universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Sensitivity study of the Z' boson in events involving taus in the final state at Belle II"

verfasst von / submitted by

Renata Donica, BSc.

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2020 / Vienna, 2020

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 066 876
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Physik
Betreut von / Supervisor:	Hon.-Prof. Dipl.-Phys. Dr. Eberhard Widmann

Acknowledgements:

Dr. Gianluca Inguglia for providing the idea, means and support for this thesis, **Florin Ignatyuk** for helping in regards to the programming aspects, and **Michel Bertemes** for practical guidance and tutoring.

Contents

1 Introduction **3**

2 Theoretical background **4**
2.1 The standard model 4
2.2 Dark Matter 6
2.2.1 Evidence 6
2.2.2 Particle candidates 8
2.3 The Z' boson in the $L_\mu - L_\tau$ model 9
2.4 Scattering theory 10

3 Tools and Methods **11**
3.1 Belle II 11
3.2 Software 15
3.3 The Z' channels considered 16
3.3.1 Generation 18
3.3.2 Simulation 19
3.3.3 Reconstruction 20

4 Analysis **21**
4.1 Reconstructing two muon systems 23
4.1.1 Pre-processing 27
4.1.2 Detection efficiency 29
4.1.3 The figure-of-merit 31
4.1.4 Z' event cross-section 35
4.2 Reconstructing the four muon system 38
4.3 g' Upper Limits 45
4.4 Other data visualisation types 47

5 Results and Discussion **50**

Abstract

The following study is centered on the hypothetical Z' boson within the L_μ-L_τ model, which may serve as a mediator between the Standard Model and the Dark Sector. The events considered involve a Z' decaying into two muons or two taus, at the Belle II detector. Simulated data sets are studied, exploring different methods to distinguish between the background and signal. For each assumed Z' mass, upper limits for the event cross-section are computed, which can be translated in terms of a new coupling constant g' withing the L_μ-L_τ model. Limits for g' as low as $1.29 \cdot 10^{-4}$ are achieved.

Zusammenfassung

Die folgende Studie konzentriert sich auf das hypothetische Z'-Boson innerhalb des L_μ-L_τ Modells, das als Vermittler zwischen dem Standardmodell und dem Dunklen Sektor dienen kann. Bei den untersuchten Ereignissen zerfällt ein Z' am Belle-II-Detektor in zwei Myonen oder Tauonen. Es werden simulierte Datensätzen berücksichtigt, wobei verschiedene Methoden zur Unterscheidung zwischen Hintergrund und Signal untersucht werden. Für jede angenommene Z' Masse werden Obergrenzen für den Wirkungsquerschnitt berechnet, die sich als neue Kopplungskonstante g' im Rahmen des L_μ-L_τ Modell übersetzen lassen. Es werden Grenzwerte so niedrig wie $1.29 \cdot 10^{-4}$ für g' erreicht.

1 Introduction

The existence of dark matter (DM) is supported by various cosmological studies, from the unusual rotation curve of spiral galaxies [1], to the displaced center of mass in galaxy clusters such as MACS J0025.4-1222 [2]. However, despite the many theoretical models proposed to account for a Dark Sector (DS), the particle nature of dark matter and/or of a dark force is yet to be discovered. The search for new phenomena in particle physics (such as the DS) has been one of the goals of collider experiments over the past decades, including the Belle and Belle II collaborations at KEKB, respectively SuperKEKB.

The SuperKEKB accelerator is attempting to breach instantaneous luminosity, while the Belle II detector provides high precision in the low-GeV phase-space. This offers great insight in the B meson sector [3], and has inspired many promising DM studies. This thesis is focused on the hypothetical $Z'_{L_\mu - L_\tau}$ boson, acting as the mediator between the Standard Model (SM) and the Dark Sector. A new U(1)' symmetry is considered, as an extension to the existing

SM symmetry group SU(3)xSU(2)xU(1). Two decay channels with final states involving Z', μ-s and τ-s are investigated, according to the geometry of the Belle II particle detector. The generation of the events and their interactions with the detector are simulated using Monte Carlo simulation methods.

The overall goal is to find under which selections the Z' data is sufficiently different from background data to be successfully detected by the Belle II. Specifically, the minimal cross-section of the Z' events is computed, such that they would be visible over the background, when searching for them in real data. Since the cross-section σ can be translated into the dark coupling constant g', upper limits for g' are obtained.

The following pages present the different steps I undertook in my analysis under the guidance of my mentors and colleagues, as well as the theory motivating this study and a description of the Belle II facility.

2 Theoretical background

2.1 The standard model

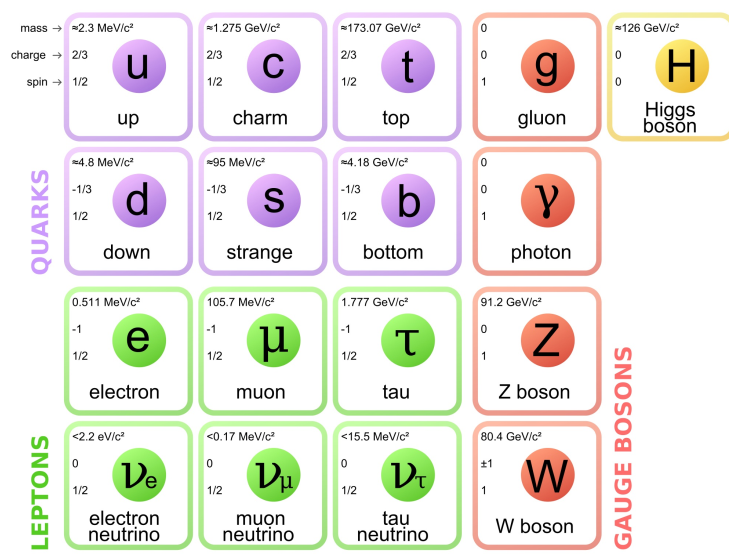


Figure 1: The Standard Model (SM) of particle physics. Figure from [4].

During the second half of the 20^{th} century, the study of elementary particles departed from nuclear and quantum physics into its own, new field. Early collision experiments revealed a much richer variety of particles than initially believed. As higher energies were achieved in experiments, and sub-structures were revealed, it became clear that the particles found could be categorized. Much like the periodic table for atoms, the Standard Model (figure 1) serves as a classifier and predictor for elementary particles. It is the theoretical framework describing three of the four fundamental forces: the strong, weak, and electromagnetic force.

The main distinction made in the SM is between fermions and bosons. Fermions have half-integer spin and anti-symmetric wavefunctions. They obey Fermi-Dirac statistics, including the Pauli exclusion principle which forbids electrons from occupying the exact same state. Elementary fermions are further classified into three generations of quarks and three generations of leptons, each with defining quantum numbers such as charge and spin. They can form more complex particles like mesons and baryons, with mesons having full integer spin and baryons half-integer spin. In a simplified outlook, fermions are the building blocks of matter, with all the stable matter being made out of fermions of the first generations (electrons, and up and down quarks).

On the other hand, bosons are force carriers. They have integer spin, symmetrical wavefuntions and obey the Bose Einstein statistics, which for example allows super-fluidity. The SM is a quantum field theory where interactions are mediated by gauge bosons (also called vector bosons).

Interaction	Mediator	Range [m]	Strength
Strong	gluon	10^{-15}	1
Electromagnetic	photon	∞	1/137
Weak	$Z\&W^{\pm}$	10^{-18}	10^{-6}
Gravity	graviton?	∞	$6{\cdot}10^{-39}$

Table 1: The four fundamental interactions and their respective mediators, range and coupling constants. The graviton is hypothesised, but not yet confirmed.

Table 1 displays the four fundamental forces and the gauge bosons mediating them. The range of the forces is related to the mass of the respective bosons and it generally decreases as mass increases. The Z and W bosons being massive, the range of the weak force is limited. The zero mass of the other bosons results in an infinite range, except in the case of the strong force, where confinement prohibits it. Because the strong force increases with distance, the further two color charged particles are from each other, the more energy is necessary to separate them. The principle of quantum confinement makes them however inseparable: the energy increases until there is enough available to create an additional particle-antiparticle pair.

The coupling constants are an indicator of the relative strength of each interaction, normalized with respect to the strong force. Despite their name, the coupling constants actually vary, depending on the energy of the phenomenon studied. For example, the electromagnetic constant is measured to be 1/137 in atomic studies, but at energies over 200 GeV such as those obtained the Large Electron Positron collider at CERN, it increases to 1/128. These dependency on other variables is why the constants are often referred to as "running coupling constants". It can be argued that as energy increases, all constants tend towards the same value. At energies comparable to the ones present in the early Universe, the values of the constants would coincide. This would imply a single, all encompassing original interaction [5].

Another type of bosons are the scalar ones, such as the Higgs. The Higgs boson can be interpreted as the excitation of the Higgs mass field. The Higgs field has a non-zero vacuum expectation valve in spontaneous symmetry breaking, upon which particles are given mass.

Each particle in the SM has an anti-matter partner with opposite quantum numbers, which figure 1 doesn't display. Studies have also proposed hypothetical fundamental particles such as the graviton, super-symmetric particles, or dark matter candidates, but despite the arguments for their existence, they are seen as beyond standard model (BSM) extensions.

2.2 Dark Matter

2.2.1 Evidence

Despite its overall success and predictability, the SM is arguably an incomplete theory of nature. It fails to fully describe gravity, including the 10^{-33} discrepancy between its coupling constant and the other coupling constants (the hierarchy problem). There are also missing puzzle pieces when it comes to the neutrino oscillations, the magnetic moment of the muon or the dark matter and dark energy content of the Universe.

The existence of DM is overall accepted due to its solid experimental evidence [6]. One of the first proofs comes from the study of galaxy rotation curves [1]. By laws of classical mechanics, it is expected that objects in a spinning system would move slower the further away from the center they are. The angular velocity should decrease as a function of the radius, and the predicted behavior is observed when considering smaller scales like that of our solar system. When studying spiral galaxies such as NGC 3198 in figure 2, the velocity flats out as distance increases. The preferred explanation for this phenomenon is the presence of unseen - or "dark" - extra mass around the edge of the galaxies.

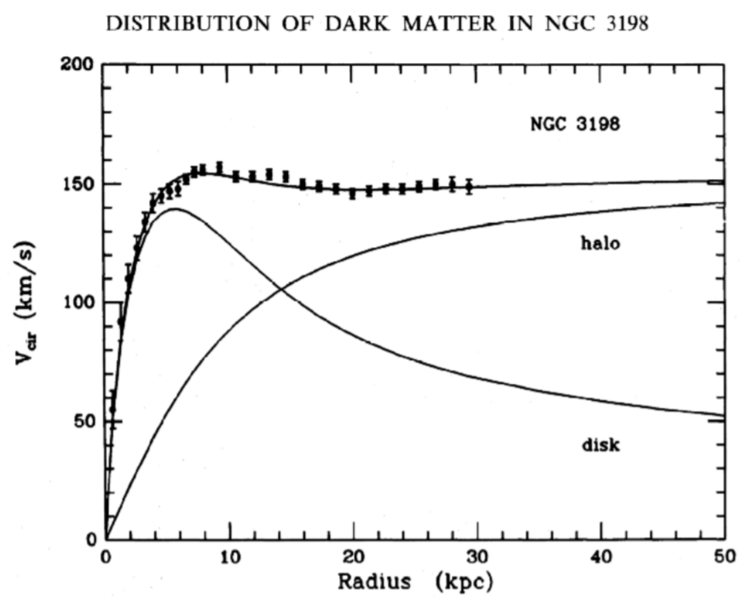


Figure 2: Rotation curve of the NGC 3198 galaxy, expected (lower curve) versus observed (data points). If a halo of DM is added, the discrepancy can be explained. Figure from [1]

Another observation pointing at the existence of DM is the study of galaxy clusters. Their mass can be calculated in various ways (such as by gravitational lensing), and in most of the cases it turns out bigger than theoretically expected. Because DM influences the density and velocity of ordinary matter, it would also affect the microwave background in a way that coincides with experimental data. The study of bullet clusters (i.e. colliding galaxies) also confirms this, since their apparent center of mass is usually far displaced from the expected one.

There are also theoretical arguments for DM, such as the need for density perturbation in the early structure formation of the universe. If there was only ordinary matter available, it couldn't "condense" into heavier objects due to the amount of radiation. Since DM doesn't interact electromagnetically, its density perturbation can grow freely; DM can start the structure formation process, which would then allow ordinary matter to follow due to gravitational attraction, and eventually form the stars and galaxies that make up the universe today [6].

All the astronomical studies mentioned reach the same conclusion: there must be matter in the universe which has mass, but no electric charge[1]. In fact, dark matter constitutes up to 85% of the mass of the universe, and yet the only proofs available to date come from cosmological studies and don't integrate well in other fields. If it were up to particle physics alone, there would be no urgency to fit dark matter into the SM, both theoretically speaking and judging by experimental observations. However, because of the overwhelming evidence for new physics, the search for BSM evidence is crucial. Since the discovery of the Higgs boson, the approach to particle physics has changed from just filling in the empty places of the SM. The main efforts nowadays are to relate the quantum world with gravity, solve the hierarchy problem, and explain DM with the tools available.

2.2.2 Particle candidates

The need to account for new physics such as the dark sector has motivated the further development of the SM to BSM. Some of the requirements a new particle needs to fulfil in order to be a dark candidate is to not interact electromagnetically (have no electric charge) and to interact gravitationally (have non-zero mass). Since

[1] There are also DM skeptical scientists, who try explaining the observations through other means, such as the Modified Newton Dynamics [7].

there are no constraints regarding the other three interactions, the most researched candidate is the weakly interacting massive particle (WIMP) [8], though strongly and gravitationally interacting particles SIMPs [9] and GIMPs [10] were also proposed.

Other possible dark particles could be the axions [11], which were originally hypothesized to solve the strong CP problem in chromodynamics. Even SM neutrinos have been on the list of possible candidates, since the discovery of neutrino oscillation which proves that they do in fact have mass, unlike previously assumed. Sterile neutrinos [12] or their supersymmetric cousins neutralinos are also considered, as well as other supersymmetric particles such as gravitinos.

Besides the fermionic candidates, there are also BSM bosons to consider. Dark gauge bosons would mediate dark forces acting between DM or even ordinary matter. The easiest way to account for dark interactions is to consider them as analogous to the SM ones. This approach gives rise to the dark photon A' [13]. The theory behind the model suggests that dark particles are charge sensitive under A', while being completely neutral when it comes to the SM photon γ.

The Z' boson is born from the same attempt to extend the SM around the already established interactions.

2.3 The Z' boson in the $L_\mu - L_\tau$ model

There are various theoretical models that predict different Z' bosons [14], such as the E6, Little Higgs, or Kaluza-Klein models, but the most simple ones result from the extension consisting of an extra U(1) symmetry group. A gauge symmetry U(1)' is added to the SM in close analogy to the U(1) symmetry defining the standard electromagnetic interaction. Along with this addition comes a new gauge boson Z' as a mediator of the new force. The hypothetical Z' boson is a color-singlet of charge 0, spin 1, and unknown, non-zero mass. Depending on which theoretical branch of the model is followed, Z' can interact with different SM particles [15].

