LHCHXSWG-2015-005

February 22, 2016

Simplified template cross sections

A. B[1,*]

[1] *C. D. University*

[*] AB@cern.ch

Disclaimer: Work in progress. Feedback and comments are very welcome.

0.1 Simplified Template Cross Sections

0.1.1 Overview

After the successful Higgs coupling measurements during the LHC Run1, which had as their main results measured signal strength and multiplicative coupling modifiers, it is important to discuss in which way the experiments should present and perform Higgs coupling measurements in the future. Simplified template cross sections were developed to provide a natural way to evolve the signal strength measurements used during Run1. Compared to the Run1 coupling measurements, the simplified template cross section framework allows to reduce in a systematic fashion the theory dependences that must be directly folded into the measurements. This includes both the dependence on the theoretical uncertainties in the SM predictions as well as the dependence on the underlying physics model (i.e. the SM or BSM models). In addition, they provide more finely-grained measurements (and hence more information for theoretical interpretations), while at the same time allowing and benefitting from the global combination of measurements in all decay channels.

The primary goals of the simplified template cross section framework are to maximize the sensitivity of the measurements while at the same time to minimize their theory dependence. This means in particular

- combination of all decay channels
- measurement of cross sections instead of signal strengths, in mutually exclusive regions of phase space
- cross sections are measured for specific production modes (with the SM production serving as kinematic template)

- measurements are performed in abstracted/simplified fiducial volumes
- allow the use of advanced analysis techniques such as event categorization, multivariate techniques, etc.

The measured exclusive regions of phase space, called "bins" for simplicity, are specific to the different production modes. Their definitions are motivated by

- minimizing the dependence on theoretical uncertainties that are directly folded into the measurements
- maximizing experimental sensitivity
- isolation of possible BSM effects
- minimizing the number of bins without loss of experimental sensitivity

These will of course be competing requirements in some cases and some compromise has to be achieved. The implementation of these basic design principles is discussed in more detail below.

A schematic overview of the simplified template cross section framework is shown in Fig. 1. The experimental analyses shown on the left are very similar to the Run1 coupling measurements. For each decay channel, the events are categorized in the analyses, and there are several motivations for the precise form of the categorization. Typically, a subset of the experimental event categories is designed to enrich events of a given Higgs production mode, usually making use of specific event topologies. This is what eventually allows the splitting of the production modes in the global fit. Another subset of event categories is defined to increase the sensitivity of the analysis by splitting events accordings to their expected signal-to-background ratio and/or invariant-mass resolution. In other cases, the categories are motivated by the analysis itself, e.g. as a consequence of the backgrounds being estimated specifically for certain classes of events. While these are some of the primary motivations, in the future the details of the event categorization can also be optimized in order to give good sensitivity to the simplified template cross sections to be measured.

The center of Fig. 1 shows a sketch of the simplified template cross sections, which are determined from the experimental categories by a global fit that combines all decay channels and which represent the main results of the experimental measurements. They are cross sections per production mode, split into mutually exclusive kinematic bins for each of the main production modes. In addition, the different Higgs decays are treated by fitting ratios of the partial widths.

The measured simplified template cross sections together with the ratios of decay widths then serve as input for subsequent interpretations. Such interpretations could for example be the determination of signal strength modifiers or coupling scale factors κ (providing compatibility with earlier results), EFT coefficients, tests of specific BSM models, and so forth. For this the experimental results should quote the full covariance among the different bins. By aiming to minimize the theory dependence that is folded into the first step of determining the simplified template cross sections from the event categories, this theory dependence is shifted into the second interpretation step, making the measurements more long-term useful. For example, the treatment of theoretical uncertainties can be decoupled from the measurements and can be dealt with at the interpretation stage. Propagating improvements in theoretical predictions and their uncertainties into the measurements itself, which is a very time-consuming procedure and unlikely to be feasible for older datasets, becomes much less important. Propagating future theoretical advances into the interpretation, on the other hand, is generally much easier.

To increase the sensitivity to BSM effects, the simplified template cross sections can be interpreted together with e.g. POs in Higgs boson decays. To make this possible, the experimental and theoretical correlations between the simplified template cross sections and the decay POs would need to be evaluated and taken into account in the interpretation. This point will not be expanded on further in this section, but would be interesting to investigate in the future.

