Study on the search for leptoquarks in lepton-quark collisions at the LHC using proton tagged events

Akrap, Petra

University of Split
Faculty of Science

Study on the search for leptoquarks in lepton-quark collisions at the LHC using proton tagged events

Master thesis

Petra Akrap

Split, October 2023

As part of Erasmus+ programme, this thesis was done at Sapienza University of Rome, Italy. I would like to give my thanks to CMS Rome group, in particular to Dr. Francesco Santanastasio and Dr. Paolo Meridiani for granting me an opportunity to work on this topic, and for all their knowledge, patience and support during my time in Rome. I'd also like to express my heartfelt thanks to the friends I made in CMS "baita". I truly appreciate your warm welcome and genuine inclusiveness.

I am very grateful to my supervisor Dr. Toni Šćulac for all the advice, encouragement and help provided throughout my 2nd year of my Master's programme.

Lastly, I am also very thankful to all my friends and family that have supported me and have always been there by my side throughout this journey.

Temeljna dokumentacijska kartica

Sveučilište u Splitu
Prirodoslovno – matematički fakultet
Odjel za fiziku
Ruđera Boškovića 33, 21000 Split, Hrvatska

Diplomski rad

Istraživanje potrage za leptokvarkovima u sudarima lepton-kvark na LHC-u koristeći događaje označene protonima

Petra Akrap

Sveučilišni diplomski studij Fizika, smjer Računarska fizika

Sažetak:

Leptokvark (LQ) je hipotetska čestica koja omogućava transformaciju kvarkova u leptone i obrnuto. Predviđen je u mnogim teorijama izvan Standardnog modela. U ovom radu se nastoji procijeniti potencijal subdetektora *Precision Proton Spectrometer* (PPS) u svrhu poboljšanja trenutačne inkluzivne potrage za LQ-ima, s posebnim naglaskom na mehanizam pojedinačne produkcije LQ-a koji je induciran leptonima. Analiza se provodi u dva dijela. Prvo se temeljito proučavaju signalni događaji LQ-a služeći se Monte Carlo generatorima, zalazeći u raspodjele mase čestice, kinematičke raspodjele njezinih finalnih stanja te presjeke za različite vrste LQ-ova. Nakon toga slijedi propitivanje rekonstrukcije raspršenih protona unutar PPS-a te definiranje kriterija za označavanje raspršenog protona. Kriterij se potom procjenjuje na uzorcima signala i pozadine.

Ključne riječi: CMS, PPS, leptokvark, gubitak količine gibanja protona

Rad sadrži: 40 stranica, 25 slika, 6 tablica, 33 literaturna navoda. Izvornik je na engleskom jeziku.

Mentor: doc. dr. sc. Toni Šćulac

Neposredni voditelj: izv. prof. dr. sc. Francesco Santanastasio

Ocjenjivači: doc. dr. sc. Toni Šćulac,
prof. dr. sc. Ilja Doršner,
doc. dr. sc. Marko Kovač

Rad prihvaćen: 10. listopad, 2023.

Rad je pohranjen u Knjižnici Prirodoslovno – matematičkog fakulteta, Sveučilišta u Splitu.

Basic documentation card

University of Split
Faculty of Science
Department of Physics
Ruđera Boškovića 33, 21000 Split, Croatia

Master thesis

Study on the search for leptoquarks in lepton-quark collisions at the LHC using proton tagged events

Petra Akrap

University graduate study programme Physics, orientation Computational Physics

Abstract:

Leptoquarks (LQ), hypothetical particles that turn quarks into leptons and vice versa, are predicted in many theories beyond the Standard Model. This research assesses the potential of the Precision Proton Spectrometer (PPS) subdetector in enhancing the existing inclusive search of the LQ, with a particular focus on single production mechanism, specifically lepton-induced processes. It utilizes proton tag by requiring a scattered proton in case of a LQ event. The analysis is done in two parts. First, a detailed study of leptoquark signal events is done on a generator level, delving into mass distributions of the particle, kinematic distributions of its final states and the cross sections for various LQ types. This is then followed by proton reconstruction test of the PPS, as well as defining a proton tag selection criteria, which is then evaluated on both signal and background samples.

Keywords: CMS, PPS, leptoquark, proton momentum loss

Thesis consists of: 40 pages, 25 figures, 6 tables, 33 references. Original language: English.

Supervisor: Assist. Prof. Dr. Toni Šćulac

Leader: Asoc. Prof. Dr. Francesco Santanastasio

Reviewers: Assist. Prof. Dr. Toni Šćulac,
Prof. Dr. Ilja Doršner,
Assist. Prof. Dr. Marko Kovač

Thesis accepted: October 10th, 2023

Thesis is deposited in the library of the Faculty of Science, University of Split.

Contents

1 Introduction

The Standard Model of particle physics, developed over several decades, gives a solid understanding of the fundamental building blocks of the universe and the forces that dictate their interactions. Many of its predictions have been confirmed successfully through various experiments, yet it is not without its limitations. For example, it does not fully explain neutrino oscillations, their non-zero masses, matter-antimatter asymmetry and gravity, making it an incomplete theory. Nevertheless, it is used as a basis for building more exotic theory models that incorporate hypothetical particles and elaborate symmetries (such as supersymmetry) to tie up loose ends in the Standard Model.

In this analysis, the focus will be on leptoquarks, hypothetical particles that bridge the gap between the fundamental constituents of matter, quarks and leptons. They emerge as significant puzzle pieces in theoretical extensions of the Standard Model, such as Grand Unified Theory, technicolor theories, and compositeness theory, offering a potential look into physics beyond the known framework. Leptoquarks are colored bosons that decay into a quark and a lepton/neutrino. Their mass is expected to be higher than the currently known elementary particles. Experimental interest for such particles dates back to the 1990s when they were first explored at the HERA lepton-proton collider. A renewed interest has grown lately as new experimental findings not explained by the Standard Model appeared, such as bottom quark anomalies and muon g-2 anomaly.

Leptoquarks are currently searched for in ATLAS and CMS experiments at the Large Hadron Collider in CERN. There are various production mechanisms of the leptoquark, and three of them are studied at the colliders: pair production, single production (gluon-initiated) and non-resonant production. In the existing single production searches, basic selection of events is applied by requiring two leptons and jets with high transversal momenta. However, with recent development in computing leptonic parton distribution functions which indicate the probability for the lepton to contain a fraction of proton's momentum, lepton-induced single production mechanism becomes relevant. With a high cross section and clean final state, which is comprised of high-p_t lepton and jet, it gives an enhancement in mass reconstruction and lower background.

A further gain in sensitivity can be achieved by requiring a single proton tag. When two protons collide, there is a possibility that one of them stays intact emmiting a photon which then produces a pair opposite sign leptons of the same flavor. Scattered protons can be detected with a special subdetector in the CMS experiment, the Precision Proton Spectrometer (PPS), specifically designed for such purposes. In this thesis, viability of the PPS subdetector was examined. First, a study of leptoquark kinematics, mass distributions and cross sections were carried out at the generator level. Afterwards, proton reconstruction in PPS is tested and a selection criteria for the LQ signal process is established. The efficiency of the selection criteria

is then studied on both signal and background, assessing the improvement in comparison to a standard inclusive analysis that makes no use of PPS informations.

2 Leptoquarks at the Large Hadron Collider

To study matter at a fundamental level, particle accelerators like the Large Hadron Collider (LHC) are constructed, capable of boosting subatomic particles to very high energies and colliding them in a controlled way. The primary mission of the LHC is testing and validating the theoretical predictions of the Standard Model, probing the properties of established particles, as well as searching for exotic particles predicted by theories beyond SM. So far, it has achieved several breakthroughs, including the discovery of Higgs boson, creation of quark-gluon plasma, observing several hadrons such as tetraquarks and pentaquarks, etc.

This chapter delves deeper into the LHC's workings and provides an overview of physics in colliders. It also shortly outlines the LHC's activities over the years. Subsequently, a detailed description of the central figure in this thesis, the leptoquark, is presented. The leptoquark's characteristics, production mechanisms employed in collider searched, and, more specifically, the lepton-induced single production of the leptoquark will be thoroughly discussed.

2.1 Large Hadron Collider

The Large Hadron Collider is the biggest particle collider ever built, situated at CERN nearby Geneva, Switzerland. It is a circular proton-proton collider at a center-of-mass energy $\sqrt{s} = 13.6$ TeV with a circumference of 27 km. It consists of superconducting magnets cooled to -271.3°C which create a bending magnetic field that guides two proton particle beams each in opposite directions. The particle beams are crossed in four interaction points with four LHC experiments located on each: ATLAS (*A Toroidal LHC Apparatus*), CMS (*Compact Muon Solenoid*), ALICE (*A Large Ion Collider Experiment*) and LHCb (*LHC beauty*). A scheme of CERN accelerator complex is shown in Figure 1.

Before colliding, protons are pre-accelerated at increasing energies before being injected into the LHC. First, protons are extracted from hydrogen and accelerated in the linear accelerator Linac2. Afterwards, they are further accelerated in ring-shaped accelerators Proton Synchrotron Booster (PSB) and Proton Synchrotron (PS). Then, protons are injected into Super Proton Synchrotron (SPS), reaching an energy value of 450 GeV, at which point they're forwarded into the main ring in opposite directions and in separate ultrahigh vacuum beam pipes. In the main ring, each beam reaches the final energy of 6.8 TeV by being injected into electromagnetic fields that are contained in 16 metallic chambers that are known as radiofrequency cavities. Once they're at their maximum energy, they collide at the four interaction points with a maximum instantaneous luminosity of about $2 \times 10^{34} cm^{-2} s^{-1}$.

Luminosity is an indicator of the quantity of collisions and can be assessed from the beam properties. It is proportional to the overlap of the colliding beam bunches moving through each

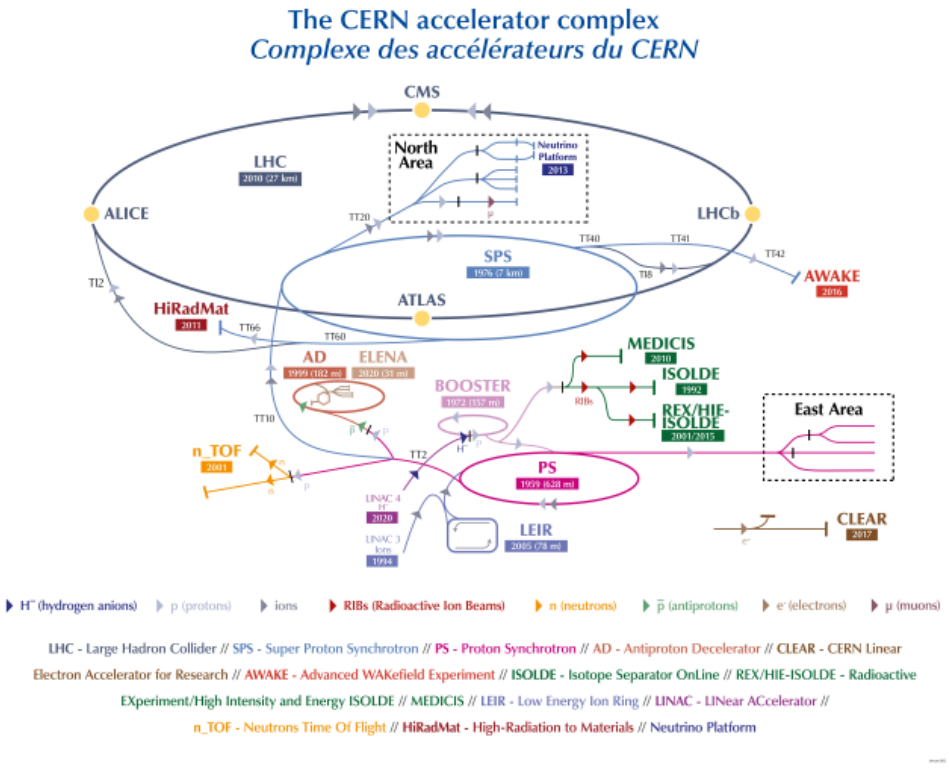


Figure 1: *The CERN accelerator complex [1].*

other. Instantaneous luminosity can be expressed in a simple way as:

$$L = f\frac{N_1 N_2 N_b}{4\pi\sigma_x\sigma_y} \tag{2.1}$$

where N_1 and N_2 are number of particles in the two colliding bunches, f is the frequency of the bunches circulating in the collider and σ_x and σ_y represent the standard deviation of the transverse beam sizes in the horizontal and vertical directions.

The number of expected events, N_{evt}, is the product of the cross section of the hadronic interaction σ_{hadr} and the time integral over the instantaneous luminosity L:

$$N_{evt} = \sigma_{hadr}\int L(t)dt = \sigma_{hadr}L_{int} \tag{2.2}$$

The unit of L is $cm^{-2}s^{-1}$. The integrated luminosity that appears in the equation above, has the same unit of measure of the inverse of a cross section, therefore is measured in inverse barns b^{-1}, where 1 b = $10^{-24}cm^2$.
The higher the luminosity, the higher the pile-up. Pile-up events are pp interactions occurring in the bunch crossings that overlap with the interaction of interest. There are two types of pileup:

In-time-pile-up. pp interactions occurring in the same bunch crossing as the collision of interest;

Out-of-time-pile-up pp interactions occurring in bunch crossings that come just before and after the collision of interest. This effect is due to the signal decay time in the detector that can be comparable with the time between two consecutive bunches.

2.1.1 Kinematics of parton scattering & hadronic cross sections

A sketch of two-to-two parton scattering $p_1 + p_2 \longrightarrow p_3 + p_4$ is presented in Figure 2, where the z-axis is oriented along the direction of the colliding partons.

