Simplified Template Cross Sections

Frank Tackmann

Deutsches Elektronen-Synchrotron

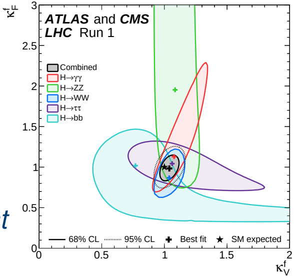
Higgs/B9 meeting, DESY
June 26, 2017

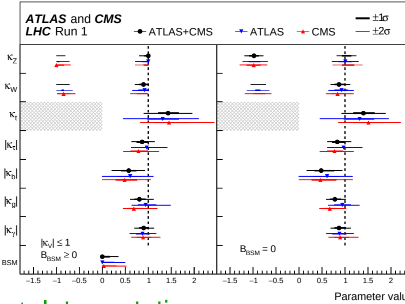
Introduction.

Direct Coupling Fits.

Data

theory-dependent





Measurement + Interpretation

Pros

- Maximum possible sensitivity
- Allows use of advanced selection techniques (MVAs, black magic, ...)
- Can benefit from kinematic correlations among production modes across all decay channels in combination

Cons

- Theory predictions and *uncertainties* maximally entangled in results
- Any nontrivial theory changes require new results from experiments

Separating Measurement from Interpretation.

Measurement — Interpretation

Data $\longrightarrow$ Simplified Cross Sections $\longrightarrow$ κ_i, g_k $\longrightarrow$ Lagrangian parameters

Minimize theory dependence — *Direct theory dependence*

Goals

- Minimize dependence on theory systematics in measurements
  - Clearer and systematically improvable treatment at interpretation level (acceptance corrections and extrapolations to total xsec)
- Minimize model dependence in measurements
  - Decouples measurements from discussions about specific models (SM, linear/nonlinear EFT, BSM models)
- Measurements stay long-term useful
- Allows easy (re)interpretation with different theory inputs/assumptions
  - Improved theory predictions/uncertainties
  - μ_i, κ_i, anomalous couplings, EFT coefficients, specific BSM scenarios

Fiducial Cross Sections.

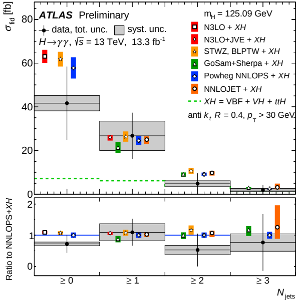
Pros: Staying as close as possible to what is actually measured

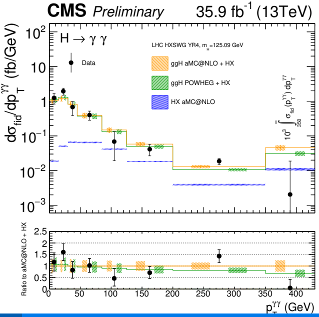
- Allows maximally theory-independent measurements
- Representation of the data that remains long-term useful

$\Rightarrow$ Of course nothing new and routinely done in other SM measurements

However, Higgs is quite different from (other) SM measurements

- Many different production and decay modes with large differences in
  - Statistics
  - Relative signal/background
  - Theory uncertainties
  - BSM sensitivity





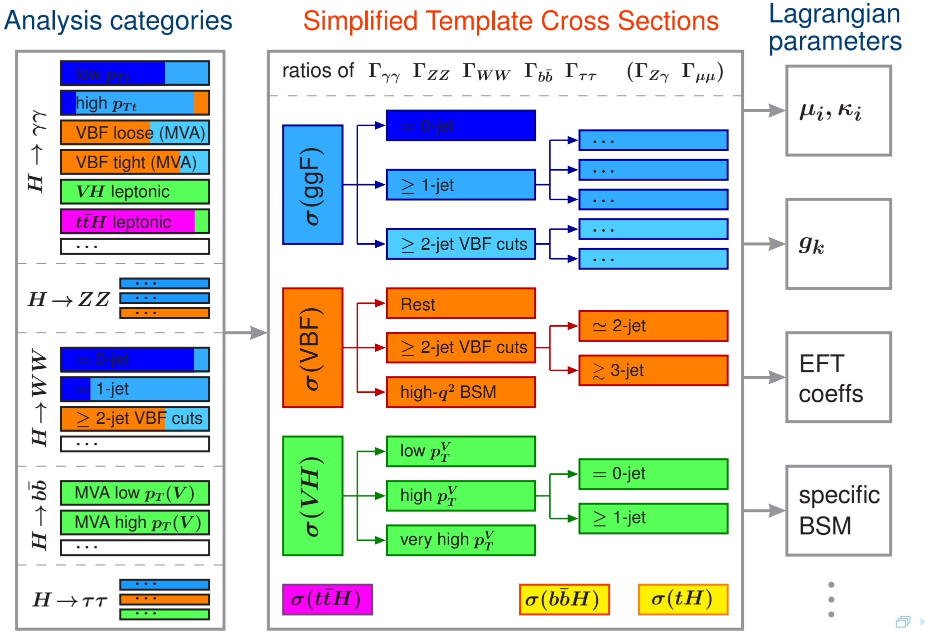
Disadvantages of Fiducial Cross Sections.

