Projected exclusion limits on the SM Higgs boson cross sections obtained by combining the $H \rightarrow WW$ and $ZZ$ decay channels

Grégory Schott
Institut für Experimentelle Kernphysik
Universität Karlsruhe (TH) – Germany

for the CMS collaboration

17th International Conference on Supersymmetry and the Unification of Fundamental Interactions
Northeastern University, Boston, 5–10 June 2009
Introduction (1)

- Many channels related in Higgs analyses
  - Overlapping regions of sensitivities
  - Correlations/uncertainties treated optimally in a simultaneous analysis
  - Combination of LHC analyses

- Many channels related in Higgs analyses

- Overlapping regions of sensitivities
- Correlations/uncertainties treated optimally in a simultaneous analysis
- Combination of LHC analyses

figs. from CMS physics technical design report vol.2 (chap. 2.2, 10.10 and 10.16)
Introduction (2)

- **Projected exclusion limits** on the Standard-Model Higgs
  - combining 2 modes $H \rightarrow ZZ$, $H \rightarrow WW$
    - both analyses presented earlier in this session
    - most sensitive for large $m_H$
    - presenting preliminary results
  - multiple statistical approaches considered

- Aims:
  - study **sensitivity reach** of Higgs at CMS
  - preparation/understanding of **cross-channel correlations**
  - implementation and test of **statistical tools**

- Implications of **decrease in beam energy** from 14 to 10 TeV
  - extrapolation from the 14 TeV results
Statistical methods

- 2 methodologies applied:
  - Bayesian
  - modified frequentist (CLs)

- parameter of interest: signal strength: \( r = \frac{\sigma_H}{\sigma_{H,SM}} \)
  - \( r=1 \): in the assumption of Standard-Model rates
  - \( r=0 \): in the assumption only background is present

- both based on a common probability density function:
  \[ \text{pdf}(\text{data}|r,\theta) \rightarrow \text{used differently} \]

- both interpretations of the data also have different meanings
Likelihood function

- **Counting analyses** (2-bin problem):
  - \( N_i \) observed events
  - expect \( b_{0i} \) background events
  - for projected exclusion limits: assume \( N_i = b_{0i} \)
  - want to say something about the signal yield \( s_i = r \cdot s_{0i} \)

\[ L_i(N_i | r, s_{0i}, b_{0i}) = \frac{e^{-r \cdot s_{0i} + b_{0i}}}{N_i!} (r \cdot s_{0i} + b_{0i})^{N_i} \]

\[ \rightarrow \text{simple likelihood:} \]

- **Multiple channels** \( L(r) = \prod_i L_i(N_i | r, s_{0i}, b_{0i}) \)

- **Likelihood function with measured observable(s)**
  \[ L(\vec{x} | r, s, b, \vec{\theta}_s, \vec{\theta}_b) = \frac{e^{-s+b}}{N!} (s + b)^N \prod_{i=1}^{N} (s f_s(\vec{x} | \vec{\theta}_s) + b f_b(\vec{x}, \vec{\theta}_b)) \]
  - \( f_s, f_b \): signal and background distributions from MC or control samples
  - simultaneous treatment of the analyses:
    \( \rightarrow \) retain full information for treating correlations and systematic uncertainties
Bayesian approach

- **Bayesian theorem:**

\[
P(N_s|\text{data}, \theta) = \frac{\text{pdf}(\text{data}|N_s, \theta) \pi(N_s)}{\int \text{pdf}(\text{data}|N_s, \theta) \pi(N_s) \, dN_s}
\]

- test compatibility of signal+background model against a given data sample
- for projected limits, assume \( n_{i,\text{obs}} = b_{0i} \)
- assume flat prior on the signal yield: \( \pi(N_s) = \text{constant} \)
- assumed correlations: 100% between all \( s_i \) and \( b_i \) (for simplification)
- 4 channels to combine, 8 systematics errors: max. error \( \sigma_{\text{max}} \) ; assume truncated Gaussian pdf \( g_T(x) \) with sigma \( \sigma_{\text{max}} \)
  \[
  L(r) = \frac{\int dx \prod_i p(n_i | b_i + r \cdot s_i) g_T(x) \pi(r)}{\int dr \int dx \prod_i p(n_i | b_i + r \cdot s_i) g_T(x) \pi(r)}
  \]
  \[
  s_i = s_{0i} \left(1 + \frac{\sigma_{si}}{\sigma_{\text{max}}} x\right) \quad b_i = b_{0i} \left(1 + \frac{\sigma_{bi}}{\sigma_{\text{max}}} x\right)
  \]
- 95% credibility upper limit on the ratio \( r_{95\text{CL}} \) of the signal yield/cross-section to the one projected
Frequentist approach

