VRIJE UNIVERSITEIT BRUSSEL

Dissertation for the Degree
Master in Physics and Astronomy

SEARCH FOR SUPERSYMMETRIC PARTNERS OF THE TOP QUARK IN EVENTS WITH TWO DISPLACED MUONS USING THE 2016 DATA FROM THE CMS EXPERIMENT

MATTHIAS STUCKENS
2016-2017

CERN-THESIS-2017-197

Promotor: Prof. Dr. Freya Blekman
Co-promotor: Dr. Denys Lontkovskyi
Sciences & Bio-Engineering Sciences

Contents

1 Introduction.

With recent excitement arising from the excesses found in the data taken by the CMS and ATLAS experiments at the LHC, the search for beyond the standard model physics (BSM) is a hotter topic than ever. Even though the excesses disappeared upon further analysis, the excitement has not. Theorists let their imagination free coming up with exotic theories while the experimentalists try to prove them wrong or right.

This paper present the results of a search for the supersymmetric partner of the top quark (stop quark) using the displaced di-muon signature arising from stop quark decay. A method for the estimation of Non-QCD background events and a data-driven method for the estimation of QCD background events is given and performed.

The event preselection was performed by Denys Lontkovskyi using tools retrieved from a previous displaced di-lepton search [11].

2 Theoretical motivation.

2.1 The standard model

The standard model is probably the greatest accomplishment in the field of physics, due to its ability to describe nature and its uncanny predictive power. The standard model describes the elementary particles as well as the fundamental interactions between them. These interactions are described by the exchange of force carrying particles called gauge bosons.

Particles are separated into two groups depending on their spin. Due the the quantum nature of physics, these particles can only have spins which are discrete values proportional to h, where h is Planck's constant. Particles with integer spin are called bosons while particles with half integer spin are called fermions.

Table 1 contains the fermions of the standard model arranged according to

Generation	$q = -1$	$q = 0$	$q = \frac{2}{3}$	$q = \frac{-1}{3}$
1	e^-	ν_e	u	d
2	μ^-	ν_μ	c	s
3	τ^-	ν_τ	t	b

Table 1: The elementary particles of the standard model arranged according to their f-physical properties.

their physical properties. The particles in the same column have the same physical properties, with the exception of their mass which increases when going to a higher generation and their weak isospin. Table 1 is further divided into the first two columns, which are called leptons, and the last two columns, which are called the quarks. The quarks have an additional property called color charge, which allows them to interact via the strong force.

In addition to the fermions, the standard model comprises five bosons, which are the photon, Z-boson, Brout-Englert-Higgs-boson, the positively and negatively charge W-boson and the gluon. The bosons act like the force carriers between the fermions and themselves creating the fundamental forces of nature.

The immense achievement of the standard model is not only knowing what the fundamental building blocks of nature are, but having found a way to describe them. From the work of Noether in classical mechanics, a theory was build around the Lagrangian and the principle of least action. The resulting quantum field theory was capable of not only describing measurements in experiment, but could predict the results as well.

The basis of the theory revolves around describing particles as a field with the particles themselves being the quanta of the field. This is done by finding a Lagrangian for every interaction and using Noether's principle of finding symmetries of the theory. Consider the simplest form of the Lagrangian which is just a kinetic and a mass term (also known as the Dirac Lagrangian),

$$\mathscr{L}_{Dirac} = i\bar{\psi}\gamma^{\mu}\partial_{\mu}\psi - m\bar{\psi}\psi\,. \tag{1}$$

Where ψ is the fermion field, γ^{μ} are the Dirac matrices and m the mass of the particle. The Lagrangian is invariant under symmetry

$$\psi \to \psi' = e^{-\vec{\alpha}(x)\cdot\frac{\vec{\tau}}{2}}\psi\,, \tag{2}$$

as long as the partial derivative is replaced by

$$\partial_{\mu} \to D_{\mu} = \partial_{\mu} - ig\frac{\vec{\tau}}{2}\cdot\vec{A_{\mu}}\,. \tag{3}$$

Where g is the coupling constant, $\vec{\tau}$ are the generators of SU(2) and $\vec{A_{\mu}}$ the newly introduced field. This results in the Lagrangian

$$\mathscr{L}_{Dirac} = i\bar{\psi}\gamma^{\mu}\partial_{\mu}\psi - m\bar{\psi}\psi\ + g\bar{\psi}\gamma^{\mu}\frac{\vec{\tau}}{2}\cdot\vec{A_{\mu}}\psi\,. \tag{4}$$

This method can now be applied to the strong force brought forth by the symmetry group SU(3). The electromagnetic and weak force, though behaving different at low energies, were found to be two aspects of the same force called the electroweak force. The electroweak force is brought forth by the SU(2)XU(1). The resulting symmetries from these group give rise to

$$\mathscr{L}_{electroweak} = i\bar{\psi}\gamma^{\mu}\partial_{\mu}\psi - m\bar{\psi}\psi\ + g'\bar{\psi}\gamma^{\mu}\frac{\sigma^{\alpha}}{2}W_{\mu}^{\alpha}\psi + g'\bar{\psi}\gamma^{\mu}\frac{Y}{2}B_{\mu}\psi \tag{5}$$

and

$$\mathscr{L}_{strong} = i\bar{\psi}\gamma^{\mu}\partial_{\mu}\psi - m\bar{\psi}\psi\ + g_s\bar{\psi}\gamma^{\mu}\frac{\lambda_{\alpha}}{2}\cdot G_{\mu}^{\alpha}\psi\,, \tag{6}$$

the Lagrangians for the electroweak and strong force respectively. Here σ_{α} are the Pauli matrices, Y the weak hypercharge, λ_{α} the Gell-Mann matrices and B_{μ}, W_{μ}^{α} with $\alpha \in [1,2,3]$, G_{μ}^{α} with $\alpha \in [1,...,8]$ the newly acquired fields.

Linear combinations of these fields represent four of the gauge bosons discussed earlier. These linear combinations are

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp (iW_\mu^2) \tag{7}$$

$$Z_\mu = W_\mu^3 cos\theta_W - B_\mu sin\theta_W \tag{8}$$

$$A_\mu = W_\mu^3 sin\theta_W + B_\mu cos\theta_W \,. \tag{9}$$

Where θ_W is given by

$$tan\theta_W = \frac{g'}{g} \,. \tag{10}$$

The $W_\mu^\pm$ corresponds to the W-bosons, Z_μ with the Z-boson and A_μ with the photon. The eight gluon fields are simply given by G_μ^α.

The final addition to the standard model is the Brout-Englert-Higgs-boson. It arises from the addition of a scaler field ϕ with a corresponding addition to the standard model Lagrangian of the form

$$\mathscr{L}_{BEH} = (D\phi)^2 - \mu^2\phi^2 - \lambda\phi^4 \,. \tag{11}$$

This addition was made in order to provided a gauge invariant way of introducing masses to the fermions and bosons of the standard model.

Despite successes of the standard model, the theory has several shortcomings to the theory. These include the absence of gravity, the zero mass of the neutrino's which according to recent discoveries have a small but non-zero mass, the absence of dark matter candidates and the fine-tuning problem. The fine-tuning problem consist of the Higgs mass getting contributions from higher order corrections. Taking these corrections into account the BEH mass near the planck scale, should be chosen extremely accurately with a precision in the order of 10^{24}. This precision matching up perfectly would be a minor miracle, however the theory of supersymmetry (SUSY) has a way of solving not only this last problem but also a way of adding possible candidates for the hypothetical dark matter particles.

2.2 Supersymmetry

Supersymmetry adds an additional symmetry to the standard model [1]. It introduces a fermion to every boson and a boson to every fermion. These are often referred to as the superpartners. This solves the hierarchy problem in the calculation of the corrections to the BEH-boson mass. The contributions of all particles are canceled out by the contributions of their superpartners.

The assumption seems easy enough but none of these proposed particles have of yet been discovered. Should supersymmetry be a complete symmetry of the theory, then the masses of the superpartners should have been the same as their respective partners and they should have been discovered. This means that SUSY has to be broken, at least to some point. The solution is therefore to introduce a SUSY Lagrangian comprised of two parts, one that conserves the symmetry and one that breaks it.The cancellation of the quadratic terms in the

BEH-boson mass correction is desired to be preserved, otherwise the point of introducing SUSY would be pointless. SUSY-breaking terms that do this are called soft breaking terms.

Many variations of supersymmetry exist. The most prominent of which is the minimal supersymmetric standard model (MSSM). This theory introduces SUSY in order to solve the hierarchy problem by introducing as few parameters as possible. The MSSM introduces a new symmetry called R-parity. R-parity of a process is obtained from the multiplication of the R-parity of all individual particles (R_p).

$$R_p = (-1)^{3(B-L)+2S} \tag{12}$$

With B the baryon number, L the lepton number and S the spin. This symmetry would cause the lightest supersymmetric particle (LSP) to be stable.

Many searches have been performed for the supersymmetric partners arising from the MSSM. Because the lack of any signal, it would seem that the MSSM is not the true model that describes nature. However, it is important to take into account the matter in which these searches are made. The detection of particles using the experiments at the LHC are usually not optimized for the detection of particles arising from interaction vertices far from the primary interaction vertex. Should the superpartners have sufficiently long lifetimes, they could travel far enough from the primary interaction vertex and be missed by standard searches. Therefore, studies dedicated to "displaced" models are important.

Fortunately, the theory of Displaced Supersymmetry allows for particles with sufficiently long lifetimes [4]. This theory introduces R-parity violating terms into the SUSY-Lagrangian. This is done carefully by only allowing the lepton number to be violated and in doing so preventing proton decay. This is important because the proton half-life has been found to be over 10^{32yr}. Introducing this extra term, allows superpartners. with lifetimes long enough to have escaped detection in standard SUSY searches.

3 The CMS experiment

The CMS experiment is an detector located at the Large Hadron Collider (LHC) at CERN in Geneva Switzerland. The LHC is a synchrotron with a circumference of 27 km which can accelerate protons up to energies of 14 TeV.

The CMS detector consists of multiple sub-detectors in different layers. This allows the detector to detect different types of particles and their properties. A schematic overview of the detector can be seen on figure 1.

The inner most layer surrounding the beampipe is the silicon tracker consisting of silicon pixels surrounded by silicon strip detectors with 66 million and 9.3 million read out channels respectively. The purpose of the tracker is the reconstruction of the particle tracks. It is also used in the determination the impact parameter which is especially important considering this analysis focused on the detection of muons with large impact parameters[5].

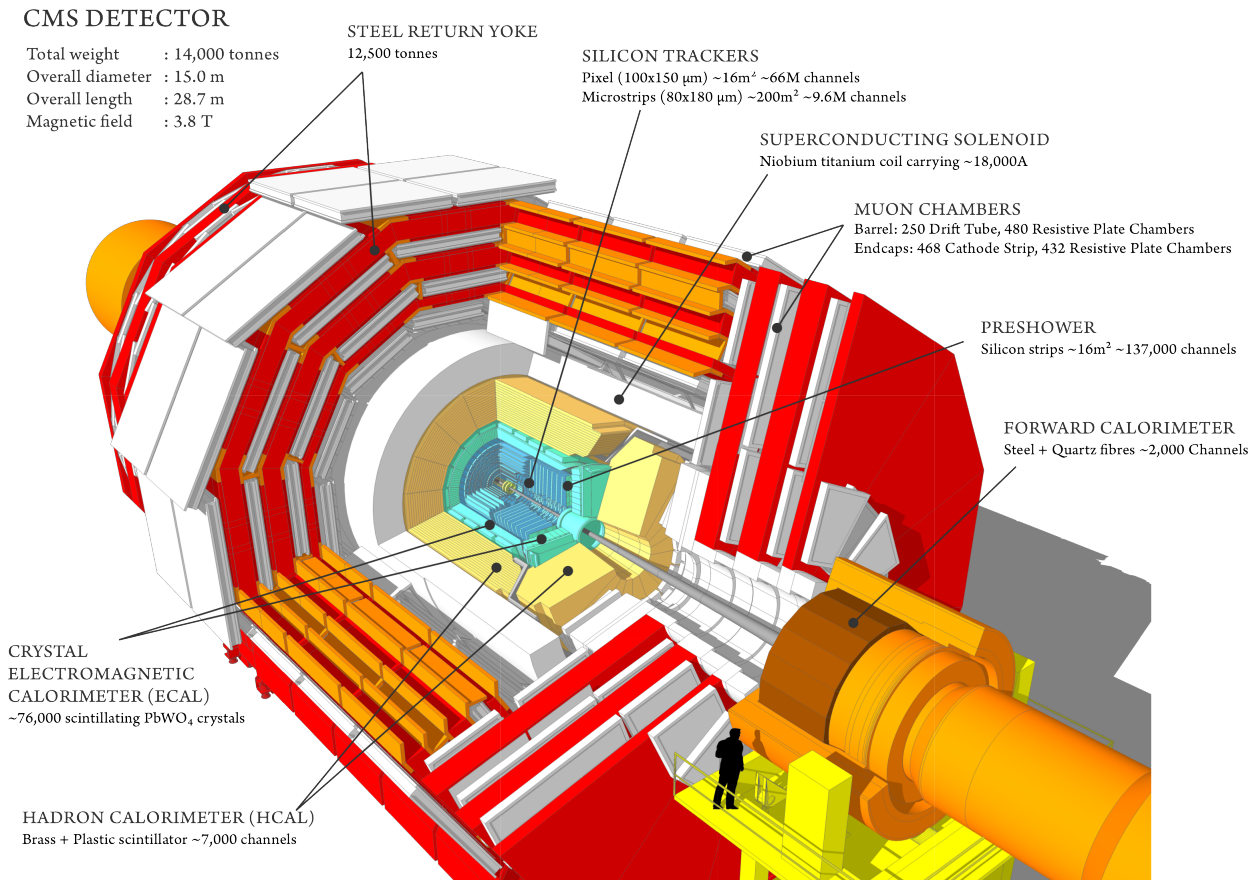

Figure 1: Schematic overview of the different layers of the CMS detector[12].

The next layer is an electromagnetic calorimeter (ECAL) which is made up of an array of lead tungstate crystals. Particles interact electromagnetically and are detected as they deposit their energy into the high density crystal. The ECAL is mainly used for the detection of electrons and photons[6].

The ECAL is followed by a hadronic calorimeter (HCAL). The HCAL is has alternating brass absorbers and plastic scintillator plates. The purpose of the brass layers is to adsorb the energy of the particles passing through it. The particles are detected as a light signal, when they pass through the scintillators[7].

Between the HCAL and the outer most layer, a superconducting solenoid is placed. It is capable of providing a magnetic field of 3.8T. It is primarily used to bend the track of charged particles, aiding in the identification and track reconstruction.

The muon chambers make up the outer layer of the detector. They are made up of cathode strip chambers and resistive plate chambers. The muons pass through

the muon chambers and are able to make multiple hits along the way[8].

4 Methodology

4.1 Signal Topology

The final state studied in this thesis consists of two top-squarks with opposite charges produced in the proton-proton collision. These stop quarks are considered to have a lifetime of sufficient length for them to travel a considerable distance from the primary interaction vertex. The stop quarks consequently decay into a lepton and a b-quark. Due to the strong force preventing quarks to appear as free states, the b-quark consequently transforms into a collection of b-hadrons which further decay. This collection of particles is called a b-jet. This results in a signature of two displaced leptons and two displaced b-jets. A lepton or jet is considered displaced if it has a significantly large impact parameter (d0), which is defined as the distance from the beamspot to the point of closest approach of the track. The definition of the impact parameter and the topology can be seen in figure 2. The analysis will disregard the b-jets in order to keep the search as general as possible and making it sensitive to other types of topologies. The final signature consists of two displaced leptons, which can be studied in ee, $\mu\mu$ and $e\mu$ channels. This analysis will focus and the $\mu\mu$ signature.

4.2 Data and MC samples

4.2.1 Data samples

The data used in this analysis consist of samples recorded using the CMS detector over all eras (B,C,D,E,F,G and H) in 2016 at $\sqrt{s} = 13$ TeV. This corresponds to an integrated luminosity of 35.7fb^{-1}. The *DoubleMuon* dataset was used as the search sample because it contains all data consistent with a di-muon signature. The *DoubleMuon* sample was also used for a data-driven QCD background estimation.

4.2.2 Standard model background and signal Monte Carlo samples

The MC samples for the background and signal estimation were produced by the CMS collaboration in the so called summer16 production campaign. All aspects of the CMS experiment are included in the simulation. These contain among others the geometry of the detector, the calibration of the detector, the alignment scenario, the performance of the various detection components and the magnetic field maps.

4.3 Event preselection

The first part of the analysis is the event preselection. This consist of reducing the data sample to only the events consistent with the displaced di-muon topology discussed in 4.1.