Embedding U(1)' in an abelian gauge group results in a quark-phobic Z' which only couples to some leptons and their corresponding neutrino species. Within the current SM, the differences in family-lepton numbers don't present any anomaly, which means they can be gauged without adding any new fermions [16]. This results in the three sub-models: the $L_e - L_\mu$, $L_e - L_\tau$ and $L_\mu - L_\tau$ model, each predicting a Z' boson interacting with the respective SM leptons.

The Z' analysed in this paper is defined according to the L_μ-L_τ model, meaning it only decays to μs, τs, or invisibly. It does so with equal likelihood, given that is has enough energy. So the Z' decay branching ratios depend on its mass in relation to the mass of the two leptons. Equation (1) depicts this relationship, where the invisible decay products are ν_μ or ν_τ. From an experimental point of view, "invisible" is anything impossible to detect. Even a decay into dark fermions $\chi\overline{\chi}$ is allowed by the model, but the case considered here only assumes invisible decay to tau and muon neutrinos.

$$\begin{gathered} M_{Z'} < 2M_\mu \Rightarrow Br(Z \rightarrow invisible) \approx 1 \\ 2M_\mu < M_{Z'} < 2M_\tau \Rightarrow Br(Z \rightarrow invisible) \approx \frac{1}{2} \\ M_{Z'} > 2M_\tau \Rightarrow Br(Z \rightarrow invisible) \approx \frac{1}{3} \end{gathered} \tag{1}$$

The Lagrangian describing Z' interactions within this model takes the form in (2), where l and $\bar{l}$ are the Dirac spinor vectors for the lepton and the respective antilepton, θ takes only ± 1 values (negative for μ,$\nu_{\mu,L}$ and positive for τ,$\nu_{\tau,L}$) and g' is the U(1)' coupling constant.

$$\mathcal{L} = \sum_l \theta g' \bar{l} \gamma^\mu Z'_\mu l \tag{2}$$

Despite not being the topic of focus in this thesis, it should be mentioned that a loop contribution involving a Z'$_{L_\mu - L_\tau}$ would explain the discrepancy between the predicted and measured magnetic moment of the muon, as well as give rise to a new, right-handed neutrino sector.

2.4 Scattering theory

One of the defining properties of a decay event is its cross-section σ. It is proportional to the probability of the event taking place, and often used interchangeably with likelihood. As its naming suggests, the cross-section is at its basis a geometrical descriptor, and is measured in area units such as barn ($1\text{b} = 10^{-28}\text{m}^2$). In particle interactions, the cross-section is the area in which the particles must "meet" in order for them to scatter on each other.

In accelerator physics, millions of decay events take place in short spans of time, with various cross-sections. In order to quantify and compare the performances of these experiments, the concept of luminosity is used. The luminosity L is defined as the number of events N produced in a certain amount of time t, divided by the total cross-section of the interaction (3). Another helpful quantity which follows in (3) is the integrated luminosity. For full efficiency, the integrated luminosity is simply the ratio between the number of events and their cross-section. Since the detection efficiency ϵ is sub-unitary in practice, formula (4) is used to calculate the cross-section.

$$L = \frac{1}{\sigma}\frac{dN_{prod}}{dt} \Rightarrow \int L dt = \frac{N_{prod}}{\sigma} = \frac{N_{det}}{\sigma \cdot \epsilon} \tag{3}$$

$$\epsilon = \frac{N_{det}}{N_{prod}} \Rightarrow \sigma = \frac{N_{det}}{\epsilon \cdot \int L} \tag{4}$$

3 Tools and Methods

3.1 Belle II

The main fields actively searching for BSM physics are cosmology and experimental particle physics. In the latter case, cross-sections for direct interactions would be too low to measure. There are three scenarios for DM detection, visualized in figure 3. A DM particle χ can either be detected itself (directly), or by detecting its SM decay products P.

Since dark matter particles do not interact with ordinary matter by their defining nature (or interact very weakly)[2], searches are mostly performed indirectly. There is also little constrain on the mass of both DM candidates and dark mediators (such as Z'), so what collider experiments do is probe for new particles within the energy limits of the accelerators.

[2]Dark force carriers such as the Z' do interact with SM particles, but Z' itself is not considered dark matter!

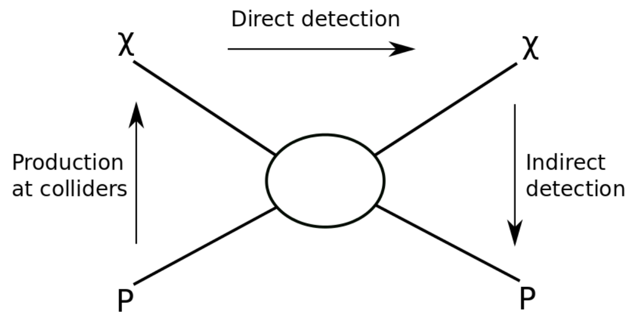


Figure 3: Scheme of possible detection methods for DM. Figure from [17].

Out of the colliding experiments, the so called B-factories allow the measurement of rare event processes with high precision: electron-positron collisions producing B-mesons have led to the measurement of CKM parameters and the observation of CP violation [18]. One of the first B-factories was the KEKB accelerator at KEK (*High Energy Accelerator Research Organisation*), in Tsukuba, Japan [19]. The coresponding particle detector Belle has gathered data from 1999 to 2010, collecting an integrated luminosity of about 710 fb^{-1} at the $\Upsilon(4S)$ resonance mass, before both the accelerator and the detector were upgraded.

The succeeding Belle II detector [3] at SuperKEKB [20] started operating in 2018, with improved conditions and equipment. Between February and July 2018, a commissioning run was performed to test the entire complex, while regular operation started in spring 2019. The design luminosity is planned to reach $8 \cdot 10^{35}\text{cm}^{-2}\text{s}^{-1}$ (40x higher than at KEKB), which makes it promising when searching for low mass dark matter. The main improvement in the luminosity comes from the nano-beam scheme, which consists of narrowing the beta function at the interaction point (IP) and increasing the horizontal crossing angle between the two colliding beams [21].

As its predecessor, Belle II collects data on asymmetric electron positron collisions at the $\Upsilon(4S)$ resonance energy of $\sqrt{s}$= 10.58 GeV. Occasional runs at energies of other Υ resonances are also performed, though not considered in this study. An overview of the SuperKEKB complex is represented in figure 4.

Both the electron and positron are linearly accelerated to their respective energies, circulate along the >3 km long circumference, then

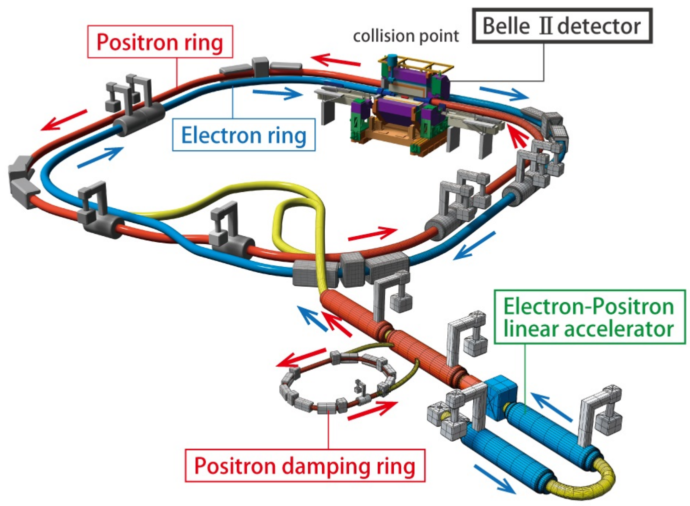


Figure 4: The SuperKEKB accelerator. Figure from [22].

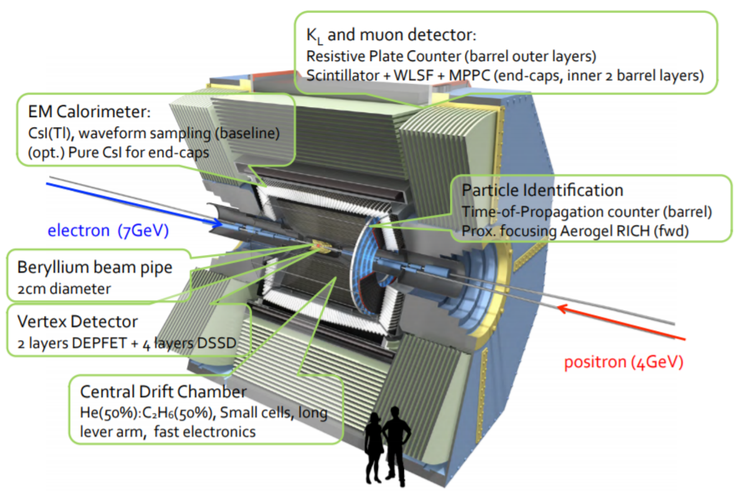


Figure 5: The Belle II detector. Figure from [23].

collided in the center of the Belle II detector. The 7 GeV electrons collide with the 4 GeV positrons, resulting in a center-of-momentum collision energy of $\sqrt{s} = 10.58$ GeV. Because of the asymmetry in the beam energy, the particles resulting from the electron-positron annihilation experience a relativistic Lorentz boost. A boosted decay time/length of the B mesons insures better measurements of CP violation. The Belle II detector was designed as a forward detector with this intend, where 'forward' is defined as the direction of the higher-energetic e^- beam.

As illustrated in figure 5, the Belle II detector has an onion-like shape, consisting of several enclosing layers of cylindrical (and forward) sub-detectors. Each of them serves different tasks in identifying and reconstructing charged and neutral particle candidates [24]. As the produced particles travel from the IP outwards and pass through each detector, pieces of information on them are gathered and eventually put together to fully describe the said particles (within the physical limitations of both the particles and the detectors).

The e^+e^- collision takes place in the Berylium beam pipe they travel in, which is considered along the z axis of the system, with the IP being the (0,0,0) point. The closest detectors to follow are the Pixel Detector (PXD) and the Silicon Vertex Detector (SVD), measuring the decay vertex position and the initial paths of the decay particles produced. They were the last components installed at Belle II, and PXD is planned to be upgraded in 2020 [25]. The SVD consists of four layers of double-sided silicon strip sensors, whereas the PXD has two layers of pixelated silicon sensors.

After the PXD and SDV, the particles travel through a wire chamber: the Central Drift Chamber Detector (CDC). The half helium, half ethane gas filling the chamber is ionized as charged particles pass through, revealing their trajectories. Due to the magnetic field applied, the particle's momentum, trajectory, and dE/dx ratio are measured.

Enclosing the CDC, there is the Time-Of-Propagation (TOP) and Aerogel Ring-Imaging Cherenkov (A-RICH) counters, which measure the timing and pattern of the Cherenkov radiation[3] emitted from charged particles. The TOP covers the barrel region and uses quartz crystals as radiators, whereas the A-RICH is located at the forward end-cap detector to account for the forward nature of the collisions, and it's based on aerogel radiator. Passing through the respective materials, charged particles emit Cherenkov radiation in the shape of a cone. The counters reconstruct the ring-image of said cones, both when it comes to the location and orientation they

[3] Cherenkov radiation is radiation emitted by a particle traveling through a certain medium faster than light travels through the same medium.

occur in, and the precise timing of their production.

Next is the Electromagnetic Calorimeter (ECL), responsible for measuring the energy of photons and electrons passing through. As other detectors, it consists of a barrel part and a forward part, both incorporating thousands of Caesium iodide crystals, which are surrounded by the superconducting solenoid coil producing the magnetic field of 1.5 Tesla [26].

Lastly, the outer most detector is the long living K_L and muon Detector (KLM), responsible for gathering information on the K_L meson and muons, starting with identifying them. Muons are 200 times heavier than electrons and are very penetrative, which is why they're usually not stopped by previous detectors. The KLM computes the muon and hadrons likelihood using the range and transverse scattering of each track [27]. It consists of alternating layers of active detectors and iron plates, the latter serving as the magnetic flux return for the solenoid.

Despite the improvement from the initial Belle, the detector does not cover the full angular region (4π), and the performance of each component varies according to the energy of the particles. Some particles escape undetected through the areas of the detector which are not sensitive, such as the water cooling components, the gas pipes, or other mechanical and electronic "blind spots". Not all of the data can be stored; the amount depending on the trigger systems. Besides this, exterior radiation passing through the detector can be falsely registered as event data, while IP particles such as neutrinos can't be detected at all.

Taking all of this into consideration, each particle parameter detected is stored. Default thresholds are typically applied to validate the data and account for possible misclassification, but each study can require specific conditions to filter its results.

3.2 Software

In the era of computing prowess, working with simulated data is largely preferred before moving the focus of the study to real data, all the more so when it comes to BSM events that might not exist. Not only the physical components of Belle had to be upgraded and replaced to match the high goals of the next generation, but the software as well. For Belle II, most data simulation and processing is carried by BASF2 (Belle Analysis Software 2).

The BASF2 documentation and files are available to members of the collaboration, including examples of code written in both Phython 3 and C++ (root), as well as tutorials and a data base of particles and detector parameters. The framework is structured in modules, each independent and already defined, which the users can arrange in so called paths, or modify depending on the desired outcome. With all the resources offered at reach, one can personalize their own studies.

When it comes to new physics studies, MadGraph is a particularly useful framework, since it is not limited by the predefined, standard conditions of other BASF2 generators (such as EVTGEN), allowing the generation of new particle processes.

The recipe for a typical analysis study consists of the following steps:

- Generating particles, according to a desired physics model;
- Simulating the effects of the multiple detector layers, obliging to their real life limitations;
- Collecting the resulting data, which for all intents and purposes is identical to the one produced in the real detector;
- Reconstructing the original event data from the available results;
- Applying various selection criteria to optimize the study and to reject possible backgrounds.

More details into the development of each step is given in the next chapters, following the Z' decay channels considered in this paper.

3.3 The Z' channels considered

For this study, two decay channels involving Z' are investigated. Their Feynman diagrams are displayed in figure 6 and 7. Both cases consider the electron positron annihilation resulting either in two taus or two muons. The Z' boson is radiated off by tau or muon, then promptly decays to two muons, respectively two taus. Figures figure 6 and 7 depict the intermediate production of a virtual photon

$\gamma*$, but the process can occur via a Higgs or Z boson as well. Given the center of mass energy of Belle II, the contributions of these additional diagrams (with Higgs and Z) is negligible.

