Status of the CMS
SM Higgs Search

Joe Incandela
UCSB/CERN
July 4, 2012
Status of the CMS SM Higgs Search

Raw $2E_T \sim 2$ TeV
4 jets with $E_T > 40$
Estimated PU \( \sim 50 \)

Joe Incandela
UCSB/CERN
July 4, 2012
On behalf of the CMS Collaboration
The Standard Model

<table>
<thead>
<tr>
<th>Measurement</th>
<th>Fit</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Delta \alpha_{\text{had}}^{(a)}(m_Z) )</td>
<td>0.02758 ± 0.00035 0.02768</td>
</tr>
<tr>
<td>( m_Z ) [GeV]</td>
<td>91.1875 ± 0.0021 91.1874</td>
</tr>
<tr>
<td>( \Gamma_Z ) [GeV]</td>
<td>2.4952 ± 0.0023 2.4959</td>
</tr>
<tr>
<td>( s^0_{\text{had}} ) [nb]</td>
<td>41.540 ± 0.037 41.479</td>
</tr>
<tr>
<td>( R_l )</td>
<td>20.767 ± 0.025 20.742</td>
</tr>
<tr>
<td>( A_{\text{lb}}^{0,1} )</td>
<td>0.01714 ± 0.00095 0.01645</td>
</tr>
<tr>
<td>( A_l(P_{\gamma}) )</td>
<td>0.1465 ± 0.0032 0.1481</td>
</tr>
<tr>
<td>( R_b )</td>
<td>0.21629 ± 0.00066 0.21579</td>
</tr>
<tr>
<td>( R_c^{0,b} )</td>
<td>0.1721 ± 0.0030 0.1723</td>
</tr>
<tr>
<td>( A_{\text{lb}}^{0,c} )</td>
<td>0.0992 ± 0.0016 0.1038</td>
</tr>
<tr>
<td>( A_{\text{lb}}^{0,b} )</td>
<td>0.0707 ± 0.0035 0.0742</td>
</tr>
<tr>
<td>( A_b )</td>
<td>0.923 ± 0.020 0.935</td>
</tr>
<tr>
<td>( A_c )</td>
<td>0.670 ± 0.027 0.668</td>
</tr>
<tr>
<td>( A_l(\text{SLD}) )</td>
<td>0.1513 ± 0.0021 0.1481</td>
</tr>
<tr>
<td>( \sin^2 \theta_{\text{eff}}^{\text{ lept}}(Q_{\text{fb}}) )</td>
<td>0.2324 ± 0.0012 0.2314</td>
</tr>
<tr>
<td>( m_W ) [GeV]</td>
<td>80.399 ± 0.023 80.379</td>
</tr>
<tr>
<td>( \Gamma_W ) [GeV]</td>
<td>2.085 ± 0.042 2.092</td>
</tr>
<tr>
<td>( m_t ) [GeV]</td>
<td>173.3 ± 1.1 173.4</td>
</tr>
</tbody>
</table>

July 2010
The Standard Model

Confirmed to better than 1% uncertainty by 100’s of precision measurements

<table>
<thead>
<tr>
<th>Measurement</th>
<th>Fit</th>
<th>$\sigma_{\text{meas}} - \sigma_{\text{fit}} / \sigma_{\text{meas}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta \alpha_{\text{had}}^{(s)}(m_Z)$</td>
<td>0.02758 ± 0.00035</td>
<td>0.02768</td>
</tr>
<tr>
<td>$m_Z$ [GeV]</td>
<td>91.1875 ± 0.0021</td>
<td>91.1874</td>
</tr>
<tr>
<td>$\Gamma_Z$ [GeV]</td>
<td>2.4952 ± 0.0023</td>
<td>2.4959</td>
</tr>
<tr>
<td>$\sigma^0_{\text{had}}$ [nb]</td>
<td>41.540 ± 0.037</td>
<td>41.479</td>
</tr>
<tr>
<td>$R_l$</td>
<td>20.767 ± 0.025</td>
<td>20.742</td>
</tr>
<tr>
<td>$A_{l\tau}^{0,1}$</td>
<td>0.01714 ± 0.00095</td>
<td>0.01645</td>
</tr>
<tr>
<td>$A_l(P_L)$</td>
<td>0.1465 ± 0.0032</td>
<td>0.1481</td>
</tr>
</tbody>
</table>

$A_\ell$ | 0.670 ± 0.027 | 0.668 |

$A_l(SLD)$ | 0.1513 ± 0.0021 | 0.1481 |

$\sin^2 \theta_{\ell\tau}^{\text{eff}}(Q_{l\tau})$ | 0.2324 ± 0.0012 | 0.2314 |

$m_W$ [GeV] | 80.399 ± 0.023 | 80.379 |

$\Gamma_W$ [GeV] | 2.085 ± 0.042 | 2.092 |

$m_t$ [GeV] | 173.3 ± 1.1 | 173.4 |

July 2010
The Standard Model

1 Missing piece: Higgs

Confirmed to better than 1% uncertainty by 100’s of precision measurements

- \( \Delta_{\text{had}}(m_Z) \) 
- \( m_Z \) [GeV] 
- \( \Gamma_Z \) [GeV] 
- \( A_{\text{had}}^0 \) [nb] 
- \( R_l \) 
- \( A_{l,b}^{0,1} \) 
- \( A_l(P_L) \) 

\[
\begin{align*}
\Delta_{\text{had}}(m_Z) &= 0.02758 \pm 0.00035 \quad 0.02768 \\
m_Z &= 91.1875 \pm 0.0021 \quad 91.1874 \\
\Gamma_Z &= 2.4952 \pm 0.0023 \quad 2.4959 \\
A_{\text{had}}^0 &= 41.540 \pm 0.037 \quad 41.479 \\
R_l &= 20.767 \pm 0.025 \quad 20.742 \\
A_{l,b}^{0,1} &= 0.01714 \pm 0.00095 \quad 0.01645 \\
A_l(P_L) &= 0.1465 \pm 0.0032 \quad 0.1481 \\
\end{align*}
\]
The Standard Model

