Combined Higgs searches: steps towards interpretation

Kyle Cranmer, New York University

Northwest Terascale Research Projects: Interpreting emerging Higgs data
Psychology / Sociology

\[ m_H = 124 \text{ GeV} \]

- Combined (68%)
- Single channel

CMS, \( \sqrt{s} = 7 \text{ TeV} \)

- \( L = 4.6-4.8 \text{ fb}^{-1} \)

- \( H \rightarrow bb \)
- \( H \rightarrow \tau\tau \)
- \( H \rightarrow \gamma\gamma \)
- \( H \rightarrow WW \)
- \( H \rightarrow ZZ \rightarrow 4l \)
Inclusive diphoton sample

\[ \sqrt{s} = 7 \text{ TeV}, \int Ldt = 4.9 \text{ fb}^{-1} \]

ATLAS Preliminary

Data - Bkg model

Kyle Cranmer (NYU)  Higgs combination & Interpretation, Good Friday / J. Conway's birthday, U Oregon
Figure 1: Left: The Higgs boson rate favoured at 1 (dark blue) and 2 (light blue) in a global SM fit as function of the Higgs boson mass. Right: assuming $m_h = 125$ GeV, we show the measured Higgs boson rates at ATLAS, CMS, CDF, D0 and their average (horizontal gray band at $±1$). Here 0 (red line) corresponds to no Higgs boson, 1 (green line) to the SM Higgs boson.

This shows the main anomalous features in current measurements. First, the $WW$ and $ZZ$ channels exhibit some excess, mainly driven by the vector boson fusion data presented at the Moriond 2012 conference. Second, there is a deficit in the vector channels. Finally, the average rate of fermionic channels lies along the SM prediction; here the new Tevatron combination for $h\rightarrow b\bar{b}$ plays an important role.

3 Reconstructing the Higgs boson properties
3.1 Reconstructing the Higgs boson branching fractions

The Higgs boson observables that can be most easily affected by new physics contributions are those that occur at loop level: the $h\rightarrow WW$ and $h\rightarrow gg$ rates. They are particularly relevant for the LHC Higgs searches because $WW$ is the cleanest final state, and because $gg\rightarrow h$ is the dominant Higgs boson production mechanism. The left panel of Fig. 2 shows, as yellow contours, the 1 and 2 significance of the observed deviation from SM expectations at $m_h = 125$ GeV.
We are in this together

"Until we hear different, it's Jersey's problem."

Kyle Cranmer (NYU)

Higgs combination & Interpretation, Good Friday / J. Conway's birthday, U Oregon
Inputs to the combination

Table 1: Summary of the individual channels contributing to the combination. The central number in the three-part mass ranges indicates the transition from a low-$m_H$ to high-$m_H$ optimised event selections.

<table>
<thead>
<tr>
<th>Higgs Decay</th>
<th>Subsequent Decay</th>
<th>Additional Sub-Channels</th>
<th>$m_H$ Range</th>
<th>L [fb$^{-1}$]</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H \to \gamma\gamma$</td>
<td>$\ell\ell\ell'$</td>
<td>${4\ell, 2\ell\mu, 2\mu2\ell, 4\mu}$</td>
<td>110-150</td>
<td>4.9</td>
</tr>
<tr>
<td>$H \to ZZ$</td>
<td>$\ell\ell\nu\nu$</td>
<td>${ee, \mu\mu} \otimes {\text{low pile-up, high pile-up}}$</td>
<td>200-280-600</td>
<td>4.7</td>
</tr>
<tr>
<td>$H \to WW$</td>
<td>$\ell\ell\nu\nu$</td>
<td>${e, \mu} \otimes {0\text{-jet, 1-jet, VBF}}$</td>
<td>110-300-600</td>
<td>4.7</td>
</tr>
<tr>
<td>$H \to \tau^+\tau^-$</td>
<td>$\ell\ell\ell\nu$</td>
<td>${ee, \mu\mu} \otimes {0\text{-jet, VBF, VH}}$</td>
<td>110-150</td>
<td>4.7</td>
</tr>
<tr>
<td>$\nu\nu$</td>
<td>$\ell\ell\ell\nu$</td>
<td>${e, \mu} \otimes {0\text{-jet, 1-jet, VBF}}$</td>
<td>110-150</td>
<td>4.7</td>
</tr>
<tr>
<td>$W \to \ell\nu$</td>
<td>$\ell\ell\ell\nu$</td>
<td>${1\text{-jet}}$</td>
<td>110-150</td>
<td>4.7</td>
</tr>
<tr>
<td>$Z \to \ell\ell$</td>
<td>$\ell\ell\ell\nu$</td>
<td>$E_T^{miss} \in {</td>
<td>120,160</td>
<td>, [160,200], \geq 200\ \text{GeV}}$</td>
</tr>
</tbody>
</table>