Optimizing for maximal theory independence requires sacrificing sensitivity

- Requires clean decay channels: $H \to \gamma\gamma, ZZ, (WW)$
- Requires signal definitions such that all experimental efficiencies are independent of production mode
  - Otherwise, efficiency corrections introduce dependence on assumed SM production mode mix
  - Often cannot use MVAs to optimize kinematic signal selection cuts but need simple (rectangular) cuts
  - Sometimes this is just not possible
- Projecting onto several single-differential spectra looses information compared to fully-differential level and introduces statistical overlap

$\Rightarrow$ Simplified template cross sections

Simplified Template Cross Section Framework.



Construction.

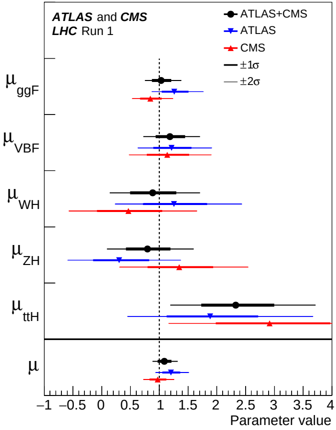
Consider schematic μ fits:

$$\sigma_1^{\mathrm{meas}} = A_1^{ggH} \times \mu_{ggH} \times \sigma_{ggH}^{\mathrm{SM}} \quad + \quad A_1^{\mathrm{VBF}} \times \mu_{\mathrm{VBF}} \times \sigma_{\mathrm{VBF}}^{\mathrm{SM}}$$

$$\sigma_2^{\mathrm{meas}} = A_2^{ggH} \times \mu_{ggH} \times \sigma_{ggH}^{\mathrm{SM}} \quad + \quad A_2^{\mathrm{VBF}} \times \mu_{\mathrm{VBF}} \times \sigma_{\mathrm{VBF}}^{\mathrm{SM}}$$

$$\sigma_3^{\mathrm{meas}} = \cdots$$

- σ_i^{meas}: measured analysis categories
- A_i^{ggH}, A_i^{VBF}: Acceptances for SM processes ($\rightarrow$ theory-dependent)



Construction.

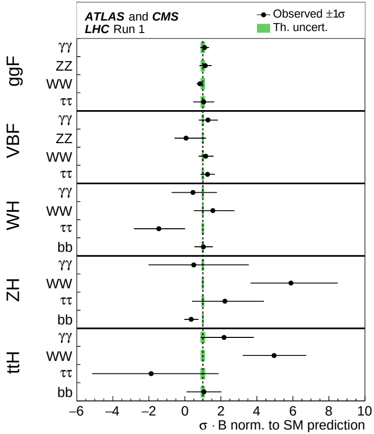
Consider schematic μ fits:

$$\sigma_1^{\mathrm{meas}} = A_1^{ggH} \times \sigma_{ggH} + A_1^{\mathrm{VBF}} \times \sigma_{\mathrm{VBF}}$$

$$\sigma_2^{\mathrm{meas}} = A_2^{ggH} \times \sigma_{ggH} + A_2^{\mathrm{VBF}} \times \sigma_{\mathrm{VBF}}$$

$$\sigma_3^{\mathrm{meas}} = \cdots$$

- σ_i^{meas}: measured analysis categories
- A_i^{ggH}, A_i^{VBF}: Acceptances for SM processes ($\rightarrow$ theory-dependent)
- First step: Fit for σ_{ggH}, σ_{VBF} rather than μ_{ggH}, μ_{VBF}
  - In the SM correspond to total ggF and VBF production cross sections
  - Can combine channels by assuming or fitting ratios of BR
  - Already available



Construction.

