Identifying Lepton Reconstruction Failures in the CMS detector for a Disappearing Tracks Search

Undergraduate Research Thesis

Presented in partial fulfillment of the requirements for graduation with honors research distinction in Physics in the undergraduate colleges of The Ohio State University

by

Luca Lavezzo

The Ohio State University

May 2021

Project Advisor: Professor Chris Hill, Department of Physics

Identifying Lepton Reconstruction Failures in the CMS detector for a Disappearing Tracks Search

Abstract

A brief overview of the CMS experiment at the LHC and of related experimental particle physics is given. A recently published search at CMS for beyond the Standard Model long lived particles that have a 'disappearing tracks' signature is summarized. A method utilized in the search, names 'fiducial maps', which addressed lepton reconstruction inefficiencies of the CMS detector is studied in more detail, and solutions are proposed to replace it for electrons and muons independently. For electrons, a deep learning classifier is developed to identify electron tracks that fail to be reconstructed as electrons by the CMS software. This is shown to improve the reach of the search by reducing the electron background by $20 \pm 5\%$ compared to the fiducial maps; the classifier also retains 6-15% more signal, except for the longest Higgsino lifetime studied in the search, for which the fiducial map is shown to have slightly better performance. Similarly, a deep learning classifier is developed to identify muon tracks that fail the CMS reconstruction. The classifier identifies $28 \pm 3\%$ of isolated tag and probe muons that failed the reconstruction; however, this classifier is not able to identify the muons that pass the full selections of this search.

Contents

1. The CMS Experiment

Particle physics experiments seek to produce and measure particles through high energy collisions and dedicated detectors. As only the electron, proton, photon, and neutrinos are stable, all other particles decay: these unstable particles travel a distance of order $\gamma v \tau$ before decaying, where τ is each particle's mean lifetime. Depending on their lifetimes, some particles travel through portions of the detector and can thus be measured directly, while some can only be detected through their decay products. These are either stable particles, or particles with relatively longer lifetimes. To measure the properties of particles, their interactions with specific materials are exploited.

1.1. Interaction of particles with matter

1.1.1 Interaction of charged particles

When a relativistic charged particle passes through matter, it interacts electromagnetically with the atomic electrons and loses energy through ionization. The mean energy loss per unit length is given by the Bethe-Boloch equation,

$$\frac{dE}{dx} \approx -4\pi^2\hbar^2c^2\alpha^2\frac{nZ}{mv^2}\left\{ln\left[\frac{2\beta^2\gamma^2c^2m^2}{I}\right]\right\}$$

Where Z is the atomic number of the medium, n its number density, I its the effective ionization potential, and v is the speed of the traversing particle [1]. The ionized atoms and liberated electrons left behind by the charged particles can be detected, which can be used to reconstruct the particles' tracks using dedicated detectors, appropriately called trackers and which will be later discussed. By applying strong magnetic fields, which curve the charged particles' trajectories in helical paths owing to the Lorentz force law, the particles' momenta can be determined by measuring their radii of curvature.

1.1.2 Interactions of electrons and photons

While at low energies ionization dominates for charged particles, high energy electrons lose energy in matter primarily through bremsstrahlung (breaking radiation), whereby a photon is radiated when the electron interacts with the electromagnetic (EM) field of a nucleus. Also, high energy photons lose energy mostly through e^+e^- pair production when in the field of the nucleus. The characteristic amount of matter traversed is called the radiation length X_0, and the average radiation energy loss for electrons is approximated by,

$$\frac{dE}{dx} = -\frac{E}{X_0}$$

where E is the energy of the electron. These two phenomena, breaking radiation and e^+e^- pair production, give rise to EM showers, where a high energy electron radiates a photon, which will then produce a e^+e^- pair and these

two processes will continue to succeed one another, thus generating a 'shower' of photons and electrons/positrons. Alternatively, a photon can begin the shower by producing a e^+e^- pair, which in turn produce breaking radiation photons, and the EM shower follows. An EM shower stops when the produced particles fall below critical energy, defined to be the energy at which ionization and radiation loss are equal, where the electrons and positrons lose energy primarily by ionization, rather than breaking radiation [2].

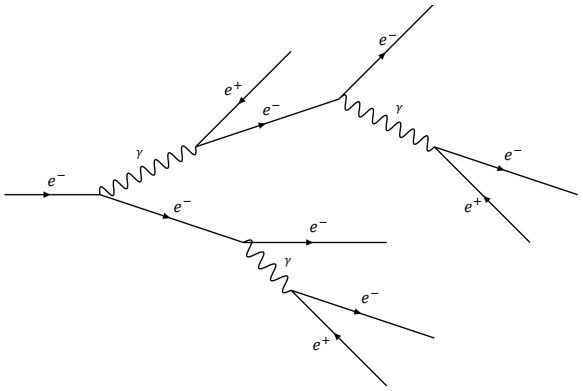

FIG. 1: Sample diagram of an electromagnetic shower. The processes will continue until the final state products fall below the critical energy.

1.1.3 The strong interactions of charged particles

High energy hadrons, on the other hand, can interact with the nucleus via the strong force, which will create several quarks and antiquarks, which combine to form mesons and baryons, a process called hadronization [2]. Hadronic showers are thus more variable than EM showers due to the many final states that are possible; these can even be electromagnetic, for example a π^0 can decay through $\pi^0 \to \gamma\gamma$, so that a fraction of the energy of the shower will be electromagnetic. Final state partons and hadrons typically appear in bunches called 'jets'.

1.2. LHC

The Large Hadron Collider (LHC) is a particle collider located 100 meters under Swiss and French territory near Geneva, Switzerland; it collides primarily proton beams, and occasionally heavy ions. Proton-proton (pp) collisions currently reach center of mass energy of $\sqrt{s} = 14$ TeV, with each beam of protons traveling around the LHC ring, of 27km circumference, with half of such energy. In order to achieve this, first hydrogen atoms are stripped of their electrons, leaving the protons. Several accelerators are used in sequence to accelerate these protons: first, Linac2 accelerates them to 50 MeV, then the Proton Synchrotron Booster (1.4 GeV), the Proton Synchrotron (25 GeV), and finally the Super Proton Synchrotron, where they reach 450 GeV. The LHC then takes them to 7 TeV using superconducting dipole magnets, cooled to 1.9 K with superfluid helium, producing a magnetic field of 8.3 T, and eight radio frequency (RF) cavities per beam. The phase oscillations of the electromagnetic fields inside these RF cavities divide the protons into 2808 bunches of approximately 1.15×10^{11} protons. They achieve this by accelerating

or decelerating the protons that have a different amount of energy or a different timing than what the RF cavity is set to, forcing them into bunches. These bunches are accelerated around the ring 25ns apart from each other. The bunches traveling in opposite directions around the main LHC ring collide at four interaction points (IPs) situated at the locations of four experiments: A Large Ion Collider Experiment (ALICE), A Toroidal LHC Apparatus (ATLAS), Compact Muon Solenoid (CMS), and Large Hadron Collider beauty (LHCb) [3].

1.3. CMS

CMS is one of the aforementioned experiments at the LHC, and one of its two general purpose detectors, along with ATLAS. CMS has a diameter of 15m and a length of 21.5m, and it is characterized by its solenoid magnet which produces a magnetic field of 4T, used to bend charged particles to measure their transverse momenta (p_T). Inside the magnet are, in order from the interaction point: the Inner Detector (ID), the electromagnetic calorimeter (ECAL), and the hadronic calorimeter (HCAL), while outside the magnet is another part of the HCAL, as well as the muon system, where another magnetic field of approximately 2T bends the muons in the opposite direction of the first magnet. Each of these subdetectors is specialized in measuring certain particles, but they are all used in conjunction to reconstruct events, as will be explained in the forthcoming sections. A full diagram of the structure of CMS is shown in Figure 2. The adopted coordinate system is with z along the beam line, ϕ as the polar angle around z, and θ as the angle from the z axis. Other useful variables that are often adopted are R, which is the distance extending radially outwards from the z axis, and η which is defined by $\eta = -ln\left[tan\left(\frac{\theta}{2}\right)\right]$ and which is useful because it is invariant under Lorenz boosts.

