0.1 Event selection

Event selection is done in order to decrease background and get a clean signal process. VBS processes have a small cross-section around 1 fb resulting in a small event count. In the 2015 and 2016 Atlas run with 36 fb^{-1} not even 100 events produce VBS processes. Therefore, one has to carefully select Events in order to get a meaningful signal process. The event selection is implemented as described in [1] using the Common Analysis Framework(CAF)[2]. The object selection however was already done in ELCore and account for the event cleaning, GoodRun-List, Trigger, Primary Vertex in Table 0.1 these can't be change in the CAF. Therefore, only the event selection done in the CAF will be discussed. ELCore produces beam reconstruction level samples. The event selection in the CAF is split into the three phase space regions lllvjj, WZjj and the VBS signal region(VBSSR). Even though the regions are different the selection criteria overlap. The WZjj region use the selection criteria of the lllvjj as base and introduces new selection criteria. The same is true for the VBSSR which builds upon the WZjj region as show in Table 0.1.

lllvjj region: Only events with 3 or more leptons that pass the Z-analysis are chosen. The leptons are then assigned to the decaying gauge boson. For this same flavour and opposite charge(SFOC) leptons pairs are chosen. The pair with invariant mass close to the Z boson mass is assigned to the Z boson. The highest transverse momentum p_T lepton from the remaining leptons is assigned to the $W^{\pm}$ boson and required to pass the W-analysis. Chosen leptons need a transverse momentum $p_T > 25(27)$ GeV for the 2015(2016) campaign for the events to pass the trigger threshold. Events have to have two or more events in order to be selected. These jets need to have $p_T(j) > 40$ GeV to be considered as tagging jets.

WZjj region: Additional cuts are applied for the WZjj region in order to maximize $W^{\pm}Z$ contribution in the lllvjj final state. Some events in ZZ diboson production are expected to produce an additional lepton therefore a four lepton baseline veto is applied. Events where the jet is considered a b-jet are discarded to minimize $t\bar{t}$ and other t-quark contributions. The transverse mass of the $W\pm$ boson is required to be greater than $M_T(W) > 30$ GeV. The Z lepton pair has to have an invariant mass within 10 GeV of the Z boson mass $m_Z = 91.1876$.

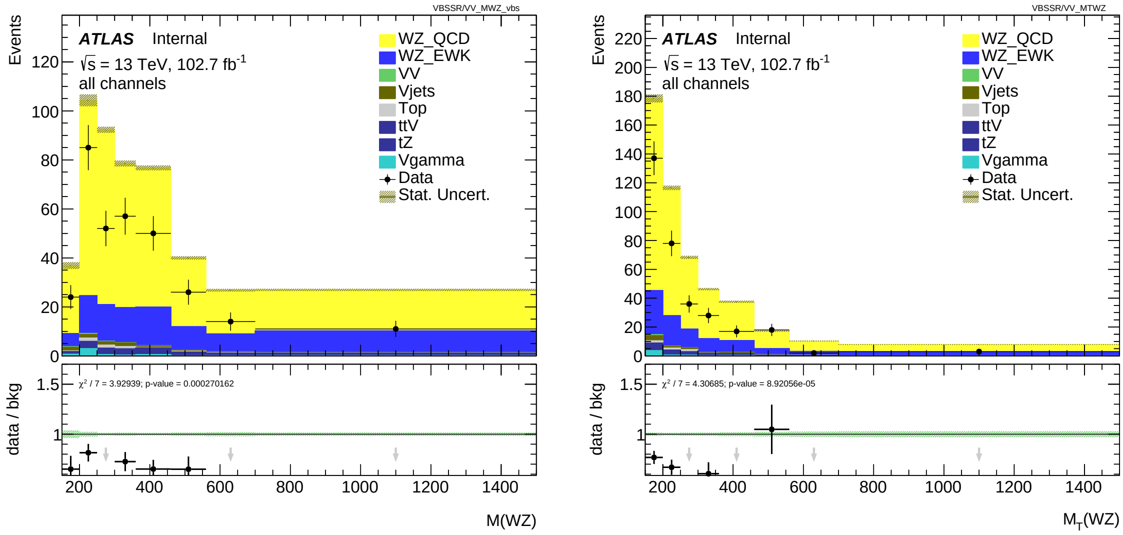
VBSSR region: The resonance fitting is done in the VBSSR region. For this additional requirements must be added increasing WZjj-EW6 contribution compared to WZjj-EW4 and WZjj-EW5. For this two cuts for the tagging jets have to be added. The invariant mass must be $M(jj) > 500$ GeV and the absolute rapidity difference hast be constraint $\Delta Y(jj) > 2$.