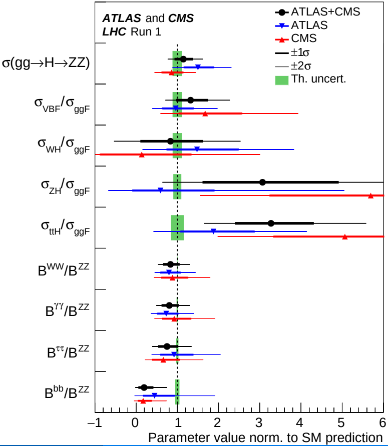
Consider schematic μ fits:

$$\sigma_1^{\text{meas}} = A_1^{ggH} \times \sigma_{ggH} + A_1^{\text{VBF}} \times \sigma_{\text{VBF}}$$

$$\sigma_2^{\text{meas}} = A_2^{ggH} \times \sigma_{ggH} + A_2^{\text{VBF}} \times \sigma_{\text{VBF}}$$

$$\sigma_3^{\text{meas}} = \cdots$$

- σ_i^{meas}: measured analysis categories
- A_i^{ggH}, A_i^{VBF}: Acceptances for SM processes ($\to$ theory-dependent)
- First step: Fit for σ_{ggH}, σ_{VBF} rather than μ_{ggH}, μ_{VBF}
  - In the SM correspond to total ggF and VBF production cross sections
  - Can combine channels by assuming or fitting ratios of BR
  - Already available



Construction.

$$\sigma_1^{\text{meas}} = A_{1a}^{ggH} \times \sigma_{ggH}^{a} + A_{1b}^{ggH} \times \sigma_{ggH}^{b} + A_{1c}^{\text{VBF}} \sigma_{\text{VBF}}^{c} + \cdots$$

$$\sigma_2^{\text{meas}} = A_{2a}^{ggH} \times \sigma_{ggH}^{a} + A_{2b}^{ggH} \times \sigma_{ggH}^{b} + A_{2c}^{\text{VBF}} \sigma_{\text{VBF}}^{c} + \cdots$$

$$\sigma_3^{\text{meas}} = \cdots$$

Next step: Split up production cross sections into kinematic regions a, b, c, ...

- Separately fit bin cross sections σ_{ggH}^{a}, σ_{ggH}^{b}, σ_{VBF}^{c}, ...
- Bin acceptances A_{ij}^{ggH}, A_{ij}^{VBF} now only need to assume/depend on SM kinematics *inside* a given bin
  - If this becomes a limitation $\rightarrow$ further split the bin

$\Rightarrow$ Direct extension of existing framework, can be implemented by experiments straightforwardly on top of existing MC samples

Defining Features.

- Measure cross sections but separated into production modes
  - Allows different efficiencies/acceptances for different production modes without incurring dependence on SM production mode mix
  - SM processes act as kinematic templates
  - Future: Can add more kinematic templates (e.g. CP-odd Higgs)
- Non-Higgs backgrounds are subtracted
  - Future: Can add templates for BSM sensitive backgrounds (e.g. $pp \to WW$)
- Inclusive over the Higgs decays
  - Can perform a global combination of channels
- “Simplified” bin definitions abstracted from the actual measurement categories
  - Allow some acceptance corrections,
  - Analyses can use optimized selections at reconstruction level, MVAs ...
  - Avoid extrapolations that are unnecessary or nontrivial (i.e. theory sensitive)

$\Rightarrow$ Maximize sensitivity while reducing theory dependence

Comparison.

Fiducial: Optimized for maximal theory independence

- Minimize acceptance corrections
- Simple (rectangular) signal cuts
- “Exact” fiducial volume
- Fiducial in Higgs decay
- Targeted object definitions

Agnostic to production modes

(Single-)differential distributions
(overlapping events)

Only $H \to \gamma\gamma, ZZ, (WW)$
(by default no combination of channels)

Simplified: Maximize sensitivity while reducing theory dependence

- Allow larger acceptance corrections
- Allow event categories, MVAs, ...
- Abstracted/simplified fiducial volumes
- Inclusive in Higgs decay
- Common idealized object definitions

Xsec split by production mode

Xsec split into mutually exclusive regions of phase space

Explicitly designed for combination of all decay channels

Basic Design Principles.