While the simplified cross section bins have some similarity to a differential cross section measurement, they aim to combine the advantages of the signal strength measurements and fiducial and differential measurements. In particular, they are complementary to full-fledged fiducial and differential measurements and are neither designed nor meant to replace these. Fully fiducial differential measurements are of course essential but can only be carried out in a subset of decay channels in the foreseeable future. They are explicitly optimized for maximal theory independence. In practice, this means that in the measurements acceptance corrections are minimized, typically, simple selection cuts are used, and the measurements are unfolded to a fiducial volume that is as close as possible to the fiducial volume measured for a particular Higgs decay channel. In contrast, simplified template cross sections are optimized for sensitivity while reducing the dominant theory dependence in the measurement. In practice, this means that simplified

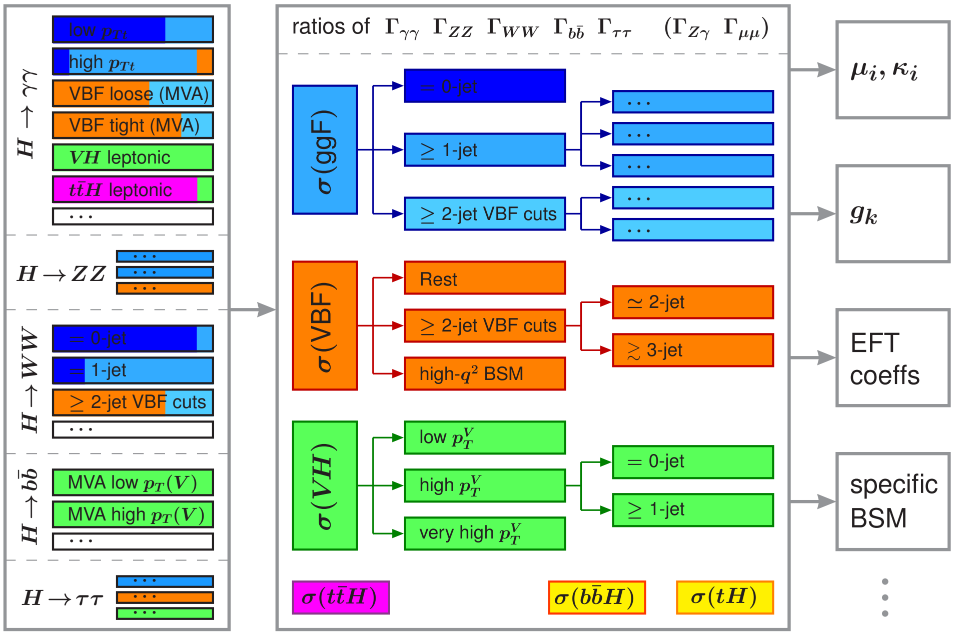

Figure 1: Schematic overview of the simplified template cross section framework.

fiducial volumes are used and larger acceptance corrections are allowed in order to maximally benefit from the use of event categories and multivariate techniques. They are also inclusive in the Higgs decay to allow for the combination of the different decay channels. The fiducial and differential measurements are designed to be agnostic to the production modes as much as possible, while the separation into the production modes is an important part of the design of the simplified template cross sections.

0.1.2 Guiding principles in the definition of simplified template cross section bins

As outlined above, several considerations have been taken into account in the definition of the simplified template cross section bins.

One important design goal is to reduce the dependence of the measurements on theoretical uncertainties in SM predictions. This has several aspects. First, this requires avoiding that the measurements have to extrapolate from a certain region in phase space to the full (or a larger region of) phase space when this extrapolation carries nontrivial or sizeable theoretical uncertainties. A example is the case where an event category selects an exclusive region of phase space, such as an exclusive jet bin. In this case, the associated theoretical uncertainties can be largely avoided in the measurement by defining a corresponding truth jet bin. The definition of the bins is preferably in terms of quantities that are directly measured by the experiments to reduce the needed extrapolation.