$\mu = \sigma/\sigma_{SM}$
Table 1: Summary of the individual channels contributing to the combination. The central number in the three-part mass ranges indicates the transition from a low-$m_H$ to high-$m_H$ optimised event selections.

<table>
<thead>
<tr>
<th>Higgs Decay</th>
<th>Subsequent Decay</th>
<th>Additional Sub-Channels</th>
<th>$m_H$ Range</th>
<th>L [fb$^{-1}$]</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H \rightarrow \gamma\gamma$</td>
<td>–</td>
<td>9 sub-channels ($p_T, \eta, \gamma$ conversion)</td>
<td>110-150</td>
<td>4.9</td>
</tr>
<tr>
<td>$H \rightarrow ZZ$</td>
<td>$\ell\ell\ell\ell$</td>
<td>${4e, 2e2\mu, 2\mu2e, 4\mu}$</td>
<td>110-600</td>
<td>4.8</td>
</tr>
<tr>
<td></td>
<td>$\ell\ell\nu\bar{\nu}$</td>
<td>${ee, \mu\mu} \otimes {\text{low pile-up, high pile-up}}$</td>
<td>200-280-600</td>
<td>4.7</td>
</tr>
<tr>
<td></td>
<td>$\ell\ell q\bar{q}$</td>
<td>${b\text{-tagged, untagged}}$</td>
<td>200-300-600</td>
<td>4.7</td>
</tr>
<tr>
<td>$H \rightarrow WW$</td>
<td>$\ell\ell\ell\ell$</td>
<td>${ee, \mu\mu} \otimes {0\text{-jet, 1-jet, VBF}}$</td>
<td>110-300-600</td>
<td>4.7</td>
</tr>
<tr>
<td></td>
<td>$\ell\ell q\bar{q}$</td>
<td>${e, \mu} \otimes {0\text{-jet, 1-jet}}$</td>
<td>300-600</td>
<td>4.7</td>
</tr>
<tr>
<td>$H \rightarrow \tau^+\tau^-$</td>
<td>$\ell\ell 4\ell$</td>
<td>${e\mu} \otimes {0\text{-jet}} \oplus {1\text{-jet, VBF, VH}}$</td>
<td>110-150</td>
<td>4.7</td>
</tr>
<tr>
<td></td>
<td>$\ell\tau_{\text{had}} 3\ell$</td>
<td>${e, \mu} \otimes {0\text{-jet}} \otimes {E_T^{\text{miss}} \leq 20\text{ GeV}}$</td>
<td>110-150</td>
<td>4.7</td>
</tr>
<tr>
<td></td>
<td>$\tau_{\text{had}} \tau_{\text{had}} 2\ell$</td>
<td>${1\text{-jet}}$</td>
<td>110-150</td>
<td>4.7</td>
</tr>
<tr>
<td>$VH \rightarrow b\bar{b}$</td>
<td>$Z \rightarrow \nu\bar{\nu}$</td>
<td>$E_T^{\text{miss}} \in {[120, 160), [160, 200]}$</td>
<td>110-130</td>
<td>4.6</td>
</tr>
<tr>
<td></td>
<td>$W \rightarrow \ell\nu$</td>
<td>$p_T^W \in {&lt;50, 50-100, 100-200, \geq 200\text{ GeV}}$</td>
<td>110-130</td>
<td>4.7</td>
</tr>
<tr>
<td></td>
<td>$Z \rightarrow \ell\ell$</td>
<td>$p_T^Z \in {&lt;50, 50-100, 100-200, \geq 200\text{ GeV}}$</td>
<td>110-130</td>
<td>4.7</td>
</tr>
</tbody>
</table>

