JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN

Search for Highly Ionizing Particles with the Pixel Detector in the Belle II Experiment

SUCHE NACH HOCHIONISIERENDEN TEILCHEN MIT DEM PIXELDETEKTOR IM BELLE II EXPERIMENT

MASTER THESIS

by
Katharina Dort

supervised by

PD Dr. Sören Lange AR
II. Physics Institute
Faculty 07
Justus-Liebig-Universität Gießen

May 2019

ABSTRACT

The Belle II experiment located at the high-energy research facility KEK in Tsukuba (Ibaraki) Japan started operation in 2018. Compared to the predecessor experiment Belle, Belle II plans to increase the peak luminosity by a factor of 40, by employing nanobeam technology in the interaction region.

A novel Pixel Detector (PXD) featuring two layers of DEPFET silicon sensors has been installed in Belle II for tracking of charged particles close to the interaction region. The DEPFET technology allows to thin down the sensitive detector region to only 75 µm. The low material budget reduces multiple scattering and eliminates the demand for active cooling. With pixel sizes ranging between 50 µm to 85 µm the PXD has an excellent spatial resolution, which is crucial for the tracking capabilities of Belle II. The PXD employs an online data reduction system in order to cope with high background occupancy. Clusters generated by charged particles, which do not leave a signal in the outer sub-detectors, are discarded allowing for a data reduction by a factor of $\sim$10.

Despite being an efficient filter against background, the PXD reduction system also prevents the detection of signal particles, which are not registered by the outer sub-detectors. Among these, Highly Ionizing Particles (HIPs) possess a characteristically severe energy loss limiting their penetration depth into the detector. In particular, magnetic monopoles and anti-deuterons as possible HIP candidates are considered in this thesis. Without a signal in the outer sub-detectors, the particles remain undetected resulting in a loss of information about HIPs.

In this thesis the possibility of identifying HIPs solely with information provided by the PXD is presented. PXD clusters generated by HIPs and background particles are compared and differences are analyzed. The separation between HIP and background clusters is performed with neural network algorithms operating in a multi-dimensional parameter space of PXD pixel information. The neural networks are found to outperform a sequence one dimensional cuts on PXD cluster properties. The influence of the input vector set is analyzed and the performance of the networks is critically reviewed. The prospect of employing neural networks trained on PXD data for offline and online applications is discussed.

ZUSAMMENFASSUNG

Das Belle II am Hochenergie-Forschungszentrum KEK in Tsukuba (Ibaraki) in Japan befindet sich seit 2018 in der Anfangsphase der Datennahme. Im Vergleich zu dem Vorgänger-Experiment Belle plant Belle II die maximale Luminosität um einen Faktor 40 zu erhöhen. Nanobeam-Technologie in der Interaction Region soll die Erhöhung der Luminosität ermöglichen.

Ein Pixeldetektor (PXD) bestehend aus zwei Lagen bestückt mit DEPFET Silizium-Sensoren wurde als innerster Sub-Detektor in Belle II installiert. Der Pixeldetektor dient zum Tracking von geladenen Teilchen in der Nähe der Interaction Region. Die DEPFET Technologie erlaubt es die Sensordicke auf 75 µm zu reduzieren. Die geringe Dicke der Sensoren reduziert Mehrfachstreuung und verhindert die Notwendigkxeit von aktiver Kühlung im sensitiven Bereich des Detektors. Mit Pixelgrößen zwischen 50 µm und 85 µm hat der PXD eine exzellente räumliche Auflösung, die wesentlich für die Tracking-Fähigkeiten des gesamten Detektors ist. Der PXD verwendet eine online Datenreduktion, um die hohen Untergrundraten zu bewältigen. Cluster, generiert von geladenen Teilchen, die kein Signal in den äußeren Sub-Detektoren hinterlassen, werden verworfen, was eine Datenreduktion um den Faktor 10 erlaubt.

Trotz der effizienten Filterung gegenüber Untergrund-Ereignissen, verhindert das PXD Datenreduktionssystem ebenso die Detektion von Signalteilchen, die nicht in den äußeren Sub-Detektoren registriert werden. Zu diesen gehören Hochionisierende Teilchen (Highly Ionizing Particles - HIPs), welche einen charakteristisch hohen Energieverlust besitzen, welcher die Eindringtiefe in den Detektor drastisch limitiert. Insbesondere magnetische Monopole und Anti-Deuteronen werden als HIP-Kandidaten in dieser Thesis behandelt. Ohne ein Signal in den äußeren Sub-Detektoren werden diese Teilchen nicht detektiert und sämtliche Informationen über HIPs werden verworfen.

Die Durchführbarkeit HIPs lediglich mit PXD Daten zu identifizieren wird in dieser Thesis untersucht. PXD Cluster von HIPs und Untergrund werden verglichen und die Unterschiede analysiert. Die Unterscheidung zwischen HIP und Untergrund Clustern wird mit neuronalen Netzen, welche in einem multidimensionalen Parameterraum von PXD Pixelinformationen arbeiten, vollführt. Die Überlegenheit von neuronalen Netzen gegenüber eindimensionalen Cuts auf Cluster-Eigenschaften wird gezeigt. Der Einfluss der Input-Vektoren auf die Performance der neuronalen Netze wird kritisch beleuchtet. Eine Aussicht bezüglich des Einsatzes von neuronalen Netzen, die mit PXD Daten trainiert wurden, für offline und online Anwendungen wird diskutiert.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

The Standard Model (SM) of particle physics has proven to be tremendously successful in describing the fundamental forces of nature. Within the reach of precision and energy currently available at modern experiments, the theoretical predictions of the Standard Model achieve an unprecedented accuracy in describing experimental results. Nevertheless, the SM as a phenomenological theory is considered to be far from complete. Fundamental questions as to the origin of matter/anti-matter imbalance, the description of gravity, the occurrence of dark matter and many more remain unanswered. Various theories suggesting physics beyond the SM have been developed. However, as of today, they suffer from a lack of experimental evidence.

The Belle II experiment, located at the SuperKEKB accelerator, is designed to put the SM to a stringent test. Belle II profits from the world-record design luminosity of the SuperKEKB accelerator, cutting-edge technology used for all detector components as well as state-of-the-art analysis techniques. Unlike the predecessor experiment Belle, a Pixel Detector (PXD) is installed as the innermost sub-detector in Belle II. The PXD profits from a short radial distance from the interaction region, a low material budget and a high spatial resolution. Its primary task is the improvement of tracking in close proximity to the interaction point.

In this thesis the prospect of exploiting the PXD for identification of Highly Ionizing Particles (HIPs) is investigated. HIPs such as anti-deuterons or magnetic monopoles are characterized by a particularly severe energy loss limiting their range in the Belle II detector to the inner vertex region. Both particle species play a fundamental role in the indirect (anti-deuterons) or direct (magnetic monopoles) detection of physics beyond the Standard Model. The confinement of HIPs to the inner region of the Belle II detector makes detection impossible. Consequently, information about potentially new physics scenarios are lost. The PXD as the innermost sub-detector can help to identify HIPs despite their short range in the Belle II detector preventing the particles from remaining undetected. Neural networks operating in a mutlidimensional parameter space of PXD data are considered to identify HIPs. In particular, the application of Feed-Forward Neural Networks and Self-Organizing Maps are presented in this thesis.

This thesis is structured as follows: A quick overview of the current status in modern particle physics is given in Chapter 2. The Standard Model of particle physics is introduced, its shortcomings and the prospect of physics beyond the Standard Model are discussed.

In Chapter 3 the Belle II experiment is presented. First, the physics motivation for the experiment is outlined. Subsequently, the SuperKEKB accelerator is introduced. Its most prominent feature, namely the world-leading design luminosity, is reviewed with regard to the techniques, which help achieve the high luminosity. The various components of the Belle II detector are presented as well. While the technical set-up and the purpose of all sub-detectors are only briefly discussed, the PXD is considered in more detail due to its relevance for this thesis. The technology of the silicon sensors, the geometry of the entire PXD as well as the read-out and data acquisition are presented. Additionally, beam-induced and luminosity-dependent background sources expected at Belle II are described and consequences for the PXD are pointed out. The Chapter closes with an overview over the Belle II software framework `basf2`.

The analysis strategy is disclosed in Chapter 4. The implementation of Feed-Forward Networks and Self-Organizing Maps and the composition of the input vectors are outlined. For both neural networks a generic network architecture is presented, which only needs small adjustments to suit the requirements of the different training sets.

The identification of anti-deuterons against background is presented in Chapter 5. One and two dimensional cluster property distributions from anti-deuteron clusters and background clusters are compared and the distribution of pixel charges in a cluster is examined. Subsequently, the training results for Feed-Forward Network and Self-Organizing Map are discussed. The training process itself is analyzed as well.

Chapter 6 is dedicated to magnetic monopoles. A brief theoretical overview focusing on the monopole's energy loss characteristics is given. After the impact of the monopole mass on cluster properties and pixel charge distributions is studied, the training results for the neural networks are examined.

Finally, in Chapter 7 a summary and outlook is given. The results of this thesis are reviewed and discussed in a broader context with special regard to future applications.

CHAPTER 2

PHYSICS BACKGROUND

2.1 Standard Model of Particle Physics

The Standard Model (SM) of particle physics is a theoretical framework describing three of the four fundamental interactions: the electromagnetic, weak and strong interaction. The model has proven to predict experimental measurements with an astonishing accuracy in an energy range from a few eV to several TeV.

The SM is based on the idea of fundamental particles. These particles can be divided into two groups: *fermions* and *bosons*. While observable matter consists of fermions, interactions are mediated by bosonic particles. Interactions are characterized by the transfer of energy and momentum. Each interaction is associated with a coupling constant determining its strength.

Mathematically, the SM is a relativistic Quantum Field Theory (QFT). Equations of motions are derived from a Lagrangian density $\mathcal{L}$ by the Euler-Lagrange formalism. Particles are represented by field operators - fermions by spinor fields and bosons by vector fields. The SM is build on the principle of local gauge invariance i.e. the action, as the time integral of the Lagrangian, is not altered by the application of a local gauge transformation.

In the following Section the most important aspects of the SM for this work are introduced. A more detailed description can be found in [9]. Since the SM is not a complete theory in the sense that it fails to describe several phenomena and observations such as gravity, non-zero neutrino masses or dark matter/dark energy, the end of this Section is devoted to extensions of the SM.

2.1.1 Particles and Interactions

As shown in Table 2.1 the fermions in the SM can be grouped together according to the charges which they carry. The magnitude of the charges assigned to a particle determine the probability to emit or absorb a mediating boson (in this context also referred to as *gauge boson*).

The strong interaction is mediated by eight different bosons called *gluons*, which couple to the color charge C of a particle. The color charge has the three eigenstates red, green and blue ($r/g/b$). The strong interaction is described by Quantum Chromodynamics (QCD), which is based on the $\mathrm{SU}(3)_C$ symmetry group. The subscript C denotes the color charge. The gauge bosons of the weak interaction are the Z^0, W^+ and W^-. The weak interaction inherits the symmetry group $\mathrm{SU}(2)_L$. The subscript L refers to the weak isospin carried by left-handed particles. Solely particles with weak isospin are able to participate in the weak interaction. The electromagnetic interaction is described by Quantum Electrodynamics (QED), which is based on the $\mathrm{U}(1)_Y$ symmetry group. The subscript Y stands for the weak hypercharge. The mediating boson - the photon γ - couples to the electric charge Q of particles. Altogether the symmetry group of the SM Lagrangian can be summarized as: $\mathrm{SU}(3)_\mathrm{C} \otimes \mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$ [10].

Fermions with color charge are referred to as *quarks*. *Leptons* on the other hand do not possess color charge and are therefore unable to participate in the strong interaction (see Table 2.1). Quarks and leptons are further subdivided into three generations. In every generation *up-type* quarks are characterized by a fractional electric charge of $+2/3\,e$ and *down-type* quarks by $-1/3\,e$. The mass of the quarks increases from generation to generation. Each lepton generation contains an electrically charged particle and a neural particle referred to as neutrino.

The twelve fermions of the SM each have a partner, for which the sign of the charges is reversed. All other physical properties of these so-called *antiparticles*

Table 2.1: Fermions of the Standard Model grouped according to the charges they carry. The color charge is abbreviated as C, the electric charge as Q. The third component of the weak isospin for left-handed fermions is $T_{3,L}$ and $Y_{L/R}$ is the weak hypercharge of left-handed/right-handed particles [11].

	Generation			Charge				
	1st	2nd	3rd	C	$Q[e]$	$T_{3,L}$	Y_L	Y_R
Quarks	u	c	t	$r/g/b$	$+2/3$	$-1/2$	$+1/3$	$+4/3$
	d	s	b	$r/g/b$	$-1/3$	$+1/2$	$+1/3$	$-2/3$
Leptons	e^-	μ^-	τ^-		-1	$-1/2$	-1	-2
	ν_e	ν_μ	ν_τ			$+1/2$	-1	

remain unaltered.

In the SM an additional scalar (spin = 0) particle H^0 appears as a result of the Higgs mechanism. The masses of the elementary particles manifest themselves as a result of their coupling to the Higgs field.

2.1.2 Perturbation Theory

In general an analytical derivation of the equations of motions from the SM Lagrangian are not feasible. A solution is nonetheless obtainable by using perturbation theory, which relies on the expansion of quantities in orders of the coupling constant.

The probability amplitude $A_{i\to f}$ for an initial state $|i\rangle$ to transition into the final state $|f\rangle$ is determined by the scattering matrix S:

$$A_{i\to f} \propto \langle i| S |f\rangle . \tag{2.1.1}$$

The square of the probability amplitude yields the transition probability $\Gamma_{i\to f}$. In perturbation theory the scattering matrix is expressed as [12]:

$$S = \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int_{-\infty}^{\infty} dx_1 \int_{-\infty}^{\infty} dx_2 ... \int_{-\infty}^{\infty} dx_n T(L_I(x_1)L_I(x_2)...L_I(x_n)) \tag{2.1.2}$$

where i is the imaginary unit, T is the time ordering operator and L_I the Lagrangian density in the interaction picture. The n-th term in the series is proportional to $\sqrt{\alpha}^n$, where α is the coupling constant.

When the probability amplitude A is evaluated, each term in the series can be represented by so-called Feynman graphs. Two generic Feynman diagrams are shown in Fig. 2.1. The first graph corresponds to a so-called *tree-level* diagram, it is also referred to as Leading Order (LO), since it is the first non-trivial term in the expansion. The diagram on the right-hand side represents the term following the LO and is therefore called Next to Leading Order (NLO). In this example the diagram contains a virtual loop. When the series is truncated after the first non-trivial term, *Fermi's golden rule* [13] is recovered. It provides an easy albeit crude estimation of the transition probability.

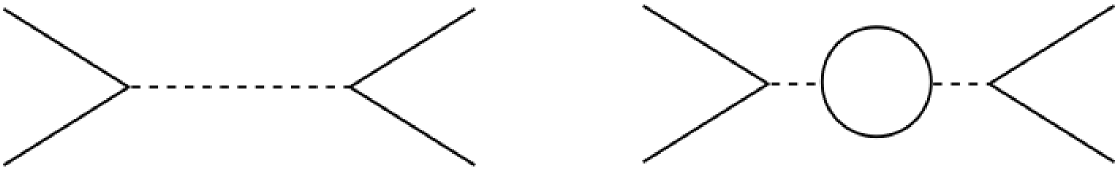


Figure 2.1: Generic Feynman diagrams - the diagram on the left-hand side is called *tree-level* diagram and represents a contribution of the Leading Order (LO) in the expansion. The diagram on the right-hand side contains a virtual loop. Since it arises from the next term in the expansion, it is referred to as Next to Leading Order (NLO).

2.1.3 Physics Beyond the Standard Model

Both experimental and theoretical evidence suggests that the SM is an incomplete theory. It does not provide a description of gravitation and does not incorporate the occurrence of dark matter as well as dark energy. The existence of dark matter is implied by the mass distribution of galaxy clusters in the universe, gravitational lensing and the rotation velocity of spiral galaxies [14]. Additional evidence is provided by anisotropies in the cosmic microwave background [15]. Furthermore, dark energy has been suggested as an explanation for the accelerating expansion of the universe, which can studied from distant type Ia supernovae [16].

Even for *conventional* matter there are unsolved questions such as the matter-antimatter asymmetry in the universe [17]. Assuming that matter and antimatter were produced in equal amounts during the Big Bang, baryon number violating processes are required to yield a matter-antimatter asymmetry [18].

Despite these fundamental issues, there is additional shallow evidence not contradicting the predictions of the SM but calling its inner consistency into question. For instance there are free parameters in the SM, which require experimental determination. From a complete theory it is expected, that no unknown parameters are present. In addition, the low mass of the Higgs boson is understood to arise from an *unnatural* cancellation of parameters in the SM (*fine tuning*) [8]. This is known as the *hierarchy problem*.

Recent decades saw the rise of novel theoretical frameworks such as Super Symmetry (SUSY) [19], String Theory [20] or the Grand Unified Theory (GUT) [21], which try to remedy the shortcomings of the SM. For each of these a multitude of sub-theories were developed over the years. Experimental evidence is required to validate or falsify the multitude of theoretical frameworks. Particle collider experiments such as Belle II offer a unique environment to probe the SM with unprecedented precision.

CHAPTER 3

THE BELLE II EXPERIMENT

The SuperKEKB accelerator is an asymmetrical electron positron collider located at the KEK (High Energy Accelerator Research Organisation) in Tsukuba (Japan). As an upgrade to the KEKB accelerator, SuperKEKB is designed to reach an instantaneous luminosity of $8 \cdot 10^{35} \mathrm{cm}^{-2}\mathrm{s}^{-1}$ and a integrated luminosity of $50\,\mathrm{ab}^{-1}$ by 2025 [22].

The increase in luminosity requires an update of the detector positioned at the interaction point of electron-positron collisions as well. The upgraded detector, referred to as Belle II, features new sub-detectors, refined tracking capabilities due to more fine-grained detector parts and many more improvements compared to its predecessor the so-called Belle experiment. Additionally, novel analysis strategies and reconstruction algorithms are developed to extent and complement the physics program of Belle. In the following Chapter the physics motivation for Belle II is outlined. Subsequently, technical details about the SuperKEKB accelerator and Belle II are given. In particular, the design of the Pixel Vertex Detector is reviewed. The Chapter closes with an introduction to the software analysis framework `basf2` employed by the Belle II collaboration.

3.1 Physics Motivation

Belle II is designed to work at the intensity frontier of particle colliders with a focus on precise measurements of rare processes, which is complementary to the high energy approach established at LHC or the Tevatron. Correspondingly, Belle II is highly sensitive to new physics scenarios manifesting themselves through small deviations from SM predictions. In the following a small excerpt of the physics studies planned for Belle II is given. The list does not claim to be complete, but nevertheless provides insight into the physics prospects of a second generation B-factory. Most notably, searches for low multiplicity final states are introduced at the end of this Section.

The search for beyond the SM particles in B decays is a major project at Belle II. A charged Higgs [23], if existing, could modify branching ratios in the decays $B \to \tau\nu$ and $B \to D^{*}\tau\nu$ leading to deviations from SM predictions. The high luminosity at Belle II allows to determine the branching ratios with unprecedented sensitivity.

The measurement of CP violation was a strong motivation for the upgrade from Belle to Belle II. Anomalies in CP violations are a clear signal for beyond the SM physics. Radiative and electroweak penguin diagrams are particularly sensitive to new physics due to loop contributions of beyond the SM particles. Additionally, the measurement of CP violation in the transition $b \to s$ and the comparison with measurements of $b \to c$ transitions will be used to probe new physics scenarios.

Another focus for Belle II analyses focusing on beyond the SM physics lies in τ meson decays. Most notably, Lepton Flavor Violation, occurring, for instance, in the decays $\tau \to \mu\gamma$ or $\tau \to \mu\mu\mu$ is heavily suppressed in the SM. Some new physics scenarios, however, predict an enhancement of Lepton Flavor Violation by several orders of magnitude making Belle II sensitive to these decay modes.

Charmonium spectroscopy at Belle II becomes possible through the decay of B mesons into $c\bar{c}$-pairs. In contrast to dedicated charm factories like the BES III experiment [24], which relies on the resonant production of charmonium, the production via B meson decays allows to access a multitude of different quantum numbers. Charmonium-like states, which are not predicted by theoretical models, are referred to as *exotic states* [25]. There are still speculations about the nature (tetraquarks, meson molecules,...) of these states. Exotic states play a significant role in the physics program of Belle II especially since measurements at Belle II are complementary to analyses presently performed at BES III and LHCb.

3.1.1 Dark Sector Physics at Belle II

The envisaged luminosity of SuperKEKB allows Belle II to improve current limits in various Dark Sector searches. Dedicated trigger settings for Dark Matter searches are employed increasing the sensitivity for specific new physics final states. A set of Dark Sector studies are briefly introduced to provide an overview of the activities in Belle II.

Dark Photons

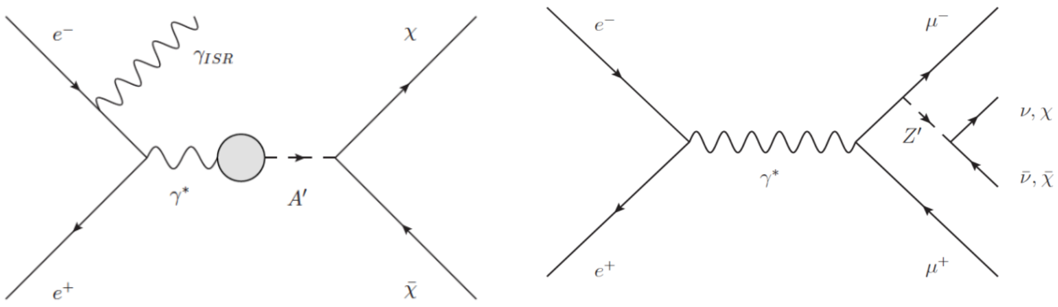
Dark photons are dark matter candidates with the ability to couple to the SM photon γ through kinetic mixing [26] as visualized in Fig. 3.1a. The production of a virtual dark photon from electron-positron annihilation is depicted. The electron emits an initial state photon γ_{ISR}

The dark photon can decay into SM fermions $e^+e^- \to \gamma_{\text{ISR}} f\bar{f}$ or into dark matter particles $e^+e^- \to \gamma_{\text{ISR}}\chi\tilde{\chi}$ provided that the dark matter decay products χ and $\tilde{\chi}$ are lighter than A'. At Belle II initial state radiation (ISR) helps to select dark photon candidates. The detection of the initial state photon is crucial, in particular for the decay of A' into dark matter particles. A single-photon trigger based on individual clusters in the electromagnetic calorimeter has been devised for this purpose [27]. Belle II is expected to lower upper limits for A' processes in the mass range of tenth of MeV to a few GeV.

Axion-like Particles

Axion-like particles (ALPs) are characterized by their coupling to SM gauge bosons. Various extensions of the SM propose the existence of ALPs. Most notably various solutions of the strong CP problem feature ALPs [28].

The coupling of ALPs to the SM photon is exploited for experimental searches at Belle II [2]. The most important production processes are photon fusion $e^+e^- \to e^+e^-a$ and ALP-strahlung $e^+e^- \to \gamma a$, where a is the axion-like particle. Feynman diagrams for the two processes can be seen in Fig. 3.2. Production via photon fusion has a higher cross section compared to the ALP-strahlung process unless the center-of-mass energy of the electron-positron collision is close to the ALP mass. Experimental identification of ALPs created via photon fusion, however, is more difficult due to the low momentum transferred to ALPs in the fusion process. The decay products of low-momentum ALPs inherit the low momentum, which complicates detection due to high QED background.



(a)

(b)

Figure 3.1: Production mechanism of a dark photon A' (left-hand side) and massive vector boson Z' (right-hand side) from [1]. (a) Coupling to the SM photon γ enables production of A' in e^+e^- collisions. The dark photon can decay into the dark matter particles χ and $\tilde{\chi}$ provided that A' is not the lightest dark matter particle. (b) In this example the vector boson Z' is emitted from a muon and decays either in neutrinos ν or dark matter particles. The two muons combined with the missing energy help in the identification.

The ALP-strahlung channel is therefore the primary focus of Belle II analyses. Depending on the mass and the lifetime of the ALP different event signatures are expected. A long-lived, low mass ALP can decay outside of the detector region leaving only the photon, which emitted the ALP, for detection. An ALP decaying into $\gamma\gamma$ inside the sensitive detector region, on the other hand, leads to a three photon final state.

Massive Vector Boson Z'

Some extensions of the SM predict the existence of a new massive vector boson referred to as Z' coupling to the $L_\mu - L_\tau$ current [29] . In Belle II the decay of the Z' into an invisible final state is considered. As shown in Fig. 3.1b, the Z' production in combination with a muon pair is the most promising search channel due to the clean muon singal. A peak in the missing recoil mass reconstructed from the two muons serves as an unmistakable signature for the Z'.

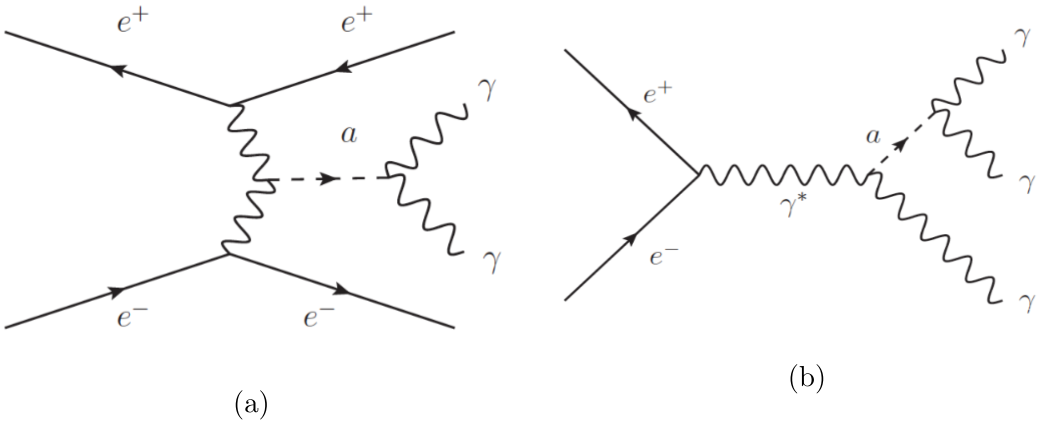


Figure 3.2: Feynman diagrams for photon fusion (left-hand side) and ALP-strahlung (right-hand side) from [2] . In the fusion process the ALP is generated by two virtual photons. ALP-strahlung involves the emission of an ALP from a virtual photon.

3.2 Status of SuperKEKB and Belle II

The commissioning and operation of the SuperKEKB accelerator and Belle II detector is divided into three phases. Phase I was held in early 2016. It was dedicated to accelerator commissioning without any collisions taking place. Beam background generated by single beams was studied. In Phase II most sub-detectors of Belle II were assembled. However, only parts of the PXD and the SVD were installed. The first collision of electrons and positrons was achieved on April 28th 2018. During Phase II data equivalent to about $L_{\text{int}} = 500\,\text{pb}^{-1}$ of integrated luminosity was gathered with which various resonances like the J/ψ or D mesons were rediscovered. The maximum peak luminosity during Phase II reached $5.5 \cdot 10^{33}\text{cm}^{-1}\text{s}^{-1}$ [30]. Phase III started in March 2019.

3.3 The SuperKEKB e^+e^- - collider

The SuperKEKB is an asymmetrical electron-positron collider with a center-of-mass energy of $10.58\,\text{GeV}$ and a target luminosity of $8 \times 10^{35}\,\text{cm}^{-2}\text{s}^{-1}$. As the successor of the KEKB collider [31], which was shut down in 2010, it is located in Tsukuba (Ibaraki) in Japan. Compared to KEKB, the target luminosity of

SuperKEKB is 40 times higher, which entails various technical requirements for all detector components and acceleration units.

In Fig. 3.3 a schematic overview of the SuperKEKB accelerator is depicted. The accelerator hosts a double-ring structure with the High Energy Ring (HER) storing electrons and the Low Energy Ring (LER) for positron storage. Before injection into the ring, the electrons are accelerated to an energy of 7 GeV and the positrons to 4 GeV in a linear accelerator. The positron beam is guided through the Damping Ring in order to reduce the emittance of the beam with radiation damping [32]. The planned beam current is 2.6 A for electrons and 3.6 A for positrons.

The two main storage rings are divided into four arc sections (D3, D6, D9 and D12), which host normal conducting dipole bending magnets in order to guide the particles around the ring, and four straight sections (Tsukuba, Oho, Fuji and Nikko). Both straight and bent sections are equipped with focusing magnets and collimators to maintain a small diameter of the beam. Superconducting and normal conducting RF cavities are installed along the straight sections to mitigate the energy loss due to synchrotron radiation.

The particle beams are brought to a collision at the Interaction Point (IP), which is surrounded by the Belle II detector. The asymmetric energies of electrons and positrons provide a boost in the laboratory system [33] in order to facilitate vertex determination of short-lived particles. The beams are tilted with respect to each other by 83 mrad in order to to avoid a head-on collision, which would introduce various difficulties and disadvantages due to beam optical effects. The center-of-mass (CMS) energy of the collision $E_{\mathrm{CMS}} = 10.58\,\mathrm{GeV}$ is chosen to coincide with the mass of the $\Upsilon(4\mathrm{S})$ resonance, which decays nearly exclusively to two B mesons [9]. The CMS energy can still be tuned in a small energy regime ranging from the $\Upsilon(1S)$ mass up to the $\Upsilon(6S)$ mass.

Particles in the beams are grouped into bunches. Each ring can store 2506 bunches with roughly 10^{11} particles per bunch. Newly accelerated bunches are continuously injected from the linear accelerator into the main storage rings. When new particles enter the storage ring, it takes a few milliseconds until their emittance is sufficiently reduced. During this time an increase in background is generated by these *noisy bunches*. Consequences for the detector are described in Section 3.8.

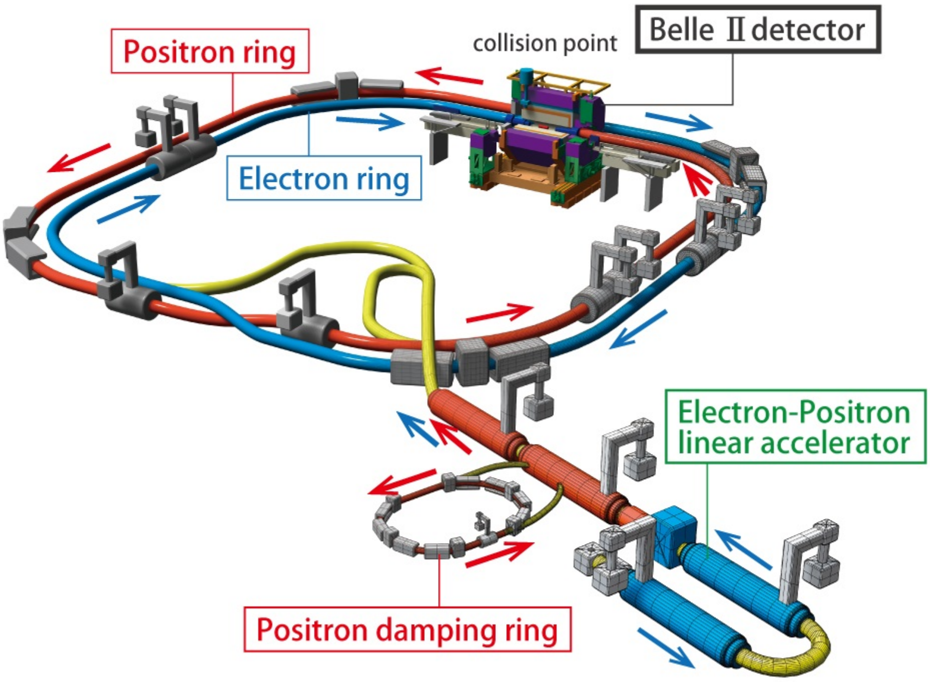


Figure 3.3: Schematic drawing of the SuperKEKB accelerator from [3]. Electrons and positrons are accelerated in the linear accelerator at the bottom of the image before entering the HER (blue) and the LER (red). The emittance of the positron beam is reduced in the damping ring before injection. The Belle II detector is located at one of the straight sections.

3.3.1 Luminosity of SuperKEKB

The instantaneous luminosity L is the proportionality constant between the number of events per unit of time $\mathrm{d}N/\mathrm{d}t$ and the cross section σ:

$$\frac{\mathrm{d}N}{\mathrm{d}t} = L \cdot \sigma. \tag{3.3.1}$$

The luminosity of two colliding equal sized beams can be expressed by [34]:

$$L \propto \frac{\gamma_{\pm}}{2er_e}\left(\frac{I_{\pm}\xi_{y\pm}}{\beta^*_{y\pm}}\right) \tag{3.3.2}$$

as long as the beam size in one direction is much larger than the size in the direction perpendicular to that. The index (+) stands for positrons and (-) for electrons, e is the electric charge and r_e the classical electron radius. The luminosity is consequently determined by the Lorentz factor $\gamma_\pm$, the vertical beam-beam parameter $\xi_{y\pm}$ describing the force on a particle exerted by the beam particles belonging to the opposing bunch, the beam current $I_\pm$ and the vertical beta function $\beta^*_{y\pm}$, which is correlated with the vertical beam size $\sigma^*_{y\pm}$.

To reach the design luminosity of $L = 8 \cdot 10^{35}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ a *nano-beam scheme* is employed. The vertical beta function $\beta^*_{y\pm}$ is reduced at the IP with a superconducting final focusing magnet. The beam size in this region is only 48 nm justifying the term *nano-beam*. The vertical beta function is 20x lower compared to KEKB and the beam current is two times higher, in total, leading to a luminosity increase of 40 according to Eq. 3.3.2.

3.4 The Beam Pipe at the Interaction Region

The Interaction Region (IR) is defined as the area around the IP with an extension of roughly 4 m along the storage rings. The IR is equipped with quadrupole magnets for the final focusing of the beams. The beam pipe around the IR consists of two Beryllium layers with Paraffin in-between them acting as a coolant. The high stiffness, low atomic number, good thermal stability as well as its diamagnetic nature validate the choice of Beryllium as beam pipe material. The low atomic number is advantageous to suppress multiple scattering of particles coming from the IP and the diamagnetic properties ensure that no interference with the magnetic field of the focusing magnets can occur. An additional gold coating at the inside of the beam pipe is used to suppress synchrotron radiation.

In Table 3.4 a summary of the material budget of the beam pipe surrounding the IP is presented. The gold coating will increase in thickness from 6.6 µm in Phase II to 10 µm in Phase III.

Table 3.1: Material budget of the beam pipe in the Belle II experiment [8]. A thin gold coating increasing in thickness from 6.6 µm in Phase II to 10 µm in Phase III serves as the innermost layer of the beam pipe. Two Beryllium (Be) layers, with a Paraffin layer acting as coolant between them, account for the majority of beam pipe material.

Gold plate	Thickness	10 µm/6.6 µm
Inner Be pipe	Inner radius	10 mm
	Thickness	0.6 mm
Paraffin coolant	Inner radius	1 mm
Outer Be pipe	Inner radius	12 mm
	Thickness	0.4 mm

3.5 The Coordinate System

The Belle II detector is conventionally described by a right-handed Cartesian coordinate system with the origin located at the nominal IP. The z axis is parallel to the beam pipe. The positive direction of the z axis points to the forward region of the detector. The positive part of the y axis points to the top of the detector and the orientation of the x axis is parallel to the radial direction towards the outside of the detector ring [6]

3.6 The Belle II Detector

The general-purpose detector Belle II surrounds the IP of the SuperKEKb accelerator. It has an asymmetric design to account for the forward boost in the direction of the electron beam. The angular acceptance of Belle II amounts to $17^\circ < \theta < 150^\circ$ for the polar angle and 2π for the azimuthal angle.

With regard to the polar angle the detector is divided into three regions:

- the forward region ($17^\circ < \theta < 30^\circ$)
- the barrel region ($30^\circ < \theta < 125^\circ$)
- the backward region ($125^\circ < \theta < 150^\circ$).

The regions differ in terms of detector set-up and material budget. A schematic drawing of Belle II is presented in Fig. 3.4. An overview over the sub-components of the detector is given in the following Section. The PXD will be reviewed in more depth in Section 3.8 due to its significance for this work.

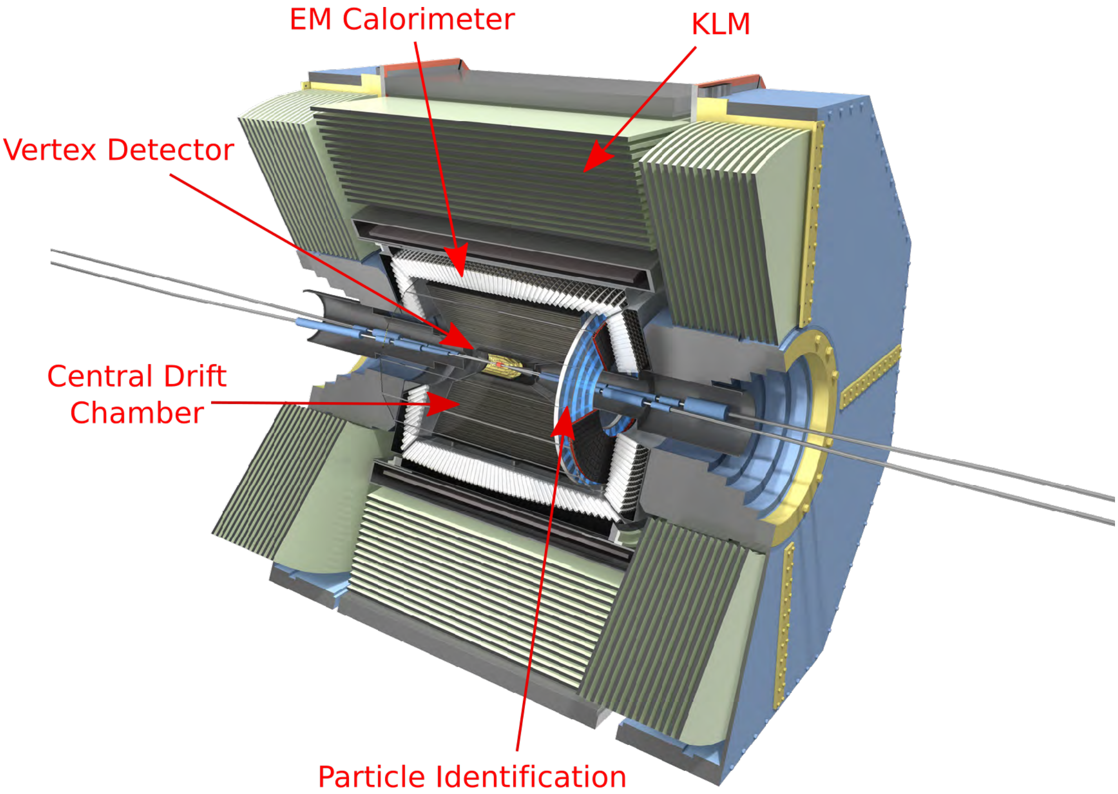


Figure 3.4: Schematic drawing of the Belle II detector from [4]. The Vertex Detector consists of the PXD and SVD intended to provide accurate tracking close to the interaction region. The Central Drift Chamber (CDC) provides tracking information as well and measures the energy loss of particles. Particle Identification is performed with the TOP and the ARICH by exploiting the Cherenkov effect. The EM Calorimeter (ECL) is designed to measure the total energy and the angular position of particles. The outermost sub-detector - the KLM - is used for muon and K_L identification.

Belle II consists of the following sub-detectors:

- **Pixel Vertex Detector (PXD)**

  The PXD is the innermost component of the Belle II detector. Its primary task is the determination of decay vertices of short-lived particles (e.g. B-mesons) from their long-lived decay products. This is accomplished by reconstructing particle tracks with the tracking detectors of Belle II and extrapolating them to the interaction region.

  The PXD is mounted directly onto the beam pipe around the IP and consists of two layers of pixelated DEPFET sensors [35]. A matrix of 767 x 249 DEPFET sensors combined with read-out ASICs form a module. The PXD consists of 40 such modules arranged in two layers positioned at radii of 14 mm and 22 mm from the IP, respectively.

- **Silicon Vertex Detector (SVD)**

  Similar to the PXD the SVD is designed to provide track information of charged particles. It allows for a track-reconstruction of low-p_t-tracks down to about 30 MeV [6]. When combined, the PXD and SVD are referred to as Vertex Detector (VXD).

  The SVD consists of Double-sided Silicon Strip Detectors (DSSDs) arranged in four layers. The DSSDs are composed of a silicon bulk sandwiched between strips of p-type and n-type doped material. While the innermost layer is entirely covered with small rectangular sensors, the outer layers host large rectangular (barrel sensors) and trapezoidal-shaped (slanted sensors) sensors for improving acceptance and precision of forward boosted particles [36, 37].

  The four SVD layers are positioned at radii between 38 mm to 140 mm from the IP. The average material budget per layer is approximately twice as much as the material budget for the PXD layers.