There will of course always be residual theoretical uncertainties due to the experimental acceptances for each truth bin. Reducing the theory dependence thus also requires to avoid cases with large variation in the experimental acceptance within one truth bin, as this would introduce a direct dependence on the underlying theoretical distribution in the simulation. If this becomes an issue, the bin can be further split into two or more smaller bins, which reduces this dependence in the measurement and moves it to the interpretation step.

To maximize the experimental sensitivity, the analyses should continue to use event categories primarily optimized for sensitivity, while the definition of the truth bins should take into consideration the experimental requirements. However, in cases where multivariate analyses are used in the analyses, it has to be carefully checked and balanced against the requirement to not introduce theory dependence, e.g., by selecting specific regions of phase space.

Another design goal is to isolate regions of phase space, typically at large kinematic scales, where BSM effects could be potentially large and visible above the SM background. Explicitly separating these also reduces the dependence of the measurements on the assumed SM kinematic distribution.

In addition, the experimental sensitivity is maximized by allowing the combination of all decay channels, which requires the framework to be used by all analyses. To facilitate the experimental implementation, the bins should be mutually exclusive to avoid introducing statistical correlations between different bins. In addition, the number of bins should be kept minimal to avoid technical complications in the individual analyses as well as the global fit, e.g. in the evaluation of the full covariance matrix. For example, each bin should typically have some sensitivity from at least one event category in order to avoid the need to statistically combine many poorly or unconstrained measurements. On the other hand, in BSM sensitive bins also limits are already very useful for the theoretical interpretation.

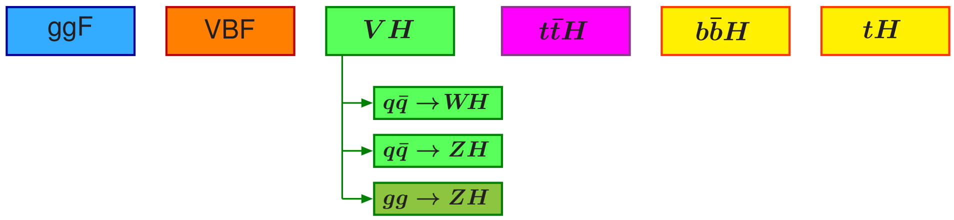

Figure 2: Stage 0 bins.

Staging

In practice, it will be impossible to define a set of bins that satisfies all of the above requirements for every analysis. Some analyses will only be able to constrain a subset of all bins or only constrain the sum of a set of bins. In addition, the number of bins that will be possible to measure increases with increasing amount of available data. For this reason, several stages with an increasing number of bins are defined. The evolution from one stage to the next can take place independently for each production mode.

Stage 0 Stage 0 is summarized in Fig. 2 and corresponds most closely to the measurement of the production mode μ in Run1. At this stage, each main production mode has a single inclusive bin, with associated Higgs production separated into $q\bar{q} \to WH$, $q\bar{q} \to ZH$ and $gg \to ZH$ channels.

Stage 1 Stage 1 defines a binning that is targeted to be used by all analyses on an intermediate time scale. In principle, all analyses should aim to eventually implement the full stage 1 binning. If necessary, intermediate stages to reach the full stage 1 binning can be implemented by a given analysis by merging bins that cannot be split. In this case, the analysis should ensure that the merged bins have similar acceptances, such that the individual bins can still be determined in an unbiased way in the global combination of all channels. In the diagrams presented below, the possibilities for merging bins are indicated by "(+)".

Stage 2 Defining the stage 2 binning in full detail is very difficult before having gained experience with the practical implementation of the framework with the stage 1 binning. Therefore, instead of giving a detailed proposal for the stage 2 binning, we only give indications of interesting further separation of bins that should be considered for the stage 2 binning.

0.1.3 Definition of leptons and jets

The measured event categories in all decay channels are unfolded by the global fit to the simplified template cross sections bins. For this purpose, and for the comparison between the measured bins and theoretical predictions from either analytic calculations or MC simulations, the truth final state particles need to be defined unambiguously. The definition of the final state particles, leptons, jets, and in particular also the Higgs boson are explicitly kept simpler and more idealized than in the fiducial cross section measurements. Treating the Higgs boson as a final state particle is what allows the combination of the different decay channels.