1 Missing piece: Higgs

Confirmed to better than 1% uncertainty by 100’s of precision measurements

<table>
<thead>
<tr>
<th>Measurement</th>
<th>Fit</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta\alpha_{\text{had}}(m_Z)$</td>
<td>$0.02758 \pm 0.00035$</td>
</tr>
<tr>
<td>$m_Z$ [GeV]</td>
<td>$91.1875 \pm 0.0021$</td>
</tr>
<tr>
<td>$\Gamma_Z$ [GeV]</td>
<td>$2.4952 \pm 0.0023$</td>
</tr>
<tr>
<td>$\sigma_0^0$ [nb]</td>
<td>$41.540 \pm 0.037$</td>
</tr>
<tr>
<td>$R_i$</td>
<td>$20.767 \pm 0.025$</td>
</tr>
<tr>
<td>$A_{0,l}^{0,l}$</td>
<td>$0.1714 \pm 0.00095$</td>
</tr>
<tr>
<td>$A_l(P_L)$</td>
<td>$0.1465 \pm 0.0032$</td>
</tr>
</tbody>
</table>

$A_0$ 0.670 ± 0.027 0.668
$A_l(SLD)$ 0.1513 ± 0.0021 0.1481
$\sin^2\theta^{\text{eff}}_{Q_l}(Q_{lf})$ 0.2324 ± 0.0012 0.2314
$m_W$ [GeV] 80.399 ± 0.023 80.379
$\Gamma_W$ [GeV] 2.085 ± 0.042 2.092
$m_t$ [GeV] 173.3 ± 1.1 173.4

July 2010
Where we stood last week

1. **$M_{\text{top}} \text{ vs. } M_W$**
   - **Tevatron $M_W$ Tour de Force!!**
     - $m_W = 80385 \pm 15 \text{ MeV (World Ave – Mar 2012)}$
   - Shifts for SM Higgs expectation

2. **Colliders leave little space**

   $W \rightarrow t \rightarrow W \bar{b}$
   $\sim M_t^2$

   $W \rightarrow H \rightarrow W W$
   $\sim \ln(m_H)$

**This is the main story of the past year**

**Eliminated $\sim 475 \text{ GeV of the mass range.}$**
- $\sqrt{s}=8$ TeV: 25-30% higher $\sigma$ than $\sqrt{s}=7$ TeV at low $m_H$
- All production modes to be exploited
  - $gg$ VBF, VH, $ttH$
  - Latter 3 have smaller cross sections but better S/B in many cases
- $\sqrt{s}=8$ TeV: 25-30% higher $\sigma$ than $\sqrt{s}=7$ TeV at low $m_H$
- All production modes to be exploited
  - $gg$ VBF VH ttH
- Latter 3 have smaller cross sections but better S/B in many cases
5 decay modes exploited

- High mass: $WW, ZZ$
- Low mass: $b\bar{b}, \tau\tau, WW, ZZ, \gamma\gamma$
- Low mass region is very rich but also very challenging:
  main decay modes ($b\bar{b}, \tau\tau$) are hard to identify in the huge background
- Very good mass resolution (1%): $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ \rightarrow 4l$
- Not-yet-excluded region: 
  ~[115-130] GeV
- The five decay modes discussed today have comparable sensitivities for exclusion.
- Most analyses used in this combination have been re-optimized. In order to avoid the possibility of an unintended bias, all selection criteria in the analyses of the 2011 and 2012 data were fixed before looking at the result in the signal region.
p-values

- Probability that background fluctuates to give an excess as large as the (average) signal size expected for a SM Higgs.
  - Takes into account all analysis steps, estimated backgrounds, etc. for the 5 search channels indicated.
- Excellent prospects for exploring properties
How is it possible to go so far so fast?

LHC performance: 2010-2011-2012

Stellar performance of the LHC enables all experiments to produce significant physics results.

Many thanks to the LHC teams and the many others who made this possible!
Total weight 14,000 t
Overall diameter 15 m
Overall length 28.7 m

Pixels & Tracker
- Pixels (100x150 µm²)
  ~ 1 m² ~66M ch
- Si Strips (80-180 µm)
  ~200 m² ~9.6M ch

ECAL
- 76k scintillating PbWO₄ crystals

HCAL
- Scintillator/brass Interleaved ~7k ch

Iron Yoke

Preshower
- Si Strips ~16 m²
  ~137k ch

Forward Cal
- Steel + quartz Fibers ~k ch

Muons

Solenoid coil

CMS

MUON ENDCAPS
- 473 Cathode Strip Chambers (CSC)
- 432 Resistive Plate Chambers (RPC)

MUON BARREL
- 250 Drift Tubes (DT) and
  480 Resistive Plate Chambers (RPC)
Current Operational Status*

- Pixels: 97.1%
- Strips: 97.75%
- Preshower: 97.1%
- ECAL Barrel: 99.16%
- ECAL Endcaps: 98.54%
- HCAL Barrel: 99.92%
- HCAL Endcaps: 99.96%
- HCAL Forward: 99.88%
- HCAL Outer: 96.88%
- Muon DT: 99.1%
- Muon CSC: 97.67%
- Muon RPC: 98.2%

*As of June 15 2012
**ECAL calibration, 2012 data**

- **W→ev E/p:** Stable E scale during 2012 run after light monitoring (LM) corrections:
  - ECAL Barrel (EB): RMS stability after corrections 0.19%

- **Z→ee:** Good resolution with preliminary energy calibration for 2012:
  - Instrumental resolution: 1.0 GeV in ECAL Barrel
Last Autumn
- cpu time for high PU >40 sec/event
- Memory usage well above 2 GB.
  - Means we cannot use all the cores!
- Even 200 Hz looked hard!

Task force started December
- Major success!

