Search for gauge-mediated supersymmetry in events with photons and a Z boson decaying to charged leptons at CMS

von
Sebastian Wuchterl

Masterarbeit in Physik

vorgelegt der
Fakultät für Mathematik, Informatik und Naturwissenschaften der
RWTH Aachen

im
xx 2018

angefertigt im
I. Physikalischem Institut B

bei
Prof. Dr. Lutz Feld

Contents

Chapter 1

Analysis strategy and background estimation

Processes producing SUSY particles have significantly smaller cross sections the most of the SM processes. Therefore, a sophisticated understanding of the relevant production and decay scenarios is needed, in order to separate the potential interesting signal events from the huge amount of SM background. Since this search is targeting SUSY scenarios with bino and wini like NLSPs, finl states with photons and Z bosons are expected, Because the the dileptonic decay of the Z boson is investigated, there are not so much SM background processes leading to this final state. Especially the requirement of a photon and missing transverse momentum $p_{\mathrm{T}}^{\mathrm{miss}}$ reduces most of the background, so that e.g. the QCD background is negligible. Leftover are processes producing two same - flavor opposite charge (SFOC) leptons, one photon and missing transverse momentum. The most important ones are $t\bar{t}(\gamma)$ production, Drell-Yan/$Z()$, ZZ, and WZ diboson production. More detal can be found in the corresponding sections in Section 1.2. The key strategy of this analysis is, to impose as few as requiremens as necessary, to obtain a inclusive event selection, so that many model scenarios can be investigated. Hence, only the existence of all the expcted final state particles is required, where the Z boson is reconstructed from the two leptons. Therefore, the event selection can be rather loose, including loose requirements on the lepton and photon energies. Not many additional requirements are needed int he following, to obtain a sensitive signal region selection.

1.1. Event Selection

In this first section, the event selection will be explained, including the definitions important for the background prediction, which is based on MC simulation, where the simulation will be handled in dedicated control regions enriched for each group of backgrounds, a validation region to examine the background prediction, and finally a signal region, where two bin counting experiment will be performed. But first, the preselection will be introduced.

1.1.1. Preselection

The preselection acts as a first rough definition of the phase space that is of interest, and rejects inefficient parts of the used triggers. The preselection imposed on the dilepton triggered events is defined as follows:

- exactly one SFOC lepton pair (ee or $\mu\mu$),
- at least one photon,
- $\Delta R(\ell_1, \gamma) > 0.3$ and $\Delta R(\ell_2, \gamma) > 0.3$,
- $81 < m_{\ell\ell} < 101\,\text{GeV}$.

The first two conditions imply the existence of the final state particles, including the definition of the physics objects and lepton pair selection explained in Section **??**. The third requirement reduces contributions of final state radiation photons, that are radiated off by the leptons in the showering process. The last invariant dilepton mass requirement ensures that both leptons originate from a onshell Z boson, and reduces different contributions of SM backgrounds that produce dileptons with an invariant mass not in agreement with a Z boson.

1.1.2. Signal region

The signal region (SR) selection is optimized to maintain high sensitivity for various SUSY scenarios both with electroweak and strong production. For the models considered here, see Section **??**, the NLSP ($\widetilde{\chi}_1^0$) can decay to a Z boson or a photon in im combination with gravitinos $\widetilde{G}$, that are undectable, leading to missing transverse momentum. Therefore, the requirement of a missing transverse momentum $p_{\text{T}}^{\text{miss}}$ largen than most of the SM background processes are capable to produce, enables a good separation between SM background and SUSY signal. The $p_{\text{T}}^{\text{miss}}$ threshold is not supposed to be to high, to maintain sensitivity to low neutralino masses as well. Additional high separation power has the stransverse mass M_{T2}. Because it yields a good estimate of the NLSP mass, and there is no known SM particle which can decay to photons or a Z boson together with neutrinos creating an momentum imbalance in the detector, M_{T2} on average is much larger for SUSY processes than for SM ones. Both the M_{T2} and $p_{\text{T}}^{\text{miss}}$ distribution of events fulfilling the preselection for the total background and signal points of each model are shown in Figure 1.1. Hence, the SR is defined trough the following criteria:

- $p_{\text{T}}^{\text{miss}} > 150\,\text{GeV}$,
- $M_{\text{T2}} > 100\,\text{GeV}$.

The final SR is binned in two search bin in the $p_{\text{T}}^{\text{miss}}$ distribution, one ranging from 150 to 200 GeV, and the second from 200 to infinity, see Section **??**

TODO October 24, 2018: More information to each background, explain here everything.

1.1.3. Control regions

Different control regions (CRs) are defined for the four main background contributions in order to study and develop a background prediction for the associated background. These CRs are built such, that they are fully orthogonal to the signal region and among each other. The main

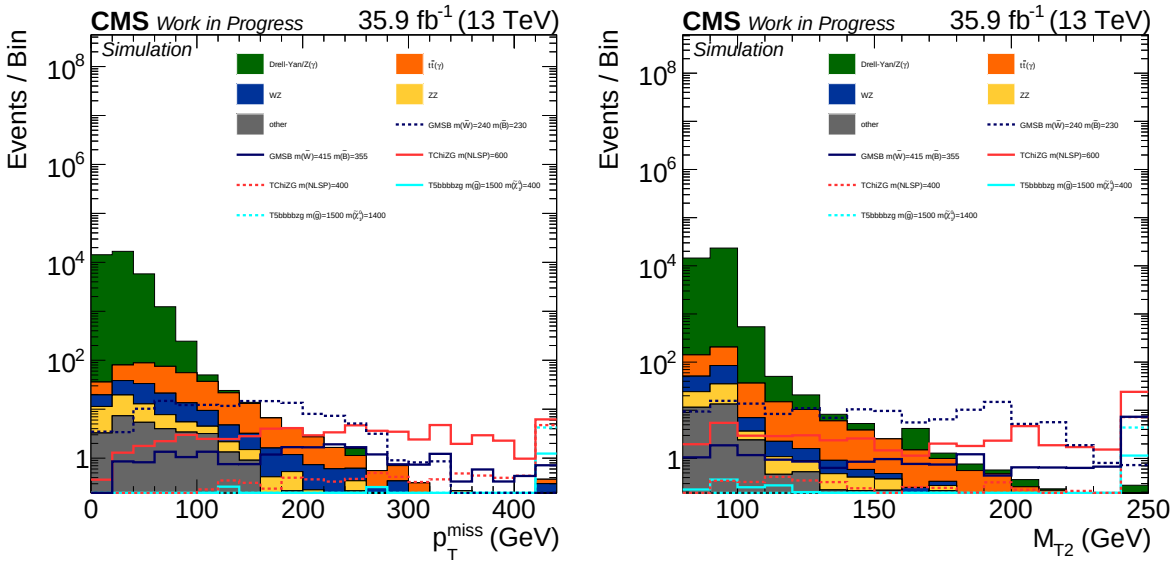

Figure 1.1.: ToDo

groups of backgrounds are $t\bar{t}(\gamma)$, DY/$Z(\gamma)$, WZ, and ZZ production.

$t\bar{t}(\gamma)$ - control region

To obtain a CR for the $t\bar{t}(\gamma)$ background with high statistics, the flavor symmetry of the procees is exploited. The two top quarks can decay independently with the same probability to an electron or a muon, resulting in an equal number of same flavor and different flavor final states in case of $t\bar{t}(\gamma)$ production. Here, the different flavor triggers are needed to guarantee a basis data set with a sufficient amount of events. In addition, the background can be studied in the same high $p_{\mathrm{T}}^{\mathrm{miss}}$ and M_{T2} regions where the SR is defined due to the underlying symmetry. The preselection criteria need to be adjusted, resulting in a CR definition of

- exactly one different flavor - opposite charge (DFOC) lepton pair (eμ),
- at least one photon,
- $\Delta R(\ell_1, \gamma) > 0.3$ and $\Delta R(\ell_2, \gamma) > 0.3$,

Of course, the invariant dilepton mass window is removed, since the reconstruced dilepton mass is not in coincidence with the Z boson mass, but has continous distribution different from the top quark mass, because the neutrinos originating from the leptonic top decays are not reconstructed individually and are therefore not considered in the calcuulation. The requirement being responsible for the orthogonality to the SR is the different dilepton composition requirement.

Drell-Yan/$Z(\gamma)$ - control region

The Drell-Yan/$Z(\gamma)$ background has nearly the same phenomenological topology as the SUSY signal, except for lower missing transverse momentum, since it is mainly due to mismeasurements of jets. Therefore, the $p_{\mathrm{T}}^{\mathrm{miss}}$ distribution of the Drell-Yan/$Z(\gamma)$ background is defined by the $p_{\mathrm{T}}^{\mathrm{miss}}$ resolution of the detector and reconstruction. The final control region definition on top of

the preselection reads only:

- $p_{\mathrm{T}}^{\mathrm{miss}}$¡100 GeV.

So this region is orthogonal to the SR per definition of the inverted $p_{\mathrm{T}}^{\mathrm{miss}}$ requirement.