Higgs boson

The simplified template cross sections are defined for the production of an on-shell Higgs boson, and the unfolding should be done accordingly. A global cut on the Higgs rapidity at $|Y_H| < 2.5$ is included in all bins. As the current measurements have no sensitivity beyond this rapidity range, this part of phase space would only be extrapolated by the MC simulation. On the other hand, it is in principle possible to use forward electrons (up to $|\eta|$ of 4.9) in $H \to ZZ^* \to 4\ell$ and extend the accessible rapidity range. For this purpose, an additional otherwise inclusive bin for $|Y_H| > 2.5$ can be included.

Jets

Truth jets are defined as anti-k_t jets with a jet radius of $R = 0.4$, including all stable particles associated to the jet, including neutrinos and leptons from hadron decays. Stable here has the usual definition, with a lifetime greater than 10 ps, i.e. those particles that are passed to GEANT in the experimental simulation chain. All decay products from the Higgs boson decay should be removed as they are accounted for by the truth Higgs boson. By default, truth jets are defined without restriction on their rapidity. A possible rapidity cut can be included in the bin definition. A common p_{T} threshold for jets at 25 or 30 GeV should be used for all truth jets. A lower threshold would in principle have the advantage to split the events more evenly between the different jet bins. Experimentally, a higher threshold at 30 GeV is favored due to pile up. **Still to be decided after final feedback from the experiments.**

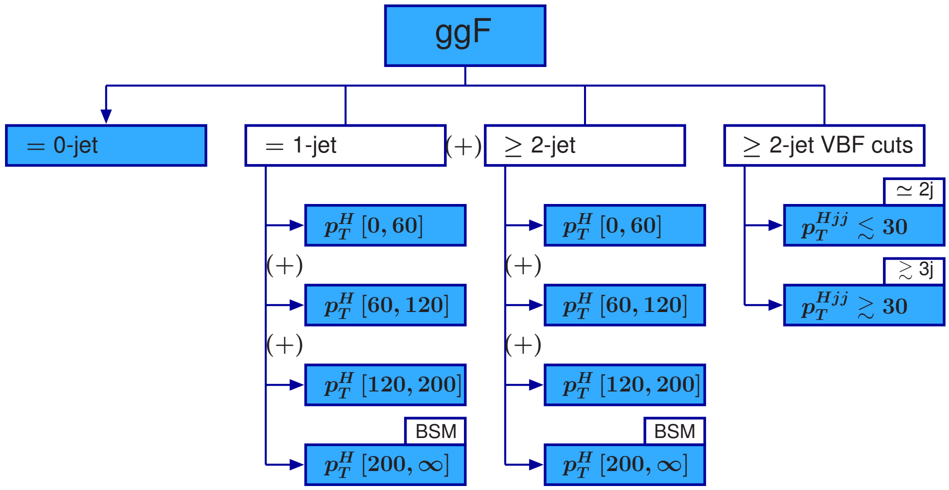

Figure 3: Stage 1 binning for gluon fusion production. **Status Jan14 as presented at the workshop.**

Leptons

Electrons and muons from vector boson decays in VH production are defined as dressed, i.e. FSR photons should be added back to the electron or muon. τ leptons are defined from the sum of their decay products. There should be no restriction on the transverse momentum or the rapidity of the leptons. That is, for a leptonically decaying vector boson the full decay phase space is included.

0.1.4 Bins for $gg \to H$ production

Stage 0 Inclusive gluon fusion cross section within $|Y_H| < 2.5$.

Stage 1 The Stage 1 binning is depicted in Fig. 3 and summarized as follows:

- Split into jet bins: $N_j = 0$, $N_j = 1$, $N_j \geq 2$, $N_j \geq 2$ with VBF selection cuts (defined with the same cuts as the corresponding bin in VBF production **(include p_T^{j1} cut here as well?)**. The jet bins are motivated by the use of jet bins in the experimental analyses. Introducing them also for the simplified template cross sections avoids folding the associated theoretical uncertainties into the measurement.

The separation of the $N_j \geq 2$ with VBF cuts is motivated by the wish to separately measure the gluon fusion contamination in the VBF selection. If the fit has no sensitivity to determine the gluon fusion and the VBF contributions with this topology, the sum of the two contributions can be quoted as result.