- **Central Drift Chamber (CDC)**

  The CDC surrounds the VXD in the barrel region of the detector. The CDC is designed to determine the momenta of charged particles from the bending radius of the particles' trajectories in the magnetic solenoid field of Belle II. Additionally, the CDC provides information about the energy loss of particles allowing for particle identification [38]. The CDC is also a important for the Belle II hardware trigger system, which is discussed in more detail in Section 3.7.

  The CDC extends over a radial region between 160 cm to 1130 cm [8]. It consists of 14,336 gold plated tungsten wires (*sense wires*) with a thickness of 30 µm. Each of the sense wires is surrounded by 8 aluminum *field wires*. The sense wires are arranged in 9 superlayers, each of which hosts 6 layers with the same wire orientation. The innermost superlayer has two additional layers to cope with the high occupancy due to background particles. The wire orientation between the superlayers alternates between parallel to the beam axis (*axial layers*) and skewed by an angle of -74 mrad to 70 mrad (*stereo layers*). The entire CDC is filled with a drift gas mixture of 50% helium and 50% C_2H_6. The low atomic number of the gas mixture serves to suppress multiple scattering. The radiation length of the mixture is approximately 640 m [39].

A charged particle traversing the CDC ionizes the drift gas. The electric field between the wires induces a drift of the liberated electrons to the sense wires, in which an electrical current is generated. Spatial information about the incident position in the xy plane is obtained by timing measurements combined with the known drift velocity of electrons. The tilted stereo layers provide additional information about the position in z. The electric pulse induced at the sense wires is proportional to the deposited energy of the incident particle, which allows for energy loss measurements.

- **Time of Propagation Counter (TOP)**

  The TOP is located in the barrel region of the detector and provides a criterion to separate $K^{\pm}$ mesons from $\pi^{\pm}$ mesons by exploiting the Cherenkov principle. A charged particle, whose velocity exceeds the speed of light in the radiator material of the TOP, emits Cherenkov light with the opening cone θ_C [40]:

  $$\cos\theta_C = \frac{1}{\beta n(\gamma)} \tag{3.6.1}$$

  where $n(\gamma)$ is the refractive index of the radiator material for a given wavelength λ and β is the particle velocity over the speed of light.

  The TOP consists of 16 rectangular 2.5 m long and 2 cm thick quartz radiators. Cherenkov phtotons, generated by the passage of a charged particle, are internally reflected at the radiator walls confining their propagation to the interior of the quartz tiles until they interact with the photo sensors at the end of each quartz module. The photo sensors are arranged as micro channel plate photomultiplier tubes (MCP-PMTs).

  For the separation of $\pi^{\pm}$ and $\mathrm{K}^{\pm}$ mesons the mass difference between the two particle species is exploited. For a given momentum $\mathrm{K}^{\pm}$ mesons emit Cherenkov light in a wider cone according to Eq. 3.6.1. The number of reflections in the quartz is lower, since the wider cone implies a smaller angle between photon track and quartz interface. The time between creation of Cherenkov photons until detection at the end of the quartz modules (*time of propagation*) is consequently smaller for $\mathrm{K}^{\pm}$ mesons compared to the lighter $\pi^{\pm}$ mesons. The measurement relies on information about the incident position and incident angle of the charged particle, which is provided by the tracking detectors.

- **Aerogel Rich Detector (ARICH)**

  The ARICH is a proximity-focusing Aerogel Ring-Imaging Cherenkov detector [41] located in the forward endcap region of the Belle II detector. Similar to the TOP, the ARICH uses the Cherenkov principle to separate $\pi^\pm$ mesons from $K^\pm$ mesons. In addition, the sub-detector provides a discrimination criterion for pions, muons and electrons with momenta below 1 GeV.

The ARICH covers a polar region between $14.78^\circ < \theta < 33.71^\circ$. It consists of an aerogel [42] serving as Cherenkov radiator.

Charged particles with velocities exceeding the speed of light in the aerogel emit Cherenkov light as described in the previous paragraph. The Cherenkov cone is projected onto a photon detector plane leaving a ring image. The radius of the ring depends on the opening angle of the Cherenkov cone. According to Eq. 3.6.1 the opening angle is determined by the incident particle's velocity, from which the particle mass can be computed for a given momentum.

- **Electromagnetic Calorimeter (ECL)**

  The ECL is designed to differentiate between photons and neutrons. Additionally, the angular coordinates and energy of particles (mainly photons) are determined with the ECL.

  The ECL covers a range of $12.4^\circ < \theta < 155.1^\circ$ and consists of 8736 thallium-doped caesium iodine CsI(Tl) crystals with a size of $6\,\mathrm{cm} \times 6\,\mathrm{cm} \times 30\,\mathrm{cm}$, which is equivalent to a radiation length of $16.1\,X_0$ [8,43]. A pair of $10\,\mathrm{mm} \times 20\,\mathrm{mm}$ Hamamatsu S2744-08 photodiodes are attached to each crystal.

  Particles penetrating the crystals of the ECL generate photons either through ionization and excitation or electromagnetic/hadronic showers. Showers are associated with a neutral particle, if they are not matched to a track provided by the tracking detectors. The lateral extent of showers supplies an additional identification criterion, since showers generated by hadrons (hadronic showers) are broader compared to electromagnetic showers generated by photons or electrons.

- **Superconducting Solenoid**

  All of the sub-detectors except for the KLM are surrounded by a superconducting magnet producing a constant solenoid field of 1.5 T. The field lines are parallel to the main detector axis. Electrically charged particles are forced onto helical trajectories due to the Lorentz force generated by the magnetic field. The curvature of the trajectories is used to reconstruct the momentum of particles [44].

- **K_L and muon detector (KLM)**

  The KLM is used to identify K_L mesons and $\mu^\pm$ with a momentum above ~ 0.6 GeV.

The KLM consists of alternating layers of 47 mm thick iron plates and active detector elements. It covers the polar region $20^\circ < \theta < 155^\circ$. The iron plates, serving as return yoke for the magnetic field from the solenoid, are exploited as a shower medium for K_L mesons. The showers are detected in the active detector layers of the KLM. Muons, on the other hand, are solely slowed down by the iron plates.

K_L mesons are identified by requiring that no track reconstructed from the inner tracking detectors points to the KLM cluster under investigation. An identification efficiency of up to 80% at 3 GeV is achieved with this method. The muon identification is performed with a likelihood ratio for the hypotheses of the particle being a muon rather than a charged hadron. The difference between the predicted and the measured last layer, in which the particle was detected as well as the goodness of a Kalman fit under the muon hypothesis are considered in the likelihood ratio. An identification efficiency of 89% above 1 GeV is achievable with the KLM [45].

3.7 Trigger System and Data Acquisition

The Belle II trigger initiates the read-out for the entire detector. The first trigger decision is made on hardware level (*Level 1 trigger*) by taking information provided by several Belle II sub-detectors into account. The Level 1 trigger follows a hierarchical structure. Information from all sub-trigger systems, each associated with a specific sub-detector, are gathered in the Global Decision Logic (GDL) [46], where the global trigger decision is made within 4.5 µs. At the target luminosity the Level 1 trigger is expected to run at an average rate of 30 kHz.

After a trigger signal is issued the data of all sub-detectors except for the PXD are read out through a common data link - the *Belle2Link* [47] - and sent to the *Event Builder*, where all data belonging to one collision are combined into an *event*. A computing farm equipped with `basf2` reconstructs each event. Upon reconstruction the High Level Trigger (HLT) decides to save or discard the event. The PXD data is added to the event in another Event Builder if the HLT decides to keep the event.

3.7.1 Random Triggers

Belle II employs two types of *random* triggers. Their main purpose is the detection of background data. The two trigger channels are referred to as trigger bit 25 and trigger bit 26. Trigger bit 25 (*Bhabha trigger*) is issued 50 µs after a Bhabha event is registered. Its reliance on Bhabha scattering makes trigger bit 25 luminosity-dependent. Trigger bit 26 (*pseudo-random trigger*) is based on a random seed generated on basis of a local independent clock. In contrast to trigger bit 25 the trigger rate is independent of the luminosity.

3.8 Pixel Vertex Detector (PXD)

As the innermost component, the PXD is the first sensitive part of the detector to interact with charged particles produced in the e^+e^--collisions. The PXD has to withstand a high exposure to background events due to its close proximity to the IR. As a consequence, the PXD must also handle a data rate, which is drastically higher compared to all other sub-detectors. For a better understanding of the PXD, the following Section outlines the working principle of the DEPFET pixel sensor as well as the geometrical set-up of the PXD.

3.8.1 DEPFET Pixel Sensor

A Depleted Field Effect Transistor (DEPFET) pixel sensor is a monolithic semiconductor sensor [37,48]. The monolithic design is characterized by the integration of the amplification circuit into the sensitive material of the device allowing to thin the pixel thickness down to 75 µm. Each PXD layer contributes only 0.21% of radiation length to the total amount of material in the Belle II detector due to the small sensor thickness. The low material budget suppresses multiple scattering, which is crucial for the measurement of low-momentum or highly ionizing particles. An additional advantage of the thin sensors is the low power consumption. No active cooling in the acceptance region of the detector is required.

A DEPFET is based on a p-channel Metal-Oxide-Semiconductor Field-Effect Transistor (MOSFET) [49] attached to a high-resistivity silicon detector substrate referred to as *bulk*. A schematic depiction of the sensor is shown in Fig. 3.5. The MOSFET is formed by two p-doped regions, serving as source and drain, in the n-doped bulk. The current flowing between source and drain is modulated by the FET gate contact.

The entire bulk can be depleted of movable charge carriers by applying a sufficiently high negative voltage to the p+ contact at the back side of the sensor. The passage of charged particles through the depleted bulk generates electron-hole pairs. While the holes drift to the negative back side of the sensor, the electrons are collected by an n-doped implant called *internal gate*, which is positioned approximately 1 µm below the transistor channel. The electrons induce image charges in the transistor increasing the drain current once the gate is activated [5]. The magnitude of the drain current is a measure for the number of captured electrons. A linear behavior is expected, if less than ∼40,000 electrons are collected, saturation is reached at ∼60,000 electrons [6].

The accumulated electrons can be removed from the internal gate by applying a positive voltage to the n-doped clear gate. A punch-through from the internal gate to the clear contact is induced clearing the captured electrons from the internal gate.

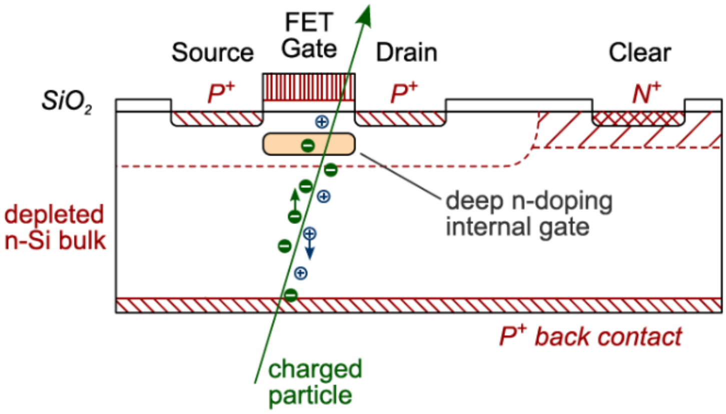


Figure 3.5: Schematic drawing of a DEPFET sensor from [5]. A negative voltage is applied to the p-doped back side depleting the silicon bulk. The n-doped internal gate serves as a collection electrode for negative charges. The collected charge carriers induce image charges in the transistor modifying the drain current once the gate is activated. After each read-out cycle the negative charges are removed from the internal gate by applying a positive voltage to the n-doped clear electrode.

3.8.2 PXD Module

The DEPFET sensors are arranged in modules as visualized in Fig. 3.6. Each module is divided into two segments. The inner segment is positioned closer to the IP and comprises 250 x 256 pixels. The outer segment consists of a matrix of 250 x 512 pixels. While the v-axis is equivalent to the global z-axis, the u-axis lies parallel to the r-ϕ direction. Thus in total each module consists of 767 pixels in v direction and 249 pixels in u direction.

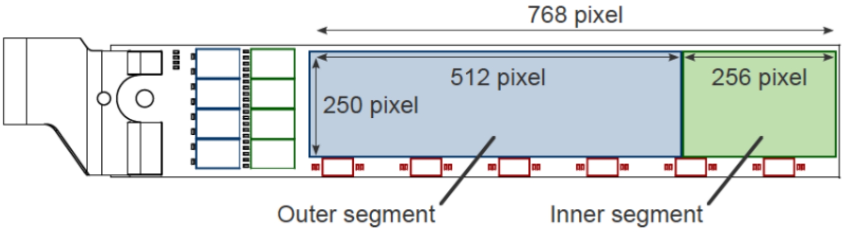
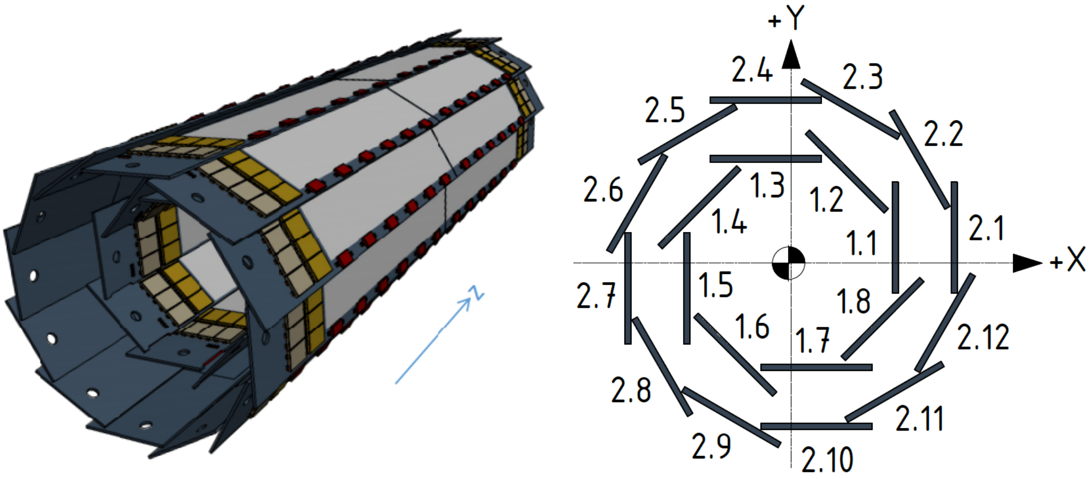


Figure 3.6: Schematic drawing of a PXD module from [6]. The local v coordinate is equivalent to the global z coordinate and the u coordinate is parallel to the r-ϕ direction. The Switchers are marked as red rectangles, ASICs on the rim of the module are responsible for the read-out.

The modules are read out in a *rolling shutter mode* meaning that all sensors along four rows are addressed in parallel. The read-out is performed by ASICs bump-bonded to the rim of each module. Six so-called *Switchers* are responsible for the selection of pixel rows. They are marked in red in Fig. 3.6. The inner segment is read out by two Switchers, the remaining four address the pixels on the outer segment. The Switchers initiate the read-out by applying a voltage to the gates of the pixel sensors. The source-drain current is monitored by Drain Current Digitizer (DCD) chips [50]. For each pixel the base current (*pedestal*) is subtracted from the analog current value. After subtraction the current is digitized in an 8 bit ADC and sent to the Digital Handling Processor (DHP) [51] chips, where it is stored in a ring buffer. When the DHP receives a trigger signal, a digital pedestal is subtracted from the 8 bit digitized current value. Additionally, a *common mode*, representing the average zero offset of a read-out block, is subtracted. Finally, only pixels with a non-zero current value are transferred to the data acquisition (*zero suppression*). After each read-out cycle the Switcher chips provide a clear voltage to the pixel sensors to remove the accumulated electrons from the internal gate.

For each row a a read-out time of 104.17 ns is required resulting in a total read-out time of 20 µs for an entire module. During this time frame all detected charges are



(a)

(b)

Figure 3.7: Schematic representation of the PXD from [7] and [6]. On the left-hand side the CAD drawing of the sub-detector is shown. It consists of 40 modules divided into two layers. Each module can be divided into two segments.
On the right-hand side the numbering scheme of the PXD is visualized. A specific module is addressed in the format $(layer).(ladder).(module)$, where *layer* is 1 for the inner and 2 for the outer layer and *ladder* is assigned according to the image. The variable *module* is not shown.

integrated. The long integration time poses a problem whenever a noisy-bunch crosses the interaction region. Noisy-bunches, resulting from the continuous injection into the storage ring, increase the PXD background level significantly for 3.5 ms every 20 ms. The PXD is unable to function properly with the high background level leading to an effective dead-time of about 20%. The effect is mitigated by operating the PXD in the *gated mode* during the noisy bunch crossings. A positive voltage is applied to the clear gate leading to an accumulation of newly created electrons at the clear contact instead of the internal gate. Electrons captured by the internal gate before the gated mode is activated remain at the internal gate , which prevents a loss of charge information. During the noisy bunch crossings the gated mode renders the PXD blind to newly created charge carriers reducing the dead-time significantly.

3.8.3 Geometrical Set-Up

The PXD consists of 40 modules, which are arranged in two layers of a windmill-like structure as presented on the left-hand side of Fig. 3.7. The inner layer hosts 16 modules, which are combined into 8 ladders by glueing together two modules. In a similar fashion the outer layer consists of 12 ladders by combining 24 modules [52]. While the inner layer is positioned 14 mm from the interaction point, the other layer is located at a radius of 22 mm.

The following numbering convention is established to address the modules unambiguously: (*layer*).(*ladder*).(*module*) Where the variable *layer* is 1 for the inner and 2 for the outer layer. The variable *ladder* refers to a specific ladder. It starts from 1 at the rightmost ladder and is increased counter-clockwise as can be seen on the right-hand side of Fig. 3.7. For the module positioned in the forward part of each ladder the variable *module* is set to 1, for the other module it is 2.

During Phase II only four of the 40 modules were installed in Belle II. The installed modules (1.1.1, 1.1.2, 2.1.1, 2.1.2) sit on two ladders, which are positioned in a line along the x axis. As of today (beginning of Phase III) all modules on layer 1 and two modules on layer 2 are assembled.

3.8.4 Data Acquisition

Once the PXD receives a trigger signal, the DEPFET Handling Hub (DHH) initiates the data acquisition. The DHH consists of 40 DHE chips and 8 Data Handling Concentrator (DHC) chips. Each Data Handling Engine (DHE) is connected to the four DHPs on one PXD module. The DHEs extract, combine and clusterize the pixel data. Subsequently, the data is sent to the DHCs, where sub-event building is performed. Each DHC is connected to two DHEs from the inner and three from the outer PXD layer to ensure that data is distributed evenly among the chips.

The PXD data is further reduced by the Online Selector Nodes (ONSEN) [53]. The ONSEN stores the PXD data for 5 s or until a HLT signal is received. The separation between signal and background in the PXD data is performed based on particle tracks reconstructed by the HLT. The tracks are extrapolated to the PXD, where Regions of Interest (ROIs) are defined in a certain radius around the intersection point between PXD and extrapolated track. The size of the region is determined by the resolution associated with the track from which a ROI is formed. Pixels located outside of ROIs are discarded leading to a reduction in PXD data by about a factor of 10.

Additionally, the Data Concentrator (DATCON) [54] provides ROIs to the PXD solely based on SVD data. While the HLT is more sensitive to high momentum particles, the DATCON focuses on low momentum tracks, which are still reconstructable with the SVD. The DATCON is, however, unable to issue a trigger signal. Events for which no HLT trigger is issued are discarded entirely without consideration whether ROIs are present or not.

A major disadvantage of the ONSEN is its blindness to particles with a radial range smaller than the outermost SVD layer. Neither HLT nor DATCON recognizes these short-range particles typically characterized by low momentum and a high energy deposit.

3.9 Background at Belle II

Compared to its predecessor Belle II has to cope with a significant increase in background due to the higher luminosity and operation characteristics of the SuperKEKB accelerator. The main background sources can be divided into two categories according to their production mechanism: *beam-induced* and *luminosity-dependent* background. While the former one originates from the interaction of the beam particles with each other or with residual gas in the beam pipe, the luminosity-dependent processes arise directly from the electron-positron collisions. In the following the dominating background processes are briefly discussed:

3.9.1 Beam-induced Background

Beam particles circling around the storage ring are constantly subject to interactions with residual gas such as H_2 and CO (carbon monoxide) [55] in the beam pipe or to inhomogeneities in the magnetic fields. As a consequence, beam particles can leave the beam's nominal trajectory. If the acceptance of the accelerator is left, a collision with the beam pipe or other accelerator components is inevitable. A particle shower can be generated by the beam particles due to their high energy. If the shower develops close to the interaction region, the shower particles can enter the sensitive part of the detector.

The high luminosity of SuperKEKB demands a high density of particles in a beam bunch. A disadvantage of this high density is the collision rate between particles within a bunch. Additionally, the Betatron oscillations of the beam, which occur perpendicular to the beam direction [33] enhance the collision probability. Particle collisions with an energy transfer high enough to divert a particle such that it leaves the bunch are referred to as *Touschek effect* [56]. For a detailed explanations as well as the techniques employed to simulate the process in the basf2 framework, see [6] and references therein.

The third class of beam-induced background is a consequence of synchrotron radiation. Since the electrons and positrons are accelerated by the bending and focusing magnets along the storage ring, they are subject to the emission of radiation known as synchrotron radiation [57]. Especially the QCS near the IP is a source of synchrotron radiation affecting the PXD. The contribution from the HER is much larger compared to the radiation stemming from the LER.

Beam-induced background can be identified by the origin of the particle flux. While the processes stated above can occur along the entire beam pipe, particles generated by the e^+e^- collision are close to the nominal IP.

3.9.2 Luminosity-dependent Background

The luminosity-dependent background scales with the luminosity of the accelerator. The background at SuperKEKB is therefore approximately 40 times larger compared to KEKB [6]. A theoretical description about the impact of luminosity-dependent background on the PXD cannot rely on experimental experience, since, as of today, no pixel detector in such a close vicinity to the IP at a collider with the luminosity reached by SuperKEKB has been employed.

A prominent source of background at Belle II is radiative Bhabha scattering due to its high cross section in electron-positron collisions. In the reaction $e^+e^- \rightarrow e^+e^-\gamma$ the scattering can either occur with a wide opening angle between the two leptons allowing them to reach the Belle II detector or the leptons are scattered under a shallow angle, in which case they continue along the beam line. An interaction between the scattered particles with the accelerator components can then lead to back-scattering into the detector. If the photon from the radiative Bhabha scattering reaction interacts with the detector material, it can generate a *giant* *dipole resonance* [58]. This can induce the emission of a neutron, which is able to create severe damages in the sensors and electronics of the PXD.

The dominating background process is the two-photon process:

$$e^+e^- \to e^+e^-\gamma\gamma \to e^+e^-e^+e^-. \tag{3.9.1}$$

The background particles are generated with a low energy. While the PXD is heavily affected from the two-photon background, the contribution to the outer sub-detectors is negligible.

3.10 The Software Framework of Belle II

The Belle II software framework `basf2` [59] is the successor of the Belle Analysis Framework (`basf`), which was the main software tool in the Belle experiment [60]. The framework `basf2` serves multiple purposes related to offline reconstruction and analysis as well as online tasks such as data acquisition, data quality monitoring and the HLT [61].

The architecture of `basf2` follows a modular set-up. A data processing chain is constructed by defining so-called `paths` serving as containers for `modules`. Each `module` is designed to fulfill a specific task (e.g. reading/writing data, tracking, performing a full detector simulation etc.). The `modules` inside a `path` are processed in a linear order one after another. The data is stored in a common storage called `DataStore` in order to transfer it between `modules`.

Input data is processed on an event by event basis. For each event all `modules` specified in the `path` are executed, while input and generated data are stored in the `DataStore`. When the last module has been executed, the `DataStore` is cleared and the next event is processed independently from the former one. It is possible to establish multiple paths branching off of each other, which allows for more flexibility, if different classes of events require different treatments.

While the core code of `basf2` is written in C++, Python is used as an interface. Users address the `basf2` framework with scripts referred to as `steering files`. The `path` and all `modules` are declared in those Python files.

3.10.1 Full Physics Simulation

A full simulation with `basf2` starts with the generation of Monte Carlo particles. For this purpose, `basf2` provides an interface to EvtGen [62], which is used to model particle production processes. EvtGen is particularly suited for the simulation of B mesons and related resonances. Alternatively, MC particles can be generated with the `particle gun module`, which allows to define customized distributions for momentum, energy, vertex position and various other properties.

The passage of particles through the Belle II detector is simulated after the production of primary MC particles. The simulation requires precise knowledge about the material and geometrical set-up of the detector. The detector geometry is built from so-called *primitives* (boxes, spheres,...), to which properties such as radiation length are assigned. Instead of storing all geometrical objects, which make up the Belle II detector, only the geometry parameters are written to XML files for look-up during the simulation. The `Gearbox module` handles the creation of the geometry from the XML files. The real detector geometry is subject to variations as a consequence of external influences such as temperature. By measuring the geometry regularly, a precise detector description can be maintained at all times.

For the full detector simulation `basf2` employs Geant4 [63]. Geant4 is a standard toolkit in particle physics to simulate the passage of particles through matter taking magnetic fields, energy loss, multiple scattering, the creation of secondary particles and various other effects into account. The passage of a particle through material is discretized into steps. In each step the interaction between particle and detector material is computed. The default step size in the PXD is set to 5 µm [6].

All sub-detectors are equipped with a `Digitizer` mimicking the response of a real detector. The digitization occurs at the end of each Geant4 step. The digitization routine is unique for each of the different sub-detectors. For the PXD the creation of electron-hole pairs, their drift and diffusion inside the silicon sensors as well as the registration of electrons at the internal gate are simulated [64].

The simulated detector response is treated like proper data in the reconstruction stage. For the PXD the clustering of individual pixels is the most prominent reconstruction step. Charged particles impinging on the PXD typically activate several adjacent pixels. In order to identify a set of pixels with the passage of a single particle they are grouped together to clusters. The `Clusterizer module` handles the identification of clusters. Additionally, the `Clusterizer` computes cluster properties such as size, total charge, seed charge and many more. The position of a cluster is reconstructed by a head-tail algorithm [65] or center-of-gravity algorithm depending on the size of the cluster. The cluster serve as an input for subsequent physics analyses and for tracking purposes.

CHAPTER 4

ANALYSIS OVERVIEW

Countermeasures for the drastic reduction of PXD data and the inevitable loss of information related to low-momentum or highly ionizing particles have already been considered in the past. Most notably, a six-layer tracking using the two PXD and four SVD layers has been proposed [66, 67] and a NeuroBayes algorithm trained on PXD clusters has been developed [68] for the identification of slow pions. Complementary to these works, the identification of highly ionizing particles is the central objective of this thesis. The application of neural networks is the main instrument for this endeavor. The following Chapter is therefore dedicated to the implementation of the two types of neural networks considered in this thesis: Feed-Forward Networks and Self-Organizing Maps. First, the input vector sets are presented and the choice of input properties is motivated. A more detailed description of the techniques employed in this thesis can be found in the additional material available on the Belle II document server or the publication page of the II Physics Institute of the University of Giessen [69].

4.1 Input Vectors

The composition of the input vector is subject to stringent and often contradictory requirements. A high information density, small number of entries to decrease computing power and strong robustness against errors, to name only a few demands, are mandatory for a stable and successful training process. In the following two sets of input variables are introduced.

4.1.1 Cluster Properties

Clusters are commonly described by a set of numbers encoding their geometrical characteristics. These *cluster properties* are computed by the `Clusterizer module` implemented in `basf2`. While properties such as the cluster size, total cluster charge and many more are part of the common computation routine for the `Clusterizer`, other properties like the minimum pixel charge, the cluster angle, etc. have to be determined manually by iterating over all pixels belonging to a cluster.

In Table 4.1 all cluster properties considered in this thesis are listed. The second column notes whether computation of a property is already included in the `Clusterizer`. All cluster properties, which are not computed by the *Clusterizer*, are determined manually by iterating over the individual pixels inside the cluster. Those properties deemed to be a suitable part of the input vector are marked with a ✓in the third column. The decision is justified in the additional material (see supporting material on Belle Document Server or publication page of II Physics Institute of University of Giessen [69]). Computability is the main criterion for all choices.

4.1.2 Pixel Charge Distributions

The pixel charges from a 5 x 5 pixel matrix around the seed charge of a cluster are considered. The 25 unprocessed pixel charges are combined into an input vector for the neural networks. Contrary to the computation of cluster properties, the information contained in the activated pixels is not reduced as long as the extension of the 5 x 5 pixel matrix captures the entire set of non-zero pixel cells. The learning process computationally more expensive due to the greater size of

Table 4.1: Cluster properties considered in this thesis. The second column marks whether a property is already computed by the `Clusterizer`. The third column indicates which properties are chosen to be part of the input vector for the neural networks.

Property	Computed by `Clusterizer`	Part of input vector
Total cluster charge	✓	✓
Cluster seed charge	✓	✓
Cluster minimum charge	×	✓
Total cluster size	✓	✓
Total cluster length	×	×
Cluster size in u	✓	✓
Cluster size in v	✓	✓
Cluster angle	×	×
Cluster eccentricity	×	×

the input vector.

4.1.3 Discussion

The input vector sets considered in this thesis do not contain any information about the PXD layer, the specific ladder or module. Further development of machine learning techniques at the PXD should consider including them for two reasons: the global position of a cluster, which is encoded in the combination of PXD layer, ladder and module variables, can contribute novel information helping in the identification of the particle, which generated the cluster. The background occupancy in the second layer, for instance, is much lower compared to the first layer, which could make information coming from the second layer more meaningful. Additionally, the global distribution of beam background clusters along the z axis is distinctly different to the distribution of clusters generated at the IP. A detailed explanation about this phenomenon will be given in the subsequent Chapters. Apart from the prospect of increasing separation power, the geometrical properties of a pixel depend on the global position in the PXD as explained in Section 3.8. A pixel located at an inner module on the first layer, is smaller in v direction than pixels positioned farther away from the IP to account for the higher occupancy. The sharing of charge carriers to adjacent cells should be inhibited on the outermost pixels in the PXD due to the larger spatial extend of an individual sensor cell. Read-out thresholds as well as noise levels also differ from pixel to pixel compromising the assumption that all pixels are equal. While these phenomena should be taken into account eventually, they are negligible for the precision required for a first feasibility study. Moreover, the assumption of uniform pixel properties to avoid the treatment of individual pixel cells is not unreasonable and generally common practice.

4.2 Artificial Neural Networks

Artifical Neural Networks (ANNs) are a form of machine learning. Machine learning in contrast to conventional static programming techniques, relies on the ability of an artificial executing system to progressively adapt to solving a certain task.

An ANN in particular is modeled after biological neural networks [70]. It consists of processing units referred to as *artificial nodes*, which are connected among each other.

ANNs can be classified into two categories according to the type of learning process, which they undergo:

- **Supervised Learning**

  Each input is associated with a desired result. Learning is performed by comparing this result with the output of the ANN for a given input. The artificial nodes are adapted to maximize the agreement.

- **Unsupervised Learning**

  No desired output is assigned to the input vectors. The ANN clusters data according to certain features. The aggregation process is mainly determined by the metric of the ANN. A Euclidean metric is a popular choice due to its compatibility with human perception of space.

Feed-Forward Neural Networks (FFNs) as a form of supervised learning and Self-Organizing Maps (SOMs) representing unsupervised learning algorithms are chosen for the identification of HIPs. The implementation of the two network types is described in the following.

4.2.1 Implementation of Feed-Forward Networks

FNNs are implemented with the PyTorch framework [71]. PyTorch offers a library with machine learning algorithms adapted for CPUs and GPUs supporting CUDA [72].

The FFN architecture used in the following Chapters is presented in Table 4.2. The abbreviation *ClsPrp* is used whenever the cluster properties serve to encode clusters for the input vector set to the network. Likewise, *PxlDis* stands for the choice of the pixel charge distribution as input vectors. The FNN consists of four fully-connected layers: one input, one output and two hidden layers. Each layer hosts 40 (ClsPrp) or 50 (PxlDis) nodes. The activation function is a Rectified Linear Unit (ReLu) and the optimizer employs Stochastic Gradient Descent (SGD). A batch size of 256 has been chosen to comply with the total number of vectors (300,000 - 500,000) in the training set. A learning rate of 1E-4 with momentum 0.9 is employed.

Table 4.2: Hyperparameters of Feed Forward Neural Networks. The abbreviation ClsPrp stands for neural networks, which are presented with cluster properties, PxlDis for those operating with pixel charge distributions.

Parameter	ClsPrp	PxlDis
Nr. of layers	4	
Nodes per layer	40	50
Activation Function	Rectified Linear Unit (ReLu)	
Loss Function	CrossEntropy	
Optimizer	Stochastic Gradient Descent (SGD)	
Batch Size	256	
Learning Rate	0.0001	
Momentum	0.9	

A *softmax* function transforms the output of the FFN such that it can be interpreted as probability. The function can be defined as:

$$\text{Softmax}(x_i) = \frac{\exp(x_i)}{\sum_j \exp(x_j)} \tag{4.2.1}$$

where the x_i are components of the FFN output vector. The sum in the denominator runs over all entries in the output vector. The elements of the output vector are transformed to values between 0 and 1. The sum of all vector components adds up to 1 allowing to treat the individual entries as probabilities. Modifications of the above definition such as the α-softmax function or related functions such as the α-quasimax are common parameterization tools as well [73].

4.2.2 Implementation of Self-Organizing Maps

The implementation of SOMs is performed with MiniSom [74]. Its minimalistic set-up and reliance on Numpy allows for quick and straightforward adaptations of the SOM algorithm to suit the requirements of the training processes performed in this thesis.

In Table 4.3 the hyperparameters of the SOM used in the following two Chapters are listed. Adaptations to these parameters are required for some training processed. Whenever deviations from the network architecture in Table 4.3 are used, it will be mentioned and discussed explicitly.

Table 4.3: Hyperparameters for Self Organizing Map. A two-dimensional node space and a Gaussian neighborhood kernel have been chosen.

Parameter	ClsPrp/PxlDis
Nr. of nodes:	15 x 15
Neighborhood function:	Gaussian
Width of Gaussian:	7
Learning Rate:	0.01

4.2.3 Receiver Operating Characteristic (ROC) Curve

A Receiver Operating Characteristic (ROC) curve is a popular measure for the quality assessment of ANNs [75]. In Fig. 4.1 a schematic representation of a ROC curve is shown. For a given efficiency the background rejection of a neural network is computed. Efficiency is defined as:

$$\frac{\#\text{ vectors correctly identified as signal}}{\#\text{ vectors identified as signal}} \tag{4.2.2}$$

and background rejection as

$$\frac{\#\text{ vectors correctly identified as background}}{\#\text{ true background vectors}} \tag{4.2.3}$$

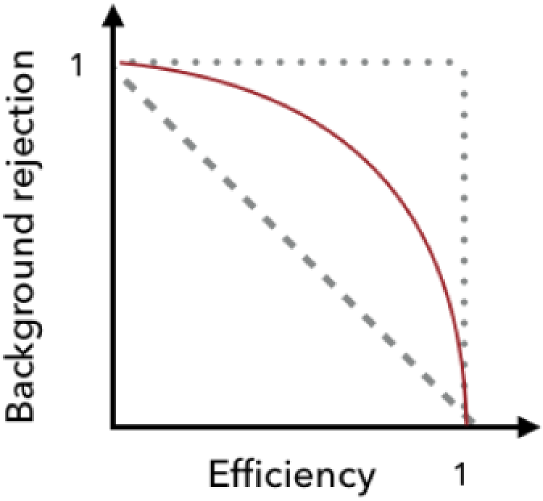


Figure 4.1: Schematic representation of a ROC curve. The two dashed lines correspond to perfect classification (*box function*) and random guessing (diagonal).

Both figures of merit are usually correlated in case of non-ideal classification: an increase in efficiency typically implies a decrease in background rejection. Conversely, a tight signal definition leads to high background rejection and a low efficiency.

In case of perfect classification a constant background rejection and efficiency can be achieved resulting in a ROC curve resembling a *box function*. In case of random guessing ROC curves follow the dashed diagonal in Fig. 4.1. In all other cases the curves lie between the two extremes. A large area under a ROC curve (Area Under Curve - AUC) is associated with a high classification power of an ANN. The shape of the curve can reveal characteristics of the classification process.

CHAPTER 5

IDENTIFICATION OF ANTI-DEUTERONS

The detection of cosmic ray anti-deuterons is a promising endeavor for the indirect search for dark matter annihilation or decay [76]. The low astrophysical background distinguishes the indirect search using anti-deuterons from typical cosmic ray candidates like γ-rays or anti-protons. In order to understand the production of anti-deuterons from dark matter decays, particle collider measurements are essential.

In dark matter annihilations or decays, the decay products can hadronize to form anti-protons and anti-neutrons, which can coalesce into anti-deuterons. Donato et al. [77] first reported the significance of low-energy anti-deuterons as a probe for indirect dark matter searches. The astrophysical background mainly arises from the interaction of anti-protons and anti-neutrons with the interstellar medium. Due to the high threshold energy of this process and the steeply falling energy spectrum of cosmic rays, the background is considered to be extremely low. In addition, background anti-deuterons produced in this manner are characterized by a large kinetic energy.

The production of anti-deuterons from anti-protons and anti-neutrons is described by the so-called *coalescence model.* While early research used an analytical model to describe the merging of anti-protons and anti-neutrons, recent publications (see e.g. [76, 78] and references therein) employ an event-by-event Monte Carlo simulation backed up with data from high energy collider experiments.

In this Chapter the capability of performing anti-deuteron studies at Belle II is briefly discussed. In particular, the range of anti-deuterons in the Belle II detector is investigated. Subsequently, PXD clusters generated by anti-deuterons are characterized based on specific cluster properties and pixel charge distributions. In the last Section the application of neural networks for the identification of anti-deuteron clusters is presented.

5.1 Anti-Deuterons at Belle II

Anti-deuteron production in $\Upsilon(2S)$ and $\Upsilon(3S)$ decays as well as in e^+e^- continuum at a CMS energy of 10.58 GeV have been measured by the CLEO [79], ARGUS [80] and BaBar collaboration [81]. The results are subject to considerable uncertainties and inconsistent with MC simulations [76]. New anti-deuteron measurements at Belle II could help to shed light on these contradictions.

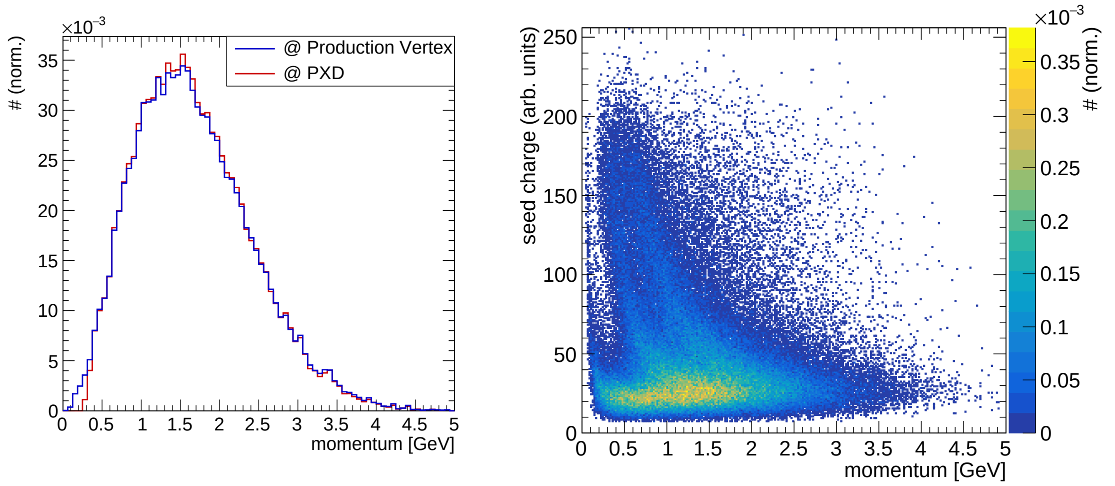
The production mechanism

$$\Upsilon(4S) \rightarrow 2\pi^+ + 2\pi^- + \pi^0 + p + n + \bar{d} \tag{5.1.1}$$

is considered to benchmark the detection capabilities of Belle II with regard to anti-deuterons [82]. The anti-deuteron $\bar{d}$ is produced from a $\Upsilon(4S)$ decay along with five pions $\pi^{\pm,0}$, a neutron n and a proton p. The event is generated with EvtGen. A PHSP (N-body phase space) decay model is used.

The momentum distribution at the production vertex and the PXD for anti-deuterons is depicted on the left-hand side of Fig. 5.1. While the most probable momentum is ~1.5 GeV, the tail of the distribution extends to ~4.5 GeV. The distributions differ only below ~0.3 GeV. Anti-deuterons with production momentum below this value are unable to reach the PXD, they decay in the beam pipe.

