Recasting Higgs Data II

Guillaume Drieu La Rochelle

IPNL

March 19, 2013

Solutions on the market

How are these issues dealt by now :

- Obtaining the efficiencies
  - Espinosa et al.(12.07.1717), Belanger et al. (1212.5244), Cacciapaglia et al.(1210.8120) ... among others
- Include correlations
  - Giardino et al. (1303.3570), Corbett at al. (1211.4580) ... among others

Obtaining efficiencies

- A need depending on the model prediction
  - Not necessary if productions scale as SM

Obtaining efficiencies

- A need depending on the model prediction
  - Not necessary if productions scale as SM
- First method : Private communications
  - Pros : Exact result. Cons : quite not practical

Obtaining efficiencies

- A need depending on the model prediction
  - Not necessary if productions scale as SM
- First method : Private communications
  - Pros : Exact result. Cons : quite not practical
- Second method : Estimate through event generator
  - Cons : Approximate results and little reliability. Somehow tedious.
  - Pros : Can be done systematically.
  - Also, some discrepancies may cancel since we are only interested in ratios of efficiencies.

Correlations

- Even if all efficiencies are known, what about correlations between subchannels?
- Systematics are no more subdominant :

year	2012	2013
f_{sys}	20%	50%

Correlations

- Even if all efficiencies are known, what about correlations between subchannels?
- Systematics are no more subdominant :

year	2012	2013
f_{sys}	20%	50%

- Split uncertainties (as in Giardino et al. 1303.3570)

$$\sigma_i = (\tilde{\sigma}_i, \sigma_i^{\text{prod}})$$

Correlations

- Even if all efficiencies are known, what about correlations between subchannels?
- Systematics are no more subdominant :

year	2012	2013
f_{sys}	20%	50%

- Split uncertainties (as in Giardino et al. 1303.3570)

$$\sigma_i = (\tilde{\sigma}_i, \sigma_i^{\text{prod}})$$

- But there are more correlations :

$\sigma_{gg\to H+0j}, \sigma_{gg\to H+1j}, \sigma_{gg\to H+2j}$

have significant uncertainties

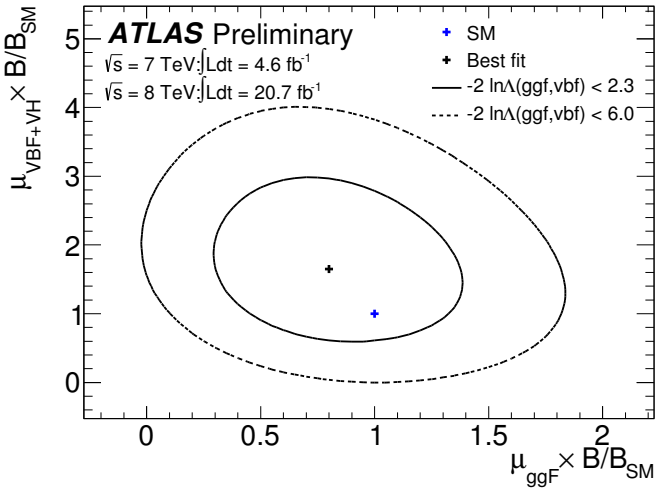
$\sigma_{gg\to H+X}$

is more precise ➡ correlation $(0j,1j,2j)$

Efficiencies and Correlations

- CMS and ATLAS provide a fit for each channel in

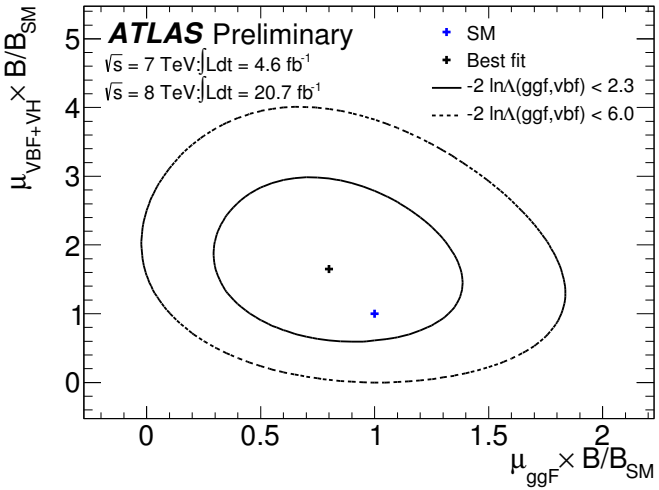
$$(\mu_{\text{ggh/tth}}, \mu_{\text{VBF/VH}})$$



Efficiencies and Correlations

- CMS and ATLAS provide a fit for each channel in

$$(\mu_{\text{ggh/tth}}, \mu_{\text{VBF/VH}})$$



- The χ^2 can be read at once.
  - It does include the efficiencies exactly.
  - It includes correlations between productions modes.

Drawbacks

- Force assumptions on the model :
  - $\bar{t}tH$ and $gg \to H$ do not contribute at the same time
    ➡ so $H \to \bar{b}b$ OK
  - Need some custodial symmetry to have $\kappa_{\mathrm{VBF}} = \kappa_{\mathrm{VH}}$.

- Misses correlations :
  $=$ production mode & $\neq$ decays do have correlations

- What about varying m_H?

Going further : the precise likelihood

- Could theorists have access to the full likelihood?
  - Not tied only to Higgs data
  - RooStats workspace ➡ Let us ask Sezen...
  - Sparse grid : we do not need that many points.

Going further : the precise likelihood

- Could theorists have access to the full likelihood?
  - Not tied only to Higgs data
  - RooStats workspace ➟ Let us ask Sezen...
  - Sparse grid : we do not need that many points.

- Or maybe just a sketch ?
  - Suggestion in Giardino et al. 1303.3570)
    ➟ Gaussian approximation in the parameter space
  - Free parameters for all productions and decays ➟ O(10) parameters

$$\mathcal{L}(\kappa) = \frac{1}{2}\kappa_i \mathbf{M}_{ij} \kappa_j$$

  - only needs a 10×10 matrix.

Trouble with full likelihood

- Requirement : We need to be in the statistical asymptotic regime.
  - Not if we have more and more subchannels

Trouble with full likelihood

- Requirement : We need to be in the statistical asymptotic regime.
  - Not if we have more and more subchannels

- One parameter per production mode may not be enough :
  - Effective lagrangian $\mathcal{L}_{\text{eff}} \supset k_W H W_\mu W^\mu + k'_W D_\mu H W^{\mu\nu} W_\nu$ $\neq$ Lorentz structure have $\neq$ efficiencies

- And what about BDT?
  - Trained on SM samples.

Main questions

- What should theorists use if not full $\mathcal{L}$?
  - Are 2D plots enough?
  - Get the mass dependence.
  - Which approximating function for $\mathcal{L}$ (Gaussian, Grid interpolation)?

- How do we assess accuracy of each method?