- The $N_j \geq 2$ with VBF cuts bin is split further into an exclusive 2-jet-like and inclusive 3-jet-like bin. The split is implemented by a cut on $p_{\mathrm{T}}^{Hjj} = |\vec{p}_{\mathrm{T}}^{H} + \vec{p}_{\mathrm{T}}^{j1} + \vec{p}_{\mathrm{T}}^{j2}|$ at around 30 GeV **(To be finalized, could be lowered to 25 GeV)**. This variable is chosen as a compromise between the different kinematic variables used by different channels to enrich VBF production. (In particular the kinematic variables $\Delta\phi_{H-jj}$ and p_{T}^{j3} are both correlated with p_{T}^{Hjj}). This cut is explicitly included here since it induces nontrivial theory uncertainties in the gluon-fusion contribution.

- The $N_j = 1$ and $N_j \geq 2$ bins are further split into p_{T}^{H} bins.
  - 0 GeV $< p_{\mathrm{T}}^{H} <$ 60 GeV: The boson channels have most sensitivity in the low p_{T}^{H} region. The upper cut is chosen as low as possible to give a more even split of events but at the same time high enough that no resummation effects are expected. The cut should also be sufficiently high that the jet p_{T} cut introduces a negligible bias.
  - 60 GeV $< p_{\mathrm{T}}^{H} <$ 120 GeV: This is the resulting intermediate bin between the low and high p_{T}^{H} regions. The lower cut here is high enough that this bin can be safely treated as a hard $H+j$ system in the theoretical description.
  - 120 GeV $< p_{\mathrm{T}}^{H} <$ 200 GeV: The boosted selection in $H \to \tau\tau$ contributes to the high p_{T}^{H} region. Defining a separate bin avoids large extrapolations for the $H \to \tau\tau$ contribution. For $N_j =$ 2, this bin likely provides a substantial part of the gluon-fusion contribution in the hadronic VH selection.
  - $p_{\mathrm{T}}^{H} >$ 200 GeV: Beyond the top-quark mass, the top-quark loop gets resolved and top-quark mass effects become relevant. Splitting off the high-p_{T}^{H} region ensures the usability of the heavy-top expansion for the lower-p_{T}^{H} bins. At the same time, the high p_{T}^{H} bin in principle offers the possibility to distinguish a pointlike ggH vertex induced by heavier BSM particles in the loop from the resolved top-quark loop.

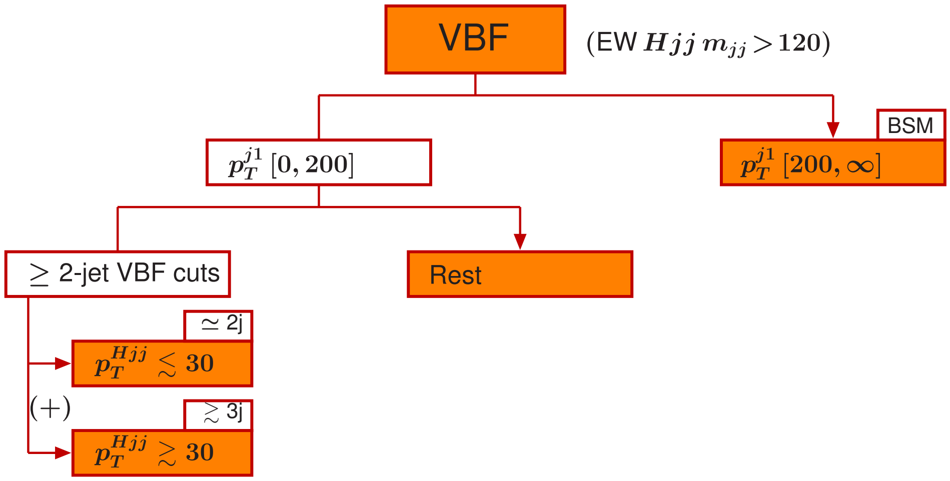

Figure 4: Stage 1 binning for vector boson fusion production.