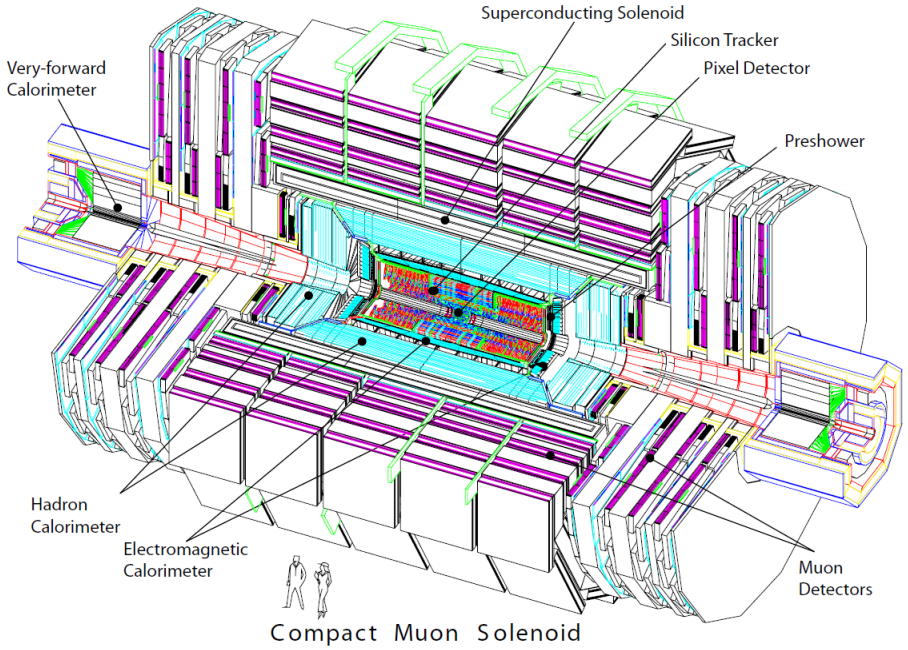

FIG. 2: A perspective view of the CMS detector, image taken from [4].

1.3.1 Inner Detector

The inner detector (ID) is designed to measure the trajectory, charge, and momentum of the particles traversing it, and to reconstruct secondary vertices. High granularity, quick response time, and resistance to high radiation are needed to make these measurements for an average of 1000 particles per bunch crossing (25 ns) [4]. The ID consists of two types of silicon detectors, the pixel detector and the silicon strip tracker, and it works by measuring the ionization of the charged particles. When the traversing charged particle ionizes the doped silicon wafer, electron-hole pairs are created, which move towards the collection electrodes due to an applied electric field [1]. These are arranged into silicon strips, or silicon pixels which give a two dimensional measurement. Several of these silicon wafers are arranged in different layers, and the hits measured in each of the layers are used to reconstruct the charged particles' tracks through the detector [1]. The pixel detector is the innermost component, and it is made of 4 barrel layers and 3 endcap disks. The silicon strip tracker sits just outside, extending to a radius of 1.1 m, and it is comprised of 15,148 strips arranged in 10 barrel layers and 12 endcap disks [4].

1.3.2 ECAL

The goal of the electronic calorimeter (ECAL) is to measure the energy and direction of electrons and photons. The ECAL is a homogenous calorimeter, so it is both an active scintillating material, to detect the electromagnetic signal, and an absorbing material, so that it initiates electromagnetic showers. Energy deposition is measured by the ECAL through the ionization or excitation of particles in lead tungstate $PbWO_4$ crystals, which are then picked up by dedicated photodetectors. The short radiation length of the crystals, $X_0 = 0.83$cm, means that the EM showers will be contained within a small region [1]. The photodetectors are designed to withstand the high radiation, high magnetic field environment, and are sufficiently fast compared to the LHC bunch crossing time. The ECAL has two main parts: the ECAL barrel (EB), covering $|\eta| < 1.479$, composed of 61200 $PbWO_4$ crystals and which uses avalanche photodiodes, and the ECAL endcaps (EE), covering $1.479 < |\eta| < 3.0$, composed of 7324 crystals in each, and which use vacuum phototriodes. The ECAL also has a preshower detector, covering $1.653 < |\eta| < 2.6$, meant to identify neutral pions, distinguish electrons against minimum ionizing particles, and improve position measurements. It is composed of two layers: a lead radiator in which showers are generated, and silicon strips sensors placed behind them which measure the deposited energy. The ECAL has a fine granularity: in the EB, the crystals achieve 360-fold granularity in ϕ and 85-fold granularity in each η direction [5].

1.3.3 HCAL

The hadronic calorimeters are designed produce hadronic showers when strongly interacting particles traverse its absorption material [4]. Since the nuclear interaction length is quite large, for example for iron $\lambda_T = 17cm$ compared

to a radiation length of 1.8cm, the HCAL occupies a larger part of the detector when compared to the ECAL [1]. The HCAL barrel rests between the ECAL, at R=1.77m, and the magnet, at R=2.95m, covering $|\eta| < 1.3$. The HCAL outer detector is placed outside the magnet in barrel region to catch the tail of the shower, and it is made of iron and plastic scintillators. The HCAL endcap covers $1.3 < |\eta| < 3$, and it is made of brass and plastic scintillators. The HCAL forward detector covers $4 < |\eta| < 5.2$ at $z = \pm 11.2$m, where it is subject to much higher radiation, and which uses a Cherenkov-based technology. The HCAL barrel and endcap are made of alternating layers of absorber, brass or steel, and tiles of plastic scintillators. The traversing particles interact in the absorber layers, producing numerous secondary particles and oftentimes showers, which the scintillators measure [4].

1.3.4 Muon System

The Muon System is composed of several subdetectors dedicated to measure muons with high precision. The p_T resolution is 1% in the barrel and about 3% in the endcaps for $p_T < 100$ GeV. The Drift Tubes (DTs) measure the muon position in the barrel $|\eta| < 1.2$. Each tube contains a wire and a gas mixture (85% Ar, 15% CO_2) that ionizes when traversed by a muon: the electrons follow the electric field to reach the wire, which records the signal. By recording the distance from the wires and the location along the wires, the DTs achieve two coordinates of the mouns' positions. The Cathode Strip Chambers (CSCs) cover the region $1.2 < |\eta| < 2.4$. They consist of positively charged anode wires crossed with negatively charged cathode panels within a gas volume (40% Ar, 50% CO_2, and 10% CF_4), which ionize when traversed by a muon: positive ions move towards the cathode and the electrons towards the anode wires. Since the strips and the wires are perpendicular, the CSCs give two spatial coordinates for each particle; they are characterized by a short drift path, which makes them suited for their location in the endcaps. Finally, the Resistive Plate Chambers (RPCs) are gaseous (96.2% $C_2H_2F_4$, 3.5% iC_4H_10, 0.3% SF_6, plus water vapor) parallel plate capacitors with high timing resolution, comparable to that of scintillators. There are 6 layers of RPCs in the barrel, and 4 layers in the endcaps [4].

1.3.5 Trigger/Particle Flow

The Level 1 (L1) Trigger uses local trigger information from all of the subdetectors except the Inner Tracker to decide whether to save or not the event using information from: the four highest E_T electrons, photons, central jets, forward jets, tau-jets, the four highest p_T muons, the missing transverse energy (MET) of the event, and the scalar sum of the jet transverse momenta H_T of the event. It achieves this using custom hardware and firmware to reduce the rate of events from 40 MHz to 100 kHz. The data is then processed by the High Level Trigger (HLT), which is relatively slower software that further selects which events to keep depending on the trigger menus, reducing the rate to approximately 1kHz [4]. The CMS offline, high level physics object reconstruction is achieved by the Particle Flow (PF) algorithm. This uses information from all the subdetectors to reconstruct all particles in the event; it's also at

this stage that MET is determined for the whole event, which is indicative of neutrinos, or possible BSM signatures. This algorithm will simply be referred to as the reconstruction employed at CMS, and will be relevant in Sections 2 and 3 where particles' tracks that fail to be reconstructed are studied, i.e. tracks in the detector that are known by other methods or by simulation to be associated but with a certain particle which are not identified as such by the Particle Flow reconstruction due to some sort of failure of the detector or the algorithm.

2. The Disappearing Tracks Search

'Disappearing tracks' searches look for long lived particles (LLPs) that decay within the CMS detector volume, whose products would go undetected due to either too little momentum or because they interact only weakly. Long lived particles obtain their name from their relatively long decay time, which allows them to measured directly in the volume of the CMS detector, rather than relying on their decay products for detection. An example of such particle is found in anomaly-mediated supersymmetry breaking (AMSB), where the chargino is long-lived and can decay within the CMS volume to a neutralino and a pion ($\tilde{\chi}_1^{\pm} \rightarrow \tilde{\chi}_1^0 \pi^{\pm}$); the pion has too little momentum to be reconstructed, and the neutralino interacts only weakly, leaving no trace in the detector (for more information about AMSB, see [6, 7]). Other searches at CMS and ATLAS excluded AMSB with a purely wino lightest supersymmetric particle (LSP), for chargino masses below 460GeV with a lifetime of 0.2 ns, and below 715GeV for a lifetime of 3 ns [8, 9]. A recent search at CMS [10] sought to expand those limits by looking at pp collision data collected between 2017 and 2018 at $\sqrt{s} = 13$ TeV, corresponding to a total of $101 fb^{-1}$. The results of this search are the topic of this section, and the remainder of this thesis will cover the improvements that were studied for the next version of this disappearing tracks search.

2.1. Simulated Data

In addition to the data collected at CMS, simulated signal events are produced at leading order (LO) precision with Pythia 8.240 using the NNPDF3.0 LO parton distribution function [11, 12]. Chargino masses are produced in the range 100–1100 GeV for the wino LSP case, and in the range 100-900 GeV for the higgsino LSP case, both in steps of 100 GeV. The mean lifetimes τ of these simulated signal events are varied from 6.67 ps to 333 ns in logarithmic steps to explore the experiment's sensitivity to a range of models.

