hh projections - CMS Status

Konstantin Androsov, Olivier Bondu, Maxime Gouzevitch
on behalf of CMS-hh community

CP3 - UC Louvain

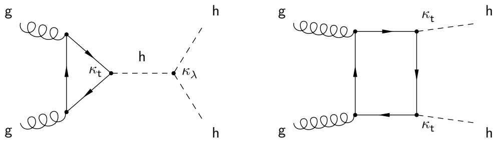
2016-09-15 - ECFA prep. meeting

DISCLAIMER

Should the slides ever end up outside of this meeting: this is an INFORMAL meeting, the content shown here represents INTERNAL WORK IN PROGRESS and as such has NOT BEEN APPROVED BY THE CMS COLLABORATION.

THESE SLIDES ARE CONFIDENTIAL AND NOT FOR DISTRIBUTION

Why hh?



The Standard Model hh production

- Mainly produced via **gluon fusion** (as for single h)
- Access to the **self-coupling λ**
  - Scalar potential structure
- **Key property measurement** of h(125)
- Major setback: **very low production cross-section**
  - $\sigma(pp \to hh)^{SM}_{NNLO+NNLL} = 33.45$ **fb** @ 13 TeV
  - Strong destructive interference of the two main diagrams: by (lack of) chance SM is almost the most destructive case...
- *Start* to be sensitive to SM at HL-LHC: **upgrade strategy is decisive**

Bibliography: existing public CMS-hh results

13 TeV nonresonant

- 2015 data (2.3-2.7 fb-1):
  - **HIG-16-026** (b$\bar{\text{b}}$b$\bar{\text{b}}$)
  - **HIG-16-024** (b$\bar{\text{b}}\ell\bar{\nu}_\ell\bar{\ell}\nu_\ell$)
  - **HIG-16-012** (b$\bar{\text{b}}\tau^-\tau^+$)
  - **HIG-16-032** (b$\bar{\text{b}}\gamma\gamma$)
- 2016 data (12.9 fb-1):
  - HIG-16-028 (b$\bar{\text{b}}\tau^-\tau^+$)

8 TeV nonresonant (19.7 fb-1)

- 1603.06896 (b$\bar{\text{b}}\gamma\gamma$)
- HIG-15-013 (b$\bar{\text{b}}\tau^-\tau^+$)

14 TeV projections (3000 fb-1)

- FTR-15-002

Resonant results (for completeness):

b$\bar{\text{b}}\tau^-\tau^+$: HIG-16-013, HIG-16-029, Phys. Lett. B 755 (2016) 217, HIG-15-013, EXO-15-008

b$\bar{\text{b}}\ell\bar{\nu}_\ell\bar{\ell}\nu_\ell$: HIG-16-011

b$\bar{\text{b}}$b$\bar{\text{b}}$: **HIG-16-002**, Phys. Lett. B 749 (2015) 560, Eur. Phys. J. C 76 (2016) 371

b$\bar{\text{b}}\gamma\gamma$: HIG-16-032, 1603.06896

FTR-15-002: Delphes and other TP studies (1)

FTR-15-002: 14 TeV studies

- Done for TP (**not new**)
- HL-LHC with 3 ab-1, <pu> = 140
- Performed to **discuss the CMS Phase-II upgrade**
- Only gluon-fusion production (90% of the total)
- Triggers assumed 100% efficient

$b\bar{b}\ell\bar{\nu}_\ell\bar{\ell}\nu_\ell$: 1500 expected events

- Study using Delphes
- Only $t\bar{t}$ background
- ***m_{jj}* selection**
- **Cut on MVA**, signal extracted via cut-and-count
- Uncertainty on signal yield in $\approx$ **180-500%** range
- Document includes sensitivity study on background uncertainty

FTR-15-002: Delphes and other TP studies (2)