“VBF” and “VH” confusing contaminated by other processes

Cuts change and channels come in / drop out
Mass-dependent cuts

Since the cuts change with the $m_H$ hypothesis, there is not just one likelihood function. Different selections have different minima associated to them, stitching them together this way is dubious.
Since the cuts change with the $m_H$ hypothesis, there is not just one likelihood function. Different selections have different minima associated to them, stitching them together this way is dubious.
Single channel model

\[
f(\mathcal{D}|\nu, \alpha) = \text{Pois}(n|\nu) \prod_{e=1}^{n} f(x_e|\alpha),
\]

\[
\nu_{\text{tot}} = \sum_{s \in \text{samples}} \nu_s \quad f(x) = \frac{1}{\nu_{\text{tot}}} \sum_{s \in \text{samples}} \nu_s f_s(x),
\]
Explicit parametrization of systematics

Important to distinguish between the source of the systematic uncertainty (e.g., jet energy scale) and its effect.

- The same 5% jet energy scale uncertainty will have different effect on different signal and background processes
  - not necessarily with any obvious functional form
  - Usually possible to decompose to independent “uncorrelated” sources

Imagine a table that explicitly quantifies the effect of each source of systematic.

- Entries are either normalization factors or variational histograms
Combined Model:
\[
f_{\text{tot}}(\mathcal{D}_{\text{sim}}, \mathcal{G}|\alpha) = \prod_{c \in \text{channels}} \left[ \text{Pois}(n_c|\nu(\alpha)) \prod_{e=1}^{n_c} f(x_{ce}|\alpha) \right] \cdot \prod_{p \in \mathcal{S}} f_p(a_p|\alpha_p).
\]

Combined dataset:
\[
\mathcal{D}_{\text{sim}} = \{\mathcal{D}_1, \ldots, \mathcal{D}_{c_{\text{max}}}\} = \{\{x_{c=1,e=1} \cdots x_{c=1,e=n_c}\}, \ldots \{x_{c=c_{\text{max}},e=1} \cdots x_{c=c_{\text{max}},e=n_{c_{\text{max}}}}\}\}
\]

`global observables` from auxiliary measurements: \(\mathcal{G}\)
Representing the PDF

PDF components

Observables and parameters

nuisance observables global observables

categories constraints

Representing the PDF

PDF components

Observables and parameters

nuisance observables global observables

categories constraints

Representing the PDF

PDF components

Observables and parameters

nuisance observables global observables

categories constraints

Representing the PDF

PDF components

Observables and parameters

nuisance observables global observables

categories constraints

Representing the PDF

PDF components

Observables and parameters

nuisance observables global observables

categories constraints

Representing the PDF

PDF components

Observables and parameters

nuisance observables global observables

categories constraints

Representing the PDF

PDF components

Observables and parameters

nuisance observables global observables

categories constraints
**Thumbnail of the statistical procedure**

\[ \lambda(\mu) = \frac{L(\mu, \hat{\theta}(\mu))}{L(\hat{\mu}, \hat{\theta})} \]

**Follow LHC–HCG Combination Procedures**

- **CL_s** to test signal hypothesis
- **p_0** to test background hypothesis
- **\(\hat{\mu}\)** to estimate signal strength

---

**Image**

- **ATLAS**
  - Integration range: \(1 \text{ L } dt \sim 1.0-4.9 \text{ fb}^{-1}\)
  - \(\sqrt{s} = 7 \text{ TeV}\)

---

**Graphs**

- **Graph (a)**: CL_s Limits
- **Graph (b)**: p_0 (local)
- **Graph (c)**: Strength Parameter (**\(\hat{\mu}\)**)
Median & bands from asymptotics

Get Median and bands in seconds, not days!