The cluster seed charge in the PXD as a function of momentum for all charged particles in the $\Upsilon(4S)$ decay is presented on the right-hand side of Fig. 5.1. Distinct bands reminiscent of Bethe-Bloch curves for the three particle species can be seen



(a)

(b)

Figure 5.1: (a) Momentum distribution at the production vertex and the PXD for anti-deuterons. While the most probable momentum value is ~1.5 GeV, the tail of the distribution extends up to ~4.5 GeV. A change in the momentum distribution between production vertex and PXD is only discernible below ~0.3 GeV (b) Cluster seed charge in the PXD as a function of momentum for all charged particles generated in the production mechanism considered in this Chapter. The Bethe-Bloch bands are well-separated above ~2.5 GeV. The highest band can be associated with anti-deuterons, the central band with protons and the lowest band with pions.

demonstrating that the cluster seed charge is a suitable measure for the particle's energy loss in the PXD. The highest band originates from anti-deuterons, the central band from protons and the lowest band can be associated with pions. Above ~2.5 GeV the bands are visually indistinguishable rendering identification based on the magnitude of the deposited energy difficult. Low-momentum particles, on the other hand, display clearly separated bands.

For the remainder of this thesis anti-deuteron events are generated with the `particle gun` instead of EvtGen. A flat distribution in ϕ from o to 2π and in θ between 17° and 150° is chosen. The momentum distribution is chosen to be flat as well. It ranges between 0 GeV to 3 GeV. Phase III beam background (MC simulation) is overlaid during the simulation.

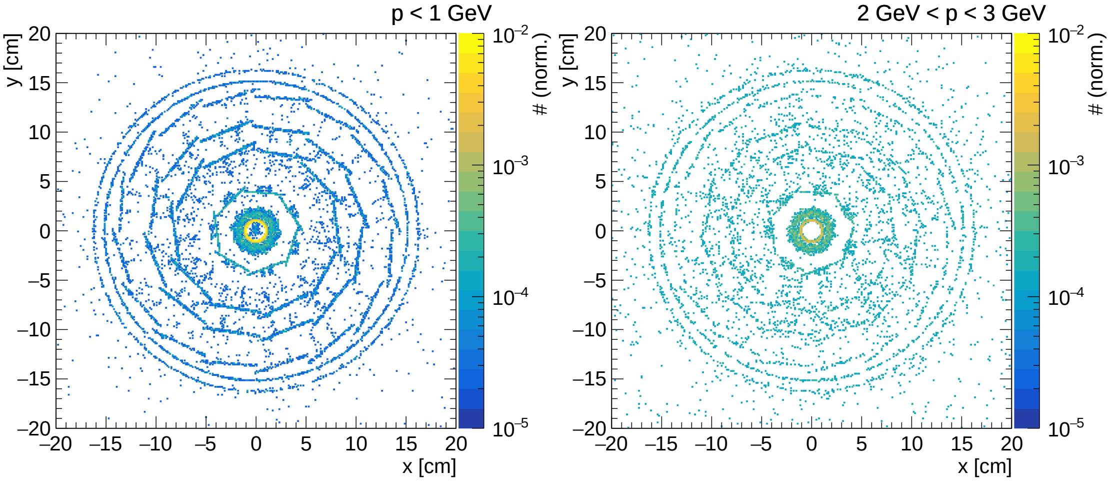
5.1.1 Range

The primary challenge in the identification of anti-deuterons is their limited range in the detector. The range is strongly correlated with the momentum of particles requiring different treatments for different momentum regimes.

The decay vertices of the MC anti-deuterons is used to estimate their range in the detector. A projection of the decay vertices onto the xy plane is shown in Fig. 5.2 for anti-deuterons with momentum lower than 1 GeV (left-hand side) and with momentum ranging from 2 GeV to 3 GeV (right-hand side). In both cases a significant amount of anti-deuterons decay in the beam pipe, which is visible as a yellow ring around the center of the plot. The green ring around the beam pipe represents the PXD layers. The two layers are not clearly distinguishable due to the resolution chosen for the plot. The density of decay vertices in the PXD is about one order of magnitude lower than the decay vertex density in the beam pipe. The individual layers and even the wind-mill structure of the SVD is clearly discernible from the distribution of decay vertices. For the low momentum anti-deuterons the number of decay vertices is about one order of magnitude lower compared to the particles with 2 - 3 GeV momentum. The low-momentum anti-deuterons decay predominantly in the beam pipe and PXD layers leaving a smaller number of particles decaying in the SVD layers compared to the higher momentum particles.

A projection of decay vertices onto the onto the ρz plane yields similar results as can be seen in Fig. 5.3 for $p < 1\,\mathrm{GeV}$ anti-deuterons and in Fig. 5.4 for $2\,\mathrm{GeV} < p < 3\,\mathrm{GeV}$ anti-deuterons. The positions of decay vertices mirror the material budget of the Belle II detector. Below approximately $\rho = 20\,\mathrm{cm}$ the layers of the vertex detectors are discernible. Above $\rho = 20\,\mathrm{cm}$ and below approximately $\rho = 115\,\mathrm{cm}$ three regions along z direction can be differentiated: the forward, barrel and backward region. Starting from the backward (negative z) region the backward endcap ECL as well as additional electronics are discernible from the density of decay vertices. In the barrel region the density of vertices is sparse due to the low material budget of the CDC. In the forward region the ARICH and the forward ECL endcap are visible. At about $\rho = 120\,\mathrm{cm}$ the TOP and at $\rho = 125\,\mathrm{cm}$ the lower part of the barrel ECL can be seen.

While for the low-momentum anti-deuterons there are no more decay vertices above the first few cm in the ECL, the high-momentum anti-deuterons penetrate more deeply into the sub-detector. The density of decay vertices, however, is approximately three orders of magnitude lower compared to the vertex region. A



(a)

(b)

Figure 5.2: Decay vertices of anti-deuterons projected onto the xy plane for anti-deuterons with momenta below 1 GeV (left-hand side) and 2 - 3 GeV (right-hand side). The beam pipe is discernible as a yellow ring indicating the high density of decay vertices in this region. The PXD can be recognized as a green ring around the beam pipe. The SVD layers are clearly distinguishable. The density of decay vertices in the SVD is about one order of magnitude higher for the higher momentum particles.

close-up of the vertex region is shown in Fig. 5.5 for the low-momentum and in Fig. 5.6 for the high-momentum anti-deuterons. In both cases the density of decay vertices exhibits a maximum at the beam pipe located at about $\rho \sim 1\,\mathrm{cm}$. The density decreases approximately by two to three orders of magnitude in the PXD and SVD layers, when the incident particles have sub-GeV momenta and only by one order of magnitude for anti-deuterons with $p = 2$ - $3\,\mathrm{GeV}$. The stronger decline in density, which is also observed in Fig. 5.2, can be attributed to the smaller number of low-momentum particles leaving the beam pipe.

The fraction of particles decaying in the vertex region can be estimated from Fig. 5.7. A projection of the decay vertices onto the ρ axis in the range $z = -3\,\mathrm{cm}$ to $z = 5\,\mathrm{cm}$ is shown. Fig 5.7a and Fig 5.7b display the same distribution, but with a different scaling of the vertical axis. Only a few percent of anti-deuterons with sub-GeV momentum are able to leave the vertex region. Nearly no sub-GeV anti-deuterons are present above $\rho \sim 140\,\mathrm{cm}$. The hard cut-off in range can be attributed to the pronounced Bragg peak for heavy particles [83]. For the high

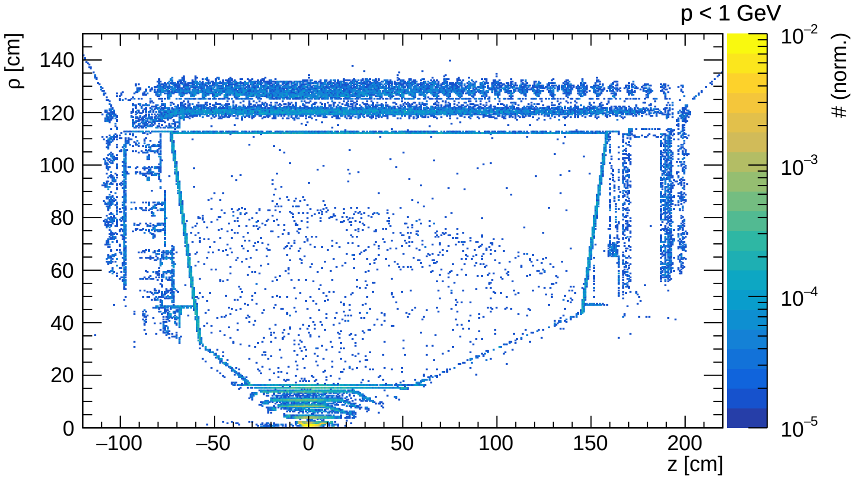


Figure 5.3: Projection of decay vertices of anti-deuterons with $p < 1\,\mathrm{GeV}$ to the $z\rho$ plane. The density of decay vertices has a maximum in the vertex region. A hard cut-off in the barrel part of the ECL ($\rho \gtrsim 120\,\mathrm{cm}$) is observable.

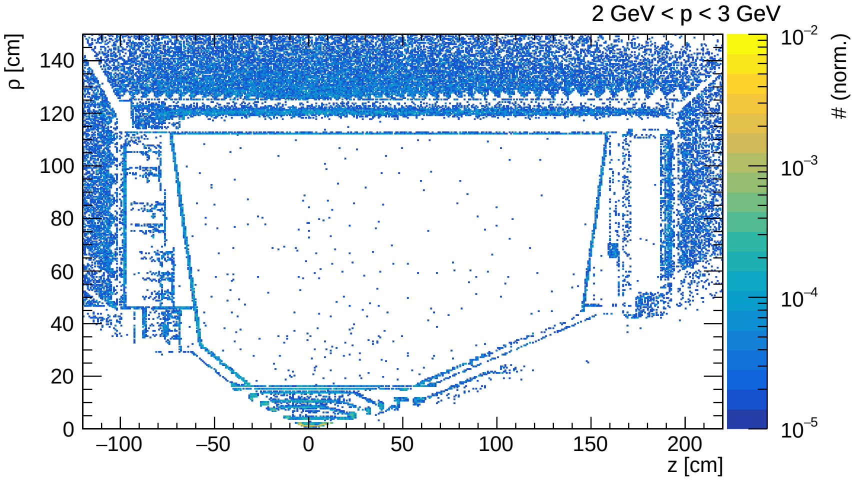


Figure 5.4: Projection of decay vertices of anti-deuterons with $p < 3\,\mathrm{GeV}$ to the $z\rho$ plane. In contrast to the low-momentum anti-deuterons, the distribution of decay vertices extends farther into the ECL. The density of decay vertices in the ECL is several orders of magnitude lower compared to the vertex region.

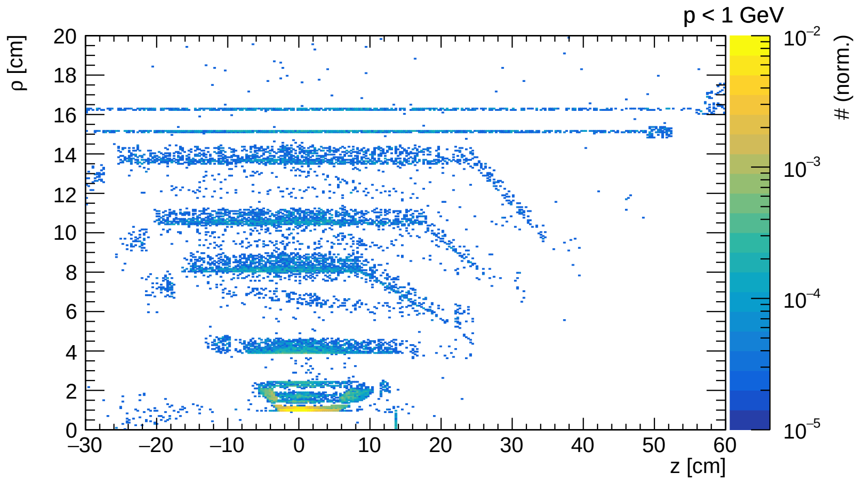


Figure 5.5: Decay vertices in the vertex region projected onto the $z\rho$ plane for anti-deuterons with $p < 1\,\mathrm{GeV}$. The density of decay vertices in the SVD layers ($\rho \gtrsim 4\,\mathrm{cm}$) is about three order of magnitude lower compared to the density in the beam pipe ($\rho \sim 1\,\mathrm{cm}$). The density in the PXD layers ($1.4\,\mathrm{cm} \lesssim \rho \lesssim 2.2\,\mathrm{cm}$) is about two orders of magnitude lower.

momentum particles the situation is slightly less severe. Nevertheless, the majority of particles are confined to the vertex region as well.

5.2 Cluster Properties

The properties of clusters generated by the anti-deuterons are studied in the following Section. To assess the impact of particle momentum on the cluster properties, anti-deuterons with momentum $p < 1\,\mathrm{GeV}$ and $p = 2$ - $3\,\mathrm{GeV}$ are considered separately. The distributions are compared with clusters produced by background particles.

5.2.1 One Dimensional Distributions

The distributions of all cluster properties for anti-deuterons and background are listed in Appendix E. As examples, the total cluster charge and total cluster size

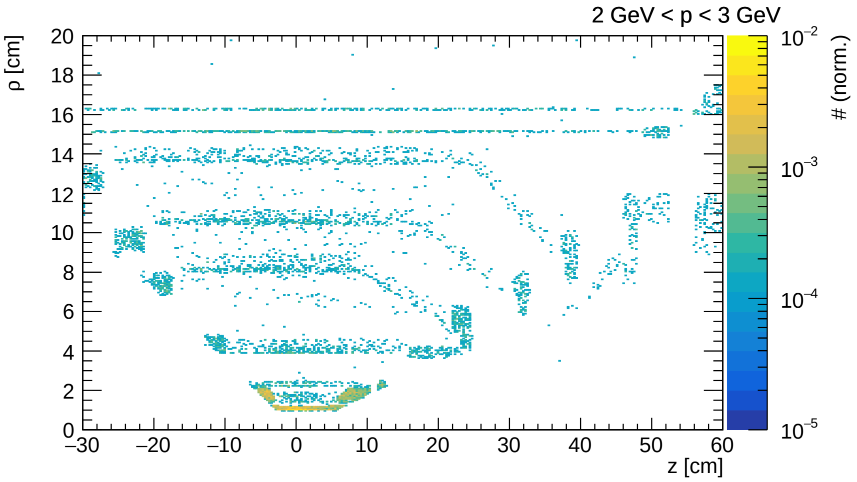


Figure 5.6: Decay vertices in the vertex region projected onto the $z\rho$ plane for anti-deuterons with $2\,\text{GeV} < p < 3\,\text{GeV}$. The density of decay vertices in PXD as well as SVD layers is about two orders of magnitude lower compared to the beam pipe.

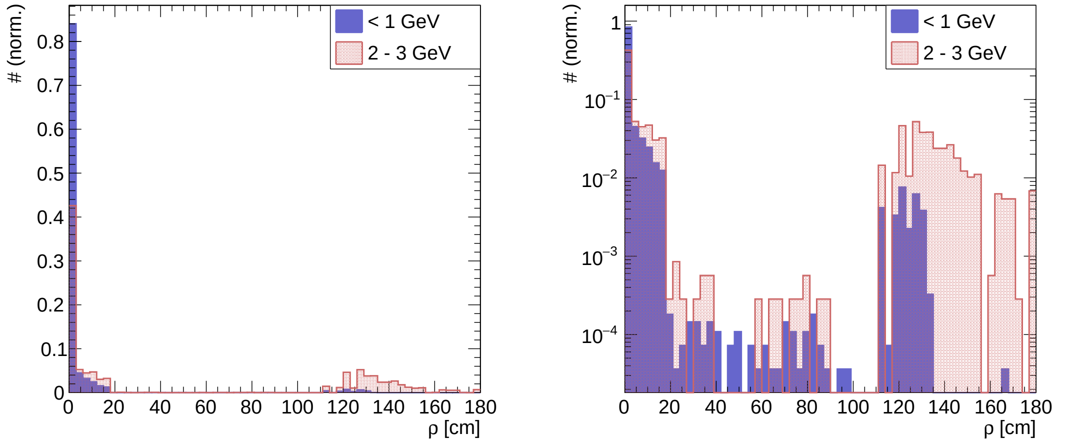
are discussed in this Section.

The cluster charge distributions for anti-deuterons with momentum $p < 1\,\text{GeV}$ (left-hand side) and $p = 2$ - $3\,\text{GeV}$ (right-hand side) as well as beam background are presented in Fig. 5.8. The most probable cluster charge value for beam background is roughly one order of magnitude lower than the most probable value of the sub-GeV anti-deuteron distribution. The background curve resembles a Landau distribution [1] characterized by a heavy tail extending into the high cluster charge region. The high-charge tail of the background distribution overlaps with the cluster charge distribution for low-momentum anti-deuterons. The distributions for anti-deuterons with 2 - 3 GeV momentum and background overlap fully. The most probable value for the high-momentum anti-deuterons is slightly higher and the high-charge tail less pronounced compared to the background distribution.

The shape of the cluster charge distribution for low-momentum anti-deuterons differs from the background and high-momentum anti-deuteron distribution. It is broader and falls less steeply with decreasing charge than the other two. Its

---

[1] For thin silicon sensors the Bichsel model [84] is a more accurate description than the Landau model [85]. For a rough description of the peak shape adopted by the cluster charge distribution the resemblances to the Landau function is a suitable comparison.

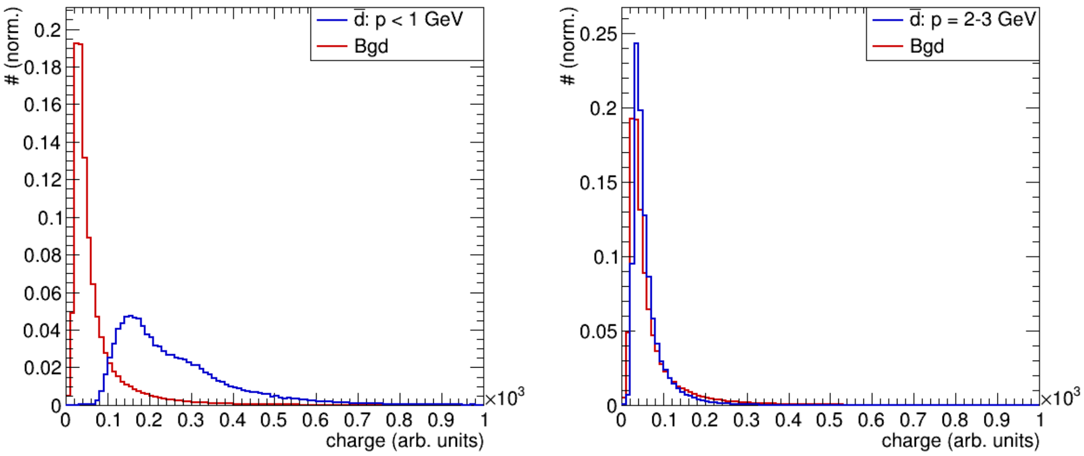


(a) (b)

Figure 5.7: Projection of decay vertices of anti-deuterons onto the ρ axis in the range $z = -3\,\text{cm}$ to $z = 5\,\text{cm}$. Anti-deuterons with sub-GeV momentum decay predominantly in the vertex region, only a sub-percent amount of particles arrive at the ECL. While the majority of anti-deuterons with $2\,\text{GeV} < p < 3\,\text{GeV}$ also decays before leaving the SVD, a few percent are able to reach the ECL.

shape deviates from a a typical Landau-Gauss distribution. The origin of this deviation is most likely related to *uncollected* charge carriers in the silicon sensors. A high energy deposition in one sensor promotes the sharing of charge carriers to neighboring cells. If the number of charge carriers collected at the internal gates of neighboring cells is not sufficient to surpass the read-out threshold the charge carriers are *lost* since they remain unregistered by a pixel cell. The total cluster charge cannot be considered equivalent to the amount of deposited energy anymore due to the loss of charge carriers. As a consequence, a low energy tail develops and an overall broadening of the distribution takes place.

The cluster size distributions for both momentum regimes are shown in Fig. 5.9. The background as well as the anti-deuteron distribution for $p = 2\text{ - }3\,\text{GeV}$ exhibit a maximum at a cluster size of two. The distributions are decreasing with increasing size. Low momentum particles, on the other hand, are less likely to generate single pixel cluster. Additionally, the distribution does not decrease monotonously but possesses a second local maximum at cluster size four. This behavior is explainable on the basis of the cluster size distributions in u and v direction as depicted in Fig. 5.10 and Fig. 5.11.

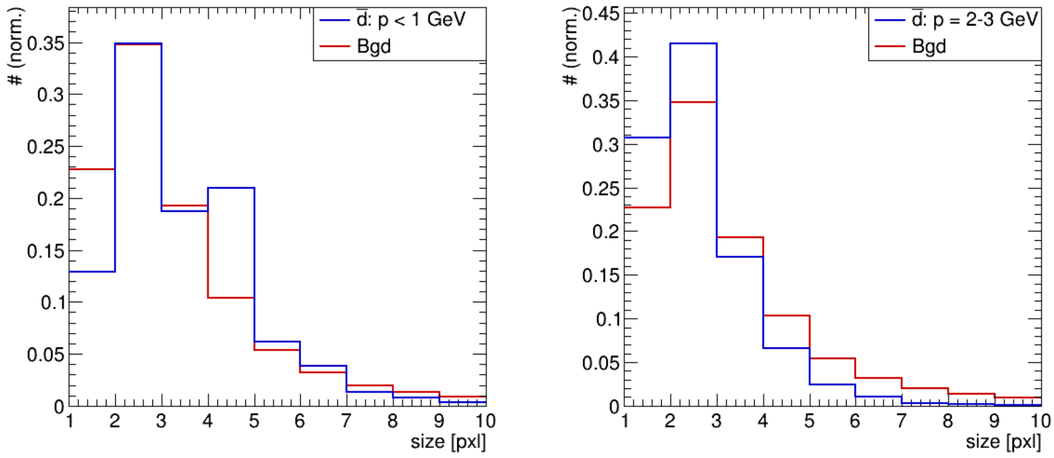


(a) (b)

Figure 5.8: Cluster charge distribution for anti-deuterons with $p < 1\,\mathrm{GeV}$ (left-hand side) and $p = 2$ - $3\,\mathrm{GeV}$ (right-hand side). The distribution for background particles is overlaid in red. The cluster charge distribution for low-momentum anti-deuterons is shifted towards higher charge values compared to background and high-momentum anti-deuteron distribution. Additionally, the distribution is broader and its shape deviates from a typical Landau-Gauss distribution.

While the cluster size in u does not exceed two for anti-deuterons, it can take higher values for background particles. The source of beam-induced background is not confined to the IP, background particles can originate from regions with a considerable distance to the detector. Consequently, they are able to enter the vertex region under a shallow angle leading to the activation of multiple pixels along the u axis. For particles generated at the IP the cluster extension in u is mainly determined by the amount of deposited energy. A high energy deposit generates a high number of electron-hole pairs in the depleted silicon increasing the possibility to have diffusion of the liberated charge carriers into neighboring pixel cells. The higher mean energy deposition of low-momentum anti-deuterons therefore leads to a lower number of clusters with a single pixel in u direction compared to high-momentum anti-deuterons. The same holds true for the cluster size in v with the only exception that the forward boost due to the asymmetrical collision energy increases the cluster size compared to the u direction.

The lack of clusters with a single pixel in u/v direction is responsible for the second local maximum in the cluster size distribution of low-momentum anti-deuterons.



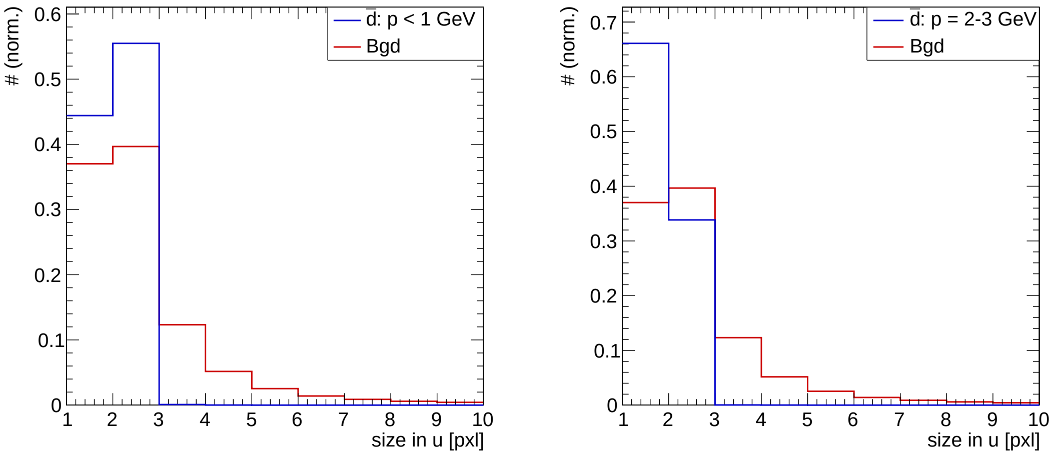
(a) (b)

Figure 5.9: Cluster size distribution for anti-deuterons with $p < 1\,\mathrm{GeV}$ (left-hand side) and $p = 2$ - $3\,\mathrm{GeV}$ (right-hand side) overlaid with the distribution for background particles. The cluster size distribution for low-momentum particles exhibits a global maximum at size two and another local one at size four. The distribution for the high-momentum particles has only one maximum at size two. The distribution is falling monotonously for increasing cluster size.

The most common value for the cluster size in u/v direction is two. The maxima in the total cluster size distribution are therefore located at size two and size four. The reason for the higher number of single pixel clusters for high momentum particles follows the same line of argumentation.

The cluster seed charge increases with decreasing anti-deuteron momentum as shown in Fig. 5.1. Analogously, the mean cluster size increases with decreasing momentum as well (see Fig. 5.12a). The increase is less pronounced compared to the increase in total cluster charge due to *threshold effects*. For an increase in cluster size the signal of an additional pixel has to be larger than the readout threshold. Depending on the threshold, the increase in cluster size can be considerably suppressed.

The lack of clusters with a single pixel in v direction for the low-momentum anti-deuterons can be confirmed as well by considering the cluster size in v as function of momentum (see Fig. 5.12b). The number of single pixel clusters declines continuously with decreasing momentum since the higher energy deposit tends to



(a) (b)

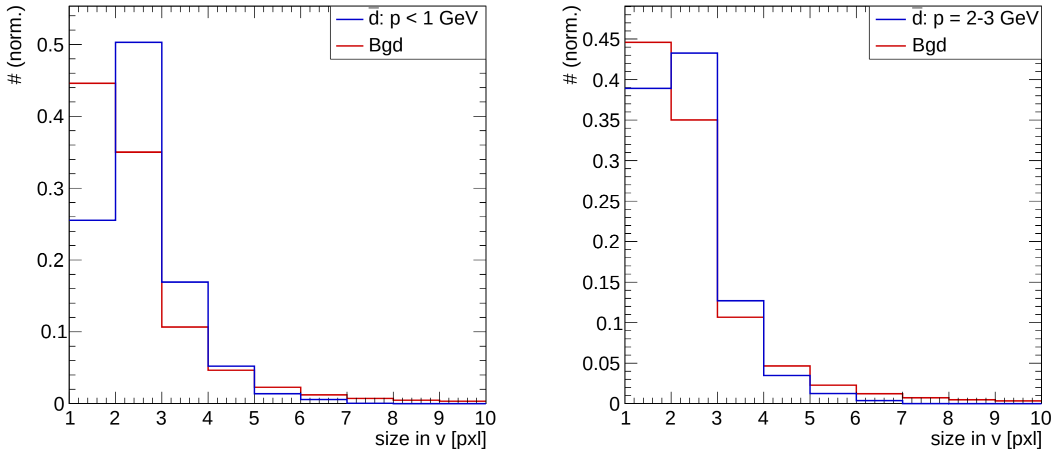
Figure 5.10: Cluster size distribution in u for anti-deuterons and background. While the distributions for low-momentum anti-deuterons exhibits a maximum at size two, high-momentum particles are more likely to generate clusters with size one in u direction. The size in u for background particles can be much higher compared to the anti-deuterons.

increase charge sharing to adjacent cells. The impact of particle momentum on all other cluster properties is shown in Appendix D.

5.2.2 Two Dimensional Distributions

A two dimensional representation of cluster properties can reveal correlations between the variables and allows to investigate the applicability of two dimensional cuts. In this Section two exemplary two dimensional distributions are presented, the entire set of distributions are listed in Appendix F.

The cluster charge against cluster size is shown in Fig. 5.13a for anti-deuterons with $p < 1\,\mathrm{GeV}$, in Fig. 5.14a for anti-deuterons with $p = 2$ - $3\,\mathrm{GeV}$ and in Fig. 5.15a for background. All distributions exhibit a band-like shape. Large low-charge cluster and small high-charge cluster are missing indicating a correlation between cluster charge, which should be proportional to the amount of deposited energy, and number of charge carriers diffusing to adjacent pixel cells as explained in the



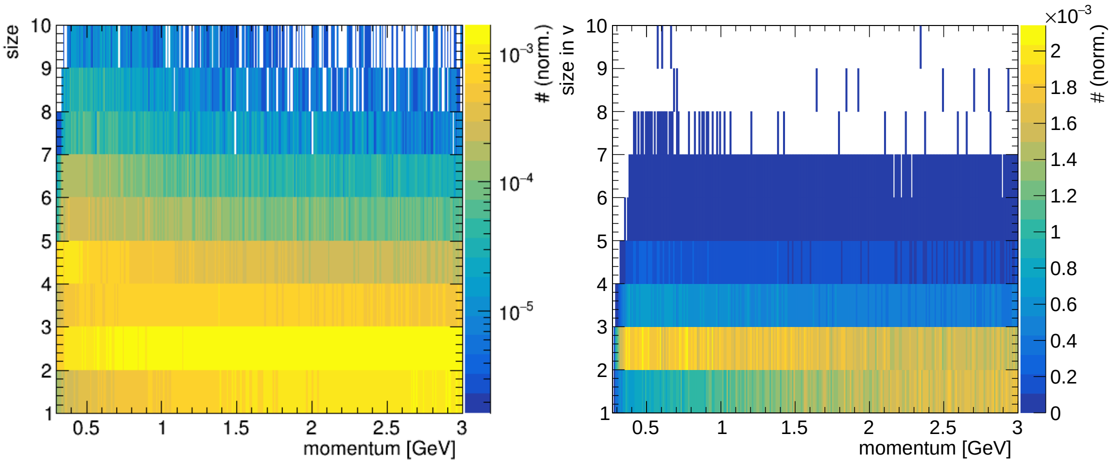
(a) (b)

Figure 5.11: Cluster size distribution in v for anti-deuterons and background. The background distribution is falling monotonously with increasing size in v. Both anti-deuteron distributions have a maximum at size two. The maximum is more pronounced for low-momentum anti-deuterons. The high-momentum particles exhibit a greater number of clusters with size one in v direction.

previous Section. While the sub-GeV anti-deuteron distribution extends along the diagonal, the slope of the distributions for the 2 - 3 GeV particles as well as the background is much steeper.

The low-momentum anti-deuteron distribution covers parameter regions inaccessible by the background distribution. The inability of background particles to generate high-charge clusters has already been discussed in the previous Section. Additionally, the cluster size in the medium charge region (∼300) assumes values between two to nine for anti-deuterons and four to ten for background particles confirming that the correlation between cluster size and cluster charge is blurred for background particles due to the shallow angle under which they enter the PXD. A two dimensional cut, however, would only deliver fair results in the high-cluster charge region.

In Fig. 5.13b the cluster size in u against size in v is depicted for anti-deuterons with $p < 1$ GeV. The anti-deuterons generate clusters in a narrow parameter range with respect to the size in u. The size in u and v appear to be largely uncorrelated: a large cluster size in v does not entail a big extension in u and vice versa. The



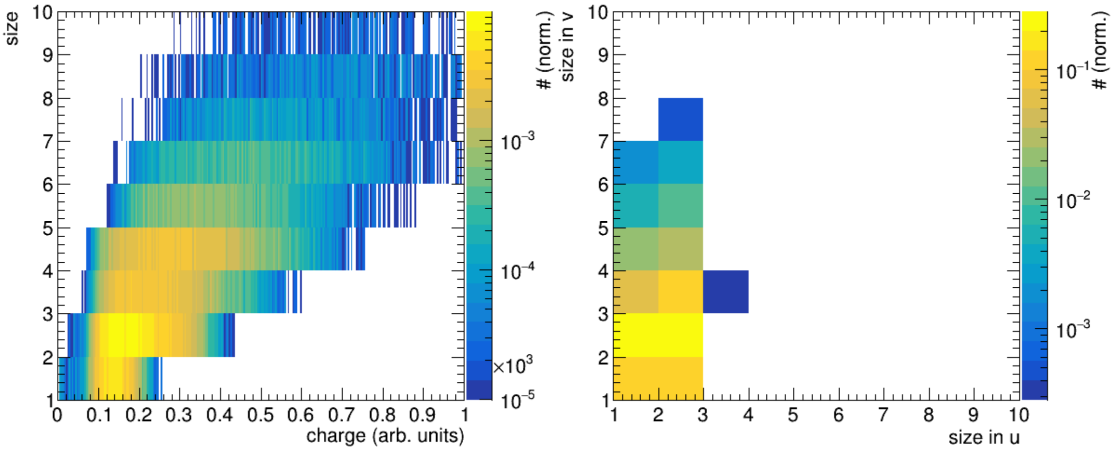
(a) (b)

Figure 5.12: Cluster size and cluster size in v for anti-deuterons as a function of momentum. Note that the z axis is logarithmic for the total cluster size. The cluster size increases with decreasing momentum. A similar behavior is observable for the cluster charge (see Fig. 5.1). The number of clusters with a single pixel in v direction declines monotonously with decreasing momentum.

same distribution for anti-deuterons in the 2 - 3 GeV momentum range, shown in Fig. 5.14b, only differs slightly from the low-momentum particles. While the maximum for the low-momentum particles is at a size of 2 in v and independent of u, the high-momentum anti-deuterons are most likely to generate clusters with size 1 in u and size 1 or 2 in v. The shift is ascribable to the lower energy deposition, which has been discussed above. The weak impact of particle momentum on the two-dimensional cluster size in u/v is an effect of the threshold, which needs to be surpassed to increase the cluster size by one unit (see previous Section).

Background particles, in contrast to the anti-deuteron distributions, cover the entire parameter region shown in the plot. The source of the background particles has a significant impact on the size and shape of the cluster making it difficult to entangle the influence of energy deposit and charge sharing from geometrical effects like the angle under which they enter the IR.

In Appendix F the entire set of two dimensional distributions is shown. All distributions for anti-deuterons and background differ in several aspects. Reasons for this difference are generally relatable to the higher energy deposit of anti-deuterons



(a) (b)

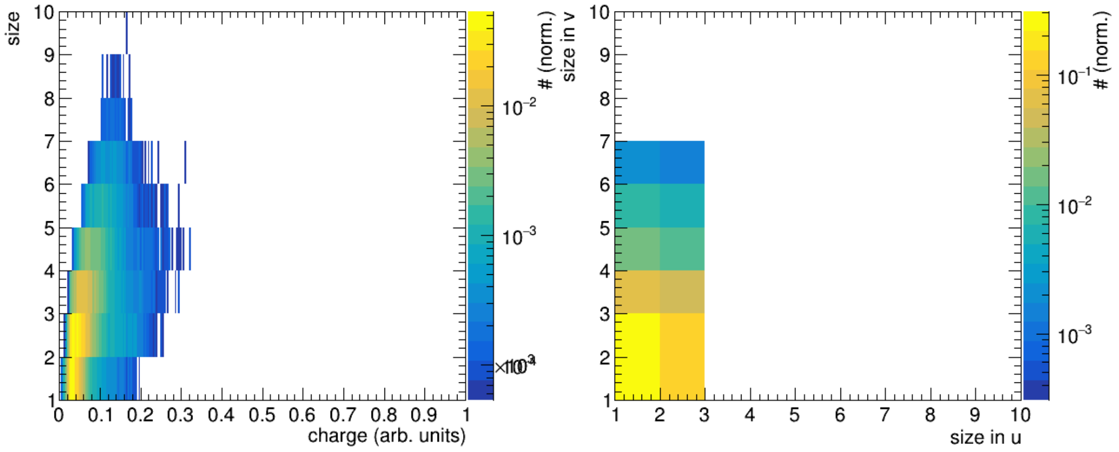
Figure 5.13: Total custer charge against size and size in u against size in v for anti-deuterons with $p < 1\,\text{GeV}$. Small high-charge and large low-charge clusters are missing from the distribution, whose shape resembles a band stretching along the diagonal of the plot. While the size in u is confined to a few values, the size in v extends up so seven.

and the off-vertex source of background particles.

5.3 Pixel Charge Distribution

The average pixel charge distribution in a 5 x 5 pixel matrix around the seed charge of a cluster for $p < 1\,\text{GeV}$ (left-hand side) and $p = 2$ -3 GeV (right-hand side) anti-deuterons is depicted in Fig. 5.16. Both pixel charge distributions are symmetric with respect to a reflection in u/v direction at the origin. Additionally, the inner 3 x 3 matrix around the seed charge is point symmetric making the distribution invariant under exchange of u and v. Considering the full 5 x 5 matrix the pixel charges in v direction remain higher with increasing distance to the seed charge than in u direction in accordance with the higher mean size in v.

Both distributions exhibit a cross-like shape indicating that charge sharing along the coordinate axes is more likely compared to diffusion to diagonal cells due to the larger interface. The u/v symmetry in the inner part of the pixel charge



(a)

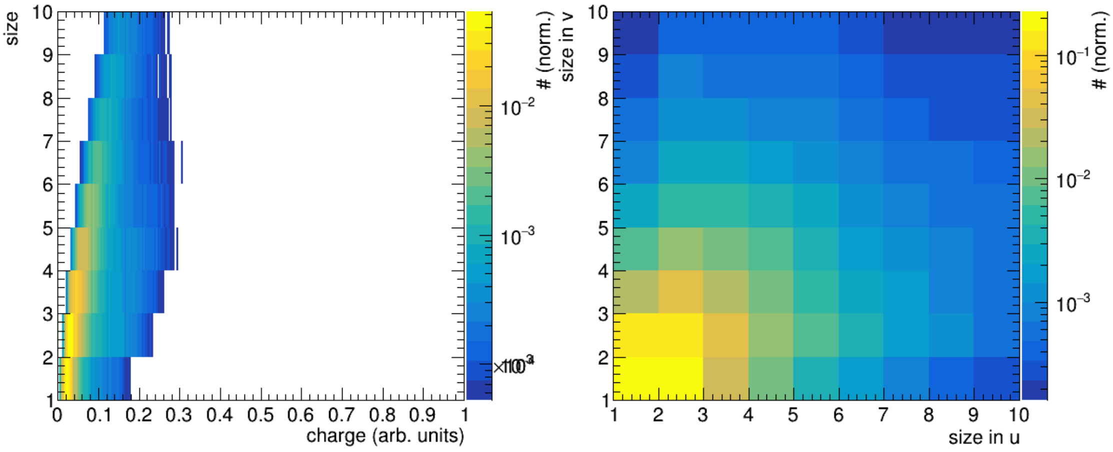
(b)

Figure 5.14: Total custer charge against size and size in u against size in v for anti-deuterons with $2\,\text{GeV} < p < 3\,\text{GeV}$. High-charge clusters are missing altogether. Compared to the low-momentum anti-deuterons, the slope of the distribution is steeper. Small cluster charges are more common. The size in u against size in v distribution appears similar to the low-momentum anti-deuterons except for a slight shift in the maximum towards clusters with size 1 in v direction.

distribution suggests the dominance of charge sharing effects over the boost in v direction, which is responsible for the overall larger size in v. Nevertheless, a comparison between u and v direction should be treated very cautiously due to the different extensions of the DEPFET sensors in both spatial directions, which are even subject to variation depending on the location of the sensor in the PXD (see discussion in Section 4.1).

The average charge per pixel around the seed pixel is higher for low-momentum anti-deuterons compared to the high-momentum particles confirming the increase in charge sharing with high energy deposition once again.

The pixel charge distribution for background particles, which is shown in Fig. 5.17, displays total symmetry with respect to u and v direction due to the lack of forward boost. Charge sharing effects manifest themselves in the cross-shape charge distribution as explained above. The amount of charge per pixel is smaller compared to both anti-deuteron distributions indicating that charge sharing is suppressed due to the lower energy deposition.



