Measurements of Higgs boson production and properties in the di-photon decay channel using the CMS detector

Vittorio Raoul Tavolaro on behalf of the CMS collaboration

38th International Conference on High Energy Physics, 4 August 2016, Chicago IL
Introduction

- Clean final state with two highly energetic photons
- Final state fully reconstructed with high resolution
- Very small branching fraction (~0.2%)
- Large backgrounds (γγ, γj, jj)
- Exclusive categories targeting: gluon-gluon fusion (ggH), vector boson fusion (VBF) and ttH production modes
- Search for a narrow peak on a falling background in mass distribution
- 2016 dataset analysed, 12.9/fb collected at 13 TeV (HIG-16-020)
Analysis strategy

- **Photons** energy reconstructed and corrected, rejection of fake photons
- **Events categorised** into classes (S/B, mass resolution, additional particles) to **improve the analysis** sensitivity
- Extraction of signal through fit of **di-photon invariant mass** spectrum in each category
$m_{\gamma\gamma} = \sqrt{2E_1E_2(1 - \cos \theta)}$

- Electro-magnetic calorimeter (ECAL) response:
  - corrected for **change in time**
  - inter-calibrated to be uniform in $\eta/\phi$
  - adjustment of **absolute scale**

- **Energy and its uncertainty** corrected for local and global shower containment:
  - **regression** targeting $E_{\text{true}}/E_{\text{reco}}$
  - **Scale** vs time and **resolution** calibration: $Z \rightarrow ee$ peak used as reference
  - **Corrected** energies and resolutions used in the analysis

---

**V.R. Tavolaro**

**04.08.2016**
**Vertex identification**

\[ m_{\gamma\gamma} = \sqrt{2E_1E_2(1 - \cos \theta)} \]

- Vertex assignment correct within **1 cm → negligible** impact on mass resolution

- No ionization in the tracker for photons

- **Multi-variate approach** for vertex identification
  - exploit kinematic correlations and **track** distribution imbalance
  - direction of **conversion tracks**, when present

- Second MVA estimates **probability of correct vertex** choice, used for di-photon classification

- Method validated on \( Z \rightarrow \mu\mu (\gamma+j \text{ for converted } \gamma) \) events, where vertex found after removing muon tracks
Photon selection

- **Double-photon trigger** selection based on transverse energy, $m_{\gamma\gamma}$, isolation and electromagnetic shower shapes variables

- Minimal **pre-selection**, similar to but tighter than trigger selection

- **MVA based** photon ID classifier:
  - goal: discriminate between **prompt** and **fake** photons
  - inputs are shower shape variables, particle-flow isolations, kinematic of photon, median energy density ($\rho$)

![Graph showing Photon ID BDT score with Data, Z→ee MC, and MC syst.]

CMS Preliminary 12.9 fb$^{-1}$ (13 TeV)
Tagged events

**ttH tags**

**ttH leptonic tag**

\[ t\bar{t} \rightarrow bl\nu_l\bar{b}q\bar{q} \quad t\bar{t} \rightarrow bl\nu_l\bar{b}l' \nu_l' \]

- (sub)lead pT/m_\gamma > 1/2(1/4)
- at least one lepton (\(\ell = e, \mu\)), away from Z peak
- \(\geq 2\) jets
- \(\geq 1\) b-jet

**ttH hadronic tag**

\[ t\bar{t} \rightarrow bq\bar{q} \quad t\bar{t} \rightarrow bq\bar{q} \]

- (sub)lead pT/m_\gamma > 1/2(1/4)
- 0 leptons
- \(\geq 5\) jets
- \(\geq 1\) b-jet
Tagged events

**ttH tags**

**ttH leptonic tag**

\[ \bar{t} \rightarrow b \nu \bar{t} b q \bar{q}' \]

- (sub)lead \( p_T/m_{\gamma\gamma} > 1/2(1/4) \)
- at least one lepton \((\ell = e, \mu)\), away from Z peak
- \( \geq 2 \) jets
- \( \geq 1 \) b-jet

**ttH hadronic tag**

\[ \bar{t} \rightarrow b q \bar{q}' \bar{b} q \bar{q}' \]

