

Additional Information on the RA2 “Data vs. Combined Background” Plots

August 18, 2011

This documentation refers to the search for Supersymmetry in the jets and missing momentum final state¹ (“RA2”). In Fig. 9, the H_T and \cancel{E}_T distributions of the 986 candidate events observed after the baseline selection are compared to the Standard Model background expectation derived from data. The individual background contributions as well as the ratio data to total background expectation are shown.

The integrated expectations from the individual background contributions and their statistical and systematic uncertainties are listed in Tab. 10. The total integrated background expectation 927.5 ± 103.2 has been computed taking into account the actual probability densities of the individual contributions.

However, the available input for Fig. 9 consists of one histogram per background contribution containing the number of predicted events per bin; uncertainties or correlations between bins or different backgrounds are not provided. This is too little information to properly combine the different backgrounds and their uncertainties per bin. Therefore, certain approximations were made to obtain Fig. 9. They are aimed at providing an as accurate as possible statement on how the agreement between data and background as well as the uncertainties evolve with H_T and \cancel{E}_T , respectively.

The idea is, in each bin, to estimate a statistical component of the uncertainty, which reflects the number of entries, and a constant relative systematic component. The overall normalisation of both should be such that, if integrated, the total uncertainty from Tab. 10 is regained. The procedure is as follows:

1. the total background expectation in each bin is computed as the linear sum of the individual background contributions. The resulting total integrated background of 913.7 events is scaled to the properly combined yield from Tab. 10, i. e. 927.5;
2. The total uncertainty from Tab. 10., $\sigma_{\text{tot}} = 103.2$, is split into a statistical, $\sigma_{\text{tot}}^{\text{stat}}$, and a systematic, $\sigma_{\text{tot}}^{\text{syst}}$, component. This is done by summing up the listed statistical uncertainties on the individual backgrounds b in

¹CMS PAS SUS-11-004, *Search for supersymmetry in all-hadronic events with missing energy* (2011)

quadrature and subtracting them in quadrature from the total uncertainty,

$$\begin{aligned}\sigma_{\text{tot}}^{\text{stat}} &= \sqrt{\sum_b (\sigma_b^{\text{stat}})^2} = 43.0 , \\ \sigma_{\text{tot}}^{\text{syst}} &= \sqrt{(\sigma_{\text{tot}})^2 - (\sigma_{\text{tot}}^{\text{stat}})^2} = 93.7 ,\end{aligned}\tag{1}$$

where σ_b^{stat} are the statistical uncertainties on the individual background components. Hence, linear combination of the backgrounds and uncorrelated statistical and systematic uncertainties are assumed;

- the statistical uncertainty of the total background prediction in each bin i is set to

$$\sigma_i^{\text{stat}} = s \cdot \sqrt{n_i} ,\tag{2}$$

where n_i denotes the total number of predicted background events in that bin. s is a global scale factor which assures that the integrated total statistical uncertainty from summing up the bin contributions quadratically amounts to $\sigma_{\text{tot}}^{\text{stat}}$ from (1),

$$\begin{aligned}s &= \frac{\sigma_{\text{tot}}^{\text{stat}}}{\sqrt{\sum_i n_i}} \\ \sigma_i^{\text{stat}} &= \sqrt{\frac{n_i}{\sum_i n_i}} \cdot \sigma_{\text{tot}}^{\text{stat}} ,\end{aligned}\tag{3}$$

- the systematic uncertainty of the total background prediction in each bin i is set to the constant relative value

$$\sigma_i^{\text{syst}} = \frac{n_i}{\sum_i n_i} \cdot \sigma_{\text{tot}}^{\text{syst}} .\tag{4}$$

In conclusion, the total number of predicted background events per bin in Fig. 9 as well as their statistical and systematic uncertainties have been approximated from several sources: the average number of events predicted in each bin by the individual background methods, the integrated statistical and systematic uncertainties on each background prediction, as well as the properly combined, total integrated background yield and its total uncertainty. For the approximation, essentially the proper profile likelihood based combination is ignored and replaced by a linear combination assuming uncorrelated backgrounds. However, the differences are small in the baseline region. While the size of the total uncertainty in each bin might be underestimated slightly, in particular at medium H_T and \cancel{H}_T , respectively, the total integrated uncertainty corresponds to $\sigma_{\text{tot}} = 103.2$ as cited in Tab. 10 and a realistic impression of the evolution of the expected background and its uncertainty is gained.