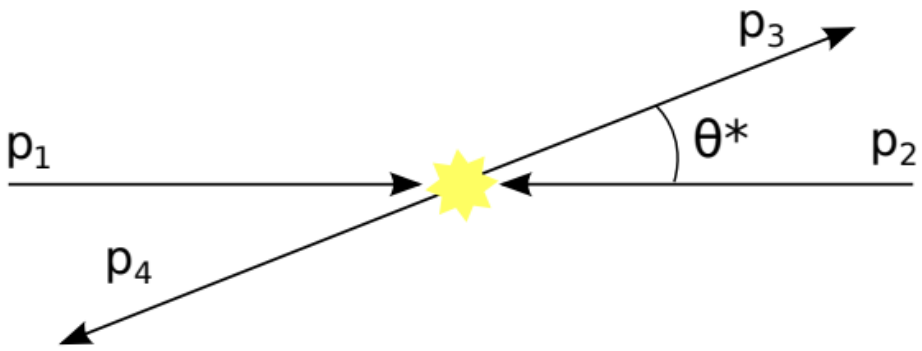


Figure 2: *A sketch of parton scattering [2].*

The energy available for the interaction is determined by the central-of-mass energy $\sqrt{s}$. To compute $\sqrt{s}$, four-momenta of the incoming particle are summed up, represented as **P**, using the respective energy (E_k) and momentum (p_k) of each k-th incoming particle:

$$\sqrt{s} = \sqrt{\mathbf{P} \cdot \mathbf{P}} = \sqrt{(\sum_k E_k)^2 - |\sum_k \vec{p_k}|^2} \tag{2.3}$$

The natural units system is used where the speed of light c=1. The center-of-mass energy remains constant regardless of the reference system, making it an invariant quantity. At the LHC, two protons are accelerated to the same energy level, which largely surpasses their masses. These high-energy protons then collide head-on. Under these conditions, the center-of-mass energy of the entire two-proton system is approximately given by $\sqrt{s} \approx 2E = 2 \cdot 6.8$ TeV $= 13.6$ TeV. At such high energies, the interaction occurs between two subatomic constituents called partons residing within each proton. These partons are typically quarks and gluons. Gluons themselves carry color charge and therefore lead to quantum chromodynamics (QCD) processes that lead to parton showers. Quarks can be categorised into valence and sea quarks. Valence quarks are proton's main constituents that contribute to its quantum numbers. Sea quarks, on the other hand, appear within protons due to quantum fluctuations. For the same reason leptons can also be found inside a proton, as later discussed in Chapter 2.2.1. These partons contribute a fraction to the overall proton's momentum. Consequently, the center-of-mass energy available for the interaction between two partons is generally lower than $\sqrt{s}$. The

fractional momentum carried by a parton inside a proton is referred to as Bjorken-x and is simply denoted as x. Given the values of x for two specific partons (x_1 and x_2), the square of their center-of-mass energy ($\hat{s}$) is linked to that of the proton-proton system by the following:

$$\hat{s} = x_1 x_2 s \tag{2.4}$$

The probability of a parton possessing a specific value of x is described by parton distribution function (PDF, indicated with $f_i(x, \mu^2)$). The PDF enables calculating the probability of encountering a specific type of parton i. This parton carries a fraction x of the momentum at a squared energy scale of μ^2. Figure 3 shows PDFs for two different energy scales, with $xf(x)$ on vertical axis. The valence quark u and d PDFs have a peak at about x = 0.1, while sea quarks along with gluons carry stronger contribution for lower x values. The total cross-section

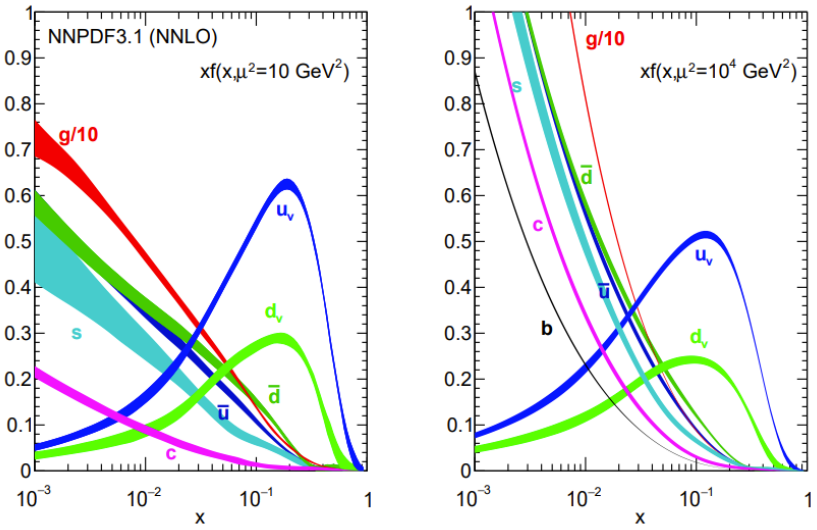


Figure 3: *PDFs for quarks and gluons [3].*

of the hard scattering process can be divided into two components. The first component is a normalization factor that relies on the PDFs. The second component is determined by the parton-parton scattering cross-section, $\hat{\sigma}_{ij}(x_1 p_A, x_2 p_B, \mu^2)$, where p_A and p_B are the momenta of incoming hadrons:

$$\sigma_{hadr} = \sum_{ij} \int dx_1 dx_2 f_i(x_1, \mu^2) f_j(x_2, \mu^2) \hat{\sigma}(x_1 p_A, x_2 p_B, \mu^2) \tag{2.5}$$

2.1.2 History

LHC has first started operating on 10th September, 2008. Run 1, or the first research run, took place from 2010 to 2013 and collided beams up to $\sqrt{s} = 8$ TeV. It resulted with the discovery of the Higgs boson in July 2012. This research run was followed by the first long

shutdown (LS1) which was a two-year detector upgrade enabling collisions at 13 TeV. In 2015 LHC restarted its operation with Run 2 in which a total integrated luminosity of 140 fb^{-1} was collected. The second long shutdown (LS2) took place from 2018 to 2022 with another upgrade and maintenance completed. Currently, the LHC is in Run 3 phase since April 2022 with a maximum center-of-mass energy of 13.6 TeV and an integrated luminosity of 67 fb^{-1}. After the third long shutdown planned from 2026 until 2029, the luminosity is intended to be increased further by upgrading the LHC to the High Luminosity-Large Hadron Collider (HL-LHC). The HL-LHC data taking period will continue for about 10 years with target of collecting 3000 fb^{-1} by 2040.

2.2 Leptoquarks

Leptoquarks (LQs) are hypothetical particles that emerge in extentions of the Standard Model. They are predicted by various Grand Unified Theories, such as the Pati-Salam model [4], and some supersymmetry models. In technicolor models, scalar leptoquarks are anticipated to exist at the TeV energy scale [5]. Additionally, compositeness of quarks and leptons also provide scenarios in which leptoquarks may exist [6]. Recently, anomalies observed in the R_K and R_D ratios [7] in the semi-leptonic *B* decays and in muon g-2 measurement [8] have renewed interested in the models with TeV scale leptoquarks (although the newest measurement of R_K [9] is now compatible with the Standard Model).

LQs couple to a quark and a lepton, leading to a fractional charge which is different depending on the LQ flavor. They carry both lepton and baryon numbers, as well as other quantum numbers that vary among theories. The spin of a LQ state is either 1 (vector) or 0 (scalar). A leptoquark can decay into any of the six SM quarks with any of the three charged leptons or three neutrinos. LQs are often referred to being of the first, second or third generation type, depending on the generation of the charged lepton and the quark it couples to. For example, if a LQ couples to tau and top/bottom quark, it's regarded to be of third generation type. However, cross-generational mixing is possible with model dependent constraints.

In this work, scalar LQ models with a 100% branching ratio in one type of charged lepton and one type of quark are considered, independent on the generation type. There could exist LQs with different masses, but for simplicity it is assumed that only the lightest one can be probed at LHC energy levels. In the limit of large scalar LQ masses, the decay width is given by

$$\Gamma_{scalar} = \frac{\lambda^2 M}{16\pi} \tag{2.6}$$

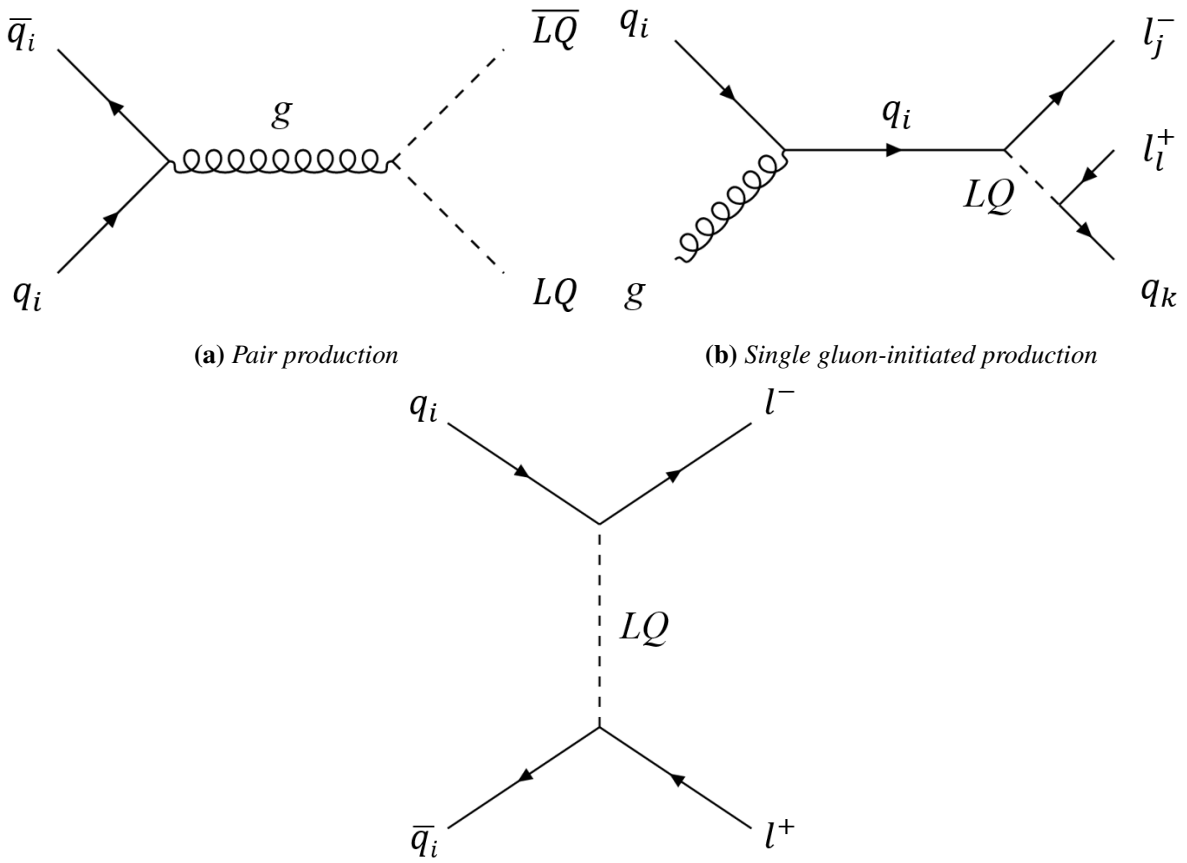
where λ is the coupling constant. For $\lambda = 2$, the relative width is $\frac{\Gamma}{M} \approx 8\,\%$, resulting in relatively narrow resonances.

Given that LQs are colored, they can be copiously produced at hadron colliders. There are many searches at hadron colliders for all three LQ generations using various production mechanisms. In Figure 4a, a Feynman diagram of scalar LQ pair production due to the quark-antiquark annihilation ($q\bar{q} \longrightarrow LQ\bar{LQ}$) is depicted. Such process has a large cross section and is independent on the coupling constant. However, for higher masses it gets harder to produce pairs of LQs and the cross section reduces compared to other processes. In such case, single production mechanism is more suitable for LQ searches, which has cross section proportional to λ^2. Figure 4b shows said mechanism. The subsequent decay of the LQ production yields one jet and two leptons in final state, with the second lepton having a typically high p_t. Along with resonant production mechanisms, non-resonant ones in *t*-channel Drell-Yan process (Figure 4c) also carry important contributions in LQ searches, as their cross section scales as λ^4, giving

		e-channel			
Single production (gluon)	M_{LQ}<1755 GeV	19.6 fb^{-1}	CMS	8 TeV	[10]
Pair production	M_{LQ}<1435 GeV	35.9 fb^{-1}	CMS	13 TeV	[11]
Pair production	M_{LQ}<1800 GeV	138 fb^{-1}	ATLAS	13 TeV	[12]
		μ-channel			
Pair production	M_{LQ}<1810 GeV	138 fb^{-1}	CMS	13 TeV	[13]
Pair production	M_{LQ}<1700 GeV	138 fb^{-1}	ATLAS	13 TeV	[12]
Single production (gluon)	M_{LQ}<660GeV	19.6 fb^{-1}	CMS	8 TeV	[14]
		τ-channel			
Single production (gluon)	M_{LQ}<800 GeV	138 fb^{-1}	CMS	13 TeV	[17]
Pair production	M_{LQ}<1490 GeV	138 fb^{-1}	ATLAS	13 TeV	[16]
Pair production	M_{LQ}<1220 GeV	138 fb^{-1}	CMS	13 TeV	[17]
Single production (lepton)	700 GeV<M_{LQ}<900 GeV (λ = 2)	138 fb^{-1}	CMS	13 TeV	[18]

Table 1: *Recent LQ mass limit results at the LHC. All gluon-initiated single production results are obtained for* $\lambda = 1$.

bigger cross sections for large λ. Current mass limit results at the LHC for *e*-, μ- and $\tau-$ channels are given in table 1. LQs with masses below the 1 TeV scale are essentially excluded by all previous searches.



(a) *Pair production*

(b) *Single gluon-initiated production*

(c) *Non-resonant production*

Figure 4: *Different leptoquark production mechanisms.*

2.2.1 Lepton-proton collisions

Presenting itself as an alternative in collider search is resonant lepton-induced single production. Due to quantum fluctuations in the proton, presence of charged leptons is allowed. In this process, the collision consists of a lepton from one proton and a quark from the other, resulting in resonant production of a leptoquark state as shown in Figure 5. This production mechanism became recently relevant thanks to precise calculations of leptonic parton distribution functions (PDF) inside the proton [20]. As mentioned earlier, the gluon-initiated LQ single production process yields a second lepton in its final state with a high transversal momentum, which is a drawback as it is hard to tell which of the two detected leptons is actually coming from the produced leptoquark. The lepton-induced process instead gives a clean final state made up of leptons and one jet in final state. While most of the energy is in the leading lepton and jet which makes them easy to pair to make a LQ candidate, the second lepton coming from the pair production produced by the emitted photon is low-p_t and often out of tracker acceptance due to large η.

Like in the gluon-induced process, the leptoquark appears in the detector as a resonance in the invariant mass of a lepton and a quark jet. A feasibility study was done by theorists to estimate the sensitivity of this channel compared to existing analyses [21]. Figure 6 displays a 95% confidence level (CL) limits on the magnitude of the $\lambda_{\mu u}$ couplings as a function of mass for scalar LQs. The black, yellow and green lines represent pair production, single production and DY non-resonant production limits respectively, while the red and orange ones correspond to the search using lepton-initiated mechanism for different luminosities. With respect to the existing searches, the lepton-initiated one predicts higher sensitivity at high mass and high coupling (M > 1.5 TeV & λ > 0.2 for integrated luminosity of 36 fb^{-1}).