$b\bar{b}\gamma\gamma$: 320 expected events

- Study using gen level info smeared to model Phase-II performance
- Lepton veto ($t\bar{t}h$ rejection)
- Photon categories, no b-tag categories
- 2D **($m_{\gamma\gamma}$, m_{jj}) fit** signal extraction
- Expected significance of 1.6 σ
- Uncertainty on signal yield of 67%
- Document includes sensitivity studies (eg Phase-I aged)

$b\bar{b}\tau^-\tau^+$: 9000 expected events

- $\tau_\mu\tau_h$ and $\tau_h\tau_h$ channels
- Study using Delphes (most backgrounds) and gen-smearing ($t\bar{t}$)
- **$m_{\tau\tau}$ and m_{jj} selection**
- Signal extraction on
  - **m_{T2}** for $\tau_h\tau_h$
  - MVA output for $\tau_\mu\tau_h$
- Expected significance of 0.9 σ
- Unc. on signal yield of 105%
- Document includes a trigger study

Combination of $b\bar{b}\gamma\gamma$and $b\bar{b}\tau^-\tau^+$

- Expected **significance of 1.9 σ**
- **Uncertainty on signal yield of 54%**

What is being projected in these slides?

Projections at 13 TeV for 3000 fb-1

- **From currently existing data analyses: only SM case**
  - No λ_{hhh} scan (neither for your other favorite anomalous coupling)
- Delphes: **new feasibility study for hh $\rightarrow b\bar{b}WW \rightarrow b\bar{b}jj\ell\bar{\nu}_\ell$**
- No λ measurement performance projection based on 1607.04251
  - Admittedly this is not either directly 'hh' projection anyway
- Resonant analysis: **WED exclusion bounds from $b\bar{b}b\bar{b}$**

Some considerations:

- $\sigma^{13\text{TeV}}_{SM\,HH} = 33.41^{+7.3\%}_{-8.4\%}$ fb at 13 TeV for $m_{\text{H}} = 125.09$ GeV
- $\sigma^{14\text{TeV}}_{SM\,HH} = 39.51$ fb at 14 TeV for $m_{\text{H}} = 125.09$ GeV (18% increase)
- **Naively** expect a factor $\frac{\sigma^{14\text{TeV}}_{SM\,HH}/\sigma^{13\text{TeV}}_{SM\,HH}}{\sqrt{14\text{TeV}/13\text{TeV}}} = 1.15$ boost in sensitivity by going to 14 TeV

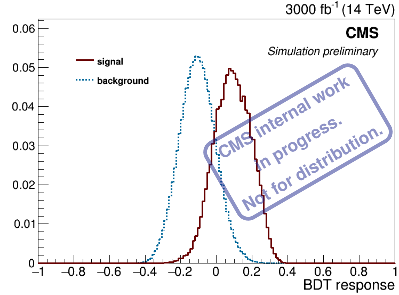
Projections: Delphes analyses

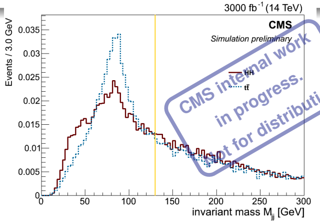
- $b\overline{b}jj\ell\bar{\nu}_\ell$

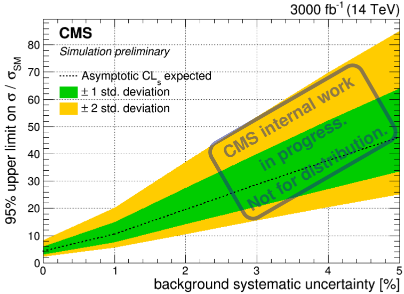
hh → $b\bar{b}VV$ → $b\bar{b}\ell\bar{\nu}_\ell jj$, 14 TeV, 3000 fb-1, <pu> = 140

New feasibility study at 14 TeV

- Delphes (CMS Phase-II)
- **Challenge of jet combinatorics**
- $t\bar{t}$ background only
- PUPPI for PU jet removal
- MVA trained with 34 variables (jet multiplicities, p_T's, $\not{E}_T$, $\Delta\phi$'s, ΔR's, invariant masses)
- MVA cut and count
- **Upper limit ≈ $4 - 46\times$ SM**