WZ - control region

In order to obtain a high purity WZ control region, the SFOC dilepton criteria is adjusted such, that the fully leptonic decay of the diboson system is studied. So, as in the preselection, a SFOC lepton pair is required, but the additional lepton veto is removed. Hence, the existence of third lepton, which can be either an electron or an muon, is demanded. This ensures exlusivity between the SR and this CR. Since the W boson decays to a lepton and the corresponding neutrino for the selected events, with an additional $p_{\mathrm{T}}^{\mathrm{miss}}$ and m_T requirement, which is calculated using the third lepton, assumed to come from the W boson, and $\vec{p}_{\mathrm{T}}^{\mathrm{miss}}$, and is therefore an estimate for the W boson mass, a further selection designated for WZ diboson production is achieved. In the determination of the lepton pair belonging to the decayed Z boson, all combinations fulfilling flavor and charge requirements are tested, and in the case of multiple valid combinations, the combination with the invariant dilepton mass closest to the Z boson mass is chosen. To ensure a selection with a suitable amount of data, because the cross section for diboson production is rather low, the existence of at least photon from the preselection is removed. The final region selection reads:

- exactly one SFOC lepton pair (ee or $\mu\mu$),
- exactly one additional third lepton (e or μ),
- $\Delta R(\ell_1, \gamma) > 0.3$ and $\Delta R(\ell_2, \gamma) > 0.3$,
- $81 < m_{\ell\ell} < 101\,\mathrm{GeV}$,
- $p_{\mathrm{T}}^{\mathrm{miss}} > 70\,\mathrm{GeV}$,
- $m_T(\vec{p}_{\mathrm{T}}^{\mathrm{miss}}, \mathrm{p}_{\ell_3}) > 50\,\mathrm{GeV}$.

ZZ - control region

The strategy to achieve a pure ZZ diboson selection that is not overlapping with the SR selection is very similar to the WZ CR selection described above. If events are selected, where both Z bosons decay leptonically to charged leptons (ee or $\mu\mu$), per definition an exclusive control region is obtained. As in the WZ CR selection, flavor and charge requirements are considered to construct Z boson candidates from the four selected leptons. In cases, where multiple possibilities to reconstruct two Z bosons exist, the possibility with the first Z boson candidate mass closest to the nominal Z boson mass is chosen. The second Z boson candidate has to fulfill a looser mass agreement. Also, the existance of a photon in the event is not required. In total the selection criteria read:

- exactly two SFOC lepton pairs (ee or $\mu\mu$),
- $\Delta R(\ell_1, \gamma) > 0.3$ and $\Delta R(\ell_2, \gamma) > 0.3$ for the first pair,
- $76 < m_{\ell\ell} < 106\,\mathrm{GeV}$ for the mass of the first Z boson candidate,

- $50 < m_{\ell\ell} < 130\,\text{GeV}$ for the mass of the second Z boson candidate.

1.1.3.1. Validation region

An additional validation region (VR) is determined, where the developed background predictions are verified by perfoming data - prediction comparisons. Therefore, it is convinient to choose the VR in phasespace close to the SR. This is achievied by applying the same selection requirements as for the SR, but either the $p_{\text{T}}^{\text{miss}}$ requirement, or the M_{T2} requirement needs to fail. Hence, the VR is a sideband to the SR. The VR may also not overlap with the DY/$Z(\gamma)$ CR, thus a minimal $p_{\text{T}}^{\text{miss}}$ requirement needs to be imposed. The selection requirements for the VR in addition to the preselection are the following:

- $p_{\text{T}}^{\text{miss}} > 100\,\text{GeV}$,
- either $p_{\text{T}}^{\text{miss}} < 150\,\text{GeV}$ or $M_{\text{T2}} > 100\,\text{GeV}$.

A visualization of the signal, validation, and DY/$Z(\gamma)$ control region definitions in the M_{T2}-$p_{\text{T}}^{\text{miss}}$ plane can be found in Figure 1.2. Two dimensional histograms for the number of expected events motivating the region definitions are shown in Figure 1.3 for the all background processes, and each one benchmark point for all three considered signal models. These distributions show also

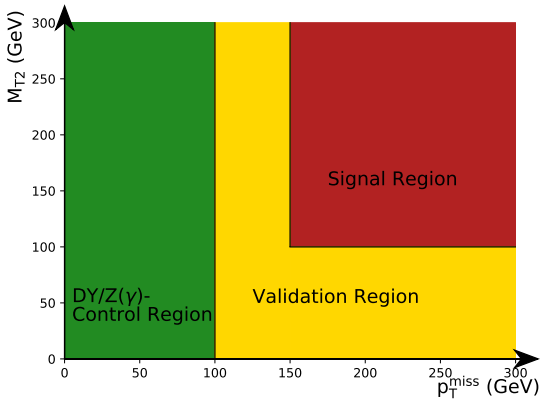

Figure 1.2.: Visualization of the definition of the signal, validation and DY/$Z(\gamma)$ control region in the M_{T2}-$p_{\text{T}}^{\text{miss}}$ plane.

specific feautures of the considered signal models regarding the important kinematic observables. Because the `TChiZG` model is a one parameter scan, and the only free paramater is the NLSP mass, which directly determines the maximum amount of transverse momentum for the decay products, ans thus directly the boson p_{T} and the $p_{\text{T}}^{\text{miss}}$ due to the gravitinos. Thus, $p_{\text{T}}^{\text{miss}}$ and M_{T2} calculated from the bosons and $p_{\text{T}}^{\text{miss}}$ are strongly correlated per definition. Also, the endpoints of the $p_{\text{T}}^{\text{miss}}$ and M_{T2} distributions are of the same order, and coincide with the NLSP mass. In case of the strong SMS, this is not directly the case. Since this is a two dimensional pramater scan, the event kinematic depends on the mass difference between the gluino and the NLSP mass, as discussed in Section **??**. In cases where the mass difference is small, the kinemtics evolve similar as for the EWK SMS, while for larger mass differences, as shown in Figure 1.3,

the correlation between M_{T2} and $p_{\mathrm{T}}^{\mathrm{miss}}$ is much weaker and the distributions are much broader. The distributions of the GMSB model behave like a mixture of the two SMSs, depending on the wino and bino masses.
And most importantly, the signal events mainly populate the SR, with small contributions in the CRs and the VR, while the majority of the background events contribute to the DY/$Z(\gamma)$ CR and a minor part to the VR. Only a minority contributes to the SR, as it is desired. So in total a good separation between background and signal is achieved.

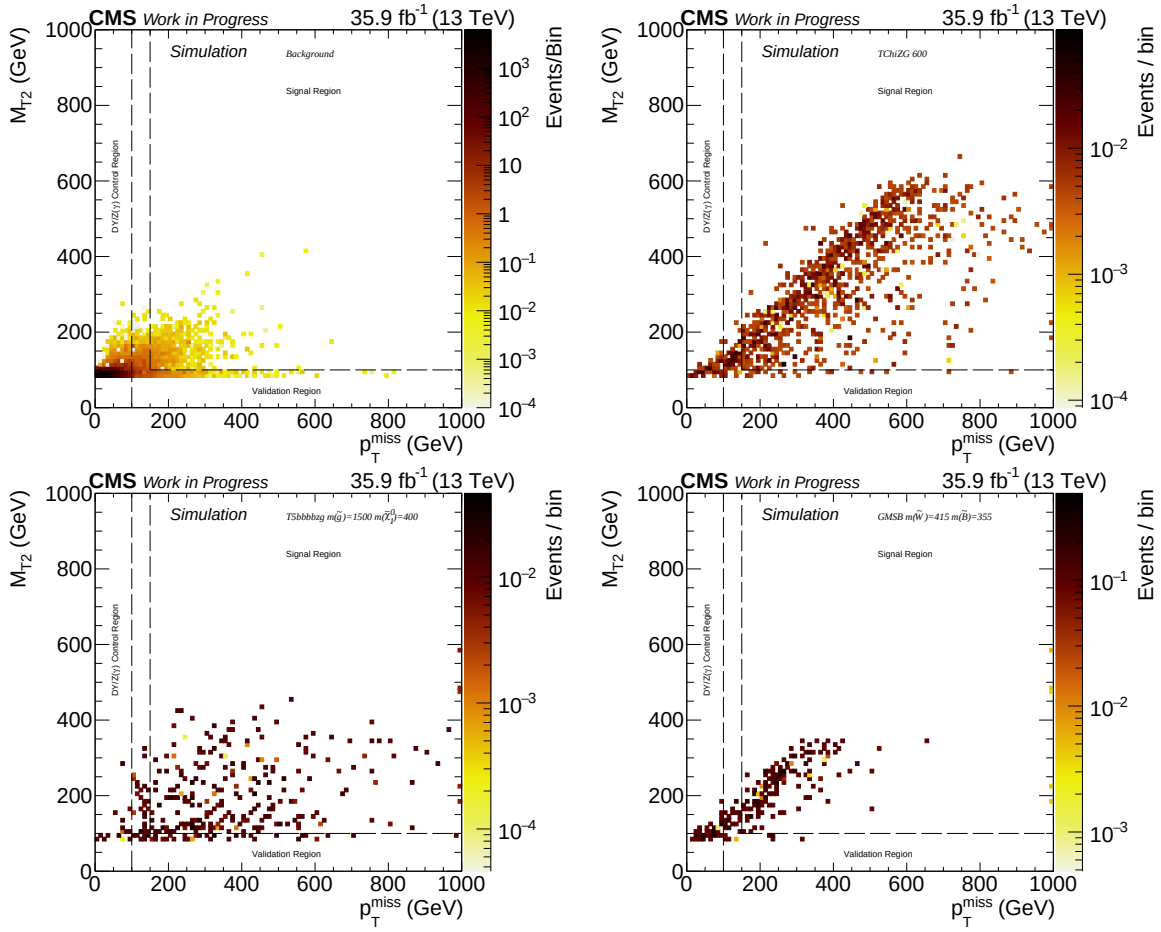

Figure 1.3.: Distribution of the total number of expected background events in the M_{T2}-$p_{\mathrm{T}}^{\mathrm{miss}}$ plane (upper left). Distributions for the number of expected signal events for the `TChiZG` SMS with an NLSP mass of 600 GeV (upper right), the `T5bbbbZG` SMS with a gluino mass of 1.5 TeV and a neutralino mass of 400 GeV (bottom left), and the GMSB model with $m(\widetilde{W} = 455\,\mathrm{GeV})$ and $m(\widetilde{B} = 355\,\mathrm{GeV})$ (bottom right).

