

---

# Beam Background Monitoring and Characterization of CMOS Pixel Detectors for Belle II at SuperKEKB

---

Dissertation

zur Erlangung des mathematisch-naturwissenschaftlichen Doktorgrades  
„Doctor rerum naturalium“  
der Georg-August-Universität Göttingen

im Promotionsstudiengang Physik  
der Georg-August University School of Science (GAUSS)

vorgelegt von

Yannik Buch

Göttingen, 2026

Betreuungsausschuss

Prof. Dr. Ariane Frey

II. Physikalisches Institut, Georg-August-Universität Göttingen

PD Dr. Benjamin Schwenker

II. Physikalisches Institut, Georg-August-Universität Göttingen

Prof. Dr. Stan Lai

II. Physikalisches Institut, Georg-August-Universität Göttingen

Mitglieder der Prüfungskommission:

Referentin: Prof. Dr. Ariane Frey

II. Physikalisches Institut, Georg-August-Universität Göttingen

Koreferent: Prof. Dr. Stan Lai

II. Physikalisches Institut, Georg-August-Universität Göttingen

Weitere Mitglieder der Prüfungskommission:

Jun.-Prof. Dr. Thibaud Humair

II. Physikalisches Institut, Georg-August-Universität Göttingen und Deutsches Elektronen-Synchrotron (DESY), Hamburg

Prof. Dr. Claus Ropers

Max-Planck-Institut für Multidisziplinäre Naturwissenschaften Göttingen

apl. Prof. Dr. Jörn Große-Knetter

II. Physikalisches Institut, Georg-August-Universität Göttingen

Prof. Dr. Steffen Schumann

Institut für Theoretische Physik, Georg-August-Universität Göttingen

Tag der mündlichen Prüfung:

17.04.2026

---

# Beam Background Monitoring and Characterization of CMOS Pixel Detectors for Belle II at SuperKEKB

---

## Abstract

The SuperKEKB storage ring at KEK in Japan aims to deliver an instantaneous luminosity of up to  $6 \times 10^{35} \text{ cm}^{-2} \text{ s}^{-1}$ . To achieve this, a redesign of the interaction region will be necessary. The projected beam backgrounds that the Belle II experiment will experience are substantial. The Belle II experiment plans an upgrade of its vertex detector to cope with the increasing hit rates and accommodate the changing geometry of the interaction region. The proposed vertex detector (VTX) consists of 5 fully pixelated barrel-shaped layers using the specifically developed OBELIX sensor. The OBELIX sensor is currently in the design stage and is derived from the TJ-Monopix2 sensor. Test beam results of irradiated TJ-Monopix2 chips in a temperature controlled environment to test the feasibility of operation at room temperature are presented in this thesis. The results will impact the thermal design and choice of front-end for the final design of OBELIX. Further, the results will provide a reference for the expected impact of irradiation on the performance of OBELIX.

Another important task is the mitigation of beam backgrounds. To support this effort, a diagnostic beam background monitoring tool, named BGNet, was developed. It is a neural network for predicting the beam background decomposition of the hit rate of Belle II sub-detectors based on the state of SuperKEKB. The neural network is trained using 1 Hz time series of diagnostic variables from the SuperKEKB slow control system as input features and measured hit rates of Belle II sub-detectors as regression targets. This thesis presents performance evaluations of BGNet that were performed for its commissioning. Further, a real-time application and a chat bot were developed to increase the accessibility and applicability of BGNet predictions.





---

# Strahluntergrundüberwachung und Charakterisierung von CMOS-Pixelsensoren für Belle II am SuperKEKB Speicherring

---

## Zusammenfassung

Der SuperKEKB-Speicherring am KEK in Japan soll eine instantane Luminosität von bis zu  $6 \times 10^{35} \text{ cm}^{-2} \text{ s}^{-1}$  liefern. Die erwarteten Strahluntergründe, denen das Belle II Experiment ausgesetzt sein wird, sind groß. Das Belle II-Experiment plant ein Upgrade seines Vertex-Detektors, um den steigenden Trefferraten gerecht zu werden. Der vorgeschlagene Vertex-Detektor (VTX) besteht aus 6 Lagen, die den speziell entwickelten OBELIX-Pixelsensor verwenden. Der OBELIX-Sensor befindet sich derzeit in der Entwicklungsphase und ist vom TJ-Monopix2-Sensor abgeleitet. In dieser Arbeit werden die Ergebnisse von Teststrahlungsmessungen mit bestrahlten TJ-Monopix2-Chips in einer temperaturkontrollierten Umgebung vorgestellt, um die Realisierbarkeit eines Betriebs bei Raumtemperatur zu testen. Die Ergebnisse werden sich auf das thermische Design und die Wahl des Frontends für das endgültige Design von OBELIX auswirken und als Referenz für die zu erwartenden Auswirkungen der Bestrahlung auf die Leistung von OBELIX dienen.

Ein weiterer wichtiger Schritt ist die Minderung der Strahluntergründe. Um diese Bemühungen zu unterstützen, wurde ein diagnostisches Werkzeug zur Überwachung des Strahluntergrunds, namens BGNet, entwickelt. Es handelt sich um ein neuronales Netzwerk zur Vorhersage der Strahluntergrundzerlegung der Trefferrate von Belle II-Subdetektoren aufgrund des Zustands von SuperKEKB. Das neuronale Netzwerk wird mit Hilfe von 1 Hz Zeitreihen diagnostischer Variablen aus dem SuperKEKB-Slow-Control-System als Eingangsdaten trainiert. Gemessenen Trefferraten der Belle II Subdetektoren dienen als Regressionsziele für das Training. Diese Arbeit präsentiert Leistungsbewertungen von BGNet, die für dessen Inbetriebnahme durchgeführt wurden. Dazu wurden eine Echtzeitanwendung und ein Chatbot entwickelt, um die Zugänglichkeit und Anwendbarkeit der BGNet-Vorhersagen zu verbessern.



---

## Contents

---

|          |                                                         |           |
|----------|---------------------------------------------------------|-----------|
| <b>1</b> | <b>Introduction</b>                                     | <b>1</b>  |
| <b>2</b> | <b>The Standard Model of Particle Physics</b>           | <b>3</b>  |
| <b>3</b> | <b>Acceleration of Charged Particles</b>                | <b>13</b> |
| 3.1      | Basic Principles of Particle Accelerators . . . . .     | 13        |
| 3.2      | Beam Parameters and Beam Focusing . . . . .             | 17        |
| 3.3      | Beam Collimation . . . . .                              | 22        |
| <b>4</b> | <b>Silicon Sensor Technology</b>                        | <b>25</b> |
| 4.1      | Silicon Semiconductors . . . . .                        | 25        |
| 4.2      | Particle Interactions with Matter . . . . .             | 29        |
| 4.3      | Signal Generation . . . . .                             | 34        |
| 4.4      | Radiation Damage . . . . .                              | 35        |
| 4.5      | Depleted Monolithic Active Pixel Sensors . . . . .      | 37        |
| 4.6      | Charge Collection, Digitization and Threshold . . . . . | 40        |
| <b>5</b> | <b>SuperKEKB</b>                                        | <b>45</b> |
| 5.1      | Luminosity . . . . .                                    | 48        |
| 5.2      | LINAC and Injection Scheme . . . . .                    | 49        |
| <b>6</b> | <b>Belle II</b>                                         | <b>51</b> |
| 6.1      | Vertex Detector . . . . .                               | 51        |
| 6.2      | Central Drift Chamber . . . . .                         | 53        |
| 6.3      | Cherenkov Detectors . . . . .                           | 54        |
| 6.4      | Electromagnetic Calorimeter . . . . .                   | 55        |
| 6.5      | $K_L$ and Muon Detector . . . . .                       | 56        |
| 6.6      | Trigger System and Data Acquisition . . . . .           | 56        |

|           |                                                                           |            |
|-----------|---------------------------------------------------------------------------|------------|
| 6.7       | Radiation Monitors . . . . .                                              | 58         |
| 6.8       | Slow Control System . . . . .                                             | 59         |
| <b>7</b>  | <b>VTX Upgrade Proposal</b>                                               | <b>61</b>  |
| 7.1       | The TJ-Monopix2 DMAPS Chip . . . . .                                      | 64         |
| 7.2       | The OBELIX DMAPS Chip . . . . .                                           | 65         |
| 7.3       | Mechanics . . . . .                                                       | 67         |
| <b>8</b>  | <b>Measurements of the TJ-Monopix2 Chip</b>                               | <b>71</b>  |
| 8.1       | Test Beam Setup . . . . .                                                 | 71         |
| 8.2       | Analysis Setup . . . . .                                                  | 77         |
| 8.3       | Test Beam Results . . . . .                                               | 78         |
| 8.4       | Comparison to Simulation . . . . .                                        | 98         |
| <b>9</b>  | <b>BGNet: Beam Background Monitoring</b>                                  | <b>109</b> |
| 9.1       | Machine Learning - Deep Neural Networks . . . . .                         | 110        |
| 9.2       | Beam Backgrounds . . . . .                                                | 112        |
| 9.3       | Structure of BGNet . . . . .                                              | 118        |
| 9.4       | Explainability of BGNet Models . . . . .                                  | 125        |
| 9.5       | Beam Loss Predictions . . . . .                                           | 131        |
| 9.6       | Real-time Predictions . . . . .                                           | 135        |
| 9.7       | BGNet Chat Bot . . . . .                                                  | 140        |
| 9.8       | Performance of BGNet . . . . .                                            | 147        |
| <b>10</b> | <b>Conclusion</b>                                                         | <b>173</b> |
|           | <b>Bibliography</b>                                                       | <b>176</b> |
|           | <b>Glossary</b>                                                           | <b>185</b> |
|           | <b>List of Figures</b>                                                    | <b>187</b> |
|           | <b>List of Tables</b>                                                     | <b>203</b> |
| <b>A</b>  | <b>Appendix: Test Beam Results</b>                                        | <b>205</b> |
| A.1       | Cluster Properties . . . . .                                              | 205        |
| A.2       | Simulated Cluster Properties . . . . .                                    | 209        |
| <b>B</b>  | <b>Appendix: BGNet</b>                                                    | <b>211</b> |
| B.1       | Input Variables for Hit Rate Models . . . . .                             | 211        |
| B.2       | Input Variables for Beam Loss Models . . . . .                            | 214        |
| B.3       | Requirements for Training Data . . . . .                                  | 216        |
| B.4       | Decompositions of Diamond Radiation Monitors for $\phi$ -Profile Analysis | 216        |
|           | <b>Acknowledgements</b>                                                   | <b>221</b> |

# CHAPTER 1

---

## Introduction

---

The pursuit of knowledge and the desire to understand the world around us has driven scientific endeavors since the emergence of mankind. One such area of research that arose is the field of particle physics with its aim to describe the fundamental building blocks of the universe and the forces governing their interactions. The Standard Model of particle physics is currently providing the most precise theory. It is constantly tested against measurements with increasingly high precision and at energy scales reaching above 1 TeV. Yet, there are observations of phenomena that cannot be described by the Standard Model, such as the apparent matter-antimatter asymmetry in the observable universe, the oscillation of neutrinos and the nature of dark matter.

Particle colliders have allowed experiments to conduct systematic measurements of particle physics. They operate either at the *energy frontier* or the *intensity frontier*. Accelerators at the energy frontier strive to increase the center-of-mass energy of the particle collisions to find novel heavy particles and interactions at higher energy scales. At the intensity frontier, the targets are high precision measurements by providing high luminosities and therefore a large number of collision events. This is done in order to find rare and forbidden processes and anomalies that implicate the existence of new physics.

One of the currently operating accelerators at the intensity frontier is the SuperKEKB collider. It collides electrons and positrons with a center-of-mass energy at the  $\Upsilon(4S)$  resonance, which decays into pairs of B mesons. SuperKEKB is therefore considered a B-factory. The Belle II experiment is a general-purpose particle detector and is built around the interaction point of SuperKEKB. It aims to collect a data set of  $50 \text{ ab}^{-1}$  enabling a diverse physics program that spans measurements of the bottom quark, charm quark, the  $\tau$  lepton, time-dependent CP violation and searches for new light particles as potential candidates for dark matter constituents.

The large amount of radiation produced by the accelerator, the beam background, creates a challenging environment for detectors, especially for the vertex detectors as they are closest to the interaction point. Vertex detectors provide measurements of the particle trajectories of charged particles in order to reconstruct their four-momenta and the position of the production vertex. With the increasing luminosity delivered by SuperKEKB, the requirements on the capabilities of the vertex detector increase. To that end, an upgrade of the current vertex detector is planned for 2032. The upgraded vertex detector will utilize the Optimized BeLle Ii piXel sensor (OBELIX), which is a depleted monolithic active pixel sensor (DMAPS). The OBELIX chip is based on the design of the TJ-Monopix2 chip and is currently in the design stage.

Even before the scheduled upgrade, the luminosity is planned to be increased further, from the current highest achieved luminosity of  $5.1 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$  towards the target luminosity of  $6 \times 10^{35} \text{ cm}^{-2}\text{s}^{-1}$ . This shall be achieved by large increases to the beam currents and by reducing the beam size at the interaction point, leveraging the nano-beam scheme of SuperKEKB. However, the beam backgrounds are already large and approach the limits of Belle II sub-detectors. This has an impact on the lifetime of detector components. Mitigation strategies for the beam background depend on the explicit contribution to the beam background that shall be reduced. While the contributing background sources are known, the exact composition of the beam background is not immediately known and is always changing, depending on the state of SuperKEKB. A background monitoring tool, called BGNet, has been developed to predict the beam background composition in real-time. It uses neural networks that predict the individual contributions of beam background sources to the hit rate of a Belle II sub-detector, based on variables describing the state of SuperKEKB.

The work presented in this thesis focuses on the characterization of the TJ-Monopix2 DMAPS chip as part of the Belle II upgrade efforts and the beam background monitoring tool BGNet. The theoretical structure and motivation for the Belle II measurements are presented in Chapter 2. Then, accelerators and silicon sensor technology are introduced in Chapter 3 and 4. In Chapter 5, the SuperKEKB accelerator complex is presented and, in Chapter 6, the Belle II detector is presented. The upgrade plans for Belle II with focus on the planned vertex detector upgrade are described in Chapter 7. The work of the author is described in Chapter 8, relating to the test beam measurements of irradiated samples of TJ-Monopix2, and in Chapter 9, relating to the beam background monitoring with BGNet.

## CHAPTER 2

---

### The Standard Model of Particle Physics

---

The predictions of the Standard Model (SM) of particle physics [1–4] have been rigorously tested since its emergence in the 1970s and is to date the best description of elementary particles and their interactions. One of the latest major tests of the SM was the discovery of the Higgs Boson in 2012 and the subsequent investigation of the SM predictions for the Higgs couplings and Higgs potential [5–7]. Currently, the measurements show no deviation from the SM description of the Higgs Boson [8].

The SM is a quantum field theory (QFT) and is based on the local gauge symmetry group  $SU(3)_C \times SU(2)_L \times U(1)_Y$  with the respective gauge groups corresponding to the strong interactions, the weak interactions and the electromagnetic interactions, respectively. The elementary particles are either fermions with half-integer spin described by the Fermi-Dirac statistics [9] or bosons with integer spin described by the Bose-Einstein statistics [10].

A mass hierarchy is present among fermions, which is reflected by an ordering into three generations of matter. The bosons are either vector-bosons with spin-1 which are associated with the mediation of their respective force or scalar-bosons with spin-0. The Higgs boson is the only known spin-0 particle. The particles of the SM are displayed in Figure 2.1. Interactions are governed by the respective QFT, the boson’s properties and their coupling to fermions. The vertices allowed in the SM of vector-bosons and fermions are displayed in Figure 2.2.

A brief motivation on how the respective vertices emerge from the QFT is given using Quantum Electrodynamics (QED) with its  $U(1)_Q$  gauge symmetry, where  $Q$  refers to the electric charge. QED is a relativistic quantum field theory that was developed in the 1940s by Tomonaga, Schwinger and Feynman [12–14]. The associated

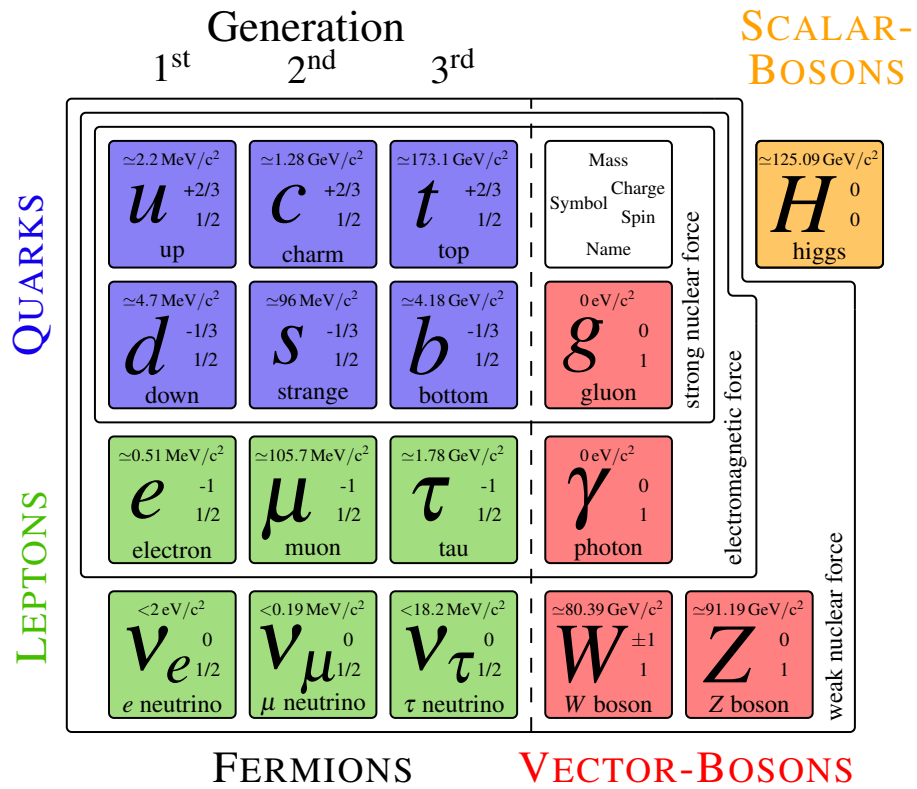


Figure 2.1: Artistic depiction of the particles in the Standard Model. Measurements are taken from Reference [11].



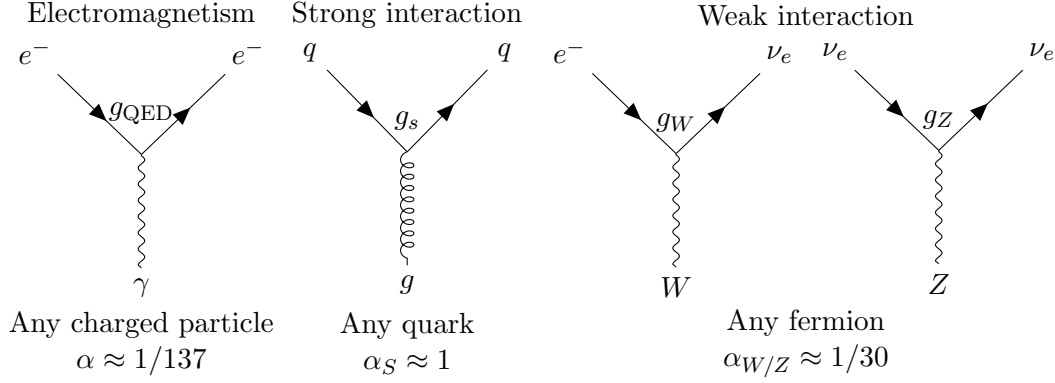


Figure 2.2: Standard Model vertices of fermions coupling with bosons. Each vertex, and associated force, has an associated coupling strength denoted as  $g_{\text{QED}}$ ,  $g_s$ ,  $g_W$ ,  $g_Z$ , which can be expressed as dimensionless coupling constants  $\alpha$ ,  $\alpha_s$ ,  $\alpha_{W/Z}$ . Coupling constant values are at momentum transfers of  $q^2 = 1 \text{ GeV}$ [11].

Lagrangian is:

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi - e\bar{\psi}\gamma^\mu\psi A_\mu \quad (2.1)$$

It describes the propagation of electrically charged fermions, corresponding to  $\bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi$ , and the massless photon, corresponding to  $-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$ , in spacetime. The interaction vertices correspond to  $-e\bar{\psi}\gamma^\mu\psi A_\mu$ . Since the incoming and outgoing fields can be reordered arbitrarily, the vertices describe emission, absorption as well as pair creation and annihilation processes. Similarly, the possible vertices for each QFT emerge from the respective Lagrangian expression. The fields described in the Lagrangian are the fundamental piece of the SM from which the particles arise as excitations.

From the Lagrangian terms, Feynman rules can be derived that can describe any physically allowed particle interaction. A Feynman diagram is an expressive representation of particle physics processes which yields the Lorentz-invariant *matrix element*,  $M^{i,f}$ , of the interaction Hamiltonian  $\hat{H}_{\text{int}}$  operator acting on an initial state  $i$  causing it to transition into a final state  $f$ ,

$$\langle f | \hat{H}_{\text{int}} | i \rangle = \langle \psi_1 \psi_2 \dots | \hat{H}_{\text{int}} | \psi_a \psi_b \dots \rangle = M^{i,f}. \quad (2.2)$$

The associated transition probability for this process is proportional to  $|M^{i,f}|^2$  and is usually obtained using Fermi's golden rule. Here, it is assumed that the wavefunctions  $\psi$  describing the initial and final states have the appropriate Lorentz-invariant normalization to  $2E$  particles per unit volume

$$\int_V \psi^* \psi d^3\mathbf{x} = 2E. \quad (2.3)$$

The states  $i$  and  $f$  are eigenstates of the free-particle Hamiltonian  $\hat{H}_0$ , that, together with the interaction Hamiltonian, describes the total Hamiltonian  $\hat{H} = \hat{H}_0 + \hat{H}_{\text{int}}$ . Particle physics experiments allow for the observation of the transition of an initial state into a final state and cross-sections for these transitions can be measured. Different transitions with the same initial and final states cannot be differentiated and interfere with each other, which is reflected by the measured cross-sections.

Quantum Chromodynamics (QCD) describes the interactions of gluons and quarks with the associated color charge [15]. The gluon is a massless boson which itself carries the color charge, including anti-color charge, that it couples to, allowing for self-interactions. The SU(3) symmetry associated with QCD has 8 linearly independent generators, which implies the existence of 8 gluons in an octet each with a different charge configuration. One can identify terms in the QCD Lagrangian that are similar to terms in the QCD Lagrangian. Although the quark propagation term in the QCD Lagrangian refers to singular quarks, free quarks have not yet been observed. Instead they appear in colorless states, which are either baryons or meson with 3 or 2 valence quarks in the bound state, respectively. Bound states with more quarks can also exist, like the pentaquark, which is a bound state of four quarks and one antiquark [16]. This observation is called color-confinement. The potential  $V_c$  that develops between two color charges has a constant energy density at large distances  $x$  due to the linear term

$$V_c(r) \propto \kappa r. \quad (2.4)$$

The constant  $\kappa$  has been measured to be  $\kappa \approx 1 \text{ GeV fm}^{-1}$  [11]. Thus, the energy stored in the color field approaches infinity for any pair of two color charges that are separated towards infinity. Through the hadronization process, modeled for example by the Lund-String model [17], a pair of quarks can be created from the energy stored in the color field in order to form colorless states again. A second implication of color-confinement is that gluons, carrying a color charge themselves, are also affected by that same color field, leading to the short-ranged nature of the strong interaction.

Quantum Flavordynamics (QFD) describes the weak nuclear force. It is mediated by the massive  $W^\pm$  and  $Z$  bosons. The  $W^\pm$  boson couples to the weak isospin  $T_3$ , while the  $Z$  boson couplings are expressed in separate vector and axial-vector couplings, however, both terms are still dependent on  $T_3$ . Left-handed particles and right-handed antiparticles have  $T_3 = \pm \frac{1}{2}$  and are placed in doublets as up-type ( $T_3 = +\frac{1}{2}$ ) and down-type ( $T_3 = -\frac{1}{2}$ ) particles. The handedness of the particles refers to the chirality eigenstates, which coincides with the helicity eigenstates for massless particles, or in highly relativistic approximations  $E \gg m$  for massive particles. The particle doublets are:

$$\begin{pmatrix} u \\ d' \end{pmatrix}_L, \begin{pmatrix} c \\ s' \end{pmatrix}_L, \begin{pmatrix} t \\ b' \end{pmatrix}_L, \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L. \quad (2.5)$$

Where  $d', s'$  and  $b'$  refer to the weak eigenstates of the respective quarks. Right-handed particles and left-handed antiparticles are placed in singlets and have  $T_3 =$

0. The transitions between the leptons and quarks of one doublet are typically mediated by the  $W^\pm$  bosons, the flavor-changing charged current. The SM also allows for a flavor-changing neutral current at next-to-leading-order (NLO), however these higher order interactions are highly suppressed by the GIM-mechanism [18]. The transition probabilities attached to these quark transitions are given by the unitary Cabibbo–Kobayashi–Maskawa (CKM) Matrix [19], that relates the weak eigenstates and the mass eigenstates of the quarks. The CKM Matrix is explained in more detail below. Interactions of the  $W^\pm$  bosons violate parity maximally, while Z bosons violate parity as well, but not maximally and thus still couple to the right-handed chirality states of particles, albeit not with the same strength.

The electromagnetic and weak interaction can be unified in one gauge group  $SU(2)_L \times U(1)_Y$  known as the electroweak theory. The electroweak gauge symmetry can be broken by introducing two complex scalar fields, the Higgs doublet. The vector boson fields  $B_\mu$  and  $W_\mu^{1,2,3}$  are related to  $U(1)_Y$  and  $SU(2)_L$ , respectively. From linear combinations of these fields, the photon field  $A_\mu$  and the weak interaction boson fields  $W_\mu^\pm$  and  $Z_\mu$  emerge, bearing also explicit mass terms for the  $W^\pm$  and Z bosons. The parameters of the introduced potential can be chosen such that any vacuum state sees an asymmetric potential, revealing the physical Higgs field. The masses of fermions are a direct result of their interaction with the Higgs field via the Yukawa coupling. The strength of the Yukawa coupling is proportional to the mass and its strength is not predicted by the SM [5–7].

## CP Violation

The transitions between different quark flavors via the flavor-changing charged current of the weak interaction was first described by Cabibbo [20]. The weak eigenstates  $(d', s')$  are related to the mass eigenstates  $(d, s)$  via

$$\begin{pmatrix} d' \\ s' \end{pmatrix} = \begin{pmatrix} \cos \theta_C & \sin \theta_C \\ -\sin \theta_C & \cos \theta_C \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix}, \quad (2.6)$$

where  $\theta_C = 13.02^\circ$  is the Cabibbo angle [11]. With the discovery of a third quark generation [21, 22] this formalism was extended and the CKM matrix was introduced. The CKM matrix  $V^{\text{CKM}}$  is represented by

$$V^{\text{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \quad (2.7)$$

and is unitary in order to preserve the norm, i.e.  $\sum_{j,k} V_{ij} V_{ik} = \delta_{jk}$ . This allows for  $|V_{ij}|^2$  to be interpreted as a probability. An alternative representation is the Wolfenstein parametrization

$$V(A, \lambda, \rho, \eta) = \begin{pmatrix} 1 - \frac{1}{2}\lambda^2 & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{1}{2}\lambda^2 & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4). \quad (2.8)$$

This parametrization explicitly depends on a complex phase  $\eta$  that is non-reducible. CP violation (CPV) arises in processes involving the CKM matrix when the value of this complex phase is different from zero. The values of the parameters  $A$ ,  $\lambda$ ,  $\rho$  and  $\eta$  need to be measured experimentally. The CKM matrix can be represented as a triangle in the complex plane through the unitarity conditions like

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0 \quad (2.9)$$

$$1 - \frac{|V_{ud}|}{|V_{cd}||V_{cb}|}V_{ub}^* - \frac{|V_{tb}|}{|V_{cd}||V_{cb}|}V_{tb} = 0. \quad (2.10)$$

The sides and angles of this triangle can be independently measured, over constraining it, and testing the CKM mechanism in the SM. One example for such a measurement can be derived from  $V_{td} = |V_{td}|e^{-i\beta} = A\lambda^3(1 - \rho - i\eta)$  using the Wolfenstein parametrization, which gives  $\tan\beta = \eta/(1 - \rho)$ . From Equation 2.9 it can be seen that the angle  $\beta$  corresponds to an internal angle of the unitarity triangle. The unitarity triangle is displayed in Figure 2.3, also indicating the angle  $\beta$ . A fit of the unitarity triangle based on currently available data is displayed in Figure 2.4. Measuring the unitarity triangle is a key part of the Belle and Belle II physics programs. Below the measurement of the angle  $\beta$  is discussed in more detail.

The CPV then appears as three different types. Direct CPV occurs when the branching fraction of the decay of an initial state to a final state differs from the branching fraction of the decay of a CP transformed initial state into a CP transformed final state.

Indirect CPV, or CPV in mixing, can be observed in neutral mesons which can transition to their CP transformed counterpart via a box loop. The transition is described via a  $2 \times 2$  Hamiltonian. If the off-diagonal elements of the Hamiltonian differ from the off-diagonal elements of the complex conjugate of the Hamiltonian, i.e. if the off-diagonal element has an imaginary part, the mixing induces CPV. As a result, the transition time is different between the two directions. CPV also occurs in the interference of the aforementioned modes of CPV when a neutral meson and the corresponding antimeson can decay into the same CP eigenstate. The CPV in the interference of direct and indirect CPV is the biggest contribution of CPV in the B meson system and is therefore of special importance to Belle II. Presently, CPV is a potential explanation for the apparent matter-antimatter asymmetry in our universe. This motivates measurements and experiments relating to CPV, like the Belle II experiment [23].

An example measurement of CPV at Belle II will be introduced here briefly and serves as motivation for the work performed by the author, which is described in this thesis.

The example concerns the measurement of the angle  $\beta$ , or  $\phi_1$ , which is an interior angle of the unitary triangle, that was performed by Belle [26]. The angle is defined as:

$$\beta = \arg\left(-\frac{V_{cd}V_{cb}^*}{V_{td}V_{tb}^*}\right), \quad (2.11)$$

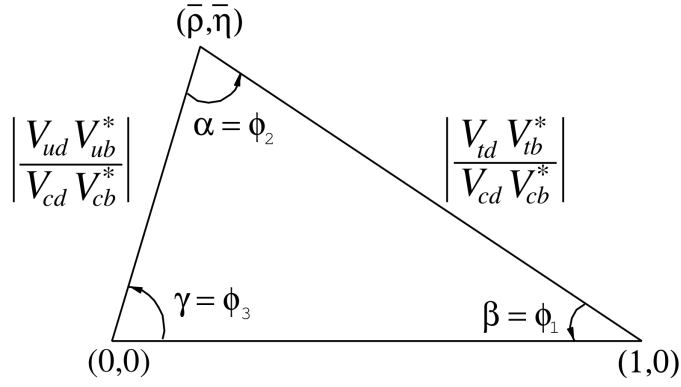


Figure 2.3: A representation of the unitarity triangle in the complex plane [11].

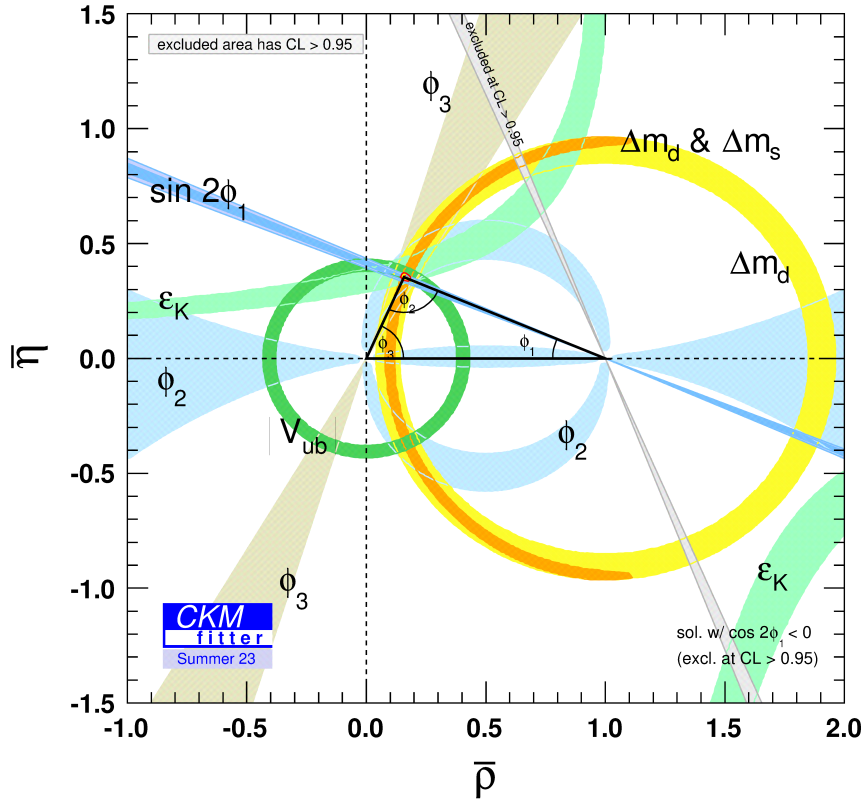


Figure 2.4: The current unitarity triangle fit determined with CKMFITTER [24]. The figure is taken from Reference [25].

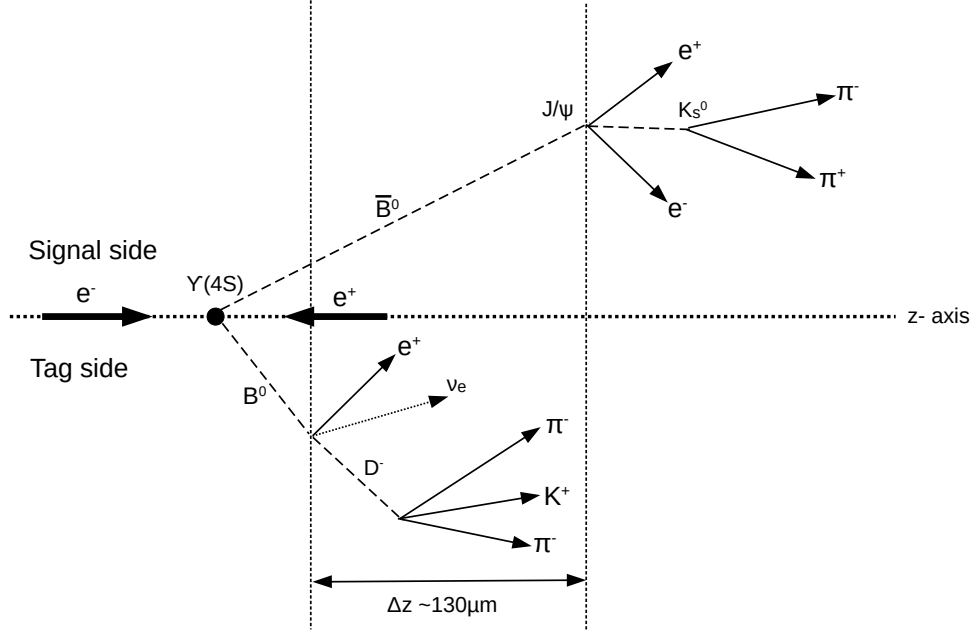


Figure 2.5: Drawing of two B meson decays with displaced vertices and their distance  $\Delta z$ . The event has a tag side and a signal side. The flavor of the tag side B meson is determined by measuring the charge of the electron, allowing to implicitly infer the flavor of the signal side B meson. The apparent spatial separation and the directional bias of the particles of the signal and tag side benefit only the comprehensibility of the drawing and are not due to actual physical effects separating the two decays.

where  $\arg(\cdot)$  denotes the complex phase of its argument with  $\arg(z) = \theta$  for a complex number  $z = |z|e^{i\theta}$ .

The measurement requires the production of an entangled pair of neutral B mesons, where one B meson decays into a CP eigenstate at time  $t_{\text{CP}}$ , the signal side, and the other decays into a final state that allows the reconstruction of the flavor of that B meson at time  $t_{\text{tag}}$ , the tag side. Here, the CP eigenstate is taken to be  $J/\psi K_s^0$  and a visualization of the decay is displayed in Figure 2.5.

The time-dependent CP violation is measured from the decay rate asymmetry:

$$A_{\text{CP}}(\Delta t) = \frac{\Gamma(\bar{B}^0(\Delta t) \rightarrow J/\psi K_s^0) - \Gamma(B^0(\Delta t) \rightarrow J/\psi K_s^0)}{\Gamma(\bar{B}^0(\Delta t) \rightarrow J/\psi K_s^0) + \Gamma(B^0(\Delta t) \rightarrow J/\psi K_s^0)} \quad (2.12)$$

$$= \mathcal{S}_f \sin(\Delta m_d \Delta t) + \mathcal{A}_f \cos(\Delta m_d \Delta t), \quad (2.13)$$

where  $\Delta m_d$  is the mass difference between the two neutral B meson mass eigenstates and  $\Delta t = t_{\text{CP}} - t_{\text{tag}}$  is the decay time difference. The parameters  $\mathcal{S}_f$  and  $\mathcal{A}_f$  are CPV parameters. The SM predicts that  $\mathcal{S}_f = -\xi_f \sin(2\beta)$  and  $\mathcal{A}_f = 0$ . Experimentally,

the extraction of  $\sin(2\beta)$  requires the reconstruction of the signal decay into the CP eigenstate and flavor tagging the initial state of the B meson on the tag-side and the measurement of the decay time difference  $\Delta t$ . Since the mass difference between the  $\Upsilon(4S)$  and its decay products, the two B mesons, is small, the B mesons are approximately at rest in the rest frame of the  $\Upsilon(4S)$ . When the energies of the colliding electrons and positrons are asymmetric, the  $\Upsilon(4S)$  receives a Lorentz boost along the  $z$ -axis, and the decay time difference  $\Delta t$  of the B mesons can be determined from their decay vertex displacement  $\Delta z = z_{\text{CP}} - z_{\text{tag}}$ . The decay vertex displacement is indicated in Figure 2.5. The decay time difference can be obtained via  $\Delta t \simeq \Delta z / \beta \gamma c$  [26], where  $\beta \gamma = 0.28$  is the boost factor of SuperKEKB [27].

Several measurements of  $\sin(2\beta)$  and a world average of those measurements is displayed in Figure 2.6. The discussion of errors in the Belle measurement [26] shows that the vertex resolution is a major contributor to the systematic error. The vertex resolution, in part, motivates the vertex detector upgrade of Belle II from a physics perspective and is presented in Chapter 7. Measurements by the author on silicon pixel detectors considered for the upgrade are presented in Chapter 8. Further, the statistical uncertainty is shown to be larger than the systematic uncertainty. To obtain a larger data set within a sensible time the instantaneous luminosity needs to be large. The luminosity and the means available to increase the luminosity are discussed in Section 5.1. An important criterion to be able to increase the luminosity is to control the beam background in order to stay within hardware limitations and guarantee safe and efficient data taking. The author took part in the development of a beam background monitoring tool that is presented in Chapter 9. It delivers predictions of the beam background decomposition and can prompt operators to use appropriate mitigation strategies.

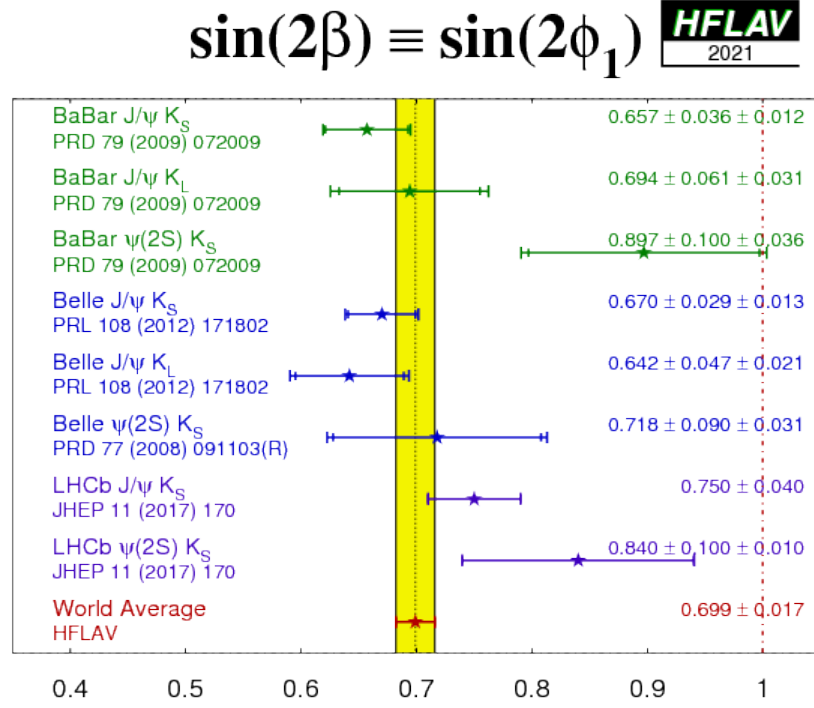


Figure 2.6: Averages of  $\sin 2\beta = \sin 2\phi_1$  from different experiments. In cases where two errors are quoted for measurements, the first error is the statistical uncertainty and the second error is the systematic uncertainty. In cases, where only error is provided, it is identified as the combined error from statistical and systematic uncertainties [28].



## CHAPTER 3

---

### Acceleration of Charged Particles

---

This chapter introduces particle accelerator concepts in a historic order to showcase challenges and the subsequent developments, leading up to the description of SuperKEKB in the following Chapter 5. Further, beam dynamics and beam parameters are discussed in order to create a basis of understanding necessary for the beam background description and the explanation of the beam background monitor described in Chapter 9. This chapter is based on [29, 30].

#### 3.1 Basic Principles of Particle Accelerators

The principle equation to apply a force on charged particles and accelerate them is the Lorentz force

$$\vec{F}_L = q(\vec{E} + \vec{v} \times \vec{B}) \quad (3.1)$$

with the electric charge  $q$ , the particle velocity  $\vec{v}$ , the electric field  $\vec{E}$  and the magnetic field  $\vec{B}$ . The deposited kinetic energy can be calculated with a line integral

$$\Delta E_{\text{kin}} = \int \vec{F}_L \, d\vec{s} = q \int \vec{E} \, d\vec{s}. \quad (3.2)$$

Two observations can be made from this calculation. First, the acceleration due to an electric field happens in the direction of the electric field and is independent of the particle's velocity. Second, the line integral over the contribution of the magnetic field is zero, the magnetic field does not contribute to a change in the kinetic energy of a charged particle. However, charged particles with  $\vec{v} \neq 0$  will be deflected inside a magnetic field into the direction that is normal to both the direction of the magnetic field as well as the direction of motion. The change of direction of  $\vec{v}$  due to a

magnetic field is also considered an acceleration, despite the magnitude of  $\vec{v}$  staying unchanged.

The path of a charged particle in a magnetic field is described by the equilibrium of the centripetal force  $F_C$  and the Lorentz Force

$$\frac{\gamma m v^2}{r} \vec{n} = q(\vec{v} \times \vec{B}) \quad (3.3)$$

with the particle's mass  $m$ , the radius  $r$  of the curved motion, the Lorentz factor  $\gamma$  and the unit vector  $\vec{n}$  that points in the direction of the centripetal force. Assuming that only the magnetic field has an orthogonal contribution  $B_\perp$  with respect to the particles motion this equation reduces to

$$r = \frac{p}{qB_\perp} \quad (3.4)$$

with the particle momentum  $p$ . For accelerators that keep particles on a fixed nominal orbit, this equation becomes a condition with

$$r = \frac{p}{qB_y(p)} = \text{const.} \quad (3.5)$$

As a result the magnetic field strength must become a function of the particle's momentum  $B_y = B_y(p)$  and change during the acceleration of a charged particle to bend the particles trajectory on a constant orbit. The coordinate system used defines  $\vec{y}$  to point in axial direction,  $\vec{x}$  to point in radial direction and  $s$  is the longitudinal coordinate parallel to the direction of motion of particles along a nominal orbit.

