Title
Performance and Calibrations of the Compact Muon Solenoid Muon System and a Search for Sphalerons at the Large Hadron Collider

Author
Bravo, Cameron

Publication Date
2018

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Performance and Calibrations of the Compact Muon Solenoid Muon System and a Search for Sphalerons at the Large Hadron Collider

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Physics

by

Cameron Bily Bravo

2018

ABSTRACT OF THE DISSERTATION

Performance and Calibrations of the Compact Muon Solenoid Muon System and a Search for Sphalerons at the Large Hadron Collider

by

Cameron Bily Bravo

Doctor of Philosophy in Physics

University of California, Los Angeles, 2018

Professor Jay Hauser, Co-Chair

Professor David Saltzberg, Co-Chair

This thesis first presents a detailed study of the Compact Muon Solenoid (CMS) Cathode Strip Chamber track stubs used to build muon tracks in the first layer of filtering, done in hardware, of the data. We study the spatial and angular resolutions of the track stubs with respect to the reconstructed hits using the full granularity data. We then move to a general description of a new technology, Gas Electron Multipliers (GEM), which is to be used in a near-future upgrade to the system. The characterization and calibration of the front-end electronics of the GEMs is discussed in detail. It is demonstrated that the noise of the electronics is sufficiently low to allow high efficiency operation of the technology with the first prototype GEM detector for the new subsystem. We then discuss the phenomenology of sphaleron-induced transition in proton-proton collisions. A new

Monte Carlo generator is built to study this phenomenology in detail. We then perform a search analysis of the 35.9 fb^{-1} of data collected by the CMS detector in 2016. We set the first experimental upper limit on sphaleron production in proton-proton collisions.

The dissertation of Cameron Bily Bravo is approved.

Michael McNitt-Gray

David Saltzberg, Committee Co-Chair

Jay Hauser, Committee Co-Chair

University of California, Los Angeles

2018

To my parents . . .

who—among so many other things—

always told me I could achieve whatever

I wanted, as long as I worked hard.

Table of Contents

List of Figures

List of Tables

Acknowledgments

Thanks to Aaron Dominguez for giving me the opportunity to begin doing particle physics research. Thanks to Jose Lazo-Flores for all his support while I was at PSI and ETH. Thanks to Roland Horisberger for teaching me so much about electronics. Thanks to Pieter Everaerts for teaching me many foundational analysis skills. Thanks to Eric Cotner, Ji-Haeng Huh and Doo Jin Kim for having so many theoretical discussions with me. Thanks to Graciela Gelmini for the original inspiration to work on experimental sphaleron physics. Thanks to Steve Mrenna for his help with interfacing BaryoGEN with PYTHIA. Thanks to Nick McColl for many interesting discussions related to the search analysis. Thanks to Martin Kwok and Greg Landsberg for collaborating with me to set the first ever experimental limit on the sphaleron pre-exponential factor. Thanks to David Saltzberg for the initial push needed to begin working on writing the BaryoGEN generator, and all of the wonderful advice on most of the things I worked on. A special thanks goes to Jay Hauser for his sage advice throughout my entire graduate experience, giving me the freedom to work on Sphalerons, and teaching me how to truly be a scientist.

Vita

2013 B.S. (Physics and Mathematics), University of Nebraska-Lincoln.

2015 M.S. (Physics), UCLA, Los Angeles, California.

CHAPTER 1

Introduction

Elementary particle physics is the study of the fundamental constituents of matter and how they interact. Quantum Field Theory (QFT) is the mathematical framework of modern particle physics. Conceptually, QFT models all of matter and interactions as particles. The particles which make up matter are called fermions and the force carrying particles are called gauge bosons. Elementary particle physics works on characterizing particles in terms of how strongly they interact with other particles, stability, and abundance in the universe. Any two particles are considered of the same species if they interact identically. An electron is an example of one species of particle. All particles have an anti-particle, most of which are unique from the particle; however, some particles are their own anti-particle. Many particles are composite, meaning they are made of multiple particles. Throughout this document, the word particle will be used to mean a specific species. Every particle has an associated field in which it propagates. This field can be thought of as a vacuum in which excitations can be produced, and all the individual excitations of the field are the individual particles of the species to which the field belongs. The field is the collection of all particles and anti-particles of an individual species and the potential for more or less number of that particular species to exist.

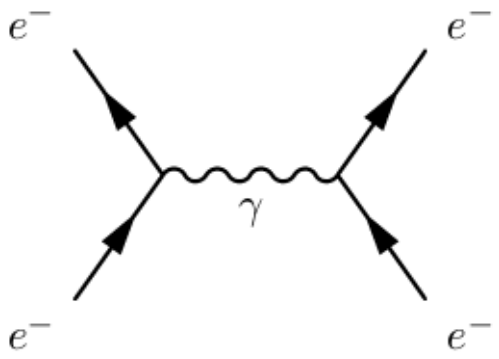

Figure 1.1: A Feynman diagram depicting the interaction of two electrons via the electromagnetic force. The wavy line is a photon propagator. The photon is the boson which carries the electromagnetic force. In this particular type of interaction, the species of the incoming and outing particles are identical, specifically two electrons. The two points where two electrons and the photon connect are called a vertices.

The interactions of each particle are studied experimentally via scattering experiments. In such an experiment we accelerate particles to a known amount energy and steer them to minimize the distance between the particles. When the particles get closer together, they become more likely to interact via the exchange of additional particles. After an interaction occurs, the particles that leave the point of interaction will not necessarily be the same species of particles that entered the interaction. Characterization of the experiments are made by recording the species and direction of the particles leaving the interaction. It is impossible to know exactly the nature of any single scattering event, so the nature of the particles are studied over some large number of scattering events.

QFT categorizes different types of interactions via what are known as Feynman diagrams. Figure 1.1 is an example of a Feynman diagram, which shows

two electrons interacting via the exchange of a photon. This diagram does not completely characterize all electromagnetic (EM) interactions between two electrons; however, it does describe a large fraction of them. A full description of EM interactions is given by the QFT called Quantum Electrodynamics (QED), which is fully described in [Fey90]. QED is the most stringently tested theory of physics.

1.1 The Standard Model

The most complete theory of our understanding of all particles and how they interact is known as The Standard Model (SM) of particle physics. This theory names all fundamental fermions and gauge bosons and precisely describes all possible interactions. The theory also explains all possible composite particles. There are three forces described by the SM, which does not include those of gravitation. Figure 1.2 is a graphic presenting all of the fundamental particles included in the SM.

1.1.1 The Fermions

The fermions in the SM are split into two groups called quarks and leptons. These groups correspond to which of the gauge bosons the particle can potentially interact with them. The quarks can potentially interact with any of the gauge bosons while the leptons can not interact with the gluon. Each of these two groups are split into three "doublets", each of which corresponds to a different generation of matter. The three generations of matter are electron, muon, and tau. The quark doublet directly above a lepton doublet is of the same generation.

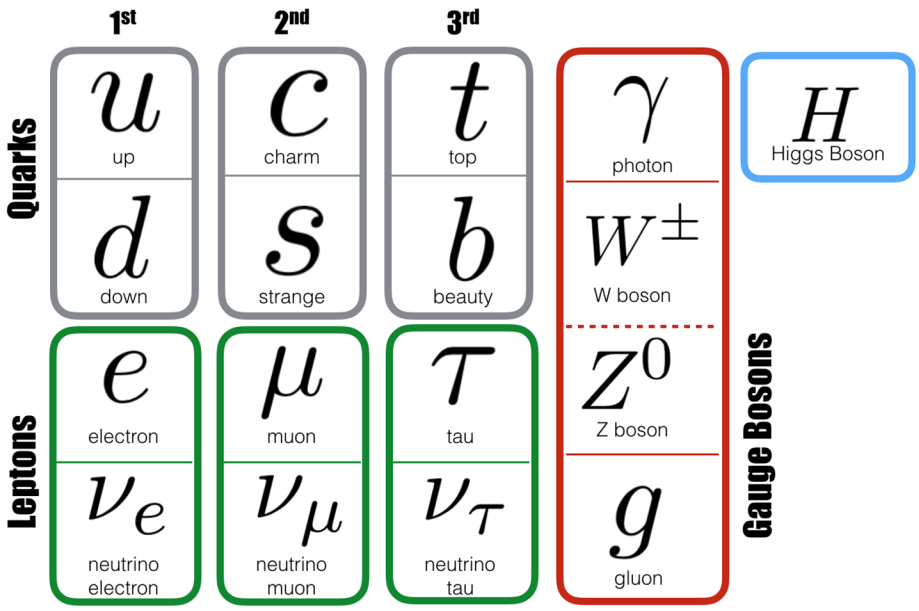

Figure 1.2: A diagram showing all fundamental particles in the SM. The fermions are on the left in grey and green. The gauge bosons are the right in red. The Higgs boson is in blue on the far right.

1.1.2 The Gauge Bosons

The fundamental forces of the universe are carried by gauge bosons. The forces described by the SM are electromagnetic, weak, and strong interactions. This gauge theory has specifically a $\mathrm{SU}(3)_C \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ internal symmetry, where C refers to color, L to the left chiral nature of SU(2), and Y is the weak hypercharge. Quantum Chromodynamics (QCD) is the theory of the strong interaction, coming from the SU(3) portion of the gauge group. The $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ portion is the symmetry of the Electroweak (EW) interaction. Noether's Theorem guarantees each of the symmetries of the SM have an associated conserved charge. This charge for QCD is called color, and the conserved values of EW interactions are known as the weak isospin and weak hypercharge. The gauge bosons directly from the Electroweak symmetry are massless.

1.1.3 The Higgs Boson

An important property of particles is the intrinsic angular momentum, more commonly known as spin. Particles can any multiple of $\frac{1}{2}$ spin. Fermions always have an odd multiple of half spin. Bosons have an even multiple of $\frac{1}{2}$ spin, including zero. In general, not all bosons are gauge bosons, as some are composite of some even number of fundamental fermions. The carriers of the strong and electroweak interactions are all spin-1 gauge bosons.

There is one additional particle in figure 1.2 which has yet to be mentioned. The Higgs boson is spin-0 and plays a special role in the SM. This particular field has a special potential associated with it, that allows the vacuum of the universe to have some non-zero amount of energy. When this vacuum energy becomes non-

zero a transition of the vacuum occurs, and a new local minimum of the vacuum energy is realized by the universe. This vacuum transition however rearranges the how all the particles interact, and in particular separates the EW interactions into two orthogonal forces, the EM force and the weak force. When EW force breaks into two separate forces, the new gauge bosons associate with the weak force are no longer massless. The gauge bosons in figure 1.2 represent those in the SM after the vacuum transition induced by the presence of the Higgs field.

1.2 Matter and Antimatter

Every particle's field carries both the matter and antimatter particles of that species. There are several different interpretations of what antimatter is, I will describe one that is essentially the Feynman-Stueckelberg interpretation. The mathematical solutions of the particle excitation of the fields of matter and antimatter are identical up to the sign of the oscillations in time. This minus sign typically lead to an interpretation of antimatter being matter moving backwards in time. Using some mathematical tricks one can move the minus sign around in the equations and make the solution appear to be a particle moving forward in time with opposite charge, but still equal spin and mass. Aside from this, matter and antimatter both interact identically in the SM. The existence of antimatter is a direct consequence of combining the theory of relativity with quantum mechanics. [CDS12]

There are several quantum numbers used to characterize the species of a particle. Two of these numbers are Baryon number (B) and Lepton number (L). Quarks have a B of one-third while antiquarks have a B of minus one-third. Lep-

tons like the electron have an L of 1 while the positron (anti-electron) has an L of minus one. None of the interactions of fermions with gauge bosons can change the total B or L in the universe.

This brings us to one of the biggest mysteries of physics. The results of cosmic-ray experiments, measurements of the abundance of all particles flying into the earths atmosphere, indicate there is nearly no antimatter in the universe compared to the amount of matter observed. How is it that the best theory we have of particles and all their interactions indicate symmetric interactions of matter and antimatter, yet we see such a large asymmetry in their abundance? No physics has every been observed capable of providing a satisfactory explanation of the mechanism of the matter over antimatter asymmetry, though many theories provide potential solutions. One general class of theories which explains the asymmetry through the generation of a baryon asymmetry is known as baryogenesis. Baryogenesis models of the universal matter asymmetry have three necessary conditions, first described by Sakharov in 1967 [Sak91]: baryon number violation, C and CP violation, and a deviation from thermal equilibrium.

1.3 EW Baryogenesis

Perturbative solutions of the SM are not capable of providing a mechanism of baryogenesis. This however is not true, as the SM does have a lesser known set of solutions which are capable of providing an explanation! EW baryogenesis is the particular potential answer to the mystery which requires no BSM physics. The SM already provides C and CP violation with weak interactions. Baryon number violation would be potentially provided by sphalerons and/or instantons in the

SM. Instanton-induced vacuum transitions are inherently a tunneling process, so they are heavily suppressed. These are non-perturbative solutions to the SM which violate $B + L$ while conserving $B - L$ via an axial-anomaly in the theory. The final requirement for any baryogenesis scenario is an out of equilibrium thermal state. The main topic of this thesis on the phenomenology of sphaleron-induced vacuum transition, in particular in scattering type experiments. These types of solutions to EW gauge theories seem to imply that B and L are somehow related to the topology of the vacuum of the EW gauge field.

1.4 Experimental Methods

There are three main experimental methods commonly used when directly studying the scattering of particles. All three involve understanding how particles interact with a large amount of matter. The first method is to build a large number of small sensitive detectors that can be used to localize the precision positions within the path of the particle. These methods are in general called tracking. Particles coming from a scattering event will fly through many different detectors and the position of small energy deposits during this flight are recorded and used to reconstruct the precise path traversed by the particle. The second method is to absorb and measure all of the energy carried by a particle, which is called calorimetry. The final method used, typically to aid in identification of the particle species, is measurement of photons emitted by particles traveling greater than the phase velocity of light in that material. When the particle does so, it makes a "sonic boom" type of effect which emits photons in a cone. This cone of photons is known as Cherenkov radiation, where the size of the cone is related to the velocity

of the particle. This can be combined with a measurement of the momentum of the particle to calculate the mass of the particle. It is common for an experiment to have some combination of these methods to measure the properties of interest of the particles of interest in the experiment.

1.4.1 Interactions of Particles with Matter

All experimental techniques used in particle physics utilize our understanding of how particles interact with matter. All the methods are reliant on our understanding of the EM force in general. The types of interactions that can happen depend heavily on the energy of the particle being measured. All materials are characterized by how different particles will interact with them at different energies.

1.4.2 Tracking

Some of the most important tools in experimental particle physics are trackers. The general idea with these types of systems is to measure many points in space through which the particle passed. Software algorithms are then used with the set of position measurements to reconstruct a path, or track the particle took through the entire system. Tracking is a complicated game of connect the dots.

The most important characteristic of a track is its geometry. While there are a few different types of detection material used in tracking, they all work on the same principle. Minimum ionizing particles (MIPs) are detected via deposition of small amounts of energy via ionization of atoms in a small localized area. The ionization is measured via either the observation of current through a reversed biased diode

or through scintillation photons, generated as a result of the ionization, or via the induced current in a thin layer of metal by the drift of electrons in a gas from a cascade seeded by the initial ionization event.

The geometric design of the tracker should reflect several important considerations. An understanding of the rate per unit area expected to pass through the detector is important to ensure the detector will not become saturated with signals. The position resolution necessary for a successful track projections to be reconstructed is also an important characteristic of tracking detectors. The geometries can be categorized as either strip or pixel trackers. Strip trackers will have long-thin channels in hermetic geometries and pixel trackers will have a large number of small, nearly square, channels in usually simpler geometries. A full tracker system can also commonly be built of several different technologies.

Track reconstruction algorithms tend to be combinatorial in nature. One will first assume some shape parametrization for track segments between each "hit" seen in the detector for the event. Then the parameter space will be searched for the track which best fits the hits. The specific parameterization of the track shape and method of finding the best fit are common topics in the discussions of a specific analysis of the data. Hits can also commonly be shared between multiple neighboring channels. In this case, one must also decide how to interpolate the signal of neighboring channels to obtain more precise hit resolution.

1.4.3 Calorimetry

1.4.3.1 EM Calorimeters

Electromagnetic showers are a common phenomenon used to measure the energy of the primary particle initiating the shower. There are two types of interactions which cascade to make an EM shower. Bremsstrahlung is a photon coming from the deceleration of a charged particle. This radiative effect must recoil off of a nucleus to conserve momentum in the rest frame of the radiating particle. Figure 1.3 is a depiction of Bremsstrahlung radiative process. The other process in EM showers is called pair production. This is when a high energy photon interacts with a nucleus to produce and electron and a positron. The photon will no longer exist after the production. Figure 1.4 is a graphic of the pair production process.

A cascade of bremsstrahlung and pair-production will result in a conversion of the primary particles energy into a large number of ultraviolet photons. A sensor is used to measure this flash of photons, which is calibrated with particles on known energy.

One of the most important characteristics of a material for particle physics measurements is the radiation length, which is both the mean distance over which a high-energy electron is reduced to $\frac{1}{e}$ of its initial energy by bremsstrahlung, and $\frac{7}{9}$ of the mean free path for a high-energy photon to produce an electron positron pair. Electrons and photons incident traversing any material will produce an EM shower, and the radiation length characterizes the average length of the cascade in the material of interest. The Moliere radius is related to the width of showers at its maximum number of charged particles, on average. The characteristics of the showers is a driving factor in the choice of material and geometry of an

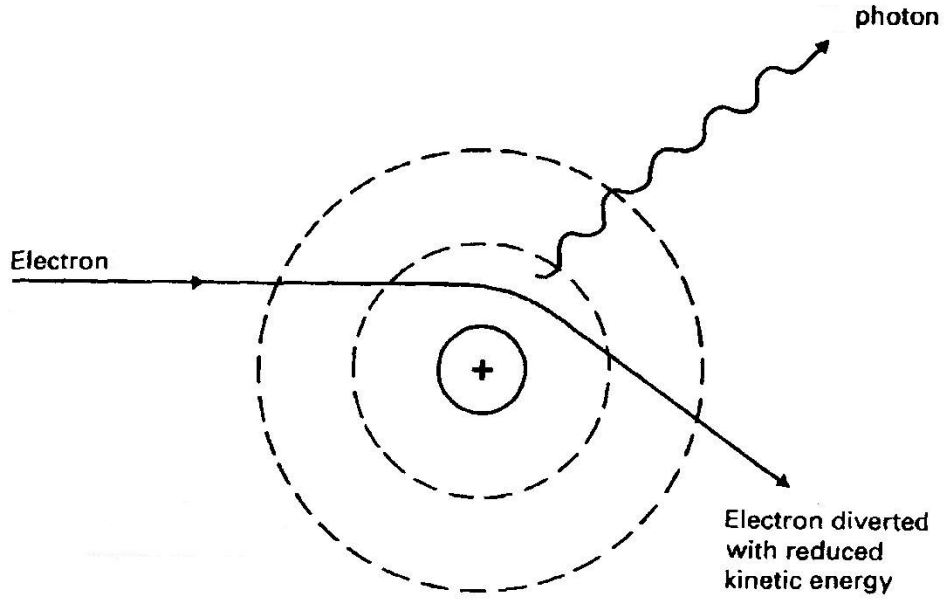

Figure 1.3: A graphic representing the bremsstrahlung process. The electron interacts with a nearby nucleus in the lattice of the material and radiates a high energy photon. After the radiative event the electron will have less kinetic energy.

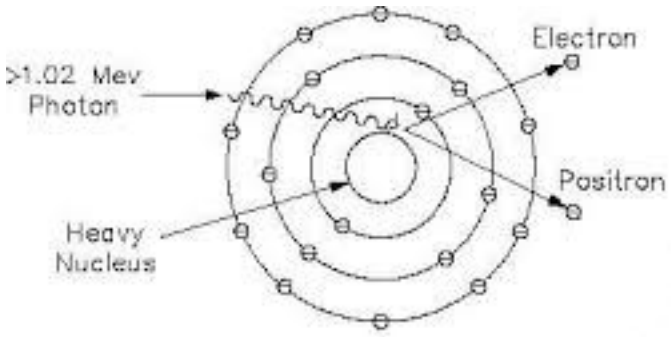

Figure 1.4: A graphic representing the pair-production process. The photon interacts with a nearby nucleus in the lattice of the material and converts into an electron-positron pair.

EM Calorimeter (ECAL), used to measure the energy of high-energy electrons, positrons, and photons.

1.4.3.2 Hadronic Calorimeters

Many particles are less likely to induce an EM shower than they are to induce a hadronic shower because they couple to the strong force. A different type of instrument is needed to measure the energy of these particles, called a hadronic calorimeter (HCAL). Any particle which will interact with the strong force, all particles composite of quarks, is called a hadron. Some hadrons are unstable and will decay immediately, but many are stable enough to travel some reasonable amount of distance into the detector. These particles will interact, via the strong force, with the nuclei of the material to produce a cascade of interactions forming a hadronic shower. The total number of charged particles in the shower is closely related to the energy of the primary particle. There exists a special type of plastic called scintillator. When charged particles traverse scintillator it will exhibit luminescence driven by the ionization of the plastic. The light emitted by this process will typically be in the UV range of the spectrum, and wil not be reabosrbed by the scintillator material. This light is then detected by some sort of photosensor, similar to EM calorimeters, and the amount of this light is then calibrated to the energy of the particle which started the hadronic shower. This works well because the scintillating material does not reabsorb the emitted UV light.

1.4.4 Detector Systems

Modern-day particle physics detectors will typically consist of many subsystems of several technologies. The exact choice of technologies and geometry will depend on the specific scientific goals of the experiment and conditions it will need to operate under. It is typical for a minimal system to consist of a tracking system and one of the two types of calorimeters. The system of different technologies is ultimately used to provide momentum and energies of particles while being able to identify the particular species of each particle. This thesis describes an experiment performed using the Compact Muon Solenoid (CMS) detector system.

1.4.4.1 The Compact Muon Solenoid

The central feature of the CMS apparatus is a superconducting solenoid of 6 m internal diameter, providing a magnetic field of 3.8 T. Within the solenoid volume are a silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter (HCAL), each composed of a barrel and two endcap sections. Forward calorimeters extend the pseudorapidity coverage provided by the barrel and endcap detectors. Muons are detected in gas-ionization chambers embedded in the steel flux-return yoke outside the solenoid.

In the barrel section of the ECAL, an energy resolution of about 1% is achieved for unconverted or late-converting photons that have energies in the range of tens of GeV. The remaining barrel photons have a resolution of about 1.3% up to a pseudorapidity of $|\eta| = 1$, rising to about 2.5% at $|\eta| = 1.4$. In the endcaps, the resolution of unconverted or late-converting photons is about 2.5%, while the

remaining endcap photons have a resolution between 3 and 4% [Kha15c]. When combining information from the entire detector, the jet energy resolution amounts typically to 15% at 10 GeV, 8% at 100 GeV, and 4% at 1 TeV, to be compared to about 40%, 12%, and 5% obtained when the ECAL and HCAL calorimeters alone are used.

Jets are reconstructed offline from the energy deposits in the calorimeter towers, clustered using the anti-k_t algorithm [CSS08, CSS12] with a distance parameter of 0.4. In this process, the contribution from each calorimeter tower is assigned a momentum, the absolute value and the direction of which are given by the energy measured in the tower, and the coordinates of the tower. The raw jet energy is obtained from the sum of the tower energies, and the raw jet momentum by the vectorial sum of the tower momenta, which results in a nonzero jet mass. The raw jet energies are then corrected to establish a relative uniform response of the calorimeter in η and a calibrated absolute response in transverse momentum p_T.

The global event reconstruction (also called particle-flow event reconstruction [Sir17]) aims to reconstruct and identify each individual particle in an event, with an optimized combination of all subdetector information. In this process, the identification of the particle type (photon, electron, muon, charged hadron, neutral hadron) plays an important role in the determination of the particle direction and energy. Photons are identified as ECAL energy clusters not linked to the extrapolation of any charged particle trajectory to the ECAL. Electrons (e.g. coming from photon conversions in the tracker material) are identified as a primary charged particle track and potentially many ECAL energy clusters corresponding to this track extrapolation to the ECAL and to possible bremsstrahlung photons emitted along the way through the tracker material. Muons are identified

as a track in the central tracker consistent with either a track or several hits in the muon system, associated with an energy deficit in the calorimeters. Charged hadrons are identified as charged particle tracks neither identified as electrons, nor as muons. Finally, neutral hadrons are identified as HCAL energy clusters not linked to any charged hadron trajectory, or as ECAL and HCAL energy excesses with respect to the expected charged hadron energy deposit.

For each event, hadronic jets are clustered from these reconstructed particles using the infrared and collinear safe anti-k_T algorithm [CSS08, CSS12] with a distance parameter of 0.4. The jet momentum is determined as the vectorial sum of all particle momenta in the jet, and is found from simulation to be within 5 to 10% of the true momentum over the whole p_T spectrum and detector acceptance. Additional proton-proton interactions within the same or nearby bunch crossings can contribute additional tracks and calorimetric energy depositions to the jet momentum. To mitigate this effect, tracks identified to be originating from pileup vertices are discarded and an offset correction is applied to correct for remaining contributions. Jet energy corrections are derived from simulation to bring measured response of jets to that of particle level jets on an average. In situ measurements of the momentum balance in dijet, photon + jet, Z + jet, and multijet events are used to account for any residual differences in jet energy scale in data and simulation [Kha17b]. The jet energy resolution amounts typically to 15% at 10 GeV, 8% at 100 GeV, and 4% at 1 TeV. Additional selection criteria are applied to each jet to remove jets potentially dominated by anomalous contributions from various subdetector components or reconstruction failures.

Muons are measured in the pseudorapidity range $|\eta| < 2.4$, with detection planes made using three technologies: drift tubes (DT), cathode strip chambers

(CSC), and resistive plate chambers (RPC). Matching muons to tracks measured in the silicon tracker results in a relative transverse momentum resolution for muons with $20 < p_T < 100$GeV of 1.3–2.0% in the barrel and better than 6% in the endcaps. The p_T resolution in the barrel is better than 10% for muons with p_T up to 1 TeV [Cha12].

A more detailed description of the CMS detector, together with a definition of the coordinate system used and the relevant kinematic variables, can be found in Ref. [Cha08].

CHAPTER 2

Study of CSC Cathode Trigger Primitive Performance

2.1 Introduction to CMS Cathode Strip Chambers

The outermost detector system of CMS is used to identify and measure the momentum of muons coming from the interaction point. The muon system is built using multiple different gas detector technologies. In the region closer to the beam pipe, the main technology used is known as cathode strip chambers (CSC), a gas detector with wire anodes and cathode strips etched from a large plane of copper. The wires and strips run nearly perpendicular to each other so a 2D hit can be reconstructed in the plane of the detector layer. Each CSC chamber has 6 layers of wires and strips.

The anode wire signals are made into binary signals by requiring the signal to be over a configurable threshold. These binary signals are then routed to the Anode Local Charged Track (ALCT) board, which is located on the chamber. The ALCT uses a simple pattern recognition algorithm which looks for signals in 4 or more layers which line up in a "track". When one of these stubs of anode hits are found, what is called an ALCT is formed and sent to the Trigger Mother Board (TMB).

The cathode signals are amplified, buffered, and digitized by the on chamber board called the Cathode Front End Board (CFEB). This board uses an ADC to digitize the signals from the cathode strips, to be used in the reconstruction. The CFEB also uses an array of comparators to digitize the signals quickly to half-strip resolution for use in the trigger. These half-strip signals are sent to the TMB, where a pattern recognition algorithm is used to form a Cathode Local Charged Track (CLCT) from four or more layers of hits within a time window. When the TMB receives an ALCT and forms an in-time CLCT it will form a trigger object known as a Local Charged Track (LCT) which is sent along a fixed latency path to the Track Finder to be correlated from LCTs in other stations of the system.

The wire hits and digitized strip signals are read out when a Level One Accept (L1A) signal is received from the trigger computer. These signals, for each layer, are reconstructed into 2D combinations of cathode and anode hits called recHits, which are the lowest level unit used in the full offline reconstruction. These are built by sophisticated fits to the ADC data recorded by the CFEBs and the wire data recorded by the ALCT. The algorithm to build CSC recHits is a standard sequence in the CMS software package (CMSSW). Every recHit is assigned an energy which is the sum of the ADC counts after pedestal subtraction. The recHits without a comparator peak below the matched distribution because of the baseline threshold used in the comparator circuit. The recHits of each the layers in a chamber are then combined into segments. The segments are then used in the track fitting procedure during the muon reconstruction. Segments require a minimum number of three layers of recHits to be built.

