

Study of Higgs bosons decaying to bottom quarks at CMS

Sara Alderweireldt
for the CMS collaboration

Universiteit Antwerpen

SUSY 2015
Lake Tahoe, California

23-29 August 2015



Part 1: SM searches with $H \rightarrow b\bar{b}$

VH

ttH

VBF

Part 2: BSM searches with $H \rightarrow b\bar{b}$

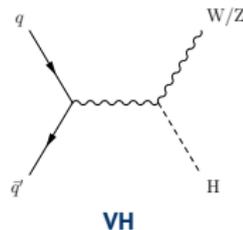
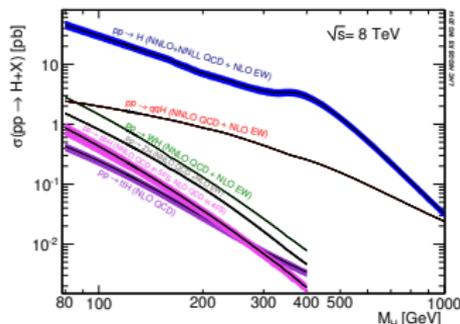
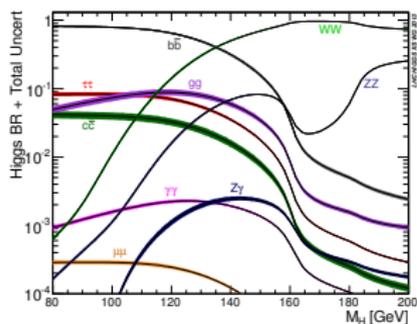
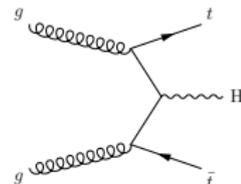
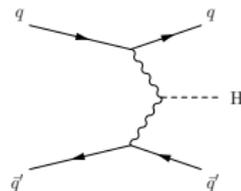
tHq

$A \rightarrow Zh$

ϕ

h_1

- $H \rightarrow b\bar{b}$ has the **largest branching fraction @ 125 GeV: 58%**
- **Large QCD-induced $b\bar{b}$ backgrounds** (overwhelming for ggH)
 - **VH and ttH:** (semi)leptonic decay channels
→ lower cross section, but pure QCD background is not dominant
 - **VBF:** $2q2b$ final state
→ higher cross section, but QCD background irreducible
→ approach with dedicated triggers
- Perform **combination of all $H \rightarrow b\bar{b}$ channels**


VH

ttH

VBF

Associated production with vector bosons

1. Analysis strategy
2. Multivariate discrimination
3. Results

VH

ttH

VBF

CMS-HIG-13-012

PRD 89 (2014) 012003

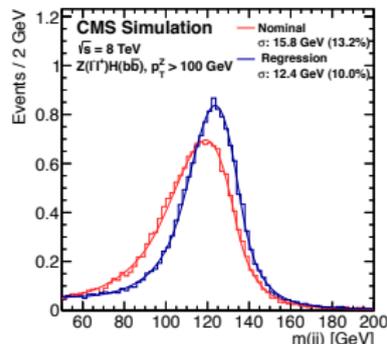
- ▶ Search in 6 decay channels: $W(e,\mu,\tau_h)H$, $Z(ee/\mu\mu)H$, $Z(\nu\nu)H$
- ▶ Backgrounds: W/Z + jets, $t\bar{t}$, diboson, QCD

Analysis strategy

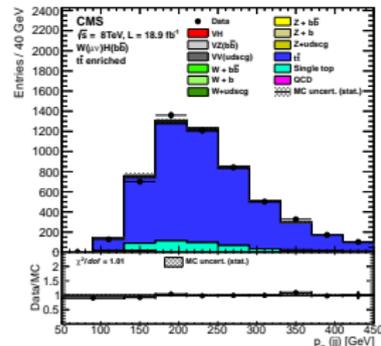
- Set **b-tag requirements** to identify $H \rightarrow b\bar{b}$
- Classify events in $p_{\perp}(V)$ **boost categories** to exploit the VH recoil
- Build a **multivariate discriminant** using the reconstructed $b\bar{b}$ invariant mass, b-tag values, kinematic variables, ...
- **Fit the discriminant shape** and extract limits

Optimization

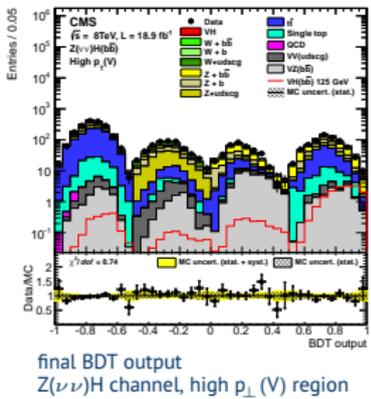
- Multivariate regression to derive jet energy correction on top of the average CMS one $\rightarrow b\bar{b}$ invariant mass resolution improves $\sim 15\%$
- Background calibration in control regions:
 - $\rightarrow t\bar{t}$, W/Z +jets (light flavour, b or $b\bar{b}$ enriched)
 - \rightarrow shape and normalization in signal region better constrained



$Z(l(l)H(bb))$ dijet invariant mass distribution before and after energy correction.