Charged particles gain kinetic energy by traversing an electric field. The Lorentz force expressed only in the presence of an electric field reduces to

$$\vec{F}_L = q\vec{E} \text{ with } \vec{E} = \vec{E}_0 \exp(i(\omega t - ks)), \quad (3.6)$$

where  $\vec{E}$  is expressed as a plane wave equation with the phase  $\Phi = \omega t - ks$  in the particle system. The phase is described by the frequency  $\omega$ , the time  $t$  and the wavenumber  $k$  with the spatial coordinate  $s$ . The generation of electrical fields for the acceleration of particles can be divided into two categories. Electrostatic fields can be generated by a voltage difference between two electrodes. Charged particles will be accelerated towards an electrode and can pass the electrode through a hole. Electrodes with holes for particles are also referred to as *iris*. Prominent examples for electrostatic acceleration stages are the Cockroft-Walton accelerator with cascade voltage generators [31] and Van de Graaff accelerators [32].

The practical limitation of such machines lies in the linear scaling of the electrical field strength and the applied voltage difference  $E \propto \Delta V$  and the subsequent necessity of handling very high voltages. Voltage breakdowns of any part in the machine then will not only cause the loss of built up charges and voltage but also pose risks to damage the machine. Many considerations about the inefficiency of such machines are connected to the fact that they were used as linear accelerators,

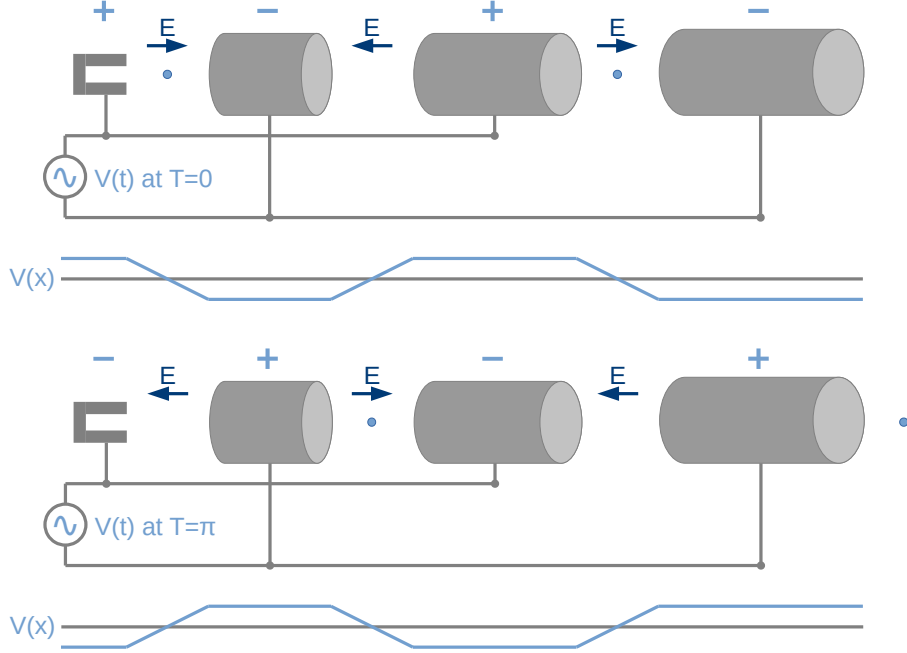


Figure 3.1: Schematic of Wideroe drift tubes accelerating a charged particle. The changing voltage  $V(t)$  is connected to the accelerating segments in an alternating pattern and at  $T = 0$  (top) and  $T = \pi$  (bottom), assuming that one period is equal to  $2\pi$ . The voltage as a function of the position  $x$  along the structure is indicated.

where each accelerating segment is only traversed once by the particle. Below these high voltage limits, however, these accelerating structures are still successful with Cockroft-Walton setups still being in use today as initial acceleration stages in modern acceleration facilities.

There are two major design changes that will be discussed here that are used by modern accelerators. These are generation of electric fields by radio frequency waves and the change to circular accelerators.

Changing electric fields had their beginnings in Wideroe drift tubes [33]. There, segmented hollow cylinders act as electrodes and neighboring tubes always carry the opposite charge. The emerging electric fields between the segments point in alternating directions and particles need to pass the gaps only when the electric field points in the proper direction to accelerate the particle further. The inside of each cylindrical segment acts as Faraday cage and no electric fields act on the particles. During this time the polarity of the electrodes switches and once the particles leave the field-free domain of the Faraday cage the electric field points in the proper direction again, accelerating the particle. A schematic of such a setup is displayed in Figure 3.1.

The energy gain at each crossing is  $\Delta E = q\Delta V$ , where  $\Delta V$  is the potential

difference between two neighboring tubes. The condition on the timing of the particles reaching the electric field and the period of the changing electric field can be expressed as a synchronicity condition

$$L \approx vT \quad (3.7)$$

with the tube length  $L$ , the particle velocity  $v$  and the voltage period  $T$ . As the particle accelerates, the time needed to traverse a tube decreases and the tube length needs to be increased or the period of the voltage needs to decrease. In addition, this condition can only be fulfilled by a single particle or a group of particles that occupy a similar region in phase space, commonly referred to as a bunch, rather than a continuous beam of particles. Nevertheless, a series of bunches in particle accelerators is still referred to as a beam, be it with a bunched structure. The typical frequency of the voltage generators is in the regime of radio-frequencies (RF). The advantage of RF accelerators is that they avoid the electrical breakdown limitations that electrostatic accelerators are subjected to. In theory a particle's energy can be raised indefinitely using a single RF generator, given an accelerator of infinite length, to fulfill the synchronicity condition.

One approach to reduce the length of accelerators is to reuse accelerating elements by constructing a circular accelerator. The cyclotron employs a single RF cavity with a circular geometry, where each half of the RF cavity is a semi circle. These halves are usually referred to as D's of the cyclotron. Similar to the linear counterpart, the D's act as Faraday cages and particles see no electric field while inside, and are accelerated in the gap between the D's. They do, however, experience a uniform magnetic field that spans the entire volume of the cyclotron, and explicitly inside the D's, to force the charged particles on a circular orbit, according to the Lorentz force as discussed above. As the particles are accelerated and the momentum increases the radius in a constant magnetic field will increase according to Eq. 3.4. So, as the particles move faster, the path length inside a D grows proportionally and the revolution frequency stays constant, thus the RF generator frequency can also be held constant. Due to relativistic limits this linear scaling only holds for non-relativistic particles at  $\gamma \approx 1$ , where  $\gamma$  refers to the Lorentz factor. The synchronicity condition follows as [29]

$$f_{\text{rf}} = \frac{qB_y}{2\pi m_0 c \gamma} = \text{const.} \quad (3.8)$$

assuming that the revolution frequency is equal to the frequency of the RF generator. The condition of non-relativistic velocities results in an energy limit, which for protons is typically already reached at around 20 MeV. Lighter particles, like electrons, are even more severely limited. Once the radius of the particles' flight path approaches the extent of the D's, the particles are ejected through an opening in the D's and leave the accelerating structure. Since in general  $f_{\text{rf}} = f_{\text{rf}}(\gamma)$  and  $\gamma$  will increase as the energy of the accelerated particle increases, the cyclotron can also be adapted to vary the RF and the resulting structure is called synchro-cyclotron.

Synchro-cyclotrons thus do not require  $\gamma \approx 1$  and allow to accelerate particles to higher energies compared to the cyclotron.

The synchrotron follows a different design choice by keeping the radius constant, rather than the magnetic field. This allows to build the accelerating structures along the designated orbit of the particles. Magnets and accelerating structures can be much smaller and put in series along the ring, rather than operating a singular large magnet over the entire area like the cyclotrons. The condition on the magnetic field when requiring a constant radius is described by Equation 3.5. From this condition the maximum achievable particle momentum follows immediately and depends on the radius and the maximum magnetic field. In case of ion colliders the charge multiplicity can also be considered as lever arm to increase the maximum possible momentum, although there are other considerations and limitations when accelerating particles with high charge multiplicities [29, 30]. The synchronicity condition is [29]

$$f_{\text{rf}} = \frac{qB_y}{2\pi\gamma mc} h \quad (3.9)$$

with the harmonic number  $h$ , an integer number that relates the revolution frequency  $f_{\text{rev}}$  of the particles around the ring with the RF frequency

$$h = \frac{f_{\text{rf}}}{f_{\text{rev}}}. \quad (3.10)$$

The revolution frequency for a particle with velocity  $v$  is

$$f_{\text{rev}} = \frac{v}{2\pi r} = \frac{B_y q}{2\pi\gamma mc} = \frac{B_y qc}{2\pi pc} \beta, \quad (3.11)$$

where Equation 3.5 was inserted. Due to the RF frequency's implicit dependence on the velocity  $\beta$  through the revolution frequency, it is necessary to modulate the frequency as long as the particles mass is sufficiently large compared to its energy. For particles close to the speed of light at  $\beta \approx 1$ , the velocity will barely change with an increase in energy and  $\beta$  will be approximately constant and the frequency modulation needed decreases significantly. A particular type of synchrotrons are storage rings. Here, the particles are injected into the synchrotron ring with their nominal energy and the storage ring stores the particles on a stable orbit, without explicitly increasing the particles energy. SuperKEKB is such a storage ring. The particle beam's structure depends on the RF frequency and determines the maximum length of a particle bunch and, in conjunction with the radius of the synchrotron, also sets an upper limit on the total number of bunches along the ring.

## 3.2 Beam Parameters and Beam Focusing

Already during the early days of accelerator development it was found that there is a need to focus particle beams. Particle trajectories are sensitive to slight deviations

in phase space and easily spiral out of the design orbit without correcting forces. First proposed for the design of the Betatron, a gradient in the magnetic field was introduced such that the restoring force (compare to Equation 3.5)

$$F_x = \frac{\gamma m v^2}{r} - q v B_y \quad (3.12)$$

with the gradient of the magnetic field

$$B_y = B_{0y} + \frac{\partial B_y}{\partial x} x \quad (3.13)$$

becomes

$$F_x \approx \frac{\gamma m v^2}{R} \left(1 - \frac{x}{R}\right) - e v B_{0y} \left(1 - n \frac{x}{R}\right) \quad (3.14)$$

for particles at  $r = R + x$ , where  $R$  corresponds to the nominal orbit, with a small deviation  $x$  with  $x \ll R$ . The field index  $n$  is defined as

$$n = -\frac{R}{B_{0y}} \frac{\partial B_y}{\partial x}. \quad (3.15)$$

The equation of motion

$$\ddot{x} + \omega_x^2 x = 0 \quad (3.16)$$

is, due to the restoring force in the horizontal plane, described by an harmonic oscillator with the frequency

$$\omega_x = \frac{v}{R} \sqrt{1 - n}. \quad (3.17)$$

The field index is thus constrained to  $n < 1$ , otherwise the differential equation will evaluate to an exponential function due to the imaginary part of  $\omega_x$  causing the amplitude to grow exponentially.

The magnetic field in the horizontal direction is obtained using Maxwell's curl equation

$$\frac{\partial B_x}{\partial y} - \frac{\partial B_y}{\partial x} = 0, \quad (3.18)$$

which can be inserted into Equation 3.13 to obtain [29]

$$B_x = \int \frac{\partial B_y}{\partial x} dy = - \int n \frac{B_{0y}}{R} dy = -n \frac{B_{0y}}{R} y. \quad (3.19)$$

Similar to the motion in the horizontal plane, the equation of motion in the vertical plane can be found to be

$$\ddot{y} + \omega_y^2 y = 0, \quad (3.20)$$

which is again a harmonic oscillator with a frequency of

$$\omega_y = \omega_0 \sqrt{n}, \quad (3.21)$$

constraining the field index to  $n > 0$ . The parameter  $\omega_0$  is the orbital revolution frequency. In this calculation an acceleration of the particles, i.e. terms with  $\dot{y}$ , were neglected in the context of SuperKEKB being a storage ring. If acceleration is taken into account, a reduction of the particles amplitude with growing energy, or adiabatic damping, can be observed. Combining the constraints on the field index for the horizontal and vertical plane Steinbeck's stability criterion  $0 < n < 1$  is obtained [29]. The magnetic field  $B_x$  increases vertically providing beam stability and focusing in the vertical direction. However, looking at the allowed field indices, one can observe that the field component  $B_y$  decreases horizontally resulting in a defocusing in the horizontal plane. This mechanism is called weak focusing and a focusing in both planes cannot be achieved by this effect [29]. The horizontal defocusing effect can be compensated by the focusing provided by the primary magnet of the Betatron, thus providing a net positive focusing in both transverse directions. The focusing by the betatron magnet, or any other multipole magnet, is referred to as strong focusing.

Strong focusing is implemented via explicit magnets to steer and focus the beam. Dipole magnets keep the particles on the design orbit, while quadrupole magnets focus the beam. Magnets of higher order, like sextupole and octupole magnets, correct for chromaticity, which describes the momentum-dependence of the focusing effects of quadrupole magnets, and other higher-order effects. The working principle of a dipole was discussed above by applying the Lorentz force exerted by a magnetic field on a moving charged particle. While a quadrupole operates on the same physical principle, its application for focusing particle beams is briefly summarized. The field of a quadrupole is chosen such that it produces a linear gradient magnetic field that focus the beam particles depending on their position in phase space, while the field in the center is zero. A schematic of a quadrupole field is displayed in Figure 3.2.

A quadrupole can only be focusing in one plane transverse to the particles' motion, while it acts defocusing in the other. The correcting force that focuses or defocuses the beam is derived from the Lorentz force with  $F = -evB$ . Combining two quadrupole magnets that are turned by  $90^\circ$  with respect to each other arranges a FODO cell, that is shown to have a net positive effect on the focusing in both planes for all possible ray paths [29]. A FODO cell is depicted in Figure 3.3. The collection of all magnets of the accelerator is referred to as the magnet lattice.

The motion of particles in a periodic system can be described with Hill's equation [29]

$$x'' + K(s)x = 0. \quad (3.22)$$

The periodicity is accounted for by  $K(s)$  by requiring

$$K(s) = K(s + L) \quad (3.23)$$

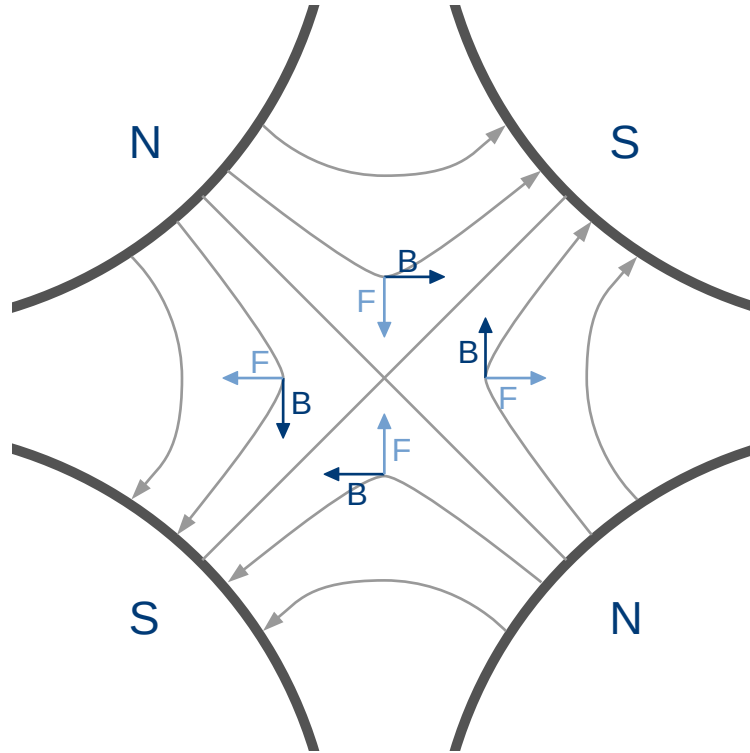


Figure 3.2: Sketch of the magnetic field produced by a quadrupole magnet. The direction of the magnetic field is indicated at selected points and the direction of the Lorentz force for an electron traversing the magnet vertically into the plane of the drawing.



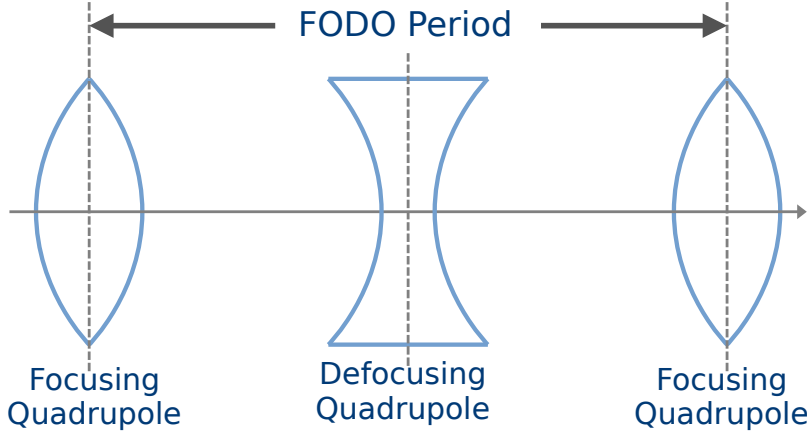


Figure 3.3: FODO cell consisting of focusing and defocusing quadrupole magnets. The strength of the focusing magnet is typically chosen at half the strength of the defocusing magnet. The spaces in between the quadrupole magnets can be occupied by accelerating cavities.

where  $L$  is the period of the system. The general solution of Hill's equation is given by [29]

$$x(s) = \sqrt{\varepsilon\beta(s)} \sin(\phi(s) + \phi_0) \quad \text{with} \quad \phi(s) = \int_{s_0}^s \frac{1}{\beta} dt. \quad (3.24)$$

The solution is an oscillator but the amplitude and phase shift  $\phi(s)$  depend on the position in  $s$ . The amplitude can be identified as the transverse beam envelope  $E(s) = \sqrt{\varepsilon\beta(s)}$ , with the beta function  $\beta(s)$ , that is defined by the magnet lattice and its geometry, and the emittance of the beam  $\varepsilon$ . The transverse beam size  $\sigma_T(s)$  can include a correction term to account for the energy spread  $\sigma_E = \Delta E/E$  and the resulting variation in the deflection in electromagnetic fields is defined by the dispersion function  $\eta(s)$ . The function for the transverse beam size follows as [29]

$$\sigma_T(s) = \sqrt{\beta(s)\varepsilon + (\eta(s)\sigma_E)^2}. \quad (3.25)$$

As a particle moves along  $s$ , it moves in phase space. This movement in phase space is constrained to an elliptical path. The number of revolutions on the ellipse per turn in the ring is defined by the particle tune. The emittance is a property of an ensemble of particles and defines the area of an ellipse in phase space  $A_{\text{Ellipse}} = \pi\varepsilon$  that encloses all particles of that ensemble [29]. The emittance is dependent on the particle energy, which can be disregarded for storage rings but for the linear injectors, an energy independent formulation is necessary, thus the normalized emittance  $\varepsilon_n = \gamma\beta_s\varepsilon$  is defined, with  $\beta_s$  being the normalized beam velocity. According to Liouville's theorem all particles inside this ellipse will remain there and the area of the

ellipse will be constant over time. Liouville's theorem holds for particle ensembles under conservative forces that are, for example, exerted by dipole and quadrupole magnets, such that revolutions around a synchrotron and beam injection and extraction from beam lines are subject to Liouville's theorem. However, there are limitations to applying this theorem in that the phase space area changes under synchrotron radiation, as it is not a closed system anymore.

Synchrotron radiation is electromagnetic radiation that is emitted when charged particles are accelerated. Both the acceleration in longitudinal direction and transverse direction cause synchrotron radiation. The radiated power  $P$  for both contributions is [29]

$$P_{\parallel} = \frac{2}{3} \frac{r_c}{mc} \left( \frac{d\vec{p}_{\parallel}}{dt} \right)^2 \quad \text{and} \quad (3.26)$$

$$P_{\perp} = \frac{2}{3} \frac{r_c}{mc} \gamma^2 \left( \frac{d\vec{p}_{\perp}}{dt} \right)^2, \quad (3.27)$$

where  $r_c$  is the classical particle radius,  $m$  its mass,  $\vec{p}$  its momentum and  $\gamma$  the Lorentz factor. The radiated power  $P_{\perp}$  is much larger due to  $\propto \gamma^2$  and therefore  $P_{\parallel}$  is disregarded in further calculations. In circular accelerators, particles are constantly subject to acceleration transverse to its direction of motion to keep particles on orbit and synchrotron radiation is tangentially emitted from the beam in the horizontal plane. When inserting the Lorentz force into Equation 3.27 and assuming singly charged particles, the total radiation power becomes [29]

$$P_{\gamma} = \frac{4\pi}{\mu_0} \frac{2r_c^2 c}{3(mc^2)^2} B^2 E^2. \quad (3.28)$$

Using Equation 3.4 the magnetic field  $B$  can be expressed using the bending radius  $r$  to obtain

$$P_{\gamma} = \frac{2c}{3} \frac{r_c}{(mc^2)^3} \frac{E^4}{r^2}, \quad (3.29)$$

where  $E$  is the energy of the particle.

### 3.3 Beam Collimation

The transverse particle distribution of a particle beam is characterized by two regions. A high-density beam core with its particles being on or close to the nominal beam orbit and a beam halo with particles that have a large betatron amplitude or are off-momentum and are traveling at distance to the nominal orbit. For Gaussian particle distributions the beam core is typically defined as all particles within  $\pm 3-4\sigma$ . Mechanisms that lead to a formation of beam halos include particle scattering, like Touschek and beam-gas scattering (see Section 9.2), magnet field errors and misalignments of optical systems. Essentially any component or effect that influences

the collective beam or individual particles adversely can promote the formation of a beam halo. Particles in the beam halo can be lost due to collisions with accelerator components producing a particle shower. This excess radiation can affect the lifetime and performance of detectors and activate accelerator or detector components. This also directly affects the beam life as this is a major contribution to the loss of stored beam particles. For the safety and performance of the accelerator and the detector it is therefore important to control the beam halo.

A collimator is a block, or series of blocks, of material to intentionally intercept halo particles and synchrotron photons from halo particles to avoid collisions with, or in the area of, sensitive components. Collimators define the aperture in a way that detector and accelerator components with the smallest aperture in the region of the interaction point (IP) are in the shadow of the collimators. Since the collision of a halo particle with a collimator can produce secondary particles, a series of collimator heads is used, where consequent collimators are designed to stop these secondary particles.

A collimation system usually consists of two parts, the betatron collimation system and the energy collimation system. The betatron collimation system specifically targets the first term of r.h.s of the beam size  $\sigma_T(s)$  in Equation 3.25 related to the betatron function, while the energy collimation system is targeting the second term that relates to the dispersion.

The aperture, or collimation depth  $n(s)$ , is defined as [29]:

$$n(s) = \frac{a_T}{\sigma_T(s)} = \frac{a_T}{\sqrt{\beta(s)\varepsilon + (\eta(s)\sigma_E)^2}}, \quad (3.30)$$

where  $a_T$  is the transverse half aperture. To effectively act on the betatron component, the betatron collimation system is placed in a region along  $s$  where the dispersion function is sufficiently small and can be neglected. Additionally, the collimator can be placed at a point where the betatron function is large since the same collimation depths can be reached with a larger aperture. The betatron function along a FODO cell is displayed in Figure 3.4. Similarly the energy collimation system should be positioned at a point where the dispersion term and betatron term are large.

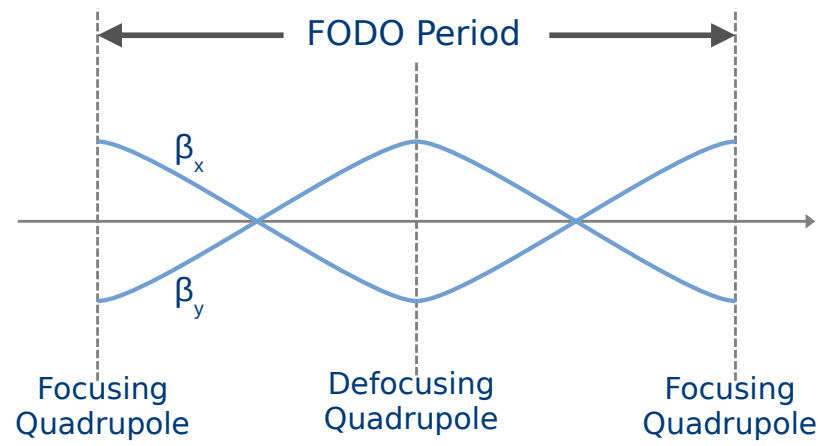


Figure 3.4: The betatron functions along a FODO cell.

## CHAPTER 4

---

### Silicon Sensor Technology

---

Semiconductors play a crucial role in modern electronics and are the basis for integrated circuits, diodes and many more applications. The use of imaging sensors in high energy collider experiments has been established since its first use in the DELPHI experiment in 1997 [34]. Typically, semiconductor pixel sensors exhibit improved spatial resolution and can cope with higher hit rates compared to previous technologies, e.g. wire chambers, albeit at higher costs. However, due to commercial processes for complementary metal–oxide–semiconductor (CMOS) devices, the costs associated to production have significantly reduced. This chapter focuses on physical concepts relating to the use of silicon semiconductor sensors in collider experiments and is based on [35–37].

#### 4.1 Silicon Semiconductors

Electrons in atoms are bound by a coulomb potential and occupy discrete energy levels. For molecules, the potential is a superposition of all potentials of the atoms in the molecule. The allowed states of electrons appear as orbitals that are defined by the superposed electric field of the molecule. When atoms are arranged in fixed lattices with a periodic structure, the electric field will also have a periodic structure and the allowed states align into bands across the lattice. The highest energy band below the Fermi energy is the valence band and the lowest energy band above the Fermi energy is the conduction band. Nominally, at temperatures  $T = 0$  K, all bands up to the Fermi energy are filled, while all bands above are empty. In this configuration, the charges in the valence band cannot move since all states are occupied and there are no charges in the conduction band. Once an electron is lifted into the conduction band, a current can flow there by virtue of the electron moving along the conduction band, and along the atomic lattice. Similarly, the vacancy in

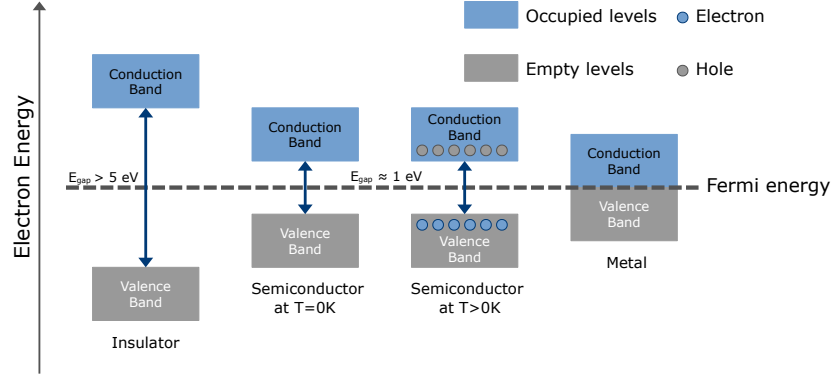


Figure 4.1: Band structure of insulators, semiconductors and metals. The semiconductor band structure is shown for temperatures of  $T = 0 \text{ K}$  and  $T > 0 \text{ K}$ . At  $T > 0 \text{ K}$  electrons can be lifted from the valence band into the conduction band by thermal excitation.

the valence band can propagate along the lattice and appear as a positive charge. The vacancy is commonly referred to as *hole*. The energy difference  $E_{\text{gap}}$  between the conducting and valence band differentiates the materials into insulators with  $E_{\text{gap}} > 5 \text{ eV}$ , semiconductors with  $E_{\text{gap}} \approx 1 \text{ eV}$  and conductors (metals), where the valence and conducting bands touch or overlap. At temperatures of  $T > 0 \text{ K}$  electrons in semiconductors can be lifted into the conduction band via thermal excitation and conduct current, while at  $T = 0 \text{ K}$  semiconductors act as insulators. The valence and conduction bands for different materials are visualized in Figure 4.1.

Silicon, being a group IV element, has four valence electrons and silicon crystals have a regular pattern with equidistant atoms. Each valence electron is covalently bound to two atoms and neighboring atoms share two covalently bound electrons. By introducing a group III element into the crystal structure a hole is introduced. The process of introducing such an impurity is called doping and doping with a group III element is called p-doping. At  $T = 0 \text{ K}$  the hole is stationary, while at  $T > 0 \text{ K}$  the hole can be filled by a valence electron, moving the hole. The crystal structure of a p-doped silicon crystal is shown in Figure 4.2. The p-type impurity introduces an additional energy level, the acceptor level, between the valence band and the Fermi energy. The distance between the acceptor level and the valence band is very small at  $\approx 0.005 \text{ eV}$ . Additionally, the Fermi level moves closer to the acceptor level. The main charge carrier in p-type silicon are therefore holes and the concentration of holes in the valence band is increased as at room temperature most acceptor states are filled.

Introducing a group V element into a silicon crystal is called n-doping and leads to an additional, weakly bound, electron. The crystal structure of an n-doped silicon crystal is shown in Figure 4.3. An additional donor level is introduced with a small distance of  $\approx 0.005 \text{ eV}$  to the conduction band and the Fermi energy, again, moves closer to the new energy level. The acceptor and donor levels and the shifted Fermi

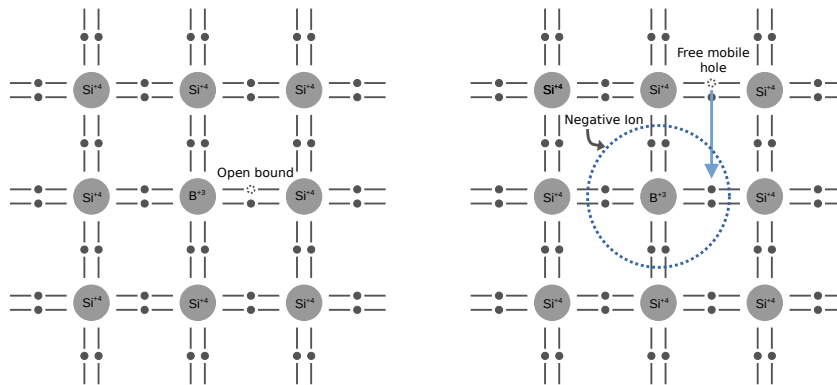


Figure 4.2: Left: Silicon crystal at  $T = 0$  K with an implanted boron atom introducing an open bound. Right: At  $T > 0$  K an electron from a covalent silicon bound may fill the hole. The hole is mobile.

energy for p- and n-type silicon are visualized in Figure 4.4.

#### 4.1.1 The pn Junction

When coupling p-type silicon and n-type silicon, electrons can be observed to diffuse into the p-type silicon and filling empty holes, leaving stationary negative ions due to the difference in the Fermi energy of both sides. Similarly, holes diffuse into the n-type silicon, leaving stationary positively charged atoms. This process stops once the negative charge density in the p-type silicon is large enough to stop further diffusion of electrons into the p-type silicon. During this process the Fermi energy of both sides approach each other and once the diffusion stops the Fermi energy of both sides is identical. The area where the ions formed is now devoid of charge free charge carriers and is called *depletion region*. The positive ions on the n-side and the negative ions on the p-side generate a contact potential  $V_0$  between them and an electric field forms. The potential  $V_0$  is proportional to the size of the depletion region. Due to the electric field, electrons in the conduction band can drift from the p-side back to the n-side and holes in the valence band may drift in the opposite direction. Additionally, the electrons and holes move due to thermal diffusion, which is inherently an undirected movement and they may move opposite to their respective drift direction.

The depletion region allows for a deeper penetration of the electric field into the bulk material as free charge carriers can screen out electric fields [37]. When a charged particle traverses the depletion region, atoms are ionized and charge are liberated. The negative and positive charges are immediately separated due to the drift created by the electric field and kept from immediate recombination. The absence of free charge carriers in the depletion region reduces the chance of recombination of the liberated charges with free charges present due to thermal excitations.

By applying an additional bias voltage, the size of the depletion region can be

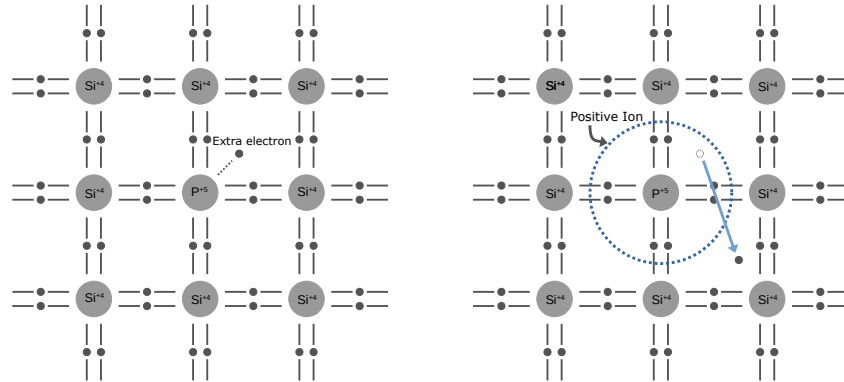


Figure 4.3: Left: Silicon crystal at  $T = 0\text{ K}$  with an implanted phosphorus atom introducing an extra, weakly bound electron. Right: At  $T > 0\text{ K}$  the electron may be lifted into the conduction band, creating a mobile charge.

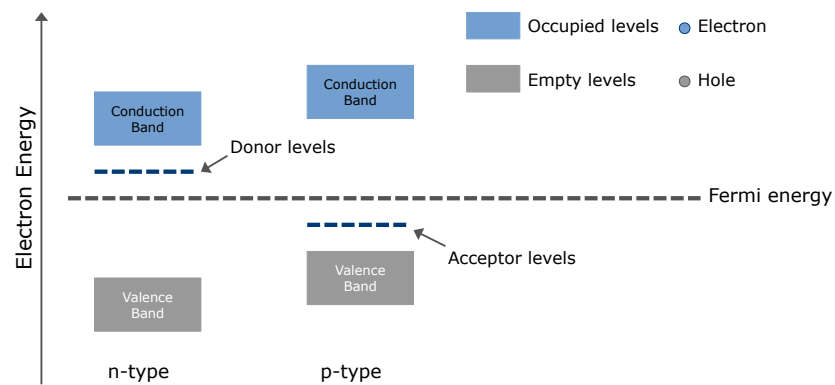


Figure 4.4: Band structure for p-type and n-type silicon semiconductors with the introduced acceptor and donor levels.



manipulated. Applying a voltage that counteracts the intrinsic voltage difference across the depletion region is called forward bias and will cause the depletion region to shrink. The depletion region can be reduced to small thicknesses allowing charges to tunnel through the depletion region. This effect is exploited in diodes. When applying a voltage that is aligned with the intrinsic voltage, a reverse bias, the depletion region can be increased. Depletion regions in the order of 20  $\mu\text{m}$  to 30  $\mu\text{m}$  are used in current day particle sensors (see TJ-Monopix2 and OBELIX chips in Chapter 7).

The geometry of the depletion region can further be manipulated by changing the doping concentrations of the n- and p-doped regions. The depletion region will reach further into the low-doped region and will have only a small depth in regions with large doping concentrations. The doping concentrations are chosen such that the n-doped region has a small doping concentration, creating the largest part of the sensitive volume in the n-doped region.

## 4.2 Particle Interactions with Matter

A relativistic particle that traverses a medium typically scatters with the bound electrons of atoms. While the atom is ionized or excited, the traversing particle loses energy. The average energy loss per path length, or stopping power, for heavy charged particles  $dE/dx$  is dependent on the medium and is described by the Bethe-Bloch formula [11, 35, 38, 39] for  $0.1 \lesssim \beta\gamma \lesssim 1000$ :

$$\left\langle -\frac{dE}{dx} \right\rangle = Kz^2 \frac{Z}{A} \frac{1}{\beta^2} \left[ \frac{1}{2} \ln \frac{2m_e c^2 \beta^2 \gamma^2 W_{\max}}{I^2} - \beta^2 - \frac{\delta(\beta\gamma)}{2} \right] \quad (4.1)$$

The parameters of the Bethe-Bloch formula are displayed in Table 4.1. The formula is generally only valid for particles heavier than electrons,  $m > m_e$ .

The stopping power of a heavy charged particle in the Bethe-Bloch regime  $0.1 \lesssim \beta\gamma \lesssim 1000$  therefore depends both on the particle properties as well as the properties of the medium. At  $\beta\gamma \approx 3$  a minimum can be found. Particles that fall into the minimum are called minimum ionizing particles (MIPS). The Bethe-Bloch regime is the most relevant for particle detection in particle collision experiments as the energy of the produced particles typically lies within this regime. Thus, silicon sensors are often tested using MIPS as they leave the smallest signals in the sensor and are more difficult to detect, but make up a large fraction of the produced particles. At  $\beta\gamma \gtrsim 1000$  the dominant process becomes bremsstrahlung, while at lower energies  $\beta\gamma \lesssim 0.01$  the Lindhard-Scharff approximation and the phenomenological parameterization by Andersen and Ziegler describe the stopping power. The stopping power over the different regimes for muons in copper is displayed in Figure 4.5.

For test beam measurement with thin silicon sensors the instantaneous energy loss is important. The instantaneous energy loss distribution for thin media is not symmetrical and described by a Landau distribution [35, 40]. Single collisions with very large momentum transfers are possible, although their probability is small. The

| Variable              | Description                                                          |
|-----------------------|----------------------------------------------------------------------|
| $c$                   | Light speed                                                          |
| $Z$                   | Atomic number                                                        |
| $A$                   | Number of protons and neutrons                                       |
| $m_e$                 | Mass of the electron                                                 |
| $v$                   | Particle velocity                                                    |
| $\beta$               | $v/c$                                                                |
| $\gamma$              | Lorentz factor $1/\sqrt{1 - \beta^2}$                                |
| $K$                   | $4\pi N_A r_e^2 m_e c^2$                                             |
| $N_A$                 | Avogadro's number                                                    |
| $W_{\max}$            | Maximum transferred energy loss to an electron in a single collision |
| $z$                   | charge number of incident particle                                   |
| $I$                   | Mean excitation energy                                               |
| $\delta(\beta\gamma)$ | Density effect correction to ionization energy loss                  |

Table 4.1: Variables, and their description, of the Bethe-Bloch formula in Equation 4.1.

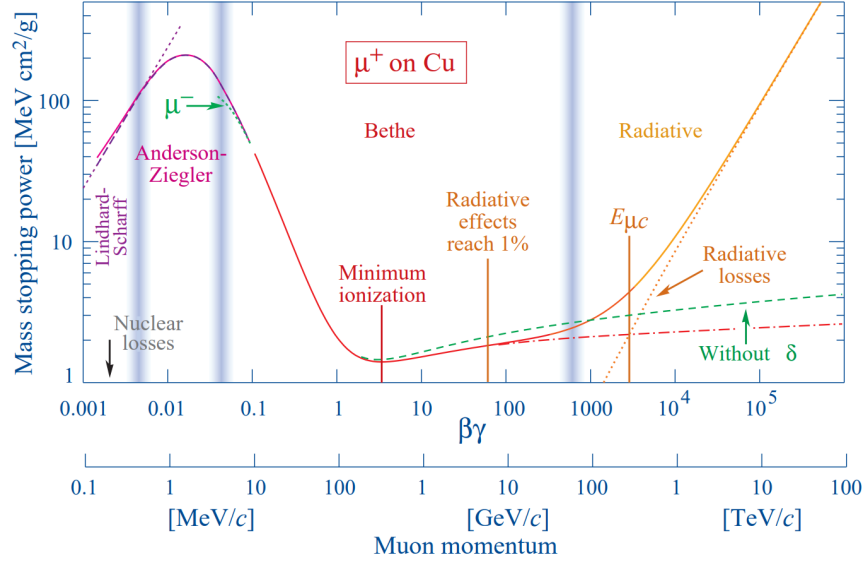


Figure 4.5: Linear stopping power for muons in copper [11].

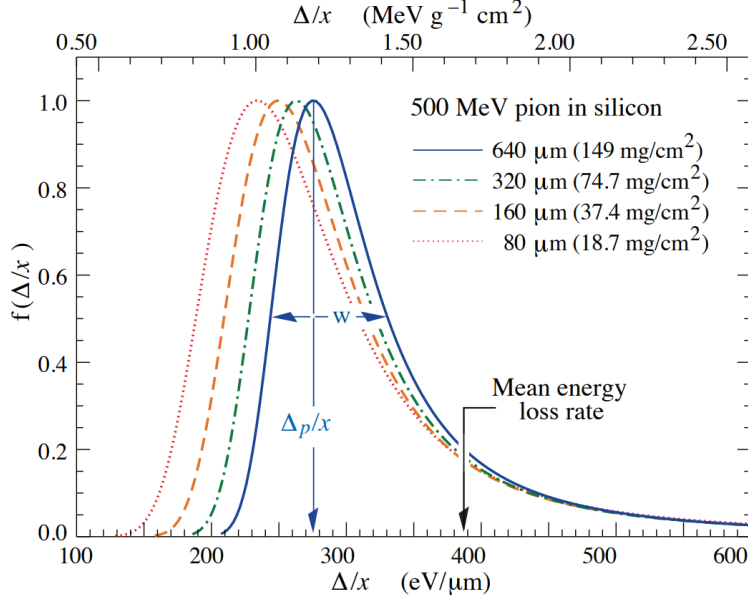


Figure 4.6: Energy loss distributions for 500 MeV pions in silicon at different thicknesses. The distributions are Landau-shaped and are normalized to unity at the most probable value  $\Delta_p/x$ . The width  $w$  is the full width at half maximum [11].

average energy loss described by the Bethe-Bloch formula is affected by the large weight of these collisions. A description of the most probable energy loss becomes a more practical description for the energy loss. Examples for a Landau-shaped energy loss distributions in silicon for different thicknesses are displayed in Figure 4.6.

For electrons with energies  $\mathcal{O}(1 \text{ GeV})$  the primary process of energy loss is bremsstrahlung. Below that threshold at energies  $\mathcal{O}(10 \text{ MeV})$ , the electrons, and similarly positrons, lose energy primarily through ionization and excitation. The energy loss due to different processes with respect to the electron energy is displayed in Figure 4.7. The losses through ionization happen at all energies, the fraction of energy lost may become very small compared to bremsstrahlung, as depicted in Figure 4.7. Ionization of high energy electrons is necessary for the test beam measurements discussed in Chapter 8. Here, electrons with energies of 4.2 GeV are used and their ionization of silicon is necessary to detect them with the sensors used and the energy loss due to ionization is similar to what is expected for MIPS. Bremsstrahlung photons are typically not measured in test beams.