At intermediate stages, all lower three p_{T}^H bins, or any two adjacent bins, can be merged. Alternatively or in addition the $N_j = 1$ and $N_j \geq 2$ bins can be merged by individual analyses as needed, and potentially also when the combination is performed at an intermediate stage.

Stage 2 In Stage 2, the high p_T^H bin should be split further, in particular if evidence for new heavy particles arises. In addition, the low p_{T}^H region can be split further to reduce any theory dependence there. If desired by the analyses, another possible option is to further split the $N_j \geq 2$ bin into $N_j = 2$ and $N_j \geq 3$.

0.1.5 Bins for VBF production

To avoid any potential confusion between VBF production and associated production with a hadronically decaying vector boson, VBF production is defined as electroweak production of Hjj with $m_{jj} > 120$ GeV.

Stage 0 Inclusive vector boson fusion cross section within $|Y_H| < 2.5$.

Stage 1 The Stage 1 binning is depicted in Fig. 4 and summarized as follows:

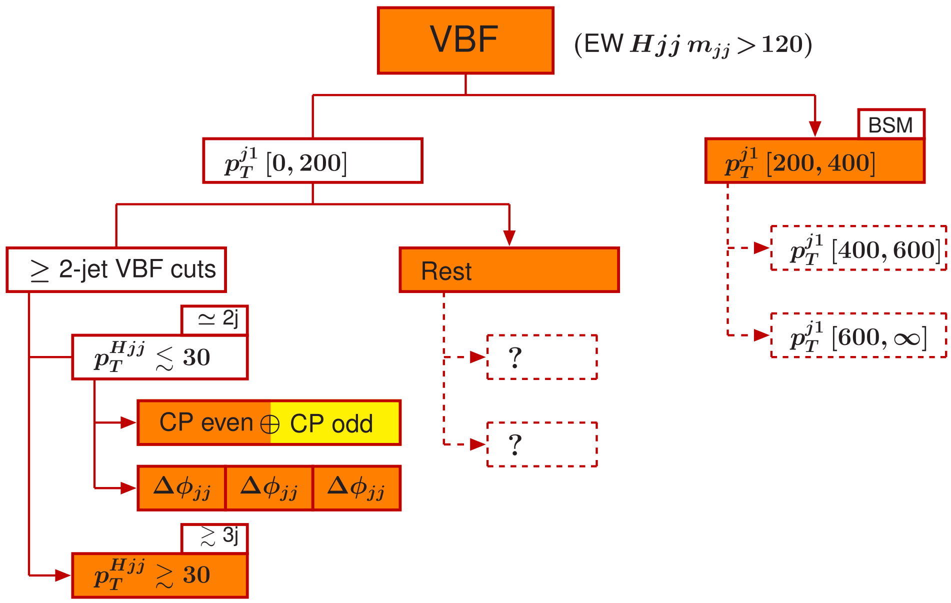

Figure 5: Possible Stage 2 binning for vector boson fusion production.

- VBF events are split by p_T^{j1}, the transverse momentum of the highest-p_T jet. The lower p_T^{j1} region is expected to be dominated by SM-like events, while the high-p_T^{j1} region is sensitive to potential BSM contributions. The suggested cut is at 200 GeV, to keep the fraction of SM events in the BSM bin small.

- The $p_\mathrm{T}^{j1} < 200$ GeV bin is split further into a bin with typical VBF topology and all remaining events ("rest"). The proposed VBF selection cuts are $m_{jj} > 400$ GeV, $\Delta\eta_{jj} > 2.8$, **(okay? explicit p_T^j and η_j cuts?)** which should provide a good intermediate compromise among the various VBF selection cuts employed by different channels. The "rest" bin can be sensitive to certain BSM contributions that do not follow the typical SM signature with two forward jets.

- The bin with typical VBF topology is split into an exclusive 2-jet-like and inclusive 3-jet-like bin using a cut on p_T^{Hjj} at around 30 GeV **or maybe lower at 25? To be finalized**. See the corresponding discussion for gluon-fusion production.