Improvements
- A factor 2.5 in speed
  - Under ~15” per event on average
- Much reduced memory use
  - Well under 2 GB

Physics performance unchanged
- Kept our AAA rating:
  - E.g. no explicit $p_T$ threshold on tracks

Prompt Reconstruction at Tier-0:
*Limit on our data-taking rate versus event processing time for low and high memory use cases*
- Sustain 400M/Month
  - 900M, 600M past 2 months but had help from Tier-1’s

- Sustain 1B/month
  - Peaks high as 2.3B
Reconstruction
Optimal combination of information from all subdetectors

- Returns a list of reconstructed particles
  - $e, \mu, \gamma$, charged and neutral hadrons
    - Used in the analysis as if it came from a list of generated particles
    - Used as building blocks for jets, taus, missing transverse energy, isolation and PU particle identification

Made possible by CMS granularity and high magnetic field
- Cluster reconstruction in ECAL
  - Common for both electrons and photons (Electrons also reconstructed as photons)
  - Designed to collect bremsstrahlung and conversions in extended phi region
- Dedicated track reconstruction for electrons
  - Gaussian Sum Filter allows for tracks with large bremsstrahlung
- Photon identification specific to $H \rightarrow \gamma\gamma$
- Energy scale and resolution
  - Extensive control with $Z \rightarrow e^+e^-$ and $J/\psi \rightarrow e^+e^-$ for both electrons and photons
- Multivariate electron identification in 2012
  - ECAL, tracker, ECAL-tracker-HCAL matching, impact parameter
  - 30% efficiency improvement in $H \rightarrow ZZ \rightarrow 4e$ wrt cut based ID
- Multivariate training against background in data
- Start with particle flow muons
- Efficiency above 96% down to $p_T = 5$ GeV
  - Above 99% efficiency for $p_T > 10$ GeV
- Efficiency in data using $J/\Psi$ and $Z$ peak

Tighter quality criteria applied in some analyses
Particle-based isolation

Sum energy of particles in $\Delta R$ cone around the lepton

- Global event description eliminates double counting

Efficiency is stable in high PU environment

Pile-up contribution:
- Negligible for charged hadrons (vertexing)
- Neutrals corrected with global energy density ($\rho$)

Detector vs Particle Flow Isolation

Detector Isolation

Particle Flow Isolation

Endcaps $p_T < 10$ GeV
- **Tau identification:**
  - Reconstruct individual decay modes
  - Charged hadrons + electromagnetic obj
    - EM strips account for material effects

- **Tau isolation:**
  - Multivariate discriminator using sum of energy deposits in dR rings around the tau

---

**Real Tau Charged Isolation**

**Fake Tau Charged Isolation**

**CMS Simulation, s=7 TeV**

**Z\rightarrow\tau\tau MC**

**Gen |\eta| < 2.3**

- HPS Loose Comb d\beta
- HPS Med Comb d\beta
- HPS MVA Loose

---

Real taus

Fake taus
Jet reconstruction

- Reconstruction with particle flow objects

- Pileup jets structure differs wrt regular jets:
  - Pileup jets originate from several overlapping jets which merge together
  - Likelihood grows rapidly with high pileup

- Discriminant exploits shape and tracking variables
  - discrimination both inside and outside tracker acceptance
Standard Model: Precision Jets, W, and γ*/Z

Inclusive jet and dijets. 2-4% JES. Constrains gluon PDF up to x=0.6

Differential Drell-Yan cross section: 2.5M µµ pairs tests NNLO cross sections and PDFs

CMS-PAS-QCD-11-004

CMS-PAS-EWK-11-007
**Standard Model at 7 TeV 2010-2011**

**CMS Preliminary, √s=7 TeV**

- CMS e/μ+jets+btag: 164 ± 3 ± 12 ± 7 (val ± stat. ± syst. ± lum)
- CMS dilepton (ee,μμ,μe): 170 ± 4 ± 16 ± 8 (val ± stat. ± syst. ± lum)
- CMS all-hadronic: 136 ± 20 ± 40 ± 8 (val ± stat. ± syst. ± lum)
- CMS dilepton (μτ): 149 ± 24 ± 26 ± 9 (val ± stat. ± syst. ± lum)
- CMS 2010 combination: 154 ± 17 ± 6 (val ± tot. ± lum.)
- CMS e/μ+jets+btag: 150 ± 9 ± 17 ± 6 (val ± stat. ± syst. ± lum)
- CMS dilepton (ee,μμ,μe): 168 ± 18 ± 14 ± 7 (val ± stat. ± syst. ± lum)
- CMS e/μ+jets: 173 ± 14 ± 36 ± 29 ± 7 (val ± stat. ± syst. ± lum)

**Production Cross Section, σ_{tot} [pb]**

- CMS 95%CL limit
- CMS measurement (stat+syst)
- theory prediction

- **W**
  - E_{T} > 30 GeV
  - |η_{JET}| < 2.4
  - ΔR(γ,γ) > 0.7
  - CMS 95%CL: 36 pb⁻¹

- **Z**
  - CMS 95%CL: 36 pb⁻¹

- **H(127) → ZZ**
  - CMS 95%CL: 1.1 fb⁻¹
  - CMS 95%CL: 4.7 fb⁻¹

- **CMS Preliminary, √s=7 TeV**

- **CMS e/μ+jets+btag**
  - TOP-11-003 (L=0.8-1.09/pb)
  - CMS Preliminary, √s=7 TeV

- **CMS dilepton (ee,μμ,μe)**
  - TOP-11-005 (L=1.14/fb)
  - CMS Preliminary, √s=7 TeV

- **CMS all-hadronic**
  - TOP-11-007 (L=1.09/fb)
  - CMS Preliminary, √s=7 TeV

- **CMS dilepton (μτ)**
  - TOP-11-006 (L=1.09/fb)
  - CMS Preliminary, √s=7 TeV

- **CMS 2010 combination**
  - arXiv:1108.3773 (L=36/pb)
  - CMS Preliminary, √s=7 TeV

- **CMS e/μ+jets+btag**
  - arXiv:1108.3773 (L=36/pb)
  - CMS Preliminary, √s=7 TeV

- **CMS dilepton (ee,μμ,μe)**
  - arXiv:1105.5661 (L=36/pb)
  - CMS Preliminary, √s=7 TeV

- **CMS e/μ+jets**
  - arXiv:1106.0902 (L=36/pb)
  - CMS Preliminary, √s=7 TeV

**Theory:** Langenfeld, Moch, Uwer, Phys. Rev. D80 (2009) 054009
MSTW2008(N)NLO PDF, scale ⊗ PDF(90% C.L.) uncertainty

**Fabulous agreement**

**Lots of data**

... on to the Higgs...
$H \rightarrow XX$
Main analysis is a Multi-Variate-Analysis (MVA)
  - MVAs for photon ID and event classification
    - Fit mass distribution in 4 event classes based on a diphoton MVA output + 2 di-jet categories
  - Improvement in expected limit ~15% over cut-based analysis
  - Cross-checked with an alternative background model extraction:
    - Fit output of a 2\textsuperscript{nd} MVA combining diphoton MVA and $m_{\gamma\gamma}$ using data in mass sidebands to construct the background model