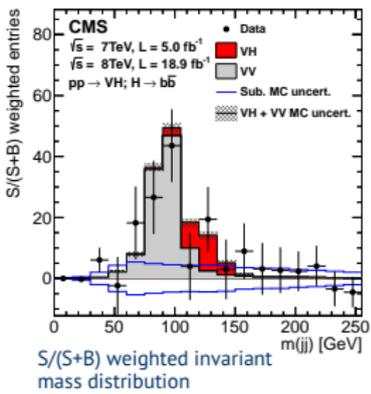


$W(l\nu)$ channel p_{\perp} distribution in the $t\bar{t}$ control region.



Multivariate discrimination

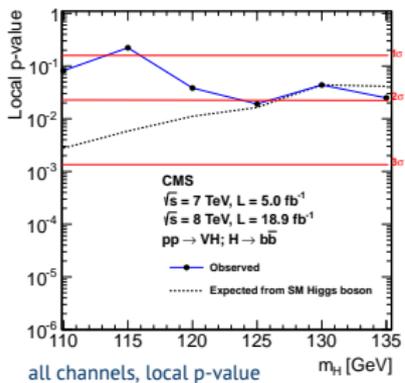
- Several **background-specific BDTs** are trained to separate individual backgrounds from signal ($t\bar{t}$, V+jets and VV).
- **Events are classified** according to these specific BDTs before being run through the final BDT (per mass and per channel).
- A **fit** is performed to the shape of the final BDT output distribution, composed of four subsets of events (simultaneously in all channels).



Results

update including ggZH contribution

- An excess is observed, compatible with the SM expectation.
- observed (expected) limit 1.68 (0.85) at $m_H = 125$ GeV
- significance 2.1σ (2.5σ)
- best-fitted $\mu = 0.89 \pm 0.43$



$t\bar{t}$ associated production (1)

using BDTs

1. Analysis strategy
2. Multivariate discrimination
3. Results

VH

ttH

VBF

CMS-HIG-13-029
JHEP 09 (2014) 087

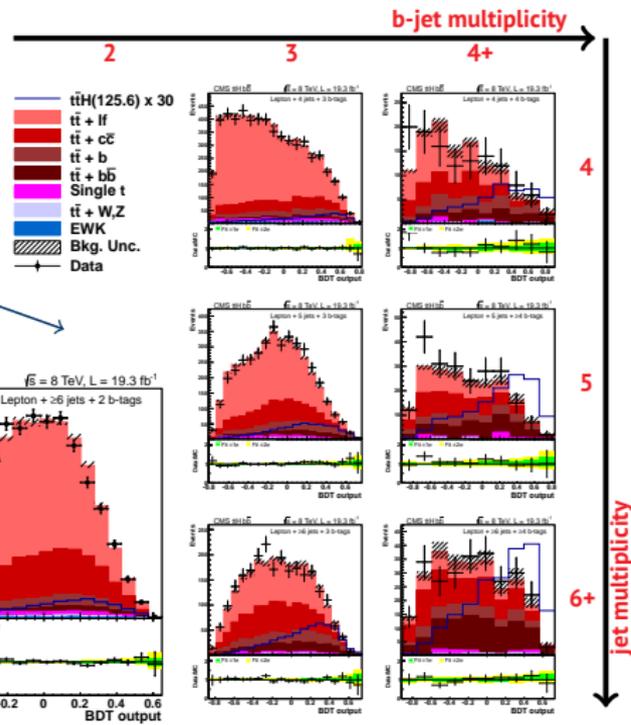
- ▶ Search in lepton+jets and dilepton channels
- ▶ Main backgrounds: $t\bar{t} + (b\bar{b}/c\bar{c}/\text{light quarks}/b/t)$, $t\bar{t} + V, W/Z + \text{jets}$, diboson & QCD

Analysis strategy

- Look for **high jet multiplicity** and **many b tags**
- Classify events in N_{jet} and $N_{b\text{-jet}}$ bins
- Build **multivariate discriminants** using event kinematics, shape and b-tag variables
- **Fit the discriminant shape** and extract limits

Results

- observed (expected) limit 4.1 (3.5) at $m_H = 125.6$ GeV
- significance 0.4σ (0.6σ)
- best-fitted $\mu = 0.7 \pm 1.9$