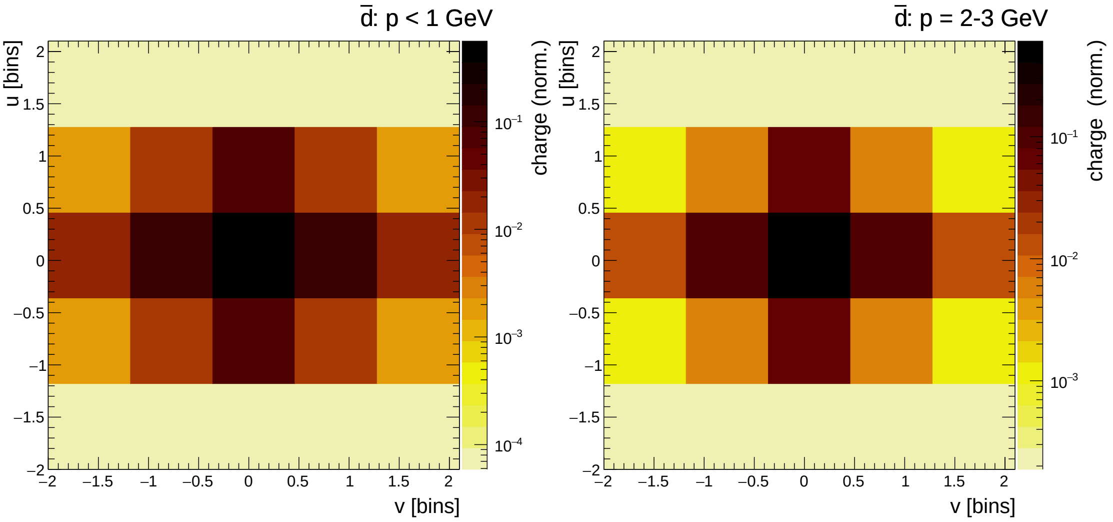
(a)

(b)

Figure 5.15: Total custer charge against size and size in u against size in v for background particles. The charge and size of clusters generated by background is similar to the clusters formed by high-momentum anti-deuterons. The cluster size, however, can reach higher values. Cluster size in u against size in v is not confined to a small parameter region. While small size values are most common, clusters can adopt all size combinations shown in this plot. The distribution is largely symmetric with respect to an exchange of u and v.

5.4 Training Results

Various FFNs and SOMs are trained with different input sets. Anti-deuterons in a small ($p < 1\,\mathrm{GeV}$) and in a wide ($p < 3\,\mathrm{GeV}$) momentum range are trained separately. The training sets consist of 360,000 clusters and the testing sets of 140,000 clusters. For both momentum ranges training is performed with cluster properties and with pixel charge distributions. The ROC curves for a few selected results are compared. At the end of this Section a tabulated overview over all results is shown. The training process itself is described only once to give an example. The other training sequences do not differ noticeably.



(a) (b)

Figure 5.16: Average pixel charge distibution for anti-deuterons with $p < 1\,\mathrm{GeV}$ (left-hand side) and $2\,\mathrm{GeV} < p < 3\,\mathrm{GeV}$ (right-hand side). Both distributions exhibit a cross-like shape. They are point-symmetric in the inner 3 x 3 matrix. If the full 5 x 5 matrix is considered, the pixel charge values in v direction are higher. The charge per pixel falls more rapidly with increasing distance to the seed pixel for the high-momentum anti-deuterons.

5.4.1 Anti-Deuterons with $p < 1\,\mathrm{GeV}$

Feed-Forward Neural Network

The output after each training epoch can be represented in a classification distributions similar to the ones shown in Fig. 5.18. The horizontal axis encodes the classification according to the network. The output of the softmax function (see Section 4.2) is used as classifier. A low number corresponds to background-like events, a high number to anti-deuterons. The real background is colored in blue and the real anti-deuterons in orange. A cut on the horizontal axis determines whether an event is classified as background or anti-deuteron. For a cut at 0.5, for instance, all vectors with a softmax value greater than 0.5 are associated with anti-deuterons and below 0.5 with background.

The entire training process is overseen by monitoring the loss function and the accuracy every ten epochs for a training and a testing set. The accuracy is defined

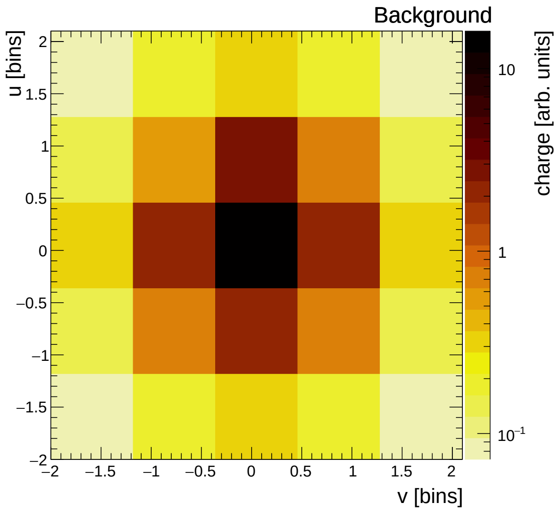
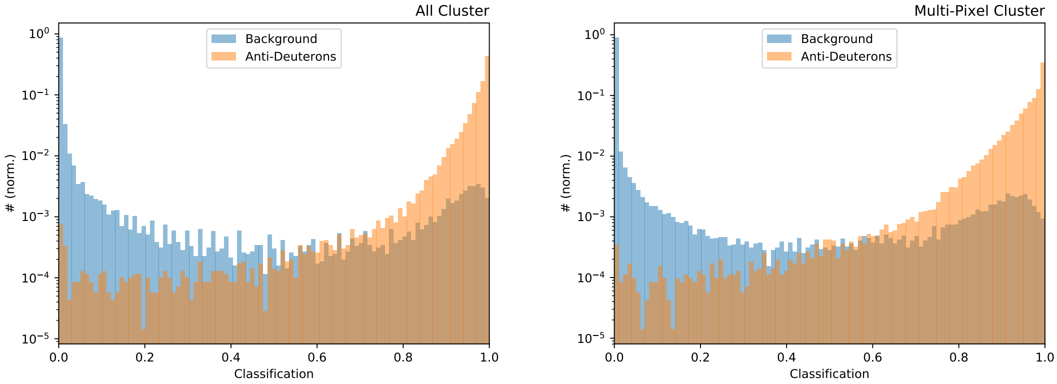


Figure 5.17: Average pixel charge distibution for background. A cross-like shape is observable. The distribution is invariant under an exchange of u and v. The decrease in pixel charge with increasing distance to the origin is stronger compared to the anti-deuteron distributions.

as number of correctly identified events over the number of total events for a fixed cut on the classification axis. The testing set is a collection of input vectors, which are not part of the training set. They serve as an independent validation tool and can help to prevent specialization on the training set (see discussion about overtraining in Appendix B).

In Fig. 5.19 loss and accuracy as a function of epoch for the training of anti-deuterons with sub-GeV momentum against background are depicted. The cut on the classification axis is chosen to be 0.5. The loss function for the training set falls slightly in the first few epochs, but remains constant during the rest of the training. Overall a small decrease in loss with increasing epoch is observable for the testing set as well, however, the loss function is significantly more volatile and exhibits an offset compared to the training set.

The accuracy function reflects the volatility of the loss function as well. Contrary to the loss function, no offset between training and testing set is observable. Both functions appear to saturate after ∼400 epochs at an accuracy slightly above 97.5%. However, a sub-percent increase in accuracy in the remaining 600 epochs is still noticeable. The training is truncated after 1000 epochs. This cut-off is chosen to achieve high accuracy through a large number of training epochs while maintaining



(a) (b)

Figure 5.18: Classifier of FFN trained to separate anti-deuterons with $p < 1\,\mathrm{GeV}$ from background for all clusters (left-hand side) and only multi-pixel cluster (right-hand side). A high classification value corresponds to anti-deuterons, a low value to background. The true anti-deuterons are colored in orange, the true background in blue. In the anti-deuteron-like classification regime the background is suppressed by two to three orders of magnitude. The background distribution falls steeply for increasing classification values and rises again on the far-right side of the horizontal axis. The rise is less pronounced for multi-pixel cluster.

a reasonably short training duration (see Appendix B).

The trained classifier for the separation of anti-deuterons with sub-GeV momentum and background is shown in Fig. 5.18a. For classification values close to one the real anti-deuteron curve dominates, the background is suppressed by two to three orders of magnitude. Correspondingly, low classification values are adopted predominantly by real background vectors, the anti-deuterons are suppressed by nearly four orders of magnitude. While the anti-deuteron distribution falls monotonously with decreasing classification values, the background distribution exhibits an increase of nearly one order of magnitude at the right-most part of the plot. This behavior suggests a systematic misclassification of certain types of background events.

The identification power of the FFN is limited since there are cluster types which are not unique to anti-deuterons or to background. If these clusters appear in the training set once with an anti-deuteron label and a second time with a background label, the training of the network is severely inhibited. Particularly single pixel clusters are affected from this phenomenon. The activation of only a sin-

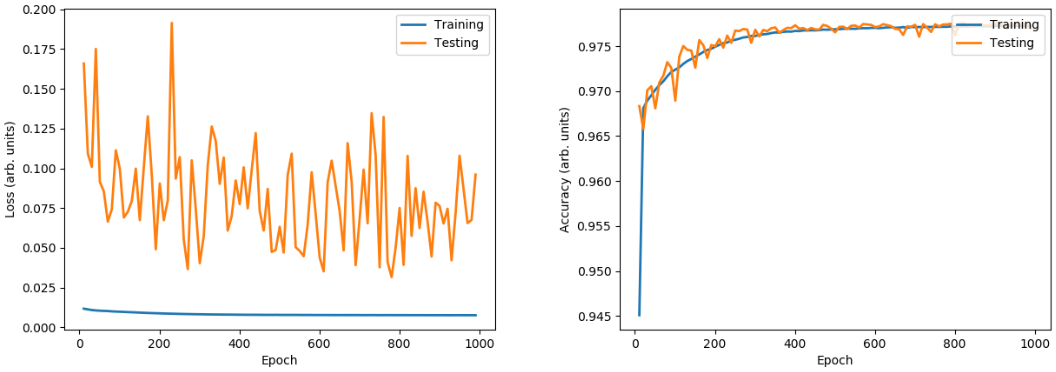


Figure 5.19: Loss and accuracy as a function of training epoch for training and testing set. The loss function for the training set falls with increasing epoch, the loss function for the testing set is more volatile and exhibits an offset compared to the training set. The volatility for the testing set also manifests itself in the accuracy function, but no offset is observable. The accuracy for the training rises steeply and begins to saturate after ∼400 epochs.

gle pixel reduces the number of independent variables required to describe the cluster unambiguously to one value: the pixel charge. The limited amount of information enhances the risk of encountering cluster types, which are generated by anti-deuterons as well as by background particles. When single pixel clusters are excluded from the training and testing set, the classification curve is slightly modified as can be seen in Fig. 5.18b. Note the different scaling of the vertical axis. The background distribution falls again for classification values close to one supporting the assumption that single pixel clusters can impair the performance of the neural network for high classification values.

For low-momentum anti-deuterons the occurrence of single-pixel cluster is less problematic compared to the identification of anti-deuterons carrying high momenta. On average, the high energy deposition of the low-momentum anti-deuterons is still represented by a single pixel. In Fig. 5.20 the minimum pixel charge in a cluster against momentum of the particle generating this cluster can be seen. Three bands characterized by a high density are distinguishable. They correspond to clusters with a specific size. Single pixel clusters are not subject to charge sharing, the entire charge deposited in the pixel is registered in one cell leading, on average, to a higher minimum charge compared to two pixel clusters. The minimum charge pixel in a two pixel cluster is usually the pixel cell, who receives the diffusing charge carriers from the adjacent pixel cell, on which the incident particle impinges. Consequently, the highest band belongs to single pixel clusters, the second one to two pixel clusters and the third one to clusters with size values above two. The resolution of the plot does not allow to distinguish additional cluster sizes in the third band. The density of the highest two bands decreases with decreasing momentum due to the simultaneous increase in cluster size with decreasing momentum (see Section 5.2.1). Below 1 GeV the single pixel clusters are characterized by an average minimum charge value (or equivalently total charge value) above 150. The overlap with the background distribution is therefore small and the classification of single pixel clusters for low-momentum particles is not significantly hindered by the limited amount of cluster information. With increasing momentum the single pixel charge falls sharply making the clusters appear more background-like. The classification of single-pixel clusters in a wider momentum range will be discussed in more detail in the next Section.

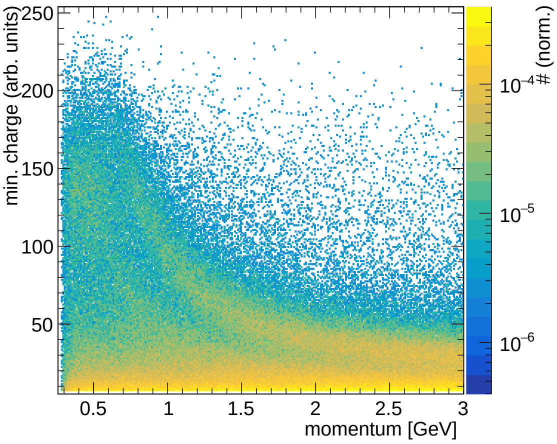
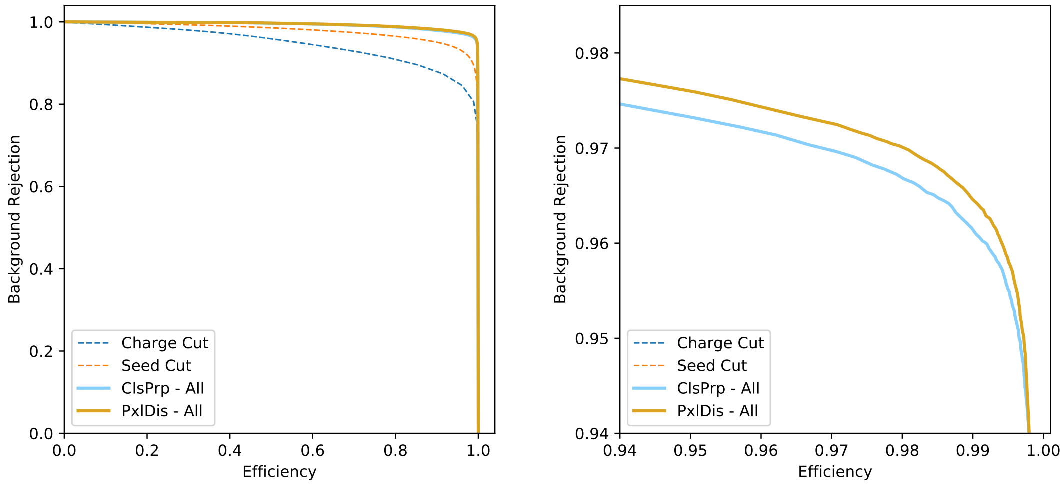


Figure 5.20: Minimum pixel charge as a function of momentum carried by the incident particle. Three bands are distinguishable. The density of the two higher bands decreases with decreasing momentum.

ROC curves can be generated from the distribution in Fig. 5.18 by varying the classification cut along the horizontal axis. A cut on the far-left side corresponds to little background suppression and high efficiency. Performing the cut on the opposing side leads to strong background suppression but low efficiency. The classification curve is sampled by 100 equidistant cuts. The resulting ROC curves are shown in Fig. 5.21. The dotted lines represent cuts on the cluster charge and the seed charge. The neural networks trained with cluster properties (ClsPrp) or pixel charge distributions (PxlDis) both perform significantly better than the one dimensional cuts. A close-up of the upper right part of the curve, as shown in Fig 5.21b, reveals that training with the pixel charge distribution provides better



(a)

(b)

Figure 5.21: ROC curves for Feed-Forward Networks trained with cluster properties and pixel charge distributions. ROC curves for a cut on the cluster charge and the seed charge are shown as well. Both neural networks deliver better results than the one dimensional cuts. On the right-hand side a close-up of the upper right corner is shown. The FFN trained with pixel charge distributions performs slightly better.

results below efficiencies of ∼99.8%. The charge and seed cut are not visible in Fig 5.21b since the curves lie below the range of the axes.

The better performance of the pixel charge distribution can be attributed to the higher amount of information encoded in the input vectors. The curves meet at an efficiency of about 0.998 suggesting that the type of input vector set is not relevant at lower background rejection values. This behavior can be partly attributed to the occurrence of single pixel cluster, for which the information content is the same in the pixel charge distribution and cluster properties. In Fig. 5.22 the difference in background rejection of both input vector sets is shown as a function of efficiency for all clusters and multi-pixel clusters. The efficiency range has been chosen in such a way as to match Fig. 5.21b. For multi-pixel clusters the difference in performance is significantly greater verifying the impact of single pixel clusters on the training outcome.

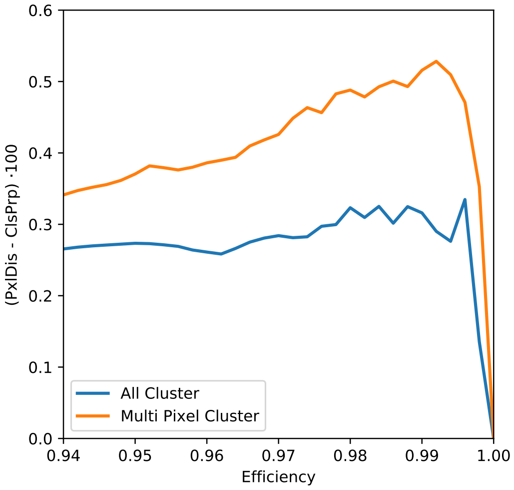
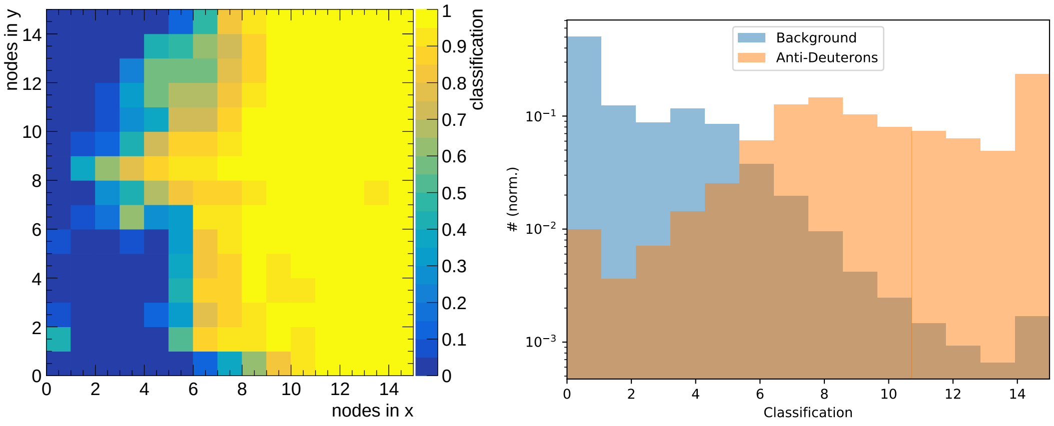


Figure 5.22: Difference in background rejection for FFNs trained with pixel charge distribution and cluster properties as a function of efficiency. If training is performed with all clusters the difference is less pronounced compared to FFNs solely trained with multi-pixel cluster.

Self-Organizing Map

The same training sets introduced in the previous Section are used to train SOMs. The classification pattern of the trained map is depicted in Fig. 5.23a. The coordinate axes represent nodes and the z axis encodes the classification according to the SOM. Zero corresponds to background clusters and one to anti-deuteron clusters as before. The plot is generated with the testing set by monitoring the winning node for each input vector. If the input vector stems from a background event, the winning node is defined to be *background-like* and the z value is shifted closer to one. For anti-deuteron events the winning nodes are called *signal-like* and the z value is shifted close to zero.

The trained map in Fig. 5.23a shows the average of this classification scheme. While background-like nodes cluster in a band at low x values, nodes at high x values are signal-like. A transitional region is observable around $x = 5$. It is not homogeneous in y direction. For further evaluation a one-dimensional cut along the x direction is made. This choice may undermine the inhomogeneity in y, but allows for an appropriate comparison with the FFN results. The result is depicted in Fig. 5.23b. Instead of a softmax function, the classifier is now represented by



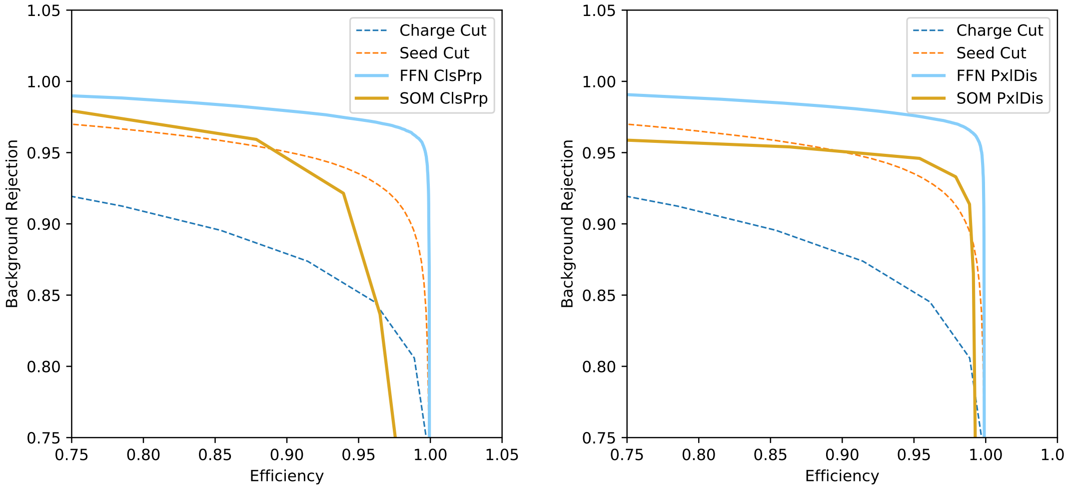
(a)

(b)

Figure 5.23: Classification map of SOM trained with cluster properties of anti-deuterons with sub-GeV momentum (left-hand side) and projection onto the x axis (right-hand side). High classification values correspond to anti-deuterons clusters and low ones to background clusters. While anti-deuteron events form a cluster in a band-like region at high node values in x, background events tend to occupy nodes with low x values. The transition region is not homogeneous in y direction. A projection on the x axis allows to comare the SOM results with the FFNs. The background is suppressed by two to three orders of magnitude in the anti-deuteron-like regime.

the nodes in x direction leading to a strong quantization of the vertical axis. In the anti-deuteron region, corresponding to high classification numbers, the background is suppressed by two to three orders of magnitude similar to the results of the FNN. In the background region, on the other hand, the suppression of anti-deuterons is less strong compared to the FFN. Similar to the classification distributions of the FFNs (see Fig 5.18) the true background first falls in the anti-deuteron-like region, but exhibits an increase on the far right of the horizontal axis. The observation of this phenomenon for both types of neural networks indicates that the reason lies in the input vectors themselves and not in the network architecture.

The classification distributions in Fig. 5.23b allow the generation of a ROC curve, which is depicted in Fig. 5.24a. The one dimensional cuts on cluster and seed charge as well as the result of the FFN from the previous Section are also shown. The ROC curve for the SOM is not smooth due to the quantization effects explained above. The overall performance of the SOM is several percent worse com-

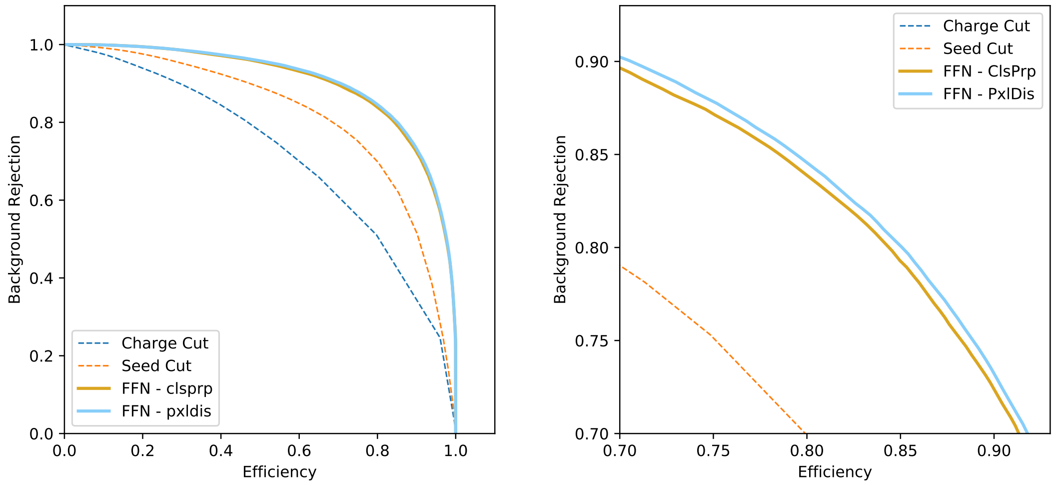


(a) (b)

Figure 5.24: ROC curves for Self-Organizing Maps, Feed-Forward Networks and one dimensional cuts on total cluster charge and seed charge. ROC curves for the SOM lie below the FFNs' curve for cluster properties as well as pixel charge distributions. In certain efficiency ranges the background rejection of the SOM is worse than the one dimensional cuts.

pared to the FFN. For efficiencies above ∼90% the SOM performs even worse than the one dimensional cuts. A SOM trained on pixel charge distributions behaves similarly. The ROC curve, shown in Fig. 5.24b, reveals that the SOM's performance is inferior to the seed charge cut in most of the efficiency regime depicted in this plot.

It should be kept in mind that these results are only valid for an evaluation of the SOM along one axis and the network architecture chosen for this thesis. The results do not allow to give a general statement about the performance of SOMs compared to FFNs. It should be noted that SOMs are originally not designed for classification purposes. While the identification of HIP clusters turned out to be feasible nonetheless, the application of the algorithm in this manner is not optimal. Additionally, a direct comparison between supervised and unsupervised classification methods is disputable as well. The two methods operate under different conditions: while a supervised neural network has access to the true classification information, an unsupervised neural network has not. This information imbalance has a crucial impact on the maximum performance, which is achievable.



(a)

(b)

Figure 5.25: ROC curves for Feed-Forward Networks trained with anti-deuterons with $p < 3\,\mathrm{GeV}$. The background rejection for a given efficiency is worse compared to the sub-GeV anti-deuterons. A close-up of the upper right part of the curves (right-hand side) reveals that training with the pixel charge distribution delivers better results. Both FFNs have a higher background rejection than the one dimensional cuts.

5.4.2 Anti-Deuterons with $p < 3\,\mathrm{GeV}$

Feed-Forward Neural Network

FFNs are trained with the cluster properties and pixel charge distributions of clusters generated by anti-deuterons in the momentum range 0 to 3 GeV. In the cluster property distributions, presented in Section 5.2.1, a strong overlap between anti-deuterons with $2\,\mathrm{GeV} < p < 3\,\mathrm{GeV}$ and background was observable. The performance of FFNs trained with the wider momentum range is therefore expected to be worse compared to the anti-deuterons with sub-GeV momentum. This assumption is confirmed by the ROC curves for FFNs trained with $p < 3\,\mathrm{GeV}$ anti-deuterons in Fig. 5.25a. For a given efficiency the background rejection is worse compared to the results from the previous Section. The dashed ROC curves representing one-dimensional cuts lie well below the FNN curves. A close-up of the upper right part of the curves is depicted in Fig. 5.25b. Training with the pixel charge distribution delivers better results once again supporting the assumption that the higher information content encoded in the charge distributions improves the FNN's performance.

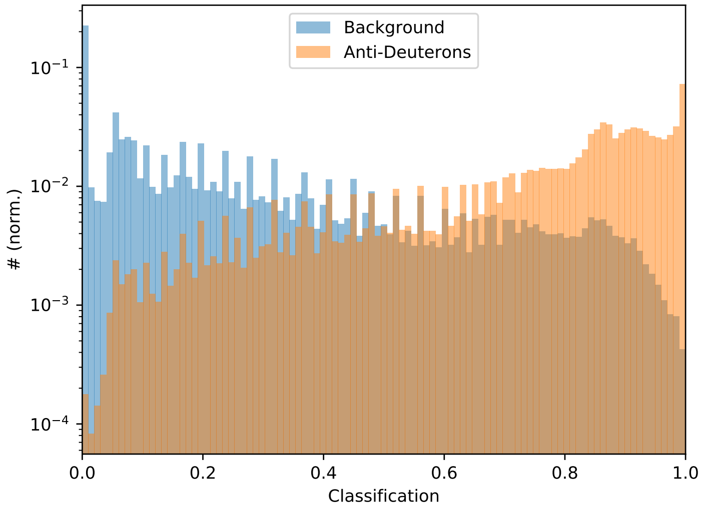
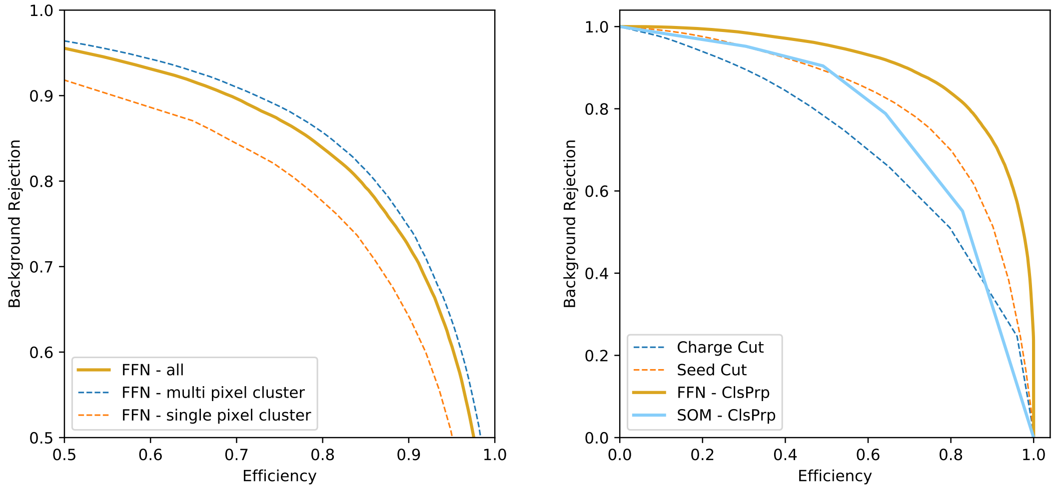


Figure 5.26: Classification distribution for anti-deuterons with $p < 3\,\mathrm{GeV}$. In contrast to the classification distributions for the low-momentum anti-deuterons the background/anti-deuteron distributions decline less steeply for increasing/decreasing classification values. The decrease of both distributions is monotonous.

The classifier for the FFN trained with pixel charge distributions is shown in Fig. 5.26. In contrast to the classifier distributions for training with the low-momentum anti-deuterons, the background distribution falls less steeply with increasing classifier values. The same holds true for the decline of anti-deuteron distribution for low classification values. Consequently, there is an extensive range of classification values, which are adopted by both particle species with nearly the same probability. There is no increase in background for classification values close to 1, which was observable for the low-momentum anti-deuterons. The fraction of background particles adopting values close to one is in the order of 10^{-3} for the FFNs presented with clusters from $p < 1\,\mathrm{GeV}$ anti-deuteorns. For the $p < 3\,\mathrm{GeV}$ anti-deuteron clusters the fraction is the same indicating that the behavior at the edges of the classification axis are dominated by the low-momentum anti-deuteron clusters. The higher-momentum anti-deuterons generate clusters, which are more background-like in terms of cluster properties and pixel charge distribution. A clear separation from background is therefore not feasible and classification values are farther away from zero and one. A discussion about the spikes in the



(a) (b)

Figure 5.27: ROC curves for Feed-Forward Networks and Self-Organizing Maps trained with clusters generated by anti-deuterons with $p < 3\,\mathrm{GeV}$. A better performance of the FFN is achieved, if the training set contains only multi-pixel clusters. The SOM trained with cluster properties performs worse than the FFN and the cut on the seed charge.

classification distribution is given in Appendix B.

The separation power is impaired by single pixel clusters as illustrated in Fig. 5.27a. The limited amount of information in a single pixel leads to a decline in separation power. The cluster charge distribution of anti-deuterons with momentum 2 -3 GeV revealed a strong overlap for background and anti-deuteron distribution. Single-pixel clusters with a charge value in this overlap region are unidentifiable. If single-pixel clusters are excluded from the training set, the FFN delivers a better performance compared to training with all cluster types.

While the ratio of single pixel clusters to all clusters is ∼13% for anti deuterons with sub-GeV momentum, the percentage increases to ∼23% for anti-deuterons in the momentum range 0 - 3 GeV. The higher number of single pixel clusters in the training set is one reason for the decline in separation power with higher anti-deuteron momentum.

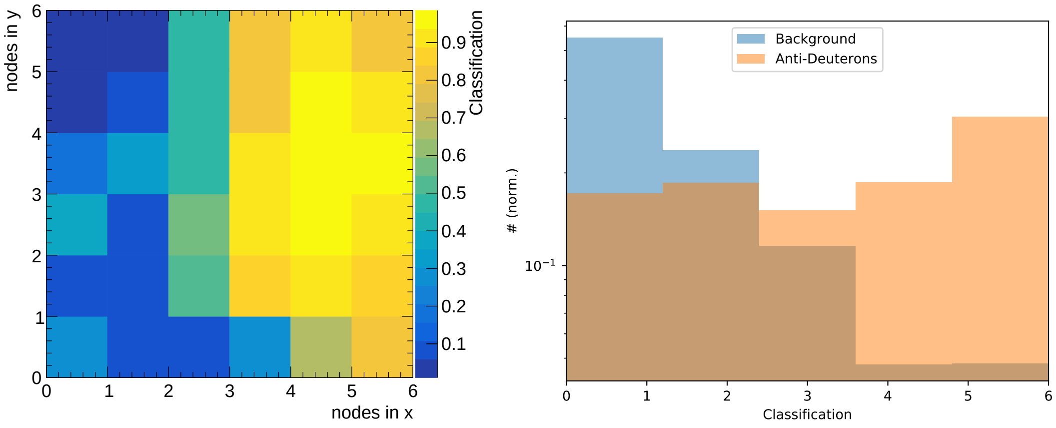
Self-Organizing Maps

SOMs are trained as explained above. The results can be compared to the FFNs by considering the respective ROC curves shown in Fig. 5.27b. Both networks are trained with the same training set, which consists of cluster properties from 0-3 GeV anti-deuterons and background. The background rejection achieved with the SOM is significantly lower compared to the FFN and for most efficiency values also lower than the cut on the seed charge. The network architecture of the SOM had to be modified with respect to the generic network introduced in Section 4.2.2. With a 15 x 15 node space convergence is not achievable. A drastic reduction of the node space to 6 x 6 nodes resolves the convergence issue, but introduces more pronounced quantization effects. The width of the Gaussian kernel is adjusted to 2 to account for the small map size.

The trained map and the one dimensional representation of the classifier are depicted in Fig. 5.28. The small extension of the map compresses the anti-deuteron-like region and the background-like region to a low number of nodes leading to a transitional region, which is proportionally larger compared to the SOM presented in the previous Section. The background suppression in the anti-deuteron region is better than the anti-deuteron suppression in the background region.

The reason for the smaller number of nodes required to reach convergence could be attributed to an inherently smaller number of independent features in the input vector set.

The linear correlations between all cluster properties for $p < 1\,\mathrm{GeV}$ (left-hand side) and $p < 3\,\mathrm{GeV}$ (right-hand side) anti-deuterons are shown in Fig. 5.29. The linear correlations between charge values (e.g. between total cluster charge and seed charge or total cluster charge and minimum pixel charge) exhibit a strong difference between the two momentum regimes. The *Seed - Charge* entry, for instance, is 56 for the sub-GeV anti-deuterons and 81 for the larger momentum range. On average, the low-momentum anti-deuterons liberate more charge carriers in the silicon sensors increasing the probability of charge sharing to neighboring pixel cells. The ratio of seed charge to total cluster charge therefore tends to decrease leading to a decrease in linear correlation between the two variables. The simultaneous increase in cluster size, however, is not reflected in the linear correlation matrix. The reason is explainable on the basis of the cluster property distributions shown in Section 5.2.1. While the cluster charge is highly sensitive to the anti-deuteron momentum, the differences in cluster size distributions are small due to the quantized nature of the size variable. Consequently, a low-momentum anti-deuteron



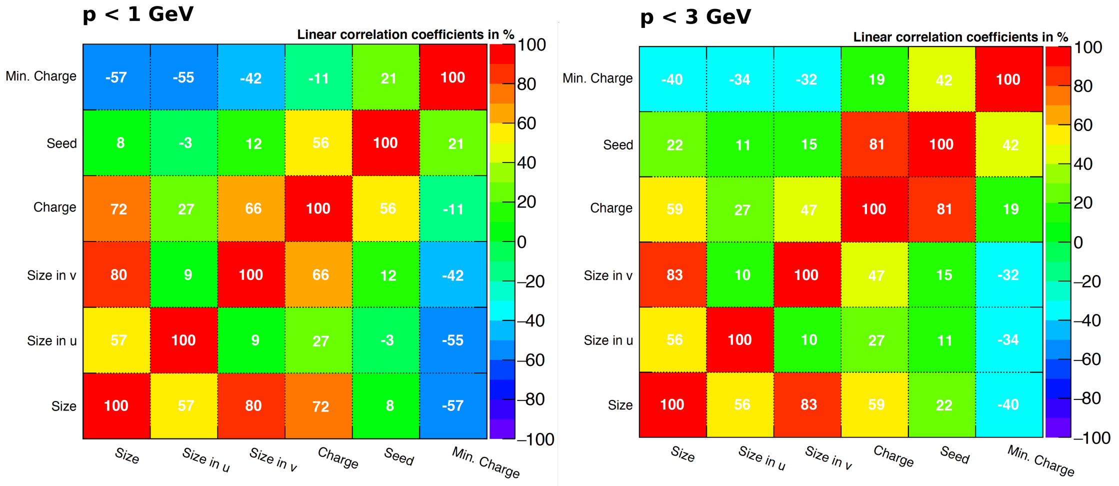
(a)

(b)

Figure 5.28: Classification map of Self-Organizing Map trained with cluster properties of anti-deuterons with $p < 3\,\mathrm{GeV}$ and projection of the map onto the x axis. The node space is reduced from 15 x15 nodes to 6 x 6 nodes to achieve convergence. High classification values correspond to anti-deuteron clusters, low values to background clusters. The transitional region between anti-deuteron-like nodes and background-like nodes is proportionally larger compared to the SOM presented in the previous Section.

depositing a high amount of energy will promote sharing of charges to different pixel cells leading to a decrease in linear correlation between cluster charge, seed charge and minimum charge since seed and minimum charge contribute less to the total cluster charge. At the same time, a threshold has to be surpassed to invoke an increase in cluster size. The impact of the high energy deposition on the cluster size variables is therefore negligibly small.

While an increase in linear correlations between some variables is indeed observable, it is scarcely able to validate the drastic reduction of SOM nodes. Higher-order correlations may contribute to the decrease in number of nodes as well. Additionally, the higher number of single-pixel clusters in the high momentum range could reduce the number of independent variables required to capture the features of all input vectors. A reliable statement about this issue would require the fine-tuning of the neural network architecture first.



p < 1 GeV

	Size	Size in u	Size in v	Charge	Seed	Min. Charge
Min. Charge	-57	-55	-42	-11	21	100
Seed	8	-3	12	56	100	21
Charge	72	27	66	100	56	-11
Size in v	80	9	100	66	12	-42
Size in u	57	100	9	27	-3	-55
Size	100	57	80	72	8	-57

p < 3 GeV

	Size	Size in u	Size in v	Charge	Seed	Min. Charge
Min. Charge	-40	-34	-32	19	42	100
Seed	22	11	15	81	100	42
Charge	59	27	47	100	81	19
Size in v	83	10	100	47	15	-32
Size in u	56	100	10	27	11	-34
Size	100	56	83	59	22	-40

(a) (b)

Figure 5.29: Linear correlation matrices for anti-deuterons with $p < 1\,\text{GeV}$ (left-hand side) and $p < 3\,\text{GeV}$ (right-hand side). Correlations between charge variables (e.g. total cluster charge and seed charge) are different for the two momentum regimes: for the anti-deuteron sample with $p < 3\,\text{GeV}$ the linear correlations between two charge parameters are generally greater compared to the sample of $p < 1\,\text{GeV}$ anti-deuterons.

5.4.3 Summary

All training results for the FFNs trained with anti-deuteron distributions are summarized in Table 5.1. In the forth and fifth columns the accuracy for training and testing set at an classification cut of 0.5 are listed. In the last column the Area Under Curve (AUC) denotes the integral of the ROC curve in percent.

There are no discrepancies between training and testing results in any entry. The differences of max. $\pm 0.1\%$ can be attributed to statistical fluctuations and rounding errors. Likewise, a high accuracy at a certain point on the ROC curve generally implies a large area under the curve.

For anti-deuterons in the momentum range $p < 3\,\text{GeV}$ training with the pixel charge distribution delivers slightly better results compared to the cluster properties. The impact of single pixel clusters on the training is observable by the better performance of the FFN when single pixel clusters are excluded. If training is performed with single pixel clusters only the result is significantly worse.

Table 5.1: Accuracy of anti-deuteron FFNs for a cut at 0.5 on the classifier axis. The accuracy is computed after 1000 training epochs for the training and the testing set. Area Under Curve (AUC) denotes the integral of the ROC curve.