Also cross-checked with a cut based analysis
  - Simple and robust
    - Cut based photon ID and event classification
      - Fit data mass distribution in 2 rapidity x 2 shower shape =4 categories with different Signal over Background (S/B) + 2 di-jet categories

Published for 2011 data
Search for a narrow mass peak with two isolated high Et photons

- Blind analysis in 2012
- Re-reco 2011 data into unchanged 2011 analysis
- Background MC only used for analysis optimization, Z->ee also to measure photon efficiencies and resolution with data
- ECAL cluster energies corrected using a MC trained multivariate regression
  - Improves resolution and restores flat response of energy scale versus pileup
    - Inputs: Raw cluster energies and positions, lateral and longitudinal shower shape variables, local shower positions w.r.t. crystal geometry, pileup estimators
- Regression also used to provide a per photon energy resolution estimate
- To measure the Energy Scale and resolution: use $Z \rightarrow e^+e^-$
- **Photon pre-selection:**
  - $E_{T\gamma_1}/m_{\gamma\gamma}>3$, $E_{T\gamma_2}/m_{\gamma\gamma}>4$
  - Photon Id a bit tighter than trigger selection and MC EM enrichment filters
    - Efficiency measured using tag and probe with $Z\rightarrow ee$
  - Electron veto: Efficiency measured using tag and probe with $Z\rightarrow \mu\mu\gamma$

- **MVA based photon ID discriminates photons from fakes:**
  - Inputs: isolation, shower shape, per event energy density, pseudorapidity

---

![Graphs showing validation with $Z\rightarrow ee$](image)

**Validation with $Z\rightarrow ee$ (inverted electron veto)**
The $\gamma\gamma$ Vertex Choice

- Mass reconstruction
  - Depends on the correct position of the primary vertex
- Interaction vertex is identified using tracks from recoiling jets and underlying event plus conversions
  - correct in ~83% of cases for pileup in 2011 sample.
  - correct in ~80% of cases for pileup in 2012 sample.
- Vertex identification with a BDT
  - Input variables: $\Sigma p_t^2$, $\Sigma p_t$ projected onto the $\gamma\gamma$ transverse direction, $p_t$ asymmetry and conversions
  - Correct vertex finding probability also estimated using a BDT

Data-MC efficiency for $Z\rightarrow\mu\mu$
After removing the $\mu$ tracks

Efficiency to identify correct vertex
Diphoton MVA trained on signal and background MC with input variables largely independent of $m_{\gamma\gamma}$:

- Kinematics: $p_T$ and $\eta$ of each photon, and $\cos \Delta \phi$ between the 2 photons
- Photon ID MVA output for each photon
- per-event mass resolution and vertex probability
- Encode all relevant information on signal vs background discrimination (aside from $m_{\gamma\gamma}$ itself) into a single di-photon MVA output to first order independent of $m_{\gamma\gamma}$

Residual data-MC disagreement:

- For BG only make analysis sub-optimal
- For signal would cause some category migration included in the systematic errors
Di-jet Tagging

- Exclusive selection of di-photon events with VBF-like topology:
  - Two high \( p_T \) jets with large pseudo-rapidity difference and invariant mass
- High S/B
- \(~80\%\)-pure VBF events for large di-jet invariant masses

**Di-jet event with:**
- diphoton mass 121.9 GeV
- dijet mass 1460 GeV
- jet \( p_T \): 288.8 and 189.1 GeV
- jet \( \eta \): -2.022 and 1.860
Analysis improvements in 2012:

- Split di-jet tagged events in two categories based on $M_{jj}$ and jet $p_T$
  - ~15% improvement in sensitivity for dijet category
  - better sensitivity to separate different Higgs production modes
- Removal of jets from pileup events
  - Based on the jet shape variables, tracks in jet and vertexing
  - Cross-checked using $Z$+jet and $\gamma$+jet events

**Dijet selection cuts**

<table>
<thead>
<tr>
<th>Variable</th>
<th>2011</th>
<th>2012</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Loose</td>
<td>Tight</td>
</tr>
<tr>
<td>$p_T(j_1)$</td>
<td>$&gt; 30$ GeV</td>
<td></td>
</tr>
<tr>
<td>$p_T(j_2)$</td>
<td>$&gt; 20$ GeV</td>
<td>$&gt; 30$ GeV</td>
</tr>
<tr>
<td>$\Delta\eta(j_1,j_2)$</td>
<td>$&gt; 3.5$</td>
<td>$&gt; 3.0$</td>
</tr>
<tr>
<td>$</td>
<td>\eta_{\gamma\gamma} - \frac{1}{2}(\eta_{j1} + \eta_{j2})</td>
<td>$</td>
</tr>
<tr>
<td>$\Delta\phi(jj,\gamma\gamma)$</td>
<td></td>
<td>$&gt; 2.6$</td>
</tr>
<tr>
<td>$m_{jj}$</td>
<td>$&gt; 350$ GeV</td>
<td>$&gt; 250$ GeV</td>
</tr>
</tbody>
</table>
7 TeV Mass Distribution in Categories

- Background model is entirely from data.
- Fit to mass distribution in each category with polynomial functions (3rd to 5th degree)
  - keep bias below 20% of fit error.
  - causes some loss of performance due to number of parameters in fit function.
8 TeV Mass Distribution in Categories

- **Untagged 0**
- **Untagged 1**
- **Untagged 2**
- **Untagged 3**
- **Dijet tight**
- **Dijet loose**
- Sum of mass distributions for each event class, weighted by S/B
- \( B \) is integral of background model over a constant signal fraction interval
- Expected 95% CL exclusion 0.76 times SM at 125 GeV
- Large range with expected exclusion below $\sigma_{SM}$
- Largest excess at 125 GeV
Minimum local p-value at 125 GeV with a local significance of 4.1 σ
Similar excess in 2011 and 2012
Independent cross check analyses give similar results
Global significance in the full search range (110-150 GeV) 3.2 σ
Combined best fit signal strength
\( \sigma/\sigma_{SM} = 1.56 \pm 0.43 \times SM \), consistent with SM.

Best fit signal strength consistent between different classes
H $\rightarrow$ ZZ$^*$
H \rightarrow ZZ^{(*)} \rightarrow 4l \; (l = e, \mu): \text{the golden channel}

Clean signature: narrow peak, low background

Background: irreducible ZZ^{(*)}; reducible Z+jets, ttbar, WZ

One of the best performing channels in the whole mass range ...