The performance of the CSC trigger primitives are studied in detail in this thesis. Comparisons of trigger data are made with full granularity data. The

comparator data from the CFEBs is compared to the recHit positions. The LCTs are compared with the reconstructed segments. The set of patterns used to form CLCTs has never been optimized. The CSC segments used to fit selected global muon tracks are used to select the chambers from which the raw data is stored for further analysis. The use of each of the patterns is carefully studied. A simplified emulator of the CLCT algorithm is written to study the performance of alternative pattern sets.

These studies are done using the CMS Charmonium dataset from the 2016F run period. Opposite-sign global muon pairs are preselected and the muons from preselected pairs are selected if their invariant mass is within 0.1 GeV of the J/ψ mass and have a transverse momentum over 2.0 GeV. This sample has a large signal-to-background ratio and produces a large number of "real" muons. The invariant mass spectrum of the preselected muons can be seen in figure 2.1 (left). The transverse momentum spectrum of the selected muons can be seen in figure 2.1 (right).

2.2 CFEB Comparator Performance

The most fundamental unit of CSC cathode trigger data are the comparators. An array of comparators is used to compare the signal of each strip with the nearest and next-to-nearest neighbor strips and a baseline threshold. The output of these comparators are then put into an asynchronous logic circuit which creates pulses with half-strip hit resolution [Hau99]. A simplified block-diagram of this circuit is shown in figure 2.2. The recHit energy distribution is shown in figure 2.3. The distribution of recHit energy with missing energy own can be explained as

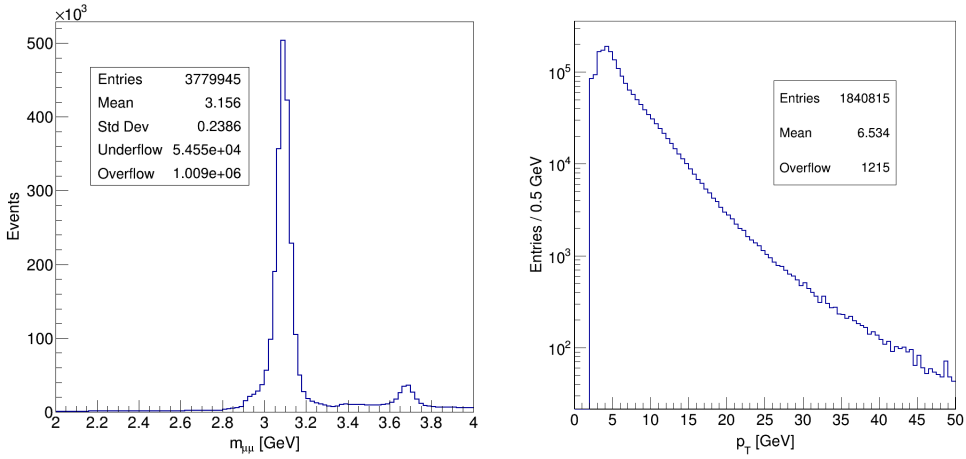

Figure 2.1: (Left) The invariant mass spectrum of preselected opposite-sign global muon pairs in the 2016F Charmonium data sample. (Right) The transverse momentum spectrum of the selected muons in the J/ψ mass peak.

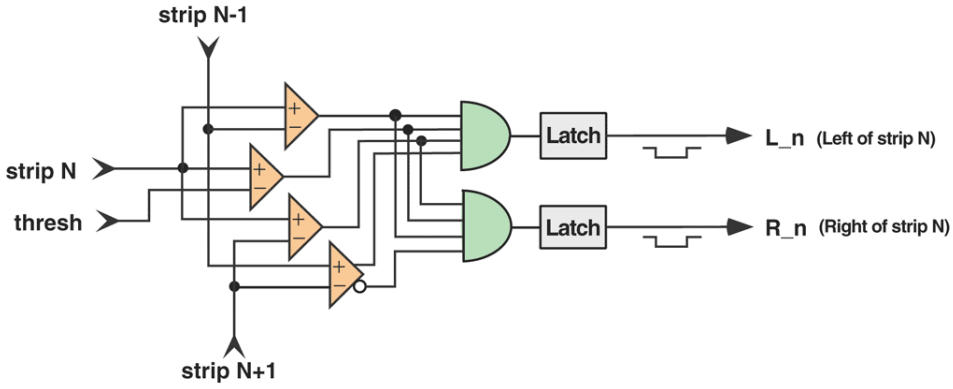

Figure 2.2: A simplified block diagram of the comparator array circuit used to generate digital half-strip trigger signals.

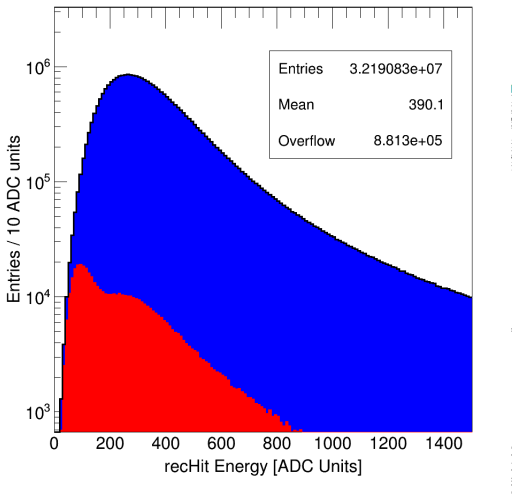

Figure 2.3: The distribution of recHit energies for recHits with and without matching comparators. The blue histogram is for recHits with a comparator, and accounting for 1.9% of the total the red is without a matching comparator.

the sum of two distributions. comparator performance by lowering the baseline threshold. The first being the same shape as the total distribution is due to events with missing TMB data blocks, investigated further in 2.3.2. The second distribution is the peak on the low end of the spectrum due to signals failing to go over the baseline threshold of the comparator array circuit. The current threshold is configured to be 30 mV in all chambers. It was observed that there was some potential room for improvement of Taking cosmic-ray data with an ME4/2 test stand revealed that thresholds down to 17 mV can be used without seeing any significant increase in noise. A similar test was done with the ME1/1 cosmic-ray test stand and changing the threshold did not seem to change anything. After some investigation it was discovered that the comparator threshold was hard coded in the control software to be set to 50 mV. This bug was fixed and similar results to what was observed in ME4/2 were produced. It was decided at this point to

implement this bug fix for official software release. The thresholds in ME1/1 were recuded to 30 mV in 2017 for CMS operation at Point 5.

The recHits were used to study the performance of the comparators by taking the difference of the positions of the two different position measurements. Each recHit is matched to the comparator which is closest to it in the relevant layer of the CSC chamber. RecHits in layers without any comparator data are categorized to be without a match. The difference in position of the comparator and recHit are shown in figure 2.4. The standard deviation of this distribution should be compared to $\frac{0.5}{\sqrt{12}} \approx 0.144$ which is the ideal for a square distribution of width 0.5. If the recHits measured the muon position perfectly, and the comparators measured muon position to on half-strip, the standard deviation of the distribution in figure 2.4 would be expected to be $\frac{0.5}{\sqrt{12}}$. The performance of the comparators is observed to be almost ideal in this case. Spikes seen at -0.25 and 0.25 corresponding to the cases where only one strip is over the pedestal, forcing the recHit to be fixed in the center of the strip. The offset of the mean from zero is due to the bias purposely built into the comparator circuit to avoid oscillation of the comparators while strips do not have any signal. RecHits were then put into three categories dependent on the energy deposit associated with them. RecHits with energy less than or equal to 200 ADC counts are called low energy (15.3%), energy from 200 to 500 ADC counts are called medium energy (61.2%), and above 500 ADC counts are high energy (23.5%). The comparator resolution for low energy recHits is shown in figure 2.5, medium energy in figure 2.5, and high energy shown in figure 2.6. It appears the single strip recHits are predominantly low energy. As the recHit energy increases the comparator resolution becomes more symmetric, the mean shifts towards zero, and the standard deviation decreases. The comparators are

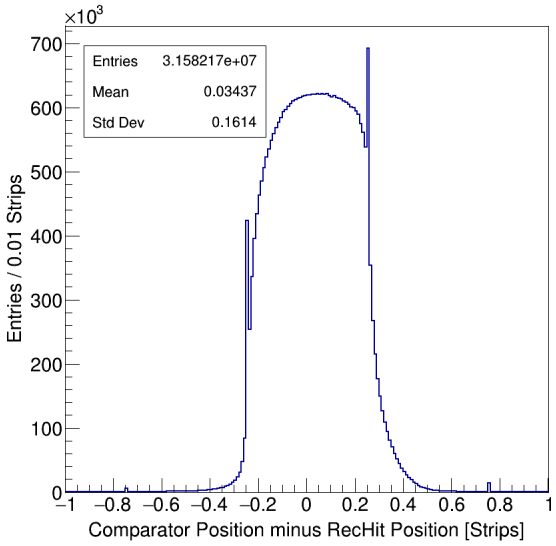

Figure 2.4: The difference in position of matched recHits and comparators. The recHit resolution is much better than the comparator resolution, so this can be used to study the performance of the comparator data.

performing as expected.

2.2.1 Comparison to Simulation

The comparator performance was compared to the simulated performance, as a check of the integrity of the simulation. A simulated sample was used which is based on a muon gun at the LHC interaction point set to launch muons with $p_T = 100$ GeV and flat in η and ϕ. The muons are simulated using a model of the CMS detector in Geant. These events undergo the standard reconstruction algorithms. Muons are selected if they pass the standard CMS tight Muon ID and have a transverse momentum over 20 GeV. The reconstructed transverse momentum of the selected muons is shown in figure 2.7. Then the comparator resolution was analyzed just as before with the Charmonium data. The CFEB

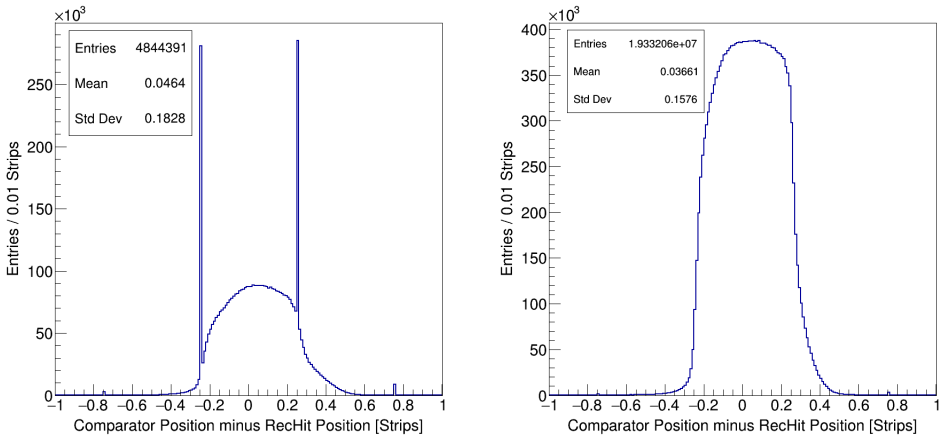

Figure 2.5: (Left) The comparator resolution for recHits with low energy deposit (ADC < 200). (Right) The comparator resolution for recHits with medium energy deposit (ADC 200-500).

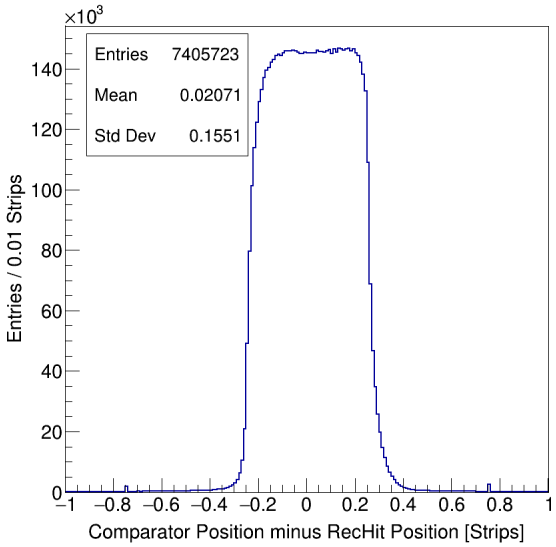

Figure 2.6: The comparator resolution for recHits with high energy (ADC $>$ 500).

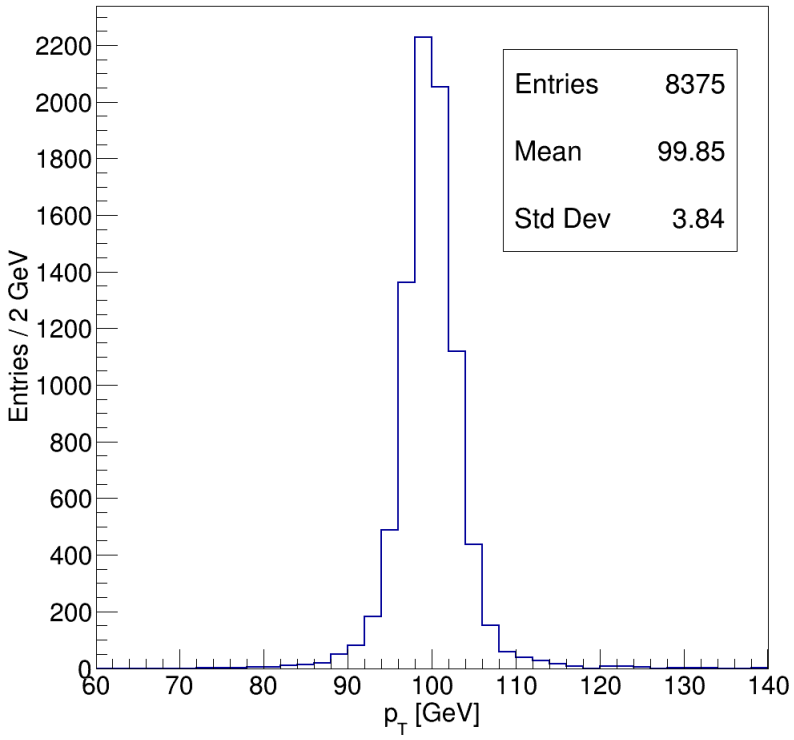

Figure 2.7: The transverse momentum spectrum of the selected muons from the muon gun data sample used for studying the comparator performance in simulation.

comparator resolution from the simulation is shown in figure 2.8. The RMS of this distribution is not far off from what is observed in data. It appears that the comparator bias has not been implemented in the simulation.

2.3 LCT Performance

The CSC recHits within a single chamber are then used to build CSC segments. Segments are ideally built from one recHit in each of the six layers. The segments are later used to build muon tracks. The trigger analog of the segment is the LCT. Pretriggers are generated by having two or more layers of cathode hits in one of

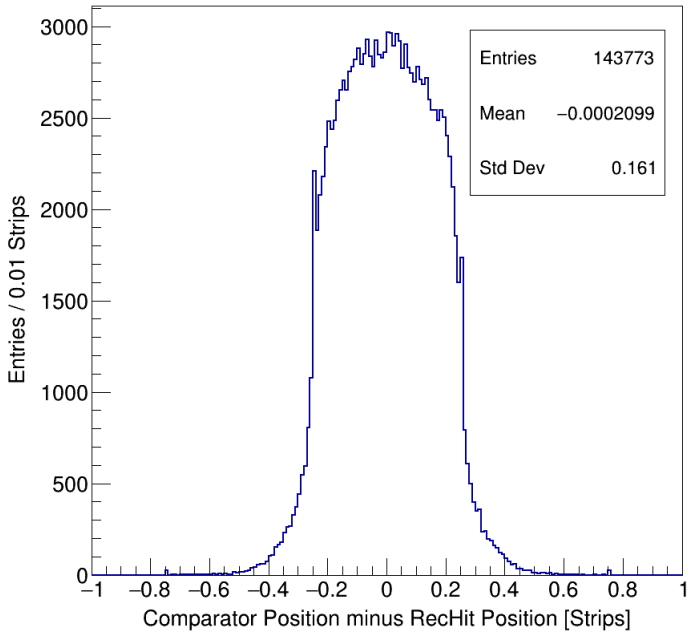

Figure 2.8: The simulated CFEB comparator resolution.

predefined the patterns. The pretriggers determine if the wire data will be buffered for the entire L1A delay or not. The wire data is sent from the AFEBs to the ALCT board where ALCT signals are generated by four or more time-coincident hits fitting within one of two predefined patterns. The ALCT signals are then sent to the TMB for further processing. The comparator data is transmitted from the CFEB to the TMB every bunch crossing. The TMB then looks for four or more layers of hits in one of the nine patterns shown in figure 2.9. The LCT is assigned the pattern ID (PID) of the CLCT used to build it. Figure 2.10 shows the PID distribution for the chambers associated with the selected muons from the J/ψ selection applied to the 2016F Charmonium data sample.

```
Hit pattern LUTs for 1 layer: - = don't care, xx= one hit or the other or both
Pattern       id=2        id=3        id=4        id=5        id=6        id=7        id=8        id=9        idA
Bend dir      bd=0        bd=1        bd=0        bd=1        bd=0        bd=1        bd=0        bd=1        bd=0
|           |           |           |           |           |           |           |           |
ly0      --------xxx xxx-------- -------xxx- -xxx------- ------xxx-- --xxx------ -----xxx--- ---xxx----- ----xxx----
ly1      ------xx--- ---xx------ ------xx--- ---xx------ -----xx---- ----xx----- -----xx---- ----xx----- -----x-----
ly2 key  -----x----- -----x----- -----x----- -----x----- -----x----- -----x----- -----x----- -----x----- -----x-----
ly3      ---xxx----- -----xxx--- ---xx------ ------xx--- ----xx----- -----xx---- ----xx----- -----xx---- -----x-----
ly4      -xxx------- -------xxx- -xxx------- -------xxx- ---xx------ ------xx--- ---xxx----- -----xxx--- ----xxx----
ly5      xxx-------- --------xxx -xxx------- -------xxx- --xxx------ ------xxx-- ---xxx----- -----xxx--- ----xxx----
|           |           |           |           |           |           |           |           |
// Extent   0123456789A 0123456789A 0123456789A 0123456789A 0123456789A 0123456789A 0123456789A 0123456789A 0123456789A
// Avg.bend - 8.0 hs    + 8.0 hs    -6.0 hs     +6.0 hs     -4.0 hs     +4.0 hs     -2.0 hs     +2.0 hs      0.0 hs
// Min.bend -10.0 hs    + 6.0 hs    -8.0 hs     +4.0 hs     -6.0 hs     +2.0 hs     -4.0 hs      0.0 hs     -1.0 hs
// Max.bend - 6.0 hs    +10.0 hs    -4.0 hs     +8.0 hs     -2.0 hs     +6.0 hs      0.0 hs     +4.0 hs     +1.0 hs
```

Figure 2.9: These are the nine patterns used in firmware to generate cathode pretriggers and CLCTs.

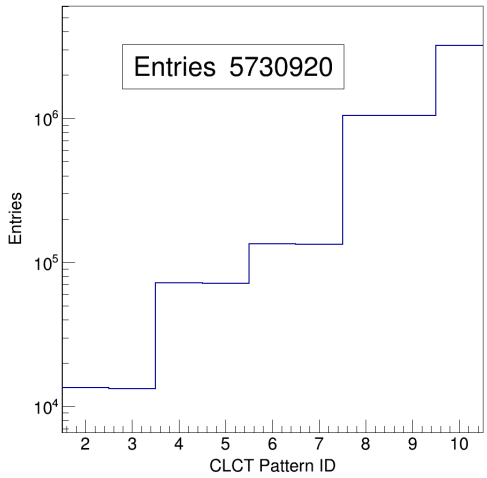

Figure 2.10: The distribution of PIDs for LCTs in chambers associated to muons from J/ψ decays in the 2016F Charmonium data sample.

2.3.1 CLCT Pattern ID Performance

When a CLCT is generated in time with an ALCT, the TMB forms an LCT which is then transmitted to the trigger processor to be built into tracks with the LCTs from all of the chambers in the muon system. Each TMB can send up to two LCTs per bunch crossing. The CLCT bending angle should be correlated with the muon transverse momentum. Every segment on the selected muons is matched to the closest LCT in the same chamber for that event. These matches are then categorized by the CLCT PID associated with that LCT. A set of efficiencies, defined as number of segments with matched LCTs with a CLCT PID greater than or equal to N, is than used to study the capability of the CLCT patterns to measure transverse momentum. Figure 2.11 shows the result for ME1/2 chambers, where the bending power of the solenoid is relatively large, and the most ideal performance was observed. Patterns eight and above are sufficient to give efficiency for transverse momentums down to about 8.0 GeV and patterns six and above are3 sufficient to give good efficiency down to about 5.0 GeV. The more typical pattern activity is shown in figure 2.12. It is seen that patterns eight and above are sufficiently efficient down to 2.0 GeV in most CSC chambers. The distribution of the segment slope was also studied for each CLCT PID. The units of the segment slope and angle were changed from cm to strip and layer units. The distribution of segment slope in x-z plane, for ME1/2, is shown in figure 2.13. The segment slope distribution for ME2/1 is shown in figure 2.14. The larger angular distribution in ME1/2 is expected because of the result in efficiency as a function of transverse momentum. A summary of all the segment slope distributions is shown in figures 2.15 and 2.16. The position resolution of the LCTs is studied by comparing the key half-strip (KHS) of the CLCT to the segment position. The

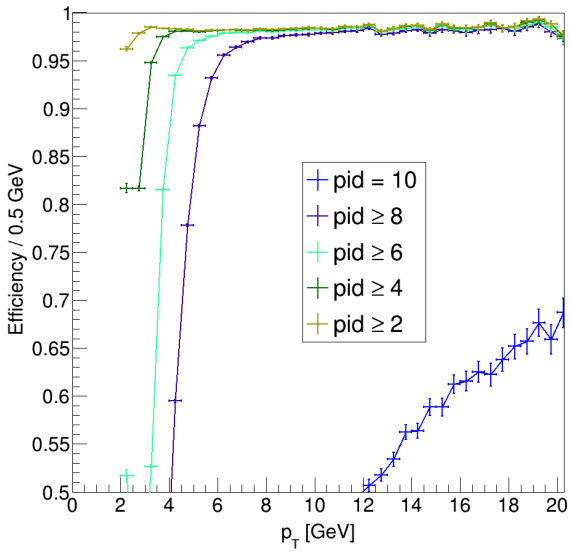

Figure 2.11: The LCT efficiency on segments with different PID thresholds for ME1/2, where the best pattern performance is observed. Patterns six and above alone have high efficiency down to a transverse momentum of about 5.0 GeV.

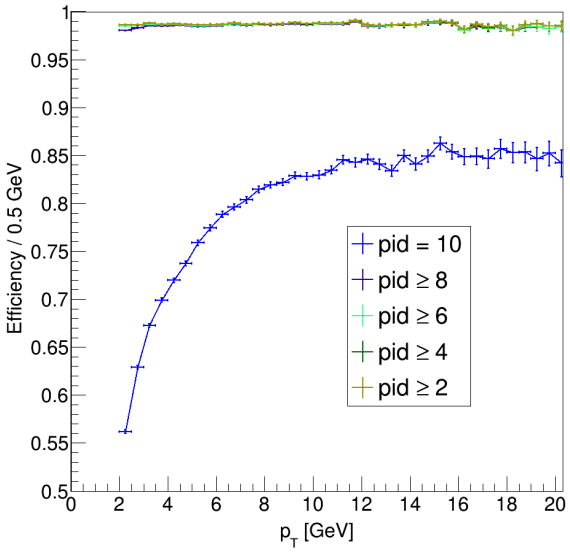

Figure 2.12: The LCT efficiency on segments with different PID thresholds for ME2/1, comparable to most CSC chambers.

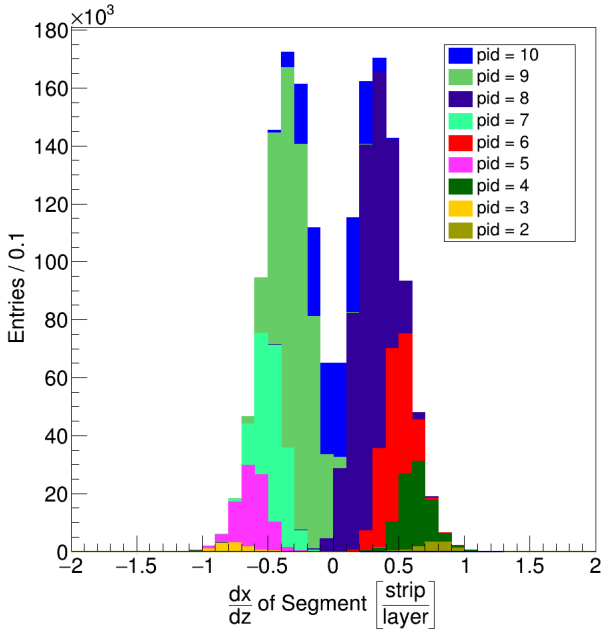

Figure 2.13: The segment slope distribution in ME1/2 broken up by associated CLCT PID.

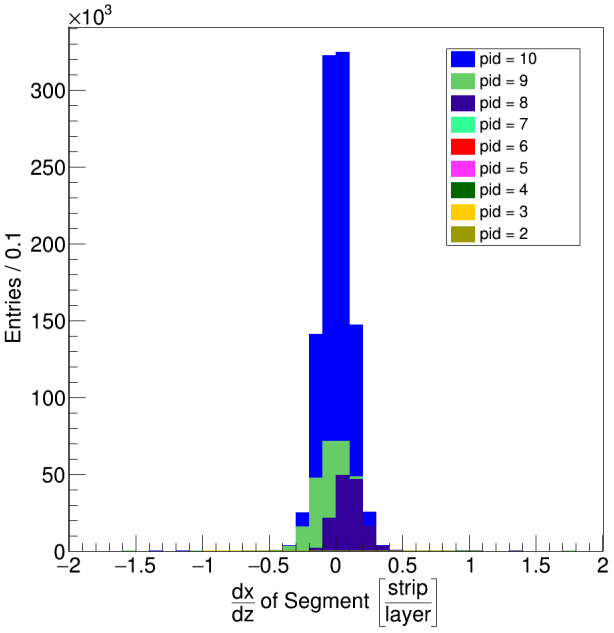

Figure 2.14: The segment slope distribution in ME2/1 broken up by associated CLCT PID.

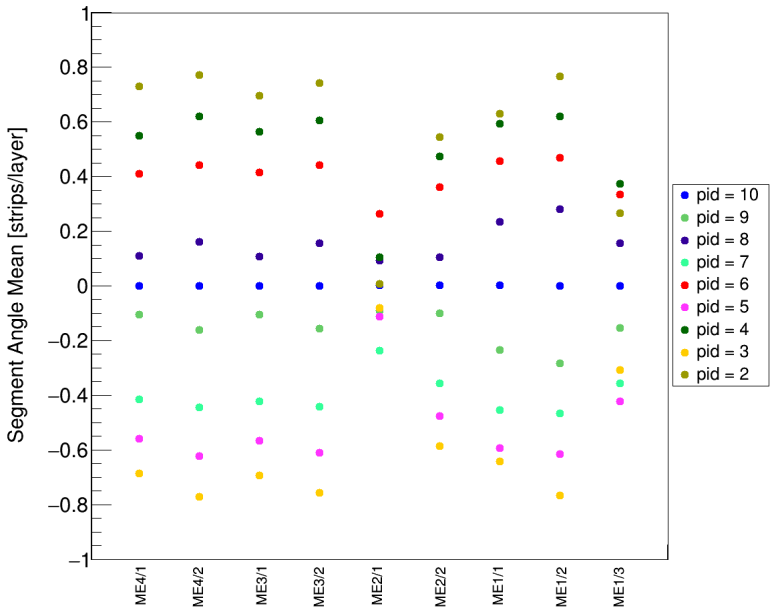

Figure 2.15: The mean of the segment slope distributions of each CLCT PID separated by chamber type.