Figure 0.1 shows the cuts for the VBSSR region applied for the combined data from mc16d and mc16e with a combined luminosity of 102.7 fb^{-1}. These histograms will be combined to

WZ + Background in the following plots. Ideally one would use all campaigns mc16a, mc16d and mc16e with combined luminosity of 139 fb^{-1} to achieve the best statistic. This however is currently not possible due to limits in CAF.

Event selection criterion	lllvjj region	WZjj region	VBSSR
Event cleaning	$\surd$	$\surd$	$\surd$
GoodRunList	$\surd$	$\surd$	$\surd$
Trigger	$\surd$	$\surd$	$\surd$
Primary Vertex	$\surd$	$\surd$	$\surd$
lllvjj finale state			
$\geq$ 2 jets	$\surd$	$\surd$	$\surd$
$\geq$ 3 Z-Analysis leptons	$\surd$	$\surd$	$\surd$
One SFOC pair	$\surd$	$\surd$	$\surd$
l_W is in W-Analysis selection	$\surd$	$\surd$	$\surd$
Transverse momentum of leading leptons			
$p_T(l) > 25(27) GeV$	$\surd$	$\surd$	$\surd$
Transverse momentum of subleading jet			
$p_T(j_2) > 40 GeV$	$\surd$	$\surd$	$\surd$
$M_T(W) > 30 GeV$		$\surd$	$\surd$
$\lvert M(ll) - 91.1876 GeV \rvert < 10 GeV$		$\surd$	$\surd$
four Baseline leptons veto		$\surd$	$\surd$
b-jet veto		$\surd$	$\surd$
$M(jj) > 500 GeV$			$\surd$
$\Delta Y(jj) > 2$			$\surd$

Table 0.1: Schematic view of cuts applied in different phase space regions. [1]



(a) **(b)**

Figure 0.1: VBSSR region for a) invariant mass, b) transverse mass. Showing reproduction of SM as well as the in section **??** discussed dominance of QCD and EW processes. VV, Vjets, Top, ttV, Vgamma, WZ_EWK, WZ_QCD are combined to $'WZ\_SM + Background'$ in later plots. For the fits WZ_EWK is used as signal.

In order to get a valid result the SM limits have to be reproduced. The samples used for this thesis where already produced before starting the thesis and no samples where produced specifically for this thesis. Therefore, samples used for the analysis did not use the same settings as the samples used for the SM limit setting([3],[4]) resulting in differences in the number of events in tabular 0.2. The EFT limits for the SM for the ATLAS detector only use mc16a data while this analysis uses mc16d and mc16e data resulting in smaller limits since more data is used. The limit comparison is shown in tabular 0.3.

VBSSR	mc16d+mc16e	mc16a	SM
WZ QCD	356.8912 ± 2.4118	121.5475 ± 1.1721	144 ± 41
WZ EWK	91.2393 ± 0.4231	32.4363 ± 0.2522	24.9 ± 1.4
Background	27.2161 ± 2.0335	11.9691 ± 0.4701	31.1 ± 3.9
WZ SM + Background	475.3466 ± 3.2912	165.9529 ± 1.4050	200 ± 41

Table 0.2: Overview of events in VBSSR region. mc16a is compared to the [5] for the reproduction of the SM EFT Limits. While the resonance fitting is done with the mc16e+mc16d data. The differences between the mc16a and the SM is like cause by some ELCore settings since the cuts possible in the CAF tool where replicated.

Operator	EFT Optimized binning $M(WZ)$	EFT Optimized binning $M_T(WZ)$	SM Optimized Binnnig $M_T(WZ)$
S0	[-49 , 47.4]	[-46 , 44.1]	[-78 , 78]
M0	[-9.1 , 11.5]	[-9.3 , 9.65]	[-15 , 15]
M1	[-16 , 15.7]	[-15 , 14.4]	[-23 , 23]
T0	[-0.43 , 0.457]	[-0.38 , 0.325]	[-1.4 , 1.4]
T1	[-0.82 , 0.713]	[-0.72 , 0.682]	[-0.97 , 0.97]
T2	[-2.4 , 2.16]	[-2.2 , 1.96]	[-2.8 , 2.8]

Table 0.3: Asmiov fit limits for SM data with 95% CL. Differences arise from the use of mc16e, mc16d data compared to the mc16a data and the use of a BSM search optimized binning [150,200,250,300,360,460,560,700,1500] compared to the SM optimized binning [150,200,250,300,400,1500] used in [5]

0.2 Maximum Likelihood Method

In order to achieve a scientifically relevant result a statistical framework that fits the analysis has to be used. The maximum likelihood method builds the base of the statistics used. [6], [7] and [8] provide an in depth view into the topic, but a short introduction is given here to understand the fitting results. The maximum likelihood method takes observed data and an assumed probability distribution in order to give the best estimate for a parameter.