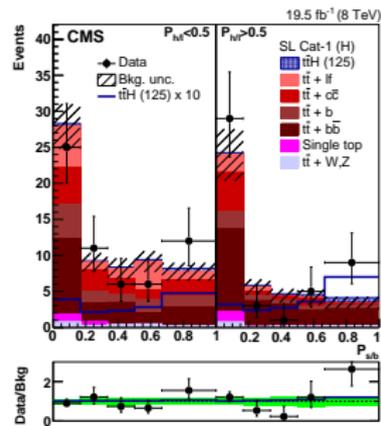


Analysis strategy

- Alternative approach using a **matrix element method**,
→ increased sensitivity over the BDT analysis.
- Define two **likelihood ratios** based on LO matrix elements,
→ **discriminate between $\bar{t}\bar{t}H$ (sig) and $\bar{t}\bar{t}+b\bar{b}$ (bkg)**.

$$P_{s/b} = \frac{w(\vec{y}|\bar{t}\bar{t}H)}{w(\vec{y}|\bar{t}\bar{t}H)+k_{s/b}w(\vec{y}|\bar{t}\bar{t}+b\bar{b})} \quad P_{h/l} = \frac{f(\vec{\xi}|\bar{t}\bar{t}+hf)}{f(\vec{\xi}|\bar{t}\bar{t}+hf)+f(\vec{\xi}|\bar{t}\bar{t}+lf)}$$

- **Classify events into one dilepton and three semi-leptonic categories based on (b)jet multiplicities.**
- Perform a **2D fit** of the ratios to extract limits.



Distribution of the $P_{s/b}$ discriminant, in two $P_{h/l}$ bins

MEM good for dealing with

- combinatorics
- small S/B ratios

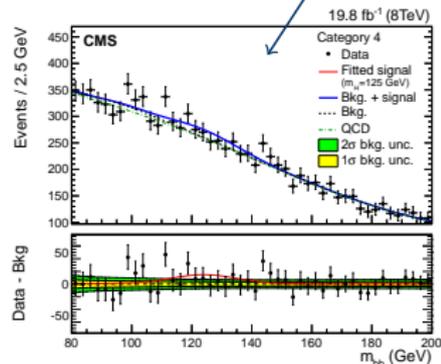
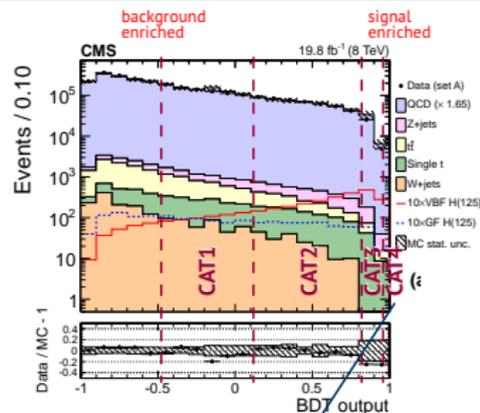
Results

- observed (expected) limit
4.2 (3.3) at $m_H = 125$ GeV
- best-fitted $\mu = 1.2^{+1.6}_{-1.5}$

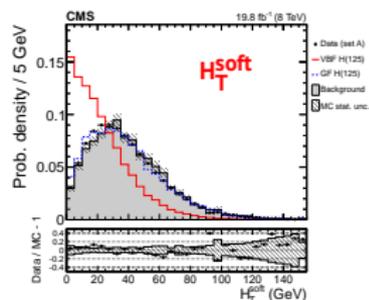
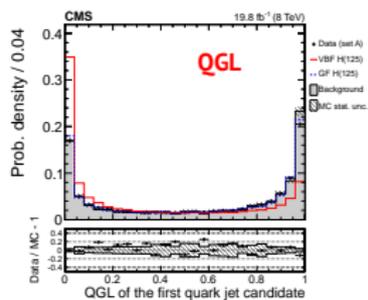
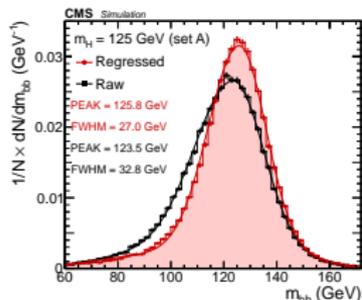
- ▶ Search for 4-jet final state ($q\bar{q}b\bar{b}$)
- ▶ Main backgrounds: QCD, W/Z+jets, $t\bar{t}$, single-t

Analysis strategy

- Two data samples: **nominal set A** (dedicated trigger) and **parked set B** (general purpose trigger)
- Event selection: **forward VBF light jets** and **central b jets**
- **Multivariate discriminant** using kinematics, b-tagging & quark/gluon likelihood, **leaving $m_{b\bar{b}}$ distribution unbiased**
- **Divide events into categories** based on BDT output value
→ 4 CATs in set A, 3 CATs in set B
- **Fit $m_{b\bar{b}}$ spectrum** simultaneously in all categories
- **Validation of fit strategy** by fitting $Z \rightarrow b\bar{b}$ resonance:
obs (exp) significance 3.6σ (3.3σ)



Fit of the invariant mass distribution in CAT4



Optimization

- **jet p_{\perp} regression** similar to VH analysis
→ improved $m_{b\bar{b}}$ resolution
- **quark/gluon discrimination** using jet shape properties (CMS-PAS-JME-13-002)
- no colour flow in rapidity gap between VBF jets
→ discriminating variable **soft H_T**
- improved event interpretation
→ **b jet identification using BDT**
- **QCD template derivation** in bkg enriched categories (CAT1 & CAT5) + transfer to other ones with $1^{\text{st}}/2^{\text{nd}}$ order polynomial multiplicative functions