1.2. Background Estimation

In this section the used background estimation methods are introduced, including the study of systematic effects. The total background is composited of $t\bar{t}(\gamma)$ production, Drell-Yan and $Z(\gamma)$

events, WZ and ZZ diboson production, and a remaining group of backgrounds, composited of e.g. singletop and triboson production (See Section **??** and Section 1.1.3). The relative amount of backgrounds contributing to the two SR bins are depicted in Figure 1.4. The most dominating background is $t\bar{t}(\gamma)$, followed by WZ, ZZ, and Drell-Yan/$Z(\gamma)$.
The leptonic decays of the top pairs generated in association with a photon significantly contribute to the SR, because the final state signatures appear very similiar to those of the SUSY models and the cross section is very high. If both W bosons decay leptonically, a SFOC lepton pair in connection with neutrinos are produced, that generate a large amount of $p_{\mathrm{T}}^{\mathrm{miss}}$ in the detector. Also real photons can be produced in the hard interaction, or can be radiated off in the initial or the final state by the charged participating particles. Since the total mass of the top pair cannot be reconstruced due to the momentum imbalance, there is a suffienct probability to measure a invariant dilepton mass close to the Z boson mass.
In the Drell-Yan process, in context with FSR or ISR, and in $Z\gamma$ production, directly a on-shell Z boson is produced, thus the $\mathrm{m}_{\ell\ell}$ requiremet for the SR is fulfilled. Although a photon is also produced, ths background does not contribute significantly to the SR, because only nongenuine $p_{\mathrm{T}}^{\mathrm{miss}}$ is generated in the process due to mismeasurements and therefore the $p_{\mathrm{T}}^{\mathrm{miss}}$ distribution drops steeply.
WZ production is also an important background, since the W boson can decay leptonically and therefore can generate a sufficient amount of genuie $p_{\mathrm{T}}^{\mathrm{miss}}$ to pass the SR requirements. While a photon can be generated in all possible radiation processes, an additional electron or jet originating from the W boson can fake a photon signature in the detector. In cases of real photons, the additional lepton can also be lost in the reconstruction.
Lastly, the ZZ event signature can be very similar to the one expected from various SUSY signals. One boson can decay to a pair of charged leptons, while the other decays to neutrinos leading to a large amount of $p_{\mathrm{T}}^{\mathrm{miss}}$ in the event. Together with FSR or ISR of a photon, it mimics the signal topology in many ways.
If the different background processes are compared, the diboson processes may have a more similar event kinematic than the $t\bar{t}(\gamma)$ process, but due to the lower diboson cross sections, and the small branching fraction for a Z or W boson to decay to charged leptons, the $t\bar{t}(\gamma)$ background dominates the other contributions.

The strategy to perform a sophisticated background prediction is based on the MC simulation for all backgrounds. As mentioned in Section **??**, different CRs are developed, which are enriched with corresponding background events. In those CRs, scale factors (SF) α_i are calculated by scaling the total simulation to the measured data yield. Contributions of other backgrounds in the distinct CR are considered in the calculation by fixing them, thus only one background is scaled. The SFs are calculated via the following formula:

$$1 = \frac{\#Data}{FixedBackground + \alpha_i \cdot \#BackgroundToScale} \tag{1.1}$$

$$\Longrightarrow \alpha_i = \frac{\#Data - \#FixedBackground}{\#BackgroundToScale}. \tag{1.2}$$

Due to the used method, which is performed similar in every CR, possible uncertainties in the normalization cancel, only uncertainties in the shape need to be considered, see Section 1.3. The

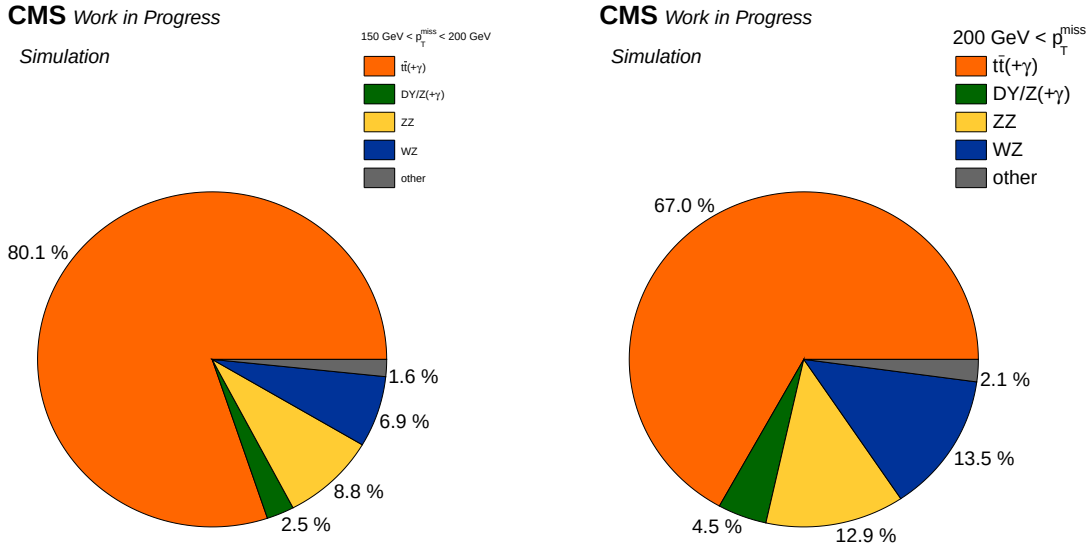

Figure 1.4.: Relative amounts for all group of considered backgrounds for the two search bins in the SR.

uncertainty arising from the SF calculation is purely of statistical origin, and is calculated via error propagation

$$\sigma_{\alpha_i}^2 = \sigma_D^2 \cdot \left(\frac{D}{B}\right)^2 + \sigma_F^2 \cdot \left(\frac{F}{B}\right)^2 + \sigma_B^2 \cdot \left(\frac{D-F}{B^2}\right)^2, \tag{1.3}$$

where $B = \#BackgroundToScale$, $F = \#FixedBackground$, $D = \#Data$, and σ_i the corresponding statistical uncertainties. These are assumed to be uncorellated Poissonian uncertainties. Since the event yields are high enough, the uncertainties can be calculated via $\sigma_i = \sqrt{N_i}$, where N is the associated total integral in the selection. In the case of MC simulation, for the calculation of the SF the normalized event yield is used, whereas in the error calculation the pure MC event count is considered to reflect the underlying precision of the simulation. These method is referred to as "integral method" in the following.
After determining all SFs, the background prediction will be tested in the VR. Therefore, the simulated backgrounds in the VR will be scaled accordingly, and the result will be investigated, see Section 1.2.6.

1.2.1. Top pair production in association with a photon

As shown in Figure 1.4, the main background contribution is $t\bar{t}(\gamma)$, which is composited of standard top pair production with and without photon radiation, and top-pair production in association with a photon, while the latter dominates. Both compositions will be determined together by rescaling the simulation such, that the integral of events of the MC simulation in the CR matches the integral of events in data as explained in Section 1.2. The raw and rescaled yields of the simulation, together with the observation in the CR are quoted in Table 1.1. The overall purity of the selection is of the order of 84.62%. For the scaled event yields, only weights to account for the trigger efficiency correction and a global weight as mentioned in Section **??**

are applied. With the integral method applied, the SF given in Equation (1.4) is obtained.

Table 1.1.: My caption

process	raw simulation	simulation	data
$t\bar{t}$	12077	416.93	
$t\bar{t}\gamma$	128396	1420.19	
sum	140473	1837.12	1750
other	14631	269.21	

$$\alpha_{t\bar{t}(\gamma)} = 0.806 \pm 0.324(stat.)[\hat{=}4.02\%]. \tag{1.4}$$

After application of $\alpha_{t\bar{t}(\gamma)}$, the number of events in data and total simulation match per definition.
Since the analyzed final state in this CR is very sensitive to higher order effects, such as jet and photon radiation, and complex matrix elements to describe the hard interaction, and these are very hard to model in the simulation, a SF away from one is not unexpected. Because the generation of MC simulation in general is time and ressource consuming, simplifcations are made on e.g. radiation effects, and not all aspects of a specific procees are calaculated and generated at NLO, altough was used for the generation of the $t\bar{t}\gamma$ sample. NNLO electroweak effects can cause negative corrections for example due to destructive terms in higher order corrections. The uncertainty of the order of 4% indicates a sufficient amount of events in the corresponding CR. The resulting prediction shows a good agreement with data, as can be seen from the $p_{\mathrm{T}}^{\mathrm{miss}}$ and M_{T2} distributions in the CR after rescaling in Figure 1.5. The uncertainty of the fit method is of statistical origin, and is treated as a systematic in the following. The agreement for additional important observables, such as the transverse momentum of the photon and both leptons, is also investigatd, and the plots can be found in the appendix in Figure **??**TODO October 24, 2018: ref. To strenghen the confidence in the outcome, Kolmogorov-Smirnov [1] tests are performed to study the shape agreement. No significant disagreement was found, and the KS-values from the tests results are given written on the plot for each individual distribution. KS-values much larger than < 0.1 indicate agreement between the input distributions.