A summary of the different cuts applied in the preselection is given in ta-

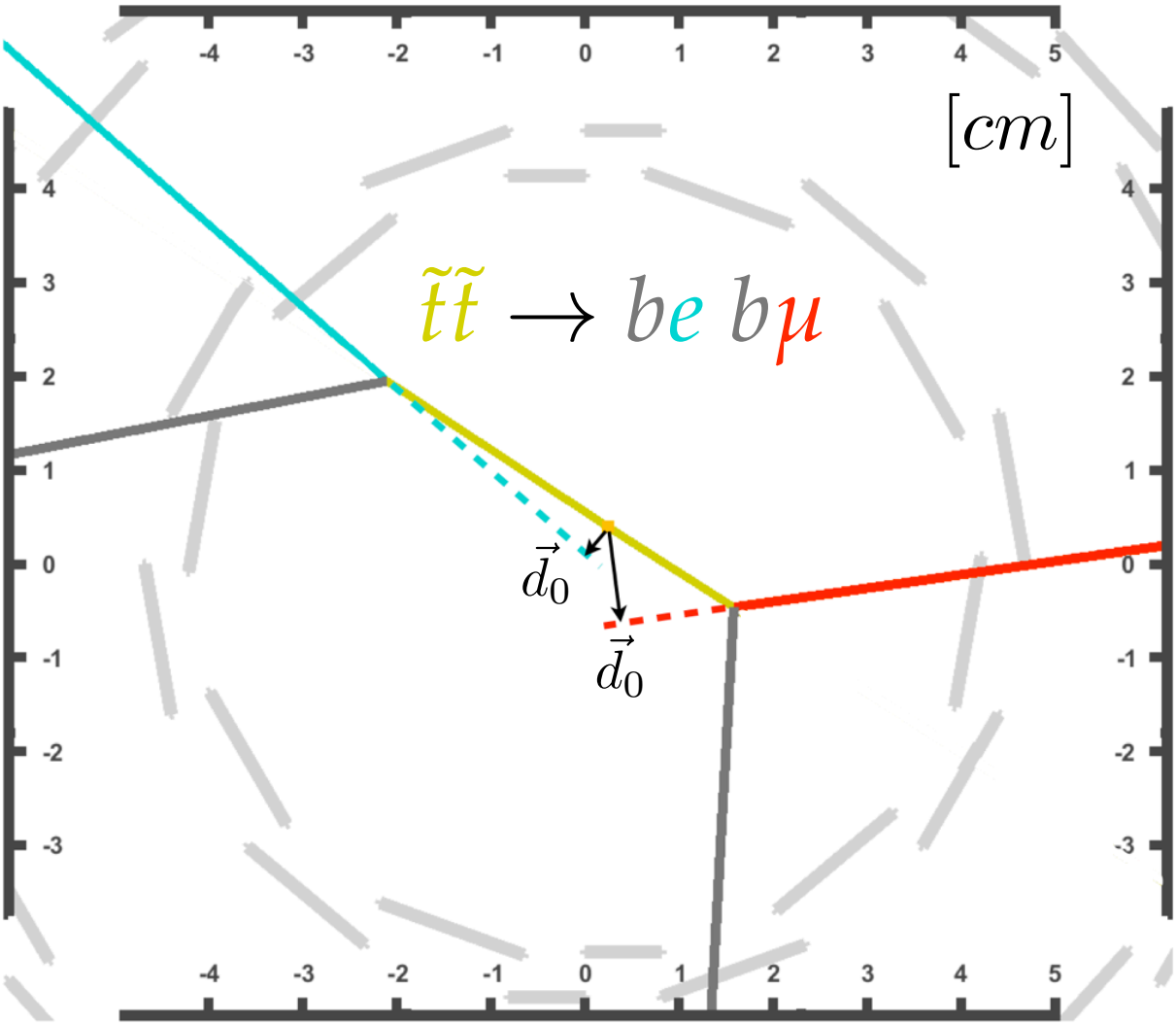

Figure 2: Transverse view of the inner part of the CMS detector, demonstrating MC simulation of a $\tilde{t}\tilde{t}$ event with displaced topology.[9]

ble 4.3. A cut is defined as the removal of certain events of parameters which do not satisfy certain criteria. Here η is the pseudorapidity, pT the transverse momentum, $\Delta R = \sqrt{\eta^2 + \phi^2}$, q_μ the charge of the muon and $\Delta\beta - correctedIso$ the isolation observable defined as the ratio of the transverse momentum of the particles around the muon ($\Delta R < 0.4$) to the transverse momentum of the muon [9] corrected with the $\Delta\beta$ method [10]. The isGlobalMuon and isPFmuon are muon identification requirements that depend on which part of the detector measured hits that were used in the muons track reconstruction. The $\chi 2$ is a measure of the track fit quality.

The cuts correspond with the so-called tight muon identification requirement.

$Cut - parameter$	$Cut - value$
$\|\eta\|$	< 2.4
pT	> 40
$\Delta\beta - correctedIso$	< 0.15
ΔR	> 0.5
$q_\mu * q_\mu$	-1
$isGlobalMuon$	$True$
$isPFMuon$	$True$
$\chi 2$	< 10
$nofTrackerLayersWithMeasurement$	> 5
$nofValidMuHits$	> 0
$nofValidPixelHits$	> 0
$nofMatchedStations$	> 1

Table 2: Summary of the muon cuts in the event preselection.

They are chosen in order to ensure the maximal quality of the reconstructed tracks and to reject fake reconstructed muons. The η, pT, $\Delta\beta - correctedIso$, ΔR $q_\mu * q_\mu$ requirements are used to make sure the corresponding particle is a muon arising from prompt processes and not processes like $\pi \to \mu X$, $K \to \mu X$ or $B \to \mu X$. The others ensure the high quality of the reconstructed track.

4.4 Regions

4.4.1 Prompt Control Region (PCR)

The prompt control region is comprised of events for which both muons pass the preselection and have an impact parameter ($|d0|$) smaller than 100 μm. The region is primarily used to verify that the Monte Carlo (MC) simulation was performed correctly and that the necessary corrections have been applied. Furthermore, this regions can be used for the calculation of additional normalization factors to further correct the MC simulation. The verification of the MC simulation is performed by comparing the distributions of the event parameters between data and MC. The prompt control region can be seen on figure 3 as CR I.

4.4.2 Displaced Control Region (DIR)

The displaced control region is comprised of events for which both muons pass the preselection and have an impact parameter between 100 μm and 200 μm. The main purpose of the displaced control region is the elimination of prompt sources like Z and W bosons. This region is further used to obtain a data driven sample of QCD background events. The displaced control region can be seen in figure 3 as CR II.

4.4.3 Signal regions

The three signal regions will be used to determine if there is any excess in data over the determined background.
There are two definitions of the three signal regions. The inclusive signal regions

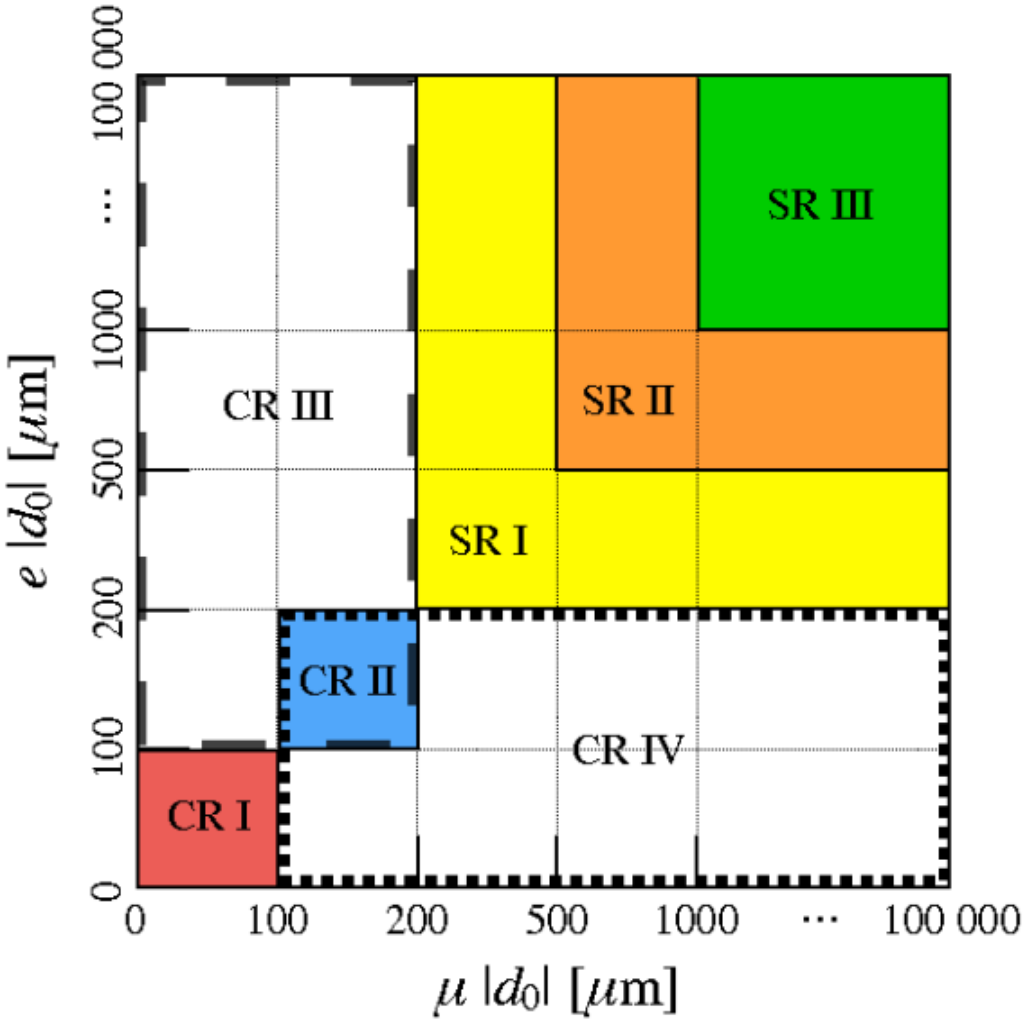

Figure 3: Definition of the di-muon regions. The Prompt control region (CRI), the Displaced control region(CRII) and the three exclusive signal regions (SRI, SRII and SRIII). [3]

are not mutually exclusive. The first signal region is defined as the region in which the muons pass the preselection and have a $d0 > 0.02\ \mu m$. The second has a $d0 > 0.05\ \mu m$ and the third a $d0 > 0.1\ \mu m$.
The exclusive signal regions are defined to be mutually exclusive. The definition is similar to the inclusive regions with the exception that events which are present in other regions are removed. This results in the first, second and third exclusive signal region which can be seen in figure 3 as the L shaped regions SR I, SR II and SR III respectively. The first inclusive region consists of the sum of SRI, SRII and SRIII, while the second inclusive signal region can be seen

as the sum of SRII and SRIII. The third inclusive signal region can seen as SRIII.

The inclusive signal regions will be used to estimated the expected number of events coming from known standard model processes while the exclusive signal regions will be used for the limit setting. The regions are constructed in this manner to eliminate the contamination of most of the prompt decaying processes making the analysis sensitive to displaced event signatures.

4.4.4 B-Enriched regions

Due to the limited statistics in the QCD MC simulation the resulting estimates will be highly inaccurate with large statistical errors. Therefore a data driven method is used to estimate the QCD contribution to the background in the signal regions. This data driven method uses the so called B-Enriched regions as input.

The B-Enriched region is dominated by QCD background events and is therefore an ideal candidate to study the QCD d0 distribution. The B-Enriched region is further divided in the Isolated and Non-Isolated B-Enriched regions.

These regions consider a different topology than the previously discussed dilepton regions. The events contained in this regions are comprised of a b-jet (hereafter referred to as the tag jet), an additional jet (hereafter referred to as the probe jet), which has a $\Delta\phi > 2.5$ with the tag jet, and a muon contained within the probe jet ($\Delta R < 0.5$). Where

$$\Delta\phi = |\phi_{TagJet} - \phi_{ProbeJet}| \tag{13}$$

and

$$\Delta R = \sqrt{(\eta_\mu - \eta_{ProbeJet})^2 + (\phi_\mu - \phi_{ProbeJet})^2}\,. \tag{14}$$

With η_μ and $\eta_{ProbeJet}$ denoting the pseudorapidity of the muon and probe jet respectively and ϕ_μ, ϕ_{TagJet} and $\phi_{ProbeJet}$ the azimuthal angle of the muon, tag jet and probe jet respectively. The jets are required to have $\eta < 2.4$ and a $p_T \geq 30GeV$. A summary of the cuts is given in table 3. The last two cuts in table 3 ensure the tag- and probe-jet to be back to back and the muon to be contained within the probe-jet.

4.4.5 B-Enriched Non-Isolated region

The B-Enriched Isolated region has the topology and cuts mentioned in section 4.4.4 with the additional requirement that the Isolation has a value between 0.15 and 1.5. This ensures that no signal is expected in this region.

4.4.6 B-Enriched Isolated region

The B-Enriched Non-Isolated region has the topology and cuts mentioned in section 4.4.4 with the additional requirement that the Isolation has a value lower than 0.15. This might cause data in the B-Enriched Isolated region to be

$Cut - parameter$	$Cutappliesto$	$Cut - value$
$\|\eta\|$	$muon$	< 2.4
pT	$muon$	> 40
$\Delta\beta - correctedIso$	$muon$	< 0.15
$isGlobalMuon$	$muon$	$True$
$isPFMuon$	$muon$	$True$
$\chi 2$	$muon$	< 10
$nofTrackerLayersWithMeasurement$	$muon$	> 5
$nofValidMuHits$	$muon$	> 0
$nofValidPixelHits$	$muon$	> 0
$nofMatchedStations$	$muon$	> 1
$\|\eta\|$	jet	< 2.4
pT	jet	> 30
$\Delta\phi$	$probe - jet, tag - jet$	> 2.5
ΔR	$muon, probe - jet$	< 0.5

Table 3: Summary of the muon and jet cuts in the B-Enriched region.

the same as in the signal regions. However due to the extremely low relative rate of potential events from the signal regions, the contribution of such events is negligible.

4.4.7 Isolated Inverted regions

The Isolated inverted regions are the same regions discussed in sections 4.4.1, 4.4.2 and 4.4.3. The exception being that the isolation criteria is inverted from < 0.15 to $0.15 < Iso < 1.5$. This region is useful because, due to the inverted isolation criteria, it has a much larger QCD contribution than the isolated signal regions have.

4.5 Background and signal estimation methods

The non-QCD background estimation will consist of obtaining data-MC agreement in the prompt control region. Once the prompt control region has given us the confidence to trust the MC simulation, the number of events in the signal region can be counted. Due to limited statistics in the signal regions, a method using muon efficiencies will be developed.

The QCD estimation method will consist of using the B-Enriched Non-Isolated region to calculated a transfer factor from the displaced control region to the signal regions. Multiplying this transfer factor with the number of estimated events in the displaced control region will produce a prediction for the amount of events in the signal regions. The B-Enriched Isolated and Isolated Inverted regions will be used to verify this result.

The signal estimation will be done by direct counting the events in the Monte Carlo sample after the necessary correction have been applied.

5 Non-QCD background estimation

The Non-QCD background consists of

- $Drell - Yan$ production, which is the annihilation of a quark with an antiquark producing a electroweak gauge boson (γ and Z^0) which decays into two muons.
- $t\bar{t}$ decay to two muons, two quarks, a neutrino and an anti-neutrino through two intermediate bosons
- *Diboson*, two W or Z bosons decay to muons and (anti)neutrinos.
- *SingleTop*: a top quark is produced in a process which produces an additional boson. The top quark decays to a muon and a (anti)neutrino while the boson generates at least one more muon.
- $W \rightarrow \mu\nu$: A W-boson decays to a muon and an (anti)neutrino. In the diagram of this process an additional muon is produced through a subprocess. For example leptonic decays of heavy-flavour hadrons.

5.1 Prompt Control region

5.1.1 Normalization to luminosity

The first act to be performed on the MC simulation is to normalize it to the integrated luminosity of the data taking period. The number events of a certain process is given by

$$N = \mathscr{L} \cdot \sigma \tag{15}$$

Where N is the number of events, $\mathscr{L}$ the integrated luminosity and σ the cross-section of the process in question. From this a normalization factor can be calculated. This normalization factor is given by

$$NF = \frac{\mathscr{L}_d}{\mathscr{L}_{mc}} = \frac{\mathscr{L}_d \cdot \sigma}{N_{mc}} \tag{16}$$

Where $\mathscr{L}_d$ and $\mathscr{L}_{mc}$ is the integrated luminosity of the data and MC respectively and N_{mc} the number of event in the MC sample. Scaling the number of events generated in the MC sample with the normalization factor will result in a MC sample with the same integrated luminosity as the data. The result can be seen in figure 4. As can be seen the MC seems to match the data on the Z peak, however the off-Z-peak MC does not model the data well.

5.1.2 Application of pile-up, Isolation and ID scale factors

After the normalization to the luminosity, the pile-up, Isolation and ID scale-factors are applied. These scalefactors have to be applied in order to account for the differences between data and MC of the pile-up, isolation and ID distributions.

The MC simulation is made using various assumption about the distribution of the physical parameters that control the proton-proton collisions. When

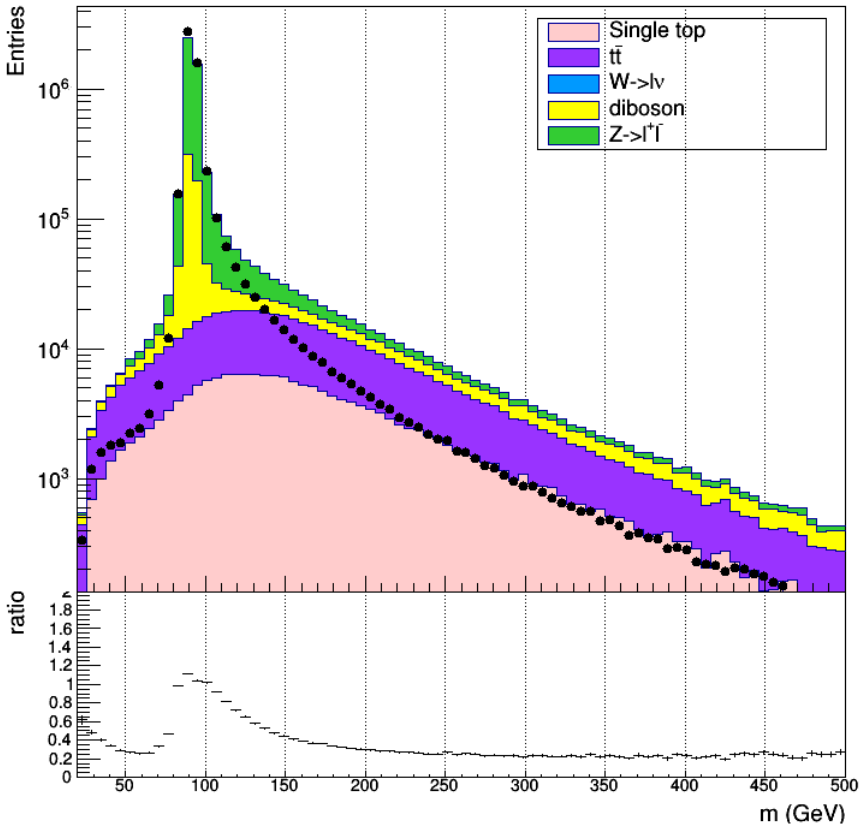

Figure 4: $\mu\mu$ invariant mass histogram after renormalization of the MC samples to luminosity. No Pile-up, Iso and ID scale factors are applied.

comparing the mean of the distribution between data and MC, there are differences. In order to take this into account, the ratio is taken from the two normalized distributions and the resulting histogram can be used to scale the MC events in order to obtain the same distribution as seen in data.

This is done for the pile-up; isolation and ID distributions. The MC samples are consequently scaled using the obtained histograms. The result can be seen in figure 5. A major improvement can be seen as the MC seems to follow the distribution of the data. This can be seen as the lower plot, which shows the ratio of the two, is flat and centered around the value of one.