Results

- observed (expected) limit 5.5 (2.5) at $m_H = 125$ GeV
- significance 2.2σ (0.8σ)
- best-fitted $\mu = 2.8^{+1.6}_{-1.4}$

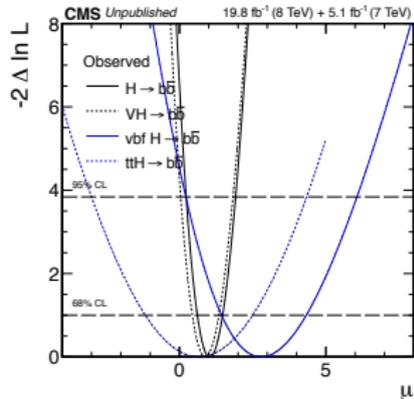
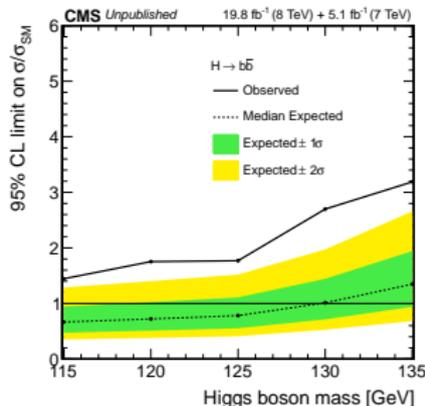
The three public CMS $H \rightarrow b\bar{b}$ results have been combined:

- **VH**: updated results (EPJC 75 (2015) 212)
- **ttH**: BDT analysis (JHEP 09 (2014) 087)
- **VBF** (arXiv:1506.01010)

H \rightarrow bb Channel	Best fit (68% CL)		Upper limits (95% CL)		Signal significance	
	Observed		Observed	Expected	Observed	Expected
VH	0.89 ± 0.43		1.68	0.85	2.08	2.52
ttH	0.7 ± 1.8		4.1	3.5	0.37	0.58
VBF	$2.8^{+1.6}_{-1.4}$		5.5	2.5	2.20	0.83
Combined	$1.03^{+0.44}_{-0.42}$		1.77	0.78	2.56	2.70

Results:

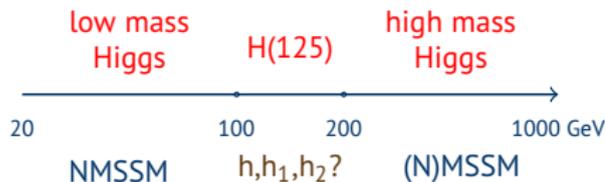
- observed significance 2.6σ
- best-fitted $\mu = 1.03^{+0.44}_{-0.42}$



model	structure	CP even	CP odd	charged
MSSM	2 doublets	h, H	A	H^\pm
NMSSM	2 doublets + 1 singlet	h_1, h_2, h_3	a_1, a_2	h^\pm

($\phi = h, H, A$)

- ▶ **Direct searches** for additional Higgs bosons with decays to bottom quarks
- ▶ **Both at low and at high mass** (~ 20 -1000 GeV)
- ▶ Set both **model dependent** and **model independent** limits



- ↔
- 1) $A \rightarrow Zh$ with $Z \rightarrow ll$ and $H \rightarrow b\bar{b}$
 - 2) **neutral MSSM** $b\bar{b}\phi$ with $\phi \rightarrow b\bar{b}$ (4b FS)
 - 3) **light NMSSM** h_1 with $h_1 \rightarrow b\bar{b}$ and MET from SUSY particle decays (\tilde{g}/\tilde{q})

- ▶ **Indirect searches** through tests of Higgs boson couplings

- ↔
- 4) $bq \rightarrow tHq$ with $H \rightarrow b\bar{b}$ and $t \rightarrow bl\nu$

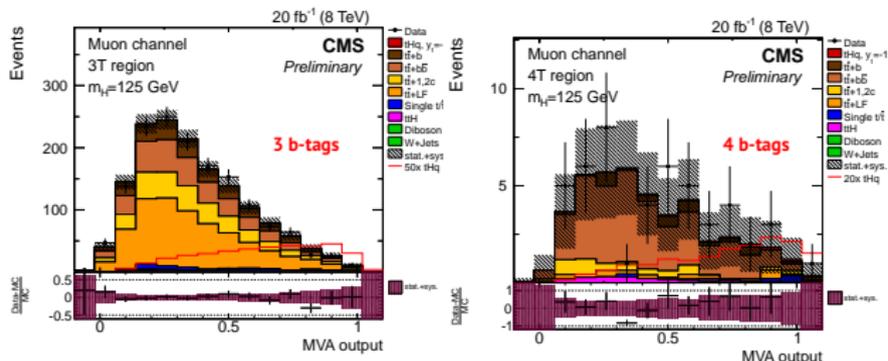
- ▶ **Search for H produced in association with single top:** $tHq \rightarrow 3bql\nu$
 → SM cross section is small
 → BSM cross section could be enhanced through anomalous couplings (here: $\kappa_t = -1$)
- ▶ **Main background:** $t\bar{t} \rightarrow 2b2ql\nu$