Furthermore, in order to study a possible bias of observables not well modeled within the simulation, additional cross checks are performed. Similar to the SF calculation in the integral method, SFs are determined via ^{2} minimization in the same CRs. Albeit this method is somewhat correlated to the method explained above, it enables the possibility to study influences of different distributions in different binnings. Therefore, $\alpha_{t\bar{t}(\gamma)}$ was varied in a range in small steps of 0.01 for each binning and distribution, and for every iteration the simulation was scaled and the χ^2 was calculated, which is defined as:

$$\chi^2 = \sum_{i}^{N_{bins}} \frac{\left(N_{i,bin}^{obs} - N_{i,bin}^{predicted}\right)^2}{N_{i,bin}^{predicted}}. \tag{1.5}$$

The best fit value is determined to be at the minimum of the χ^2 distribution, which is obtained by fitting a polynomial of fourth grade to the measured points. The uncertainty of the χ^2-fit

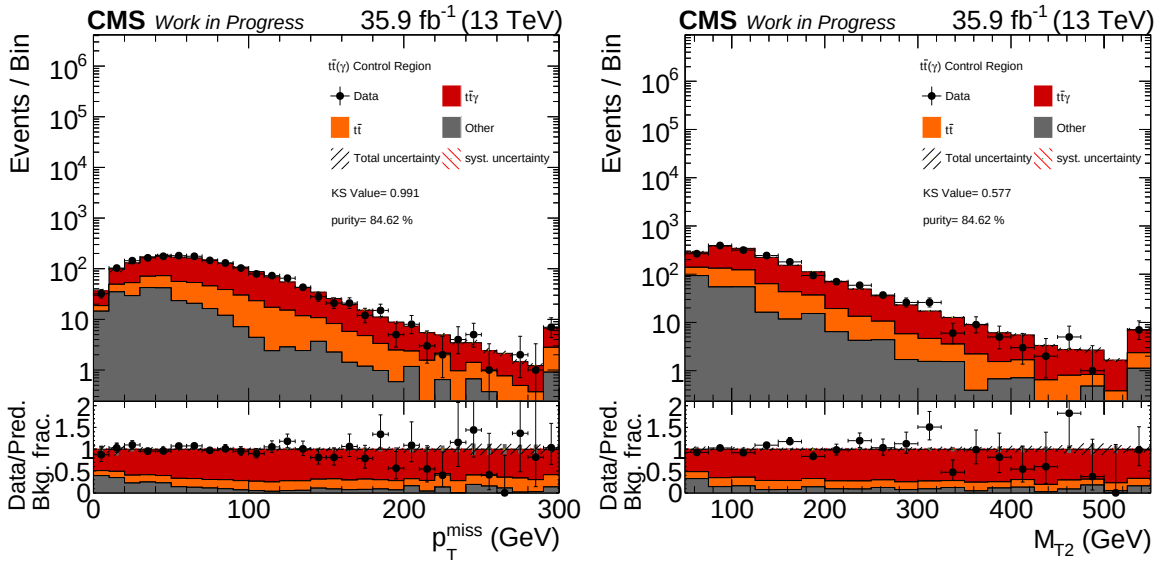

Figure 1.5.: ToDo

method, that is also of statistical origin, is calculated by taking the difference between the best fit value and the values, where $\chi^2 = \chi^2_{BestFit} + 1$. An example fit in the $p_{\mathrm{T}}^{\mathrm{miss}}$ distribution is shown in Figure 1.6 left. The number of degrees of freedom (ndf) is given by the number of bins used to obtain the χ^2 distribution reduced by the number of parameters that are being optimized. The only free parameter is $\alpha_{t\bar{t}(\gamma)}$, so *ndf* equals the bin number subracted by one. The resulting curves are smooth, indicating a proper performance of the fit. A comparison of all obtained SFs for the different setups and the SF obtained by the integral method is shown in Figure 1.6 right. No significant deviation can be observed, and all SFs are in agreement within each other. Hence, the scaling of the $t\bar{t}(\gamma)$ background seems stable and well described in the CR.

Further studies

Since the $t\bar{t}(\gamma)$ background is the most dominating one, further studies are realized, in which the underlying MC simulation samples and their combination are examined. Hence, different overlap removal procedures, see Section **??**, and different available cross sections are used, while the fit procedure itself is varied. In addition, available corrections to improve the modeling of jet radiation in the initial state, and correctios to enhance the the description of the top quark p_{T} distribution, that both are used to correct the LO samples and the samples that are generated with POWHEG (see Section **??**), are applied in different combinations. The used fit methods include a χ^2-template fit, where the ratio between the fractions of the $t\bar{t}$ and $t\bar{t}\gamma$ simulation is left free as an additional paramater, and a fit where only the simulation is made used. Eventually, no substantial differences are observed, and the best performance and stability is provided by the inital approach explained above in detail.

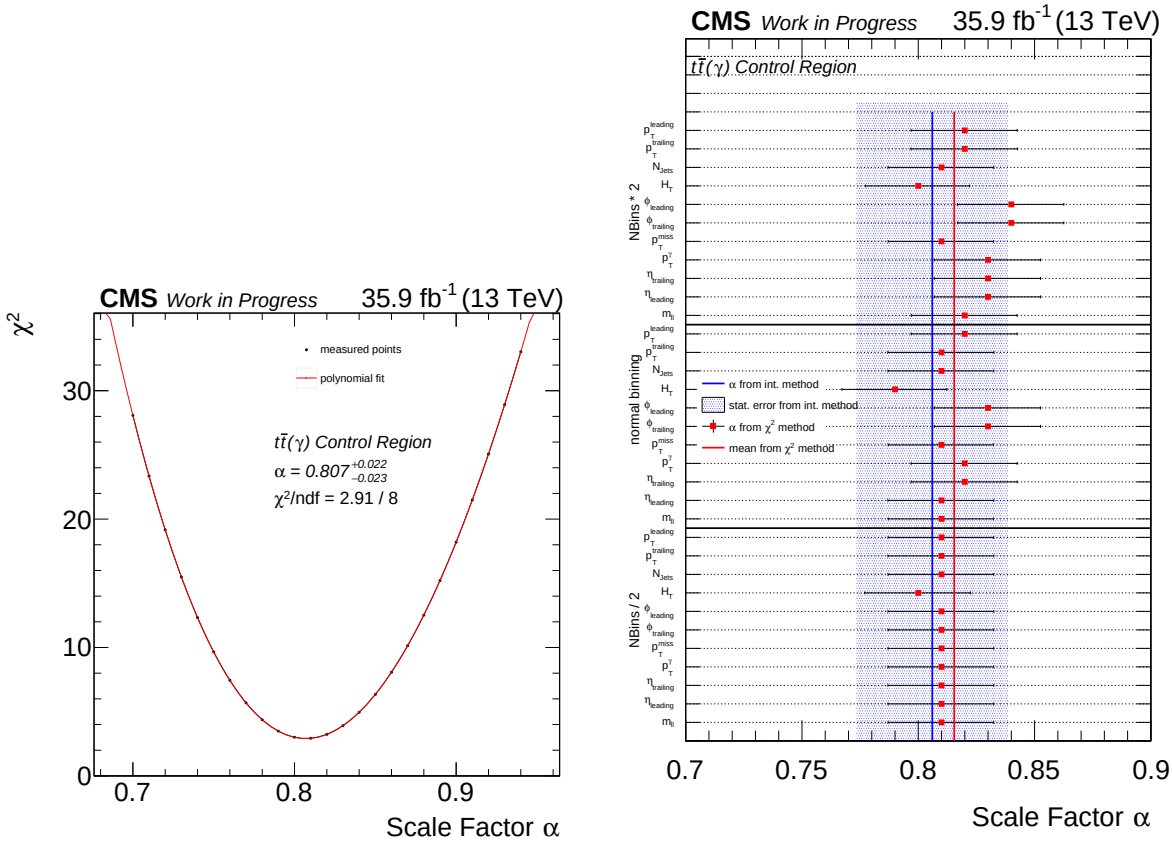

Figure 1.6.: ToDo

1.2.2. Drell-Yan and $Z\gamma$ diboson production

The Drell-Yan and dibosonic $Z\gamma$ background is admittedly small in the two SR bins, but its contribution is the fourth largest, and its needs to be assured, that the background is modeled well. As in the case of the $t\bar{t}(\gamma)$ background, its composited of two major parts, the Drell-Yan process, where quarks annihalte to offshell photons or Z bosons and generate leptons in their decays, and the diboson production of $Z\gamma$ pairs. The same integral method explained above, is used to determine the SF $\alpha_{DY/Z(\gamma)}$ in the dedicated CR, as introduced in Section 1.1.3. The relevant event yields are quoted in Table 1.2, and the resulting SF is stated in Equation (1.6). With a purity of about 99%, and a total statistics that large, a precise SF determination is feasible, as can be concluded also from the small statistical uncertainty of $\approx 1\%$. The agreement between simulation and data is very good, also before application of $\alpha_{DY/Z(\gamma)}$, since it it nearly one. The post-scaling distributions of $p_{\mathrm{T}}^{\mathrm{miss}}$ and M_{T2} are shown in Figure 1.7. They show overall a good agreement, only in the high M_{T2} region there are some fluctuations due to the limited statistics being present both in data and simulation. Further distributions are looked at, and the can be found in the appendix in Figure **??**. The same cross checks as for the $t\bar{t}()$ background are made, including the Kolmogorov-Smirnov tests, and the additional χ^2-fits with variable bin size in different observables. All in all the KS-values indicate a very good matching

between predicted and observed shape, albeit the Kolmogorov-Smirnov test provides a very small KS value for the consitency between the $p_{\mathrm{T}}^{\mathrm{miss}}$ distibutions. But, this is mainly due to the high statistics in data and therefore a higher absolute discrepancy, contradicting with the lower statistics in simulation, leading to larger fluctuations. These discrepancy is not visible in the ratio shown under the plot.