Another big advantage of this process that leaves one proton intact is that it allows for proton tagging, thus adding a low-background event category which can potentially further improve sensitivity. This could be done with a CMS subdetector described in Chapter 3. The focus in this work will be the specific case in which one proton is scattered, emitting a photon, as shown in Figure 5. The current estimation of the proton survival probability is affected by theory uncertainties as there hasn't been a dedicated study addressing this issue.

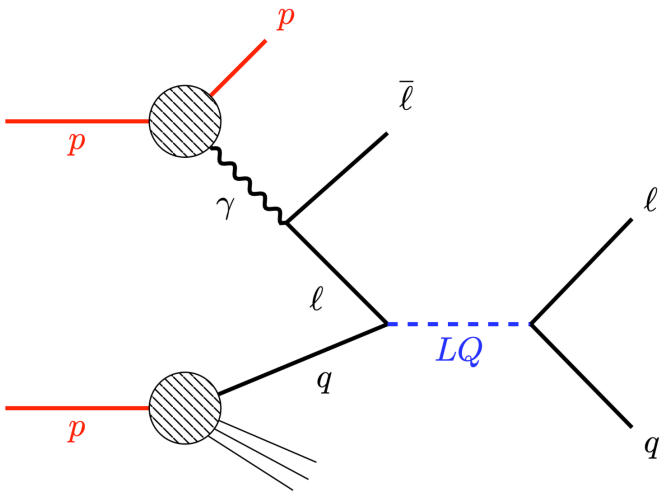


Figure 5: *Resonant lepton-induced single LQ production in pp collisions [21].*

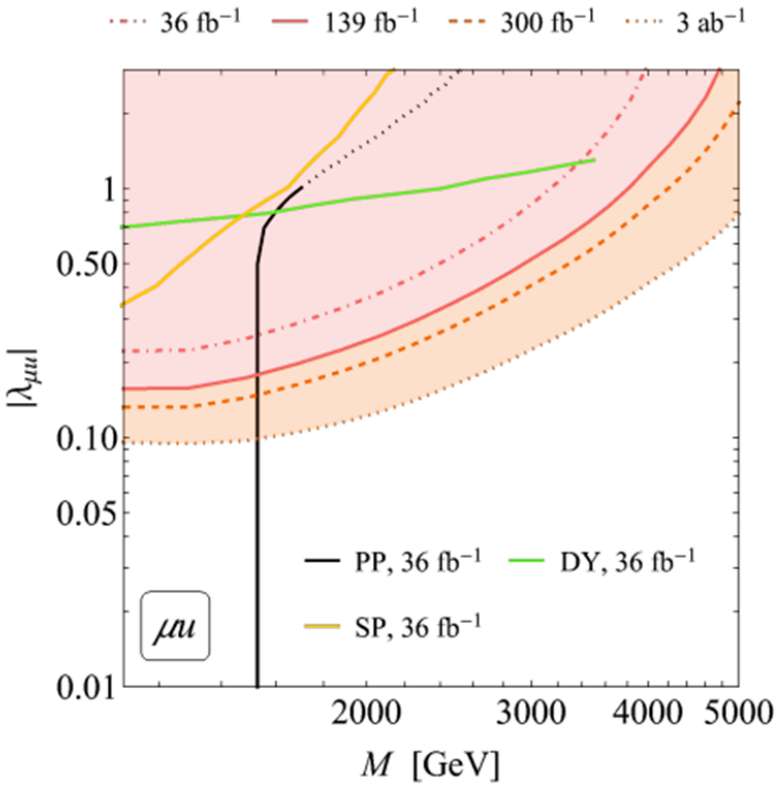


Figure 6: *95% CL limits on the parameter space of LQ in case of μu coupling based on LHC Run 2 [21].*

3 CMS Experiment

The Compact Muon Soleonid, or simply CMS, is a multi-purpose detector. It is designed to be able to detect most particles, resulting in an onion-like structure, with each layer specialised in different measurements. A display of layers is shown in Figure 7.

Once a hard interaction between two partons in the primary interaction vertex takes place,

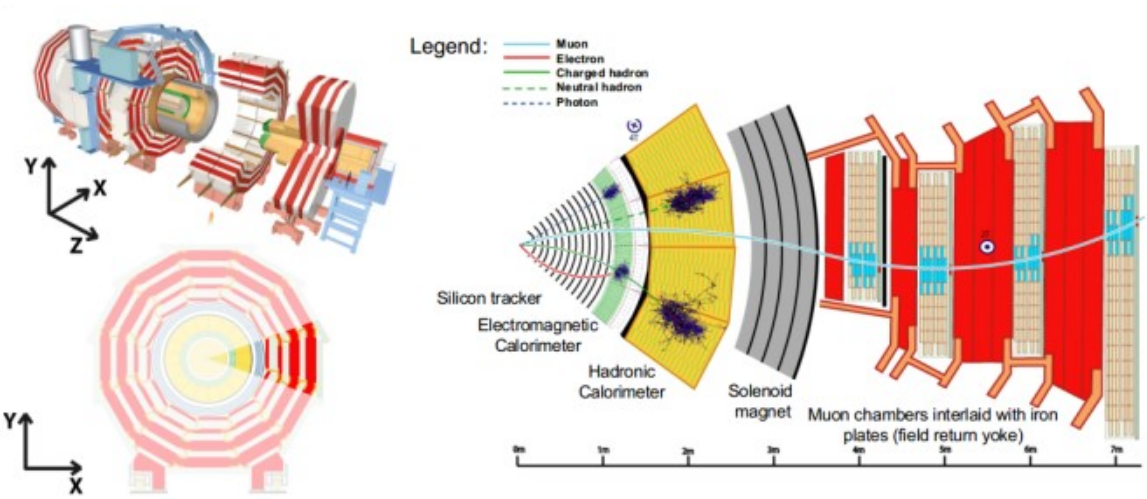


Figure 7: *A view of the layered structure of the CMS [22].*

many different outgoing particles burst. They encounter the following components, starting from the interaction point:

Tracking system. The CMS tracker system has a pixel vertex detector and a tracker. It is responsible for trajectory reconstruction of charged particles produced in pp interactions along with interaction points. It is made out of several layers of silicon detectors that identify positions of the particles passing through them, thus calculating their momentum from their tracks. In order to achieve high granularity to separate close tracks, the inner layer is formed by very small pixel detectors. The total area of silicon detectors covers up to $|\eta| = 2.4$.

Electromagnetic calorimeter. The electromagnetic calorimeter (ECAL) is made out of a cylindrical inner part (barrel) and two disks on the ends (endcaps) which are covered with lead-tungstate ($PbWO_4$) scintillating crystals. Its main objective is to measure energy of incoming electrons, positrons and photons with a high resolution. The crystals are high in density and produce light in short and fast bursts that allow for a precise detection. The light is then sent to photodetectors and signal is read as a measure of the particle's energy.

Hadronic calorimeter. Similar to the ECAL, the hadronic calorimeter (HCAL) measures energy of hadrons produced in pp collisions (as well as the missing transverse energy in combination with ECAL). It is formed by alternating layers of absorbing material and fluorescent scintillator material. The HCAL needs to contain and measure large cascades of particles that are produced once a hadron hits the absorber material. In order to fit such a massive part into CMS, it's organised into barrel, endcaps and forward sections. The forward sections are positioned at either end of CMS to pick up particles coming out of the collision at shallow angles relative to the beam line.

Superconductive magnet. The superconducting Niobium-Titanium solenoid magnet is a crucial part of the CMS, with a 3.8 T uniform magnetic field. It bends the paths of particles that emerge from pp collisions. The less a particle's path is curved by the field, the more momentum it has. The magnetic field is returned by a saturated iron yoke, which is also mechanical support of the detector.

Muon system. Muons have a separate detector as they are not absorbed by any of the calorimeter detectors. A particle is measured by fitting a curve to the hits registered in the four muon stations, which are outside of the magnet coil. The path is measured by tracking its position through the layers of each station. The reconstruction of muon tracks involves aligning data from both the muon system and the tracker to form a trajectory characterized by two opposing curves in the transverse plane. This kind of trajectory serves as a sign for identifying the muon. The muon system is built of Drift tube chambers (DT), Cathode strip chambers (CSC) and Resistive plate chambers (RPC).

Due to massive amount of pp interaction readouts, a trigger system is developed to save potentially interesting events and discard others, as saving all events would not be necessarily useful since most of them are soft collisions. The maximum amount of data that can be recorded is $\sim$ 2 GB/s. The Trigger system reduces the event rate from 40 MHz to $\sim$ 1 kHz (equivalent to an average event size of $\sim$ 2 MB). It is organized in two levels:

Level-1 Trigger. The first trigger that the signal encounters is Level-1 (L1) which reduces the initial event rate to $\sim$ 100 kHz. It features several algorithms to store a general description of the event that are passed to the HLT for further processing, if the event is accepted. The L1 is made by custom hardware processors designed to perform logical operations on the signals generated by the subdetectors.

High-Level Trigger. The High-Level Trigger (HLT) is a software system made out of a set of algorithms which select specific events. It reduces output rate to about 1 kHz.

3.0.1 The Particle Flow algorithm & jet reconstruction

A "jet" in particle physics refers to a spray of particles that originates from a quark or a gluon. The quarks and gluons carry a color charge and need to recombine into color-neutral hadrons. The process of this recombination, known as hadronization, leads to the production of many particles. These particles subsequently lose their energy through a cascade of decays within the CMS calorimeters. The ECAL and HCAL measure the energy released by the showering particles. The silicon tracker detects charged particles and provides measurements of their transverse momentum.

To identify and reconstruct the various particles effectively, the CMS employs the Particle Flow (PF) reconstruction algorithm, which utilizes information from all these sources. The PF algorithm first collects reconstructed hits in each subdetector and creates a list of basic reconstructed elements. These elements are then matched topologically to build PF particle candidates (electron, hadrons, photons, etc.). This algorithm attempts to mirror the true particle composition of the event as best as possible. All the information on the momentum and the energy of the PF candidates are saved and available for analyses.

To form jets, the PF candidates are grouped using jet clustering algorithm. The standard used at the LHC is the anti-k_t algorithm [23] with distance parameters R = 0.4 (AK4) or R = 0.8 (AK8) that indicate the size of the jet. This method combines two particles if a suitable defined distance, which is a function of the inverse of the transverse momentum k_t, is less than a given threshold. The high-energy particles are clustered first, and then the softer ones are added in decreasing energy. The four-momenta of the reconstructed jet would be ideally equal to that of the originating parton, but the energy and momentum of jets have to undergo corrections due to various effects. These effects are the inclusion of particles not coming from the primary interaction (pile-up) into the jet and the non-uniform jet energy response across the detector.

3.1 Precision Proton Spectrometer

Since 2016, a CMS subdetector called Precision Proton Spectrometer (PPS) has been measuring protons scattered at small angles which carry between about 84 to 97% of the incoming beam momentum [24]. PPS is a a near-beam magnetic spectrometer that uses the LHC magnets about 210 m away from the interaction point, on both sides of CMS. Since the scattered protons have lost a small fraction of their initial momentum, they're bent out of the beam path and their trajectories are measured, specifically the displacement of the proton with respect to the beam line. The higher the displacement, the larger is the loss of proton's momentum. The detector planes are inserted horizontally into the beam pipe by means of movable near-beam devices ("Roman Pots" or RPs) that allow the planes to be brought to only a few milimeters away from the beam without affecting any aspects of accelerator operation.

A scheme of the beam line on one of the two sides of CMS (usually referred to as "arms") where the RPs are positioned is shown in Figure 8.

Left arm corresponds to positive z direction and to LHC sector 45, while the right arm corresponds to negative z (LHC sector 56). The tracking RPs are each equipped with 10 planes of silicon strip sensors with a spatial resolution of about 12 μm and each plane contains 512 individual strips.

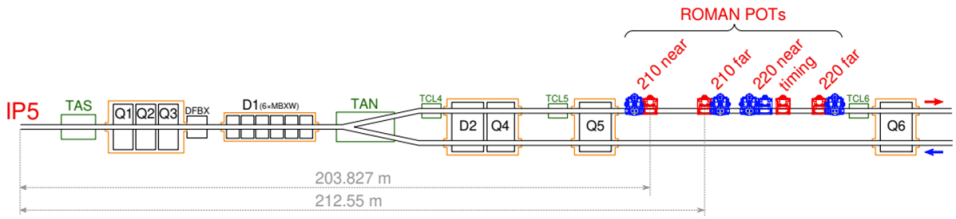


Figure 8: *A layout of the beam line between the interaction point (IP5) and the RPs in LHC sector 56 [24].*

3.1.1 Proton Reconstruction

Figure 9 depicts protons arriving at the PPS detectors after being warped by the magnetic fields of the accelerator. The detectors measure x,y positions and then from there find the fractional momentum loss ξ and four-momentum transfer t of the proton at collision point. The scattered

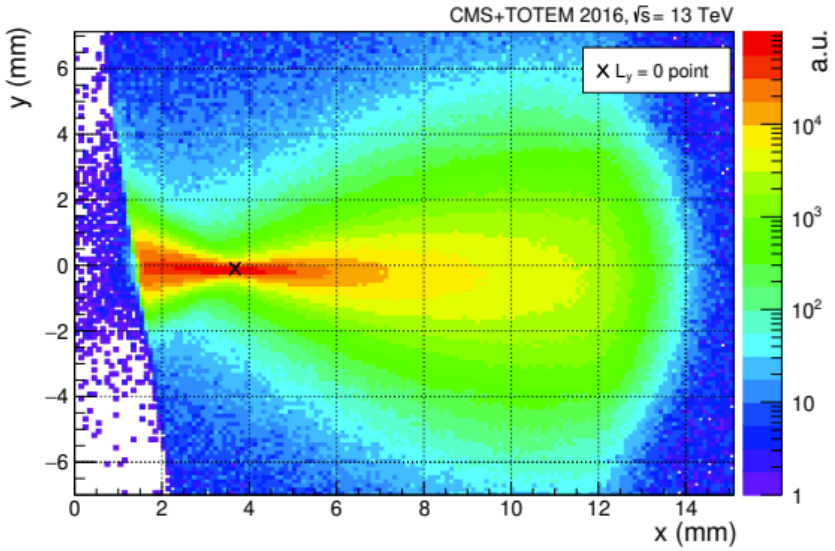


Figure 9: *A hit map of track impact points measured in a RP positioned in sector 45. The beam center is at x = y = 0 [24].*

proton reconstruction with PPS has two elements:

Alignment. PPS alignment is essential for determining the position of the sensors with respect to each other inside a RP, the relative position of the RPs and the overall position of the spectrometer with respect to the beam.

