CMS Compact Muon Solenoid

Emmy Noether-Programm
Deutsche Forschungsgemeinschaft DFG

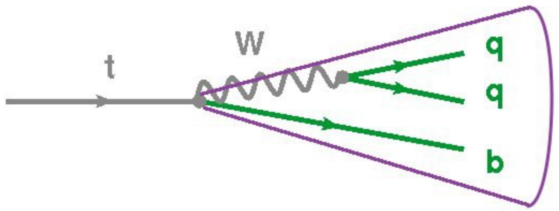
UHH Universität Hamburg

# Search for vector-like T and B quarks in all-hadronic final states at CMS

***Ivan Marchesini***
***on behalf of the CMS Collaboration***
***VLQ 2014***

# Outline

- Heavy mass of the VLQ quark:
  - → **boosted** hadronic decay products, merged jets
  - → **challenging** hadronic final states



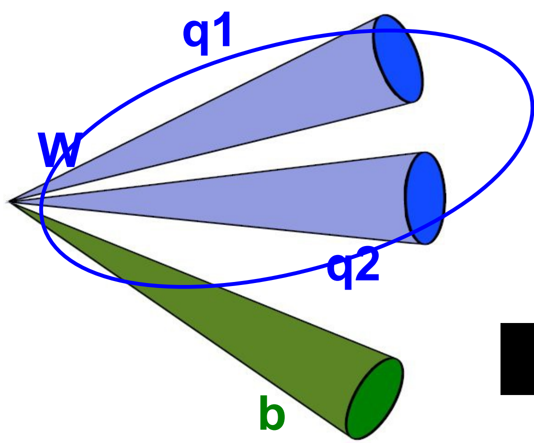

- Accessing boosted hadronic final states:
  - → **substructure**:
    - PAS-B2G-14-002: T→tH, H→bb, full hadronic
    - PAS-B2G-14-001: B→bH, H→bb
  - → clean **signatures with photons**:
    - PAS-B2G-14-003: T→tH, H→γγ

    (this search includes a leptonic category, included for completeness)

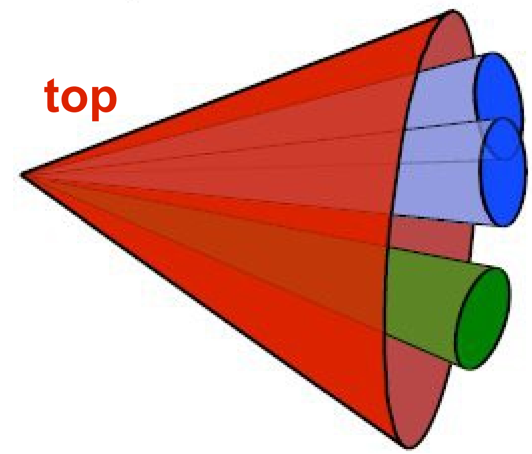
# Searches with substructure

# Boosted regime

**Non-boosted top:**
**→b and W-decays in different jets**



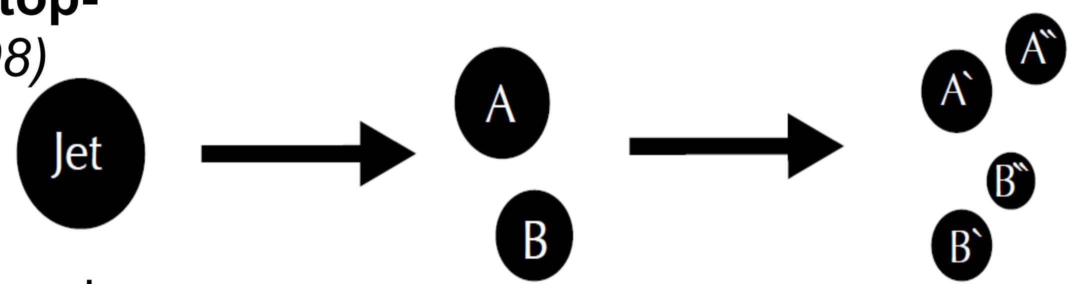

**Boosted top:**
**→decays in one jet**



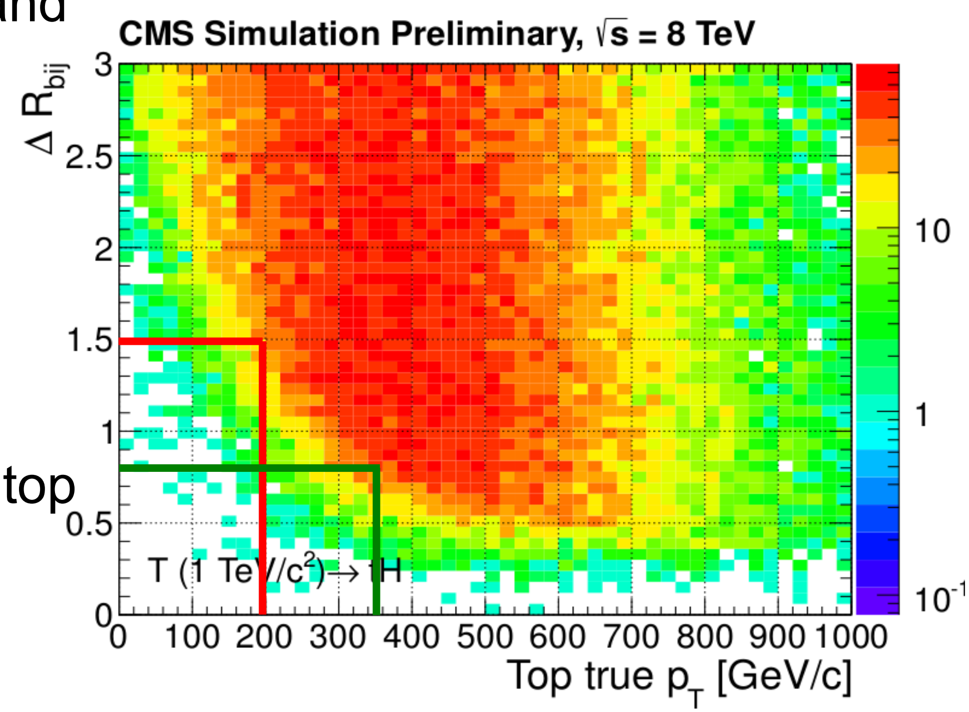

- **Resolved regime**:
  - → low-$p_T$ analyses (SM)
  - → multiple small jets (**AK5**) from decays of heavy particles
- **Boosted regime**:
  - → decay products merged in larger **fat-jets** (**CA8**, **CA15**)
  - → **substructure** tools to identify particle originating the jet

# Top-Tagging
## *[CMS-PAS-JME-10-013, CMS-PAS-JME-13-007]*

**CMSTopTagger**. Based on **JHU top-tagger** (*Kaplan et al [PRL 101 (2008) 142001]*):




- start with **CA8 jets**
- **reverse clustering sequence** and identify subjets
- requirements on jet/subjets masses and subjet multiplicty




**HEPTopTagger** (*Plehn et al., [arXiv:1006.2833]):*

- move to larger jets: **CA8→CA15**
- larger cone: sensitive to less boosted top jets, $p_T > 200$ GeV
- seamless transition between resolved and boosted regime

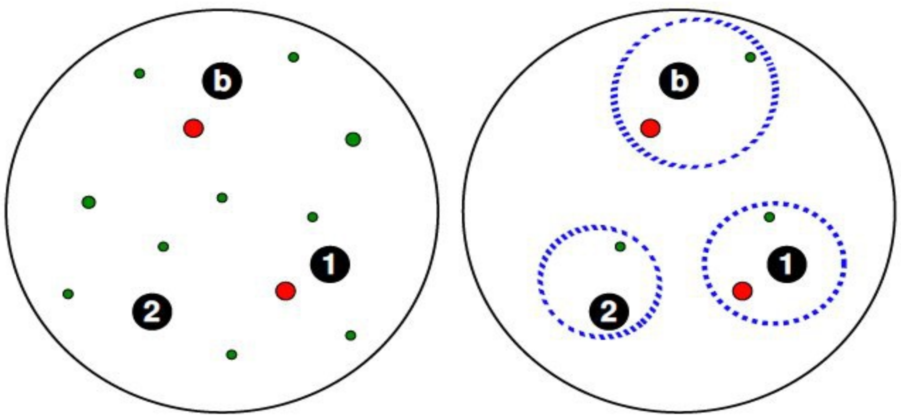
# HepTopTagger
## *[CMS-PAS-JME-13-007]*

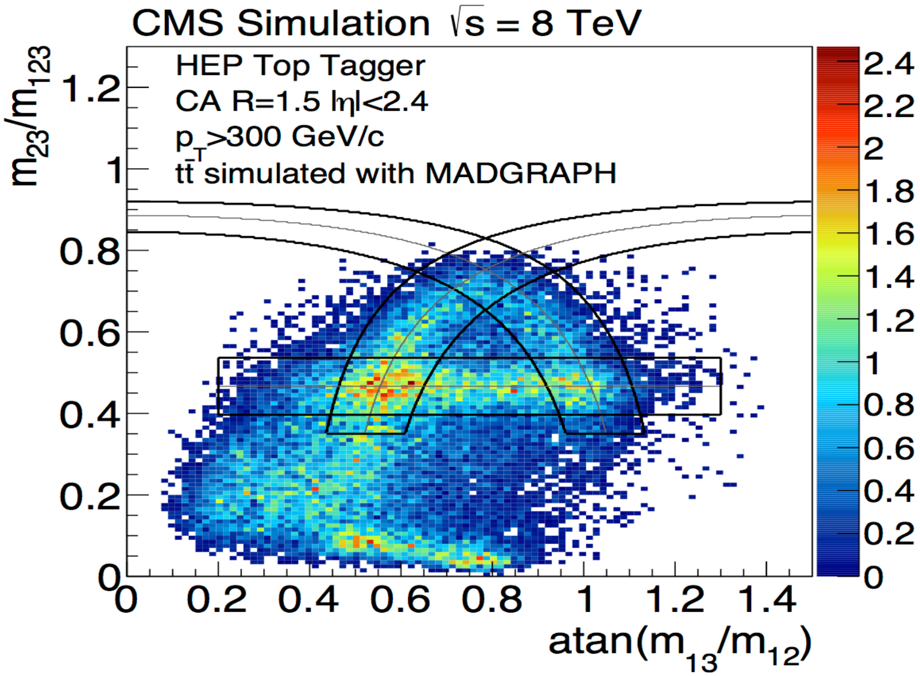
- First step is **cleaning** of the large fat-jet (mass-drop+**filtering**):
  - **soft-radiation** dilutes substructure of the fat-jet
  - large-jet: more from **contamination** from pile-up

- Built exactly **3 subjets** from remaining constituents, using smaller CA cone

- Use Dalitz distributions based on **subjet masses** to identify tops (select A region)

