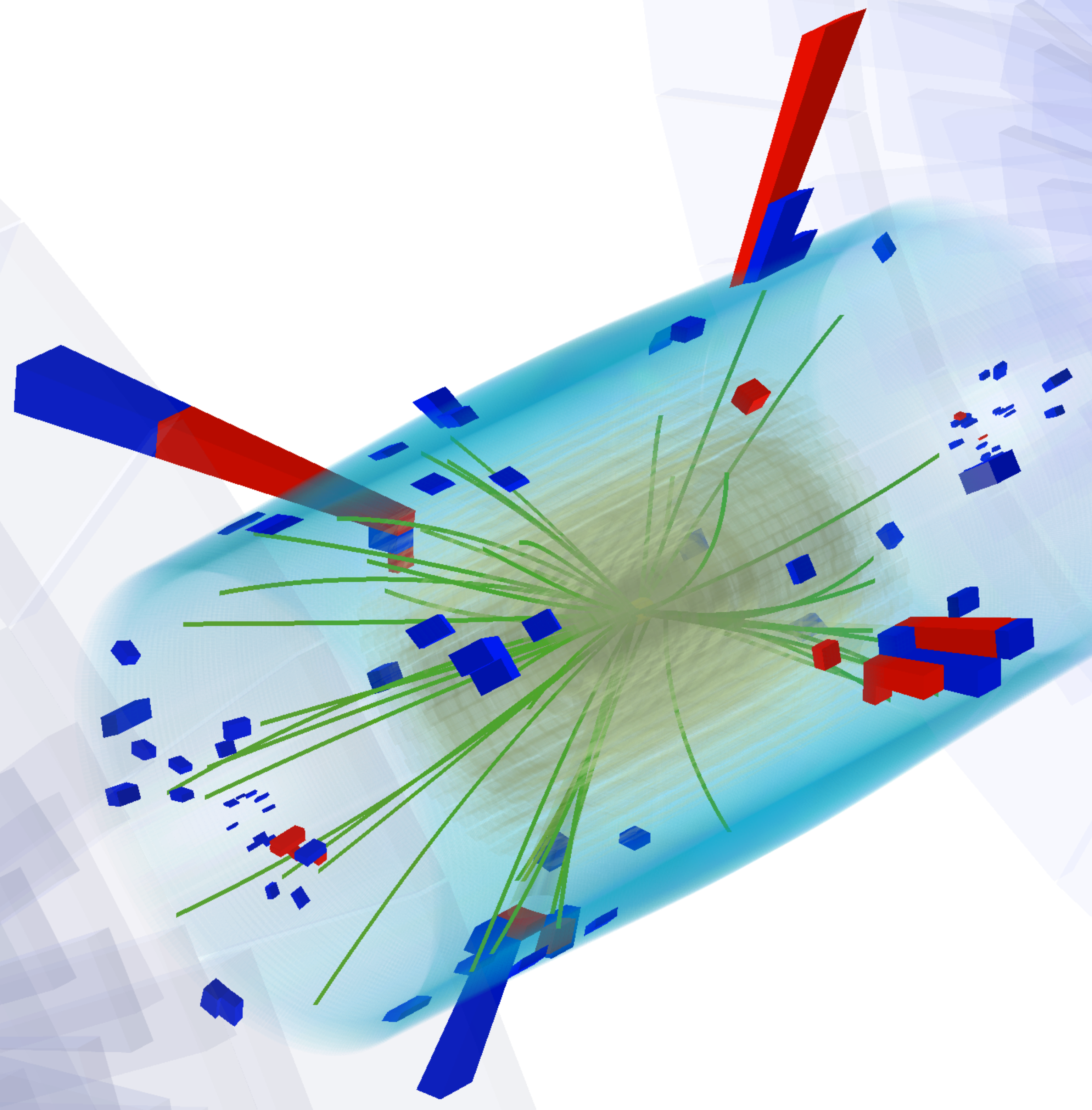
Resonant production - boosted



- 900 GeV - 4 TeV mass range explored
  - low edge corresponding to the transition from resolved to boosted topologies
- Both spin 0 and spin 2 resonances studied
- The HH analysis is part of a broader program of study of HH, WH and ZH production that is sensitive to extra dimensions and heavy vector triplet models
  - first 13 TeV LHC results in this final state and mass range

Conclusions

- The bb$\tau\tau$ decay channel is one of the key elements in the HH exploration at the LHC
  - favourable BR and background contamination
  - experimentally challenging final states
- CMS has developed a broad program of bb$\tau\tau$ analyses
  - resonant and nonresonant production
  - dedicated analysis for the high boost regime
- Resonant and nonresonant production investigated
  - SM and anomalous coupling signals
  - Resonant production and reinterpretations
- Current results based on 2016 data are 1.5 years old!
  - with the largest datasets collected in 2017 and 2018 (~ 4× luminosity) expect large improvements



HH→bb$\tau\tau$ event candidate selected in 2016 data