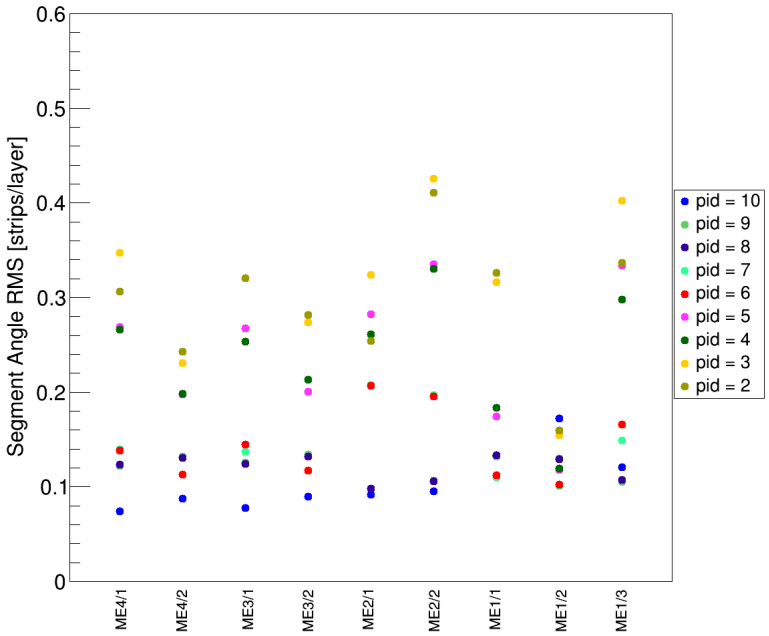

Figure 2.16: The RMS of the segment slope distributions of each CLCT PID separated by chamber type.

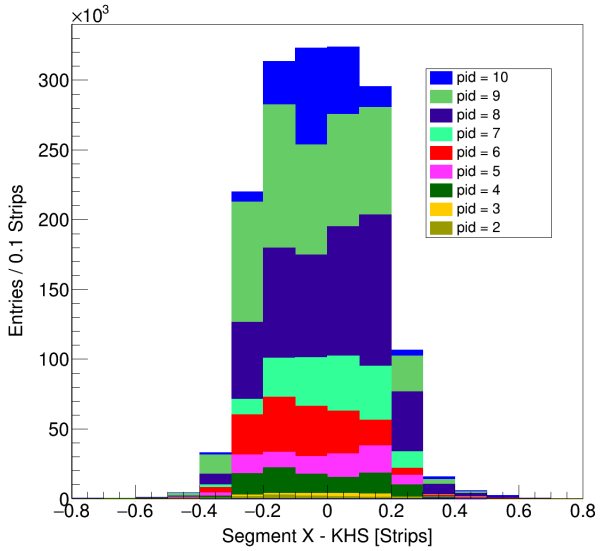

Figure 2.17: The difference of CLCT KHS and segment position in ME1/2 broken up by associated CLCT PID.

difference in position of LCT and segment is shown in figure 2.17 for ME1/2. The resolution looks reasonable overall, but one might expect better than half strip resolution with six layers of hits going into this trigger primitive. The position resolution in ME2/1 is shown in figure 2.18. The overall resolution here looks comparable to ME1/2, but patterns 8 and 9 do seem to be concentrated at the edge of the half-strip more than the middle. In comparison, pattern ten seems to be concentrated at the center of the half-strip. Similar behavior of these patterns is observed in most other chambers. A summary of all of the position resolution distributions is shown in figures 2.19 and 2.20.

2.3.2 LCT Multiplicities

The LCT data is transmitted to the emTF to be built into tracks, and both systems record the LCTs upon receipt of an L1A. Ideally, both systems would

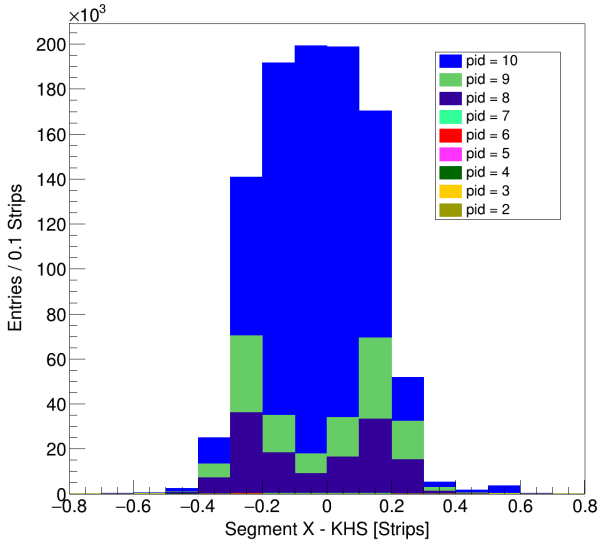

Figure 2.18: The difference of CLCT KHS and segment position in ME2/1 broken up by associated CLCT PID.

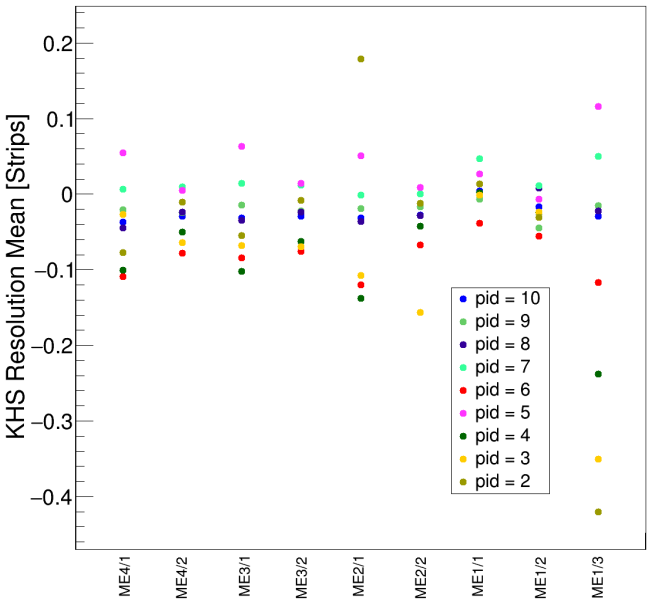

Figure 2.19: The mean of the LCT position resolutions distributions for each CLCT PID separated by chamber type.

Table 2.1: A summary of the performance of the LCTs broken up by chamber type and Pattern IDs for Station 1.

Ring	**PID**	$\langle$**dx/dz**$\rangle$	$\mathbf{dx/dz_{RMS}}$	$\langle$**x**$\rangle$	$\mathbf{x_{RMS}}$	**%**
1	10	0.0014	0.13	0.0041	0.13	30
1	9	-0.23	0.13	-0.0069	0.18	28
1	8	0.23	0.13	0.027	0.19	27
1	7	-0.46	0.11	0.047	0.15	5.1
1	6	0.46	0.11	-0.039	0.16	5.1
1	5	-0.59	0.17	0.027	0.2	2.3
1	4	0.59	0.18	0.0023	0.2	2.3
1	3	-0.64	0.32	-0.001	0.27	0.34
1	2	0.63	0.33	0.014	0.24	0.35
2	10	0.00018	0.17	-0.017	0.13	13
2	9	-0.28	0.13	-0.045	0.17	29
2	8	0.28	0.13	0.0079	0.17	29
2	7	-0.47	0.1	0.011	0.14	9.5
2	6	0.47	0.1	-0.055	0.16	9.5
2	5	-0.62	0.12	-0.0066	0.17	4.9
2	4	0.62	0.12	-0.024	0.18	4.9
2	3	-0.77	0.15	-0.023	0.19	0.6
2	2	0.77	0.16	-0.031	0.2	0.63
3	10	-0.00061	0.12	-0.029	0.14	53
3	9	-0.15	0.11	-0.015	0.2	23
3	8	0.16	0.11	-0.022	0.2	23
3	7	-0.36	0.15	0.05	0.18	0.27
3	6	0.33	0.17	-0.12	0.24	0.32
3	5	-0.42	0.33	0.12	0.53	0.083
3	4	0.37	0.3	-0.24	0.54	0.089
3	3	-0.31	0.4	-0.35	0.68	0.022

Table 2.2: A summary of the performance of the LCTs broken up by chamber type and Pattern IDs for Station 2.

Ring	**PID**	$\langle dx/dz \rangle$	dx/dz_{RMS}	$\langle x \rangle$	x_{RMS}	%
1	10	0.0014	0.091	-0.031	0.14	73
1	9	-0.09	0.098	-0.019	0.21	13
1	8	0.093	0.098	-0.036	0.2	13
1	7	-0.24	0.21	-0.0014	0.39	0.074
1	6	0.26	0.21	-0.12	0.38	0.077
1	5	-0.11	0.28	0.051	0.86	0.039
1	4	0.11	0.26	-0.14	0.98	0.04
1	3	-0.081	0.32	-0.11	0.84	0.011
1	2	0.0074	0.25	0.18	1.1	0.0097
2	10	0.0012	0.095	-0.028	0.14	71
2	9	-0.1	0.11	-0.017	0.2	14
2	8	0.11	0.11	-0.029	0.2	15
2	7	-0.36	0.2	2.5e-05	0.25	0.19
2	6	0.36	0.2	-0.067	0.29	0.21
2	5	-0.48	0.34	0.0087	0.51	0.11
2	4	0.47	0.33	-0.042	0.51	0.11
2	3	-0.59	0.43	-0.16	0.45	0.033
2	2	0.54	0.41	-0.012	0.46	0.026

Table 2.3: A summary of the performance of the LCTs broken up by chamber type and Pattern IDs for Station 3.

Ring	**PID**	$\langle \mathbf{dx/dz} \rangle$	$\mathbf{dx/dz_{RMS}}$	$\langle \mathbf{x} \rangle$	$\mathbf{x_{RMS}}$	%
1	10	0.00039	0.077	-0.032	0.15	75
1	9	-0.1	0.13	-0.014	0.2	12
1	8	0.11	0.12	-0.035	0.2	12
1	7	-0.42	0.14	0.015	0.2	0.43
1	6	0.41	0.14	-0.085	0.23	0.46
1	5	-0.57	0.27	0.064	0.38	0.26
1	4	0.56	0.25	-0.1	0.37	0.27
1	3	-0.69	0.32	-0.068	0.34	0.059
1	2	0.7	0.32	-0.054	0.34	0.068
2	10	0.00089	0.09	-0.029	0.15	64
2	9	-0.16	0.13	-0.022	0.2	16
2	8	0.16	0.13	-0.023	0.19	16
2	7	-0.44	0.12	0.012	0.17	1.2
2	6	0.44	0.12	-0.076	0.18	1.2
2	5	-0.61	0.2	0.014	0.28	0.58
2	4	0.6	0.21	-0.062	0.3	0.61
2	3	-0.76	0.27	-0.07	0.29	0.12
2	2	0.74	0.28	-0.0078	0.3	0.13

Table 2.4: A summary of the performance of the LCTs broken up by chamber type and Pattern IDs for Station 4.

Ring	**PID**	$\langle \mathbf{dx/dz} \rangle$	$\mathbf{dx/dz_{RMS}}$	$\langle \mathbf{x} \rangle$	$\mathbf{x_{RMS}}$	%
1	10	-0.0012	0.074	-0.037	0.15	74
1	9	-0.1	0.12	-0.021	0.2	13
1	8	0.11	0.12	-0.045	0.2	12
1	7	-0.42	0.14	0.0068	0.2	0.43
1	6	0.41	0.14	-0.11	0.2	0.46
1	5	-0.56	0.27	0.055	0.39	0.23
1	4	0.55	0.27	-0.1	0.39	0.24
1	3	-0.69	0.35	-0.027	0.41	0.056
1	2	0.73	0.31	-0.077	0.39	0.062
2	10	0.0012	0.088	-0.029	0.15	62
2	9	-0.16	0.13	-0.024	0.2	16
2	8	0.16	0.13	-0.024	0.2	17
2	7	-0.44	0.11	0.0099	0.17	1.3
2	6	0.44	0.11	-0.078	0.18	1.4
2	5	-0.62	0.2	0.0051	0.26	0.71
2	4	0.62	0.2	-0.05	0.27	0.76
2	3	-0.77	0.23	-0.064	0.25	0.18
2	2	0.77	0.24	-0.01	0.26	0.18

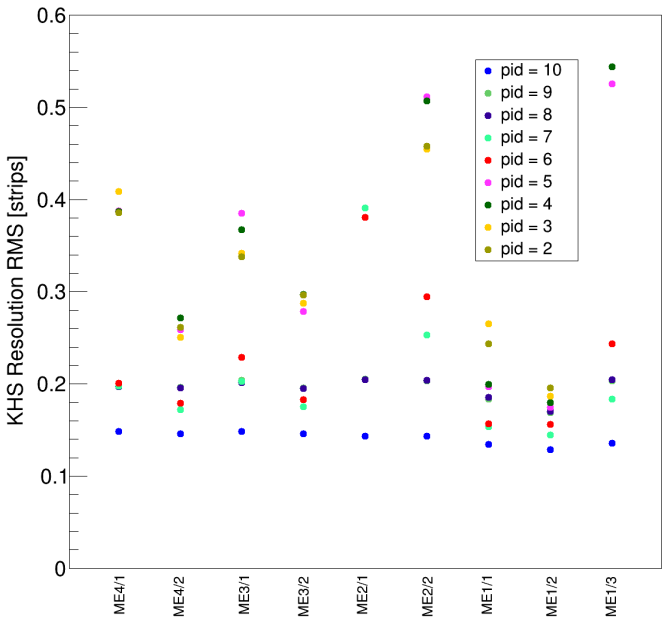

Figure 2.20: The RMS of the LCT position resolution distributions for each CLCT PID separated by chamber type.

record identical LCT data for every event. The number of LCTs in chambers with segments in the selected muon tracks were counted from the TMBs and the emTF. It was discovered that these sets do not match as expected, shown in figure 2.21. Any off-diagonal entry is considered an error. The errors are categorized into three main categories. The first category, affecting 0.07% of the selected muons, is when the number of LCTs from the emTF is two or more and the number of LCTs from the TMBs is one less. The second category, affecting 0.685% of the selected muons, is when the number of LCTs from the TMBs is two or more and the number of LCTs from the emTF is one less. The third main error category, affecting 0.248% of the selected muons, is when the number of LCTs from the emTF is two or more and the number of LCTs from the TMBs is zero. The third category is the most interesting case because it is the least understood failure

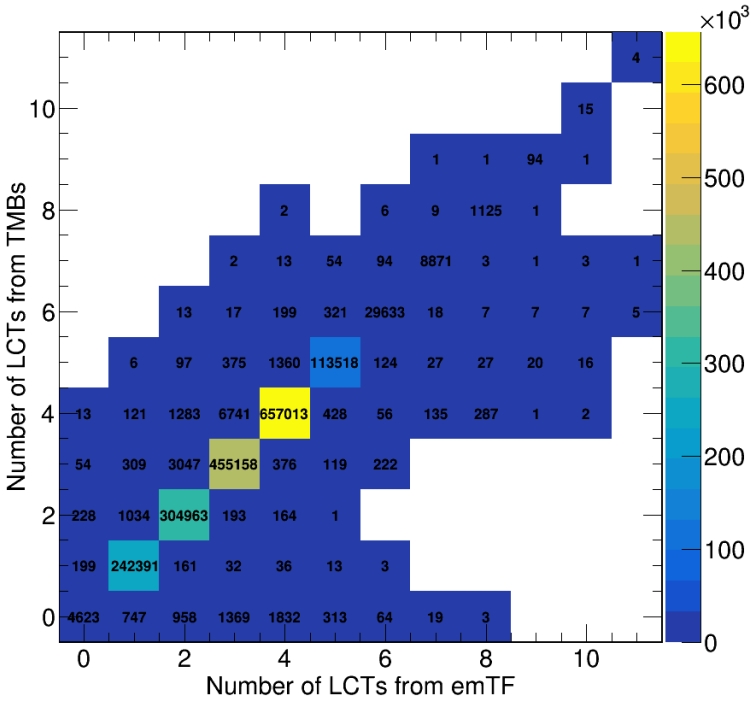

Figure 2.21: A comparison of the LCT data as recorded by the TMBs and by the emTF.

mode.

Category I errors appear to happen more in inner ring chambers than outer ring. The errors have no correlation in muon transverse momentum or ϕ, so these errors could potentially be due to the higher rates in these chambers. Category II errors are observed to have no correlation with muon transverse momentum or ϕ. Category III errors are the most concerning of the three types of errors being made in recording LCTs. They appear everywhere in the CSC system with equal probability, checked by looking at the muon kinematic distributions in this case normalized by the entire set of selected muons. In an attempt to understand what could be causing the Category III errors, the timing distributions of the LCTs in

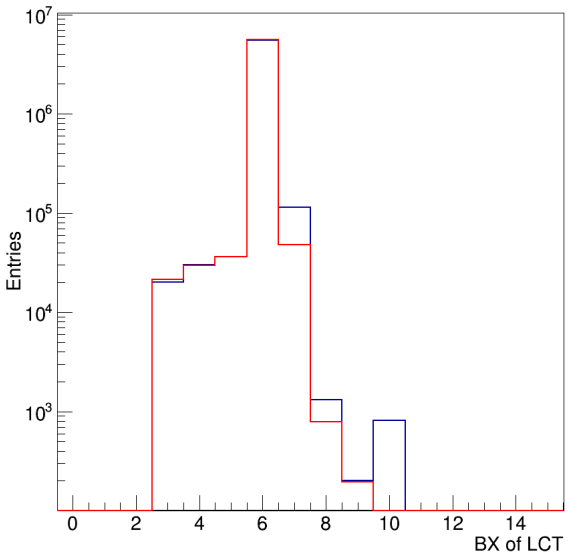

Figure 2.22: LCT timing distribution from the emTF in blue and cscTF in red in units of LHC clock cycles (BX) of 25 ns.

the emTF and legacy cscTF are shown in figure 2.22 and for Category III error muons in figure 2.23. Carefully examining the raw data for these events revealed that the TMB data blocks were missing entirely. It is observed that these events specifically come 3 bunch crossings after a bunch crossing also with an L1A. The TMB is only capable of reading out any given LCT for a single L1A. The proximity in time of the earlier L1A means it is "stolen" from the correct, later L1A. This also explain the 0.248% of muons where this is observed, since the occupancy of an L1A in a bunch crossing is 100 kHz divided by 40 MHz. The timing of the LCTs in the emTF is one bunch crossing later than in the cscTF for the Category III error muons, as seen in figure 2.23. This could be due to a bug in the emTF firmware when the LCTs are duplicated for multiple L1A readouts.

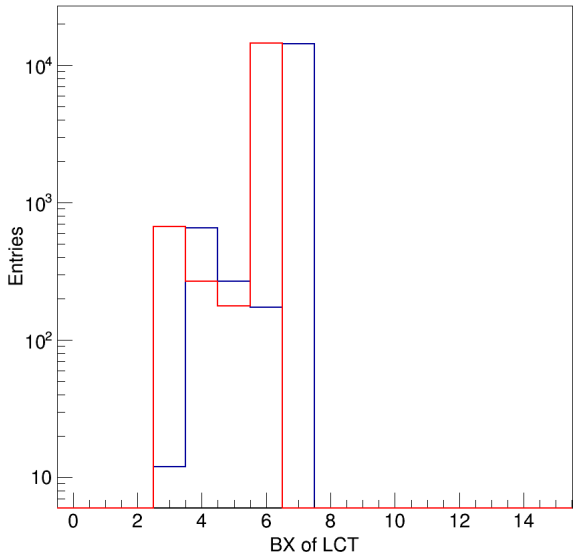

Figure 2.23: LCT timing distribution, for category III error muons, from the emTF in blue and cscTF in red in units of LHC clock cycles (BX) of 25 ns.

CHAPTER 3

Improvement of the Endcap Muon Trigger with GEM Detectors

In an effort to increase the total integrated luminosity that can be potentially collected from the LHC, the injector chain will be upgraded during the second long shutdown (LS2, 2019-2020). The instantaneous luminosity is expected to reach upwards of $2 \cdot 10^{34} cm^{-2} s^{-1}$. After the third long shutdown (LS3, 2024-2026), the high-luminosity (HL-LHC) upgrade is planned to reach instantaneous luminosities reaching $10^{35} cm^{-2} s^{-1}$. CMS plans to install Gas Electron Multiplier (GEM) chambers in the forward region as shown in figure 3.1 in order to keep up with the increased demands of operating at these higher luminosities.

GEM detectors utilize metal-insulator-metal foil with tiny holes etched through it in a lattice. Figure 3.2 shows an SEM image of a hole and a schematic view of the electric field lines within a hole. These foils provide the electric field configuration necessary to achieve a gas-gain in signal from the initial electrons ionized by the primary charged particle. The volume inside of the holes in the foils contain a large electric field which accelerate electrons and create an avalanche of ionization of the gas. This process is known as a Townsend avalanche, the same process by which wire chambers achieve a gas-gain.

Multiple GEM foils can be used in the same chamber to achieve a higher gas-

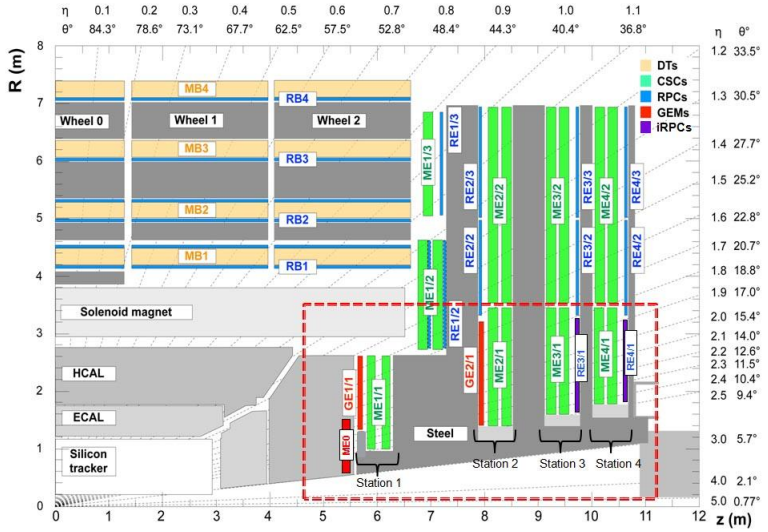

Figure 3.1: Shown inside the dased box are the new muon detectors to be installed in CMS in preparation for the HL-LHC upgrades. [CSS15]

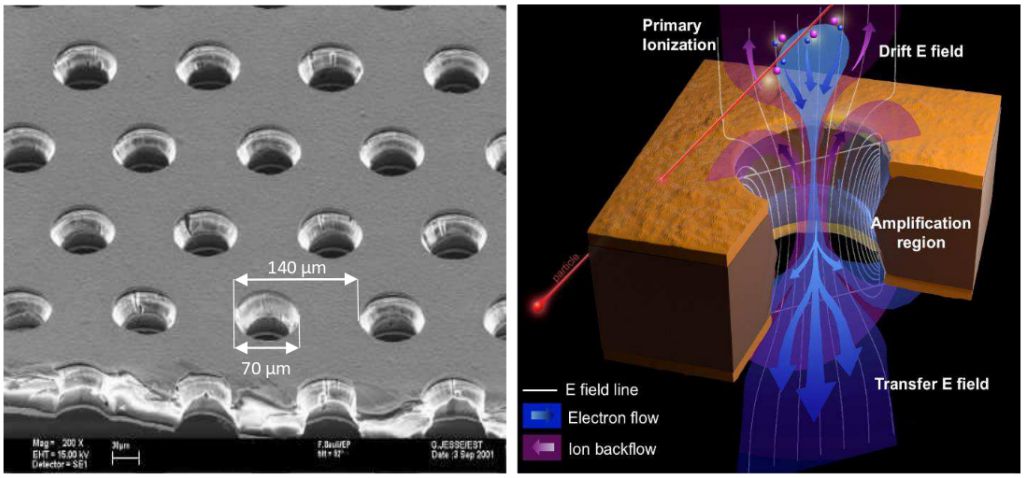

Figure 3.2: Scanning Electron Microscope (SEM) picture of a GEM foil (left) [Sau97] and schematic view of the electric field lines (white), electron flow (blue), and ion flow (purple) through a bi-conical GEM hole (right). The outer diameters of the hole are 70 microns and the inner diameter is 50 microns; the hole pitch is 140 microns. [CSS15]

gain in stages. The CMS experiment has designed chambers which utilize three foils to achieve a gas-gain on the order of 10^4. The triple-GEM design is shown in more detail in figure 3.3. The electrons are emitted from the final foil towards a copper layer segmented into strips. Each strip is individually connected to a read-out chip known as a VFAT. As the electrons approach the strip, a current is induced into the read-out channel of the VFAT. Each channel of the chip has a charge-sensitive preamplifier, a shaper, and a constant fraction discriminator. The readout of a channel is binary, latched for a configurable number of 25 ns bunch crossing intervals. The VFAT also transmits data with half of the resolution for immediate use by the trigger, and buffers the data in memories for sequential readout to the data acquisition system, as detailed below.

3.1 GE1/1 Detectors

The first station of GEM chambers planned to be installed is GE1/1. The location of this station in CMS can be seen in figure 3.1. The design of these detectors is nearing completion, and production of the chambers has already begun. This station of GEM detectors is planned to be installed during Long Shutdown two of the LHC, which will be 2019-2020. These chambers could become key to improving the forward muon momentum resolution at low momentum, which would be key to decreasing the muon trigger rate. Each chamber will cover a 10 degree portion of the ring, and will have 3072 channels segmented into eight η sectors. An exploded view of the mechanical design is shown in figure 3.4. A picture of the first prototype detector for GE1/1 is shown in figure 3.6. These chambers are read out by a sophisticated mixed-signal electronics system. A block diagram of

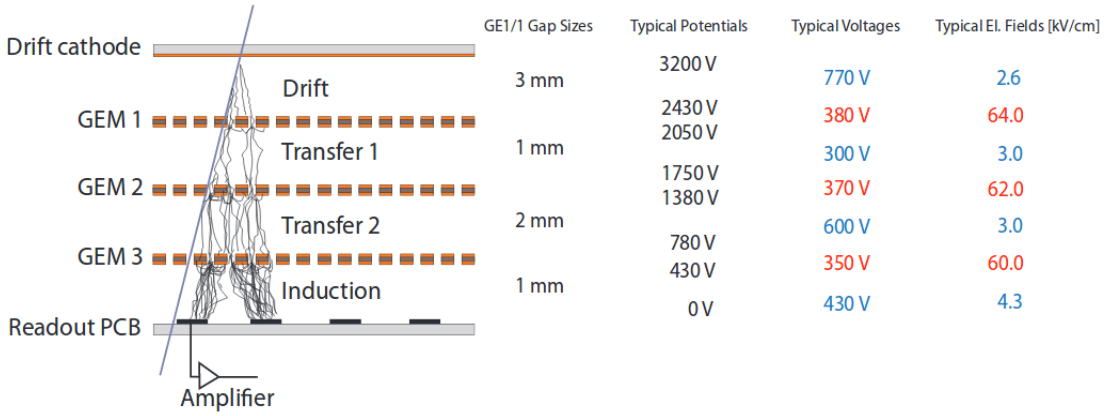

Figure 3.3: Principle of operation of a generic triple-GEM chamber and definition of drift, transfer, and signal induction gap regions within the detector [Sau97]. The columns on the right give the actual gap sizes in the GE1/1. They also list typical values for electric potentials on the seven electrodes and typical values for voltages and electric fields across the four gaps (blue) and the three foils (red) if the nominal potential of 3200 V for operation in Ar/CO_2 70:30 is applied to the drift cathode. [CSS15]

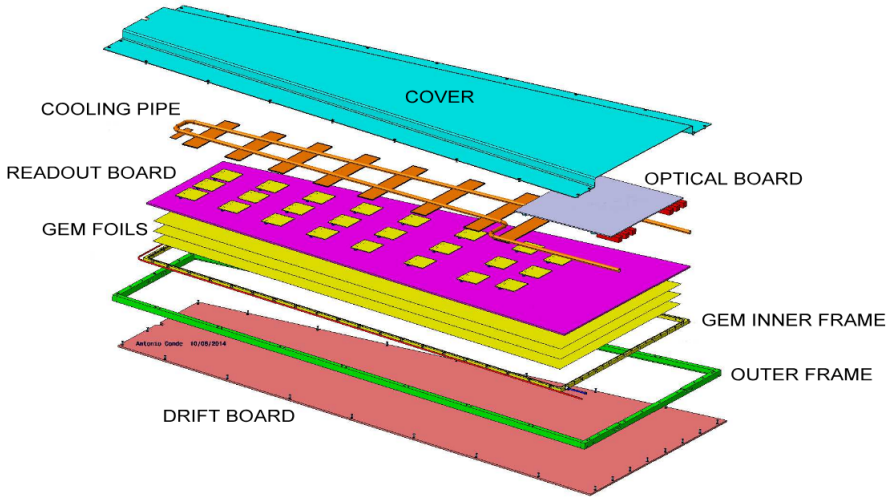

Figure 3.4: Exploded view of the mechanical design of a single GE1/1 chamber. [CSS15]

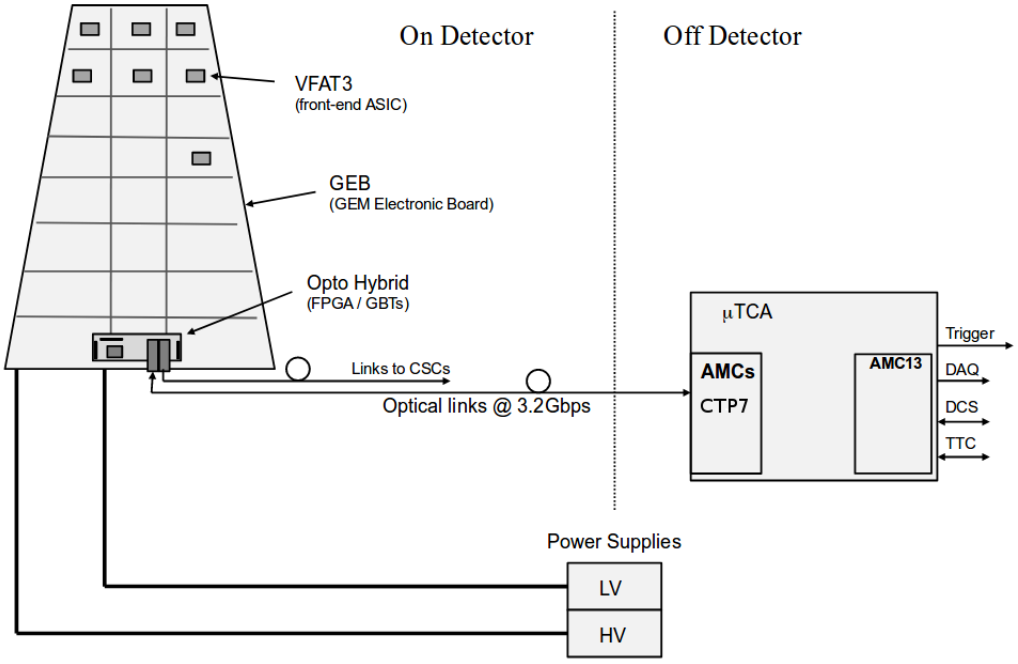

Figure 3.5: The GEM electronics readout system. [CSS15]

this system is shown in figure 3.5.