**top fat-jet cleaning**

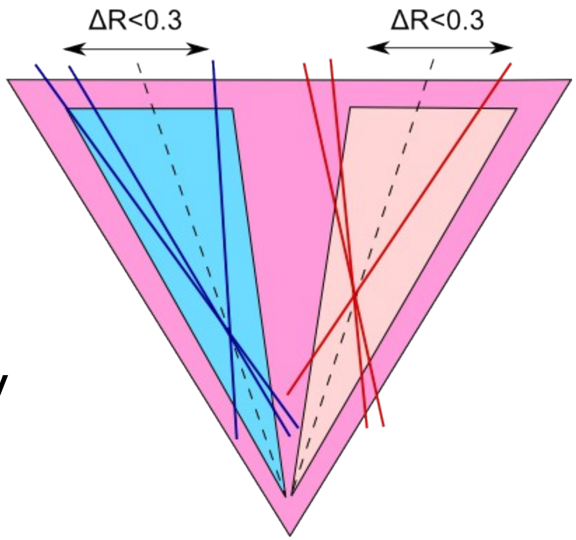


# Subjet b-tagging and Higgs-tagging

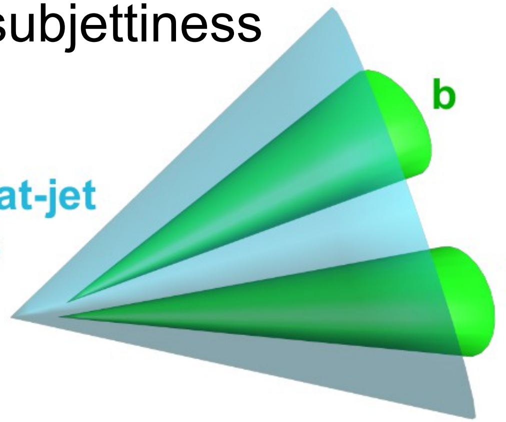
*[CMS-PAS-BTV-13-001, CMS-DP-2014/031]*

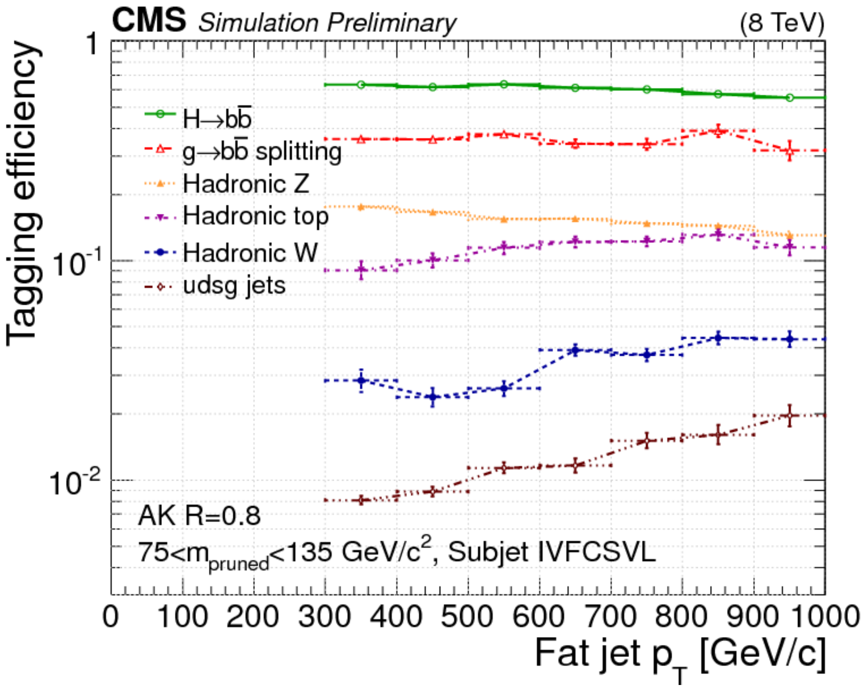
- **Combined Secondary Vertex** CSV tagger:
  - well established for isolated AK5 jets
  - likelihood-ratio combination of track displacement + secondary vertex (SV) information




- Subjet b-tagging: apply CSV to tracks and SV reconstructed within subjets

- **Higgs→bb tagger** based on 2 subjet b-tags:
  - eccellent performance
  - can be further improved by mass cut, n-subjettiness



# Additional Variables
## *[CMS-PAS-JME-13-007]*

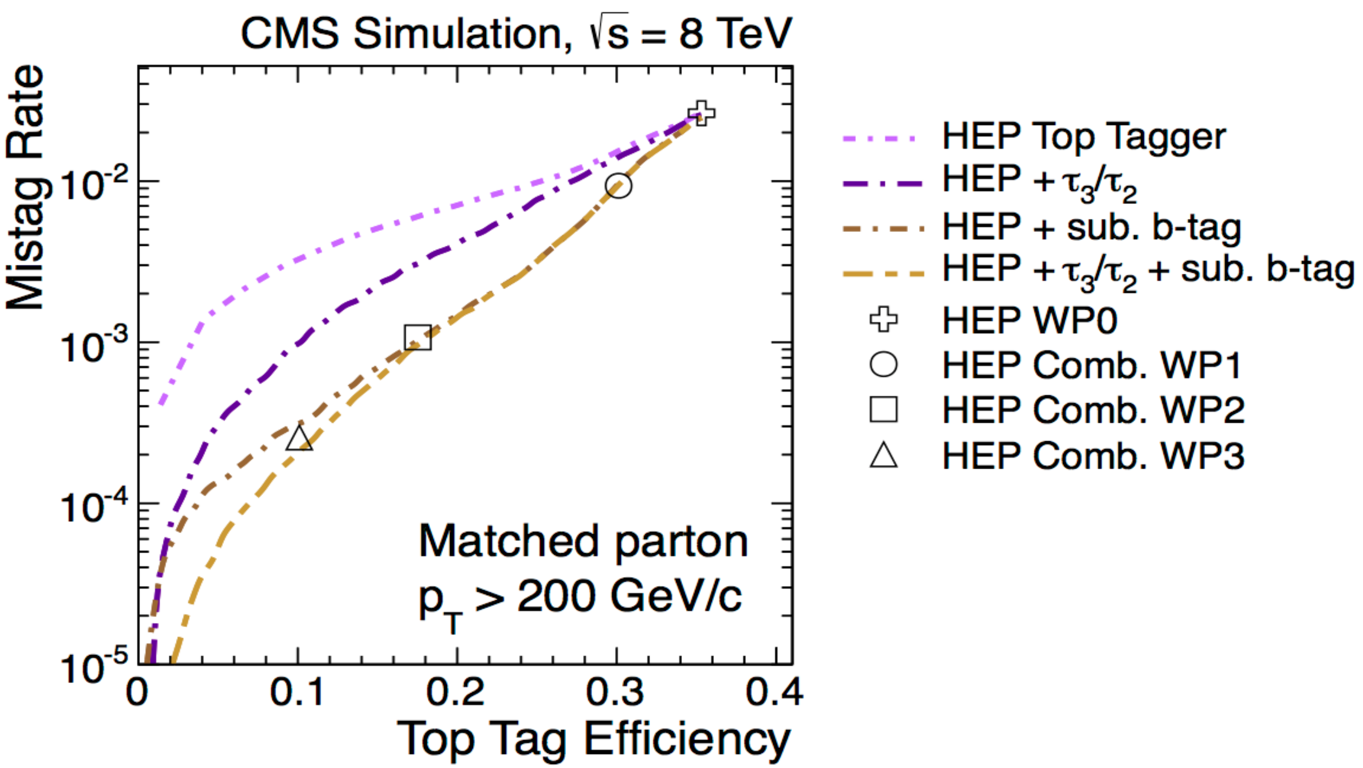
- **N-subjettiness $\tau_N$**: probability to have N-subjets

$$\tau_N = \frac{1}{d_0} \sum_k p_{T,k} \min\{\Delta R_{1,k}, \Delta R_{2,k}, \cdots, \Delta R_{N,k}\}$$

$d_0 = \sum_k p_{T,k} R_0$, and $R_0$ is the original jet radius

$\tau_2/\tau_1$: H/W/Z-tagging

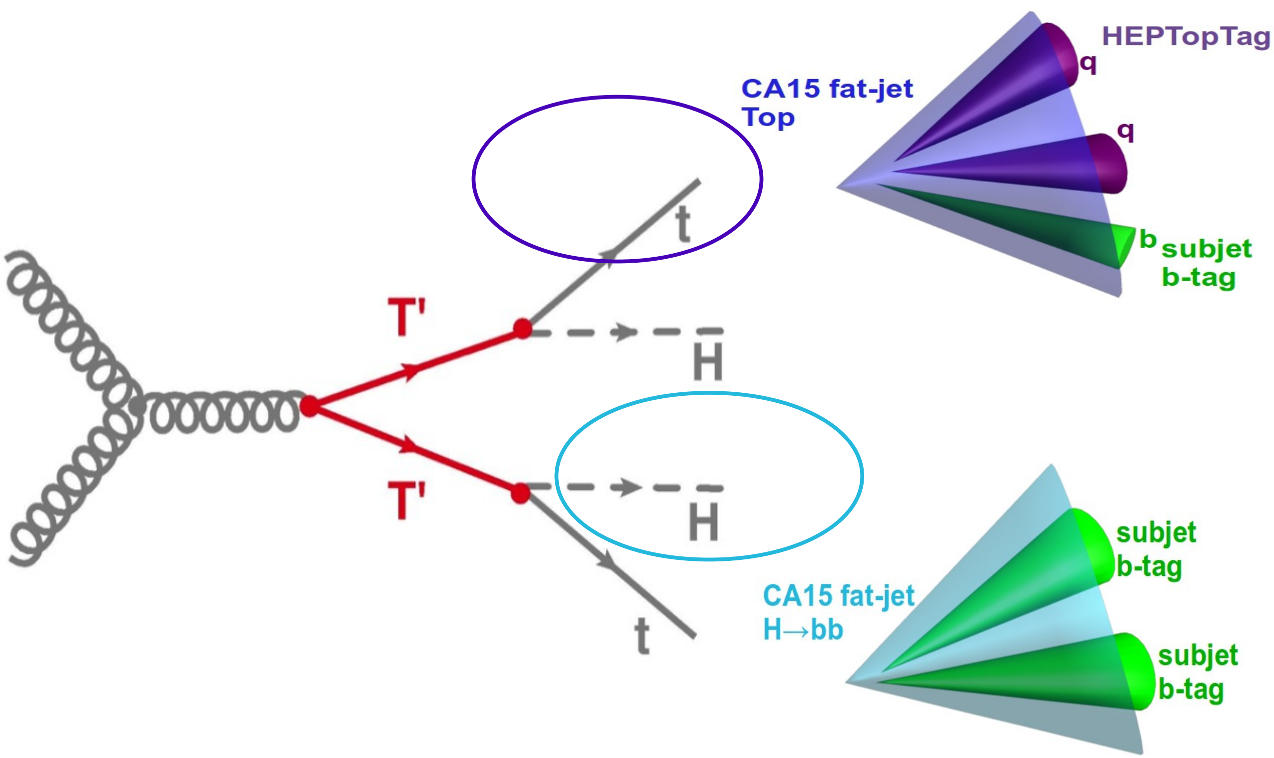
$\tau_3/\tau_2$: top-tagging




improved top-tagging performance with subjet b-tagging and $\tau_3/\tau_2$

# Fully Hadronic T→tH, H→bb

*[CMS-PAS-B2G-14-002]*




- Large $H_T$ region
- Use of **most advanced substructure**:
  - **HEPTopTagger**
  - **subjet b-tagging** (**first time**)
- Increase acceptance:
  - ≥ 1 top-tag and ≥ 1 Higgs-tag