2.2. Event Selection

A disappearing track signature, produced by LLPs that would decay within the tracking detectors and whose products would be undetected, would appear as having missing outer hits in the tracking detector, little energy deposited in the calorimeters, and no associated hits in the muon systems.

Since there is no tracking information available for the L1 trigger, L1 and HLT triggers were set to have certain thresholds of p_T^{miss}, which would be produced by an initial-state-radiation (ISR) jet recoiling against an electroweakino pair, and occasionally also had tracking requirements such as isolated tracks above a certain p_T and associated tracker hits.

The offline selection starts by requiring $\vec{p}_T^{\,miss,\not{\mu}} > 120$ GeV and at least one jet with $p_T > 110$ GeV and $|\eta| < 2.4$, to be consistent with the topology of an IRS jet. To reject events with spurious $\vec{p}_T^{\,miss}$ from mismeasured jets, $\Delta\phi$ between direction of highest p_T and $\vec{p}_T^{\,miss}$ must be greater than 0.5. Also, for events with at least two jets, the maximum $\Delta\phi$ between them is required to be less than 2.5 radians. A loss of power to the HEM for a period in 2018

forced a veto on $-1.6 < \vec{p}_T^{miss} < -0.6$ during that time.

After the basic selection, isolated tracks are selected with $p_T > 55$ GeV and $|\eta| < 2.1$, and are required to have no nearby large p_T jets and tracks. They must also must be associated with the primary pp interaction vertex, defined by requiring the transverse impact parameter $|d_0| < 0.02$ cm and the longitudinal impact parameter $|d_z| < 0.5$ cm.

The two main backgrounds of the search are isolated leptons whose tracks seem to disappear because of reconstruction failures, and spurious tracks which aren't associated to any real particle, and which are just failures of the reconstruction algorithm. Spurious tracks arise when missing hits in the tracker are allowed by the reconstruction algorithm, which, especially for shorter tracks, increases the combinatorial probability of the algorithm reconstructing a fake track. This background is suppressed by requiring no missing inner and middle hits, and at least four hits in the pixel detector.

The other source of backgrounds are isolated leptons. These events can seem to have disappearing tracks because of high breaking radiation in the case of electrons, nuclear interaction in the case of hadronically decaying taus (τ_h), the failure to associate energy from nonfunctional or noisy channels in the calorimeters, or other reconstruction failures due to either a malfunctioning part of the detector or the reconstruction algorithm. This background is suppressed by rejecting tracks within $\sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} < 0.15$ of a reconstructed lepton (electron, muon, τ_h). Regions of the detector with lower lepton reconstruction efficiency are vetoed (such as areas of incomplete coverage of the Muon System), as are tracks that pass near known noisy or faulty ECAL channels.

Regions of the detector of lower lepton reconstruction efficiency are identified using tag and probe studies for electrons, muons, and hadronically decaying taus. Efficiencies of regions $\eta \times \phi = 0.1 \times 0.1$ are calculated, and those regions which are below two standard deviations of the mean efficiency of the detector are rejected. This method was named "Fiducial Map" method: in Figure 3 the inefficiency of the regions are laid out in an η vs. ϕ plane, and the areas of excessive inefficiency are circled in green. There were several drawbacks to the fiducial map method. Firstly, the method causes a 4% loss of signal acceptance, reducing the impact of the search. Secondly, these fiducial maps had to be created for each of the three leptonic backgrounds, and for each of the data-taking periods. This can only be done once all the data is collected in order to discover inefficient areas of the detector, meaning that the analysis cannot be carried out until the end of of the data taking. Finally, this method does not address why certain leptonic events are not reconstructed, and does not identify unique failures of the reconstruction, but only regions that have a higher average inefficiency over the whole data taking period.

Finally, disappearing tracks are selected by requiring at least three missing outer hits in the tracker, as well as $E_{calo}^{\Delta R<0.5} < 10$ GeV. Three signal categories are defined depending to the number of tracker layers associated with the track n_{lay}: $n_{lay} = 4, n_{lay} = 5, n_{lay} \geq 6$.

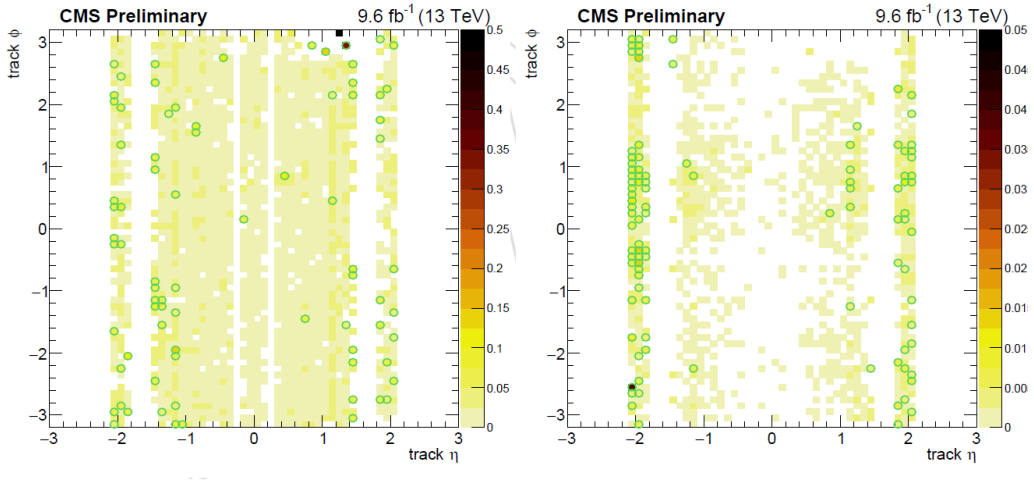

FIG. 3: Example of two fiducial maps from the previous analysis. Electrons are shown on the left and muons on the right, from the 2017C dataset. The green circles indicate the regions with at least $2\sigma_{mean}$ inefficiency.

2.3. Background Estimation

2.3.1 Leptons

For the lepton backgrounds, the reconstruction fails in a way that a track is observed but no lepton candidate is produced. To estimate this background, the selections are split into four: the probability for a lepton candidate to fail to be identified as a lepton (P_{veto}), the conditional probability of a single-lepton to pass the offline requirements given it passes the veto requirements (P_{off}), the conditional probability of single-lepton event passes the trigger requirement, given the veto and offline requirements (P_{trig}), and finally the conditional probability that a single-lepton event survives the HEM veto, given veto, offline, and trigger requirements (P_{HEM}). These conditional probabilities are estimated individually for each lepton flavor using tag and probe samples $Z \to e^-e^+$, $Z \to \mu^-\mu^+$, and $Z \to \tau^-\tau^+$. Their product gives the probability for a lepton to pass the selections, and this number is adjusted by the efficiencies of each of the single lepton triggers and the number of each lepton events with $p_T > 55GeV$, thus giving the total amount of estimated background from the charged leptons for each of the signal categories.

2.3.2 Spurious Tracks

To estimate the spurious tracks component of the background, $Z \to e^-e^+$, $Z \to \mu^-\mu^+$ are used as control samples. A 'sideband' region of $0.05 < |d_0| < 0.50$ cm is used to measure the frequency of spurious tracks. The sideband region is fit to the signal region by fitting to a Gaussian plus a constant, and everything is normalized by the number of events passing the basic selection. This procedure is repeated separately for each signal category.

2.4. Uncertainties

2.4.1 Background Estimates

The assumption is made that no energy is deposited in calorimeters by leptons that aren't identified; this was tested for electrons and τ_h to provide a systematic uncertainty. The background estimation for τ_h in the signal regions $n_{lay} = 4$ and $n_{lay} = 5$ are not robust, so values are averaged for $n_{lay} \geq 4$. The effect of this averaging is measured for electrons, where enough data is available, and the difference in the values range between 1 and 11%. The spurious track estimate rests on several assumptions: (1) that the probability of a spurious track is independent of the physics content of the event, (2) that the sideband method accurately describes the signal region, (3) the estimate is independent of the definition of the sideband region. The first two are tested using $Z \to e^- e^+$ and $Z \to \mu^- \mu^+$ control samples, and the third is tested by variating the sideband region. The fit parameters of the sideband method are varied by ± 1 standard deviation to obtain further systematic uncertainties.

2.4.2 Signal Selection Efficiencies

Several uncertainties were quantified and taken into account: theoretical uncertainties in the chargino production cross sections; uncertainty in the total integrated luminosity of the LHC; uncertainties in the signal yields due to corrections or scale factors, evaluated by varying each correction by ± 1 standard deviation of their measured uncertainties; uncertainty of the track reconstruction efficiency; uncertainty of the selection criteria; uncertainty in the trigger efficiency.

2.5. Results

Table I displays the observed number of events and the expected background in each category for each data-taking period, which are taken as independent counting experiments. The observations are consistent with the expected total background. Upper limits are set at 95% confidence level (CL) on the product of the cross section and branching fraction for each signal model.