- Evaluate frequentist compatibility of data (in toy-MC experiments)
  - Method based on a test statistics $t$ (for example: $t = -2 \ln \frac{L_{SB}}{L_B}$)
  - Systematic uncertainties taken into account by Bayesian marginalization

- Play out toy experiments ($N_{i,\text{obs}}$) in the background-only hypothesis
  - In each pseudo-experiment vary $b_i$’s according to assumed systematic errors
  - Play out toy experiments again, now assuming $b_i + r \times s_i$
  - Again vary $b_i$’s and $s_i$’s according to assumed systematic errors

Assume that observation is $t_{\text{data}} = <t>$

- $p_{SB} = CL_{SB} = \text{Prob} (t > t_{\text{data}})$
- $p_B = 1 - CL_B = \text{Prob} (t < t_{\text{data}})$

$CL_S$ prescription (modified-frequentist):
  - $CL_S = CL_{SB} / CL_B$ (conservative)
  - cures background downward fluctuations
  - for a 95% CL upper limit:
    - scan for $r$ such that $CL_S = 0.05$

Distribution of the test statistics
H → ZZ and H → WW inputs

- Combination based on CMS AN-2008/039 and CMS AN-2008/050
- Ratio $s/\sqrt{b}$ in:
  - ZZ sub-modes: similar → yields lumped together
  - WW sub-modes: varies → considered separately (80% correlation assumed in CLs)

- Projected yields at 14 TeV

- Projected yields at 10 TeV
Results at 14 TeV

- Results with $L = 1 \text{fb}^{-1}$ at 14 TeV
  - expect to exclude: $140 < m_H < 250 \text{ GeV/c}^2$ (with those inputs)
  - in ranges 120–150 & 185–200 both analyses are complementary
14 \rightarrow 10 \text{ TeV approximations}

- **Change in cross section** for signal and main backgrounds
  - in $H \rightarrow WW^* \rightarrow 2l2\nu$

- studied at generator level (MCFM):
  - signal: 55%
  - background: 65–70%
    - $tt$ background: 40%

- **Main impact** from cross-section change:
  - processes originating from primary gluons ($gg \rightarrow H$) suffer more than those involving $qq$ initial states (predominant backgrounds)
  - processes with higher invariant mass get hit harder (see next slide)
14 → 10 TeV approximations

- Change in cross section for signal and main ZZ background
  - in $H \rightarrow ZZ^* \rightarrow 4l$
    - signal: 50–55%
    - background: 65–70%

- Change of detector acceptance
  - at lower $\sqrt{s}$: objects get less boosted
  - studied at generator level (driven by kinematic cuts)
  - acceptance change is a second order effect
  - expect also negl. $H \rightarrow WW$ acceptance change
Results at 10 TeV

- **Bayesian / CLs limits in good agreement**
  - although different assumptions on correlations and methods
  - deviations O(10%)

- **Expected excluded range at 10 TeV of SM Higgs**
  - 150–190 GeV/c² (with L = 1 fb⁻¹)
  - loss of factor 1.5 in sensitivity
    → 2 times larger integrated luminosity needed to reach the same sensitivity

- Exclusion of SM Higgs boson with $m_H \sim 160$–$170$ GeV/c² requires $\sim 200$ pb⁻¹ at 10 TeV
Conclusion

- First combination of Higgs analyses in CMS
  - preparation of tools (Atlas-CMS-ROOT collaborative tool RooStats), RSC
  - comparison of statistical methods
  - basic account for cross-channel correlations

- Results:
  - good agreement of the frequentist and Bayesian results
  - 14 $\rightarrow$ 10 TeV extrapolations
  - expected exclusion of $150 < m_H < 190$ (with 1 fb$^{-1}$ @ 10 TeV)

- Plans:
  - improvement in the $H \rightarrow ZZ$ and $H \rightarrow WW$ analyses (use of additional information in the likelihood model, ...)
  - combination with other Higgs channels ($H \rightarrow \gamma\gamma$, VBF $H \rightarrow WW$, ...)
  - refined treatment of correlations and systematic uncertainties
  - projections for 200 pb$^{-1}$ @ 10 TeV