**top-tag**:
HEPTopTagger
subjet b-tagging
$p_T$>200 GeV

**Higgs-tag, H→bb**:
CA15 filtered jet
2xsubjet b-tagging
mass > 60 GeV
$p_T$ > 150 GeV

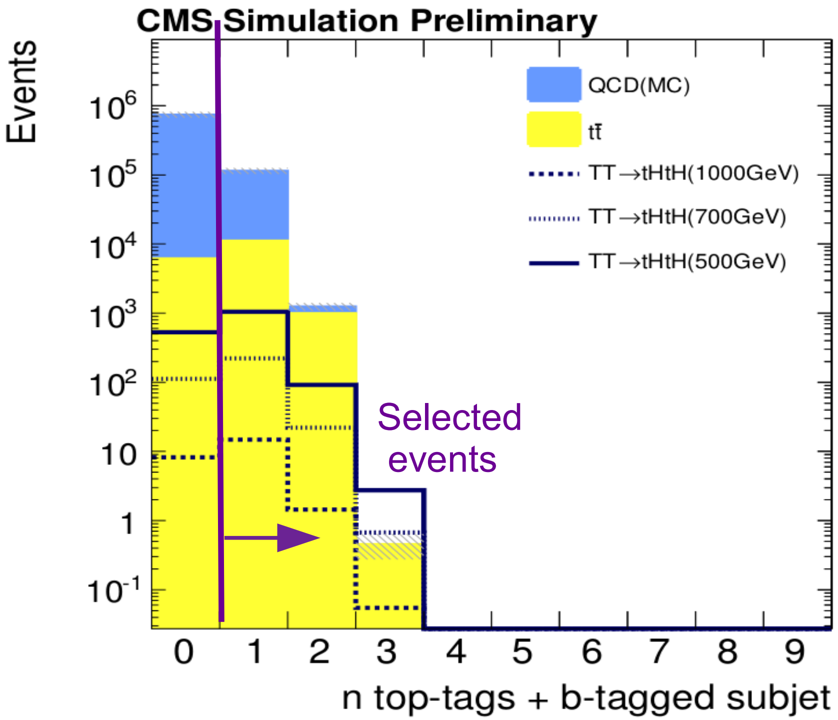
# Event Selection

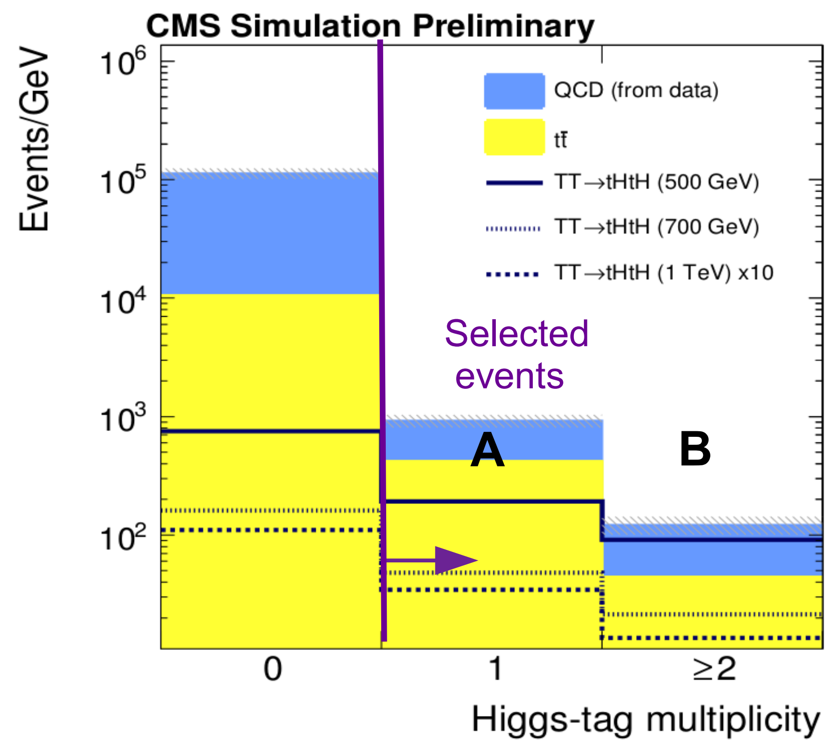
- Use of substructure: **QCD background** reduced at the level of **ttbar**

- Two categories for analysis:

**A**) single Higgs tag
**B**) multi Higgs tag

exploiting excellent signal/background ratio in multi Higgs-tag category



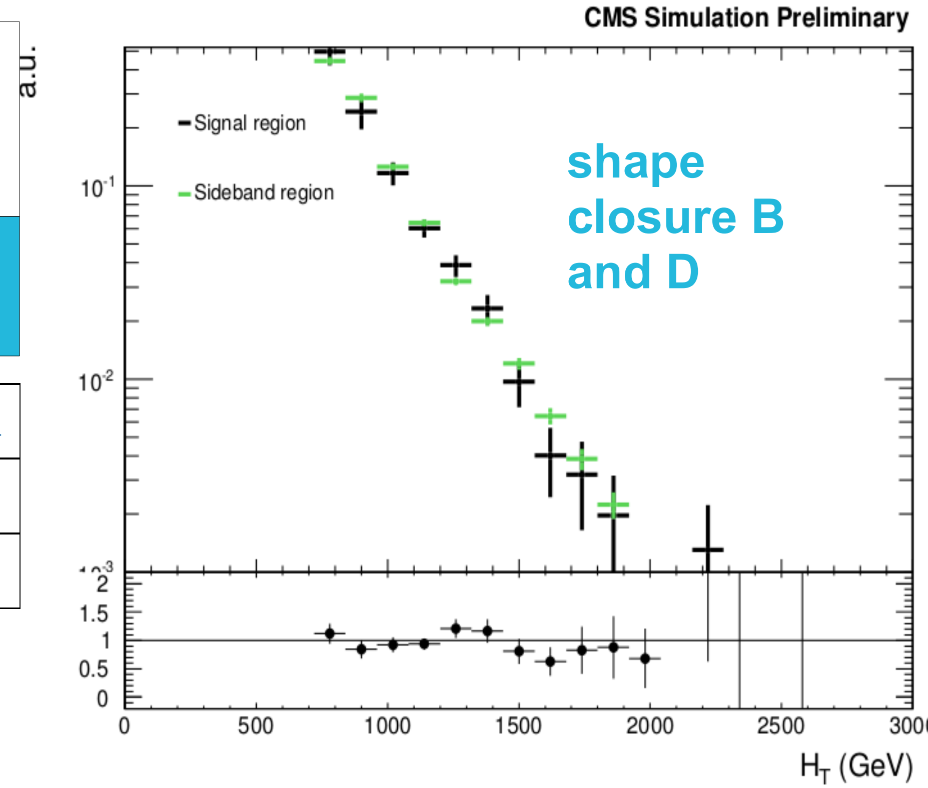



# Background Estimation

- ttbar background from MC. QCD contribution derived from control regions using **ABCD method**:
  - sidebands obtained **inverting substructure** selections
  - if inverted criteria are uncorrelated: rates **A/B=C/D**

| | |
|---|---|
| Region A:<br>Invert top & Higgs tag | Region B:<br>Invert top tag |
| Region C:<br>Invert Higgs tag | Region D:<br>Signal region |

| | Inclusive selection | Single Higgs-tag bin | Multi Higgs-tag bin |
|---|---|---|---|
| A/B | 164±4 | 185±5 | 1417±97 |
| C/D | 160± 13 | 185±17 | 1203 ± 250 |



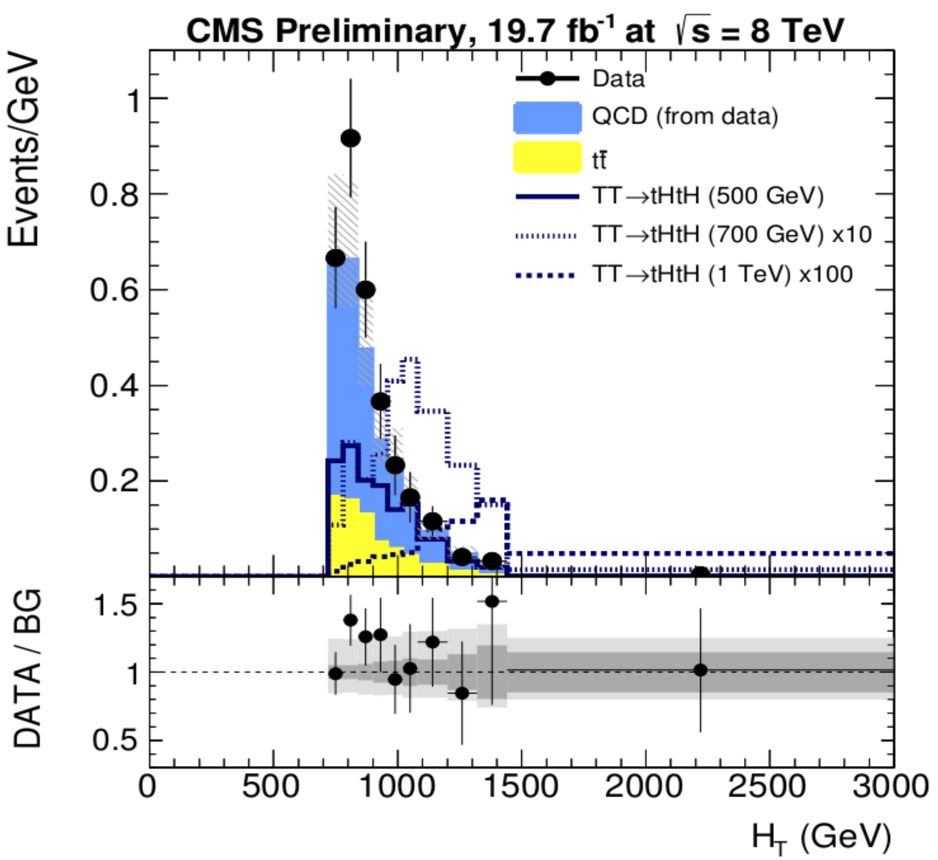

- **Signal region**:
  - rate D=C*B/A
  - shape from region B: **good closure** for all relevant variables in MC

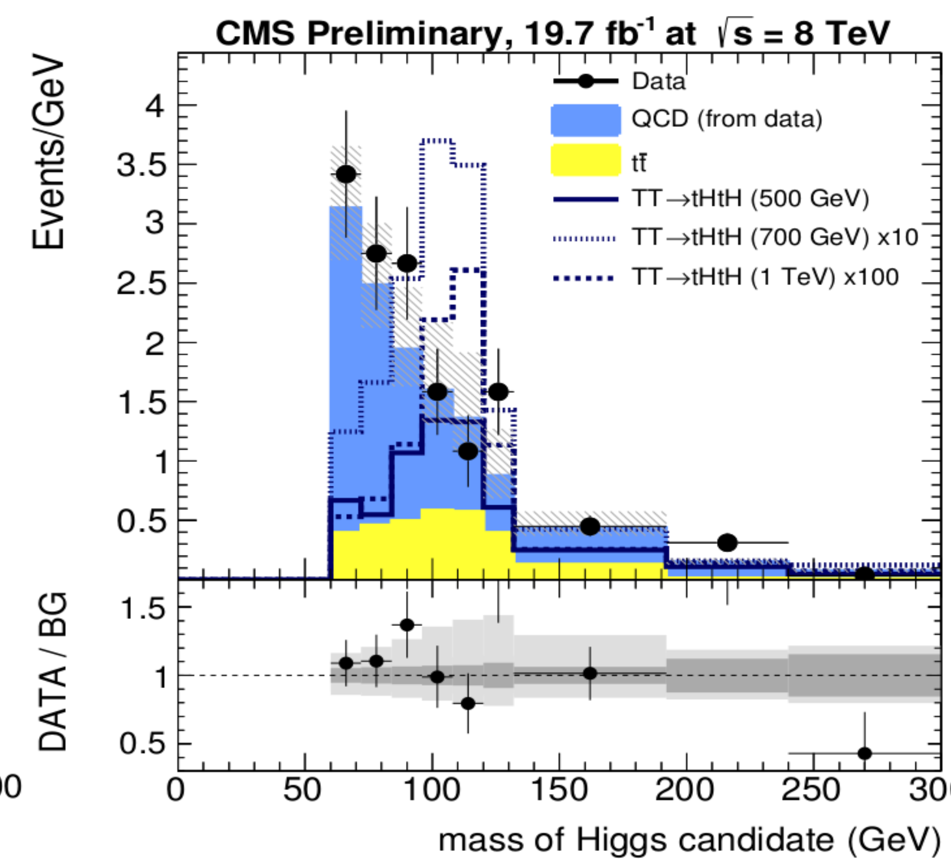
# Final Distributions

- Two variables used for limit setting, with visible shape difference signal/background (here for multi Higgs-tag category)