Identify phase-space regions most important to separate out from theory side

- Where are largest theory systematics? BSM sensitivity?
- Try to minimize residual theory dependence in measurements
  - Avoid non-constant signal acceptance within one bin
  - Try to align cuts with dominant channel/categories to reduce extrapolations

Impossible to define one set of bins perfect for every analysis and theory, so aim to find a good compromise

- Only add additional bins with "sufficiently good reason" (see above)
- Some decay channels will only be able to constrain sum of certain bins and must be able/allowed to combine bins
  - Bins are defined to be mutually exclusive and sum up to parent bin
  - If merged bins have similar acceptance $\rightarrow$ Bins can be split in the combination (unbiased, only some loss in sensitivity)
  - If merged bins have different acceptance $\rightarrow$ Split the bins if at all possible, otherwise combine and assign uncertainty in measurement
- Bin definitions can evolve with statistics $\rightarrow$ Staging

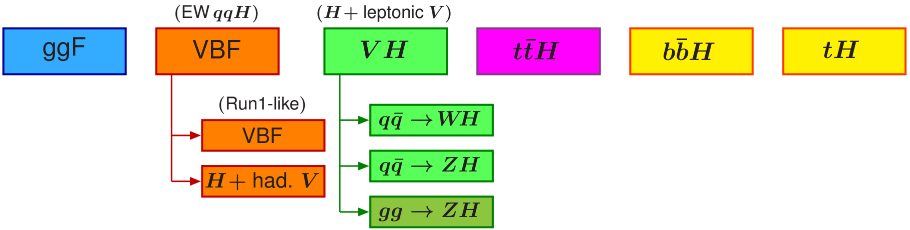
Staging.

Define different “stages” for each production mode

- Each analysis implements the binning according to the appropriate stage
- Evolution of different production modes can take place independently
- Bin definitions can evolve with statistics
  - Individual analyses can quote sum of bins while sensitivity is still limited
  - In BSM “overflow” bins even limits are very interesting
  - Can split into more fine-grained bins as required and allowed by statistics (previous determinations remain useful)

- Stage 0: closest correspondence to Run1
- Stage 1
  - All “minimally hoped-for” splits
  - Intermediate steps to get there indicated by “$(+)$” for possible bin merging
  - Early measurements will show if adjustments are needed (will not make any changes unless serious problems arise)
- Stage 2: to be defined (after gaining more real-life experience)

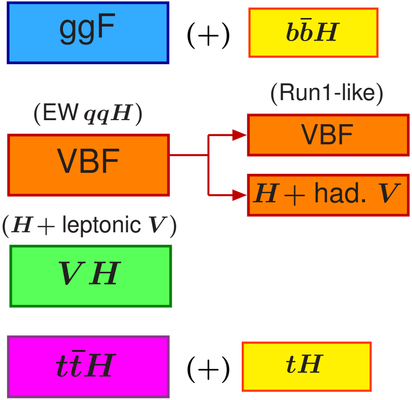
Stage 0.

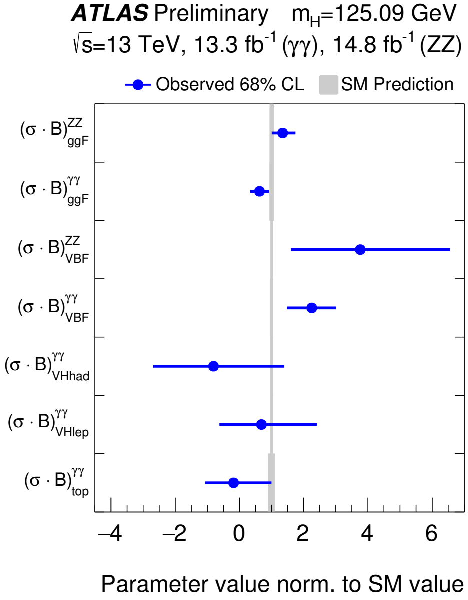


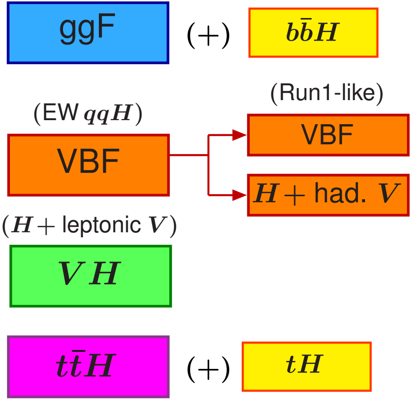
Inclusive cross section per production mode