\[ f(\bar{q}_\mu | \mu') = \Phi \left( \frac{\mu' - \mu}{\sigma} \right) \delta(\bar{q}_\mu) \]

\[ + \begin{cases} \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{\bar{q}_\mu}} \exp \left[ -\frac{1}{2} \left( \sqrt{\bar{q}_\mu} - \frac{\mu - \mu'}{\sigma} \right)^2 \right] & 0 < \bar{q}_\mu \leq \mu^2/\sigma^2 \\ \frac{1}{\sqrt{2\pi}} \exp \left[ -\frac{1}{2} \frac{(\bar{q}_\mu - (\mu^2 - 2\mu' \mu')/(2\mu/\sigma)^2)^2}{(2\mu/\sigma)^2} \right] & \bar{q}_\mu > \mu^2/\sigma^2 \end{cases} \]

G. Cowan, KC, E. Gross, O. Vitells
[arXiv:1007.1727]
CLs w/ toys & asymptotics, Bayesian
This is a global scaling of $\sigma$ BR for all production and decay modes.

\[ \mu \text{ plot} \]

\[ \text{Signal strength} \]

\[ \text{ATLAS Preliminary} \]

- Best fit
- $-2 \ln \lambda(\mu) < 1$

\[ \text{2011 Data} \]

\[ \int Ldt = 4.6-4.9 \text{ fb}^{-1} \]

\[ \sqrt{s} = 7 \text{ TeV} \]

\[ \text{ATLAS Preliminary} \]

\[ \text{Best fit} \]

\[ -2 \ln \lambda(\mu) < 1 \]

\[ \text{2011 Data} \]

\[ \int Ldt = 4.6-4.9 \text{ fb}^{-1} \]

\[ \sqrt{s} = 7 \text{ TeV} \]
Individual $\mu$ plots

Figure 8: The best-fit signal strength $\mu = \sigma / \sigma_{SM}$ as a function of the Higgs boson mass hypothesis for the $H \rightarrow \gamma\gamma$ (a), the $H \rightarrow \tau\tau$ (b), the $H \rightarrow ZZ^*(\rightarrow \ell^+\ell^-\ell^+\ell^-)$ channel in the low mass region (c), the $H \rightarrow ZZ^*$ across the full search range (d), the $H \rightarrow bb$ (e), and $H \rightarrow WW^*$ (f) individual channels. The $\mu$ value indicates by what factor the SM Higgs boson cross section would have to be scaled to best match the observed data. The band shows the interval around $\hat{\mu}$ corresponding to a variation of $-2 \ln \lambda(\mu) < 1$. 

ATLAS Preliminary 

$\int L dt = 4.7 \pm 0.3 \text{ fb}^{-1}$ 

$\sqrt{s} = 7 \text{ TeV}$
Moving beyond a single $\mu$

In practical terms, we need a parameter for each branching ratio and a parameter for each production cross-section.

- We need signal contributions for different production mode separately
  - some work, need efficiencies and signal shapes for each
  - Call them $\mu_i$ and $\mu_j$ where $i$ is production index and $j$ is decay index

Only consider SM-like interactions at first

- if distributions change, much more work needed

However, the production and decay really are not independent

- better to re-parametrize in terms of couplings
Moving beyond a single $\mu$

\[ \vec{\mu} = (\mu_{WW}, \mu_{ZZ}, \mu_{YY}, \mu_{ZZ}, \mu_{WW}, \mu_{YY}) = \mu_i \]

\[ \equiv |\vec{\mu}| \vec{p} = \mu_0 \]

Call $|\vec{\mu}| = \mu$, thus $\vec{p} = \mu_i = \frac{\mu_i}{\mu}$

This line is what we test now in SM Higgs search.

Our best fit $\hat{\mu}_i = \hat{\mu}$

Profile Construction on $n$ parameters of interest

\[ \lambda(\vec{\mu}) = \frac{L(\vec{\mu}, \hat{\mu}(\vec{\mu}))}{L(\vec{\mu}, \hat{\mu})} \]
**μ_i contours for individual channels**

Above drawn with equal efficiencies, different efficiencies will lead to a tilted contour.