Charginos in the wino LSP case with a lifetime of 3 (0.2) ns are excluded up to a mass of 884 (474)GeV at 95% CL, the most stringent constraints to date. In the higgsino LSP case, charginos 431 with a lifetime of 3 (0.05) ns are excluded up to a mass of 750 (175)GeV at 95% CL. The expected and observed upper limits on the product of cross sections of electroweak production and branching fractions in the wino LSP case are shown in Fig. 4 for four chargino lifetimes. Two-dimensional constraints derived from the intersection of the theoretical predictions with the expected and observed upper limits, for each chargino mass and mean proper lifetime considered, are shown in Fig. 5 for a purely wino LSP (left) and for a purely higgsino LSP (right).

TABLE I: Summary of the estimated backgrounds and the observation. The first and second uncertainties shown are the statistical and systematic contributions, respectively. Table obtained from [10]

Data-taking period	n_{lay}	Expected backgrounds			Observation
		Leptons	Spurious tracks	Total	
2017	4	$1.4 \pm 0.9 \pm 0.2$	$10.9 \pm 0.7 \pm 4.7$	$12.2 \pm 1.1 \pm 4.7$	17
	5	$1.1 \pm 0.4 \pm 0.1$	$1.0 \pm 0.2 \pm 0.6$	$2.1 \pm 0.4 \pm 0.6$	4
	$\geq$6	$6.7 \pm 1.1 \pm 0.7$	$0.04 \pm 0.04^{+0.08}_{-0.04}$	$6.7 \pm 1.1 \pm 0.7$	6
2018 A	4	$1.1^{+1.0}_{-0.6} \pm 0.1$	$6.2 \pm 0.5 \pm 3.5$	$7.3^{+1.1}_{-0.8} \pm 3.5$	5
	5	$0.2^{+0.6}_{-0.2} \pm 0.0$	$0.5 \pm 0.1 \pm 0.3$	$0.6^{+0.6}_{-0.2} \pm 0.3$	0
	$\geq$6	$1.8^{+0.6}_{-0.5} \pm 0.2$	$0.04 \pm 0.04^{+0.06}_{-0.04}$	$1.8^{+0.6}_{-0.5} \pm 0.2$	2
2018 B	4	$0.0^{+0.8}_{-0.0} \pm 0.0$	$10.3 \pm 0.6 \pm 5.4$	$10.3^{+1.0}_{-0.6} \pm 5.4$	11
	5	$0.4^{+0.7}_{-0.3} \pm 0.1$	$0.6 \pm 0.2 \pm 0.3$	$1.0^{+0.7}_{-0.3} \pm 0.3$	2
	$\geq$6	$5.7^{+1.2}_{-1.1} \pm 0.6$	$0.00^{+0.04}_{-0.00} \pm 0.00$	$5.7^{+1.2}_{-1.1} \pm 0.6$	1

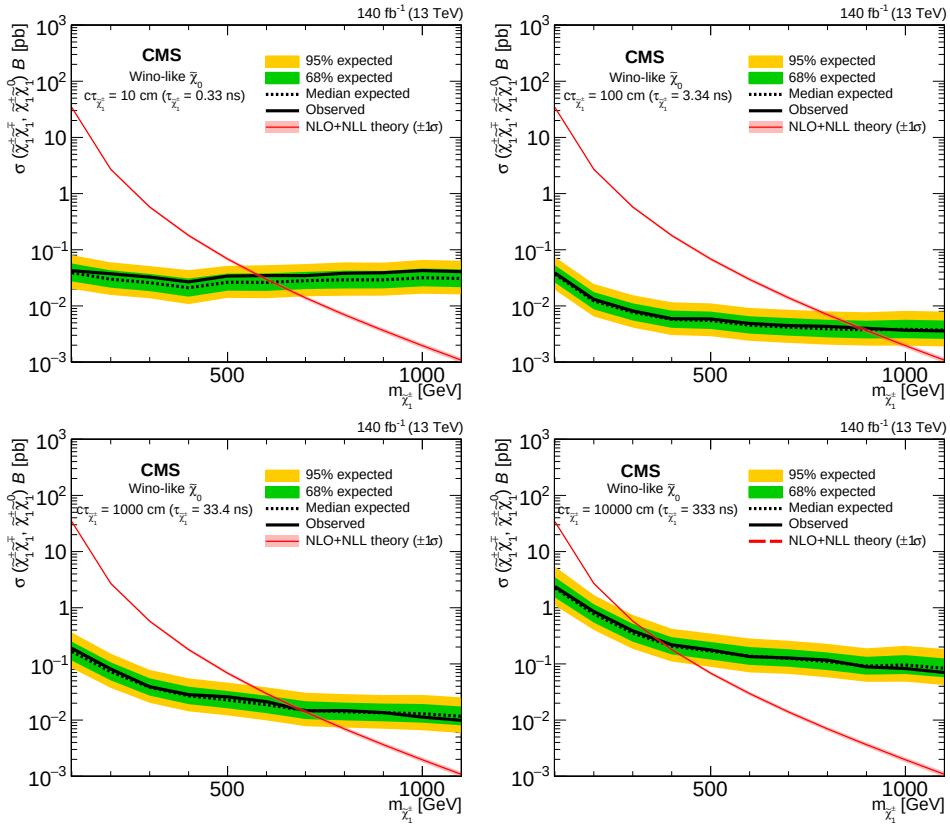

FIG. 4: The expected and observed 95% CL upper limits on the product of cross section and branching fraction for direct production of charginos as a function of chargino mass for chargino lifetimes of 0.33, 3.34, 33.4, and 333 ns, for a purely wino LSP. Shown are the full Run 2 results, derived from the results of the search in the 2017 and 2018 data sets combined with those of Ref. [13], obtained in the 2015 and 2016 data sets. Figures obtained from [10]

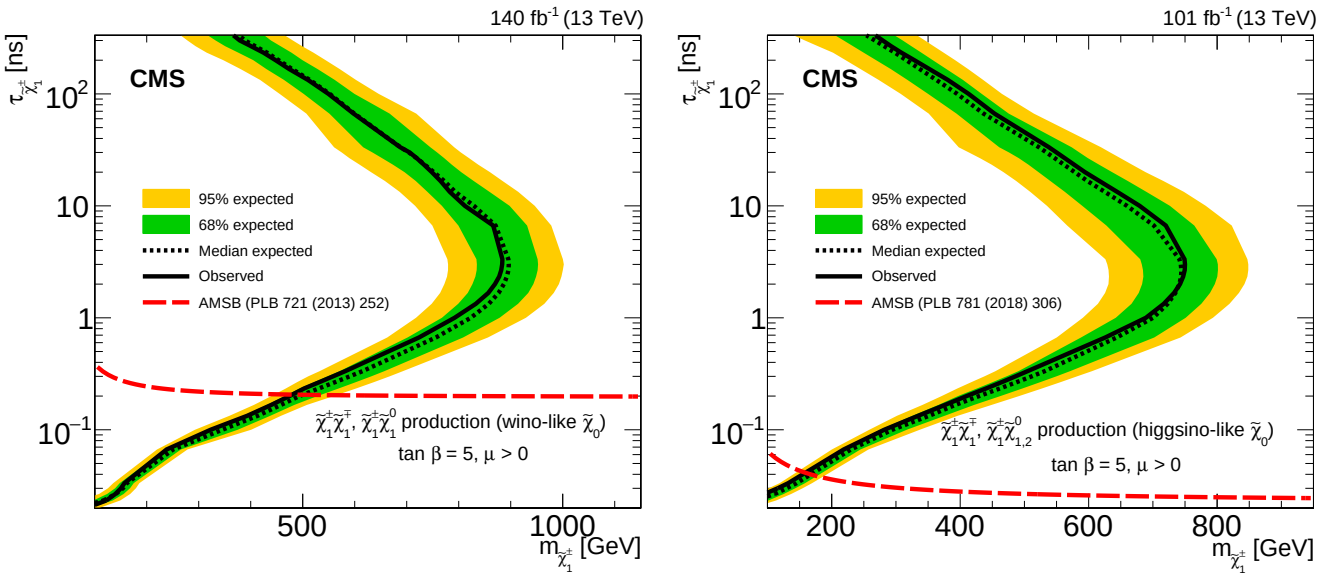

FIG. 5: The expected and observed constraints on chargino lifetime and mass for a purely wino LSP on the left, and a purely higgsino LSP on the right, in the context of AMSB, where the chargino lifetime is explicitly varied. For the wino LSP, the chargino branching fraction is set to 100% for $\tilde{\chi}_1^{\pm} \to \tilde{\chi}_1^0\pi^{\pm}$. For the higgsino LSP, following Ref. [14], the branching fractions are taken to be 95.5% for $\tilde{\chi}_1^{\pm} \to \tilde{\chi}_{1,2}^0\pi^{\pm}$, 3% for $\tilde{\chi}_1^{\pm} \to \tilde{\chi}_{1,2}^0 e\nu$, and 1.5% for $\tilde{\chi}_1^{\pm} \to \tilde{\chi}_{1,2}^0\mu\nu$ in the range of chargino masses of interest, with equal branching fractions and production cross sections between $\tilde{\chi}_1^0$ and $\tilde{\chi}_1^0$. For the purely wino LSP, the data was combined with those from Ref. [13], obtained in the 2015 and 2016 data sets. The regions to the left of the curves are excluded at 95% CL. The prediction for the chargino lifetime from Ref. [15]([16]) is indicated as the dashed line on the left (right) plots. Figures obtained from [10].