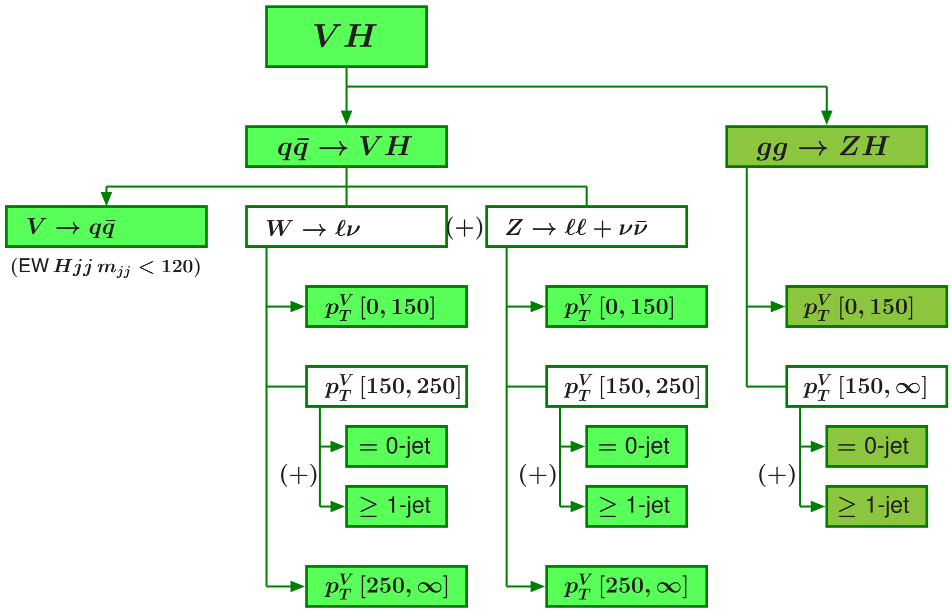

Figure 6: Stage 1 binning for vector boson fusion production.

Stage 2 More splits are introduced at Stage 2 as illustrated in Fig. 5. While the details require more discussion and cannot be finalized at the present, this could include

- The high-p_{T}^{j1} bin can be split further by separating out very high-p_{T}^{j1} events for example with additional cuts at 400 GeV and 600 GeV.
- The "rest" bin can be split further, e.g., by explicitly separating out a looser VBF selection.
- For the $N_j = 2$ VBF topology bin can be split further to gain sensitivity to CP odd contributions, e.g. by splitting into sub-bins of $\Delta\phi_{jj}$ or by measuring a continuous parameter.

0.1.6 Bins for VH production

Stage 0 Inclusive associated production with vector bosons cross section within $|Y_H| < 2.5$.

Stage 1 The Stage 1 binning is depicted in Fig. 6 and summarized as follows:

- VH production is first split into production from $q\bar{q}$ and gg initial state.
  - Production from $q\bar{q}$ initial state is split into three V final states: hadronic $V \to jj$, defined as electroweak Hjj production with $m_{jj} < 120$ GeV, $W \to \ell\nu$ and $Z \to \ell\ell + \nu\bar{\nu}$
  - $W \to \ell\nu$ and $Z \to \ell\ell+\nu\bar{\nu}$ are split further into bins of p_T^V, aligned with the quantity used in the $H \to b\bar{b}$ analysis, which is one of the main contributors to the VH bins.
    - $p_T^V < 150$ GeV receives contributions from the bosonic decay channels and from $H \to b\bar{b}$ with $W \to \ell\nu$ and $Z \to \ell\ell$, which do not rely on E_T^{miss} triggers
    - $150 \text{ GeV} < p_T^V < 250$ GeV receives contributions from $H \to b\bar{b}$ with $Z \to \nu\bar{\nu}$ due to the high threshold of the E_T^{miss} trigger
      - This bin is split further into a $N_j = 0$ and a $N_j \geq 1$ bin, reflecting the different experimental sensitivity and to avoid the corresponding theory dependence.
    - $p_T^V > 250$ GeV is sensitive to BSM contributions.
  - Production from gg initial state is split in analogy to production from the $q\bar{q}$ initial state, apart from the $p_T^V > 250$ GeV bin, which is not split out.