Projections: 13TeV analyses extrapolations

- $b\bar{b}b\bar{b}$
- $b\bar{b}\ell\bar{\nu}_\ell\bar{\ell}\nu_\ell$
- $b\bar{b}\tau^-\tau^+$
- $b\bar{b}\gamma\gamma$

Procedure

- **Kinematics are SM-like**
- **Production only through gluon-fusion** (90% of the total)
- **Simplified datacards from the data analysis teams**
  - Enforced requirement at preapproval time
- Harmonization of assumptions on common uncertainties
- Asimov dataset with signal strength fixed to the SM value
- For the final numbers: **SM BR are assumed** (from LHCHXSWG)
- **NO combination** of channels is attempted

Common uncertainties: the $H \to b\bar{b}$ leg

theory Uncertainties divided by 2 for Scenario 2

lumi Assumed to be 1.5%

JES **Assumed 1% accuracy**

- We had typically 2% in 2015 data considering the p_{T}range

JER Expected to **degrade by 10%** (considering the p_{T}range)

- **Applied to $b\bar{b}\gamma\gamma$** analysis which exploits directly $m_{b\bar{b}}$ resolution
- **Other final states:** either do not exploit directly $m_{b\bar{b}}$ or suffers from other larger uncertainties: assuming **unchanged JER**

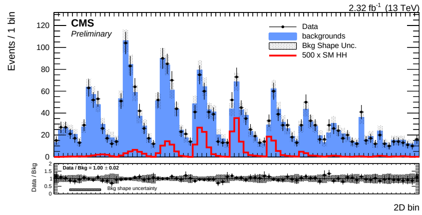
b-tag : **Assume tagging uncertainty of 1%, unchanged WPs**

- Tracker upgrade planned for HL-LHC (compensating performance loss due to larger PU)
- Nearly all selected signal jets are b-jets, typical b-tag uncertainty in 2015 was 1.5%: divide by a factor 1.5
- MC-driven background estimates (eg $t\bar{t}$): similarly to signal: eff. unc. scaling on a per-jet basis

HIG-16-026: hh → $b\bar{b}b\bar{b}$

Analysis in a nutshell

- Analysis of 2015 data
- **Huge QCD multijet background, data-driven** estimate with event mixing
- **B-tagging at trigger level**
- MVA for signal extraction
- Upper limit at 3490 fb on $\sigma \times BR$ ($\approx$ 308$\times$ SM)



Projection

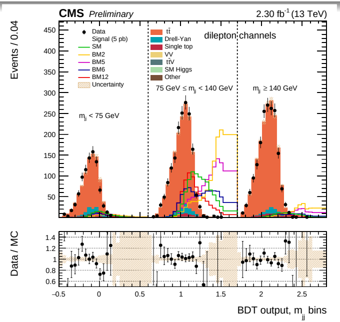
- Main uncertainty: background statistics, scale with stat
- All other backgrounds: assumed unchanged wrt 2015 analysis
- **Worry about trigger thresholds**: may alter improvements of better b-tag or better knowledge of trigger efficiencies

Channel	Median expected limits in μ_r		Z-value		Upward uncertainty as fraction of $\mu_r = 1$	
	Scenario 2	Stat. Only	Scenario 2	Stat. Only	Scenario 2	Stat. Only
$b\bar{b}b\bar{b}$	7.0	2.9	0.39	0.67	2.5	1.5

HIG-16-024: $hh \to b\bar{b}VV \to b\bar{b}\ell\bar{\nu}_\ell\bar{\ell}\nu_\ell$

Analysis in a nutshell

- Analysis of 2015 data
- Large **$t\bar{t}$, DY bkg, MC-driven**
- Dilepton triggers
- TH2D (MVA, m_{jj}) signal extract.
- Exp. upper limit at $92.8^{+59.9}_{-33.4}$ fb on $\sigma \times BR$ ($\approx 400\times$ SM)