3. Deep Learning Approach to the Lepton Background

The fiducial map method, whereby regions of the detector with low lepton reconstruction efficiency are vetoed in the event selection of the disappearing tracks search, did not seek to understand why the tracks failed the lepton reconstruction, and had several drawbacks to the analysis already mentioned in Section 2.2. The following section will cover the methods that were developed to understand and identify these lepton reconstruction failures, with the goal of replacing the fiducial maps in the search. The results of these methods are considered in the context of the background and AMSB signal specific to the disappearing tracks search. These methods are also evaluated more generally as tools to identify leptons that failed the traditional reconstruction without the selections specific to the disappearing tracks search, demonstrating their utility to other analyses. Electrons and muons are treated as separate classification tasks due their different signatures in the various subdetectors of CMS, their different reconstruction algorithms, and thus different types of reconstruction failures. The other type of leptons present in the background of the search are the hadronically decaying taus, which have not been studied yet. For both classifications, deep learning-based classifiers were developed with an architecture based on deepSets [17], which was selected for its performance on Monte Carlo data independently for both electrons and muons.

3.1. DeepSets

The deepSets architecture was introduced by Zaheer et Al. [17]. Its defining characteristic is the ability to take sets as inputs; namely, the model's output is invariant to permutations of the elements of the input set. The main idea rests on the following theorem from the Zaheer et Al. paper:

Theorem 1 *A function $f(X)$ operating on a set X having elements from a countable universe, is a valid set function, i.e., invariant to the permutation of instances in X, iff it can be decomposed in the form $\rho(\sum_{x\in X}\phi(x))$, for suitable transformations ϕ and ρ.*

The standard neural network layer is represented as $\boldsymbol{f}_\Theta = \boldsymbol{\sigma}(\Theta x)$ where $\Theta \in \mathbb{R}^{M\times M}$ is the weight vector and $\sigma : \mathbb{R} \to \mathbb{R}$ is a nonlinearity function such as the sigmoid. Since neural networks act as universal approximators of functions, the technique of deepSets is to replace ϕ and ρ by neural network layers. Thus, each element x_m of the set is transformed by a number of layers into some latent representation $\phi(x_m)$. These representations are then added $\Sigma_m\phi(x_m)$ and this summation is processed by the ρ network for any task of regression or classification, which is a typical densely connected layer. Since the available and relevant information for the classification of the leptons' tracks are the hits in the various subdetectors of CMS, the permutation invariance of deepSets is theoretically useful, since there is no preferred order to those inputs.

3.2. Data Generation

For training the classifiers, Drell-Yan events are generated using samples `DYJetsToLL_M-5to50_TuneCP5_13TeV-madgraphMLM-pythia8` and `DYJetsToLL_M-50_TuneCP5_13TeV-madgraphMLM-pythia8`. Drell-Yan processes occur

when a quark and antiquark of two different hadrons annihilate to create a virtual photon or Z boson, which then pair produces a lepton and anti-lepton pair. The same reconstruction as used for the real data is applied to the simulated

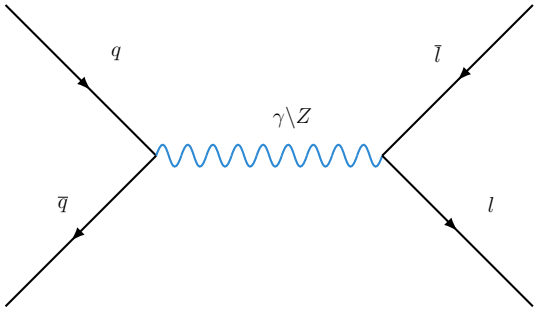

FIG. 6: Feynman diagram of sample Drell-Yan process.

events for 2017 and 2018 conditions. The cross section is calculated at next to next to leading order (NNLO) and is 71,310pb for 5 $GeV < M_{ll} < 50$ GeV, 18,610pb for 10 $GeV < M_{ll} < 50$ GeV, and 5,765.4pb for $M_{ll} > 50$ GeV. For validation of the classifiers, the same tag and probe $Z \rightarrow \bar{l}l$ samples used to estimate the background in the search are used, as well as the same simulated Higgsino samples.

3.3. Electrons

As more attention was turned toward studying the electron reconstruction failures, the energy deposition patterns in the ECAL showed that many appeared as normal electrons, but with parts of the shower missing when compared to the prototypical electromagnetic shower, as can been seen in Figure 7. This prompted the idea to apply techniques in computer vision and deep learning to learn to identify leptons that failed the reconstruction. With this motivation, a deep learning-based classifier was developed with an architecture based on deepSets [17], which was selected for its performance, with the goal of identifying electron tracks out of all the tracks that fail the electron reconstruction.

3.3.1 Training

Each track was associated with the 100 most energetic hits in the ECAL, HCAL, and MUO within a window of $\eta \times \phi = 0.25 \times 0.25$ centered around the track. Each track was then represented as a set of hits, and each hit was represented as a vector of the form,

$$(\Delta\phi_{(recHit-track)},\ \Delta\eta_{(recHit-track)},\ detType,\ E)$$

where $\Delta\phi_{(recHit-track)}$, $\Delta\eta_{(recHit-track)}$ are the differences in coordinates between the hit and the track, $detType$ indicates whether the hit was in the ECAL, HCAL, or MUO, and E is the energy of the hit for the ECAL or HCAL, while it was set to 1 for hits in the MUO where energy measurments aren't available. Since the hits do not contain any information regarding the track itself, track-level information was also added as input to the ρ network, which

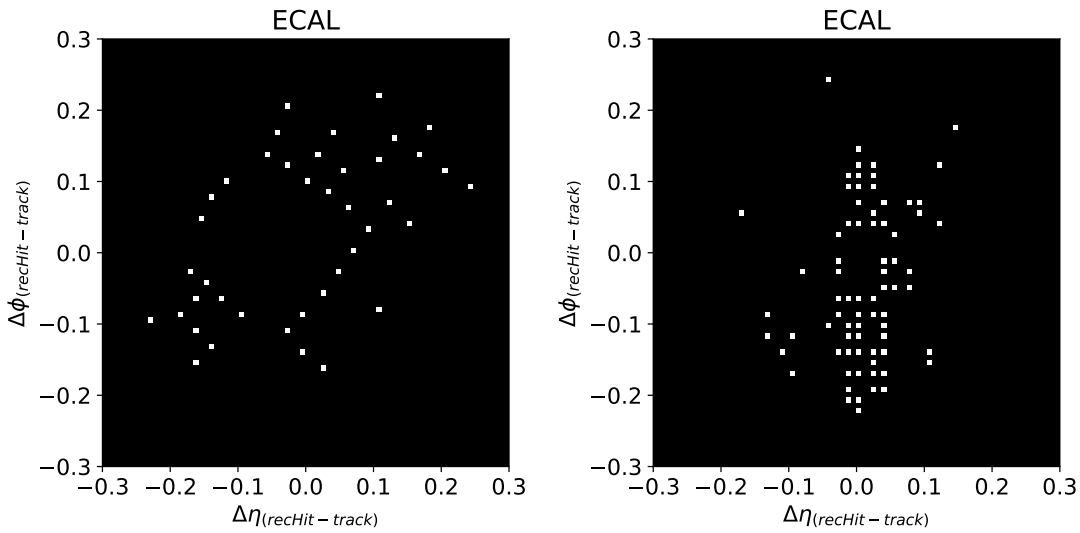

FIG. 7: Reconstructed hits in the ECAL associated with an electron plotted w.r.t. to the coordinates of the track they are matched to. Note the large missing 'chunks' in the middle of both of these showers, indicating malfunctioning parts of the ECAL.

included the track's absolute η and ϕ coordinates, the number of hits in the tracker associated to that track, and the amount of pileup in the event.

For training, tracks were selected that failed the electron, muon, and hadronically produced taus reconstruction vetoes, which are the three possible products of a Drell-Yan process which aren't neutrinos; this is the same selection as in the disappearing tracks search. Thus, tracks with incomplete detector coverage, such as between the ECAL barrel and endcap sections, are rejected. Furthermore, tracks are required to have at least $\sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} > 0.5$ from any jet and have $|\eta| < 2.4$. This minimal selection of isolated, non-reconstructed leptons is aimed at learning the different ways in which the electrons fail the reconstruction throughout the detector, and doesn't constrain the classifier to be applicable only to the disappearing tracks search.

The training data was split into two classes: tracks that were matched to simulated electrons versus tracks that were distant from simulated electrons. Since only tracks that failed the electron reconstruction are considered for this classification and the reconstruction is generally highly efficient ($> 85\%$ depending on the energy of the electron, see [18]), meaning that most tracks that are matched to simulated electrons are also reconstructed as electrons, after the selections are applied the Drell-Yan sample contained approximately 7 million non-electron tracks and 50,000 electron tracks. This imbalanced dataset made the classification task more challenging, and several techniques had to be attempted to address the issue: initial biases, class weights, and under-sampling of non-electron events were attempted in the training. Several models were tested on this data: among others, deepSets, convolutional neural networks, random forests, and anomaly detection methods. As mentioned already, deepSets was chosen due to its performance. The model's hyperparameters were chosen using a k-fold cross validation over the Drell-Yan sample. The following features were optimized this way: ϕ and ρ networks architectures, the number of hits in each set, the ratio of electron to non-electron events to train the classifier with, the window size in ϕ and η, and the inclusion and

selection of the track-level information.