Electrons that undergo bremsstrahlung emit a photon in presence of an electric field of an atomic nucleus. Depending on the energy of the emitted photon, different processes dominate. At photon energies  $< 1 \text{ MeV}$  the photoelectric effect dominates the cross-section. The photon interacts with a bound electron in the shell of an atom and is absorbed. In a similar energy regime the photon also undergoes Thomson scattering, which describes elastic scattering between the photon and an unbound elec-

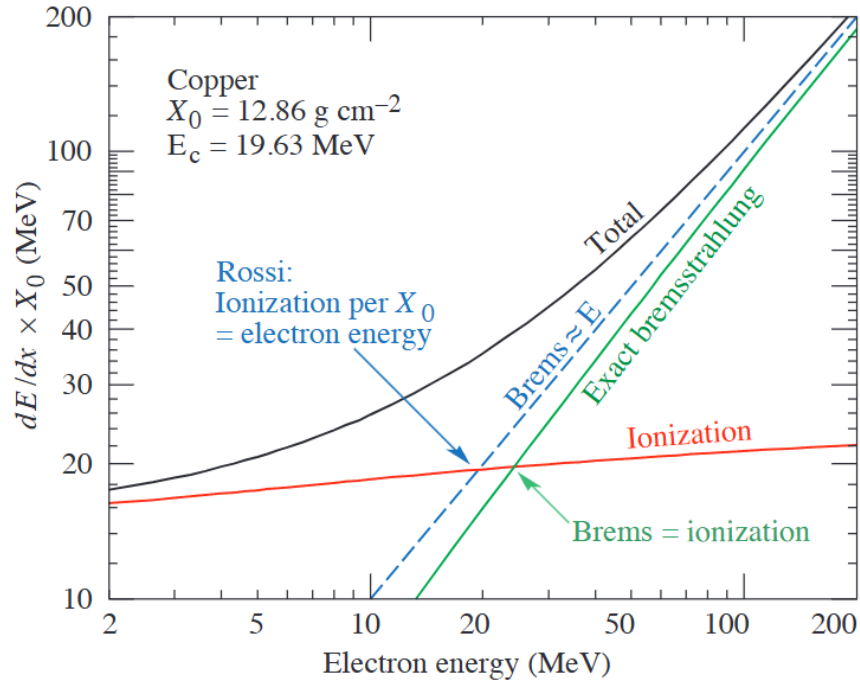


Figure 4.7: Energy loss for electrons in copper due to bremsstrahlung and ionization. The electron energy at which the energy loss due to ionization and bremsstrahlung is equal is indicated for different definitions. This critical energy is at  $E_e \approx 20 \text{ MeV}$  [11].

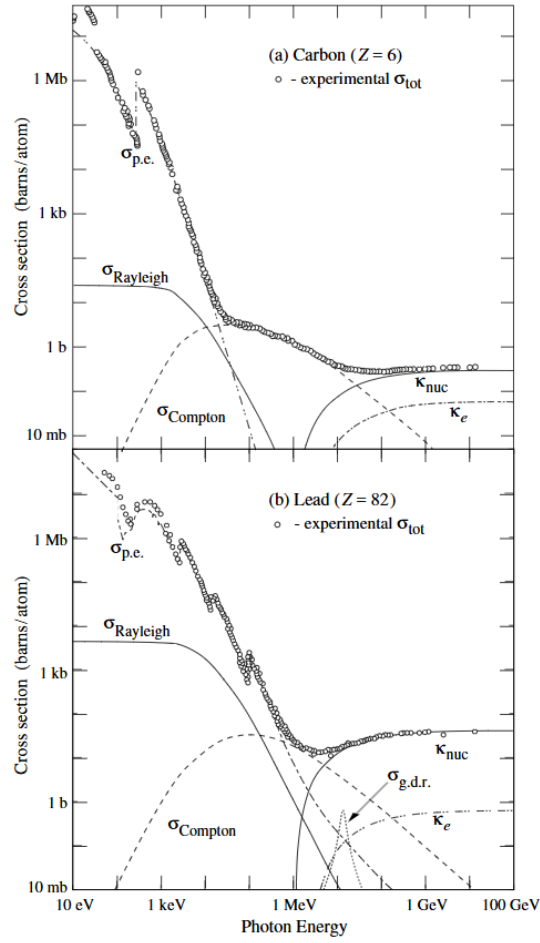


Figure 4.8: Photon cross-sections for interactions with matter depending on the photon energy in carbon and lead [11].

tron and Rayleigh scattering between the photon and an atom. At photon energies  $\mathcal{O}(1 \text{ MeV})$  the dominant process is inelastic scattering, called Compton scattering. Beyond energies of  $> 10 \text{ MeV}$  electron-positron pair production is the dominant process, although the production threshold is at  $E_\gamma \leq 2m_e c^2 = 1.022 \text{ MeV}$ . To conserve momentum the photon splits into an electron-positron pair in the presence of a nucleus. The resulting high energy electron and positron undergo bremsstrahlung due to the deflection in matter. The photons emitted may continue this cycle, given enough energy. The result of this chain reaction, or avalanche, is an electromagnetic shower. This process can similarly be started by an initial photon. An overview of cross-sections for processes of photons in matter depending on the photon energy is displayed in Figure 4.8.

### 4.3 Signal Generation

Once a charged particle traverses the depletion region, charges are liberated by ionization. The produced charges, that are electrons and holes, move apart due to the presence of the electric field  $E = |\vec{E}|$ . The movement due to an electric field is typically referred to as *drift*. The drift velocities for electrons and holes are described by [35]:

$$v_d^e = \mu_e E \quad (4.2)$$

$$v_d^h = \mu_h E. \quad (4.3)$$

The  $\mu_e$  and  $\mu_h$  are the mobility of electrons and holes, respectively. The mobility is dependent on the electric field  $\mu_{e,h} = \mu_{e,h}(E)$ . The mobility is empirically described for typical values of the field strength  $E$  [37] by:

$$\mu = \frac{\mu_0}{\left[1 + \left(\frac{\mu_0 E}{v_{\text{sat}}}\right)^\beta\right]^{\frac{1}{\beta}}}, \quad (4.4)$$

where  $\mu_0$  is the low field mobility,  $v_{\text{sat}}$  is the saturation drift velocity and  $\beta$  is an empirically determined parameter. Electrons typically have a higher mobility with  $\beta_e \approx 2$  compared to holes, which have  $\beta_h \approx 1$  [41].

An additional mode of movement is described by *diffusion*. Here, it describes the thermal motion of charge carriers along the charge density gradients. Diffusion is a stochastic process and does not describe an inherently directed movement. The longitudinal motion due to diffusion along the drift direction of charge carriers is small compared to the drift velocity and thus not of particular importance. The transverse motion, however, causes charge clouds to spread in the transverse direction, affecting the measurement in pixel detectors. The transversal spread is [35]

$$\sigma_T = \sqrt{2D_e t_D} \quad (4.5)$$

with  $D_e$  being the diffusion coefficient of electrons and  $t_D$  is the drift time. The diffusion coefficient is dependent on the mobility  $\mu_e$  and the temperature  $T$  via [35]

$$D = \mu_e k_B T. \quad (4.6)$$

At sufficient transverse sizes, the charge cloud can induce signals at multiple electrodes in pixel detectors, which can have an immediate effect on the spatial resolution. This is discussed in detail in Section 8.3.2.

A current  $I_s$  at the charge collecting electrode  $i$  is induced by moving charges according to the Shockley-Ramo theorem [42, 43]:

$$I_{s,i} = q \vec{E}_{w,i} \vec{v}, \quad (4.7)$$

where  $q$  is the charge of the moving charge carrier,  $v$  its drift velocity (diffusion is neglected) and  $\vec{E}_{w,i}$  is the electric weighting field. The electric weighting field is related to the weighting potential  $\vec{E}_{w,i} = -\vec{\nabla}\phi_{w,i}$ . The weighting potential is obtained when applying a potential  $U_i = 1$  to the electrode  $i$ , while all other electrodes are at a potential of zero. Thus, the weighting potential is a geometric property of the sensor and its electrodes. The charge at the charge collecting electrode  $Q_{s,i}$  due to the induced current of a charge moving from  $\vec{r}(t_0)$  to  $\vec{r}(t)$  from time  $t_0$  until  $t$  can be obtained via integration

$$Q_{s,i}(t) = -q \int_{t_0}^t \vec{E}_{w,i} \vec{v} dt' \quad (4.8)$$

$$= q [\phi_{w,i}(\vec{r}(t)) - \phi_{w,i}(\vec{r}(t_0))] \quad (4.9)$$

This charge is the signal measured by a charge collecting electrode in a pixel detector.

## 4.4 Radiation Damage

When particles interact with a silicon sensor, i.e. particles scattering with silicon atoms, reversible and irreversible changes to the silicon crystal lattice can be introduced. The damages can be divided into two categories, the bulk damage and the surface damage, which are discussed in the following sections. This section is based on References [44–46].

### 4.4.1 Bulk Damage

When massive particles traverse the silicon bulk, they can create defects in the silicon lattice. A model that describes these effects is the non-ionizing energy loss (NIEL) hypothesis [47]. When a hadron with sufficient energy traverses the silicon, it can kick a silicon atom out of the lattice, leaving a vacant slot and an interstitial atom in the lattice, called a Frenkel pair. The energy threshold to create a Frenkel pair in a silicon lattice is 25 eV [45]. The interstitial silicon atom will carry most of the remaining energy of the initial hadron. Depending on the energy of the interstitial silicon atom it can scatter with other silicon atoms in the lattice and create cluster damage. It is assumed that the damage to the lattice is linearly dependent with the energy loss due to the displacement. Each particle type has a different defect distribution and “hardness factor”. To obtain a unified scale to express bulk damage caused by radiation, the 1 MeV neutron equivalent fluence is used, by scaling each type of radiation with its hardness factor.

The changes in the lattice affect the electrical properties of the silicon by shifting energy levels and creating new energy levels in the band gap. Explicitly, the leakage current, the depletion voltage and the charge collection efficiency are affected. This is visualized in Figure 4.9.

**Leakage current:** In the bulk of silicon sensors there is a constant current that is caused by the thermal generation of electron-hole pairs. The threshold for

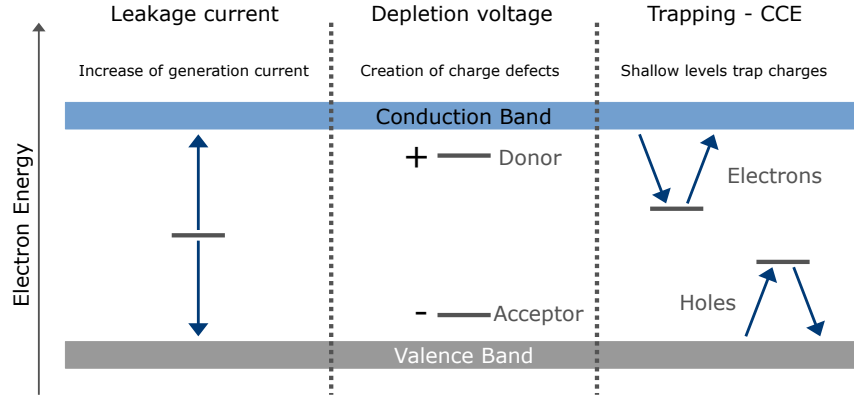


Figure 4.9: Effects of bulk damage and the effects on the leakage current, depletion voltage and charge collection efficiency (CCE).

generating electron-hole pairs is  $\approx 3.6\text{eV}$ . Defects with energies in the middle of the band gap need less energy to be ionized, thus the energy needed to generate an electron-hole pair is reduced and the leakage current increases. The increase of leakage current is proportional to the 1 MeV neutron equivalent fluence the sensor has received.

**Depletion Voltage:** When displaced by irradiation, donor atoms may not be able to further act as donors, and di-vacancy defects (a pair of holes) may be introduced, increasing the density of acceptors. The space charges created in the band gap and the change of concentration of charge carriers changes the effective doping concentration and the depletion voltage. In an n-doped substrate, the depletion voltage drops with irradiation until a type inversion into a p-type silicon occurs. After the type inversion to a p-type substrate, the depletion voltage rises again [48].

**Charge collection deficiency:** Electrons and holes may be trapped by defects in the band gap, while drifting towards the charge collection electrode. The drifting charges will contribute to the charge induced at the electrode until they are trapped and stationary. The charges may eventually be released before the integration time is over and can again contribute to the collected charge. If the charges are released outside a integration time window and the signal is already formed, they will contribute to the leakage current only. The trapping of charges is a main contributor to the loss of efficiency in irradiated sensors. The amount of defects and therefore the impact of charge trapping scales with the NIEL fluence.

Most defects, such as interstitial atoms, space charges or impurities, may travel through the lattice. Given enough time, some defects, such as vacancies and interstitial atoms, may recombine in a process called annealing [44]. This effectively reverts defects back to their original, undamaged state. The annealing rate depends on the temperature and will also happen during the operation of silicon sensors in the experiments.



#### 4.4.2 Surface Damage

Surface damage describes damage to the  $\text{SiO}_2$  surface and the bulk-surface interface, mostly due to gamma irradiation. Photons with energies below 1 MeV undergo Compton scattering or the photoelectric effect. The secondary electrons produced ionize the material, creating electron-hole pairs. The pairs are separated by electric fields when the sensor is operated and cannot recombine. Some electrons and holes may be trapped in the  $\text{SiO}_2$  or in the  $\text{Si-SiO}_2$  interface, creating oxide charges. Due to their low mobility, holes are more likely to remain in the material, forming a positive oxide charge barrier between the surface and the bulk. Below the oxide charge barrier, electrons accumulate, preventing a full depletion of the pixel volume. Due to the charges, the electrical field is deformed, causing localized high electric fields, reducing the voltage breakdown threshold. Surface damage effects that are caused by photon irradiation scale with the total ionizing dose (TID) deposited in the material, i.e. the energy deposited by ionization per unit mass. Also the additional charges at the bulk interface can affect the working point of transistors, possibly pushing them out of their operational ranges. Oxide charges may also recombine over time and anneal.

### 4.5 Depleted Monolithic Active Pixel Sensors

Since both the TJ-Monopix2 chip and the OBELIX chip are depleted monolithic active pixel sensors (DMAPS) this technology is presented in this section. This section is based on [35–37, 49] for general information on sensors and [50, 51] for specific information on the OBELIX and TJ-Monopix2 design.

In monolithic active pixel sensors (MAPS) the readout electronics are integrated into the same monolithic piece of silicon used for the sensor. However, the electronics need to be shielded from the sensor to not affect each other, i.e. avoid crosstalk between the digital and analog domain of the sensor. This electrical separation is achieved by placing the readout electronics in a deep p-well. Once the integration of an extra deep p-well was possible using commercial and cheap CMOS processes, MAPS gained traction. TJ-Monopix2 was produced in 180 nm CMOS processes. A cross-section of the TJ-Monopix2 pixel is displayed in Figure 4.10 The OBELIX chip is similarly designed in the same 180 nm CMOS process. In early implementations an n-well was placed in the p-well to act as a charge collection electrode, however, the depletion region was very limited. The charges were generated in a p-type epitaxial layer below the p-well. One issue regarding this design was the radiation hardness. Its advantages were small material budgets, high spatial resolution, affordability and low noise [52, 53].

The design was improved by introducing a dedicated small collection electrode next to the p-well containing the readout electronics. To increase the size of the depletion region, an n-doped layer across the full length of the pixel was introduced in the top part of the sensor above the p-type epitaxial layer. The small size of the

charge collection electrode leads to a small capacitance and reduces the noise. The TJ-Monopix2 chip and OBELIX chip have small collection electrodes. A competing design uses large collection electrodes that surround the p-well and also deplete the full volume of the pixel. The large collection electrode leads to mostly linear electric field and short drift paths through the pixel, improving the radiation hardness.

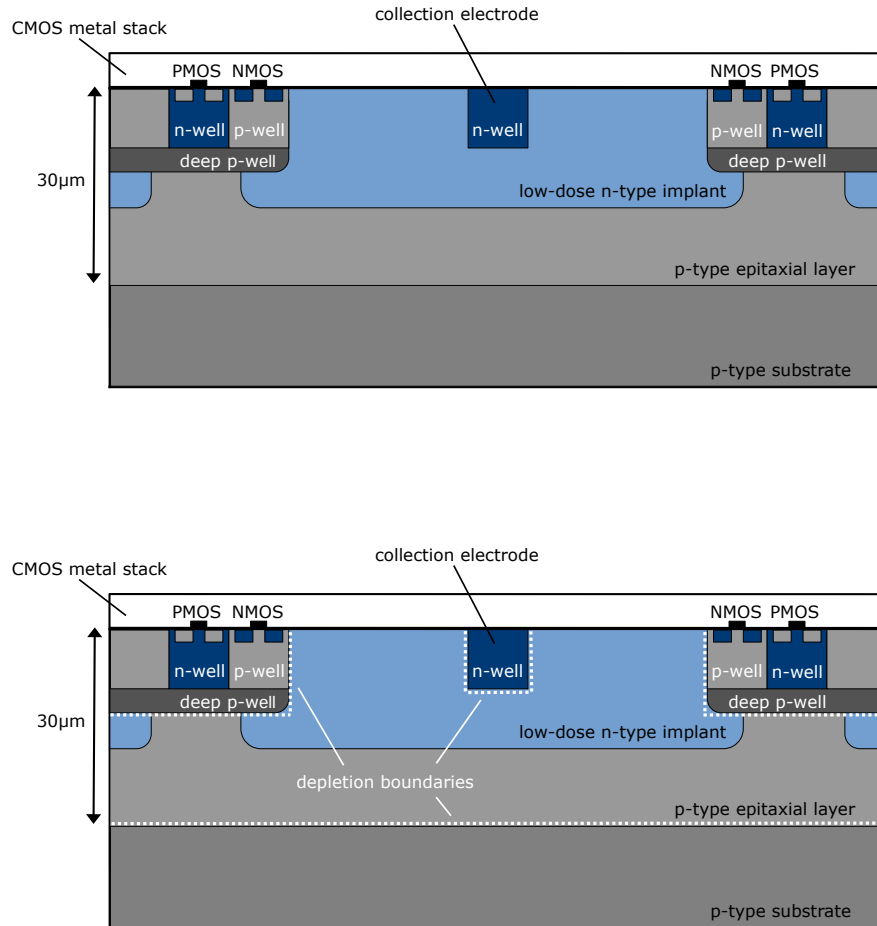


Figure 4.10: Top: Sensor geometry of a TJ-Monopix2 DMAPS pixel. Bottom: Sensor geometry with indicated boundaries for the depletion region. Modified from [53, 54].

## 4.6 Charge Collection, Digitization and Threshold

The analog signal in the pixel is amplified, modified (or shaped), and digitized. This and related concepts are discussed in this section. The collected charge  $Q_{\text{sig}}$  produces a voltage  $V_{\text{sig}}$  at the input of the amplifier circuit depending on the capacitance  $C$  of the charge collection electrode like

$$V_{\text{sig}} = \frac{Q_{\text{sig}}}{C}. \quad (4.10)$$

An example of an incoming pulse is displayed in Figure 4.11. The front-end circuit implemented in TJ-Monopix2 and indications of the signal shape is displayed in Figure 4.12.

The incoming signal raises the voltage from a baseline and over a constant discriminator threshold. The discriminator outputs a digital signal and is 1 as long as the signal is above the threshold. The length of the digital signal is the Time-Over-Threshold (ToT). The transition of the digital signal from 0 to 1 signifies the position of the leading edge (LE) and the transition from 1 to 0 signifies the position of trailing edge (TE) with  $\text{ToT} = \text{TE} - \text{LE}$ . The analog charge pulse is therefore converted into a digital signal and the length of the digital signal is proportional to the collected charge. Since the ToT of a signal is dependent on the front end settings, the ToT units only provide a relative measurement of the charge, which needs to be calibrated (see Section 8.1.1). To increase the precision of the ToT, a quick rise time of the signal is desired, as well as a longer signal, to cover a larger range of ToT values. The latter is done by shaping the signal in the amplifier circuit.

The threshold can be measured by injecting charge into the charge collection electrode. This is done applying a configurable voltage on the charge collecting electrode through an injection capacitance. When the charge is injected, the front end response can be probed. If the injected charge is too small the signal pulse will not cross the threshold and no hit will be registered. The injected charge is increased continuously in small steps until the signal crosses the threshold and a non-zero ToT can be read out. When disregarding noise, this would produce a step-function where each injection below the threshold is not registered, whereas every injection above the threshold is registered. However, noise fluctuations will push some injections over the edge and registering a hit, even though the injected charge alone would have been too small and vice versa. This smears out the step-function and it becomes an Gaussian error-function, which is called s-curve in this context. This assumes that the noise is gaussian distributed. The threshold is then determined to be the value at which a pixel fires 50% of the times. The noise is determined by the slope of the s-curve at the point of the threshold and called equivalent noise charge (ENC). An example for an s-curve is shown in Figure 4.13.

The threshold is now determined in units of the configurable voltage applied to inject the charge. To determine the threshold in terms of a charge in units of electron charges  $e^-$ , the capacitance of the injection capacity needs to be determined. Then, the Equation 4.10 can be applied. The absolute value of the capacitance can be

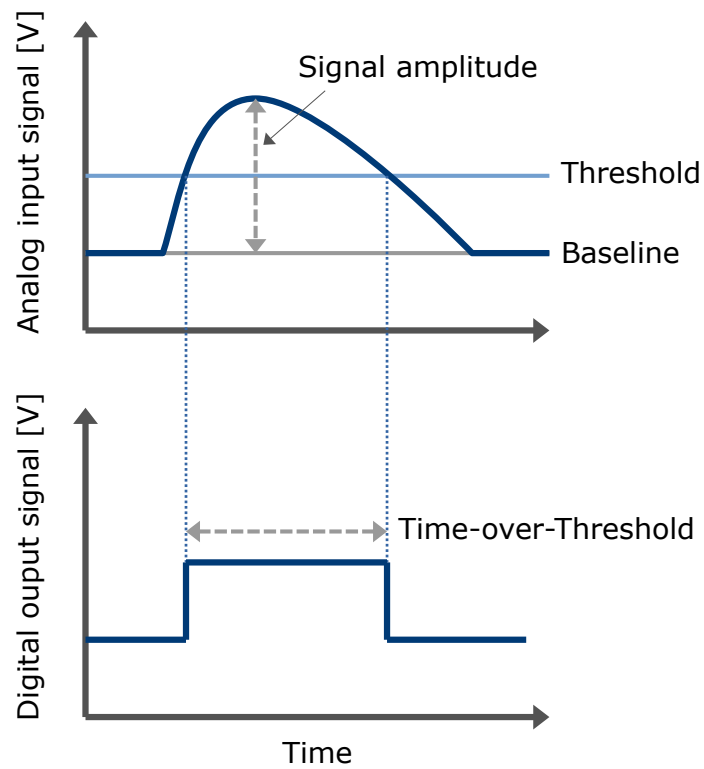


Figure 4.11: Analog input signal with a signal amplitude larger than the discriminator threshold and the corresponding digital output signal with a length equal to the Time-over-Threshold.

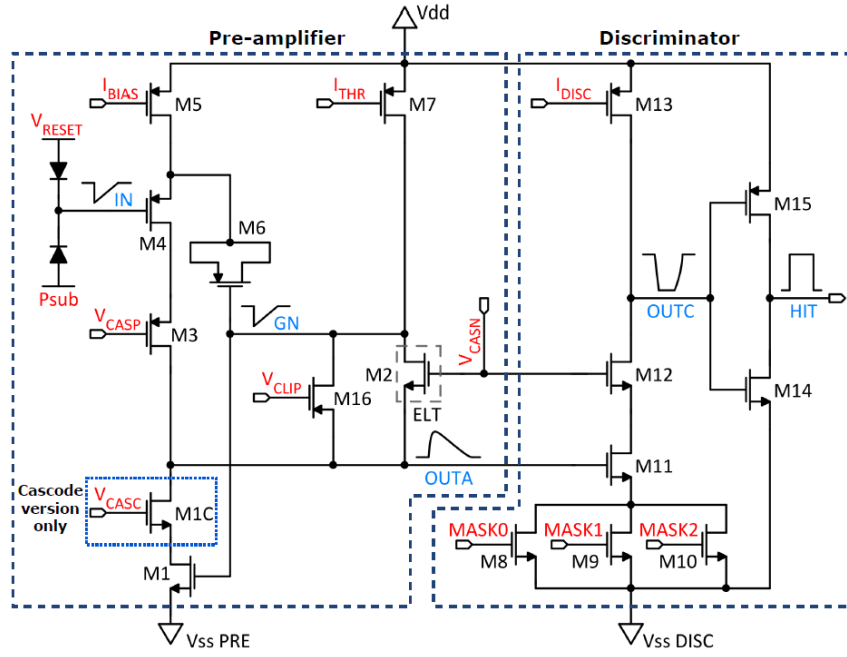


Figure 4.12: Schematic of the TJ-Monopix2 front-end circuit. The shape of the signal is indicated throughout the circuit [54].

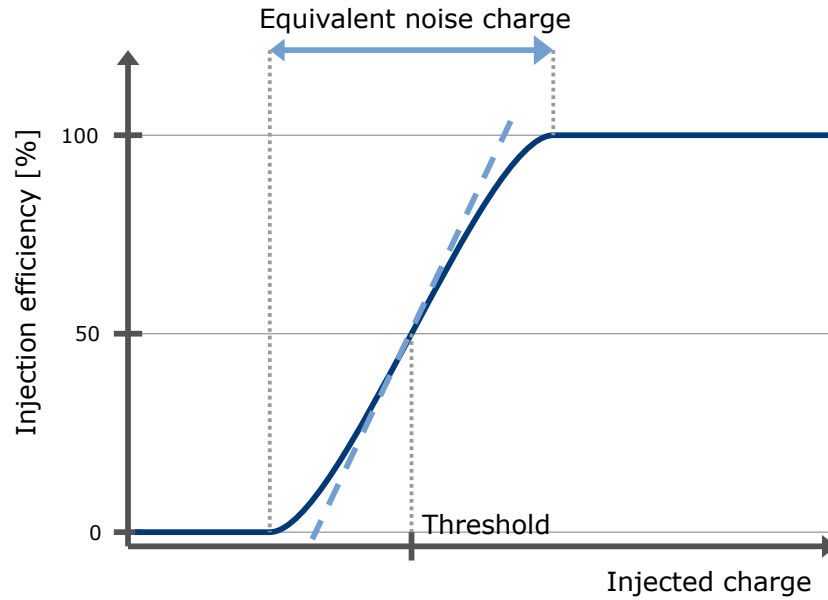


Figure 4.13: S-curve measurement to determine the threshold using injections.

determined by reference measurements with x-rays at discrete energies, for example using an  $\text{Fe}^{55}$  source with a characteristic peak at 5.89 keV.





## CHAPTER 5

---

### SuperKEKB

---

The SuperKEKB collider is located in Tsukuba, Japan, at the KEK facility [27]. SuperKEKB is comprised of two separate storage rings with counter-rotating, asymmetric beams of electrons and positrons, respectively. The electrons in the high energy ring (HER) have an energy of 7 GeV, while the positrons in the low energy ring (LER) have an energy of 4 GeV. The center-of-mass energy  $E_{CM} = 10.48$  GeV is at the  $\Upsilon(4S)$  resonance, which decays into a pair of B mesons with a branching fraction of  $\text{BR}(\Upsilon(4S) \rightarrow B\bar{B}) > 0.96$  [11]. The collider is therefore commonly referred to as a B-factory. The beams could be brought to collision at four potential points along the straight sections of SuperKEKB. Only one such interaction region is used and is occupied by the Belle II detector.

The asymmetric energies are necessary to create a decay vertex displacement between the pair of B mesons, originating from the  $\Upsilon(4S)$ . With a B meson lifetime of  $1.5 \times 10^{-12}$  s [11] and a boost of  $\beta\gamma = 0.28$ , the distance is approximately 130  $\mu\text{m}$ . The  $\Upsilon(4S)$  meson's lifetime of  $10^{-20}$  s is sufficiently small to assume that its decay products originate from the IP. The necessity for creating this displacement was motivated in Chapter 2 and a schematic of the vertex displacement is displayed in Figure 2.5.

The collider was designed to deliver an instantaneous luminosity of  $8 \times 10^{35} \text{ cm}^{-2}\text{s}^{-1}$  but due to operational difficulties this expectation was since reduced to  $6 \times 10^{35} \text{ cm}^{-2}\text{s}^{-1}$ . Since the start of operation in 2018 the highest instantaneous luminosity reached is  $5.1 \times 10^{34} \text{ cm}^{-2}\text{s}^{-1}$ . A number of key variables of SuperKEKB and a comparison to the predecessor machine KEKB are displayed in Table 5.1.

Each beam consists of approximately 2500 bunches with  $\mathcal{O}(10^{10})$  particles per bunch. The beam life in each ring is  $\mathcal{O}(20 \text{ min})$  due to particle loss by collisions and beam background effects (see Section 9.2). To keep the beam currents and the instantaneous luminosity up, particles are continuously injected at a rate of 50 Hz.

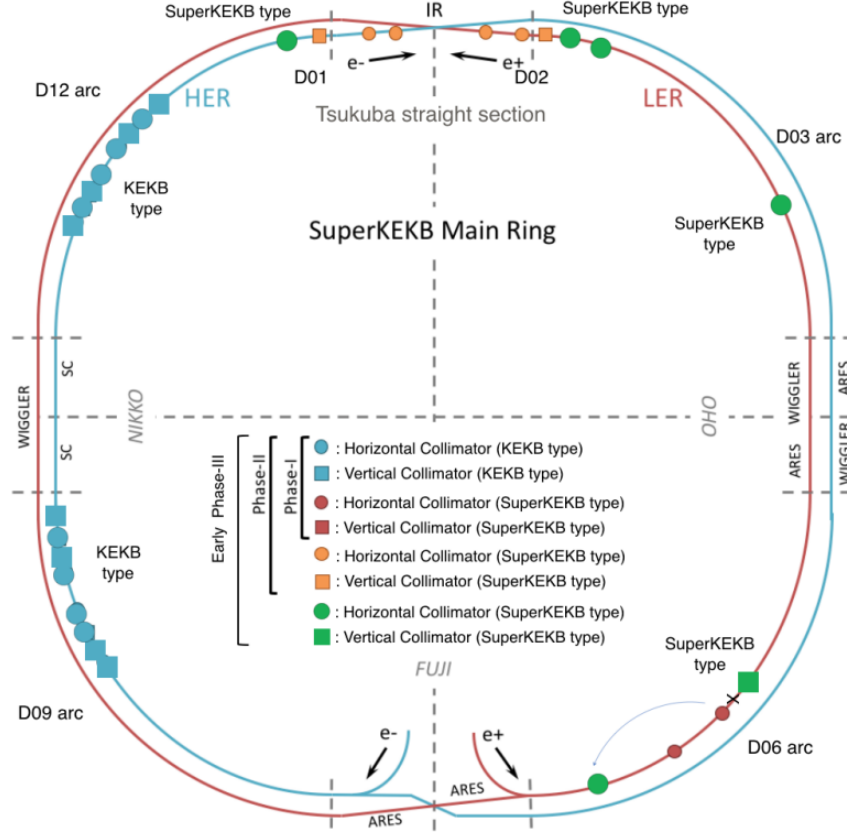


Figure 5.1: Map of collimators in SuperKEKB. The arc sections D03, D06, D09, D12 are indicated. They, together with section D01, the Tsukuba straight section and location of Belle II, house most of the collimators installed in SuperKEKB [55].

To suppress backgrounds collimators are placed around both rings. Collimators are one of the most important tools to mitigate beam backgrounds in Belle II. The layout of the collimators is displayed in Figure 5.1. The collimators used by SuperKEKB are put into two categories for the purposes of this work: Horizontal and vertical collimators. Each collimator has two collimator jaws on opposing sides along their respective orientation. The position of a collimator jaw can either be given as an absolute position relative to its mounting on the beam pipe or as a relative distance to the beam. The latter is the relevant measure to mitigate beam backgrounds. The collimators have systematic names according to their orientation, section and number in section.

A schematic depiction of SuperKEKB including the LINAC is displayed in Figure 5.2.

| Parameter             | KEKB                       | SuperKEKB                      | Unit                                     |
|-----------------------|----------------------------|--------------------------------|------------------------------------------|
| Beam energy           | 8.0( $e^-$ ), 3.5( $e^+$ ) | 7.0( $e^-$ ), 4.0( $e^+$ )     | GeV                                      |
| Beam current          | 1.2( $e^-$ ), 1.5( $e^+$ ) | 2.6*( $e^-$ ), 3.6*( $e^+$ )   | A                                        |
| Number of bunches     | 1584                       | 2500                           | -                                        |
| Horizontal beam size  | 170                        | 10.1*( $e^+$ ), 10.7*( $e^-$ ) | $\mu\text{m}$                            |
| Vertical beam size    | 940                        | 48*( $e^+$ ), 62*( $e^-$ )     | nm                                       |
| Inst. luminosity      | 0.21                       | 6*                             | $10^{35} \text{ cm}^{-2} \text{ s}^{-1}$ |
| Integrated luminosity | 1.04                       | 50*                            | $\text{ab}^{-1}$                         |

Table 5.1: Comparison of performance parameters of KEBB and SuperKEKB [56].  
 (\*) marked values are nominal or design values that have not been reached yet.

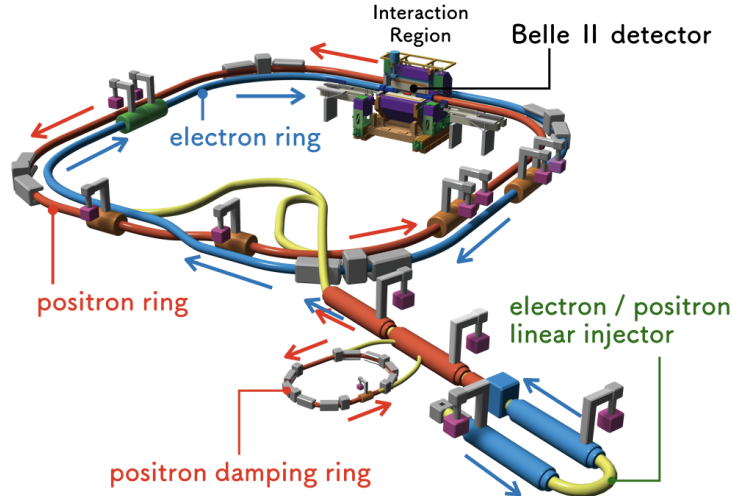


Figure 5.2: Schematic depiction of the SuperKEKB collider, as well as the linear accelerator and the positron damping ring [57].

## 5.1 Luminosity

The counter rotating beams of electrons and positrons are brought to collision in the interaction region, which is enclosed by the Belle II detector. The number of physics events  $n$  produced with a cross-section  $\sigma$  of the respective physical process

$$n = \int_0^T \mathcal{L} \sigma \, dt \quad (5.1)$$

provides an effective definition of the instantaneous luminosity. A cross-section is a parameter of a stochastic process in units of an area and provides a measure of probability for a quantum mechanical process to occur, however, it is not a normalized quantity and cannot be directly interpreted as a probability. The cross-section depends on the physics involved in the process as it depends on the transition element obtained from Fermi's golden rule (see Chapter 2).

The aggregate of all particles, and their measurements, that stem from a single collision are considered an *event*. Measuring many events, i.e. collecting a data set with a large integrated luminosity, and counting their occurrences and structure allows to measure the cross sections of the observed events. The total delivered, or integrated, luminosity  $L_{\text{int}}$  over a time frame  $T$  can be obtained by integration

$$L_{\text{int}} = \int_0^T \mathcal{L} \, dt. \quad (5.2)$$

Assuming the same number of bunches  $n_b$  and a Gaussian shaped beam profile of the bunches in both beams yields

$$\mathcal{L} = \frac{N_1 N_2 n_b f_{\text{rev}}}{2\pi \sqrt{\sigma_{1x}^2 + \sigma_{2x}^2} \sqrt{\sigma_{1y}^2 + \sigma_{2y}^2}}, \quad (5.3)$$

where  $N_1$  and  $N_2$  are the total number of particles in each beam,  $\sigma_{1x}$  ( $\sigma_{2x}$ ) and  $\sigma_{1y}$  ( $\sigma_{2y}$ ) are the beam sizes for bunch 1 (2) in the transverse coordinates  $x$  and  $y$  and  $f_{\text{rev}}$  is their revolution frequency. The Equation 5.3 disregards several effects like beam-beam effects and the actual bunch geometry arising from the nano-beam scheme [27]. A simplified equation for the instantaneous luminosity becomes [56]

$$L = \frac{\gamma_{\pm}}{2er_e} \left( \frac{I_{\pm} \xi_{y\pm}}{\beta_{y\pm}^*} \right), \quad (5.4)$$

where  $\gamma$  is the Lorentz factor,  $e$  the elementary charge and  $r_e$  the classical electron radius, while the index  $\pm$  denotes the reference to either the positrons (+) or electrons (−). The parameter  $\xi_y$  is the vertical beam-beam parameter and  $\beta_y^*$  is the vertical beta function at the IP. This representation of the equation shows how the luminosity is predominantly determined by modifying the beam currents  $I_{\pm}$  or  $\beta_y^*$ , while  $\xi_y$  is usually set to a value of  $\xi_y \approx 0.09$ .

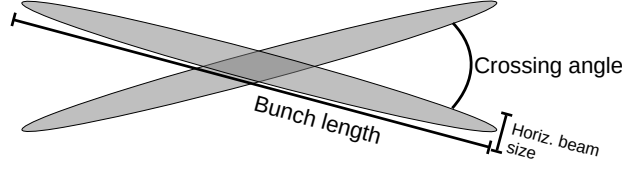


Figure 5.3: Schematic illustration of the nano-beam scheme in the horizontal plane. The crossing angle is 83 mrad, the bunch length is  $\approx 6$  mm and the horizontal bunch size is  $\approx 10 \mu\text{m}$  [27].

The nano-beam scheme's main purpose is to reduce the beam size of colliding bunches to increase the luminosity. This is realized by squeezing the vertical beta function at the IP  $\beta_y^*$  to 0.27 mm. The beam cannot be squeezed along its entire length, but only for a small section, which needs to be larger than the effective bunch length. The effective bunch length is obtained by projecting the horizontal beam size of one bunch onto the other under their crossing angle. A schematic depiction can be seen in Figure 5.3.

## 5.2 LINAC and Injection Scheme

The LINAC [56] at KEK was first commissioned in 1982 as 2.5 GeV electron accelerator for the Photon Factory. A source for positrons was added for the TRISTAN experiment and in 1995 the LINAC was upgraded to accommodate the need for 8 GeV electrons by KEKB. Today, the LINAC serves three experiments with continuous top-up injections: SuperKEKB, Photon Factory and the Photon Factory Advanced Ring. To accommodate the different needs of all facilities, the LINAC supports pulse-to-pulse-modulations. A schematic depiction of the LINAC is displayed in Figure 5.4.

The LINAC has three sections, two of which are linear sections and a rounded J-arc section with a combined length of 600 m. The energy values attached to the sectors refer to nominal values for particles dedicated to be injected into SuperKEKB.

The acceleration chain begins in section A with the creation of free electrons using a laser driven photo cathode RF gun [56, 58]. For the electrons the RF gun aims to create electron bunches of up to 4 nC with a vertical emittance of 10 mm mrad. With an acceptable vertical emittance of 20 mm mrad at the point of injection into the HER, the LINAC can afford an additional emittance dilution of 10 mm mrad. In contrast, the positrons are created using a thermionic electron gun that creates electron bunches with a charge of up to 10 nC that are accelerated to 3.2 GeV and shot onto a 14 mm tungsten plate in sector 1 with a conversion efficiency of 10%. The emittance of the resulting positron bunches is larger than the acceptable vertical emittance of 15 mm mrad and is reduced using a 1.1 GeV damping ring located between sections 2 and 3. After re-injection into the LINAC the positrons are accelerated to the nominal energy of 4 GeV for injection into the

| Parameter                       | Electron | Positron | Unit          |
|---------------------------------|----------|----------|---------------|
| Normalized horizontal emittance | 40       | 100      | $\mu\text{m}$ |
| Normalized vertical emittance   | 20       | 15       | mm mrad       |
| Energy spread                   | 0.16     | 0.07     | %             |
| Bunch charge                    | 4        | 4        | nC            |
| No. of bunches/pulse            | 2        | 2        |               |
| Repetition rate                 | 50       | 50       | Hz            |

Table 5.2: Parameters of the electron and positron bunches at the moment of injection into the SuperKEKB main ring [56].

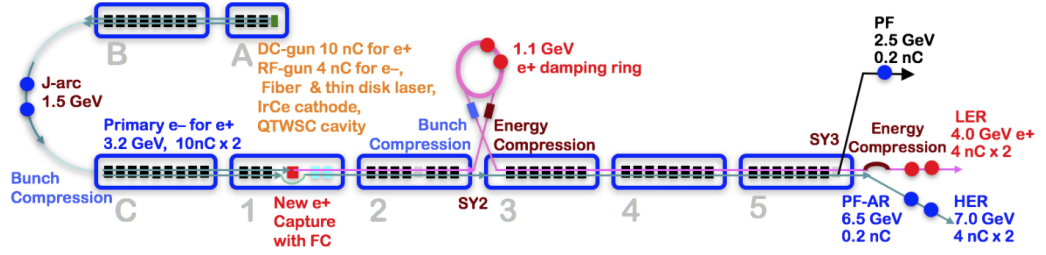


Figure 5.4: Schematic depiction of the linear accelerator and the damping ring illustrating the different stages and the respective energy and charge of the electron and positron bunches [56].

main ring in sections 3 to 5. A list of bunch parameters at the time of injection is displayed in Table 5.2.

The injection system uses the betatron injection scheme. There, the stored bunch receives a fast closed orbit bump by two kicker magnets and is brought close to the septum magnet. The septum magnet minimizes interference between the injected and stored bunch while guiding the injected beam towards the desired position. However, the septum magnet also demands a minimum spatial separation of the stored and injected bunch to physically accommodate the septum magnet. The injected bunch is purposely mismatched and placed next to the bumped stored beam, leading to betatron oscillations. Over the course of a few turns, the betatron amplitude dampens due to synchrotron radiation and the bunch-by-bunch feedback system [59] and the injected bunch merges with the stored bunch.

Belle II is a general purpose particle detector with a near  $4\pi$  coverage [60]. Its subsystems are arranged cylindrically around the IP and due to the asymmetric energies of the beam the detector is elongated in the direction of the boost. A schematic view of Belle II and its coordinate system are displayed in Figure 6.1. This chapter is based on [60].

## 6.1 Vertex Detector

The purpose of the vertex detector is to reconstruct the vertex position, which is the space point where a particle decayed. The vertex is reconstructed by finding the space point where the tracks, i.e. the reconstructed flight paths, of the decay products intersect. For vertices of B-mesons, that decay before reaching the innermost detector layer, it is necessary to extrapolate the tracks inwards to reconstruct the vertex.

The 6-layer vertex detector (VXD) consists of the Pixel Detector (PXD), contributing the two innermost layers, and the Silicon Vertex Detector (SVD), contributing the four outer layers. The PXD is a silicon pixel detector based on DEpleted p-channel Field Effect Transistors (DEPFET). The layers are at radii of 14 mm and 22 mm and have a windmill structure. Specifications for the pixel pitches on the respective PXD modules are displayed in Table 6.1. A schematic of the structure of the PXD is displayed in Figure 6.2.

The layers have eight and twelve ladders, respectively, with each ladder consisting of two modules. Each module is equipped with application-specific integrated circuits (ASICs) that facilitate the pixel readout and are mostly located at the outward-pointing ends of the modules, placing them outside of the detector acceptance, allowing for active cooling of these ASIC components while the 75  $\mu\text{m}$  thick active sensor area is air-cooled. The PXD can achieve a single track vertex

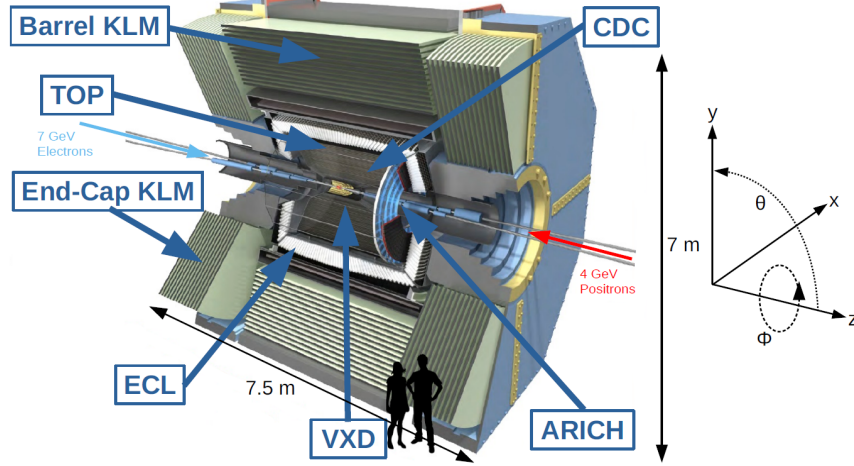


Figure 6.1: The Belle II detector and its coordinate system. The direction of the beams and the sub detectors are indicated. Modified from [61].