- Closest correspondence to Run1 production-mode μ measurements, *but* expressed in terms of cross sections and restricted to $|Y_H| < 2.5$
- "VBF" defined as electroweak qqH
  - Split into Run1-like VBF and hadronic VH
- "VH" defined as H + leptonic V
  - Split into WH and ZH, and/or $q\bar{q}\to ZH$ and $gg\to ZH$
- Once meaningful, $b\bar{b}H$ and tH

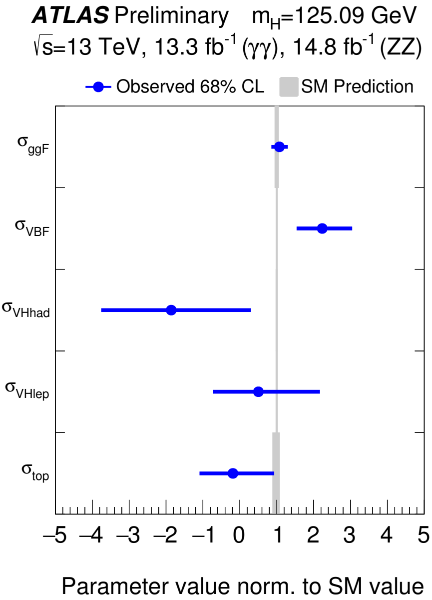
First Stage-0 Measurements at Run 2.



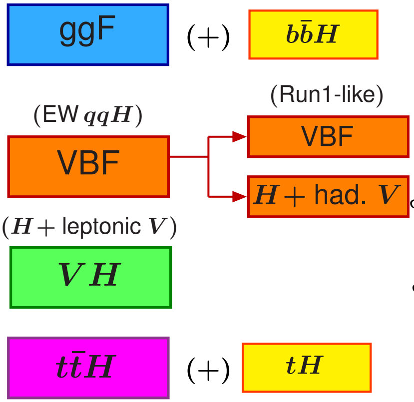


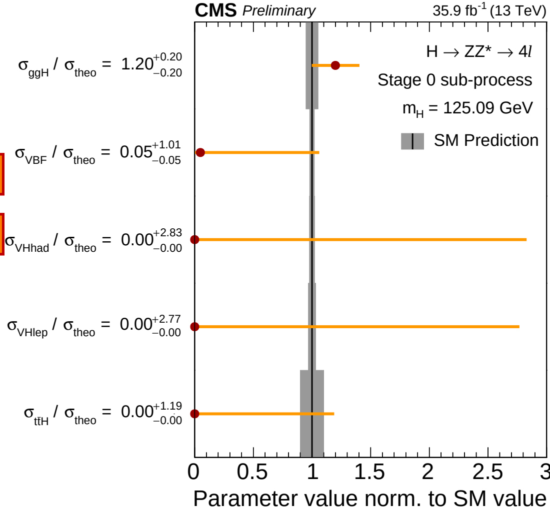


(Combined $\gamma\gamma$ and ZZ assuming SM BR)

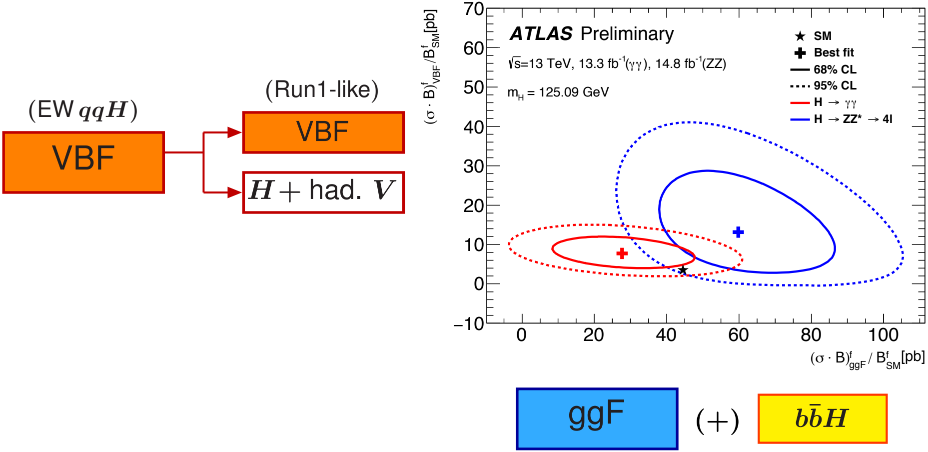


First Stage-0 Measurements at Run 2.