Table 1.2.: My caption

process	raw simulation	simulation	data
Drell-Yan	11710	13008.83	
$Z\gamma$	170161	22692.88	
sum	181871	35701.41	38419
other	87337	374.00	

$$\alpha_{DY/Z(\gamma)} = 1.066 \pm 0.001(stat.)[\hat{=}0.87\%]. \tag{1.6}$$

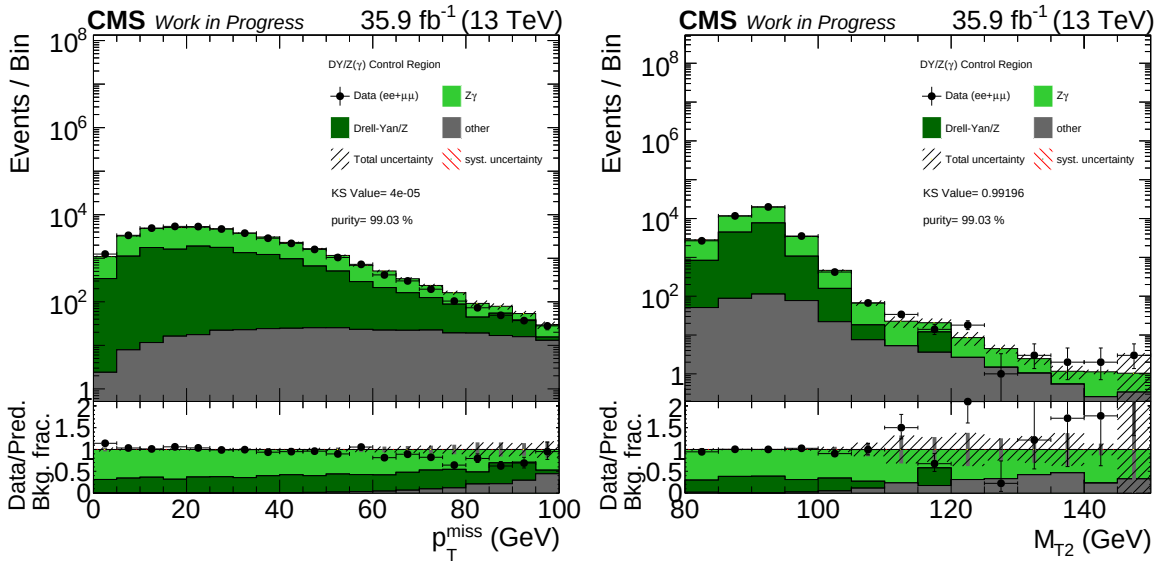

Figure 1.7.: ToDo

The χ^2-fit studies, shown in Figure 1.8 together with an example fit in the $p_{\mathrm{T}}^{\mathrm{miss}}$ distribution, show excellent agreement over all variations. Hence, the background prediction seems to work fine also in the DY/$Z(\gamma)$ CR.

1.2.3. WZ diboson production

The diboosn production of WZ pairs is the second important background contribution for this analysis. It is also tuned to agree with the measurement in the WZ CR, while the SF α_{WZ} is calculated with the integeal method. The expected purity of selected WZ events in this CR is about 84%, and the number of events given in Table 1.3 point out, that the number of events

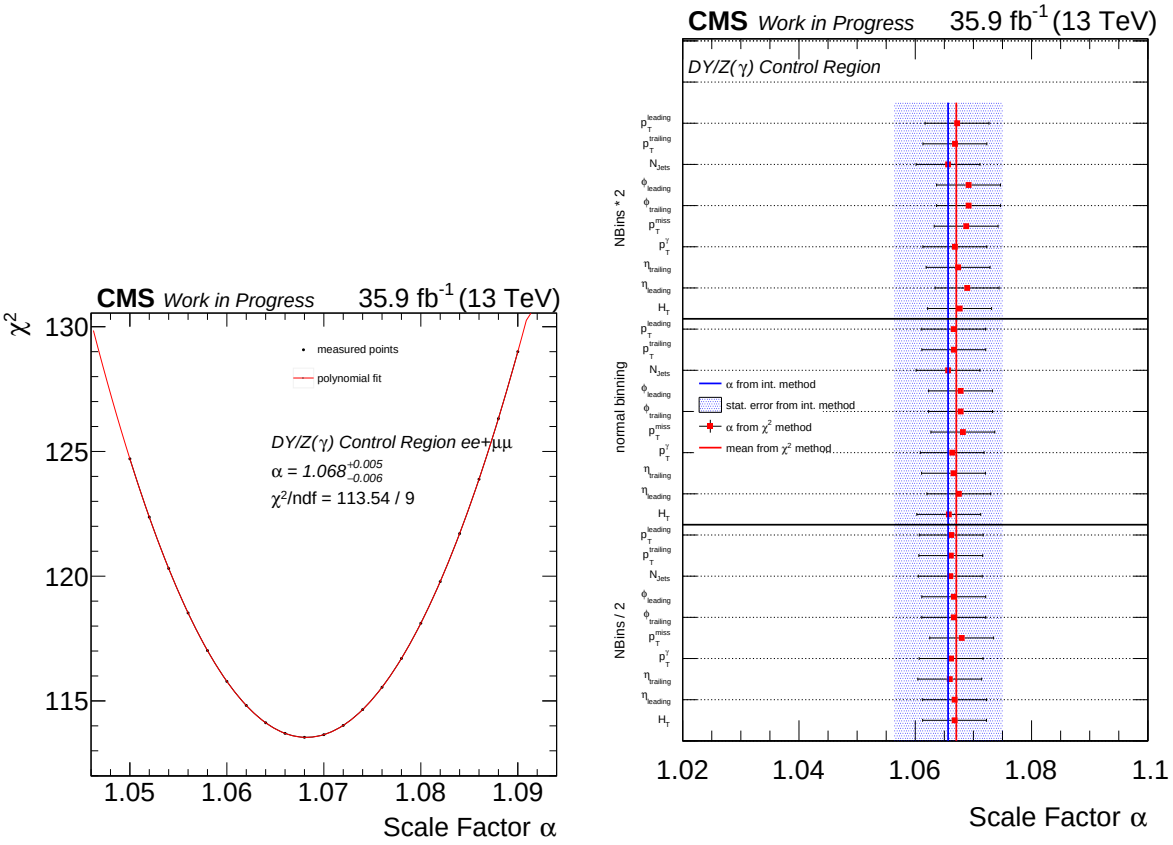

Figure 1.8.: ToDo

Table 1.3.: My caption

process	raw simulation	simulation	data
WZ	113744	895.74	
sum	113744	895.74	1193
other	93914	186.94	

is suffienct to determine a SF with adequate precision. The agreement beforehand is also very good, altough this background is a higher order process and is therefore simulated at NLO, see Section **??**, which includes several complicated effects the event generator needs to handle properly. the SF $\alpha_{DY/Z(\gamma)}$ is determined to be

$$\alpha_{\mathrm{W}Z} = 1.123 \pm 0.037(stat.)[\hat{=}3.29\%], \tag{1.7}$$

which indicates that usage of NNLO cross sections for the WZ samples is nearly suffienct enought to describe the background. Of course higher order effects can lead to higher cross sections, thus it is not unexpected to obtain a SF varying $\approx 12\%$ from unity. Resulting distributions of the dicription of the $p_{\mathrm{T}}^{\mathrm{miss}}$ and M_{T2} variables are shown in Figure 1.9. As can be seen, the predicted shape is in consitency over the whole range with the oberserved data, also in the very high $p_{\mathrm{T}}^{\mathrm{miss}}$

and M_{T2} regime. This is further confirmed by the additonal cross checks, that are also realized in the other two CRs described above. The KS-values close to one strenghen the trust in the background estimation method. The comaprison plots for other distributions can be found in the appendix in Figure **??**. The χ^2-fit studies lead to the same outcome, the uncertainties seem