3.3.2 Results

The final architecture included the track-level information mentioned in Section 3.1.3, and was trained using 50% of the training events as electron tracks, which represented a significant under-sampling of the non-electrons. The architecture for the ϕ networks was 3 fully connected layers with sizes [400, 256, 128], while the ρ network had 4 fully connected layers with sizes [128, 128, 64, 32]. Figure 8 shows the classifier output for the tag and probe samples

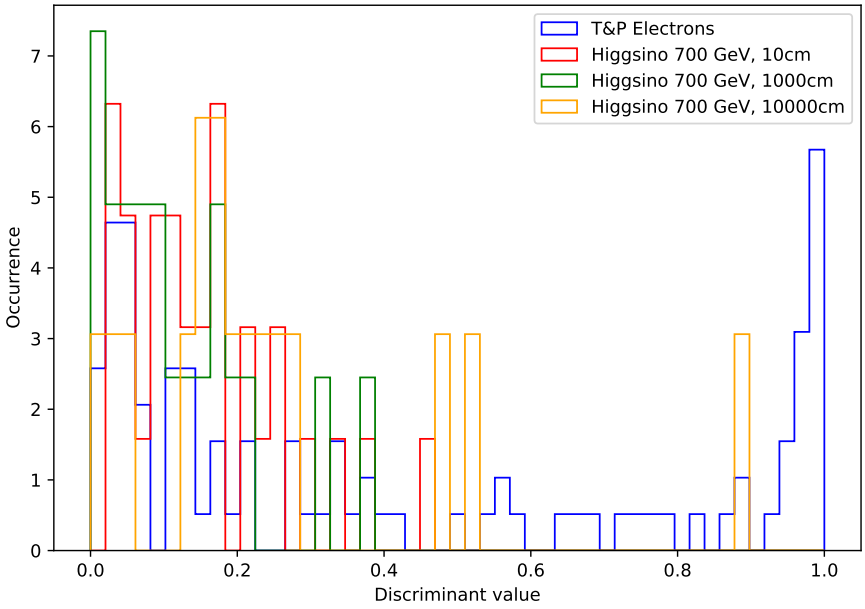

FIG. 8: Normalized histograms of the classifier output on some of the Higgsino samples and the tag and probe electrons, both of which passed the full analysis selection.

and the simulated Higgsinos of various lifetimes, both of which passed the full disappearing tracks analysis selections discussed in Section 2.2. The classifier was trained using "1" as the label for the electrons and "0" as the label for the non-electrons.

Sample	Number of Tracks	Pass deepSets	Pass Fiducial Maps	Eff. deepSets/ Eff. Fiducial Map
Tag & Probe Electrons	95	59	75	0.8±0.1
Higgsino 700 GeV, 10cm	31	31	29	1.1±0.3
Higgsino 700 GeV, 100cm	120	113	105	1.1±0.1
Higgsino 700 GeV, 1000cm	20	20	17	1.2±0.4
Higgsino 700 GeV, 10000cm	16	14	15	0.9±0.3

TABLE II: Performance of electron deepSets model on events that pass the full disappaering tracks analysis selection.

Table II shows the performance of the classifier and the fiducial maps on the Higgsino samples and the tag and probe electrons, with the full analysis selection applied. The number of events passing of each method is interpreted

as in the analysis, meaning that the events that pass each method are those labelled as non-electrons. For the classifier discriminant must be smaller than 0.5 to be labeled as a non-electron, while for the fiducial map method tracks are kept only in regions with an inefficiency less than $2\sigma_{mean}$, exactly as in the analysis. The last column in the table compares the performance of the fiducial map method and that of the classifier by taking the ratio of the number of events that pass each method. The classifier is thus able to cut out approximately an additional $20 \pm 5\%$ of the electron background compared to the fiducial map. Also, the classifier increases signal acceptance by 6-15% depending on the lifetime compared to the fiducial map, with exception of the 10000cm Higgsino where there is a 1 event difference between the methods, though the small sizes of these samples mean that the deepSets and the fiducial map methods efficiencies are within uncertainty. Therefore, the deepSets classifier is better able to identify the lepton reconstruction failures, as well as to retrieve some of the lost signal acceptance from the fiducial maps. It is interesting to note the bimodal peak of the deepSets ouput when evaluated on the tag and probe electrons (blue histogram in Figure 8). A future analysis of this data might try to develop a better understanding of the events in each of the peaks. Since the tag and probe sample is not perfectly pure, there is a non-electron background present in the sample, which would be expected to have a low discriminant value and might explain the peak around 0 in the classifier output.

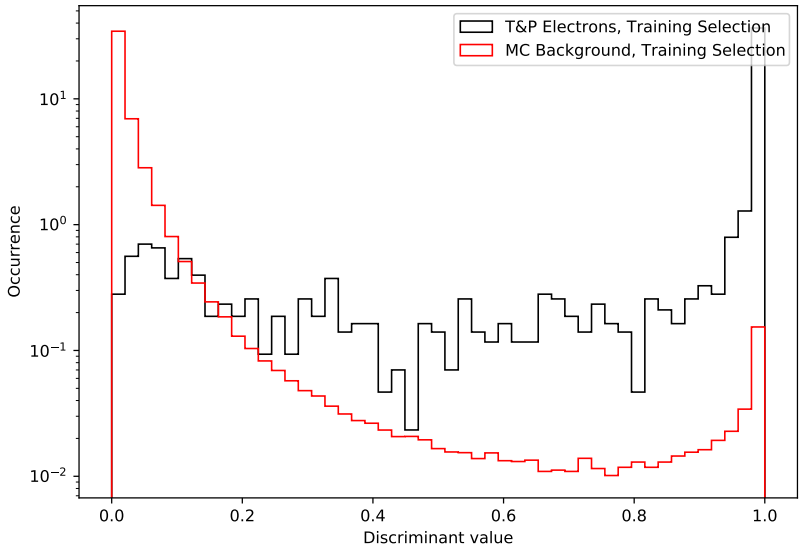

FIG. 9: Normalized histograms of the classifier output on the various samples with the training selections applied.

The classifier can also be validated on the tag and probe electrons that passed the training selection outlined in Section 3.1.3. This corresponds to a broader sample of real, isolated electrons that failed the reconstruction, without the selections specific to the disappearing tracks analysis, and is thus indicative of the utility of this model in other analyses. Since the model was trained on this larger set of events to learn electron reconstruction failures in general, the performance is expected to be better than on the subset of events that passed the disappearing tracks analysis selections. In addition, the classifier is evaluated on the sample of Monte Carlo tracks that were not matched or reconstructed to electrons, in order to compare its performance to the electron tracks. Only the training selection was applied to both of these samples. Figure 9 shows the classifier output for the two classes. Table III shows the

performance of the model using a cut on the classifier output at 0.5. As can be seen from Table III, the classifier

Sample	Classified as $e^{-\backslash+}$ (% of total)	Classified as not $e^{-\backslash+}$ (% of total)
Tag & Probe Electrons	1825 (87%)	274 (13%)
Drell-Yan Monte Carlo Non-Electrons	7632 (1%)	736549 (99%)

TABLE III: Performance of electron deepSets model on events that pass the training selection.

demonstrates a high level accuracy in identifying electrons that failed the reconstruction, especially compared to the fiducial map method, and presents an increase in the veto power when compared to the traditional reconstruction, and can thus have applications for analyses in which these type of reconstruction failures are present as a background.

3.4. Muons

The CMS reconstruction software relies mostly on the information from the Inner Tracker and the Muon Systems to identify muon tracks [19]. Thus, although the goal is similar as in the previous Section, namely to identify non-reconstructed muon tracks and thereby replace the fidual maps in the analysis, the reconstruction failures will be different given the different subdetectors responsible for identifying muons and electrons and the information they provide about the traversing particles. Nevertheless, the best performing method was again found to be a deep learning-based classifier with an architecture based on deepSets [17] which used the hits only from the Muon Systems along with information from the other subdetectors to classify the tracks.

3.4.1 Training

Reconstructed hits in the Muon Systems within a window of $\eta \times \phi = 1.0 \times 1.0$ centered around the track were considered for each track. The hits in all the subdetectors of the Muon Systems (CDCs, DTs, and RPCs) were then stored as a set, with each hit being of the form,

$$(\Delta\phi_{(recHit-track)},\ \Delta\eta_{(recHit-track)},\ \text{Station Number},\ \text{Time},\ detType)$$

where $\Delta\phi_{(recHit-track)},\ \Delta\eta_{(recHit-track)}$ are the difference in coordinates between the hit and the track; the station number refers where the hit was within a subdetector, which gives information about the third spatial coordinate radially outwards from the interaction point; the time is obtained from the segments in the MUO and is given in nanoseconds; finally, detType is the one-hot encoded index of whether the hit was in a CDC, DT, or RPC. Each track was represented by the 20 hits closest to the track. Track-level information was also added as input to the ρ network, which included the track's absolute η and ϕ coordinates, the amount of pileup, the sum of the energy deposited in the ECAL and HCAL (separately) within $\sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} < 0.1$ of the track, the p_T error, and the reduced χ^2. The last two variables reflect how well the reconstructed track fits the hits.