Optics. Accurate reconstruction of the scattered proton momentum based on tracked data from the RPs relies on a precise understanding of the magnetic fields encountered by the protons as they travel from the IP to the RPs. This is described in terms of "beam optics", in which the components of the beamline are treated as optical lenses. The proton trajectory is characterized using transport matrices, which transform the initial kinematics of protons scattered at the IP into the kinematics measured at the RP positions.

The trajectory of the protons coming from the IP is described by $\vec{d*} = (x*, \Theta_x*, y*, \Theta_y*, \xi)$, and is transformed into the kinematics detected at the RP approximately by:

$$\vec{d} = T(s, \xi)\vec{d*}, \tag{3.1}$$

where *s* is the distance from the IP along the beam orbit, $T(s, \xi)$ is the transport matrix whose elements are optical functions and ξ is the proton momentum loss. (x, y) represent transverse vertex position, while (Θ_x, Θ_y) are vertical components of the scattering angle. In the horizontal plane, the leading term is

$$x = D_x(\xi)\xi, \tag{3.2}$$

and in the vertical it is

$$y = L_y(\xi)\Theta_y* \tag{3.3}$$

$D_x(\xi)$ is the dispersion which mildly depends on ξ and $L_y(\xi)$ is the vertical effective length. At any location *s* in the RP region there is a value of ξ (ξ_0) where $L_y = 0$, resulting in a 'pinch' of track impact points distribution at $x_0 \approx D_x\xi_0$. From there, D_x is estimated as $D_x = \frac{x_0}{\xi_0}$. ξ is reconstructed by simply inverting Eq. 3.2. The subleading terms in the proton transport are ignored and instead included in the systematic uncertainties.

Two proton track reconstruction algorithms are available, singleRP and multiRP. SingleRP method simply treats each RP as a separate detector, reconstructing tracks locally in each, which gives simpler derivation and systematics and maximizes efficiency/acceptance. However, this method does not give full proton kinematics and the resolution degrades for larger proton angles. MultiRP, which is used in this analysis, combines measurements of both tracking RPs in a given arm if both have a singleRP track reconstructed and the tracks align in terms of their coordinates. The result is a global track, with full proton kinematics at IP and much better resolution/uncertainties. A downside is some loss of efficiency due to radiation damage, acceptance difference, etc.

3.1.2 Proton kinematics matching

Assuming a central exclusive process as shown in Figure 10, the scattered proton kinematics are the following:

$$\xi = 1 - \frac{|p_f|}{|p_i|} \tag{3.4}$$

$$M_X = \sqrt{s\xi_1\xi_2} \tag{3.5}$$

$$y_X = \frac{1}{2}\log(\frac{\xi_1}{\xi_2}) \tag{3.6}$$

ξ is the fractional momentum loss of a particle, p_i and p_f are proton's initial and final momentum respectively, M_X is the mass and y_X is the rapidity of the final state X. Assuming

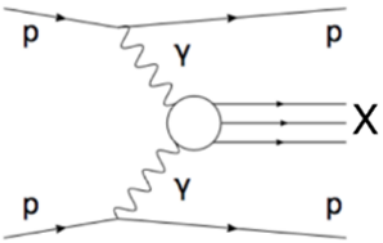


Figure 10: *Feynman diagram of a central exclusive process [25]*

both protons are scattered (Figures 10 & 11), PPS can directly measure ξ for both protons. However, if the kinematics of the final state X is known, then ξ can be calculated by simple kinematics calculations derived from energy and momentum conservation laws. This can be measured instead by CMS:

$$\xi_1 = \frac{E_X + P_{X,z}}{2P_{beam}}$$
$$\xi_2 = \frac{E_X - P_{X,z}}{2P_{beam}} \tag{3.7}$$

where E_X is the energy of the produced particle X and $P_{X,z}$ is its longitudinal momentum (see Appendix A for calculation). If the two ξ measured in both PPS and CMS match, it suggests that there was a diffractive scattering. However, usually there are multiple reconstructed protons in PPS primarily due to the presence of pile-up interactions. They often come from diffractive processes that overlap with the hard scattering event of interest.

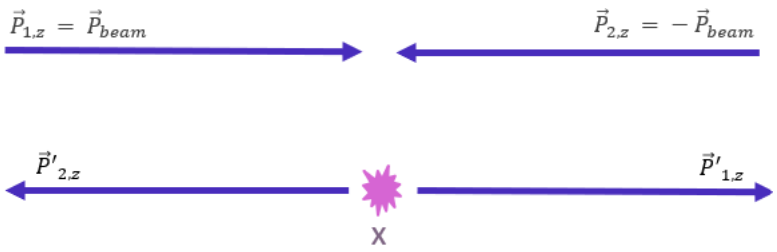


Figure 11: *Schematic view of two protons colliding in CMS and producing particle X, assuming they are scattered afterwards.*

4 Generator level studies for LQ signal

4.1 Leptoquark event generation

The first task of this work is to study and properly understand lepton-induced resonant leptoquark production at NLO. For this purpose, a Monte Carlo tool is used to generate such events. The event generation consists of two steps, with the first one being leptoquark generation and its subsequent decay and the second being parton showering.

The first step is done with POWHEG [26]. It can handle all leptoquark flavor combinations with different masses and coupling constants [27]. The output is given in the form of a Les Houches event (LHE). Each particle is assigned with an ID number in accord to the PDG particle numbering scheme [28]. Along with the ID, each particle has its own four-vector and mass values and the event is denoted with its cross section.

The obtained LHE file is later being processed by HERWIG which is a parton showering program that gives the complete picture of the process [29, 30]. For now, only HERWIG can handle lepton-induced processes such as the one discussed in this work [27]. The final result are ROOT files ready for analysis. Later on, these files are passed to the full CMS simulation thorugh *Geant* [31], which is then used for further studies described in Chapter 5.

In the final list of generated particles, there is no final proton, i.e. proton after the scattering happened. To obtain the four-vector of the final state proton, the initial proton and emitted photon from the scattering (Figure 5) are identified in the list of generated particles. From there, the four-momentum is calculated by simply subtracting the four-momenta of the initial proton and the emitted photon. In the MC generator used for this analysis, events where there is an emitted γ is about 70%. Hence, the assumption is that in 70% of the events, a scattered proton will be present. However, this percentage is subject to large theory uncertainties and the final results heavily rely on it. That being said, a rescaling factor can be easily applied to provide results with a different assumed probability.

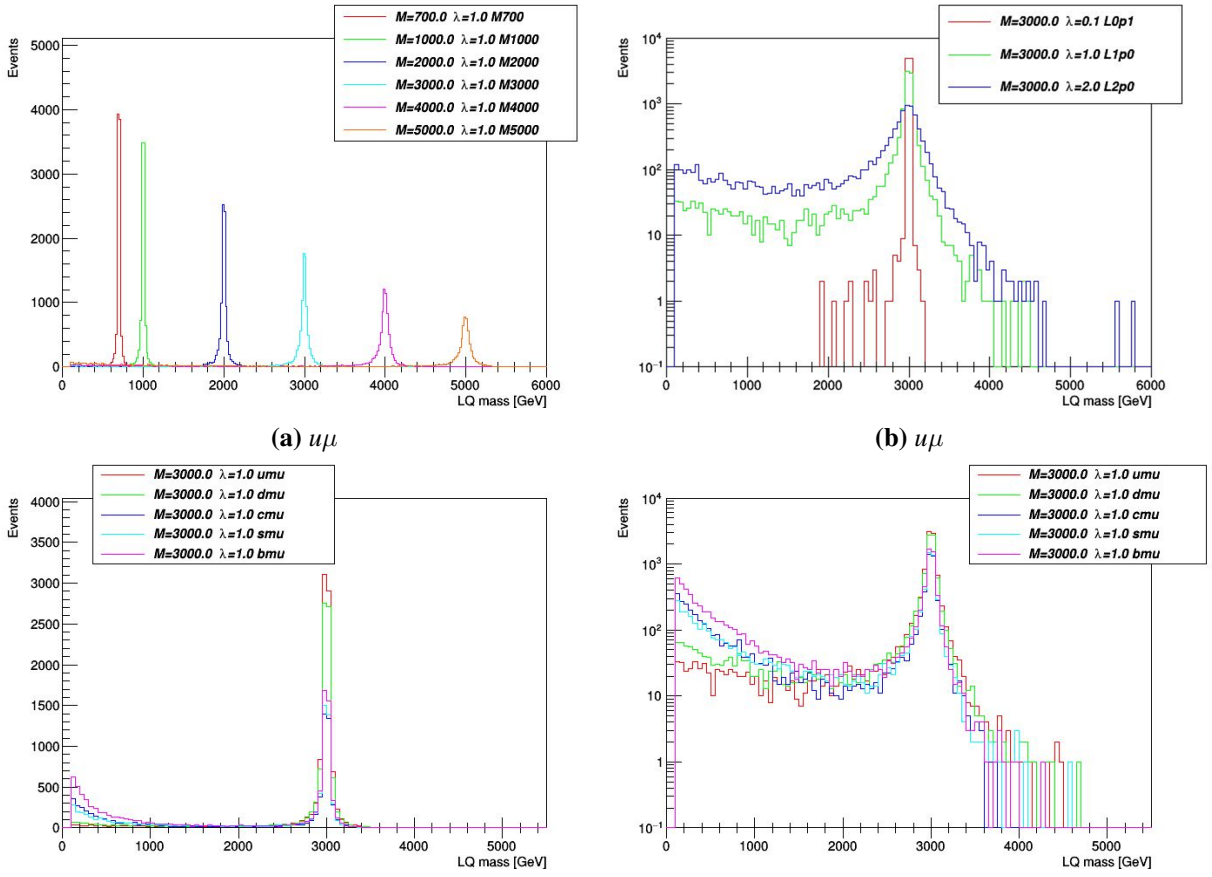
4.2 Kinematic distributions

In order to understand the kinematics of the leptoquark process, different leptoquark models are generated, with varying masses and coupling constants. Much of the focus are final states with electrons and muons.

4.2.1 LQ mass distributions

In Figure 12, distributions of the LQ generated mass are shown for different nominal LQ mass, coupling and flavor of initial and final state leptons and quarks. In general, the LQ mass

distribution along with a Breit-Wigner signal peak also has a tail at lower mass. The reason is due to the convolution of the PDF (manifesting itself as a tail) and the cross-section of a hard-scattering process (manifesting itself as a peak) as described in Eq. 2.5.
In Figure 12a, it is depicted how the signal peak as well as its width change with increasing nominal LQ mass. This is due to the fact that width scales with mass (Eq. 2.6). Another factor is that for larger mass there's higher probability to have events at lower masses, far from the nominal ones. When looking at distributions depending on coupling (Figure 12b), it can be further seen how the width is consistently increasing along with λ^2 (Eq. 2.6). In Figure 12c and d, mass distribution for different quark flavor is shown. As expected, due to their differing PDFs, valence quarks (up and down) have a more prominent peak than sea quarks (charm, strange and bottom). On the other hand, at lower energies the opposite is occurring, as the off-shell production is more dominant for the sea quarks and less so for valence quarks.

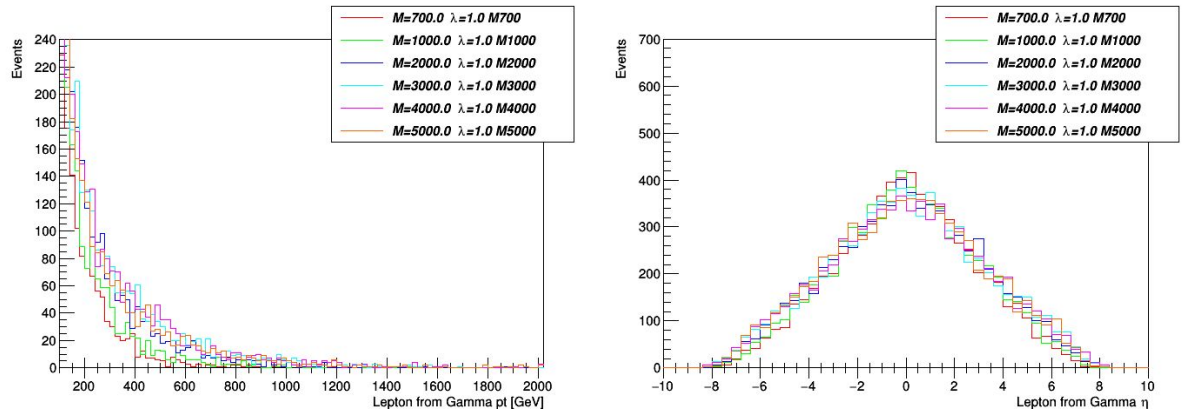


(a) $u\mu$ **(b)** $u\mu$

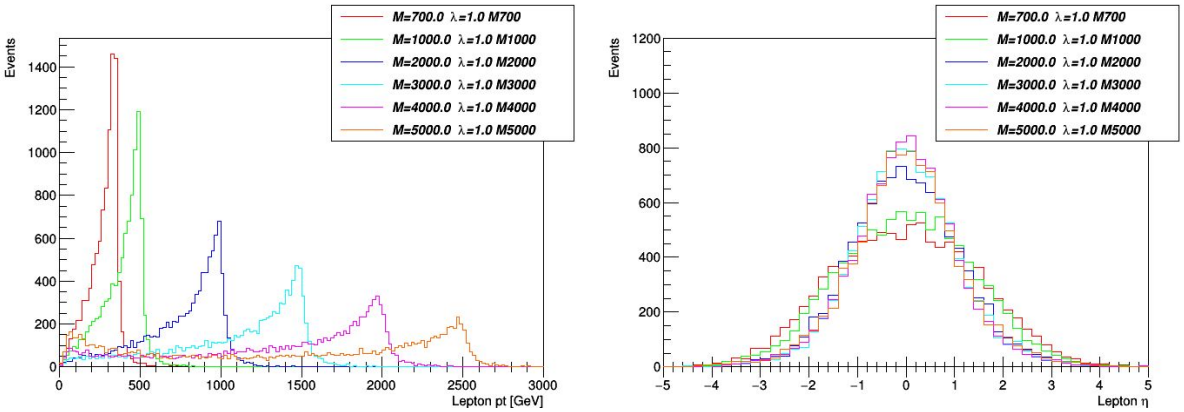
(c) *different quark flavors* **(d)** *different quark flavors in logarithmic scale*