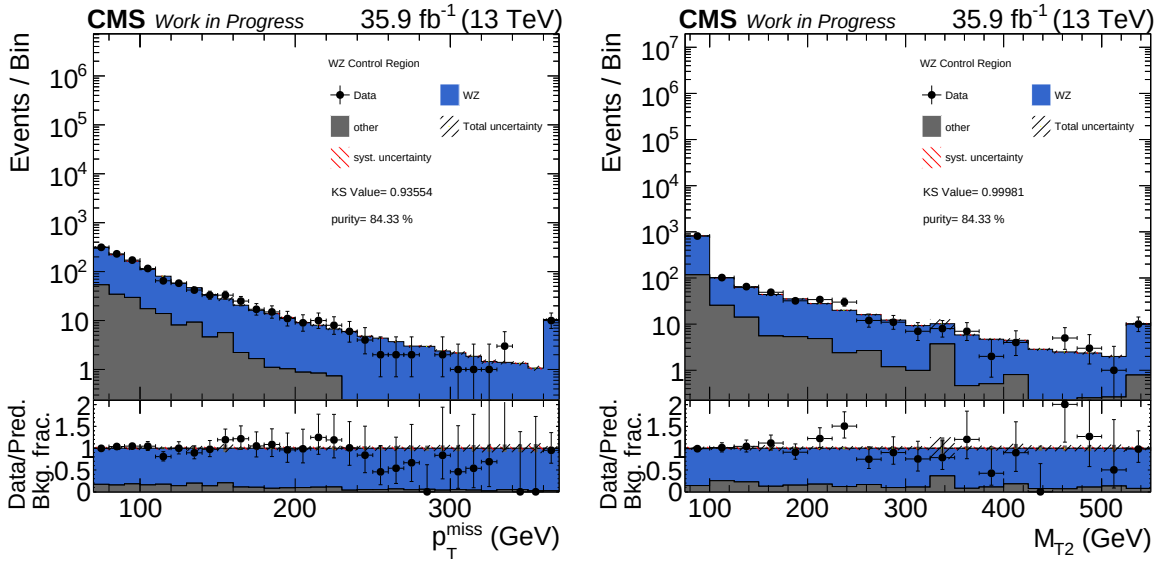

Figure 1.9.: ToDo

to be well estimated and of statistical origin, since no large shape deviation can be observed in the performed fits, as can be seen in Figure 1.10 right. Each individual χ^2-fit seems to behave properly, see Figure 1.10 left for an example in the $p_{\mathrm{T}}^{\mathrm{miss}}$ distribution.

1.2.4. ZZ diboson production

The last important and fourth biggest contribution is ZZ diboson production, where both Z bosons decay leptonically, one to charged leptons such as electrons or muons, and the other one to neutrinos. To simplify the contstruction of a dedicated CR, as explained in Section 1.1.3, the assumption is made, that there should be no difference fo the generator to generate neutral leptons or charged ones. Only the detector response should differ. Therefore, ZZ can be well estimated in a four lepton CR. So a pure selection of ZZ can be established, although the statistics in data is quite low due to small cross section and the low branching fraction for the $Z \to \ell\ell$ decay. Event counts are quoted in Table **??**, with a selection purity of nearly 100%. With the integral method applied, a SF α_{ZZ} stated in Equation (1.8) is determined, being relatively close to one.

$$\alpha_{ZZ} = 1.109 \pm 0.064(stat.)[\hat{=}5.74\%]. \tag{1.8}$$

Post-fit distibutions to show the good agreement in the $p_{\mathrm{T}}^{\mathrm{miss}}$ and M_{T2} distribution can be found in Figure 1.11. As mentioned, the precision is limited by the low statistis of the selected sample, but nevertheless is suffienct enough to establish a working background prediction. The shapes agree very well, as it is additionally indicated by the KS-values printed on the plots, also in the high M_{T2} region. The properties of the $ZZ \to 4\ell$ procees does not allow to study the

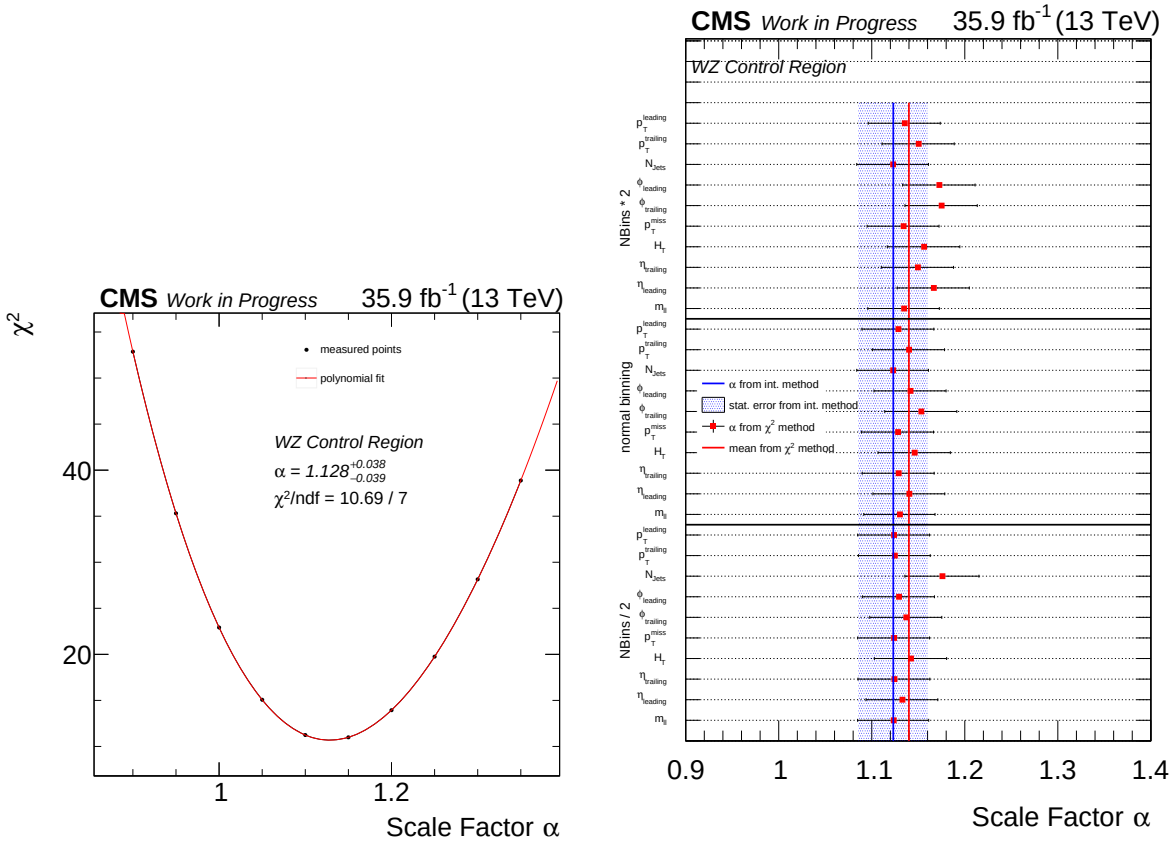

Figure 1.10.: ToDo

Table 1.4.: My caption

process	raw simulation	simulation	data
$ZZ(\to 2\ell 2\nu)$	5	0.0005	
$ZZ(\to 4\ell)$	459221	226.94	
sum	459226	226.94	251
other	79	0.07	

prediction in the high $p_{\mathrm{T}}^{\mathrm{miss}}$ regime, since only nongenuine $p_{\mathrm{T}}^{\mathrm{miss}}$ is produced, but agreement of the p_{T} distributions of the leptons, that are listed in Figure **??** in the appendix, can be translated into the agreement of the $p_{\mathrm{T}}^{\mathrm{miss}}$ in the $ZZ \to 2\ell 2\nu$ background, that is generated with the same event generator. The same χ^2-fit studies, see Figure 1.12, are performed as for the other three background estimations shown above, and yield the same conclusion, although showing more fluctuations in the choice of the variable or binning due to the lower total event counts in the observed data. But, all in all also this background estimation method behaves properly.

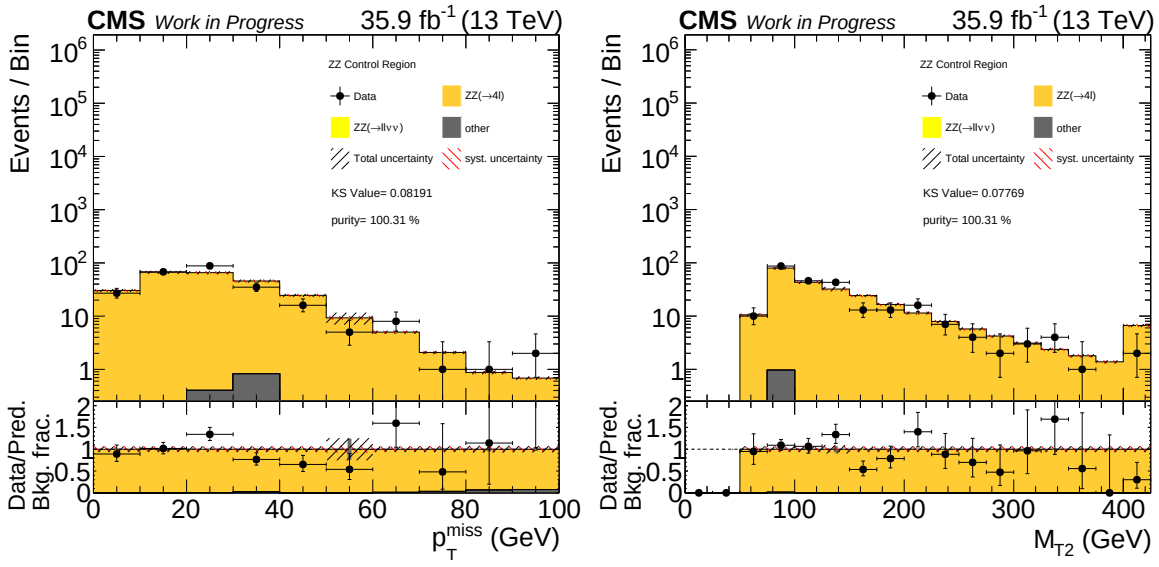

Figure 1.11.: ToDo

1.2.5. Other standard model backgrounds

Additional minor backgrounds, such as triboson processes as W$Z\gamma$, WWγ, diboson WW and Wγ, W + *jets* production, and single top processes as listed in Table **??**, are taken from plain simulation after reweighting them as described in Section **??** to the meaureed luminosty and NLO and NNLO cross sections.