Need significant differences in shape or relative efficiency across sub-channels to break degeneracy among production modes

If we introduce additional σ_X 2-d profile will fill plane
The total width of the Higgs boson is too small to be measured directly at the LHC. Therefore we have to make one single model assumption about how to treat the total width, which we take as

\[ \Gamma_{\text{tot}} = \sum_{\text{obs}} \Gamma_i(g_{iiH}) + \text{generation universality} \]
The re-parametrization can be done as a post-processing step on the combined model if the inputs are broken down by production mode and decay.

- Recent paper with a convenient parametrization in terms of couplings

Interpreting LHC Higgs Results from Natural New Physics Perspective

Dean Carmi\textsuperscript{a}, Adam Falkowski\textsuperscript{b}, Eric Kuflik\textsuperscript{a}, and Tomer Volansky\textsuperscript{a}

The decay widths of the Higgs relative to the SM predictions are modified approximately as,

\[
\frac{\Gamma(h \rightarrow bb)}{\Gamma_{SM}(h \rightarrow bb)} = |c_h|^2, \quad \frac{\Gamma(h \rightarrow WW^*)}{\Gamma_{SM}(h \rightarrow WW^*)} = \frac{\Gamma(h \rightarrow ZZ^*)}{\Gamma_{SM}(h \rightarrow ZZ^*)} = |c_\gamma|^2, \quad \frac{\Gamma(h \rightarrow gg)}{\Gamma_{SM}(h \rightarrow gg)} \simeq |c_g|^2, \quad \frac{\Gamma(h \rightarrow \gamma\gamma)}{\Gamma_{SM}(h \rightarrow \gamma\gamma)} = \left| \frac{\hat{c}_\gamma}{\hat{c}_{\gamma,SM}} \right|^2, \tag{2.5}
\]

where $\hat{c}_\gamma$ includes also the one-loop contribution due to the triangle diagram with the W boson\textsuperscript{1},

\[
\hat{c}_\gamma(\tau_t, \tau_W) = c_\gamma(\tau_t) - \frac{c_\gamma}{\sqrt{8\tau_W^2}} \left[ 3(2\tau_W - 1)f(\tau_W) + 3\tau_W + 2\tau_W^2 \right], \tag{2.6}
\]

where $\tau_W = m_h^2/4m_W^2$. For $m_h = 125$ GeV one finds $\hat{c}_\gamma \simeq c_\gamma - 1.04c_\nu$, and thus $\hat{c}_{\gamma,SM} \simeq -0.81$. 

\[c_\nu = c_d = c_\tau = 1\]

\[\delta c_\gamma, \delta c_d, \delta c_\tau\]
**To-do list**

Efficiency and acceptance for each production mode vs. $M_H$ for each channel
- not really enough b/c shapes etc.

$\mu_i$ contours per channel (next slide)

Form combined likelihood with $L(\mu_i, \mu_j; m_H)$
- no relationship between production and branching ratio
  - punt on other coupling structures, higher dim operators etc.

Reparametrize w/ D5 ($\Delta_g$ and $\Delta_y$) no constraint on total width
- ie. no generational universality and include unobserved $\Gamma_{\text{hidden}}$
- no sensitivity, could make plots with limited range on $\Delta_g$, $\Delta_y$, $\Gamma_{\text{hidden}}$
  assuming generational universality
- make plots on ratios with respect to HWW coupling

Reparameterize w/ $\Delta_g = \Delta_y = \Gamma_{\text{hidden}} = 0 +$ generational universality

Add additional assumptions, ie. W/Z fixed or THDM relationships?

All couplings fixed to SM, but add $\Delta_g$, $\Delta_y$, $\Gamma_{\text{hidden}}$
- would give sensitivity to counterfeit like scenarios?
Daniele Alves pointed out a potentially clean way to measure $WW$ coupling: $pp \rightarrow WH \rightarrow WWW \rightarrow l^\pm l^\pm jj$

Curious that I never see separately cross-section for $pp \rightarrow W^+H$ and $pp \rightarrow W^-H$ in LHC Higgs Cross section working group.

- new physics and backgrounds can break the charge symmetry