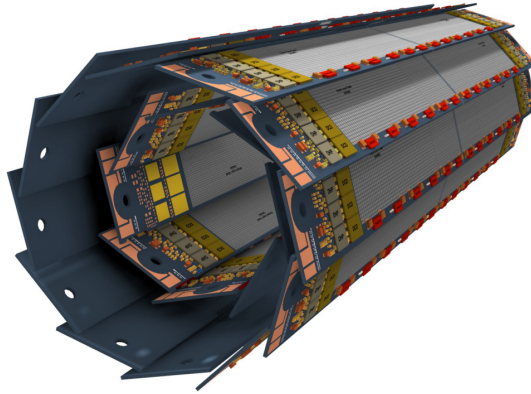


Figure 6.2: CAD drawing of the current Belle II PXD with all 20 ladders installed [62].



| Layer | Radius<br>(mm) | Ladders | Sensors | Pixels/Sensor<br>$u \times v$ | Pitch<br>$u \times v$ ( $\mu\text{m} \times \mu\text{m}$ ) |
|-------|----------------|---------|---------|-------------------------------|------------------------------------------------------------|
| 1     | 14             | 8       | 16      | $250 \times 768$              | $50 \times (55 \text{ to } 60)$                            |
| 2     | 22             | 12      | 24      | $250 \times 768$              | $50 \times (70 \text{ to } 85)$                            |
| Sum   |                | 20      | 40      | 7680000                       |                                                            |

Table 6.1: Specifications of the Belle II PXD.

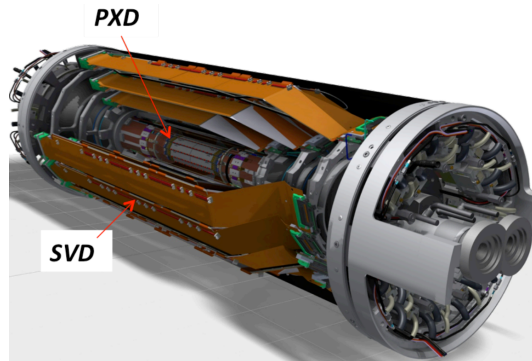


Figure 6.3: CAD drawing of the current Belle II VXD. [62]

resolution better than  $15 \mu\text{m}$ . The slow  $20 \mu\text{s}$  rolling shutter read-out, however, leads to large occupancies that limit track-building capabilities and its performance in high-background scenarios.

The SVD comprises the four outer layers of the VXD and is made from double-sided strip detectors (DSSDs). The SVD excels at time resolution with  $3 \text{ ns}$  allowing for a strong fake rate reduction. At occupancies above  $6 \%$ , meaning that  $6 \%$  of all pixels are registering a hit, the track finding rapidly degrades giving an effective upper limit of the acceptable occupancy of the inner SVD layer. The CAD drawing of the VXD is displayed in Figure 6.3.

## 6.2 Central Drift Chamber

The Central Drift Chamber (CDC) envelops the VXD and provides track reconstruction capabilities and precision momentum measurements of charged particles with the solenoidal magnetic field using

$$p = qBR. \quad (6.1)$$

A typical charged particle leaves  $\approx 40$  hits in the CDC volume with a spatial resolution of  $100 \mu\text{m}$  and a time resolution of  $1 \text{ ns}$ . The resolution of the transverse momentum using the CDC and the SVD is

$$\frac{\sigma_{p_t}}{p_t} = \sqrt{(0.1 \% p_t)^2 + (0.3 \% / \beta)^2}. \quad (6.2)$$

The magnetic field is produced by a superconducting solenoid and has a strength of 1.5 T and envelops all detectors except the  $K_L$  and muon detector, which is discussed below. Further, The CDC provides track information to the first level trigger within the 5  $\mu$ s trigger latency of Belle II. By determining the path length and the deposited energy within the CDC volume, a measurement of the energy loss  $dE/dx$  can provide a particle identification via the Bethe-Bloch formula, discussed in more detail in chapter 4. This is especially important for particles that decay before reaching the subsequent Cherenkov detectors or have momenta that are too small to escape the CDC volume given the strong magnetic field.

The CDC features 14336 sense wires arranged in 56 layers grouped in 9 super-layers that determine the wire orientation. The wires have two orientations and are either parallel to the  $z$ -axis (axial wires) or tilted forwards/backwards with respect to the  $z$ -axis (stereo wires) to avoid position and direction ambiguities in the measurements.

### 6.3 Cherenkov Detectors

The Time-Of-Propagation counter (TOP) and Aerogel Ring imaging Cherenkov Counter (ARICH) provide measurements of the Cherenkov light. Cherenkov light is emitted by massive charged particles traversing a medium, with a refractive index  $n$ , faster than the speed of light  $v_c(n) = c/n$  in that medium. The charged particle temporarily polarizes the atoms of the (dielectric) medium. Since the particle is faster than the propagation of the electric field, the “cloud” of polarized atoms trails behind the particle and is aligned along the flight path, so that they do not cancel each other out. Short electromagnetic pulses are created by the creation and relaxation of the dipoles. The elementary waves interfere constructively in the case of particle velocities greater than  $v_c(n)$ , creating the Cherenkov radiation in a cone-like shape. The emission angle of the Cherenkov light is related to the velocity  $\beta$  of the incidence particle:

$$\cos \theta_C = \frac{1}{\beta n}. \quad (6.3)$$

Combined with a momentum measurement, this provides a measurement of the rest mass of the particle:

$$m = \frac{p}{\beta \gamma c}. \quad (6.4)$$

This allows, for example, to distinguish between pions and kaons by their difference in rest masses.

The TOP counter is positioned in the barrel region around the CDC and consists of 16 massive synthetic fused silica quartz bars. The photons are internally reflected until they reach the expansion prism with the attached micro channel plate photo multiplier tubes (MCP-PMTs).

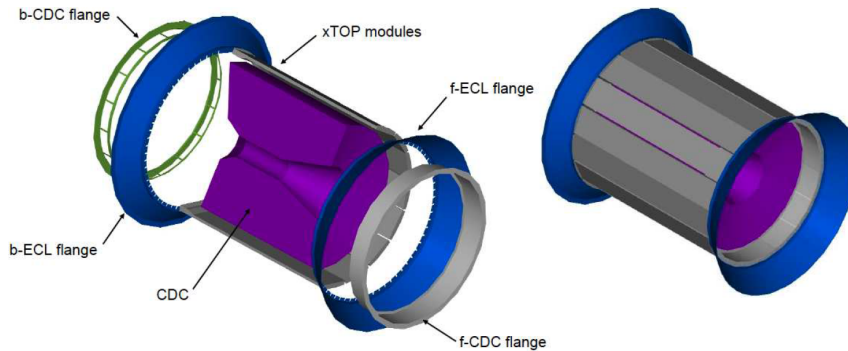


Figure 6.4: THE ARICH and TOP sub detectors in an exploded view.

The ARICH provides particle identification measurements in the forward end-cap region. It has a 2 cm thick aerogel radiator in order to produce enough Cherenkov photons, typically in the order of  $\mathcal{O}(10)$ , to make a reliable measurement. The aerogel radiator consists of two layers with different refractive indices. The refractive indices of the aerogel layers are chosen such that the two cones with different Cherenkov angles exactly overlap at the end of the expansion volume. The photon detectors consist of a vacuum tube that encloses an avalanche photo-diode. The ARICH and TOP detectors are displayed in Figure 6.4

## 6.4 Electromagnetic Calorimeter

The electromagnetic calorimeter (ECL) provides crucial measurements of non-charged particles, that are predominantly neutral pions and photons. At sufficient energies, photons and electrons create electromagnetic showers when traversing the scintillator material of the ECL. After enough shower generations, the energy of all partaking particles is small enough and their energy is absorbed by the atoms of a scintillation material, exciting the atoms (see Section 4.2). Upon de-excitation the atom emits a photon, called scintillation light. The produced light is then measured and the number of photons is proportional to the initial energy of the particle.

The ECL covers the end-cap and barrel regions, using over 8500 CsI(Tl) crystals with a depth of 30 cm, equaling  $\approx 16 X_0$ . The scintillation light is detected by photodiodes that are glued to the back of the crystals. The ECL measurements also contribute to the luminosity measurements,  $K_L^0$  detection, electron identification and to the level-1 trigger. Also, the injection background veto timing is tuned using the ECL signal. Due to the large injection background, the Belle II data taking is paused during the passage of a noisy bunch that was recently topped-up [63]. The geometry of the ECL is displayed in Figure 6.5.

## 6.5 $K_L$ and Muon Detector

The  $K_L$  and muon detector (KLM) is the outermost detector and envelops the other sub detectors in the barrel and end-cap regions. It is outside the magnetic field of the solenoid as the iron plates serve as a return yoke for the magnetic field. The KLM provides additional PID measurements for  $K_L$  mesons for uncharged tracks or muons and charged pions for charged tracks.

The iron plates alternate with active detection layers of resistive plate chambers (scintillators with photomultipliers) in the outer barrel region (inner barrel region and end-cap region). The dead-time of the resistive plate chambers, which were inherited from Belle, is large and the inner barrel layers are unable to cope with the larger backgrounds due to increasing luminosity. The end-cap region suffers from a high rate of back-scattered neutrons. The geometry of the KLM is displayed in Figure 6.5.

## 6.6 Trigger System and Data Acquisition

The Belle II trigger system uses a level-1 hardware trigger and a high level trigger (HLT) that is software-based. The trigger system has a high efficiency to select hadronic events from  $\Upsilon(4S) \rightarrow B\bar{B}$  and from events like  $e^+e^- \rightarrow q\bar{q}$  that involve light quarks (in this context light quarks refer to up-, down-, strange- and charm-quarks) [60]. The trigger chain starts with the level-1 trigger. The level-1 trigger has a 30 kHz Poisson distributed trigger rate with a maximum latency of 5  $\mu$ s, after which a trigger decision must be made. The length of the latency is defined by the speed of the level-1 trigger Field-Programmable Gate Arrays (FPGAs), the complexity of the algorithms to reach a trigger decision and the amount of background in the events as the events become larger and more complicated. The maximum possible latency is defined by the ring buffers of the detector front-end electronics that need to keep the data until a trigger decision is made. If the trigger latency is too long, then data in the events will be lost as it will be overwritten by the following data. In Belle II all detectors are designed to buffer data up to 5  $\mu$ s, and if an event is selected, send the related data to the HLT. The PXD is an exception to this rule, as it does not take part in any trigger decision and data needs to be buffered until the final trigger decision by the HLT is made. The data is temporarily saved on the Online Selection Nodes (ONSEN) system [64].

The level-1 trigger uses data provided by the CDC, ECL, TOP, ARICH and KLM. The CDC provides track information in the form of track count, 2D and 3D tracks. This information is sensitive to events with a high multiplicity, i.e. events with many tracks, which are mostly hadronic events. The 3D tracks, that additionally reconstruct the position along the  $z$ -axis, can effectively reject background not originating from the IP. The ECL provides cluster counts as well as timings and sums over the deposited energy. The ECL information is sensitive to events with charged and neutral particles with high and low energy clusters. Additionally, Bhabha and

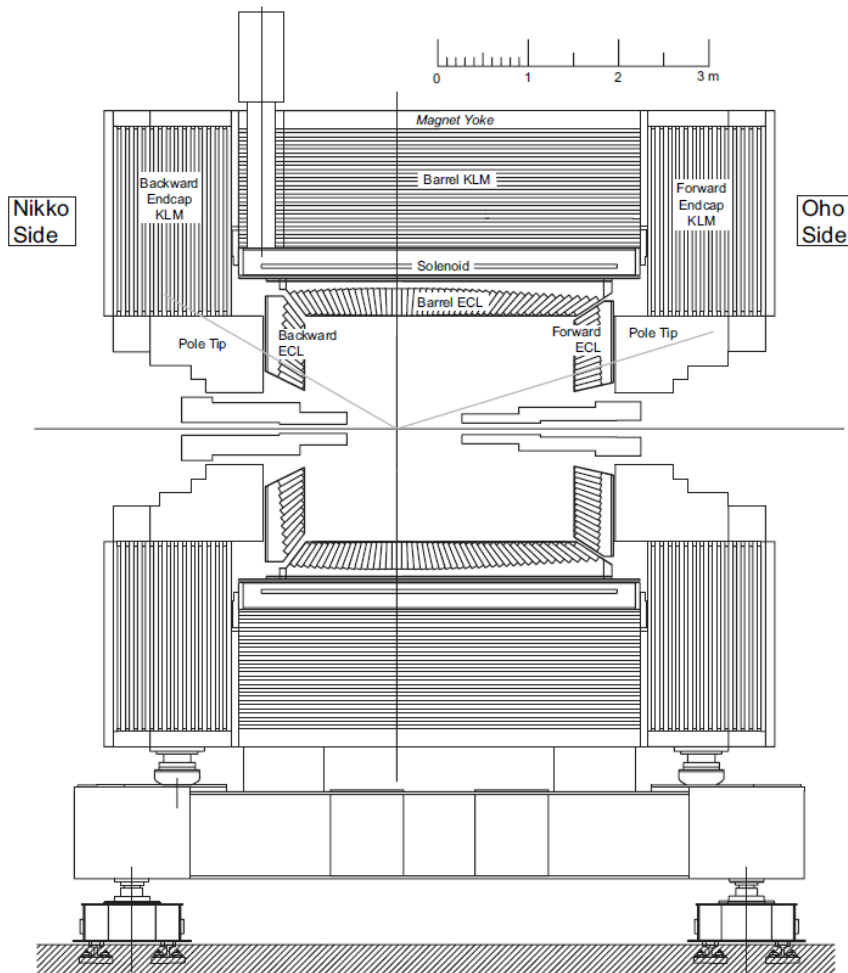


Figure 6.5: Cut-view of the ECL and KLM sub systems as well as components of the magnet system.

two photon events can be identified due to the characteristic signature they leave in the ECL. The KLM similarly provides cluster information that may be matched with the CDC tracks. The TOP and ARICH detectors provide PID information from the barrel and end-cap regions. The information is bundled and forwarded to the global reconstruction logic that performs a low-level reconstruction of the event. The output of this reconstruction is used by the global decision logic to make the level-1 trigger decision. If the result is positive, the trigger signal is sent back to trigger the readout of the buffered data of the detector front-ends, marking the first step of the Belle II DAQ. Data that is outside a triggered time windows will be overwritten and permanently lost.

The data is collected by the common pipeline platform for electronics readout (COPPER) via the Belle2Link [60, 65]. The Belle2Link is the data transmission system via optical fibers that connects the front-end electronics and the backend systems. The detector data is sent to the event builder that collects and merges the data into one block. This block of data is not reduced as for the level-1 trigger but contains every information read out from COPPER. The block of data is input to the HLT that performs a full event reconstruction with the Belle II Analysis Software Framework (BASF2) [66]. Upon the full event reconstruction, a final trigger decision is made, leading to the event to be saved or discarded. As opposed to the level-1 trigger, the HLT does not utilize FPGAs but is a CPU farm. The HLT trigger decision can take seconds and has no fixed latency. If the event is selected, the corresponding data in the reconstructed ROIs from the PXD is pulled and added to the event. The HLT reduces the trigger rate from 30 kHz to 10 kHz. The events are processed again offline by the same full event reconstruction but to a higher precision and including the PXD data.

## 6.7 Radiation Monitors

Diamond-based radiation monitors, or diamond sensors, were installed in 2018 and are placed around the IP in the VXD volume. A schematic of the positions of the diamond sensors are displayed in Figure 6.6. They monitor the integrated dose Belle II is subjected to and can trigger a beam abort signal in case there is a rapid increase in radiation to protect detector components. The integrated dose is the energy deposited per sensitive mass by radiation into the diamond sensor. Later, the system was expanded by an additional 8 sensors that were placed around the final focusing quadrupole magnet system (QCS) to protect against potential QCS quenches. If the integrated dose over a time of 10  $\mu$ s (1 ms) surpasses 8 mrad (80 mrad), the sensor triggers a very fast (fast) beam abort. A beam abort takes about 20 to 30  $\mu$ s which is equivalent to 2-3 revolutions around SuperKEKB.

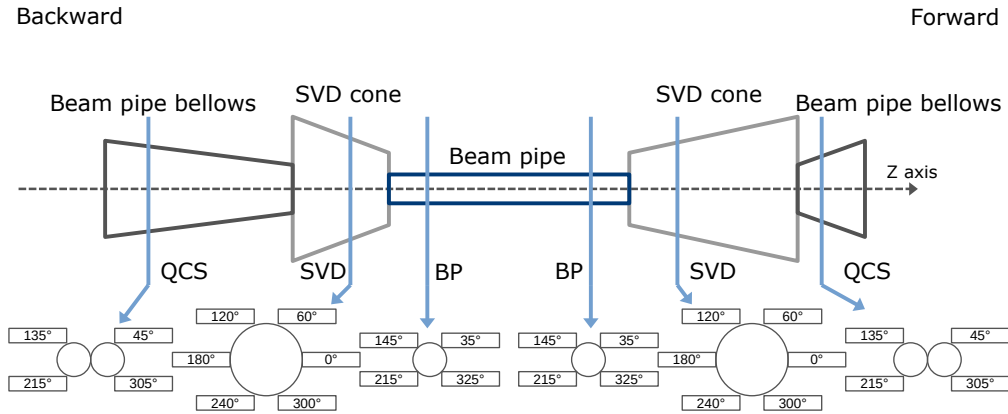


Figure 6.6: Positions of the diamond radiation monitors at the beam pipe bellows close to the QCS final focusing magnets of SuperKEKB, on the SVD support cone and the beam pipe. The angle indicates the position of the diamond in terms of the angle  $\phi$  of the Belle II coordinate system, which is indicated in Figure 6.1. Modified from [67].

## 6.8 Slow Control System

Belle II, similar to SuperKEKB and the LINAC, uses the experimental physics and industrial control system (EPICS) [68, 69] to manage the slow control. It supports a large number of Input/Output controllers (IOCs) that span a distributed system. The function of each IOC can be divided into two roles, server and client. Servers read and write from and into the network, for example, by publishing parameters reported from a physical hardware component into the network. Clients can display and store the data, for example, by providing a dashboard to visualize the status of a system.

Process variables (PVs) are the standard high-level objects that are used for communication between systems. Each PV has a unique name in the network and can store additional metadata. The PVs are distributed using User Datagram Protocol (UDP) and thus a client can fetch any amount of data without making an explicit connection with a handshake to the publishing server, although data integrity is not guaranteed and data is not automatically recorded. PV data will only be saved by the *archiver appliance* [70], which is a client system that fetches the values of a configured list of PVs and writes their values to disk.

The data published in EPICS includes:

- Environmental sensor data.
- Diagnostic data from Belle II sub detector systems.
- Variables reporting on the status of systems.

- Interactive variables used by operators to steer Belle II and related systems.

The leakage current of the CDC is one of the diagnostic variables that are available in EPICS. It is monitored to ensure safe operation, as large leakage currents can damage the CDC and recording this data allows to identify trends and mitigate problems in advance. Since the leakage current is proportional to the hit rate of the CDC, it is also used in the background monitoring tool BGNet, presented in Chapter 9. Similarly, the dose rates of the diamond radiation monitors are available in EPICS and used in BGNet. This motivates the value of recording such data and making it available beyond the initial purpose of monitoring the machine safety.

Environmental sensor data includes, for example, temperature and humidity sensors. These sensors typically report their measurements at a rate of 1 Hz.

In EPICS, thresholds on the numerical values of PVs can be defined to trigger warning signals that can be connected to appropriate actions, e.g. shutting systems off in case of unsafe temperatures. These warnings can be considered EPICS variables themselves, reporting on the status of Belle II. Another type of status variables is, for example, the PV that indicates if Belle II is powered on.

Interactive variables are set by operators. Operators can use them to set the jaw positions of collimators or control the power supplies, turning systems on or off.



## CHAPTER 7

---

### VTX Upgrade Proposal

---

The luminosity projections for SuperKEKB suggest that only  $2 \times 10^{35} \text{ cm}^{-2}\text{s}^{-1}$  can be reached with the current interaction region (IR) design, a factor 3 below the design instantaneous luminosity. To reach the full data set of  $50 \text{ ab}^{-1}$  during the 2040s, achieving the design luminosity will be necessary. An IR redesign, and an accompanying change of the QCS, is planned with an installation foreseen during Long Shutdown 2 (LS2), around 2032.

Since an IR redesign will change the spatial envelope of the inner detectors that are structurally dependent on the QCS, a redesign of the vertex detector is required. This also opens the opportunity to construct a new vertex detector with an upgraded technology to benefit the physics performance and increased resilience against beam background. The VTX upgrade proposes to replace the VXD with a 5 or 6 layer, fully pixelated detector [71]. While physics performance simulations favor a 6 layer geometry, the geometry will ultimately be decided by the mechanical feasibility. Preliminary studies of the VTX mechanics suggest that a 6 layer geometry will not fit into the available volume and a 5 layer geometry is preferred. All layers will use the DMAPS chip OBELIX (Optimized BELle II pIXel sensor). It is based on the modified 180 nm CMOS process by TowerSemi.

The VTX is expected to improve on the vertex position resolution, as well as the track reconstruction of low momentum particles that do not reach into the CDC volume. Other performance figures are expected to have minor improvements or be similar to the current VXD system.

The Belle II tracking [72] relies heavily on the CDC. In Belle II, track reconstruction is either directed inwards or outwards. The CDC standalone track finding can seed a track in the CDC and extend the track inwards. Similarly, a track can be found and reconstructed by the SVD standalone tracking and extended outwards. The tracks that are found by both tracking algorithms are then merged and PXD

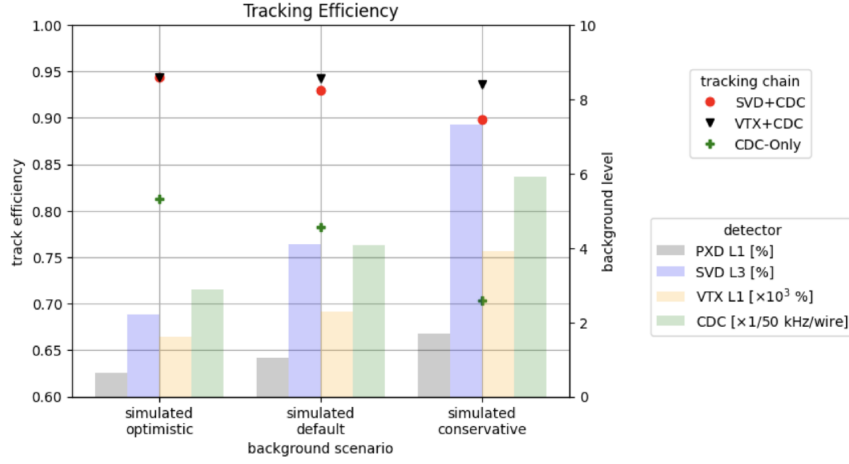


Figure 7.1: Comparison of the occupancies of the PXD, SVD, CDC and VTX and the track efficiency for different tracking setups at different background scenarios [71].

hits are added. The CDC track finding performance is especially affected by large hit rates due to high backgrounds. An upgrade of the CDC read-out electronics is planned, recovering some of the lost performance. Also, the VTX is expected to have a better standalone track finding over the SVD, recovering overall tracking performance by improving the outward directed tracking, which is less affected by higher CDC hit rates.

To simulate the performance during future operation, the background conditions need to be extrapolated to reflect the expected increase in backgrounds over the current conditions. Three background scenarios were considered by scaling the single beam backgrounds: Optimistic scenario 1 (SC1) scaling by factor 2, intermediate/default scenario 2 (SC2) scaling by factor 5, conservative scenario 3 (SC3) scaling by factor 10. The factors were chosen to be large since the errors on the background extrapolation are large. The comparison of the track finding efficiency of CDC standalone, CDC+SVD and CDC+VTX for different background extrapolations is shown in Figure 7.1.

The improved physics performance for events with soft pions was evaluated using  $B^0 \rightarrow D^{*-}(\bar{D}^0\pi_{\text{soft}}^-)\ell^+\nu$  with either  $\bar{D}^0 \rightarrow K^+\pi^-$  or  $\bar{D}^0 \rightarrow K^+\pi^-\pi^+\pi^-$ , although both final states show the same performance. Pions are considered soft when their transverse momentum  $p_T$  is small. The efficiency and purity for the  $\bar{D}^0 \rightarrow K^+\pi^-$  final state are shown in Table 7.1.

Further the reconstruction efficiency as a function of the transverse momentum and the polar angle  $\theta$  is shown in Figure 7.2.

Above transverse momenta of  $p_T > 0.3 \text{ GeV}$  the performance increases vanish as the CDC performance dominates the performance figures.

|                                | Belle II SC2 | VTX SC1 | VTX SC2 | VTX SC3 |
|--------------------------------|--------------|---------|---------|---------|
| Generated events               | 32533        | 32559   | 32559   | 30255   |
| Correctly reconstructed signal | 10059        | 16913   | 16848   | 15583   |
| Self-crossfeed                 | 28495        | 51375   | 51826   | 47527   |
| Efficiency                     | 30.9 %       | 51.9 %  | 51.7 %  | 51.5 %  |
| Purity                         | 26.1 %       | 24.8 %  | 24.5 %  | 24.7 %  |

Table 7.1: Reconstruction efficiency and purity for the  $\bar{D}^0 \rightarrow K^+\pi^-$  final state. The current Belle II detector with background SC2 is compared with the upgraded VTX and all background scenarios [71].

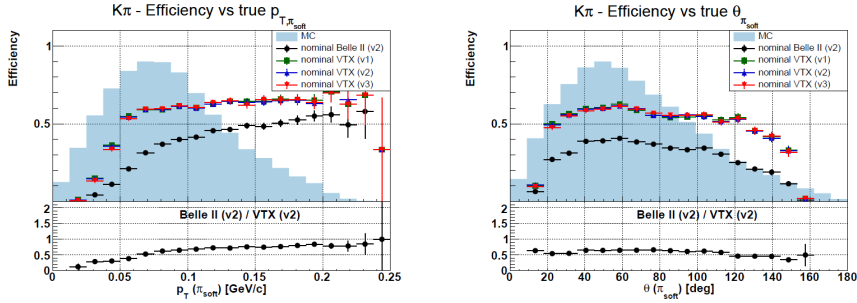


Figure 7.2: Reconstruction efficiency of  $\bar{D}^0 \rightarrow K^+\pi^-$  final state as function of the transverse momentum of the soft pion  $p_T(\pi_{\text{soft}}^-)$  and the cosine of the polar angle  $\theta$ . The current Belle II detector with background SC2 is compared with the upgraded VTX and all background scenarios. The bottom plot shows the the ratio of the efficiency for the Belle II geometry and the VTX geometry for background SC2 [71].

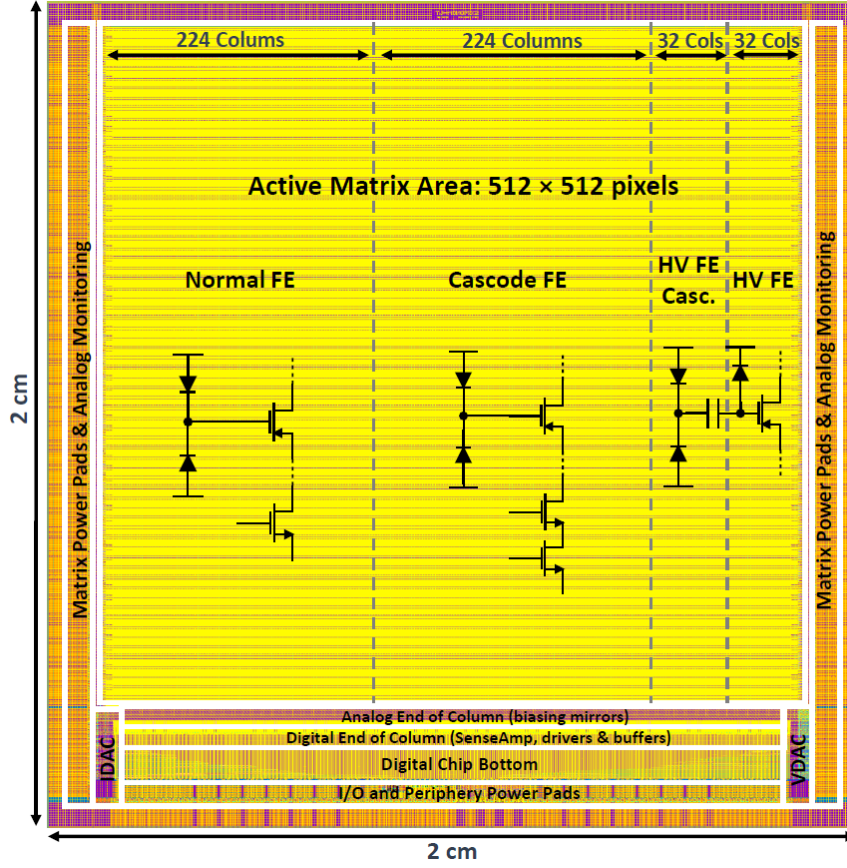


Figure 7.3: Layout of the TJ-Monopix2. The  $512 \times 512$  pixels active matrix is divided into four areas with different FEs [54].

## 7.1 The TJ-Monopix2 DMAPS Chip

The TJ-Monopix2 is a DMAPS CMOS chip and the successor of the TJ-Monopix chip [52]. They stem from discontinued developments for the ATLAS outer tracker and were chosen as basis to develop the OBELIX chip as the requirements were similar. The initial requirements and the OBELIX specifications are described in more detail in the next section.

The pixels are quadratic with a pitch of  $33.04 \mu\text{m}$  arranged in a  $512 \times 512$  pixel matrix with a size of  $1.7 \text{ cm} \times 1.7 \text{ cm}$  for the active area and  $2 \text{ cm} \times 2 \text{ cm}$  for the entire chip.

The TJ-Monopix2 features active pixels with four variations concerning the coupling to the front-end amplifier circuit and the amplifier circuit itself. The layout of the four different FEs is displayed in Figure 7.3. The cross section of the TJ-Monopix2 pixel is shown in Figure 4.10 and the full front-end circuit is shown in Figure 4.12

The analog pixel is either DC-coupled to the amplifier circuit and called “normal” front end or AC-coupled and called “HV” (High Voltage) front end. The front-end amplifier circuit is varied by an addition of a transistor to allow for higher gain. The DC-coupled front end allows for biasing of the substrate (PSUB) and the deep p-well (PWELL) with negative voltages. The voltage of PWELL is limited to  $-6\text{ V}$  since the deep p-well hosts the electronics of the active pixel and the potential will affect the transistors adversely at higher voltages. The PSUB voltage has to match the PWELL voltage to avoid punch-through and damage to the pixel. This is especially critical during the power up of the sensor for unirradiated sensors. After irradiation, however, the punch-through effect is delayed and a potential of  $-20\text{ V}$  can be applied to PSUB. Possibly higher voltages could be applied, but tests showed that the performance at  $-20\text{ V}$  was already satisfactory. The AC-coupled front end is biased by applying a positive voltage to the charge collecting electrode (HV). The AC coupling to the front end allows for larger potentials, however the additional capacitor leads to a charge loss of around 30 % [54]. The maximum voltage that can be applied is around  $40\text{ V}$ . A chip was tested until a bias of  $62\text{ V}$  and the breakdown voltage was reached at  $63\text{ V}$ , breaking the chip.

A tuning of the threshold can be achieved by either manipulating the signal produced by the front-end amplifier or the discriminator. The global threshold can be tuned using registers, manipulating the front-end amplifier. In the case of TJ-Monopix2 this is achieved by changing the baseline output of the amplifier. This cannot be tuned on a per-pixel basis. The bias current of the discriminator  $I_{\text{DISC}}$  is obtained by [54]:

$$I_{\text{DISC}}[i] = I_{\text{DB}} + (i - 1)I_{\text{Tune}}, \quad (7.1)$$

where the currents  $I_{\text{DB}}$  and  $I_{\text{Tune}}$  are obtained as a product of their 8 bit register setting, an integer value between 0 and 255, and the current for the least significant bit, which depends on the Digital-to-Analog Converter (DAC) supplying the current. The value  $i$  is a three bit register setting and takes values between 0 and 7. It can be set individually for each pixel, allowing for a local tuning of the behavior of the discriminator. Inhomogeneities in the pixel response can be compensated. The threshold tunings applied during measurements in Chapter 8 include local tunings and thresholds can be assumed to be homogeneous over a sensor front end.

## 7.2 The OBELIX DMAPS Chip

The development of OBELIX started with the TJ-Monopix and TJ-Monopix2 chips that were developed for the ATLAS outer tracker as they had similar specifications. A summary of the initial requirements on the design of new vertex detector chip are shown in Table 7.2.

The OBELIX chip will inherit the matrix design of TJ-Monopix2, i.e. the analog design of the pixels and the column drain read-out architecture. Currently, it is

|                    | Requirements                                               | TJ-Monopix2                                            | OBELIX                                               |
|--------------------|------------------------------------------------------------|--------------------------------------------------------|------------------------------------------------------|
| Pixel pitch        | $< 40 \mu\text{m}$                                         | $33.04 \mu\text{m}$                                    | $33.04 \mu\text{m}$                                  |
| Spatial resolution | $< 12 \mu\text{m}$                                         | $< 9.5 \mu\text{m}$                                    | $< 9.5 \mu\text{m}$                                  |
| Sensor thickness   | $< 100 \mu\text{m}$                                        | -                                                      | thinning to $40 \mu\text{m}$                         |
| Hit rate burst     | $> 600 \text{ MHz}$                                        | $> 600 \text{ MHz}$                                    | $> 600 \text{ MHz}$                                  |
| Hit rate sustained | $> 120 \text{ MHz}$                                        | $\gg 100 \text{ MHz}$                                  | $> 120 \text{ MHz}$                                  |
| Trigger rate       | $30 \text{ kHz}$                                           | -                                                      | $30 \text{ kHz}$                                     |
| Trigger delay      | $> 10 \mu\text{m}$                                         | -                                                      | $> 10 \mu\text{m}$                                   |
| Integration time   | $< 100 \text{ ns}$                                         | -                                                      | $< 100 \text{ ns}$                                   |
| TID                | $100 \text{ Mrad}$                                         | -                                                      | $100 \text{ Mrad}$                                   |
| NIEL fluence       | $5 \times 10^{14} \text{ n}_{\text{eq}}/\text{cm}^2$       | $1.5 \times 10^{15} \text{ n}_{\text{eq}}/\text{cm}^2$ | $5 \times 10^{14} \text{ n}_{\text{eq}}/\text{cm}^2$ |
| SEU tolerance      | frequently<br>( $\text{min}^{-1}$ ) flash<br>configuration | -                                                      | 10 Hz flash config-<br>uration                       |
| Powering           | through voltage<br>regulators                              | -                                                      | LDOs $\geq 0.2 \text{ V}$<br>drop-out                |
| Output             | one at $< 200 \text{ MHz}$                                 | one at $160 \text{ MHz}$                               | one at $339 \text{ MHz}$                             |

Table 7.2: Initial chip requirements for a tracking detector upgrade and the TJ-Monopix2 specifications [71] compared to the final OBELIX specifications [50].

planned to have two front ends, the AC-coupled cascode front end and the DC-coupled cascode front end, in the first iteration of the design, called OBELIX-1 [71]. However, previously it was considered that the OBELIX-1 could feature only one front end, the DC-coupled cascode front end, potentially forgoing another production cycle that would have been OBELIX-2, if no issues in the implemented design changes are found in OBELIX-1. The decision for this potential front end was reliant on the test beam results presented in Chapter 8.

The matrix of OBELIX will be extended to  $896 \times 464$  pixels while the total chip size increases to  $30.2 \times 18.8 \text{ mm}^2$ . Low dropout (LDO) regulators were implemented for power distribution across the enlarged chip area.

The digital periphery needed to be redesigned to be compatible with the Belle II trigger system. A trigger unit was added featuring a dual-stage memory buffer that can be read-out once a trigger signal is produced. At the required  $120 \text{ MHz}$  continuous hit rate, a trigger latency of  $10 \mu\text{s}$  at a  $30 \text{ kHz}$  trigger rate can be sustained. The trigger latency is doubled compared to the current Belle II trigger latency [60], allowing for more elaborate trigger decisions. The VTX can also provide data for the level-1 trigger decision using the developed Track Trigger Transmission unit (TTT) [50, 73]. For the purposes of the TTT, multiple pixels are grouped into “macropixels” and only binary hit data is considered. The OBELIX chips send this hit data as an additional output to FPGAs that compare the hit data to pattern tables. This feature will only be available for the outer VTX (oVTX) as the performance degrades

at high occupancies with the required computation time and the power dissipation in the digital periphery increasing. The fast time stamping using only the binary hit information of “macropixels” can also be used as a standalone feature allowing for precision time stamping of tracks with a time resolution  $< 3$  ns. This feature will also only be available for the oVTX due to decreased performance at large hit rates and increased power dissipation.

### 7.3 Mechanics

Both the 5 and 6 layer VTX geometry will have a detector acceptance of  $17^\circ < \theta < 150^\circ$  and the layers will be positioned between radial distances of 14 mm and 140 mm similar to the current VXD. A considered upgrade of the QCS magnets could move the tips of the magnets closer to the IP. This would reduce the geometric envelope of the VTX, reducing its angular acceptance. Since there are no new QCS geometries yet, current simulations assume the present geometry. Preliminary studies suggest, that the new QCS geometry only permits a 5 layer geometry, however, a final decision on the VTX geometry has not been reached at the time of writing.

The inner VTX (iVTX) consists of the two innermost layers at radii of 14 mm and 22 mm. It is designed to reduce material budget by having self-supporting, all silicon ladders and a minimal cooling solution. The ladders are arranged in a windmill structure with an overlap of about 10 pixels. Each ladder has four OBELIX chips that are diced out together from the silicon wafer allowing for a monolithic and rigid all-silicon structure. The chips are then electrically connected using a redistribution layer (RDL) that is added in a post-production process. The RDL adds metal and polymer layers using a thin film process with photo lithography, providing connections for both power and data between the chips of a ladder. The area underneath the active sensor matrix is then thinned to  $40\ \mu\text{m}$  while the edges of the ladder remain unthinned at a thickness of  $400\ \mu\text{m}$  for structural support. The material budget for an iVTX layer is then about  $0.2\% X_0$  in case the ladders are solely air-cooled. However, since the release of the conceptual design report (CDR) [71] it was understood that only air-cooling would not suffice with the targeted temperature and the projected power dissipation. The power dissipation is currently expected to be  $200\ \text{mW}/\text{cm}^2$  up to  $300\ \text{mW}/\text{cm}^2$  for hit rates up to  $120\ \text{MHz}/\text{cm}^2$  [50]. The test beam results shown in this thesis were part of this decision. The addition of a thin cooling pipe increases the material budget to  $0.3\% X_0$ .

The oVTX comprises the outer 3 (4) layers for the 5 layer (6 layer) design. The properties of the layers for both scenarios are displayed in Table 7.3 and 7.4. Drawings of the mechanical structure are displayed in Figure 7.4.

The oVTX mechanical structure is closely related to the developments for the ALICE ITS2 [75]. The sensors are cut from the wafer individually and the sensor backside is glued onto a carbon fiber cold plate. The cold plate is thermally coupled to a polyamide pipe for the liquid coolant. For additional stiffness, a 3D carbon fiber structure in a V or  $\Omega$  shape is brought onto the cold plate. The OBELIX chips of a

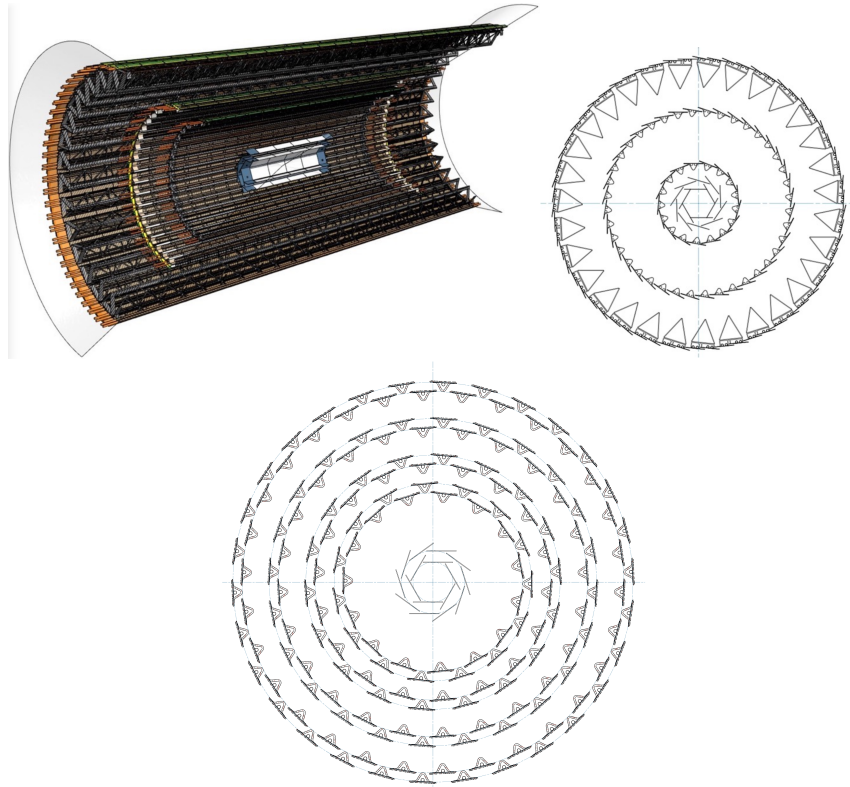


Figure 7.4: Top left: 3D cut-view of the 5 layer VTX [71]. Top right: Cross section of the 5 layer VTX [71]. Bottom: Cross section of the 6 layer VTX.



|                  | L1   | L2   | L3               | L4               | L5            | Unit    |
|------------------|------|------|------------------|------------------|---------------|---------|
| Radius           | 14.1 | 22.1 | 39.1             | 89.5             | 140           | mm      |
| # Ladders        | 6    | 10   | 17               | 40               | 31            |         |
| # Sensors/ladder | 4    | 4    | 7                | 16               | $2 \times 24$ |         |
| Material budget  | 0.2* | 0.2* | 0.3 <sup>†</sup> | 0.5 <sup>†</sup> | 0.8           | $\%X_0$ |

Table 7.3: VTX layer parameters for the 5-layer design [71]. \*: the addition of a necessary cooling component will increase the material budget to  $0.3\%X_0$  [74] over the previous value in the CDR [71]. <sup>†</sup>: the material budget of the oVTX layers was increased to  $0.8\%X_0$  [74] over the previous value in the CDR [71].

|                  | L3               | L4               | L5               | L6               | Unit    |
|------------------|------------------|------------------|------------------|------------------|---------|
| Radius           | 62.5/69.0        | 82.5/89.0        | 108/114.5        | 133.5/140        | mm      |
| # Ladders        | 30               | 36               | 48               | 60               |         |
| # Sensors/ladder | 12               | 16               | 20               | 24               |         |
| Material budget  | 0.6 <sup>†</sup> | 0.6 <sup>†</sup> | 0.6 <sup>†</sup> | 0.6 <sup>†</sup> | $\%X_0$ |

Table 7.4: VTX outer layer parameters in the alternative 6-layer design [50, 74]. The two radii for each layer reflect the radii of the staggered sub-layers. <sup>†</sup>: the material budget of the oVTX layers was increased to  $0.8\%X_0$  [74] over the previous value in the CDR [71].

ladder are electrically interconnected by wire bonding to a flex PCB that is glued on the top side of the sensors. Due to possible space constraints the PCBs of layer 3 will only be connected in the backward side (towards negative  $z$ -axis) of the detector. All other layers are connected from both sides.