5.1.3 Z-peak renormalization

The result obtained in section 5.1.2 can still be approved upon. When taking the ratio of the amount of data to MC, the value $91.15 \pm 0.07\%$ is obtained. This discrepancy can be explained by effects of the detector and reconstruction which were not taken into account during the MC production and the application of

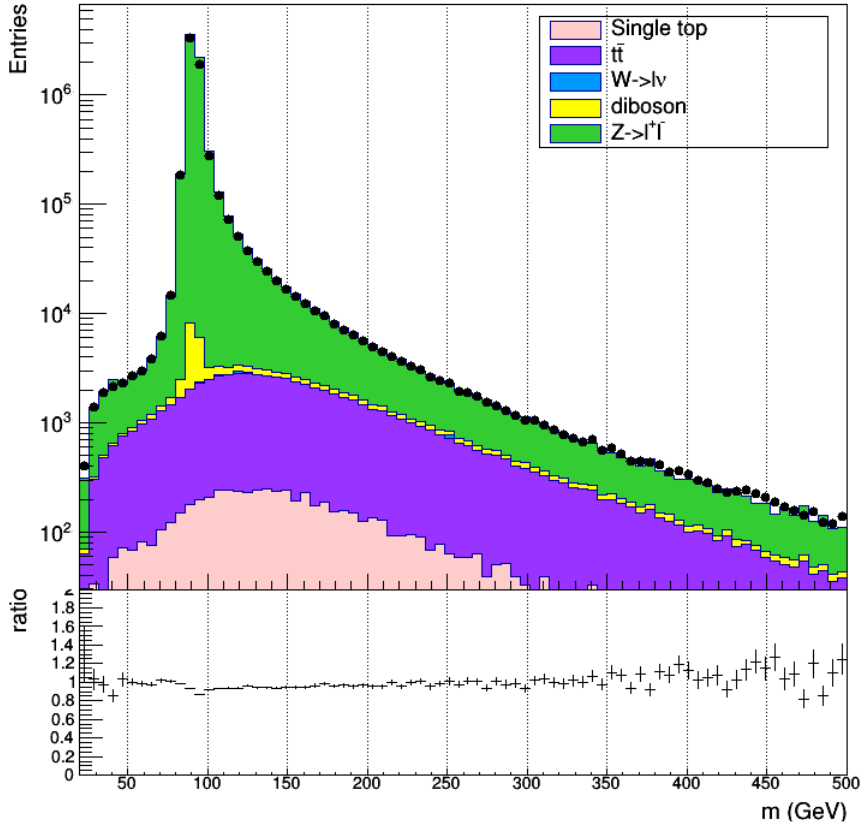

Figure 5: $\mu\mu$ invariant mass histogram after renormalization of the MC samples to luminosity and the application of the ISO-SF, ID-SF and pile-up-SF.

scale factors.

In order to improve the result, an additional normalisation factor is calculated using the Z peak in the invariant mass histogram. Using the fact that the ratio-plot is flat, the ratio between the number of events in data to MC can be used as a scale factor.
A 40 GeV window around the Z-boson mass of $91.1876 \pm 0.0021\%$ is taken and the Z peak normalization factor is calculated as

$$Z_{NF} = \frac{N_d}{N_{MC}} \tag{17}$$

Z_{NF} is the Z peak renormalization factor and N_d and N_{MC} are the number of events in data and MC respectively. This is done for every era (B,C,D,E,F,G and H) separately The final Z peak renormalization factor is then calculated as the luminosity weighted average

$$Z_{NF} = \Sigma_{era} Z_{NF,era} \frac{\mathscr{L}_{era}}{\mathscr{L}} \tag{18}$$

Era	Z_{NF}
B	0.88 ± 0.0010
C	0.83 ± 0.0014
D	0.86 ± 0.0011
E	0.83 ± 0.0012
F	0.80 ± 0.0013
G	0.97 ± 0.0010
H	1.03 ± 0.0010
$Total$	0.92 ± 0.0004

Table 4: Z peak renormalization factors for the different eras.

Process	Yield
$Z \to l^+l^-$	310298.43 ± 738.78
$Diboson$	5696.02 ± 33.80
$t\bar{t}$	47848.43 ± 96.86
$SingleTop$	4762.30 ± 76.74
$W \to l\nu$	271.17 ± 137.94
$Data$	387256.00 ± 622.30
Data-MC-Comparison	$Data/MC$
$Data/MC$	$104.29 \pm 0.27\%$

Table 5: Non-QCD MC background contribution and data yields with data-MC agreement in the prompt control region.

Where the sum is made over the different eras, $Z_{NF,era}$ is the Z peak renormalization factor in the era and $\mathscr{L}_{era}$ is the integrated luminosity in the era. The Z peak renormalization factors for the different eras can be found in table 4.

The Z peak renormalization factor is consequently applied to the MC samples. The result can be seen on figure 6. The ratio of data to MC is now 104.29 ± 0.27. The remaining discrepancy can be attributed due to the absence of the QCD background contribution. This was left out due to the limited statistics giving inaccurate results.

The yields of the different Non-QCD MC samples and the data MC comparison can be seen in table 5.1.3.

Seeing the agreement between data and MC on figure 6 gives confidence to the fact the data and MC should agree in the Displaced control and signal regions.

In figure 7 can be seen the pT distribution. Agreement can be seen between data and MC in the pT distribution.

In figure 8 and 9 the distribution of the η and ϕ distributions, respectively can be seen. On both the structure of the detector can be seen. On figure 8 this can been seen around the value of 0.3 corresponding with a gap between the muon chambers. On figure 9 the structure of the CMS detector is even more

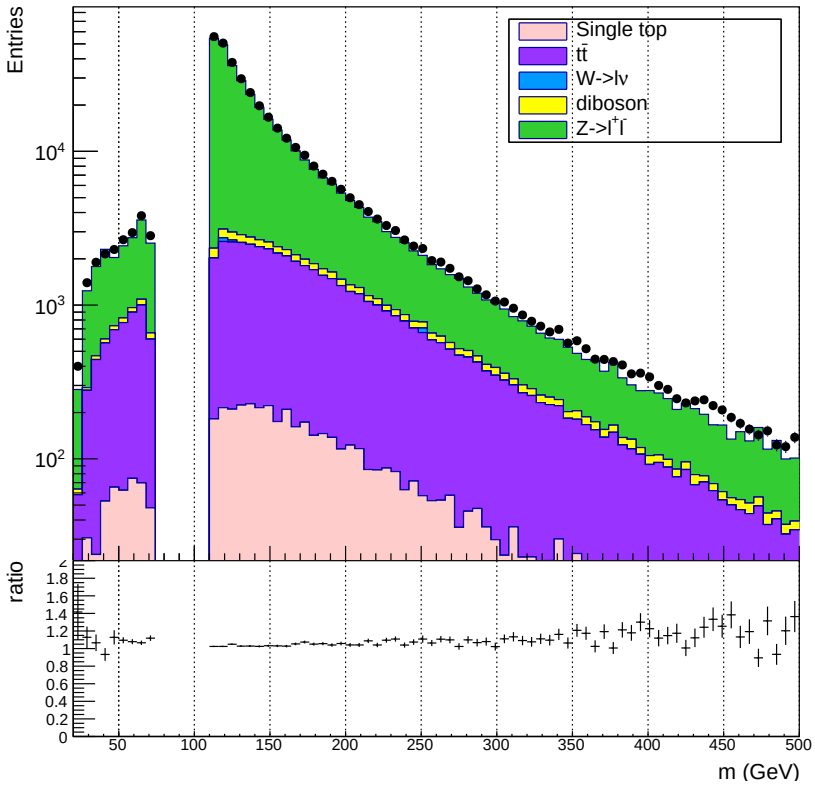

Figure 6: $\mu\mu$ invariant mass histogram after renormalization of the MC samples to luminosity, the application of the ISO-SF, ID-SF and pile-up-SF and the Z peak renormalization factor.

apparent as the twelve sections that can be seen on figure 1 are visible as small peaks. This is due to the fact that less events are measured in the regions where the different sections connect.

On figure 10 can be seen the distribution of the impact parameter. In contrast to the previous distributions, the MC does not appear to model the data very well. First of all the absence of the QCD background contribution must be noted. This contribution will mostly be contained at higher values in the d0 distribution. This would account for the underestimation of the data at higher values of d0. The discrepancy at lower values can be accounted for by the fact that the alignment scenario used in the MC simulation is only an estimate of the true alignment scenario. However this effect only influences the impact parameter up to a value of about 100 μm. The discrepancy will therefore not effect the signal regions which are found at higher values of the impact parameter[9].

Figure 11 shows the distribution of the muon isolation. The MC seems to

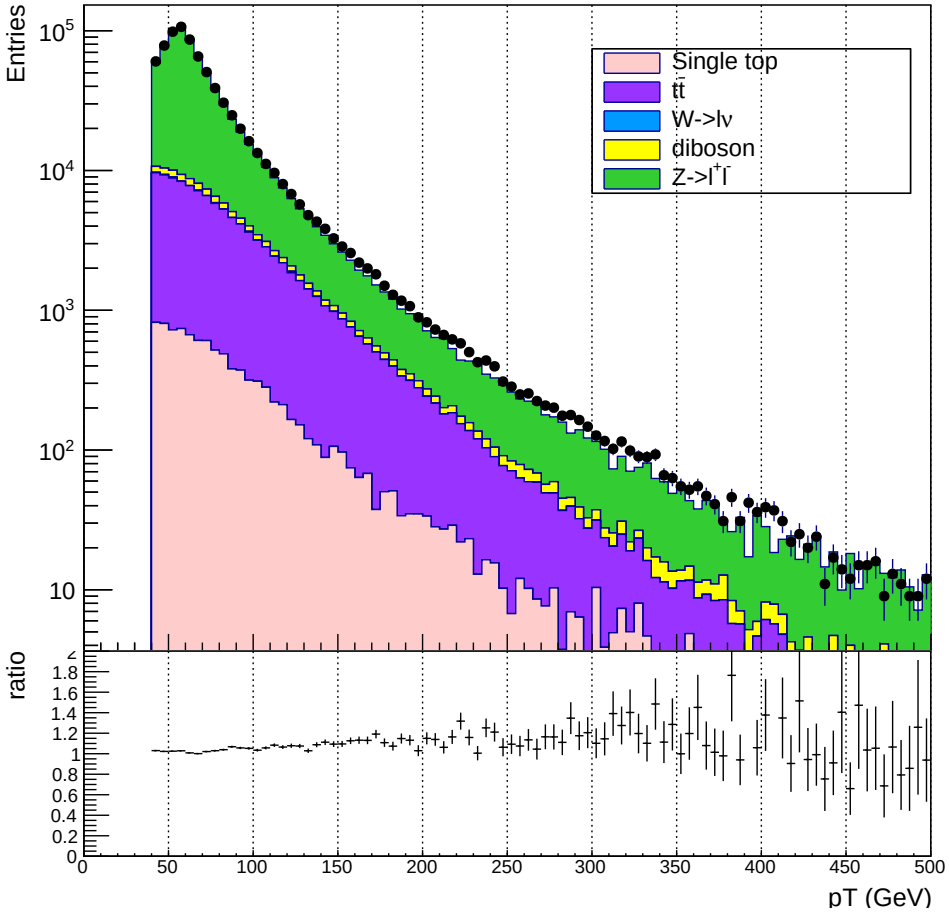

Figure 7: Muon pT histogram after renormalization of the MC samples to luminosity, the application of the ISO-SF, ID-SF and pile-up-SF and the Z peak renormalization factor.

match the data quite well. Only small discrepancies at higher values which can be accounted for by the absence of the QCD background contribution in simulation.

In figures 12 and 13 can be seen the Non-QCD MC distributions of the η and ϕ distributions for each of the muons of the di-muon signature separately. The positively charged muon is set on the horizontal axis while the negatively charged muon is set on the vertical axis. On both figures a correlation between the two muons can be seen. On figure 12, the correlation is due to the fact that the values of eta are higher at zero. This can also be seen on figure 8. Once more the structure of the CMS detector can be seen around values of ± 0.3 which on figure 12 can be seen as a double cross. On figure 13 the correlation is even

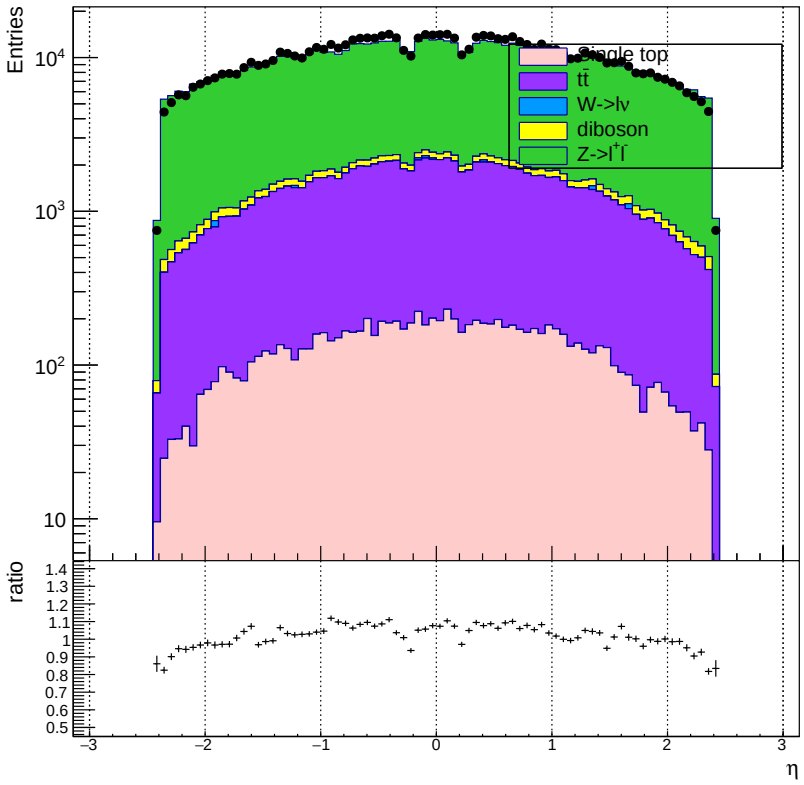

Figure 8: η distribution after renormalization of the MC samples to luminosity, the application of the ISO-SF, ID-SF and pile-up-SF and the Z peak renormalization factor.

more apparent. This is due to the fact that the muons are mainly produced back to back as the Z-boson is much heavier than the muons.

Figure 14 shows the d0 Non-QCD MC distribution for both muons separately. As is apparent, most of the muons arise from prompt production processes. When going to higher values of the impact parameter of both muons the number of expected events is quite low. This will be even more apparent in the displaced control region and will result in a low Non-QCD background contamination in the signal regions.

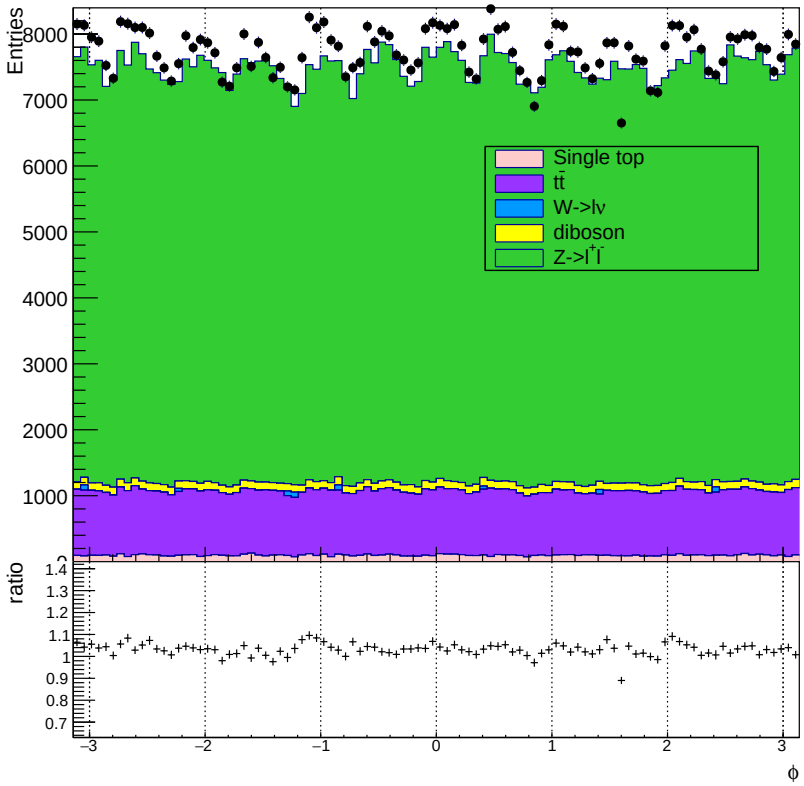

Figure 9: ϕ distribution histogram of the muons after renormalization of the MC samples to luminosity, the application of the ISO-SF, ID-SF and pile-up-SF and the Z peak renormalization factor.

5.2 Displaced Control region

Now that confidence has been established in the data MC agreement using the prompt control region, it is time to look at the displaced control region. The displaced control region is mainly used to estimate the QCD contribution in this region, in order to get an estimate for the QCD background contribution in the signal regions using a transfer factor method. Because signal is expected to be found at higher values of d0 and the Non-QCD contributions are known from the MC simulation, the excess in data can be explained as coming from QCD processes. The amount of QCD background events in the displaced control region can thus be estimated as the data with the Non-QCD MC contribution subtracted.

5.2.1 Z peak discrepancy

Figure 15 shows the invariant mass distribution of the muons in the displaced control region. There is an excess of events in data. The ratio of data to MC

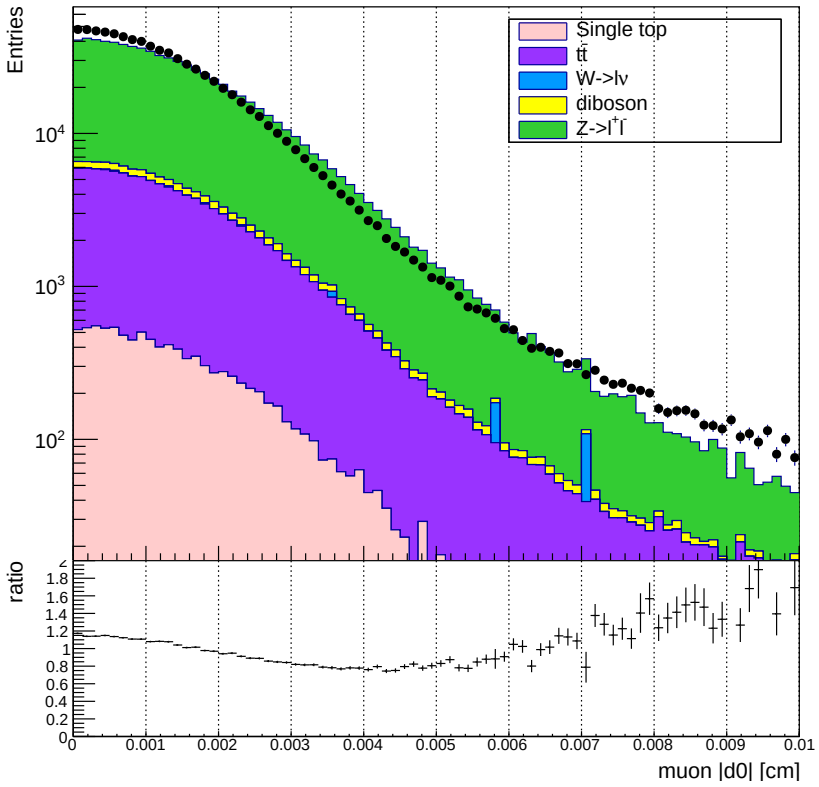

Figure 10: Muon d0 distribution histogram after renormalization of the MC samples to luminosity, the application of the ISO-SF, ID-SF and pile-up-SF and the Z peak renormalization factor.

is $451.45 \pm 299.76\%$. However the excess seems to peak around the value of the Z-boson mass. This indicates that the excess contains events coming from Z bosons. Therefore the Z peak is removed from the displaced control regions and the signal regions. The result of this can be seen on figure 16. The yields of the different Non-QCD MC samples and the data MC comparison can be seen in table 6.