Projection

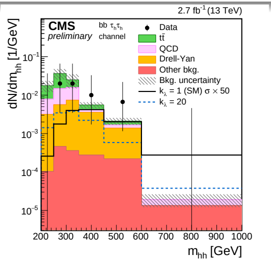
- All **backgrounds** would **move to data-driven**: neglecting unc. (too aggressive?)
- Assume unchanged trigger unc. (in absence of complete study)
- Preexisting 14 TeV result: **similar sensitivity**
- Analysis will be improved with further (lepton flavour) categories: current **estimate is a bit optimistic**

Channel	Median expected limits in μ_r		Z-value		Upward uncertainty as fraction of $\mu_r = 1$	
	Scenario 2	Stat. Only	Scenario 2	Stat. Only	Scenario 2	Stat. Only
$b\bar{b}\ell\bar{\nu}_\ell\bar{\ell}\nu_\ell$	4.7	4.6	0.45	0.47	2.2	2.3

HIG-16-012: hh $\rightarrow b\bar{b}\tau^-\tau^+$

Analysis in a nutshell

- Ana. of 8TeV, **2015**, 2016 data
- $\boldsymbol{\tau_e\tau_h}$, $\boldsymbol{\tau_\mu\tau_h}$ **and** $\boldsymbol{\tau_h\tau_h}$ channels
- Large **$t\bar{t}$, DY bkg, MC-driven**
- QCD background, data-driven
- kinematic fits, MVA cut, signal extracted on m_{hh}
- Upper limit at 7.2 pb on $\sigma \times BR$ ($\approx$ 200$\times$ SM)



Projection

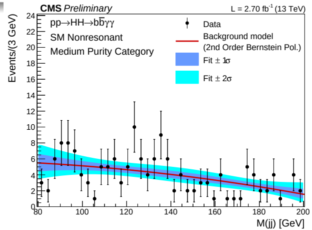
- 2016: SiStrip Tracker ineff.: safer projection with 2015 analysis
- Neglect QCD unc. (stat. dom.)
- Other bkg: from MC
- $t\bar{t}$ shape unc.: divided by 3 (understood p_T disagreement)
- Lepton uncertainties unchanged
- Preexisting 14 TeV result is better: **understood**
  - Uses MT2, which is not so great for sensitivity to BSM
  - Assumes 0 QCD, while here *uncertainty* is 0 (higher bkg yield)
  - Change in center of mass energy

Channel	Median expected limits in μ_r		Z-value		Upward uncertainty as fraction of $\mu_r = 1$	
	Scenario 2	Stat. Only	Scenario 2	Stat. Only	Scenario 2	Stat. Only
$b\bar{b}\tau^-\tau^+$	7.0	4.6	0.30	0.44	3.5	2.3

HIG-16-032: hh → $b\bar{b}\gamma\gamma$

Analysis in a nutshell

- Analysis of 2015 data
- **$\gamma\gamma + jj$, $\gamma j + jj$ bkgs, data-driven**
- Same photons as in $h(\gamma\gamma)$ analysis (trigger, ID, iso, unc.)
- Two b-tag categories, cut on $\tilde{M}_X$
- 2D $(m_{\gamma\gamma}, m_{jj})$ fit signal extraction
- Exp. upper limit at 7.85 fb on $\sigma \times BR$ ($\approx 90\times$ SM)



Projection

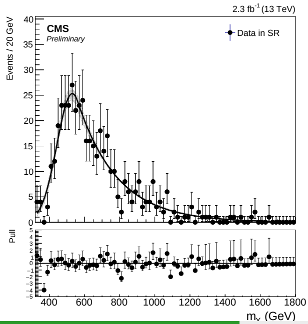
- **Assume same bkg shape**
- Bkg unc. scale with $1/\sqrt{L}$
- Unc. for $h(\gamma\gamma)$ leg follow $h(\gamma\gamma)$ projection
  - Signal photons are 90% in the barrel: neglect endcap eff. degradation
- Preexisting 14 TeV result: **similar sensitivity**