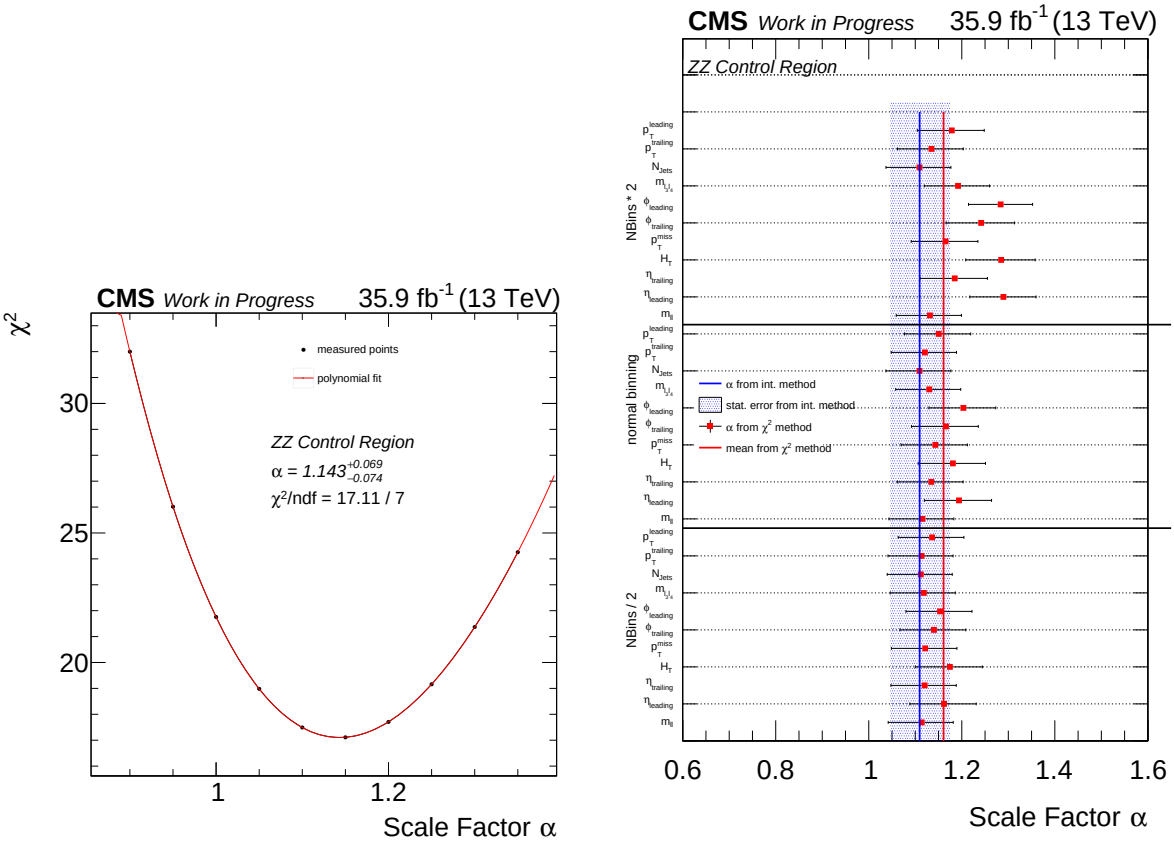

Figure 1.12.: ToDo

1.2.6. Validation of the background estimation

After all main backgrounds are tuned in various corresponding CRs and SFs are determined, and the rest is taken from simulation, the background prediction is validated in the VR defined in Section **??**. It is very close to the SR where the background prediction shall work suffienctly well, and therefore is ideal to study the obtained background prediction. The VR orthogonal VR selection is achievied by requiring nealry the same requirements as for the SR, but only one or both of the $p_{\mathrm{T}}^{\mathrm{miss}} > 150\,\mathrm{GeV}$ and $M_{\mathrm{T2}} > 100\,\mathrm{GeV}$ needs to fail.
Resulting comparisons between prediction and oberved data in different distributions are shown in Figure **??** for the most important variables, such as $p_{\mathrm{T}}^{\mathrm{miss}}$ and M_{T2}, together with p_{T} distributions of the leading and trailing lepton. Comparisons for the photon p_{T} and the invariant dilepton mass $m_{\ell\ell}$ are shown in Figure 1.14.
Overall the agreement is very good in all distributions, although being limited by the lower statsitics in some kinematic regions. The same Kolmogorov-Smirnov tests as for the different CRs are performed to gain more trust in the working of the background prediction. As can be read from the resulting KS-values, each qouted on the corresponding plot, the shape agreement considering statistical uncertainties and the systsemaic uncertainties originating from the statistical ones from the integral method, is very good. Therefore, the background prediction is

assumed to work fine in the SR due to the similar kinematics in the VR based on the design as a kinematic sideband in $p_{\mathrm{T}}^{\mathrm{miss}}$ and M_{T2}.

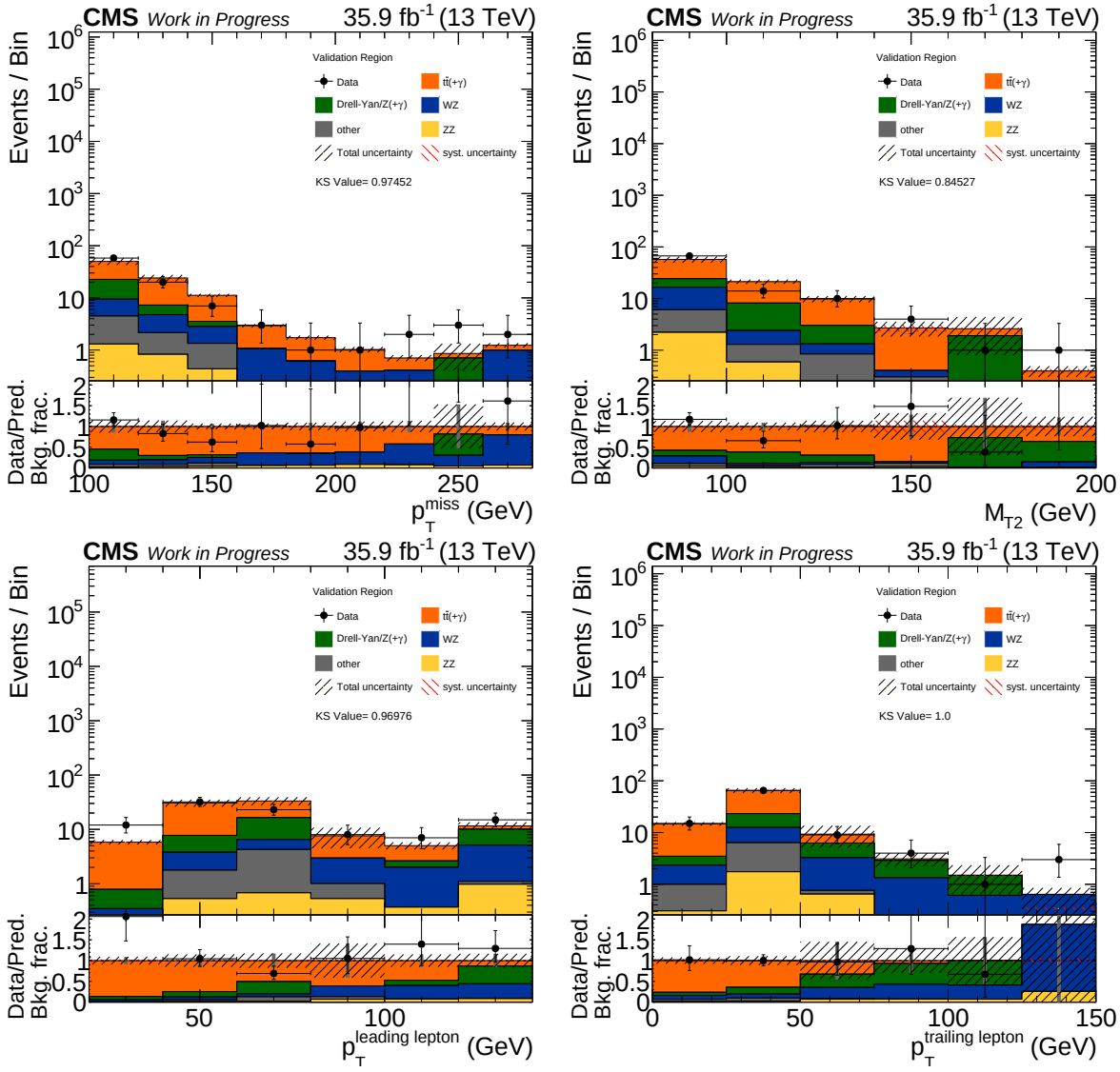

Figure 1.13.: ToDo

1.2.7. Signal contamination

If SUSY would exist, it would of course not only produce events being obervable in the SR, but also some of them may appear in the CRs, that are used to determine a proper background prediction, under the asumption that there is no influence of possible BSM signals. So to account for this effect, the so-called signal contamination of the background CRs is not considered in the background estimation itself, but is rather translated into a reduction of the predicted signal yield in the SR. Hence, at first the signal contamination needs to be measured for each signal