Stage 2 More splits are introduced at Stage 2 as illustrated in Fig. 5. While the details need more discussion, this could include

- Split of the $Z \to \ell\ell + \nu\bar{\nu}$ into $Z \to \ell\ell$ and $Z \to \nu\bar{\nu}$
- Split of the $p_T^V < 150$ GeV into into a $N_j = 0$ and a $N_j \geq 1$ bin, potentially apart from in the $Z \to \ell\ell$, which will suffer from the low $Z \to \ell\ell$ branching ratio
- Split of the $p_T^V > 250$ GeV bin into $p_T^V < 400$ GeV and $p_T^V > 400$ GeV, to increase the sensitivity to BSM contributions with very high p_T^V, potentially apart from the $Z \to \ell\ell$
- Potentially analoguous splits for the production from gg initial state

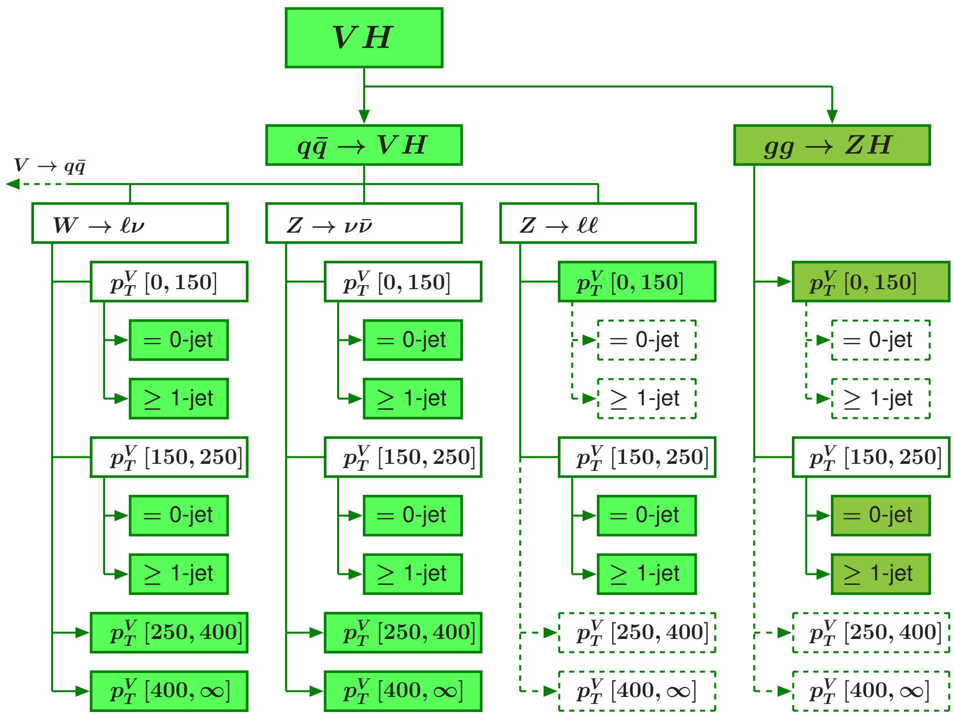

Figure 7: Simplified template cross section definitions for production in association with vector bosons (stage 2).

0.1.7 Treatment of $t\bar{t}H$ production

Stage 0 Inclusive $t\bar{t}H$ production with $|Y_H| < 2.5$.

Stage 1 Currently no additional splits beyond Stage 0 are foreseen. One option might be to separate different top decay channels.

Stage 2 In the long term it could be useful to split into bins with 0 and ≥ 1 additional jets or one or more bins tailored for BSM sensitivity.

0.1.8 Treatment of $b\bar{b}H$ and tH production

In the foreseeable future, there will only be one inclusive bin for $b\bar{b}H$ production and only one inclusive bin for tH production.

0.1.9 Summary

Simplified template cross sections provide a way to evolve the signal strength measurements that were performed during LHC Run1, by reducing the theoretical uncertainties that are directly folded into the measurements and by providing more finely-grained measurements, while at the same time allowing and benefitting from the combination of measurements in many decay channels. Several stages are proposed: stage 0 essentially corresponds to the production mode measurements of Run1 and stage 1 defines a first complete setup, with indications for potential bin merging when a given channel cannot yet afford the full stage 1 granularity. A complete proposal for the stage 2 binning will need to be based on experience of using the simplified template cross section framework in real life, but some indications of what could be interesting are given here.