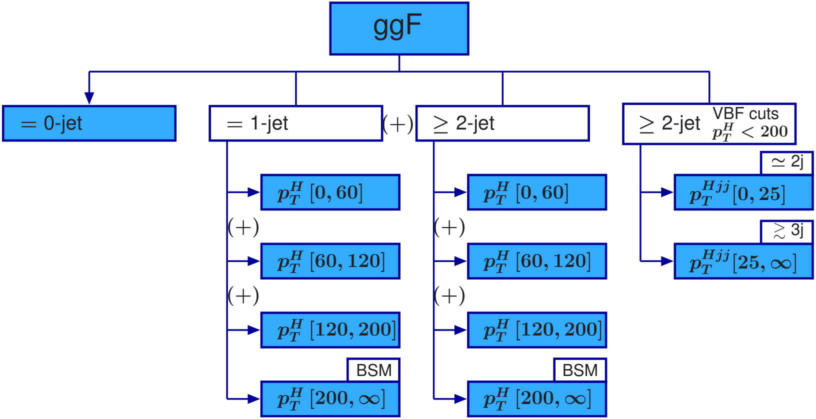


First Stage-0 Measurements at Run 2.



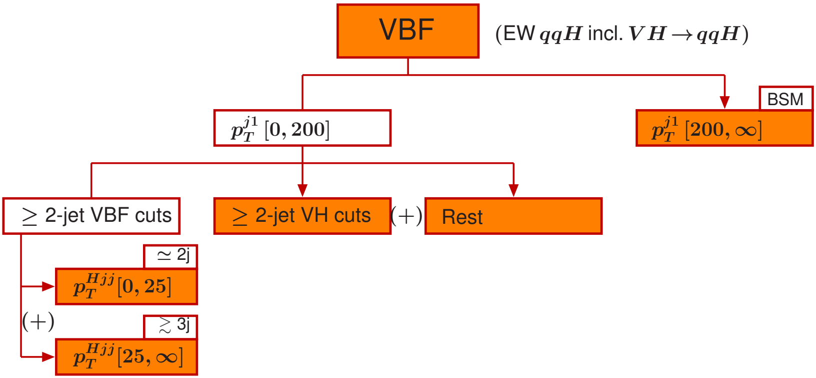
- Normalized to SM branching ratios for plotting purposes only

Gluon Fusion – Stage 1.



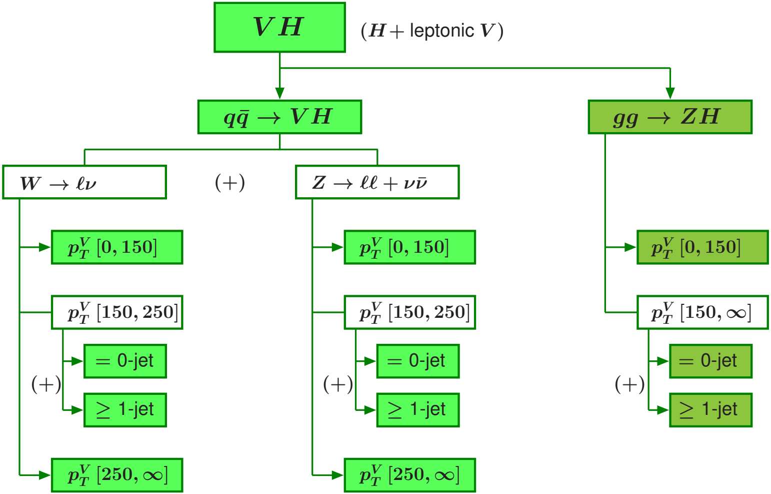
- Jet bins motivated by experimental analyses
- High p_T^H bins target boosted categories ($\tau\tau$) and BSM overflow
- VBF-like cuts to constrain ggF contribution in VBF categories

VBF – Stage 1.