Search strategy

- **select:** 3 or 4 b-tags, 1 fwd light q, 1 lepton
- **2 multivariate discriminants for jet assignment**
 → 2 event interpretations: 1x tHq & 1x $t\bar{t}$
 → 2 sets of observables
- **1 multivariate discriminant for event classification**
 → tHq, $t\bar{t}$ and global inputs
- **fit the discriminant shape** to extract limits

Results

- observation agrees with both b-only and s+b hypothesis
- limits on the anomalous coupling model
 - observed 7.6
 - expected 5.1



Pseudoscalar A boson in 2HDM

$$A \rightarrow Z h \rightarrow l^+ l^- b \bar{b}$$

tHq

A \rightarrow Zh

ϕ

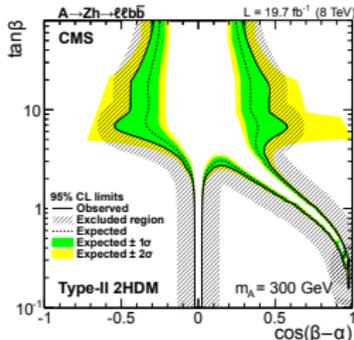
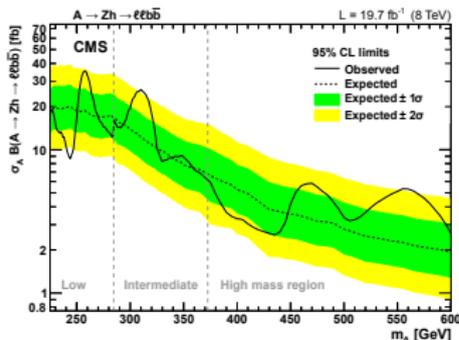
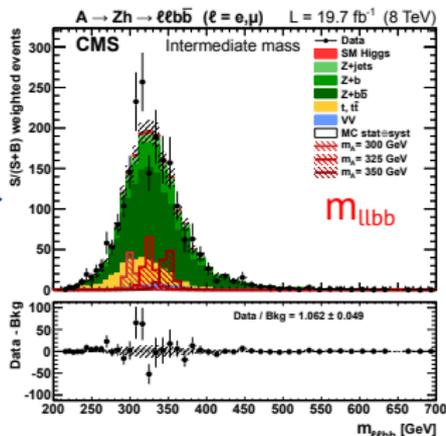
h_1

CMS-HIG-14-011
PLB 748 (2015) 221

- ▶ Search for A in mass range 225-600 GeV
- ▶ Main backgrounds: V + jets, $t\bar{t}$, VH, QCD

Search strategy

- reconstruct m_{llbb} invariant mass spectrum
- improve m_{llbb} resolution with **kinematic fit** to known m_h
- normalize $Z \rightarrow b\bar{b}$ and $t\bar{t}$ backgrounds in **control regions**
- train **BDT per mass region** (low/intermediate/high)
- perform **2D fit to $m_{llbb} \times$ BDT in signal region** and extract limits on cross section times branching fraction



Results

- **model independent limits** ranging from 30 fb (low mass) to 3 fb (high mass)
- exclusion limits for **Type-I and Type-II 2HDM** with $m_A = 300$ GeV

Neutral MSSM boson ϕ (=A,H,h)

tHq

A \rightarrow Zh

ϕ

h_1

$\phi \rightarrow b\bar{b}$ **[NEW]**

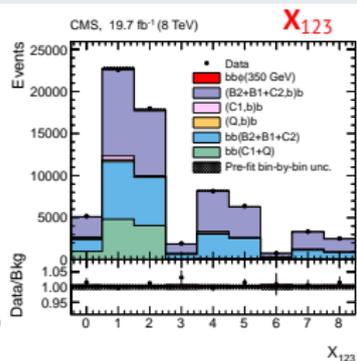
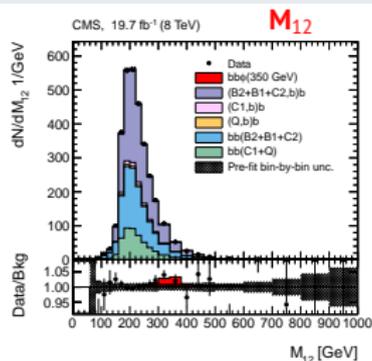
CMS-HIG-14-017

Submitted to JHEP – arXiv:1506.08329

- **Search for many b jets:** ϕ produced in association with 1 or 2 b quarks.
- **Main backgrounds:** heavy flavour QCD multijets, $t\bar{t}$ + jets