---

## Measurements of the TJ-Monopix2 Chip

---

A total of four test beams were conducted to characterize the TJ-Monopix2 chip, debug its operation and provide feedback for the OBELIX design. The test beam campaigns conducted by the upgrade collaboration of Belle II are summarized in Table 8.1. The author took part in all listed test beams, taking on more responsibilities with each test beam. At the latest test beam in April 2025, the author was responsible for the computing hardware, networking and the data analysis during and after the test beam.

The following sections focus on the test beam results obtained at DESY in March 2025, focusing on systematic studies of irradiated TJ-Monopix2 chips at different temperatures. The test beam results were a necessary input to the design of the VTX by quantifying the effect of different temperatures on the performance, establishing operational temperature ranges. Further, the systematic comparison of the two remaining considered front ends contributes to the decision process for which front end is considered to be used for OBELIX. This test beam was also used as an opportunity to prove that the TJ-Monopix2 system could be used in a synchronized read-out mode (AIDA mode), which will be necessary in the Belle II environment. The test beam results of initial tests of unirradiated samples can be found in references [51, 76]. Parts of the following results have been previously published in conference proceedings by the author [77].

### 8.1 Test Beam Setup

The test beam in 2025 was conducted at the DESY II test beam facility [78]. The DESY II synchrotron circulates electron beams of up to 7 GeV. A fiber target that can be brought into the main electron beam produces bremsstrahlung photons. Using a secondary target, pair production is induced and the resulting electrons and

| Date       | Length                    | Location | Scientific program                                                                                                                                      |
|------------|---------------------------|----------|---------------------------------------------------------------------------------------------------------------------------------------------------------|
| July 2022  | 2 Weeks                   | DESY     | 4 unirradiated samples, understanding sensor properties/tuning and test DAQ system                                                                      |
| July 2023  | 1 Week + 1 Week parasitic | DESY     | 2 irradiated samples (1 p-irradiated, 1 n-irradiated), systematic tuning, low threshold operation, timing tests                                         |
| July 2024  | 1 Week                    | DESY     | 2 irradiated samples (1 p-irradiated, 1 X-ray irradiated), high temperature tests, compare front end performances                                       |
| April 2025 | 2 Weeks                   | DESY     | 4 irradiated samples (3 p-irradiated, 1 e-irradiated), temperature controlled environment, compare front ends at different temperatures, test AIDA mode |

Table 8.1: Summary of the TJ-Monopix2 test beam campaigns.

positrons pass a magnet that selects both the particle type and the particles energy. The electrons available from this double conversion have an energy of 1 GeV - 6 GeV. The energy can be selected by the user via a bending magnet with an adjustable magnet strength. The particle rate is dependent on the selected energy. Additionally, the particle rate can be reduced by changing the aperture of a collimator that is positioned after the magnet. An energy of 4.2 GeV was chosen for all measurements. Electrons at this energy are considered minimum ionizing particles (MIPS), referring to the Bethe-Bloch Equation 4.1. This energy minimizes multiple scattering of the electrons used for the measurements, while maintaining a sufficient particle rate. After the collimator the electron beam enters the user area, where experiments can be conducted. At the end of the beam nozzle the beam size can be manipulated by putting different lead cylinders at the end of the beam nozzle, that act as an additional collimator.

The DURANTA EUDET-type telescope [79], available at DESY, was used as a beam telescope. The beam telescope provides reference measurements to determine, for example, time and spatial resolution and efficiency. A small material budget of the telescope is crucial to reach good spatial resolutions. The DURANTA telescope consists of 6 planes of 50  $\mu\text{m}$  thick MIMOSA26 MAPS sensors [80] with a  $18.4 \times 18.4 \mu\text{m}^2$  pixel pitch. The sensors are read out using a 115  $\mu\text{s}$  rolling-shutter style readout. The long rolling shutter readout limits the time resolution that can be achieved by the DURANTA telescope. Therefore, an additional timing plane is installed to allow for fine time stamping. The timing plane uses a TelePix2 sensor

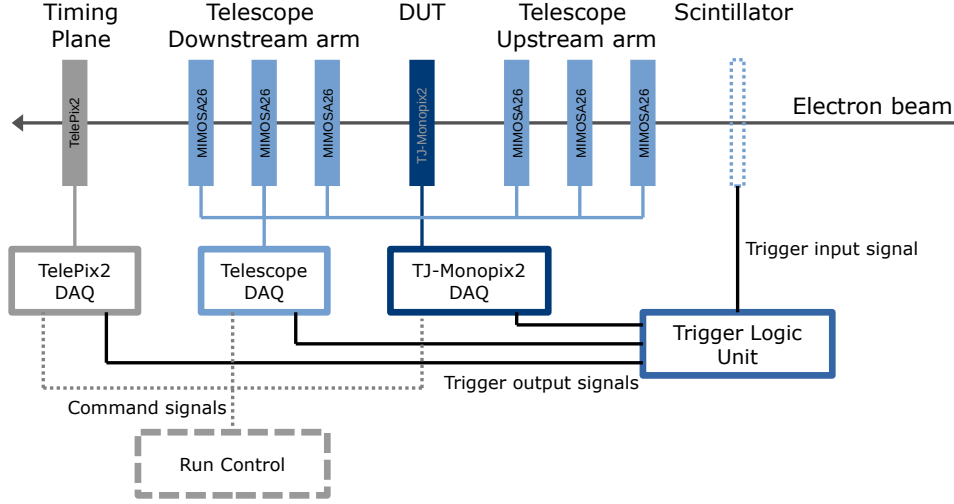


Figure 8.1: Schematic of the test beam setup with the DURANTA EUDET-type telescope, the TJ-Monopix2 DUT and a TelePix2 timing reference plane. The scintillator in front of the telescope produces a trigger signal once a particle passes, which is then distributed by the TLU to the respective DAQ systems. From the run control machine the setup is configured and controlled.

[81] that is placed downstream of the telescope as the last sensor plane. The TJ-Monopix2 chip, the Device Under Test (DUT), is mounted between the third and fourth telescope plane, so that the upstream and downstream arms of the telescope each consists of 3 planes. In front of the first sensor are two scintillators with PMTs that send a trigger signal to the Trigger Logic Unit (TLU) [82] once an electron passes them. The schematic of the setup is displayed in Figure 8.1.

During the test beam two synchronization modes were used, that are referred to as EUDET mode and AIDA mode [82].

**EUDET mode** Once the TLU receives the trigger input signal from the PMT upon a passage of a particle through the scintillator, the TLU distributes the trigger signal to all the respective Data Acquisition Systems (DAQ) triggering the readout. Once the trigger signal is received by the DAQs, they raise a busy signal for the duration of the data read-out and the distribution of the trigger number from the TLU, blocking the distribution of further trigger signals. This triggered read-out of all DAQs is associated with a common trigger number distributed by the TLU. After all DAQs are ready, they each lower their respective busy signal and the TLU returns to its original state, ready to accept another trigger input signal. The maximum trigger rate is thus limited by the slowest sensor read-out in the setup. Similarly, the time resolution that can be measured using the EUDET mode is limited to the time resolution of the slowest sensor, which is the MIMOSA26 telescope sensor with

| Serial number | Particle type | Fluence [ $n_{eq}/cm^2$ ] | Substrate      |
|---------------|---------------|---------------------------|----------------|
| W14R12        | -             | unirradiated              | Czochralski    |
| W8R4          | proton        | $1 \times 10^{14}$        | 30 $\mu m$ EPI |
| W2R17         | proton        | $2.5 \times 10^{14}$      | 30 $\mu m$ EPI |
| W8R6          | proton        | $5 \times 10^{14}$        | 30 $\mu m$ EPI |
| W2R5          | electron      | $5 \times 10^{14}$        | 30 $\mu m$ EPI |

Table 8.2: TJ-Monopix2 samples available for the 2025 test beam campaign.

its 115  $\mu s$  long rolling shutter readout. The data of all sensors are aligned by virtue of the common trigger number associated with the data.

**AIDA mode** The TLU distributes a common clock that all DAQs are synchronized to. Once a trigger input signal is received by the TLU, it raises the trigger line in sync with the TLU common clock. The trigger signal is associated with a timestamp by the TLU. After the trigger is sent, a DAQ can raise a busy signal for the duration of its readout. The TLU does not send any data or trigger number to the DAQs. The data of each sensor is associated with a timestamp based on the TLU common clock by each DAQ. The maximum trigger rate depends on the TLU's capability of issuing trigger signals, or the length of busy signals send by the DAQs, although these can be ignored. The data streams need to be aligned offline during reconstruction using the timestamps of the recorded triggers. The time resolution is dependent on the timestamp precision.

The EUDAQ2 [83] framework was used to manage the components of the test beam setup, triggering initialization, propagating settings and starting/stopping the data taking. The DURANTA telescope operates in EUDET mode and uses the EUDAQ2 data collector to obtain data. The TelePix2 and TJ-Monopix2 were using their respective DAQ systems and were operated in AIDA mode allowing for the timestamped track reconstruction. For TJ-Monopix2, this is handled by a BDAQ53 FPGA readout board [84] running the basil software [85].

To allow for temperature control of the DUT, a Peltier cell is thermally coupled to the TJ-Monopix2 chip/carrier PCB. Since there is no on-chip temperature probe, the carrier PCB negative-temperature-coefficient (NTC) resistor is used to gauge the chip temperature with  $T_{chip} = T_{NTC} + 7^\circ C$ . All temperatures that are referred to during the following sections of the test beam analysis refer to  $T_{NTC}$  unless explicitly stated otherwise. The temperature range used was  $10^\circ C - 50^\circ C$ . To prevent condensation on the electrical components, the chip was flushed with nitrogen.

The TJ-Monopix2 samples available for the 2025 test beam are summarized in Table 8.2.

### 8.1.1 Charge Calibration

Before the analysis began, preprocessing steps were performed. For this analysis, the only preprocessing step was applying the charge calibration to the raw data of the

TJ-Monopix2, if it is available. The charge calibration allows to convert the relative ToT units into an absolute unit of charge.

At the end of the amplifier circuit, the charge is digitized and given in ToT as the difference of 7-bit wide timestamps of the trailing edge and the leading edge (see Section 4.6). The sample clock for the timestamps has a frequency of 40 MHz leading to a time resolution of 25 ns. The charge-to-digital conversion is dependent on the state of the pixel front end, e.g. how the signal is shaped. Thus, the ToT values cannot be used to infer an absolute value of the collected charge and cannot be compared between two measurements with different front end configurations. To obtain the absolute collected charge in units of  $e^-$ , the ToT values need to be calibrated.

The calibration is done by utilizing the charge injection feature of the TJ-Monopix2 [54]. Two voltages  $V_{\text{low}}$  and  $V_{\text{high}}$  are propagated through the pixel matrix. The value of each voltage is set in DAC units. The voltage difference between the two propagated voltages is then applied to the charge collection electrode through an injection capacitance  $C_{\text{inj}}$ . The capacitance of  $C_{\text{inj}}$  is not immediately known but can be determined by calibration with an  $^{55}\text{Fe}$  source that has a characteristic X-ray peak at 5.90 keV. The conversion factors are  $9 e^-/\text{DAC}$  for the DC-coupled front end and  $18 e^-/\text{DAC}$  for the AC-coupled front end. The applied voltage is known, but the internal buffers of  $V_{\text{low}}$  and  $V_{\text{high}}$  are only linear up to 140 DAC units. Beyond the linear regime, quantitative measurements are not possible. This sets a limit on the injected charge to  $9 e^-/\text{DAC} \times 140 \text{ DAC} = 1260 e^-$ .

The charge calibration is then performed by injecting known charges into the charge collection electrode, recording the ToT response and fitting a calibration curve of the form

$$f_{\text{cal}}(x) = \left( \frac{a}{x} + \frac{1}{b} \right) (x - d). \quad (8.1)$$

The parameters  $a$ ,  $b$  and  $d$  are the fit parameters. The same calibration is applied to all pixels of a front end. This calibration needs to be performed for each set of settings of the amplifier circuit and under the same conditions as were present during the test beam measurement. The charge calibration curves were produced centrally after the test beam campaign at INFN Pisa by R. Mammone and were kindly provided to the author.

An example for such a calibration curve is displayed in Figure 8.2. At each injection step, the ToT response of all pixels is recorded. The ToT response is not the same for all pixels. As a result, each ToT is associated with a distribution of pixel responses at different injected charges that resulted in the respective ToT value. The number of pixel hits recorded at a given injected charge is indicated by the color of each bin. The mean charge of each distribution is identified and used as data point for the fit. The means are indicated by the red dots. The fit of the calibration curve on the pixel response to injected charges converges well in the range of available injection charges.

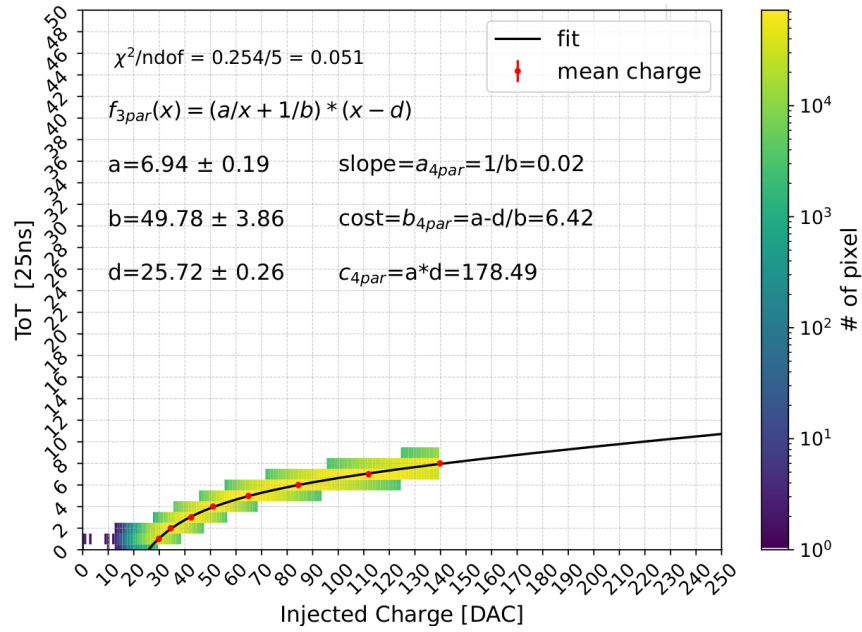


Figure 8.2: The plot shows the ToT value in correspondence to the injected charge in DAC units. A fit was performed using the Equation 8.1. The fit parameters are indicated in the top left. The calibration corresponds to the DC-coupled front end of TJ-Monopix2 sensor W8R6.



## 8.2 Analysis Setup

Two test beam analysis tools have been used to obtain the presented results. The results for the cluster properties and efficiencies were obtained with the Corryvreckan framework [86]. A dedicated analysis for the spatial resolution of an irradiated TJ-Monopix2 chip was performed with TBSW (Test Beam SoftWare) [87–89]. Although the methods used in both frameworks differ, the analysis objectives are the same. In the following, the analysis is described on the example of Corryvreckan. The analysis has two steps: Alignment and Analysis.

**Alignment:** The data files for each sensor and the TLU are loaded into Corryvreckan and events are defined based on the TLU trigger modes. The individual pixel hits are then clustered. Here, it is assumed that an electron traversing the sensor can cause multiple neighboring pixels to record a hit (or fire) by either directly traversing multiple pixels due to a non-normal flight path (here, normal refers to the geometric relation of a vector and a plane) with respect to the sensor plane, or by charge sharing between the pixels due to lateral diffusion of the charges produced in the pixels. Through clustering, these pixel hits that originate from the same incident particle are grouped together. The reconstructed hit position of the particle through the cluster is then obtained through a center-of-gravity approach, taking the pixel charge into account [90]. The hit position  $x^{\text{hit}}$  ( $y^{\text{hit}}$ ) along  $x$  ( $y$ ) of a cluster with  $n$  pixels and the signal charge  $q_i$  and position  $x_i$  ( $y_i$ ) of the  $i^{\text{th}}$  pixel is given by

$$x^{\text{hit}} = \frac{\sum_1^n q_i x_i}{\sum_1^n q_i} \quad (8.2)$$

and  $y^{\text{hit}}$  is analogously determined. The calibrated pixel charge is used if it is available. A general-broken-lines (GBL) [91] track model is then fit onto the clusters to reconstruct the particle tracks through the telescope. Using these tracks the telescope can be aligned in software. This is realized by fitting the alignment parameters of the sensor planes, i.e. the positions and rotations, and the track parameters together. This is an iterative process that continuously updates the geometry and refits the tracks until the telescope is sufficiently aligned. The algorithm that is mainly used is the Millipede algorithm [92].

**Analysis:** The analysis follows similar steps as the alignment procedure until the tracking step. Here, the tracking explicitly excludes the DUT in order to not bias the reconstructed track position with respect to the DUT. The tracking uses the aligned geometry that was the result of the alignments step and, again, the GBL track model. Each cluster on the telescope planes can only be associated to one track. The interpolated track intersects on the DUT are then associated with cluster hits of the DUT. Using this association, time and spatial resolution, efficiency and other metrics can be measured. For the analysis each track needs to have a hit on every tracking plane, including the TelePix2. For the association of DUT hits, a spatial cut of 100  $\mu\text{m}$  and a timing cut of 50  $\mu\text{s}$  are applied.

The order of steps for both the alignment and analysis is visualized in Figure 8.3. Explicit configuration files can be found in the public analysis repository [93].

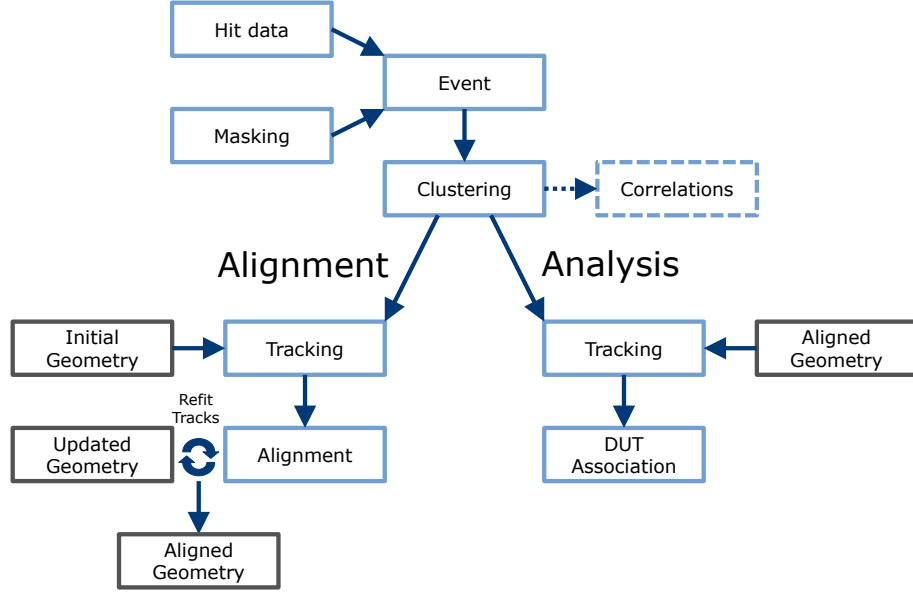


Figure 8.3: Schematic of the test beam analysis flow.

### 8.3 Test Beam Results

The results of the four irradiated TJ-Monopix2 samples are presented and compared in the following. The measurements were performed for 2 of the 4 front ends available on the TJ-Monopix2. The front ends considered are the cascaded AC- and DC-coupled front ends. The non-cascaded versions of the front ends were already excluded as potential front ends for the OBELIX chip in earlier measurements. Hereafter, the cascaded label is omitted in favor of better readability. Some measurements will refer to the maximum bias of the respective front end of the sensor. The maximum bias that could be applied to a sensor and front end varied and is summarized in Table 8.3.

| Serial number | Maximum bias AC FE | Maximum bias DC FE |
|---------------|--------------------|--------------------|
| W8R4          | 35 V               | 10 V               |
| W2R17         | -                  | 15 V               |
| W8R6          | 40 V               | 20 V               |
| W2R5          | 40 V               | 20 V               |

Table 8.3: The maximum bias that could be applied to each irradiated sample for the AC-coupled front end (AC FE) and the DC-coupled FE (DC FE). The AC-coupled front end of W2R17 was not operational.

### 8.3.1 Cluster Properties

Three main cluster properties are discussed here: Cluster size, cluster charge and seed charge. The cluster size is the number of pixels in a cluster. The cluster charge is the sum of all charges of all pixels that are part of a cluster. The seed charge is the highest single pixel charge contributing to the cluster. For 1-pixel clusters the seed charge and cluster charge are identical. In order to suppress noise, only clusters that were matched to a track are selected for the results provided here.

The cluster size and cluster charge distributions for an unirradiated sample (W14R12) and a sample at the maximum available fluence (W8R6) are shown in Figure 8.4 for the DC-coupled front end. The W14R12 sensor has a Czochralski substrate that, in theory, allows for deeper depletion compared to the EPI substrate of W8R6. The larger depletion volume would result in larger signals and larger clusters, but previous measurements determined the depletion depth to be similar [76]. The most probable cluster size for the unirradiated sample is 2. About 20 % of the clusters have a size  $\geq 3$ . For the irradiated sample the most probable cluster size is 1 with negligible contributions of clusters with size  $\geq 3$ . In this example the charge is not calibrated and is given in units of ToT. Results of calibrated charges are given below. The distribution of ToT values for the unirradiated sample is much broader, using a larger range of ToT values. In contrast, the range of ToT values used for the irradiated sample is compressed, reducing the charge resolution. While the uncalibrated charges cannot be directly compared, the collected charge will be smaller for an irradiated sample due charge trapping. Charges are trapped when the drift of liberated electrons in the depleted region is interrupted by defects. The charges are released after some time, that may be longer than the integration time. If the charges are released outside the integration window, the charges will contribute to the leakage current, rather than the signal. The distributions of the pixel size and cluster charge in ToT for the other samples are shown in Appendix A.1.

The mean of the cluster size is for all irradiated sensors depending on the bias voltage are displayed in Figure 8.5. The cluster size can be observed to increase with the bias voltage as the depleted volume increases and charges deposited at the corners and edges contribute to the signal. The likelihood for pixels with a small deposited charge to detect a sufficiently large signal increases and multi-pixel clusters appear

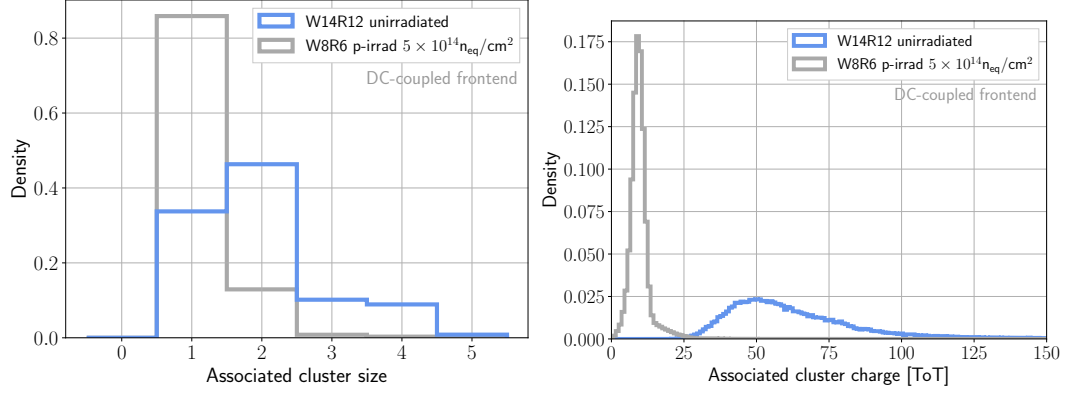


Figure 8.4: Comparison of the histograms of the cluster size and cluster charge for associated clusters in the DC-coupled front end. It compares an unirradiated sample (W14R12) and an irradiated sample (W8R6). Both measurements were taken at similar thresholds of  $\approx 300 e^-$ .

more often. The cluster size reaches a maximum that is dependent on the irradiation fluence. For the proton-irradiated sample with the highest fluence, the maximum is reached at 12 V while for the sample with the least fluence the maximum is  $< 6$  V. The cluster size reduces as the charge cloud drifts faster due to the larger electric field and the transversal spread due to diffusion reduces according to Equation 4.5. This reduces charge sharing between neighboring pixels, reducing the cluster size.

Charge calibrations are available for TJ-Monopix2 for samples W8R6 with a fluence of  $5 \times 10^{14} \text{ n}_{\text{eq}}/\text{cm}^2$  and W8R4 with a fluence of  $1 \times 10^{14} \text{ n}_{\text{eq}}/\text{cm}^2$ . The mean of the calibrated cluster charge and seed charge for W8R6 and W8R4 depending on the bias voltage are displayed in Figure 8.6 for the DC-coupled front end and in Figure 8.7 for the AC-coupled front end.

For the DC-coupled front end of W8R6, the range of bias voltages that can be applied is larger due to the large fluence and the reduced punch-through effect (see Section 7.1), compared to W8R4. The collected charge increases over the range of bias voltages for W8R6 as the drift velocity increases and more volume is depleted. Once the volume is fully depleted, the increase of the collected charge diminishes. This cannot be observed with W8R4, where the increase in collected charge already happened at smaller voltages and has already leveled out, similar to W8R6 after a bias voltage of 12 V. From the effects of NIEL damage, it is expected that a lower bias voltage already depletes the full volume at smaller fluences. A difference of the cluster charge for W8R6 would be expected due to the larger bulk damage and the additional charge trapping. However, the charge collected by both samples is similar with  $\approx 1700 e^-$ . The effect of larger fluences on the collected cluster charge appears to be small, once both samples are fully depleted. The difference between the seed charge and cluster charge is small as the clusters are predominantly of size 1.

A bias voltage of up to 40 V was applied to the AC-coupled front end for W8R6.

The collected cluster charge increased steadily and the trend suggests that the sensor was still not fully depleted at 40 V, however bias voltages much beyond 40 V are not considered safe for operation. The highest bias voltage applied to the AC-coupled front end of W8R4 is 35 V with negligible changes of the collected charge with different bias voltages, suggesting that the volume is already mostly depleted. The difference between the collected charge between the two samples is more pronounced in the AC-coupled front end compared to the DC-coupled front end. The cluster charge difference between W8R6 and W8R4 is  $\approx 500\ e^-$ , or  $\approx 30\%$  more charge collected by W8R4 over W8R6. The charge collected by the AC-coupled front end at respective maximum bias is smaller than the charge collected by the DC-coupled front end. While the higher bias voltage that applied to the AC-coupled front end should result in a larger collected charge, the additional capacitive coupling causes a significant loss of charge [54].

Comparisons for the uncalibrated charges are displayed in Appendix A.1 for the DC-coupled front end and for the AC-coupled front end.

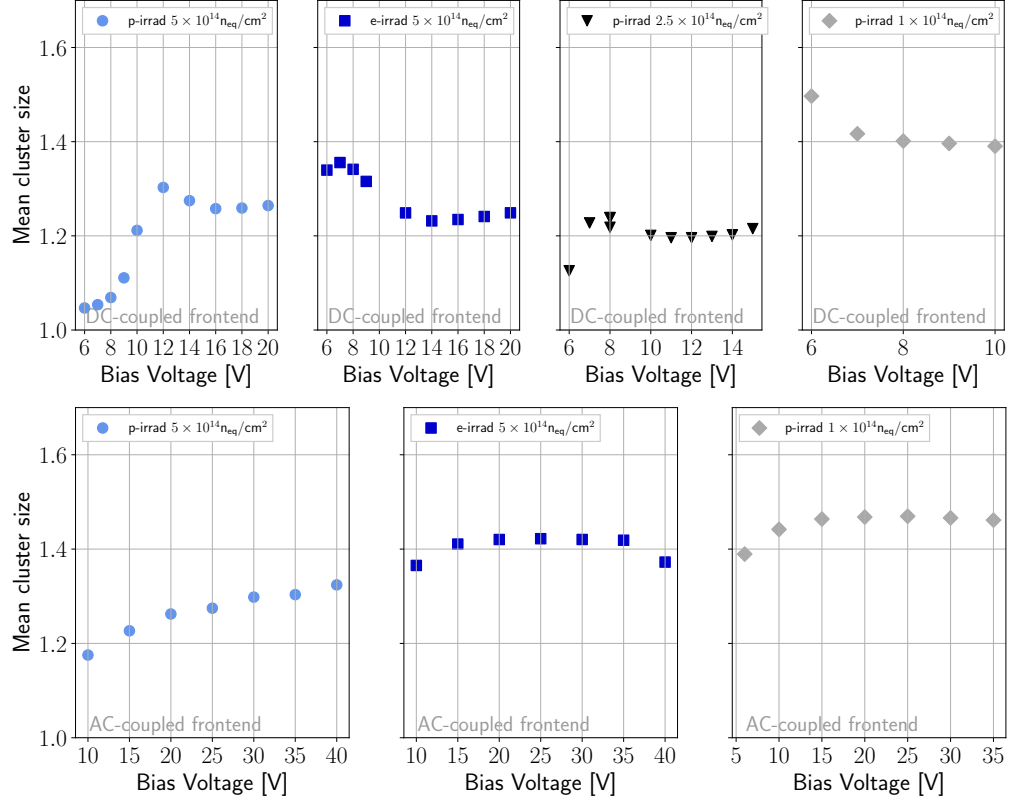


Figure 8.5: Comparison of the cluster size for all irradiated samples against the bias voltage. The threshold during these measurements was  $\approx 250 e^-$ . Top: Cluster size of the DC-coupled front end. Bottom: Cluster size of the AC-coupled front end. The statistics for these measurements are large and errors are not visible.

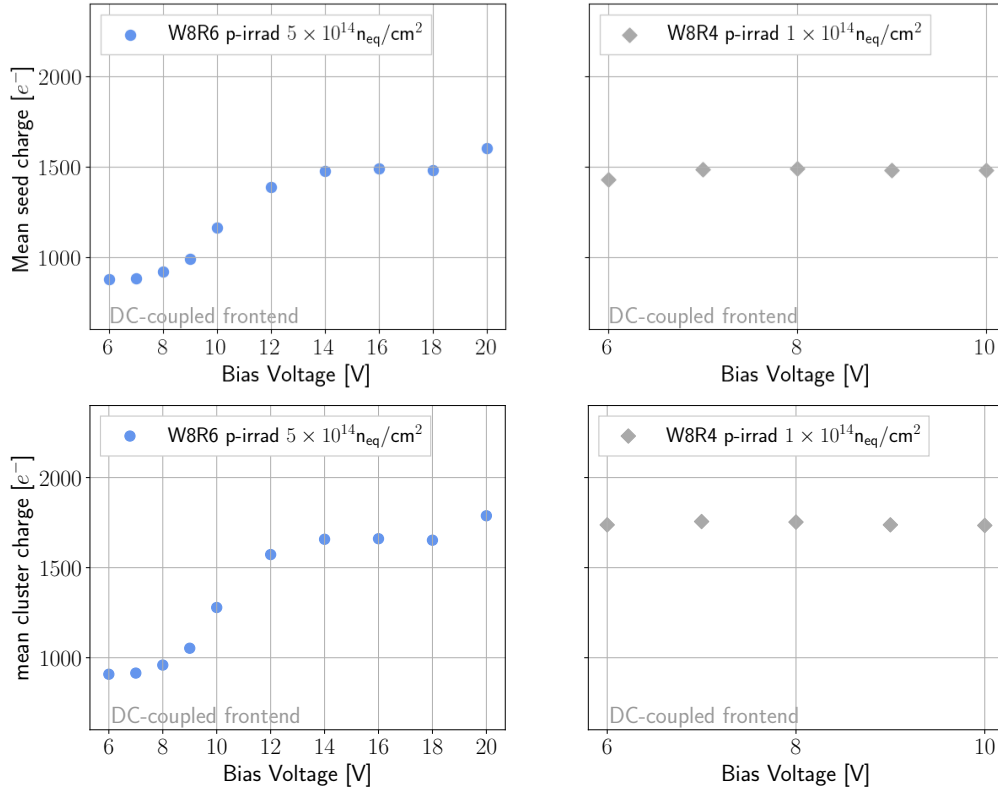


Figure 8.6: Comparison of the calibrated seed charge and cluster charge against the bias voltage for samples W8R6 and W8R4. The threshold during these measurements was  $\approx 250 e^-$ . The statistics for these measurements are large and errors are not visible. Top row: Seed charge of the DC-coupled front end. Bottom row: Cluster charge of the DC-coupled front end.

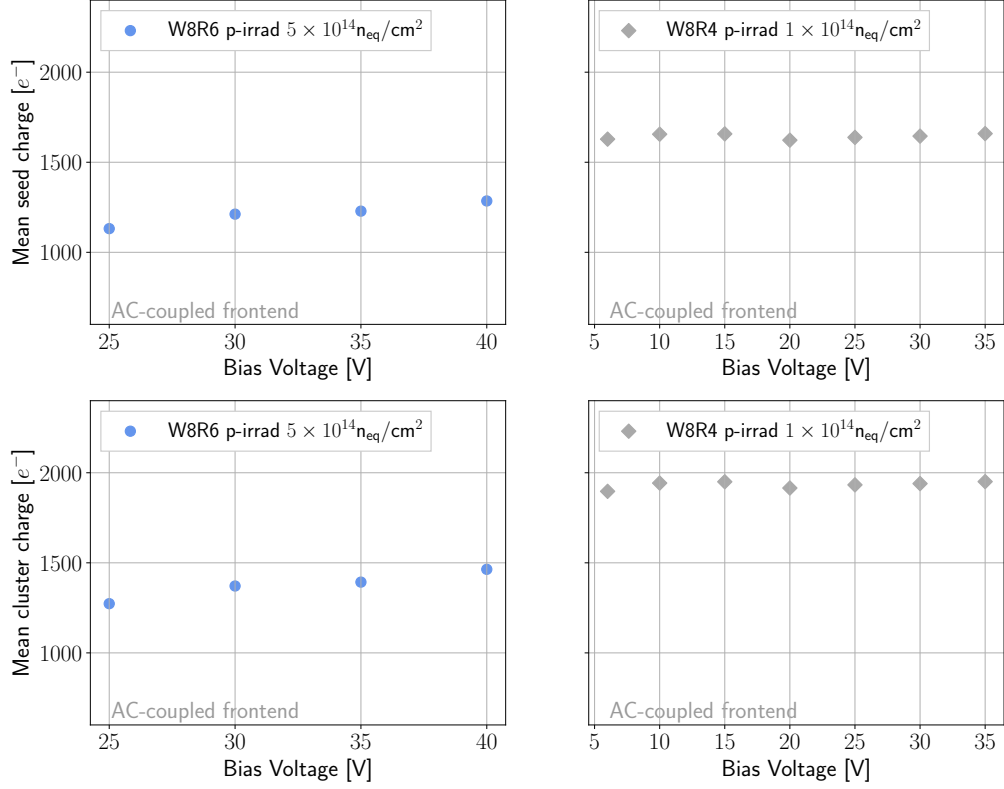


Figure 8.7: Comparison of the calibrated cluster charge and cluster charge against the bias voltage for samples W8R6 and W8R4. The threshold during these measurements was  $\approx 250 e^-$ . The statistics for these measurements are large and errors are not visible. Top row: Seed charge of the AC-coupled front end. Bottom row: Cluster charge of the AC-coupled front end.



### 8.3.2 Spatial Resolution

The 1-dimensional intrinsic binary spatial resolution for a pixel detector producing only single pixel clusters can be obtained from the standard deviation of a box function [36]

$$b(x) = \begin{cases} \frac{1}{d-c} & \text{for } c \leq x \leq d \\ 0 & \text{for } x < c \text{ or } x > d \end{cases} \quad (8.3)$$

with the standard deviation  $\frac{d-c}{\sqrt{12}}$ . The pixel boundaries are  $c$  and  $d$  and the pixel pitch is identified as the distance  $d - c$ . For the TJ-Monopix2 this simple binary intrinsic resolution is  $\sigma_{\text{intrinsic}} = \frac{33.04 \mu\text{m}}{\sqrt{12}} = 9.54 \mu\text{m}$ . An improvement over this resolution is expected due to multi-pixel clusters and information of the measured charge. The effective pixel pitch for 1-pixel clusters is then reduced by the by the length of the region, where multi-pixel clusters are produced. A visual representation of this effect along a 1D axis is given in Figure 8.8. In practice, the border between a single and a multi pixel response is not sharp. This can be visualized in the cluster size in dependence of the in-pixel hit position. For the measurement that is used here for the determination of the spatial resolution this visualization is shown in Figure 8.9.

The reconstructed hit position needs to be unbiased. The position finding algorithm used by TBSW uses a data-driven cluster shape correction that produces unbiased hit positions for multi-pixel clusters [87]. With the effective pixel pitch for single pixel clusters, the resolution of 1-pixel clusters measured for the TJ-Monopix2 should be smaller than  $9.54 \mu\text{m}$ . The multi-pixel resolution given the measured charge information approaches zero, if the charge distribution is known, and assuming no noise in the charge measurement of the pixels contributing to the cluster [36]. The Gaussian term in the fit function that is applied below, is a combination of the measurement noise of the pixels of the DUT and uncertainties arising from the finite resolution of the telescope.

The Gaussian noise term that originates from the finite telescope resolution and the resulting uncertainty of the track intersection position with the DUT, needs to be subtracted. The track positions  $x_{\text{Track}}$  ( $y_{\text{Track}}$ ) are statistically independent from the hit positions  $x_{\text{Hit}}$  ( $y_{\text{Hit}}$ ), since the DUT hit was not used for the track fitting. The measured spatial resolution  $\sigma_x^{\text{DUT}}$  ( $\sigma_y^{\text{DUT}}$ ) in  $x$  ( $y$ ) can then be obtained from the standard deviation  $\sigma_x^{\Delta}$  ( $\sigma_y^{\Delta}$ ) of the spatial residual distributions for all associated DUT clusters along  $x$  ( $y$ ) coordinates  $\Delta_x^{\text{DUT}} = x_{\text{Track}} - x_{\text{Hit}}$  ( $\Delta_y^{\text{DUT}} = y_{\text{Track}} - y_{\text{Hit}}$ ) and the telescope resolution, or mean track uncertainty,  $\sigma_x^{\text{tel}}$  ( $\sigma_y^{\text{tel}}$ ) via

$$\sigma_{x/y}^{\text{DUT}} = \sqrt{\left(\sigma_{x/y}^{\Delta}\right)^2 - \left(\sigma_{x/y}^{\text{tel}}\right)^2}. \quad (8.4)$$

The measurement of the spatial resolution after irradiation is performed for the DC-coupled front end for the TJ-Monopix2 chip W8R6, which was proton irradiated to a fluence of  $5 \times 10^{14} \text{ n}_{\text{eq}}/\text{cm}^2$ . The spatial residual distributions for all cluster

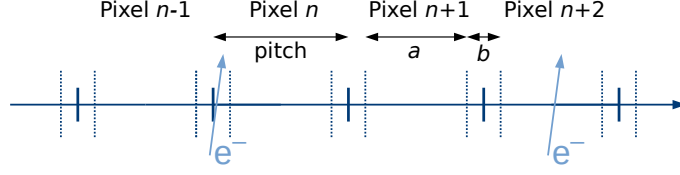


Figure 8.8: Visualization of a simplified 1D model for the effective pixel pitch of single-pixel clusters in the presence of multi-pixel clusters. The solid lines indicate the physical pixel borders. The dotted lines indicate the boundaries between the inner region of the pixel and the edge region of the pixel. An incidence particle traversing a pixel in the inner region with a length  $a$  of pixel  $n + 2$  produces a response only in pixel  $n + 2$ . If it traverses the pixel in the edge region with length  $b$  between the pixels  $n$  and  $n - 1$ , a response is measured in the two adjacent pixels, producing a multi-pixel cluster. The effective pixel pitch for a single-pixel cluster is equal to the length  $a$ . The pixel pitch is equal to the length of  $a + b$ .

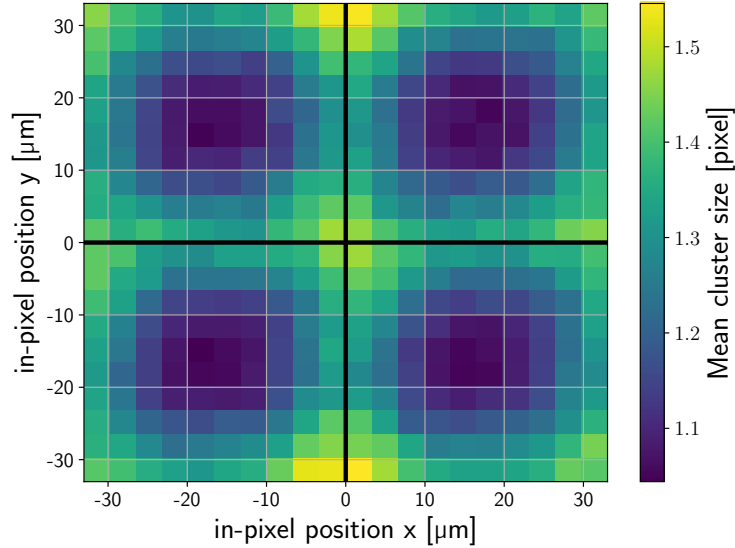


Figure 8.9: Mean cluster size depending on the reconstructed in-pixel hit position from measurements for the proton irradiated sample W8R6. Pixels are grouped in  $2 \times 2$  cells and mapped onto a common cell. The boundaries between the cells is indicated by solid black lines. This result was obtained using TBSW with cluster shape-based position reconstruction.

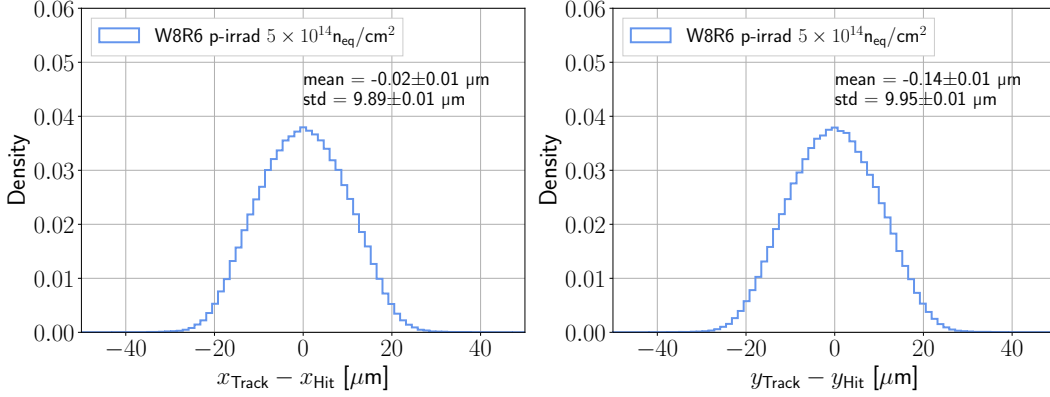


Figure 8.10: Residual distributions for W8R6. The statistical mean and standard deviation are drawn into the plot. Left: Residual distribution along the  $x$ -axis. Right: Residual distribution along the  $y$ -axis. This result was obtained with TBSW.

sizes along  $x$  and  $y$  and the statistical mean and standard deviation are displayed in Figure 8.10. The spatial residual distribution in for all clusters is a statistical mixture of residual distributions of 1-pixel, 2-pixel and larger ( $> 2$ ) clusters.