5.2.2 Estimation of QCD in the displaced control region

Because the remaining excess in data over the Non-QCD MC background contributions is considered to be coming from the QCD Background contribution, the amount of QCD events can be estimated by subtracting the Non-QCD MC from the data in the displaced control region. This results in an estimated 11.68 ± 3.62 QCD events.

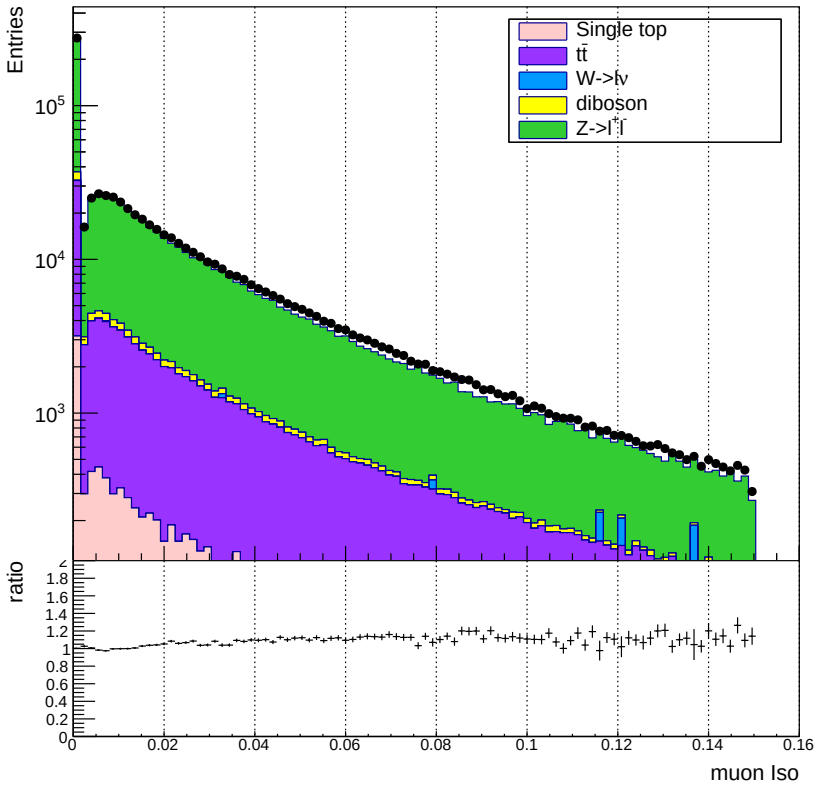

Figure 11: Muon isolation histogram after renormalization of the MC samples to luminosity, the application of the ISO-SF, ID-SF and pile-up-SF and the Z peak renormalization factor.

5.3 Signal region

5.3.1 Estimation the Non-QCD background using direct counting

The first method to estimate the number of events, arising from Non-QCD contributions, is a direct counting method. The MC samples corrected with all the normalization and scale factors, calculated in previous sections and with the Z peak removed, are used. In the respective regions, the number of events are retrieved directly from the histogram.

The result can be seen in table 7. Due to the limited number of MC events in the signal regions (two $t\bar{t}$ events and one $Z \to l^+l^-$ event), the results have huge relative uncertainties.

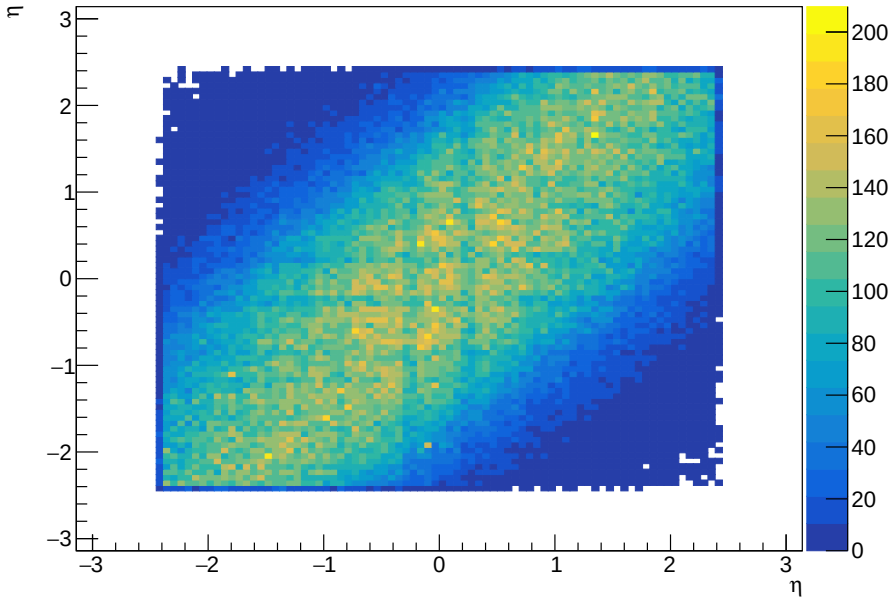

Figure 12: Muon η vs η histogram after renormalization of the MC samples to luminosity, the application of the ISO-SF, ID-SF and pile-up-SF and the Z peak renormalization factor. The positively charged muon is placed on the horizontal axis and the negatively charged muon on the vertical axis.

Process	Yield
$Z \to l^+l^-$	2.85 ± 2.01
$Diboson$	0.032 ± 0.014
$t\bar{t}$	0.44 ± 0.27
$SingleTop$	0.0 ± 0.0
$W \to l\nu$	0.0 ± 0.0
$Data$	15.00 ± 3.87
Data-MC-Comparison	$Data/MC$
$Data/MC$	$451.45 \pm 299.76\%$

Table 6: Non-QCD MC background contribution and data yields with data-MC agreement in the displaced control region.

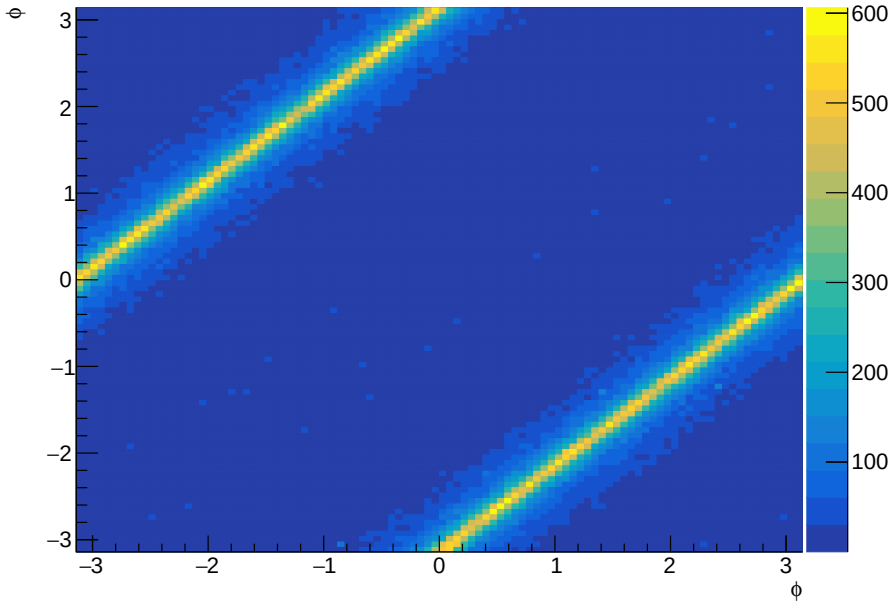

Figure 13: Muon ϕ vs ϕ histogram after renormalization of the MC samples to luminosity, the application of the ISO-SF, ID-SF and pile-up-SF and the Z peak renormalization factor. The positively charged muon is placed on the horizontal axis and the negatively charged muon on the vertical axis.

Inclusive regions	$d0 \in [0.02, 2.0]$	$d0 \in [0.05, 2.0]$	$d0 \in [0.1, 2.0]$
$Z \to l^+l^-$	1.05 ± 1.05	0.0 ± 0.0	0.0 ± 0.0
$Diboson$	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
$t\bar{t}$	0.36 ± 0.27	0.0 ± 0.0	0.0 ± 0.0
$SingleTop$	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
$W \to l\nu$	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Exclusive regions	$d0 \in [0.02, 0.05]$	$d0 \in [0.05, 0.1]$	$d0 \in [0.1, 2.0]$
$Z \to l^+l^-$	1.05 ± 1.05	0.0 ± 0.0	0.0 ± 0.0
$Diboson$	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
$t\bar{t}$	0.36 ± 0.27	0.0 ± 0.0	0.0 ± 0.0
$SingleTop$	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
$W \to l\nu$	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0

Table 7: Non-QCD MC background contribution in the inclusive and exclusive signal regions using the direct count method.

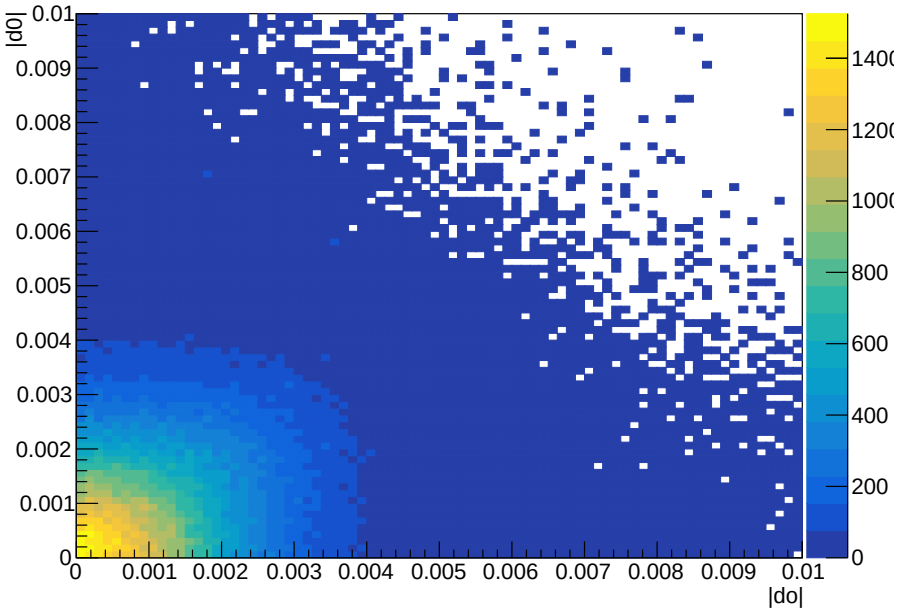

Figure 14: Muon d0 vs d0 histogram after renormalization of the MC samples to luminosity, the application of the ISO-SF, ID-SF and pile-up-SF and the Z peak renormalization factor. The positively charged muon is placed on the horizontal axis and the negatively charged muon on the vertical axis.

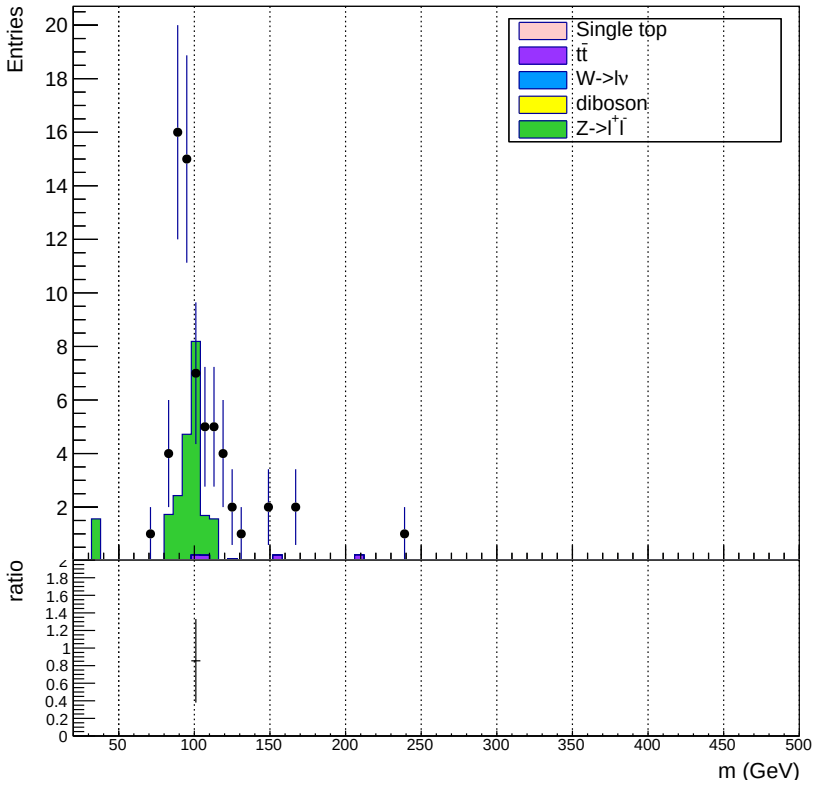

Figure 15: $\mu\mu$ invariant mass histogram in the displaced control region.

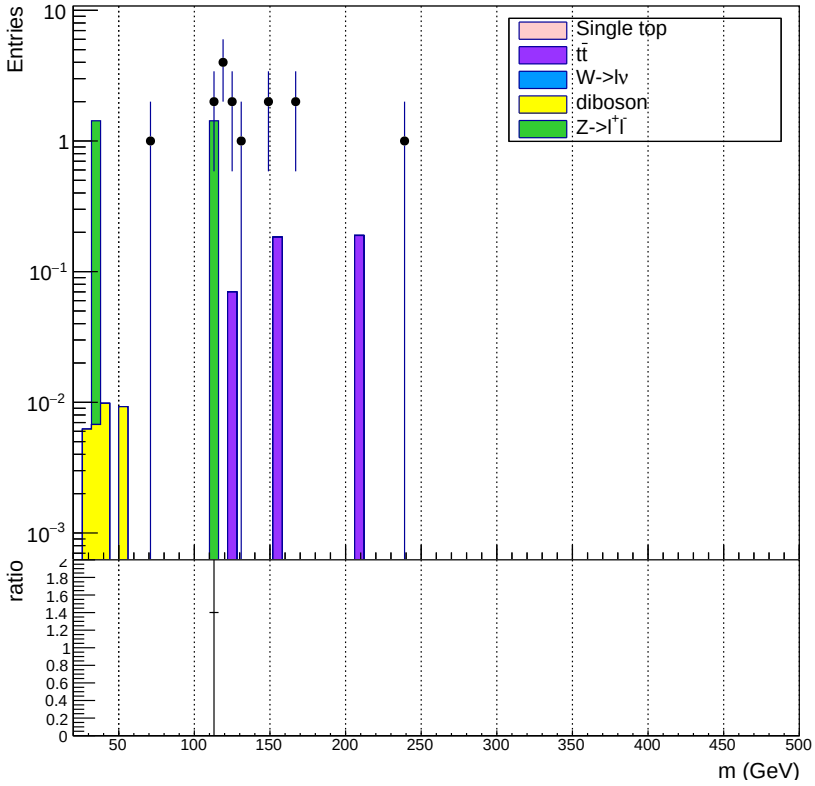

Figure 16: $\mu\mu$ invariant mass histogram in the displaced control region with the Z peak removed. The dominant contribution can now be assumed to be QCD events

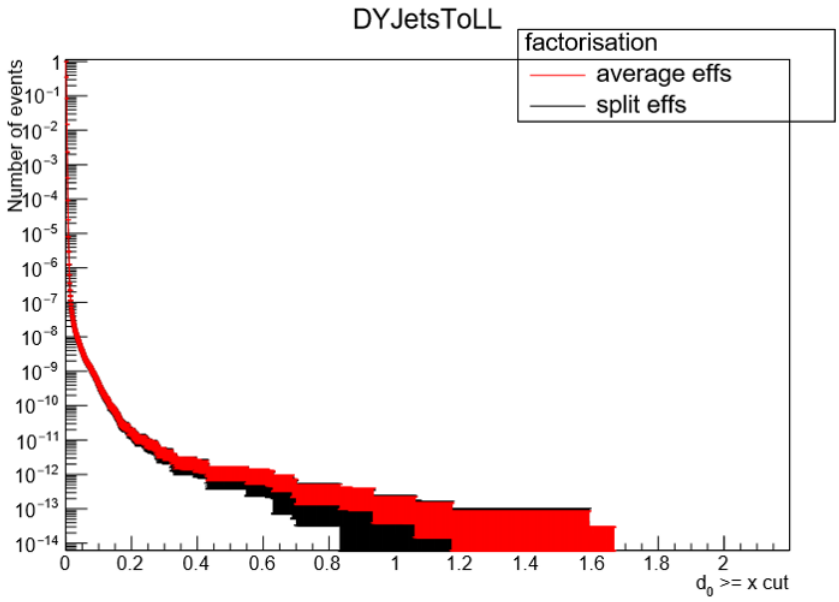

Figure 17: Muon efficiency for the Drell-Yan sample for both the average and split method.

5.3.2 Estimating the Non-QCD background using the Muon efficiencies

A second method to estimate the number of Non-QCD events in the signal regions is using muon efficiencies. The muon efficiency is defined as the number of muons passing a certain cut on the impact parameter divided by the total number of muons in the sample. The number of events in a certain region with both muons with an impact parameter larger than a certain d0 would be

$$N = N_{tot} * \epsilon_{\mu 1} * \epsilon_{\mu 2} \tag{19}$$

Where N is the number of predicted events, N_{tot} is the total number of events and $\epsilon_{\mu 1}$ and $\epsilon_{\mu 2}$ are the muon efficiencies for the first and second muon respectively

Two methods to calculate the efficiencies were used. The average method calculates the efficiencies using a single d0 distribution made up of both muons. The split method calculates a separate efficiency for the positive and negative muon separately using the respective d0 distributions. In the average method $\epsilon_{\mu 1}$ is the same as $\epsilon_{\mu 2}$ and in the split method $\epsilon_{\mu 1}$ and $\epsilon_{\mu 2}$ are the efficiency of the positive and negative muon respectively. Both methods are expected to give the same result.