Figure 12: *Leptoquark mass distributions at generator level.*

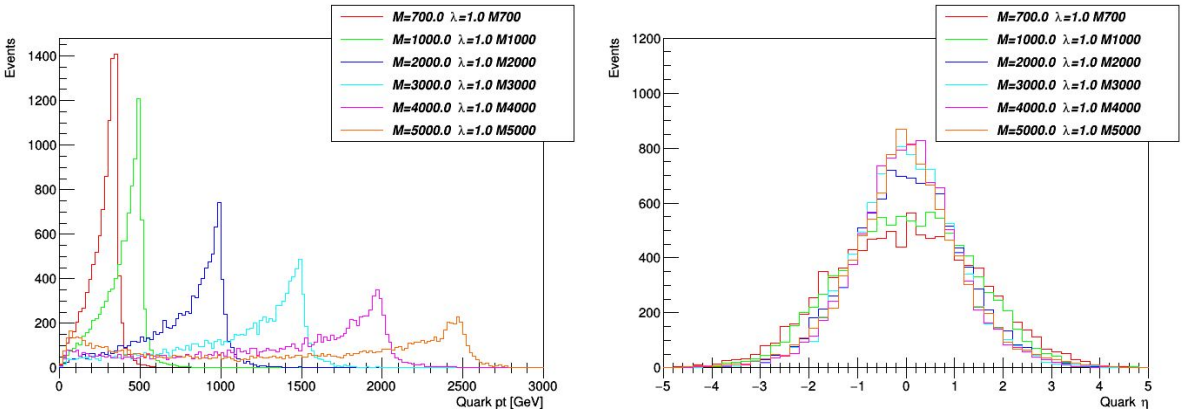
4.2.2 Kinematics of final state particles



(a) *Transversal momentum of lepton emitted from photon* **(b)** *Pseudorapidity of lepton emitted from photon*



(c) *Transversal momentum of a lepton emitted from leptoquark* **(d)** *Pseudorapidity of a lepton emitted from leptoquark*



(e) *Transversal momentum of a quark emitted from leptoquark* **(f)** *Pseudorapidity of a quark emitted from leptoquark*

Figure 13: *$u\mu$ leptoquark distributions at generator level for $\lambda = 1$*

In Figure 13, final states of LQ kinematics for $u\mu$ flavor with a fixed coupling constant $\lambda = 1$ are shown. As seen in Figure 5, from the scattered proton, a photon is emitted which then produces a pair of leptons. In Figures 13a&b it is shown that the final state lepton coming from γ carries a low transverse momentum and that it also has a high pseudorapidity which is often out of the tracker's acceptance of $|\eta| < 2.4$. The other lepton from the pair production produces

a leptoquark with a quark coming from the fragmented proton, which then decays into final state lepton and quark, both mostly within detector's acceptance (Figures 13d&f). Given that sometimes the lepton from γ won't be detected, it is fitting then to have two signal categories in the analysis, corresponding to the number of detected leptons. The lepton from the LQ decay has a substantially higher p_t than the one from γ as seen in Figure 13c, making it much more simpler to identify the leading lepton compared to gluon-initiated process. As expected, it is shown in Figures 13c&e that the final state lepton and quark from the decay have a transversal momentum which is roughly half of LQ's mass.

These results are identical for LQs with *ue* final state. Even though $u\mu$ and *ue* flavored LQs have a different initial and final state leptons, the kinematics stays the same as the difference between their lepton PDFs isn't as notable as compared to quarks (Figure 3). After all, the only difference between the leptons are their masses, which are very small compared to proton energy.

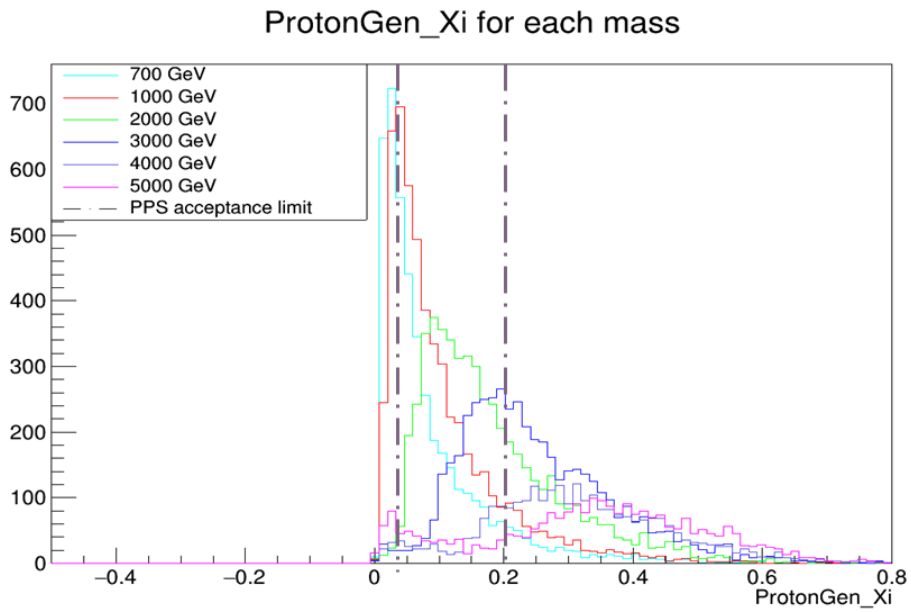
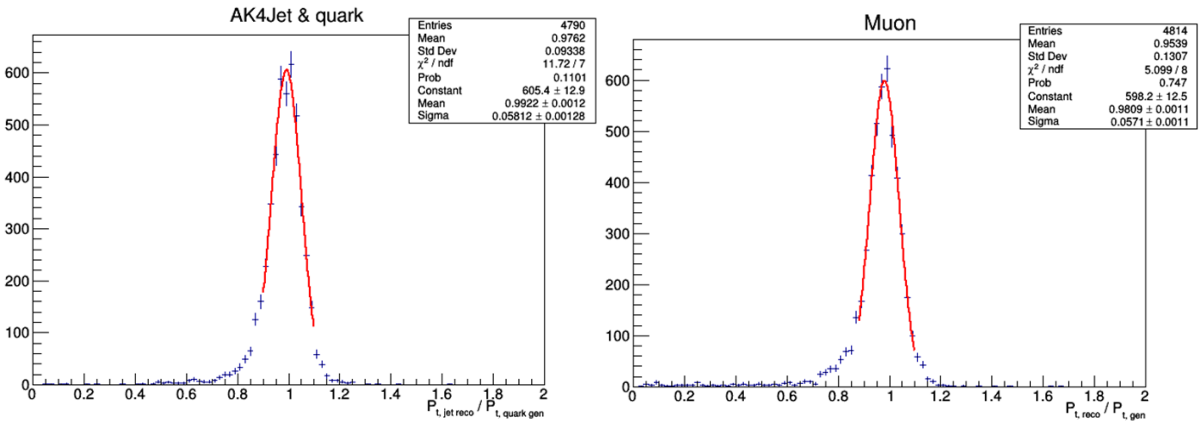


Figure 14: *ξ distribution for different LQ masses on generator level.*

Figure 14 shows distribution of ξ for each generated LQ mass. The PPS acceptance ranges from about 0.04 to 0.2, and a significant fraction of protons fall in PPS acceptance, making it possible to measure them.

The resolution of particle reconstruction in CMS is tested by simply plotting the ratio of reconstructed quantity and generated one, as shown in Figure 15. Considered are the lepton and the quark combination from LQ decay. For the quark match, the AK4 jet is taken into account. To obtain the mean and the standard deviation, a gaussian function is fitted on the peak of the distribution. The mean is in both positioned roughly around 1, with a standard deviation of about 6%, indicating a well made reconstruction.



(a) *Reconstructed AK4 jets and generated quarks.* **(b)** *Reconstructed and generated muons.*

Figure 15: *Ratio of transversal momenta of reconstructed final state particles and generated particles for* $m_{LQ} = 3$ *TeV.*

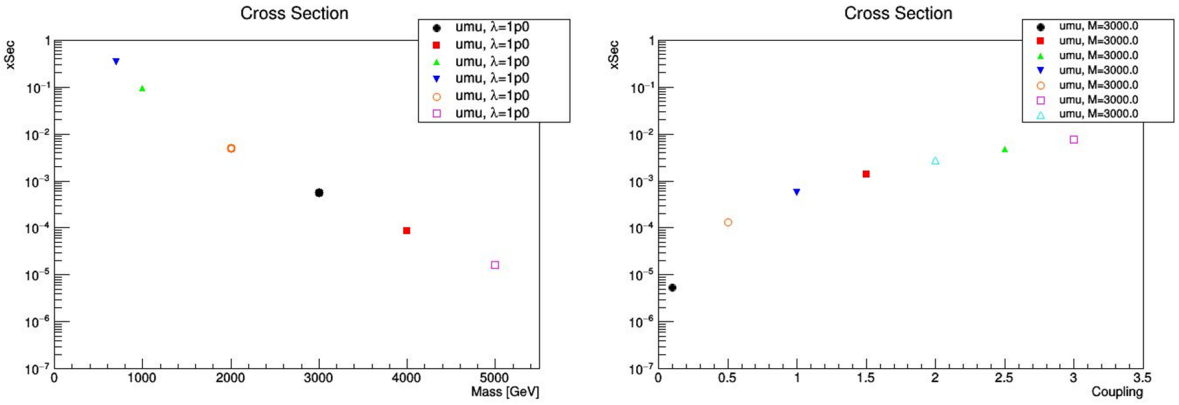
4.3 Cross section of generated samples

Along with before mentioned distributions, the cross section of generated samples was also studied. In Table 2, cross sections for different leptoquark flavors are listed. As anticipated, it's highest for valence quarks, while for sea quarks it's declining with higher quark mass. Up quark is the highest out of the valence quarks due to the fact that inside a proton there are two up and one down quark. As the sea quarks, the reason for a steep decline in values is the fact that $m_s < m_c < m_b$, making it less probable for them to interact as mass increases.

$m_{LQ} = 3$ **TeV,** λ**=1**	σ **[fb]** $\pm$ **5% [27]**
uμ	0.572
dμ	0.242
sμ	0.0419
cμ	0.0251
bμ	0.0123

Table 2: *Cross section depending on LQ flavor*

Figures 16a and 16b show expected behaviors for σ, an increase with coupling and a decrease with mass. It's worth comparing cross sections of lepton-initiated to gluon-initiated processes (diagrams shown in Figures 5 and 4b respectively). As seen in Figure 17, the lepton-initiated process surpasses gluon-initiated one for masses above 1 TeV. For example, at 3000 GeV has up to 30 times larger cross section than the gluon-initiated one, giving similar result for different λ and e, μ and τ initial states [27].



(a) *σ as a function of mass at $\lambda = 1$*

(b) *σ as a function of coupling at 3 TeV.*

Figure 16: *Cross sections (in pb) of uμ LQ production.*

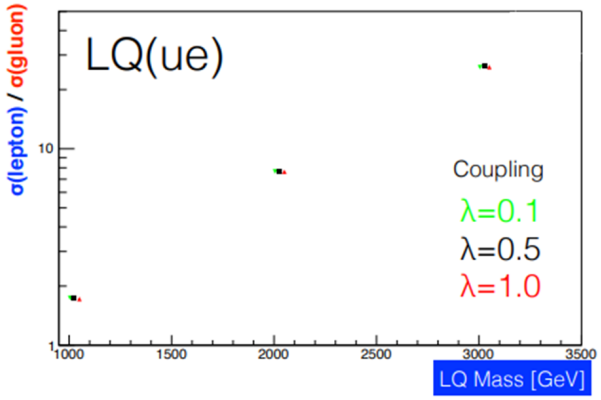


Figure 17: *Cross section ratio of lepton-initiated and gluon-initiated leptoquark production processes for different couplings and masses.*

5 Feasibility study

Before describing the work done in this thesis, it should be remarked that there is an ongoing CMS inclusive analysis for LQs in which a peak in the invariant mass spectrum of the lepton-jet system is searched for, ignoring the possible presence of scattered protons. In such analysis, the signal region is divided in two categories, depending on the number of leptons detected (1 or 2 leptons). The final states includes very high p_t objects in TeV scale. Let's assume that the LQ signal is uμ. The full event selection is presented in Table 3, along with few signal versus background plots in Figure 18. In 2μ category, there's an additional demand that the mass of the two muons don't surpass 110 GeV, to make sure that the background events with a Z boson produced in association with jets are excluded. Furthermore, one of the requirements involves the missing transverse energy (*met*), a quantity that is attributed to undetected neutrinos.

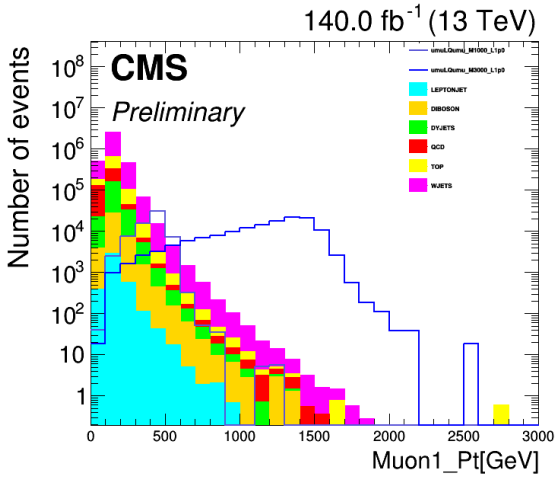
Once the full selection is applied, a search for a bump in the invariant mass distribution of leading muon and leading jet is done.