The gas mixture used in the chamber is 70% Ar and 30% CO_2. It was measured at a test beam with an early prototype of the system at an operational gas-gain of about $8,000$, about 4% of hits have less than 2 fC of charge deposited on a strip, with a most probable value of 4 fC. Thus 2 fC is a good first order approximation of the threshold which much be achieved to operate the detector efficiently.

3.2 VFAT3 ASIC: Calibration

We must calibrate and characterize the performance of the front end ASIC to determine if sufficient threshold value is achievable in the system for efficient operation of the detector. The strips of the read-out board are connected to the input channels of a VFAT3 ASIC. Each of these chips can handle 128 channels,

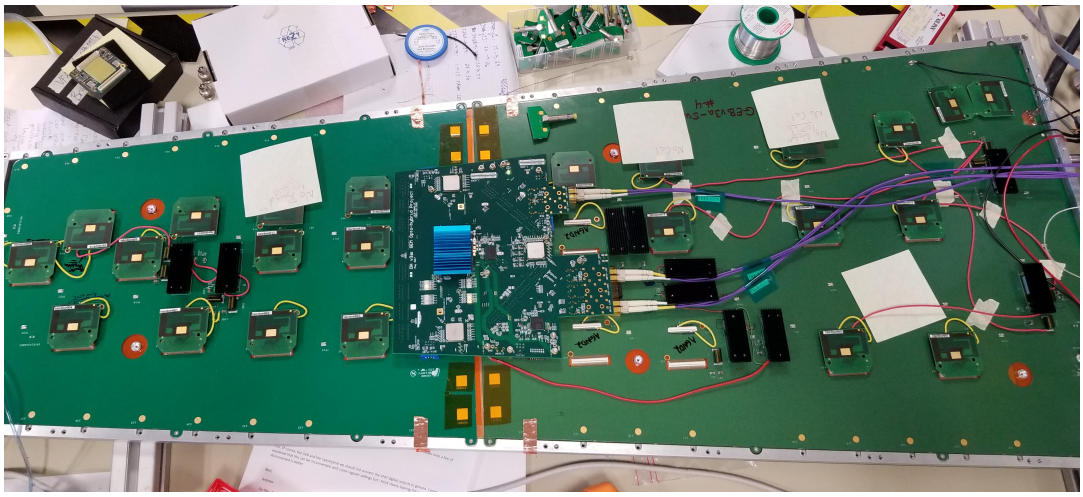

Figure 3.6: The first prototype of the GEM detectors that will be installed as the GE1/1 system. This is not the final version of electronics to be used, but building this system revealed several issues to be addressed in the next iteration of the electronics design.

so 24 VFATs are used in total to read out an entire chamber. A block diagram of the ASIC is shown in figure 3.7. The VFAT3 has two pads to take critical voltages and currents off of the chip for external calibration. There also exist two internal analog-to-digital converters (ADC) which need calibration to aid in configuration of the chip.

The bias block of the VFAT3 is shown in figure 3.8 and the bias block in figure 3.9. The first step in calibration of the chip is to determine the operational setting of the I_{ref} digital-to-analog converter (DAC). This is the reference current used to bias all other DACs on the chip. The nominal value of this current is 10 µA. First, the chip is configured to push this current out of a pad. This DAC has an output scale factor of 0.5 (scaled between the calibration and bias blocks), therefore we search for the 6-bit value which sets the current out of this pad closest to 5 µA,

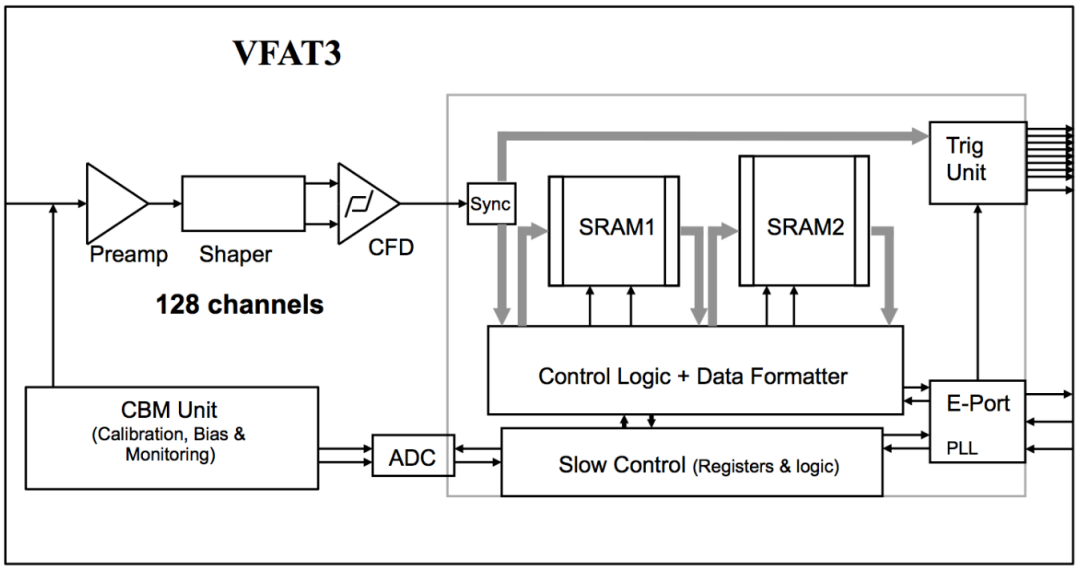

Figure 3.7: VFAT3 block diagram. [CSS15]

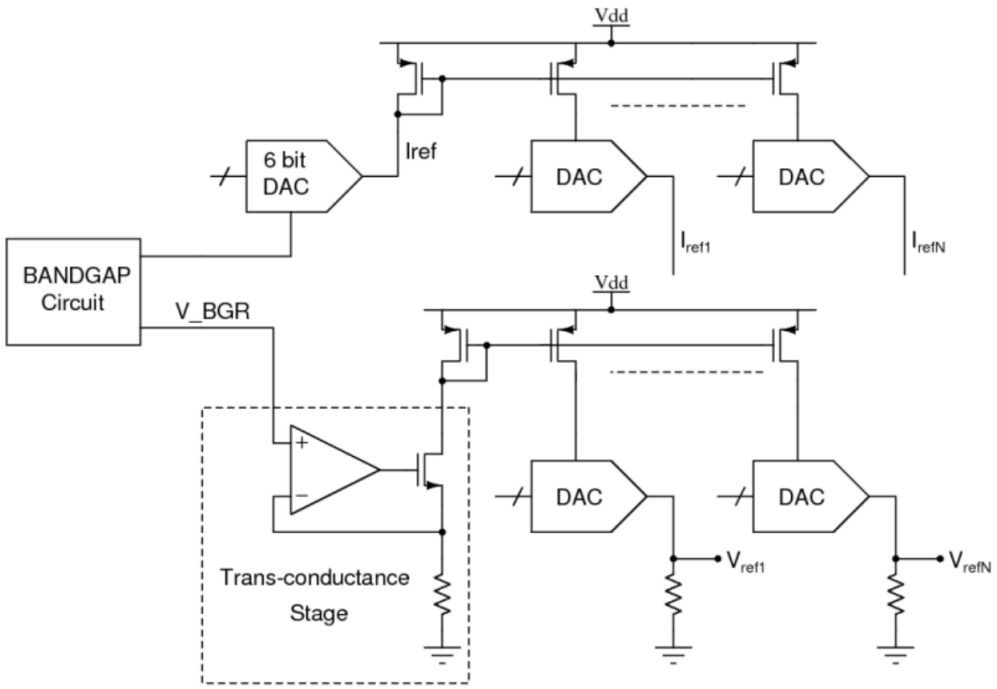

Figure 3.8: The architecture of the VFAT3 bias block.

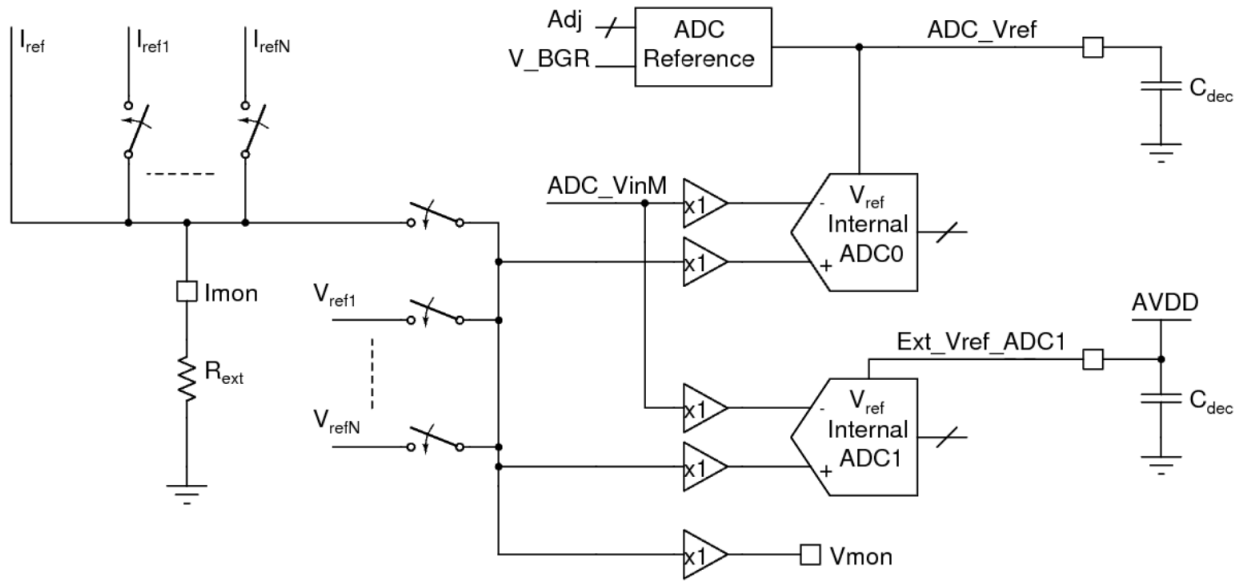

Figure 3.9: The architecture of the VFAT3 monitoring block.

typically around a DAC value of 30.

The next step is to configure the internal voltage reference for the first ADC (ADC0). This is done by configuring the chip to output this voltage on the monitoring pad. Then the value of this 2-bit DAC is found which brings the voltage closest to 1 V. Then a DAC is chosen with a large range of output voltages. This DAC voltage is configured to be routed to the monitoring circuit. All of the DAC values are scanned and both ADC values from the chip are recorded as well as the voltage using an external ADC. An example of a result from this procedure is shown in figure 3.10. The second internal ADC (ADC1) is referenced with a voltage provided externally.

The VFAT3 has an internal pulse generator which can be calibrated to deliver a known amount of charge to any channel. The circuit used to generate calibration pulses is shown in figure 3.11. The two voltages of the step are measured via the

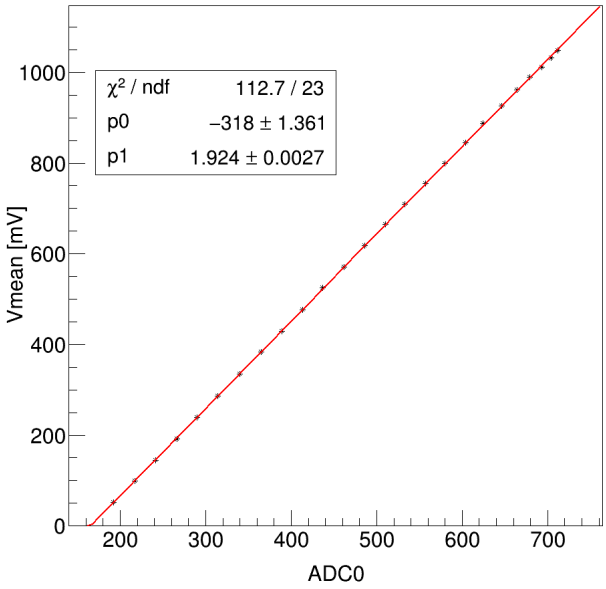

Figure 3.10: The data and fit used to calibrate the VFAT3 internal ADC response. The parameters from the linear fit are then used as calibration constants for this chip. No point deviates from the fit more than a few percent.

internal ADC for all valid inputs of the CAL DAC. The low and high voltage of the step are sampled and the difference in voltage is calculated. This voltage is then used with an assumption of a 100 fF capacitor to calculate the amount of charge delivered for a particular configuration of the CAL DAC. A linear fit is then performed to characterize the behavior of the pulse circuit. An example of a result of this procedure is shown in figure 3.12. The technology used to build the chip offers metal-insulator-metal capacitor structures which can provide a capacitance known to 1%. This allows us to accurately know the amount of charge delivered to a channel, as long as only a single channel is configured to receive a pulse.

Now that we can accurately deliver a configurable amount of charge to any channel, we can begin to characterize the performance of the VFAT3. The first step is to time in the read-out of SRAM1 into SRAM2 (see figure 3.7) with respect

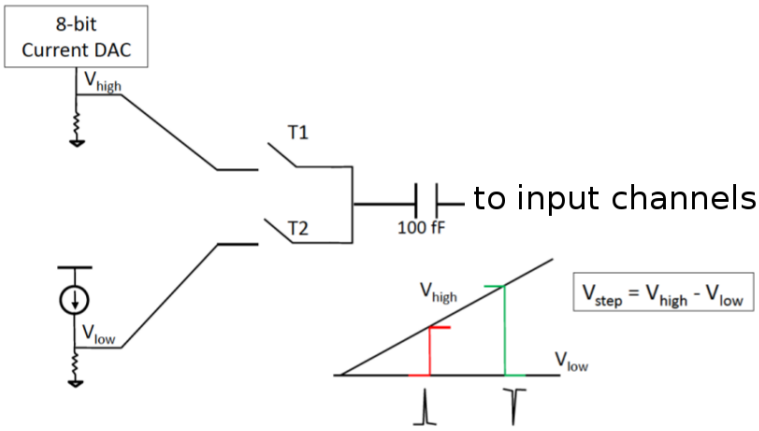

Figure 3.11: The operational principle used to generate calibrated pulses for the input channels of the VFAT3 ASIC.

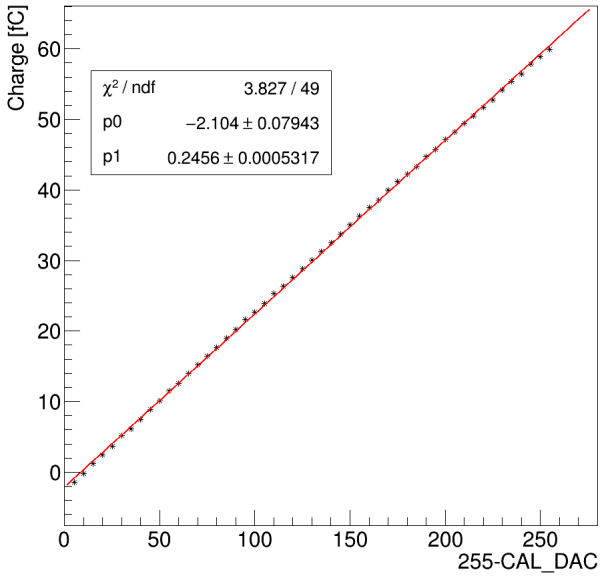

Figure 3.12: The result of calibrating the internal pulse injection circuitry of the VFAT3 ASIC. The chip was purposely designed to guarantee the zero charge crossing within the output range of the CAL DAC. No point deviates from the fit more than a few percent.

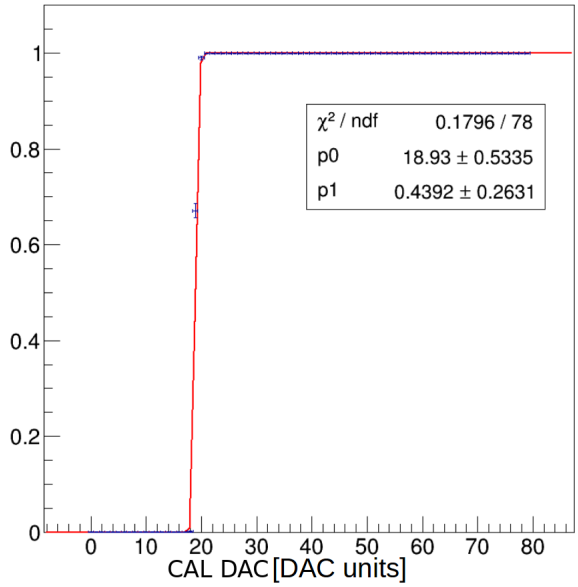

Figure 3.13: An s-curve for a single channel of a VFAT3 ASIC. The y-axis is the fraction of 1,000 pulses read out as a hit.

to arrival of the calibration pulse and level one accept (L1A) signal. This timing is configured with the "latency" register of the chip, and is the difference between the write and read pointer values. This is functionally the number of 40 MHz clock cycles between entry of data into SRAM1 and its read-out with an L1A signal. Now that the Data Acquisition (DAQ) block of the chip is timed in, we inject 1,000 pulses into a channel and count the number of times the DAQ reports a hit in that channel. This is repeated for every value of the global threshold DAC. The result of this procedure, known as an s-curve in figure 3.13. The data is fit with an error function. The 50% turn on point of the resulting error function is the threshold and the width is the equivalent noise charge (ENC). After populating all 24 of the VFAT3 slots on a detector, an s-curve is taken for all channels on every chip. The results of such a scan are shown in figure 3.14. This tool is useful for diagnosing the overall health of the detector. Dead channels will appear as

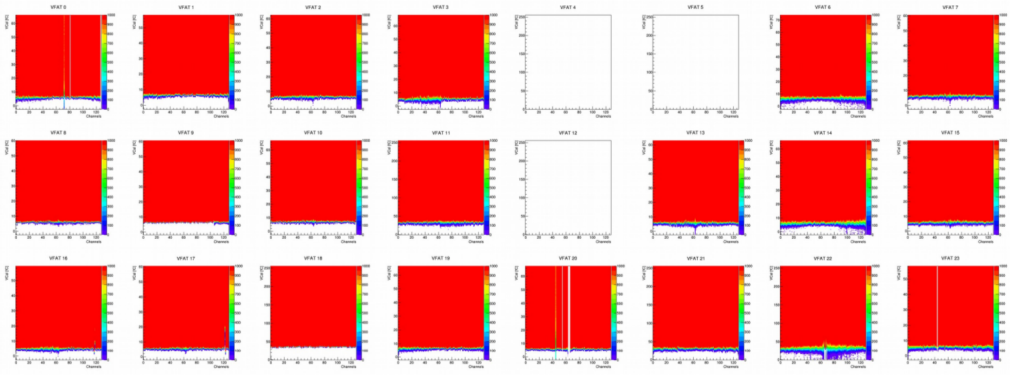

Figure 3.14: A graphic summary of the data from a full detector s-curve scan. The x-axis of each block is the VFAT channel number. The y-axis is the amount of charge injected in fC. The color represents the total number of hits seen out of 1,000 injections. The 8x3 arrangement of the VFATs matches their geometrical arrangement on the detector.

white vertical strips in these plots and noisy chips will have larger green bands.

Initially, two resistors on each of the 24 hybrid boards carrying the VFAT chip shorted the analog ground to the read-out board (ROB) of the detector. This created a ground loop in which some amount of the return current from powering the VFAT could take a path through the ROB, near the strips. The ENC was observed to be as high as 10 fC in this configuration which would make the detector inoperable because a threshold would need to be set to around 40 fC, much higher than the 2 fC needed to be reasonably efficient. The 48 resistors creating this ground loop were removed, and the results of before and after the removal of the ground loop can be seen in figure 3.15. An ENC of about 0.5 fC was observed across the entire detector.

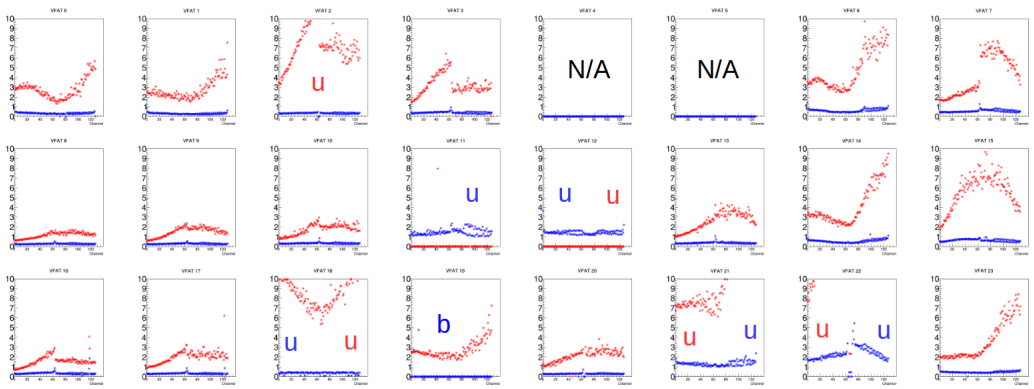

Figure 3.15: A comparison of the ENC across the entire detector before (red) and after (blue) removal of a ground loop. The y-axis is the ENC in fC, extacted from the fits to s-curves. The x-axis is the channel of the vfat. A u indicates that particular VFAT was not calibrated at the time the measurement was taken. The b means that particular VFAT had broken communication during the measurement. VFAT slots 4 and 5 were not populated during the measurement because these GEB slots had design bugs in this particular iteration of the electronics.

3.3 GEB and OH

The main printed circuit board of the detector is the GEM electronics board (GEB) which provides power and the interconnectivity of the other boards in the system. A single DC voltage is provided to the GEB, which uses multiple boards using FEAST DC-DC converters to generate the multiple voltage needed throughout the system. The GEB also provides connection of the digital signals between the Optohybrid (OH) and the 24 VFATs.

The OH is the communication hub of the system. It utilizes a total of ten 3.2 GHz links: seven transfer and three receive links. The Virtex 6 [HK12] field programmable gate array (FPGA) is used to receive and zero-suppress the trigger data from the VFATs by packing them into clusters. The packed trigger clusters are sent independently to the back-end and the OTMB in the CSC system for further processing.

3.4 DAQ

The back-end DAQ board used is called a CTP7 [hr16]. The modern DAQ block of the VFAT3 allows near direct connection of the front-end chips to the back-end. The same path which provides the interface for fast-operation commands (reset, calibration pulse, L1A, etc.) also provides the slow control interface. The CTP7 is therefore the main control unit of the detector. When an L1A arrives at the CTP7 from the main Level One CMS trigger system, the signal is first forwarded to all 24 VFATs on the chamber. The VFATs then build a packet from the buffered data and begin sending it to the CTP7. It is here they are concatenated with a header/trailer and sent onwards to the central CMS DAQ system.

The same link that is used for the full DAQ is also used for slow control, so the CTP7 is also responsible for running routines for calibration and configuration. These routines are implemented in middleware running on the embedded system of the CTP7. The middleware architecture exposes important routines directly to the DAQ software via a remote procedure call. The CTP7 has as deamon which listens, catches, serializes, and executes all remote procedure calls made to the system.

CHAPTER 4

The Phenomenology of Sphaleron-Like Transitions in Proton-Proton Collisions

As discussed in Chapter 1, the electroweak sphaleron of the Standard Model provides violation of baryon and lepton numbers, and could be relevant to the generation of the matter-antimatter asymmetry of the universe. Several recent publications have argued that the pre-exponential factor in sphaleron-induced transitions in high energy proton-proton collisions might not be much less than one, leading to an observable rate of spectacular events at the LHC or future colliders. We created and developed BaryoGEN, an event generator that facilitates the exploration of such transitions with $\Delta N_{CS} = \pm 1$ in proton-proton collisions with minimal assumptions. BaryoGEN outputs standard Les Houches Event [Alw07] files that can be processed by PYTHIA [SAC15], and the code is publicly available. We discuss various approaches to searches for the sphaleron-induced transitions in proton-proton collisions.

4.1 The Sphaleron: A Gauge Field Configuration

The class of solutions of gauge field theories to which the sphaleron belongs were first proposed in 1976 by 't Hooft [Hoo76]. These solutions are nonperturbative,

so the cross-sections for processes mediated by the sphaleron cannot be calculated perturbatively, e.g. by using Feynman diagrams. The solutions are high-energy but are unstable and decay immediately.

The electroweak (EW) sphaleron was first described in 1984 [KM84]. It is also a critical piece of a leading cosmological model for the generation of the matter-antimatter asymmetry in the universe known as Electroweak Baryogenesis [Tro99] as well as thermal Leptogenesis [FY86]. The critical aspect of the sphaleron used in this model is the violation of baryon number (B) and lepton number (L) with conservation of $B - L$. Classically there exist 12 globally conserved U(1) currents, corresponding to the conservation of baryon and lepton numbers. Under reasonable approximations the $(B+L)$-violating phenomenology is summarized as follows [KT90]. An anomaly breaks the conservation of the U(1) currents [Hoo76]:

$$\partial_\mu J^\mu = \frac{g^2}{16\pi^2}\mathrm{Tr}[F_{\mu\nu}\tilde{F}^{\mu\nu}], \tag{4.1}$$

where $F_{\mu\nu}$ is the $\mathrm{SU}(2)_L$ field strength tensor. The integral of this term can be non-zero:

$$N_{CS} = \frac{g^2}{16\pi^2}\int d^4x \mathrm{Tr}[F_{\mu\nu}\tilde{F}^{\mu\nu}]. \tag{4.2}$$

This integral is known as a Chern-Simons [CS74] or winding number. The parameter N_{CS} is a continuous characteristic of the gauge field with a periodic vacuum potential. The local minima of this potential are at integer values of N_{CS} while the sphaleron solutions exist at the local maxima, where N_{CS} is a half-integer. Sphaleron-like transitions are characterized by the ΔN_{CS} of the vacuum transition they mediate. The anomaly exists equally for each fermion doublet. This means that, due to the vacuum transition, L changes by ΔN_{CS} for each of the three lepton doublets, since each lepton has $L = 1$ or $L = -1$, while B changes

by $3\Delta N_{CS}$ because each quark doublet has $B = 1/3$ and there are three colors and three families. This results in two important relations that are essential to the phenomenology of sphalerons [KT90]:

$$\begin{aligned} \Delta(B+L) &= 6\Delta N_{CS} \\ \Delta(B-L) &= 0. \end{aligned} \tag{4.3}$$

The sphaleron cannot exist without enough energy to overcome the periodic potential in N_{CS}, which can be described given all of the electroweak couplings. Since the discovery and mass measurement of the Higgs boson [al12, Col12], all of these couplings have been measured, enabling the calculation of the minimum energy required for sphaleron-like transitions as $E_{Sph} = 9$ TeV [KM84, TW15].