$H_T = \Sigma p_T$(subjets) of all CA15 jets with $p_T$ > 200 GeV

Invariant **mass** of b-tagged subjets in **Higgs-tagged CA15 jet**

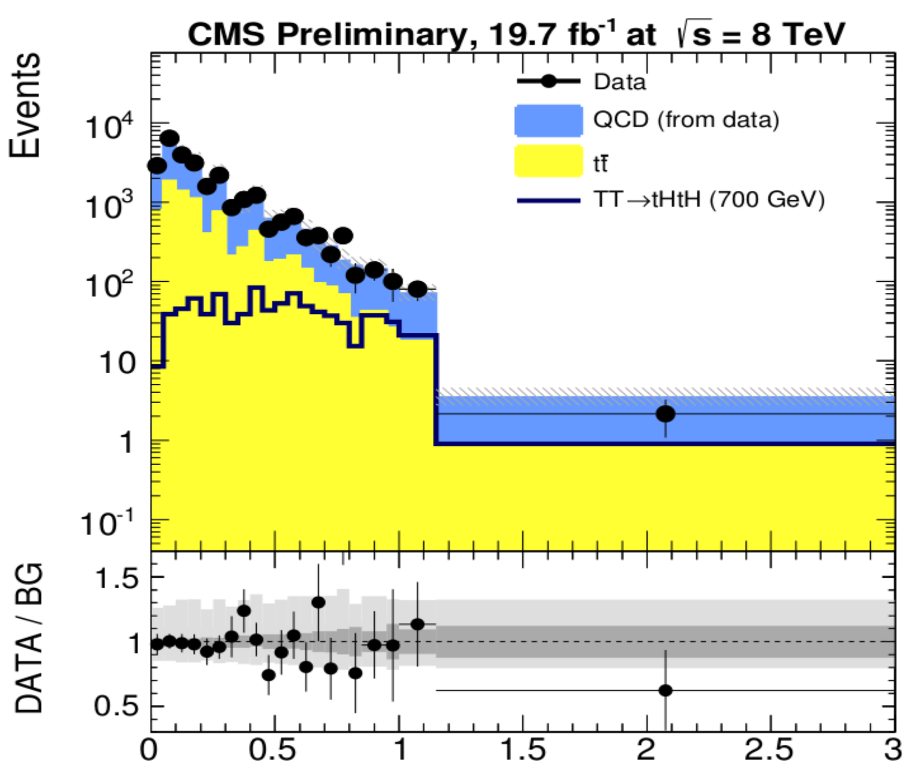


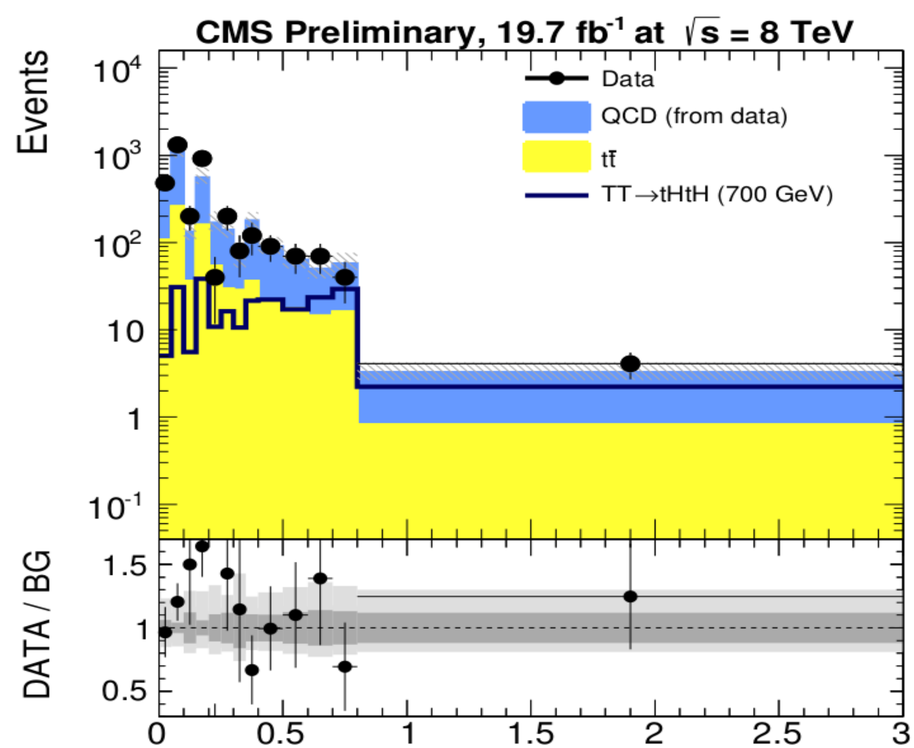
# Likelihood

- $H_T$ and Higgs-candidate mass observables **uncorrelated**: combined in likelihood

- Background at lower L values and signal at higher L values: good **discrimination power**



single Higgs-tag

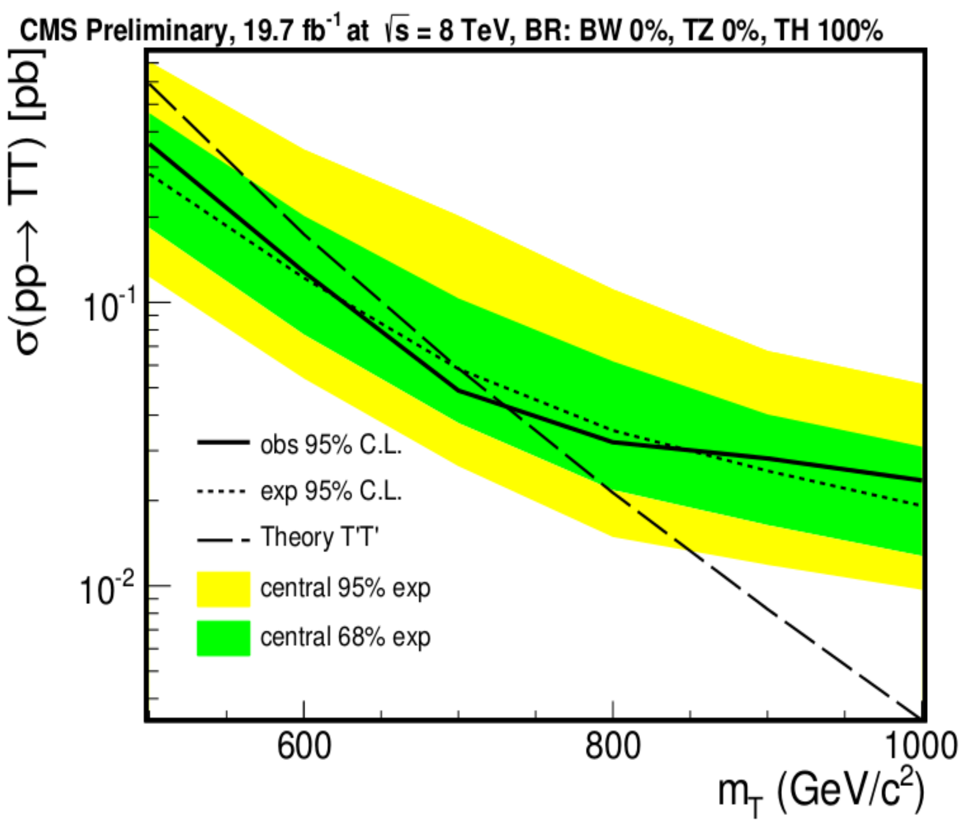


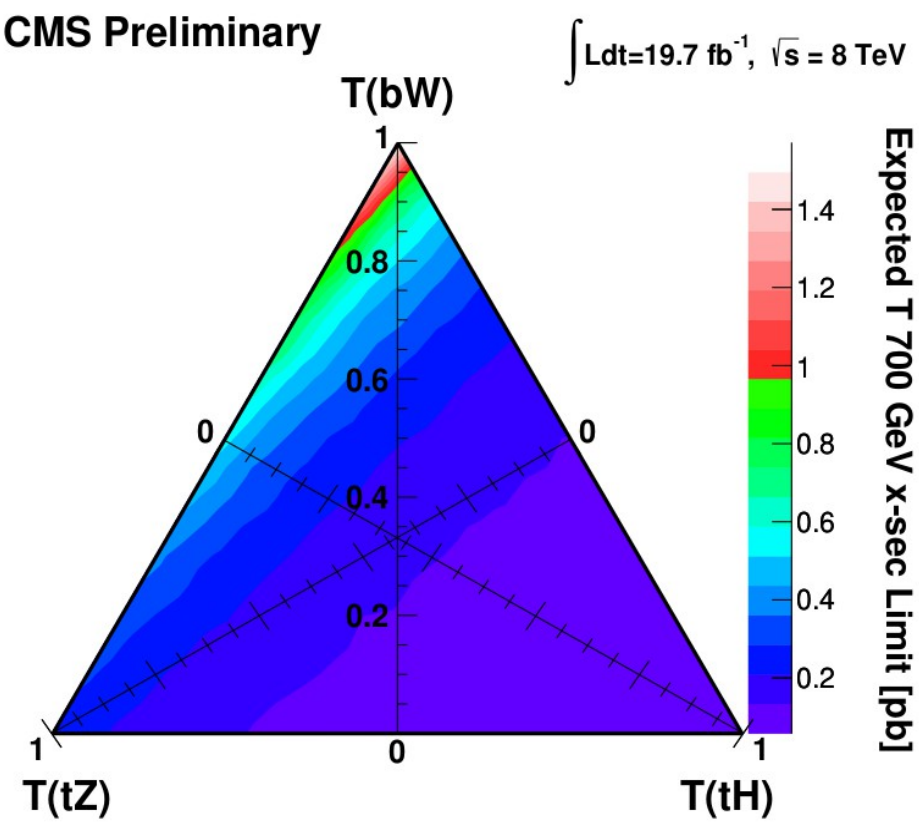
multi Higgs-tag

# Limits

- Search targeting at **T→tH** decay: (expected) observed limit (**701**) **747** GeV