... but extremely demanding channel for selection, requiring the highest possible efficiencies (lepton Reco/ID/Isolation).

Production cross sections

\[ \sigma(pp \rightarrow H \rightarrow ZZ^{(*)} \rightarrow 4l) \; [fb] \]

H \rightarrow l^+l^-l^+l^- \; (l = e, \mu)

H \rightarrow e^+e^-\mu^+\mu^- \n
H \rightarrow e^+e^-e^+e^- \n
Small branching ratio

m_H [GeV]

p_T of the 4 leptons

Low p_T leptons

p_T of the 4 leptons

H \rightarrow ZZ^{(*)} \rightarrow 4\mu

m_H = 126 \text{ GeV}

Before the selection

After the selection

p_T^1

p_T^2

p_T^3

p_T^4
Blinding policy: analysis optimized blindly for 2012, applied to 2011 reoptimization

Do NOT look at $110 < m_{4l} < 140$ GeV, and $m_{4l} > 300$ GeV

Main changes:

- New lepton ID (MVA + PFlow)
- New lepton PFlow isolation
- Final State Radiation (FSR) recovery
- 2D analysis: $m_{4l}$ + Kinematic Discriminant

>20% improvement @ $m_H = 126$ GeV wrt 2011 analysis

Expected exclusion range 121–540 GeV
Final State Radiation recovery algorithm

- Applied on each Z for photons near the leptons

\[\Delta R(\ell, \gamma)_{\text{min}} < 0.5\]

- Associates photon with Z if:
  - \(M(ll+\gamma) < 100\) GeV
  - \(|M(ll+\gamma)-M_Z| < |M(ll)-M_Z|\)
- Removes associated photons from lepton isolation calculation

Expected Performance for \(M_H = 126\) GeV

- 6% of events affected
- Average purity of 80%
- 2% added in analysis

Particle Flow ID

- \(E_T > 2\) GeV
- \(|\eta| < 2.4\)

Isolation

CMS Simulation, \(\sqrt{s} = 8\) TeV

Events affected by FSR

\(M_H = 126\) GeV
- Irreducible background $ZZ \to 4l$
  - Estimated using simulation
  - Phenomenological shape models
  - Corrected for data/simulation scale

- Reducible backgrounds estimated from data
  - Extrapolation from control samples enriched with misidentified leptons
  - Total uncertainty ~50%

Validation in data

- Control sample
- Signal Region (WFC)
  - Irreducible (sim.): 2.6
  - Reducible (data): 11.3
  - Observed: 12
Matrix Element Likelihood Analysis: uses kinematic inputs for signal to background discrimination

\[
\{ m_1, m_2, \theta_1, \theta_2, \theta^*, \Phi, \Phi_1 \}
\]

\[
\text{MELA} = \left[ 1 + \frac{P_{\text{bkg}}(m_1, m_2, \theta_1, \theta_2, \Phi, \theta^*, \Phi_1 | m_{4\ell})}{P_{\text{sig}}(m_1, m_2, \theta_1, \theta_2, \Phi, \theta^*, \Phi_1 | m_{4\ell})} \right]^{-1}
\]
2D analysis using \( \{m_{4l}, \text{MELA}\} \)

MELA offers powerful discrimination of background technique applicable for signal hypothesis testing

CMS simulation Higgs, 126 GeV
8 TeV DATA

4-lepton Mass: 126.9 GeV

$\mu^-(Z_1) p_T : 24$ GeV

$\mu^+(Z_1) p_T : 43$ GeV

$e^-(Z_2) p_T : 10$ GeV

$e^+(Z_2) p_T : 21$ GeV

$\mu^-(Z_1)$

$\mu^+(Z_1)$

$e^-(Z_2)$

$e^+(Z_2)$
July 4th 2012  The Status of the Higgs Search

J. Incandela for the CMS COLLABORATION

7 TeV DATA

4μ+γ Mass : 126.1 GeV

γ(Z₁) Eₜ : 8 GeV

μ⁻(Z₁) pₜ : 28 GeV

μ⁻(Z₂) pₜ : 14 GeV

μ⁺(Z₁) pₜ : 67 GeV

μ⁺(Z₂) pₜ : 6 GeV
4-lepton Mass: 125.8 GeV

μ⁺(Z₂) p_T: 12 GeV
μ⁻(Z₂) p_T: 15 GeV

e⁺(Z₁) p_T: 28 GeV
e⁻(Z₁) p_T: 14 GeV

7 TeV DATA
Results: $m(4\ell)$ spectrum

Yields for $m(4\ell)=110..160$ GeV

<table>
<thead>
<tr>
<th>Channel</th>
<th>4e</th>
<th>4$\mu$</th>
<th>2$e2\mu$</th>
<th>4$\ell$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ZZ background</td>
<td>2.65 $\pm$ 0.31</td>
<td>5.65 $\pm$ 0.59</td>
<td>7.17 $\pm$ 0.76</td>
<td>15.48 $\pm$ 1.01</td>
</tr>
<tr>
<td>Z+X</td>
<td>1.20$_{-0.78}^{+1.08}$</td>
<td>0.92$_{-0.55}^{+0.65}$</td>
<td>2.29$_{-1.30}^{+1.81}$</td>
<td>4.41$_{-1.60}^{+2.21}$</td>
</tr>
<tr>
<td>All backgrounds</td>
<td>3.85$_{-0.84}^{+1.12}$</td>
<td>6.58$_{-0.81}^{+0.88}$</td>
<td>9.46$_{-1.56}^{+1.96}$</td>
<td>19.88$_{-1.95}^{+2.33}$</td>
</tr>
<tr>
<td>$m_H = 126$ GeV</td>
<td>1.51 $\pm$ 0.48</td>
<td>2.99 $\pm$ 0.60</td>
<td>3.81 $\pm$ 0.89</td>
<td>8.31 $\pm$ 1.18</td>
</tr>
</tbody>
</table>

164 events expected in [100, 800 GeV]
172 events observed in [100, 800 GeV]

Event-by-event errors
Perform 2D fit

- MELA discriminant versus $m_{4l}$
  - Data points shown with per-event mass uncertainties
Perform 2D fit