In general, as described below, sphaleron-like transitions in proton-proton scattering will produce spectacular events containing large numbers of energetic particles that are similar to signatures expected from hypothetical microscopic black holes. Tye and Wong [TW15] argue that the "pre-exponential factor" governing the rate of these transitions might not be much less than one. This argument is currently under debate and has been criticized in [FFS16, BT16] and followed up by Tye and Wong [TW16, TW17]. Ellis and Sakurai [ES16] present a study of such transitions in proton-proton collisions by reinterpreting the results of a search for microscopic black holes by ATLAS [Ae16]. Beyond Standard Model (BSM) fermions can also lead to significant rates [CRS18]. In this paper, we describe BaryoGEN, a new and more general event generator for sphaleron-like transitions in proton-proton collisions. BaryoGEN interfaces to general-purpose tools such as PYTHIA [SAC15] via Les Houches Event (LHE) files [Alw07], and the code has been made public [Bra]. BaryoGEN is relevant in general for any model violating baryon and lepton-number via a $\Delta N_{CS} = \pm 1$ vacuum transition due to a chiral

anomaly, even for BSM searches at LHC energies.

4.2 Sphaleron Phenomonology

Sphaleron-like transitions in proton collisions mediate processes such as:

$$u + u \to e^+\mu^+\tau^+\bar{t}\bar{t}\bar{b}\bar{c}\bar{c}\bar{s}\bar{d} + X$$

In general the baryon and lepton number violating part of the interaction can be written as [KT90]:

$$\epsilon_{ab}\epsilon_{cd}\epsilon_{ef}\epsilon_{gh}\epsilon_{ij}\epsilon_{kl}D^a_\alpha D^b_\beta D^c_\gamma D^d_\delta D^e_\zeta D^f_\mu D^g_\theta D^h_\eta D^i_\iota D^j_\nu D^k_\kappa D^l_\lambda$$

Where D represents a fermion doublet which is chosen as one of the 12 doubets of a transition by lower indices. The upper indices are the SU(2) indices. The 6 epsilons guarantee conservation of the electroweak charges. To determine the number of quatum mechanically unique final states the first step is to assign a number (lower indices) to each of the 12 doublets in a sphaleron transition. These doublets consist of a single lepton doublet from each generation and 3 quark doublets from each generation (one for each color). The epsilons then require all of the doublets to be paired. The number of ways to pair 12 unique objects is:

$$\frac{\prod_{n=1}^{6}\binom{2n}{2}}{6!} = 10395$$

There is then a factor of two for each epsilon because the SU(2) indices can be chosen in two ways. There is then one final factor of two for the sign of N_{CS} of the transition. This makes a total number 1,330,560 quantum mechanically unique sphaleron-like transitions. Many of these final states would look identical in CMS

though, so we need to determine the number of phenomenologically different types of transitions. Each of the eight configurations of leptons is phenomenologically distinct. All quarks not in the third generation are effectively identical so only the four unique configurations of the third generation quarks count as distinct configurations. This gives a total of 32 different phenomenologically distinct final states of sphaleron transitions.

The structure of the decay phase space of these multi-particle final states is unknown. After the fermion content of the final state is determined, the fermions are given random momenta according to phase space using ROOT [BR97] class **TGenPhaseSpace**, which in turn invokes the GENBOD function from CERN-LIB. These output of TGenPhaseSpace are deweighted via rejection sampling so all events in the end have a weight of one. The kinematics of top quarks from sphaleron decays can be seen in figures 4.1, 4.2, and 4.3.

The cross section of sphaleron production at the LHC is typically written as the cross section of parton interactions with center of mass energy above 9 TeV times $p \in [0, 1]$, which is known as the pre-exponential factor. This has been written down as [ES16]:

$$\sigma = \frac{1}{m_W^2} \sum_{ab} \int dE \frac{d\mathcal{L}_{ab}}{dE} p \exp\left(c \frac{4\pi}{\alpha_W} S(E)\right)$$

where

$$\frac{d\mathcal{L}_{ab}}{dE} = \frac{2E}{E_{CM}^2} \int_{\ln\sqrt{\tau}}^{-\ln\sqrt{\tau}} dy f_a(\sqrt{\tau}e^y) f_b(\sqrt{\tau}e^{-y})$$

and

$$S(E) = \begin{cases} \frac{(1-a)}{E_{Sph}} E + \frac{a}{E_{Sph}^2} E^2 - 1 & : 0 \leq E \leq E_{Sph} \\ 0 & : E \geq E_{Sph}. \end{cases}$$

This integral is numerically evaluated by the Monte Carlo generator via the

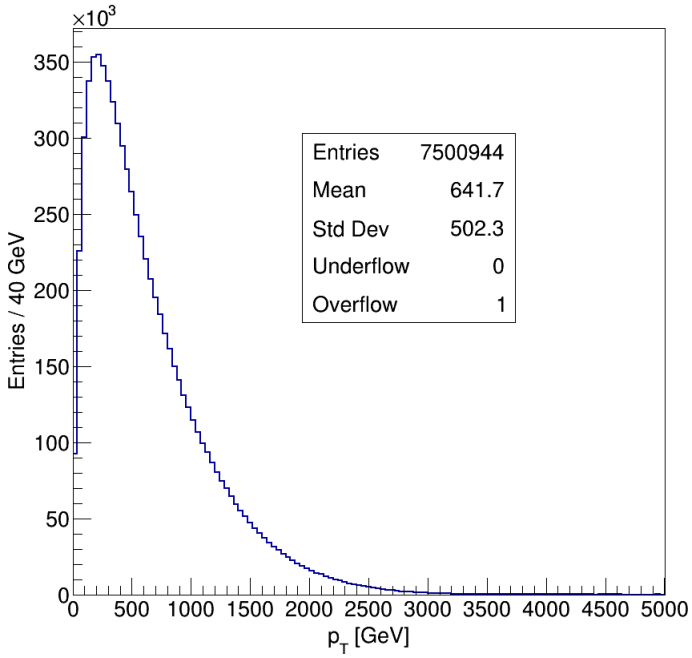

Figure 4.1: The distribution of the p_T of top quarks from sphaleron transitions in the sample of 10,000,000 events.

acceptance-rejection method. The LHAPDF [BFL15] library of parton distribution functions (PDF) is used to access the data of the chosen PDF grid. The grid used is called CT10 [NGG12]. The initial state kinematics of the sphalerons are shown in figures 4.4 and 4.5.

4.3 Properties of the Final State

The states produced through sphaleron transitions have a rich phenomenology which can be interpreted in many ways since nothing can be calculated explicitly regarding the relative rates of the many possible final states. These transistions include population of 12 additional fermion states, and we must first discuss how the initial state of partons effects the final state of fermions. Then we will look at

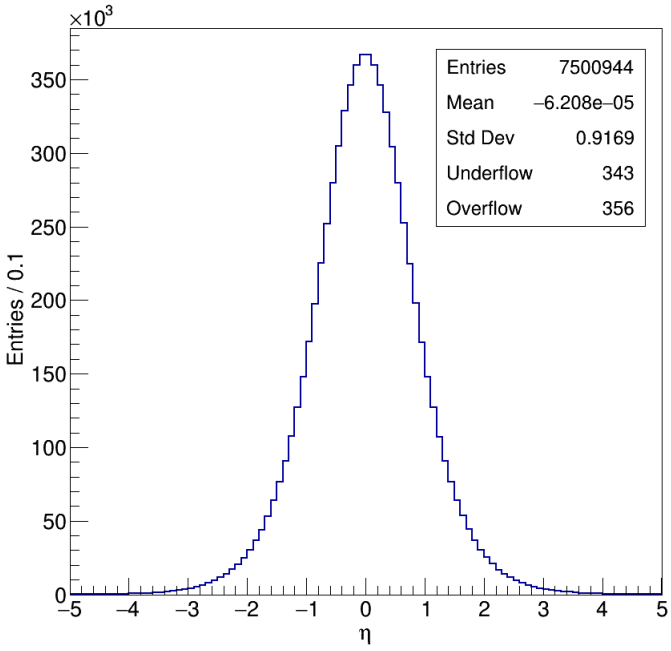

Figure 4.2: The distribution of the η of top quarks from sphaleron transitions in the sample of 10,000,000 events.

the details of color flow from the inital state to the final state, given the possibility of quark cancellations. Finally, we present the distribution of various kinematic quantities and particle multiplicities of the final states. A seperate sample of 100,000 events was produced and decayed and hadronized using PYTHIA [SAC15].

4.3.1 Incoming Partons and Cancellations

When producing the set of outgoing fermions the two incoming partons are added to the generated set of 12 outgoing fermions. If this addition results in any quark-antiquark pairs, they are removed. This generator therefore cancels zero, one, or two quarks, whereas the treatment in Ellis and Sakurai [ES16] only allowed for zero or two cancellations for $\Delta N_{CS} = 1$ or -1, respectively. The prior treatment implicitly assumes negligible sea quark interactions but this less reasonable at

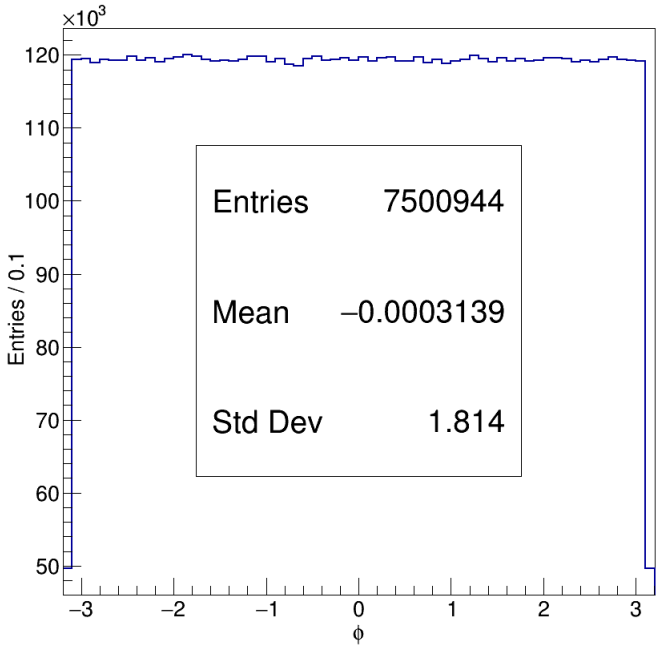

Figure 4.3: The distribution of the ϕ of top quarks from sphaleron transitions in the sample of 10,000,000 events.

higher energies such as 100 TeV.

Zero quark cancellations yield a 14-fermion final state, including three leptons, three third-generation quarks, and eight light quark jets. With one cancellation there are 12 outgoing fermions, of which six are light quark jets, and with two cancellations there are 10 outgoing fermions, of which only four are light quark jets. We make a simple ansatz that $N_{CS} = -1$ and $N_{CS} = +1$ transitions are equally likely. We also use the CT10 [NGG12] Parton Distribution Function (PDF) set throughout this article. Under these assumptions, the relative rates of these three cases are shown in Table 4.1. It is observed that double cancellations are rare when $\Delta N_{CS} = +1$, and zero cancellations are uncommon when $\Delta N_{CS} = -1$, but single cancellations are actually common for $\Delta N_{CS} = -1$ as shown in table 4.1. The distribution of the PDG ID [Oli14] of the fermions directly from the

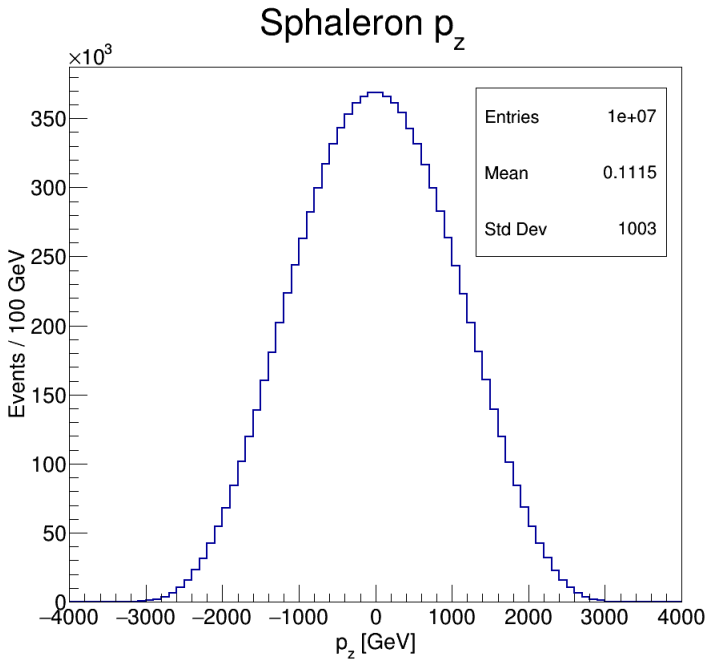

Figure 4.4: The distribution of the z-momentum of the produced sphaleron events.

transition are shown in figure 4.6. The excess for up quarks (PDG ID $+2$) and deficit for up antiquarks (PDG ID -2) is due to the up quarks being the most common incoming partons, which are thus the most probable to be added as an extra outgoing fermion in the $\Delta N_{CS} = +1$ case, or to cancel an outgoing up antiquark in the $\Delta N_{CS} = -1$ case.

It may be noted that the values in Table 4.1 depend on the relative levels of sea and valence quarks at large x and at extremely high values of q^2 not yet directly probed by experiment; in particular a large variation is seen in NNPDF3.0 [Bal15] with respect to other choices of PDF sets. Table 4.2 was prepared using NNPDF3.0 to illustrate the difference the use of different PDF sets can make.

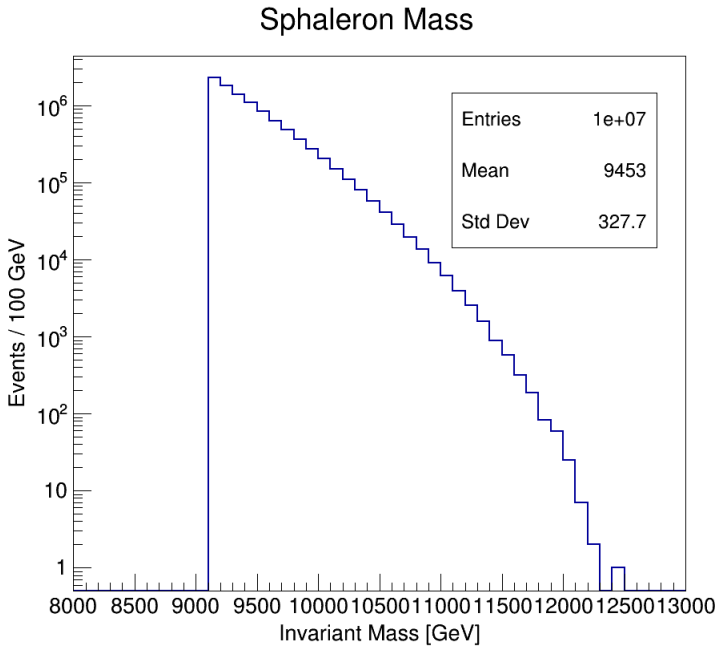

Figure 4.5: The distribution of the invariant mass of the produced sphaleron events.

Multiplicity	$N_{CS} = -1$	$N_{CS} = +1$
10	27.9%	0.0%
12	19.1%	0.5%
14	3.0%	49.5%

Table 4.1: The fractions of each fermion multiplicity of sphaleron-like transitions, with $\sqrt{s} = 13$ TeV, and equal probabilities for $N_{CS} = -1$ and $N_{CS} = +1$, from a generated sample of 10^7 events and CT10 PDFs.

Multiplicity	$N_{CS} = -1$	$N_{CS} = +1$
10	20.6%	0.8%
12	25.4%	14.5%
14	3.9%	34.7%

Table 4.2: The fractions of each fermion multiplicity of sphaleron-like transitions, with $\sqrt{s} = 13$ TeV, and equal probabilities for $N_{CS} = -1$ and $N_{CS} = +1$, from a generated sample of 10^7 events and NNPDF3.0 PDFs.

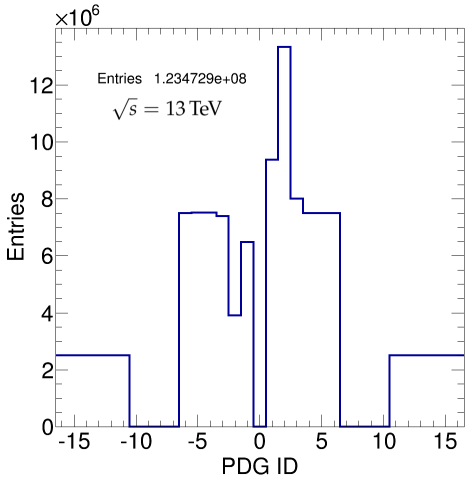

Figure 4.6: The number of each of the fermion types (PDG IDs [Oli14]) coming directly from 10^7 sphaleron-like transitions and CT10 PDFs.

4.3.2 Color Flow

In order to use PYTHIA to decay and hadronize the hard processes made by this generator, color flow lines must be drawn and enumerated. PYTHIA can only handle baryon number violation of one at a given vertex, and since sphaleron-like transitions violate baryon number by 3, the transition is factorized by introducing fake mediators in the LHE file. These mediators are defined for convenience with PDG IDs of 1000022 ($\tilde{\chi}_1^0$), 1006213 (R^+), and 1006223 (R^{++}). Note that HERWIG [BGG16] must be configured to recognize these particle IDs in order to use it with the output LHE files created by this generator.

The outgoing quarks, after cancellations, are first grouped together by generation. An example of the graph represented by an event in the outgoing LHE file is shown in figure 4.7. This particular transition vertex has two incoming quarks, the outgoing leptons, the outgoing uncanceled incoming quarks, and a fake mediator for each generation of outgoing quarks.

Each generation of outgoing quarks must be a color singlet, otherwise it is possible to get two identical fermion states, which is not allowed by the Pauli exclusion principle. Each generation of quarks is given a fake mediator to ensure that PYTHIA builds this set of quarks into a color singlet. In the case of quark cancellation, the mediator will share a color line with the canceled quark. If it is ever the case that two quarks are canceled from the same generation, the mediator is retracted and the only remaining quark from that particular generation comes directly off the central vertex.

After the fermion content of the final state is determined, the fermions are given random momenta according to phase space using ROOT [BR97] class **TGen-**

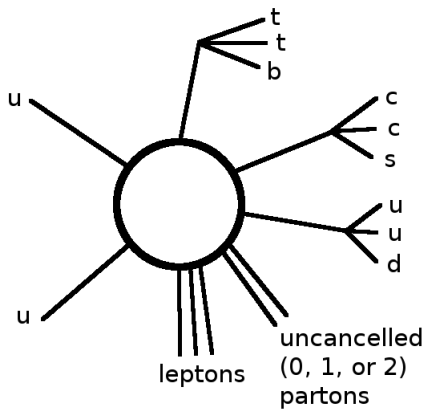

Figure 4.7: An example of a sphaleron-induced transition event starting from an initial state of two up quarks. The diagram is a representation of the event as written to the LHE file in order to allow correct determination of color flow by PYTHIA for decay and hadronization.

PhaseSpace, which in turn invokes the GENBOD function from CERNLIB.

4.3.3 Simulation Results

Simulations were conducted at the center of mass energies of 13, 14, 28, and 33 TeV. The corresponding cross sections for a pre-exponential factor of one for these beam energies are, respectively, 10.05, 50.72, 111,958, and 285,053 fb. A total of 10^5 events were produced and then processed with PYTHIA at each energy. These simulations were made with the nominal $E_{Sph} = 9$ TeV and, as previously mentioned, CT10 PDFs and the ansatz that $N_{CS} = -1$ and $N_{CS} = +1$ transitions are equally likely. We count the number N of quarks and charged leptons with transverse momentum $p_T > 20$ GeV and pseudo-rapidity $|\eta| < 2.5$, so that the results can be easily compared with those in [ES16].

Distributions of observable invariant mass are presented in figure 4.11, com-

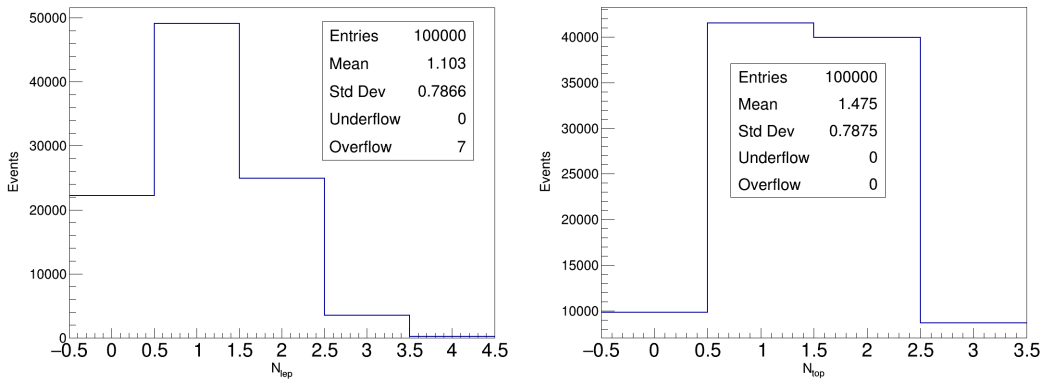

Figure 4.8: Distributions of lepton (left) and top quark (right) multiplicities, requiring them to have $p_T > 20$ GeV and $|\eta| < 2.5$. We observe sphaleron-like transitions commonly have a few of these physics objects, which are uncommon in proton-proton collisions.

parable to figure 3 (left) of [ES16]. The total transverse energy of all jets H_T, and the missing transverse energy $\not{E}_T$, are shown in figure 4.12, comparable to figure 5 (top left and bottom left, respectively) of [ES16].

Various characteristic quantities were also studied in the simulations. The distributions of the lepton and top multiplicities are shown in figure 4.8. Distributions of jet multiplicity and charged particle multiplicity are shown in figure 4.9. Distributions of the p_T of the two leading jets in each event are shown in figure 4.10.

The distributions of $N \cdot N_{CS}$ at center-of-mass energy 13 TeV, after top quark and W decays, for each of the number of potential incoming parton cancellations, is shown in figure 4.13. This plot separates the distributions of N by the sign of N_{CS} (separating fermion from antifermion-type transitions), and corresponds to figure 4 (top left) of [ES16].

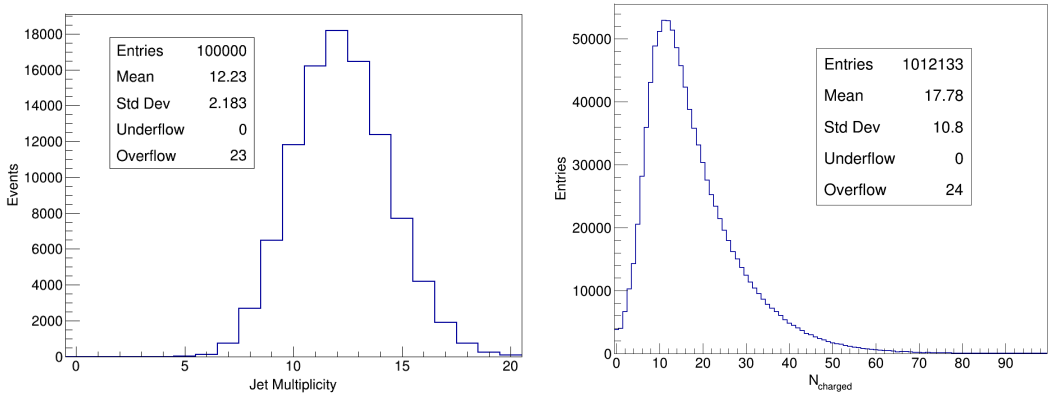

Figure 4.9: Distributions of jet multiplicity (left) and the multiplicity of charged particles in each jet (right), requiring jets to have $p_T > 50$ GeV and $|\eta| < 5.2$.

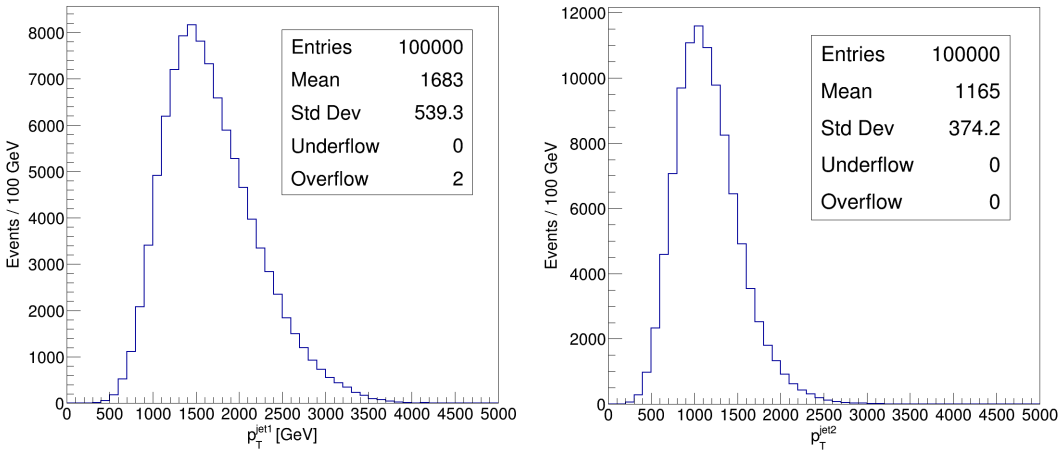

Figure 4.10: Distributions of leading (left) and the second leading (right) jet p_T, requiring jets to have $p_T > 50$ GeV and $|\eta| < 5.2$.

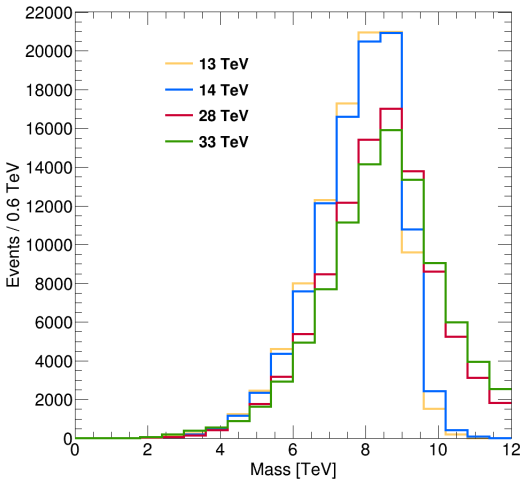

Figure 4.11: Invariant-mass distributions of observable final-state particles in simulated LHC collisions at 13, 14, 28, and 33 TeV. These simulations are made with the nominal $E_{Sph} = 9$ TeV.

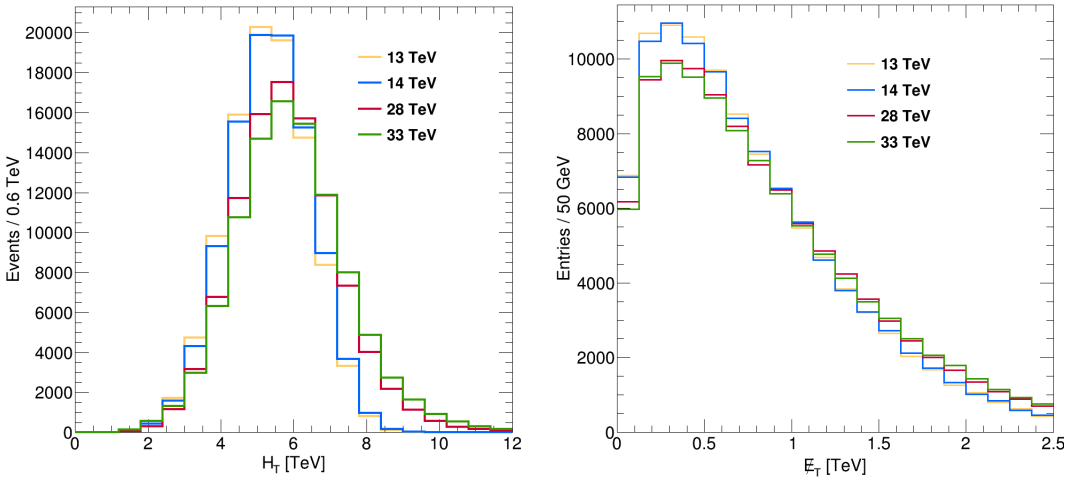

Figure 4.12: Left Panel: Distributions of the scalar sum of p_T^{jet} of events in simulated LHC collisions at 13, 14, 28, and 33 TeV. Right Panel: Distributions of $\not{E}_T$ in simulated LHC collisions at all four energies. These simulations are made with the nominal $E_{Sph} = 9$ TeV.