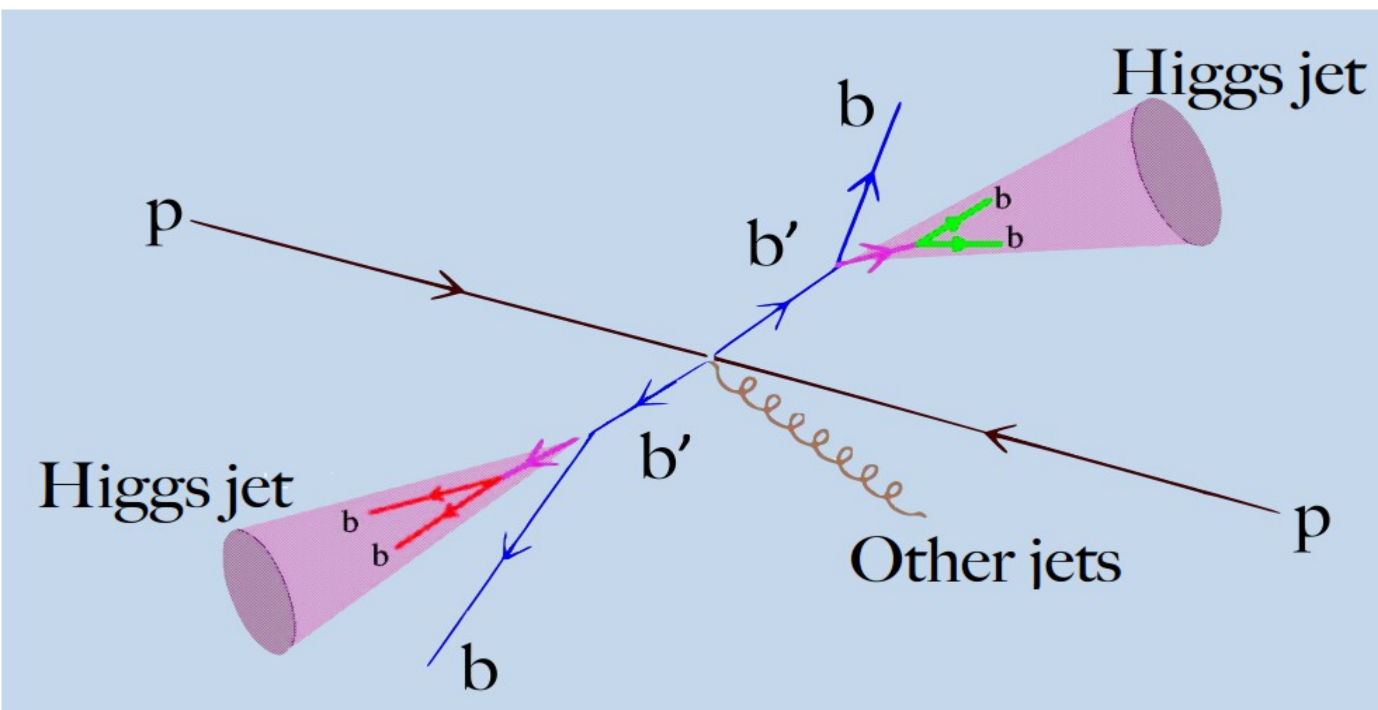
- Cross-section limits provided for **all possible BR** to three possible decays: **bW, tZ, tH**



# B→bH, H→bb
## *[CMS-PAS-B2G-14-001]*

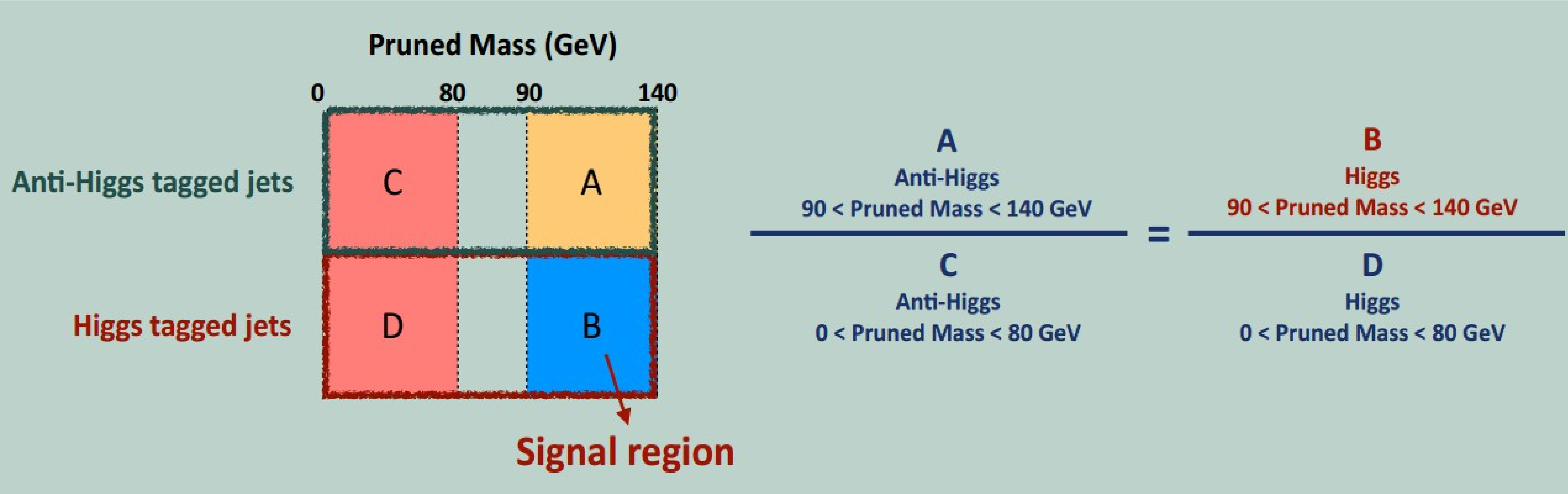
- **Higgs-tagger** here given by:
  - **pruned CA8** jet $p_T$ > 300 GeV
  - 2 **subjet b-tags**
  - N-subjettines $\tau_2/\tau_1$ < 0.5
  - 90 < **mass pruned** < 140 GeV




- Large $H_T$ region
- ≥ 1 Higgs-tag
- ≥ 1 regular b-tagged AK5 jet
- **Event categories**: 1 AK5 b-tag, ≥ 2 AK5 b-tags

# Background Estimation

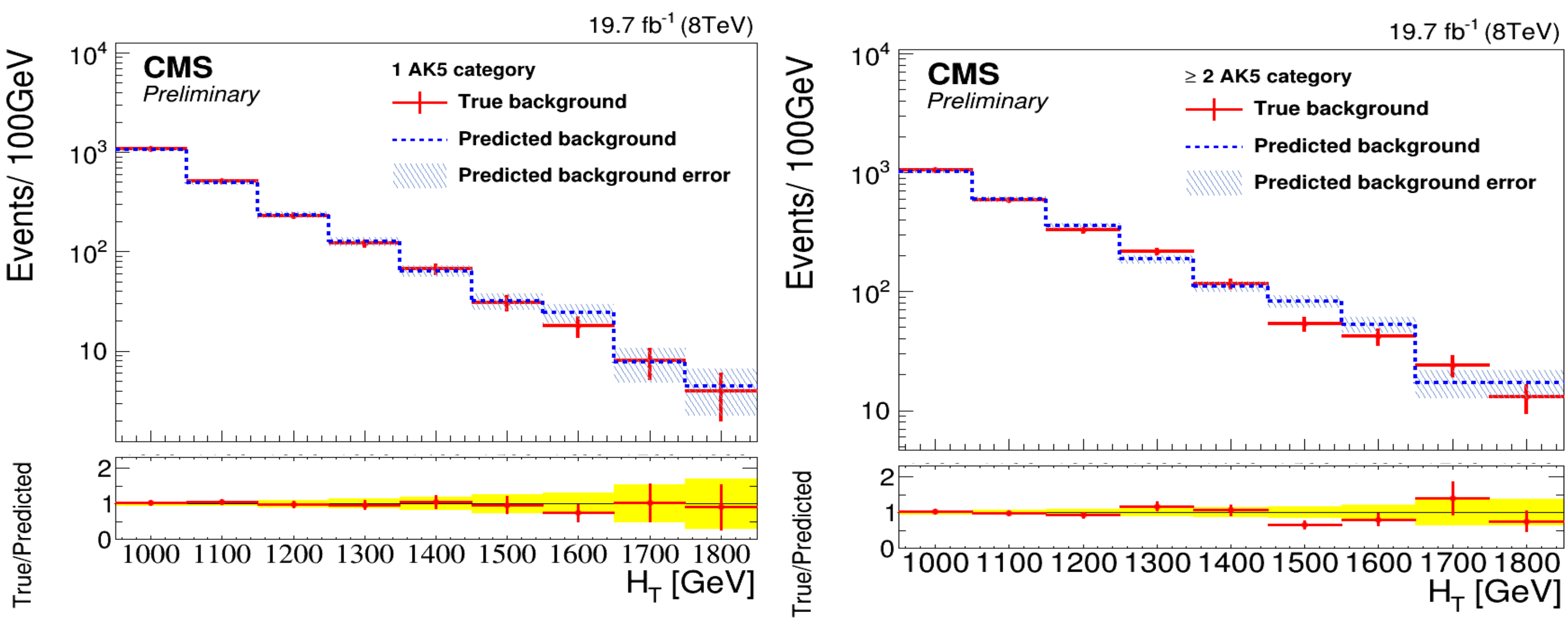
- ttbar background from MC. QCD contribution evaluated using **ABCD method**
- Control region from inverted substructure requirements




- **Closure test** of the method done in MC and also in data, using a **signal-free control sample**: standard selection, but no b-tagged AK5 jet

# Closure Test in Data

- Performed in **two event categories**:
  - ==1 AK5 jet with $p_T$ > 80 GeV
  - ≥2 AK5 jets with $p_T$ > 80 GeV
- Good agreement between true and predicted background by ABCD method for the observable used for limit setting **$H_T = \sum p_T$ (AK5 jets)**



# Event Selection

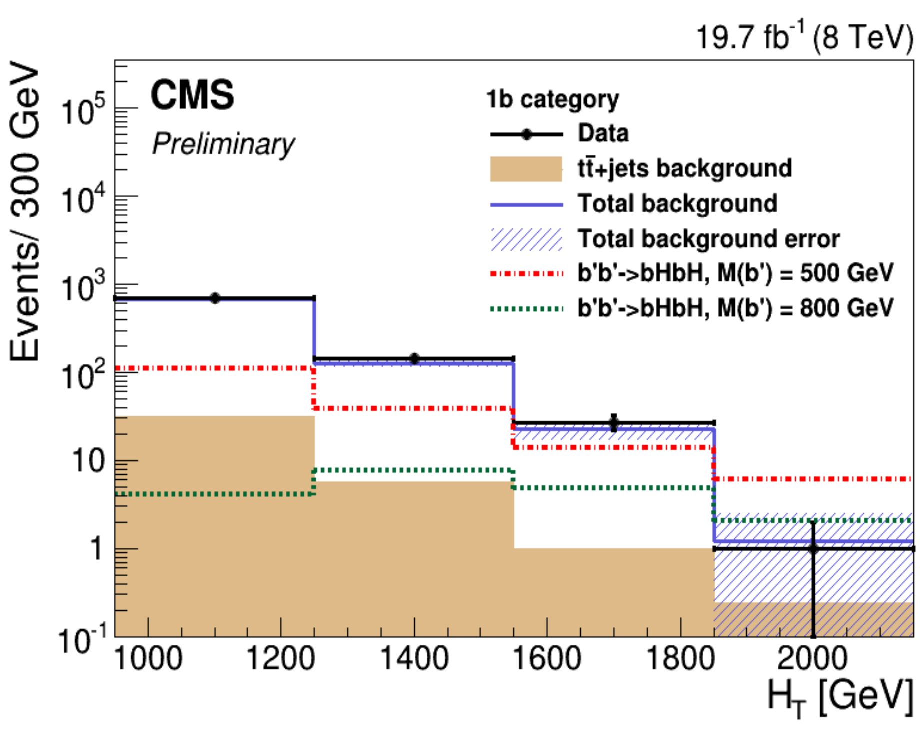
- Good signal selection efficiency and background rejection

