Statistical Procedure for $H \to ZZ \to lljj$ and $H \to ZZ \to ll\nu\nu$

1 Statistical Procedure

For each Higgs mass hypothesis m_H a one-sided upper-limit is placed on the standardized cross-sections $\mu = \sigma/\sigma_{SM}$ at the 95% confidence level. The upper limit is based on a binned distribution of m_{llqq} or m_T, and the statistical modeling of the distributions is similar in both cases. The likelihood function is based on a product of Poisson probabilities for each bin based on the expected number of events, which is parametrized by μ and several nuisance parameters α_i corresponding to the various systematic effects. The likelihood function is given by

$$\mathcal{L}(\mu, \alpha_i) = \prod_{m \in \text{bins}} \text{Pois}(n_m|\mu_m) \prod_{i = \in \text{Syst}}^{5} N(\alpha_i) \tag{1}$$

where m is an index over the bins of the template histograms, i is an index over systematic effects, μ_m is the expected number of events in bin m given by

$$\mu_m = \mu L \eta_1(\boldsymbol{\alpha})\, \sigma_{1m}(\boldsymbol{\alpha}) + \sum_{j \in \text{Bkg Samp}} L \eta_j(\boldsymbol{\alpha})\, \sigma_{jm}(\boldsymbol{\alpha}), \tag{2}$$

where $\eta_j(\boldsymbol{\alpha})$ parametrizes uncertainty in the overall normalization and $\sigma_{jm}(\boldsymbol{\alpha})$ parametrizes uncertainties in the shape of the distribution of the discriminating variable. The nuisance parameters α_i are associated to the source of the systematic effect (eg. the jet energy scale uncertainty), while $\eta_j(\boldsymbol{\alpha})$ and $\sigma_{jm}(\boldsymbol{\alpha})$ represent the effect of that uncertainty. The α_i are scaled so that $\alpha_i = 0$ corresponds to the nominal expectation and $\alpha_i = \pm 1$ correspond to the $\pm 1\sigma$ variations of the source. The effect of these variations is quantified by dedicated studies that provide $\eta_{ij}^{\pm}$ and $\sigma_{ijm}^{\pm}$, which are then used to form

$$\eta_j(\boldsymbol{\alpha}) = \prod_{i \in \text{Syst}} I(\alpha_i; \eta_{ij}^0, \eta_{ij}^+, \eta_{ij}^-) \tag{3}$$

and

$$\sigma_{jm}(\boldsymbol{\alpha}) = \prod_{i \in \text{Syst}} I(\alpha_i; \sigma_{ijm}^0, \sigma_{ijm}^+, \sigma_{ijm}^-) \tag{4}$$

with

$$I(\alpha; I^0, I^+, I^-) = \begin{cases} I^0 + \alpha(I^+ - I^0) & \text{if } \alpha > 0 \\ I^0 & \text{if } \alpha = 0 \\ I^0 - \alpha(I^- - I^0) & \text{if } \alpha < 0 \end{cases} \tag{5}$$

enabling piece-wise linear interpolation in the case of asymmetric response to the source of systematic.

Explicit dependence on the nuisance parameters is eliminated by forming the profile likelihood ratio

$$\lambda(\mu) = \frac{\mathcal{L}(\mu, \hat{\hat{\alpha}}_i(\mu))}{\mathcal{L}(\hat{\mu}, \hat{\alpha}_i)}, \tag{6}$$

where $(\hat{\mu}, \hat{\alpha}_i)$ are the unconstrained maximum likelihood estimates and $\hat{\hat{\alpha}}_i(\mu)$ is the conditional maximum likelihood estimate of the nuisance parameters with μ fixed. Since we are interested in one-sided upper-limits, we would not consider $\hat{\mu} > \mu$ evidence against μ. Thus, we employ the test statistic

$$q_\mu = \begin{cases} -2\log\lambda(\mu) & \text{if } \hat{\mu} \leq \mu \\ 0 & \text{if } \hat{\mu} > \mu \end{cases} \tag{7}$$

and use the asymptotic distributions of this test statistic to infer the 95% confidence limits.

We have compared the limits from the asymptotic distributions to those from the equivalent calculation calibrated with pseudo-experiments and found consistency. In addition to the observed limits, we also present the median, $\pm 1\sigma$ and $\pm 2\sigma$ bands of the upper limits obtained from background pseudo-experiments. Note, that the background-only pseudo-experiments were generated assuming $(\mu = 0, \hat{\hat{\alpha}}_i(\mu = 0))$, where the profiling is based on the observed data.