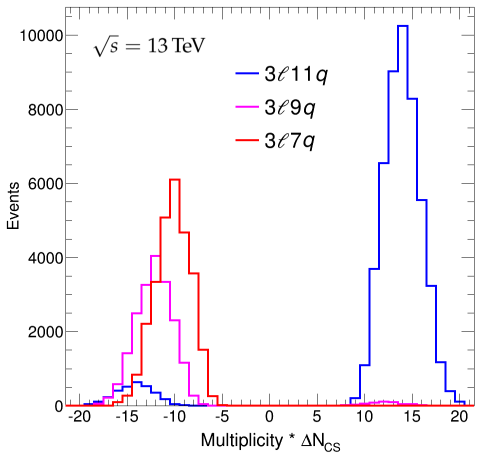

Figure 4.13: The distributions of multiplicities of energetic ($p_T > 20$ GeV) jets and leptons within a nominal fiducial detector acceptance of $|\eta| < 2.5$, after processing with PYTHIA, including top quark and W decays, as well as jet fragmentation and hadronization. The multiplicities are plotted separately for $N_{CS} = \pm 1$ by multiplying them by N_{CS}, and are shown separately for each of the possibilities for number of quark-antiquark cancellations (0, 1, and 2).

4.4 Search Strategies

It may be sufficient at current LHC energies and integrated luminosities to search for sphaleron-like transitions by simply looking for excesses of high multiplicity, high transverse-energy events, since the very high multiplicity and very high energy predicted for these transitions leads to a very low level of QCD, top quark, and electroweak backgrounds. Such searches are similar to those done for microscopic black holes [CMS17a, Ae16]. This is the strategy used later this dissertation. An ansatz that $N_{CS} = -1$ and $N_{CS} = +1$ transitions are equally likely is simple and seems sufficient in the absence of positive signals.

It may become necessary to reduce the background levels further in future searches for sphaleron-like transitions, in order to achieve maximum sensitivity. Such circumstances include a) greatly increased amounts of LHC data, b) higher proton-proton energies, c) searches for Beyond Standard Model sphaleron-like transitions with lower transition energies [ST17, KN17], or d) if positive signals are found. In such cases, an attractive, but more complex method would be to classify the events based on lepton content and missing transverse energy. Those channels containing charged leptons and/or large missing transverse energy would have greatly reduced QCD background; and many of the channels would contain same-sign charged leptons that are rare in the Standard Model.

4.5 Using the Generator

The BaryoGEN code is available as a public github repository [Bra]. It is recommended that the general user download the most recent tag. The code has dependencies that are noted in the README included in the source code. It is

assumed that ROOT [BR97], LHAPDF [BFL15], and CT10 [NGG12] have been correctly installed. In general, the program is run with the syntax:

```
BaryoGEN sqrtS threshold maxweight Nevents pNCS bCancel Filename
```

If the number of arguments is not correct, the program will output this syntax reminder and terminate. The **sqrtS** is the proton-proton center-of-mass energy in GeV. The **threshold** is the minimum energy in GeV required for a transition. The parameter **maxweight** must be set to a value that is greater than any possible probability given by the PDF set used, and also should be changed any time the threshold energy is varied. The **pNCS** and **bCancel** parameters will be described in the following paragraph, and the **Filename** is the name of the output files.

There are a few options for configuring the types of output events. The **pNCS** parameter sets the probability of events to have $\Delta N_{CS} = +1$. The probability of getting $\Delta N_{CS} = +1$ is one minus this probability. In the results presented in this paper, we have set **pNCS** to 0.5 as mentioned at the end of section 4.3.3. The generator also gives the ability to turn off the parton cancellations by setting **bCancel** to 0 or to turn them on by setting it to 1. The configurations that correspond most closely to Ellis and Sakurai [ES16] are with **pNCS** set to 0.0 or 1.0 with parton cancellations on.

The following commands will build the executable and generate $10,000$ events. The generated output files will be called testRun.root and testRun.lhe.

```
git clone https://github.com/cbravo135/BaryoGEN.git
cd BaryoGEN
make
./BaryoGEN 13000 9000 5e-4 10000 0.5 1 testRun
```

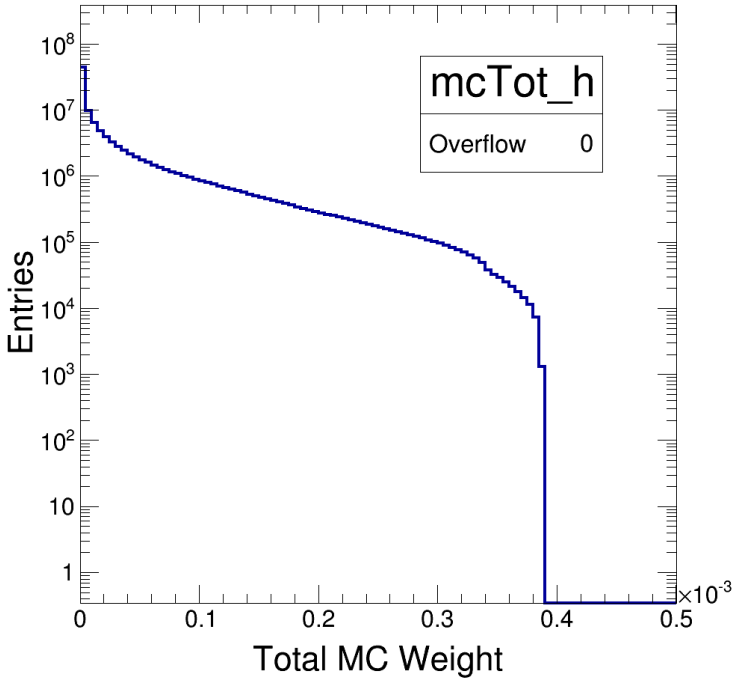

Figure 4.14: An example of mcTot_h (pp energy 13 TeV, $E_{Sph} = 9$ TeV) that has a high enough maximum weight for the sample produced.

These commands will output two files: testRun.root and testRun.lhe. The LHE file can be used to further process the events with the user's software of choice for hadronization and fragmentation. The file testRun.root contains histograms and an Ntuple of the generated events, to make it easy for the end user to analyze the generated events. The histogram mcTot_h, as seen in figure 4.14, is the most important histogram to check of these. If this histogram has any events in the overflow bin, the second argument given to the program must be increased so that the output is correctly weighted. This maximum weight parameter only needs to change if the proton-proton energy or the transition energy threshold is changed.

4.6 Conclusions

We have a presented a new Monte Carlo generator, BaryoGEN, for simulating baryon and lepton-number violating sphaleron-like transitions in proton-proton collisions. The BaryoGEN output uses LHE files to interface smoothly with PYTHIA for hadronization and fragmentation. Kinematic distributions have been produced for comparison with the performance of other generators.

It may be sufficient at current LHC energies and integrated luminosities to search for sphaleron-like transitions by simply looking for excesses of high multiplicity, high transverse-energy events, since the very high multiplicity and very high energy predicted for these transitions leads to a very low level of QCD, top quark, and electroweak backgrounds. Such searches are similar to those done for microscopic black holes [CMS17a, Ae16]. An ansatz that $\Delta N_{CS} = -1$ and $\Delta N_{CS} = +1$ transitions are equally likely is simple and seems sufficient in the absence of positive signals.

It may become necessary to reduce the background levels further in future searches for sphaleron-like transitions, in order to achieve maximum sensitivity. Such circumstances include a) greatly increased amounts of LHC data, b) higher proton-proton energies, c) searches for BSM sphaleron-like transitions with lower transition energies [ST17, KN17], or d) if positive signals are found. In such cases, an attractive, but more complex method would be to classify the events based on lepton content and missing transverse energy. Those channels containing charged leptons and/or large missing transverse energy would have greatly reduced QCD background; and many of the channels would contain same-sign charged leptons that are rare in the Standard Model.

BaryoGEN is an event generator that faciliates the study of a sphaleron-like class of $(B + L)$-violating transitions at LHC or future proton colliders. It is uncertain whether rates due to Standard Model sphaleron transitions are large enough to be observable, but such transitions could also arise from beyond Standard Model physics. In either case, these transitions would give rise to spectacular signatures, and could be the first direct evidence of baryon-number and lepton-number violation.

CHAPTER 5

A Search for Sphalerons in CMS 2016 Data

5.1 The Data Sample

This analysis is based on a data sample recorded with the CMS detector in proton-proton collisions at a center-of-mass energy of 13 TeV in 2016, corresponding to an integrated luminosity of 35.9 fb^{-1}. Since typical signal events may contain jets, leptons, photons, and missing transverse momentum from undetected particles, to be as inclusive as possible we employ a trigger based on the H_T variable, defined as the scalar sum of the p_T of all jets in an event reconstructed at the HLT. We require $H_T > 800$ GeV for the earlier portion of the data taking and $H_T > 900$ GeV for the later one; the increase in the threshold value being driven by the need to keep the trigger rate within an acceptable limit as the LHC luminosity steadily increased. For the later portion of the run, to recover some of the residual inefficiency of the H_T trigger, we also used a logical OR with several single-jet triggers with the E_T thresholds of 450–500 GeV. The resulting trigger selection is fully efficient for events that subsequently satisfy the offline requirements used in the analysis.

5.2 Event reconstruction

The particle-flow (PF) algorithm [Sir17] aims to reconstruct and identify each individual particle in an event with an optimized combination of information from the various elements of the CMS detector. The energy of photons is directly obtained from the ECAL measurement, corrected for zero-suppression effects. The energy of electrons is determined from a combination of the electron momentum at the primary interaction vertex as determined by the tracker, the energy of the corresponding ECAL cluster, and the energy sum of all bremsstrahlung photons spatially compatible with originating from the electron track. The energy of muons is obtained from the curvature of the corresponding track. The energy of charged hadrons is determined from a combination of their momentum measured in the tracker and the matching ECAL and HCAL energy deposits, corrected for zero-suppression effects and for the response function of the calorimeters to hadronic showers. Finally, the energy of neutral hadrons is obtained from the corresponding corrected ECAL and HCAL energies.

The reconstructed vertex with the largest value of summed physics-object p_T^2 is taken to be the primary proton-proton interaction vertex. The physics objects are the jets, clustered using the anti-k_T jet finding algorithm [CSS08, CSS12] with the tracks assigned to the vertex as inputs, and the associated missing transverse momentum, taken as the negative vector sum of the p_T of those jets. Events are required to have at least one reconstructed vertex within 24 (2) cm of the nominal collision point in the direction parallel (perpendicular) to the beams.

For each event, hadronic jets are clustered from the PF candidates using the anti-k_T algorithm with a distance parameter of 0.4. The jet momentum is de-

termined as the vectorial sum of all particle momenta in the jet, and is found from simulation to be within 5 to 10% of the true momentum over the whole p_T spectrum and detector acceptance. Additional proton-proton interactions within the same or neighboring bunch crossings (pileup) can contribute additional tracks and calorimetric energy depositions to the jet momentum. To mitigate this effect, tracks originating from pileup vertices are discarded and an offset correction is applied to correct for the remaining contributions. Jet energy corrections are derived from simulation, to bring the measured response of jets to that of particle-level jets on average. In situ measurements of the momentum balance in dijet, multijet, γ+jet, and leptonically decaying Z+jet events are used to account for any residual differences in the jet energy scales in data and simulation [Kha17b]. The jet energy resolution amounts typically to 15% at a jet p_T of 10 GeV, 8% at 100 GeV, and 4% at 1 TeV. Additional selection criteria are applied to each jet to remove those potentially dominated by anomalous contributions from various subdetector components or reconstruction failures. All jets are required to have $p_T > 70$ GeV and be within $|\eta| < 5$. For the leading p_T jet in each event, the energy fraction carried by muon candidates failing the standard identification [Sir18] is required to be less than 80%. This requirement removes events where a low-momentum muon is misreconstructed with very high momentum and misidentified as a high-energy jet. We further require the leading jet in an event to have a charged-hadron fraction of less than 0.99 if this jet is found within $|\eta| < 2.4$ [CMS17b].

The missing transverse momentum, p_T^{miss}, is defined as the magnitude of the vectorial sum of transverse momenta of all PF candidates in an event. The jet energy corrections are further propagated to the p_T^{miss} calculation.

Details of muon reconstruction can be found in Ref. [Sir18]. The muon candi-

date is required to have at least one matching energy deposit in the pixel tracker and at least six deposits in the silicon strip tracker, as well as at least two track segments in the muon detector. The transverse impact parameter and the longitudinal distance of the track associated with the muon with respect to the primary vertex are required to be less than 2 and 5 mm, respectively, to reduce contamination from cosmic ray muons. The global track fit to the tracker trajectory and to the muon detector segments must have a χ^2 per degree of freedom of less than 10. Muon candidates are required to have $p_T > 70$ GeV and to be within $|\eta| < 2.4$.

Details of electron and photon reconstruction can be found in Refs. [Kha15b] and [Kha15c], respectively. Electron and photon candidates are required to have $p_T > 70$ GeV and $|\eta| < 2.5$, excluding the $1.44 < |\eta| < 1.57$ transition region between the ECAL barrel and endcap detectors where the reconstruction is suboptimal. We use standard identification criteria, corresponding to an average efficiency of 80% per electron or photon. The identification criteria include a requirement that the transverse size of the electromagnetic cluster be compatible with the one expected from a genuine electron or photon, and that the ratio of the HCAL to ECAL energies be less then 0.25 (0.09) for electrons and less than 0.0396 (0.0219) for photons in the barrel (endcap). In addition, photon candidates are required to pass the conversion-safe electron veto requirements [Kha15c], which disambiguates them from electron candidates.

Muons, electrons, and photons are required to be isolated from other energy deposits in the tracker and the calorimeters. The isolation $\mathcal{I}$ is defined as the ratio of the p_T sum of various types of additional PF candidates in a cone of radius $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ of 0.4 (muons) or 0.3 (electrons and photons), centered on the lepton or photon candidate, to the candidate's p_T. For muons, the numerator

of the ratio is corrected for the contribution of neutral particles due to pileup, using one half of the p_T carried by the charged hadrons originating from pileup vertices. For electrons and photons, an average area method [CS08]], as estimated with FastJet [CSS12], is used. The details of the isolation requirements are the same as used in an earlier 13 TeV analysis [CMS17d], except that for electrons we use a slightly tighter isolation requirement of $\mathcal{I} < 0.07$.

To avoid double counting, we remove jets that are found within a radius of $\Delta R = 0.3$ from a muon, electron, or photon, if the latter object contributes more than 80, 70, or 50% of the jet p_T, respectively.

5.3 Analysis strategy

We closely follow the approach for semiclassical BH searches originally developed by CMS for Run 1 analyses [CMS11,CMS12,CMS13] and subsequently used in the studies of early Run 2 [CMS17d] data. This type of analysis is less sensitive to the details of signal process and the relative abundance of various particles produced, as it considers all types of particles in the final state. We use a single discriminating variable S_T, defined as the scalar sum of p_T of all energetic objects in an event (which we define as jets, electrons, muons, and photons with p_T above a certain threshold, of which there are N), plus p_T^{miss} in the event: $S_T = p_T^{miss} + \sum_{i=1}^{N} p_T^i$.

This definition of S_T is robust against variations in the signal model, and is also sensitive to the cases when there is large p_T^{miss} due to multiple high energy neutrino production in the transition. This makes the analysis sensitive to all possible final states of sphaleron-induced vacuum transitions.

The S_T distributions are then considered separately for various inclusive ob-

ject multiplicities (*i.e.*, $N \geq N_{min} = 3, \ldots, 11$). The background is dominated by SM QCD multijet production and is estimated exclusively from control samples in data. The observed number of events with S_T values above a certain threshold is compared with the background and signal+background predictions to either establish a signal or to set limits on the signal production. This approach does not rely on the Monte Carlo (MC) simulation of the backgrounds, and it also has higher sensitivity than exclusive searches in specific final states, *e.g.*, lepton+jets [Aad13, Aad14].

The main challenge of the search is to describe the inclusive multijet background in a robust way, as sphaleron signals correspond to a broad enhancement in the high tail of the S_T distribution, rather than to a narrow peak. Since these signals are expected to involve a high multiplicity of final-state particles, one has to reliably describe the background for large jet multiplicities, which is quite challenging theoretically as higher-order calculations that fully describe multijet production do not exist. Thus, one cannot rely on simulation to reproduce the S_T spectrum for large N correctly.

To overcome this problem, a dedicated method of predicting the QCD multijet background directly from collision data has been developed for the original Run 1 analysis [CMS11] and used in the subsequent Run 1 [CMS12, CMS13] and Run 2 [CMS17d] searches. It has been found empirically, first via simulation-based studies, and then from the analysis of data at low jet multiplicities, that the shape of the S_T distribution for the dominant QCD multijet background does not depend on the multiplicity of the final state, above a certain turn-on threshold. This observation reflects the way a parton shower develops via nearly collinear emission, which conserves S_T. It allows one to predict the S_T spectrum of a multijet

final state using low-multiplicity QCD events, *e.g.*, dijet or trijet events. This "S_T invariance" provides a powerful method of predicting the dominant background for BH production by taking the S_T shape from low-multiplicity events, for which the signal contamination is expected to be negligible, and normalizing it to the observed spectrum at high multiplicities at the low end of the S_T distribution, where signal contamination is negligible even for large multiplicities of the final-state objects. The method has been also used for other CMS searches, *e.g.*, a search for stealth supersymmetry [Kha15a] and a search for multijet resonances [Kha17a].

5.4 Sphaleron Signal Samples

We simulate the sphaleron signal for three values of the transition energy $E_{sph} = 8$, 9, and 10 TeV. The parton-level simulation is done with the CT10 LO PDF set [NGG12]. In the process of studying various PDF sets, we found that the NNPDF3.0 yields a significantly larger fraction of sea quarks in the kinematic region of interest than all other modern PDFs. While the uncertainty in this fraction is close to 100%, we chose the CT10 set, for which this fraction is close to the median of the various PDF sets we studied. The PDF uncertainties discussed in Section 5.6 cover the variation in the signal acceptance between various PDFs due to this effect.

The typical final-state multiplicities for the $N_{CS} = \pm 1$ sphaleron transitions resulting in 10, 12, or 14 parton-level final states are shown in Fig. 5.1. The $N_{CS} = 1$ transitions are dominated by 14 final-state partons, as the proton mainly consists of valence quarks, thus making the probability of cancellations small.

The cross section for sphaleron production is given by [ES16]: $\sigma = \mathrm{PEF}\,\sigma_0$,

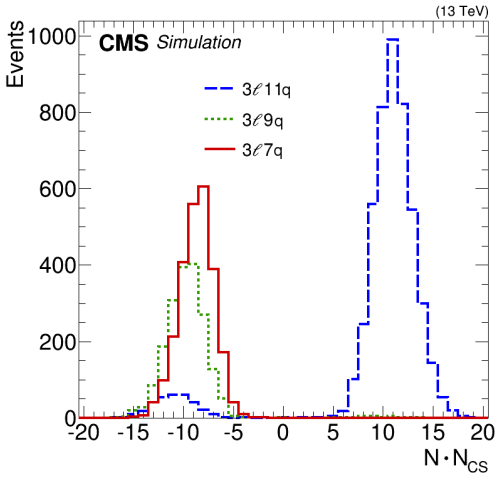

Figure 5.1: Observed final-state particle multiplicity N distributions for $N_{CS} = \pm 1$ sphaleron transitions resulting in 10, 12, and 14 parton-level final-state multiplicities. The relative numbers of events in the histograms are proportional to the relative probabilities of these three parton-level configurations.

where σ_0 = 121, 10.1, and 0.51 fb for E_{sph} = 8, 9, and 10 TeV, respectively, and PEF is the pre-exponential factor, defined as the fraction of all quark-quark interactions above the sphaleron energy threshold E_{sph} that undergo the sphaleron transition. The distributions of the 2 analysis variables of the 9 TeV signal sample are shown in figure 5.2.

5.4.1 Background samples

In addition, we use simulated samples of W+jets, Z+jets, γ+jets, $t\bar{t}$, and QCD multijet events for auxiliary studies. These events are generated with the MadGraph5 v2.2.2 [AFF14] event generator at LO or next-to-LO, with the NNPDF3.0 PDF set of a matching order.

The fragmentation and hadronization of parton-level signal and background

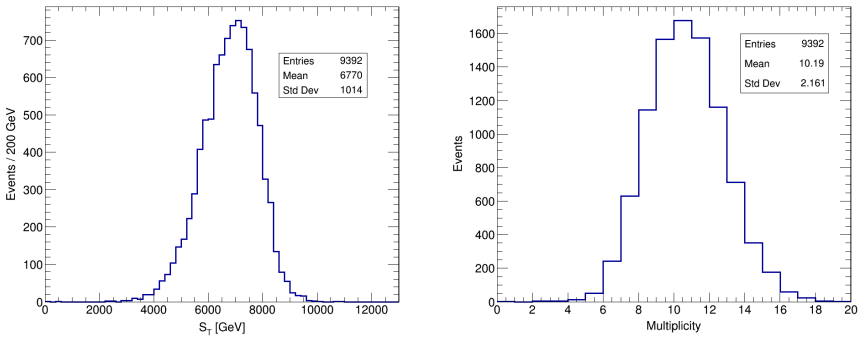

Figure 5.2: Distributions of the two analysis variables after full simulation of the CMS detector and reconstruction of the physics objects.

samples is done with PYTHIA v8.205 [SAC15], using the underlying event tune CUETP8M1 [Kha16]. All signal and background samples are reconstructed with the detailed simulation of the CMS detector via GEANT4 [Iva03]. The effect of pileup interactions is simulated by superimposing simulated minimum bias events on the hard-scattering interaction, with the multiplicity distribution chosen to match the one observed in data.

5.5 Background estimate

5.5.1 Background composition

The main backgrounds in the analyzed multi-object final states are: QCD multijet, V+jets (where V = W, Z), γ+jets, and $t\bar{t}$ production, with the QCD multijet background being by far the most dominant. Figure 5.3 illustrates the relative importance of these backgrounds for the inclusive multiplicity $N \geq 3$ and 6 cases, based on simulated background samples. To reach the overall agreement with the data, all simulated backgrounds except for the QCD multijets are normalized

to the most accurate theoretical predictions available, while the QCD multijet background is normalized so that the total number of background events matches that in data. While we do not use simulated backgrounds to obtain the main results in this analysis, Fig. 5.3 illustrates an important point: not only is the QCD multijet background at least an order of magnitude more important than other backgrounds, for both low- and high-multiplicity cases, but also the shape of the S_T distributions for all major backgrounds is very similar, so the method we use to estimate the multijet background, discussed below, provides an acceptable means of predicting the overall background as well.

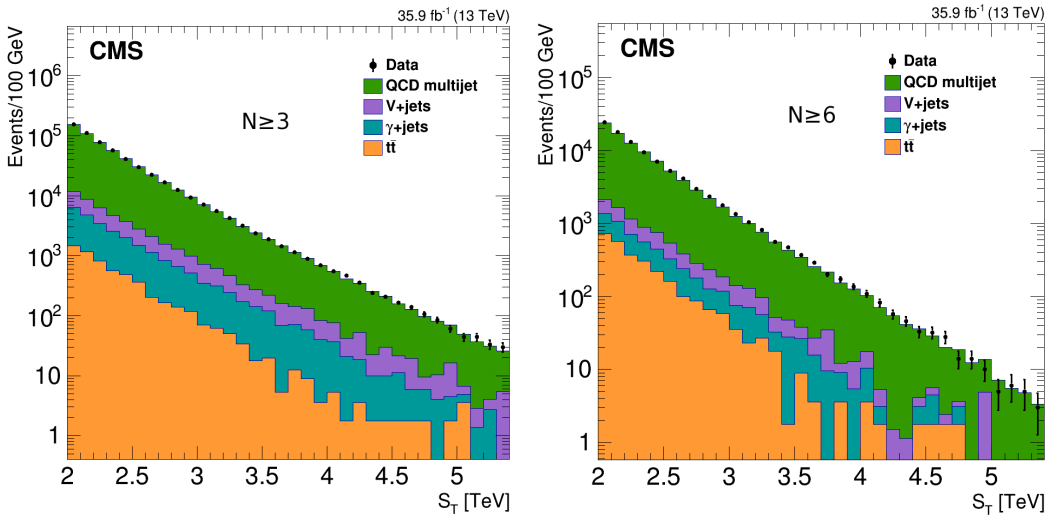

Figure 5.3: The S_T distribution in data for inclusive multiplicities of (left) $N \geq 3$ and (right) $N \geq 6$, compared with the normalized background prediction from simulation, illustrating the relative contributions of major backgrounds.

5.5.2 Background shape determination

The background prediction method used in the analysis follows closely that in previous similar CMS searches [CMS11, CMS12, CMS13, CMS17d]. As discussed

in Section 5.3, the central idea of this method is that the shape of the S_T distribution for the dominant multijet backgrounds is invariant with respect to the final-state object multiplicity N. Consequently, the background shape can be extracted from low-multiplicity spectra and used to describe the background at high multiplicities. The S_T value is preserved by the final-state radiation, which is the dominant source of extra jets beyond LO $2 \to 2$ QCD processes, as long as the additional jets are above the p_T threshold used in the definition of S_T. At the same time, jets from initial-state radiation (ISR) change the S_T value, but they contribute a relatively small fraction of S_T and also typically change the multiplicity N by just one unit. Consequently, we extract the background shape from the $N = 3$ S_T spectrum, which already has a contribution from ISR jets, and therefore reproduces the S_T shape at higher multiplicities better than the $N = 2$ spectrum used in earlier analyses. To estimate any residual noninvariance in the S_T distribution, the $N = 4$ S_T spectrum, normalized to the $N = 3$ spectrum in terms of the total number of events, is also used as an additional component of the background shape uncertainty. Furthermore, to be less sensitive to the higher instantaneous luminosity delivered by the LHC in 2016, which resulted in a higher pileup, and to further reduce the effect of ISR, the p_T threshold for all objects was raised to 70 GeV, compared to 50 GeV used in earlier analyses. The reoptimization that has resulted in the choice of a new exclusive multiplicity to be used for the baseline QCD multijet background prediction and a higher minimum p_T threshold for the objects counted toward S_T was based on extensive studies of MC samples and low-S_T events in data.