- (sub)lead \( p_T/m_{\gamma\gamma} > 1/2(1/4) \)
- 0 leptons
- \( \geq 5 \) jets
- \( \geq 1 \) b-jet

**VBF tags**

- Identify events with **2 jets** through a MVA
  - inputs: \( p_T/m_{\gamma\gamma} \) of both photons, \( p_T \) of both jets, \( m_{jj} \), \( \Delta\eta_{jj} \), Zeppenfeld variable, \( \Delta\phi_{\gamma\gamma} \)
  - 2 jets with \( p_{T1} > 30 \text{GeV}, p_{T2} > 20 \text{GeV}, |\eta| < 4.7, m_{jj} > 250 \text{ GeV} \)
- **VBF Classification** BDT combines di-jet and diphoton BDT: 2 data categories (**VBF tag 0-1**)
“Untagged” events

Categorise events into classes (S/B, resolution) to improve the analysis sensitivity

- **Multivariate discriminator** is used to separate diphoton pairs with signal-like kinematics, high photon ID scores and good mass resolution from background.

- Input variables: pT of the two photons rescaled by pair mass, $\eta$, $\cos(\Delta\phi)$, photon ID scores, relative mass resolutions under correct/incorrect vertex selection hypotheses, vertex probability estimate.

- 4 event categories (“Untagged” events)
• **Fully parametric** signal model from simulation
  • physical nuisances allowed to float
  • **corrections** and data/MC efficiency **scale factors** applied

• Background model **data driven**:
  • background functional form treated as **discrete nuisance** parameter
  • for each category, use different functional forms
    (sums of exponentials, sums of power law terms, Laurent series and polynomials)
Mass spectra, tagged events

V.R. Tavolaro

04.08.2016
Mass spectra, untagged events

Untagged 0

Untagged 1

Untagged 2

Untagged 3

V.R. Tavolaro

04.08.2016
Mass spectra, all categories

\[ \text{CMS Preliminary} \]

\[ H \rightarrow \gamma\gamma \]
\[ \hat{m}_H = 126.0 \text{ GeV}, \hat{\mu} = 0.95 \]

Events / GeV

0
1000
2000
3000
4000
5000
6000
7000
8000
9000

12.9 fb\(^{-1}\) (13 TeV)

All categories

- Data
- S+B fit
- B component

\[ \pm 1 \sigma \]
\[ \pm 2 \sigma \]

B component subtracted

S/(S+B) weighted

\[ \text{CMS Preliminary} \]

\[ H \rightarrow \gamma\gamma \]
\[ \hat{m}_H = 126.0 \text{ GeV}, \hat{\mu} = 0.95 \]

All categories

- Data
- S+B fit
- B component

\[ \pm 1 \sigma \]
\[ \pm 2 \sigma \]

B component subtracted

S/(S+B) weighted

\[ \mu = 126.0 \text{ GeV} \]

Preliminary

CMS (13 TeV)

12.9 fb\(^{-1}\)

V.R. Tavolaro

04.08.2016
• **Significance** at 125.09 GeV: **5.6σ observed** (6.2σ expected)
• **Maximum observed** significance is **6.1σ** at 126.0 GeV
• Best-fit signal strength  $\hat{\sigma}/\sigma_{SM} = 0.95^{+0.21}_{-0.19} = 0.95 \pm 0.17\,(\text{stat.})^{+0.08}_{-0.05}\,(\text{theo.})^{+0.10}_{-0.07}\,(\text{syst.})$
Results

- Signal strength measured is measured in **bosonic and fermionic** components

\[
\hat{\mu}_{VBF,VH} = 1.59^{+0.73}_{-0.45} \\
\hat{\mu}_{ggH,t\bar{t}H} = 0.80^{+0.14}_{-0.18}
\]

- **Production mechanism** signal strengths measurements **compatible with SM**

- More details on **mass**: see Ulascan Sarica's talk
  “Combined results of the 125 GeV Higgs boson on the mass, tensor structure, and couplings measured by the CMS detector”, **today@1:10pm**
**Fiducial cross section**