The efficiencies can be seen on figures 17, 18, 19, 20 and 21 for the $Drell - Yan$, $Diboson$, $t\bar{t}$, $SingleTop$ and $W \rightarrow l\nu$ samples respectively As expected, both methods give similar results. The Non-QCD background estimation will be made using the split method.

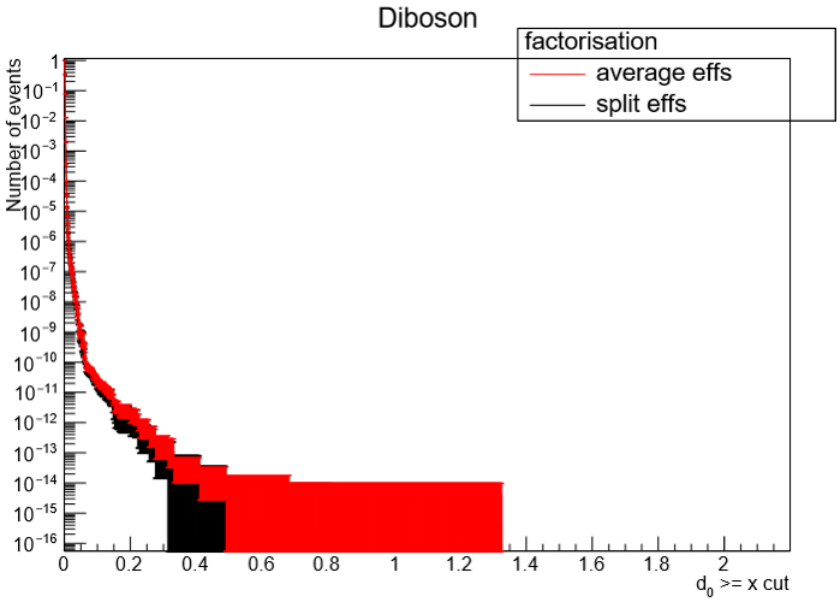

Figure 18: Muon efficiency for the Diboson sample for both the average and split method.

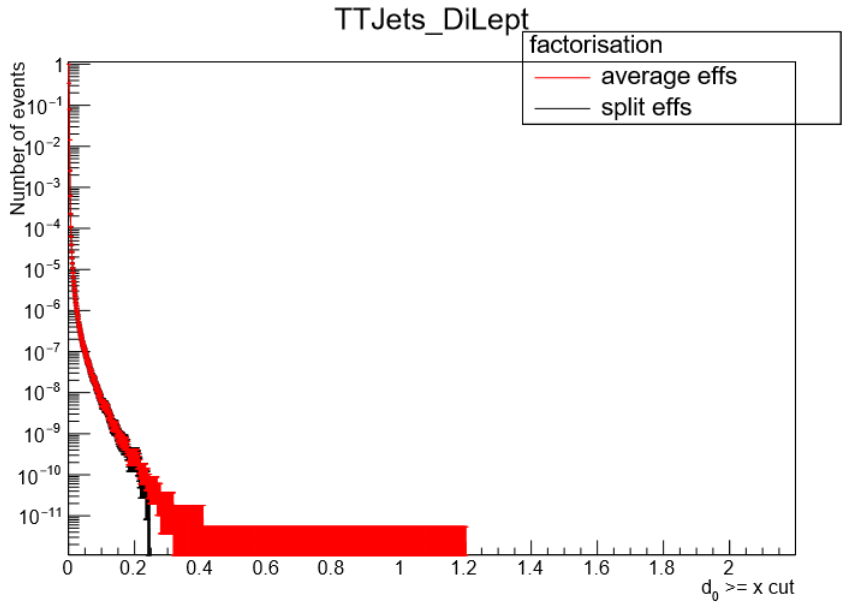

Figure 19: Muon efficiency for the $t\bar{t}$ sample for both the average and split method.

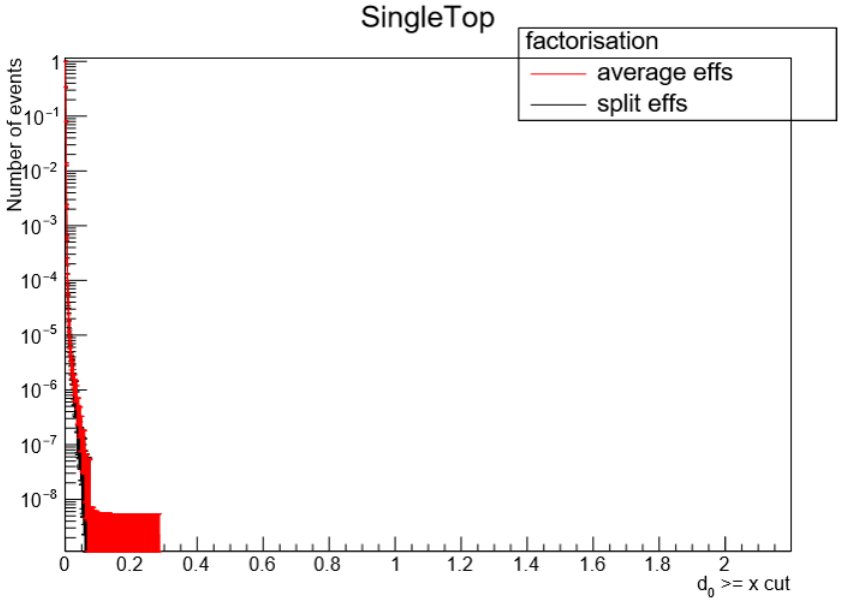

Figure 20: Muon efficiency for the SingleTop sample for both the average and split method.

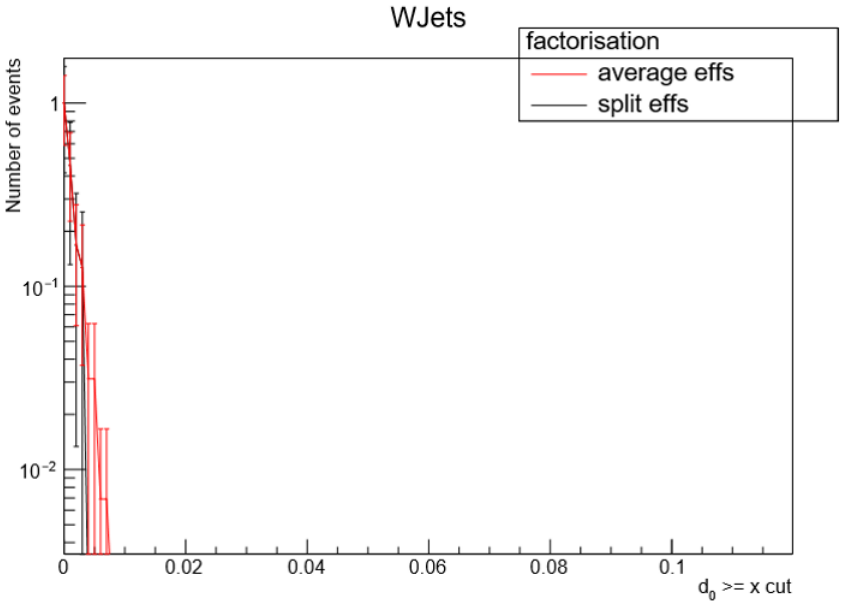

Figure 21: Muon efficiency for the $W \rightarrow l\nu$ sample for both the average and split method.

Inclusive regions	$d0 \in [0.02, 2.0]$	$d0 \in [0.05, 2.0]$	$d0 \in [0.1, 2.0]$
$Z \to l^+l^-$	0.28 ± 0.03	0.11 ± 0.01	$(0.23 \pm 0.05) \cdot 10^{-1}$
$Diboson$	$(0.35 \pm 0.14) \cdot 10^{-2}$	$(0.71 \pm 0.46) \cdot 10^{-5}$	$(0.32 \pm 0.20) \cdot 10^{-6}$
$t\bar{t}$	$(0.81 \pm 0.07) \cdot 10^{-1}$	$(0.36 \pm 0.07) \cdot 10^{-2}$	$(0.26 \pm 0.11) \cdot 10^{-3}$
$SingleTop$	$(0.12 \pm 0.07) \cdot 10^{-1}$	$(0.48 \pm 0.49) \cdot 10^{-3}$	0.0 ± 0.0
$W \to l\nu$	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0
Exclusive regions	$d0 \in [0.02, 0.05]$	$d0 \in [0.05, 0.1]$	$d0 \in [0.1, 2.0]$
$Z \to l^+l^-$	0.17 ± 0.03	$(0.83 \pm 0.14) \cdot 10^{-1}$	$(0.23 \pm 0.05) \cdot 10^{-1}$
$Diboson$	$(0.34 \pm 0.14) \cdot 10^{-2}$	$(0.67 \pm 0.46) \cdot 10^{-5}$	$(0.32 \pm 0.20) \cdot 10^{-6}$
$t\bar{t}$	$(0.78 \pm 0.07) \cdot 10^{-1}$	$(0.34 \pm 0.07) \cdot 10^{-2}$	$(0.26 \pm 0.11) \cdot 10^{-3}$
$SingleTop$	$(0.11 \pm 0.07) \cdot 10^{-1}$	$(0.48 \pm 0.50) \cdot 10^{-3}$	0.0 ± 0.0
$W \to l\nu$	0.0 ± 0.0	0.0 ± 0.0	0.0 ± 0.0

Table 8: Non-QCD MC background contribution in the inclusive and exclusive signal regions using the efficiency method.

The results of the Non-QCD background estimation using the efficiency method can be seen in table 8. The values seen are consistent with the numbers found in table 7. The value of 0.0 ± 0.0 for the $W \to l\nu$ can be understood when looking at the graph in figure 21. The sharp drop in efficiency means that the probability of finding a muon in the signal regions is extremely low. Therefore the estimate of 0.0 ± 0.0 is acceptable. Due to the low values in table 8, the contribution of Non-QCD background to the signal regions is minimal.

6 QCD Background estimation

Due to the limited statistics, the estimation method for the Non-QCD background described in section 5 won't work for the QCD estimation method. The direct counting method, described in 5.3.1, will result in an estimation of 0.0 ± 0.0 for all signal regions due to no MC events being present. The efficiency method, described in 5.3.2, will result in 0.0 ± 0.0 due to not a single muon in the di-muon events having an impact parameter larger than 0.02 cm. The number of QCD background events in the signal region is expected to be considerably higher. Therefore the reasoning that made the estimate of 0.0 ± 0.0 for the $W \to l\nu$ sample acceptable, can not be used. In order to get an accurate estimate, a data-driven method will be used.

6.1 Estimation of QCD using the B-Enriched Non-Isolated Control region

The data driven method consist of using the impact parameter distribution of the BEnriched Non-Isolated region. Given that the isolation criteria is not consistent with that of the signal regions, a check will be made that the impact parameter distribution between the non-isolated and isolated BEnriched regions are indeed the same. The QCD background will be estimated by calculating a

transfer factor for each inclusive signal region as

$$TF = \frac{N_{signal}}{N_{Displaced}}\,, \tag{20}$$

where TF is the transfer factor, $N_{displaced}$ is the amount of data events in the region with an impact parameter between 0.01 cm and 0.02 cm and N_{signal} is the amount of data events in the with an impact parameter between the boundaries of the respective signal region.

The amount of QCD in the signal region is consequently determined as the amount of QCD estimated in the displaced control region, calculated in section 5.2.2, multiplied by the transfer factor. The amount of QCD events in the exclusive regions is determined by taking the corresponding inclusive region and subtracting the amount of QCD events from the inclusive signal regions with larger $|d0|$.

On figure 22 can be seen the distribution of the impact parameter in the BEnriched Non-Isolated region. The discrepancy between the amount of data and the amount of MC events is of no concern because only the shape of the d0 distribution is of value to the QCD background estimation. The similarity in shape can be seen as the ratio plot being flat. As can be seen on figure 22, the QCD is by far the dominant contribution to the background in the BEnriched Non-Isolated region. The data distribution containing only QCD is obtained by taking the amount of data events in each bin and subtracting from it the Non-QCD MC which has been corrected with all the normalization and scale factors discussed in section 5. This distribution is used for the calculation of the transfer factors.

The transfer factors for the three inclusive regions can be seen in table 9.

Region	TF
$d0 \in [0.02, 2.0]$	$1.20 \pm 0.23 \cdot 10^{-2}$
$d0 \in [0.05, 2.0]$	$0.45 \pm 0.15 \cdot 10^{-2}$
$d0 \in [0.1, 2.0]$	$0.25 \pm 0.10 \cdot 10^{-2}$

Table 9: Transfer factors for the inclusive signal regions using the BEnriched Non-Isolated region.

These are multiplied with the QCD estimate in the displaced control region made in section 5.2.2. The resulting estimates are given in table 10.

Inclusive regions	$d0 \in [0.02, 2.0]$	$d0 \in [0.05, 2.0]$	$d0 \in [0.1, 2.0]$
QCD	14.01 ± 3.52	5.27 ± 1.32	2.93 ± 0.74
Exclusive regions	$d0 \in [0.02, 0.05]$	$d0 \in [0.05, 0.1]$	$d0 \in [0.1, 2.0]$
QCD	8.74 ± 3.76	2.33 ± 1.51	2.93 ± 0.74

Table 10: QCD background contribution in the inclusive and exclusive signal regions using the BEnriched NonIsolated region.

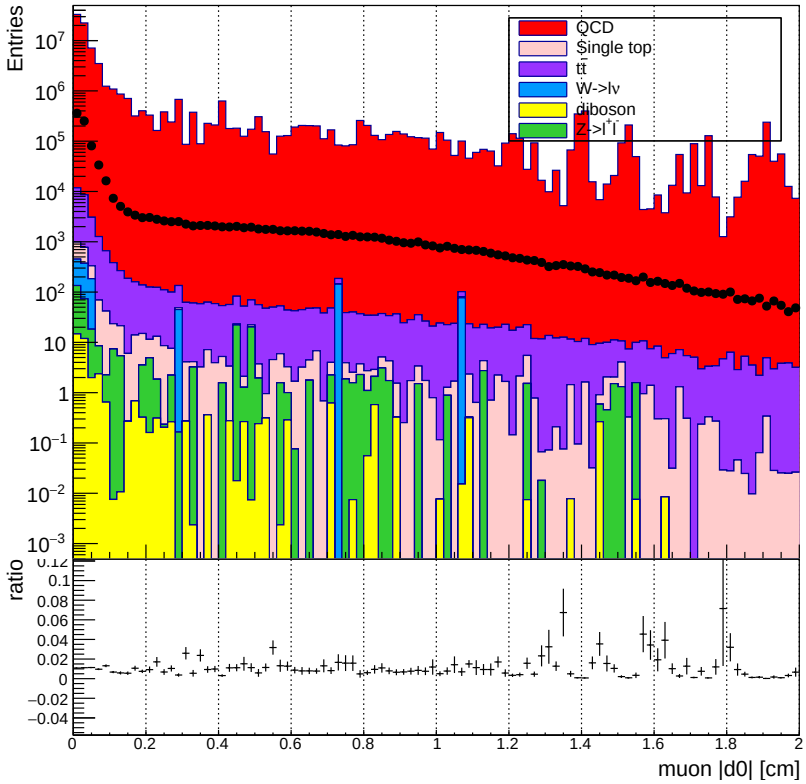

Figure 22: Muon impact parameter distribution the de BEnriched Non-Isolated region. The lack of correct normalization is of no concern since only the shape of the impact parameter distribution is important. The ratio plot proves the similarity in shape.

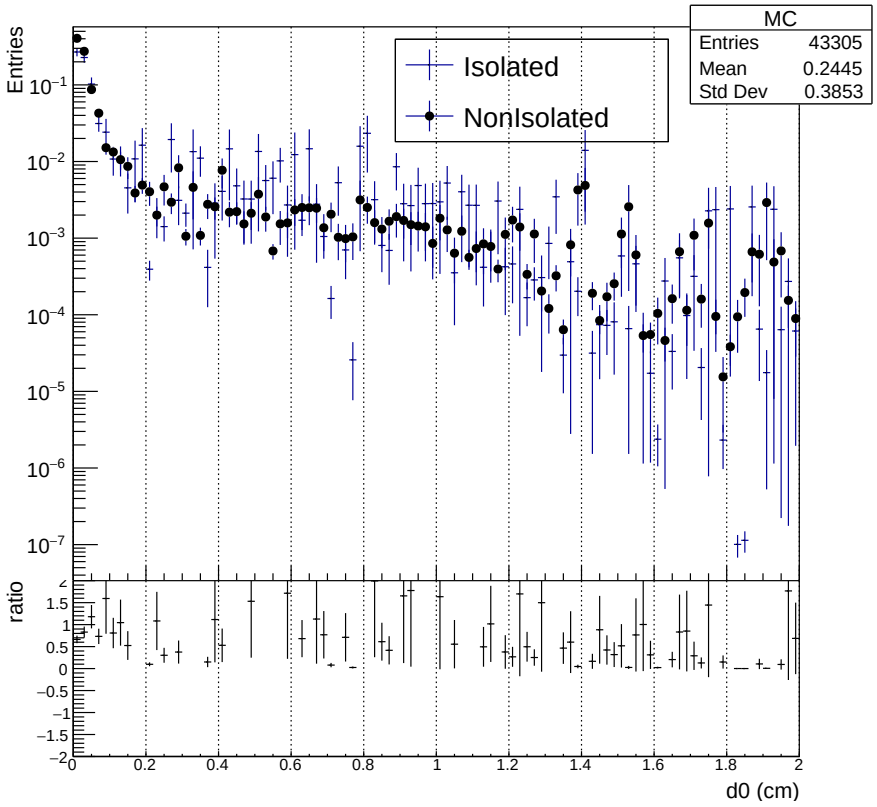

Figure 23: Comparison of the muon impact parameter distribution of the MC samples between the BEnriched Non-Isolated and BEnriched Isolated region.

6.2 Discrepancy in the impact parameter distribution between the Isolated and Non-Isolated B-Enriched regions

One of the main assumptions for the QCD data driven method to work, is that the impact parameter distribution has the same shape in the BEnriched Non-Isolated as in the BEnriched Isolated region, because in the signal region the isolation criteria $Iso < 0.15$ is applied rather than $0.15 < Iso < 1.5$. In order to check this, a comparison is made between the Isolated and Non-Isolated region. The result for the MC sample can be seen on figure 23. The distributions of the isolated and Non-Isolated regions have similar shape within uncertainties. However when looking at the distributions, a large fluctuation of the number of QCD events can be seen over the d0 interval, rather than a smooth transition. This can be attributed to inaccuracies in the production of the QCD MC samples. This is partly the reason a data-driven method is used for the QCD prediction.