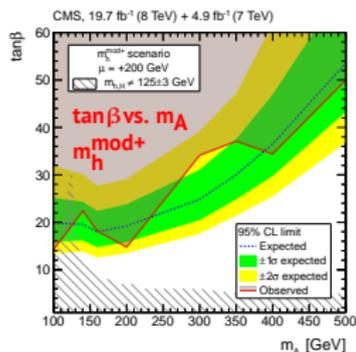
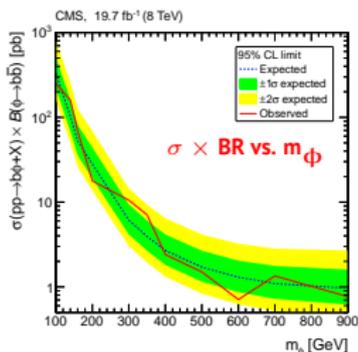
Search strategy

- QCD makes **dedicated trigger** essential \rightarrow tight online b-tagging
- **select:** 3b = signal, 2b = background
- **two discriminating variables** chosen:
 - M_{12} : inv. mass of the 2 leading jets
 - X_{123} : b-tag estimator built from Σm_{3V}
- **derive templates:**
 - $\text{sig}(M_{12}, X_{123}, m_\phi)$ from simulation
 - $\text{bkg}(M_{12}, X_{123})$ from 2b category
- perform **2D fit to** (M_{12}, X_{123}) and extract limits



Results

- limits on $\sigma \times \text{BR}$ from 300pb (low mass) to 1pb (high mass)
- limits on $\tan\beta$ in various MSSM benchmark scenarios $\rightarrow m_h^{\text{max}}, m_h^{\text{mod+}}, m_h^{\text{mod-}}, \text{light-}\tilde{\tau}, \text{light-}\tilde{t}$



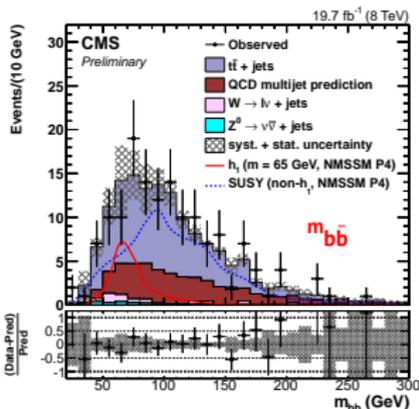
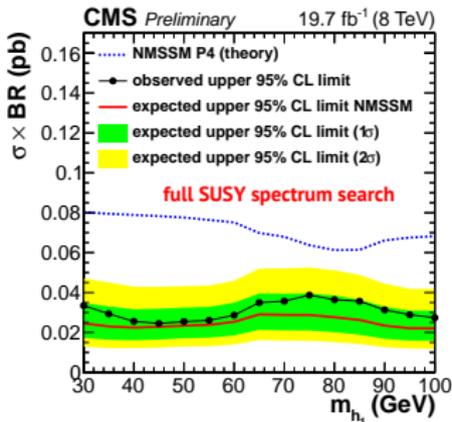
- ▶ **Search for light h_1 :**
 - identify h_2 with SM H(125)
 - h_1 possibly suppressed in conventional production modes
 - could be present in SUSY cascades
- ▶ **Main background:** $t\bar{t}$ + jets, QCD

Search strategy

- **select:** large H_T , large MET, 2 hard q, 2 soft b
- **two approaches**
 - 1) search h_1 peak over bkg
 - 2) search for full SUSY spectrum
- background (QCD, $t\bar{t}$) validated in control regions
- **fit $m_{b\bar{b}}$ spectrum** and extract limits

Results

- observation agrees with both b-only and s+b hypothesis
- NMSSM P4 excluded for $m_{h_1} \in [30, 100 \text{ GeV}]$ at SUSY scale $\sim 1 \text{ TeV}$
- NMSSM parameter space scanned → further exclusion limits set



Summary

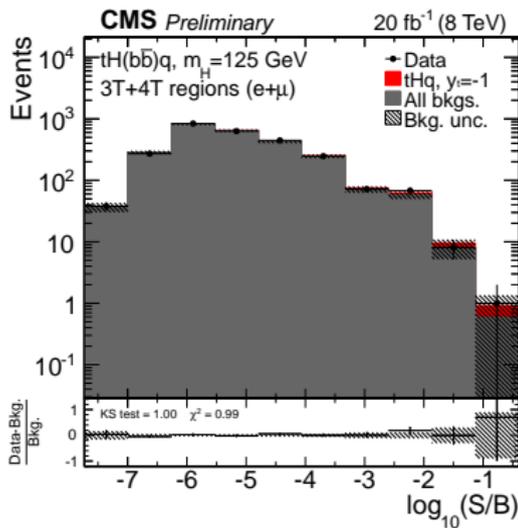
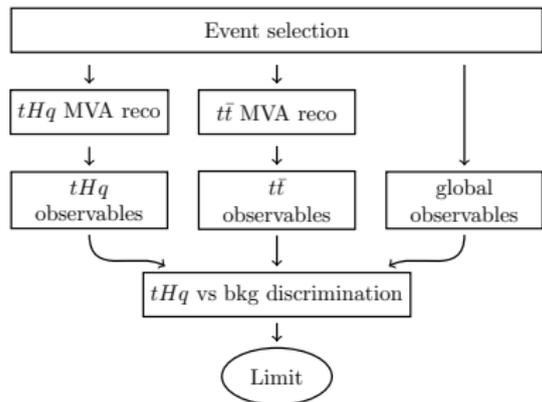
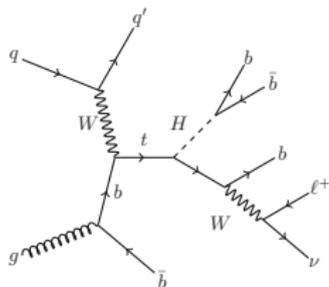
Legacy of Run I