- **Different event categorisation:** 3 mass resolution categories

- **Fiducial cross section measured profiling** $m_H$:
  \[ \hat{\sigma}_{fid} = 69^{+16}_{-22} (\text{stat.})^{+8}_{-6} (\text{syst.}) \text{fb} \]

- **Theoretical prediction for** $m_H = 125.09$ GeV
  \[ \sigma_{fid}^{th.} = 73.8 \pm 3.8 \text{fb} \]
Summary

• CMS $H \rightarrow \gamma \gamma$ results using 12.9/fb of 13 TeV collision data collected in 2016 have been presented

• Observation ($6.1\sigma$ peak significance) of the Higgs boson in the di-photon channel

• Best fit signal strength is $\hat{\sigma}/\sigma_{SM} = 0.95^{+0.21}_{-0.19}$

• **Bosonic and fermionic** components of signal strength are observed

  $\hat{\mu}_{VBF, VH} = 1.59^{+0.73}_{-0.45}$ \hspace{1cm} $\hat{\mu}_{ggH, t\bar{t}H} = 0.80^{+0.14}_{-0.18}$

• **Fiducial cross-section** is measured to be $\hat{\sigma}_{fid} = 69^{+16}_{-22}(\text{stat.})^{+8}_{-6}(\text{syst.})$ fb
BACKUP
# Event yields per category

<table>
<thead>
<tr>
<th>Event Categories</th>
<th>SM 125GeV Higgs boson expected signal</th>
<th>Bkg (GeV⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>ggh</td>
</tr>
<tr>
<td>Untagged Tag 0</td>
<td>11.92</td>
<td>79.10%</td>
</tr>
<tr>
<td>Untagged Tag 1</td>
<td>128.78</td>
<td>85.98%</td>
</tr>
<tr>
<td>Untagged Tag 2</td>
<td>220.12</td>
<td>91.11%</td>
</tr>
<tr>
<td>Untagged Tag 3</td>
<td>258.50</td>
<td>92.35%</td>
</tr>
<tr>
<td>VBF Tag 0</td>
<td>9.35</td>
<td>29.47%</td>
</tr>
<tr>
<td>VBF Tag 1</td>
<td>15.55</td>
<td>44.91%</td>
</tr>
<tr>
<td>TTH Hadronic Tag</td>
<td>2.42</td>
<td>16.78%</td>
</tr>
<tr>
<td>TTH Leptonic Tag</td>
<td>1.12</td>
<td>1.09%</td>
</tr>
<tr>
<td>Total</td>
<td>647.77</td>
<td>87.93%</td>
</tr>
</tbody>
</table>
Event yields per category

- Measurement of coupling modifiers to vector bosons and fermions ($k_V, k_f$) and to photons and gluons ($k_\gamma, k_g$)

---

**CMS Preliminary**

12.9 fb$^{-1}$ (13 TeV)

**H$\rightarrow$γγ**

$m_H$ Profiled

- Best Fit
- 1σ
- 2σ
- SM
Analysis flowchart

- Energy regression
- Vtx ID BDT
- Vtx prob BDT
- Pho ID BDT
- Scale and res. corr.
- Dijet BDT
- DiPho BDT
- VBF BDT
- VBF cats.
- Untagged cats.
- ttH req.
- ttH cats.
Vertex BDTs inputs

- **Vertex ID BDT:**

\[
\sum_i |\vec{p}_T^i|^2 , -\sum_i (\vec{p}_T^i \cdot \frac{\vec{p}_T^{\gamma \gamma}}{|\vec{p}_T^{\gamma \gamma}|}) \text{ and } (|\sum_i \vec{p}_T^i| - p_T^{\gamma \gamma}) / (|\sum_i \vec{p}_T^i| + p_T^{\gamma \gamma})
\]

- if conversions are present

and the pull \(|z_{vtx} - z_e|/\sigma_z\) between the longitudinal position of the reconstructed vertex, \(z_{vtx}\), and the longitudinal position of the vertex estimated using conversion track(s), \(z_e\). The variable