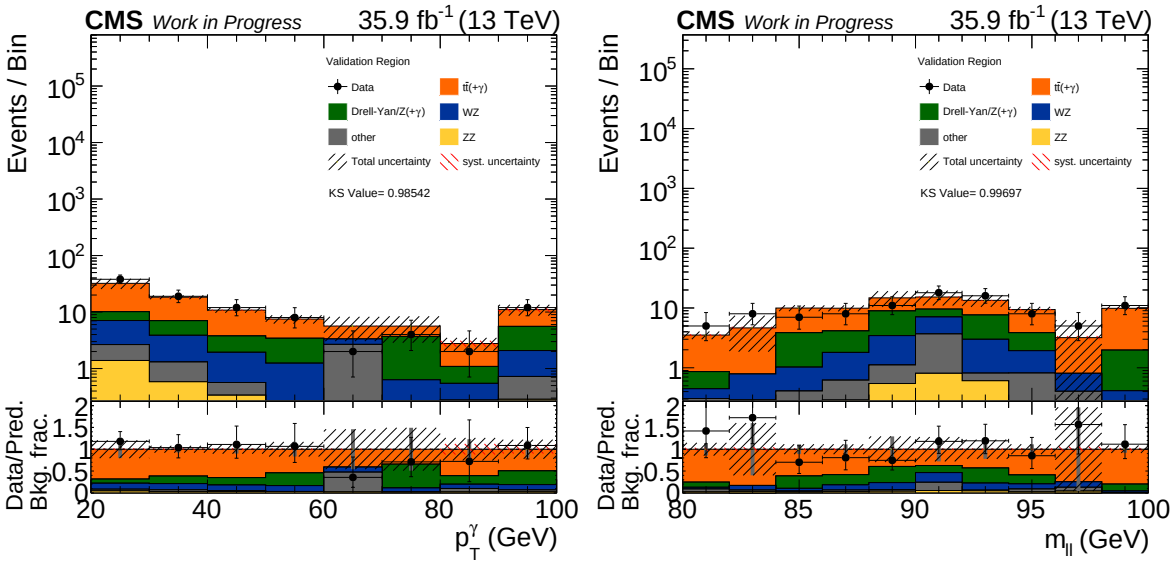

Figure 1.14.: ToDo

point of all samples individually in all CRs.

1.3. Study of systematic uncertainties

Bibliography

[1] “Kolmogorov–Smirnov Test”, pp. 283–287. Springer New York, New York, NY, 2008. doi:10.1007/978-0-387-32833-1_214.

Appendix A

Appendix

Table A.1.: All SM MC samples used in the analysis with their cross section. In the case k-factors are applied, they are given separately. {..} stands for RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV_v6 in abbreviation. All samples are of the MINIAODSIM format.

process	data set	$\sigma \cdot k$[pb]
ttbar		
$t\bar{t} \to \ell^+\nu b + \ell^-\bar{\nu}\bar{b}$	/TTTo2L2Nu_Tune*_ttHtranche3_13TeV-powheg-pythia8/{..}-v1	87.31
ttbarGamma		
$t\bar{t}\gamma$	/TTGamma_Dilept_Tune*_13TeV-amcatnlo-pythia8/{..}-v2	1.679
	/TTGamma_Hadronic_Tune*_13TeV-amcatnlo-pythia8/{..}-v2	3.482
	/TTGamma_SingleLeptFromT_Tune*_13TeV-amcatnlo-pythia8/{..}-v2	2.509
	/TTGamma_SingleLeptFromTbar_Tune*_13TeV-amcatnlo-pythia8/{..}-v2	2.509
DrellYan		
$Z/\gamma^* \to 2\ell$	/DYJetsToLL_M-50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8/{..}_ext2-v1	5765.4
diboson		
$Z\gamma \to 2\ell\gamma$	/ZGTo2LG_Tune*_13TeV-amcatnloFXFX-pythia8/{..}_ext1-v1	$117.864 \cdot 1.06$
	/ZGTo2LG_Tune*_13TeV-amcatnloFXFX-pythia8/{..}-v1	$117.864 \cdot 1.06$
	/ZGTo2LG_PtG-130_Tune*_13TeV-amcatnloFXFX-pythia8/{..}-v1	$0.1404 \cdot 1.06$
WZ	/WZTo3LNu_Tune*_13TeV-powheg-pythia8/{..}-v1	$4.42965 \cdot 1.109$
	/WZTo3LNu_Tune*_13TeV-powheg-pythia8/{..}_ext1-v3	$4.42965 \cdot 1.109$
$ZZ \to 2\ell2\nu$	/ZZTo2L2Nu_13TeV_powheg_pythia8_ext1/{..}-v1	$0.5644 \cdot k$
	/ZZTo2L2Nu_13TeV_powheg_pythia8/{..}-v1	$0.5644 \cdot k$
$ZZ \to 4\ell$	/ZZTo4L_13TeV_powheg_pythia8/{..}-v1	$1.212 \cdot k$
	/ZZTo4L_13TeV_powheg_pythia8_ext1/{..}-v1	$1.212 \cdot k$
$WW \to 2\ell2\nu$	/WWTo2L2Nu_13TeV-powheg/{..}-v1	12.178
$Wg \to \ell\nu g$	/WGToLNuG_Tune*_13TeV-amcatnloFXFX-pythia8/{..}_ext3-v1	489
W jets		
$W + jets$	/WJetsToLNu_Tune*_13TeV-amcatnloFXFX-pythia8/{..}_ext2-v2	61526.7
	/WJetsToLNu_Tune*_13TeV-amcatnloFXFX-pythia8/{..}-v1	61526.7
triboson		
WWg	/WWG_Tune*_13TeV-amcatnlo-pythia8/{..}_ext1-v1	0.2147
WZg	/WZG_Tune*_13TeV-amcatnlo-pythia8/{..}-v1	0.04123
single top		
$W^+ \to t\bar{b}$	/ST_s-channel_4f_leptonDecays_13TeV-amcatnlo-pythia8_*/{..}-v1	3.36
$q\bar{b} \to q'\bar{t}$	/ST_t-channel_antitop_4f_incl*Decays_13TeV-powhegV2-*-pythia8_*/{..}-v1	80.95
$qb \to q't$	/ST_t-channel_top_4f_incl*Decays_13TeV-powhegV2-*-pythia8_*/{..}-v1	136.02
$\bar{b} \to W^+\bar{t}$	/ST_tW_antitop_5f_NoFullyHadronicDecays_13TeV-powheg_*/{..}_ext1-v1	11.7
$b \to W^-t$	/ST_tW_top_5f_NoFullyHadronicDecays_13TeV-powheg_*/{..}_ext1-v1	11.7

Table A.2.: All SUSY MC samples used in the analysis. {..} stands for 80X_mcRun2_asymptotic_2016 in abbreviation. All samples are of the MINIAODSIM format.

signal	data set
electroweak	
TChiNG	/SMS-TChiNG_BF50N50G_*/RunIISpring16MiniAODv2-PUSpring16Fast_{..}_miniAODv2_v0-v1
GMSB model	/GMSB_GravitinoLSP_N1decays_*/RunIISummer16MiniAODv2-PUSummer16Fast_{..}_TrancheIV_v6-v1
strong production	
T5bbbbZG	/SMS-T5bbbbZg_*/RunIISummer16MiniAODv2-PUSummer16Fast_{..}_TrancheIV_v6-v2

Table A.3.: Trigger paths used in the analysis.

trigger path
dielectron trigger
HLT_Ele17_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v*
HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v*
HLT_DoubleEle33_CaloIdL_GsfTrkIdVL_v*
HLT_DoubleEle33_CaloIdL_GsfTrkIdVL_MW_v*
dimuon trigger
HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_v*
HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_v*
HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_v*
HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ_v*
HLT_TkMu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ_v*
HLT_Mu27_TkMu8_v*
HLT_Mu30_TkMu11_v*
electron-muon trigger
HLT_Mu17_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_v*
HLT_Mu23_TrkIsoVVL_Ele8_CaloIdL_TrackIdL_IsoVL_v*
HLT_Mu23_TrkIsoVVL_Ele8_CaloIdL_TrackIdL_IsoVL_DZ_v*
HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_v*
HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v*
HLT_Mu8_TrkIsoVVL_Ele17_CaloIdL_TrackIdL_IsoVL_v*
HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_v*
HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ_v*
HLT_Mu12_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_v*
HLT_Mu12_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ_v*
HLT_Mu30_Ele30_CaloIdL_GsfTrkIdVL_v*
HLT_Mu33_Ele33_CaloIdL_GsfTrkIdVL_v*
HT trigger
HLT_PFHT200_v*
HLT_PFHT250_v*
HLT_PFHT300_v*
HLT_PFHT350_v*
HLT_PFHT400_v*
HLT_PFHT475_v*
HLT_PFHT600_v*
HLT_PFHT650_v*
HLT_PFHT800_v*
MET trigger
HLT_PFMET110_PFMHT110_IDTight_v*
HLT_PFMET120_PFMHT120_IDTight_v*
HLT_PFMET170_NoiseCleaned_v *
HLT_PFMET170_HBHECleaned_v*
HLT_PFMET170_JetIdCleaned_v*
HLT_PFMET170_NotCleaned_v*
HLT_PFMET300_v*
HLT_PFMET400_v*
HLT_PFMET500_v*