CHAPTER 9

SYSTEMATICS

This chapter describes the systematic uncertainties included in the analysis. As will be discussed in Chapter 10, the systematics are handled as nuisance parameters in the likelihood fit. Systematics can affect either the normalization of the contributions (i.e. a flat scaling of all of the bins in a histogram) or both the normalization and shape of the contributions (i.e. each bin in the histogram may be affected differently). As introduced in Section 10.2, the `Combine` tool is used to perform the statistical analysis; systematics that only impact the normalization are handled as "rate" systematics (i.e. numbers in the `Combine` datacard), while template histograms (for the up and down variations) are used to handle the "shape" systematics.

A discussion of the EFT dependence of the shape systematics is provided in Section 9.1. The full set of systematic uncertainties included in this analysis are listed in the subsequent sections. In Section 9.2 the systematics arising from experimental sources are discussed; these include uncertainties on the data-to-MC corrections described in Chapter 7, uncertainties on the background estimation, and various other sources. The systematics associated with the theoretical aspects of the modeling are listed in Section 9.3.

9.1 EFT dependence of the systematic uncertainties

Like the histograms that correspond to the nominal contribution, the histograms that correspond to the up and down variations of the shape systematics will carry 26-dimensional quadratic dependence on the WCs. However, it should be noted that the quadratic dependence of the up and down histograms may be different from the nominal histogram. To understand why this is true, let us consider an "up" variation of a SF. We recall from Chapter 3 that each generated event is characterized by a unique 26-dimensional quadratic function. In the case of a nominal SF, the function for a given event will be scaled by the nominal SF value for that event; in the case of an "up" variation, the function will be scaled by the "up" value of that SF for that event. This will be similarly true for each of the subsequent events, and in general the difference between nominal and "up" will vary from event to event. For a given bin, the quadratic parametrization corresponds to the sum of the quadratics of each of the events that pass the selection criteria for the bin. Since each of the quadratics in the nominal sum will differ from each of the quadratics in the "up" sum, the nominal and "up" quadratic parameterization will in general have different shapes.

For the systematic uncertainties associated with the generation of the MC samples (described in Section 9.3 below), there is an additional complication to consider. The up and down weights for these systematics are calculated by MadGraph at the starting point of the sample. In principle, there is no reason to assume that the up and down weights at any other point in the WC space must be the same as at the starting point. For this reason, it would be most correct to generate a dedicated sample at every point in space that we are interested in (or to modify MadGraph to calculate the up and down weights at each reweight point). These approaches are currently infeasible at scale, but small-scale studies suggest that the dependence of the systematics on the WC space is not large [31].

A more detailed discussion of the EFT dependence of the systematic uncertainties is provided in Ref. [42]; this presentation includes a discussion of how the approach implemented in this analysis differs from the approach utilized in the predecessor to this analysis (Ref. [54]), and how both of these approaches differ from the fundamentally correct approach.

Luminosity: The total uncertainty on the total luminosity for the 2016, 2017, and 2018 data-taking periods is 1.6% [22, 23, 53]. This uncertainty affects only the rate (not the shape). The systematic is correlated across years.

PU correction: The uncertainty for the PU correction (Section 7.1) is obtained by varying the proton-proton cross section used to estimate the data PU histograms by 4.6% (which corresponds to a 1σ variation). This systematic is correlated across years.

Trigger efficiency correction: The uncertainty on the trigger efficiency corrections (Section 7.2) is taken to be 2%, which is a conservative estimate that includes the effects of the dependence on the phase space used to perform the measurement and the effects of the correlation between the $E_{\mathrm{T}}^{\mathrm{miss}}$ triggers and the analysis triggers. This systematic is uncorrelated across years.

Lepton identification efficiency: As described in Section 7.3 lepton identification efficiencies are computed with the tag and probe approach. There are several sources of uncertainty that contribute to this measurement, including the statistical uncertainty, the uncertainty on the signal modeling (which is estimated by comparing the results obtained with LO and NLO DY samples) and the uncertainty on the functions used in the fitting (which is estimated by performing the fits with alternative models). The total uncertainty for this correction is the quadrature sum of these sources. There is a separate uncertainty for electrons and for muons; the systematics are correlated across years. The full details regarding the systematic uncertainties on the lepton SFs are available in Appendix A of [9].

b-tagging correction: The per-jet b-tagging SFs (Section 7.4) are different for heavy (b,c) and light jets, so a separate systematic uncertainty is included for each. For both heavy and light uncertainties, a component that is correlated and uncorrelated (across years) is considered, bringing the total number of b-tagging uncertainties to 10.

Prefire correction: The uncertainty on the L1 prefire correction (Section 7.5) takes into account the uncertainties on the prefire probabilities (20%) and the statistical uncertainty of the given bin. The uncertainty is correlated across years.

Jet energy corrections: There are numerous uncertainties associated with the corrections to the jet energy scale described in Section 7.6. These uncertainties can be grouped into categories defined by the CMS JERC group [19]. Of these groupings of uncertainty sources, we include the `Absolute` uncertainty (a combination of sources from the PU offset correction and simulated response correction), the `BBEC1` uncertainty (a combination of sources from the PU offset correction, relative p_T differences, relative resolution differences, and statistical uncertainty), the `RelativeBal` uncertainty (which accounts for differences in different methods of the p_T balance calculations), the `RelativeSample` uncertainty (which accounts for differences in the residual data to simulation differences obtained with dijet and Z+jet approaches), and the `FlavorQCD` uncertainty (which accounts for differences in responses to different jet flavors). We treat these sources of uncertainty as correlated across all data-taking periods. In total, we thus include five nuisance parameters for the jet energy scale corrections.