The goal of this work is to study the effect of adding another category to this one, made of

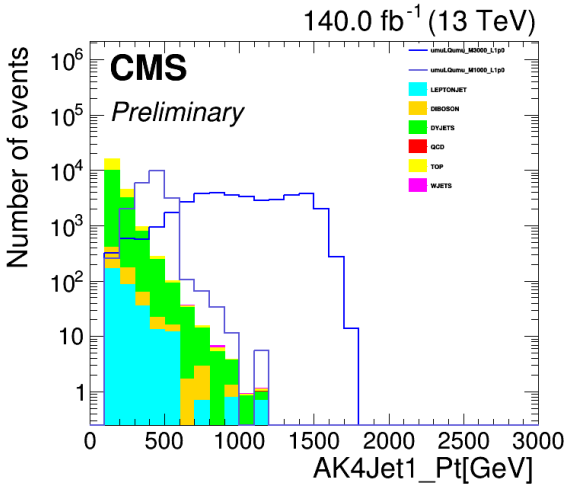
Full event selection	
pass HLT - Single Muon trigger with $p_t > 50$ GeV	
TrkIso < 0.05	
$p_t^{\mu} > 55$ GeV	
$p_t^{jet} > 100$ GeV	
$m_{\mu jet} > 300$ GeV	
$\lvert\eta\rvert < 2.5$	
$\frac{met}{\sqrt{\sum E_t}} < 4$	
$0.8 < \frac{p_t^{\mu}}{p_t^{jet}} < 1.4$	
$\frac{p_t^{\mu}}{m_{\mu jet}} > 0.2$	
One μ category	Two μ category
$\Delta\phi_{\mu jet} > 3.05$	$\Delta\phi_{\mu jet} > 2.8$
	$m_{\mu\mu} > 110$ GeV

Table 3: *Full signal event selection for both categories.*

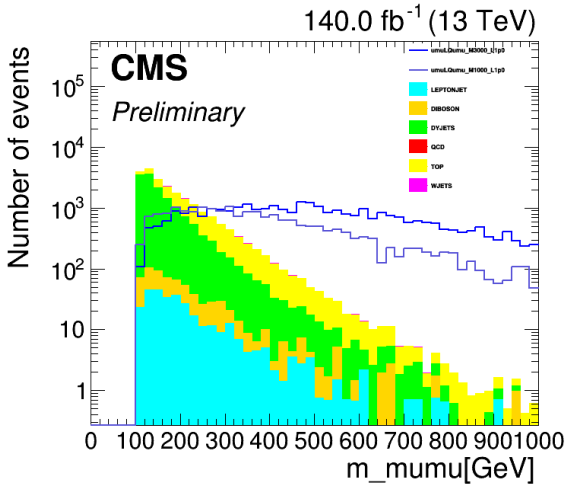
events with a proton tag in PPS, and see if it could enhance sensitivity compared to the inclusive analysis. With PPS taken into account, there will be a total of 4 categories (one μ with proton tag, one μ with no proton tag, two μ with proton tag and two μ with no proton tag).



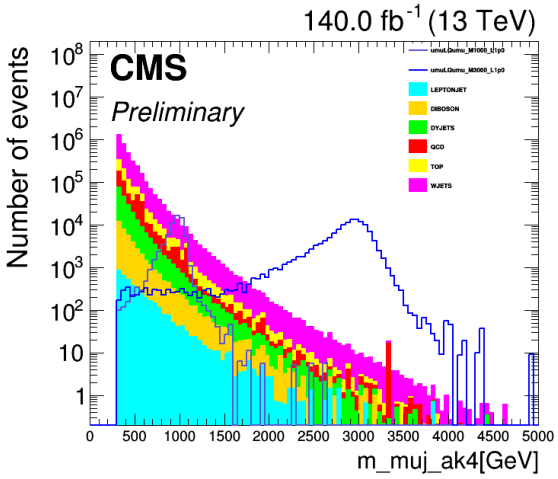
(a) *Transversal momentum of the leading muon in signal and background (1μ category)*



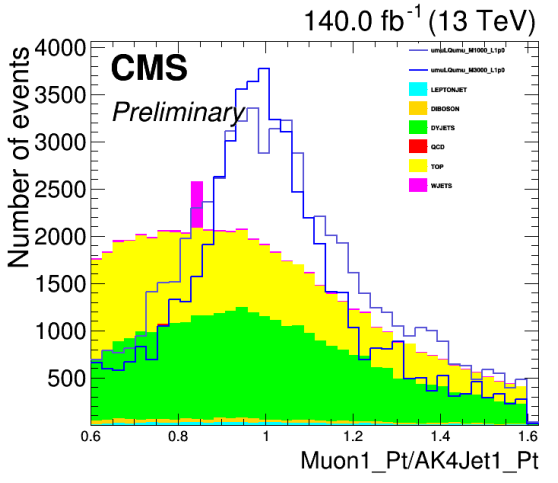
(b) *Transversal momentum of the jet in signal and background (2μ category)*



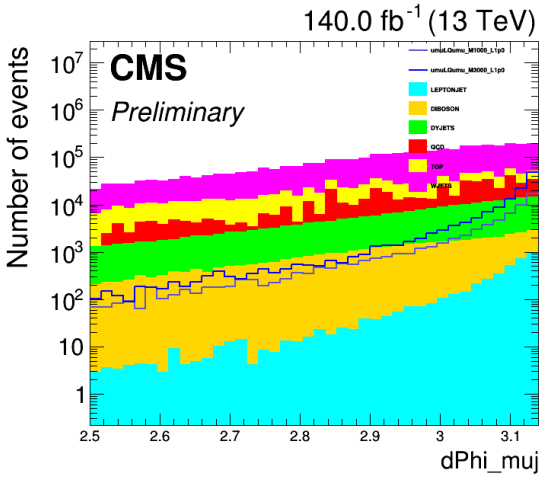
(c) *Invariant mass of the two leptons in signal and background (2μ category)*



(d) *Invariant mass of the leading muon and jet in signal and background (1μ category)*



(e) *Ratio of transversal momentums of the leading muon and jet in signal and background (2μ category)*



(f) *Azimuthal difference between leading muon and jet in signal and background (1μ category)*

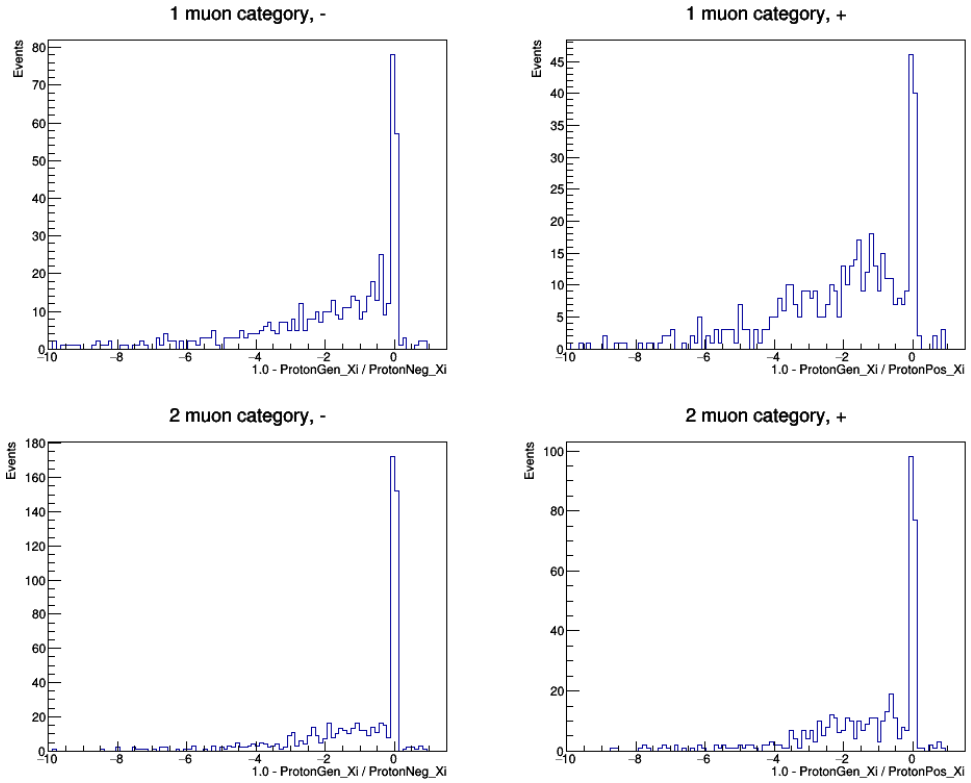
Figure 18: $u\mu$, $\lambda = 1$, $m_{LQ} = 1,3$ *TeV*

5.1 Proton reconstruction in PPS

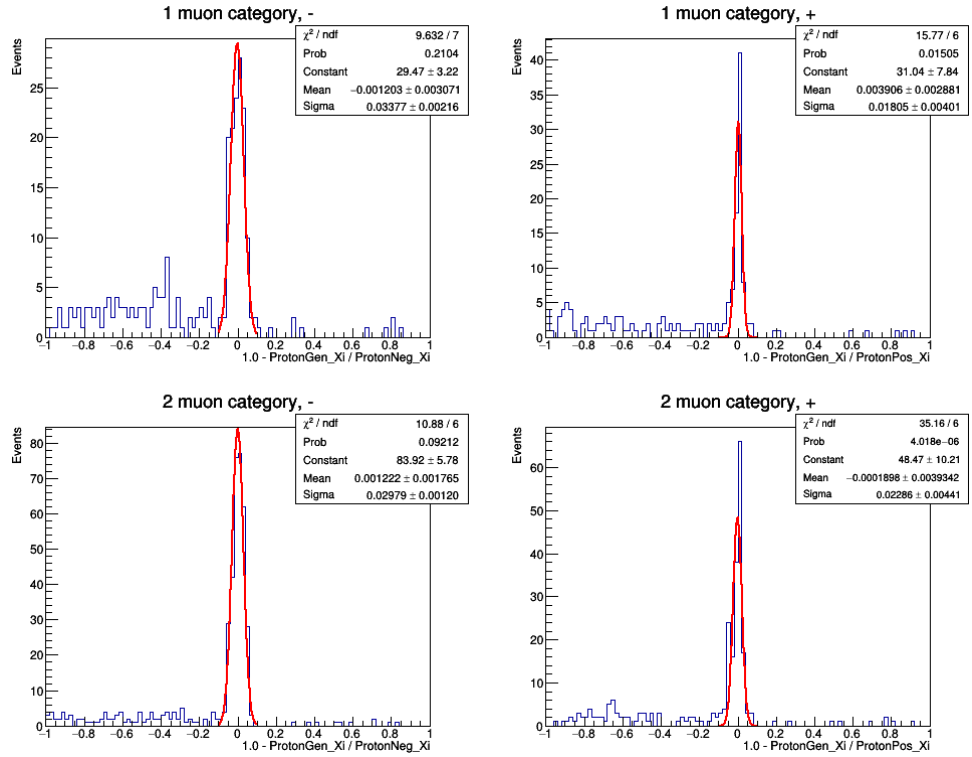
The main variable of this study is the fractional momentum loss of the scattered proton (Eq. 3.4), which is measured directly by the PPS detector. The LQ signal in CMS is reconstructed with final state lepton and jet if there's only one lepton detected (otherwise it's two leptons and jet), so consequently ξ is reconstructed with the same objects:

$$\xi_{CMS} = \frac{E_{\mu+jet} \pm P_{\mu+jet,z}}{\sqrt{s}} \tag{5.1}$$

The four-momentum is the sum of either two or three objects, depending on signal category. PPS reconstructs scattered signal protons, with a bit lower resolution compared to CMS (Figure 15). Possible signal candidates were picked by proximity to ξ calculated with generated final state particles. In Figure 19a, distributions of $1 - \frac{\xi_{GEN}}{\xi_{PPS}}$ are shown. Usually there is only one proton per event on z- and z+ side, if any. Most of the generated protons are compatible with the ones reconstructed in PPS, as there is a peak around 0. In general, events where there is a generated proton from the LQ event are taken. If the proton is out of PPS acceptance, there can still be another proton reconstructed, however the ξ will be different from the one measured in CMS, meaning it's coming from pile-up. The left tail in the distribution indicates that there are protons generated with a ξ out of the PPS acceptance, which ranges from 0.04 to 0.2, as shown in Figure 20. The asymmetry between the two sides of PPS is usually due to fixed geometry of the subdetector. Figure 19b shows the same distributions, but zoomed into the peak which is fitted with a gaussian to obtain the mean and standard deviation. The same procedure was done for all masses, each category and each z side of the PPS. Figure 21 shows mean values and resolutions for all six LQ masses. It is evident that there is no bias as the mean values are positioned around 0 and that the standard deviation goes up to 5% in ξ. The z+ side has a standard deviation that is roughly half of the one on the z- side, since it covers a narrower ξ region as seen in Figure 20. The effect is not that notable later on the fits and histograms are used mainly to match reconstructed PPS protons with the generated ones and the background is quite small. Once assured that the reconstruction is agreeable, a 3σ cut is determined from the same gaussian fit in order to match the reconstructed protons with generated signal protons, i.e. to identify if a reconstructed proton is indeed coming from a LQ. This will be used in the next section.



(a) *Distributions of* $1 - \frac{\xi_{GEN}}{\xi_{PPS}}$



(b) *Gaussian fit on the peak of the same distributions, zoomed in.*

Figure 19: *PPS resolution for* m_{LQ} *= 3 TeV for both categories and both z- and z+ sides of PPS.*

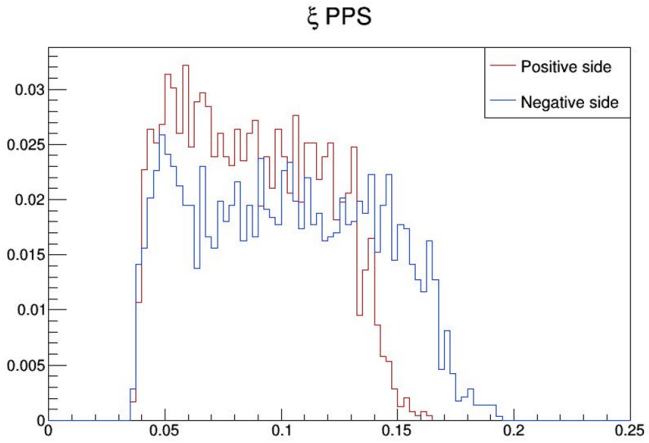


Figure 20: *Acceptance range of ξ in PPS.*

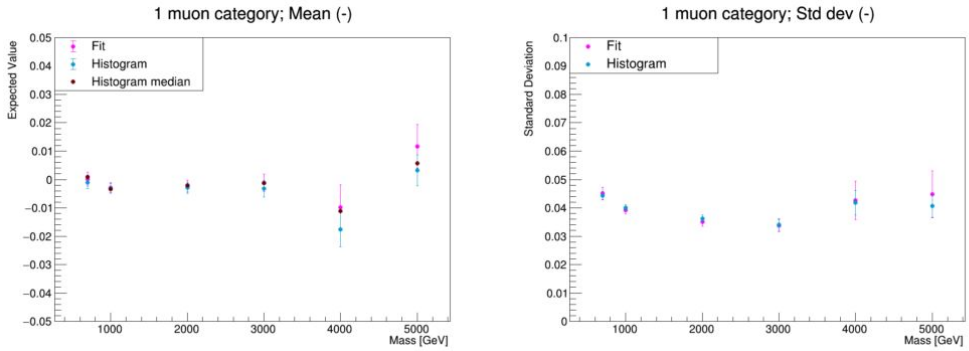
5.2 Proton tag selection criteria

An important step in this analysis is matching ξ_{CMS} (Eq. 3.7) and ξ_{PPS} (Eq. 3.4). There are numerous events in which more than one proton is detected by the PPS, as seen for example in Figure 22. Most events could have a proton tag, which means it's not enough to just demand for the presence of a scattered proton. It needs to be assured that the proton has the kinematics compatible with the ξ_{CMS} in presence of a diffractive process. The goal is to clean the sample as much as possible of the pile-up protons, which is done by applying a restriction on the variable $1 - \frac{\xi_{CMS}}{\xi_{PPS}}$, a quantity that can be computed in data and therefore applied in physics analysis.