- VBF defined as electroweak $qq'H$ production
  - including usual VBF process and VH with hadronic V decays
- First split by p_T^{j1}
  - VBF topology cuts: $m_{jj} > 400\,\mathrm{GeV}$ and $\Delta\eta_{jj} > 2.8$ (no other cuts)
  - $V(\to jj)H$ topology cuts: $60\,\mathrm{GeV} < m_{jj} < 120\,\mathrm{GeV}$
  - Rest: Everything not passing above (including events with < 2 jets)

VH – Stage 1.



- VH defined as Higgs in association with leptonically decaying V
  - $q\bar{q} \to V(\to q\bar{q})H$ part of VBF $(gg \to Z(\to q\bar{q})H$ part of ggF)
- Binning in p_T^V aligned with $H \to b\bar{b}$ (which is main contributor)

Les Houches: How to Provide Results.

Stage 1 measurements are in progress

- What information should be provided by the experiments and how, so that the results can be used for interpretations?
  - EFT fits, specific BSM models, ... as long as the acceptance used in the measurements is not too different
- Several bins that are weakly constrained and/or strongly correlated
  - Highest p_T bins (most senstive to BSM) will always be statistics limited
  - VBF-like ggF bins and true VBF cannot be easily disentangled
  - Important to take into account correlations but providing full O(30)-dimensional likelihood is unfeasible

First: Assume Gaussian behavior is a good approximation

- Provide full O(30)-dimensional covariance matrices separated by
  - statistical uncertainties
  - total experimental systematic uncertainties
  - combined* theoretical uncertainties
- *Separate out specific theoretical uncertainties that might have to be correlated between the measurement and the interpretation

Les Houches: How to Provide Results.

Second case (next simplest)

- Still assume Gaussian behavior is a good approximation
- Keep nuisance parameters of those theory uncertainties that might have to be correlated between measurement and interpretation unprofiled
  - Report extended covariance matrix (or Hessian)
  - Precise definition of corresponding uncertainty source and variation

Third case

- In case Gaussian approximation is not good enough for some bins, provide (parametrized) likelihood for these bins

In all cases

- Experiments will need to test if interpretations with these inputs possible
  - Can compare results for reference SM interpretation to using full likelihoods

Les Houches: Planned Binning Updates.

ggF

- Consider to split the highest p_T^H bins into [200, 500] and [500, ∞] to have a dedicated bin for boosted $H \to b\bar{b}$ analyses

Treatment of $gg \to Z(\to q\bar{q})H$

- $q\bar{q} \to Z(\to q\bar{q})H$ is considered part of VBF (EW qqH)
- For same reason $gg \to Z(\to q\bar{q})H$ is considered EW correction to ggF (Caused some confusion, conclusion of discussion is to keep it like this)
- In the future could split out ggF 2-jet bins the region $60 < m_{jj} < 120\,\mathrm{GeV}$

$b\bar{b}H$ and $t\bar{t}H$

- Include $b\bar{b}H$ in ggF
  - Currently no distinction without dedicated b-tagged reco categories
- Stage 1 has one inclusive bin for $t\bar{t}H$
  - Includes also tH
  - Consider split by p_T^H: [0,200] and [200, ∞] to separate nonboosted and boosted analyses
  - If useful could split [0, 200] into =0j and $\geq$1j (in addition to $t\bar{t}H$ signal jets)

Les Houches: Other Topics.

Providing a reference SM interpretation

- Parametrization of theoretical uncertainties for ggF, VBF, VH

Including final state observables

- Angular information from $H \to ZZ^*$ and $H \to WW^*$ currently not available in STXS measurements by construction
- In the measurements, can measure sensitive quantities (e.g. decay angles) in (some) STXS bins
- In principle it should be possible to extend STXS framework to include final state observables (e.g. angular coefficients/POs)

Backup Slides

Object Definitions.

- Definitions of “truth” final-state objects (adapted to current scope)
- Explicitly kept simpler and more idealized than in fiducial cross section measurements
  - Allow comparison with theoretical predictions from both analytic calculations and MC simulations

Higgs boson

- All bins are for an on-shell Higgs boson with a cut $|Y_H| < 2.5$
  - Current measurements have no sensitivity beyond this
  - Once sensitivity to higher rapidity (e.g. using forward leptons in $H \to ZZ \to 4\ell$) add an additional otherwise inclusive bin for $|Y_H| > 2.5$
- Treating Higgs as final-state particle is what allows combination of decay channels

Object Definitions.