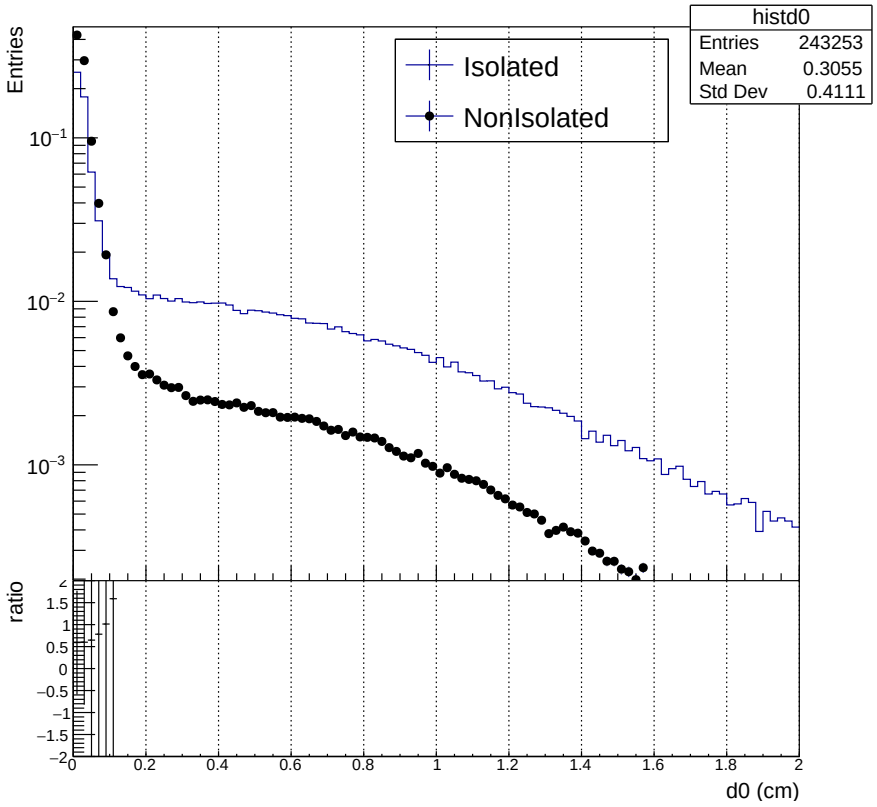

Figure 24: Comparison of the muon impact parameter distribution of the data between the BEnriched Non-Isolated and BEnriched Isolated region.

Figure 24 displays the comparison between the impact parameter distributions of the data in the Isolated and Non-Isolated regions. When comparing with figure 23, the impact parameter distribution is much smoother. However the distribution between the two regions do not appear to match. This is problematic as the amount of events in the displaced control region is higher for the Non-Isolated region, while the amount of events in the signal regions is lower. This means that the transfer factor will be different when the isolation criteria is changed. The BEnriched Non-Isolated region has a different criteria than the signal region, the background predictions were checked using an alternative approach in order to get a prediction with a higher level of confidence, additional predictions using different regions are made.

6.3 Estimation of QCD using the B-Enriched Isolated Control region

A second method to estimate the number of QCD events in the signal region is employing the same method as described in section 6.1 using the BEnriched

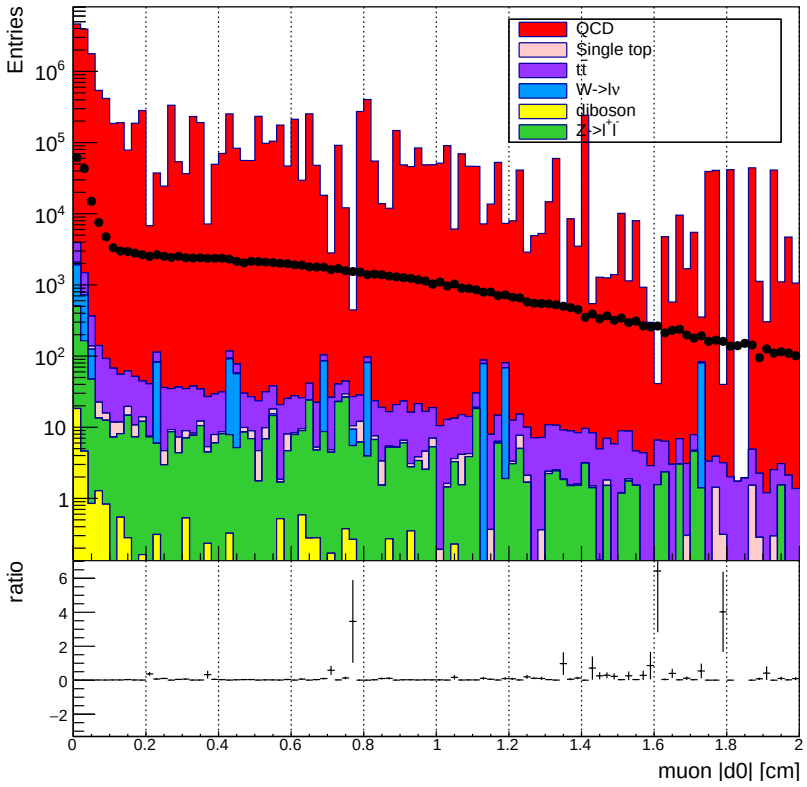

Figure 25: Muon impact parameter distribution in the BEnriched Isolated region. The shape of the distribution is the important factor for the determination of the transfer factors. The normalization is therefore of no concern. Large fluctuations can be seen in the QCD MC sample. This is due to an inaccurate simulation and is the reason a data driven method is used.

Isolated region. This region has the same isolation criteria as the signal regions, therefore it can be assumed that the impact parameter distribution is the same as in the signal regions. A possible drawback of this method is potential signal contamination in the regions utilized for the background estimation. However given the large number of events in this region with respect to the expected number of signal events, the potential contribution to the signal would be negligible.

Figure 25 shows the distribution of the impact parameter in the BEnriched Isolated region. MC simulations reasonably well describe the shape of the data. The data has smaller fluctuations which will result in a more accurate estimate using the data distribution rather the the MC sample.

Using the distribution of data displayed in figure 25, the transfer factors

Region	TF
$d0 \in [0.02, 2.0]$	2.72 ± 0.02
$d0 \in [0.05, 2.0]$	1.92 ± 0.01
$d0 \in [0.1, 2.0]$	1.65 ± 0.01

Table 11: Transfer factors for the inclusive signal regions using the BEnriched Non-Isolated region.

are calculated. The result can be seen in table 11. These transfer factors are

Inclusive regions	$d0 \in [0.02, 2.0]$	$d0 \in [0.05, 2.0]$	$d0 \in [0.1, 2.0]$
QCD	31.59 ± 7.94	22.35 ± 5.62	19.15 ± 4.81
Exclusive regions	$d0 \in [0.02, 0.05]$	$d0 \in [0.05, 0.1]$	$d0 \in [0.1, 2.0]$
QCD	9.24 ± 9.72	3.20 ± 7.39	19.15 ± 4.81

Table 12: QCD background contribution in the inclusive and exclusive signal regions using the BEnriched Isolated region.

consequently multiplied with the QCD estimate in the displaced control region given in table 12. When comparing the results in tables 10 and 12, one sees the difference caused by the discrepancy seen in figure 24. The values are considerably higher, however due to the large uncertainties they are still within three standard deviations of each other.

Initially it would seem that since the Isolated region possesses the same isolation criteria, it is kinematicaly closer to the signal regions and therefore the results obtained using this region would be more correct. However it is expected that the impact parameter distribution of both the Isolated and Non-Isolated region is the same. Because this is not the case, it is not clear whether the impact parameter distribution of the Isolated or Non-Isolated region is to be found incompatible with the impact parameter distribution of the signal regions. Because of the data being blinded, it is impossible to see which of the two distributions follows the one of the signal region most accurately. In order to cross-check result, an additional estimate is made using the Isolation Inverted region.

6.4 Estimation of QCD in Isolated inverted regions

The method using the isolation inverted region differs from the BEnriched regions in that it does not calculate a transfer factor from the displaced control region to the signal regions but from the isolation inverted signal regions to the signal regions. The impact parameter distribution in the isolation inverted region can be seen in figure 26. The data and MC show a similar shape with much less fluctuation in the data than in the MC. Using this distribution and the impact parameter distribution in the displaced control region, the transfer factor is calculated as

$$TF = \frac{N_{DCR}}{N_{DCRIsoInverted}}\,. \tag{21}$$

Where N_{DCR} is the amount of data events corresponding to QCD (The amount of data events with the amount of Non-QCD MC subtracted) in the displaced

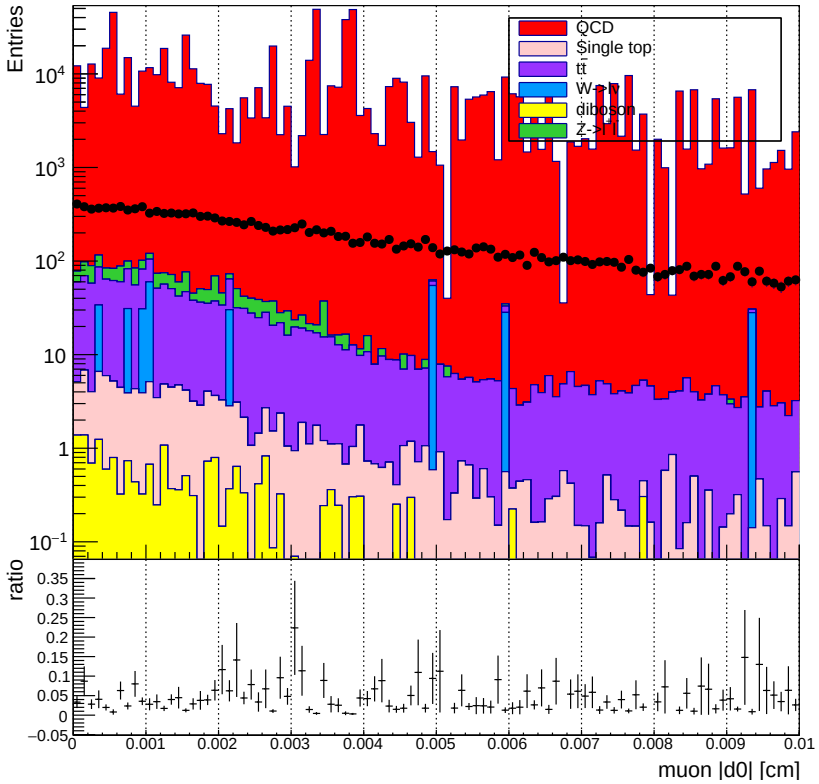

Figure 26: Muon impact parameter distribution in the Isolation Inverted region. Only the shape of the distribution is of concern for the calculation of the transfer factors. The normalization is therefore of no concern. The similarity in shape can be seen in the ratio plot, where the fluctuations can be attributed to the large variations in the QCD MC simulation rather than a possible difference in shape of the impact parameter distribution of the data.

control region and $N_{DCRIsoInverted}$ the amount of data events corresponding to QCD in the Isolated Inverted Displaced Control region. This transfer factor is determined to be $(0.98 \pm 0.41) \cdot 10^{-3}$. The transfer factor is multiplied with

Region	TF
$d0 \in [0.02, 2.0]$	28505.93 ± 6476.63
$d0 \in [0.05, 2.0]$	10971.77 ± 3828.09
$d0 \in [0.1, 2.0]$	638.24 ± 420.90

Table 13: The amount of QCD MC events in the Isolated Inverted Signal regions.

the amount of QCD events in the Isolation Inverted signal regions which can be seen in table 13. This results in QCD background contribution estimates which

Inclusive regions	$d0 \in [0.02, 2.0]$	$d0 \in [0.05, 2.0]$	$d0 \in [0.1, 2.0]$
QCD	28.00 ± 13.23	10.78 ± 5.84	0.63 ± 0.49
Exclusive regions	$d0 \in [0.02, 0.05]$	$d0 \in [0.05, 0.1]$	$d0 \in [0.1, 2.0]$
QCD	17.23 ± 14.47	10.15 ± 5.86	0.63 ± 0.49

Table 14: QCD background contribution in the inclusive and exclusive signal regions using the Isolation Inverted region.

can be seen in table 14.

6.5 QCD background prediction

When comparing table 14 with table 10 and 12, it can not be claimed that the method using the isolated inverted region follows the estimates of one of the other methods. It can therefore not be used to determine which of the two previous results is the most accurate.

The QCD background estimate will be chosen to be the most conservative of the three methods. Because this analysis look for an excess in data, the results of the BEnriched Isolated region which are the highest will be used. In order to make the analysis more sensitive, a study of the behavior of the impact parameter distribution can be performed. Once the discrepancy between the shapes of the impact parameter distribution is understood, one of the less conservative methods could be chosen or corrections to the methods could be applied.

This analysis however choses the most conservative estimate which can be found in table 12.

7 Systematic uncertainties

The final part of the background estimation process is the determination of systematic uncertainties arising from decisions made during the analysis. The two most prominent decisions where the width of the Z peak used for the determination of the Z peak renormalization factor and the Btag working point. In the BEnriched regions the required topology required one b-jet. The determination of whether a certain jet a indeed a b-jet is contingent on the value of the Btag

working point. This can be seen as the likelihood that a certain jet is indeed a b-jet.

The systematic uncertainties are determined by running the entire analysis for different values of the Z peak width and the Btag working point. The systematic uncertainties are than calculated by taking the standard deviation of the obtained predicted number of events for each of the different background contributions. The values used can be seen in table 15.

Zpeak Width	Btag Working point
30GeV	0.5426
40GeV	0.8484
50GeV	0.9535

Table 15: Values of the different Z peak widths an BTag working points. Combinations of both were used in the determination of the systematic uncertainties.

8 Signal estimation

The signal estimation will be performed using a direct count method. Signal samples in the prompt control, displaced control and inclusive signal regions are produced and the yields derived. The amount of signal events in the exclusive signal regions is calculated by subtracting from the inclusive region, the amount of signal events corresponding with the inclusive signal regions which is not part of the exclusive signal region being considered

8.1 Signal prediction in the Prompt Control region

The result of the signal estimation in the prompt control region can be seen in table 16. There are events expected for several of the signal samples, those corresponding too lower $c\tau$ and thus with a shorter path traveled by the stop quark and thus a smaller d0. However when comparing the amount of events with those in table 5.1.3, it is obvious that the amounts are to low to cause a significant contribution in this region. Therefore all steps in this analysis taken in the prompt control region, will not be influenced by potential signal contamination.

8.2 Signal prediction in the Displaced Control region

In table 17 the estimation for the signal in the displaced control region can be found. In this region, the amount of signal can play a significant role as the amount of QCD events estimated using the BEnriched regions is dependent on the amount of QCD in the displaced control region. This was calculated in section 5.2.2 as the data with the Non-QCD MC subtracted. However potential signal in this region will cause the amount of QCD estimation in the displaced control region to be increased and thus the corresponding predictions. Since the analysis looks for an excess in data, this only causes the estimates to be more conservative. Furthermore when looking at table 17, the samples which

$Mass_{C}\tau$	$Yield$
$M200c\tau1$	$16374.59 \pm 443.48(stat.) \pm 538.66(sys.)$
$M200c\tau10$	$645.65 \pm 88.66(stat.) \pm 10.38(sys.)$
$M200c\tau100$	$0.00 \pm 0.00(stat.) \pm 0.00(sys.)$
$M200c\tau1000$	$0.00 \pm 0.00(stat.) \pm 0.00(sys.)$
$M300c\tau1$	$2687.01 \pm 65.61(stat.) \pm 43.14(sys.)$
$M300c\tau10$	$112.98 \pm 13.43(stat.) \pm 2.71(sys.)$
$M300c\tau100$	$0.57 \pm 0.44(stat.) \pm 0.00(sys.)$
$M300c\tau1000$	$0.00 \pm 0.00(stat.) \pm 0.00(sys.)$
$M400c\tau1$	$707.04 \pm 15.59(stat.) \pm 5.81(sys.)$
$M400c\tau10$	$36.33 \pm 3.51(stat.) \pm 0.29(sys.)$
$M400c\tau100$	$0.60 \pm 0.46(stat.) \pm 0.00(sys.)$
$M400c\tau1000$	$0.00 \pm 0.00(stat.) \pm 0.00(sys.)$
$M500c\tau1$	$211.83 \pm 4.53(stat.) \pm 1.55(sys.)$
$M500c\tau10$	$11.73 \pm 1.07(stat.) \pm 0.00(sys.)$
$M500c\tau100$	$0.50 \pm 0.22(stat.) \pm 0.00(sys.)$
$M500c\tau1000$	$0.00 \pm 0.00(stat.) \pm 0.00(sys.)$
$M600c\tau1$	$71.25 \pm 1.52(stat.) \pm 0.14(sys.)$
$M600c\tau10$	$3.91 \pm 0.36(stat.) \pm 0.00(sys.)$
$M600c\tau100$	$0.12 \pm 0.06(stat.) \pm 0.00(sys.)$
$M600c\tau1000$	$0.06 \pm 0.05(stat.) \pm 0.00(sys.)$
$M700c\tau1$	$28.95 \pm 0.60(stat.) \pm 0.07(sys.)$
$M700c\tau10$	$1.88 \pm 0.15(stat.) \pm 0.00(sys.)$
$M700c\tau100$	$0.07 \pm 0.03(stat.) \pm 0.00(sys.)$
$M700c\tau1000$	$0.00 \pm 0.00(stat.) \pm 0.00(sys.)$
$M800c\tau1$	$11.97 \pm 0.25(stat.) \pm 0.01(sys.)$
$M800c\tau10$	$0.62 \pm 0.06(stat.) \pm 0.00(sys.)$
$M800c\tau100$	$0.01 \pm 0.01(stat.) \pm 0.00(sys.)$
$M800c\tau1000$	$0.00 \pm 0.00(stat.) \pm 0.00(sys.)$
$M900c\tau1$	$5.45 \pm 0.11(stat.) \pm 0.00(sys.)$
$M900c\tau10$	$0.33 \pm 0.03(stat.) \pm 0.00(sys.)$
$M900c\tau100$	$0.01 \pm 0.01(stat.) \pm 0.00(sys.)$
$M900c\tau1000$	$0.00 \pm 0.00(stat.) \pm 0.00(sys.)$
$M1000c\tau1$	$2.60 \pm 0.05(stat.) \pm 0.00(sys.)$
$M1000c\tau10$	$0.13 \pm 0.01(stat.) \pm 0.00(sys.)$
$M1000c\tau100$	$0.00 \pm 0.00(stat.) \pm 0.00(sys.)$
$M1000c\tau1000$	$0.00 \pm 0.00(stat.) \pm 0.00(sys.)$
$M1100c\tau1$	$1.34 \pm 0.03(stat.) \pm 0.00(sys.)$
$M1100c\tau10$	$0.08 \pm 0.01(stat.) \pm 0.00(sys.)$
$M1100c\tau100$	$0.00 \pm 0.00(stat.) \pm 0.00(sys.)$
$M1100c\tau1000$	$0.00 \pm 0.00(stat.) \pm 0.00(sys.)$
$M1200c\tau1$	$0.70 \pm 0.01(stat.) \pm 0.00(sys.)$
$M1200c\tau10$	$0.05 \pm 0.00(stat.) \pm 0.00(sys.)$
$M1200c\tau100$	$0.00 \pm 0.00(stat.) \pm 0.00(sys.)$
$M1200c\tau1000$	$0.00 \pm 0.00(stat.) \pm 0.00(sys.)$