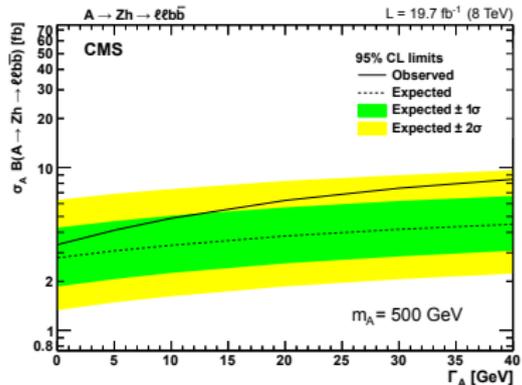
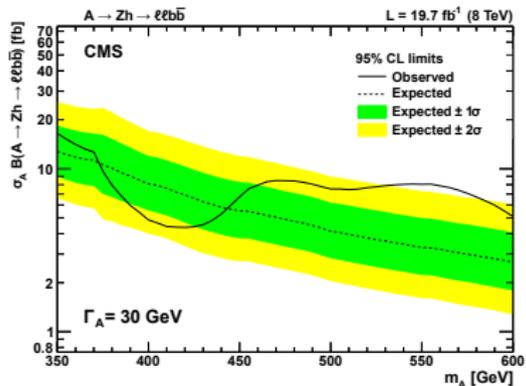
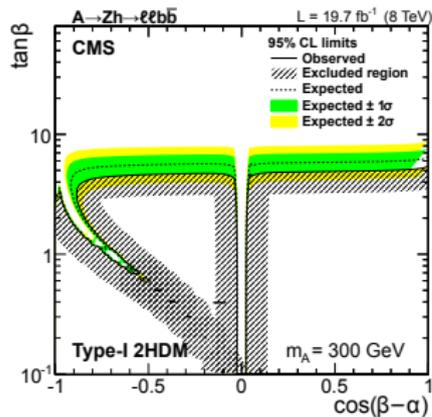
- **searches for SM H** \rightarrow $b\bar{b}$, also in the VBF channel
- **SM combination** with observed significance 2.6σ
- direct and indirect **searches for physics beyond the standard model**
 - \rightarrow large phase space covered, but no evidence yet
 - \rightarrow limits set in a wide mass range (light NMSSM h_1 , heavier MSSM ϕ)
 - \rightarrow interpretation of limits in various benchmark scenarios

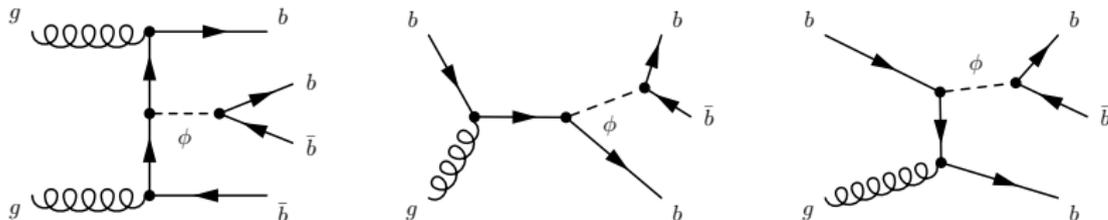
Outlook for Run II

- increased sensitivity should bring **evidence of H** \rightarrow $b\bar{b}$
- continue exploring with **extended reach for BSM searches**