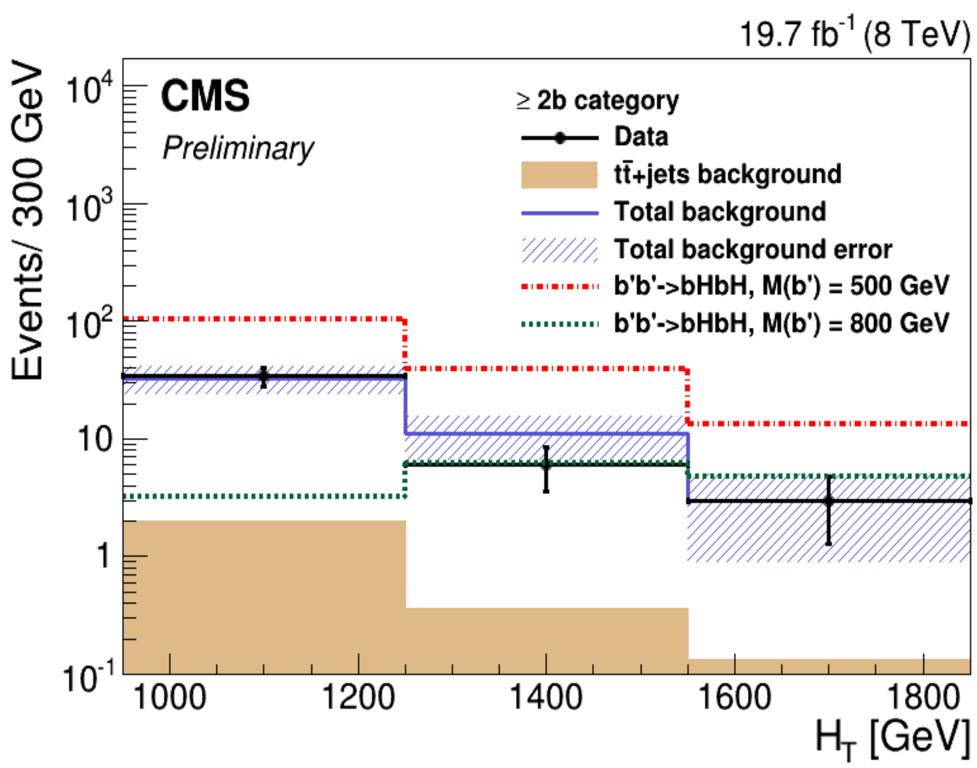
| Signal: $b'\bar{b}' \to bHbH$ | | | | |
|---|---|---|---|---|
| b' quark mass (GeV) | Yields after full selection | Yields in 1 b-tagged category | Yields in $\geq$ 2 b-tagged category | Selection efficiency |
| 500 | 326.9 | 169.2 | 157.7 | 0.037 |
| 600 | 170.6 | 90.2 | 80.4 | 0.055 |
| 700 | 77.6 | 42.1 | 35.5 | 0.070 |
| 800 | 32.8 | 18.5 | 14.3 | 0.079 |
| 900 | 13.4 | 8.0 | 5.5 | 0.082 |
| 1000 | 5.3 | 3.3 | 2.0 | 0.082 |
| 1200 | 0.9 | 0.6 | 0.3 | 0.078 |

| Process | Yields after full selection | Yields in 1 b-tagged category | Yields in $\geq$ 2 b-tagged category |
|---|---|---|---|
| Data-driven background | $872^{+49}_{-55}$ | $825^{+47}_{-52}$ | $46^{+4}_{-11}$ |
| Data | 903 | 860 | 43 |

# $H_T$ Observable

- $H_T$ observable in two b-tag categories used for limit setting
- Very good signal/background ratio

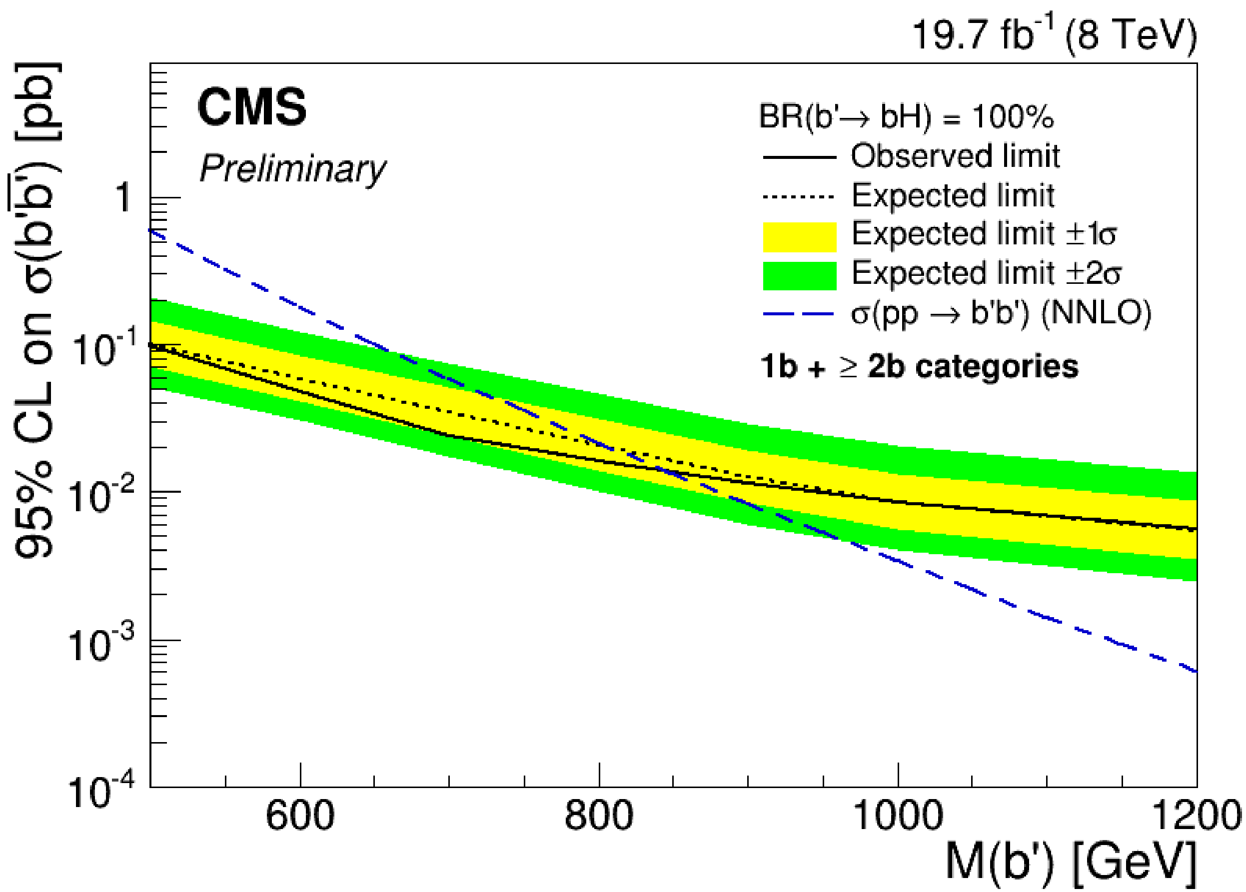


# Limits

- Search targeting at **B→bH** decay: (expected) observed limit (**846**) **811** GeV

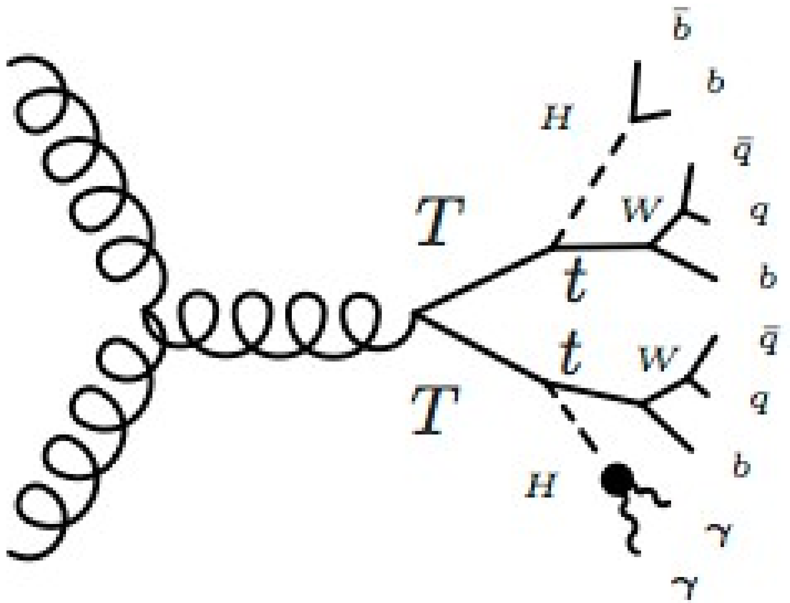
- Thanks to substructure hadronic analyses obtain limits **competitive with searches in leptonic final states**



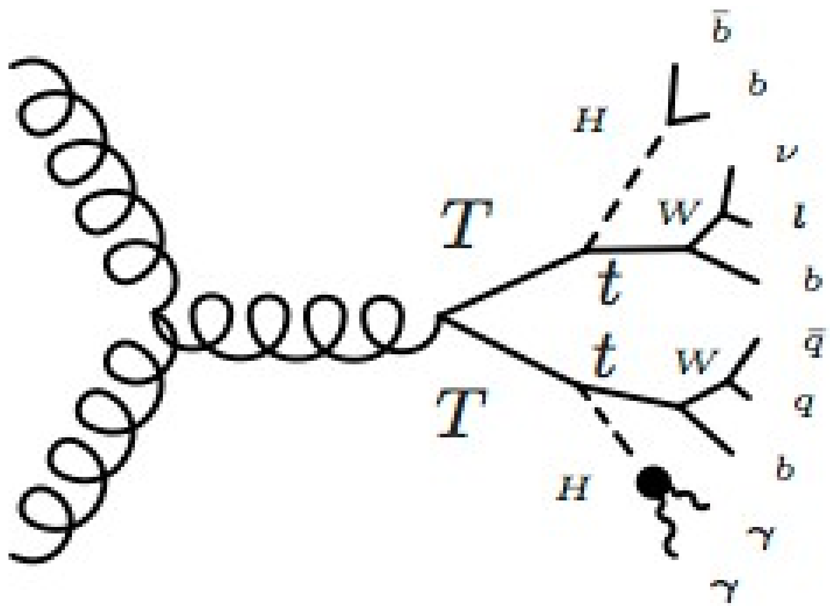
# A search with photons

# T→tH, H→γγ
*[B2G-14-003]*

- Exploit **narrow H→γγ resonances** (as done in H→γγ analysis)
- Both hadronic and leptonic final states to improve sensitivity:
  - hadronic: higher cross-section
  - letponic: better signal/background