Table 16: Signal yield in the prompt control region

$M200c\tau1$	$3089.81 \pm 193.03(stat.) \pm 83.79(sys.)$
$M200c\tau10$	$245.40 \pm 54.49(stat.) \pm 9.80(sys.)$
$M200c\tau100$	$0.00 \pm 0.00(stat.) \pm 0.00(sys.)$
$M200c\tau1000$	$0.00 \pm 0.00(stat.) \pm 0.00(sys.)$
$M300c\tau1$	$495.00 \pm 27.98(stat.) \pm 11.74(sys.)$
$M300c\tau10$	$88.17 \pm 11.79(stat.) \pm 2.72(sys.)$
$M300c\tau100$	$2.48 \pm 1.90(stat.) \pm 0.00(sys.)$
$M300c\tau1000$	$0.00 \pm 0.00(stat.) \pm 0.00(sys.)$
$M400c\tau1$	$130.09 \pm 6.67(stat.) \pm 0.54(sys.)$
$M400c\tau10$	$13.75 \pm 2.17(stat.) \pm 0.00(sys.)$
$M400c\tau100$	$0.28 \pm 0.21(stat.) \pm 0.00(sys.)$
$M400c\tau1000$	$0.00 \pm 0.00(stat.) \pm 0.00(sys.)$
$M500c\tau1$	$35.17 \pm 1.85(stat.) \pm 0.21(sys.)$
$M500c\tau10$	$4.10 \pm 0.64(stat.) \pm 0.00(sys.)$
$M500c\tau100$	$0.16 \pm 0.13(stat.) \pm 0.00(sys.)$
$M500c\tau1000$	$0.18 \pm 0.14(stat.) \pm 0.00(sys.)$
$M600c\tau1$	$14.01 \pm 0.67(stat.) \pm 0.08(sys.)$
$M600c\tau10$	$1.34 \pm 0.21(stat.) \pm 0.00(sys.)$
$M600c\tau100$	$0.06 \pm 0.05(stat.) \pm 0.00(sys.)$
$M600c\tau1000$	$0.00 \pm 0.00(stat.) \pm 0.00(sys.)$
$M700c\tau1$	$5.37 \pm 0.26(stat.) \pm 0.02(sys.)$
$M700c\tau10$	$0.51 \pm 0.08(stat.) \pm 0.00(sys.)$
$M700c\tau100$	$0.05 \pm 0.02(stat.) \pm 0.00(sys.)$
$M700c\tau1000$	$0.00 \pm 0.00(stat.) \pm 0.00(sys.)$
$M800c\tau1$	$2.12 \pm 0.11(stat.) \pm 0.00(sys.)$
$M800c\tau10$	$0.31 \pm 0.04(stat.) \pm 0.00(sys.)$
$M800c\tau100$	$0.00 \pm 0.00(stat.) \pm 0.00(sys.)$
$M800c\tau1000$	$0.00 \pm 0.00(stat.) \pm 0.00(sys.)$
$M900c\tau1$	$0.94 \pm 0.05(stat.) \pm 0.00(sys.)$
$M900c\tau10$	$0.12 \pm 0.02(stat.) \pm 0.00(sys.)$
$M900c\tau100$	$0.00 \pm 0.00(stat.) \pm 0.00(sys.)$
$M900c\tau1000$	$0.00 \pm 0.00(stat.) \pm 0.00(sys.)$
$M1000c\tau1$	$0.40 \pm 0.02(stat.) \pm 0.00(sys.)$
$M1000c\tau10$	$0.06 \pm 0.01(stat.) \pm 0.00(sys.)$
$M1000c\tau100$	$0.00 \pm 0.00(stat.) \pm 0.00(sys.)$
$M1000c\tau1000$	$0.00 \pm 0.00(stat.) \pm 0.00(sys.)$
$M1100c\tau1$	$0.22 \pm 0.01(stat.) \pm 0.00(sys.)$
$M1100c\tau10$	$0.03 \pm 0.00(stat.) \pm 0.00(sys.)$
$M1100c\tau100$	$0.00 \pm 0.00(stat.) \pm 0.00(sys.)$
$M1100c\tau1000$	$0.00 \pm 0.00(stat.) \pm 0.00(sys.)$
$M1200c\tau1$	$0.10 \pm 0.01(stat.) \pm 0.00(sys.)$
$M1200c\tau10$	$0.01 \pm 0.00(stat.) \pm 0.00(sys.)$
$M1200c\tau100$	$0.00 \pm 0.00(stat.) \pm 0.00(sys.)$
$M1200c\tau1000$	$0.00 \pm 0.00(stat.) \pm 0.00(sys.)$

Table 17: Signal yield in the displaced control region

have high enough signal contamination are the samples with lower masses for the stop quarks. These have already been excluded by other analysis'[11].

8.3 Signal prediction in the Signal region

In tables 18, 19, 20, 21, 22 can be found the estimates for the signal in the various inclusive and exclusive signal regions. When looking at the values for the different signal samples it can be concluded that the samples with lower mass would be found, however these have already been excluded. Samples with the highest mass have an expected number of events lower than the upper limit estimates of the background sources. The analysis will therefore not be sensitive to them. Finally, between the highest and lowest values of the mass, there are several samples which have not been excluded but have a number of expected events higher than the number of events estimated from background processes. This means that even with the conservative QCD background estimate, the analysis is still sensitive to potential signal samples.

9 Results

The resulting estimated number of events arising from background sources in the di-muon channel can be found in tables 23 and 24, for the inclusive and exclusive signal regions respectively.

Due to further analysis on the distribution of the impact parameter in the different regions of the QCD estimation, it was decided not to unblind the data. Therefore, no comparison between data in the signal region and expected number of events can be made and no potential excess in data can be be discovered.

The predicted number of events will however be used the set limits on the hypothesis. Given the fact that this research was performed at nearly ten times the luminosity than previous searches [11]. Therefore the resulting limits are expected to be more strict.

10 Conclusion

Results are presented for the prediction of the background sources in the search for the top squark production with displaced di-muon signature using the sample that was recorded during the 2016 data taking period of the CMS detector.

The Non-QCD background contributions, which include the $Drell-Yan$, $Diboson$, $t\bar{t}$, $W \rightarrow \nu\mu$ and $SingleTop$, were estimated using the muon efficiencies in passing a certain cut on the impact parameter. The results can be seen in table 8. It was found that the contribution are within uncertainty much smaller than one. The contribution from Non-QCD background sources will therefore be negligible.

Due to the limited statistics in the QCD MC samples, the efficiency method would underestimate the amount of QCD events. Therefore several data driven methods are developed in order to estimate the QCD background contribution.

$M200c\tau1$	$18669.26 \pm 438.97(stat.) \pm 361.10(sys.)$
$M200c\tau10$	$70781.70 \pm 860.99(stat.) \pm 1952.83(sys.)$
$M200c\tau100$	$23796.32 \pm 498.74(stat.) \pm 825.34(sys.)$
$M200c\tau1000$	$1069.69 \pm 105.59(stat.) \pm 19.43(sys.)$
$M300c\tau1$	$3080.88 \pm 65.33(stat.) \pm 45.79(sys.)$
$M300c\tau10$	$11266.98 \pm 124.26(stat.) \pm 150.12(sys.)$
$M300c\tau100$	$4590.93 \pm 79.19(stat.) \pm 74.03(sys.)$
$M300c\tau1000$	$193.49 \pm 16.43(stat.) \pm 2.14(sys.)$
$M400c\tau1$	$755.93 \pm 14.97(stat.) \pm 6.09(sys.)$
$M400c\tau10$	$2737.91 \pm 28.42(stat.) \pm 21.35(sys.)$
$M400c\tau100$	$1080.01 \pm 17.96(stat.) \pm 7.73(sys.)$
$M400c\tau1000$	$57.57 \pm 4.12(stat.) \pm 0.54(sys.)$
$M500c\tau1$	$217.51 \pm 4.26(stat.) \pm 1.25(sys.)$
$M500c\tau10$	$789.50 \pm 8.18(stat.) \pm 3.58(sys.)$
$M500c\tau100$	$327.82 \pm 5.25(stat.) \pm 1.42(sys.)$
$M500c\tau1000$	$18.69 \pm 1.25(stat.) \pm 0.15(sys.)$
$M600c\tau1$	$72.45 \pm 1.43(stat.) \pm 0.20(sys.)$
$M600c\tau10$	$270.70 \pm 2.76(stat.) \pm 0.74(sys.)$
$M600c\tau100$	$118.49 \pm 1.82(stat.) \pm 0.24(sys.)$
$M600c\tau1000$	$5.78 \pm 0.40(stat.) \pm 0.00(sys.)$
$M700c\tau1$	$27.48 \pm 0.55(stat.) \pm 0.07(sys.)$
$M700c\tau10$	$102.25 \pm 1.05(stat.) \pm 0.22(sys.)$
$M700c\tau100$	$43.84 \pm 0.69(stat.) \pm 0.10(sys.)$
$M700c\tau1000$	$2.03 \pm 0.15(stat.) \pm 0.01(sys.)$
$M800c\tau1$	$11.79 \pm 0.23(stat.) \pm 0.02(sys.)$
$M800c\tau10$	$45.20 \pm 0.46(stat.) \pm 0.07(sys.)$
$M800c\tau100$	$19.96 \pm 0.30(stat.) \pm 0.02(sys.)$
$M800c\tau1000$	$0.88 \pm 0.06(stat.) \pm 0.00(sys.)$
$M900c\tau1$	$5.22 \pm 0.10(stat.) \pm 0.00(sys.)$
$M900c\tau10$	$20.35 \pm 0.21(stat.) \pm 0.02(sys.)$
$M900c\tau100$	$8.90 \pm 0.14(stat.) \pm 0.01(sys.)$
$M900c\tau1000$	$0.52 \pm 0.03(stat.) \pm 0.00(sys.)$
$M1000c\tau1$	$2.46 \pm 0.05(stat.) \pm 0.00(sys.)$
$M1000c\tau10$	$9.43 \pm 0.10(stat.) \pm 0.01(sys.)$
$M1000c\tau100$	$4.18 \pm 0.06(stat.) \pm 0.01(sys.)$
$M1000c\tau1000$	$0.21 \pm 0.01(stat.) \pm 0.00(sys.)$
$M1100c\tau1$	$1.24 \pm 0.02(stat.) \pm 0.00(sys.)$
$M1100c\tau10$	$4.78 \pm 0.05(stat.) \pm 0.00(sys.)$
$M1100c\tau100$	$2.22 \pm 0.03(stat.) \pm 0.00(sys.)$
$M1100c\tau1000$	$0.13 \pm 0.01(stat.) \pm 0.00(sys.)$
$M1200c\tau1$	$0.64 \pm 0.01(stat.) \pm 0.00(sys.)$
$M1200c\tau10$	$2.50 \pm 0.03(stat.) \pm 0.00(sys.)$
$M1200c\tau100$	$1.14 \pm 0.02(stat.) \pm 0.00(sys.)$
$M1200c\tau1000$	$0.07 \pm 0.00(stat.) \pm 0.00(sys.)$

Table 18: Signal yield in the signal region ($d0 \in [0.02, 2.0]$)

$M200c\tau1$	$11375.09 \pm 342.81(stat.) \pm 217.86(sys.)$
$M200c\tau10$	$63294.14 \pm 813.91(stat.) \pm 1760.60(sys.)$
$M200c\tau100$	$22909.06 \pm 489.41(stat.) \pm 804.57(sys.)$
$M200c\tau1000$	$1069.69 \pm 105.59(stat.) \pm 19.43(sys.)$
$M300c\tau1$	$1841.18 \pm 50.44(stat.) \pm 30.26(sys.)$
$M300c\tau10$	$10111.13 \pm 117.70(stat.) \pm 134.08(sys.)$
$M300c\tau100$	$4426.90 \pm 77.76(stat.) \pm 72.99(sys.)$
$M300c\tau1000$	$189.12 \pm 16.23(stat.) \pm 2.14(sys.)$
$M400c\tau1$	$445.22 \pm 11.48(stat.) \pm 3.73(sys.)$
$M400c\tau10$	$2457.44 \pm 26.92(stat.) \pm 19.22(sys.)$
$M400c\tau100$	$1041.22 \pm 17.63(stat.) \pm 7.40(sys.)$
$M400c\tau1000$	$55.55 \pm 4.05(stat.) \pm 0.54(sys.)$
$M500c\tau1$	$132.82 \pm 3.33(stat.) \pm 0.86(sys.)$
$M500c\tau10$	$703.36 \pm 7.72(stat.) \pm 3.01(sys.)$
$M500c\tau100$	$315.34 \pm 5.15(stat.) \pm 1.32(sys.)$
$M500c\tau1000$	$18.44 \pm 1.24(stat.) \pm 0.15(sys.)$
$M600c\tau1$	$42.75 \pm 1.10(stat.) \pm 0.13(sys.)$
$M600c\tau10$	$240.90 \pm 2.60(stat.) \pm 0.68(sys.)$
$M600c\tau100$	$114.44 \pm 1.79(stat.) \pm 0.22(sys.)$
$M600c\tau1000$	$5.67 \pm 0.40(stat.) \pm 0.00(sys.)$
$M700c\tau1$	$16.09 \pm 0.42(stat.) \pm 0.04(sys.)$
$M700c\tau10$	$91.60 \pm 0.99(stat.) \pm 0.20(sys.)$
$M700c\tau100$	$42.21 \pm 0.67(stat.) \pm 0.10(sys.)$
$M700c\tau1000$	$2.00 \pm 0.15(stat.) \pm 0.00(sys.)$
$M800c\tau1$	$6.85 \pm 0.18(stat.) \pm 0.01(sys.)$
$M800c\tau10$	$40.29 \pm 0.43(stat.) \pm 0.06(sys.)$
$M800c\tau100$	$19.24 \pm 0.30(stat.) \pm 0.02(sys.)$
$M800c\tau1000$	$0.86 \pm 0.06(stat.) \pm 0.00(sys.)$
$M900c\tau1$	$3.08 \pm 0.08(stat.) \pm 0.00(sys.)$
$M900c\tau10$	$18.16 \pm 0.19(stat.) \pm 0.02(sys.)$
$M900c\tau100$	$8.54 \pm 0.13(stat.) \pm 0.01(sys.)$
$M900c\tau1000$	$0.50 \pm 0.03(stat.) \pm 0.00(sys.)$
$M1000c\tau1$	$1.45 \pm 0.04(stat.) \pm 0.00(sys.)$
$M1000c\tau10$	$8.39 \pm 0.09(stat.) \pm 0.01(sys.)$
$M1000c\tau100$	$4.03 \pm 0.06(stat.) \pm 0.01(sys.)$
$M1000c\tau1000$	$0.20 \pm 0.01(stat.) \pm 0.00(sys.)$
$M1100c\tau1$	$0.73 \pm 0.02(stat.) \pm 0.00(sys.)$
$M1100c\tau10$	$4.25 \pm 0.05(stat.) \pm 0.00(sys.)$
$M1100c\tau100$	$2.15 \pm 0.03(stat.) \pm 0.00(sys.)$
$M1100c\tau1000$	$0.13 \pm 0.01(stat.) \pm 0.00(sys.)$
$M1200c\tau1$	$0.37 \pm 0.01(stat.) \pm 0.00(sys.)$
$M1200c\tau10$	$2.21 \pm 0.02(stat.) \pm 0.00(sys.)$
$M1200c\tau100$	$1.10 \pm 0.02(stat.) \pm 0.00(sys.)$
$M1200c\tau1000$	$0.07 \pm 0.00(stat.) \pm 0.00(sys.)$

Table 19: Signal yield in the signal region ($d0 \in [0.05, 2.0]$)

$M200c\tau1$	$3808.81 \pm 198.59(stat.) \pm 68.58(sys.)$
$M200c\tau10$	$48945.99 \pm 715.56(stat.) \pm 1301.40(sys.)$
$M200c\tau100$	$20790.37 \pm 466.32(stat.) \pm 743.35(sys.)$
$M200c\tau1000$	$1009.47 \pm 102.65(stat.) \pm 19.43(sys.)$
$M300c\tau1$	$588.81 \pm 28.52(stat.) \pm 10.29(sys.)$
$M300c\tau10$	$7796.27 \pm 103.34(stat.) \pm 105.05(sys.)$
$M300c\tau100$	$3992.83 \pm 73.84(stat.) \pm 67.49(sys.)$
$M300c\tau1000$	$181.55 \pm 15.91(stat.) \pm 1.71(sys.)$
$M400c\tau1$	$129.67 \pm 6.19(stat.) \pm 1.38(sys.)$
$M400c\tau10$	$1892.56 \pm 23.63(stat.) \pm 14.94(sys.)$
$M400c\tau100$	$946.07 \pm 16.81(stat.) \pm 7.16(sys.)$
$M400c\tau1000$	$51.56 \pm 3.90(stat.) \pm 0.54(sys.)$
$M500c\tau1$	$42.91 \pm 1.89(stat.) \pm 0.26(sys.)$
$M500c\tau10$	$537.78 \pm 6.74(stat.) \pm 2.22(sys.)$
$M500c\tau100$	$286.46 \pm 4.91(stat.) \pm 1.14(sys.)$
$M500c\tau1000$	$16.93 \pm 1.19(stat.) \pm 0.15(sys.)$
$M600c\tau1$	$13.79 \pm 0.62(stat.) \pm 0.05(sys.)$
$M600c\tau10$	$183.98 \pm 2.27(stat.) \pm 0.58(sys.)$
$M600c\tau100$	$104.55 \pm 1.71(stat.) \pm 0.18(sys.)$
$M600c\tau1000$	$5.36 \pm 0.39(stat.) \pm 0.00(sys.)$
$M700c\tau1$	$5.06 \pm 0.23(stat.) \pm 0.01(sys.)$
$M700c\tau10$	$69.34 \pm 0.87(stat.) \pm 0.14(sys.)$
$M700c\tau100$	$38.19 \pm 0.64(stat.) \pm 0.10(sys.)$
$M700c\tau1000$	$1.89 \pm 0.14(stat.) \pm 0.00(sys.)$
$M800c\tau1$	$2.00 \pm 0.10(stat.) \pm 0.01(sys.)$
$M800c\tau10$	$30.32 \pm 0.37(stat.) \pm 0.05(sys.)$
$M800c\tau100$	$17.44 \pm 0.28(stat.) \pm 0.02(sys.)$
$M800c\tau1000$	$0.82 \pm 0.06(stat.) \pm 0.00(sys.)$
$M900c\tau1$	$0.90 \pm 0.04(stat.) \pm 0.00(sys.)$
$M900c\tau10$	$13.73 \pm 0.17(stat.) \pm 0.01(sys.)$
$M900c\tau100$	$7.73 \pm 0.13(stat.) \pm 0.01(sys.)$
$M900c\tau1000$	$0.48 \pm 0.03(stat.) \pm 0.00(sys.)$
$M1000c\tau1$	$0.43 \pm 0.02(stat.) \pm 0.00(sys.)$
$M1000c\tau10$	$6.41 \pm 0.08(stat.) \pm 0.01(sys.)$
$M1000c\tau100$	$3.66 \pm 0.06(stat.) \pm 0.01(sys.)$
$M1000c\tau1000$	$0.19 \pm 0.01(stat.) \pm 0.00(sys.)$
$M1100c\tau1$	$0.24 \pm 0.01(stat.) \pm 0.00(sys.)$
$M1100c\tau10$	$3.23 \pm 0.04(stat.) \pm 0.00(sys.)$
$M1100c\tau100$	$1.95 \pm 0.03(stat.) \pm 0.00(sys.)$
$M1100c\tau1000$	$0.12 \pm 0.01(stat.) \pm 0.00(sys.)$
$M1200c\tau1$	$0.11 \pm 0.01(stat.) \pm 0.00(sys.)$
$M1200c\tau10$	$1.67 \pm 0.02(stat.) \pm 0.00(sys.)$
$M1200c\tau100$	$1.00 \pm 0.02(stat.) \pm 0.00(sys.)$
$M1200c\tau1000$	$0.07 \pm 0.00(stat.) \pm 0.00(sys.)$