$p_{\max}$ [GeV]	Input set	Cluster types	Training [%]	Testing [%]	AUC [%]
1	ClsPrp	all	97.5	97.4	99.24
1	ClsPrp	multi pixel	97.3	97.4	99.26
1	ClsPrp	single pixel	97.7	97.7	98.82
3	ClsPrp	all	82.2	82.1	90.19
3	ClsPrp	multi pixel	83.1	83.2	91.34
3	ClsPrp	single pixel	78.7	78.7	85.94
1	PxlDis	all	97.7	97.7	99.27
1	PxlDis	multi pixel	97.7	97.7	99.32
1	PxlDis	single pixel	97.7	97.7	98.81
3	PxlDis	all	82.6	82.6	90.58
3	PxlDis	multi pixel	83.6	83.6	91.70
3	PxlDis	single pixel	78.7	78.7	86.09

For low-momentum anti-deuterons classification is facilitated due to the higher energy deposition. The AUC shows that the FFN performs best for multi pixel cluster. The accuracy for training and testing set at a fixed softmax value does not display the same result. At certain points on the ROC curve the FFN trained with single pixel clusters is able to outperform the other two networks.

While there is a significant impact of single pixel clusters on accuracy and AUC for the FFNs trained with clusters generated by $p < 3\,\mathrm{GeV}$ anti-deuterons, the impact on the FFNs trained with $p < 1\,\mathrm{GeV}$ anti-deuteron clusters is less pronounced. For low-momentum anti-deuterons even single pixel clusters tend to possess a characteristically high cluster/seed charge, which is uncommon for background clusters. Therefore, the identification power of a neural network trained on single pixel clusters stays nearly unaltered compared to training with all cluster sizes in contrast to the high-momentum anti-deuterons, where it decreases strongly. The same statements hold true if the pixel charge distributions instead of cluster properties are used as input vectors.

A longer training duration can alter the results presented in Table 5.1. For a discussion the reader is referred to Appendix B.

To put the accuracy achieved with FFNs into perspective it is worth mentioning that background levels at the PXD are estimated to be approximately four orders of magnitude higher than signal per event. It is therefore reasonable to place a classification cut on the far right of the classification axis. This would translate into a high background rejection while trading off efficiency.

CHAPTER 6

MAGNETIC MONOPOLES

A magnetic monopole is defined as an object with isolated magnetic charge. Magnetic monopoles appear in various models of particle physics beyond the SM. Inspired by predictions from theorists such as Dirac, Schwinger, 't Hooft - to name only a few - experimental research set out to detect these obscure particles. Even though the theoretical predictions appear promising, they are overshadowed by a lack of experimental evidence.

The following Chapter provides an overview about Dirac magnetic monopoles. In particular, the experimental signature of magnetically charged particles is investigated. A strong focus is placed on the energy loss of monopoles. Past and ongoing monopole studies are introduced and the search strategy at Belle II is outlined. Subsequently, the simulation of magnetic monopoles with `basf2` are presented. Analogous to the previous Chapter, the range in the detector and the characteristics of monopole clusters are analyzed. At the end of this Chapter the application of neural networks for the identification monopole clusters is presented.

6.1 Magnetic Monopoles in Modern Theories

Classical electromagnetism as well as the interaction of charged and electromagnetic fields are fully described by the Maxwell equations combined with the Lorentz force law and the Newton equations of motion. The slight asymmetry of the Maxwell equations due to the existence of electric but the absence of magnetic charges follows no underlying principle other than empirical evidence.

Under the assumption that magnetic charges exist, the Maxwell equations take the form:

$$\nabla\mathbf{D} = 4\pi\rho_e \tag{6.1.1}$$

$$\nabla\mathbf{B} = 4\pi\rho_m \tag{6.1.2}$$

$$-\nabla\times\mathbf{E} = \frac{1}{c}\frac{\partial}{\partial t}\mathbf{B} + \frac{4\pi}{c}\mathbf{j}_m \tag{6.1.3}$$

$$\nabla\times\mathbf{H} = \frac{1}{c}\frac{\partial}{\partial t}\mathbf{D} + \frac{4\pi}{c}\mathbf{j}_e. \tag{6.1.4}$$

To distinguish the fully symmetric Maxwell equations from the conventional ones, they receive the name *dual Maxwell equations.* In accordance with the common nomenclature **E** is the electric field, **D** is the displacement field, **B** the magnetic induction, **H** the magnetic field, ρ_e/ρ_m is the electric/magnetic charge density and j_e/j_m the electric/magnetic current density [86].

Despite the novel formulation of the Maxwell equations, it can be shown that the conventional Maxwell equations still hold for classical electrodynamics [87]. Theoretical models involving the existence of magnetic monopoles are generally an extension and not a reformulation of electrodynamics.

The first prominent advocate for magnetic monopoles is believed to be Paul. A. M Dirac. The monopoles appearing in his theory are therefore called *classical* or *Dirac-type* magnetic monopoles to distinguish them from monopoles arising in other theories. According to Dirac the existence of magnetic monopoles implies the quantization of the electric charge [87–89]. As the electric charge is in fact quantized, the existence of magnetic monopoles is permitted, but not required. The quantization condition was discovered, when Dirac established a quantum mechanical treatment of an electron in the magnetic field of a magnetic monopole [87]. A semi-classical treatment of his theory is easily derived and can be found in [90] for instance. As a consequence of the Dirac condition the charge of a magnetic monopole is given by [87]:

$$eg = \frac{n\hbar c}{2} \tag{6.1.5}$$

where e is the electric and g the magnetic charge. n is an arbitrary integer greater than zero. The unit Dirac charge for $n = 1$ is therefore $g = g_D = \hbar c/2e \approx 137e/2 = 68.5\,e$.

In the Dirac theory a rough estimate for the monopole mass yields values of several GeV. The estimation is based on the assumption that the classical electron radius is equal to the classical monopole radius [86].

There are several theories extending and altering the notion of the Dirac quantization condition. Some theories imply the emergence of additional unobserved particles. Fryberger for instance introduces a magnetic field potential distinct from the electric one. As a consequence, his theory predicts the existence of an additional *magnetic photon* [91].

The search for magnetic monopoles gathered momentum in the 1970ies, when 't Hooft [92] and Polyakov [93] proposed that magnetic monopoles appear in GUTs due to the spontaneous symmetry breaking of the unified group. However, they predict monopole masses m in the order of $m \gtrsim 10^{17}$ GeV. Some supersymmetric models expect magnetic monopoles in the intermediate mass range $m \sim 10^5$-10^{12} GeV [86]. Lower mass monopoles in the range $m \sim$ 3-7 TeV were proposed in the theory of electroweak interactions [94].

This thesis focuses primarily on the predictions for magnetic monopoles arising from the classical Dirac theory. When extensions or deviations from this model are considered, it will be mentioned explicitly.

6.1.1 Production

In the Dirac interpretation of magnetic monopoles, the production process in e^+e^--collisions is mediated via a virtual photon. The following Section first introduces a model for the coupling between monopoles and photons as an intermediate step for monopole production. Subsequently, a monpole pair production mechanisms is introduced and the possibility of monopole loops is discussed

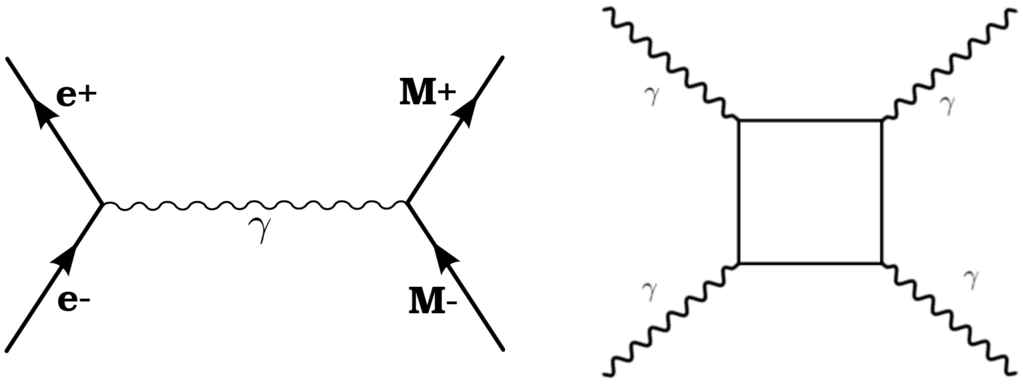


Figure 6.1: Feynman diagram to visualize monopole pair production from e^+e^--annihilation (left-hand side) and Feynman diagram for light-to-light scattering $\gamma\gamma \to \gamma\gamma$ via a magnetic monopole loop (right-hand side). The four-photon interaction is enhanced due to the presence of the monopole loop. No mathematical/physical meaning can be assigned to this Feynman diagram due to the non-perturbative behavior of monopole-photon coupling

Monopole-Photon Coupling

Under the assumption of duality as introduced in Section 6.1 the substitution $e^2 \to (g_D n)^2$ yields an expression for the magnetic coupling constant α_m:

$$\alpha_e = \frac{e^2}{\hbar c} \to \alpha_g = \frac{g_D^2 n^2}{\hbar c}. \tag{6.1.6}$$

With $n = 1$ the coupling constant becomes $\alpha_g \gtrsim 34$, which prohibits a perturbative treatment. Non-perturbative attempts were made in the past [95], however no universally accepted prediction about the coupling between magnetic monopoles and photons can be given yet. Therefore, production models are limited to first-order scenarios. The most popular model is a Drell-Yan-like pair production mechanism.

Production Mechanism

In most publications authors resort to the simple replacement $e \to g\beta$ [96], which suggests a monopole pair-production as depicted in the tree-level Feynman diagram on the left-hand side of Fig. 6.1. Note that the Feynman diagram simply serves to visualize the production process in a familiar manner. It has no physical/mathematical meaning, due to the non-perturbative behavior of the production mechanism.

The reaction $e^+e^- \rightarrow \gamma^* \rightarrow M^+M^-$ involves a virtual photon γ^*, which decays in a magnetic north and south pole. In the absence of a proper field theoretical description, the analogy to the reaction $e^+e^- \rightarrow \gamma^* \rightarrow \mu^+\mu^-$ is used to derive the production cross section. The QED cross section $\sigma(e^+e^- \rightarrow \mu^+\mu^-)$ is given by:

$$\sigma(e^+e^- \rightarrow \mu^+\mu^-) = \frac{4\pi\alpha^2}{3(\sqrt{s})^2} \tag{6.1.7}$$

where $\sqrt{s}$ is the center-of-mass energy. For monopole pair production the QED cross section is altered by replacing the electric coupling constant $\alpha_e \approx 1/137$ with the magnetic one α_m from Eq. 6.1.6. In addition, the cross section is convoluted with a threshold phase space factor and the velocity dependence of the interaction, which yields a factor of β^3 [97], where β is the velocity over the speed of light:

$$\sigma(e^+e^- \rightarrow M^+M^-) = \beta^3 \left(\frac{ng}{e}\right)^2 \frac{4\pi\alpha_e^2}{3(\sqrt{s})^2}. \tag{6.1.8}$$

The cross section ratio between QED pair production and monopole pair production as a function of the monopole mass is shown in Fig. 6.2. The monopole velocity β is obtained with:

$$\beta = \frac{p}{\sqrt{s}/2} = \frac{\sqrt{(\sqrt{s}/2)^2 - m^2}}{\sqrt{s}/2} \tag{6.1.9}$$

where $\sqrt{s} = 10.58\,\text{GeV}$, p is the momentum and m the monopole mass. It is assumed that each of the two pair produced monopoles receive half of the center-of-mass energy $\sqrt{s}/2$ on average.

The dashed gray line in Fig. 6.2 represents the QED cross section with the magnetic instead of the electric coupling constant. It marks the maximum production cross section in the limit of $\beta \rightarrow 1$.

The cross section for monopole pair production is several orders of magnitude higher than the cross section for lepton pair production. It falls rapidly with increasing mass of the monopoles, due to the simultaneous decrease in β.

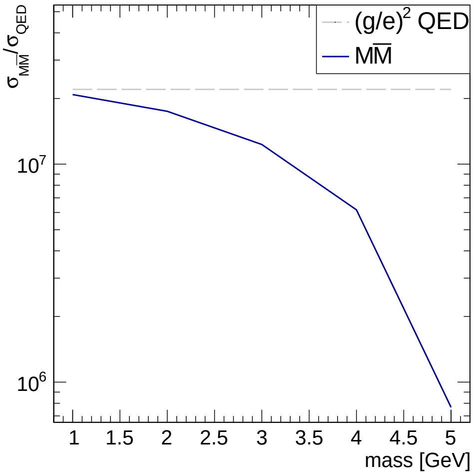


Figure 6.2: Ratio between cross sections for monopole pair production $\sigma_{M^+M^-}$ and conventional QED lepton pair production σ_{QED}. The dashed gray line represents the cross section in the limit $\beta \to 1$. The cross section falls with increasing monopole mass due to the simultaneous decrease in velocity β.

Monopole Loops

Apart from direct production, it is conceivable that magnetic monopoles can be produced in a virtual *monopole loop* [98]. The Feynman diagram for this process is presented on the right-hand side of Fig. 6.1. The virtual monopoles enhance the light-to-light scattering $\gamma\gamma \to \gamma\gamma$ due to their interaction with the photons. As an indirect evidence for magnetic monopoles, the effect is even observable at monopole masses much higher than the available production energy. The existence of monopole loops for monopoles with a mass below 120 GeV is disfavored based on calculations on the contribution of monopole loops to the anomalous magnetic moment of the muon [99].

6.1.2 Energy Loss

Magnetic monopoles interact via their magnetic field $\mathbf{B}_M$ with the magnetic dipole moment μ of ordinary matter [100]. The energy loss of magnetically charged particles in matter is therefore not identical to the loss experienced by electrically charged particles. Depending on the velocity of a magnetic monopole, three energy loss mechanisms have to be differentiated:

Monopoles with $\beta \gtrsim 10^{-1}$

In general, the description for fast magnetically charged particles is much simpler compared to projectiles with low velocities, since the binding energies of electrons to nuclei in matter is small compared to the energy of the projectile. Hence, the electrons can be approximated as free particles. In addition, at fast velocities the impact parameter of the projectile is large. The projectile can therefore interact with a vast amount of particles enabling the usage of the dipole approximation or classical macroscopic electrodynamics [101].

It can be shown that the energy loss $-dE/dx$ of fast monopoles $\beta \gtrsim 10^{-1}$ is describable by a modified Bethe-Bloch formula [101]:

$$-\frac{\mathrm{d}E}{\mathrm{d}x} = C\frac{Z}{A}g^2\left[\ln(\frac{2m_ec^2\beta^2\gamma^2}{I}) - \frac{1}{2} + \frac{k}{2} - \frac{\delta}{2} - B_m\right]\mathrm{MeVg^{-1}cm^2} \qquad (6.1.10)$$

with $C = \frac{e^4}{4\pi\epsilon_0^2 m_u m_e c^2}$ and m_u the unified atomic mass unit, I the mean ionization potential, δ the density effect correction, k the QED correction and B_m the Bloch correction.

The density correction has to be applied for particles with a sufficiently high velocity to account for the polarization forming in the medium. For $\beta\gamma \ll 10$ the density correction can be neglected [102]. The Kazama, Yang and Goldhaber cross-section correction and the Bloch correction are taken from Derkaoui et al. [103]. For $g = g_D$ they are given by $K(|g|) = 0.406$ and $B(|g|) = 0.248$.

For high velocities β the energy loss is approximately constant [102]:

$$\frac{\mathrm{d}E}{\mathrm{d}x} \approx n^2(\frac{137}{2})^2\frac{\mathrm{d}E}{\mathrm{d}x} \qquad (6.1.11)$$

The energy loss computed with Eq. 6.1.10 for a magnetic monopole with mass $m = 1\,\mathrm{GeV}$ and charge $g = g_D$ is shown in Fig. 6.3. The standard Bethe-Bloch formula for a charged particle with $g_e = 1e$ and the same mass as the monopole is added to the graph. In the $\beta\gamma$ regime for minimal ionizing particles the energy loss for magnetic monopoles is approximately three orders of magnitude higher. In contrast to electrically charged particles, monopoles are subject to a decrease in energy loss for decreasing velocity.

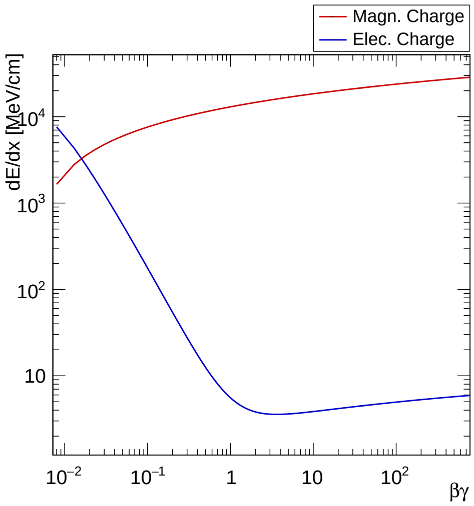


Figure 6.3: Energy loss for a magnetic monopole (red curve) computed with Eq. 6.1.10 and an electrically charged particle according to the standard Bethe-Bloch formula. Both particles have mass $m = 1\,\mathrm{GeV}$ and the minimum charge i.e. $g = g_D$ for the magnetic monopole and $g_e = 1e$ for the electrically charged particle. The energy loss for the magnetic monopole in the minmum ionizing $\beta\gamma$ regime is approximately three orders of magnitude higher.

Monopoles with $10^{-4} \lesssim \beta \lesssim 10^{-2}$

For monopoles with a velocity of $10^{-4} \lesssim \beta \lesssim 10^{-2}$ the average energy transfer is in the same order of magnitude as the binding energy of atomic electrons to the nuclei. Additionally, the electrons cannot be treated as being stationary anymore.

For this velocity range Lindhard [104] has proposed a model, in which the absorbing material is treated as an degenerate Fermi gas of noninteracting electrons. With a semi-classical approach Ahlen and Kinoshita [105] have extended this model to describe magnetic Dirac monopoles. The energy-loss is given by:

$$\frac{\mathrm{d}E}{\mathrm{d}x} = \frac{2\pi N_e g^2 e^2 \beta}{m_e c v_F} \left[\ln \frac{2 m_e v_F \Lambda}{\hbar} - \frac{1}{2} \right] \tag{6.1.12}$$

where

$$v_F = (\hbar / m_e)(3\pi^2 N_e)^{1/3} \tag{6.1.13}$$

is the Fermi velocity, N_e the electron density and Λ the Bohr radius.

In conductors two contributions have to be considered: from bulk and conduction electrons:

$$\frac{\mathrm{d}E}{\mathrm{d}x} = \left(\frac{\mathrm{d}E}{\mathrm{d}x}\right)_{\text{bulk}} + \left(\frac{\mathrm{d}E}{\mathrm{d}x}\right)_{\text{conduction}} \tag{6.1.14}$$

For non-conductors or bulk electrons in conductors the Bohr radius is given by $\Lambda = 53 \times 10^{-10}\,\text{cm}$ [106]. For conductors the Bohr radius is computed from the ratio of the fusion temperature T_m to the temperature of the nucleus T:

$$\Lambda = \frac{50a \cdot T_m}{T} \tag{6.1.15}$$

where $a = \sqrt[3]{\frac{A}{N_A \rho}}$ is the lattice constant, A is the atomic number and $N_A \approx 6.022 \times 10^{23}\,1/\text{mol}$ the Avogadro constant. It can be assumed that $T_m \approx T$, which corresponds to a partially solid and partially liquid nucleus [103]

The number density of nucleons N is obtained from the atomic number A and the Avogadro constant:

$$N = \frac{\rho \cdot N_A}{A}. \tag{6.1.16}$$

Assuming that every nucleus in a conductor contributes one electron to the conduction band, one finds:

$$N_{e,\text{cond}} = N \tag{6.1.17}$$

for the number density of conduction electrons $N_{e,\text{cond}}$. The remaining electrons in the atom contribute to the number density of the bulk electrons $N_{e,\text{bulk}}$, which therefore reads:

$$N_{e,\text{bulk}} = (A - 1) \cdot N. \tag{6.1.18}$$

In Table 6.1 the densities as well as the melting temperature T_m for the beam pipe material at the Belle II IR are summarized. With these values the Fermi velocity v_F and the Bohr Radius Λ for conduction and bulk electrons is computed. The result is given in Table 6.2.

Monopoles with $10^{-5} \lesssim \beta$

For $10^{-5} \lesssim \beta$ the dominating source of energy loss are elastic collisions with atoms. The interaction occurs between the magnetic field generated by the monopole

Table 6.1: Density of conduction $N_{e,\text{cond}}$ and bulk electrons $N_{e,\text{bulk}}$, fusion temperature T_m [107, 108] and ionization potential I [109, 110] for Beryllium, Paraffin ($C_{10}H_{22}$) and Gold.

Material	$N_{e,\text{cond}}$ $[10^{23}\,\text{cm}^{-3}]$	$N_{e,\text{bulk}}$ $[10^{23}\,\text{cm}^{-3}]$	T_m [K]	I [eV]
Beryllium	1.23	3.69	1338	9.22
Paraffin	-	2.53	-	10
Gold	0.59	4.60	9.32	1560

Table 6.2: Fermi velocity for conduction $v_{F,\text{cond}}$ and bulk electrons $v_{F,\text{bulk}}$ and Bohr radius for conduction electrons Λ_{cond} for Beryllium, Paraffin and Gold.

Material	$v_{F,\text{cond}}$ $[10^{8}\,\text{cm/s}]$	$v_{F,\text{bulk}}$ $[10^{8}\,\text{cm/s}]$	Λ_{cond} $[10^{-6}\,\text{cm}]$
Beryllium	1.78	2.57	1.01
Paraffin	-	2.27	-
Gold	1.40	5.96	1.28

and the magnetic moment of the atomic electron. Approximations for diamagnetic [111] and paramagnetic [112] materials have been proposed. Monopoles in this energy regime deposit their remaining energy in the form of phonons and magnons [113]. Since these types of energy deposition are not detectable with the PXD, monopoles in this momentum regime are not considered any further.

6.1.3 Multiple Scattering

The deflection of a charged particle in matter is induced by the Coulomb interaction between the incident particle and the nuclei of the medium. Multiple scattering events with small deflection angles add up to the total deflection θ_0, which is given by [90]:

$$\theta_0 = \frac{13.6\,\text{MeV}}{\beta c p} z \sqrt{\frac{x}{X_0}} \left(1 + 0.083 \ln\left(\frac{x}{X_0}\right)\right) \tag{6.1.19}$$

for an electrically charged particle. The thickness of the scattering medium in terms of radiation length X_0 is given by x/X_0 and z is the electric charge of the particle. Both the quantum mechanical and the classical treatment of scattering between electrically and magnetically charged particles, which is required for the description of multiple scattering for monopoles, is complex and exhibits complicated oscillatory behavior as well as backward scattering [114]. Nevertheless, for all practical purposes the substitution $e \rightarrow \beta n g_D$ is applied [97]. The additional factor β arises from the electric field implied by the Lorentz transformation of the propagating magnetic particle into the rest frame of the scattering partner. A detailed mathematical derivation can be found in [90].

6.1.4 Motion in a Magnetic Field

The duality of electromagnetic quantities implies a Lorentz force law for magnetic charges:

$$\mathbf{F}_m = g(\mathbf{B} - \beta \times \mathbf{E}) \tag{6.1.20}$$

where g is the magnetic charge and β is the velocity over the speed of light.

In this thesis a magnetic north pole is defined as a magnetically charged particle, which is accelerated parallel to the magnetic field lines. A south pole is accelerated anti-parallel to the field lines.

In general the equations of motion are solved numerically. An analytical solution is obtainable in the special case of a magnetically charged particle in a uniform magnetic field and an initial velocity perpendicular to that field [115]. If the magnetic field points in z direction and the particle starts off in x direction at time $t = 0$ the time evolution of the position $\mathbf{r}(t)$ reads:

$$\begin{aligned}\mathbf{r}(t) = {} & \frac{p_0}{gB_0}\left(\sinh^{-1}(\frac{gB_0(t+\Delta t)}{E_0}) - \sinh^{-1}(\frac{gB_0\Delta t}{E_0})\right)\mathbf{z} \\ & + \frac{E_0}{gB_0}\left(\sqrt{1+\frac{(gB_0(t+\Delta t))^2}{E_0^2}} - \sqrt{1+\frac{(gB_0\Delta t)^2}{E_0^2}}\right)\mathbf{x}\end{aligned} \tag{6.1.21}$$

where p_0 and E_0 are the initial momentum and the initial energy, respectively. The infinitesimal time steps are represented by the variable Δt.

The projection of the monopole trajectory for a north pole in a magnetic field aligned with the z axis is shown on the bottom of Fig. 6.4. On the left-hand side the xy plane and on the right-hand side the rz plane is presented. The trajectories of electrically charged particles can be seen on the top of the Figure for comparison. Contrary to electrically charged particles in a magnetic field, magnetic charges are accelerated parallel to the magnetic field. In the rz plane the trajectory resembles a parabola. In the plane perpendicular to the magnetic field the magentic monopole propagates along a straight line. An electrically charged particle, on the other hand, follows a straight line in the rz plane and curves in the plane perpendicular to the magnetic field.

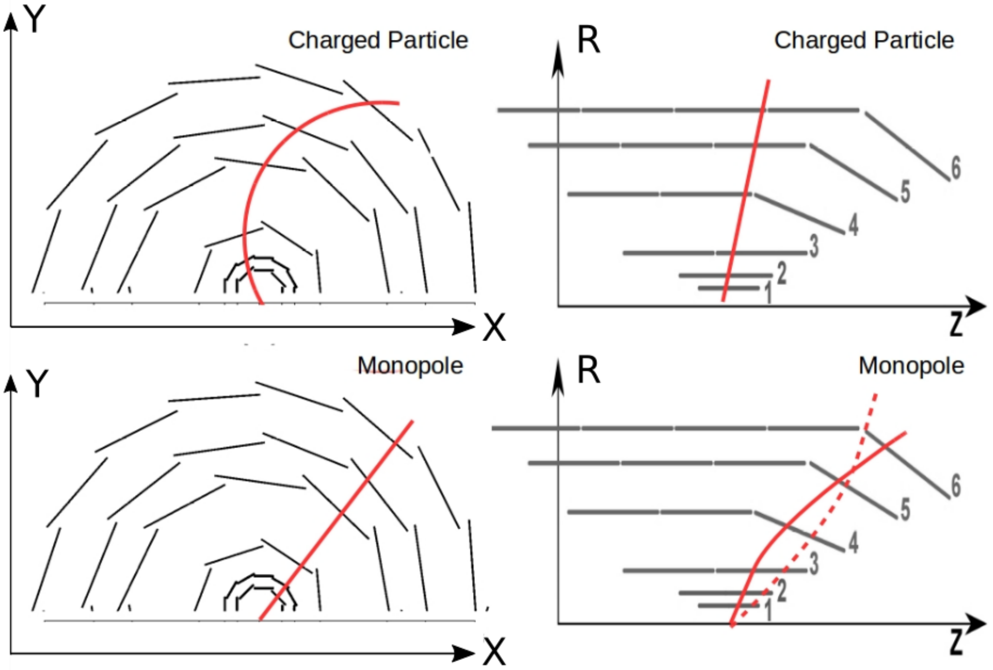


Figure 6.4: Trajectories of electrically (top) and magnetically (bottom) charged particles in a homogeneous magnetic field parallel to the z direction. In the plane perpendicular to the magnetic field - the xy plane - (left-hand side) electrically charged particles are bend and magnetic monopoles propagate along a straight line. In the rz plane (right-hand side) the monopoles follow a parabola track and the electrically charged particles a straight line. Source [116]

Monopole trajectories in the xz plane according to Eq. 6.1.21 are shown in Fig. 6.5 for magnetic north poles. On the left-hand side the monopole mass is varied between 3 GeV to 5 GeV while a unit Dirac charge is maintained. On the right-hand side the charge in multiples of the Dirac unit is varied and the mass is kept at 4 GeV. The magnetic field in both cases has a magnitude of 1.5 T and is aligned with the z direction. The initial velocity in x direction is computed with Eq. 6.1.9 with $\sqrt{s}$ fixed at 10.58 GeV.

The acceleration in z direction due to the magnetic field increases non-linearly with

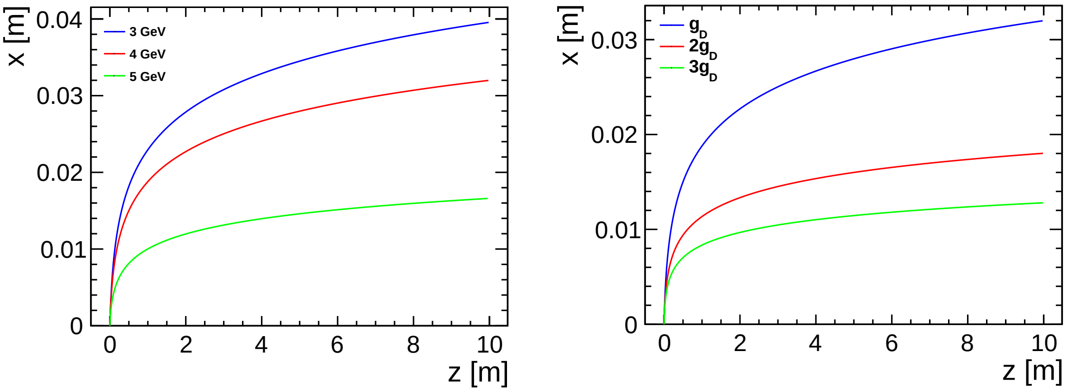


Figure 6.5: Trajectories of magnetic monopoles in a 1.5 T magnetic field aligned parallel to the z axis. The trajectories are computed with Eq. 6.1.21 with the initial velocity determined by Eq. 6.1.9. The curvature increases non-linearly with increasing mass and increasing magnetic charge.

increasing mass and increasing magnetic charge. The latter is a direct consequence of the modified Lorentz force law. A higher mass, on the other hand, is linked to a lower initial velocity in x. Therefore the trajectory curves more strongly towards the z direction.

6.1.5 Past and Ongoing Searches

The following Section gives a brief overview over the most relevant past and ongoing experimental searches for Dirac magnetic monopoles. In particular, searches at electron-positron colliders are presented. Against this background, the lack of experimental evidence is discussed and the search strategy at Belle II introduced.

Dedicated techniques for the detection of magnetically charged particles have been developed. Among these, the *Induction Technique in Superconductive Coils* [117] is characterized by superconducting coil connected to a Superconducting Quantum Interferometer Device (SQUID). If a magnetic monopole traverses the superconducting medium a current change corresponding to one unit of the flux quantum of superconductivity is induced. SQUIDs are also used for the search of magnetic monopoles trapped in matter. Most notably, the HERA collaboration used SQUID magnetometers to investigate the beam pipe to search for magnetic monopoles in e^+p-collisions [118].

Other experiments exploit the light yield in scintillators and/or the ionization in gas detectors in order to detect the characteristic energy loss of monopoles and/or their motion in a magnetic field [119]. Nuclear Track Detectors (NTDs), for instance, are a popular method to detect heavily ionizing particles. Particles depositing energy in a NTD leave a track in the medium. After chemical etching of the NTD material, the track can be examined microscopically. These experiments rely either on an astrophysical source for magnetic monopoles or they are performed at particle colliders.

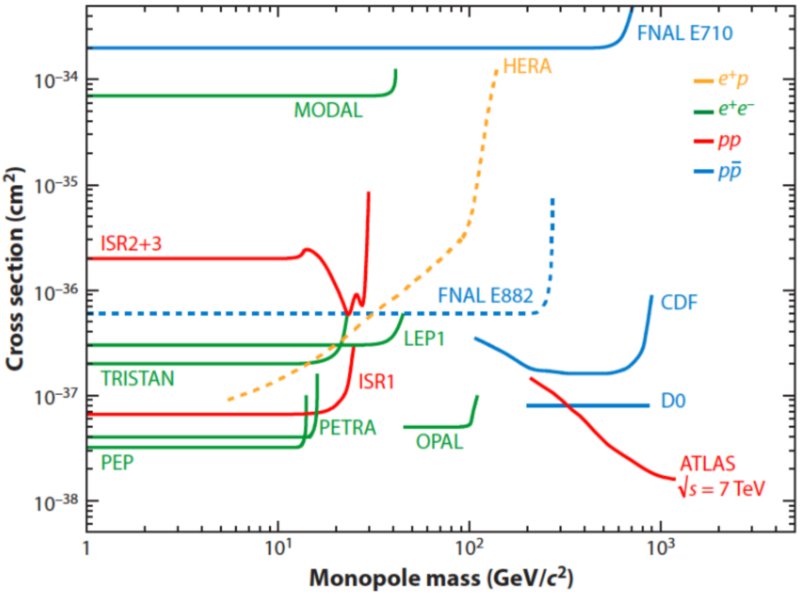


Figure 6.6: Exclusion limits for magnetic monopoles for various experiments. Electron-positron colliders are especially sensitive in the low-mass regime. Source [105]

So far e^+e^-, e^+p, $p\bar{p}$ and pp collisions have been investigated as sources for magnetic monopoles [105, 120]. Fig. 6.6 visualizes the parameter space scanned so far. Searches in e^+e^- collisions account for entries in the low-mass region. In Table 6.3 a list of studies performed in e^+e^- collisions is shown. High energy searches in e^+e^- collisions were performed at a center-of-mass energy of 206.3 GeV with the OPAL detector at LEP [121]. In their analysis they focused on high energy deposition in the tracking chambers. The TRISTAN ring at KEK [122], PETRA at DESY [102], PEP at SLAC [123] as well as the MODAL experiment at LEP [124] tried to find evidence for magnetic monopoles with NTDs.

Not shown in Fig. 6.6 are the exclusion limits from TASSO at DESY [125], whose analysis strategy was based on finding tracks related to magnetic charges, and the

Table 6.3: Past searches for magnetic monopoles at electron positron colliders. The second column specifies the technique used to identify magnetically charged particles.

Experiment	Instrument
MODAL at LEP	NTDs
TRISTAN at KEK	NTDs
CLEO at CESR	NTDs
PETRA at DESY	NTDs
TASSO at DESY	Tracking
OPAL at LEP	Wire Chamber

results from CLEO, which operated at a center-of-mass energy of 10.6 GeV [126]. In terms of beam energy CLEO is comparable to Belle II. In their search, the CLEO collaboration relied on NTDs.

6.1.6 Short-Comings of Particle Collider Experiments and Strategy at Belle II

At particle collider experiments a general-purpose detector at the IP is usually employed for the detection and identification of particles. Without the proper arrangements, these detectors are not suited for the search of highly ionizing particles and/or particles carrying a magnetic charge.

The identification of highly ionizing particles is hindered since detector calibration is usually performed with Minimum Ionizing Particles (MIPs). Detector performance studies rely on MIPs as well leading to a lack of knowledge about the detector response to HIPs. Additionally, the identification of particles is based on the combination of data recorded by multiple sub-detectors. The short range of HIPs implies that only a limited number of sub-detectors can detect the particle. This can lead to misidentification or a complete loss of information about the HIP. The radical proposal to perform HIP identification at Belle II with the innermost sub-detector - the PXD - only, would remedy this issue. The prospect of using Phase II data to perform a monopole analysis based on PXD information is discussed in Appendix C.

The presence of a magnetic field alters the trajectory of magnetically charged particles as described in Section 6.1.4. The acceleration along the magnetic field lines impedes the reconstruction of the particle's trajectory, when conventional tracking algorithms are used. Dedicated tracking algorithms for the identification of magnetically charged particles are currently developed at Belle II [127].

6.2 Simulations with `basf2`

Dirac magnetic monopoles are implemented in `basf2` simulations based on the Geant4 code provided by V. Ivanchenko [128]. A monopole pair production (see Section 6.1.1) is simulated implying the creation of a MC north and south pole in each event. Beam background (MC simulations) is overlaid in each event. The number of simulated events is adjusted to the monopole mass. The larger range of low-mass monopoles compared to the high-mass particles allows to simulate less events to reach the same number of clusters in the PXD. Each simulation is set up such that **at least** 40,000 clusters in the PXD are registered.

In the following Sections the range of magnetic monopoles in the detector is studied first. Subsequently, one and two dimensional distributions of cluster properties are shown. Pixel charge distributions of monopole clusters are analyzed as well. Finally, the training of neural networks to separate monopole from background clusters is described and the results are presented.

6.3 Range in Belle II Detector

The range of magnetic monopoles in the detector is limited due to two effects related to their magnetic charge: the high energy loss and the acceleration along magnetic field lines. The trapping of monopoles in matter as well as their energy loss in the low energy regime is model-dependent and still heavily disputed. `basf2` employs a very crude energy cut-off to simulate the trapping (or equivalently the *decay*) of monopoles in matter. This model makes reliable statements about the stopping of monopoles in matter difficult. Consequently, all `basf2` results related to the range of monopoles in the detector should be treated with caution.

Analogous to anti-deuterons, the decay vertices of the simulated monopoles serves as an estimate for their range. In Fig. 6.7 the projection of decay vertices onto

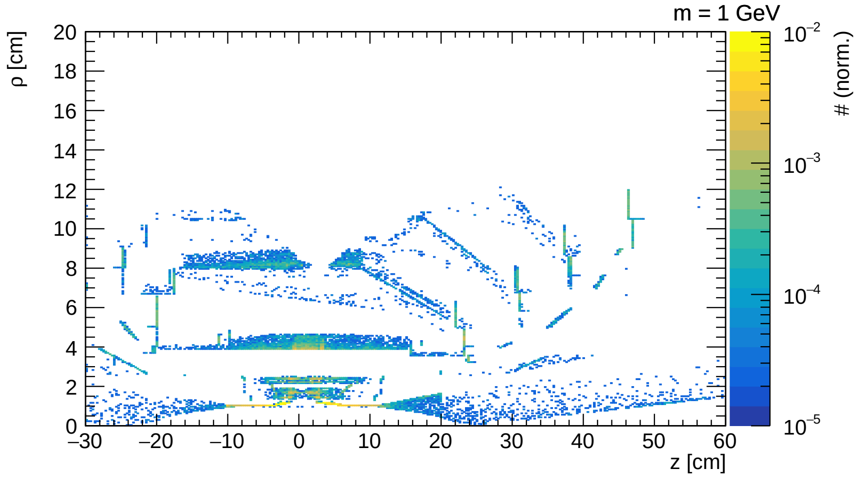
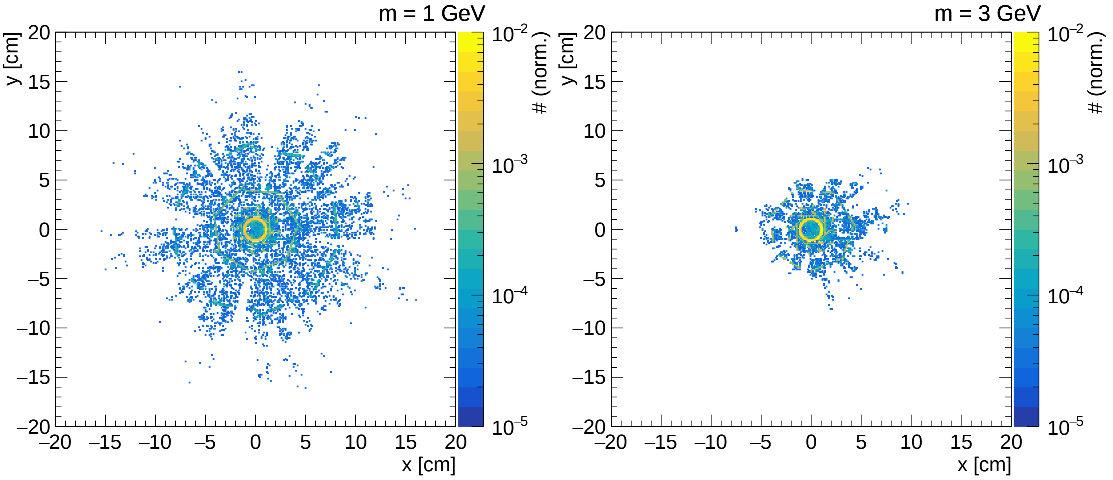


Figure 6.7: Decay vertices of 1 GeV monopoles in ρz plane. While monopoles can reach the forth SVD layer in forward direction, they are confined to the inner three SVD layers in backward direction. In the barrel part around $z \sim 2$ cm there is a gap at the second SVD layer. The density of decay vertices at the beam pipe is orders of magnitude higher than in the PXD and SVD layers.

the ρz plane for monopoles with mass $m = 1$ GeV and unit magnetic charge is presented. The outline of the beam pipe, the two PXD layers and a few SVD layers can be seen. The range of 1 GeV particles is confined to the inner vertex region. The distribution of decay vertices is asymmetrical. Monopoles can reach the forth SVD layer in forward direction due to the additional boost in the laboratory frame. In backward direction the third SVD layer marks the maximum range of monopoles. In the region around $z = 0$ cm only the second layer is accessible for monopoles and at $z \sim 2$ cm there is a gap in the distribution of decay vertices on the second SVD layer.

The interaction of monopole and the magnetic solenoid field is responsible for the lack of decay vertices in the region perpendicular to the z axis. The range of monopoles is increased in forward and backward direction and decreased in the barrel part of the detector due to the acceleration along the magnetic field lines. North poles are accelerated in positive and south poles in negative z direction.