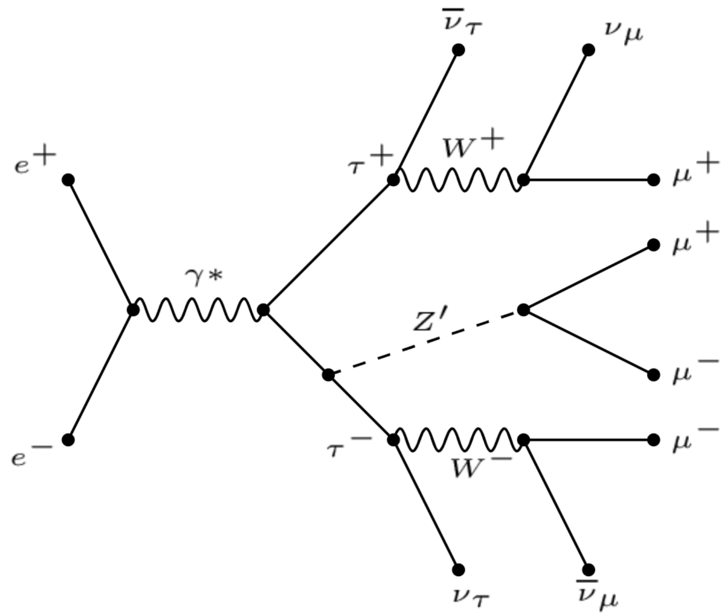


Figure 6: $e^-e^+ \to \tau^+\tau^-Z'[\to \mu^+\mu^-]$

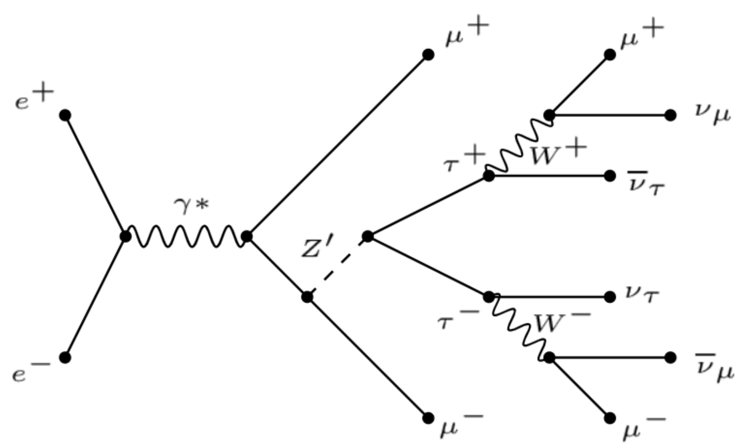


Figure 7: $e^-e^+ \to \mu^+\mu^-Z'[\to \tau^+\tau^-]$

The scenario considered here assumes all the taus involved to only decay to muons and neutrinos, which happens with a branching ratio of $\Gamma_{\tau\to\mu\mu\nu} = (17.39 \pm 0.04)\%$ [28]. This means that the final state in both cases consists of four muons and missing energy from the neutrinos: the two events look identical as far as the final data is concerned.

3.3.1 Generation

Because the processes involve a non standard particle that needs to be defined internally, the generation of the events involving Z' is executed with MadGraph (MG) [29]. The $L_\mu - L_\tau$ model is already defined within MG, so importing it is all it takes for the Z' particle to have meaning and find its place in the theoretical framework the decay operates under.

A couple of parameters must be set, such as the mass of the Z' boson, the coupling constant g' and the width of the decay. Each channel has kinematic constrains depending on the initial available energy of $\sqrt{s}$ = 10.58 GeV, and the invariant mass of the muon and tau, M_μ=0.105658 GeV/c^2, respectively M_τ=1.77686 GeV/c^2. Considering this, the following mass points are generated, in steps of 0.5 GeV/c^2:

- $e^-e^+ \to \tau^+\tau^-Z'[\to \mu^+\mu^-] \Rightarrow 2M_\mu < M_{Z'} < 10.58 - 2m_\tau \Rightarrow 14$ mass points, between 0.5 and 7 GeV (Z' $\to \mu^+\mu^-$ case);
- $e^-e^+ \to \mu^+\mu^-Z'[\to \tau^+\tau^-] \Rightarrow 2M_\tau < M_{Z'} < 10.58 - 2m_\mu \Rightarrow 13$ mass points, between 4 and 10 GeV (Z'$\to \tau^+\tau^-$ case).

For each mass point, 10^4 Z' events are generated.

The coupling constant g' needs to be positive and subunitary. A g'>1 would imply a dark interaction stronger than the strong force and this scenario is to be avoided. Other than this constraint, g' is initially given an arbitrary value. The events here are generated using g'=0.01. One of the later goals of the study is setting boundaries for the coupling constant.

The width $W_{Z'}$ of the decay is set on *auto* at this stage, since MG computes it according to the model. The value of the width doesn't influence the data generated, but it does affect the value of the event cross-section. For general data analysis, this is not a concern. Later, in order to find the precise values for the cross-sections of the dark events, formula (5) is used to compute the width for each Z' mass and g' constant used. The masses of the Z' daughters M_μ or M_τ are used for each of the two cases.

$$W_{Z'} = \frac{M_{Z'} \cdot g'^2}{12\pi} \cdot (1 + \frac{2M^2_{\mu/\tau}}{M^2_{Z'}}) \cdot \sqrt{1 - 4 \cdot \frac{M^2_{\mu/\tau}}{M^2_{Z'}}} \tag{5}$$

The generated Z' signal needs to be compared against a background (BG). BG arises from all e^+e^- decays resulting in the same final state of four muons. As established, the final state looks the same for both Z' cases, so there is only one type of BG that needs to be considered. Because the purpose is to distinguish the dark signal from the abundance of similar SM events, the BG is generated within the SM framework. When declaring the $e^-e^+ \to \mu^+\mu^-\mu^+\mu^-$ decay, the electron and positron can result in the four muons by any intermediate process possible. The *AAFH Generator* from BASF2 is used to generate one million BG events. They do not include the so called beam background, which means that 100% of the generated BG data comes form the colliding particles and their decay products.

The total cross-section for the $e^-e^+ \to \mu^+\mu^-\mu^+\mu^-$ events is given in (6), computed by the *AAFH Generator*. Considering this and using formula (3), the value for the integrated luminosity corresponding to the $N = 10^6$ BG events generated is obtained. At its full capacity, the Belle II detector should collect this luminosity in around six weeks.

$$\sigma_{e^-e^+\to\mu^+\mu^-\mu^+\mu^-} = (342 \pm 0.0285) \text{ fb} \overset{(3)}{\Longrightarrow} \int L = 2.9234 \text{ ab}^{-1} \tag{6}$$

All generated data is eventually stored in root files and can be accessed by other programs in the steps which follow. They contain information on each parameter of each particle produced, such as the energy or momentum. These values are attributed according to the statistics of the decay.

3.3.2 Simulation

Once the decay events are created (according to the theory and models defined by each Generator), the next step is simulating the Belle II detector. The decay products start in the IP, then their trajectories are determined by the generated variables attributed to each particle. The passing of the particles through each detector layer is simulated, and the result of their interactions saved. As mentioned, the outcome data is at this point indistinguishable from real data (provided that statistical enough events are considered). If the Z' boson would exist precisely under the conditions assumed, its decay would be as well detected by the simulation as it would by the detector.

3.3.3 Reconstruction

The reconstruction phase is using the available data to gather information about the original event - a step which has to be done in both simulation and real data studies. The advantage of working with simulated events is that all the decay products and their properties are known before the detection. In a real data study, the only information available is the reconstructed one, whereas an MC study also stores the generated "actual" information on each particle. These MC properties are never discarded and can be used as a comparison for the reconstructed properties. Because it is known where the Z' peaks are expected, the decay can be reconstructed in the most efficient and informative way.

There is a long list of variables which can be reconstructed, from the basic kinematics of the decay, to information about the vertex, tracking, flavor tagging, trigger or particle identification. As stated above, the original generation-level variables can also be accessed, though this is not recommended since they can not be used for real data reconstruction. In the two Z' cases here, a lot of the initial energy is lost to neutrinos, and the dark boson itself is also undetectable. However, the information available on the four muons can be used to trace it back to Z'. The particles detected can be combined, and the variables of the resulted particle system calculated. Considering this, data on the Z' particle can be reconstructed in both cases:

- $e^-e^+ \to \tau^+\tau^- Z'[\to \mu^+\mu^-] \Rightarrow$ Z' mass peaks at the invariant mass $M_{\mu^+\mu^-}$ of the $\mu^+\mu^-$ system;
- $e^-e^+ \to \mu^+\mu^- Z'[\to \tau^+\tau^-] \Rightarrow$ Z' mass peaks at the recoil mass $M^{rec}_{against\mu^+\mu^-}$ of the $\mu^+\mu^-$ system.

Both the invariant and the recoil mass can be stored directly by declaring the $\mu^+\mu^-$ system as a particle and saving the already defined *InvM* and *mRecoil* variables. Alternatively, they be calculated according to formula (7) and (8), given information on the energy E and momentum $\vec{p}$ of each muon. As a reminder, $\sqrt{s} = 10.58$ GeV is the center-of-momentum collision energy, and s is its square.

$$M_{\mu^+\mu^-} = \sqrt{2(M_\mu^2 + E_{\mu^+} \cdot E_{\mu^-} - \vec{p_{\mu^+}} \cdot \vec{p_{\mu^-}})} \tag{7}$$

$$M^{\text{rec}}_{\text{against}\mu^+\mu^-} = \sqrt{s + M^2_{\mu^+\mu^-} - 2\sqrt{s} \cdot E^{\text{CMS}}_{\mu^+\mu^-}} \tag{8}$$

The recoil mass is defined as the invariant mass of the particle(s) recoiling against the investigated particle(s). In (8), E^{CMS} is the center of mass energy of the $\mu^+\mu^-$ system, which is the sum of the center of mass energies of the individual muons. As for the invariant mass in (7), it results from equating the square of the four-momenta before and after the $\mu^+\mu^-$ mother particle decayed to its daughters μ^+ and μ^-. Knowing the energy and momenta of the daughter particles, the mother can be reconstructed (and the other way around) as per equation (9). All throughout, natural units are used, such that the reduced Planck constant $\hbar$ and the speed of light in vacuum c are equal to 1.

$$E^2 = M^2 + p^2 \tag{9}$$

Reconstructing the Z' candidates is therefore achievable. The only aspect complicating the process is that the taus decay to muons as well. The charge of the muons is detectable, but their precise origin isn't: the detector cannot differentiate between the muons coming from the Z' boson and the ones coming from the taus. Considering there are four muons available, and that they need to be paired two by two, there is four unique combinations or $\mu^+\mu^-$, which result in the additional "combinatorial" background.

4 Analysis

Belle II collision energy is 10.58 GeV, so only dark events hypothetically fitting this region can be considered. The construction of the detectors also imposes limitations on measurements, because of the equipment's sensitivity range. The focus of the search is further narrowed to fit within the constrains of both the theoretical model considered and of the decay of interest. Even after limiting the study to this search window, there is still an abundance of decay events besides the ones of interest: the background.

The events of interest (here, the Z' signal) must be separated as much as possible from the rest. A full separation is hardly achievable, but the background can be suppressed. The data can be reconstructed with certain constrains, or the constrains applied accordingly later on - personally or by the help of artificial intelligence. The most efficient way to do so is by "cutting" on the variables stored, i.e. applying selection criteria ("cuts") based on kinematic or angular variables reconstructed as properties of the candidates. The cuts can be used on individual variables or on (usually linear) combination of variables.

An overview of the reconstructed variables and their column statistics is given in figure 8. The data exemplified belongs to the Z'$\rightarrow \mu^+\mu^-$ case, and the values are describing the $\mu^+\mu^-$ system, but the variables presented are reconstructed in all the other decays and methods.

	__event__	InvM	mRecoil	E	px	py	pz	p	theta
count	402235.000000	402235.000000	402235.000000	402235.000000	402235.000000	402235.000000	402235.000000	402235.000000	402235.000000
mean	-7.618017	2.494214	6.713324	3.587607	0.154851	-0.006185	0.982125	2.294398	1.135552
std	3.974533	1.646230	1.688931	1.585163	1.352299	1.353269	1.343081	1.090059	0.613134
min	-14.000000	0.211453	-6.267357	0.261832	-4.431681	-4.677401	-3.397584	0.004399	0.002262
25%	-11.000000	1.237284	5.400372	2.358064	-0.711923	-0.865407	0.032756	1.444526	0.659171
50%	-8.000000	2.063395	6.721047	3.500464	0.162294	-0.005426	0.870738	2.190623	1.062860
75%	-4.000000	3.334733	8.110357	4.613932	1.006709	0.851059	1.896618	3.060018	1.548547
max	-1.000000	8.621480	10.303317	9.041639	4.727587	4.673973	6.684905	7.267556	3.139461

Figure 8: Examples of variables stored in the signal data frame, by reconstructing the $\mu^+\mu^-$ system in the Z'$\rightarrow \mu^+\mu^-$ case.

The *event* column is a classifier, taking values from -14 to -1 to represent each Z' mass, in case they need to be studied separately. Both the invariant mass and the recoil mass are reconstructed, even if only one of them is useful for each Z' case (in figure 8, it's the invariant mass which is related to the mass of the Z'). The energy E and the absolute value of three-momentum p are also stored, as well as each component of the $\vec{p}$ vector. *Theta* as a variable is the polar angle of the reconstructed particle, measured in radians. So in figure 8, *theta* doesn't refer to the angle between the two muons forming the $\mu^+\mu^-$ system, but at the polar angle of the said system, that is the angle between the sum of momenta vectors of the two muons and the z-axis.

Besides applying cuts on the stored variables, it is also crucial to apply selection conditions at a reconstruction level. Here, the only reconstruction level cut used is regarding the production vertex and the muon selection criteria. As explained previously, during the simulation/detection phase, each muon is stored with various parameters. These parameters need to fit certain criteria in order for the muon to be selected in the reconstruction phase. The criteria used in this paper is given in equation (10).

$$|dr| < 0.5 \text{ cm}, |dz| < 2 \text{ cm and } \mu_{\text{ID}} > 0.5 \tag{10}$$

The first two variables in equation (10) are the default vertex production criteria, indicating the boundaries within which the selected muons must have been produced. All the spacial coordinates used are cylindrical: *dr* is the distance to the IP on the *xOy* plane, and *dz* on the z-axis. Muon ID uses different sub-detector information to compute the likelihood of identification (ID), representing how likely it is that the particle detected is indeed a muon.

4.1 Reconstructing two muon systems

Single muon candidates are reconstructed, then combined. As already clarified, the Z'$\rightarrow \mu^+\mu^-$ and Z'$\rightarrow \tau^+\tau^-$ data sets are obtained from pairs on two muons. Outside of the wrong combinations, the variables describing the $\mu^+\mu^-$ system correspond to the Z' boson itself.