**hadronic final state**



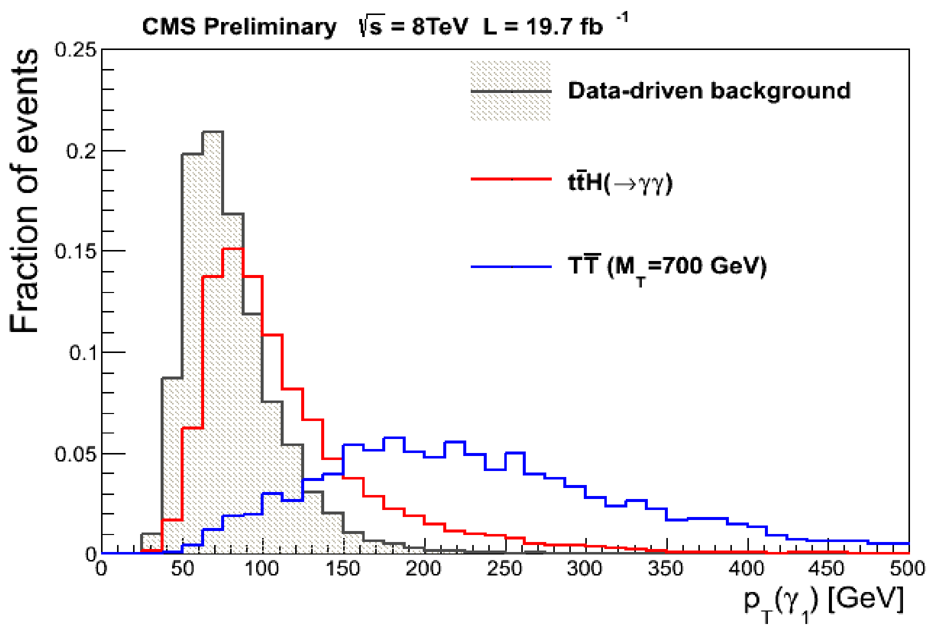

**single lepton final state**

# Analysis Optimization

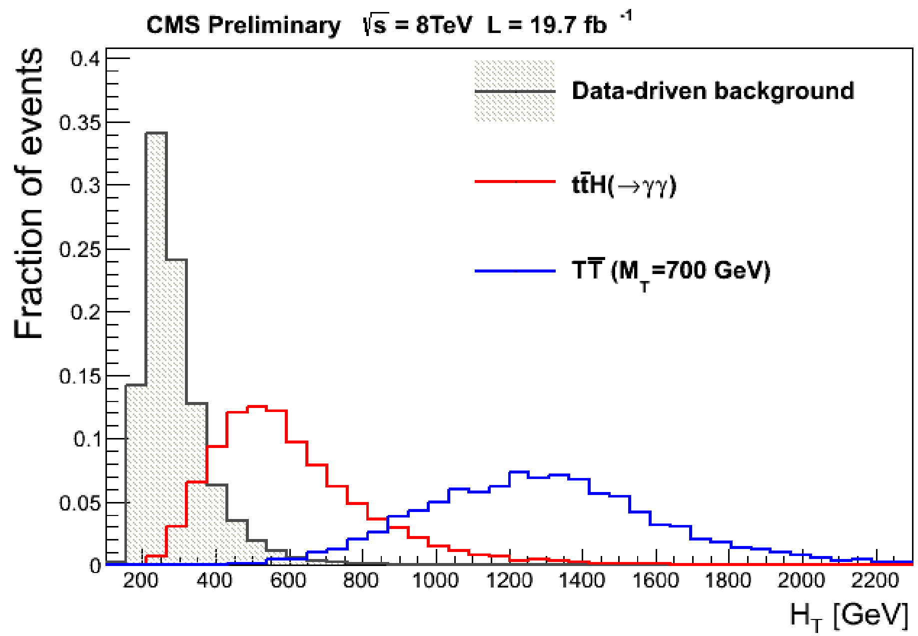
- Main backgrounds:
  - **resonant background ttH**: from MC
  - non-resonant background, from **diphoton + jets**; tt+γγ, t+γγ, tt+jets
- Non-resonant background MC not reliable:
  - for optimization studies: use **control sample CS**
  - for limit setting: sideband regions in $M_{\gamma\gamma}$ distribution (later)
- Control sample:
  - obtained **inverting photon identification**
  - kinematics re-weighted to match signal region
- Selection optimization:
  - based on CS and signal with a 700 GeV mass
  - separate from hadronic/leptonic channels

# Discriminating Variables

- Several discriminating variables considered
- Selection cuts optimized through MVA




$p_T$ leading photon

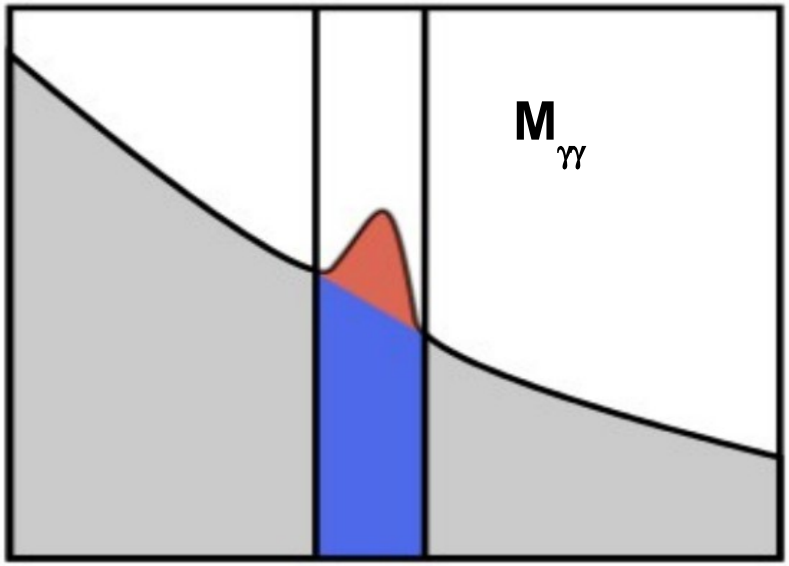



$H_T = p_T$ jets + $p_T$ photons + $p_T$ leptons

# Final Selection

| Variable | Hadronic channel | Leptonic channel |
|---|---|---|
| $p_T(\gamma_1)$ | $> \frac{3}{4} m_{\gamma\gamma}$ GeV | $> \frac{1}{2} m_{\gamma\gamma}$ GeV |
| $p_T(\gamma_2)$ | 35 GeV | 25 GeV |
| $n_{jets}$ | $\geq 2$ | $\geq 2$ |
| $H_T$ | $\geq 1000$ GeV | $\geq 770$ GeV |
| leptons | 0 | $\geq 1$ |
| b tags | $\geq 1$ | - |

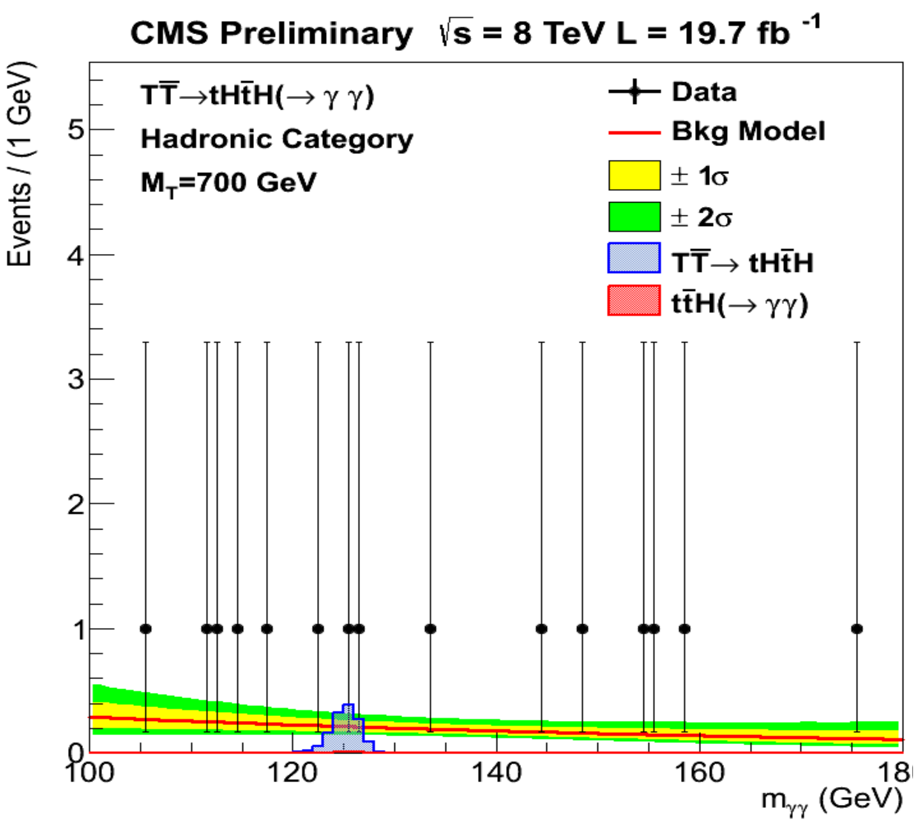
**selection cuts from MVA optimization**

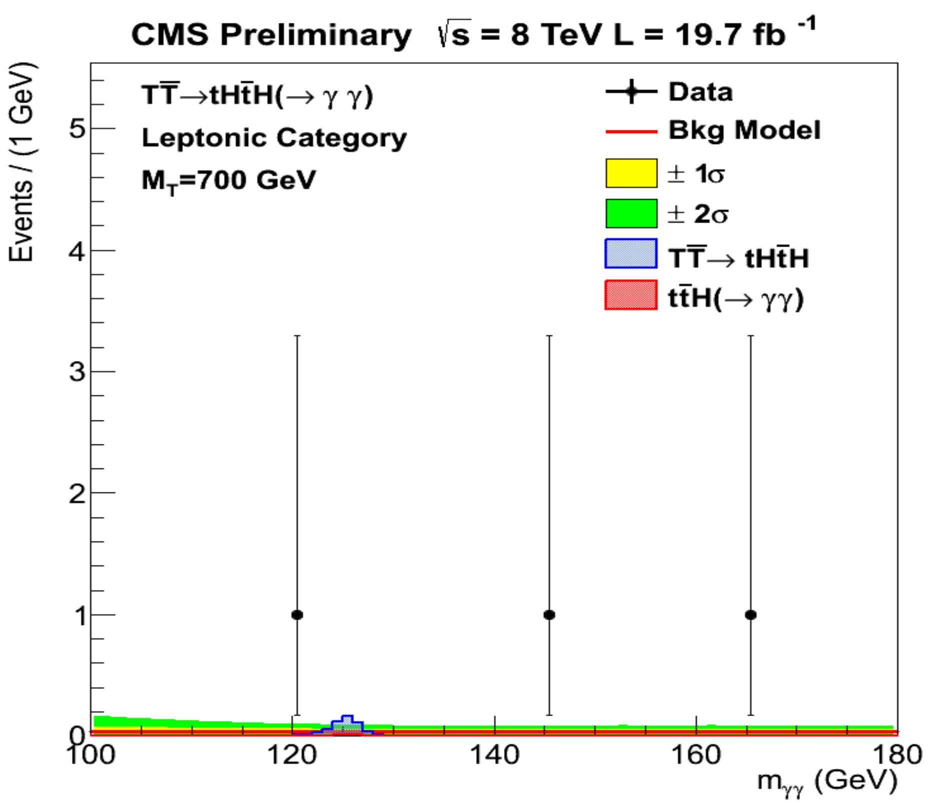



- **Background evaluation**:
  - signal gives a peak in $M_{\gamma\gamma}$ distribution
  - fit background in **sideband regions outside $M_{\gamma\gamma}$ peak** (125±3 GeV) for $M_{\gamma\gamma}$ between [100,180] GeV
  - different fit functions tested: **simple exponential** chosen