The projection of decay vertices onto the xy plane, as shown in Fig. 6.8, confirms the confinement of monopoles to the inner detector region. A comparison between the decay vertices from monopoles with $m = 1$ GeV (left-hand side) and $m =$



(a) (b)

Figure 6.8: Decay vertices of 1 GeV (left-hand side) and 3 GeV (right-hand side) monopoles in xy plane. The range of magnetic monopoles decreases with increasing monopole mass. While 1 GeV monopoles decay predominantly in the beam pipe and PXD layers, a small number of particles can still reach all four layers of the SVD. The 3 GeV monopoles are hardly able to reach the first two layers of the SVD.

3 GeV (right-hand side) shows that an increase in mass constricts the spatial range even more. Monopoles with $m = 3\,\text{GeV}$ decay mostly in the beam pipe and the PXD, only a few reach the first or second SVD layer. The decrease in range with increasing mass is a consequence of the lower amount of kinetic energy available to high-mass monopoles. For pair production of 3 GeV magnetic monopoles, 6 GeV of the available 10.58 GeV is already consumed by the production process alone.

In Table 6.4 the average number of clusters generated in the PXD per MC monopole is listed for different monopole masses. A charged particle traversing the PXD is expected to produce two clusters on average, one in the inner and one in the outer layer. The spatial overlap of PXD modules, mistakes in the clustering or insufficient energy deposition may modify the number of clusters on an event to event basis, the mean number of clusters, however, is not expected to deviate considerably from two. While 1 GeV and 2 GeV monopoles generate approximately half the amount of clusters expected, the number of clusters falls drastically for masses above 2 GeV. The majority of 3.5 GeV monopoles do not leave a signal at all and only about 1% of 4 GeV monopoles are registered in the PXD.

Table 6.4: Average number of clusters generated in the PXD for monopoles of a certain mass. On average, an electrically charged particle is expected to create two clusters in the PXD. While a 1 GeV monopole generates rougly half of that, the number of clusters falls rapidly with increasing monopole mass.

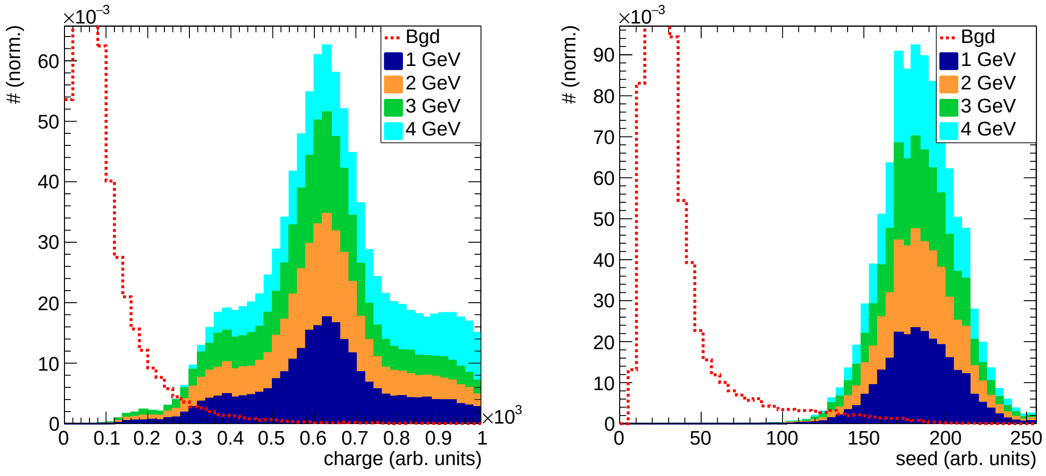
Mass [GeV]	1	2	3	3.5	3.75	4
Clusters per particle	1.15	0.97	0.62	0.34	0.13	0.014

6.4 Cluster Properties

The property of PXD clusters generated by magnetic monopoles is studied in the following Section. First, one dimensional distributions of cluster properties are presented. The distribution of background cluster properties are overlaid to allow for a comparison. Two dimensional distribution are displayed as well for two different monopole masses. To keep the following Section at an appropriate length only examples are shown, the entire set of one dimensional distributions is depicted in Appendix E. Two dimensional distributions for monopoles with mass of 1 GeV and with 4 GeV are listed in Appendix F.

6.4.1 One Dimensional Distributions

The cluster charge for background and magnetic monopoles with different masses is shown in Fig 6.9a. Note that the monopole distributions are stacked to allow for an easier comparison. The most probable cluster charge for monopoles is ~650, which is about one order of magnitude higher compared to the background distribution. The high-charge tail of the background distribution exhibits a marginal overlap with the monopole distributions. The peaks of the monopole distributions are less pronounced compared to the background peak. Additionally, the shape of the distributions differ from a Landau-Gauss function. Instead of an asymmetrical Landau-Gauss, an extensive plateau-like region around a nearly symmetrical peak is observable. A similar peak shape characterized the cluster charge distribution for the low-momentum anti-deuterons (see Section 5.2.1). A broadening of the peak and the low-energy part of the peak were attributed to the high energy deposition of the particles. The large number of charge carriers created by the high energy deposition can be lost due to diffusion to neighboring cells and the application of a read-out threshold. The energy deposition of magnetic monopoles tends to be even higher compared to anti-deuterons leading to additional charge loss effects:



(a) (b)

Figure 6.9: Total cluster charge (left-hand side) and cluster seed charge (right-hand side) for magnetic monopoles and background. The distributions for different monopole masses are stacked for better comparison. The cluster charge distribution for all monopole masses is shifted to higher charge values compared to the background. The shape of the distributions do not resemble Landau-Gauss functions. The seed charge for clusters generated by magnetic monopoles tends to adopt high adc values. The maximum adc value (255) is easily attainable from the energy deposition of the magnetically charged particles.

The cluster seed charge distribution, depicted in Fig 6.9b, indicates that it is possible for individual pixels to abandon the linear charge regime or to go into full saturation. The tail of the seed charge distributions extends up to the maximum adc value confirming that charge saturation of the seed pixel is a relevant issue. In Section 3.8.1 the capacity of the internal gate was estimated to be $\sim 60,000$ electrons. The linear regime between number of collected charge carriers and modulation of drain current is abandoned at about 40,000 electrons. Consequently, it is not unreasonable to assume that a non-negligible part of the cluster seed charge distribution is affected by the non-linear charge collection regime. The non-linear relationship between amount of deposited energy and read-out current could manifest itself in the unconventional cluster charge distribution. When full saturation of a pixel cell is reached, the loss of liberated charge carriers is even more dire. An excess of charge carriers cannot be registered due to the limited range of the analog to digital converter.

The cluster charge distributions for monopoles in the mass range 1 GeV to 3 GeV do not differ noticeably. The 4 GeV monopole distribution, on the other hand, exhibits a slightly more pronounced high-charge tail. In the cluster seed charge distribution this shift towards high-charge values is not observable indicating that the increase in total cluster charge does not arise from the seed pixel. The high-charge tail correlates with an increase in cluster size as shown on the left-hand side of Fig. 6.10, where the cluster size for different monopole masses is depicted. In general, clusters generated by magnetic monopoles tend to be larger compared to background clusters as expected from the higher energy deposition of the magnetically charged particles. The cluster size distribution is largely independent of the mass for masses between 1 GeV to 3 GeV, but smears out towards higher cluster sizes for 4 GeV monopoles. This increase in cluster size agrees with the simultaneous increase in total cluster charge and the invariance of the cluster seed charge. While the energy deposition in the seed pixel stays unaltered, the total cluster charge tends to be higher for the 4 GeV monopoles due to the contribution of more pixels in the cluster. The origin of the increase in cluster size can be traced back to the stronger influence of the magnetic field on the high-mass monopoles (see Section 6.1.4). The acceleration along the magnetic field lines decreases the incident angle between monopole track and PXD modules. The cluster size distributions in u and v direction, which are depicted in Fig. 6.11, confirm this assumption. The size in u direction for high-mass monopoles does not differ from the lower masses. The size in v, on the other hand, is increased for high-mass monopoles implying an energy deposition into several pixel cells along the v direction, which is only feasible under a shallow incident angle. The cluster seed charge remains largely unaltered since the amount of energy deposited in the first pixel cell encountered by the monopole [1] is not affected by its incident angle.

The difference in cluster properties of magnetic monopoles with different mass appears to be predominantly affected by the particle's acceleration along the beam pipe. The lower momentum of high-mass monopoles is also accompanied by a lower energy loss (see Section 6.1.2), which does not seem to alter the cluster property distributions. The non-linear behavior between deposited energy and cluster charge variables is believed to contribute partly to this effect. Additionally, the momentum distribution of monopoles at the PXD can blur the correlation between monopole mass and amount of energy loss in the PXD. The direction and magnitude of momentum for 1 GeV monopoles is less decisive for the distribution of cluster properties compared to the 4 GeV monopoles. Only a small fraction of

---

[1] This statement implies that the highest charge deposition occurs into the first pixel cell, which is a reasonably good assumption given the energy loss curve of magnetically charged particles.

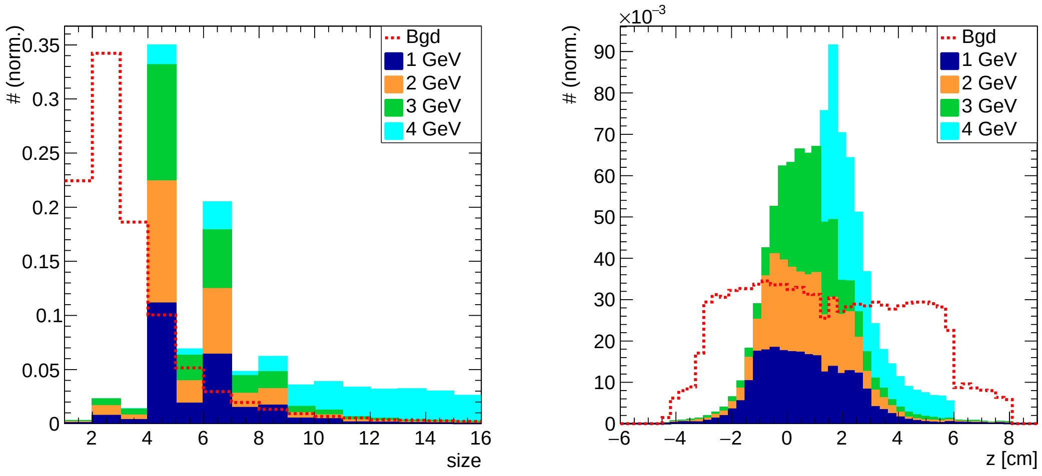


Figure 6.10: Total cluster size and cluster z position for clusters generated by magnetic monopoles and background. The mean cluster size for monopole clusters is higher compared to background clusters. The distribution is independent of the monopole mass except for the 4 GeV monopoles, which tend do generate larger clusetrs. The distribution of cluster z positions shifts in positive z direction with increasing mass.

4 GeV monopoles with an appropriate momentum properties can reach the PXD. A north pole with initial momentum parallel to the magnetic field, for instance, profits from the acceleration parallel to its initial trajectory and the forward boost in the laboratory frame. A south pole under the same conditions is accelerated in the opposite direction, which is detrimental for its overall momentum. Hence, the highest monopole momentum at the PXD is achievable for north poles whose initial momentum is parallel to the magnetic field, since the acceleration along the field, the initial velocity and the forward boost all add up constructively. The cluster z position for magnetic monopoles and background, depicted in Fig. 6.10, confirms this assumption. A shift towards positive z values with increasing monopole mass is observable. The peak originates from the north poles with initial forward momentum. A second peak for south poles with background momentum is not discernible due to the destructive effect of the forward boost on the particle's momentum. The background particles, in contrast to the monopoles, impinge on the PXD nearly homogeneously along the z axis since their production vertex is not fixed to the IP.

The entire set of one dimensional distributions are listed in Appendix E. Additionally, the monopole distributions against muons are depicted, since muon pair

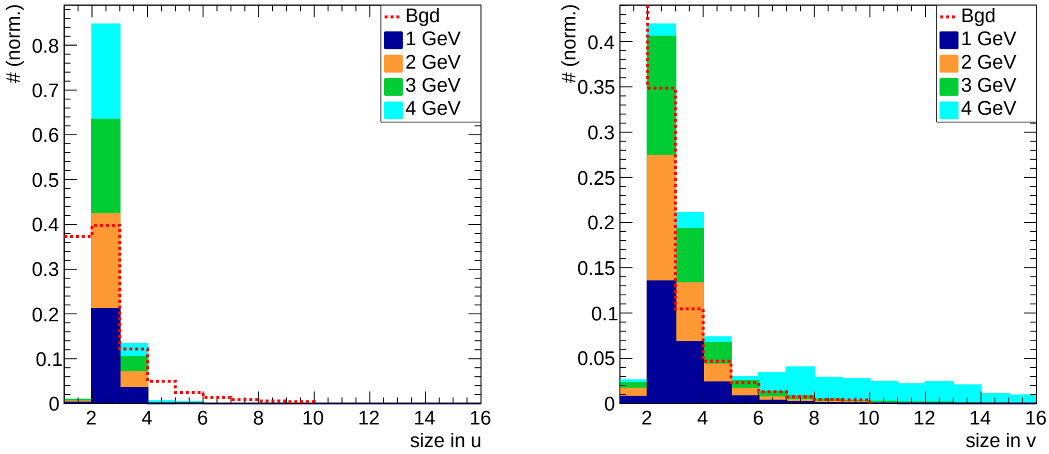


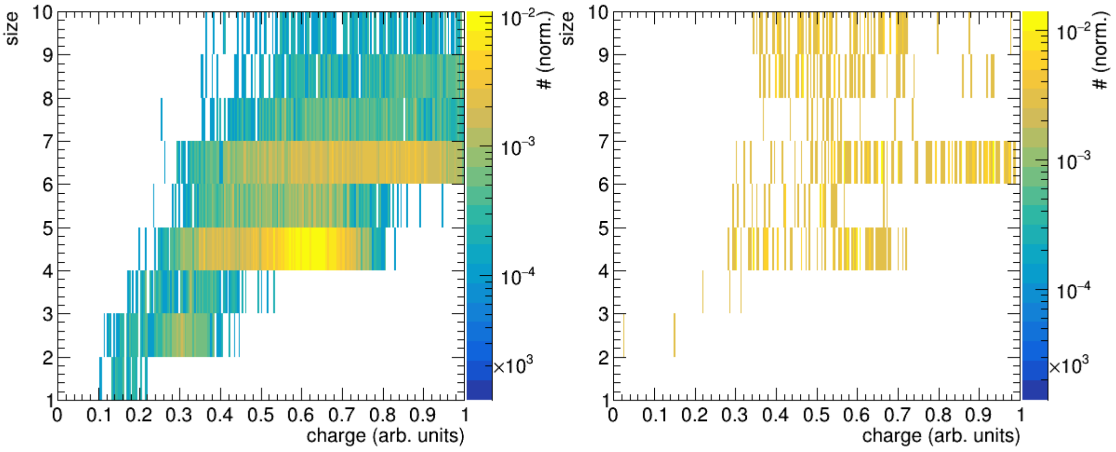
Figure 6.11: Cluster size in u and v direction for magnetic monopoles and background. While the cluster size in u is independent of the monopole mass, the size in v is increased for 4 GeV monopoles.

production can be considered as background source due to the similar event structure. The plots in the Appendix demonstrate that muons generate small low-charge clusters clearly distinguishable from magnetic monopoles.

6.4.2 Two Dimensional Distributions

In Fig. 6.12 the total cluster charge against cluster size for 1 GeV (left-hand side) and 4 GeV (right-hand side) magnetic monopoles are presented. The shape of the two dimensional distribution for the 1 GeV monopoles is similar to the sub-GeV anti-deuteron distribution. The center of gravity of the distribution is shifted towards higher charge and higher size values compared to the anti-deuterons. Additionally, the variance appears to be greater: the monopole distribution occupies a wider parameter space. In comparison to the background distribution shown in Section 5.2.2 there is some overlap in the low-charge region, the high charge region is largely free of background clusters.

The charge-size distribution for the 4 GeV monopoles exhibits an absence of small low-charge clusters. The lack of density fluctuations in the entries of the plot indicates that only the fringe of the distribution is visible. The range chosen for the axes is certainly inappropriate for the extend of the distribution, however it



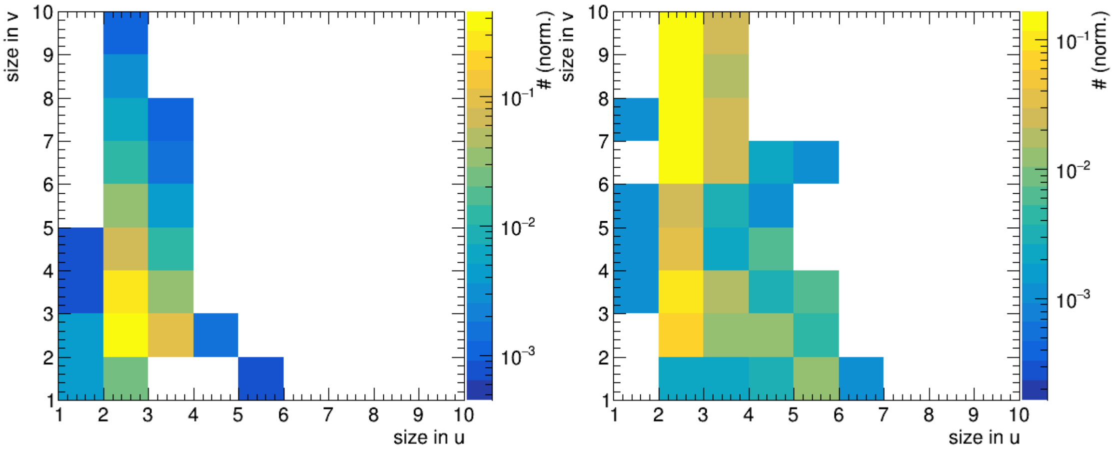
(a)

(b)

Figure 6.12: Total cluster charge against cluster size for 1 GeV and 4 GeV magnetic monopoles. A correlation between cluster charge and size is observable: an increase in cluster charge is accompanied by an increase in cluster size. While the 1 GeV monopole distribution resembles the low-momentum anti-deuteron distribution, the distribution for the 4 GeV monopoles does not contain small and/or low-charge clusters.

allows for an ideal comparison with the other two dimensional distributions. The overlap with the background distribution is negligible due to the absence of clusters in the low-charge region.

The cluster size in u against cluster size in v for low mass (1 GeV) and high mass (4 GeV) monopoles is depicted in Fig. 6.13. While the background distribution extends over the entire parameter region shown in the plot, the 1 GeV monopoles mostly generate clusters with an extension of two to three pixels in u direction. Contrary to anti-deuterons the occurrence of clusters with a single pixel in u direction is suppressed as described in the previous Section. The cluster size in u and v are anti-proportional: an increase in u is accompanied by a decrease in v due to the allocation of available charge carriers. If there is a strong diffusion of liberated charge carriers in u direction, less charge carriers are available for neighboring pixel cells in v direction. The direction in which the increase in cluster size occurs is strongly influenced by the incident position of a particle on the silicon sensor. Charge sharing to neighboring cells is suppressed, if particles deposit energy in the center of a pixel cell. If particles impinge, for instance, on the outer edge in u



(a)

(b)

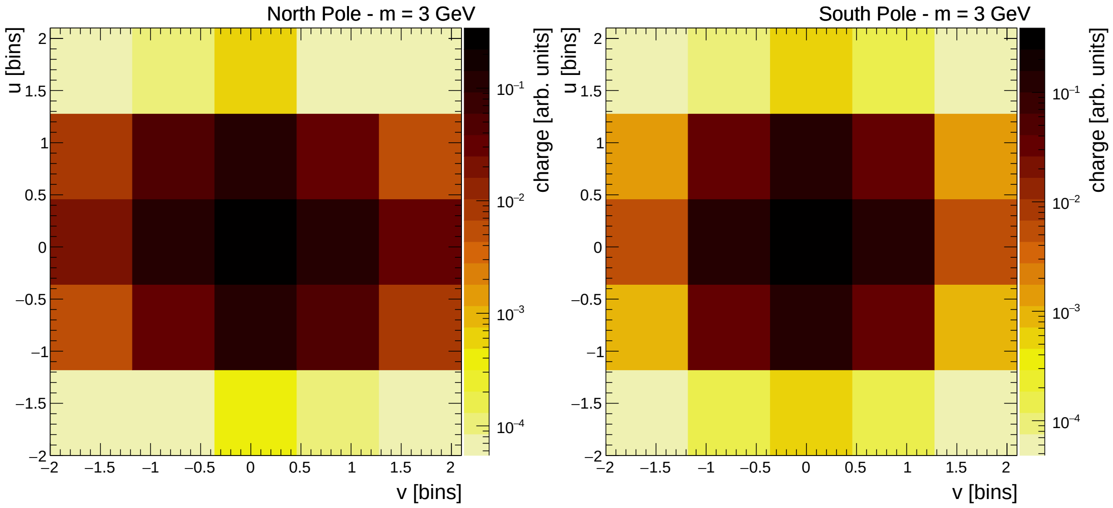
Figure 6.13: Cluster size in u against cluster size in v for 1 GeV and 4 GeV magnetic monopoles. The cluster size in u and v are anti-proportional: an increase in u size is accompanied by a decrease in v and vice versa. While the cluster size in both coordinate directions increases with increasing monopole mass, the increase in v is stronger.

direction, charge carriers are likely to diffuse to the nearest neighboring cell in u increasing the size by one unit in this direction.

For the high mass monopoles the occurrence of large clusters is more common, which also manifests itself in an increase in cluster size in u and v direction. The increase in v is more pronounced due to the acceleration along the magnetic field lines as explained above. For both distributions a strong overlap with the background distribution is inevitable since background clusters occupy the entire parameter range.

6.5 Pixel Charge Distribution

In Fig. 6.14 the pixel charge distribution for a 3 GeV magnetic noth pole (left-hand side) and a magnetic south pole (right-hand side) with unit magnetic charge are shown. Many features in the pixel charge distribution are similar to the anti-deuteron distributions (see Section 5.3): a cross-like structure, uv symmetry in the



(a) (b)

Figure 6.14: Average pixel charge distributions for magnetic north poles (left-hand side) and magnetic south poles (right-hand side) with $m = 3\,\mathrm{GeV}$. The pattern is similar to the pixel charge distribution of low-momentum anti-deuterons. While the inner 3 x 3 matrix exhibits uv symmetry, the full 5 x 5 matrix is *stretched* in v direction. The asymmetry is more pronounced for the magnetic north poles. The charge per pixel for the south pole distribution is lower in the outermost pixel row and column of the matrix.

innermost pixel region and higher charge values along the v direction compared to the u direction. Contrary to the anti-deuterons, the outermost pixels in u also receive small amounts of charge carriers confirming that the average energy deposition of magnetic monopoles is higher compared to anti-deuterons.

The pixel charge distributions for north and south pole are visibly different. The charge distribution for magnetic north poles has a stronger asymmetry. The clusters are stretched along v direction in accordance with the higher size values in v, which were observed in the previous Section. The momentum of magnetic north poles adopts a maximum value at the PXD when the initial momentum is parallel to the v/z axis. In this case, the north poles benefit from the forward boost, the acceleration along the field lines, which are parallel to the v/z direction as well, and their initial momentum. These high-momentum monopoles generate clusters with a large extension in v direction since their momentum aligns with the v axis. For south poles, on the other hand, the forward boost and the acceleration along the field lines are anti-parallel leading to a lower momentum in v and consequently a smaller cluster size in the this direction.

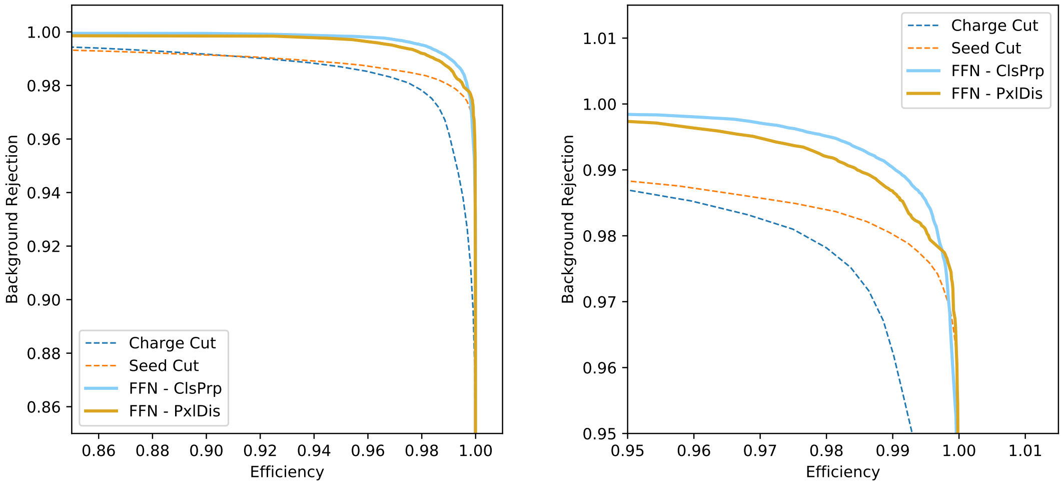
Magnetic south poles with a mass of 4 GeV do not reach the PXD at all. The low initial momentum of the heavy monopoles requires the constructive interference of forward boost and acceleration along the field lines to leave a signal in the innermost detector. Consequently, solely 4 GeV north poles are registered by the PXD.

6.6 Training Results

FFNs and SOMs are trained with equal amounts of 1 GeV, 2 GeV, 3 GeV and 4 GeV monopoles in the training set. 30,000 vectors per mass are available for training and 10,000 per mass for testing. In the following the FFN results are presented first. Subsequently, the performance of SOMs is presented. The influence of the monopole mass on the training is analyzed as well. At the end of this Section the performance of all FFNs is summarized.

6.6.1 Feed-Forward Network

Training of FFNs with cluster properties and pixel charge distributions is performed as described in Section 5.4. The ROC curves for the trained FFNs are shown in Fig. 6.15. The one dimensional cuts on seed and total cluster charge are included as well. The FFNs perform better than the cuts once again. Contrary to the anti-deuteron results, the cluster property input appears to deliver better results compared to the pixel charge distribution. A close-up of the upper right side of the curve is depicted in Fig. 6.15a . While the ROC curve belonging to the cluster property input set is smooth, the ROC curve associated with the pixel charge distributions exhibits an oscillatory behavior hinting at issues with the network architecture or input variables. The size of the pixel matrix is assumed to be responsible for the inferiority of the input set containing pixel charge distribution. Clusters generated by magnetic monopoles can easily exceed a size of 5 in v direction (see Section 6.4). Consequently, there is a significant amount of clusters, which are not entirely captured by the 5 x 5 pixel matrix leading to a loss of information. Since cluster properties are not limited to a certain pixel range, the FNN trained with the cluster properties performs better than the one trained with pixel charge distributions. Additionally, the pixel charge distributions are different for



(a) (b)

Figure 6.15: ROC curves for the identification of magnetic monopoles based on cluster properties and pixel charge distributions. One dimensional cuts on the total cluster charge and seed charge are shown as dashed curves. The FFNs perform better than the one dimensional cuts. A close-up of the upper right corner of the curve (right-hand side) reveals that the ROC curve for the FFN trained with cluster properties lies above the one trained with pixel charge distributions. The latter one also exhibits an oscillatory behavior in an efficiency region between ~0.98 and ~1.0.

north and south poles (see Section 6.5). The south pole distribution is more symmetric under u/v exchange, which is also characteristic for background particles. Consequently, the occurrence of south poles in the training set could impair the training with pixel charge distributions as well..

The presence of single pixel clusters in the input set is not an issue for magnetic monopoles. In Table 6.5 the amount of single pixel clusters for monopole masses between 1 GeV to 4 GeV is listed. The occurrence of single pixel clusters appears to be largely independent of the monopole mass except for the 4 GeV particles in accordance with the cluster size distributions discussed in Section 6.4. The effect of single pixel clusters on the training is not investigated any further since their contribution to the input set is in the sub-percent regime.

The impact of the monopole mass on the training is discernible from Fig. 6.16. FFNs are trained with vector sets containing pixel charge distributions of clusters

Table 6.5: Amount of single pixel clusters in the input vector set for different monopole masses. While the amount of single pixel clusters appears to be independent of the monopole mass between 1 GeV to 3 GeV, the amount decreases strongly for 4 GeV monopoles. For all masses the amount of single pixel clusters is in the sub-percent regime.

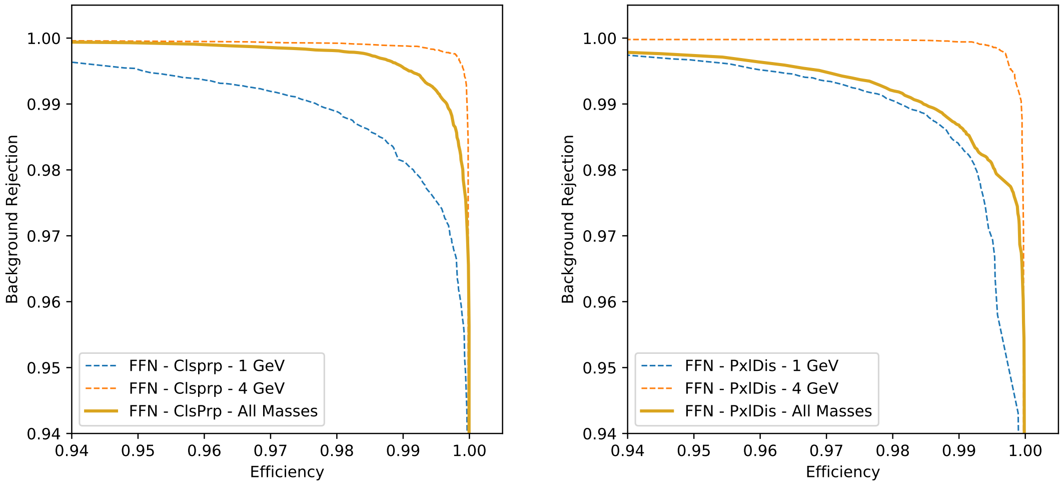
Mass [GeV]	1	2	3	4
Single Pixel Clusters [%]	0.33	0.44	0.36	1E-3

generated by magnetic monopoles with a mass of 1 GeV, 4 GeV or the entire mass range from 1 GeV to 4 GeV as shown before. The ROC curves demonstrate that identification of high-mass monopoles proves to be more successful than the identification of low-mass monopoles. The higher cluster size as well as the complete lack of low-charge clusters facilitates the identification of clusters generated by the 4 GeV monopoles (see Section 6.4).

The small size of the pixel matrix should affect the identification of 4 GeV monopoles the most, since the majority of clusters generated by the high-mass monopoles ia larger than size 5 in v direction. The ROC curve in Fig. 6.16b demonstrates that the performance of the FFN trained with pixel charge distributions of 4 GeV monopoles is barely affected by the irregularities, which arise for training with 1 GeV monopoles or with the entire mass range. The high pixel charges within the pixel matrix could overcompensate the loss of information from pixels outside the 5 x 5 matrix. Additionally, there are no south poles in the training set of 4 GeV monopoles. The pixel charge distribution of south poles is more symmetric and therefore more background-like. Consequently, a lack of south poles could be responsible for improvements in the training results.

6.6.2 Self-Organizing Map

The ROC curve for a SOM trained with cluster properties is shown in Fig. 6.17a. A comparison between SOM and FFN is complicated due to the shape of the transitional region between background and signal in node space (see Fig. 6.17b). A cut along the x axis is not appropriate for the curved boundary between monopole and background region. If the parameterization for background and monopole region is adjusted to achieve better classification results, special care should be taken to avoid over-paramterization. A paramterization following the boundary between the two classification regions as closely as possible can result in a loss of generality.

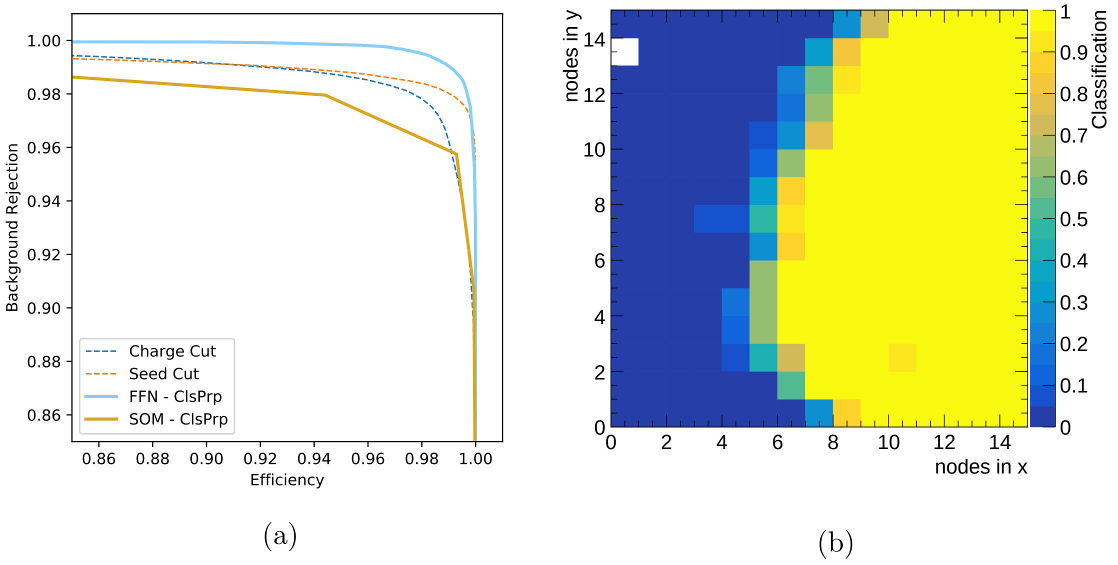


(a) (b)

Figure 6.16: ROC curves for the identification of magnetic monopoles based on cluster properties (left-hand side) and pixel charge distributions (left-hand side) for monopoles with $m = 1\,\text{GeV}$ and $m = 4\,\text{GeV}$ and all masses. The ROC curve for training with 4 GeV monopoles lies above the curves for training with all masses and 1 GeV monopoles.

It would deliver an excellent performance with the vector set used to determine the classification map, but would malfunction with a new unknown input set.

The classifier distributions for SOM and FFN are shown in Fig. 6.18. The projection of the convex boundary onto the x axis in node space manifests itself in a large transition region in the one dimensional representation. High classification values correspond to monopoles and low values to background. Background suppression in the monopole region is in the order to $\sim 10^{-2}$ to $\sim 10^{-4}$ for the SOM and $\sim 10^{-3}$ to $\sim 10^{-4}$ for the FFN. The background distribution falls with increasing classification values, but exhibits an increase at the far-right of the horizontal axis for both networks. The increase is more pronounced for the SOM, where the background suppression rises to $\sim 10^{-2}$, which is one order of magnitude higher than the $\sim 10^{-3}$ for the FFN.



(a)

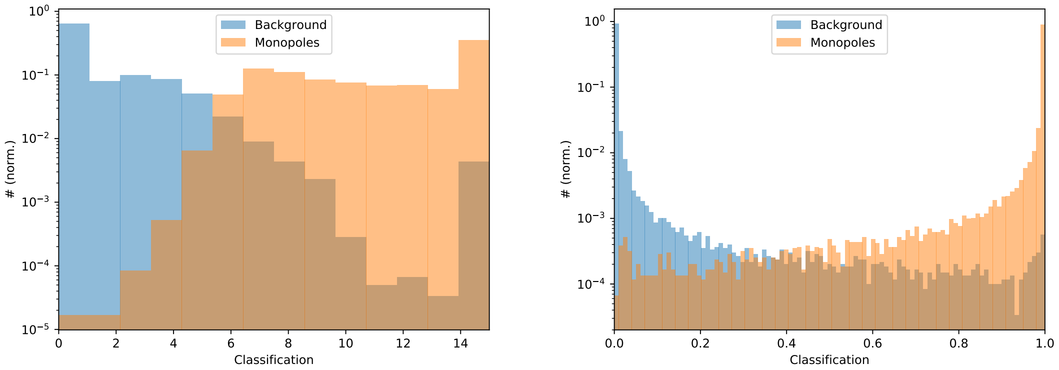
(b)

Figure 6.17: ROC curve and classification map for Self-Organizing Map trained with cluster properties. High classification values correspond to monopoles, low ones to background. The ROC curve is computed based on a one dimensional cut along the x axis in node space. The curve lies below the ROC curves for FFN as well as charge and seed cut. The transition region between background and monopole region is curved in node space. The monopole region is larger than the background region.

6.6.3 Summary

A summary of all FFNs trained with magnetic monopole clusters is given in Table 6.6. The cut on the classification axis is set to 0.5. When the entire mass range from 1 GeV to 4 GeV is considered the FFN trained with pixel charge distributions performs slightly worse than the one trained with cluster properties. The ROC curve for the FFN trained with pixel charge distributions is not smooth hinting at an issue in the learning process, which emerges from the network architecture or the limited size of the 5 x 5 pixel matrix.

The identification of 4 GeV monopoles is more accurate than the identification for lower masses due to the higher energy deposition and the significantly larger mean cluster size (see Section 6.4). The large size and lack of low-charge clusters improves the performance of the FFN compared to the low-mass monopoles. Additionally, there are no 4 GeV magnetic south poles. The neural network only has



(a) (b)

Figure 6.18: Classifier distributions for SOM and FFN trained with cluster properties of magnetic monopoles and background. High classification values correspond to monopoles, low values to background. The background distribution decreases with increasing classification values. A sharp rise of the background distribution is observable for the highest classification values. It is more pronounced for the SOM classifier.

Table 6.6: Accuracy of FFNs trained with magnetic monopole clusters for a cut at 0.5 on the classifier axis. The accuracy is computed after 1000 training epochs for the training and the testing set. The FFNs are structured according to the input set and the mass of the monopoles generating the clusters.

Input set	Mass [GeV]	Training [%]	Testing [%]	AUC [%]
ClsPrp	1 - 4	99.0	99.0	99.932
PxlDis	1 - 4	98.8	98.8	99.873
ClsPrp	1	98.6	98.6	99.877
PxlDis	1	98.8	98.7	99.890
ClsPrp	4	99.7	99.8	99.966
PxlDis	4	99.8	99.7	99.972

to learn one cluster type (north pole clusters) instead of two (north and south pole clusters). This is particularly beneficial for performance since south pole clusters are less distinct from clusters generated by background than north pole clusters.

CHAPTER 7

SUMMARY AND PROSPECTS

In this work the feasibility of identifying HIPs with PXD data was presented. Neural networks were employed to separate HIPs from background particles. Anti-deuterons and magnetic monopoles were chosen to test the application of two types of neural networks: Feed-Forward Networks and Self-Organizing Maps. The performance of the Feed-Froward Networks was shown to be better in terms of performance (background rejection and efficiency). The input variables consisted of cluster properties such as the cluster size, total cluster charge, etc. or a 5 x 5 pixel matrix around the cluster seed pixel. Training with the pixel charge distribution delivered better results. Exceptions to this observation were linked to the insufficient size of the pixel matrix.

The impact of particle momentum on the cluster properties and the pixel charge distributions was investigated for anti-deuterons. The energy deposition of low-momentum anti-deuterons is generally higher. The generated clusters are less *background-like*, due to the increase in cluster size and cluster charge properties with increasing energy deposition. Neural networks for anti-deuterons with momenta below 1 GeV were trained and excellent separation between background and signal particles was achieved. An AUC of up to 99.27% for training with pixel charge distributions was accomplished. The neural networks generally outperform one-dimensional cuts with the exception of Self-Organizing Maps, which can be worse in certain efficiency regimes. The poorer performance of Self-Organizing Maps is strongly correlated with the technique employed in this thesis to train and evaluate the map. The two machine learning algorithms were compared under fixed conditions and for a certain type of application in this work. An overall comparison between the two neural networks is not meaningful since the two algorithms are not formulated for the same class of problems.

If the momentum range of anti-deuterons is extended to include all particles with a momentum below 3 GeV, the performance of neural networks deteriorates. The training with pixel charge distributions reached an AUC of up to 90.58%. The impairing effect of single-pixel cluster on the training was identified as a significant contribution to the decline in separation power.

The experimental signature of Dirac magnetic monopoles was reviewed and the possibility to lower upper production limits with the Belle II detector was discussed. Clusters generated by particles with unit magnetic charge and a masses between 1 GeV to 4 GeV were studied. They exhibited distinctive characteristics, which facilitated the separation from background clusters. Most notably, the acceleration of monopoles along the magnetic field lines was found to manifest itself in the v-size of clusters. The application of Feed-Forward Networks for monopole-background separation yielded excellent results. AUCs up to 99.932% were reached when the magnetic monopoles were allowed to have a mass between 1 GeV to 4 GeV. When the mass was restricted to 4 GeV, an AUC of 99.972% was achieved. When training is performed with the pixel charge distributions, some irregularities were encountered. Their origin are most likely connected to the limited size of the pixel matrix and the different pixel charge distributions for north and south poles.

Prospects

While the PXD data reduction system is absolutely crucial to cope with the high data rate generated by the DEPFET silicon sensors, it has the disadvantage of discarding low-momentum or highly ionizing particles, which are not detected by the outer sub-detectors. A rescue system is required to counteract the rigorous elimination of data stemming from short-range particles. In this thesis it was shown that highly ionizing particles can be identified with PXD data only. A straight-forward implementation of a Feed-Forward Network is already sufficient to achieve excellent results. The integration of a neural network into the online reduction system could accomplish the cluster rescue.