Figures 9 and 10 display the histograms of the Z' mass, deduced from the invariant mass of the $\mu^+\mu^-$ system (Z'$\rightarrow \mu^+\mu^-$ case), respectively from the recoil mass of the $\mu^+\mu^-$ system (Z'$\rightarrow \tau^+\tau^-$ case). Out of the events registered, only the peaks correspond to the Z' boson (for each generated mass). In figure 9, it is noticeable that the events counted from the wrong combinations bundle up separately from the signal, as the Z' mass increases. From a mass higher than 5 GeV/c^2, the peaks are clearly distinguishable from the extra combinations (black, blue, red and purple histograms).

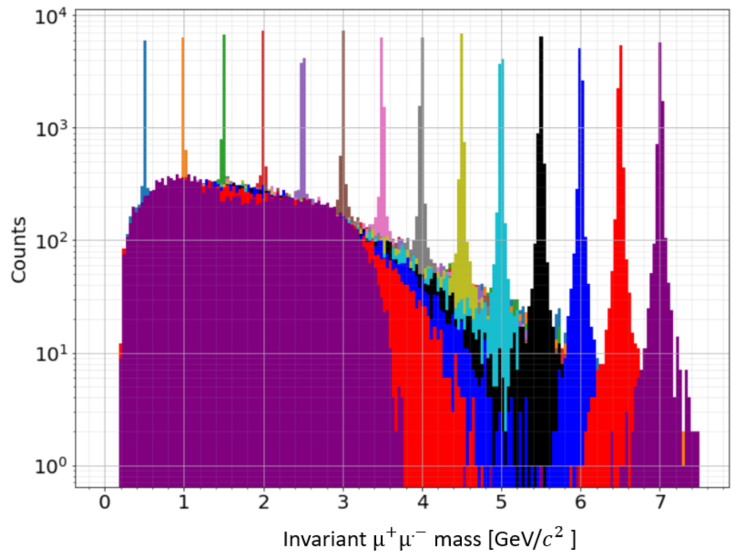


Figure 9: Z'$\rightarrow$ $\mu^+\mu^-$ peaks from $\mu^+\mu^-$ invariant mass (logarithmic scale). Peaks correspond to the 14 mass points generated, the extra counts come from the combiantoric background.

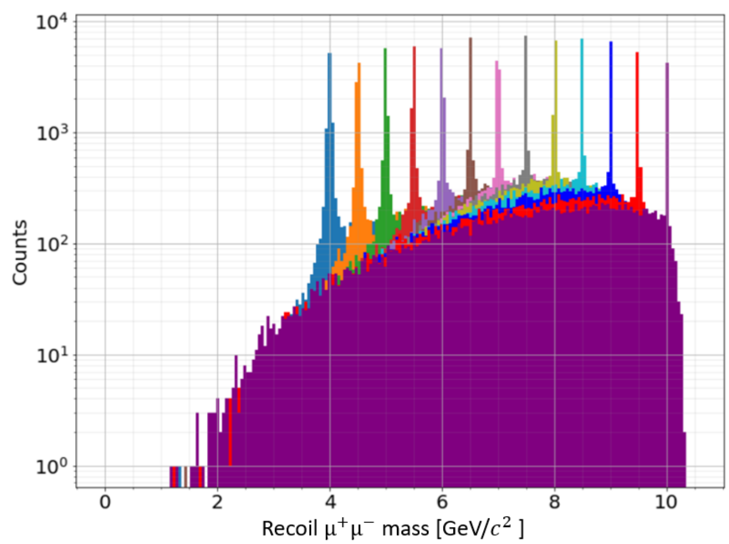


Figure 10: Z'$\rightarrow$ $\tau^+\tau^-$ peaks from $\mu^+\mu^-$ recoil mass (logarithmic scale). Peaks correspond to the 14 mass points generated, the extra counts come from the combiantoric background.

If the muons coming from the Z' boson take most of the 10.58 GeV/c^2 collision energy, as it is the case for M$_{Z'}$=7 GeV/c^2 in figure 9, then the taus are left with virtually no momentum. This implies that muons resulting from the taus could not have an invariant mass (much) higher than the mass of the two taus, which is 3.55372 GeV/c^2. The lower the Z' mass, the more momentum is left available for the taus, implicitly the muons.

In the Z'$\rightarrow$ $\tau^+\tau^-$ case (figure 10), the separation between the Z' signal and the false muon combinations is not visually clear for any of the mass peaks. What is always recoiling against two muons is another two muons plus the four neutrinos. The neutrinos have no mass, but they contributes to the wide distribution of the combinatorial background by taking away energy, regardless of the Z' mass.

The changing shape of the mass histograms is better visualized in figure 11 and 12, which compare only the lowest and highest Z' masses generated. As the Z' mass increases, the combinatorial background distribution transitions smoothly between the two extremes. As mentioned, for the Z' $\rightarrow$ $\mu^+\mu^-$ in figure 11, the combinatorial background is significantly affected by the Z' mass. This is not the case for Z' $\rightarrow$ $\tau^+\tau^-$ in 12, where shape of the bump coming from the wrong combinations stays the same.

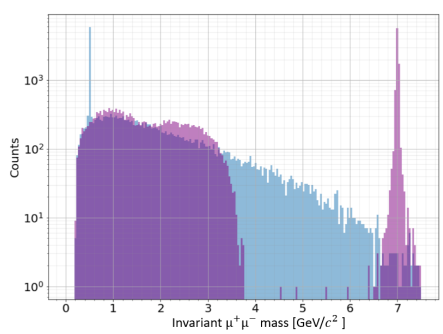


Figure 11: Z'$\rightarrow$ $\mu^+\mu^-$ peaks from $\mu\mu$ invariant mass, for $M_{Z'}$=0.5 GeV/c^2 (blue) and $M_{Z'}$=7 GeV/c^2 (purple).

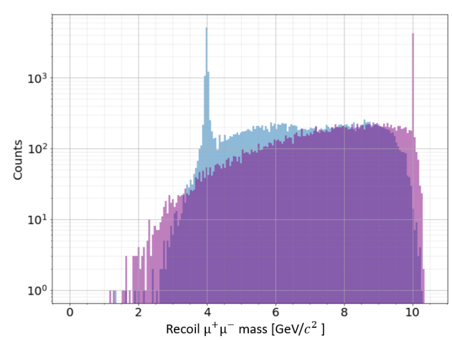


Figure 12: Z'$\rightarrow$ $\tau^+\tau^-$ peaks from $\mu\mu$ recoil mass, for $M_{Z'}$=4 GeV/c^2 (blue) and $M_{Z'}$=10 GeV/c^2 (purple).

Figure 11 and 12 also hint that the closer the Z' mass gets to the kinematic margin of the $\tau^+\tau^-$ production, the wider the base of the peak is, so the larger the decay width. This effect appears exaggerated because of the logarithmic scale used, but it is nevertheless present. Low energy muons are detected poorly and with less likelihood, due to modest interaction with the detectors. So the low E peak, respectively the high E recoil peak are affected the most. The two Z' decay channels display many mirrored properties like this, though the presence of the neutrinos can make it harder to see. If there were no invisible decay products involved, the recoil mass of one pair of muons would be precisely the invariant mass of the other pair. The fact that the standard deviation tends to increase with mass for Z' $\rightarrow \mu^+\mu^-$ and decrease as Z' mass increases for the other case can also be observed by fitting each Z' mass peak gaussianly. The means and standard deviations for some of the peaks are noted in table 2.

Decay	$M_{Z'}^{gen}$ [GeV/c^2]	Mean [GeV/c^2]	Std [GeV/c^2]
Z' $\rightarrow \mu^+\mu^-$	0.5	0.50003	0.00865
	3.5	3.49969	0.03621
	7	7.00114	0.05769
Z' $\rightarrow \tau^+\tau^-$	4	3.99611	0.13405
	7	6.99977	0.04188
	10	10.00002	0.01753

Table 2: The mean Z' masses detected and their standard deviation.

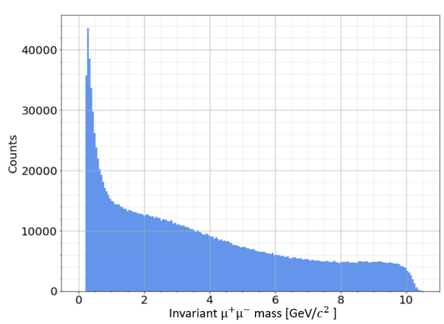


Figure 13: The invariant mass of two muons in the $e^-e^+ \rightarrow \mu^+\mu^-\mu^+\mu^-$ background.

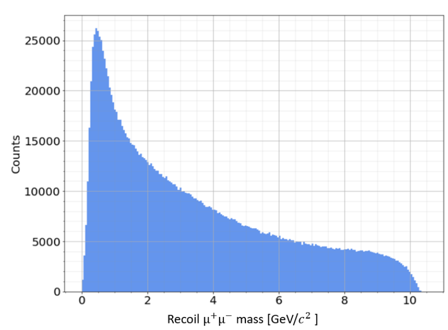


Figure 14: The recoil mass of two muons in the $e^-e^+ \rightarrow \mu^+\mu^-\mu^+\mu^-$ background.

When plotting the invariant mass or the recoil mass of the $\mu^+\mu-$ system constructed in from the background BG data, no significant pattern is present, as expected (figure 13 and figure 14). The majority of the counts are for $\mu^+\mu-$ masses bellow 1 GeV/c^2, but otherwise no peak is present.

4.1.1 Pre-processing

There is a way to separate the $\mu^+\mu^-$ pairs coming only from Z' and the ones resulting partially or entirely from the tau decays. To arrive to this conclusion, variables on each individual muon involved in the formation of the $\mu^+\mu^-$ system are reconstructed, besides the data on the $\mu^+\mu^-$ system itself. The already presented variables are stored, as well as the Monte Carlo production vertices of each particle.

Looking into the *mcProdVertex* variable of each muon, it's noticeable that the Z' daughter particles only have 0 Monte Carlo (MC) decay vertices. This means that according to the simulation, the muons coming from the Z' are produced in the interaction point - more presicely, closer to the IP than the sensitivity of the vertex detectors. By plotting the signal data which satisfies the condition that *mcProdVertex = 0*, the peaks from figures 9 and 10 are singled out. This is valid for both Z'$\rightarrow$ $\mu^+\mu^-$ and Z'$\rightarrow$ $\tau^+\tau^-$ because of the short life time of the Z' boson and its immediate intervention in the e^+e^- decay process. Based on this information, the signal data frames are separated into two: a "clean" Z' signal data, and the remaining combinatorial background.

Although registered as part of the signal data, the wrong combinations do not actually represent the Z' boson. They must not be cut away entirely, but they can be moved to the general background data BG. This adds an error to further calculations, which is addressed in chapter 4.2. The BG should only contain SM decays, but the background events coming from the BSM case studied are an exception: in real studies, the Z' peaks would be searched over all the other counts registered, including its own background. Considering that Z' needs to be identified, the total background can include everything that is not a Z' peak, so the merging can be justified.

Figure 15 and 16 is how the complete data looks like, after applying the vertex cut. To avoid confusion, the "new" background, formed by concatenating the initial BG and the combinatorial background is referred to as BKG.

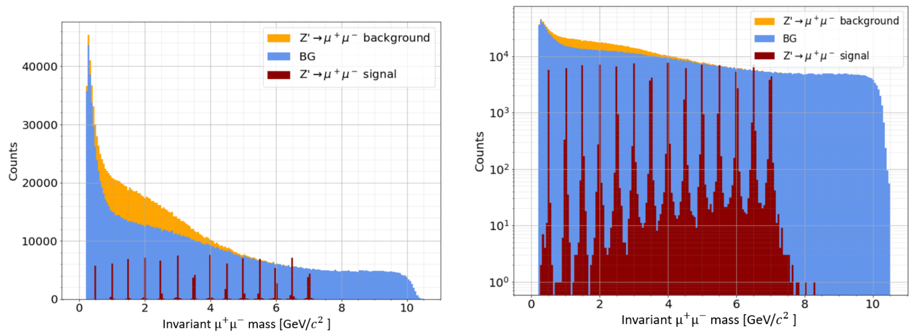


Figure 15: The invariant mass of the $\mu^+\mu-$ system, for Z'$\rightarrow$ $\mu^+\mu^-$ + BG, in logarithmic scale on the right. The combinatorial background histogram (orange) is stacked above the SM background histogram (blue). The two backgrounds are treated as indistinguishable.

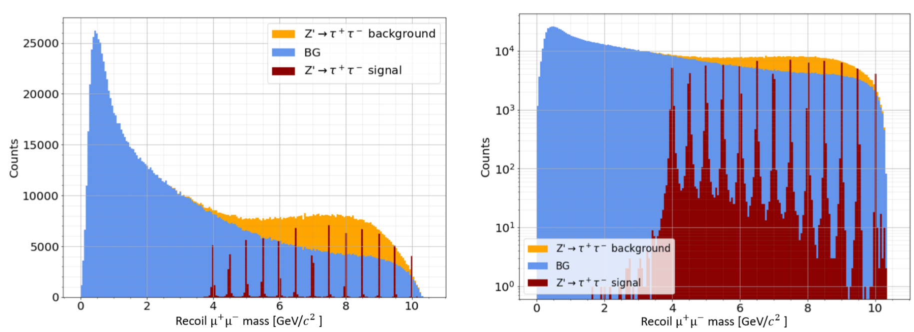


Figure 16: The recoil mass of the $\mu^+\mu-$ system for Z'$\rightarrow$ $\tau^+\tau^-$ + BG, in logarithmic scale on the right. The combinatorial background histogram (orange) is stacked above the SM background histogram (blue). The two backgrounds are treated as indistinguishable.

It should be reminded that the MC prefix in *mcProdVertex* refers to the decay vertices the particles were generated with, not the ones they were reconstructed with. In this study, none of the muons can be traced back at a reconstruction level; the only values available are the MC ones, which are stored when the decays are generated. As previously mentioned, generation level variables cannot be used in real data studies. The wrong combinations can't be simply deleted because singling them out in a real data study wouldn't be possible in this way. However, they can be redistributed to better resemble real life scenarios, as was done when joining the combinatorial background to the SM one.

Furthermore, a clean signal allows for a better understanding of the data and is necessary for computing variables such as the efficiency of the detection. The MC vertex aside, having access to information on each muon offers more opportunities when it comes to applying cuts. Generally, the more variables are available for both the BG and the signal, the easiest it is to compare the two data sets.

Another observation to make is that some of the taus could have been generated with 0 decay vertex and their daughter muons interpreted as pure Z' signal. That is, some of the events composing the red histograms in 15 and 16 could in reality belong to the orange histograms. Even so, the signal data is significantly improved using this separation.

4.1.2 Detection efficiency

Before the signal is separated depending on the MC production vertex, a total of 402235 events are reconstructed for the 14 Z' masses generated, for Z' decaying to muons (figure 9, data from table in figure 8). It results that around 28731.07 events were reconstructed per Z' mass, while only 10000 were generated. This is due to the facts that:

1. Around 25% of the events actually come from Z' and not the wrong $\mu^+\mu^-$ combinations;
2. Not all events are detected.