- **MELA discriminant versus $m_{4l}$**
  - Data points shown with per-event mass uncertainties

Results: MELA 2D plots

Data w.r.t 126 GeV Higgs Expectation
Two-lepton invariant mass plots

Grey – is simulation (expectation) for Higgs (126 GeV)
- Enrich the signal content
  - Cut: MELA > 0.5
    - Cut value chosen such that signal probability > background probability
Limits and p-values

Expected exclusion at 95% CL:
- 121-550 GeV
- Observed exclusion at 95% CL:
  - 131-162 GeV and 172-530 GeV

Expected significance at 125.5 GeV:
- 3.8 σ
- Observed significance at 125.5 GeV:
  - 3.2 σ
Expected exclusion at 95% CL:

- Observed 131-162 GeV and 172-530 GeV
- Expected 121-550 GeV

Expected significance at 125.5 GeV:

- Observed 3.2 σ
- Expected 3.8 σ
Characterization of excess near 125 GeV

- high sensitivity, high mass resolution channels: $\gamma\gamma + 4l$
  - $\gamma\gamma$: 4.1 $\sigma$ excess
  - 4 leptons: 3.2 $\sigma$ excess
  - near the same mass 125 GeV
  - comb. significance: 5.0 $\sigma$

- expected significance for SM Higgs: 4.7 $\sigma$
$H \rightarrow WW \rightarrow l\nu l\nu$
$H \rightarrow WW \rightarrow l\nu l\nu$ Signature

- **Signature:**
  - 2 high $p_T$ leptons
  - Large missing $E_T$

- **Main backgrounds:**
  - WW, top

- **Other backgrounds:**
  - $W\text{+jet}$, $Z/\gamma^*$, $WZ$, $ZZ$, $W\gamma$

- **qq $\rightarrow$ WW + gg $\rightarrow$ WW**
  - Non-resonant

- **H $\rightarrow$ WW**
  - Large BR
  - Small $\Delta\phi(ll)$

- $\mu \ p_T \ 32 \text{ GeV}$
- $e \ p_T \ 34 \text{ GeV}$
- $ME_T \ 47 \text{ GeV}$
Data-driven background estimation

- **W+jets**
  Fake rate measured in QCD enriched data sample
- **Z/γ***
  Normalised in Z mass
- **Top**
  b-tagging efficiency measured in top control region in data

Split in categories

- 0/1-jet and VBF
- Final state lepton flavors

Cut-based approach for the first 2012 result
Kinematics at Final Selection

One step before the final selection (no cuts on $\Delta\phi(ll)$ and $m(ll)$)

Final selection on $m(ll)$ (all other selection applied)
Combination of 7 TeV + 8 TeV

- Combined limits from 2011 and 2012
  - 7 TeV result using a multivariate discriminant and updated with the final luminosity measurement
  - 2012 → 5% improvement in sensitivity coming from new object definitions and selection optimized for 2012 condition

Exclusion range
expected: 123-450 GeV
observed: 129-520 GeV
Combination of 7 TeV + 8 TeV

- **Combined limits from 2011 and 2012**
  - 7 TeV result using a multivariate discriminant and updated with the final luminosity measurement
  - 2012 → 5% improvement in sensitivity coming from new object definitions and selection optimized for 2012 condition

Exclusion range

**expected:** 123-450 GeV

**observed:** 129-520 GeV
Signal injection using prediction at 5.1/ fb 8 TeV
- Average background prediction
- Signal injection for $m_H = 125$ GeV with toys
Signal injection using prediction at 5.1/fb 8 TeV
- Average background prediction
- Signal injection for $m_H = 125$ GeV with toys
VH $\rightarrow Vb\bar{b}$

$V \rightarrow l\nu, ll, \nu\nu$
- Characteristics and importance
  - By far, largest BR for $m_H < 130$ GeV
  - Key piece of the observation puzzle
    - Tests specific production & decay couplings
  - But $\sigma_{bb}(\text{QCD}) \sim 10^7 \sigma \times \text{BR}(H \rightarrow bb)$!

$\Rightarrow$ Search in associated production with $W$ or $Z$
Analysis Strategy

Associated Production => final states with leptons, MET and b-jets

Reducible Backgrounds: QCD, top, W/Z+ light jets

Less reducible: V+bb, ZZ(bb), WZ(bb)

Boosted vector bosons: 
- $p_T(V)$, 
- 2 b-tagged jets (H->bb) 
Back-to-back V and H, reconstruct $m_{bb}$

Main backgrounds estimated from data in control regions

5 channels
- $Z(\ell\ell)H(bb)$
- $Z(\nu\nu)H(bb)$
- $W(l\nu)H(bb)$
- Here 5 fb-1 @ 7 TeV (2011) + 5 fb-1 @ 8 TeV (2012)
  - Improvements OF ~ 50% in sensitivity:
  - Two Pt(V) bins: “low” and “high” – see backup
  - Fit the shape of the MVA output distribution (vs cut and count)
  - Improved b-jet energy resolution [MVA regression]

![Event selection](image.png)

**M_{bb^-} = 128 GeV**

**p_T(bb) = 181 GeV**

**ZH->\mu\mu bb candidate**
Examples of final MVA distributions

• Observed limits:
  • Compatible with either background or signal from a 125 GeV Higgs
Characterization of excess near 125 GeV

adding high sensitivity, but low mass resolution WW

comb. significance: \(5.1 \sigma\)

expected significance for SM Higgs: \(5.2 \sigma\)
$H \rightarrow \tau\tau \rightarrow \mu\tau_{h'}, e\tau_{h'}, e\mu, \mu\mu$
### Characteristics
- High $\sigma \cdot $BR at low mass
- Sensitive to all production modes
- Probes coupling to leptons
- Enhanced $\sigma \times $BR in MSSM
- Challenging large backgrounds:
  - DY→$\tau\tau$, W+Jets, QCD
• Search performed in 4 tau-pair final states: $\mu \tau_h$, $e \tau_h$, $e\mu$, $\mu\mu$
• Analysis divided into 5 categories: mass resolution, S/B
• All categories are fit simultaneously

Jets $p_T > 30$ GeV

- $\tau_h$ or $\mu p_T$:
  - 0 Jet, Low $p_T$:
    - High background
  - 1 Jet, Low $p_T$:
    - Enhancement from jet requirement
  - 1 Jet, High $p_T$:
    - Enhancement from $p_T$ and jet requirement
  - VBF:
    - 2 jets, no jets in rapidity gap
      - MVA based selection