The network architecture of Feed-Forward Networks and Self-Organizing Maps still require tuning. The hyperparameters have to be chosen to maximize performance and the prospect of using other types of neural networks (e.g. convolutional neural networks) has to be investigated. If implementation on hardware (e.g. FPGAs) is considered, the neural networks have to be adjusted accordingly. Additional cluster properties should be considered as well. Cluster properties such as cluster angle, cluster eccentricity etc. were tested in the scope of this thesis (documentation available on the publication page of the II Physics Institute of the University of Giessen [69] or the Belle II Document Server), but were considered impractical for the final input vector to the neural networks. It is, nevertheless, worth testing new combinations of cluster properties for the input vector in order to gain a better understanding of their effect on the separation power of neural networks.

A feasibility study of identifying HIPs with the PXD should be repeated on real data. In particular, if identification is targeted on exotic particles like magnetic monopoles, the capabilities of the PXD should be evaluated with a reference signal. Particles like anti-protons for instance possess an average energy deposition which is estimated to be higher than the average energy deposition of background particles. Additionally, the range of anti-protons in the detector suffices to detect them with the outer sub-detectors. A *conventional* Belle II analysis on anti-protons could therefore serve to validate the identification capabilities of the PXD investigating the same channel. With layer 1 completely assembled, early Phase III could be used for this endeavor, once the PXD read-out is confirmed to run smoothly.

APPENDIX

A Evaluation of Self-Organizing Maps

A.1 U-matrix algorithm

The evaluation of SOMs requires the identification of clusters or regions within the trained map. Most commonly, a U-matrix algorithm is used to visualize the emergence of clusters [129]. For each weight vector associated with a node the distance to its neighbors is computed. Minima in the distance function correspond to clusters in feature space.

The normalized distance function for the training of anti-deuterons with momentum $p < 3\,\mathrm{GeV}$ and a node space of 12 x 12 nodes is shown in Fig. 1. Note the modification of the size of node space, which was reduced from 15 x 15 nodes to 12 x 12 nodes for visualization purposes. In Fig. 2 the classification as function of nodes for the same SOM is depicted. The classification map is a familiar representation of the training results, which allows to assess the utility of the U-matrix approach.

Similar patterns in the distance and classification map are observable. Background-like regions, characterized by low classification values, correspond to high values in the distance map. For signal-like regions low distance values are adopted. On

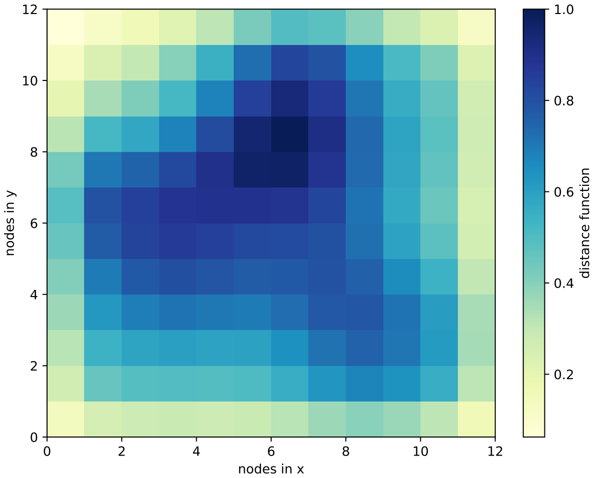


Figure 1: Distance function of SOM trained with $p < 3\,\mathrm{GeV}$ anti-deuteron cluster properties. High values correspond to a large distance between adjacent nodes, low values to small distances. Small distance values at the edges of the map are observable.

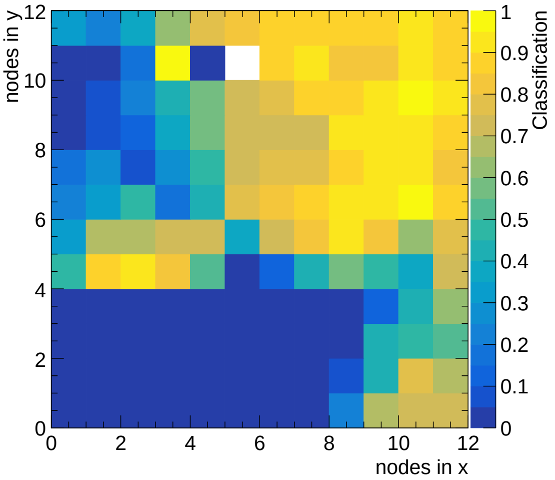


Figure 2: Classification map of SOM trained with $p < 3\,\mathrm{GeV}$ anti-deuteron cluster properties. High classification values are associated with anti-deuterons, low values with background. When compared with Fig. 1 similar patterns can be identified.

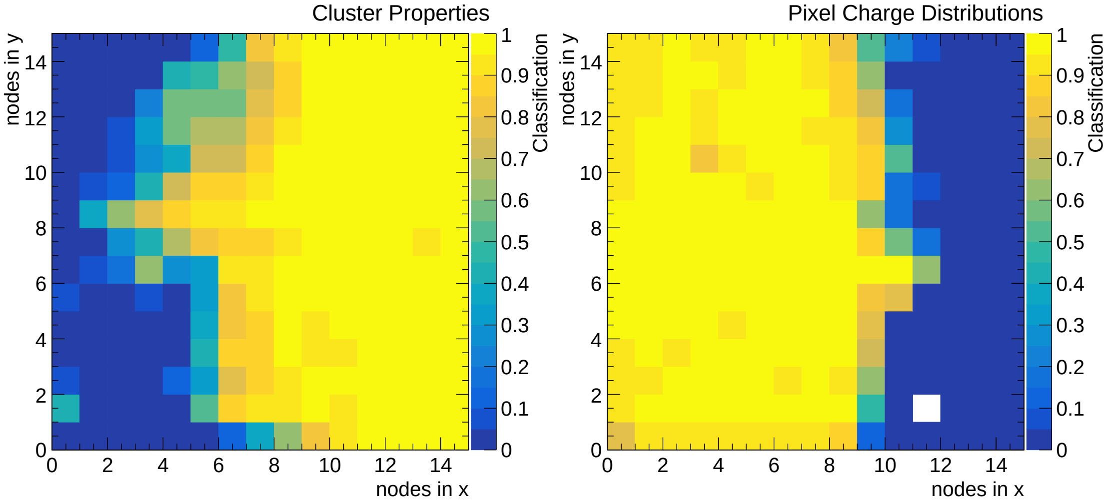
close inspection, however, there are severe discrepancies between the two representations. In particular, boundary effects are observable arising from a reduced number of neighbors for nodes on the edges of the map. The small node spaces employed in this thesis promote the appearance of boundary issues rendering evaluation based on the distance map difficult.

While distance maps based on the U-matrix are generally a valuable tool for the evaluation of SOMs, they are inadequate for the approach chosen in this thesis. Classification maps deliver more reliable results.

A.2 SOM Cluster Size

Nodes as well as feature vectors of the SOM are initialized randomly before every training. In principle, the initial SOM state should not affect the final trained state except for the specific location, at which clusters form. The size of clusters/regions for background and signal (either anti-deuterons or magnetic monopoles) is not equal. Classification maps of the SOMs trained for this thesis consist of 1/3 background-like region and about 2/3 signal-like region as shown in Fig. 3. On the left-hand side the classification map for sub-GeV anti-deuterons trained with cluster properties and on the right-hand side the map for the same particle species trained with pixel charge distributions is depicted. In both cases the anti-deuteron region is significantly larger. If the training process is repeated with a different initial map state, the regions can form at other locations of the map, their size, however, stays unaltered indicating that the number of nodes allocated to the regions is determined by the input vectors and not the network architecture itself. If solely the network architecture was responsible, the size distribution would not be reproducible and the background region would dominate in size just as often as the signal region.

It stands to reason that more nodes are required to capture the parameter space covered by anti-deuteron clusters. This holds true for certain cluster properties such as the total cluster charge, but fails for other parameters like the cluster size, which exhibits much more variety in background clusters. The occurrence of linear correlations could diminish the number of processing units required to describe a class of clusters. Similar to the results shown in Section 5.4 linear correlations of cluster properties are slightly different for background and anti-deuterons. However, the differences cannot be correlated with the size mismatch in the classification map.



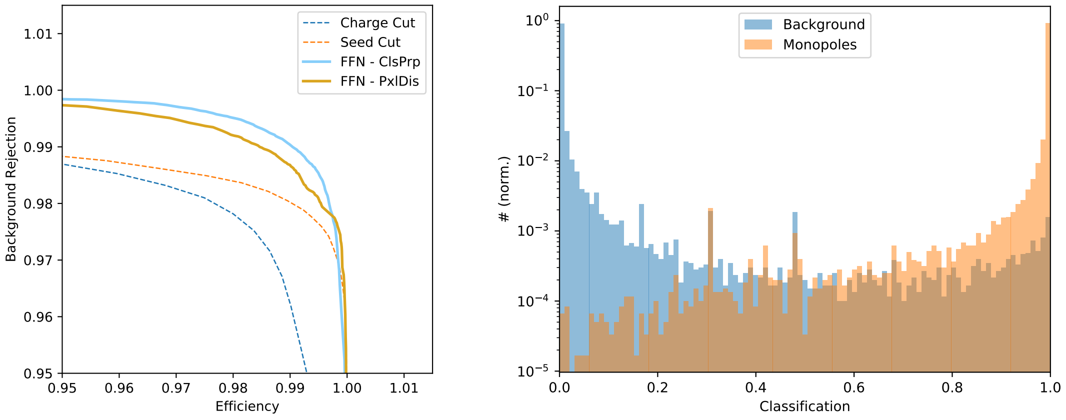
(a) (b)

Figure 3: Classification maps for SOMs trained with cluster properties (left-hand side) and pixel charge distributions (right-hand side) of clusters generated by $p < 1\,\mathrm{GeV}$ anti-deuterons and background. The size of the anti-deuteron-like region is larger than the background region in both cases. The position of the regions (right- or left-hand side of the map) is determined by the initial conditions and can therefore vary from training to training.

B Evaluation of Feed-Forward Networks

While the FFNs presented in this thesis deliver satisfactory results, there are some irregularities, which are believed to arise either from the network architecture or the input vector set.

In Fig. 4a the oscillatory behavior of the ROC curve belonging to the FFN trained with pixel charge distributions of monopole clusters is shown once again. The classification curve for the FFN is depicted in Fig. 4b. Three prominent spikes are observable. Presumably the emergence of these spikes is responsible for the oscillations in the ROC curve. In general, all classification distributions presented in this thesis exhibit a lack of smoothness. The distributions appear to suffer from quantization issues. The hyperparameters have been modified one at a time and the issue was found to remain. Systematic variations of all network parameters at once are required to investigate whether an interplay of certain hyperparameters induces the quantization problem or if the issue emerges from the properties of the



(a)

(b)

Figure 4: ROC curve and classifier distribution for pixel charge distribution of magnetic monopole clusters. (a) The ROC curve for training with the pixel charge distribution is below the curve for training with cluster properties. The pixel charge curves exhibits an oscillatory behavior in an efficiency range of about 1%. (b) The classification distributions have prominent *spikes* around classification values of 0.1 to 0.5.

input vector sets.

The training process for all FNNs is halted after 1000 epochs as discussed above. Instead of applying a generic truncation limit, it would be more suitable, to adjust the cut-off to for each training individually. In Fig. 5 the accuracy monitored during training with an input vector set consisting of cluster properties from 1 GeV monopoles and background is depicted. The accuracy for the training set exhibits a shallow rise. The testing set, despite being much more volatile, follows the rising trend of the training set indicating that the truncation after 1000 epochs is too early. A similar behavior is observable for training with anti-deuteron clusters. A training process involving 1000 epochs is finished in ~1 day and yields an accuracy of 97.5%/97.4% for training/testing set and a cut at 0.5 on the classification axis. If training is truncated after 10,000 epochs the accuracy is 97.6%/97.6%, but the training takes over one week.

A neural network is said to be in *overtraining*, when the accuracy for the training set increases, while the accuracy of the testing set decreases. The neural network becomes specialized in the training set, while suffering from a lack of generalization ability. The absence of overtraining in all FNNs trained for this thesis hints at

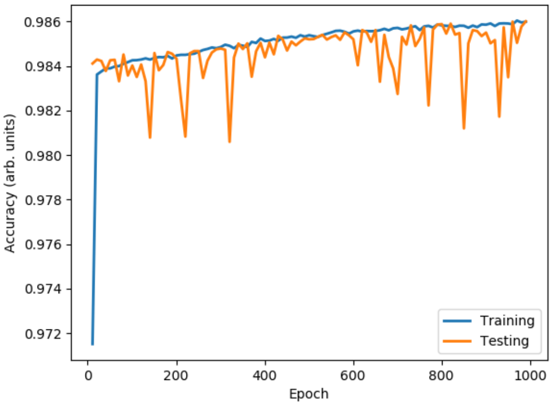


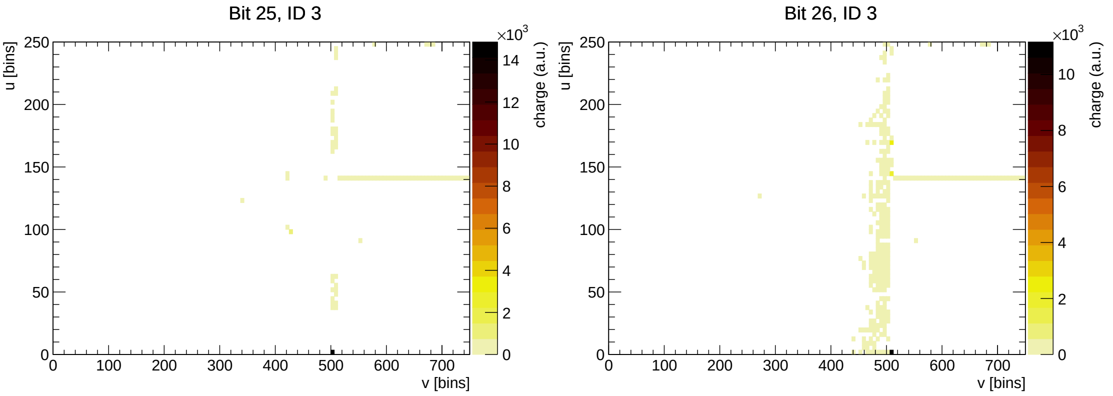
Figure 5: Accuracy of FFN trained with cluster properties of 1 GeV magnetic monopoles as a function of training epoch. The classification cut is made at 0.5. The accuracy of the training set exhibits a shallow increase with increasing epochs. The accuracy for the testing set also depicts a shallow increase, but is more volatile.

the existence of a counter-mechanism preventing overtraining. Certain types of clusters are not unique to background or signal. Consequently, there are vectors in the training set, which appear more than once but with different classification labels. In particular, single pixel clusters are affected since the pixel charge is the only identification criterion. Learning is impaired by the ambiguity of certain clusters preventing the onset of overtraining.

An optimization of the network hyperparameters and truncation limit is certainly necessary to give reliable statements as to the performance of the FFNs. Nevertheless, the excellent performance of FFNs despite the less than optimal network architecture, already demonstrates the applicability of neural networks for the separation of highly ionizing particles from background.

C PXD Data from Phase II

The analysis of PXD data from Phase II is impaired by the low geometrical coverage of the available PXD modules. Only four of the forty modules were installed rendering the detection of a monopole event signature (two clusters associated with monopoles at opposite sides of the PXD) impossible.



(a)

(b)

Figure 6: Activated pixels in PXD module 1.1.2. during two randomly triggered Phase II events. Trigger Bit 25 corresponds to the Bhabha trigger and Bit 26 to the pseudorandom trigger (see Section 3.7.1). Both events feature a single-pixel band stretching from $v \sim 500$ to the end of the module. Another broader band along the entire sensor in u direction is observable. The band is more pronounced for trigger Bit 26.

Data from random trigger events (see Section 3.7.1) are searched for large clusters with a high cluster charge. The resulting set of events is *contaminated* by pixel charge distributions shown in Fig. 6. The activated pixels of PXD module 1.1.2 (labeled as ID 3 during Phase II) for a Bhabha triggered (Bit 25) and a pseudorandom triggered (Bit 26) event are depicted. In both events a band of activated pixels stretching from ~500 to the end of the module is observable. In addition, a broader band along the u direction at $v \sim 500$ can be seen, which is more pronounced for the pseudorandom event . Similar distributions of activated pixels can be found multiple times in the data set indicating that there is an issue with the read-out electronics or the pixel cells themselves. In particular, the frequent occurrence of the thin band in v direction at the same position is unlikely to originate from a charged particle.

As long as the response and the systematic uncertainties are not well-understood a search for HIPs is not feasible. A HIP study has to be postponed until the PXD is fully operational in Phase III. Despite the missing modules in the second layer, the first layer is fully assembled. The geometrical coverage with the first layer is therefore optimal.

D Cluster Properties as Function of Particle Momentum

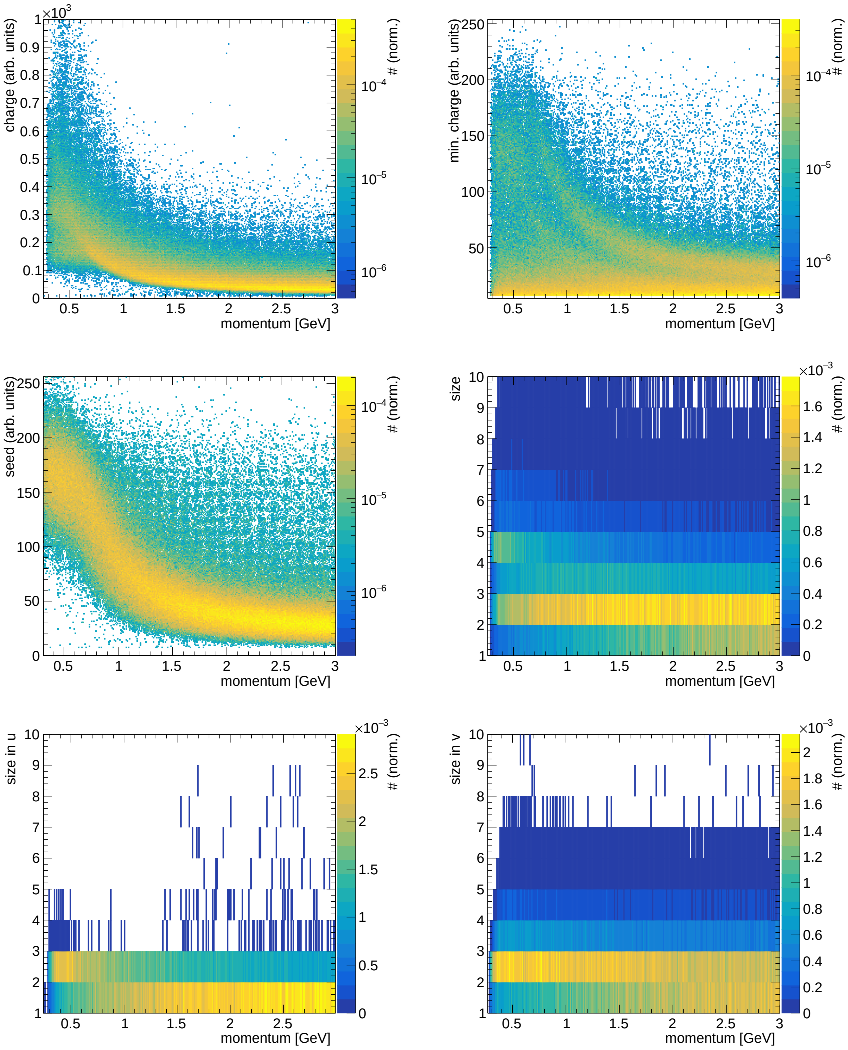


Figure 7: Cluster properties as a function of momentum of the incident particle. Charge properties (total cluster charge, seed charge and minimum pixel charge) display bands, which are reminiscent of Bethe-Bloch curves. The impact of particle momentum on size properties (total size, size in u, size in v) is less pronounced.

E One Dimensional Cluster Property Distributions

E.1 Anti-Deuterons with $p < 1\,\mathrm{GeV}$

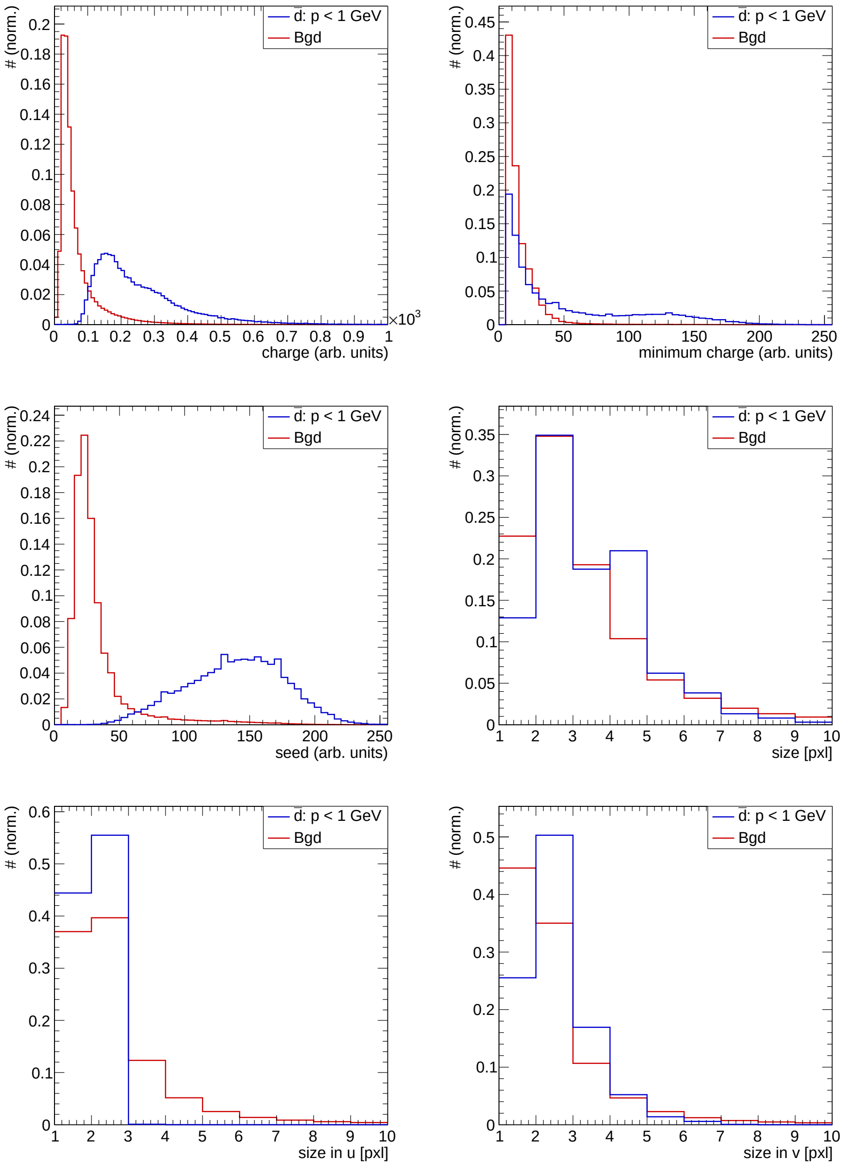


Figure 8: Cluster property distributions for anti-deuterons with $p < 1\,\mathrm{GeV}$ and background.

E.2 Anti-Deuterons with $2\,\mathrm{GeV} < p < 3\,\mathrm{GeV}$

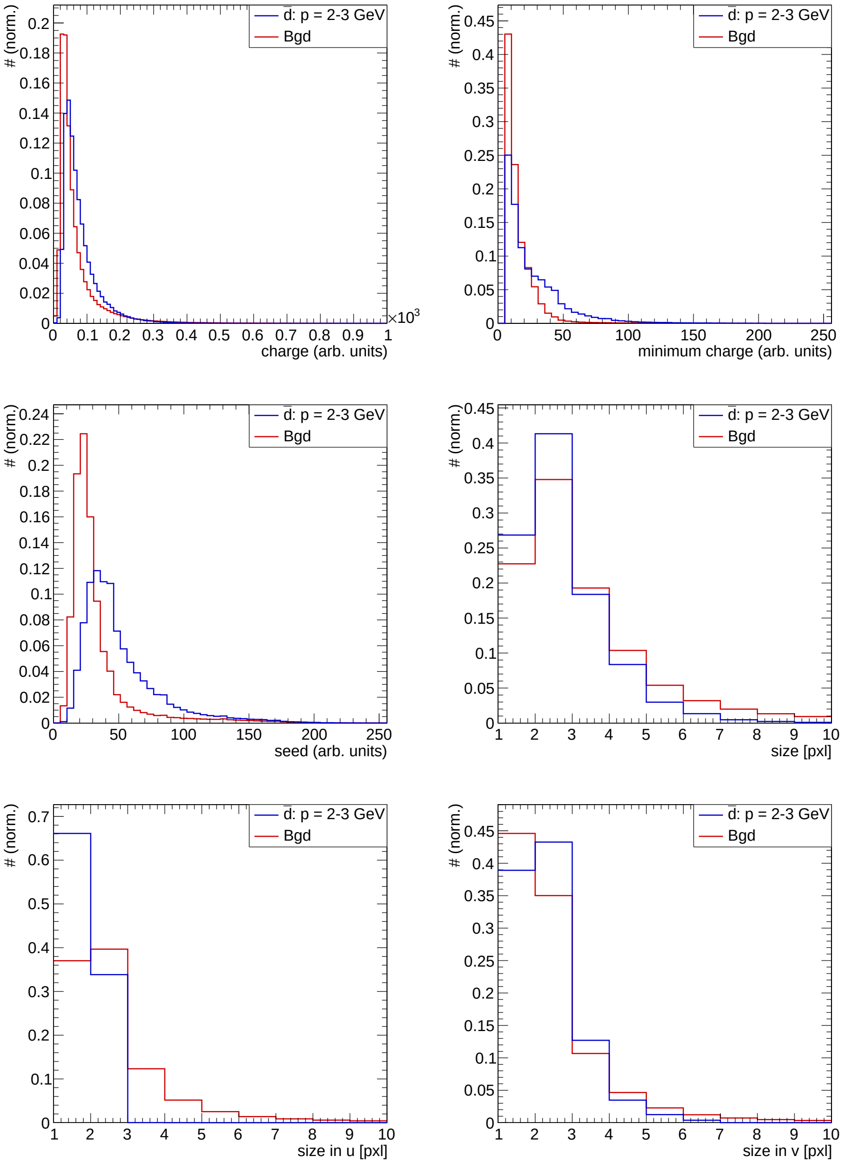


Figure 9: Cluster property distributions for anti-deuterons with $2\,\mathrm{GeV} < p < 3\,\mathrm{GeV}$ and background.

E.3 Magnetic Monopoles vs Background

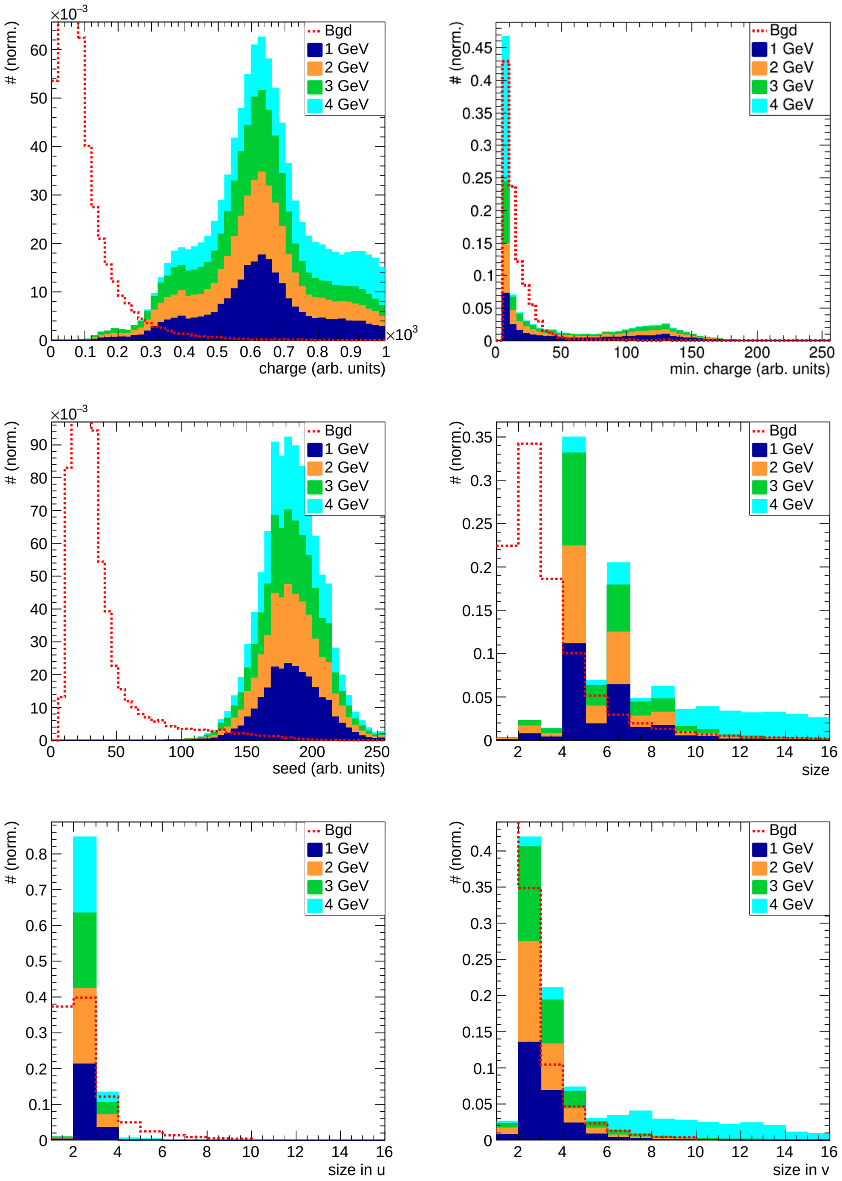


Figure 10: Cluster property distributions for magnetic monopoels and background. The monopole distributions for different masses is stacked.

E.4 Magnetic Monopoles vs Muons

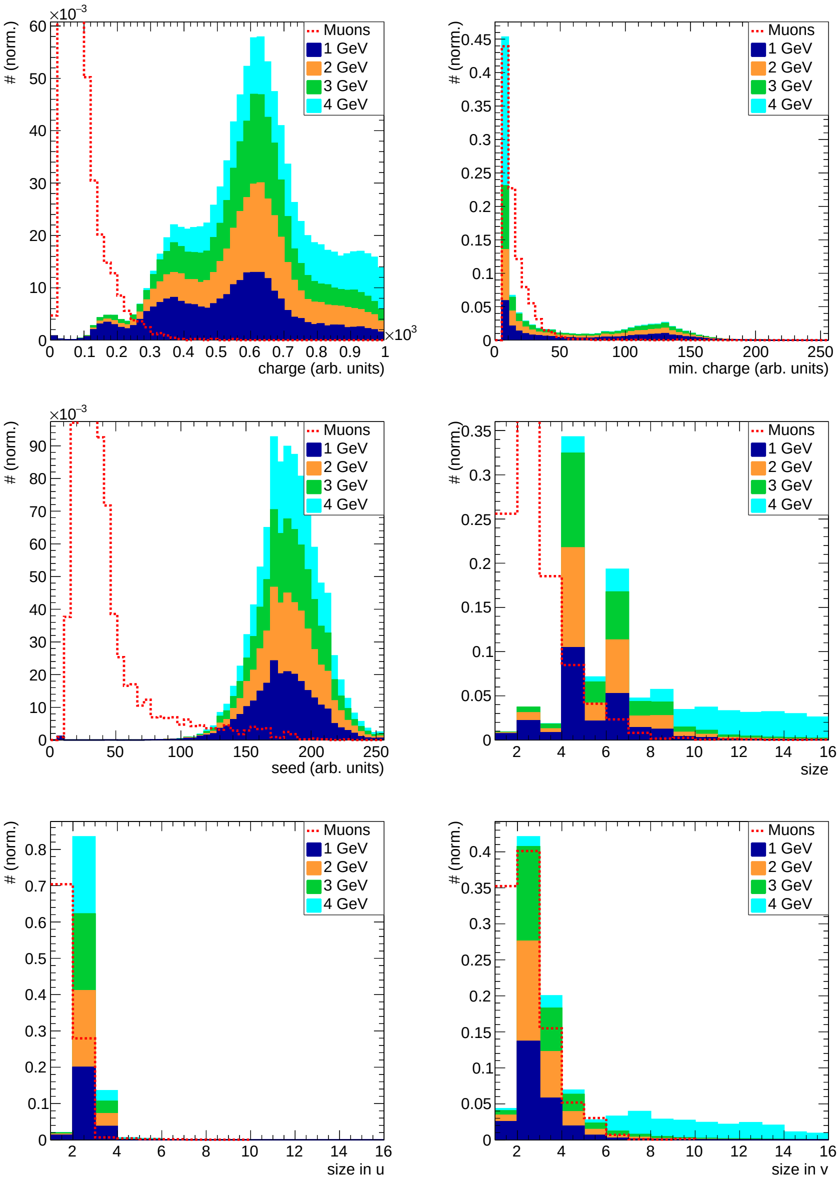


Figure 11: Cluster property distributions for magnetic monopoels and muons.

F Two Dimensional Cluster Property Distributions

F.1 Anti-Deuterons with $p < 1\,\mathrm{GeV}$

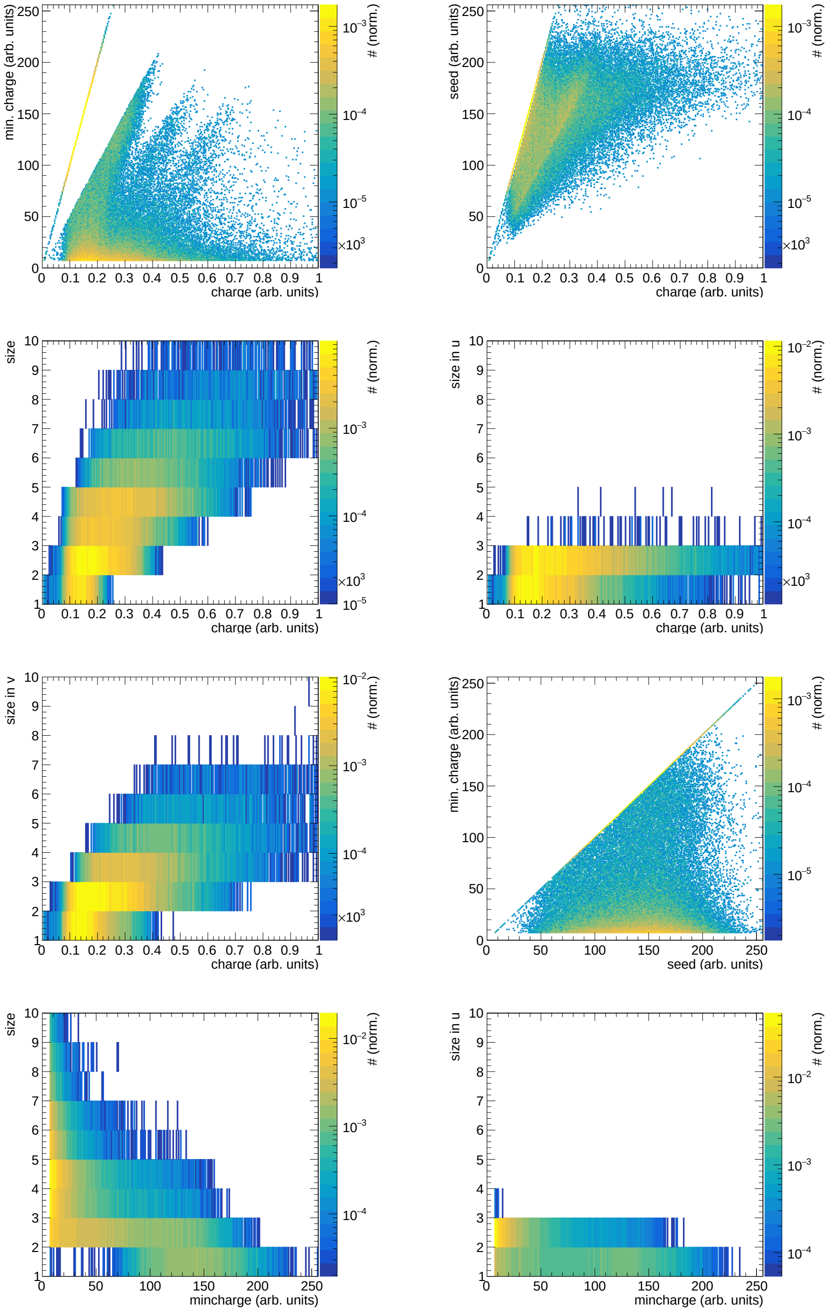


Figure 12: Two dimensional distributions for anti-deuterons with $p < 1\,\mathrm{GeV}$.

F.2 Anti-Deuterons with $2\,\mathrm{GeV} < p < 3\,\mathrm{GeV}$

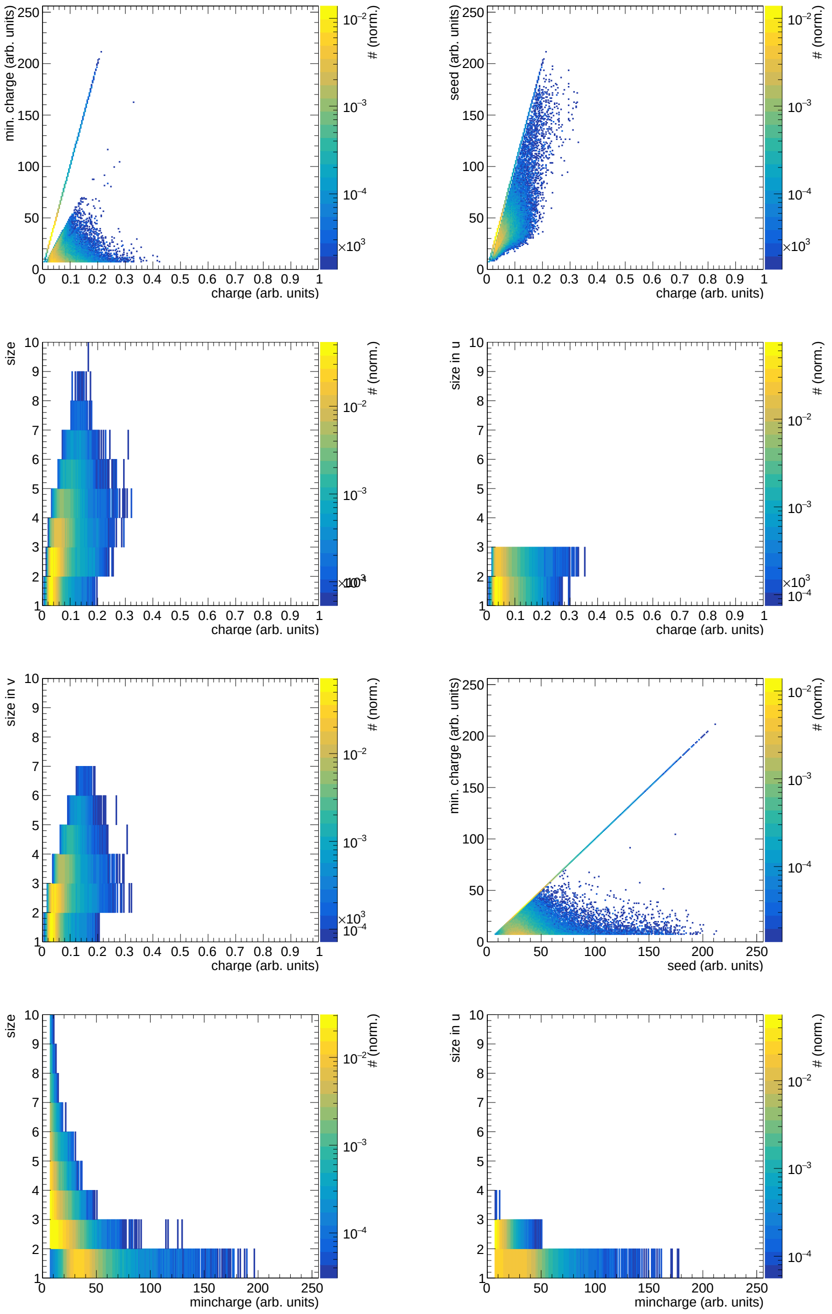


Figure 13: Two dimensional distributions for all combinations of cluster properties for anti-deuterons with momentum $2\,\mathrm{GeV} < p < 3\,\mathrm{GeV}$.