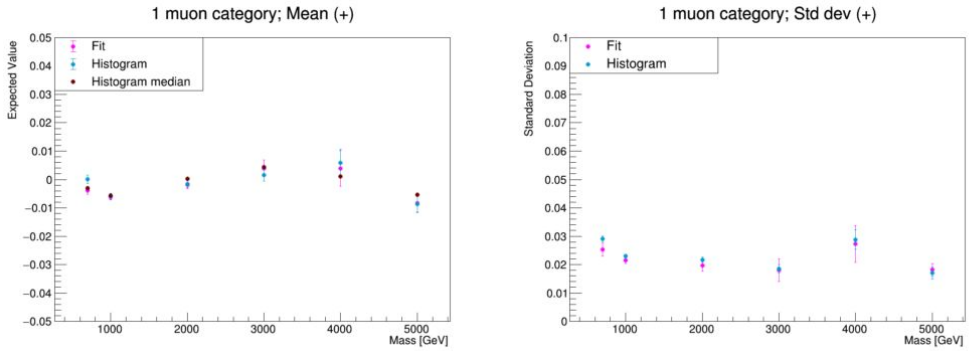
Figure 23 shows distributions of $1 - \frac{\xi_{CMS}}{\xi_{PPS}}$ for 1μ category (upper plots) and 2μ category (lower plots). The shape of the distribution greatly depends on whether there is a 2nd muon. Given that there is no 2nd muon detected in 1μ category, the shape of the peak is wider and positioned roughly above 0. A part of the energy in the final state X is not reconstructed and therefore Eq. 3.7 is not valid anymore, resulting in a less effective tag requirement. On the other hand, two muon category has a clear, narrow peak around 0.

As mentioned previously, a 3σ cut on the $1 - \frac{\xi_{GEN}}{\xi_{PPS}}$ distribution was estimated and then applied on $1 - \frac{\xi_{CMS}}{\xi_{PPS}}$ in order to distinguish matched and non-matched reconstructed and generated protons. That way it is shown that the peak indeed is due to the signal protons (as depicted in Figure 23 with magenta), while the rest are pile-up protons (depicted with blue). The shapes on both PPS sides for both categories are very similar to each other as the CMS resolution dominates over the PPS one. Next, based on the matching results, a proton tag selection is derived for each of the two categories to encompass the signal protons and discard the others:

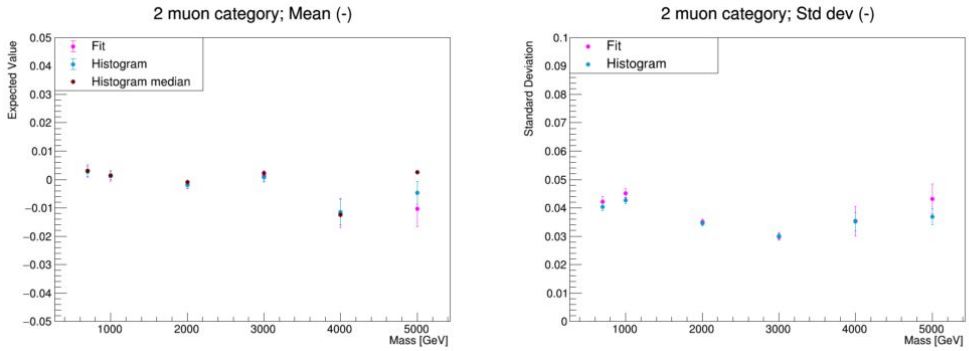
$$\begin{aligned} 1\mu \text{ category} &: 0 < 1 - \frac{\xi_{CMS}}{\xi_{PPS}} < 0.85 \\ 2\mu \text{ category} &: -0.15 < 1 - \frac{\xi_{CMS}}{\xi_{PPS}} < 0.2 \end{aligned} \tag{5.2}$$



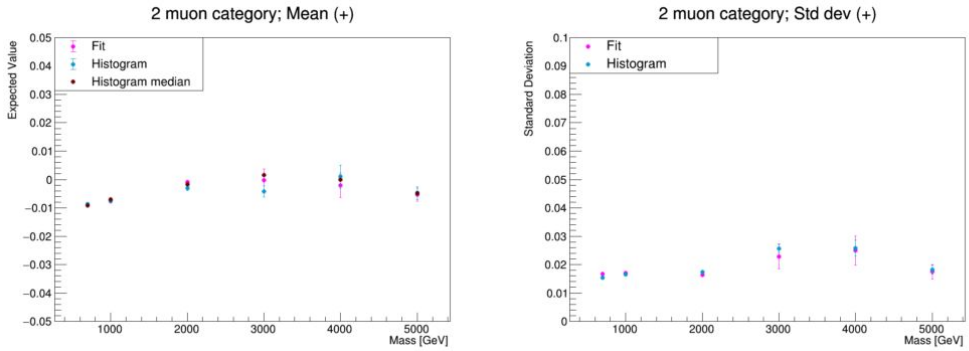
(a) *Results for 1μ category & z- side of PPS*



(b) *Results for 1μ category & z+ side of PPS*



(c) *Results for 2μ category & z- side of PPS*



(d) *Results for 2μ category & z+ side of PPS*

Figure 21: *Expected values and standard deviations from histograms and gaussian fit functions for all LQ masses from* $1 - \frac{\xi_{GEN}}{\xi_{PPS}}$ *distribution.*

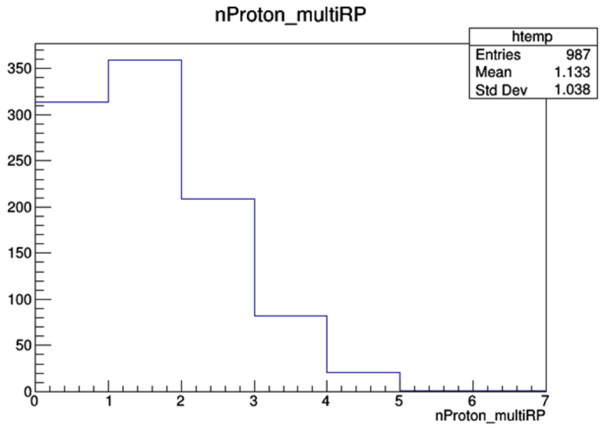


Figure 22: *Distribution of the number of MultiRP protons detected for signal LQ events in PPS for* $m_{LQ} = 2TeV$.

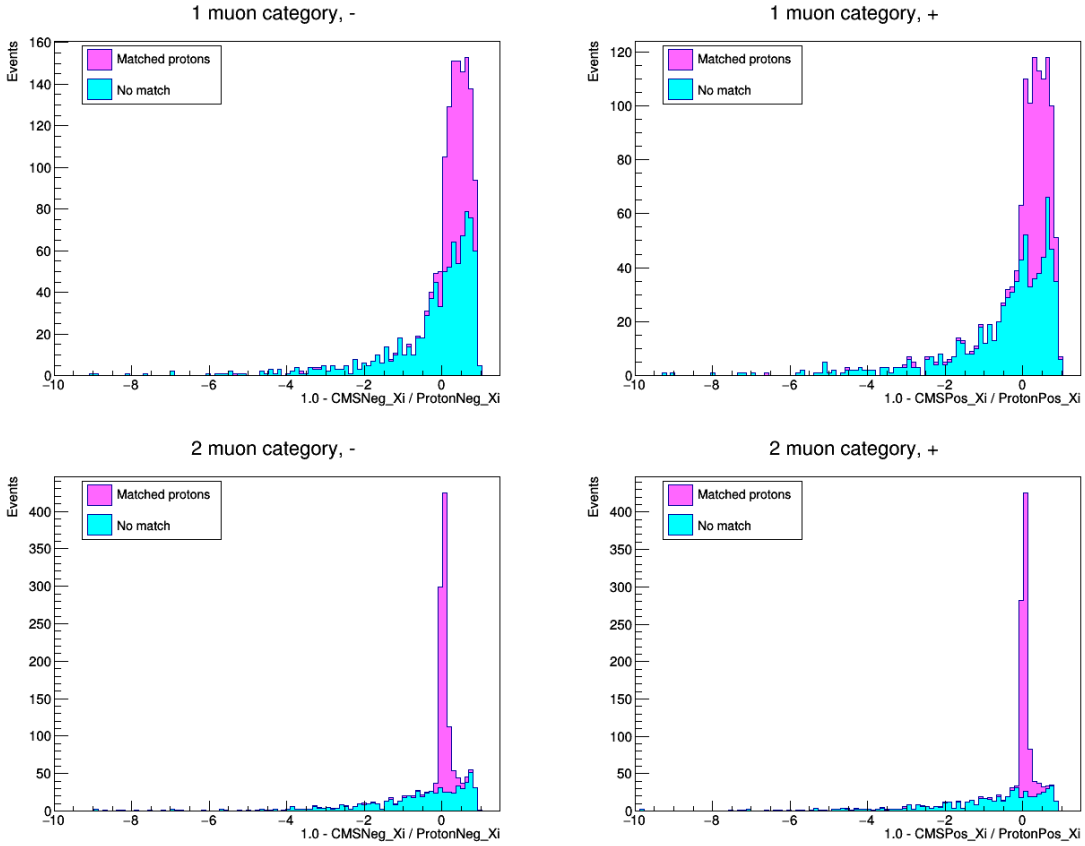


Figure 23: *Signal protons and pileup protons in* $1 - \frac{\xi_{CMS}}{\xi_{PPS}}$ *distribution for* m_{LQ} *= 1 TeV for the case when a proton is detected on z+ side of PPS.*

As a result, a signal selection criteria is formed:

1. demand a certain number of detected muons (either 1 or 2)
2. demand one proton detected per event on either z+ or z- side of the PPS with the aforementioned cuts in Eq. 5.2

5.3 Selection efficiency

Next step is to check the effect of the proton tag selection criteria on the signal and background. Figure 24 shows the signal efficiency for proton tag and no proton tag categories. The full event selection that were described in table 3 are applied along with the proton tag requirement. In the middle plot it is seen that for higher masses there is a decrease in efficiency, as expected due to the shift of the ξ out of the PPS acceptance (Figure 14). The efficiency at 700 GeV is lower than anticipated, which is due to the requirement that $m_{\mu jet} > 300$ GeV, cutting off some of the softer muons and jets. The right plot shows efficiency when the proton tag is requirement is not applied, depicting a different behavior. The efficiency increases with mass, with a slight dip at 5000 TeV due to the fact that the signal has tails and it's required a minimum μ-jet mass.

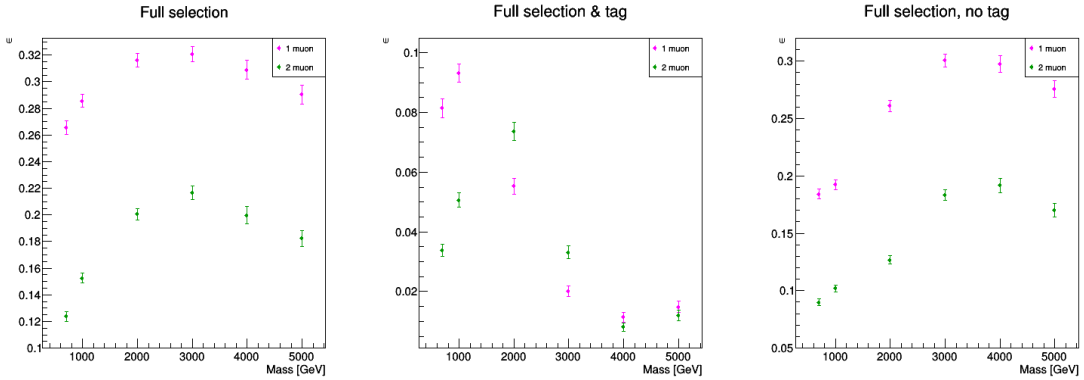


Figure 24: *Signal selection efficiency for full selection only (left), added proton tag (middle) and no proton tag (right) categories. The sum of efficiencies in middle and right give the left plot.*

m_{LQ} [GeV]	1μ	proton tag	no tag	2μ	proton tag	no tag
700	0.266±0.005	0.081±0.003	0.184±0.004	0.124±0.004	0.034±0.002	0.090±0.003
1000	0.285±0.005	0.093±0.003	0.192±0.004	0.153±0.004	0.051±0.002	0.102±0.003
2000	0.316±0.005	0.055±0.003	0.261±0.005	0.200±0.004	0.074±0.003	0.127±0.004
3000	0.321±0.005	0.020±0.002	0.300±0.006	0.217±0.005	0.033±0.002	0.184±0.005
4000	0.309±0.007	0.011±0.001	0.298±0.007	0.200±0.006	0.008±0.001	0.192±0.006
5000	0.290±0.007	0.015±0.002	0.278±0.007	0.182±0.006	0.012±0.002	0.170±0.006

Table 4: *Values of signal efficiencies and their errors for each mass and all categories.*

In order to study the effect of the tagging on the background, it has to be done on real data (luminosity of 139 fb^{-1}) with full selection applied instead, considering that PPS simulation is not available for all MC background events. Only 1% of total data is taken since it is not expected to find any signal in such case. The μ+jet mass distribution from the fraction of data is shown in Figure 25. The distribution is fitted with a function that approximately describes it, in this case a four-parameter empirical function that is usually used in dijet searches [32]:

$$f(m_{\mu jet}) = \frac{p_0(1-\frac{m_{\mu jet}}{\sqrt{s}})^{p_1}}{(\frac{m_{\mu jet}}{\sqrt{s}})^{p_2+p_3 ln\frac{m_{\mu jet}}{\sqrt{s}}}} \tag{5.3}$$