The detection efficiency for any process considered is defined in equation (11) as the ratio between the number of generated events and the number of reconstructed events. Clearly, formula (11) can be used only with simulated data, while in real data analysis theoretical information of the decay statistics is necessary, such as the branching ratio and the cross-sections of the events studied. When investigating the dark sector, the cross-sections are often unknown, which is another reason why simulation studies are performed.

$$\epsilon = \frac{N_{\text{reconstructed}}}{N_{\text{generated}}} \tag{11}$$

At a first glance, for around 28731/4 events reconstructed per 10000 Z' generated, an average detection efficiently of 71.8% is deduced, for Z'$\rightarrow \mu^+\mu^-$. The reason the values are approximated loosely is because the logic used is not entirely correct. As it is revealed once the pure Z' signal is obtained, the total events registered are not distributed evenly for each Z' mass.

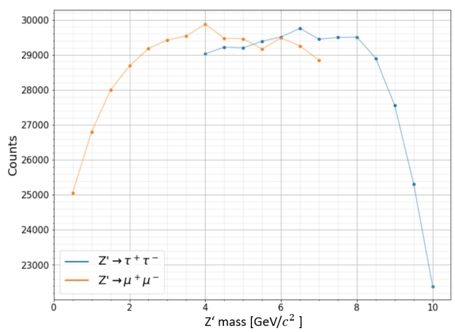


Figure 17: Total number of events registered for each Z' mass generated, including the wrong combinations.

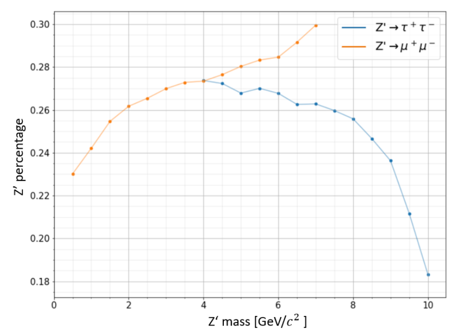


Figure 18: Percentage of the events actually coming from Z' and not the wrong combinations.

Figures 17 and 18 illustrate how the detection process depends on the Z' mass. For both decay cases considered, the amount of signal events detected (figure 17), as well as the percentage of those which actually originate from Z' (figure 18) vary depending on the mass of Z'. This is especially clear as the Z' mass approaches the energy limits that allow the pair of taus to be created, that is 0.5 GeV/c^2 for Z'$\rightarrow \mu^+\mu^-$ and 10 GeV/c^2 for Z'$\rightarrow \tau^+\tau^-$. This is another indicator that lower energy muons escape undetected in higher rates, which leads to the low detection of their invariant mass and the high detection of their recoil mass.

The detection efficiency is computed for each mass and plotted in figure 19, with the correct average of 77.88% for Z'$\rightarrow \mu^+\mu^-$ and 71.85% for Z'$\rightarrow \tau^+\tau^-$, both being higher than initially anticipated. The best results occur when the collision energy is distributed uniformly to the four muons produced in the decay, so when the average energy of each muon is highest. For a Z' mass between between 3 and 6.5 GeV/c^2, the detection efficiency is above 78% for both cases considered.

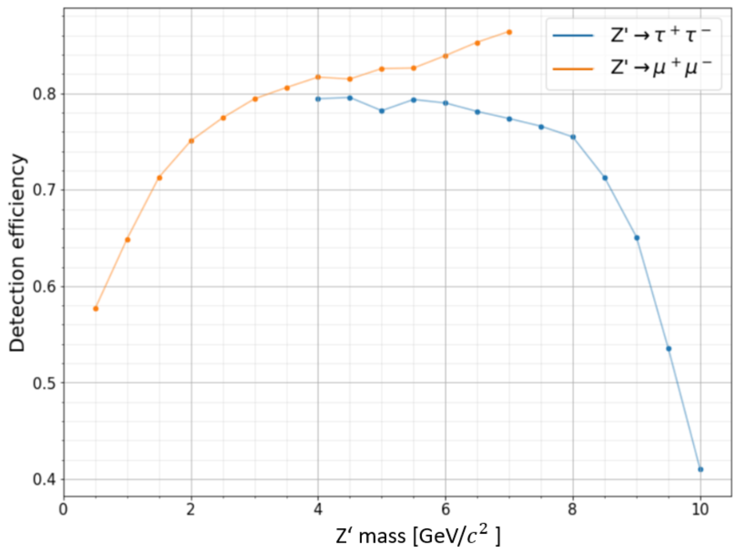


Figure 19: The detection efficiency for the two Z' processes considered, depending on the Z' mass generated, calculated according to formula (11).

4.1.3 The figure-of-merit

Once the signal events are prepared and studied by themselves, they need to be analysed together with the background. As established, the purpose is enhancing the signal/background ratio, which is achieved by applying conditions (cuts) on the variables stored and saving only the data satisfying these condition. Regarding the nature of the cuts, three important observation have to be taken into account:

1. Any cut used must be used on both the signal and the background;
2. There must be no cut made on the Z' mass, in the region where Z' can exist;
3. There should always be events left from each individual Z' mass generated.

The observations above are based on the fact that the analysis must be reproducible for real life studies. The Z' events would be collected as part of the $e^+e^- \to \mu^+\mu^-\mu^+\mu^-$ events, so the conditions could only be applied collectively for the full data, as the first point says. The second point is a reminder that the Z' mass can be anything kinematically allowed by each decay. For example, in the Z' $\to \mu^+\mu^-$ case, the minimum Z' mass generated was of 0.5 GeV/c^2, whereas in practice it can go as low as $2 \cdot M_\mu$=0.212 GeV/c^2. Cutting the invariant mass to 0.49 GeV/c^2 would be tempting for the data available, but it would exclude the possibility of many Z' mass points. Point three expresses the same rationality: for any cut applied, it should be verified that no possible Z' invariant mass is implicitly disposed of.

The goal of a "good" cut is to eliminate as many background events as possible, while simultaneously keeping as many signal events as possible. As an initial step, the BKG data is "trimmed" around the signal: all the variable values bigger/smaller than the respective maximum/minimum of the signal are thrown away[4].

Next, the invariant/recoil mass variable is cut between the minimum and maximum allowed by each of the two decays studies. In both cases, the last Z' masses generated are close to the kinematic limit, so by applying the mass condition, a couple of the signal events from the Gauss tail are also cut away. It is still a profitable cut, as table 3 proves.

After these two conditions are imposed, further one-dimensional cuts are applied - that is, boundaries on one variable at a time instead of combination of variables. In order to choose the cuts objectively, the figure of merit function is used to measure the quality of a cut. Here, the figure of merit *fom* is defined as in formula (12), where N is the counts detected for the signal, respectively background (with the reminder that BKG now includes combinatorial background as well).

$$fom = \frac{N_{sig}}{\sqrt{N_{BKG}}} \tag{12}$$

The *fom* function shows a maximum for N_{sig} maximal and N_{BKG} minimal, which is precisely the desired outcome. Linear cuts are applied on each variable, such as the condition that $E_{\mu^+}>L$ and $E_{\mu^+}<R$ in the example used in figure 20 and 21, where L and R are left and right cut values.

[4]The energy and mass of Z' are not included in this cut.

Decay	Cut	BKG^{thrown} [%]	Sig^{thrown} [%]
Z' $\rightarrow \mu^+\mu^-$	Trimming	25.48	0
	M$_{Z'}^{allowed}$	31.45	0.77
	fom	49.62	10.67
Z' $\rightarrow \tau^+\tau^-$	Trimming	33.01	0
	M$_{Z'}^{allowed}$	54.52	0.05
	fom	68.64	10.35

Table 3: The types of cuts applied sequentially, and the amount of data cut away, in percentage of the initial amount.

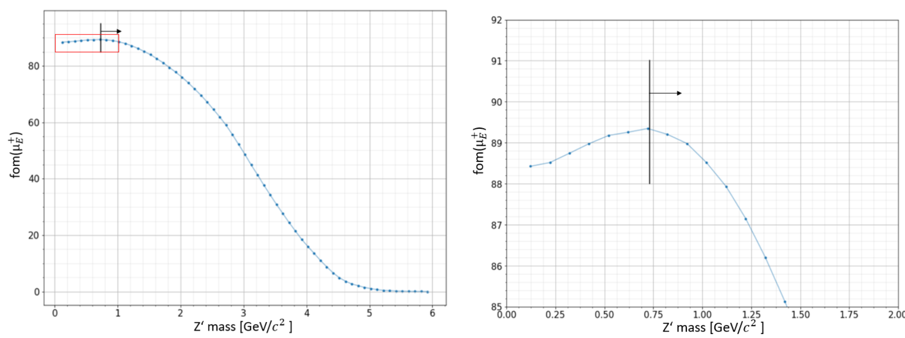


Figure 20: The L *fom* function (including zoom in) for E_{μ^+} cuts, in the Z'$\rightarrow$ $\mu^+\mu^-$ case, with the maximum marked by the black line. Data left to the cut is discarded.

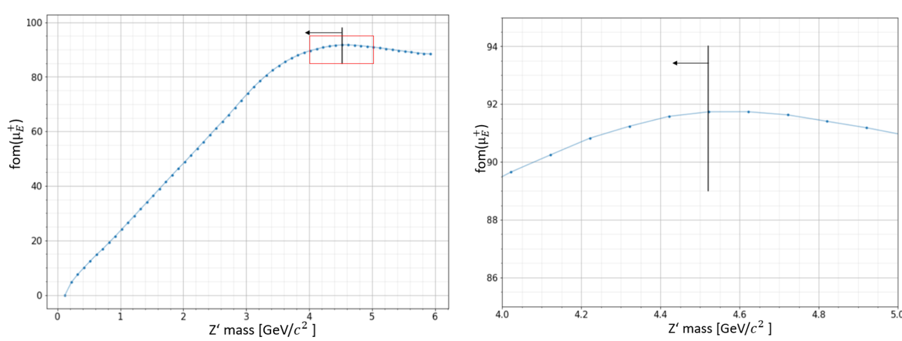


Figure 21: The R *fom* function (including zoom in) for E_{μ^+} cuts, in the Z'$\rightarrow$ $\mu^+\mu^-$ case, with the maximum marked by the black line. Data right to the cut is discarded.

As the initial step, L and R take the value of the minimum, respectively maximum of the variable considered (which is the energy E_{μ^+} of the positive muon in the example given in figure 20 and 21). The amount of events left in the signal and BKG after setting the boundaries are counted, and the *fom* calculated. Then the L is increased (R decreased) and the new *fom* value calculated. The point where the function reaches a maximum is where the ideal cut is. If the function has no local maxima, no cuts are necessary for the specific variable.

This process is automated and repeated successively for each variable except the invariant mass in the Z' $\rightarrow \mu^+\mu^-$ case, respectively the recoil mass in the Z' $\rightarrow \tau^+\tau^-$ case. At each iteration, the surviving events from the previous cuts are counted, which means the data changes at each step. However, the order in which the variables are selected for a cut hardly influences the outcome.

The formula chosen in (12) has it's downside of potentially reaching infinity values. It can run until all background is eliminated and some irrelevant amount of signal remains. A way to avoid this is having the denominator be replaced with $N_{sig} + N_{BKG}$, but this doesn't give the optimal outcome of minimal signal, which is why formula (12) is preferred. The reason the square-root of the background count is used is because keeping the signal is prioritized over minimizing the BKG count.

Regardless of how good a cut is according to the *fom* function, it shouldn't be applied if it would throw away potential Z' masses. This can be checked by plotting the Z' mass against the variable considered. In figure 22 only the data between the two lines is saved, which includes all the possible Z' mass lines.

Applying first the L *fom* cut and then the R one results in virtually identical outcomes (affecting just the decimal of the percentage thrown). The order of execution, as well as other constrains are chosen depending on the resulting cross-section value calculated in chapter 4.1.4. For Z'$\rightarrow \tau^+\tau^-$, RL is preferred to LR. In the Z'$\rightarrow \mu^+\mu^-$ however, the automatic *fom* function method is abandoned all-together. Applying the function iteratively on all variables eliminates the low Z' masses entirely, and leaves very few in the medium mass region. A possible cause for this can be the fact that all the variables saved are directly connected to Z'[5]. Applying

[5] This is not the case for Z'$\rightarrow \tau^+\tau^-$, where only the recoil mass is directly describing Z'.

cuts on all of them ultimately cuts on the Z' masses. To make sure the cuts are chosen wisely, a monitored step by step execution of the *fom* function is necessary. The *fom* plot is investigated individually for each variable, as presented in figure 20 and 21. The only variables worth cutting on turn out to be the energy and the z-axis momentum component of one muons in the $\mu^+\mu^-$ system. Applying conditions on one muon inevitably does so on the other as well, so the choice is arbitrary (here, it's μ^+ variables which are cut).

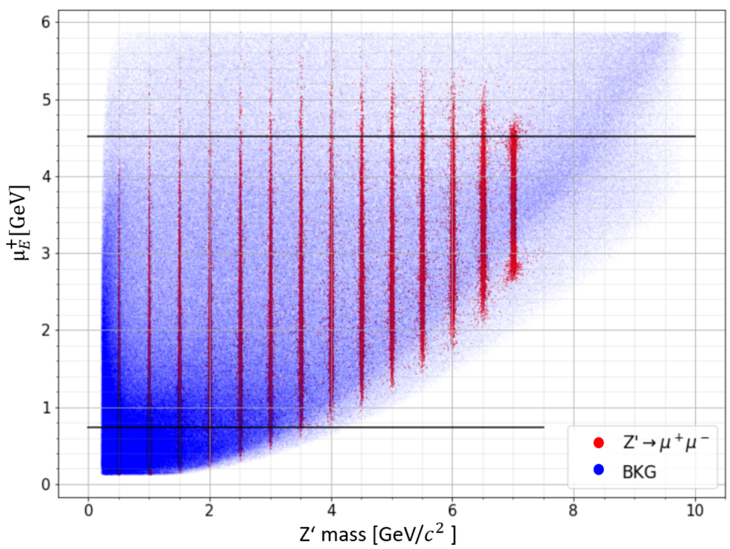


Figure 22: Scatter plot of the invariant mass of the $\mu^+\mu-$ system, against the energy of μ^+, in the Z'→ $\mu^+\mu^-$ case, with the *fom* cuts marked by the black lines. Data outside the cuts is discarded.

4.1.4 Z' event cross-section

The plots presented until now are not a fully accurate portrayal of how a signal of this kind would look compared to the BKG. In this study, the amount of dark signals generated was a hundredth of the background count (ten thousands out of a million), which is a very big ratio considering BSM events are expected to have small cross-sections.