In order to obtain the background template, we use a set of 16 functions employed in earlier searches for BSM physics in dijets, VV events, and multijet

events at various colliders. These functions typically have an exponential or power-law behavior with S_T, and are described by 3–5 free parameters. Some of the functions are monotonously falling with S_T by construction; however, some of them contain polynomial terms, such that they are not constrained to have a monotonic behavior. In order to determine the background shape, we fit the $N = 3$ S_T distribution and the $N = 4$ S_T distribution, normalized to the same total event count as the $N = 3$ distribution, in the range of 2.5–4.3 TeV, where any sizable contributions from BSM physics have been ruled out by earlier versions of this analysis, with all 16 functional forms:

$$f_{dijet1}(x) = \frac{p_0(1-x)^{p_1}}{x^{p_2}} \qquad f_{dijet2}(x) = \frac{p_0(1-x)^{p_1}}{x^{p_2+p_3\log(x)}}$$

$$f_{dijet3}(x) = \frac{p_0(1-x)^{p_1}}{x^{p_2+p_3\log(x)+p_4\log^2(x)}} \qquad f_{ATLAS1}(x) = \frac{p_0(1-x^{1/3})^{p_1}}{x^{p_2}}$$

$$f_{ATLAS2}(x) = \frac{p_0(1-x^{1/3})^{p_1}}{x^{p_2+p_3\log^2(x)}} \qquad f_{UA2_1}(x) = p_0 x^{p_1} e^{p_2 x}$$

$$f_{UA2_2}(x) = p_0 x^{p_1} e^{p_2 x+p_3 x^2} \qquad f_{UA2_3}(x) = p_0 x^{p_1} e^{p_2 x+p_3 x^2+p_4 x^3}$$

$$f_{cmsBH1}(x) = \frac{p_0(1+x)^{p_1}}{x^{p_2\log x}} \qquad f_{cmsBH2}(x) = \frac{p_0(1+x)^{p_1}}{x^{p_3+p_2\log x}}$$

$$f_{ATLASBH1}(x) = p_0(1-x)^{p_1} x^{p_2\log(x)} \qquad f_{ATLASBH2}(x) = p_0(1-x)^{p_1}(1+x)^{p_2\log(x)}$$

$$f_{ATLASBH3}(x) = p_0(1-x)^{p_1} e^{p_2\log(x)} \qquad f_{ATLASBH4}(x) = p_0(1-x^{1/3})^{p_1} x^{p_2\log(x)}$$

$$f_{ATLASBH5}(x) = p_0(1-x)^{p_1} x^{p_2 x)} \qquad f_{ATLASBH6}(x) = p_0(1-x)^{p_1}(1+x)^{p_2 x}$$

The fits performed on the full dataset are shown in figure 5.5. The lowest masses of the signal models considered, which have not been excluded by the previous analysis [CMS17d], contribute less than 2% to the total number of events within the fit range, as shown in figure 5.4. Any functional form observed not to be monotonically decreasing up to $S_T = 13$ TeV after the fit to both multiplicities is discarded. The largest spread among all the accepted functions in the $N = 3$ and

$N = 4$ fits is used as an envelope of the systematic uncertainty in the background template. The use of both $N = 3$ and $N = 4$ distributions to construct the envelope allows one to take into account any residual S_T noninvariance in the systematic uncertainty in the background prediction. We observe a good closure of the method to predict the background distributions in simulated QCD multijet events.

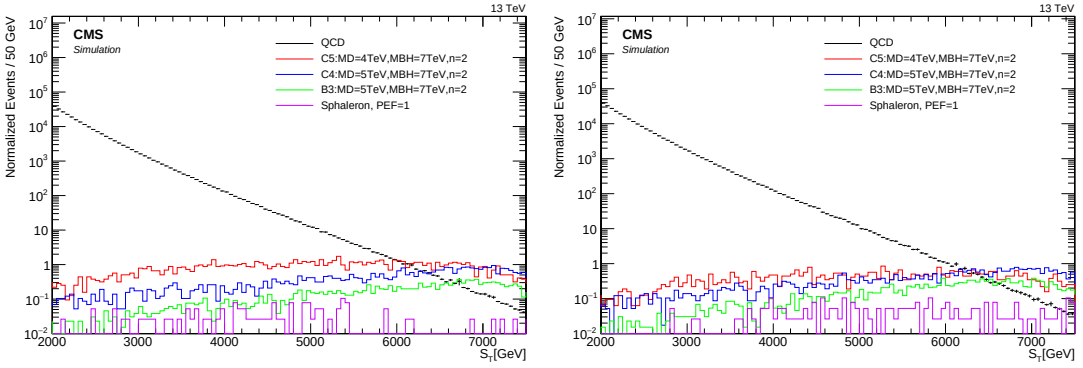

Figure 5.4: Signal contamination studies showing the QCD MC. Several different signal models that the analysis is sensitive to are shown. Sphaleron signal with PEF=1 is the most relavant for this thesis. We chose the upper bound of the fit region to be at 4.3 TeV which has less than 2% signal contamination. The $N = 3$ spectra are on the left and $N = 4$ is on the right.

The best fits (taking into account the F-test criterion [Fis22] within each set of nested functions) to the $N = 3$ and $N = 4$ distributions in data, along with the corresponding uncertainty envelopes, are shown in the two panels of figure 5.5. In both cases, the best fit function is $f(x) = p_0(1 - x^{1/3})^{p_1}/(x^{p_2+p_3 \log^2(x)})$, where $x = S_T/\sqrt{s} = S_T/(13\text{ TeV})$ and p_i are the four free parameters of the fit. The envelope of the predictions at large S_T ($S_T > 5.5$ TeV, most relevant for the present search) is given by the fit with the following 5-parameter function: $\phi(x) =$

$p_0(1-x)^{p_1}/(x^{p_2+p_3\log(x)+p_4\log^2(x)})$ to the $N = 4$ (upper edge of the envelope) or $N = 3$ (lower edge of the envelope) distributions. For S_T values below 5.5 TeV the envelope is built piecewise from other template functions fitted to either the $N = 3$ or $N = 4$ distribution.

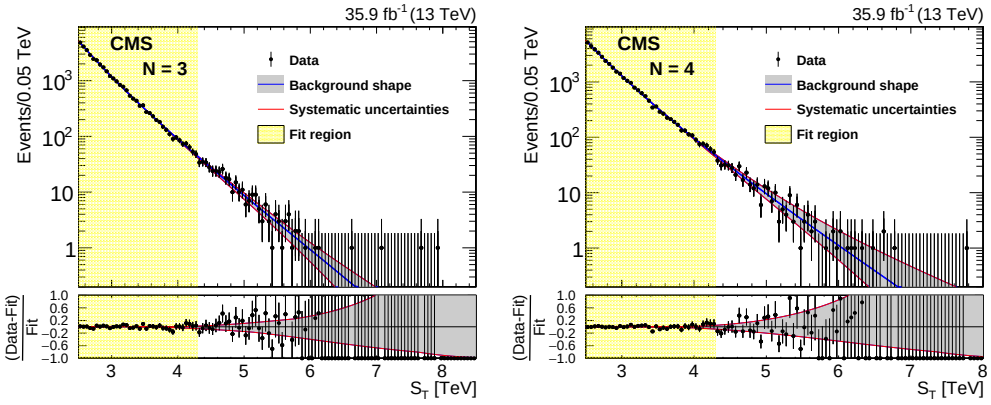

Figure 5.5: The results of the fit to data with $N = 3$ (left) and $N = 4$ (right), after discarding the functions that fail to monotonically decrease up to $S_T = 13$ TeV. The description of the best fit function and the envelope are given in the main text.

5.5.3 Background normalization

The next step in the background estimation for various inclusive multiplicities is to normalize the template and the uncertainty envelope, obtained as described above, to low-S_T data for various inclusive multiplicities. This has to be done with care, as the S_T invariance is only expected to be observed above a certain threshold, which depends on the inclusive multiplicity requirement. Indeed, since there is a p_T threshold on the objects whose transverse energies count toward the S_T value, the minimum possible S_T value depends on the number of objects in the final state, and therefore the shape invariance for an S_T spectrum with $N \geq N_{min}$

is only observed above a certain S_T threshold, which increases with N_{min}. In order to determine the minimum value of S_T for which this invariance holds, we find a plateau in the ratio of the S_T spectrum for each inclusive multiplicity to that for $N = 3$ in simulated multijet events. The plateau for each multiplicity is found by fitting the ratio with a sigmoid function. The lower bound of the normalization region (NR) is chosen to be above the 99% point of the corresponding sigmoid function. The upper bound of each NR is chosen to be 0.4 TeV above the corresponding lower bound to ensure sufficient event count in the NR. Since the size of the simulated QCD multijet background sample is not sufficient to reliably extract the turn-on threshold for inclusive multiplicities of $N \geq 9$–11, for these multiplicities we use the same NR as for the $N \geq 8$ distribution. Table 5.1 summarizes the turn-on thresholds and the NR boundaries obtained for each inclusive multiplicity.

The normalization scale factors are calculated as the ratio of the number of events in each NR for the inclusive multiplicities of $N \geq 3, \ldots, 11$ to that for the exclusive multiplicity of $N = 3$ in data, and are listed in Table 5.1. The relative scale factor uncertainties are derived from the number of events in each NR, as $1/\sqrt{N_{\mathrm{NR}}}$, where N_{NR} is the number of events in the corresponding NR.

5.5.4 Monte Carlo Closure Tests

The S_T shape invariance is checked with MC by applying the background estimation procedure to QCD MC and checking that the uncertainty encapsulates any shape non-invariance observed. It should be noted that in previous versions of this analysis, data has always been observed to be more invariant than MC, and the expectation is empirical. The ratio of the inclusive S_T spectra to the exclusive

Multiplicity	99% turn-on point (TeV)	Normalization region (TeV)	Normalization scale factor (data)
$\geq$3	2.44 ± 0.06	2.5–2.9	3.437 ± 0.025
$\geq$4	2.47 ± 0.06	2.5–2.9	2.437 ± 0.019
$\geq$5	2.60 ± 0.07	2.7–3.1	1.379 ± 0.016
$\geq$6	2.75 ± 0.11	2.9–3.3	0.652 ± 0.012
$\geq$7	2.98 ± 0.13	3.0–3.4	0.516 ± 0.015
$\geq$8	3.18 ± 0.21	3.2–3.6	0.186 ± 0.011
$\geq$9	3.25 ± 0.28	3.2–3.6	0.055 ± 0.006
$\geq$10	3.02 ± 0.26	3.2–3.6	0.012 ± 0.003
$\geq$11	2.89 ± 0.24	3.2–3.6	0.002 ± 0.001

Table 5.1: The S_T invariance thresholds from fits to simulated QCD multijet background spectra, normalization region definitions, and normalization scale factors in data for different inclusive multiplicities.

three spectrum are overlaid with the background prediction in figures 5.7 and 5.8. The background estimate using the QCD MC is shown in figures 5.9 and 5.10. We observe acceptable agreement of the predictions with the spectra in the QCD MC.

5.5.5 Comparison with data

The results of the background prediction and their comparison with the observed data are shown in Figs. 5.6 and 5.11 for inclusive multiplicities $N \geq 3, \ldots, 11$. The data are consistent with the background predictions in the entire S_T range probed, for all inclusive multiplicities.

5.6 Systematic uncertainties

There are several sources of systematic uncertainty in this analysis. Since the background estimation is based on control samples in data, the only uncertainties affecting the background predictions are the modeling of the background shape via template functions and the normalization of the chosen function to data at low S_T, as described in Section 5.5. They are found to be 1–130% and 0.7–50%, depending on the values of S_T and N_{min}, respectively.

For the signal, we consider the uncertainties in the PDFs, jet energy scale (JES), and the integrated luminosity. For the PDF uncertainty, we only consider the effect on the signal acceptance, while the PDF uncertainty in the signal cross section is treated as a part of the theoretical uncertainty and therefore is not propagated in the experimental cross section limit. The uncertainty in the signal acceptance is calculated using PDF4LHC recommendations [Ale11, Bot11] based on the quadratic sum of variations from the MSTW2008 uncertainty set (≈0.5%), as well as the variations obtained by using three different PDF sets: MSTW2008, CTEQ6.1 [Nad08], and NNPDF2.3 [Bal15] (up to 6% based on the difference between the default and CTEQ6.1 sets) for one of the benchmark models (non-rotating BH with $M_D = 3$ TeV, $M_{BH} = 5.5$ TeV, and $n = 2$, as generated by BlackMax); the size of the effect for other benchmark points is similar. To be conservative, we assign a systematic uncertainty of 6% due to the choice of PDFs for all signal samples. The JES uncertainty affects the signal acceptance because of the kinematic requirements on the objects and the fraction of signal events passing a certain S_T^{min} threshold used for limit setting, as described in Section 5.7. In order to account for these effects, the jet four-momenta are simultaneously shifted

up or down by the JES uncertainty, which is a function of the jet p_T and E_T, and the largest of the two differences with respect to the use of the nominal JES is assigned as the uncertainty. The uncertainty due to JES depends on M_{BH} and varies between < 1 and 5%; we conservatively assign a constant value of 5% as the signal acceptance uncertainty due to JES. Finally, the integrated luminosity is measured with an uncertainty of 2.5% [CMS17c]. Effects of all other uncertainties on the signal acceptance are negligible.

The values of systematic uncertainties that are used in this analysis are summarized in Table 5.2.

Uncertainty source	Effect on signal acceptance	Effect on background
PDF	±6%	NA
JES	±5%	NA
Integrated luminosity	±2.5%	NA
Shape modeling	NA	±(1–130)%, depending on S_T
Normalization	NA	±(0.7–50)%, depending on N_{min}

Table 5.2: Summary of systematic uncertainties in the signal acceptance and the background estimate.

5.7 Results

As shown in Figs. 5.6 and 5.11, there is no evidence for a statistically significant signal observed in any of the inclusive S_T distributions. The null results of the search are interpreted in terms of model-independent limits on BSM physics in energetic, multiparticle final states, and as a model-specific limit for EW sphalerons.

Limits are set using the CL_s method [Jun99, Rea02, ATL11] with log-normal

priors in the likelihood to constrain the nuisance parameters near their best estimated values. We do not use an asymptotic approximation of the CL_s method [CCG11], as for most of the models the optimal search region corresponds to a very low background expectation, in which case the asymptotic approximation is known to overestimate the search sensitivity.

5.7.1 Model-independent limits

The main result of this analysis is a set of model-independent upper limits on the product of signal cross section and acceptance ($\sigma \cdot A$) in inclusive $N \geq N_{min}$ final states, as a function of the minimum S_T requirement, S_T^{min}, obtained from a simple counting experiment for $S_T > S_T^{min}$. These limits can then be translated into limits on the M_{BH}^{min} in a variety of models, or on any other signals resulting in an energetic, multi-object final state. We start with the limits for the inclusive multiplicities $N \geq 3, 4$, which can be used to constrain models resulting in lower multiplicities of the final-state objects. Since part of the data entering these distributions are used to determine the background shape and its uncertainties, the limits are set only for S_T^{min} values above the background fit region, *i.e.*, for $S_T > 4.5$ TeV. For other multiplicities, the limits are shown for S_T values above the NRs listed in Table 5.1. These limits at 95% confidence level (CL) are shown in Figs. 5.12 and 5.13. When computing the limits, we use systematic uncertainties in the signal acceptance applicable to the specific models discussed in this paper, as documented in Section 5.6. It is reasonable to expect these limits to apply to a large variety of models resulting in multi-object final states dominated by jets. The limits on the product of the cross section and acceptance approach 0.08 fb at high values of S_T^{min}.

5.7.2 Model-specific limits

To determine the optimal point of S_T^{min} and the minimum multiplicity of the final-state objects N_{min} for setting an exclusion limit for a particular model, we calculate the acceptance and the expected limit on the cross section for a given model for each point of the model-independent limit curves, for all inclusive multiplicities. The optimal point of (N_{min}, S_T^{min}) is chosen as the point that gives the lowest expected cross section limit. In most of the cases this point also maximizes the significance of an observation, for the case of a nonzero signal present in data [CMS17d].

For the sphaleron signal, the optimal (N_{min}, S_T^{min}) point is chosen by scanning for the lowest expected limit and is found to be $(9, 6.2$ TeV) for $E_{sph} = 10$ TeV, and $(9, 5.5$ TeV) for $E_{sph} = 8\ TeV$ and 9 TeV. Consequently, the exclusion limit on the sphaleron cross section can be converted into a limit on the PEF, defined in Section 5.4. Following Ref. [ES16] we calculate the PEF limits for the nominal $E_{sph} = 9$ TeV, as well as for the modified values of $E_{sph} = 8$ and 10 TeV. The observed and expected 95% CL upper limits on the PEF are shown in Fig. 5.15. The observed (expected) limit obtained for the nominal $E_{sph} = 9\ TeV$ is 0.021 (0.012), which is an order of magnitude more stringent than the limit obtained in Ref. [ES16] based on the reinterpretation of the ATLAS result [Aad15].

We observe two events in the optimized signal region. A display of one of these events is shown in figure 5.14. Both events appear to be consistent with QCD background, having two back-to-back clusters of jets. The details of the jets in these events can be seen in tables 5.3 and 5.4.

Table 5.3: A table of the jets in the $N = 10$ and $S_T = 6.01$ TeV event in the sphaleron signal region.

Jet	$\mathbf{E_T}$ [GeV]	η	ϕ
1	1898.8	-0.20	-1.69
2	1464.2	0.90	2.16
3	731.7	-0.30	0.89
4	712.5	0.23	0.75
5	299.0	0.42	-0.86
6	254.9	0.32	-2.30
7	161.6	-0.17	2.04
8	137.4	0.90	-1.06
9	92.9	2.03	-0.98
10	91.9	-1.48	-2.01

Table 5.4: A table of the jets in the $N = 9$ and $S_T = 6.10$ TeV event in the sphaleron signal region.

Jet	$\mathbf{E_T}$ [GeV]	η	ϕ
1	2526.3	0.29	1.60
2	1377.9	-0.12	-1.73
3	746.7	0.46	-1.01
4	671.9	-0.16	-2.15
5	249.0	-0.63	-1.33
6	170.9	0.63	0.79
7	119.3	-1.00	1.66
8	107.3	-0.15	1.81
9	90.8	-0.15	-0.65

5.8 Summary

A search has been presented for generic signals of beyond the standard model physics resulting in energetic multi-object final states, such as would be produced by electroweak sphalerons. The search was based on proton-proton collision data at a center-of-mass energy of 13 TeV, collected with the CMS detector in 2016 and corresponding to an integrated luminosity of 35.9fb^{-1}. The background, dominated by QCD multijet production, is determined solely from low-multiplicity samples in data. Comparing the distribution of the total transverse momentum S_T of the final-state objects in data with that expected from the backgrounds, we set 95% confidence level model-independent upper limits on the product of the production cross section and acceptance for such final states, as a function of the minimum S_T for minimum final-state multiplicities between 3 and 11. These limits reach 0.08 fb at high S_T thresholds. By calculating the acceptance values for the sphaleron signal model, we convert these model-independent limits into the first experimental upper limit on the electroweak sphaleron pre-exponential factor of 0.021 for the sphaleron transition energy of 9 TeV.

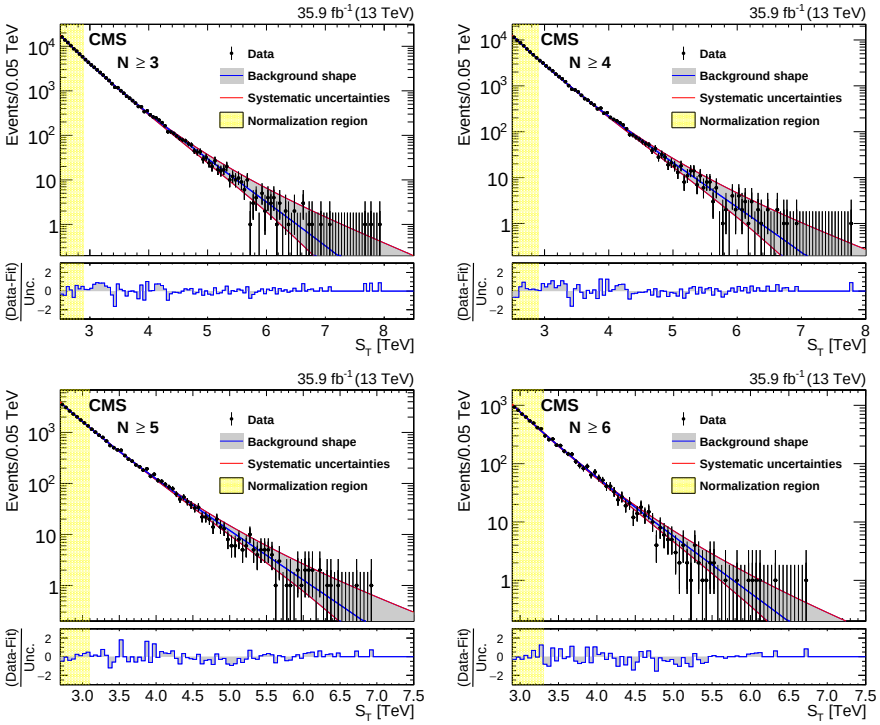

Figure 5.6: The background predictions after the normalization for inclusive multiplicities $N \geq 3, \ldots, 6$ (left to right, upper to lower). The gray band shows the background shape uncertainty alone and the red lines also include the normalization uncertainty. The bottom panels show the difference between the data and the background prediction from the fit, divided by the overall uncertainty, which includes the statistical uncertainty of data as well as the shape and normalization uncertainties in the background prediction, added in quadrature.

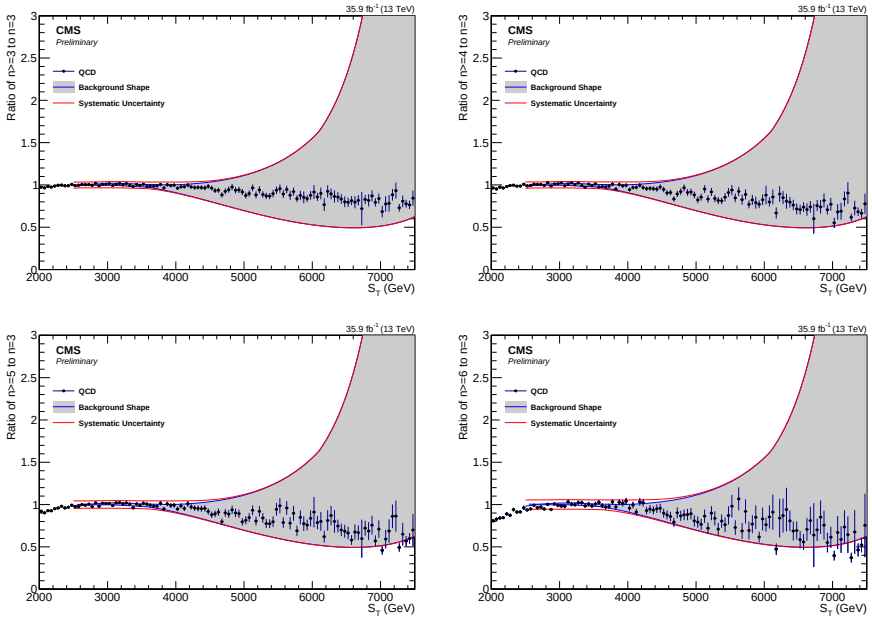

Figure 5.7: The ratio of the S_T spectrum in QCD simulation between inclusive multiplicity of 3-6 to exclusive multiplicity 3, which is normalized to the region where the curves first reach the plateau region. Grey band shows the shape uncertainty from full fitting procedure and red lines includes statistical uncertainty of normalization factor in quadrature.

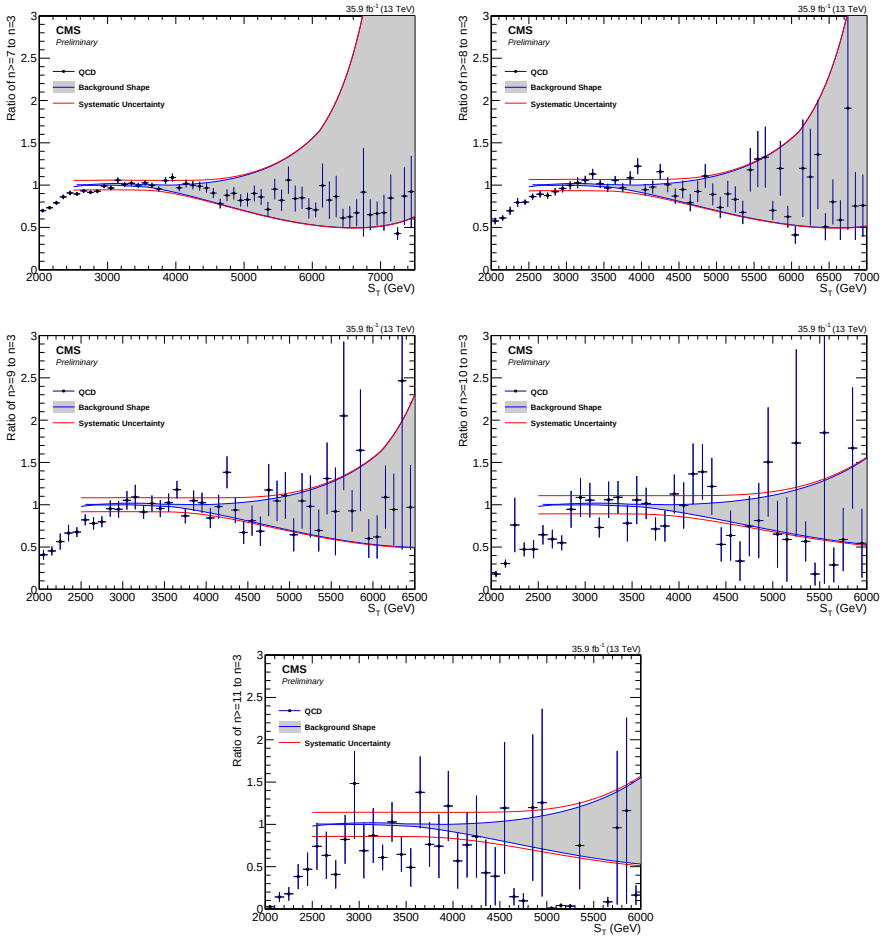

Figure 5.8: The ratio of the S_T spectrum in QCD simulation between inclusive multiplicity of 7-11 to exclusive multiplicity 3, which is normalized to the region where the curves first reach the plateau region. Grey band shows the shape uncertainty from full fitting procedure and red lines includes statistical uncertainty of normalization factor in quadrature.

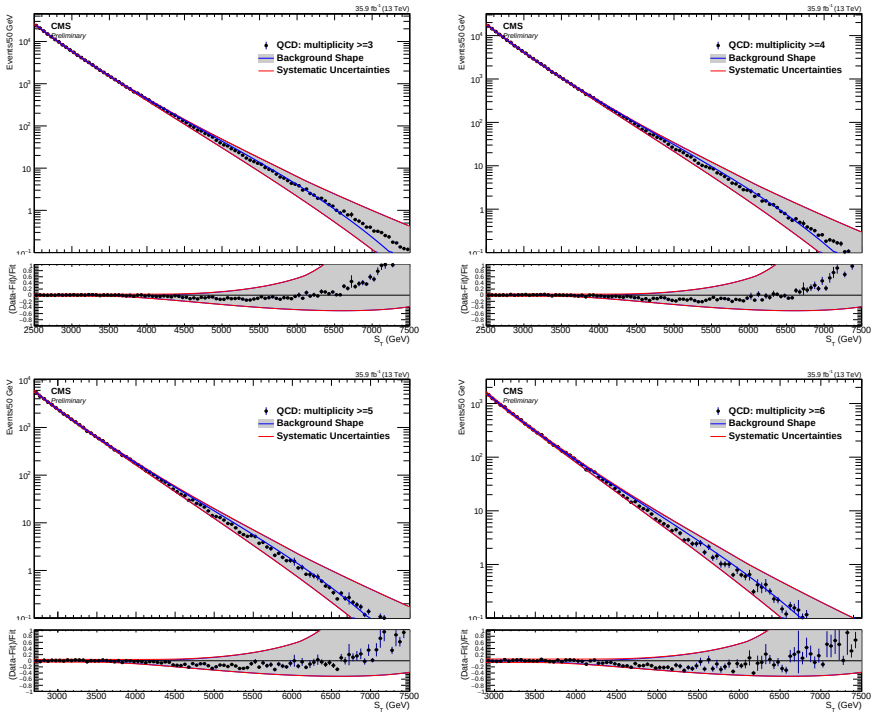

Figure 5.9: The distributions of the total transverse energy, S_T for inclusive multiplicities of objects (photons, muons, photons or jets) $N \geq 3, 4, 5, 6$. Observed data are shown by points with error bars, the solid blue lines along with the grey shaded band show the main background estimation (central blue line), along with the uncertainty band (outer blue lines). The deviation of the fit from the data is shown in the lower panes. The red lines are the normalization uncertainty and shape uncertainties added in quadrature.