Channel	Median expected limits in μ_r		Z-value		Upward uncertainty as fraction of $\mu_r = 1$	
	Scenario 2	Stat. Only	Scenario 2	Stat. Only	Scenario 2	Stat. Only
$b\bar{b}\gamma\gamma$	1.3	1.3	1.6	1.6	0.77	0.64

HIG-16-002: $X \to hh \to b\bar{b}b\bar{b}$

Analysis in a nutshell

- Analysis of 2015 data
- **Huge QCD multijet background**, data-driven estimate from a fit in side-bands
- **B-tagging at trigger level**
- kin. fit, m_{bbbb} for signal extraction



Projection

- Same systematic scaling as for $b\bar{b}\gamma\gamma$ analysis
- Results are systematics-limited
- Compare to radion production in bulk WED
- **Mass scale Λ_R**, interpreted as the ultraviolet cutoff of the model, **excluded up to 9 TeV!**

$m_X(TeV)$	Median expected limits on σ (fb)			$\sigma_R(\Lambda_R = 1\,TeV)$ (fb)	Λ_R (TeV) excluded
	2.3 fb-1	Stat. Only	Scenario 2		
0.7	129.4	3.4	7.3	584	8.9
1.0	81.5	2.4	4.4	190	6.6

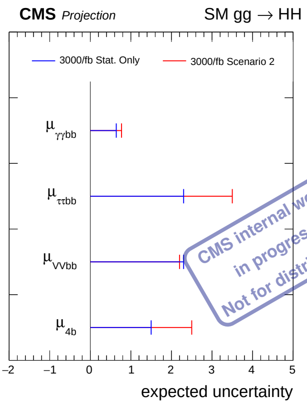
CMS internal work in progress. Not for distribution.

Summary

Channel	Median expected limits in μ_r		Z-value		Upward uncertainty as fraction of $\mu_r = 1$	
	Scenario 2	Stat. Only	Scenario 2	Stat. Only	Scenario 2	Stat. Only
$b\bar{b}\gamma\gamma$	1.3	1.3	1.6	1.6	0.77	0.64
$b\bar{b}\tau^-\tau^+$	7.0	4.6	0.30	0.44	3.5	2.3
$b\bar{b}\ell\bar{\nu}_\ell\bar{\ell}\nu_\ell$	4.7	4.6	0.45	0.47	2.2	2.3
$b\bar{b}b\bar{b}$	7.0	2.9	0.39	0.67	2.5	1.5

At this stage:

- $b\bar{b}\gamma\gamma$: similar to TP, ok
- $b\bar{b}\tau^-\tau^+$: worse than TP: **understood**, ok
- $b\bar{b}\ell\bar{\nu}_\ell\bar{\ell}\nu_\ell$: similar to TP, but may be a bit optimistic, being checked
- $b\bar{b}b\bar{b}$: tough to say (trigger + QCD), but **analysis and projection are conservative**: ok
- $b\bar{b}jj\ell\bar{\nu}_\ell$: looks challenging
- X $\rightarrow b\bar{b}b\bar{b}$: serious dent in WED models