The likelihood function for a joint probability distribution(pdf) is defined as:

$$L(\overrightarrow{x}, \overrightarrow{\Theta}) = \prod_{i=1}^{n} f(x_i, \overrightarrow{\Theta}) \tag{0.1}$$

With distribution $f(x; \overrightarrow{\Theta})$, $\overrightarrow{\Theta} = (\Theta_1, \Theta_2, \ldots, \Theta_m)$ and a measurement of n independent values $\overrightarrow{x} = (x_1, x_2, \ldots, x_n)$. The maximum of the likelihood function is the best estimate for the parameter $\overrightarrow{\Theta}$. It is often the case in particle physics to maximize the log-Likelihood function lnL. The derivative can be written as follows since the logarithm is a monotonically increasing function.

$$\left.\frac{\partial L}{\partial \Theta_i}\right|_{\overrightarrow{\Theta}=\hat{\Theta}} = 0 \qquad \Longrightarrow \qquad \left.\frac{\partial lnL}{\partial \Theta_i}\right|_{\overrightarrow{\Theta}=\hat{\Theta}} = 0 \tag{0.2}$$

Where $\hat{\Theta}$ is the maximum likelihood estimate. This can be applied to a Gaussian distribution with $E[x_i] = \mu(y_i, \overrightarrow{\Theta})$ and $V[x_i] = \sigma_i^2$

$$f(x, \overrightarrow{\Theta}) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu(y_i, \overrightarrow{\Theta}))^2}{2\sigma^2}} \tag{0.3}$$

resulting in a log-Likelihood function

$$ln\ L(\overrightarrow{\Theta}) = \sum_{i=1}^{n} \left(ln\frac{1}{\sqrt{2\pi}} - ln\ \sigma - \frac{(x_i - \mu(y_i, \overrightarrow{\Theta}))^2}{2\sigma^2} \right) \tag{0.4}$$

Only μ is dependent in the parameter on $\overrightarrow{\Theta}$ and therefore only the third therm is relevant. One can now compare the result to the chi-squared method.

$$\chi^2(\overrightarrow{\Theta}) = \sum_{i=1}^{n} \frac{(x_i - \mu(y_i, \overrightarrow{\Theta}))^2}{\sigma^2} \tag{0.5}$$

The log-likelihood function can therefore be written as $-2ln\ L = \chi^2$. If x_i is not a Gaussian distribution than the chi-squared method is an "ad hoc" method. This form has the advantage that if a Gaussian distribution is assumed the parameters can be easily determined but can give asymmetric uncertainties. The standard deviation can be calculated by setting the derivative of L with respective to σ and derivative of L with respective to μ equal to zero.

$$\frac{\partial lnL}{\partial \mu} = \sum_{i=1}^{n} \frac{x_i - \mu}{\sigma^2} = 0 \tag{0.6}$$

$$\frac{\partial lnL}{\partial \sigma} = \sum_{i=1}^{n} \frac{(x_i - \mu)^2}{\sigma^3} - \frac{1}{\sigma} = 0 \tag{0.7}$$

These two equations can now be solved simultaneously leading to

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i = \overline{x} \qquad \text{and} \qquad \sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \overline{x})^2 = n. \tag{0.8}$$

With this the one sigma and two sigma environment can be calculated for a negative log-likelihood function as $1\sigma = 1$ and $2\sigma = 4$ [9]

0.3 Likelihood ratio test

The likelihood-ratio test[8] is used to estimate the goodness of fit based on their likelihoods. In general the denominator contains the likelihood for the entire parameter space, while numerator introduces constrains on the parameter space.

$$\lambda(\epsilon) = \frac{L(\epsilon, \hat{\hat{\Theta}}_0)}{L(\hat{\epsilon}, \hat{\Theta})} \tag{0.9}$$

where $\hat{\hat{\Theta}}_0$ donates the value $\overrightarrow{\Theta}$ for which L is maximized for a given ϵ. ϵ is the conditional maximum-likelihood estimator of $\overrightarrow{\Theta}$. $\hat{\epsilon}$ and $\hat{\Theta}$ are the maximum likelihood estimators for the unconditional likelihood. In case of the log-likelihood function this can be written as:

$$-2\lambda = -2ln\left(\frac{L(A)}{L(B)}\right) = -2(lnL(A) - lnL(B)) = -2\Delta L \tag{0.10}$$

In the context of fitting in EFT $L(A)$ donates a fixed parameter while $L(B)$ donates the floating parameter. The likelihood-ratio is than plotted against the parameter values **??**.