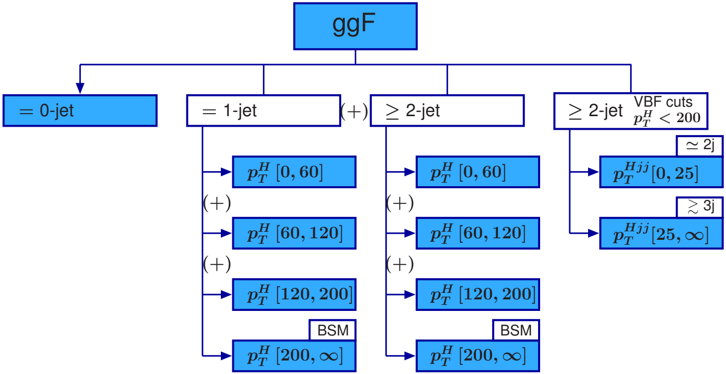
Leptons from decays of signal vector bosons (i.e. VH)

- Electrons and muons are defined as dressed
- τ defined from sum of decay products (for any decay mode)
- No restriction on lepton p_T or rapidity

Signal jets

- Anti-k_t jets with $R = 0.4$
  - built from all stable particles, including neutrinos, photons, leptons from hadron decays
  - All particles arising from Higgs decay are removed
  - All particles from leptonic decays of signal V bosons are removed
  - Decay products from hadronic decays of signal V are included
- Common p_T^j threshold of 30 GeV
- Truth jets are defined with no restriction on jet rapidity
  - Rapidity cuts can be included in bin definitions if needed

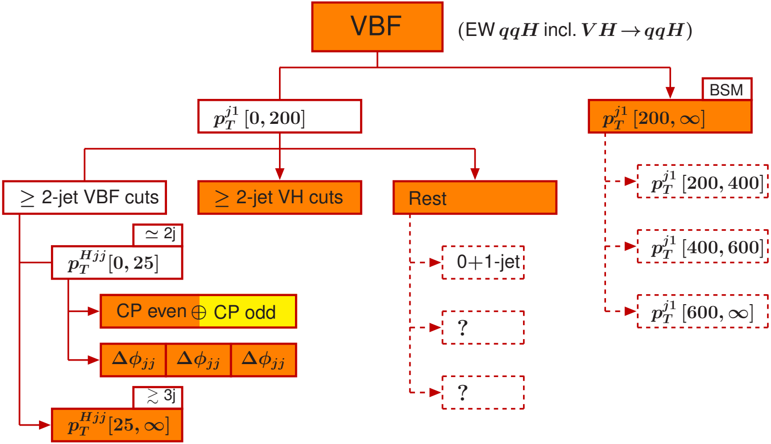
Gluon Fusion – Stage 2.



Possible options for stage 2

- High p_T^H bin can be split further (in particular if evidence for new heavy particles arises)
- Low p_T^H region can be split further to further reduce any theory dependence there
- Further split $N_j \geq 2$ into $N_j = 2$ and $N_j \geq 3$

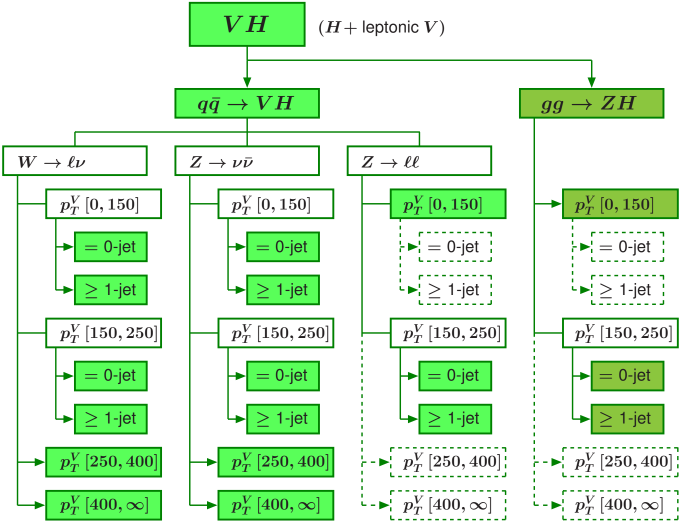
VBF – Stage 2.



Possible options for stage 2

- Add sensitivity to CP odd contributions
- Rest: Further separate out looser VBF cuts and/or 0+1 jet
- Further separate high p_T^{j1}

VH – Stage 2.



Possible options for stage 2

- Separate Z decays, further split high p_T^V