The goal is to find out how many Z' events N would have to occur, in order for its mass peak to be visible above the BKG. More precisely, the search is for the minimum cross-section of the Z' events in order for them to be detectable: the σ upper limits. The use of "upper" for minimal values can be confusing, but the limits are set on the exclusion region: up to the upper limits, no detection would be possible using the study in question. To obtain the σ values, formula (4) is used, for an integrated luminosity calculated in (3) to be 2.9234 ab^{-1}. The efficiency ϵ is computed for each Z' mass in the two cases studied, as in chapter 4.1.2. In order to cover the 95% confidence interval, the minimum number of Z' events needed is $N \geq 1.645 \cdot std(N_{det}^{BKG})$, where *std* is the standard deviation of the Gauss fit. In a normal distribution, 95% of the data can be found in the (mean $\pm$ 1.645 *std*) region, so the Z' counts exceeding this limit implies a high chance of visibility. Since the detection of particles is a counting experiment, the standard deviation of N_{det}^{BKG} is its square root.

The final form of the function used is given in equation (13), with the only variables for each mass being the BKG counts and the signal counts. They vary because they're not the total counts of the respective data, but only the counts in each signal region. When calculating the cross-section for a certain Z' mass, the search must be limited around the respective mass. Gauss fits are applied on each mass peak, and the Z' is counted in the interval where 90% of the data lies (N_{det}^{sig}). Then the BKG is counted in the (mean $\pm$ $2std$) region of the respective fit (N_{det}^{BKG}).

$$\sigma_{Z'}[fb] = \frac{1.645\sqrt{N_{det}^{BKG}}}{2923.4\frac{N_{det}^{sig}}{10000}} \tag{13}$$

Until this point, both the signal and the background data have been modified for the function (13) to yield better results. The smaller the cross-sections are, the better. On one hand because it implies a higher chance of discovery, if using the same analysis steps in a real study. On the other hand, if no Z' signal is found, then lower upper limits means a narrower remaining possible window for its cross-section. Figure 23 shows how the cross-section function decreases by applying successive cuts on the data.

Due to correlations, the order of the cuts affected the resulted cross-section, especially for certain mass point. Figure 23 displays the overall lowest results obtained. Smaller values can be obtained for certain mass region, at the expense of others. The same conclusion can be drawn from the other decay considered, whose overall lowest σ upper limits are presented in figure 24.

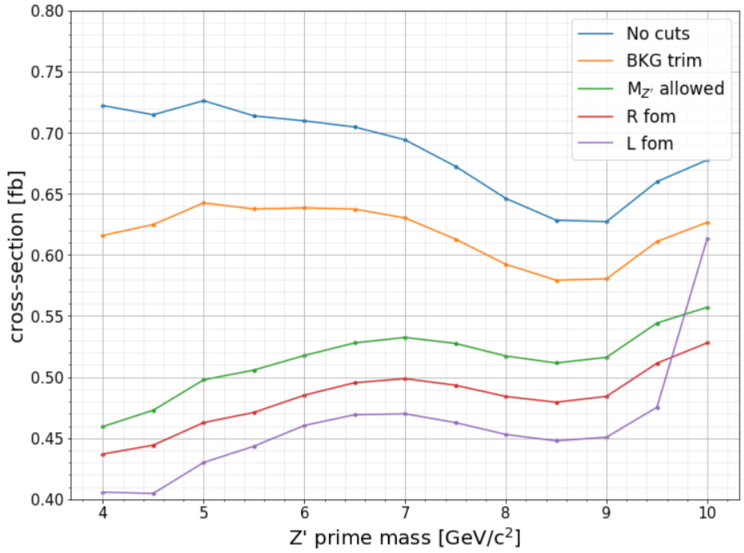


Figure 23: The cross-section upper limits of the $e^+e^- \to \mu^+\mu^- Z'[\to \tau^+\tau^-]$ decay. The cuts are applied successively, starting with the upper one listed.

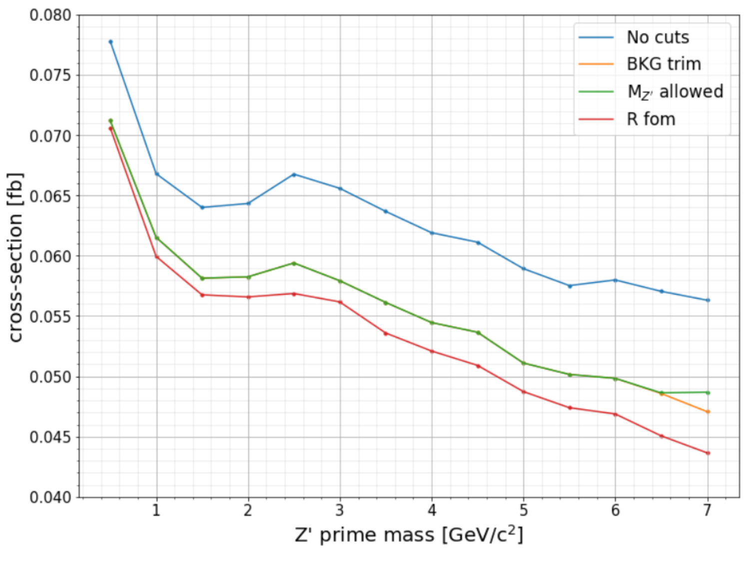


Figure 24: The cross-section upper limits of the $e^+e^- \to \tau^+\tau^- Z'[\to \mu^+\mu^-]$ decay. The cuts are applied successively, starting with the upper one listed.

In principle, it would be profitable to divide the data in two or three subsets, depending on the Z' mass, and to adjust the analysis accordingly. Also, two dimensional visualisation of the variables reveals further possible cuts. There are two reasons the segmentation of data and scatter plot cuts are not further examined here. First, a 2D or 3D *fom* function for all the cuts possible would be too time consuming, and choosing the cuts visually not fully reliable. Second and more important, much better cross-section values (as low as 0.8 ab) are obtained using another reconstruction method.

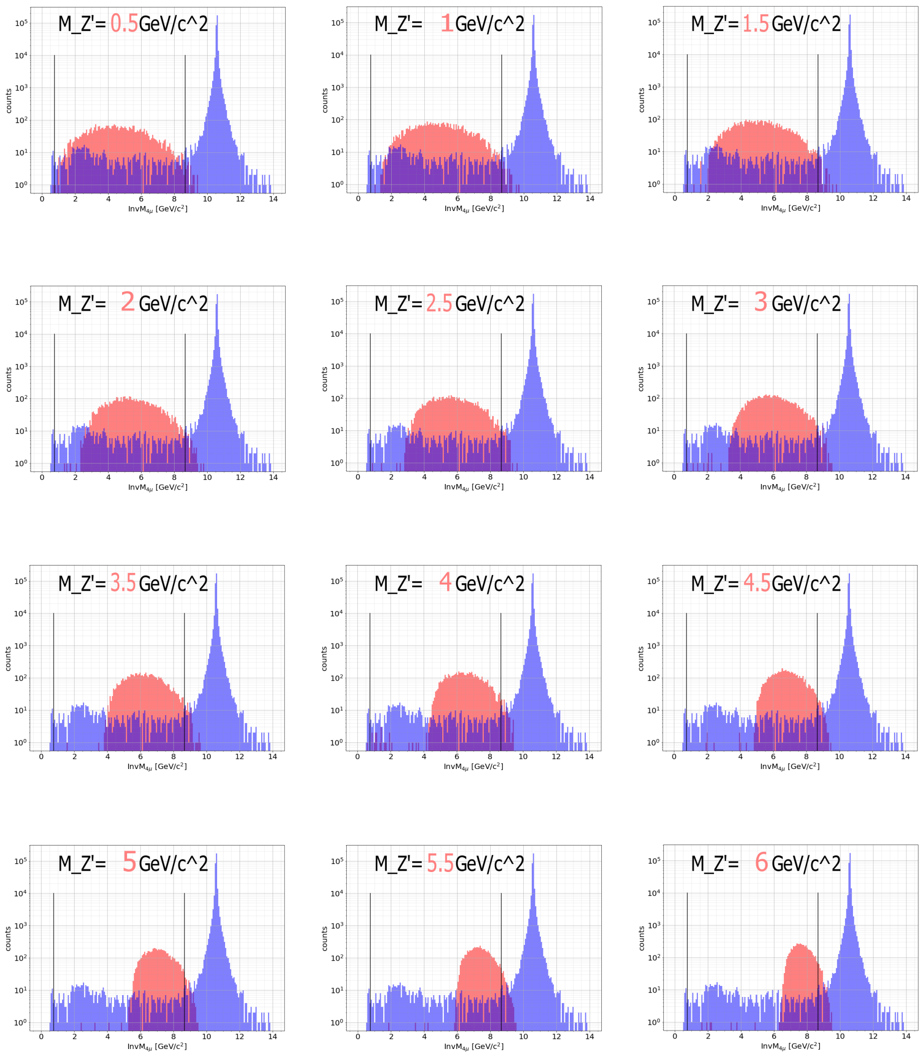
Chapter 4.2 presents the different approach to reconstructing Z' and the advantages and disadvantages coming with it. It includes the single most successful cut applied, and the resulting cross-section values, which are used to compute the upper limits of the g' coupling constant.

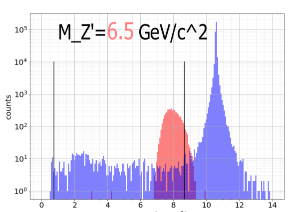
4.2 Reconstructing the four muon system

Collecting data on pairs of two muons is virtually equivalent to collecting informaion on the Z' boson itself, therefore it seems like the preferable option. However, it is not the only one offering helpful insight into the decay: reconstructing all four muons as a single particle comes with a significant advantage. In this chapter, variables of the four muon system were stored, as well as the variables of each of the four muons detected.

Most of the $e^+e^- \rightarrow \mu^+\mu^-\mu^+\mu^-$ processes occur without the involvement of invisible particles which take away from the available energy but escape undetected. This means that in the vast majority of the cases, the 10.58 GeV of the initial energy is entirely passed on to the four final state muons. When it comes to the Z' decays however, four out of the resulting 8 final state particles are neutrinos. As a result, a certain amount of the initial energy is always lost for the signal decay. With the invariant mass of the $\mu^+\mu^-\mu^+\mu^-$ system available, a single cut can take out most of the background, while keeping the signal virtually intact.

Figure 25 illustrates the results as predicted: when comparing the histograms of the signal and background, the BG counts pile up close to the collision energy, while the signal's maximum is lower. Figure 39 shows the plots for all fourteen Z' masses individually, visualizing the way the phase-space changes as the Z' mass increases.





The phase space of the 4μ invariant mass, for the background (blue) and the 14 signal mass points (pink), in the Z' $\rightarrow \mu^+\mu^-$ case.

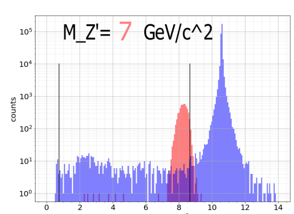


Figure on page 39 is for the Z' $\rightarrow \mu^+\mu^-$ decay, while similar but mirrored behavior is observed for the Z' $\rightarrow \tau^+\tau^-$ decay, with the 4μ mass distribution being more "compact" for the lower Z' masses.

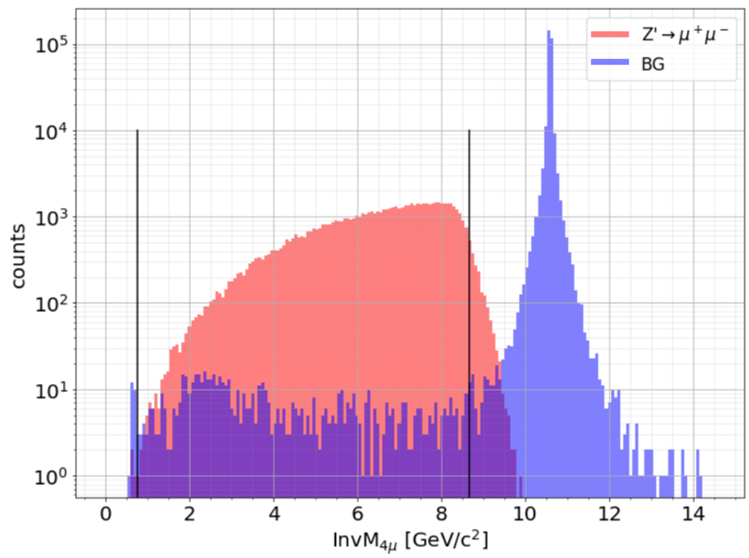


Figure 25: The invariant $\mu^+\mu^-\mu^+\mu^-$ mass, with the L and R *fom* cuts marked in black. Everything outside the cuts is discarted. Here, the signal (pink) includes combinatorial background.

The cut applied on the four muon invariant mass was deduced using the same figure of merit function defined in equation (12). Other than this, all the BG variable values were "trimmed" to be between the minimum and maximum values of the signal, as was done in the two muon system case considered previously. Table 4 displays the results of the initial cuts.

Decay	Cut	BG^{thrown} [%]	Sig^{thrown} [%]
Z' $\rightarrow \mu^+\mu^-$	4μTrim	99.8	0
	InvM$_{4\mu}$	99.89	0.03
Z' $\rightarrow \tau^+\tau^-$	4μTrim	99.81	0
	InvM$_{4\mu}$	99.9	1.73

Table 4: The types of cuts applied sequentially, and the amount of data thrown away in percentage of the initial amount. Here, the signal includes combinatorial background from the wrong muon combinations.

Despite providing the means to apply the most successful cut, there is three disadvantages when using this method.

1. The signal can not be separated before applying the cut on the four muon invariant mass;

2. The Z' mass must be computed separately;

3. There is a risk of eliminating all background in certain mass points.

Because all four muons are registered, all of the lines in the data frame (the table such as the one present in figure 8) contain both the muons coming from the Z', and the remaining ones. So there is no way of cleaning the signal and separating the muons resulting in the Z' peaks from the wrongly combined ones, while still having access to data on the 4μ system. When applying the cuts mentioned, they remove less background than it appears, since they consider the combinatorial background as part of the signal. Since almost all of the BG can be cut out, it is advantageous to continue working with this method.

Regarding the second point made, having only the four muon data saved, no peak for the Z' is visible. In both the decay studied, the mass of the Z' has to be computed using formula (7) and respectively (8). The energies and momenta on each muon are registered, so they can be combined in pairs of two such that they result in the invariant mass of the Z' boson. For the recoil mass necessary in the Z' $\rightarrow \tau^+\tau^-$ case, the four muons are reconstructed with their CMS energies, wich are added in different combinations to result in the CMS energy of the $\mu^+\mu^-$ systems.

For both the BG data and the signal ones, the positive muons are paired with the negative ones, resulting in the four combinations possible discussed. Once the pairs are formed, they are stored in new BG and signal data frames, which contain around four times more events. All the variables of the respective muons are also moved from the 4μ data frames to the new ones, including the energy and momenta of each muons, as well as their MC production vertices. As it was done in previously, the MC production vertex condition is used to obtain the clean Z' signal and a full BKG background.