$categorization based on$ $\tau_h p_T$ for $\mu \tau_h$, $e \tau_h$; $\mu p_T$ for $e\mu$; leading $\mu p_T$ for $\mu\mu$
Full $m(\tau\tau)$ Reconstruction

- **SVFit**
  - Event-by-event estimator of true $m(\tau\tau)$ likelihood
    - Matrix Element used for $\tau \rightarrow l\nu\nu$
    - Phase-Space is used for $\tau \rightarrow \pi$
    - Nuisance parameters are integrated out

- Mass peaks at true value
  - 20% improved resolution
    - With respect to 2011
  - Better separation of H from Z
Mass Distributions in Event Categories

- Constrains energy scales and efficiencies
  - Large Drell-Yan background
  - Sensitivity boosted by low/high $p_T$ split

- Enhanced sensitivity to gluon fusion
  - Improved mass resolution
  - Increased sensitivity by splitting into low and high $p_T$ categories

- 1 Jet

- VBF
  - Enhanced sensitivity to VBF production
    - Highest sensitivity for $m_H < 130$ GeV
- ~2x improvement in sensitivity in 2011 data alone
- => 70% improvement in sensitivity on the same data
- 40% improvement with the additional luminosity
- No significant departure from SM background-only expectation
- Observed limit of $1.06 \times \text{SM}$ at $m_H = 125 \text{ GeV}$
- ~2x improvement in sensitivity in 2011 data alone
  - => 70% improvement in sensitivity on the same data
  - 40% improvement with the additional luminosity
- No significant departure from SM background-only expectation
  - Observed limit of $1.06 \times \text{SM}$ at $m_H = 125 \text{ GeV}$
- ~2x improvement in sensitivity in 2011 data alone
  - => 70% improvement in sensitivity on the same data
  - 40% improvement with the additional luminosity
- No significant departure from SM background-only expectation
  - Observed limit of $1.06 \times \text{SM}$ at $m_H = 125$ GeV
Full result
**SM Higgs exclusion: confidence level**

*Expected* in absence of SM Higgs boson:

- 110 – 600 GeV at 95% CL
- 110 – 580 GeV at 99% CL
- 110 – 520 GeV at 99.9% CL
**SM Higgs exclusion: confidence level**

**Observed:**
- Observed: 110 – 122.5
- 110—112 .. 113 – 121.5

**Expected:**
- 127 – 600 GeV at 95% CL
- 128 – 600 GeV at 99% CL
Observe:  \(110 - 122.5\) \(\text{GeV}\) at \(95\%\) CL
Characterization of excess near 125 GeV

- high sensitivity, high mass resolution channels: $\gamma\gamma + 4l$
  - $\gamma\gamma$: 4.1 $\sigma$ excess
  - 4 leptons: 3.2 $\sigma$ excess
  - near the same mass 125 GeV
- comb. significance: 5.0 $\sigma$
- expected significance for SM Higgs: 4.7 $\sigma$
Characterization of excess near 125 GeV

adding high sensitivity, but low mass resolution WW

comb. significance: $5.1 \sigma$

expected significance for SM Higgs: $5.2 \sigma$
Characterization of excess near 125 GeV

- all channels together: comb. significance: $4.9 \sigma$

- expected significance for SM Higgs: $5.9 \sigma$
Characterization of excess near 125 GeV

- Observed significance: $4.9 \sigma$
- Excess seen in both
  - 7 TeV data ($3.0 \sigma$)
  - 8 TeV data ($3.8 \sigma$)
- near the same mass 125 GeV
Characterization of the excess: mass

- Likelihood scan for mass and signal strength in three high mass resolution channels
- Results are self-consistent and can be combined
Characterization of the excess: mass

To reduce model dependence, allow for free cross sections in three channels and fit for the common mass:

$$m_X = 125.3 \pm 0.6 \text{ GeV}$$
- We have observed a state decaying to di-photon and four-lepton final state with statistical significance of $5 \sigma$

- The observed state has mass near $125.3 \pm 0.6$ GeV

- Next we look at the extent to which the observed state is compatible, within the current uncertainties, with the SM Higgs boson
Compatibility with SM Higgs boson

Signal strength

- Overall best-fit signal strength in the combination:
  \( \sigma / \sigma_{SM} = 0.80 \pm 0.22 \)

- Signal strength in 7 and 8 TeV data are self-consistent
Event yields in different production times decay modes are self-consistent

- albeit many modes have not yet reached sensitivity to distinguish SM from Background
Event yields in different decay modes are self-consistent
Event yields in different production topologies are self-consistent
The measurement of the $H \rightarrow WW/H \rightarrow ZZ$ ratio is mostly driven by the ratio of the Higgs couplings to WW and ZZ, which is protected by custodial symmetry.

Combination of “inclusive” WW and ZZ yields gives

$$R_{WW/ZZ} = 0.9^{+1.1}_{-0.6}$$
Group the Higgs couplings into “Vectorial” and “Fermionic” sets.

Attach a modifier to the SM prediction to each of those ($C_V$ and $C_F$).

Use LO theoretical prediction for loop-induced $H \rightarrow \gamma\gamma$, $H \rightarrow gg$ couplings.

In agreement with the SM within the 95% confidence range

→ Need more data!
In summary

Events / 3 GeV

Data

Z+X

ZY, ZZ

m_Y = 126 GeV

CMS Preliminary

√s = 7 TeV, L = 5.05 fb⁻¹

√s = 8 TeV, L = 5.26 fb⁻¹

CMS Preliminary

√s = 7 TeV, L = 5.1 fb⁻¹

√s = 8 TeV, L = 5.3 fb⁻¹

S/B Weighted Data

S+B Fit

Bkg Fit Component

±1 σ

±2 σ

m_{γγ} (GeV)
In summary

CMS Preliminary

\[ m_{\gamma\gamma} \text{ (GeV)} \]

1200
1400
1600
1800
2000

Weighted Events \((1.67 \text{ GeV})\)

Data

S/B Weighted Data

S+B Fit

Bkg Fit Component

\pm 1 \sigma

\pm 2 \sigma

\sqrt{s} = 7 \text{ TeV}, L = 5.1 \text{ fb}^{-1}

\sqrt{s} = 8 \text{ TeV}, L = 5.3 \text{ fb}^{-1}

Local p-value

Combined obs.