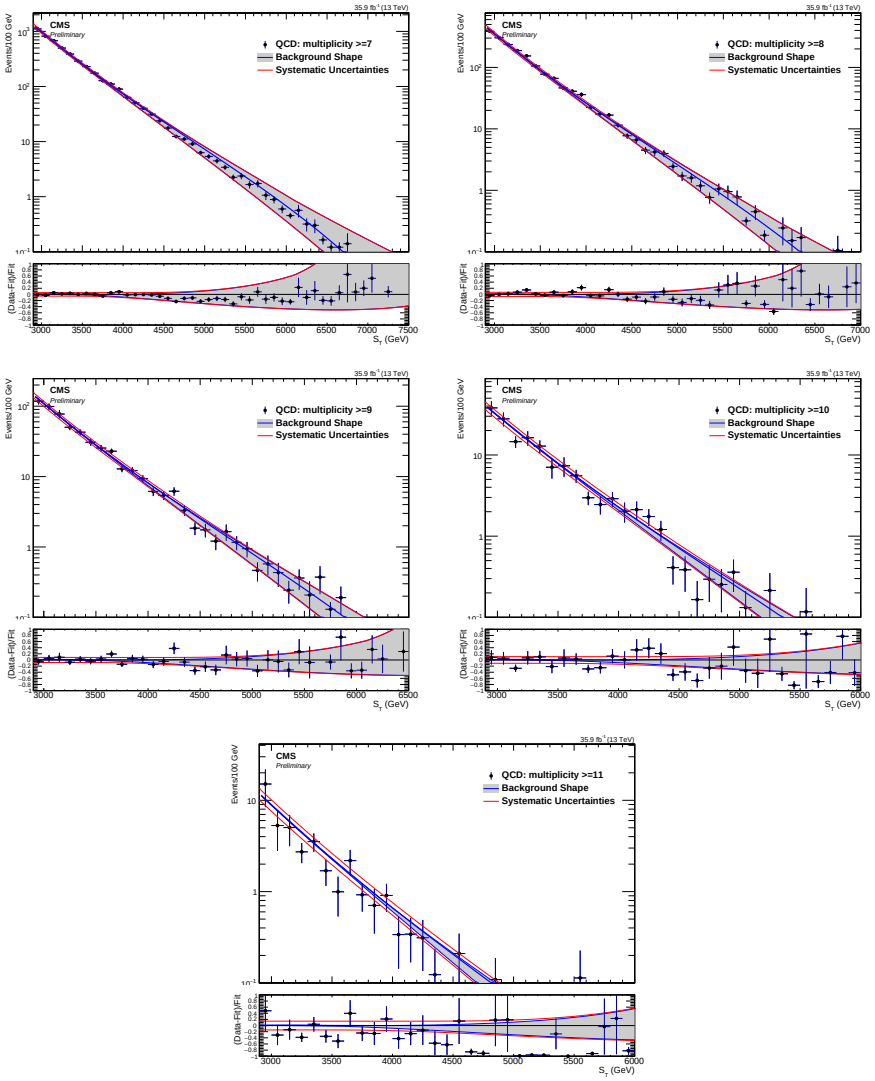

Figure 5.10: The distributions of the total transverse energy, S_T for inclusive multiplicities of objects (photons, muons, photons or jets) $N \geq 7, 8, 9, 10, 11$. Observed data are shown by points with error bars, the solid blue lines along with the grey shaded band show the main background estimation (central blue line), along with the uncertainty band (outer blue lines). The deviation of the fit from the data is shown in the lower panes. The red lines are the normalization uncertainty and shape uncertainties added in quadrature.

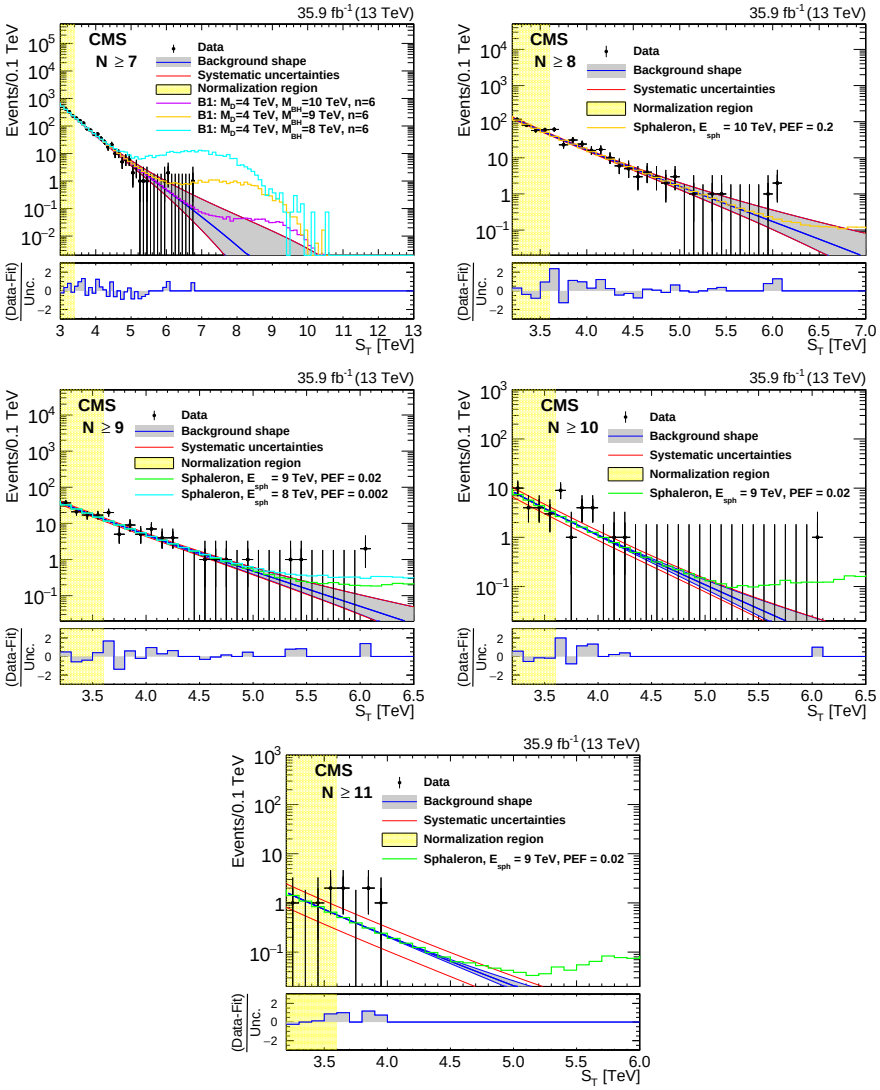

Figure 5.11: The background predictions after normalization for inclusive multiplicities of $N \geq 7, \ldots, 11$ (left to right, upper to lower). The gray band shows the shape uncertainty and the red lines also include the normalization uncertainty. The bottom panels show the difference between the data and the background prediction from the fit, divided by the overall uncertainty, which includes the statistical uncertainty of data as well as the shape and normalization uncertainties in the background prediction, added in quadrature. The $N \geq 7$ ($N \geq 8, \ldots, 11$) distributions also show contributions from benchmark BH and sphaleron signals added to the expected background.

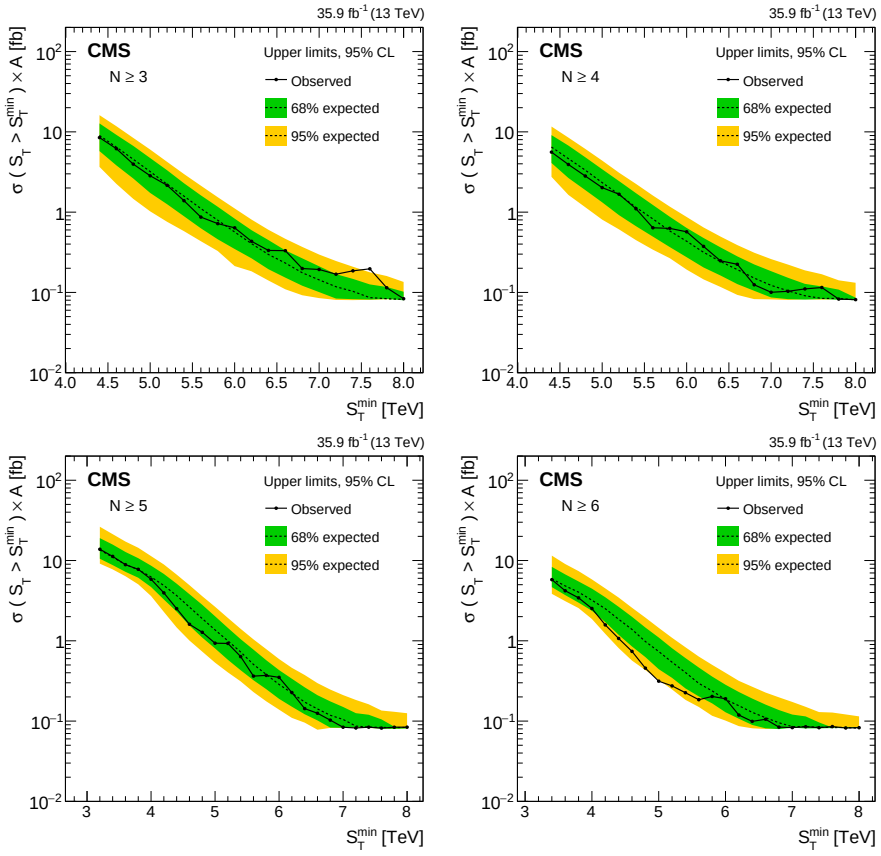

Figure 5.12: Model-independent upper limits on the cross section times acceptance for four sets of inclusive multiplicity thresholds, $N \geq 3, \ldots, 6$ (left to right, upper to lower). Observed (expected) limits are shown as the black solid (dotted) lines. The green (yellow) band represents the ± 1 (± 2) standard deviation uncertainty in the expected limit.

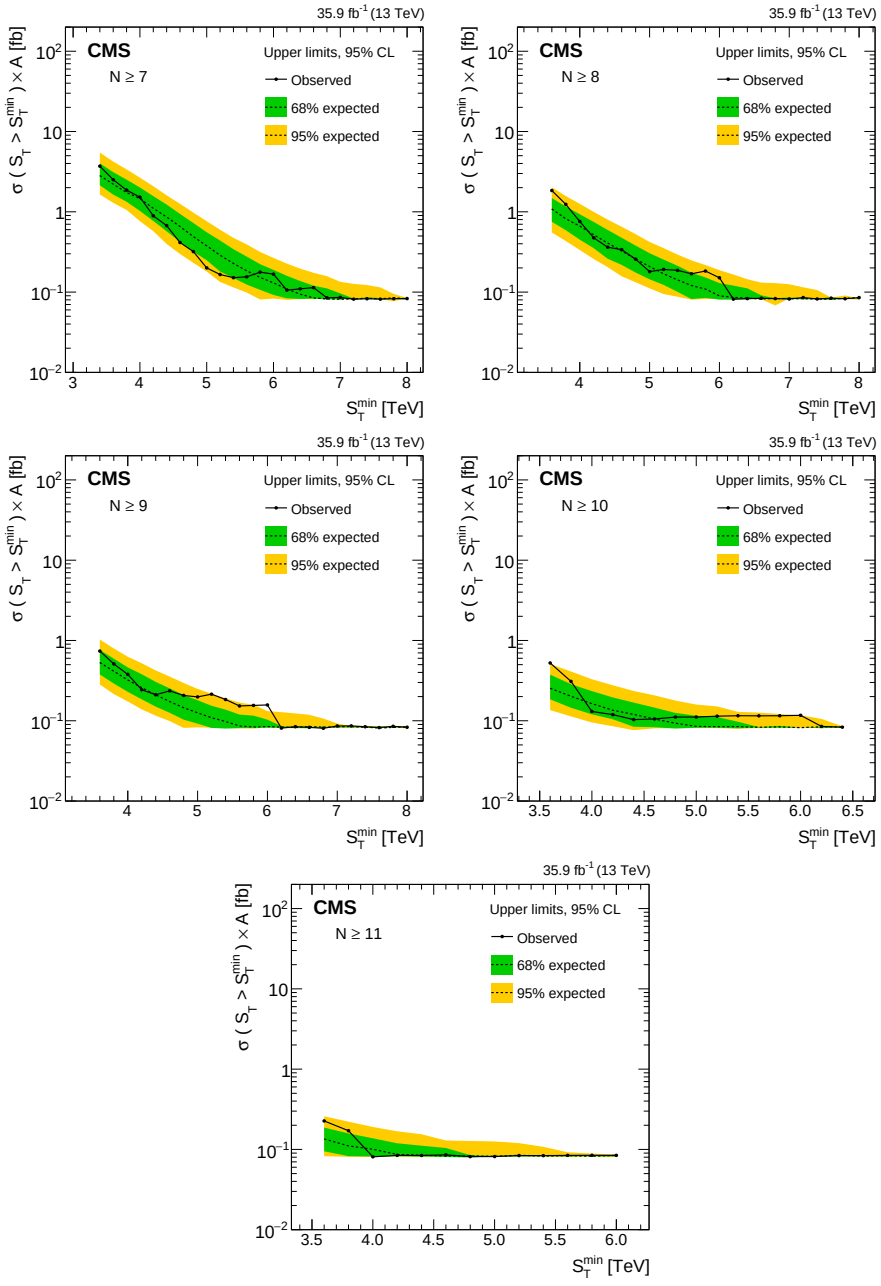

Figure 5.13: Model-independent upper limits on the cross section times acceptance for five sets of inclusive multiplicity thresholds, $N \geq 7, \ldots, 11$ (left to right, upper to lower). Observed (expected) limits are shown as the black solid (dotted) lines. The green (yellow) band represents the ± 1 (± 2) standard deviation uncertainty in the expected limit.

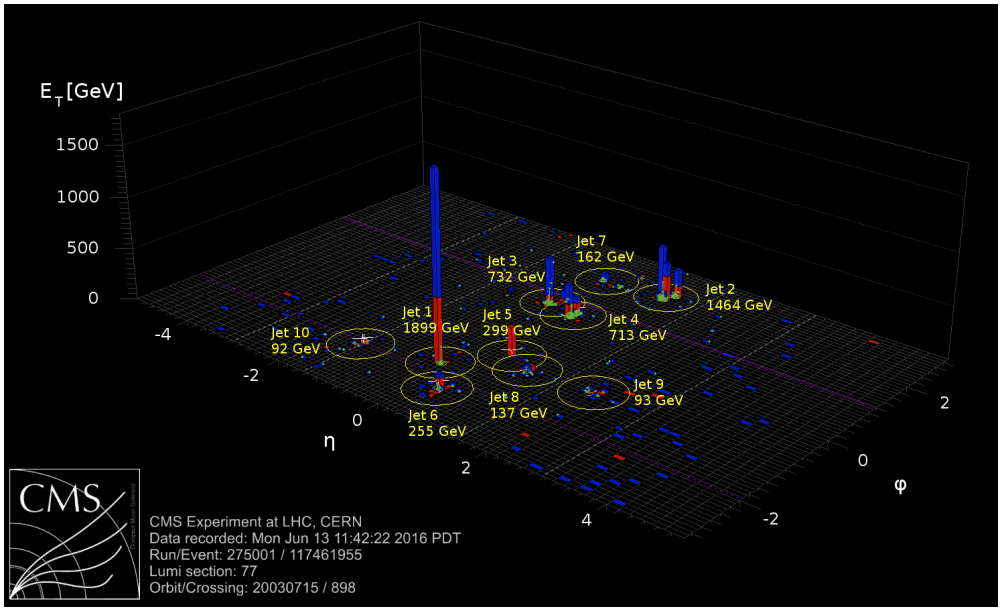

Figure 5.14: Lego display of one of two events in the signal region.

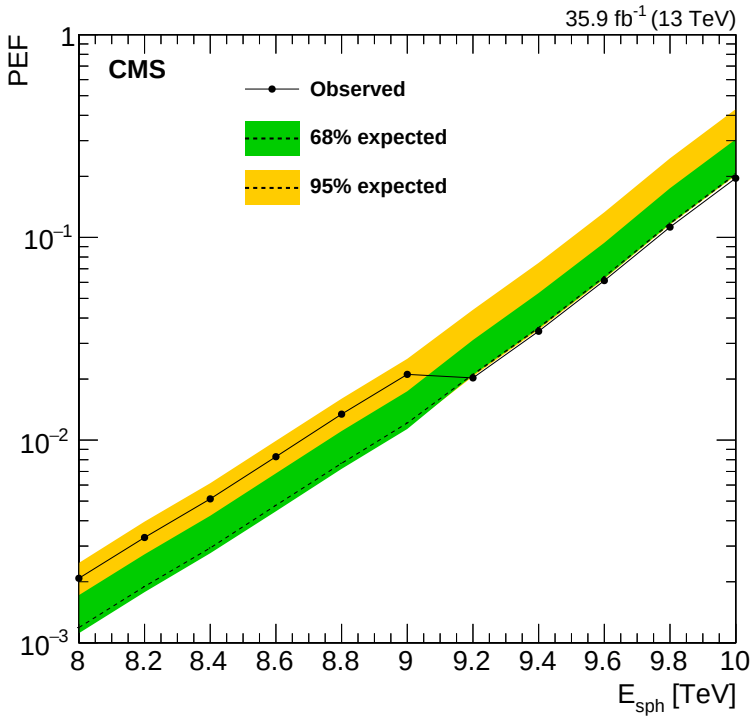

Figure 5.15: Observed (solid curve) and expected (dashed black curve) 95% CL upper limit on the pre-exponential factor PEF of the sphaleron production as a function of E_{sph}. The green (yellow) band represents the ± 1 (± 2) standard deviation uncertainty in the expected limit. The area above the solid curve is excluded by this search.

References

[Aad13] Georges Aad et al. "Search for TeV-scale gravity signatures in final states with leptons and jets with the ATLAS detector at $\sqrt{s} = 7$ TeV." *Phys. Lett.*, **B716**:122–141, 2013.

[Aad14] Georges Aad et al. "Search for microscopic black holes and string balls in final states with leptons and jets with the ATLAS detector at sqrt(s) = 8 TeV." *JHEP*, **08**:103, 2014.

[Aad15] Georges Aad et al. "Search for strong gravity in multijet final states produced in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC." 2015.

[Ae16] G. Aad and etal. "Search for strong gravity in multijet final states produced in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC." *Journal of High Energy Physics*, **2016**(3):26, Mar 2016.

[AFF14] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations." *JHEP*, **07**:079, 2014.

[al12] S. Chatrchyan et al. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC." *Physics Letters B*, **716**(1):30 – 61, 2012.

[Ale11] Sergey Alekhin et al. "The PDF4LHC Working Group Interim Report." 2011.

[Alw07] Johan Alwall et al. "A Standard format for Les Houches event files." *Comput. Phys. Commun.*, **176**:300–304, 2007.

[ATL11] ATLAS and CMS Collaborations. "Procedure for the LHC Higgs boson search combination in Summer 2011." ATL-PHYS-PUB-2011-011, CMS NOTE-2011/005, 2011.

[Bal15] Richard D. Ball et al. "Parton distributions for the LHC Run II." *JHEP*, **04**:040, 2015.

[BFL15] Andy Buckley, James Ferrando, Stephen Lloyd, Karl Nordström, Ben Page, Martin Rüfenacht, Marek Schönherr, and Graeme Watt. "LHAPDF6: parton density access in the LHC precision era." *The European Physical Journal C*, **75**(3):132, 2015.

[BGG16] Johannes Bellm, Stefan Gieseke, David Grellscheid, Simon Plätzer, Michael Rauch, Christian Reuschle, Peter Richardson, Peter Schichtel, Michael H. Seymour, Andrzej Siódmok, Alexandra Wilcock, Nadine Fischer, Marco A. Harrendorf, Graeme Nail, Andreas Papaefstathiou, and Daniel Rauch. "Herwig 7.0/Herwig++ 3.0 release note." *The European Physical Journal C*, **76**(4):196, Apr 2016.

[Bot11] Michiel Botje et al. "The PDF4LHC Working Group Interim Recommendations." 2011.

[BR97] Rene Brun and Fons Rademakers. "ROOT An object oriented data analysis framework." *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, **389**(12):81 – 86, 1997. New Computing Techniques in Physics Research V.

[Bra] Cameron Bravo. "https://github.com/cbravo135/BaryoGEN.".

[BT16] Constantin Bachas and Theodore Tomaras. "Band structure in Yang-Mills theories." *Journal of High Energy Physics*, **2016**(5):143, May 2016.

[CCG11] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. "Asymptotic formulae for likelihood-based tests of new physics." *Eur. Phys. J. C*, **71**:1554, 2011.

[CDS12] Laurent Canetti, Marco Drewes, and Mikhail Shaposhnikov. "Matter and antimatter in the universe." *New Journal of Physics*, **14**(9):095012, 2012.

[Cha08] S. Chatrchyan et al. "The CMS experiment at the CERN LHC." *JINST*, **3**:S08004, 2008.

[Cha12] Serguei Chatrchyan et al. "Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV." *JINST*, p. P10002, 2012.

[CMS11] CMS Collaboration. "Search for Microscopic Black Hole Signatures at the Large Hadron Collider." *Phys. Lett. B*, **697**:434, 2011.

[CMS12] CMS Collaboration. "Search for microscopic black holes in *pp* collisions at $\sqrt{s} = 7$ TeV." *JHEP*, **04**:061, 2012.

[CMS13] CMS Collaboration. "Search for microscopic black holes in pp collisions at sqrt(s) = 8 TeV." *JHEP*, **07**:178, 2013.

[CMS17a] "Search for black holes and other new phenomena in high-multiplicity final states in proton-proton collisions at $\sqrt{s} = 13$ TeV." *Physics Letters B*, **774**:279 – 307, 2017.

[CMS17b] "Jet algorithms performance in 13 TeV data." Technical Report CMS-PAS-JME-16-003, CERN, Geneva, 2017.

[CMS17c] CMS Collaboration. "CMS Luminosity Measurements for the 2016 Data Taking Period." CMS Physics Analysis Summary CMS-PAS-LUM-17-001, CERN, 2017.

[CMS17d] CMS Collaboration. "Search for black holes in high-multiplicity final states in proton-proton collisions at $\sqrt{s} = 13$ TeV." *Phys. Lett. B*, **774**:279, 2017.

[Col12] ATLAS Collaboration. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC." *Phys. Lett. B*, **716**:1, 2012.

[CRS18] David G. Cerdeño, Peter Reimitz, Kazuki Sakurai, and Carlos Tamarit. "B+L violation at colliders and new physics." *Journal of High Energy Physics*, **2018**(4):76, Apr 2018.

[CS74] Shiing-Shen Chern and James Simons. "Characteristic Forms and Geometric Invariants." *Annals of Mathematics*, **99**(1):48–69, 1974.

[CS08] Matteo Cacciari and Gavin P. Salam. "Pileup subtraction using jet areas." *Phys. Lett. B*, **659**:119, 2008.

[CSS08] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "The anti-k_t jet clustering algorithm." *JHEP*, **04**:063, 2008.

[CSS12] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "FastJet user manual." *Eur. Phys. J. C*, **72**:1896, 2012.

[CSS15] A Colaleo, A Safonov, A Sharma, and M Tytgat. "CMS Technical Design Report for the Muon Endcap GEM Upgrade." Technical Report CERN-LHCC-2015-012. CMS-TDR-013, Jun 2015.

[ES16] John Ellis and Kazuki Sakurai. "Search for sphalerons in proton-proton collisions." *Journal of High Energy Physics*, **2016**(4):86, 2016.

[Fey90] R.P. Feynman. *QED*. Penguin Books. Penguin, 1990.

[FFS16] Koichi Funakubo, Kaori Fuyuto, and Eibun Senaha. "Does a band structure affect sphaleron processes?" 2016.

[Fis22] R. A. Fisher. "On the interpretation of χ^2 from contingency tables, and the calculation of p." *J. Roy. Stat. Soc.*, **85**:87, 1922.

[FY86] M. Fukugita and T. Yanagida. "Barygenesis without grand unification." *Physics Letters B*, **174**(1):45 – 47, 1986.

[Hau99] J Hauser. "Primitives for the CMS Cathode Strip Muon Trigger." (CERN-OPEN-2000-070):5 p, Oct 1999.

[HK12] D. M. Hiemstra and V. Kirischian. "Single Event Upset Characterization of the Virtex-6 Field Programmable Gate Array Using Proton Irradiation." In *2012 IEEE Radiation Effects Data Workshop*, pp. 1–4, July 2012.

[Hoo76] G. 't Hooft. "Symmetry Breaking through Bell-Jackiw Anomalies." *Phys. Rev. Lett.*, **37**:8–11, Jul 1976.

[hr16] A. Svetek and M. Blake and M. Cepeda Hermida and S. Dasu and L. Dodd and R. Fobes and B. Gomber and T. Gorski and Z. Guo and P. Klabbers and A. Levine and I. Ojalvo and T. Ruggles and N. Smith and W.H. Smith and J. Tikalsky and M. Vicente and N. Woods. "The Calorimeter Trigger Processor Card: the next generation of high speed algorithmic data processing at CMS." *Journal of Instrumentation*, **11**(02):C02011, 2016.

[Iva03] V. N. Ivanchenko. "Geant4 toolkit for simulation of HEP experiments." *Nucl. Instrum. Meth.*, **A502**:666–668, 2003.

[Jun99] Thomas Junk. "Confidence level computation for combining searches with small statistics." *Nucl. Instrum. Meth.*, **A434**:435–443, 1999.

[Kha15a] Vardan Khachatryan et al. "Search for stealth supersymmetry in events with jets, either photons or leptons, and low missing transverse momentum in pp collisions at 8 TeV." *Phys. Lett. B*, **743**:503, 2015.

[Kha15b] Vardan Khachatryan et al. "Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV." *JINST*, **10**:P06005, 2015.

[Kha15c] Vardan Khachatryan et al. "Performance of photon reconstruction and identification with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV." *JINST*, **10**:P08010, 2015.

[Kha16] Vardan Khachatryan et al. "Event generator tunes obtained from underlying event and multiparton scattering measurements." *Eur. Phys. J. C*, **76**:155, 2016.

[Kha17a] Vardan Khachatryan et al. "Search for new physics in events with high jet multiplicity and low missing transverse momentum in proton-proton collisions at $\sqrt{s} = 8$ TeV." *Phys. Lett. B*, **770**:257, 2017.

[Kha17b] Vardan Khachatryan et al. "Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV." *JINST*, **12**:P02014, 2017.

[KM84] F. R. Klinkhamer and N. S. Manton. "A saddle-point solution in the Weinberg-Salam theory." *Phys. Rev. D*, **30**:2212–2220, Nov 1984.

[KN17] F. R. Klinkhamer and P. Nagel. "$SU(3)$ sphaleron: Numerical solution." *Phys. Rev. D*, **96**:016006, Jul 2017.

[KT90] Edward W. Kolb and Michael S. Turner. "The Early Universe." *Front. Phys.*, **69**:1–547, 1990.

[Nad08] P.M. Nadolski et al. "Implications of CTEQ global analysis for collider observables." *Phys. Rev. D*, **78**:013004, 2008.

[NGG12] Pavel Nadolsky, Jun Gao, Marco Guzzi, Joey Huston, Hung-Liang Lai, et al. "Progress in CTEQ-TEA PDF analysis." 2012.

[Oli14] K. A. Olive et al. "Review of Particle Physics." *Chin. Phys.*, **C38**:090001, 2014.

[Rea02] A L Read. "Presentation of search results: the CL(s) technique." *Journal of Physics G: Nuclear and Particle Physics*, **28**(10):2693, 2002.

[SAC15] Torbjörn Sjöstrand, Stefan Ask, Jesper R. Christiansen, Richard Corke, Nishita Desai, Philip Ilten, Stephen Mrenna, Stefan Prestel, Christine O. Rasmussen, and Peter Z. Skands. "An introduction to PYTHIA 8.2." *Comput. Phys. Commun.*, **191**:159, 2015.

[Sak91] Andrei D Sakharov. "Violation of CP invariance, C asymmetry, and baryon asymmetry of the universe." *Soviet Physics Uspekhi*, **34**(5):392, 1991.

[Sau97] F. Sauli. "GEM: A new concept for electron amplification in gas detectors." *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, **386**(2):531 – 534, 1997.

[Sir17] Albert M Sirunyan et al. "Particle-flow reconstruction and global event description with the CMS detector." *JINST*, **12**:P10003, 2017.

[Sir18] Albert M Sirunyan et al. "Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV." Submitted to JINST, 2018.

[ST17] Michael Spannowsky and Carlos Tamarit. "Sphalerons in composite and nonstandard Higgs models." *Phys. Rev. D*, **95**:015006, Jan 2017.

[Tro99] Mark Trodden. "Electroweak baryogenesis." *Rev. Mod. Phys.*, **71**:1463–1500, Oct 1999.

[TW15] S.-H. Henry Tye and Sam S. C. Wong. "Bloch wave function for the periodic sphaleron potential and unsuppressed baryon and lepton number violating processes." *Phys. Rev. D*, **92**:045005, Aug 2015.

[TW16] S. H. Henry Tye and Sam S. C. Wong. "The Chern-Simons Number as a Dynamical Variable." 2016.

[TW17] S.-H. Henry Tye and Sam S. C. Wong. "Baryon number violating scatterings in laboratories." *Phys. Rev. D*, **96**:093004, Nov 2017.