However, cutting away such a big percentage of the SM background also complicates the study, which is why the third point was made. After applying the cuts on the four muon variables and forming the two muon pairs as explained, there is virtually no real background left in the signal regions. With less than twenty BG events surviving the cut for each Z' mass point, there is less background from SM processes than the combinatoric background from the dark process. As a result, when calculating the cross-section of the dark events using the same steps as before, the pure Z' signal is primarily compared against its own background.

The data is visualized in figure 26. Comparing it to figures 15 and 16 from the previous method, it is clear that the SM background is drastically reduced. As much as this is precisely what is desired, it is not useful when calculating the cross-sections. The combinatorial background is dominating the BKG to the extend that for some signal regions there is no SM background left.

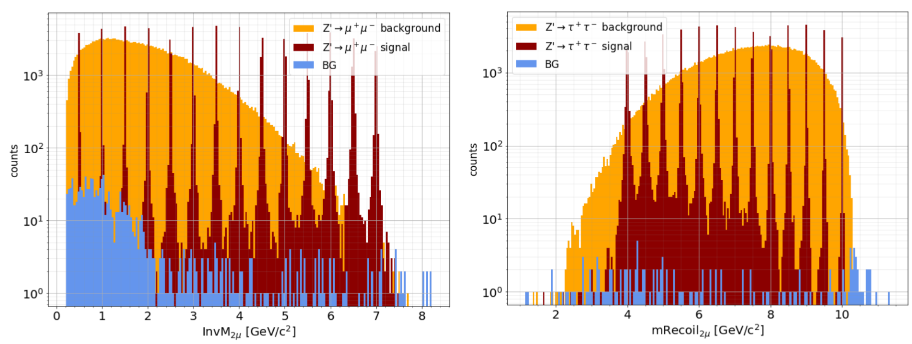


Figure 26: The computed invariant $\mu^+\mu^-$ mass (left) and the recoil mass of the $\mu^+\mu^-$ system (right), for the SM background, combinatorial background, and signal.

With this much discrepancy between the real background and the combinatorial one, treating them as the same is not as accurate. In formula (13), the integrated luminosity of 2.9234 ab^{-1} is computed using the cross-section of the $e^+e^- \to \mu^+\mu^-\mu^+\mu^-$, which is a known value corresponding to SM processes. When the majority of the background is in fact not SM, normalizing it this way to obtain the dark cross-section is not valid. Ideally, more $e^+e^- \to \mu^+\mu^-\mu^+\mu^-$ background should be generated. For Z' decaying to muons for example, there is 295,996 background events detected (out of the 1,000,000 generated), and only 325 of those survive the initial cuts on the 4μ variable. By applying further cuts on 2μ variables, the size is reduced even more than displayed in figure 26. If significantly more SM background would be generated, merging it with the combinatorial would barely affect the overall luminosity and implicitly the resulting error of the dark cross-sections.

For the data available in this part of the study, the combinatorial background can not be safely assumed to be part of BG, which leaves two options:

1. Discard the combinatorial background entirely;
2. Consider the combinatorial background as part of the signal.

Adjusting formula (13) accordingly, the two case are equivalent to equation (14), respectively (15). The background counts considered is reduced in both scenarios, and for the second, there is also the enhancement on the signal count.

$$\sigma_1[fb] = \frac{1.645\sqrt{N_{det}^{BKG-combBG}}}{0.29234 N_{det}^{sig}} \tag{14}$$

$$\sigma_2[fb] = \frac{1.645\sqrt{N_{det}^{BKG-combBG}}}{0.29234 N^{sig+combBG}} \tag{15}$$

Using equation (14) would be the most correct approach, if the signal separation would be done using something other than MC variables. If there is a way to separate the pure Z' signal from its combinatorial background that can be applied to real data, then throwing away the wrong combinations would be the ideal solution. The second approach considered is more fitting for this thesis. Using equation (15) to calculate cross-sections of the dark events results in lower values than realistically. Since the signal is always visible against its own background however, counting them both as Z' events doesn't invalidate the study.

Figure 27 and 28 display all the resulting cross-sections, using formulas (13), (14) and (15) to compute them. For Z'$\rightarrow \mu^+\mu^-$, the last two graphs overlap for the higher Z' mass points, bringing out again the fact that the signal peaks separate from the combinatorial background around the $M_{Z'}$=5.5 GeV/c^2 mass point. After the cuts on the for muon variables are applied and the two muon systems formed, further background "trimming" and *fom* cuts can be taken into consideration. In the Z'$\rightarrow \mu^+\mu^-$ case, the type and order of the cuts is chosen as to minimize the final cross-sections. For Z'$\rightarrow \tau^+\tau^-$ however, there is too little SM background left at this point, as is clear from figure 26 (right).

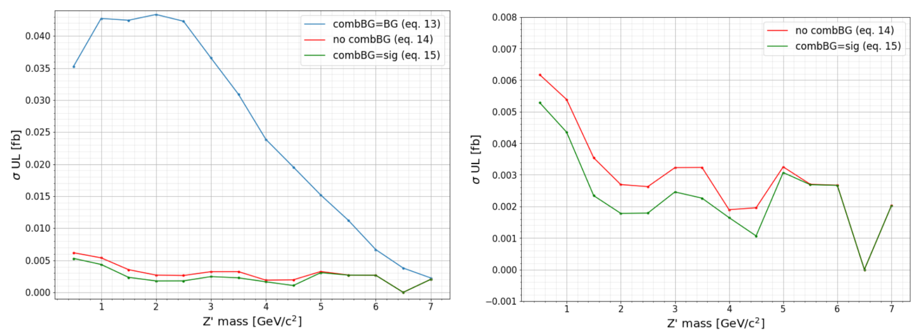


Figure 27: The cross-section upper limits for the $e^+e^- \rightarrow \tau^+\tau^- Z'[\rightarrow \mu^+\mu^-]$ decay, for different treatment of the combinatorial background. The lowest graph (green) is the one used to compute the g' upper limits.

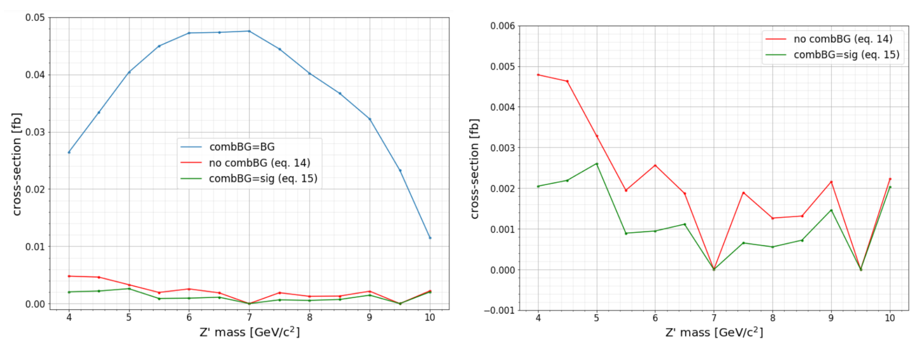


Figure 28: The cross-section upper limits for the $e^+e^- \rightarrow \mu^+\mu^- Z'[\rightarrow \tau^+\tau^-]$ decay, for different treatment of the combinatorial background. The lowest graph (green)is the one used to compute the g' upper limits.

In order to not eliminate the background counts entirely, no other cuts than the initial 4μ cuts are applied. Even so, there is two mass points where no background remains and the resulting cross-section is 0 - for the 7 and 9.5 GeV/c^2 signal regions. These points are not considered when computing the g' upper limits. There is also an abandoned mass point for Z'$\rightarrow \mu^+\mu^-$ (of 6.5 GeV/c^2). To make sure this does not occur for more mass points, in both decays considered the events are counted in the $\pm 3\sigma$ signal region, which covers 99.73% certainty. Furthermore, instead of 90% of the signal region, the entire Z' peak is fitted.

The final cross-section upper limits for the two decays are presented in figure 29. Being the lowest obtained, they are the values used to compute the g' coupling constant. Again, the goal is obtaining low g' upper limits, and the cross-section σ values are directly related (inverse proportional) to the g' coupling constant.

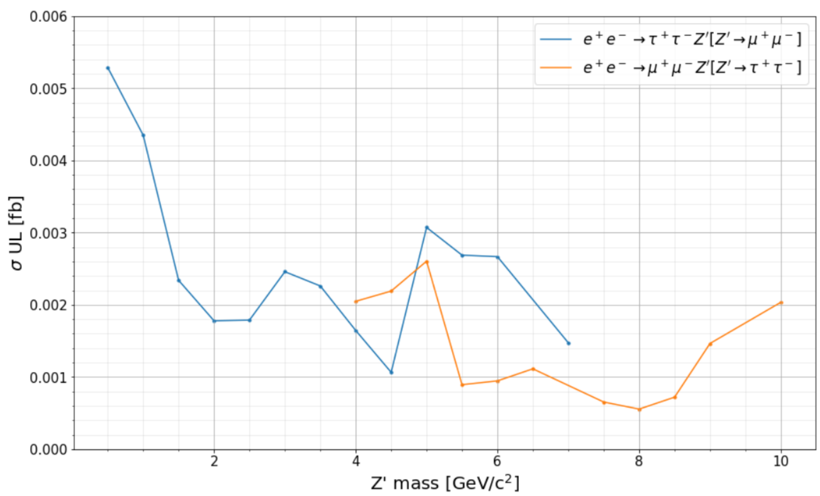


Figure 29: The cross-section upper limits of the two decays.

4.3 g' Upper Limits

As the other coupling constants describing the four fundamental forces, g' indicates the strength of the dark force, present in an U(1)' interaction such as the one between the Z' and the muon/tau. The coupling constant affects the Lagrangian describing the interaction, and its value is established experimentally.

In this case, the g' value is calculated iteratively using MadGraph and the cross-sections calculated previously, and plotted in figure 29. For any decay process generated, MG computes its corresponding cross-section. The value of g' needs to be specified at the generation level, and it influences the outcome of the cross-section value. By setting different g' values and running the code for each of them, the optimal (minimum) one can be found. That is, the g' value which results in a σ value as close as possible to, but bigger than the upper limits displayed in figure 29. The g' upper limits are plotted in figure 30, and listed in table 5.

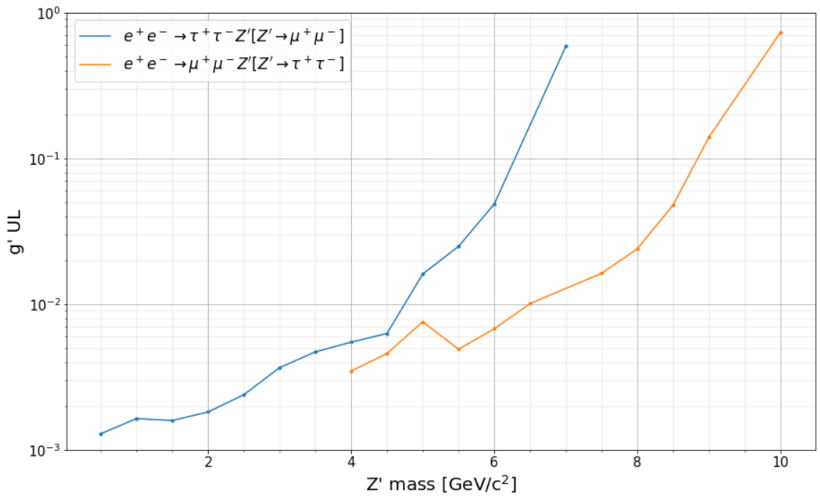


Figure 30: The coupling constant g' upper limits of the two decays.

$e^+e^- \rightarrow \tau^+\tau^- Z'[Z' \rightarrow \mu^+\mu^-]$	
$M_{Z'}$	g'_{UL}
0.5	0.00129
1	0.00164
1.5	0.00159
2	0.00182
2.5	0.00239
3	0.00367
3.5	0.0047
4	0.00548
4.5	0.00629
5	0.01609
5.5	0.02485
6	0.04862
6.5	-
7	0.5869

$e^+e^- \rightarrow \mu^+\mu^- Z'[Z' \rightarrow \tau^+\tau^-]$	
$M_{Z'}$	g'_{UL}
4	0.00347
4.5	0.00458
5	0.00756
5.5	0.0049
6	0.00677
6.5	0.01008
7	-
7.5	0.01625
8	0.02397
8.5	0.04786
9	0.13967
9.5	-
10	0.7308

Table 5: The upper g' limits for each Z' mass considered.

The g' results increase exponentially as the Z' mass increases. Even in the highest Z' mass points allowed, the g' remains sub-unitary, as desired and expected for a coupling constant. This implies the study is sensitive for all mass regions whose cross-sections are computed to be non-zero.

4.4 Other data visualisation types

The previous chapters have been focused on the steps made to obtain the lowest σ and g' limits. However there is other plots offering interesting if not helpful observations about the two data sets. Visualizing the BKG and signal data in different ways offers great insight into the processes.

As mentioned, the cuts can also be applied on scatter plots, as exemplified in figure 31. There, the momentum of the $\mu^+\mu^-$ system if plotted against its projection on the z axis, and the $\mu^+\mu^-$ energy against one of the muon's energy. The cuts illustrated are chosen visually, without running a *fom* analysis. Calculating the *fom* function for each line in the plane would not be practical. However, adding scatter plot cuts in this manner does not improve (i.e. lower) the resulting cross-section upper limits.

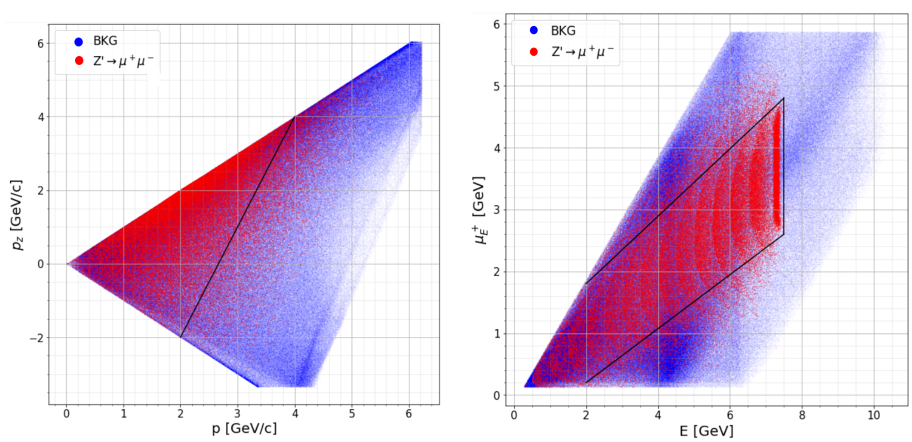


Figure 31: Possible scatter plot cuts for the $e^+e^- \rightarrow \tau^+\tau^- Z'[\rightarrow \mu^+\mu-]$ case. Data containing most of the signal events would be kept.