The spatial residual distribution is by good approximation the sum of the residual distributions for 1-pixel cluster and 2-pixel clusters. Contributions from larger clusters are small for irradiated samples, as shown in Figure 8.4. The spatial resolution can be obtained separately for 1-pixel and 2-pixel clusters. Their spatial residual distributions are displayed in Figure 8.11 for 1-pixel clusters and in Figure 8.12 for 2-pixel clusters. A convolution of a box function and a Gaussian, resembling the noise term, has been fitted to the distributions to obtain the effective pitch, as well as the parameters of the Gaussian noise. The Gaussian noise has fitted here has contributions from noise of the track positions and from the gaussian smearing in the hit positions. The fit variables are  $a$ ,  $\mu$  and  $\sigma$ , that are the half width of the box function extending from  $-a$  to  $a$  (and is defined to be centered around 0) and the mean and standard deviation of the Gaussian function. The fit functions and the fit parameters are indicated in the respective figures. As expected from the above discussed model, the effective pixel pitch for 1-pixel cluster are  $2 \times (14.96 \pm 0.11) \mu\text{m}$  and  $2 \times (14.89 \pm 0.13) \mu\text{m}$  for  $x$  and  $y$  residuals, respectively, which is smaller than the physical pixel pitch of  $33.04 \mu\text{m}$ . The width of the residuals for 2-pixel clusters is smaller compared to residuals of 1-pixel clusters, which is expected. The contribution from the box function, i.e. the fit parameter  $a$ , is small and the error of the hit position for 2-pixel clusters is dominated by the Gaussian noise, which is also expected according to the above considerations.

The unbiased mean track intersection uncertainty at the DUT is known from the track fitting by TBSW and is determined to be  $\sigma_x^{\text{tel, meas}} = (3.7942 \pm 0.0003) \mu\text{m}$  and  $\sigma_y^{\text{tel, meas}} = (3.8005 \pm 0.0003) \mu\text{m}$ . The track intersection uncertainty is determined for each track and the error on  $\sigma_x^{\text{tel, meas}}$  is only the statistical error on the mean

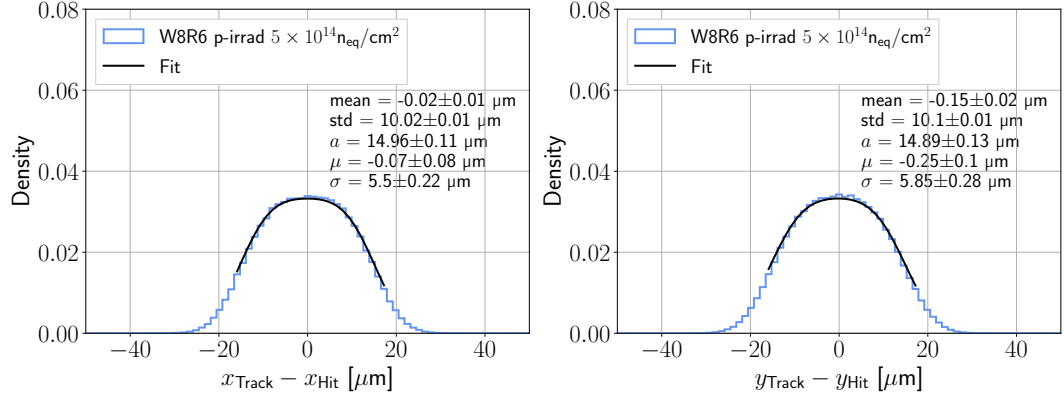


Figure 8.11: Residual distributions for single-pixel clusters for W8R6. A fit of a centered box function convoluted with a Gaussian has been applied to the distribution. The statistical mean and standard deviation, as well as the fit parameters, are drawn into the plot. Left: Residual distribution along the  $x$ -axis. Right: Residual distribution along the  $y$ -axis. This result was obtained with TBSW.

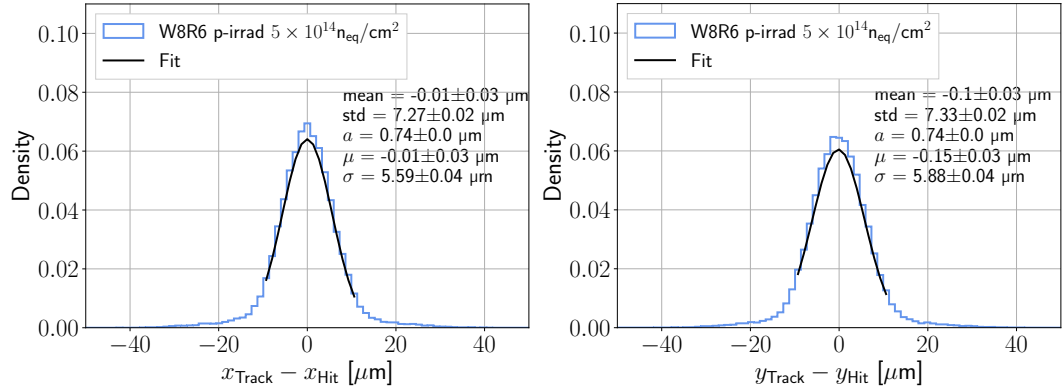


Figure 8.12: Residual distributions for 2-pixel clusters for W8R6. A fit of a centered box function convoluted with a Gaussian has been applied to the distribution. The statistical mean and standard deviation, as well as the fit parameters, are drawn into the plot. Left: Residual distribution along the  $x$ -axis. Right: Residual distribution along the  $y$ -axis. This result was obtained with TBSW.

| Cluster size    | Resolution in $x$           | Resolution in $y$           |
|-----------------|-----------------------------|-----------------------------|
| 1-pixel cluster | $(9.3 \pm 0.1) \mu\text{m}$ | $(9.4 \pm 0.1) \mu\text{m}$ |
| 2-pixel cluster | $(6.2 \pm 0.2) \mu\text{m}$ | $(6.3 \pm 0.2) \mu\text{m}$ |
| All clusters    | $(9.2 \pm 0.1) \mu\text{m}$ | $(9.2 \pm 0.1) \mu\text{m}$ |

Table 8.4: Spatial resolution achieved in measurements of W8R6 for 1-pixel, 2-pixel and all clusters.

track intersection uncertainty. Systematic errors can arise from various sources like misalignment, uncertainties of the particle energy, uncertainties in the estimation of the resolution of the telescope and uncertainties due to the non-continuous measurements of the telescope. The systematic error is estimated by determining the mean track intersection uncertainty using multiple methods and comparing their results.

The second method used is by performing TBSW track fitting using simulated data, instead of measured data (see Section 8.4 for details on the simulation). The mean track intersection uncertainties on simulated data are determined to be  $\sigma_x^{\text{tel, sim}} = (4.0466 \pm 0.0004) \mu\text{m}$  and  $\sigma_y^{\text{tel, sim}} = (4.0416 \pm 0.0004) \mu\text{m}$ . The errors quoted are statistical uncertainties as above. The third method uses a GBL track resolution calculator [94], that determines the mean track intersection uncertainty in both  $x$  and  $y$  to be  $\sigma^{\text{tel, calc}} = 3.61 \mu\text{m}$ . The difference between the first method that used measurement data and the latter two methods is  $\pm 0.2 \mu\text{m}$ , which will be used as an estimate for the systematic uncertainty. The mean track intersection uncertainty is therefore taken to be the mean of all three values. The obtained resolution, that is rounded according to the error, is  $\sigma_x^{\text{tel}} = \sigma_y^{\text{tel}} = (3.8000 \pm 0.0003 \pm 0.2000) \mu\text{m}$ , where the first error is the statistical uncertainty and the second error is the systematic error. This telescope resolution and uncertainty will be used for all following calculations, including the determination of the spatial resolution through simulation in Section 8.4.

The position resolution of 1-pixel and 2-pixel clusters, as well as the total position resolution is then obtained using the statistical standard deviations of each distribution and the determined track intersection uncertainty by TBSW. The results are given in Table 8.4. The 1-pixel cluster resolution is slightly better than the intrinsic binary resolution. The spatial resolution for 2-pixel clusters is better than for 1-pixel clusters. The total spatial resolution over all clusters is close to the 1-pixel cluster resolution, since the 1-pixel clusters dominate. The obtained results fall within in the expectations explained above.

### 8.3.3 Hit Detection Efficiency

Vertex detectors are required to produce measurements with a high efficiency as missing measurements will affect the track finding, track reconstruction and result in a decreased vertex and momentum resolution. A hit detection efficiency of  $> 99\%$  for samples irradiated to the maximum expected fluence of  $5 \times 10^{14} \text{ n}_{\text{eq}}/\text{cm}^2$  is meant

to be achieved for the VTX.

The hit detection efficiency  $\epsilon$  is defined by the number of reconstructed tracks that were matched, or associated, to DUT hits  $n_{\text{matched}}$  and the total number of reconstructed tracks  $n_{\text{total}}$  via

$$\epsilon = \frac{n_{\text{matched}}}{n_{\text{total}}}. \quad (8.5)$$

The uncertainty of the efficiency is calculated as the boundaries for the Clopper-Pearson interval [95] that takes a lower limit of 0 and an upper limit of 1 into account. The implementation of the ROOT framework is used for the computation [96]. Statistical errors are indicated in all plots for the efficiencies but are of order  $\mathcal{O}(10^{-5})$  due to the high statistics and not visible.

Dead or disabled pixels are excluded from the efficiency measurement by not counting reconstructed tracks that intersect the DUT at position of dead or disabled pixels. Noisy pixels were identified and disabled prior to the measurement, however, there is no treatment for pixels that became noisy during the measurement. The noise rate is small and the systematic bias on the efficiency measurement is neglected. The efficiencies that are presented in the following are representative of the efficiency of entire front end and no regions-of-interest have been defined within the area of the front ends. One exception is a masking of  $\approx 100$  rows on top of the sensor that were in the shadow of a copper block from the cooling system.

The efficiency in dependence of the bias voltage for both tested front ends is displayed in Figure 8.13 for all irradiated samples. For small bias voltages the efficiency drops, since the depletion region shrinks and the obtained signal charges become too small to pass the threshold. These findings are compatible with expectations arising from measurements of unirradiated samples [51, 76]. Samples with smaller NIEL fluences retain higher hit detection efficiencies  $> 99\%$  at smaller bias voltages, while the highest irradiated W8R6 chip shows a larger degradation. This is due to the increase of the depletion voltage from the irradiation, amplifying the reduction of the depletion region at smaller bias voltages. Additional charge collection inefficiencies from charge trapping further decrease the hit detection efficiency at large fluences. However, a bias voltage can be found to obtain efficiencies  $> 99\%$  for any irradiated sample.

The efficiency in dependence of the threshold for both tested front ends is displayed in Figure 8.14. The threshold was determined using the charge injection circuit as described in Section 4.6. The threshold for the TJ-Monopix2 is mainly tuned by manipulating the  $I_{\text{BIAS}}$ ,  $I_{\text{THR}}$  and  $I_{\text{CASN}}$  registers of the amplifier circuit that are indicated in Figure 4.12.

Three conclusions can be drawn from this measurement. Firstly, the efficiency degrades with larger irradiation, such that less irradiated chips have higher efficiencies, even at higher thresholds. Secondly, despite an irradiation to the same 1 MeV neutron equivalent fluence, the electron irradiated chip exhibits less performance degradation than the proton irradiated chip. This suggests that either the irradiation with electrons never reached the expected NIEL fluence or that the irradiation

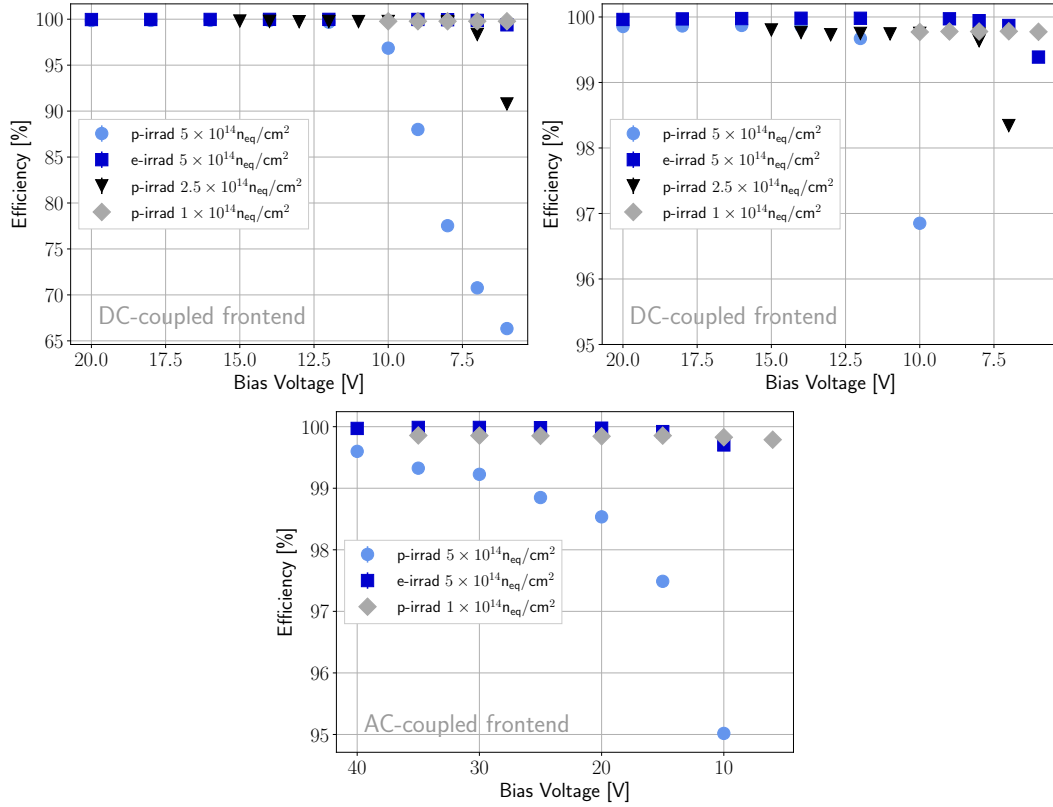


Figure 8.13: Hit detection efficiency of all irradiated samples in dependence of the bias voltage for the DC-coupled front end (top left), magnified for high efficiencies (top right), and the AC-coupled front end (bottom). The threshold during these measurements was  $\approx 250 e^-$  except the p-irradiated  $2.5 \times 10^{14} n_{eq}/cm^2$  (W2R17) which was  $\approx 400 e^-$ .

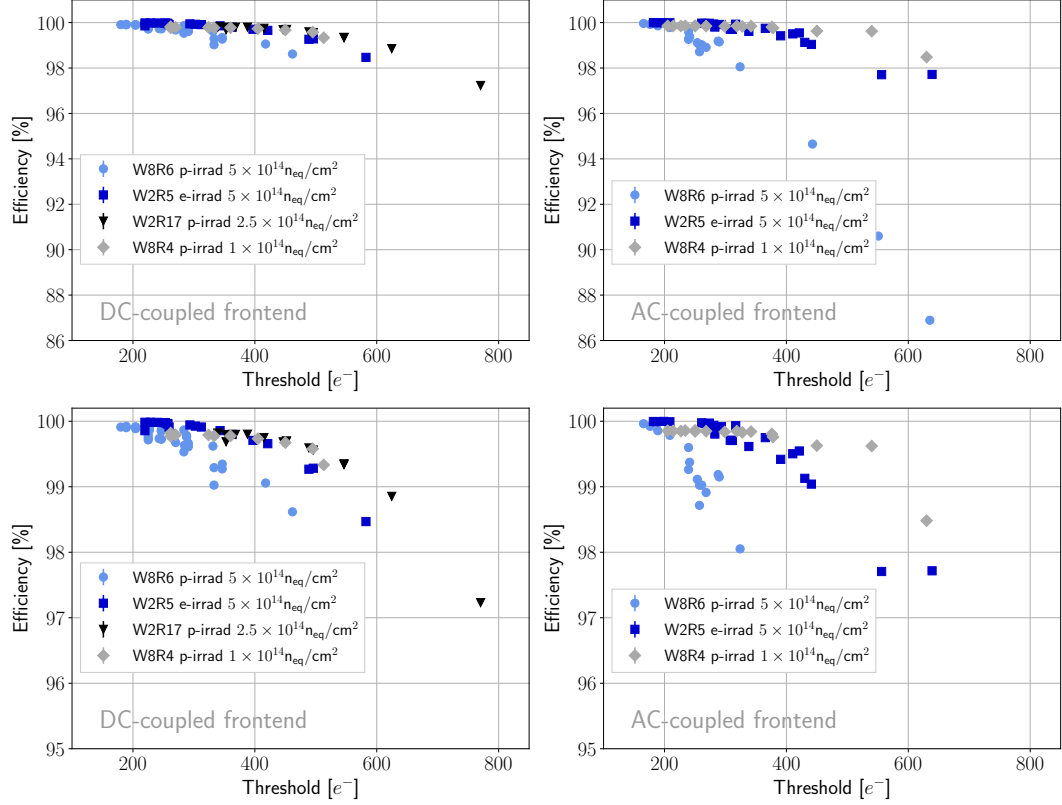


Figure 8.14: Hit detection efficiency of all irradiated samples in dependence of the threshold for the DC-coupled front end (top left) and the AC-coupled front end (top right). Plots with a smaller scale are provided for the DC-coupled front end (bottom left) and the AC-coupled front end (bottom right). All measurements were taken at maximum bias for the respective front end and chip.

with protons lead to larger bulk damage than irradiation with electrons. The irradiation procedure has since been reviewed with no apparent issues found. Thirdly, the AC-coupled front end has smaller efficiencies at the same threshold and fluence than the DC-coupled front end. The direct comparison of the AC- and DC-coupled front ends of W8R6 is displayed in Figure 8.15. The smaller efficiency of the AC-coupled front ends is assumed to originate from the AC coupling itself, where the capacitor between the charge collection electrode and the amplifier circuit leads to a charge loss of around 30 % [54].

The efficiency in dependence of the temperature is displayed in Figure 8.16. The measurements are taken at maximum bias. The efficiency is observed to degrade at higher temperatures. The effect is larger for AC-coupled front ends at large fluences. This degradation can be attributed to the increase of the threshold with increasing temperature. Importantly, the chip settings have not been changed during the



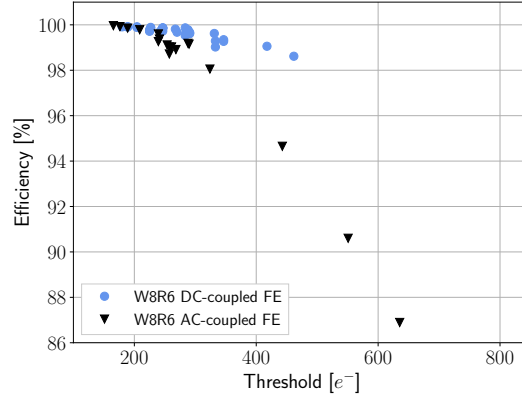


Figure 8.15: Hit detection efficiency of the AC- and DC-coupled front end of W8R6 that is proton irradiated to a NIEL fluence of  $5 \times 10^{14} \text{ n}_{\text{eq}}/\text{cm}^2$ .

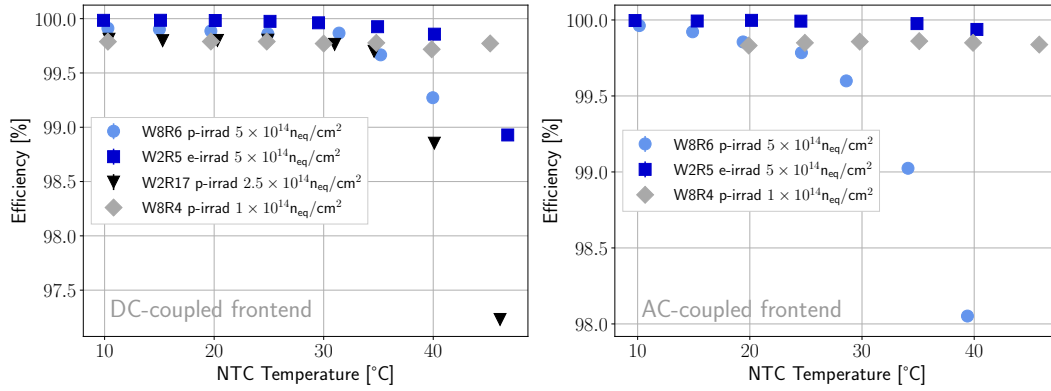


Figure 8.16: Hit detection efficiency of all irradiated samples in dependence of the NTC temperature for the DC-coupled front end (left) and the AC-coupled front end (right). All measurements were taken at maximum bias for the respective front end and chip. The threshold during these measurements was between  $170 e^-$  and  $350 e^-$  except the p-irradiated  $2.5 \times 10^{14} \text{ n}_{\text{eq}}/\text{cm}^2$  (W2R17) which was between  $350 e^-$  and  $700 e^-$ .

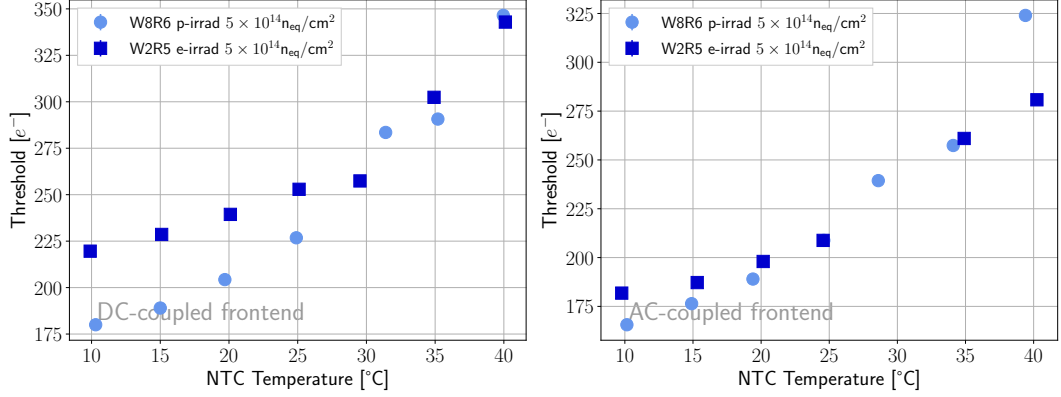


Figure 8.17: Threshold of the most irradiated samples in dependence of the NTC temperature for the DC-coupled front end (left) and the AC-coupled front end (right). All measurements were taken at maximum bias for the respective front end and chip.

temperature scan, thus the change in threshold is solely dependent on the change in temperature. The threshold in dependence of the temperature is displayed in Figure 8.17. The hypothesis for this effect is that a higher leakage current at higher temperatures causes a faster discharge of the charge collecting electrode and thus effectively increases the threshold. The leakage current in dependence of the temperature is displayed in Figure 8.18

The precise spatial resolution of the DURANTA telescope allows to study the uniformity of the pixel response within a pixel. The pixel responses are, according to the track intersection position with the TJ-Monopix2 sensor, mapped onto a single unit pixel. The unit pixel spans one pixel-pitch in each direction and has a binning of  $34 \times 34$ . The number of considered tracks for the in-pixel efficiency maps is  $> 160000$ , leading to an average of  $> 100$  tracks per bin.

The in-pixel efficiency maps at different temperatures for the AC-coupled front end of W8R6 are displayed in Figure 8.19 and for the DC-coupled front end they are displayed in 8.20. The example temperatures were chosen such that the degradation of the efficiency at high temperatures is well visualized at the same color scale for both front ends respectively.

The degradation of the efficiency is observed to be strongest in the corners. This is expected since the charge sharing between pixels is larger for cases where the incidence particle traverses the sensor at a pixel corner. At higher thresholds, which are observed for higher temperatures, the analog signal produced by the shared charge in each pixel is then too small to cross that threshold and no hit is recorded by the sensor. This effect is pronounced by the reduced charge collection in irradiated samples.

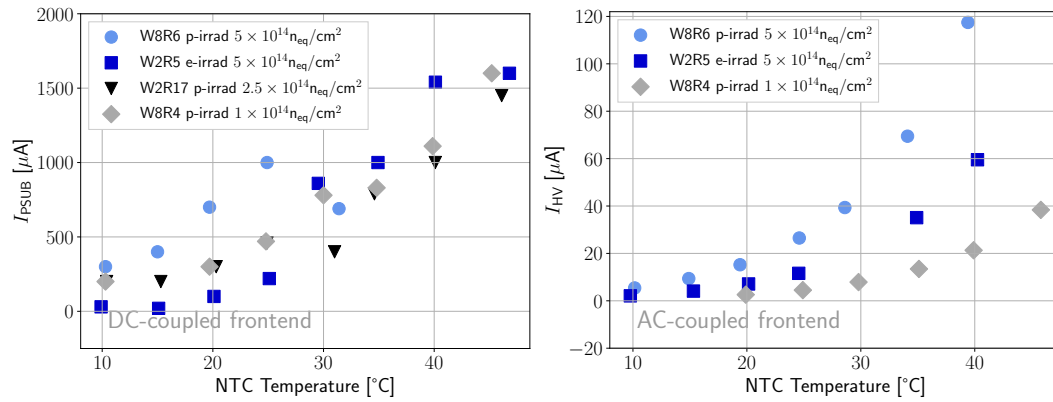


Figure 8.18: Leakage current of all irradiated samples in dependence of the NTC temperature for the DC-coupled front end (left) and the AC-coupled front end (right). All measurements were taken at maximum bias for the respective front end and chip.

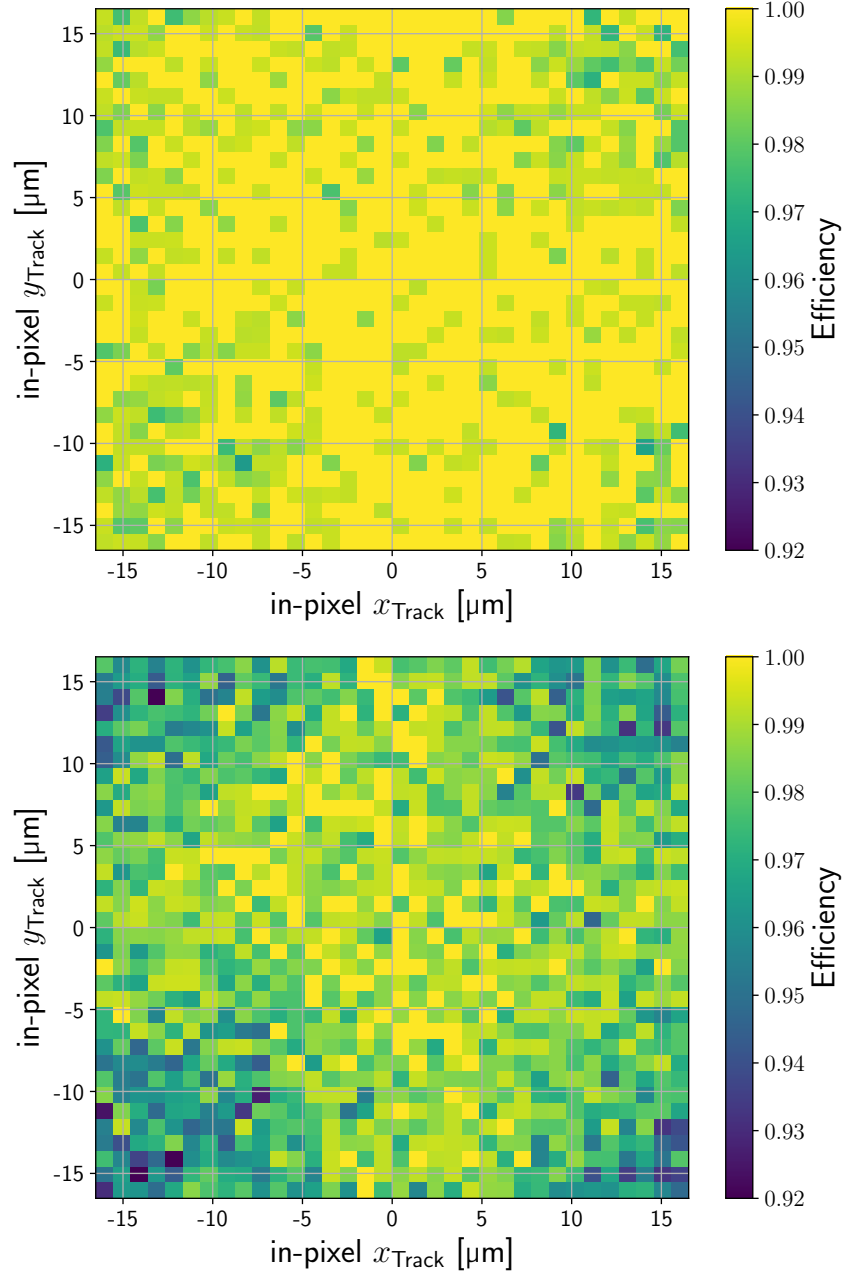


Figure 8.19: In-pixel efficiency map of the irradiated sample W8R6 with a fluence of  $5 \times 10^{14} \text{ n}_{\text{eq}}/\text{cm}^2$  for the AC-coupled front end. Top: Measurement at  $T_{\text{NTC}} = 30^\circ\text{C}$  with a total efficiency of  $99.60^{+0.02}_{-0.02} \%$  at a threshold of  $239 e^-$ . Bottom: Measurement at  $T_{\text{NTC}} = 40^\circ\text{C}$  with a total efficiency of  $98.05^{+0.03}_{-0.04} \%$  at a threshold of  $324 e^-$ .

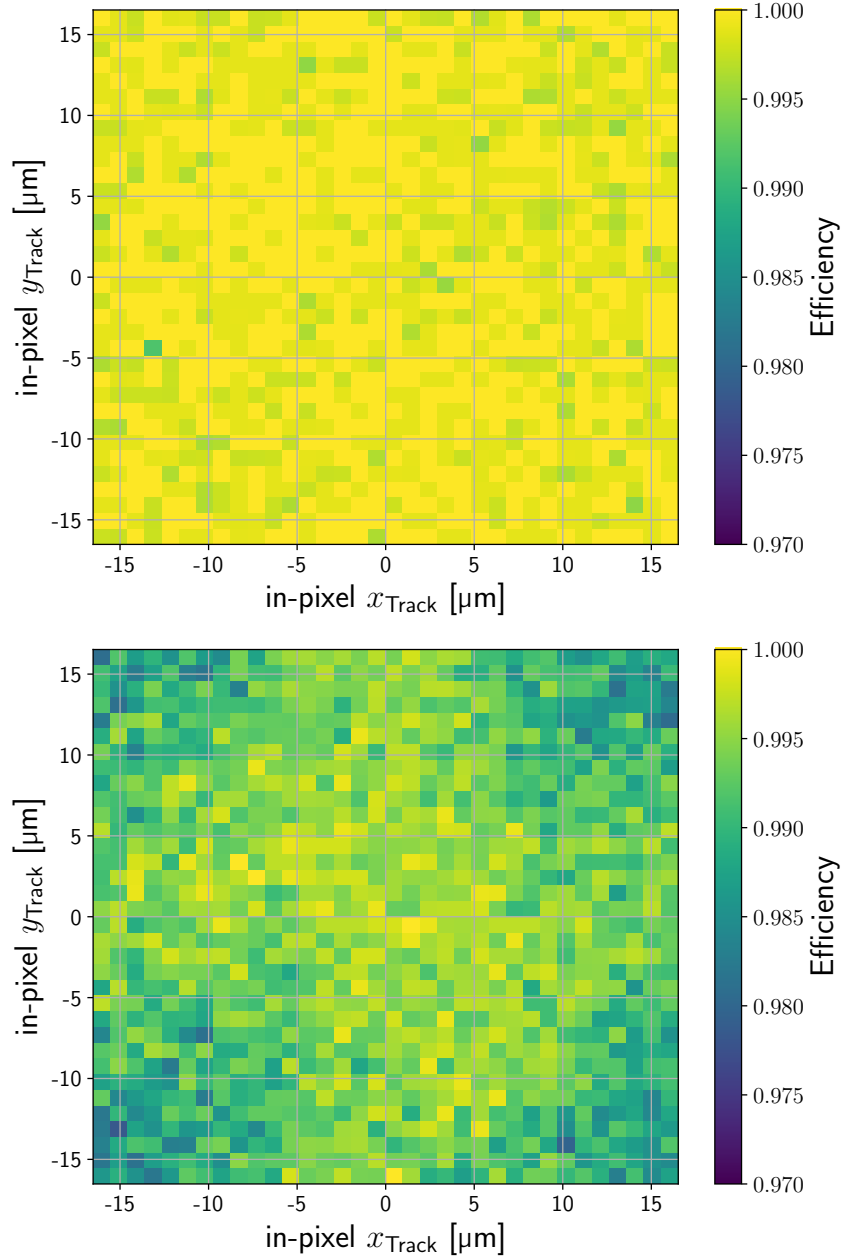


Figure 8.20: In-pixel efficiency map of the irradiated sample W8R6 with a fluence of  $5 \times 10^{14} \text{ n}_{\text{eq}}/\text{cm}^2$  for the DC-coupled front end. Top: Measurement at  $T_{\text{NTC}} = 10^\circ\text{C}$  with a total efficiency of  $99.912^{+0.003}_{-0.003} \%$  at a threshold of  $180 e^-$ . Bottom: Measurement at  $T_{\text{NTC}} = 40^\circ\text{C}$  with a total efficiency of  $99.27^{+0.01}_{-0.01} \%$  at a threshold of  $347 e^-$ .

## 8.4 Comparison to Simulation

In order to judge and confirm the conclusions of the test beam measurement a simulation of the test beam was conducted. Especially the influence of the irradiation effects should be reproduced to test if the behavior of TJ-Monopix2 is consistent with established models. The simulation is based on Allpix<sup>2</sup> [97], a modular simulation framework for semiconductor tracker and vertex detectors. Instead of simulating only the responses of the TJ-Monopix2 chip, the full DURANTA telescope consisting of six MIMOSA26 sensors is simulated as well. The simulation produces a hit table similar to the test beam measurements. This allows for reusing the same Corryvreckan and TBSW analysis setup as it was used for the measurements. The simulation aims to reproduce the measurements performed of the p-irradiated sample W8R6 which was irradiated to a fluence of  $5 \times 10^{14} \text{ n}_{\text{eq}}/\text{cm}^2$ . The simulation parameters are chosen such that they are the same as during the measurement. The simulation is performed with a threshold of  $250 \text{ e}^-$ , a bias voltage of 20 V and at an NTC temperature of  $30^\circ\text{C}$  which is approximately a chip temperature of  $37^\circ\text{C}$ . The cluster properties and spatial resolutions are compared.

A description of the simulation setup with Allpix<sup>2</sup> and the assumptions made as well as the simulation results are presented in the following.

### 8.4.1 Simulation Setup

A previously aligned geometry from the test beam analysis is used as a basis, describing the positions and rotations of all six telescope planes and a TJ-Monopix2 sensor. Only these seven components are simulated. Since there is no event building necessary for the simulation, the TLU, scintillators and the TelePix2 timing plane are not part of the simulation. An event is implicitly defined as the simulation of an electron passing through the telescope.

For each sensor the dimensions and properties of both active and passive materials are defined. From this information a Geant4 [98] geometry is constructed. An electron gun is placed in front of the center of the first telescope plane. It simulates a particle beam of electrons with an energy of 4.2 GeV and a circular Gaussian beam profile. The standard deviation of the Gaussian profile is 25 mm, which is the collimator aperture of the last collimator used before the beam reaches the sensor in the DESY test beam measurement. The passage of electrons through the sensor material is then simulated by Geant4.

In the sensitive volume of each sensor an electric field is present. The baseline assumption is a linear electric field oriented in parallel to the  $z$ -axis. The bias voltage is taken from the specific measurement that shall be reproduced by the simulation.

The deposited charges by the incidence electron are then propagated through the active sensor volume. The motion of charges due to diffusion is simulated, in order to simulate the lateral motion and charge sharing in the simplified linear electric field and with the perpendicular incidence angle of the electron beam. The temperature of the sensor is an important parameter for the simulation of the diffusion and it is

taken from the NTC temperature measurements during the test beam, accounting for the expected difference to the chip temperature. The increased trapping of charges due to irradiation is simulated. A temperature dependence is expected for trapping effects. The *Ljubljana* trapping model, also referred to as *Kramberger* trapping model, is used [99], which features explicit temperature scaling.

Once the deposited charges reach the charge collection implant, they are summed up and an analog signal is generated. The generated charge is then modified by adding an electronics noise that is drawn from a Gaussian distribution with a standard deviation of  $160\ e^-$ . This value is chosen such that it provides an effective variation of the signal charge that arises from varying pixel responses due to electric fields and the pixel front ends. A typical value for the signal charge is  $> 1000\ e^-$ . The value for the electronics noise is not realistic in the absence of a signal charge, since it would imply extremely large noise rates at thresholds of around  $200\ e^-$  that were used during the test beam measurements. The noise observed during the test beam was small. As a consequence, pixel hits that are purely generated by the electronics noise, i.e. noise hits, are suppressed. This is also necessary for the computational performance of the simulation as the number of simulated channels is significant.

The signal is compared to a set threshold and if the signal is large enough, it will be registered as a hit. A charge-to-digital conversion into ToT units is not performed by the simulation. The threshold is taken from measurements during the test beam. A threshold smearing of  $10\ e^-$  is generally assumed in the simulations. The resulting hit information is written to disk in a format that is compatible with Corryvreckan and the simulated data is reconstructed.

Since the capacitive coupling between the charge collection electrode and the amplifier circuit of the AC-coupled front end could not be explicitly simulated the simulation was only used to simulate the DC-coupled front end. The amplifier circuit and the discriminator are also not explicitly simulated and generic, effective implementations are used.

#### 8.4.2 Simulated Cluster Properties

In a first step for the comparison of the simulation to the test beam measurements, the pixel charge properties of the TJ-Monopix2 chip shall be compared. The comparison is done for the p-irradiated sample W8R6 which was irradiated to a fluence of  $5 \times 10^{14}\ n_{eq}/cm^2$ . The simulated front end is the DC-coupled front end.

The properties that are compared are the cluster size and the seed charge. The comparisons of the distributions are displayed in Figure 8.21. The cluster properties that are compared are only from clusters that are matched to tracks, both for the measurement and the simulation. The cluster size is well described by the simulation and the effect of the irradiation on the cluster size is similar to the effect seen in the measurements. Most of the clusters that are matched to tracks are 1-pixel clusters, while clusters with 3 or more pixels are rare. The ratio between 1- and 2-pixel clusters also matches with the measurements.

The pixel charges from the measurement are similarly available in units of electron charges, since charge calibrations are available for sample W8R6. However, as discussed previously, the calibrated charge has discrete values due to the conversion of the charge information into integer ToT units. The pixel charges produced by the simulation are in units of electron charges but the charge-to-digital conversion is not applied during the simulation. In order to make the two charge distributions comparable, the inverse charge calibration is applied to simulated charges, converting the units from an absolute charge in electrons into ToT units. The ToT values are then rounded to integer values and the charge conversion is applied, converting the charge back into absolute units of charge. The procedure is referred to as *double conversion* in the following. Further, only ToT values that are explicitly calibrated are considered. Higher charges, for which the charge calibration needs to extrapolate, are removed and do not contribute to the calculation of the mean values. The ToT units are equal to 25 ns due to the 40 MHz clock (see Section 8.1.1). The ToT value of the most probable value (MPV) of the seed pixel charge is 9 [25 ns] and identical between simulation and measurement resulting in a charge of  $1605 e^-$  for both. During the double conversion the ToT values in the range [8.5, 9.5) are mapped onto a value of 9 [25 ns]. The calibrated charge values of the range are [ $1432 e^-$ ,  $1782 e^-$ ). The error on the seed charge in electrons can be estimated by this range. The rounded mean difference of the edges of the range and the MPV of the seed charge is  $200 e^-$ , representing the estimated error for both the measurement and simulated MPVs due to the charge-to-digital conversion. The mean value of the measured charge is  $1604 e^-$ , while the mean of the simulated seed charges tends to slightly higher values of  $1714 e^-$ .

The cluster size and seed charge are in reasonable agreement with the measured values. Ideally, an explicit simulation of the front end on a pixel level is used to properly simulate the charge-to-digital conversion performed the amplifier circuit and its discriminator. Then a charge calibration on pixel level could be applied, similar to the measurements. Efforts by the collaboration made to implement such a simulation but is not yet available and not performed in this work. Instead, the charge conversion can only be applied after the simulation and can only be applied on the cluster level charges. The charge calibration function is not linear, especially at small charge values, and it cannot be applied on the sum of multiple charges. Thus, this comparison can only be performed on charges of single pixels, such as the seed charge. The distribution of the simulated seed charge shows a difference to the measured seed charge at small charges. The simulation produces considerably less seed charges of  $< 500 e^-$ . The reason for this is connected to the simplified front end simulation and electric fields. This difference implications for the simulation of the efficiency in dependence of the threshold. Increasing the threshold during the measurement within the typical ranges of  $200 e^-$  to  $500 e^-$  decreases the efficiency as the threshold cuts into the seed charge, the signal, distribution. When increasing the threshold in this range for the simulation, the impact on the efficiency is much smaller, as the seed charge distribution is much smaller in this range. The simulation



of the hit detection efficiency is discussed in Section 8.4.4.

Most clusters are 1-pixel clusters, where the seed charge is equal to the cluster charge and the seed charge dominates the charge contributions in multi-pixel clusters. As a result, the finding that the cluster sizes and seed charges are simulated reasonably well, allows in good approximation to assume that the cluster charges are also simulated reasonably well.

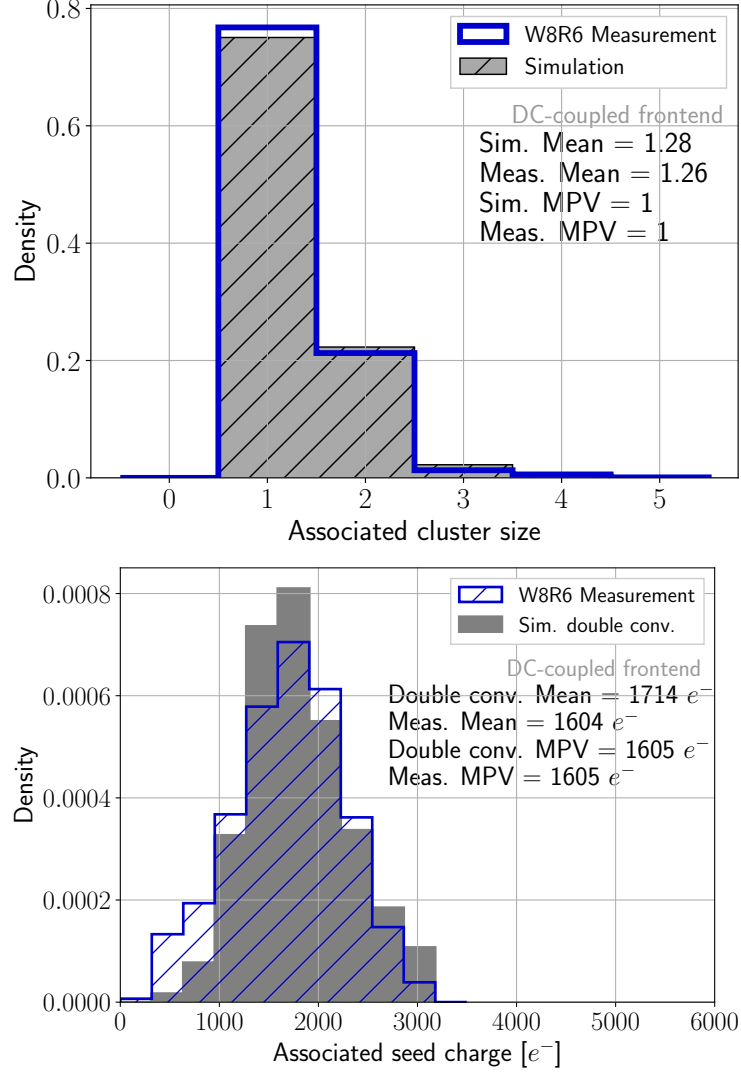


Figure 8.21: Comparison of cluster properties from simulation and measurements with the p-irradiated sample W8R6 with a fluence of  $5 \times 10^{14} \text{ n}_{\text{eq}}/\text{cm}^2$ . The measurement was taken at a threshold of  $250 \text{ e}^-$ , a bias voltage of 20 V and at an NTC temperature of  $30^\circ\text{C}$  which is approximately a chip temperature of  $37^\circ\text{C}$ . The mean and MPV are indicated in each plot for the measurement and simulation, respectively. Top: Comparison of the cluster size distribution. Bottom: Comparison of the seed charge distribution. The inverse charge conversion was applied to then apply the charge conversion again to gauge the effects of the conversion of the charge information into integer ToT values. Both distributions have been rebinned to allow for a better visualization of the discrete charge values. A version with finer binning that makes the discrete charge values visible is available in Appendix A.2.