F.3 Magnetic Monopoles with $m = 1\,\mathrm{GeV}$

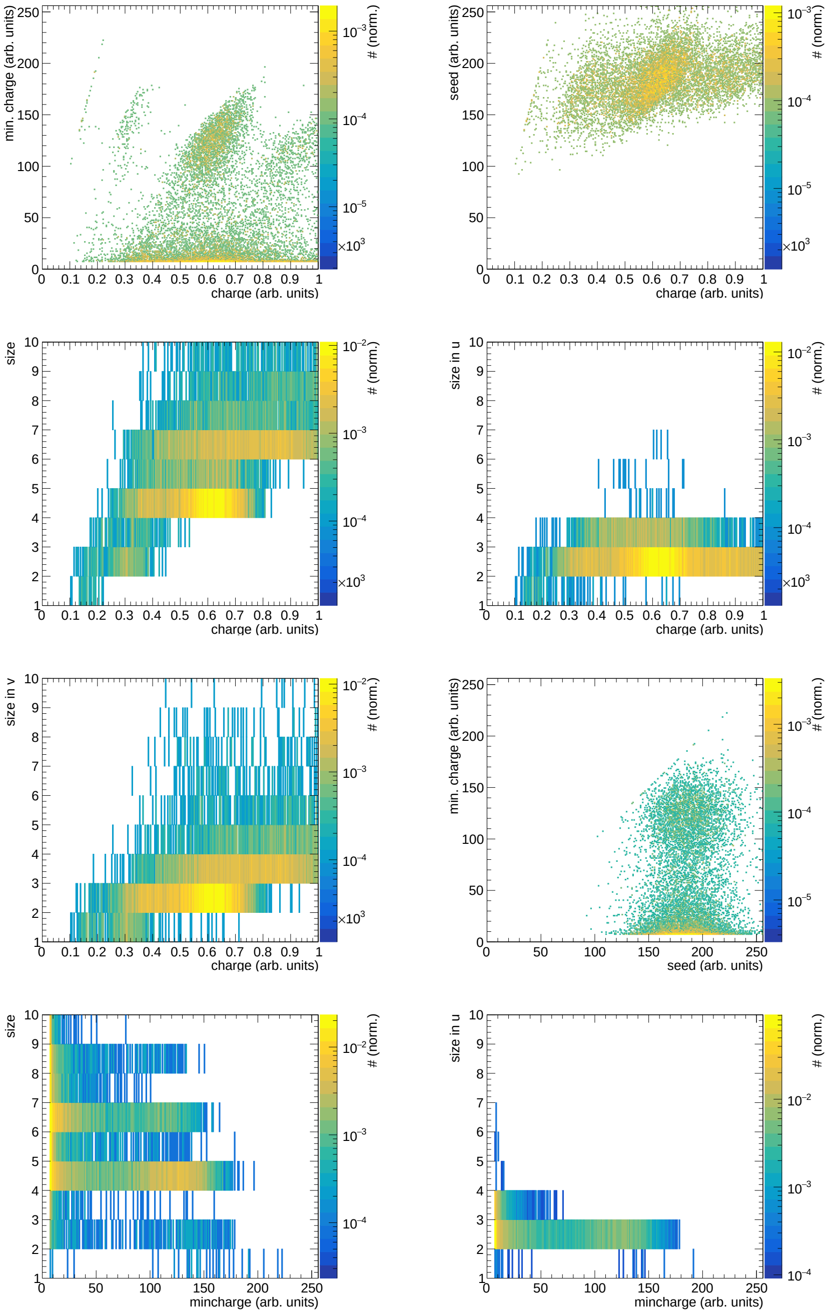


Figure 14: Two dimensional distributions for all combinations of cluster properties for magnetic monopoles with mass $m = 1\,\mathrm{GeV}$.

F.4 Magnetic Monopoles with $m = 4\,\mathrm{GeV}$

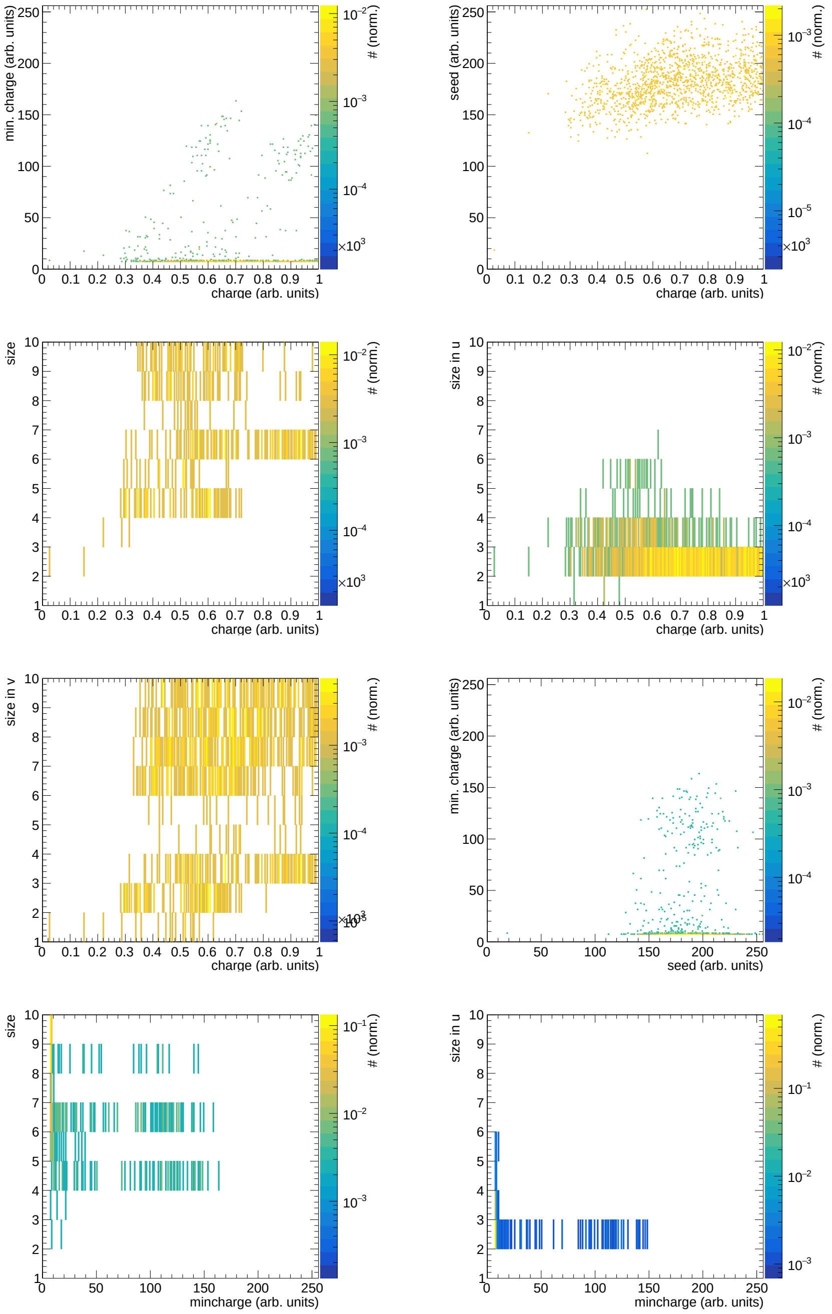


Figure 15: Two dimensional distributions for all combinations of cluster properties for magnetic monopoles with mass $m = 4\,\mathrm{GeV}$.

F.5 Background

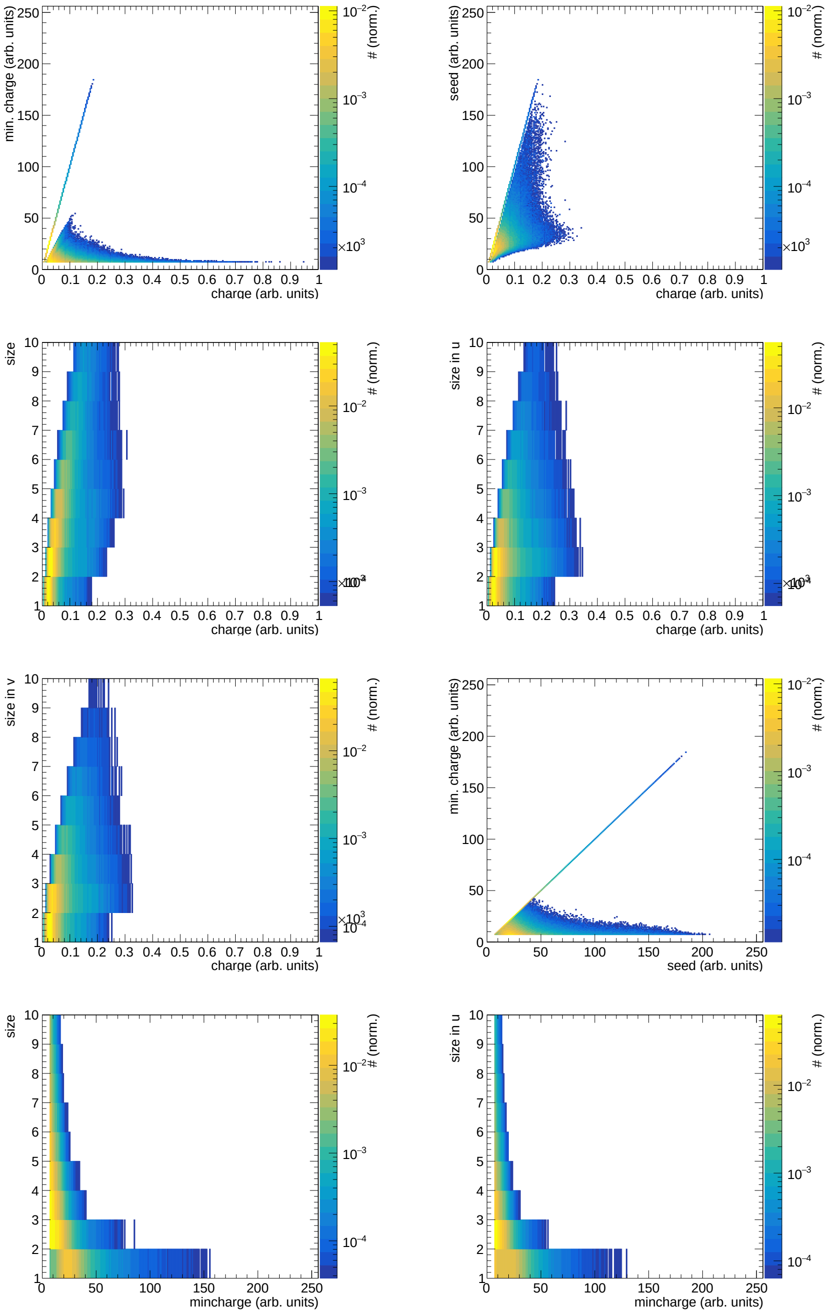


Figure 16: Two dimensional distributions for all combinations of cluster properties for background.

BIBLIOGRAPHY

[1] G. De Pietro, "First data at Belle II and Dark Sector physics," *arXiv preprint arXiv:1808.00776*, 2018.

[2] M. J. Dolan, T. Ferber, C. Hearty, F. Kahlhoefer, and K. Schmidt-Hoberg, "Revised constraints and Belle II sensitivity for visible and invisible axion-like particles," *Journal of High Energy Physics*, vol. 2017, no. 12, p. 94, 2017.

[3] "SuperKEKB and Belle II." `https://www.belle2.org/project/super_kekb_and_belle_ii`. Accessed: 2018-11-22.

[4] "Belle II Detector." `http://belle2.kek.jp/images/BelleII3D.pdf`. Accessed: 2018-11-28.

[5] K. Gärtner and R. H. Richter, "DEPFET sensor design using an experimental 3D device simulator," *Nuclear Instruments and Methods in Physics Research Section A*, vol. 568, no. 1, pp. 12–17, 2006.

[6] A. Moll, "Comprehensive study of the background for the Pixel Vertex Detector at Belle II," July 2015.

[7] H. Ye, C. Niebuhr, R. Stever, K. Gadow, and C. Camien, "Thermal mock-up studies of the Belle II vertex detector," *Nucl. Instrum. Meth.*, vol. A896, pp. 82–89, 2018.

[8] T. Abe, I. Adachi, K. Adamczyk, S. Ahn, H. Aihara, K. Akai, M. Aloi, L. Andricek, K. Aoki, Y. Arai, *et al.*, "Belle II technical design report," *arXiv preprint arXiv:1011.0352*, 2010.

[9] M. Tanabashi, K. Hagiwara, K. Hikasa, K. Nakamura, Y. Sumino, F. Takahashi, J. Tanaka, K. Agashe, G. Aielli, C. Amsler, *et al.*, "Review of particle physics," *Physical Review D*, vol. 98, no. 3, p. 030001, 2018.

[10] D. Griffiths, H. Rollnik, and T. Stange, *Einführung in die Elementarteilchenphysik.* Wiley, 1999.

[11] W. Demtroeder, "Experimentalphysik 4," 2009.

[12] F. Schwabl, R. Hilton, and A. Lahee, *Advanced Quantum Mechanics.* Advanced texts in physics, Springer, 2005.

[13] E. Fermi, *Nuclear physics: a course given by Enrico Fermi at the University of Chicago.* University of Chicago Press, 1950.

[14] D. Clowe, M. Bradač, A. H. Gonzalez, M. Markevitch, S. W. Randall, C. Jones, and D. Zaritsky, "A direct empirical proof of the existence of dark matter," *The Astrophysical Journal Letters*, vol. 648, no. 2, p. L109, 2006.

[15] N. Jarosik, C. Bennett, J. Dunkley, B. Gold, M. Greason, M. Halpern, R. Hill, G. Hinshaw, A. Kogut, E. Komatsu, *et al.*, "Seven-year wilkinson microwave anisotropy probe (WMAP) observations: sky maps, systematic errors, and basic results," *The Astrophysical Journal Supplement Series*, vol. 192, no. 2, p. 14, 2011.

[16] A. G. Riess, A. V. Filippenko, P. Challis, A. Clocchiatti, A. Diercks, P. M. Garnavich, R. L. Gilliland, C. J. Hogan, S. Jha, R. P. Kirshner, *et al.*, "Observational evidence from supernovae for an accelerating universe and a cosmological constant," *The Astronomical Journal*, vol. 116, no. 3, p. 1009, 1998.

[17] M. Dine and A. Kusenko, "Origin of the matter-antimatter asymmetry," *Reviews of Modern Physics*, vol. 76, no. 1, p. 1, 2003.

[18] A. D. Sakharov, "Violation of cp invariance, c asymmetry, and baryon asymmetry of the universe," *JETP lett.*, vol. 5, pp. 24–27, 1967.

[19] S. Dimopoulos, S. Raby, and F. Wilczek, "Supersymmetry and the scale of unification," *Physical Review D*, vol. 24, no. 6, p. 1681, 1981.

[20] K. R. Dienes, "String theory and the path to unification: a review of recent developments," *Physics Reports*, vol. 287, no. 6, pp. 447–525, 1997.

[21] G. G. Ross, *Grand unified theories.* Frontiers in physics, Benjamin/Cummings Pub. Co., 1985.

[22] J. Bennett, "The Belle II Experiment: Status and Physics Prospects," in *International Journal of Modern Physics: Conference Series*, vol. 46, p. 1860082, World Scientific, 2018.

[23] W.-S. Hou, "Enhanced charged higgs boson effects in $B^- \to \tau\bar{\nu}$, $\mu\bar{\nu}$ and $\bar{b} \to \tau\bar{\nu} + x$," *Physical Review D*, vol. 48, no. 5, p. 2342, 1993.

[24] K.-T. Chao, Y. Wang, *et al.*, *Physics at BES-III.* IHEP-Physics-Report-BES-III-2008-001-v1, 2008.

[25] S. Godfrey and S. L. Olsen, "The exotic XYZ charmonium-like mesons," *Annual Review of Nuclear and Particle Science*, vol. 58, pp. 51–73, 2008.

[26] B. Holdom, "Searching for ϵ charges and a new U(1)," *Physics Letters B*, vol. 178, no. 1, pp. 65–70, 1986.

[27] T. Ferber, "Towards First Physics at Belle II," *Acta Physica Polonica B*, vol. 46, p. 2285, 11 2015.

[28] M. Dine, W. Fischler, and M. Srednicki, "A simple solution to the strong CP problem with a harmless axion," *Physics letters B*, vol. 104, no. 3, pp. 199–202, 1981.

[29] B. Shuve and I. Yavin, "Dark matter progenitor: Light vector boson decay into sterile neutrinos," *Physical Review D*, vol. 89, no. 11, p. 113004, 2014.

[30] P. Branchini, "The Belle II Experiment: Status and Prospects," *Universe*, vol. 4, no. 10, p. 101, 2018.

[31] M. Bračko, S. Collaboration, *et al.*, "Status of the SuperKEKB factory," vol. 171, no. 1, p. 012098, 2009.

[32] M. Kikuchi, T. Abe, K. Egawa, H. Fukuma, K. Furukawa, N. Iida, H. Ikeda, T. Kamitani, K.-i. Kanazawa, K. Ohmi, *et al.*, "Design of positron damping ring for Super-KEKB," *Proceedings of IPAC*, vol. 10, p. 1641, 2010.

[33] Y. Ohnishi, T. Abe, T. Adachi, K. Akai, Y. Arimoto, K. Ebihara, K. Egawa, J. Flanagan, H. Fukuma, Y. Funakoshi, *et al.*, "Accelerator design at SuperKEKB," *Progress of Theoretical and Experimental Physics*, vol. 2013, no. 3, 2013.

[34] P. Raimondi, "New developments in super B-factories," in *2007 IEEE Particle Accelerator Conference (PAC)*, pp. 32–36, IEEE, 2007.

[35] C. Marinas, D. Collaboration, *et al.*, "The Belle II pixel detector: High precision with low material," *Nuclear Instruments and Methods in Physics Research Section A*, vol. 731, pp. 31–35, 2013.

[36] G. Caria, S. Mayekar, S. Watanuki, F. Bosi, C. Joo, D. Dutta, G. De Pietro, A. Bozek, A. Bauer, H. Park, *et al.*, "The Software Framework of the Belle II Silicon Vertex Detector and its Development for the 2016 Test-Beam at DESY," *PoS*, p. 060, 2017.

[37] G. Lutz *et al.*, *Semiconductor radiation detectors*, vol. 40. Springer, 1999.

[38] K. Emi, M. Tomoto, Y. Tamagawa, T. Takahashi, Y. Sakai, T. Yamaki, O. Nitoh, S. Uno, A. Sugiyama, H. Hirano, *et al.*, "Study of a dE/dx measurement and the gas-gain saturation by a prototype drift chamber for the BELLE-CDC," *Nucl. Instrum. Methods Phys. Res., A*, vol. 379, no. DPNU-96-04, p. 225, 1996.

[39] S. Uno, Y. Sakai, T. Tsukamoto, K. Abe, I. Adachi, T. Nozaki, M. Saitoh, O. Sasaki, H. Yamaoka, K. Emi, *et al.*, "Study of a drift chamber filled with a helium-ethane mixture," *Nuclear Instruments and Methods in Physics Research Section A*, vol. 330, no. 1-2, pp. 55–63, 1993.

[40] P. A. Cherenkov, "Visible emission of clean liquids by action of γ radiation," *Doklady Akademii Nauk SSSR*, vol. 2, p. 451, 1934.

[41] T. Matsumoto, S. Korpar, I. Adachi, S. Fratina, T. Iijima, R. Ishibashi, H. Kawai, P. Križan, S. Ogawa, R. Pestotnik, *et al.*, "Studies of proximity focusing RICH with an aerogel radiator using flat-panel multi-anode PMTs (Hamamatsu H8500)," *Nuclear Instruments and Methods in Physics Research Section A*, vol. 521, no. 2-3, pp. 367–377, 2004.

[42] S. S. Kistler, "Coherent expanded aerogels and jellies," *Nature*, vol. 127, no. 3211, p. 741, 1931.

[43] Y. Jin, H. Aihara, O. Borshchev, D. Epifanov, S. Ponomarenko, and N. Surin, "Study of a pure CsI crystal readout by APD for Belle II end cap ECL upgrade," *Nuclear Instruments and Methods in Physics Research Section A*, vol. 824, pp. 691–692, 2016.

[44] C. Grupen and B. Shwartz, *Particle detectors.* Cambridge University Press, 2008.

[45] T. Aushev, D. Besson, K. Chilikin, R. Chistov, M. Danilov, P. Katrenko, R. Mizuk, G. Pakhlova, P. Pakhlov, V. Rusinov, *et al.*, "A scintillator based endcap KL and muon detector for the Belle II experiment," *Nuclear Instruments and Methods in Physics Research Section A*, vol. 789, pp. 134–142, 2015.

[46] Y. Iwasaki, B. Cheon, E. Won, and G. Varner, "Level 1 trigger system for the Belle II experiment," in *2010 17th IEEE-NPSS Real Time Conference*, pp. 1–9, IEEE, 2010.

[47] D. Sun, J. Zhao, H. Xu, *et al.*, "Belle2link: A global data readout and transmission for Belle II experiment at KEK," *Physics Procedia*, vol. 37, pp. 1933–1939, 2012.

[48] R. H. Richter, L. Andricek, P. Fischer, K. Heinzinger, P. Lechner, G. Lutz, I. Peric, M. Reiche, G. Schaller, M. Schnecke, *et al.*, "Design and technology of DEPFET pixel sensors for linear collider applications," *Nuclear Instruments and Methods in Physics Research Section A*, vol. 511, no. 1-2, pp. 250–256, 2003.

[49] O. Neufang, "Lexikon der Elektronik," 2013.

[50] I. Perić, P. Fischer, J. Knopf, and T. H. H. Nguyen, "DCDB and SWITCHERB, the readout ASICS for Belle II DEPFET pixel detector," in *Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 2011 IEEE*, pp. 1536–1539, IEEE, 2011.

[51] H. Krüger, D. Collaboration, *et al.*, "Front-end electronics for DEPFET pixel detectors at SuperBelle (BELLE II)," *Nuclear Instruments and Methods in Physics Research Section A*, vol. 617, no. 1-3, pp. 337–341, 2010.

[52] O. Alonso, R. Casanova, A. Dieguez, J. Dingfelder, T. Hemperek, T. Kishishita, T. Kleinohl, M. Koch, H. Krüger, M. Lemarenko, *et al.*, "DEPFET active pixel detectors for a future linear electron positron collider," *IEEE Transactions on Nuclear Science*, vol. 60, no. 2, pp. 1457–1465, 2013.

[53] T. Geßler, W. Kühn, J. S. Lange, Z. Liu, D. Münchow, B. Spruck, and J. Zhao, "The ONSEN data reduction system for the Belle II pixel detector," *IEEE Transactions on Nuclear Science*, vol. 62, no. 3, pp. 1149–1154, 2015.

[54] D. Levit, I. Konorov, D. Greenwald, and S. Paul, "Fpga based data readout system of the Belle II pixel detector," *IEEE Transactions on Nuclear Science*, vol. 62, no. 3, pp. 1033–1039, 2015.

[55] S. Hashimoto, M. Hazumi, J. Haba, J. Flanagan, and Y. Ohnishi, "Letter of intent for KEK Super B Factory," tech. rep., High Energy Accelerator Research Organization, 2004.

[56] A. Piwinski, "The Touschek effect in strong focusing storage rings," *arXiv preprint physics/9903034*, 1999.

[57] J. Schwinger, "On the classical radiation of accelerated electrons," *Physical review*, vol. 75, no. 12, p. 1912, 1949.

[58] B. L. Berman and S. Fultz, "Measurements of the giant dipole resonance with monoenergetic photons," *Reviews of Modern Physics*, vol. 47, no. 3, p. 713, 1975.

[59] T. Kuhr, C. Pulvermacher, M. Ritter, T. Hauth, and N. Braun, "The Belle II Core Software," *Computing and Software for Big Science*, vol. 3, no. 1, p. 1, 2019.

[60] I. Adachi, T. Hibino, L. Hinz, R. Itoh, N. Katayama, S. Nishida, F. Ronga, T. Tsukamoto, and M. Yokoyama, "Belle computing system," *Nuclear Instruments and Methods in Physics Research Section A*, vol. 534, no. 1-2, pp. 53–58, 2004.

[61] A. Moll, "The software framework of the Belle II experiment," in *Journal of Physics: Conference Series*, vol. 331, p. 032024, IOP Publishing, 2011.

[62] D. J. Lange, "The EvtGen particle decay simulation package," *Nuclear Instruments and Methods in Physics Research Section A*, vol. 462, no. 1-2, pp. 152–155, 2001.

[63] S. Agostinelli, J. Allison, K. a. Amako, J. Apostolakis, H. Araujo, P. Arce, M. Asai, D. Axen, S. Banerjee, G. Barrand, *et al.*, "GEANT4—a simulation toolkit," *Nuclear instruments and methods in physics research section A*, vol. 506, no. 3, pp. 250–303, 2003.

[64] Z. Drasal, B. Schwenker, and K. Prothmann, "Silicon simulation code for Belle II and ILC," *PoS*, p. 027, 2011.

[65] R. Turchetta, "Spatial resolution of silicon microstrip detectors," *Nuclear Instruments and Methods in Physics Research Section A*, vol. 335, no. 1-2, pp. 44–58, 1993.

[66] L. Koch, "Study of improved K_S^0 detection for the Belle II detector," Master's thesis, University of Giessen, 2015.

[67] D. Münchow, *Development of the Online DataReduction System and Feasibility Studies of 6-Layer Tracking for the Belle II Pixel Detector*. PhD thesis, University of Giessen, 2015.

[68] S. Bähr, O. Sander, M. Heck, C. Pulvermacher, M. Feindt, and J. Becker, "Online-analysis of hits in the Belle-II pixeldetector for separation of slow pions from background," in *Journal of Physics: Conference Series*, vol. 664, p. 092001, IOP Publishing, 2015.

[69] "JLU Giessen II Physics Institute Publication Page." `http://fb07-indico.physik.uni-giessen.de:8080/wiki/index.php/Category:Publications`. Accessed: 2019-04-18.

[70] M. Hassoun, *Fundamentals of artificial neural networks*. MIT press, 1995.

[71] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, "Automatic differentiation in PyTorch," 2017.

[72] J. Sanders and E. Kandrot, *CUDA by example: an introduction to general-purpose GPU programming*. Addison-Wesley Professional, 2010.

[73] F.-M. Schleif and T. Villmann, "Miwoci workshop-2018," 2013.

[74] "Self Organizing Maps - MiniSom." `https://github.com/JustGlowing/minisom`. Accessed: 2019-01-29.

[75] J. A. Hanley and B. J. McNeil, "The meaning and use of the area under a receiver operating characteristic (roc) curve.," *Radiology*, vol. 143, no. 1, pp. 29–36, 1982.

[76] T. Aramaki, S. Boggs, S. Bufalino, L. Dal, P. Von Doetinchem, F. Donato, N. Fornengo, H. Fuke, M. Grefe, C. Hailey, *et al.*, "Review of the theoretical and experimental status of dark matter identification with cosmic-ray antideuterons," *Physics Reports*, vol. 618, pp. 1–37, 2016.

[77] F. Donato, N. Fornengo, and P. Salati, "Antideuterons as a signature of supersymmetric dark matter," *Physical Review D*, vol. 62, no. 4, p. 043003, 2000.

[78] A. Ibarra and S. Wild, "Prospects of antideuteron detection from dark matter annihilations or decays at AMS-02 and GAPS," *Journal of Cosmology and Astroparticle Physics*, vol. 2013, no. 02, p. 021, 2013.

[79] D. Asner, K. Edwards, R. Briere, J. Chen, T. Ferguson, H. Tatishvili, G .and Vogel, M. Watkins, J. Rosner, and N. Adam, "Antideuteron production," *Physical Review D*, vol. 75, no. 1, p. 012009, 2007.

[80] H. Albrecht, R. Gläser, G. Harder, A. Krüger, A. Nippe, T. Oest, M. Reidenbach, M. Schäfer, W. Schmidt-Parzefall, H. Schröder, *et al.*, "Study of antideuteron production in e+ e- annihilation at 10 GeV centre-of-mass energy," *Physics Letters B*, vol. 236, no. 1, pp. 102–108, 1990.

[81] J. Lees, "Antideuteron production in $\Upsilon(nS)$ decays and in $e^+e^- \to q\bar{q}$ at $\sqrt{s} \approx 10.58$ GeV," *arXiv preprint arXiv:1403.4409*, 2014.

[82] J. Roth and T. Kuhr, "Measurement of Antideuteron Production," 2018.

[83] D. Wagenaar, "Radiation physics principles," *Lectures Cycle. Sect*, vol. 7, 1995.

[84] H. Bichsel, "Straggling in thin silicon detectors," *Reviews of Modern Physics*, vol. 60, no. 3, p. 663, 1988.

[85] L. D. Landau, "On the energy loss of fast particles by ionization," *J. Phys.*, vol. 8, pp. 201–205, 1944.

[86] K. A. Milton, "Theoretical and experimental status of magnetic monopoles," *Reports on Progress in Physics*, vol. 69, no. 6, p. 1637, 2006.

[87] P. A. M. Dirac, "Quantised singularities in the electromagnetic field," *Proc. R. Soc. Lond. A*, vol. 133, no. 821, pp. 60–72, 1931.

[88] J. D. Jackson, "Classical electrodynamics," 1999.

[89] M. K. Sullivan and D. Fryberger, "Magnetic charge search for the Belle II detector," *arXiv preprint arXiv:1707.05295*, 2017.

[90] P. Schieferdecker, "Ludwig maximilian university munich," 2001.

[91] D. Fryberger, "On generalized electromagnetism and Dirac algebra," *Foundations of physics*, vol. 19, no. 2, pp. 125–159, 1989.

[92] G. t Hooft, "Magnetic monopoles in unified theories," *Nucl. Phys. B*, vol. 79, pp. 276–284, 1974.

[93] A. Polyakov *JETP Lett.*, vol. 20, p. 194, 1974.

[94] Y. Yang, "Dually charged particl–like solutions in the Weinberg–Salam theory," in *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 454, pp. 155–178, The Royal Society, 1998.

[95] L. P. Gamberg and K. A. Milton, "Eikonal scattering of monopoles and dyons in dual QED," in *FIFTH WORKSHOP ON QUANTUM CHROMODYNAMICS*, p. 176, World Scientific, 2000.

[96] G. Kalbfleisch, W. Luo, K. Milton, E. Smith, and M. Strauss, "Limits on production of magnetic monopoles utilizing samples from the D0 and CDF detectors at the Fermilab Tevatron," *Physical Review D*, vol. 69, no. 5, p. 052002, 2004.

[97] G. Kalbfleisch, K. Milton, M. Strauss, L. Gamberg, E. Smith, and W. Luo, "Improved experimental limits on the production of magnetic monopoles," *Physical review letters*, vol. 85, no. 25, p. 5292, 2000.

[98] I. Ginzburg and A. Schiller, "Visible effect of a very heavy magnetic monopole at colliders," *Physical Review D*, vol. 60, no. 7, p. 075016, 1999.

[99] S. Graf, A. Schaefer, and W. Greiner, "Mass limit for Dirac-type magnetic monopoles," *Physics Letters B*, vol. 262, no. 4, pp. 463–466, 1991.

[100] L. Bracci and G. Fiorentini, "Binding of magnetic monopoles and atomic nuclei," *Physics Letters B*, vol. 124, no. 6, pp. 493–496, 1983.

[101] S. Ahlen and K. Kinoshita, "Calculation of the stopping power of very-low-velocity magnetic monopoles," *Physical Review D*, vol. 26, no. 9, p. 2347, 1982.

[102] P. Musset, M. Price, and E. Lohrmann, "Search for magnetic monopoles in electron-positron collisions at 34 GeV cm energy," *Physics Letters B*, vol. 128, no. 5, pp. 333–335, 1983.

[103] J. Derkaoui, L. Patrizii, A. Margiotta, T. Lari, G. Giacomelli, V. Togo, M. Ouchrif, and V. Popa, "Energy losses of magnetic monopoles and of dyons in the earth," tech. rep., 1998.

[104] J. Lindhard *Mat. Fys. Medd. K. Dan. Vidensk. Selsk.*, vol. 28, p. 8, 1954.

[105] L. Patrizii and M. Spurio, "Status of searches for magnetic monopoles," *Annual Review of Nuclear and Particle Science*, vol. 65, pp. 279–302, 2015.

[106] S. Cecchini, L. Patrizii, Z. Sahnoun, G. Sirri, and V. Togo, "Energy losses of magnetic monopoles in aluminum, iron and copper," *arXiv preprint arXiv:1606.01220*, 2016.

[107] "PDG Atomic Nuclear Properties." `http://pdg.lbl.gov/2017/AtomicNuclearProperties/`. Accessed: 2018-10-20.

[108] P. Flory and A. Vrij, "Melting points of linear-chain homologs. the normal paraffin hydrocarbons," *Journal of the American Chemical Society*, vol. 85, no. 22, pp. 3548–3553, 1963.

[109] "NIST Atomic Spectroscopic Data." `https://www.nist.gov/pml/handbook-basic-atomic-spectroscopic-data`. Accessed: 2018-10-20.

[110] G. G. Hall, "The ionization potentials of some paraffinic molecules," *Trans. Faraday Soc.*, vol. 50, pp. 319–322, 1954.

[111] S. Nakamura. PhD thesis, University of Tokyo, 1988.

[112] L. Bracci, G. Fiorentini, and R. Tripiccione, "On the energy loss of very-slowly-moving magnetic monopoles," *Nuclear Physics B*, vol. 238, no. 1, pp. 167–180, 1984.

[113] I. Kolokolov, P. Vorob'ev, and V. Ianovski, "Slow magnetic monopole: Interaction with matter and new possibility of their detection," *arXiv preprint hep-ph/9809420*, 1998.

[114] L. Gamberg, G. R. Kalbfleisch, and K. A. Milton, "Direct and indirect searches for low-mass magnetic monopoles," *Foundations of Physics*, vol. 30, no. 4, pp. 543–565, 2000.

[115] M. J. Mulhearn, *A direct search for Dirac magnetic monopoles*. PhD thesis, Massachusetts Institute of Technology, 2004.

[116] J. S. Lange, "PXD readout and rescue system for new physics." B2GM, Oct. 2017.

[117] G. Giorgio, N. Werner, and S. Qaisar, *Theory And Detection Of Magnetic Monopoles In Gauge Theories (A Collected Set Of Lecture Notes)*. World Scientific Publishing Company, 1986.

[118] H. Collaboration *et al.*, "A direct search for stable magnetic monopoles produced in positron-proton collisions at HERA," *The European Physical Journal C-Particles and Fields*, vol. 41, no. 2, pp. 133–141, 2005.

[119] J. Derkaoui, G. Giacomelli, T. Lari, G. Mandrioli, M. Ouchrif, L. Patrizii, and V. Popa, "Energy losses of magnetic monopoles and dyons in scintillators, streamer tubes and nuclear track detectors," *Astroparticle Physics*, vol. 10, no. 4, pp. 339–352, 1999.

[120] M. Fairbairn, A. Kraan, D. Milstead, T. Sjöstrand, P. Skands, and T. Sloan, "Stable massive particles at colliders," *Physics Reports*, vol. 438, no. 1, pp. 1–63, 2007.

[121] G. Abbiendi, C. Ainsley, P. Åkesson, G. Alexander, G. Anagnostou, K. Anderson, S. Asai, D. Axen, I. Bailey, E. Barberio, *et al.*, "Search for Dirac magnetic monopoles in e+ e- collisions with the OPAL detector at LEP2," *Physics Letters B*, vol. 663, no. 1-2, pp. 37–42, 2008.

[122] K. Kinoshita, M. Fujii, K. Nakajima, P. Price, and S. Tasaka, "Search for highly ionizing particles in e+ e- annihilations at $\sqrt{s}$= 50- 60.8 GeV," *Physics Letters B*, vol. 228, no. 4, pp. 543–547, 1989.

[123] D. Fryberger, T. Coan, K. Kinoshita, and P. Price, "Search for highly ionizing particles in e+ e- collisions at $\sqrt{s}$= 29 GeV," *Physical Review D*, vol. 29, no. 7, p. 1524, 1984.

[124] J. Pinfold, R. Du, K. Kinoshita, B. Lorazo, B. Price, and M. Regimbald, "A Search for highly ionizing particles produced at the OPAL intersection point at LEP," *Physics Letters B*, vol. 316, no. 2-3, pp. 407–411, 1993.

[125] W. Braunschweig, R. Gerhards, F. Kirschfink, H. Martyn, P. Rosskamp, B. Bock, J. Eisenmann, H. Fischer, H. Hartmann, E. Hilger, *et al.*, "A search for particles with magnetic charge produced in e+ e- annihilations at 543-1543-1543-1GeV," *Zeitschrift für Physik C Particles and Fields*, vol. 38, no. 4, pp. 543–550, 1988.

[126] T. Gentile, P. Haas, M. Hempstead, T. Jensen, H. Kagan, R. Kass, S. Behrends, J. M. Guida, J. A. Guida, F. Morrow, *et al.*, "Search for magnetically charged particles produced in e+ e- annihilations at $\sqrt{s}$= 10.6 GeV," *Physical Review D*, vol. 35, no. 3, p. 1081, 1987.

[127] D. Neverov, "Dark sector physics with Belle II," in *XXXIX International Conference on High Energy Physics*, 2018.

[128] A. Bagulya, M. Vladimirov, V. Ivanchenko, and N. Starkov, "Heavy-particle energy loss simulation using the Geant4 toolkit," *Bulletin of the Lebedev Physics Institute*, vol. 36, no. 5, pp. 127–134, 2009.

[129] J. Vesanto and M. Sulkava, “Distance matrix based clustering of the self-organizing map,” in *International Conference on Artificial Neural Networks*, pp. 951–956, Springer, 2002.

Acknowledgement

I would like to thank all people, who contributed to the success of this thesis. Thank you for providing me with advice and encouragement. I cannot stress enough how much I appreciate the effort of all people, who helped me during the entire time. In particular, I would like to express my deepest gratitude to the following persons:

First and foremost, I would like thank Dr. Sören Lange not only for entrusting me with this project and giving me the opportunity to work on it under his supervision, but for his continuous advice, the organization of trips to KEK and the PhiPsi in Novosibirsk and most importantly for patiently answering the thousands of questions I posed. Thank you for all your support despite the countless other responsibilities you have to handle at the same time! I feel enormously grateful for all your help!

My deepest gratitude also goes to Prof. Dr. Claudia Höhne for her continuous support and motivating words. Thank you for encouraging me to pursue a master in sub-atomic physics and helping me wherever you could!

A thousand thanks to Dr. Hans-Georgha Zaunick, first, for spontaneously agreeing to be the second corrector of my thesis, but even more importantly for always having an open ear for questions and for selflessly helping us whenever we came to ask for advice. I appreciate and truly admire how you take time for supporting and encouraging students while staying on top of all your other duties. Thank you!

A huge thank you also to all people in the Belle II group in Giessen. Thank you for welcoming me so warmly, when I joined last year. The cheerful and relaxed atmosphere in the group was a strong source of motivation for me. It was a pleasure to join you for my thesis! Thank you Klemens for helping me set up basf2 and introducing me to the framework. Many thanks to Simon for always keeping me up to date about all PXD-related activities in Belle II. A special thanks to Dennis for all the tips (or to put it with your words 'random Japan facts') before my trip to Japan, I couldn't have been prepared any better. Thank you also for all the nice discussions over coffee and all the advice about beautiful layout styles for a thesis. Many thanks also to Gundi and Anita for managing all the bureaucratic and organizational hurdles I had to face.

A thousand thanks to the non-Belle II people of our group as well. The countless conversations over lunch and/or coffee taught me a lot, most importantly about the shocking fact that aerogel does not always shine blue as I blindly believed before I saw otherwise. You helped me to learn and understand subjects far beyond my own work. Thank you for patiently explaining to me the activities you pursue, I appreciate every minute of it.

For all the advice about Self-Organizing Maps and data processing I would like to thank Dr. Marc Strickert. Thank you for letting me profit from your vast knowledge about these topics. I am looking forward to testing more of your recommendations in the near future.

A huge thank you to the entire Belle II collaboration, which provided me with a great working environment and supported me with advice for the entire time of my master thesis. Especially I would like to thank Dr. Torben Ferber for the countless advice about neural networks, the many suggestions on how to improve my analysis and the valuable feedback. I truly appreciate your invitation to DESY and I am very grateful for all the time you took to help me, I defiantly learned a lot during this time, a thousand thanks for that!

Last but not least, I would like to thank my father, mother and sister as well as all of my friends for their continuous support and for motivating me throughout this entire time. I am very grateful to you for always keeping my spirits up.

Selbstständigkeitserklärung

Hiermit versichere ich, die vorgelegte Thesis selbstständig und ohne unerlaubte fremde Hilfe und nur mit den Hilfen angefertigt zu haben, die ich in der Thesis angegeben habe. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen sind, und alle Angaben die auf mündlichen Auskünften beruhen, sind als solche kenntlich gemacht. Bei den von mir durchgeführten und in der Thesis erwähnten Untersuchungen habe ich die Grundsätze gute wissenschaftlicher Praxis, wie sie in der ‚Satzung der Justus-Liebig-Universität zur Sicherung guter wissenschaftlicher Praxis' niedergelegt sind, eingehalten. Gemäß § 25 Abs. 6 der Allgemeinen Bestimmungen für modularisierte Studiengänge dulde ich eine Überprüfung der Thesis mittels Anti-Plagiatssoftware.

Datum	Unterschrift