The Drell-Yan Monte Carlo sample discussed in Section 3.2 was used for training. Since we are interested in the

tracks that fail the muon reconstruction, the class imbalance issue previously observed for electrons is exacerbated by the fact that the muon classification has much higher precision than the electrons' ($> 97\%$ efficient) [19]. The result is that there are very few muons that fail the reconstruction, both in the Monte Carlo and tag and probe samples. In the whole MC sample, there are 7.7 million events that pass the training selection, of which only 313 are muons. In the case of the electrons, there were enough reconstruction failures to train a deep learning model (approximately 50,000 electron tracks), but the much smaller and more imbalanced muon dataset made the classification task very challenging. Classification attempts that were based on these two classes were unsuccessful.

Since we expect muons that fail the reconstruction to resemble muon events more closely than non-muons events, a classifier well trained to identify muons, even though they may be reconstructed, might be able to identify muons that failed the reconstruction. With this assumption, the veto on the muon reconstruction was removed and the rest of the training selections discussed in Section 3.1.3 were applied. The sample was split into two classes: tracks that were matched to a generated muon and were successfully reconstructed as muons versus tracks that were distant from generated muons and that were not reconstructed as muons. Different models were tested, among which were deepSets, convolutional neural networks (CNNs), random forests, decision trees, and anomaly detection methods. DeepSets and CNNs were identified as the best performing models; further testing of each was carried out on the reconstruction failures, and eventually a model based on deepSets was selected based on its performance. The architecture of the model, the window of $\eta \times \phi$, the track-level information, and the number of hits were optimized by validating different models on the Monte Carlo data.

3.4.2 Results

The final architecture chosen included the track-level information, trained on 20 MUO hits per track, and used a window of $\eta \times \phi = 1.0 \times 1.0$. The ϕ networks consisted of 3 densely connected layers with sizes [128, 64, 32], and the ρ network consisted of 2 densely connected layers with sizes [64, 32], and was trained with the Adagrad optimizer with a learning rate of 0.001.

Figure 10 shows the performance of the classifier on the Monte Carlo data split between 3 classes: reconstructed muons (red), non-reconstructed muons (blue), and tracks not matched or reconstructed as muons (black). The classifier was trained using "1" as the label for muons and "0" as the label for non-muons. The model is successful in discriminating between the classes on which it was trained (black and red), and is able to identify $32 \pm 10\%$ of the muons that failed the reconstruction using a cut at 0.5 of the discriminant value. The classifier has thus learned to identify at least some of the muon reconstruction failures, though its sensitivity could be improved. The assumption that the muon tracks which fail the muon reconstruction are more similar to muon tracks than non-muon tracks is then true for at least some of the reconstruction failures in the Monte Carlo sample.

Figure 11 shows the performance of the classifier on the tag and probe muons that passed the training selection and those that passed the full disappearing tracks analysis selection. The performance of the classifier on the tag

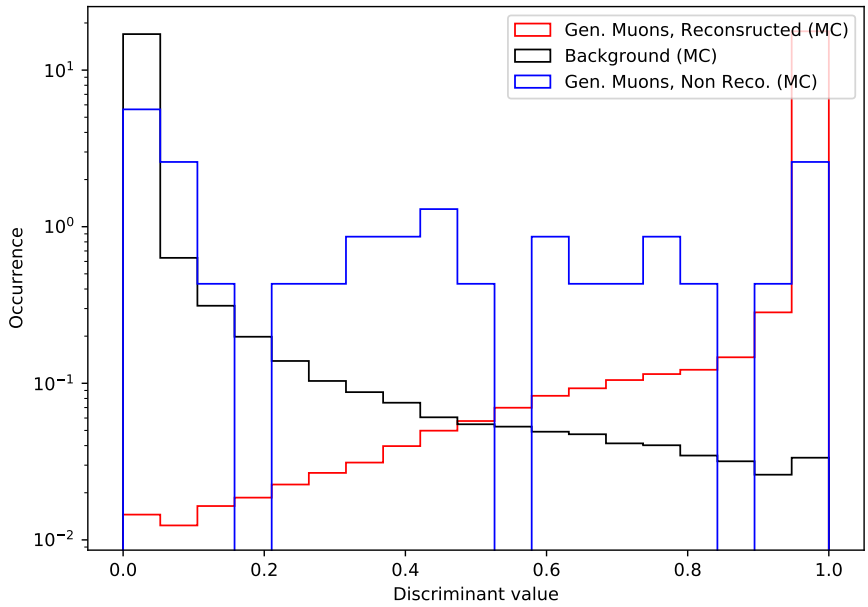

FIG. 10: Normalized histograms of the muon classifier output on Monte Carlo data passing the training selections. Note there are only 44 events in the blue histogram, compared to the 3 million tracks in each of the other two.

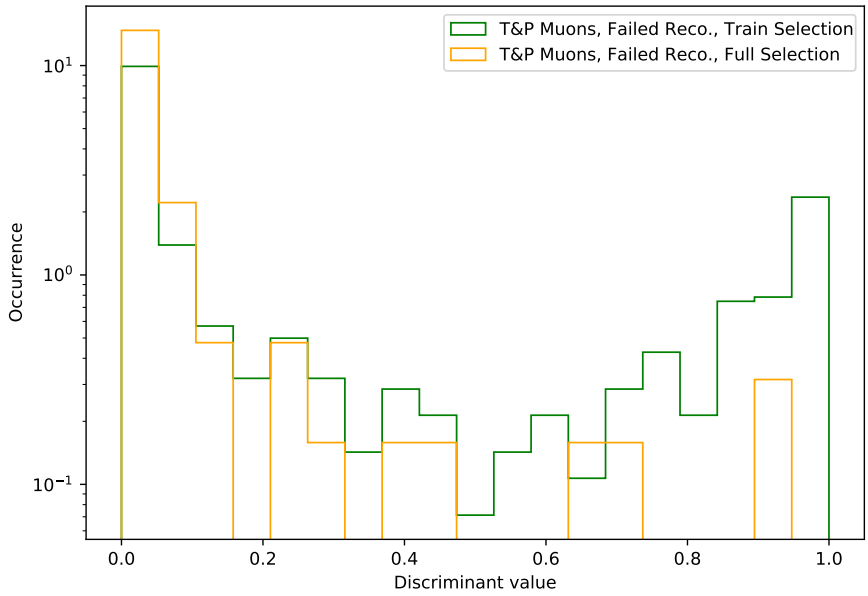

FIG. 11: Normalized histograms of the muon classifier output on tag and probe muons passing the training and full disappearing tracks analysis selections.

and probe sample with the training selection applied is comparable to the performance on the same sample from MC: the classifier identifies $28 \pm 3\%$ of the muon tracks as muons. Note that in both Monte Carlo and real data this method already represents an improvement on the veto of these isolated muons that fail the reconstruction. The poor performance of the classifier on the muons that pass the full analysis selections indicates that this subset of the tag and probe sample is fundamentally different in a way that the classifier cannot identify as muon-like. This might

mean that the reconstruction failures are different, that the tracks in the tag and probe sample that pass the full selection are not muons, or may hint at some other phenomenon that is not well understood. Its poor performance on this data prevents it from replacing the fiducial maps method in the analysis for muons.