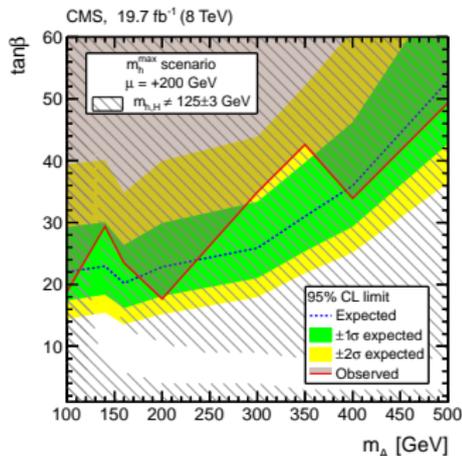
Backup



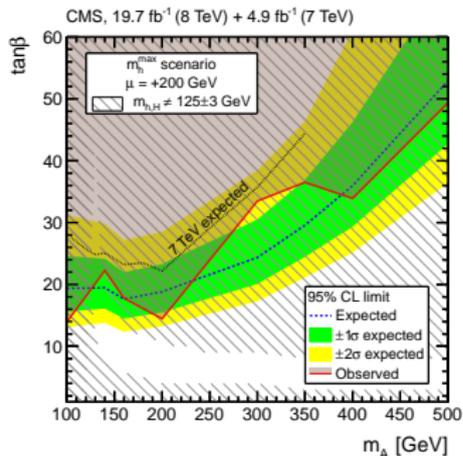


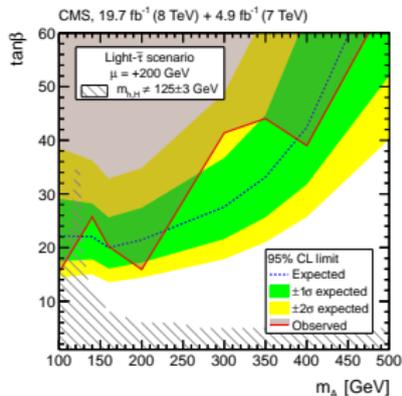
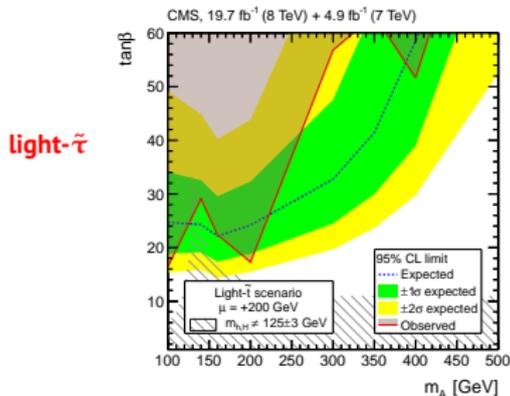
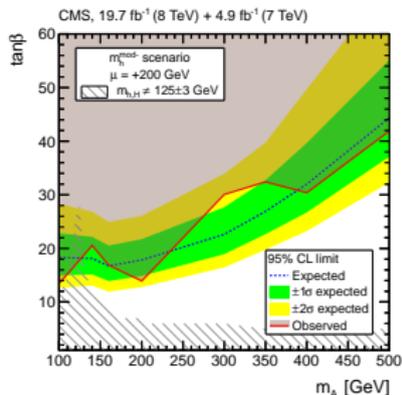
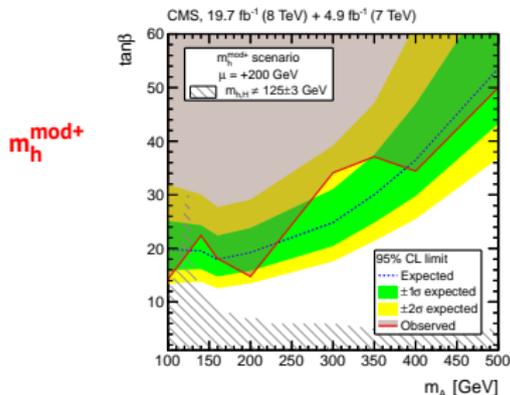


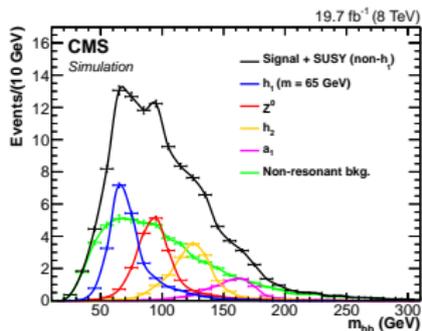
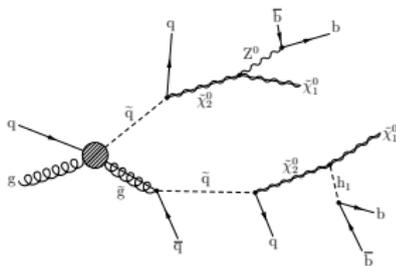
m_h^{\max}
8 TeV



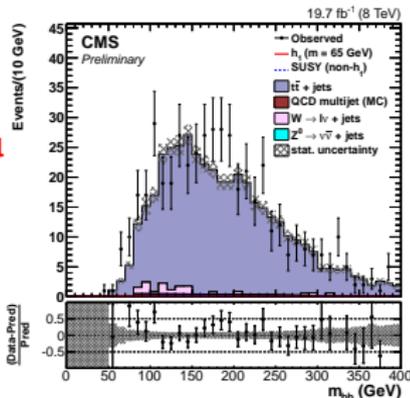
m_h^{\max}
7+8 TeV







$t\bar{t}$ control region



QCD control region