Scatter plots of various variable combinations also display the certain trends than occur. For example, the 7 GeV/c^2 Z' mass in the Z'$\rightarrow \mu^+\mu^-$ case has a distinct distribution in virtually all variables, and can be singled out from the rest of the masses the easiest. Due to the limited momentum the muons can have for this Z' mass point, the events are all distributed around the 2 GeV/c^2 value, as visible in figure 32, where pt is the transverse momentum, i.e the xOy projection of the momentum. Both graphs in figure 32 indicate that the highest Z' mass produced in the $e^+e^- \rightarrow \tau^+\tau - Z'[\rightarrow \mu^+\mu^-]$ decay is almost always propagated forward.

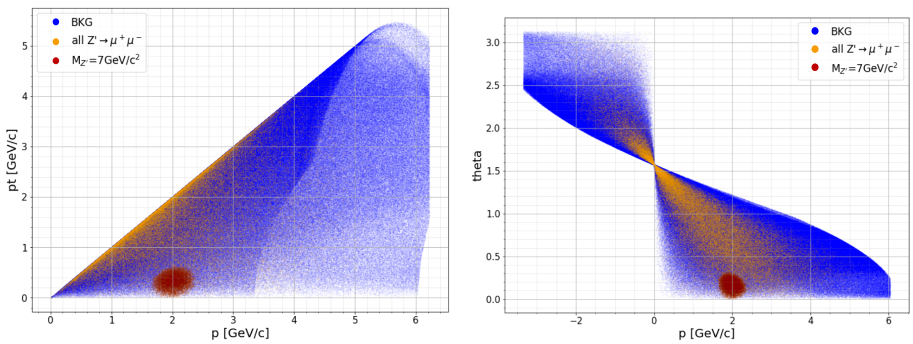


Figure 32: Z' momentum p vs transversal momentum pt (left) and versus opening angle θ, for Z'$\rightarrow \mu^+\mu^-$.

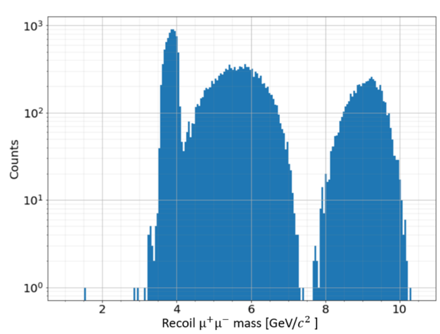


Figure 33: The $\mu^+\mu^-$ recoil mass, in the $e^+e^- \rightarrow \tau^+\tau^- Z'[\rightarrow \mu^+\mu-]$ case, for M$_{Z'}$=7 GeV/c^2, including wrong combinations.

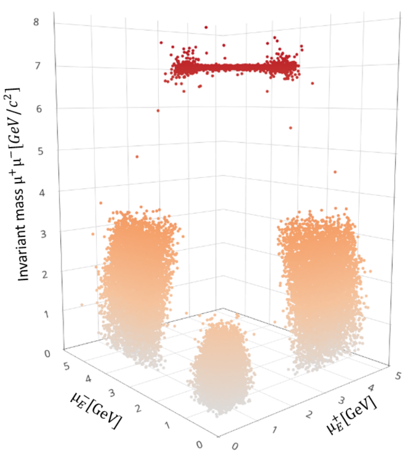


Figure 34: The $\mu^+\mu^-$ invariant mass (color coded), versus the energies of the two muons, for M$_{Z'}$=7 GeV/c^2, including wrong combinations.

Because of its constrained phase space, the 7 GeV/c^2 Z' mass is ideal for studying the distribution of each of the four $\mu^+\mu^-$ combinations. For example, plotting the $\mu^+\mu^-$ recoil mass results in three distinguishable peaks, representing the remaining combinations of muons not coming from Z' (figure 33). The same clear separation of the muon combinations is also observed in figure 34, which is a 3D scatter plot of the energies of the two muons and the invariant mass of the $\mu^+\mu^-$ system.

In the scatter plots displayed so far, the signal events are plotted with higher opacity than the BKG, for better visibility. Figure 35 gives a more realistic view of the signal and background distribution, since it plots all the event dots with the same size and opacity. This illustrates best how hard it is to make out the low masses against the densest region of the BKG.

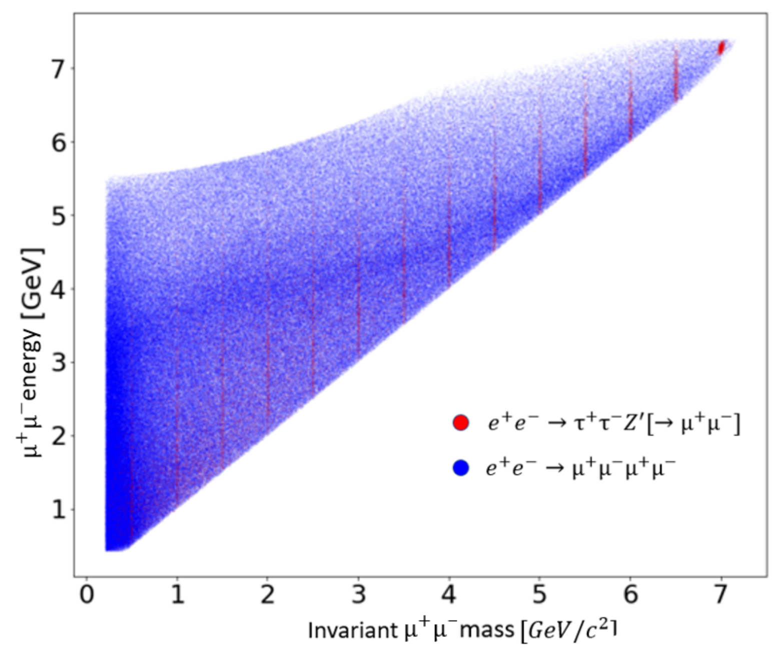


Figure 35: Energy of the $\mu^+\mu^-$ system versus its invariant mass,

5 Results and Discussion

The resulted g' upper limits values fit in the expected region between 0 and 1. The higher the Z' mass, the stronger the interaction needs to be in order for the cross-section minimal values to be met and the dark events to be detectable. This exponential behavior is better visualized in figure 37, and is present regardless which method is used to calculate the cross-sections.

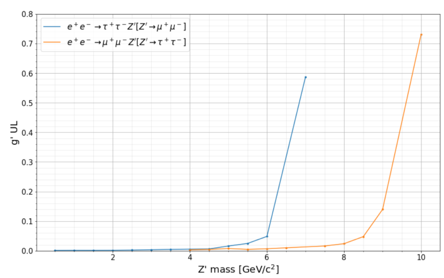


Figure 36: The g' upper limits, no axis logarithmized.

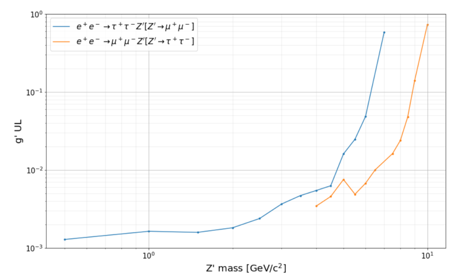


Figure 37: The g' upper limits, both axis logarithmized.

As already concluded, more $e^+e^- \to \mu^+\mu^-\mu^+\mu^-$ background should be generated in similar studies. Even if most of the four muon invariant mass peaks at the end of the spectrum, having more SM background would ensure that the sufficient counts are present in each signal region. However, since the data generated was equivalent to $\int Ldt \approx 2.9ab^{-1}$, it is enough for an initial Belle II study. In order for the combinatorial background to be grouped together with the SM one with no repercussion, the latter should numerically dominate. Another big advantage would come from discovering a way to separate the signal without using Monte Carlo variables.

One way to separate the signal from the BG or to at least suppress the BG is by letting the computer do it itself. Recent developments in the field of machine learning has made classification faster and often more accurate. There are many algorithms available for binary classification, which is what a study of this kind requires. Part of the signal and background data is used to train an algorithm to differentiate between them (at the training stage, it's known which is the signal). All variable entries are compared, and the algorithm seeks to learn what makes the ones belonging to the signal special.

In this study, Random Forest [30] was trained on the 2μ data, but it wasn't optimized further due to lack of time. Another recommended classifier for particle events is fast BDT (boosted decision trees), which has a reputation of delivering trustworthy results, though it was not used here.

Another machine learning technique looked into here was HDB-SCAN [31], which is a clustering algorithm: it's fed unlabeled data such as the background together with the signal, and it groups together objects which are similar. For a scatter plot such as in figure 35, it should be able to determine the Z' mass lines due to the higher density. It performs well for the 7 GeV/c^2 mass and then more and more poorly for lower masses. As mentioned, for the maximum M$_{Z'}$ in the Z' $\rightarrow$ $\mu^+\mu^-$ case, respectively the minimum M$_{Z'}$ in the Z' $\rightarrow$ $\tau^+\tau^-$ case, the two taus can't take momentum from the decay, resulting in tighter phase-space for the 2-muon system.

Considering that the study is performed almost entirely relying on software, the only errors that could have occurred are of BASF2 and MadGraph nature. Especially with the latter, since it was used not only to generate the dark events, but to compute the final results for the g' upper limits. The values obtained are entirely computed by MadGraph, with no human control over them other than the original Monte Carlo programming and the setting of the decay parameters. Regardless, simulation studies of this kind are proved to be reliable.

If the theoretical model considered is correct in explaining the U(1)' symmetry extension and its connection to the dark sector, and if all the other assumptions made are satisfied, then the Z' boson could be identified by the Belle II detector. The dark decays $e^+e^- \rightarrow \tau^+\tau^-Z'[\rightarrow \mu^+\mu-]$ and $e^+e^- \rightarrow \mu^+\mu^-Z'[\rightarrow \tau^+\tau-]$, for the taus decaying to muons and neutrinos, offer a promising premise for studying and observing the Z' boson. For the Z' masses and the context assumed, using the same tools and analysis methods as presented in the previous chapters, the Z' boson should peak above the Belle II background, given the g' upper limits computed for each respective mass.

References

[1] E. V. Karukes and P. Salucci, "Modeling the mass distribution in the spiral galaxy NGC 3198," *Journal of Physics: Conference Series*, vol. 566, p. 012008, dec 2014.

[2] A. Robertson, R. Massey, and V. Eke, "What does the bullet cluster tell us about self-interacting dark matter?," *Monthly Notices of the Royal Astronomical Society*, vol. 465, p. 569–587, Oct 2016.

[3] T. Abe *et al.*, "Belle II technical design report," 2010.

[4] P. Gagnon, "The standard model: a beautiful but flawed theory," *Quantum Diaries*, 2014.

[5] C. Giunti, C. W. Kim, and U. W. Lee, "Running coupling constants and grand unification models," *Mod. Phys. Lett.*, vol. A6, pp. 1745–1755, 1991.

[6] G. Bertone, D. Hooper, and J. Silk, "Particle dark matter: evidence, candidates and constraints," *Physics Reports*, vol. 405, p. 279–390, Jan 2005.

[7] M. Milgrom, "MOND theory," *Canadian Journal of Physics*, vol. 93, p. 107–118, Feb 2015.

[8] G. Arcadi, M. Dutra, P. Ghosh, M. Lindner, Y. Mambrini, M. Pierre, S. Profumo, and F. S. Queiroz, "The waning of the WIMP? a review of models, searches, and constraints," *The European Physical Journal C*, vol. 78, Mar 2018.

[9] N. Bernal and X. Chu, "Z_2 SIMP dark matter," *Journal of Cosmology and Astroparticle Physics*, vol. 2016, p. 006–006, Jan 2016.

[10] H. Kleinert, "The GIMP Nature of dark matter," *Electron. J. Theor. Phys.*, vol. 13, no. 36, pp. 1–12, 2016.

[11] D. J. Marsh, "Axion cosmology," *Physics Reports*, vol. 643, p. 1–79, Jul 2016.

[12] S. Dodelson and L. M. Widrow, "Sterile neutrinos as dark matter," *Physical Review Letters*, vol. 72, p. 17–20, Jan 1994.

[13] R. Essig *et al.*, "Dark sectors and new, light, weakly-coupled particles," 2013.

[14] J. Beringer *et al.*, "Review of particle physics," *Phys. Rev. D*, vol. 86, p. 010001, Jul 2012.

[15] I. Adachi *et al.*, "Search for an invisibly decaying Z′ boson at Belle II in $e^+e^- \to \mu^+\mu^-(e^{\pm\mp})$ plus missing energy final states," 2019.

[16] X.-G. He, G. C. Joshi, H. Lew, and R. R. Volkas, "Simplest Z′ model," *Phys. Rev. D*, vol. 44, pp. 2118–2132, Oct 1991.

[17] T. M. Undagoitia and L. Rauch, "Dark matter direct-detection experiments," *Journal of Physics G: Nuclear and Particle Physics*, vol. 43, p. 013001, Dec 2015.

[18] F. Takasaki, "The discovery of CP violation in B-meson decays," *Proceedings of the Japan Academy. Series B, Physical and biological sciences*, vol. 88, no. 7, pp. 283–298, 2012.

[19] T. I. J. National Lab for High Energy Physics, "KEKB B-factory design report," tech. rep., Japan, 1995. KEK–95-7.

[20] M. Bračko and the SuperBelle Collaboration, "Status of the SuperKEK B factory," *Journal of Physics: Conference Series*, vol. 171, p. 012098, jun 2009.

[21] Y. Ohnishi *et al.*, "Accelerator design at SuperKEKB," *Progress of Theoretical and Experimental Physics*, vol. 2013, 03 2013.

[22] S. X. staff, "First particles circulate in SuperKEKB accelerator," Apr 2016.

[23] C. Hearty, "Dark sector/dark matter searches at BaBar and outlook for Belle II," Oct 2018.

[24] "https://www.belle2.org/research/detector," 2017.

[25] "Kick-off of the Belle II phase 3 physics run," Mar 2019.

[26] V. Aulchenko *et al.*, "Electromagnetic calorimeter for Belle II," *Journal of Physics: Conference Series*, vol. 587, p. 012045, feb 2015.

[27] T. Aushev *et al.*, “A scintillator based endcap KL and muon detector for the Belle II experiment,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 789, p. 134–142, Jul 2015.

[28] M. Tanabashi *et al.*, “Review of particle physics,” *Phys. Rev.*, vol. D98, no. 3, p. 030001, 2018.

[29] O. M. Celine Degrande, “Tasi 2013: FeynRules/MadGraph tutorial,” 2017.

[30] P. C. Bhat, “Multivariate analysis methods in particle physics,” *Annual Review of Nuclear and Particle Science*, vol. 61, no. 1, pp. 281–309, 2011.

[31] L. McInnes, J. Healy, and S. Astels, “HDBscan: hierarchical density based clustering,” *The Journal of Open Source Software*, vol. 2, 03 2017.