The JER uncertainty is obtained by shifting the scale and resolution applied and is uncorrelated across data-taking periods, for a total of four nuisance parameters [18].

Charge misidentification: As described in Section 8.2, a flat 30% rate uncertainty is applied to the charge flip contribution, and the systematic is correlated across years.

Nonprompt estimation: Several sources of uncertainties are associated with the nonprompt estimation. The first source corresponds to the uncertainty in the measurement of the probability for nonprompt leptons to pass the tight selection. As described in Section 8.1, this measurement involves performing a fit to a discriminating variable; the fit is performed in three ways, and the envelope of the results (including the statistical uncertainties) is taken as the uncertainty. This envelope uncertainty is fully correlated across all p_T and η bins in the fit. To account for effects that are not fully correlated across these variables, we also include two additional sources of uncertainty that cover the most extreme variations across the p_T bins and across the η bins. After the measurement has been performed, the results are compared against results obtained with MC simulations, and the residual differences are covered with an additional closure uncertainty (a separate closure uncertainty is considered for each UL period).

In summary, the systematic uncertainty on the nonprompt estimation includes seven total nuisance parameters corresponding to the envelope uncertainty, the envelope p_T uncertainty, the envelope η uncertainty, and the four closure uncertainties. In addition to these seven systematic uncertainties, the statistical uncertainty of the AR is also included (using the `Combine` tool's `autoMCStats` functionality [24]).

Diboson N_{jet}: This systematic is derived from the 3ℓ control region to account for the under-prediction of the MC for the high-jet-bin yields. To derive the uncertainty for each jet bin, we calculate the factor by which the diboson contribution would need to be scaled in order for the prediction to match the data; a linear function is then fit to this set of points, and we evaluate the linear function to find the uncertainty factor in each jet bin. For a given bin, the up/down shifts are determined by taking the difference between the diboson contribution and the diboson contribution scaled by the uncertainty factor. This systematic is applied only to the diboson process. The systematic is correlated across years.

Missing parton: As described in Section 3.2.1, an additional additional uncertainty is applied to the single-top processes, which are not generated with an additional parton. This uncertainty is computed by comparing the N_{jet} distribution of the private LO EFT samples (reweighted to the SM) against the centrally produced NLO samples listed in Table A.10. The missing parton systematic is taken to be the up/down shift required to cover the difference between the samples (such that when the missing parton uncertainty is included in quadrature with the other systematics, the difference between the samples is fully covered by the total uncertainty). The systematic is correlated across years.

9.3 Theoretical systematic uncertainties

Renormalization and factorization: The renormalization and factorization scales (μ_R and μ_F, respectively) are fluctuated up and down by a factor of 2. The weights for each variation are computed by the event generator during the production of the sample. As described in Section 9.1, these weights are calculated at the starting point of the sample.

The effects of the $\mu_\mathrm{R}/\mu_\mathrm{F}$ systematics are handled somewhat differently than the experimental systematics described above. In order to understand this difference, let us first recall Eq. 3.1. The up and down variations of the experimental SFs affect how likely an event is to pass the given selection, so these weights are only applied to the sum in the numerator of Eq. 3.1. However, the μ_R and μ_F variations are applied to both the numerator and denominator of Eq. 3.1 since these weights correspond to how often certain parts of the phase space are populated by MC events (so these variations affect all generated events, not just the ones passing the selection). Another way to think about this is to recall that we already include a systematic uncertainty on the cross section of the simulated samples; including the μ_R and μ_F variations in both the numerator and denominator effectively cancels the overall normalization effect (which is already covered by the cross section uncertainty), leaving us with the relevant shape effect of the systematics.

In the event generation, the μ_R and μ_F are independently varied (i.e. μ_R is varied up/down while μ_F is held at nominal and vice versa) and also varied together. This results in six total variations. The envelope of the variations (i.e. the most extreme variations with respect to nominal) is taken as the $\mu_\mathrm{R}/\mu_\mathrm{F}$ systematic. The systematic is correlated across all data-taking periods, so there is in total one nuisance parameter for this uncertainty.

Parton shower: The initial and final state radiation (ISR and FSR, respectively) scales are fluctuated up and down by a factor of $\sqrt{2}$. Similar to the μ_R and μ_F systematics, these variations are applied to both the numerator and denominator of Eq. 3.1. Both the ISR and the FSR systematics are treated as correlated across all data-taking periods.

Cross section: As discussed in Section 3.2, the simulated samples are normalized to NLO or NNLO cross sections where available. The uncertainties associated with the cross section measurements are applied as rate uncertainties to the relevant processes and are correlated across data-taking periods.

An exception to the above procedure is applied for the $t\bar{t}\gamma$ samples (used to estimate the conversion contribution); these samples are normalized to LO cross sections, as there is not an appropriate NLO calculation available. In order to account for the LO cross section uncertainty, the $\mu_{\mathrm{R}}/\mu_{\mathrm{F}}$, ISR, and FSR systematic variations are applied only to the numerator of Eq. 3.1 for this process. This allows the uncertainty on both the shape and normalization to be incorporated.