- **Vertex probability BDT**

  - the number of vertices in each event,
  - the values of the BDT\(\text{VTX ID}\) output for the three most likely vertices in each event,
  - the distances between the chosen vertex and the second and third choices,
  - the transverse momentum of the diphoton system \(p_T^{\gamma \gamma}\),
  - the number of photons with an associated conversion track.
Preselection details

- \( p_T > 30 \) (20) GeV, \( p_T/m_{\gamma\gamma} > 1/3 \) (1/4) for (sub)leading-\( p_T \) photon

- \(|\eta| < 2.5\), removing \( 1.44 < |\eta| < 1.57 \), electron veto

- either \( R_9 > 0.8 \), or charged hadron isolation < 20 GeV, or charged hadron isolation relative to \( p_T < 0.3 \)

<table>
<thead>
<tr>
<th></th>
<th>H/E</th>
<th>( \sigma_{\eta\eta} )</th>
<th>( R_9 )</th>
<th>photon iso.</th>
<th>tracker iso.</th>
</tr>
</thead>
<tbody>
<tr>
<td>ECAL barrel; ( R_9 &gt; 0.85 )</td>
<td>&lt;0.08</td>
<td>–</td>
<td>&gt;0.5</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>ECAL barrel; ( R_9 \leq 0.85 )</td>
<td>&lt;0.08</td>
<td>&lt;0.015</td>
<td>&gt;0.5</td>
<td>&lt; 4.0</td>
<td>&lt; 6.0</td>
</tr>
<tr>
<td>ECAL endcaps; ( R_9 &gt; 0.90 )</td>
<td>&lt;0.08</td>
<td>–</td>
<td>&gt;0.8</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>ECAL endcaps; ( R_9 \leq 0.90 )</td>
<td>&lt;0.08</td>
<td>&lt;0.035</td>
<td>&gt;0.8</td>
<td>&lt; 4.0</td>
<td>&lt; 6.0</td>
</tr>
</tbody>
</table>
• shower shape observables;
• isolation variables based on the sums of the $p_T$ of photons and of charged hadrons, within regions of $R < 0.3$ around the candidate. Two such charged hadron isolations are used: one considering hadrons coming from the chosen vertex in the event, and one considering hadrons coming from the vertex with the largest $p_T$ sum among remaining vertices. The latter is effective in rejecting photon candidates originating from mis-identification of jets from a vertex other than the chosen one;
• the median energy density per unit area in the event, $\rho$, in order for the $\text{BDT}_{\gamma \text{ID}}$ to be independent of pileup;
• photon kinematic observables (pseudorapidity and energy), allowing the $\text{BDT}_{\gamma \text{ID}}$ to account for the dependence of the shower topology and isolation variables on $\eta$ and $p_T$. 
• BDT$\gamma$ ID of the lowest-scoring photon for data and simulation, after preselection
• the relative transverse momenta of both photons, $p_T^{1(2)}/m_{\gamma\gamma}$;
• the pseudorapidities of both photons, $\eta^{1(2)}$;
• the cosine of the angle between the two photons in the transverse plane, $\cos(\phi_1 - \phi_2)$;
• the relative diphoton mass resolution, under the hypothesis that the mass has been reconstructed using the correct primary vertex, $\sigma_m^{\text{right}}/m_{\gamma\gamma}$;
• the relative diphoton mass resolution, under the hypothesis that the mass has been reconstructed using an incorrect primary vertex, $\sigma_m^{\text{wrong}}/m_{\gamma\gamma}$;
• the per-event probability estimate that the correct primary vertex has been used to reconstruct the mass, taken from BDT$\gamma_{\text{TX PROB}}$;
• the BDT$\gamma_{\text{ID}}$ score for both photons.
Impact on BDT_{\gamma\gamma} of systematic uncertainties associated to:

- relative energy resolution (±5% relative shift)
- photon ID BDT (± 0.03 shift plus linearly increasing term)
- Z\rightarrow ee events with electrons reconstructed as photons
• leading photon \( p_T > m_{\gamma\gamma}/2 \);
• sub-leading photon \( p_T > m_{\gamma\gamma}/4 \);
• at least one lepton with \( p_T > 20 \) GeV: electrons must be within the ECAL fiducial region and pass the recommended criteria for loose requirements on the same observables as described in [27]. In addition the electron should satisfy \( |m(e, \gamma) - m_Z| > 10 \) GeV, where \( m_Z \) refers to the Z boson mass. Muons are required to have \( |\eta| < 2.4 \) and to pass a tight selection based on the quality of the track, the number of hits in the tracker and muon system, and the longitudinal and transverse impact parameters of the track with respect to the muon vertex; additionally, it has to satisfy a requirement on the relative isolation with pileup correction, based on the transverse momentum of the charged hadrons, transverse energy of the neutral hadrons and photons in a cone of \( R = 0.4 \) around the muon;
• all selected leptons (\( \ell \)) are required to have \( \Delta R(\ell, \gamma) > 0.4 \);
• at least two jets in the event with \( p_T > 25 \) GeV, \( |\eta| < 2.4 \), and \( \Delta R(\text{jet}, \gamma) > 0.4 \) and \( \Delta R(\text{jet}, \ell) > 0.4 \);
• at least one of the jets in the event has to be identified as b jet according to the CSV tagger medium requirement [28].
• BDT_{\gamma\gamma} output > -0.4. Too few events are available to optimise this selection for significance, so this choice is made simply to remove most of the events with low BDT_{\gamma\gamma} score.
- leading photon $p_T > m_{\gamma \gamma}/2$;
- sub-leading photon $p_T > m_{\gamma \gamma}/4$;
- no leptons defined according to the leptonic tag;
- at least five jets in the event with $p_T > 25$ GeV and $|\eta| < 2.4$;
- at least one of the jets in the event has to be identified as a b-jet according to the CSV tagger medium requirement [28];
- a minimum value of BDT$_{\gamma \gamma}$ output. The value is a compromise between significance optimisation and the need of a minimum number of events to fit the background.
Systematic uncertainties

- **Theory uncertainties** (PDFs, $\alpha_s$, QCD scale, underlying event and parton shower, $H \rightarrow \gamma\gamma$ branching fraction)
- **ggH contamination** in VBF and ttH tagged categories
- **Trigger** efficiency, **integrated luminosity**, **vertex** efficiency, **preselection**
- Non-uniformity of light collection, non-linearity, detector simulation, modeling of the material budget, shower shape corrections
- Photon energy **scale and resolution**
- **BDT$_\gamma$** and per-photon **energy resolution**
- **Jet** energy scale and smearings
- **b-tagging** efficiency, **gluon-splitting** fraction, **parton shower**, ID efficiency for **e and $\mu$**
Per category signal strength

### Preliminary CMS

<table>
<thead>
<tr>
<th>Category</th>
<th>Signal Strength</th>
<th>Combined ± 1σ</th>
<th>Per category ± 1σ</th>
<th>m_H Profiled</th>
</tr>
</thead>
<tbody>
<tr>
<td>Untagged 0</td>
<td>0.95 ± 0.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Untagged 1</td>
<td>1.04 ± 0.33</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Untagged 2</td>
<td>0.44 ± 0.32</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Untagged 3</td>
<td>1.33 ± 0.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VBF Tag 0</td>
<td>1.58 ± 0.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VBF Tag 1</td>
<td>0.04 ± 1.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TTH Hadronic Tag</td>
<td>2.10 ± 1.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TTH Leptonic Tag</td>
<td>1.15 ± 2.0</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

m_H Profiled

### Combined

$\hat{\mu}_{\text{combined}} = 0.95^{+0.21}_{-0.18}$

### 12.9 fb⁻¹ (13 TeV)

V.R. Tavolaro 04.08.2016
Fiducial cross section: mass spectra

Events / GeV

0
100
200
300
400
500
600
700

B component subtracted

$\gamma \gamma \rightarrow H \gamma \gamma$

$\sigma_{\text{fit}} / M_{\text{decor}}$

$\sigma_1 \pm \sigma_2$

$\pm 1 \sigma$

$\pm 2 \sigma$

$\mu_{\gamma \gamma} = 126.0$ GeV

$H_m$ -

Preliminary

CMS TeV)

$12.9$ fb$^{-1}$ (13 TeV)

V.R. Tavolaro

04.08.2016