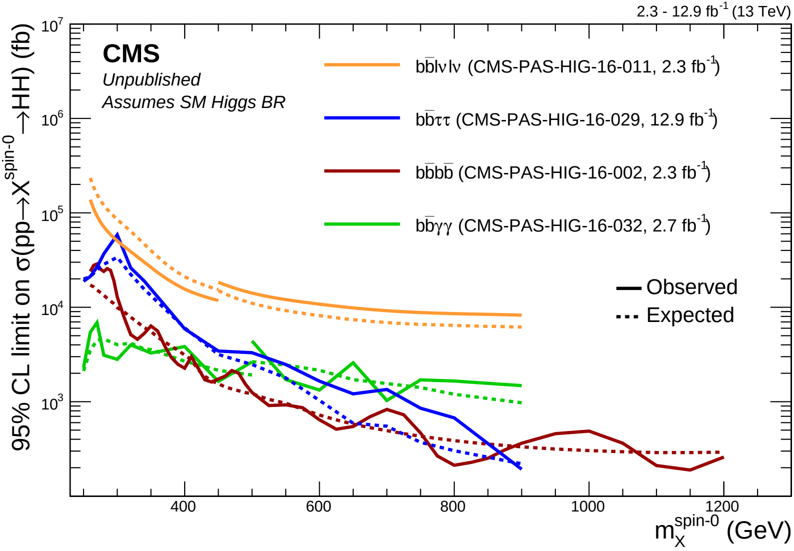


Reminder: current results

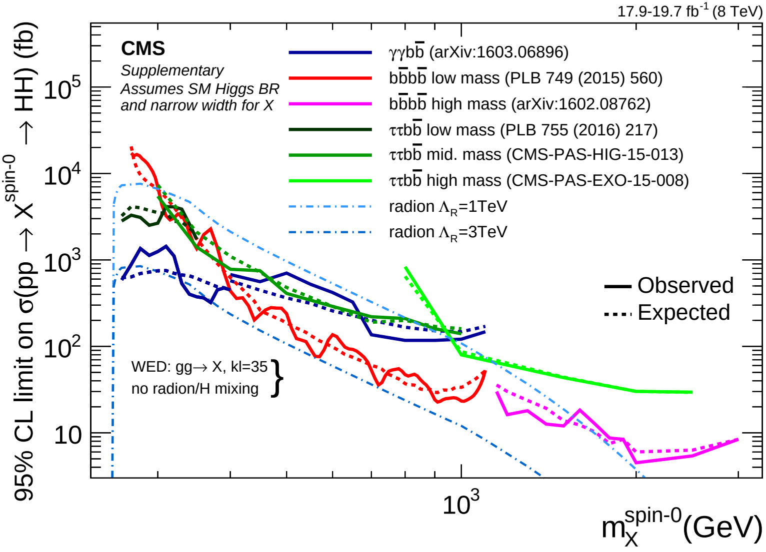
- $b\bar{b}\gamma\gamma \approx 90\times$ SM
- $b\bar{b}\tau^-\tau^+ \approx 200\times$ SM
- $b\bar{b}b\bar{b} \approx 300\times$ SM
- $b\bar{b}\ell\bar{\nu}_\ell\bar{\ell}\nu_\ell \approx 400\times$ SM

BACKUP

Resonant analyses: Run II



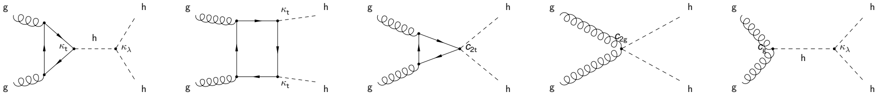
Resonant analyses: Run I



SM hh production and EFT (2)

There is hope yet: we have some leeway...

- Self-coupling λ predicted by SM, but **not constrained experimentally**
- There exist **other couplings** in BSM scenarios: c_{2t}, c_{2g}, c_g
- Some freedom on κ_t



5D parameter scan: κ_λ, κ_t, c_{2t}, c_{2g}, c_g

- **Cross-section** can vary sensibly: $[10^{-1}, 10^4] \times \sigma(pp \rightarrow hh)^{SM}$
- **Signal shape** can be significantly different from SM

Scan 5D parameter space: the clustering method

It is impractical to generate a 5D grid of signal samples

- LO generation and various theoretical arguments: most of the physics is contained in the m_{hh} spectrum
  - and somewhat in $\cos(\theta)^*_{CS}$
  - Extensive discussions in a TH-EXP workshop last year

The cluster analysis JHEP04(2016)126

- Exp. analyses **sensibility depend on the signal shape**
- **Cluster regions** of the parameter space with similar kinematics
- Define **benchmark points (BM)**: representative of a cluster
- Injection of the 12 benchmarks in the CMS full-sim MC prod.

Public results

- Used for both HIG-16-028 ($b\bar{b}\tau^-\tau^+$) and HIG-16-024 ($b\bar{b}\ell\bar{\nu}_\ell\bar{\ell}\nu_\ell$)