Table 20: Signal yield in the signal region ($d0 \in [0.1, 2.0]$)

$M200c\tau1$	$7294.16 \pm 556.88(stat.) \pm 143.71(sys.)$
$M200c\tau10$	$7487.56 \pm 1183.83(stat.) \pm 192.69(sys.)$
$M200c\tau100$	$887.26 \pm 698.56(stat.) \pm 21.84(sys.)$
$M200c\tau1000$	$0.00 \pm 149.32(stat.) \pm 0.00(sys.)$
$M300c\tau1$	$1239.70 \pm 82.51(stat.) \pm 15.54(sys.)$
$M300c\tau10$	$1155.84 \pm 170.99(stat.) \pm 16.03(sys.)$
$M300c\tau100$	$164.02 \pm 110.94(stat.) \pm 1.20(sys.)$
$M300c\tau1000$	$4.36 \pm 23.09(stat.) \pm 0.00(sys.)$
$M400c\tau1$	$310.71 \pm 18.86(stat.) \pm 2.39(sys.)$
$M400c\tau10$	$280.48 \pm 39.10(stat.) \pm 2.18(sys.)$
$M400c\tau100$	$38.79 \pm 25.16(stat.) \pm 0.33(sys.)$
$M400c\tau1000$	$2.02 \pm 5.78(stat.) \pm 0.00(sys.)$
$M500c\tau1$	$84.68 \pm 5.40(stat.) \pm 0.40(sys.)$
$M500c\tau10$	$86.14 \pm 11.23(stat.) \pm 0.58(sys.)$
$M500c\tau100$	$12.48 \pm 7.35(stat.) \pm 0.10(sys.)$
$M500c\tau1000$	$0.25 \pm 1.76(stat.) \pm 0.00(sys.)$
$M600c\tau1$	$29.70 \pm 1.80(stat.) \pm 0.08(sys.)$
$M600c\tau10$	$29.80 \pm 3.79(stat.) \pm 0.06(sys.)$
$M600c\tau100$	$4.05 \pm 2.55(stat.) \pm 0.03(sys.)$
$M600c\tau1000$	$0.12 \pm 0.57(stat.) \pm 0.00(sys.)$
$M700c\tau1$	$11.39 \pm 0.69(stat.) \pm 0.02(sys.)$
$M700c\tau10$	$10.64 \pm 1.45(stat.) \pm 0.02(sys.)$
$M700c\tau100$	$1.62 \pm 0.96(stat.) \pm 0.00(sys.)$
$M700c\tau1000$	$0.03 \pm 0.21(stat.) \pm 0.00(sys.)$
$M800c\tau1$	$4.94 \pm 0.29(stat.) \pm 0.00(sys.)$
$M800c\tau10$	$4.90 \pm 0.63(stat.) \pm 0.00(sys.)$
$M800c\tau100$	$0.72 \pm 0.42(stat.) \pm 0.00(sys.)$
$M800c\tau1000$	$0.02 \pm 0.09(stat.) \pm 0.00(sys.)$
$M900c\tau1$	$2.14 \pm 0.13(stat.) \pm 0.00(sys.)$
$M900c\tau10$	$2.19 \pm 0.28(stat.) \pm 0.00(sys.)$
$M900c\tau100$	$0.36 \pm 0.19(stat.) \pm 0.00(sys.)$
$M900c\tau1000$	$0.02 \pm 0.05(stat.) \pm 0.00(sys.)$
$M1000c\tau1$	$1.01 \pm 0.06(stat.) \pm 0.00(sys.)$
$M1000c\tau10$	$1.04 \pm 0.13(stat.) \pm 0.00(sys.)$
$M1000c\tau100$	$0.15 \pm 0.09(stat.) \pm 0.00(sys.)$
$M1000c\tau1000$	$0.01 \pm 0.02(stat.) \pm 0.00(sys.)$
$M1100c\tau1$	$0.51 \pm 0.03(stat.) \pm 0.00(sys.)$
$M1100c\tau10$	$0.53 \pm 0.07(stat.) \pm 0.00(sys.)$
$M1100c\tau100$	$0.07 \pm 0.05(stat.) \pm 0.00(sys.)$
$M1100c\tau1000$	$0.00 \pm 0.01(stat.) \pm 0.00(sys.)$
$M1200c\tau1$	$0.27 \pm 0.02(stat.) \pm 0.00(sys.)$
$M1200c\tau10$	$0.29 \pm 0.03(stat.) \pm 0.00(sys.)$
$M1200c\tau100$	$0.04 \pm 0.02(stat.) \pm 0.00(sys.)$
$M1200c\tau1000$	$0.00 \pm 0.01(stat.) \pm 0.00(sys.)$

Table 21: Signal yield in the signal region ($d0 \in [0.02, 0.05]$)

$M200c\tau1$	$7566.28 \pm 396.15(stat.) \pm 150.49(sys.)$
$M200c\tau10$	$14348.14 \pm 1082.99(stat.) \pm 462.44(sys.)$
$M200c\tau100$	$2118.69 \pm 675.82(stat.) \pm 61.47(sys.)$
$M200c\tau1000$	$60.22 \pm 147.26(stat.) \pm 0.00(sys.)$
$M300c\tau1$	$1252.36 \pm 57.94(stat.) \pm 20.12(sys.)$
$M300c\tau10$	$2314.86 \pm 156.50(stat.) \pm 29.08(sys.)$
$M300c\tau100$	$434.07 \pm 107.20(stat.) \pm 5.67(sys.)$
$M300c\tau1000$	$7.58 \pm 22.73(stat.) \pm 0.43(sys.)$
$M400c\tau1$	$315.55 \pm 13.04(stat.) \pm 2.38(sys.)$
$M400c\tau10$	$564.88 \pm 35.79(stat.) \pm 4.28(sys.)$
$M400c\tau100$	$95.16 \pm 24.35(stat.) \pm 0.25(sys.)$
$M400c\tau1000$	$3.99 \pm 5.62(stat.) \pm 0.00(sys.)$
$M500c\tau1$	$89.92 \pm 3.83(stat.) \pm 0.62(sys.)$
$M500c\tau10$	$165.58 \pm 10.24(stat.) \pm 0.79(sys.)$
$M500c\tau100$	$28.88 \pm 7.11(stat.) \pm 0.18(sys.)$
$M500c\tau1000$	$1.51 \pm 1.72(stat.) \pm 0.00(sys.)$
$M600c\tau1$	$28.95 \pm 1.26(stat.) \pm 0.08(sys.)$
$M600c\tau10$	$56.92 \pm 3.45(stat.) \pm 0.11(sys.)$
$M600c\tau100$	$9.89 \pm 2.47(stat.) \pm 0.04(sys.)$
$M600c\tau1000$	$0.30 \pm 0.56(stat.) \pm 0.00(sys.)$
$M700c\tau1$	$11.03 \pm 0.48(stat.) \pm 0.04(sys.)$
$M700c\tau10$	$22.26 \pm 1.32(stat.) \pm 0.06(sys.)$
$M700c\tau100$	$4.02 \pm 0.93(stat.) \pm 0.00(sys.)$
$M700c\tau1000$	$0.11 \pm 0.20(stat.) \pm 0.00(sys.)$
$M800c\tau1$	$4.85 \pm 0.20(stat.) \pm 0.01(sys.)$
$M800c\tau10$	$9.97 \pm 0.57(stat.) \pm 0.02(sys.)$
$M800c\tau100$	$1.79 \pm 0.41(stat.) \pm 0.00(sys.)$
$M800c\tau1000$	$0.05 \pm 0.09(stat.) \pm 0.00(sys.)$
$M900c\tau1$	$2.18 \pm 0.09(stat.) \pm 0.00(sys.)$
$M900c\tau10$	$4.44 \pm 0.26(stat.) \pm 0.00(sys.)$
$M900c\tau100$	$0.80 \pm 0.18(stat.) \pm 0.00(sys.)$
$M900c\tau1000$	$0.02 \pm 0.05(stat.) \pm 0.00(sys.)$
$M1000c\tau1$	$1.02 \pm 0.04(stat.) \pm 0.00(sys.)$
$M1000c\tau10$	$1.98 \pm 0.12(stat.) \pm 0.00(sys.)$
$M1000c\tau100$	$0.37 \pm 0.09(stat.) \pm 0.00(sys.)$
$M1000c\tau1000$	$0.01 \pm 0.02(stat.) \pm 0.00(sys.)$
$M1100c\tau1$	$0.49 \pm 0.02(stat.) \pm 0.00(sys.)$
$M1100c\tau10$	$1.02 \pm 0.06(stat.) \pm 0.00(sys.)$
$M1100c\tau100$	$0.20 \pm 0.05(stat.) \pm 0.00(sys.)$
$M1100c\tau1000$	$0.01 \pm 0.01(stat.) \pm 0.00(sys.)$
$M1200c\tau1$	$0.26 \pm 0.01(stat.) \pm 0.00(sys.)$
$M1200c\tau10$	$0.55 \pm 0.03(stat.) \pm 0.00(sys.)$
$M1200c\tau100$	$0.10 \pm 0.02(stat.) \pm 0.00(sys.)$
$M1200c\tau1000$	$0.00 \pm 0.01(stat.) \pm 0.00(sys.)$

Table 22: Signal yield in the signal region ($d0 \in [0.05, 0.01]$)

Inclusive region	$d0 \in [0.02, 2.0]$
$Z \to l^+l^-$	$0.28 \pm 0.03(stat.) \pm 0.15(sys.)$
$Diboson$	$(0.35 \pm 0.14(stat.) \pm 0.16(sys.)) \cdot 10^{-2}$
$t\bar{t}$	$(0.81 \pm 0.07(stat.) \pm 0.37(sys.)) \cdot 10^{-1}$
$SingleTop$	$(0.12 \pm 0.07(stat.) \pm 0.06(sys.)) \cdot 10^{-1}$
$W \to l\nu$	$0.0 \pm 0.0(stat.) \pm 0.0(sys.)$
QCD	$31.59 \pm 7.94(stat.) \pm 6.08(sys.)$
Inclusive region	$d0 \in [0.05, 2.0]$
$Z \to l^+l^-$	$0.11 \pm 0.01(stat.) \pm 0.05(sys.)$
$Diboson$	$(0.71 \pm 0.46(stat.) \pm 0.82(sys.)) \cdot 10^{-5}$
$t\bar{t}$	$(0.36 \pm 0.07(stat.) \pm 0.19(sys.)) \cdot 10^{-2}$
$SingleTop$	$(0.48 \pm 0.49(stat.) \pm 0.28(sys.)) \cdot 10^{-3}$
$W \to l\nu$	$0.0 \pm 0.0(stat.) \pm 0.0(sys.)$
QCD	$22.35 \pm 5.62(stat.) \pm 4.44(sys.)$
Inclusive region	$d0 \in [0.1, 2.0]$
$Z \to l^+l^-$	$(0.23 \pm 0.05(stat.) \pm 0.10(sys.)) \cdot 10^{-1}$
$Diboson$	$(0.32 \pm 0.20(stat.) \pm 0.13(sys.)) \cdot 10^{-6}$
$t\bar{t}$	$(0.26 \pm 0.11(stat.) \pm 0.12(sys.)) \cdot 10^{-3}$
$SingleTop$	$0.0 \pm 0.0(stat.) \pm 0.11 \cdot 10^{-4}(sys.)$
$W \to l\nu$	$0.0 \pm 0.0(stat.) \pm 0.0(sys.)$
QCD	$19.15 \pm 4.81(stat.) \pm 3.85(sys.)$

Table 23: Non-QCD MC background contribution in the inclusive and exclusive signal regions using the efficiency method.

The first method, in which transfer factors were calculated using the BEnriched NonIsolated region, resulting in the results found in table 10. The contribution from QCD events is much higher than the other background sources and QCD will therefore be the most dominant background contribution.

When comparing the shape of the impact parameter distribution between the BEnriched NonIsolated and Isolated region, in order to verify the results, a discrepancy in the shape was found. This discrepancy was further shown as predictions for QCD using the BEnriched Isolated region differed from the previous estimates. These results are shown in table 10.

In order to determine which of the two prediction most accurately described the actual behavior of the QCD background contribution, a third method which used the Isolated Inverted region for the calculation of the scale factors was used. The results of this method can be found in table 14.

Because the analysis looks for an excess in data, the most conservative estimate was used for the prediction of the QCD background contribution. This was the method using the BEnriched Isolated region which can be found in table 12. The decision was made to not unblind the data in order to perform further research into the discrepancies of the QCD impact parameter distributions in order to obtain a possible less conservative result.

Exclusive region	$d0 \in [0.02, 0.05]$
$Z \to l^+l^-$	$0.17 \pm 0.03(stat.) \pm 0.10(sys.)$
$Diboson$	$(0.34 \pm 0.14(stat.) \pm 0.16(sys.)) \cdot 10^{-2}$
$t\bar{t}$	$(0.78 \pm 0.07(stat.) \pm 0.35(sys.)) \cdot 10^{-1}$
$SingleTop$	$(0.11 \pm 0.07(stat.) \pm 0.05(sys.)) \cdot 10^{-1}$
$W \to l\nu$	$0.0 \pm 0.0(stat.) \pm 0.0(sys.)$
QCD	$9.24 \pm 9.72(stat.) \pm 1.71(sys.)$
Exclusive region	$d0 \in [0.05, 0.1]$
$Z \to l^+l^-$	$(0.83 \pm 0.14(stat.) \pm 0.41(sys.)) \cdot 10^{-1}$
$Diboson$	$(0.67 \pm 0.46(stat.) \pm 0.81(sys.)) \cdot 10^{-5}$
$t\bar{t}$	$(0.34 \pm 0.07(stat.) \pm 0.18(sys.)) \cdot 10^{-2}$
$SingleTop$	$(0.48 \pm 0.50(stat.) \pm 0.27(sys.)) \cdot 10^{-3}$
$W \to l\nu$	$0.0 \pm 0.0(stat.) \pm 0.0(sys.)$
QCD	$3.20 \pm 7.39(stat.) \pm 0.60(sys.)$
Exclusive region	$d0 \in [0.1, 2.0]$
$Z \to l^+l^-$	$(0.23 \pm 0.05(stat.) \pm 0.10(sys.)) \cdot 10^{-1}$
$Diboson$	$(0.32 \pm 0.20(stat.) \pm 0.13(sys.)) \cdot 10^{-6}$
$t\bar{t}$	$(0.26 \pm 0.11(stat.) \pm 0.12(sys.)) \cdot 10^{-3}$
$SingleTop$	$0.0 \pm 0.0(stat.) \pm 0.11 \cdot 10^{-4}(sys.)$
$W \to l\nu$	$0.0 \pm 0.0(stat.) \pm 0.0(sys.)$
QCD	$19.15 \pm 4.81(stat.) \pm 3.85(sys.)$

Table 24: Non-QCD MC background contribution in the inclusive and exclusive signal regions using the efficiency method.

References

[1] S. P. Martin, *A Supersymmetry Primer* (v7), 2016, arXiv:hep-ph/9709356

[2] I.J.R. Aitchison, *Supersymmetry and the MSSM: An Elementary Introduction*, 2005, arXiv:hep-ph/0505105

[3] CMS COllaboration *Search for displaced leptons in the e-mu channel*, 2016

[4] P. Graham, D. Kaplan, S. Rajendran, and P. Sarawat, *"Dispaced Supersymmetry"*, Journal if high energy physics 149(2012) arXiv:1204.6038 [hep-ph]

[5] CMS Collaboration, *The CMS tracker system project : Technical Design Report*, 1997, `http://cds.cern.ch/record/368412?ln=en`

[6] CMS Collaboration *The CMS Electromagnetic Calorimeter at the LHC*, 2008, `https://arxiv.org/ftp/arxiv/papers/0810/0810.0381.pdf` arXiv:0810.0381 [physics.ins-det]

[7] CMS Collaboration*HCAL Technical Design Report* , 1997, `https://cds.cern.ch/record/357153/files/CMS_HCAL_TDR.pdf`

[8] CMS Collaboration *The Muon Project Technical Design Report*, 1997, `https://web.physik.rwth-aachen.de/~hebbeker/cms_muon_tdr.pdf`

[9] CMS Collaboration, *Search for Displaced Suprsymmetry in Dilepton Final States at 13 TeV* 2016

[10] CMS Collaboration, *Commissioning of the Particle-Flow Reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV*

[11] Q. Python, *Search for displaced Supersymmetry in events with two leptons with large impact parameters with the CMS detector at the LHC*, 2017, Vrije Universiteit Brussel

[12] CMS detector design, `http://cms.web.cern.ch/news/cms-detector-design`