# Final Observables

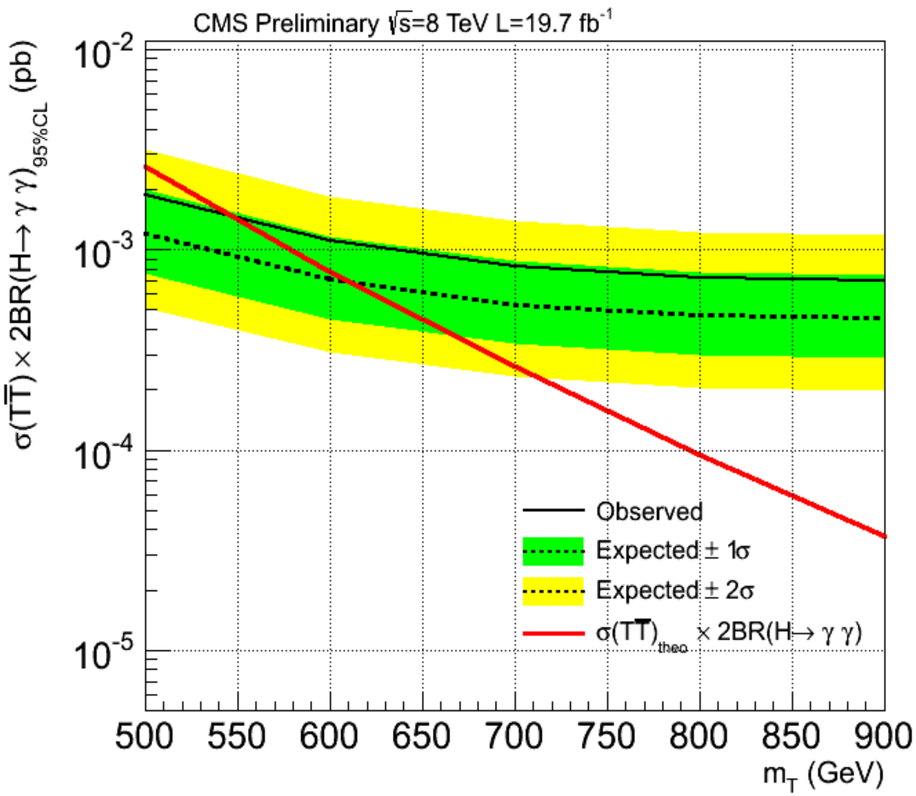
$M_{\gamma\gamma}$ in hadronic and leptonic categories, with fitted background

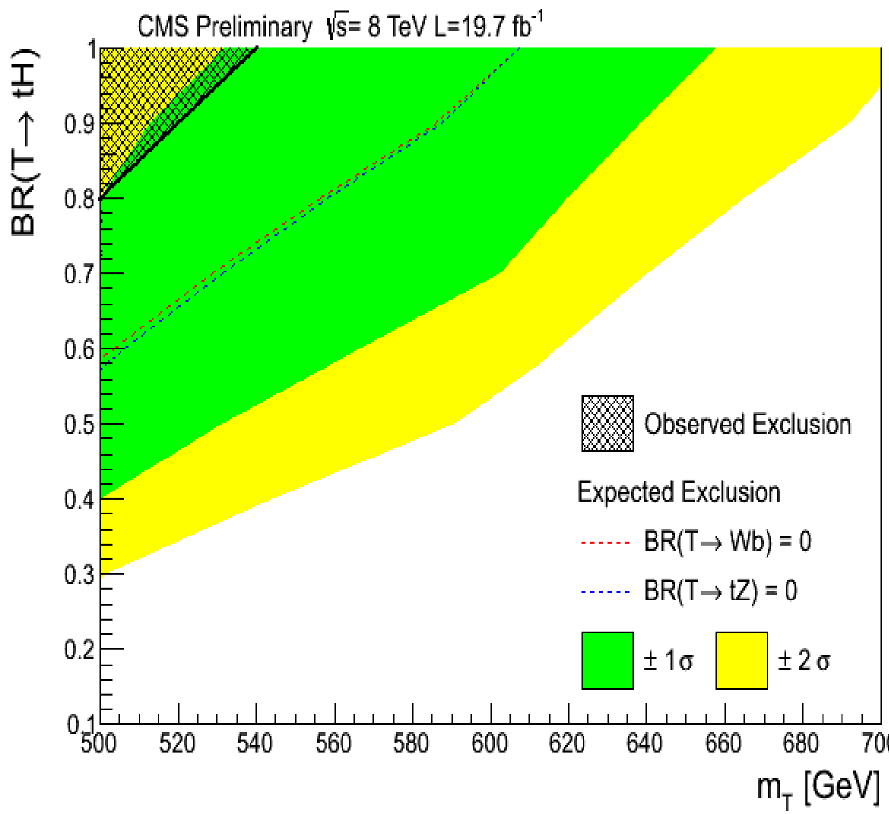


# Limits

- Search targeting at **T→tH** decay: (expected) observed limit (**607**) **540** GeV
- Limits provided for different **BR** to tH



# Conclusions

- **Development and use of substructure** in hadronic boosted analyses payed off:
  - limits competitive with those obtained by searches in clean leptonic final states

- Weaker limits obtained by the search with photons:
  - thinking in discovery mode:
    - **explicit Higgs identification** crucial to distinguish between theories proposing similar final states
    - ideal final state to **measure the key parameters** of the theory

- **Several other VLQ** searches at CMS not shown here:
  - more in talks from Gerrit and Sadia
  - **combinations** of searches underway, results expected soon

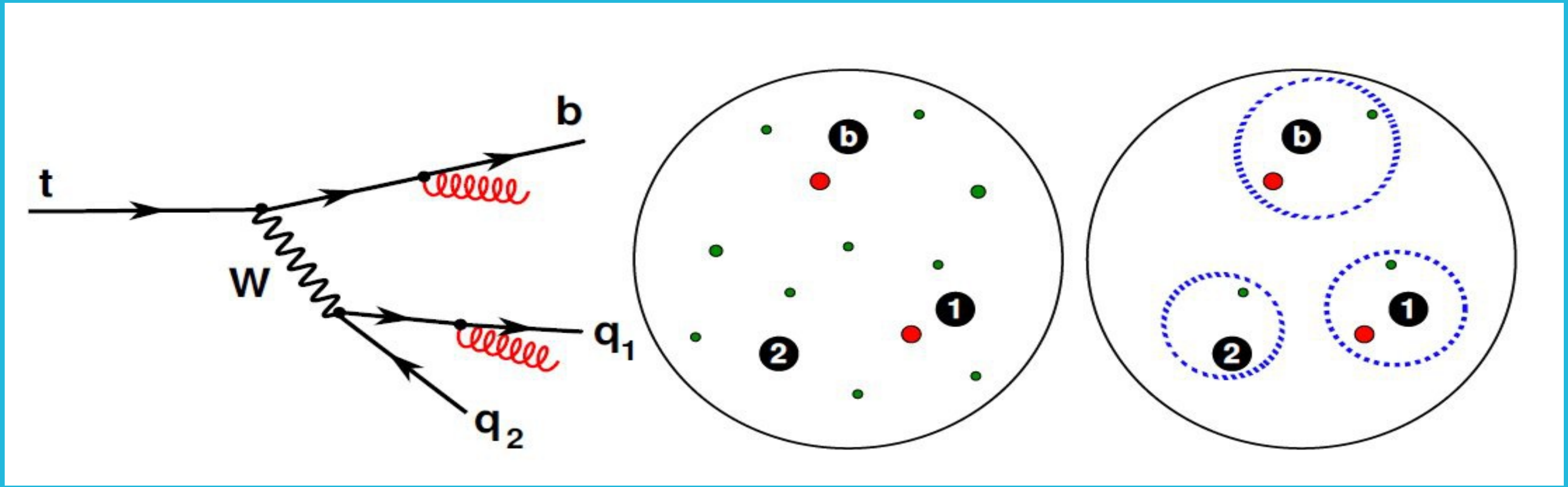
- Results shown for pair-produced VLQ: work on the **single production** searches has started

# Additional Material

# HepTopTagger
## *[CMS-PAS-JME-13-007]*

- First step is **cleaning** of the large fat-jet (mass-drop+**filtering**):
  - **soft-radiation** dilutes substructure of the fat-jet
  - large-jet: more from **contamination** from pile-up
- Built exactly **3 subjets** from remaining constituents, using smaller CA cone

**top fat-jet cleaning**

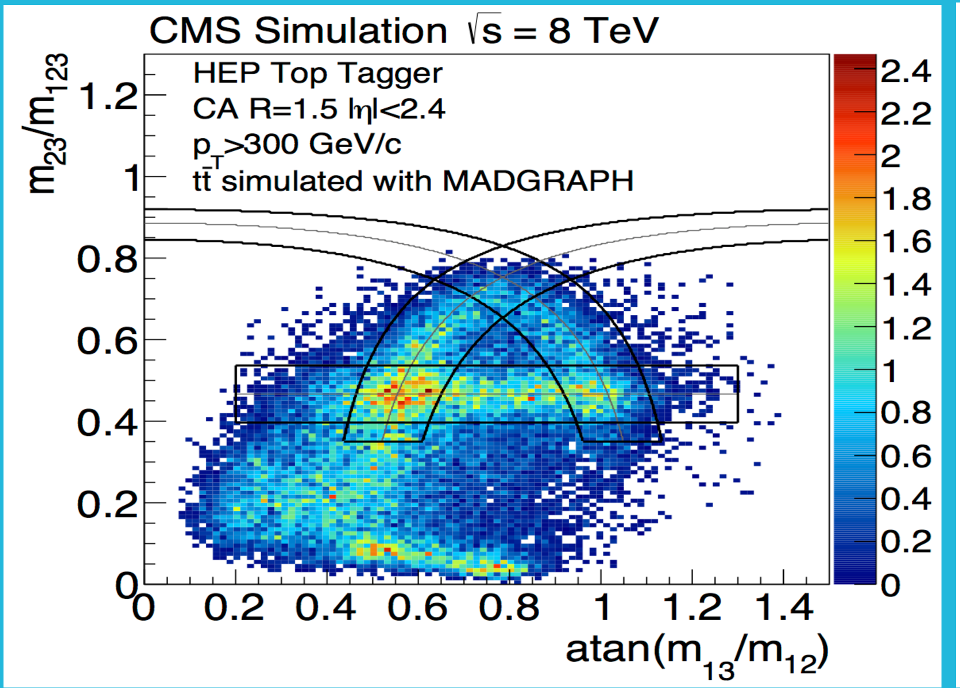


# HepTopTagger

- Use Dalitz distributions based on **subjets masses** to identify tops:
  - x axis: arctan $(m_{13}/m_{12})$
  - y axis: $m_{23}/m_{jet}$
  - $m_{ij}$ = mass *i*-subjet + *j*-mass subjet

- **Select A region**, where signal concentrates

**ttbar**




**background**