Exp. for SM Higgs

H \rightarrow bb

H \rightarrow \tau\tau

H \rightarrow WW

H \rightarrow ZZ

CMS Preliminary

\sqrt{s} = 7 \text{ TeV}, L = 5.1 \text{ fb}^{-1}

\sqrt{s} = 8 \text{ TeV}, L = 5.3 \text{ fb}^{-1}

Higgs boson mass (GeV)

\[ 116 \quad 118 \quad 120 \quad 122 \quad 124 \quad 126 \quad 128 \quad 130 \]

\[ 10^{-12} \quad 10^{-11} \quad 10^{-10} \quad 10^{-9} \quad 10^{-8} \quad 10^{-7} \quad 10^{-6} \quad 10^{-5} \quad 10^{-4} \quad 10^{-3} \quad 10^{-2} \quad 10^{-1} \quad 1 \]

\[ 1 \sigma \quad 2 \sigma \quad 3 \sigma \quad 4 \sigma \quad 5 \sigma \quad 6 \sigma \quad 7 \sigma \]

In the CMS Collaborations search for the Higgs boson, data at \( \sqrt{s} = 7 \) and 8 TeV were analyzed. The plots show the distribution of events as a function of the di-photon invariant mass, with significant deviations from the background predictions at the 5σ level, suggesting a potential discovery of the Higgs boson.
In summary

- Data
- $Z + X$
- $Z_1^+, ZZ$
- $m_t = 126$ GeV
- $\mu, 2e2\mu$ 7 TeV 4e, 4 $\mu$
- $8 T eV 4e, 4 \mu$

CMS Preliminary
- $\sqrt{s} = 7$ TeV, $L = 5.1$ fb$^{-1}$
- $\sqrt{s} = 8$ TeV, $L = 5.3$ fb$^{-1}$
- $s = 1.67$ GeV

- Local p-value
- $m_{\gamma\gamma}$ (GeV)
- Higgs boson mass (GeV)

- Combined obs.
- Exp. for SM Higgs
- $H \rightarrow bb$
- $H \rightarrow \tau\tau$
- $H \rightarrow \gamma\gamma$
- $H \rightarrow WW$
- $H \rightarrow ZZ$

- $\pm 1 \sigma, \pm 2 \sigma$

- S+B Weighted Data
- S+B Fit
- Bkg Fit Component
In summary

We have observed a new boson with a mass of $125.3 \pm 0.6$ GeV at 4.9 $\sigma$ significance!
Acknowledgements
Acknowledgements

- A very wide range of measurements have shown that SM predictions for known physics have been ~spot on.
- A tribute to a large amount of work done by our theory colleagues along with the results from the other collider experiments at LEP, Tevatron, HERA, b-factories etc.
- And the Higgs cross section WG and all those theorists who prepared the way for today!


Electroweak Theory

Electroweak Symmetry Breaking

First (single) beams circulating in the machine

Six CERN DGs, from conception to physics: Schopper, Rubbia, Llewellyn Smith, Maiani, Aymar, Heuer (from right to left) with 5-year terms!!
The LHC Project (the accelerator, the experiments, and computing) have required a long and painstaking effort on a global scale encompassing the terms of 6 Director Generals, 3 LHC Project Leaders, and 6 Research Directors.

The Project started in earnest in 1987 with Rubbia’s Long Range Planning Committee recommending the LHC as the right choice for CERN’s future.

Great appreciation of the work of teams that built and now operate the magnificent LHC accelerator

The CMS experiment is a tribute to the vision of its founders, the dedication of all of its thousands of collaborators in constructing and preparing the experiment in terms of hardware, software, computing, and physics analysis, and now the ones who operate and analyze the data (mostly young scientists!).

Acknowledgements
A small fraction of the CMS Collaboration: June 2012
Thanks to all of the CMS institutes


June 2012: 193 Institutions with ~3300 scientists and engineers
~ 2000 Signing Authors (including students)
HUGE Thanks to CERN Staff and CMS Funding Agencies

We congratulate our colleagues in the CERN accelerator departments for the excellent performance of the LHC machine. We thank the technical and administrative staff at CERN and other CMS institutes, and acknowledge support from: BMWF and FWF (Austria); FNRS and FWO (Belgium); CNPq, CAPES, FAPERJ, and FAPESP (Brazil); MES (Bulgaria); CERN; CAS, MoST and NSFC (China); COLCIENCIAS (Colombia); MSES (Croatia); RPF (Cyprus); MoER; SF0690030s09 and ERDF (Estonia); Academy of Finland, MEC, and HIP (Finland); CEA and CNRS/IN2P3 (France); BMBF, DFG, and HGF (Germany); GSRT (Greece); OTKA and NKTH (Hungary); DAE and DST (India); IPM (Iran); SFI (Ireland); INFN (Italy); NRF and WCU (Korea); LAS (Lithuania); CINVESTAV, CONACYT, SEP, and UASLP-FAI (Mexico); MSI (New Zealand); PAEC (Pakistan); MSHE and NSC (Poland); FCT (Portugal); JINR (Armenia, Belarus, Georgia, Ukraine, Uzbekistan); MON, RosAtom, RAS and RFBR (Russia); MSTD (Serbia); SEIDI and CPAN (Spain); Swiss Funding Agencies (Switzerland); NSC (Taipei); TUBITAK and TAEK (Turkey); STFC (United Kingdom); DOE and NSF (USA).

Individuals have received support from the Marie-Curie programme and the European Research Council (European Union); the Leventis Foundation; the A. P. Sloan Foundation; the Alexander von Humboldt Foundation; the Austrian Science Fund (FWF); the Belgian Federal Science Policy Office; the Fonds pour la Formation à la Recherche dans l'Industrie et dans l'Agriculture (FRIA-Belgium); the Agentschap voor Innovatie door Wetenschap en Technologie (IWT-Belgium); the Council of Science and Industrial Research, India; the Compagnia di San Paolo (Torino); Centro Siciliano di Fisica Nucleare e struttura della materia, ‘Teller Fellowship and LDRD’ ,and the HOMING PLUS programme of Foundation for Polish Science, cofinanced from European Union, Regional Development Fund.
The End

Stay tuned