Certain features of the tag and probe muons that failed the muon reconstruction are worth noting because they were different from the Monte Carlo tracks that passed the same selection, and might point to important discrepancies. The number of hits associated with each track within the $\eta \times \phi = 1.0 \times 1.0$ window are plotted in Figure 12 for the MC muons passing the training selection (blue), the tag and probe muons passing the training selection (green), and finally the tag and probe muons passing the full analysis selection (orange). Note that the tag and probe muons

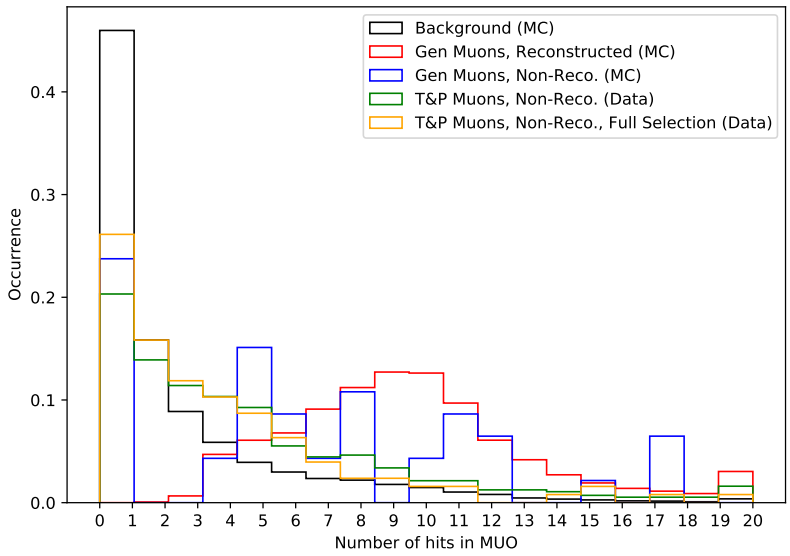

FIG. 12: Normalized distribution of the number of hits in the muon systems for different samples.

that failed the reconstruction (green, orange) have less hits associated to them when compared to well reconstructed muons (red) and to MC muons that failed the reconstruction (blue). Two hypothesis that we have formulated to explain these are: (1) the reconstruction failures are different, at least in some cases, between Monte Carlo and real data or (2) certain hits that should be associated to the tag and probe tracks are being lost when reconstructing the tracks. Another interesting feature was the η distribution of the tracks, shown in Figure 13 for the same samples as in the previous histogram. The tag and probe muons that failed the reconstruction (green, orange) peak around $\eta = \pm 2$, unlike the muons that failed the reconstruction from the MC sample (blue). This might indicate an area of low acceptance in the detector, or some particular type of failure characteristic of the subdetectors in that range of η that is discrepant between data and simulation.

More work needs to be carried out to understand these features of the data, and whether they are relevant to the classification task. Furthermore, the performance of the deepSets-based muon classifier on the events that pass the full analysis selection may not be solved by understanding these issues, and other methods should be considered. Given the difference between the muon reconstruction failures in the two samples, a model could be trained directly

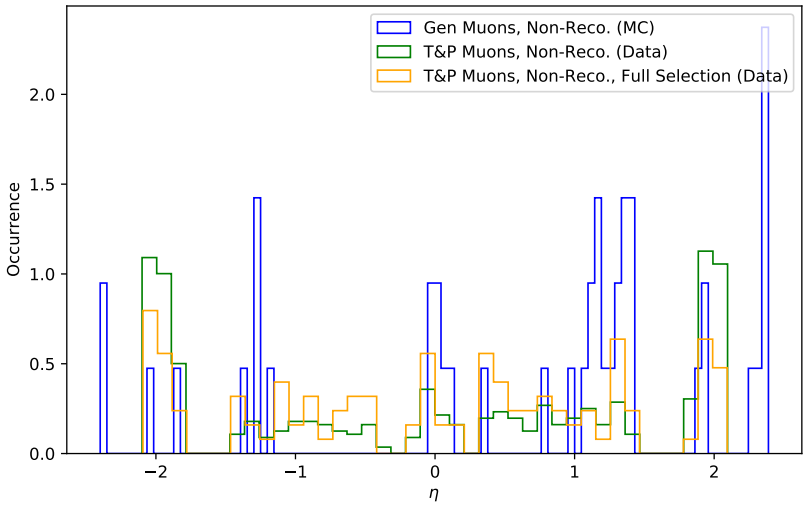

FIG. 13: Normalized η distribution of tracks for different samples.

on the tag and probe data, though it is a much smaller dataset which complicates the training of any deep learning algorithms. The highly imbalanced dataset and the fundamental differences between the reconstruction failures in Monte Carlo and real data complicate this task.

4. Conclusion

Two deep learning models are developed to identify electron and muon reconstruction failures in the CMS detector and improve a search for disappearing tracks. The first model was shown to identify $87 \pm 3\%$ of tag and probe electrons that failed the reconstruction with a cut on the discriminant at 0.5, demonstrating its utility for analyses and improvement to the reconstruction software. In the context of the disappearing tracks search, using the same cut on the discriminant value, this model was shown to veto $20 \pm 5\%$ of the electron background from the analysis, as well as retrieve 6-15% more signal, except for the longest Higgsino lifetimes, providing a suitable replacement for the fiducial maps method. The second classifier was shown to identify $28 \pm 3\%$ of the tag and probe muons that failed the reconstruction using a cut at 0.5 of the discriminant value. However, its poor performance on the tag and probe sample that passed the disappearing tracks search made it unable to replace the fiducial maps method in the analysis. The discrepancies between the muons from tag and probe and Monte Carlo samples are studied which may hint at the reason for the failure of this model on the events that pass the disappearing tracks analysis selections; further work will be needed to either improve this model or develop a new one.

Acknowledgements

The author wishes to thank Prof. Chris Hill and Dr. Brian Francis for their guidance and patience throughout the writing of this thesis, and throughout the last four years that I have spent learning and researching for the milliQan and CMS experiments. Additionally, the author would like to thank Prof. Richard Hughes for his guidance into the world of big data anlysis and machine learning which became an important element of this project, and Prof. Eric Katz for his excellent teaching of mathematics, high humours despite the COVID times, and unsolicited music recommendations. The author would like to thank Brandon Manley for all the advice and feedback on the drafts of this thesis, and for the years of collaborations and discussions in learning about this field together.

[1] M. Thomson, *Modern Particle Physics* (Cambridge University Press, 2013).

[2] M. Tanabashi and et al. (Particle Data Group), Review of particle physics, Phys. Rev. D **98**, 030001 (2018).

[3] L. Evans and P. Bryant, LHC machine, Journal of Instrumentation **3** (08), S08001.

[4] T. C. Collaboration, The CMS experiment at the CERN LHC, Journal of Instrumentation **3** (08), S08004.

[5] Energy calibration and resolution of the cms electromagnetic calorimeter in pp collisions at $\sqrt{s} = 7tev$, Journal of Instrumentation **8** (09), P09009–P09009.

[6] G. F. Giudice, R. Rattazzi, M. A. Luty, and H. Murayama, Gaugino mass without singlets, Journal of High Energy Physics **1998**, 027–027 (1998).

[7] L. Randall and R. Sundrum, Out of this world supersymmetry breaking, Nuclear Physics B **557**, 79 (1999).

[8] M. Aaboud, G. Aad, B. Abbott, O. Abdinov, B. Abeloos, S. H. Abidi, O. S. AbouZeid, N. L. Abraham, H. Abramowicz, and et al., Search for long-lived charginos based on a disappearing-track signature in pp collisions at $\sqrt{s} = 13$ tev with the atlas detector, Journal of High Energy Physics **2018**, 10.1007/jhep06(2018)022 (2018).

[9] A. M. Sirunyan, A. Tumasyan, W. Adam, F. Ambrogi, E. Asilar, T. Bergauer, J. Brandstetter, E. Brondolin, M. Dragicevic, and et al., Search for disappearing tracks as a signature of new long-lived particles in proton-proton collisions at $\sqrt{s} = 13$ tev, Journal of High Energy Physics **2018**, 10.1007/jhep08(2018)016 (2018).

[10] A. Sirunyan, A. Tumasyan, W. Adam, F. Ambrogi, T. Bergauer, M. Dragicevic, J. Erö, A. Escalante Del Valle, R. Frühwirth, M. Jeitler, and et al., Search for disappearing tracks in proton-proton collisions at s=13tev, Physics Letters B **806**, 135502 (2020).

[11] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, An introduction to pythia 8.2, Computer Physics Communications **191**, 159–177 (2015).

[12] R. D. Ball, V. Bertone, S. Carrazza, C. S. Deans, L. Del Debbio, S. Forte, A. Guffanti, N. P. Hartland, J. I. Latorre, and et al., Parton distributions for the lhc run ii, Journal of High Energy Physics **2015**, 10.1007/jhep04(2015)040 (2015).

[13] T. C. Collaboration, Search for disappearing tracks as a signature of new long-lived particles in proton-proton collisions at $\sqrt{s} = 13$ tev, JHEP **8**.

[14] S. D. Thomas and J. D. Wells, Phenomenology of massive vectorlike doublet leptons, Phys. Rev. Lett **81** (1998).

[15] S. M. M. Ibe and R. Sato, Mass splitting between charged and neutral winos at two-loop level, Phys. Rev. Lett **721**, 10.1016/j.physletb.2013.03.015 (2013).

[16] H. O. H. Fukuda, N. Nagata and S. Shirai, Higgsino dark matter or not: role of disappearing track searches at the lhc and future colliders, Phys. Rev. Lett **781**, 10.1016/j.physletb.2018.03.088 (2018).

[17] M. Zaheer, S. Kottur, S. Ravanbakhsh, B. Poczos, R. Salakhutdinov, and A. Smola, Deep sets (2018), arXiv:1703.06114 [cs.LG].

[18] Performance of electron reconstruction and selection withthe cms detector in proton-proton collisions at $\sqrt{s} = 8tev$, Journal of Instrumentation **10** (06), P06005–P06005.

[19] A. Sirunyan, A. Tumasyan, W. Adam, F. Ambrogi, E. Asilar, T. Bergauer, J. Brandstetter, E. Brondolin, M. Dragicevic, J. Erö, and et al., Performance of the cms muon detector and muon reconstruction with proton-proton collisions at s=13 tev, Journal of Instrumentation **13** (06), P06015–P06015.