### 8.4.3 Simulated Spatial Resolution

The measurement of the spatial resolution obtained for the simulation is performed similarly to the test beam measurement. The analysis of the spatial resolution is performed with TBSW, similar to the measurement. Tracks are reconstructed from the positions of pixel hits on the 6 telescope planes. The position of the intersection and its uncertainty of the reconstructed track with the DUT is computed. Then, the spatial residual is obtained from difference of the intersection position and the reconstructed hit position. The spatial resolution of the DUT is obtained from the spatial residual distribution and the track intersection uncertainty using Equation 8.4.

As for the test beam measurements, a convolution of a box function and a Gaussian, resembling the noise term, has been fitted to the distributions to obtain the effective pitch, as well as the parameters of the Gaussian noise. The fit variables are  $a$ ,  $\mu$  and  $\sigma$ , that are the half width of the box function extending from  $-a$  to  $a$  (and is defined to be centered around 0) and the mean and standard deviation of the Gaussian function. The fit functions and the fit parameters are indicated in the respective figures. The fit is performed for the simulated residual distributions. The distribution of the spatial residuals, as well as the fit and the distribution from measurements for comparison are shown in Figure 8.22 for all clusters, in Figure 8.23 for 1-pixel clusters and in Figure 8.24 for 2-pixel clusters. The spatial resolutions are given in Table 8.5.

The difference between the distributions of the residuals for 1-pixel clusters between the measurement and simulation can be analyzed using the fit parameters. The effective pixel pitches, obtained from the fit parameter  $a$ , are compatible with those obtained from the test beam measurement. The test beam measurement yielded  $a_x^{\text{meas, 1px}} = (14.96 \pm 0.11) \mu\text{m}$  and  $a_y^{\text{meas, 1px}} = (14.89 \pm 0.13) \mu\text{m}$  compared to the simulation that yielded  $a_x^{\text{sim, 1px}} = (14.99 \pm 0.08) \mu\text{m}$  and  $a_y^{\text{sim, 1px}} = (14.89 \pm 0.10) \mu\text{m}$ . A difference between the simulation and the measurement can be observed in the standard deviation of the Gaussian term. The standard deviation obtained from fits of the measured spatial residuals during the test beam are  $\sigma_x^{\text{meas, 1px}} = (5.50 \pm 0.22) \mu\text{m}$  and  $\sigma_y^{\text{meas, 1px}} = (5.85 \pm 0.28) \mu\text{m}$ , while the fit of the simulation yielded  $\sigma_x^{\text{meas, 1px}} = (4.08 \pm 0.15) \mu\text{m}$  and  $\sigma_y^{\text{meas, 1px}} = (4.88 \pm 0.19) \mu\text{m}$ . The simulation yields smaller standard deviations that are close to the contributions of the finite track intersection uncertainty, i.e. the telescope resolution, to the residual distribution. The unbiased mean track intersection uncertainties are taken to be same as during the measurement, following the reasoning and estimation of the errors in Section 8.3.2. The unbiased mean track intersection is  $\sigma_x^{\text{tel}} = \sigma_y^{\text{tel}} = (3.8000 \pm 0.0003 \pm 0.2000) \mu\text{m}$ . A smaller noise contribution for the simulation is consistent with the expectation that the simple simulation does not take all effects into account that influence the residual distribution. While, the simulation includes threshold smearing and an added electronics noise, there are no systematic uncertainties or systematic differences in the behavior between the simulated pixels. Additionally, the position and rotation of all sensor planes in the simulation is known and the alignment of the DUT sen-

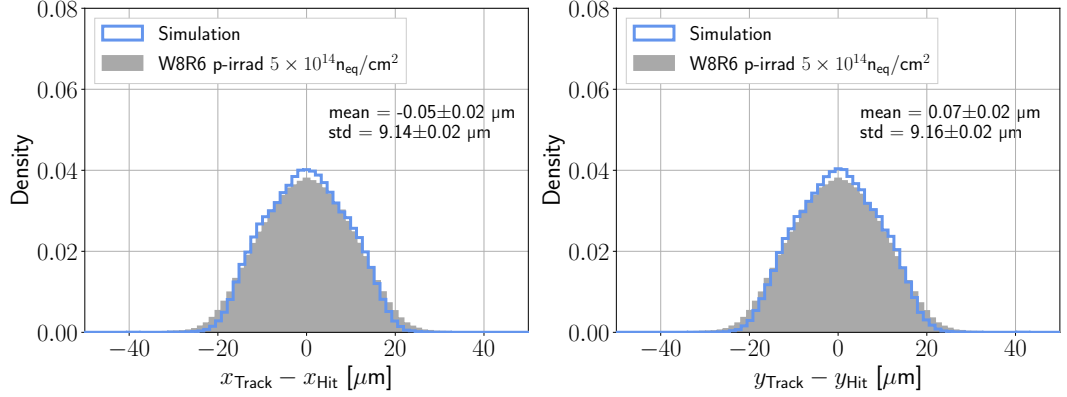


Figure 8.22: Residual distributions from simulation. The statistical mean and standard deviation are drawn into the plot for the simulation. The measured residuals are indicated for comparison. Left: Residual distribution along the  $x$ -axis. Right: Residual distribution along the  $y$ -axis. This result was obtained with TBSW.

| Cluster size    | Resolution in $x$           | Resolution in $y$           |
|-----------------|-----------------------------|-----------------------------|
| 1-pixel cluster | $(8.7 \pm 0.1) \mu\text{m}$ | $(8.7 \pm 0.1) \mu\text{m}$ |
| 2-pixel cluster | $(4.2 \pm 0.2) \mu\text{m}$ | $(4.3 \pm 0.2) \mu\text{m}$ |
| All clusters    | $(8.3 \pm 0.1) \mu\text{m}$ | $(8.3 \pm 0.1) \mu\text{m}$ |

Table 8.5: Spatial resolution achieved in simulation for 1-pixel, 2-pixel and all clusters. The simulation is performed with respect to TJ-Monopix2 sample W8R6.

sor is exact and no uncertainties of misalignments contribute to the simulation. A misalignment of the DUT can immediately affect the the width of the residual distribution of the measurement. Overall, the spatial resolution obtained from simulations is  $\approx 1 \mu\text{m}$  smaller than what was achieved in measurements at the test beam (see Table 8.4).

The same observations can be made for the residual distributions of the 2-pixel clusters. The effective pixel pitch is compatible between the simulation and the measurement, where  $a_x^{\text{meas}, 2\text{px}} = (0.74 \pm 0.01) \mu\text{m}$  and  $a_y^{\text{meas}, 2\text{px}} = (0.74 \pm 0.01) \mu\text{m}$  were obtained. The fitted standard deviation of the Gaussian is similar between the 1-pixel and 2-pixel clusters which is expected and also seen for the test beam measurements. In the simulation, the spatial resolutions along  $x$  and  $y$  are compatible, similar to the measurement, which is expected in both cases given the symmetry of the TJ-Monopix2.

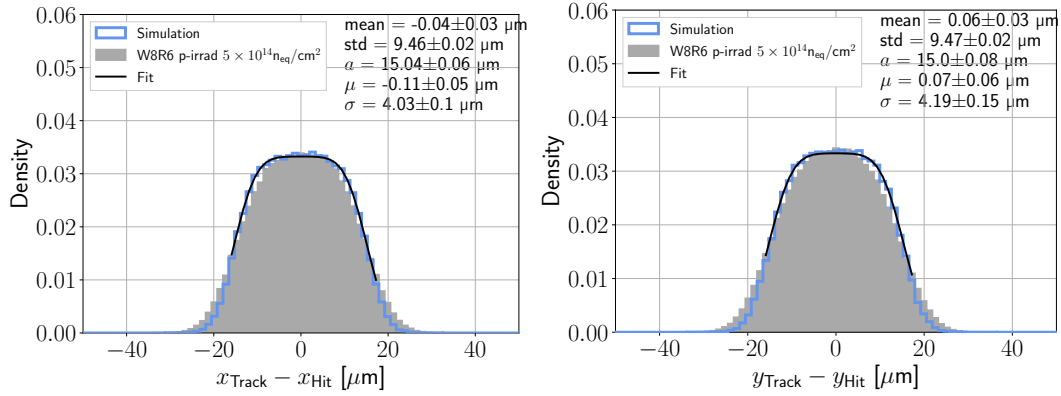


Figure 8.23: Residual distributions for 1-pixel clusters from simulation. A fit of a centered box function convoluted with a Gaussian has been applied to the distribution from simulation. The statistical mean and standard deviation, as well as the fit parameters, are drawn into the plot. The measured residuals are indicated for comparison. Left: Residual distribution along the  $x$ -axis. Right: Residual distribution along the  $y$ -axis. This result was obtained with TBSW.

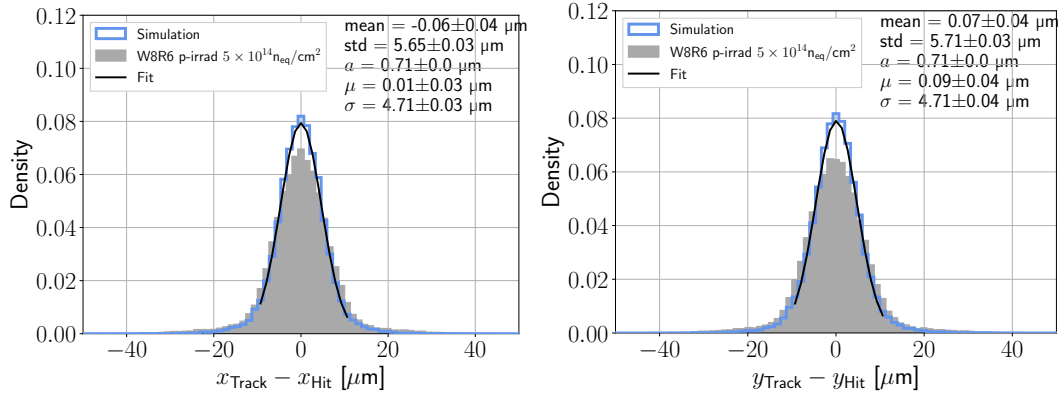


Figure 8.24: Residual distributions for 2-pixel clusters from simulation. A fit of a centered box function convoluted with a Gaussian has been applied to the distribution from simulation. The statistical mean and standard deviation, as well as the fit parameters, are drawn into the plot. The measured residuals are indicated for comparison. Left: Residual distribution along the  $x$ -axis. Right: Residual distribution along the  $y$ -axis. This result was obtained with TBSW.

#### 8.4.4 Simulated Hit Detection Efficiency

The hit detection efficiency is obtained in a similar manner as the measurement. After the tracking is performed, the tracks are matched to hits on the DUT. The same Corryvreckan analysis is applied to the measured and simulated data. The efficiency and the upper and lower errors are obtained as described in Section 8.3.3. There are no masked or noisy pixels in the simulation that would require treatment.

The efficiency in dependence of the threshold is simulated and compared to the measurement. The simulation is performed for a fluence of  $5 \times 10^{14} n_{eq}/cm^2$  and is thus compared to measurements of the TJ-Monopix2 chip W8R6. The threshold for the measurements is determined by utilizing the charge injection circuit and probing the response of the front end. For the simulation the threshold is an input value. While the threshold is varied, the other input values concerning the sensor, that are the chip temperature at  $\approx 40^\circ C$  and bias voltage at 20 V, are chosen according to the performed measurements. The measurements aggregate runs over multiple temperature points and the mean of all temperatures is used for the simulation and most measurements were performed at  $40^\circ C$ . The direct comparison of simulation and measurement is shown in Figure 8.25.

At small thresholds of  $\approx 200 e^-$ , the measured and simulated efficiency is similar and at 99.9 %. When increasing the threshold towards  $300 e^-$  the measured efficiency is in a range of 99.8 % and 99.5 %. The simulation remains at higher efficiencies very close to 100 %. Beyond  $300 e^-$  the measured efficiency exhibits a steady decrease until it reaches an efficiency of 98.6 % at a threshold of  $461 e^-$ . The efficiency of the simulation starts to decline at a  $400 e^-$  threshold and reaches the lowest measured efficiency at  $\approx 700 e^-$ . The efficiency of the simulation quickly declines after  $650 e^-$ . The simulation yields higher efficiencies at higher thresholds than the measurement and does not match the measured behavior. The reason for this difference is found in the difference of the measured and simulated seed charges in the range of  $200 e^-$  to  $1000 e^-$ . A visually adapted comparison from Section 8.4.2 is displayed in Figure 8.26.

When applying a threshold, all pixel hits with a charge that is smaller than the threshold will not be registered. Since the seed charge is the largest single-pixel charge in the cluster, no pixel will register a hit, once the threshold surpasses the seed charge. The fraction of seed charges smaller than the threshold is equal to the inefficiency of the sensor, assuming there are no other sources of inefficiencies. Since the difference of the efficiencies to 100 % at small thresholds for both measurement and simulation is much smaller than the inefficiency introduced by higher thresholds, this is assumed for the sake of this discussion. The fraction of seed charges below  $320 e^-$ , which is equivalent to the first bin, is small for both measurement and simulation. The impact on the efficiency is observed to be small for thresholds of that size. Between charges of  $320 e^-$  and  $950 e^-$ , the second and third bin, the fraction of seed charges is much higher for the measurement. The resulting inefficiency of the measurement is larger. The measurement only provides data until a threshold of  $461 e^-$ . The distribution of the simulated seed charges begins to have significant

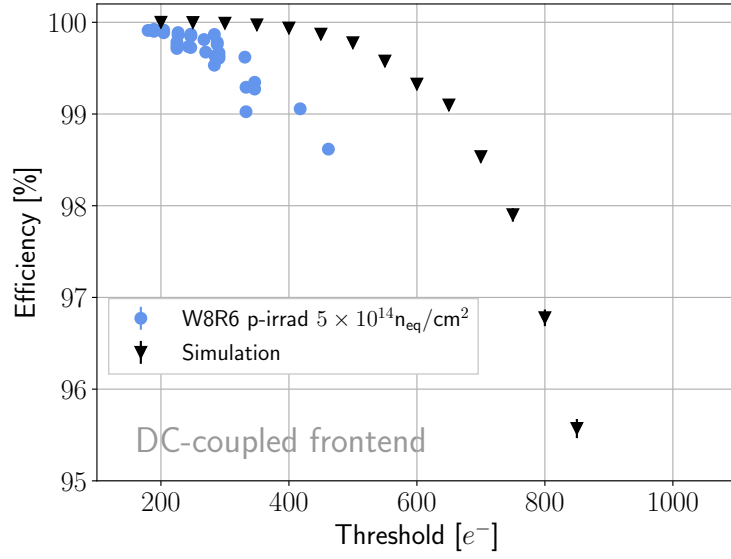


Figure 8.25: Hit detection efficiency in dependence of the threshold. The measurement is performed for the DC-coupled front end of TJ-Monopix2 sample W8R6. The parameters of the simulation are chosen with respect to the measurements. The error bars indicate the statistical error of the efficiency.

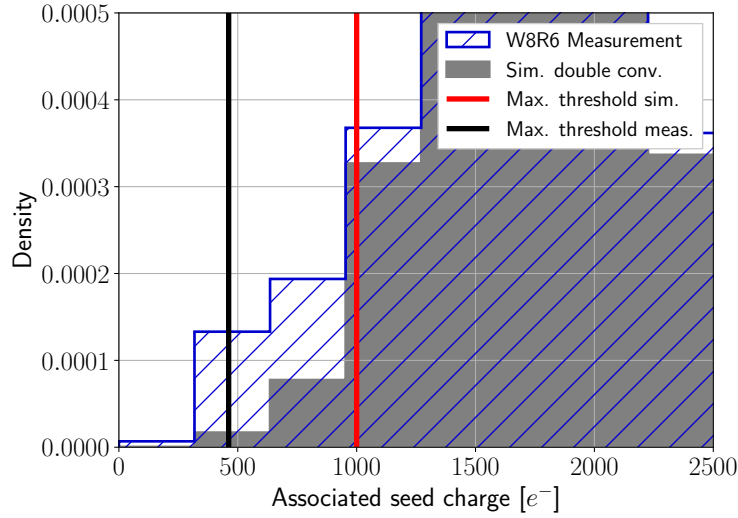


Figure 8.26: Comparison of the seed charge distribution for clusters associated to reconstructed tracks for simulation and measurement. Both distributions are normalized to an area of 1. This is a visually modified version of Figure 8.21, for details refer to Section 8.4.2. The region of the distribution relevant for effects relating to the threshold is magnified. The maximum threshold applied to TJ-Monopix2 chip W8R6 during measurements and for the simulation are indicated by vertical lines.

contributions between  $630\ e^-$  and  $950\ e^-$ , the third bin. Applying thresholds in this range leads to quickly increasing inefficiencies, which is observed.

The simulation of the hit detection efficiency suffers from the wrong simulation of small signal charges in the range of applied thresholds. The distribution of the collected charge is overall narrower (see Figure 8.21) and underestimates the fraction of small charges. Following the arguments in Section 8.4.2, this will be an effect of the missing front end simulation and the simplified electric field. This simulation will serve as a baseline for future expansions to the simulation that aim at improving the observed and discussed issues.



---

## BGNet: Beam Background Monitoring

---

BGNet is a neural network for predicting the beam background decomposition of hit rates for Belle II sub-detectors. The beam background is generated by beam particles lost near the IP. Multiple processes lead to these losses and contribute to the total measured hit rate of Belle II sub-detectors. A decomposition of the beam background can recover the individual contributions to the total measured hit rate.

Additional beam background hit rates overlay the hits from particles of signal processes, affecting the event reconstruction for physics analyses. Further, the irradiation of detector and functional components from beam backgrounds set limits on their lifetimes. In recent operation, when pushing for higher luminosities, the hit rate of the CDC reached the limits that are set to ensure safe operation. The record luminosity of  $5.1 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$  was reached, while Belle II was turned off, as large beam backgrounds did not permit safe operation, and no physics data was collected. BGNet is intended to be used as a diagnostic monitoring tool to provide insight for beam background control and mitigation.

BGNet is trained using 1 Hz time series from the EPICS slow control system [68–70]. It is implemented using TensorFlow [100] and Keras [101] machine learning framework. The source code of BGNet is publicly available and open source [102]. A paper was previously published that describes BGNet [103]. The author of this work is an author to the published BGNet paper, contributing text to the original draft, contributing to the BGNet source code and conducting tests of BGNet presented in the paper. Some parts of the paper will be summarized here for the benefit of a complete description. In addition many features and changes that were developed after the publishing of the paper will be presented and discussed. The author is part of the development team of BGNet, conducted beam background analyses with BGNet for the collaboration, performed performance analyses of BGNet and developed tools to make BGNet predictions available to the collaboration. The author

contributed to and conducted large parts of the BGNet commissioning procedure in the collaboration. Parts of the following review of BGNet were presented during a dedicated BGNet seminar that was organized during a collaboration meeting at KEK.

In the following, the basics of deep neural networks and the beam backgrounds relevant for Belle II are introduced. Then heuristic formulae and beam background measurements at Belle II are discussed. A description of the structure of BGNet is given. The real-time implementation, extension to combine multiple sub-detectors in one model, additionally predict the beam loss (and therefore the beam lifetime) is presented. Finally, the performance of BGNet is evaluated by comparing to other methods and the results are discussed.

Rather than using the internal naming conventions of variables, the naming conventions have been altered in this work and in the code of BGNet for the benefit of readability and recognizability. The naming scheme can be found in the software repository [102].

## 9.1 Machine Learning - Deep Neural Networks

The objective of supervised machine learning is to approximate some function  $f : X \rightarrow Y$  from a set of inputs  $X$  to a set of outputs  $Y$ , using an annotated data set. An annotated data set is made of  $n$  pairs of elements from  $X$  and  $Y$ :

$$((x_1, y_1), \dots, (x_n, y_n)) \in (X \times Y)^n, \quad (9.1)$$

connecting inputs with the expected outputs. The neural network provides an approximation  $f'(x, \vec{w}) = y$ , where  $\vec{w}$  are parameters chosen during the training to obtain the best approximation [104]. The weights  $\vec{w}$  are chosen such that a user-defined loss function is minimized.

This section will focus on the implementation of  $f'$  using feed forward neural networks (NN). A NN consists of layers of neurons that are connected by edges. The initial layer is the input layer, where each neuron represents an input. The last layer is the output layer, where the number of neurons defines the dimensionality of the output. The intermediate layers between the input and output layers are referred to as hidden layers. A NN with multiple hidden layers is called a deep neural network (DNN). Layers of neurons that are fully connected will be referred to as dense layers in the following sections. Feedforward networks propagate values only in one direction, from the input layer to the output layer, without providing any feedback connections [104]. There are also no connections between layers that are not directly adjacent. The structure of a feedforward neural network with dense layers, similar to the neural networks used in BGNet, is visualized in Figure 9.1.

The value  $x_i^n$  of the  $n^{th}$  neuron in a layer  $i$  is dependent on all connected neurons

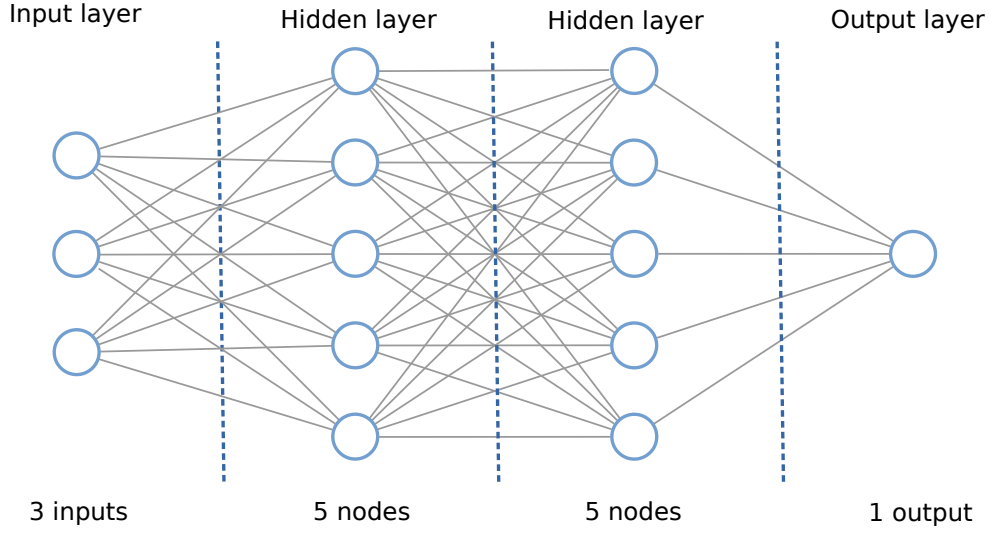


Figure 9.1: Generic layout of a deep neural network with an input layer, two hidden layers and an output layer with a single node.

$k$  in the preceding  $i - 1$  layer as

$$x_i^n = f_i \left( \sum_k w_i^{nk} x_{i-1}^k + b_i^n \right), \quad (9.2)$$

where  $w_i^{nk}$  is the weight that connects the neurons and  $b_i^n$  is a bias. The weights and biases are model parameters that are initialized with random values and the final parameters are defined by the model training. The activation function  $f_i$  modifies the output of a neuron. For binary classification tasks, the last layer would have a single neuron with a *sigmoid* function as activation function. The individual neural networks in BGNet have a single neuron in the output layer with a *softplus* activation function. In hidden layers, neurons often have non-linear *ReLU* (rectified linear unit) or *tanh* (hyperbolic tangent) activation functions. The non-linearity of the activation functions allows neural network models to better model non-linear behavior, making them more expressive, i.e. increasing the range of functions and patterns that can be approximated. BGNet uses the *tanh* activation function in its hidden layers. The mentioned activation functions are visualized in Figure 9.2.

During a supervised training the expected output, or truth value, is known and is used to define a loss function that defines a metric for the model performance during training. There are several loss functions that can be used. The loss function used in BGNet is the mean absolute error (MAE):

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|, \quad (9.3)$$

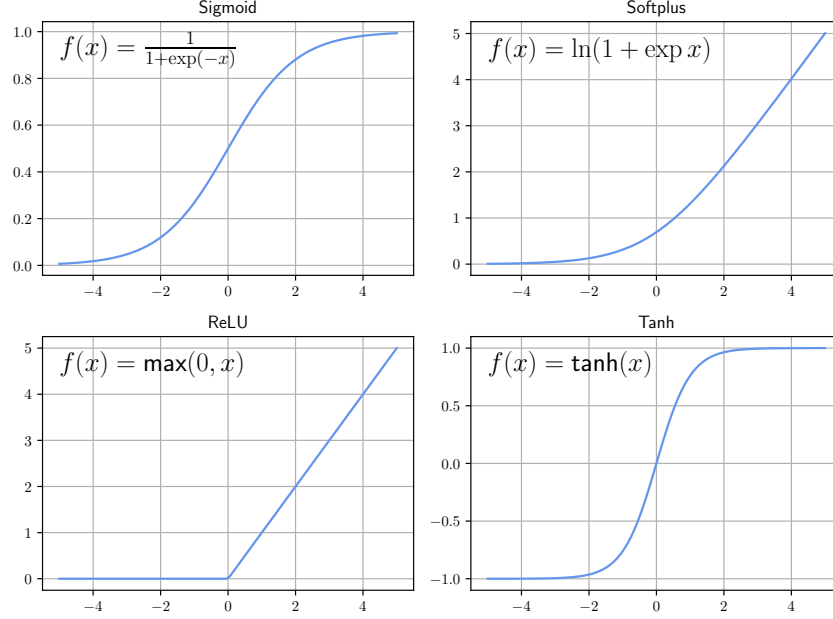


Figure 9.2: Examples for activation functions. The formula of each function is indicated in the respective plot. The tanh and softplus activation functions are used in BGNet.

where  $y_i$  are the previously defined expected outputs of the annotated data set and  $\hat{y}_i$  are the NN predictions of the corresponding input  $x_i$ . The neural network training is an iterative process, where each iteration is an epoch. At each epoch the weights and biases are updated via a back propagation scheme. During back propagation the gradients of the the loss function at each layer are computed, starting from the output layer towards the input layer [104]. The parameters of the model are then updated depending on the specific optimizing algorithm, like stochastic gradient descent [105] or Adam (Adaptive Moment Estimation) [106].

## 9.2 Beam Backgrounds

Beam-background processes are all processes that lead to particle losses near, or at the IP. Since these processes contribute to the hit rate of Belle II sub-detectors the physics performance of Belle II is negatively impacted. The radiation dose can also induce radiation damage in sensors (see Section 4.4). The beam backgrounds are expected and observed to increase at larger instantaneous luminosities and can hinder further increases of the instantaneous luminosity. The beam background has been studied during the Phase 1 commissioning [107], Phase 2 commissioning [55] and Phase 3 commissioning [108]. The major sources of beam background stem from the injection of particles into the main ring, the injection background, from

storing the particles in the main ring, the storage background with beam-gas and Touschek components and from colliding particles, the luminosity background. In the following, the considered backgrounds are explained in detail.

*Injection Background.* The electrons (positrons) for the HER (LER) are injected with their nominal energy coming from the LINAC to top-up existing bunches by merging with them using the betatron injection scheme, as outlined in Section 5.2. During the merging, particles are lost due to large betatron oscillations. Once the amplitudes of particles in the transverse plane are too large, they will collide with the beam pipe. As a result, the particles are lost and the energy is converted to an electromagnetic shower (see Section 4.2). The showers are generated along the entire ring, and if generated close to the IP the shower particles contribute to the beam-background. The bunch charge lost during the first 100 turns in SuperKEKB after injection can be up to 40% of the injected charge at the time of the initial injection. At an injection frequency of 50 Hz, elevated background level appear every 20 ms, with each spike lasting for about 4 ms. The substructure of these spikes coincides with the passing of the noisy, topped-up bunch of Belle II every  $\sim 10 \mu\text{s}$ . The mitigation strategies lie in tuning the injection parameters and the LINAC parameters to increase the injection efficiency as well as vetoing the Belle II data taking, whenever the noisy bunch passes.

*Beam-gas background.* While being stored, the electrons and positrons scatter with residual gas particles in the beam pipes. A visualization is displayed in Figure 9.3. There are four possible processes when a beam particle scatters:

- Elastic scattering with gas nuclei: The beam particle receives a kick, deflecting it in the transverse plane. The beam particle is lost if it exceeds the acceptance of aperture, i.e. collision with a collimator or beam pipe. These losses happen predominantly in the vertical plane
- Elastic scattering with hull electron: The beam particle loses energy. If the lost energy fraction is large enough it can fall out of the RF acceptance and the particle is lost.
- Bremsstrahlung induced by nuclei: In the presence of the nucleus's magnetic field, the beam particle emits a photon, reducing its energy. If the lost energy fraction is large enough, it can fall out of the RF acceptance and the particle is lost.
- Inelastic scattering with hull electron: The beam particle loses energy due to excitation and/or emission of photons as a result of the scattering. If the lost energy fraction is large enough, it can fall out of the RF acceptance and the particle is lost.

The scattering rate depends on the average pressure inside the center of the beam pipes. The pressure is measured using cold cathode gauges (CCGs) which are installed at every section of both rings. The typical pressure inside the beam pipe is

$\sim 10$  nPa ( $\sim 20$  nPa) for the HER (LER) without beam, and is called base pressure  $p_0$ . The pressure is observed to increase with the beam current  $I$ . This additional pressure is called dynamic pressure  $I \times p_1$ . The measured pressure averaged over the ring  $\bar{P}_{\text{CCG}}$  is therefore described by these two components [108]

$$\bar{P}_{\text{CCG}} = I \times p_1 + p_0 \quad (9.4)$$

which can be rewritten to

$$\bar{P}_{\text{CCG}} = I(d\bar{P}/dI)_{\text{CCG}} + \bar{P}_{\text{CCG}}(I = 0) \quad (9.5)$$

The effective pressure at the center of the ring  $\bar{P}_{\text{eff}}$ , which is seen by the beam, is found to be more sensitive to the dynamic pressure due to geometric reasons and a finite vacuum conductance. It is found to be three times larger than at the edge of the beam pipe where the CCGs are measuring [108]. The effective pressure can then be expressed as

$$\bar{P}_{\text{eff}} = 3I(d\bar{P}/dI)_{\text{CCG}} + \bar{P}_{\text{CCG}}(I = 0) = 3\bar{P}_{\text{CCG}} - 2\bar{P}_{\text{CCG}}(I = 0). \quad (9.6)$$

To mitigate this storage background the pressure needs to be reduced. This can be achieved by baking the beam pipes, removing volatile materials from the inner surface of the beam pipe and conducting beam scrubbing. During beam scrubbing the beam is stored with a defocused optic causing many particles to “scrape” on the inner wall of the beam pipe to, similar to baking, cause gas molecules to desorb the inner beam pipe surface. Scattering particles tend to leave the design orbit in the vertical directions, thus vertical collimators can effectively reduce the beam-gas background.

*Touschek background.* During storage of the beam in the rings, particles of one bunch experience Coulomb scattering. A visualization is displayed in Figure 9.3. As opposed to multiple scattering processes in the particle bunch leading to an effective increase of the bunch size, Touschek scattering describes a single scattering process that leads to both scattering particles immediately being lost. After scattering both particles have energies deviating from the nominal energy, possibly leaving the design orbit despite the synchrotron’s phase stability. The particles will then collide with the beam pipe material or a collimator producing electromagnetic showers. Since the longitudinal momentum of the scattered particles is affected, the Touschek background is mainly suppressed by horizontal collimators near the IP.

*Luminosity-dependent backgrounds.* The luminosity-dependent background, or luminosity background, describes all physics processes from the intended beam collisions of electrons and positrons at the IP. The most dominant processes are radiative Bhabha scattering  $e^-e^+ \rightarrow e^-e^+\gamma$  and two photon processes  $e^-e^+ \rightarrow e^-e^+e^-e^+$ .

At the design luminosity, the luminosity-dependent background will be the dominant background source. It cannot be mitigated as the interaction cross-sections of the relevant processes are only dependent of the underlying physics and kinematics of the interaction. Detectors that are positioned closest to the interaction point are

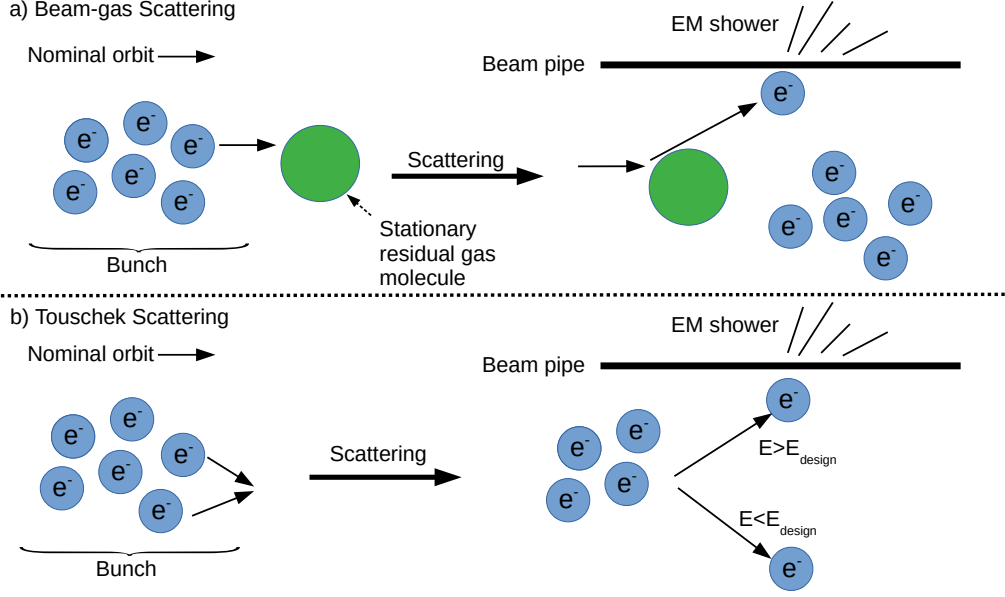


Figure 9.3: Illustration of the storage beam background processes for beam-gas (a) and Touschek (b).

more sensitive to the luminosity background. This is due to processes that produce particles with small transverse momenta that will only reach the innermost layers of the vertex detector. An example for this is the two photon process that typically produces an  $e^+e^-$  pair with small transverse momenta.

### 9.2.1 Background Prediction

Heuristic formulae for the background components have been formulated using relations known from beam dynamics and scattering theory [27, 109].

The heuristic function for the beam-gas background used in Belle II is

$$\mathcal{O}_{\text{beam-gas},i} = S_{\text{beam-gas},i} \times I_i \bar{P}_{\text{eff},i} = S_{\text{beam-gas},i} \times G_{\text{beam-gas},i}, \quad (9.7)$$

where  $\mathcal{O}_{\text{beam-gas},i}$  describes the background hit rate due to beam-gas scattering in a Belle II sub-detector,  $i$  denotes either LER or HER parameters.  $G_{\text{beam-gas},i}$  describes the heuristic term of the formula and  $S_{\text{beam-gas},i}$  is a generic sensitivity that can be obtained via fitting or using machine learning.

The units of the hit rates  $\mathcal{O}$  depend on the described sensor. The units of the variables contributing to the heuristic term are defined by their units in the EPICS slow control system and the sensitivities' units are chosen such that after r.h.s. of the equation has the same units as the described hit rate.

The heuristic function for the Touschek background is

$$\mathcal{O}_{\text{Touschek},i} = S_{\text{Touschek},i} \times \frac{I_i^2}{n_{b,i}\sigma_{x,i}\sigma_{y,i}\sigma_{s,i}} = S_{\text{Touschek},i} \times G_{\text{Touschek},i}, \quad (9.8)$$

where  $n_b$  is the number of bunches stored in a ring and  $\sigma_x\sigma_y\sigma_s$  is the bunch volume. The transverse bunch sizes  $\sigma_x$  and  $\sigma_y$  are measured continuously in SuperKEKB, but the longitudinal bunch length  $\sigma_s$  needs to be derived from other parameters. The parametrization is described in Reference [110] and the latest values, that are also used here, are given in Reference [108].

The heuristic function for the injection background is

$$\mathcal{O}_{\text{Injection},i} = S_{\text{Injection},i} \times G_{\text{Injection},i}. \quad (9.9)$$

The injection heuristics  $G_{\text{Injection},i}$  multiply the state of the respective injection gate (1 if open, 0 if closed) with the absolute of the injected charge. If the product is positive, the heuristic is set to 1, otherwise it is 0. This enforces a time-dependent structure and a vanishing injection background, when there are no injections happening. The heuristic term is an engineered feature that is used only in the context of BGNet. In previous beam background measurements the injection background was not explicitly modeled, which is discussed in more detail below. Since the perturbations of newly injected bunches happen on time scales much shorter than one second, transient effects of injections backgrounds can be disregarded when using the 1 Hz EPICS variables.

The heuristic function for the luminosity background is

$$\mathcal{O}_{\text{Lumi}} = S_{\text{Lumi}} \times \mathcal{L} = S_{\text{Lumi}} \times G_{\text{Lumi}}, \quad (9.10)$$

where  $\mathcal{L}$  is the instantaneous luminosity. A generic pedestal component is implemented using

$$\mathcal{O}_{\text{Pedestal}} = S_{\text{Pedestal}}. \quad (9.11)$$

Both  $S_{\text{Lumi}}$  and  $S_{\text{Pedestal}}$  are expected and observed to be stable over time for the considered Belle II sub-detectors. Once the various sensitivities are known, the hit rate decomposition is known. Determining the sensitivities is the main purpose of BGNet and beam background study campaigns that are presented in the following. Since  $S_{\text{Lumi}}$  and  $S_{\text{Pedestal}}$  are expected to be constant over time, they can be measured at any point during the accelerator operation, for example during beam background study campaigns, and can then be applied to physics data taking. All other sensitivities are generally dependent on the current state of SuperKEKB. The sensitivities for the storage backgrounds (LER and HER beam-gas and Touschek backgrounds) and the injection backgrounds are found to be sensitive to the collimator positions, that are constantly tuned and changed. For reasons of convenience and brevity, the indices of the sensitivities  $S$  and heuristic features  $G$  are abbreviated using numbers in figures and formulas. The numbering scheme is displayed in Table



| Background component | Index |
|----------------------|-------|
| HER Beam-gas         | 1     |
| LER Beam-gas         | 2     |
| HER Touschek         | 3     |
| LER Touschek         | 4     |
| HER Injection        | 5     |
| LER Injection        | 6     |
| Luminosity           | 7     |
| Pedestal             | 8     |

Table 9.1: Numbering scheme of background components considered by BGNet.

9.1. The total hit rate of a sub-detector is obtained by summing over all 8 hit rates of each component.

Previous approaches to measuring the sensitivities used dedicated background study campaigns to disentangle the background sources [55, 107, 108]. During these campaigns, special conditions in SuperKEKB are produced. The special condition refer to a specific state of SuperKEKB, where, for example, the collimators and optics parameters are set to the same specific values at each background study campaign. Initially, a beam is stored in only one ring and, after stopping the top-up injections, the beam decay is observed. During the beam decay, only the storage backgrounds of the respective ring can contribute to the measured hit rate. This single-beam decay is repeated with different stored beam currents and number of bunches to differentiate contributions by Touschek and beam-gas scattering. A least squares fit with the above heuristic formulae then allows for an estimation of the storage background sensitivities. The sensitivity of the luminosity contribution can be extracted by storing beams in both rings and observing the decay as they collide. Using the previously obtained sensitivities of the storage background components a second least squares fit can be performed to estimate the sensitivity to the luminosity background. The pedestal of the sub-detectors can be obtained from a dedicated time window without beams, where the measured hit rate of a sub-detector can be immediately identified as the pedestal. All variables needed to compute the heuristic formulae, as well as the sub-detector hit rates are available on the EPICS archiver appliance as 1 Hz time series (see section 6.8). Since there are no immediate models for the injection background, the injection background is explicitly excluded from the measurements and only beam decays are considered. The injection background was then defined as the difference of the measured hit rate and the determined hit rates of the other components [108]:

$$\mathcal{O}_{\text{Injection}} = \mathcal{O}_{\text{Measured}} - \mathcal{O}_{\text{Touschek}} - \mathcal{O}_{\text{Beam-gas}} - \mathcal{O}_{\text{Lumi}} - \mathcal{O}_{\text{Pedestal}}. \quad (9.12)$$

Two drawbacks of this method to determine the beam background have been identified.

Firstly, the special conditions that are used to disentangle the background sources have been kept constant since Phase 1 to provide better comparability of the background study campaigns allowing for consistent extrapolation of backgrounds and discover time-dependent long term deterioration effects. These conditions, however, diverged strongly from the conditions during physics runs, limiting its use to decrease backgrounds during data taking.

Secondly, the beam background study campaigns need considerable time, during which data taking is not possible. Thus, there are only one or two study campaigns in any given year and feedback to current operation can be slow.

Both of these points are addressed by BGNet. BGNet can produce sensitivity predictions for any SuperKEKB state during physics data taking in real-time, as will be discussed in the following. However, BGNet should not be viewed as a replacement, but as an additional tool provided to the collaboration. The traditional method of estimating the background sensitivities is well established and provides a baseline to compare results of BGNet and the extracted sensitivities are used to initialize BGNet with a set of sensible values. Further, the beam background study campaigns provide excellent data sets for the neural network training.

### 9.3 Structure of BGNet

BGNet aims to determine the background sensitivities through deep learning neural networks that learn a map from the space of collider features to the targeted sensitivities. A flow diagram is shown in Figure 9.4.

Initially, the time series for all variables needed as input or for the calculation of inputs for the neural networks and for the calculation of the heuristic features are requested from EPICS corresponding to the time  $t$ , for which the beam background is predicted. A prediction of the hit rate at time  $t$  only considers the values of the input variables at the same time  $t$ . Thus, we assume that the current beam background can be fully described by the current state of SuperKEKB and is not dependent on previous states.

The preprocessing provides an input vector with a length of  $m + 8$ , where  $m$  is the number of inputs (and 8 corresponds to heuristic features  $G_1$  through  $G_8$ ). The preprocessing also performs the calculation of variables and heuristic features, selection of suitable data, a delay correction and scaling of all inputs and heuristic features.

Suitable data is selected by requiring that Belle II and SuperKEKB are in nominal data taking conditions and in well-defined states. The explicit technical requirements are given in Appendix B.3.