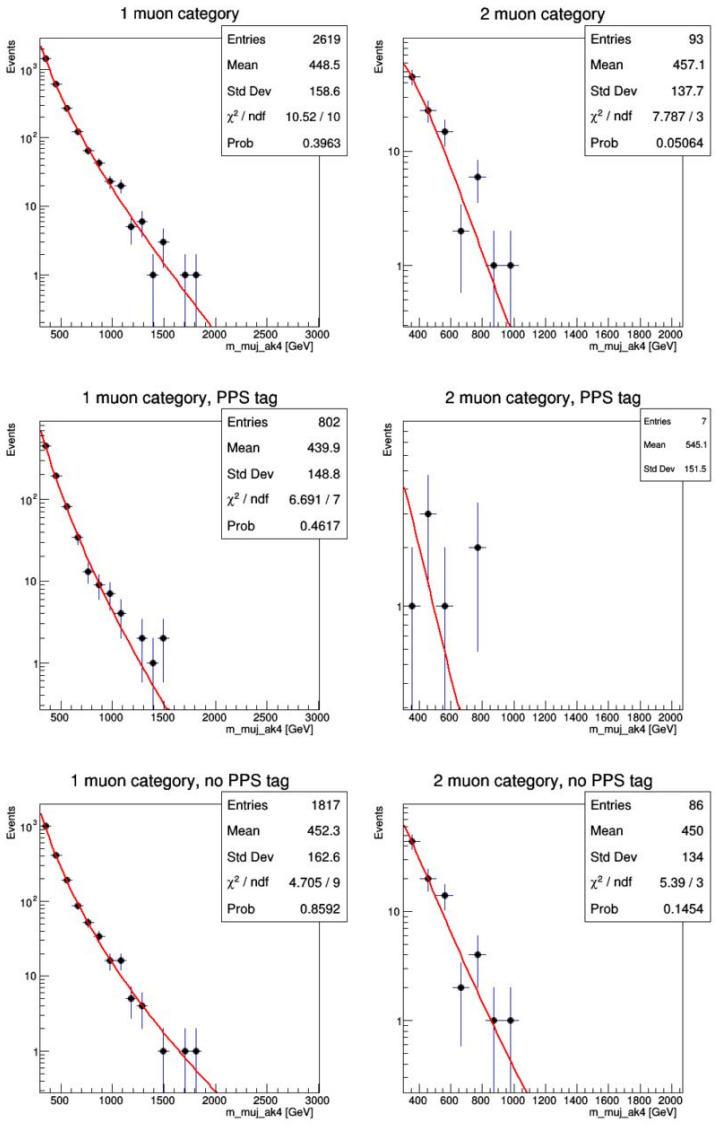


Figure 25: *μ+jet mass distribution in 1% of data for each category, fitted with Eq. 5.3.*

The 2μ proton tag category does not have enough data to perform a proper fit, so the assumption is that the proton tag and no tag categories have the same shape. This allows to rescale the no proton tag fit onto the proton tag category.

Once the fit is done on the reconstructed invariant mass distribution, it was rescaled back to full luminosity. It is then used to count the approximate number of events around each LQ mass in a window of about 10%. Table 5 shows the number of background events for each category, and it can be seen that generally there are more events that don't pass the proton tag requirement compared to the ones that do. At higher masses the background is scarce, so the effect of proton tag is not very notable.

m_{LQ} [GeV]	1μ category	proton tag	no proton tag	2μ category	proton tag	no proton tag
700	14750	3955	10795	617	39	578
1000	3957	885	3072	91	6	85
2000	70	28	42	<1	<1	<1
3000	5	4	<1	<1	<1	<1
4000	3	2	<1	<1	<1	<1
5000	<1	<1	<1	<1	<1	<1

Table 5: *Number of background events for each category, rescaled to full luminosity, rounded to one event.*

To check the improvement of the proton tag compared to the standard inclusive analysis, their respective significances are calculated, for each mass, where the significance of each category is given as $sig = \frac{S}{\sqrt{B}}$:

$$sig_{PPS} = \sqrt{sig^2_{1\mu,tag} + sig^2_{2\mu,tag} + sig^2_{1\mu,notag} + sig^2_{2\mu,notag}}$$
$$sig_{std} = \sqrt{sig^2_{1\mu} + sig^2_{2\mu}} \tag{5.4}$$

For masses 700 GeV and 1000 GeV, events are large enough to calculate said significances. The results are presented in Table 6 below:

m_{LQ} [GeV]	$\frac{sig_{PPS}}{sig_{std}}$
700	1.3
1000	1.4

Table 6: *Ratio of significances with and without the PPS tag for low LQ masses.*

The selection is an overall improvement for lower masses, having about 30-40% increase in sensitivity. For higher masses the background events enter a Poisson regime, which requires an alternative approach in computing the analysis sensitivity.

6 Conclusion

The search for LQs using lepton-induced single production mechanic is a novel analysis at the LHC, which is generally done by searching for a bump in the μ+jet mass distribution. Presented in this thesis is a feasibility study of utilizing the PPS subdetector in CMS searches for LQs in a lepton-induced single production process in which one proton is scattered after a pp collision.

A Monte Carlo signal sample of LQ events was generated for LQs with different masses, coupling constants and flavors, to study the characteristics of such event. The generation has two steps, the first being generating the central part of the process, $lq \longrightarrow LQ \longrightarrow lq$, followed by the second step, which is parton showering simulation. The resulting LQ mass distributions vary with mass and the coupling constant λ. Depending on quark flavor and values for λ, tails in very low energies become more noticeable in the distribution, as expected according to the cross-section of a hard scattering process. Additionally, a cross-section study of different LQ flavors was also carried out. As expected, the values were higher for LQs that have a valence quark and lower for the ones that have a sea quark. The cross-section is declining with increasing mass and scales with λ^2.

The $u\mu$ channel is considered onwards in the analysis. Reconstruction of scattered proton in PPS is evaluated with generated protons from LQ events by examining $1 - \frac{\xi_{GEN}}{\xi_{PPS}}$, where ξ is the proton momentum loss of the scattered proton. The reconstruction performance shows no evident bias, with the resolution being less than 5%. Furthermore, a crucial step is matching ξ_{PPS}, which is directly calculated in PPS, and ξ_{CMS}, which in CMS is determined from final state particles, to remove pile-up protons from signal sample. To achieve this, a proton tag selection criteria is determined by applying a restriction on the variable $1 - \frac{\xi_{CMS}}{\xi_{PPS}}$ that can also be computed in data.

Once that is established, signal selection efficiency is determined, and it is decreasing with mass increase due to the shift of ξ out of the PPS acceptance, ranging from about 0.02 to 0.4. The impact of the requirement is also studied on background, in order to compare to the standard inclusive search. The results give an improvement, particularly for LQs masses lower than 2 TeV, showing a 30-40% increase in sensitivity. For higher masses the number of background events becomes small and enters Poisson statistics. A different approach should be used to compute the sensitivity which was not implemented in this thesis. Furthermore, it is still expected that when the background is minimal, there will be no additional gain in splitting events into more categories.

It should be noted that the MC generator used in this analysis assumes a 70% probability of a scattered proton, a number that is subject to large theory uncertainties, needing more theory input. This approach in signal search could be useful in particular if there is a sign of excess in the inclusive search, possibly highlighting the potential signal, and it could be applied on Run 3 data in the future.

7 Bibliography

[1] CERN, *The Cern accelerator complex*, URL: https://cds.cern.ch/images/CERN-GRAPHICS-2022-001-1 (29. 6. 2023.).

[2] C. Quaranta, *Search for high-mass resonances in final states with a boosted-dijet resonance in proton-proton collisions at* $\sqrt{s} = 13$ *TeV with the CMS detector*, PhD thesis, Sapienza – University of Rome, 2021.

[3] The NNPDF Collaboration, *Parton distributions from high-precision collider data*, EPJC, 2017.

[4] J.C. Pati and A. Salam, *Lepton number as the fourth "color"*, Phys. Rev., 1974.

[5] E. Fahri and L. Susskind, *Technicolour*, Physics Reports, 1981.

[6] B. Schrempp and F. Schrempp, *Light leptoquarks*, Physics Letters, 1985.

[7] R. Aaij et al., *Test of Lepton Universality Using* $B^+ \longleftarrow K^+l^+l^+$, Phys. Rev., 2014.

[8] K. Cheung, *Muon Anomalous Magnetic Moment and Leptoquark Solutions*, Phys. Rev., 2001

[9] LHCb collaboration, *Measurement of lepton universality parameters in* $B^+ \longleftarrow K^+l^+l^+$ *and* $B^0 \longleftarrow K^{*0}l^+l^-$ *decays*, Phys. Rev., 2022.

[10] CMS Collaboration, *Search for single production of scalar leptoquarks in proton-proton collisions at* $\sqrt{s} = 8$ *TeV*, Phys. Rev., 2016.

[11] CMS Collaboration, *Search for pair production of first-generation scalar leptoquarks at* $\sqrt{s} = 13$ *TeV*, Phys. Rev., 2019.

[12] ATLAS Collaboration, *Search for pairs of scalar leptoquarks decaying into quarks and electrons or muons in* $\sqrt{s} = 13$ *TeV pp collisions with the ATLAS detector*, JHEP, 2020.

[13] CMS Collaboration, *A search for pair production of leptoquarks decaying to muons and bottom quarks at* $\sqrt{s} = 13$ *TeV*, CMS-PAS-EXO-21-019, 2023.

[14] CMS Collaboration, *Search for single production of scalar leptoquarks in proton-proton collisions at* $\sqrt{s} = 8$ *TeV*, Phys. Rev., 2016.

[15] CMS Collaboration, *Search for a singly produced third-generation scalar leptoquark decaying to a* τ *lepton and a bottom quark in proton-proton collisions at* $\sqrt{s}$*=13 TeV*, JHEP, 2018.

[16] ATLAS Collaboration, *Search for pair production of third-generation leptoquarks decaying into a bottom quark and a τ-lepton with the ATLAS detector*, EPJC, 2023.

[17] CMS Collaboration, *Search for a third-generation leptoquark coupled to a τ lepton and a b quark through single, pair, and nonresonant production in proton-proton collisions at $\sqrt{s} = 13$ TeV*, JHEP, 2023.

[18] CMS Collaboration, *Search for scalar leptoquarks produced in lepton-quark collisions and coupled to τ leptons*, Phys. Rev. Lett., 2023.

[19] I. Doršner et al., *Physics of leptoquarks in precision experiments and at particle colliders*, Physics Reports, 2016.

[20] L. Buonocore et al., *Leptons in the Proton*, JHEP, 2020.

[21] L. Buonocore et al., *Lepton-quark collisions at the Large Hadron Collider*, Physical Review Letter, 2020.

[22] David Barney, CERN document server, *CMS slice image view*, URL: http://cds.cern.ch/record/2204863 (29.6.2023.)

[23] M. Cacciari, G.P. Salam and G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP, 2008.

[24] The CMS and TOTEM Collaborations, *Observation of proton-tagged, central (semi)exclusive production of high-mass lepton pairs in pp collisions at 12 TeV with the CMS-TOTEM precision proton spectrometer*, JHEP, 2018.

[25] CMS Twiki, *Tagged protons: getting started*, URL: https://twiki.cern.ch/twiki/bin/viewauth/CMS/TaggedProtonsGettingStarted() (3.9.2023.)

[26] S. Alioli, P. Nason, C. Oleari and E. Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP, 2010.

[27] L. Buonocore et al., *Resonant leptoquark at NLO with POWHEG*, JHEP, 2022.

[28] R.L. Workman et al. (Particle Data Group), *Review of Particle Physics*, Prog. Theor. Exp. Phys., 2022.

[29] M. Bahr et al., *Herwig++ Physics and Manual*, Eur.Phys.J.C, 2008.

[30] J. Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, Eur.Phys.J.C, 2015.

[31] S. Agostinelli et al., *Geant4 - A Simulation Toolkit*, Nucl. Instrum. Meth., 2003.

[32] CMS collaboration, *Search for narrow and broad dijet resonances in protonproton collisions at $\sqrt{s} = 13$ TeV and constraints on dark matter mediators and other new particles*, JHEP, 2018.

[33] A. M. Sirunyan et al. *Search for high mass dijet resonances with a new background prediction method in proton-proton collisions at* $\sqrt{s} = 13$ *TeV*, JHEP, 2020.

A Appendix - ξ derived from kinematics

Let's assume a central exclusive process where two protons approach from opposing z directions, collide, and subsequently both particles scatter, producing final state X, as shown in Figures 10 and 11. The four-vectors of protons 1 and 2 before and after colliding are given as

$$\vec{P_1} = \begin{pmatrix} P_{beam} \\ 0 \\ 0 \\ P_{beam} \end{pmatrix}, \vec{P_2} = \begin{pmatrix} P_{beam} \\ 0 \\ 0 \\ -P_{beam} \end{pmatrix}, \vec{P_1'} = \begin{pmatrix} |P'_{1,z}| \\ 0 \\ 0 \\ |P'_{1,z}| \end{pmatrix}, \vec{P_2'} = \begin{pmatrix} |P'_{2,z}| \\ 0 \\ 0 \\ -|P'_{2,z}| \end{pmatrix}, \tag{A.1}$$

while the four-vector of state X is

$$\vec{P_X} = \begin{pmatrix} E_X \\ 0 \\ 0 \\ P_{X,z} \end{pmatrix} \tag{A.2}$$

Along the z axis, energy and momentum conservation laws give:

$$P_{beam} + P_{beam} = |P'_{1,z}| + |P'_{2,z}| + E_X \tag{A.3}$$

$$P_{beam} - P_{beam} = |P'_{1,z}| - |P'_{2,z}| + P_{X,z} \tag{A.4}$$

By summing up Eq. A.3 & A.4 and using the fact that $\xi = \frac{\Delta P}{P_{beam}}$ and $\sqrt{s} = 2P_{beam}$, ξ_1 for proton arriving from the negative z direction can be received:

$$\begin{aligned} 2P_{beam} &= 2|P'_{1,z}| + E_X + P_{X,z} \\ 1 &= \frac{|P'_{1,z}|}{P_{beam}} + \frac{E_X + P_{X,z}}{2P_{beam}} \\ \xi_1 = 1 - \frac{|P'_{1,z}|}{P_{beam}} &= \frac{E_X + P_{X,z}}{\sqrt{s}} \end{aligned}$$

In a similar way, ξ_2 for proton arriving from the positive z direction can be received by subtracting Eq. A.4 from A.3:

$$\begin{aligned} 2P_{beam} &= 2|P'_{2,z}| + E_X - P_{X,z} \\ 1 &= \frac{|P'_{2,z}|}{P_{beam}} + \frac{E_X - P_{X,z}}{2P_{beam}} \\ \xi_2 = 1 - \frac{|P'_{2,z}|}{P_{beam}} &= \frac{E_X - P_{X,z}}{\sqrt{s}} \end{aligned}$$