An empirical delay was observed between the target variables of Belle II and the accelerator input variables. The delay is, at least partially, caused by the fact that Belle II and SuperKEKB are running separate instances of EPICS. To make Belle II variables available in the SuperKEKB EPICS network, a gateway is used that publishes the Belle II variables again in the other EPICS instance with a delay,

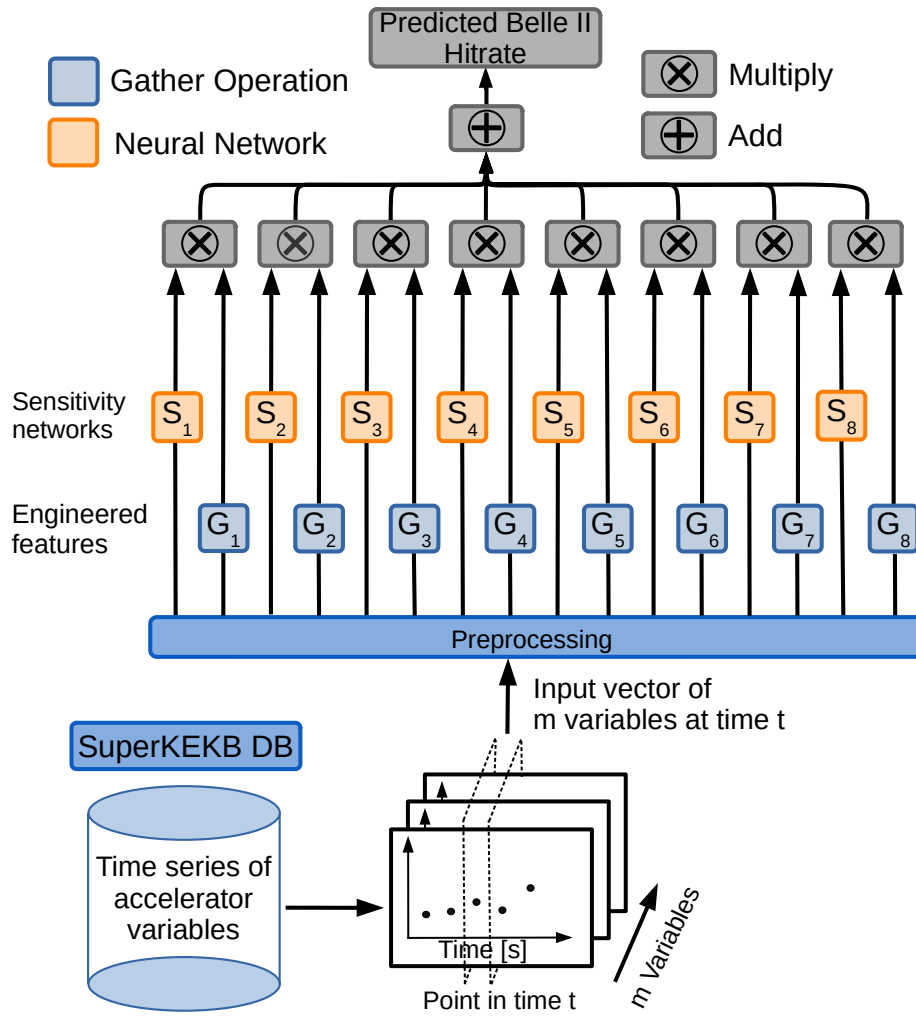


Figure 9.4: Flow diagram of BGNet. The output of the neural network is the total predicted hit rate at time  $t$  of a Belle II sub-detector.

without maintaining the original timestamp. The delay is parametrized by the shift  $d = i + w$  in units of seconds with the integer part  $i$  and the non-integer part  $w$ . The input vector  $x'$  to make a prediction at time  $t$  is then a weighted mean of the input vector  $x$  at times  $t - i$  and  $t - (i + 1)$  with the weight  $w$  [103]

$$x'(t) = (1 - w) \cdot x(t - i) + w \cdot x(t - (i + 1)). \quad (9.13)$$

The change of a variable between  $t$  and  $t + 1$  is therefore approximated as being linear. The delay varies for each sub-detector and is observed to be constant during run periods.

The last step of the preprocessing is a scaling of the input variables, heuristic features and the target hit rate. This is done to mitigate instabilities and slow training convergence due to the vastly different scales of numerical values between the variables. The input variables are individually scaled to a mean of 0 and a standard deviation of 1. The resulting distributions are clipped at values  $-5$  and  $+5$ . The heuristic features and the target hit rate are not being centered but scaled by dividing with the 95 % quantile. These scalings are calculated during the training of the model and are bundled with the model, so that data sets for predictions can be scaled in a similar manner to the training data.

One extra step of the preprocessing that only applies to the training procedure pertains to the under- and oversampling of data. The training data can be split into decay and top-up data, where there are either no injections or active injections, respectively. The data set is dominated by data with active injections with a fraction of typically  $> 90\%$ . However, decay data, especially single beam decays, are important to provide the neural network a means to disentangle the storage backgrounds. From the difference of decay data and injection data, the injection background components can be identified. To balance the data set, a random oversampling [111] is applied to the decay data to achieve an equal split of injection and decay data.

To start the neural network training, the individual data sets need to be prepared. The neural networks  $S_1$  to  $S_6$  have their own list of input variables<sup>1</sup>. The selection of input variables is based on known dependencies from beam dynamics, experience and expert opinion. The list of inputs is intentionally kept small to avoid overtraining and keep the complexity of the models low. The number of inputs into the injection background neural networks  $S_5$  and  $S_6$  are larger due to the higher complexity of describing this background component and the heuristic features that only account for the time-dependent structure and not its amplitude. The explicit list of input variables are listed in Appendix B.1.

A detailed flow diagram is shown in Figure 9.5. The respective data columns are taken from the preprocessed input vector and fed into the fully connected feed-forward networks. The networks  $S_7$  and  $S_8$  are different in that the sensitivity to the luminosity background and the detector pedestal are assumed to be constant and

---

<sup>1</sup>Depending on the context the exact meaning of the variables  $S_i$  changes. Here, the  $S_i$  refer to the neural networks as a whole, while in the previous section they referred to a one-dimensional number, that is the sensitivity and the output of the equally named neural network.

| Name            | Description                         | Unit                 |
|-----------------|-------------------------------------|----------------------|
| BP_BW_325       | Diamond radiation monitor dose rate | $\text{mrad s}^{-1}$ |
| BP_BW_215       | Diamond radiation monitor dose rate | $\text{mrad s}^{-1}$ |
| BP_FW_215       | Diamond radiation monitor dose rate | $\text{mrad s}^{-1}$ |
| BP_FW_325       | Diamond radiation monitor dose rate | $\text{mrad s}^{-1}$ |
| SVD_FW_60       | Diamond radiation monitor dose rate | $\text{mrad s}^{-1}$ |
| SVD_FW_120      | Diamond radiation monitor dose rate | $\text{mrad s}^{-1}$ |
| SVD_FW_180      | Diamond radiation monitor dose rate | $\text{mrad s}^{-1}$ |
| SVD_FW_240      | Diamond radiation monitor dose rate | $\text{mrad s}^{-1}$ |
| SVD_FW_300      | Diamond radiation monitor dose rate | $\text{mrad s}^{-1}$ |
| CDC_CUR_AVERAGE | CDC average layer leakage current*  | $\mu\text{A}$        |

Table 9.2: Currently supported target variables. The names of the diamond radiation monitors indicate their positions, compare with Figure 6.6. \*: What is here referred to as leakage current of the CDC is in the context of gaseous detectors also referred to as chamber current.

| Hyperparameter                 | Value                                                           |
|--------------------------------|-----------------------------------------------------------------|
| Number of inputs               | $S_1 : 9 \ S_2 : 8 \ S_3 : 15 \ S_4 : 12 \ S_5 : 34 \ S_6 : 27$ |
| Number of hidden layers        | $S_1$ to $S_4 : 2$ and $S_5$ to $S_6 : 2$                       |
| Number of hidden nodes         | $S_1$ to $S_4 : 8$ and $S_5$ to $S_6 : 32$                      |
| Loss                           | Mean absolute error                                             |
| Batch size                     | 32                                                              |
| Optimizer                      | Adam [106]                                                      |
| Weight and bias initialization | Glorot uniform [112]                                            |
| Early stopping patience        | 5                                                               |

Table 9.3: Hyperparameters of BGNet.

independent of the state of the accelerator. Both of these neural networks have a single constant input of  $x_7 = x_8 = 1$  and only consist of a single linear neuron. Each sensitivity network has one neuron in the output layer. It uses a softplus activation function to guarantee a positive output. The output of each neural network is then multiplied with the respective heuristic feature to obtain the background specific hit rate. The 8 contributions are summed to obtain the total hit rate. With the measured target hit rate, the absolute mean error is calculated that acts as the NN loss. The supported target variables for BGNet are listed in Table 9.2.

A table with the hyperparameters of BGNet is shown in Table 9.3

### 9.3.1 Multi-sensor BGNet

The beam background experienced by sub-detectors depends on their position in Belle II. One such example is the  $\phi$ -profile of the beam background. BGNet can

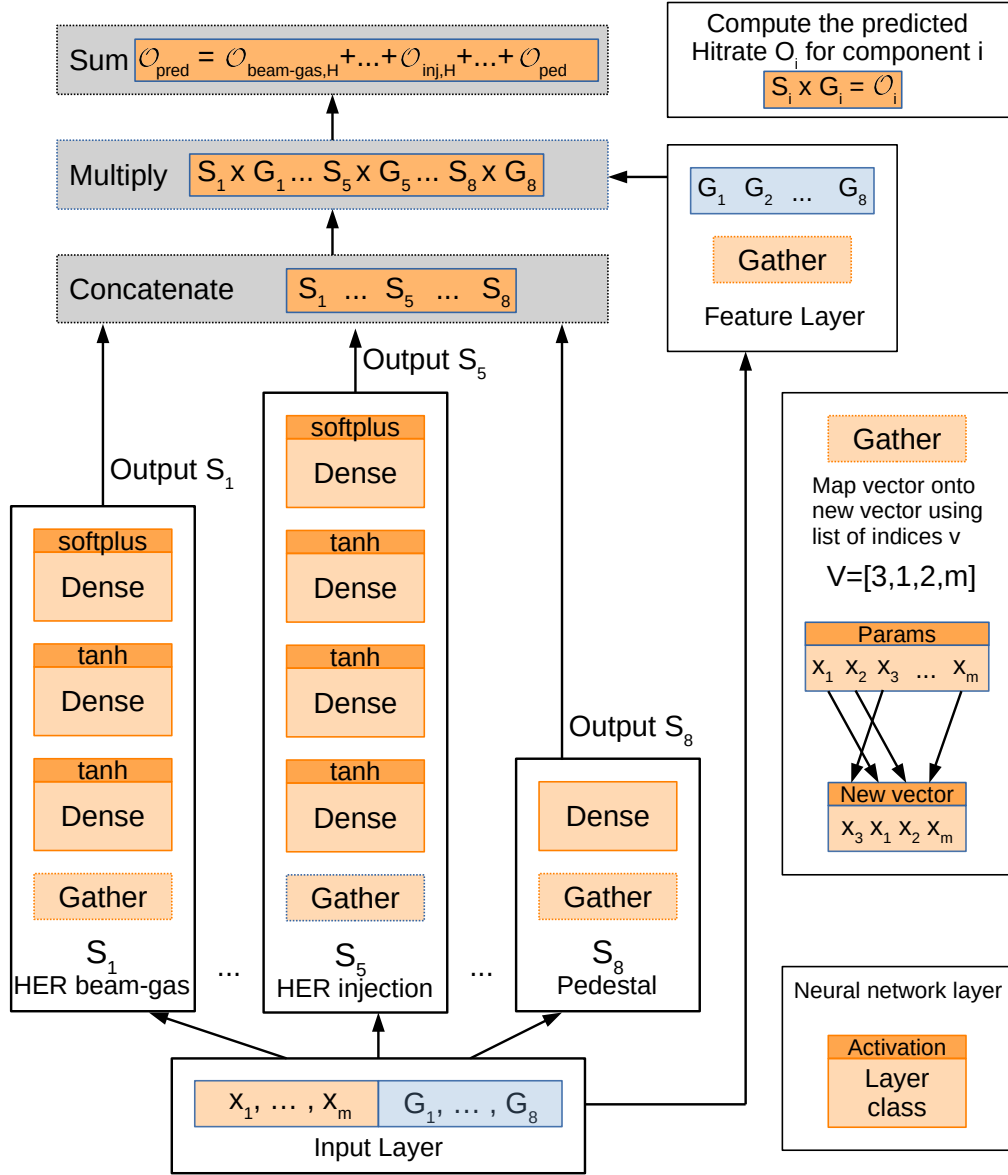


Figure 9.5: Flow diagram for BGNet applied on hit rates with explicit details on the structure of the neural networks and the operations performed. A dense layer is defined as a fully connected layer where each node is connected to every node in the previous layer. The gather operation maps a vector onto a new vector using a list of indices.

be used to make the  $\phi$ -profile of each component visible by predicting the beam background of multiple sensors with different positions in  $\phi$ . The natural choice for this were the diamond radiation monitors. In particular, the group of forward SVD diamonds was chosen, which features sensors at  $\phi = 0^\circ, 60^\circ, 120^\circ, 180^\circ, 240^\circ, 300^\circ$  (compare with Figure 6.6, showcasing the positions of all diamond radiation monitors).

The measured hit rates due to a background component for all chosen diamond sensors are related by a simple scaling factor due to the  $\phi$ -profile. This relation is exploited to avoid training a model for each diamond sensor and instead train a model that produces predictions for all diamonds simultaneously. This is realized by adding additional neurons in the output layer equal to the number of sensors described by the model. All output nodes share the preceding neural network layers. A given sensitivity network, for example HER Touschek, is now used for predictions for all sensors that are included. The different sensitivities of each sensor to a respective background source is accounted only by the weights that connect the output node. The outputs of all sensitivity networks  $S_1$  to  $S_8$  corresponding to one sensor are then collected, multiplied with the engineered heuristics  $G_1$  to  $G_8$  and summed up to obtain the total predicted hit rate. This concurrent training of multiple sensors then needs to be reflected in the loss, which is defined to be sum of all mean absolute errors between the individual measured and predicted hit rates. The input features to the sensitivity networks and their structure do not change for the multi-sensor scheme. A detailed flow diagram for BGNet with the multi-sensor scheme is shown in Figure 9.6. Results of the beam background decomposition obtained for several diamond dose rates with a multi-sensor BGNet model are displayed in Appendix B.4. The  $\phi$ -profiles are introduced in detail in Section 9.8.5 for the performance evaluation of BGNet.

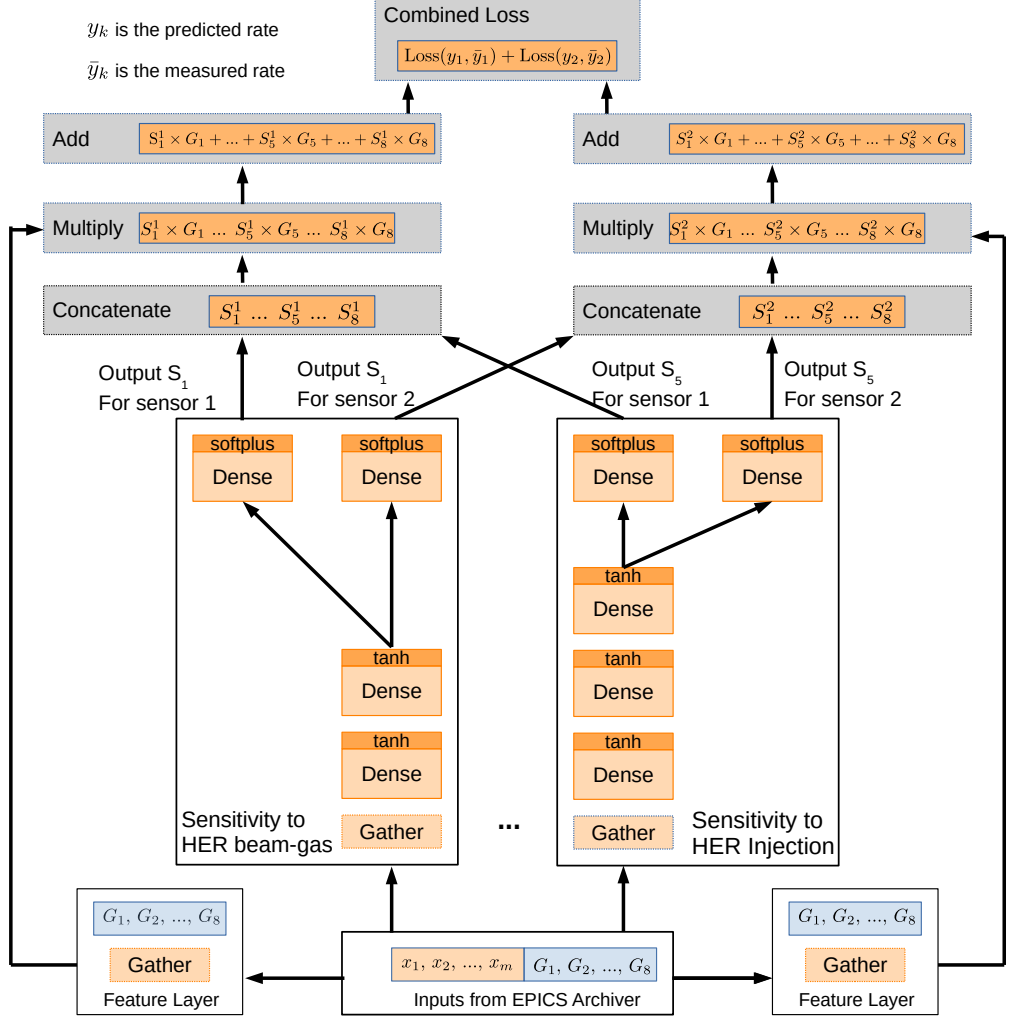


Figure 9.6: Flow diagram for BGNet with the multi-sensor scheme. Explicit details on the structure of the neural networks and the operations performed are indicated. A dense layer is defined as a fully connected layer where each node is connected to every node in the previous layer. The gather operation maps a vector onto a new vector using a list of indices.



## 9.4 Explainability of BGNet Models

The knowledge of the decomposition of the hit rates can already provide a basis for possible actionable steps to reduce the hit rate. For example a large contribution of Touschek beam backgrounds could prompt operators to prioritize closing horizontal collimators. However, the question remains if one could gain additional insight from the trained model, once it can correctly identify which variables drive the changes in hit rates. The capability of the model to correctly identify such variables and their impact on the hit rate can be tested in extrapolation studies as conducted in Section 9.8.3.

Feature attribution methods are used to quantify the impact of a change of an input variable on the neural network output. In other words, it provides a means to understand the relation between input and output of neural networks. Such understanding is crucial for all computer programs and stands in direct opposition to the stance that machine learning algorithms are perceived or treated as “black boxes”.

Integrated gradients [113] are used as a feature attribution method. They can immediately be applied on the individual networks  $S_i$  since they are real-valued and differentiable. The integrated gradient for the  $j^{\text{th}}$  input variable of the neural network for the  $i^{\text{th}}$  sensitivity is computed using [113]:

$$\phi_{ij} = (x_j - x_j^*) \times \int_{\alpha=0}^1 \frac{\partial S_i(x^* + \alpha(x - x^*))}{\partial x_j} d\alpha \quad (9.14)$$

where  $x$  is the test sample and  $x^*$  is a reference sample. The attributions are measured from the difference of output at points  $x$  and  $x^*$ . Here, the  $S_i$  refer to the mathematical functions that represent the neural network  $S_i : \mathbb{R}^{m_i+8} \rightarrow \mathbb{R}$  with  $m_i + 8$  being the length of the input vector. The partial derivative in Equation 9.14 is the gradient of the sensitivity along the  $j^{\text{th}}$  dimension. The integral represents a path integral of the gradients along the straightline path between the points  $x$  and  $x^*$ . The integrated gradients can also immediately be applied to the product of a sensitivity and a heuristic feature by replacing the  $S_i$  with  $S_i^* = S_i \times G_i$ . The advantage is that the resulting attribution values are in units of the hit rate that is the target variable of BGNet for all beam background components  $i$ . Another advantage is that the variables in the heuristic formulae are now also part of feature attribution. Using the  $S_i^*$  to perform the feature attribution is the standard use case and will be used for all examples of feature attributions in this work. The feature attribution can also be applied on the sum of  $S_i^*$ , i.e. the total predicted hit rate, to obtain the most important variables on the total prediction.

The interpretation of the attribution values is directly linked to the property of completeness. It refers to the fact that the sum of all attributions, for a given beam background component  $i$ , are equal to the difference in output for the inputs  $x$  and

$x^*$ :

$$S_i^*(x) - S_i^*(x^*) = \sum_j \phi_{ij}. \quad (9.15)$$

Since the  $S_i^*$  are used, the difference is in units of the hit rate. As a result, the attribution values for all variables  $j$  are directly comparable for a set of  $x$  and  $x^*$ . At the same time, the relative size of the attribution value for a variable  $j$  compared to the difference of output values directly indicates the importance of the variable. However, when calculating the attributions for two different test samples  $x$  and  $x'$  against the same reference sample  $x^*$ , the two resulting attribution values for the same variable  $j$  may not be directly compared as the scale of the attributions is defined by the difference in output, which is generally different for two test samples. The measure used to quantify the importance of a variable is the mean absolute attribution value over a sufficiently large set of test samples (that are compared against the same reference sample).

The size of the set of test samples should be of the scale of expected time scales to find systematic behavior between changes in input variables and the change in predicted hit rates. In practice, these changes happen on the scale of minutes, and with each test sample corresponding to a second in the 1 Hz time series, the set of test samples should contain  $> 100$  data points to be able to resolve systematic behavior. Additionally, the feature value of the variable  $j$  can be compared to the associated attribution value to test the sensibility of the attribution value and infer a systematic behavior. The choice of a reference sample is arbitrary but should represent a neutral prediction with respect to the test samples.

An example of applying the feature attributions will be discussed in the following. The feature attributions were applied on data of a beam background study campaign conducted on the 16.04.2024. The beam background study campaigns are introduced in more detail in Section 9.2.1 and the results of this campaign will be used to make a comparison with the predictions of BGNet in Section 9.8.2. During this background study campaign, the background behaved unexpectedly due to some issue in the machine. To find the cause of this issue, a BGNet model was trained on the data collected during the beam background study campaign and feature attributions were applied. The unexplained background behavior was especially large during an HER single beam decay. A single beam decay refers to a decay of beam of the HER (LER), while there is no beam in the LER (HER). The single beam decays during the background study campaign are designed so that only the number of bunches changes, in order to disentangle the storage background components. The number of bunches cannot change during a beam decay, as the beam needs to be dumped. The measured and predicted total hit rate and the predicted beam background decomposition is displayed in Figure 9.7. The first decay starting at 14:40 and the third decay starting at 16:10 behave normally. The second decay starting at 15:15 shows unexpected behavior. As the beam current decreases, the beam background is also expected to decrease according to the heuristic formulae for

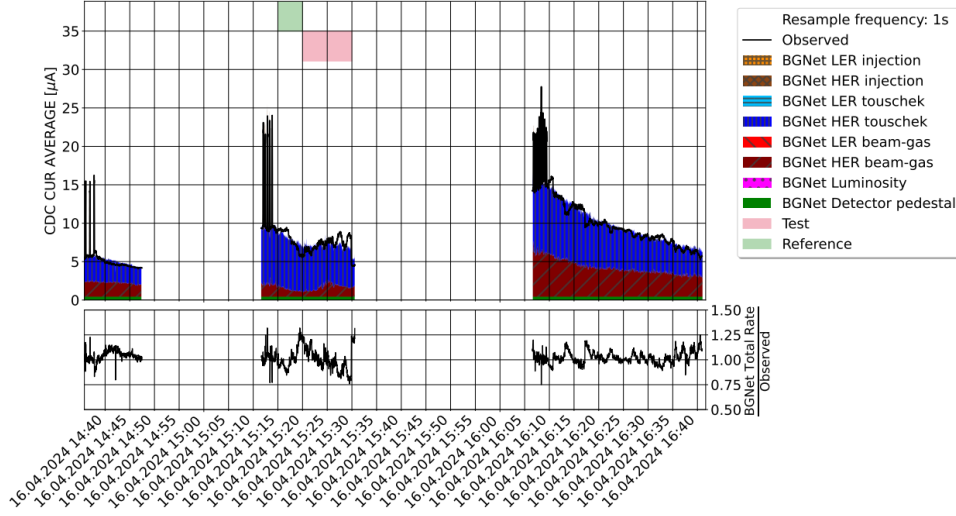


Figure 9.7: The measured and predicted total hit rate and the predicted beam background decomposition during HER single beam decays. The HER single beam decays were conducted during a beam background study campaign on 16.04.2024.

the storage background components. Instead, the beam background first falls rapidly for about 5 min until it begins to rise over the following 10 min. Over the course of the decay, the beam current steadily decreased from 250 mA down to 150 mA. This can be seen in Figure 9.8. The rise of the hit rate and the beam background can only be explained by an unexpected change of some parameter, which should explicitly be prevented during a beam background study campaign. Since the model followed the unexpected change in beam background, although not perfectly, the feature attribution method will be applied in an effort to find the cause of the increasing beam background.

The samples selected for the test and baseline, or reference, samples for the feature attribution are from time windows during the second HER single beam decay, that shows the unexpected behavior. The baseline sample is chosen to be the part that shows the decrease of the hit rate, while the test samples are from the part, where the hit rate increases. The time windows are indicated in Figure 9.7 and Figure 9.8. The attribution values for the most important input variables  $x_j$  for all samples in the test data set are displayed in Figure 9.9. The input variables are ranked by the largest mean absolute attribution values over the full test data set. The highest ranked input variable is the HER beam current with a known dependence of the hit rate. The decrease of the beam current shapes the overall beam decay and a high feature attribution value is expected. The development of the current was shown in Figure 9.8 and the mean attribution value is displayed on top. However, there is no indication that the current caused the unexpected behavior of the beam background. The second highest ranked variable is the HER vertical beam size. The time series of the HER vertical beam size compared to the measured hit rate is shown in Figure

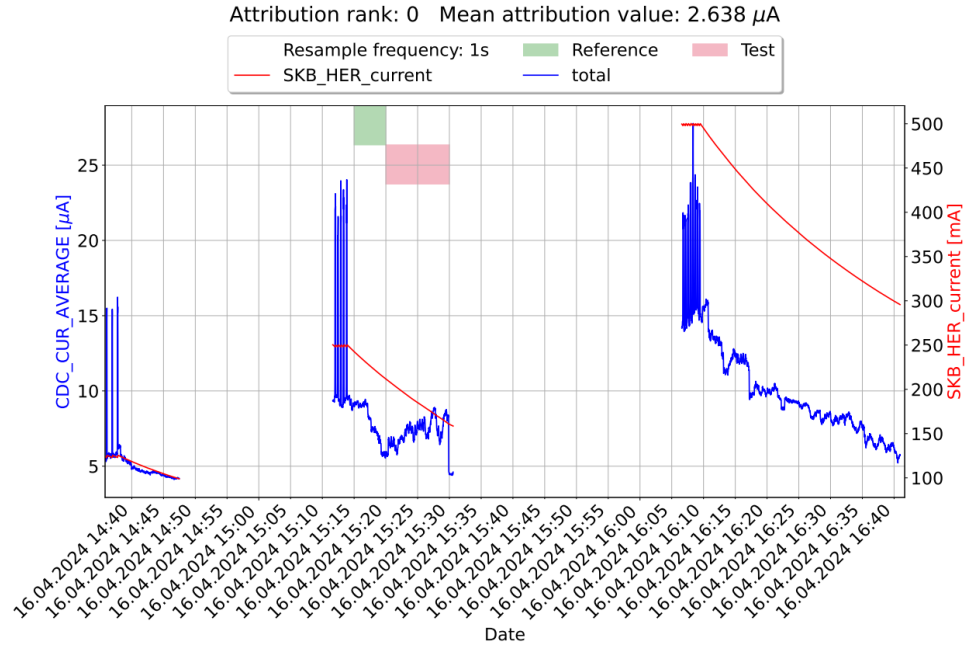


Figure 9.8: The total measured CDC average layer leakage current and the HER beam current. The measured hit rate of the CDC should decrease as the beam backgrounds should decrease as the HER beam current dissipates. The mean attribution value in units of the hit rate is indicated in the top of the plot. The attribution value refers to the neural network prediction and not the measured hit rate.

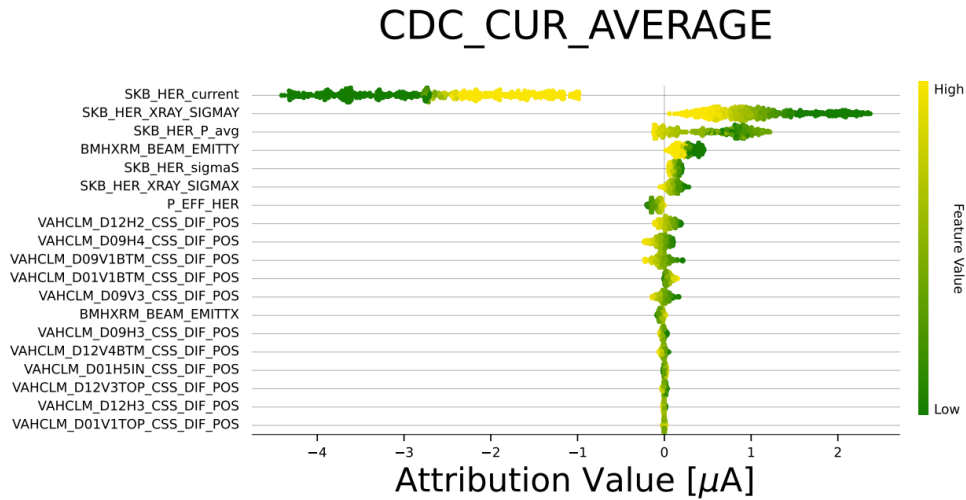


Figure 9.9: Attribution values of input variables during an HER single beam decay. The attribution values are visualized as point clouds, and the density of points is reflected by the vertical width of the point clouds. Each variable has points equal to the number of samples in the test data set. The color of each point corresponds to the relative feature value compared to all feature values present in all test samples. The input variables are ranked by their mean absolute attribution values.

9.10.

The beam size should stay approximately constant during a beam decay as in the beam decays before and after. However, it is observed to decrease rapidly from a very large value at the start of the beam decay. The beam size has a known impact on the beam background, as it is part of the heuristic for the HER Touschek background component. The impact of the decreasing beam size increased the HER Touschek background according to  $\mathcal{O}_{\text{Touschek, HER}} \propto 1/\sigma_y$ , despite the falling currents quadratic dependence on the current  $\mathcal{O}_{\text{Touschek, HER}} \propto I_{\text{HER}}^2$ . The predictions of BGNet depend on the heuristic formula to describe the impact of the beam size onto the HER Touschek background. However, it is possible that the heuristic formula is not applicable in this form during drastic transient behavior as seen from the HER vertical beam size. This could explain why the BGNet predictions were performing poorly during that event. In the following days, a larger investigation by the collaboration revealed that the bunch-by-bunch feedback system did not operate properly [27, 59]. This led to vertical tune shifts and a blow-up of the HER vertical beam size right before the beam decay. The HER vertical beam size then recovered and decreased rapidly during the beam decay. Although the root cause was not identified by BGNet, the feature attribution method was able to identify the immediate cause for the unexpected behavior of the beam background. Such a beam background analysis with BGNet using the feature attribution methods can be automatically produced within a minute using the BGNet chat bot, that is presented

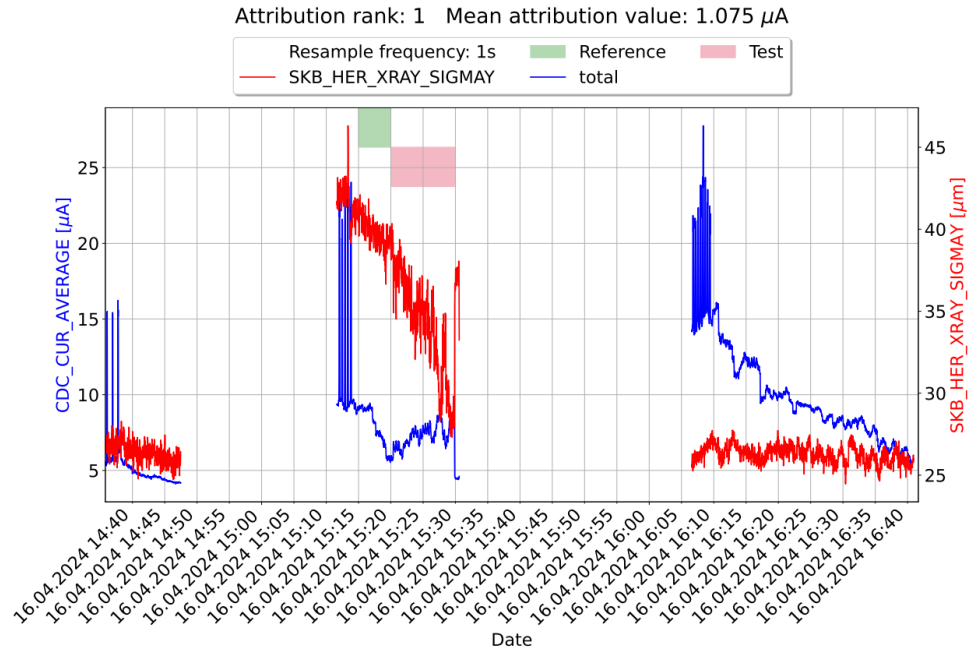


Figure 9.10: The total measured CDC average layer leakage current and the HER vertical beam size. A large change of the HER vertical beam size can be observed in the middle beam decay, while it is approximately constant during the decays before and after. The mean attribution value in units of the hit rate is indicated in the top of the plot. The attribution value refers to the neural network prediction and not the measured hit rate.

in Section 9.7.

The goal of introducing feature attributions is to provide actionable steps for operators by highlighting the most important variables. However, even when BGNet created a good surrogate model of SuperKEKB that can correctly predict the beam background decomposition, the model is a simplification. Changes in backgrounds produced by effects not described by BGNet, because the relevant variables are not part of the model and may be attributed to variables that are merely correlated or were affected by the root cause. This is not inherently bad for the performance of hit rate and decomposition predictions of BGNet and can even be intended. An example for this is the emittance of the beam. It is an input into several sub neural networks of the hit rate models. The emittance is directly related to the beam size as shown in Equation 3.25. A change of the emittance is likely to cause a change in the background decomposition. However, the emittance is a measured beam parameter and not a set parameter that can immediately be controlled or changed. It is a quantity that is sensitive to many SuperKEKB and LINAC sub-systems, settings and parameters. When a the change in backgrounds is attributed to the emittance, it does not represent an actionable step and likely does not identify the root cause. Thus, it is necessary that the results of feature attributions are interpreted by an expert. However, quickly providing the information that the hit rates have been increasing proportional to an increase in the beam emittance can be valuable by itself. Additionally, to improve the believability of the feature attributions, direct comparisons of identified variables and the measured beam background are provided, possibly showing a direct dependence between two measured quantities.

## 9.5 Beam Loss Predictions

An important parameter to increase the beam currents of SuperKEKB is the beam loss  $dI/dt$ . It describes the reduction of the beam currents due to the loss of particles. The primary mechanism of the particle loss are the beam backgrounds. Here, the beam loss will be only calculated while there are no active injections. The beam loss is then only sensitive to losses from storage backgrounds and injection backgrounds do not contribute. In contrast to the beam background contributions to the hit rate, the beam loss is sensitive to beam backgrounds along the entire ring, instead of only the particle losses close to the IP. Here, it is computed as the difference quotient  $\Delta I/\Delta t$  using the 1 Hz time series of a beam current  $I$ .

The current development in a ring can be separated into two scenarios. In the first scenario, there are no injections and the beam current steadily decreases due to particle loss of storage and luminosity backgrounds. However, the amount of beam loss due to collisions, i.e. luminosity-related beam loss, is small compared to the contributions of the storage backgrounds and is neglected. In the second scenario, there are active injections where particles are brought into the ring and the current increases. During the latter, the beam current increases, while particles are lost due to storage and luminosity backgrounds and, additionally, particles are lost due to

injection background. The ambiguities that arise between the particle losses due to injection inefficiencies and the particle losses due to storage backgrounds are difficult to resolve. Thus, only the beam loss due to storage backgrounds shall be described here and the difference quotient is only computed during times with no active particle injections.

BGNet has been applied to predict the total beam loss and its decomposition into the Touschek and beam-gas components. The beam loss  $L_i$  is computed as the difference quotient and expressed in terms of the storage background components:

$$L_i = \frac{\Delta I_i}{\Delta t} = S_{\text{beam-gas}, i} \times I_i \bar{P}_{\text{eff}, i} + S_{\text{Touschek}, i} \times \frac{I_i^2}{n_{b,i} \sigma_{x,i} \sigma_{y,i} \sigma_{z,i}}. \quad (9.16)$$

The beam-gas and Touschek components have the same form that is used in the hit rate decomposition that was introduced before, where  $i$  is either LER or HER,  $n_b$  the number of bunches,  $\sigma_x$ ,  $\sigma_y$ ,  $\sigma_z$  are the bunch sizes and  $\bar{P}_{\text{eff}, i}$  is the effective average pressure in a given ring. The beam losses of a HER and LER are independent of each other. The structure of BGNet for the prediction of the beam loss is visualized in Figure 9.11. The input variables for the Touschek and beam-gas sub neural networks for the LER and HER models are listed in Appendix B.2. The preprocessing performed on the data for the beam loss models is similar to the preprocessing performed for the hit rate models. The differences arise from the fact that the beam loss models are not dependent on data from Belle II. Thus, no delay correction between slow control variables is necessary and training data selection criteria in Appendix B.3 that are related to Belle II variables are omitted. Instead, an additional selection of training data that is outside of time windows with active injections is applied. The beam loss and its decomposition during the beam background study campaign is displayed in Figure 9.12. The beam loss can be expressed as a beam lifetime  $\tau_i$  via

$$\frac{I_i}{\tau_i} = -\frac{dI_i}{dt}. \quad (9.17)$$

The typical beam lifetimes during physics data taking runs with stable top-up injections for the LER beam is around 10 min and around 30 min for the HER beam. The feature attribution method can be applied to the beam loss models similar to the hit rate models.

The provided decomposition of the beam loss can be a helpful diagnostic tool to decrease the beam loss of both HER and LER. Decreasing the beam loss is one of the steps that can be taken to increase the stored beam currents to achieve higher instantaneous luminosities. However, this represents only one of the available lever arms to increase the currents, and other steps, like the amount of injected charge and the injection efficiency, need to be considered. Also, the current increase will lead to increased beam backgrounds that need to be mitigated at the same time to allow for an efficient and safe data taking of Belle II. The BGNet models for the hit rate decomposition and the beam loss decomposition can be used in tandem to monitor multiple aspects of the SuperKEKB and Belle II operation at the same time.



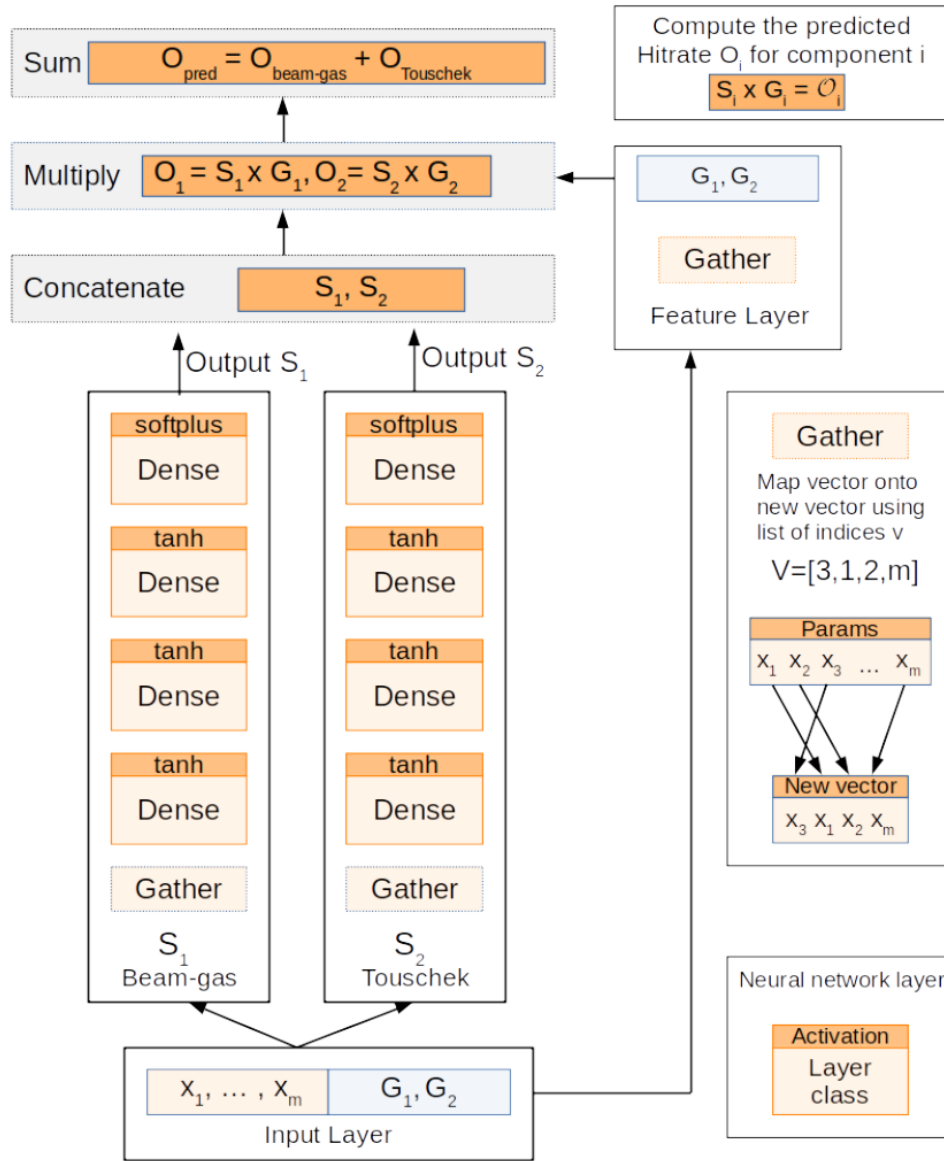


Figure 9.11: Flow diagram for BGNet applied on beam losses with explicit details on the structure of the neural networks and the operations performed. A dense layer is defined as a fully connected layer where each node is connected to every node in the previous layer. The gather operation maps a vector onto a new vector using a list of indices.

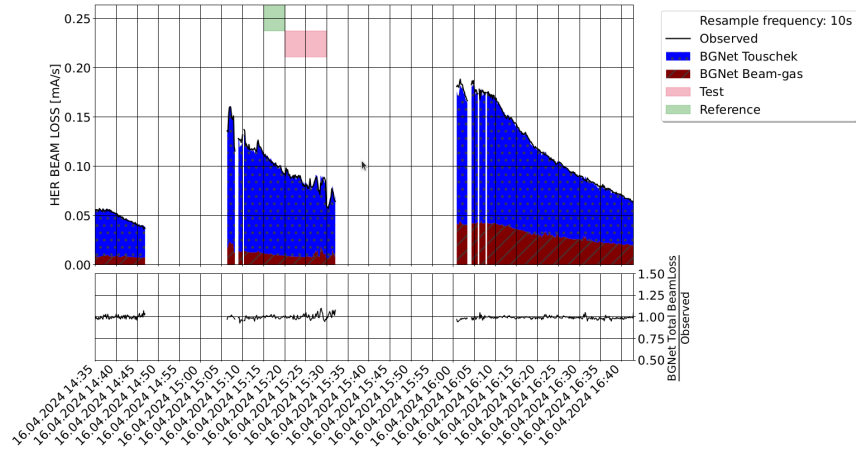


Figure 9.12: The measured and predicted beam loss of the HER beam current and the predicted beam background decomposition during HER single beam decays. The HER single beam decays were conducted during a beam background study campaign on 16.04.2024

The performance of the beam loss predictions is analyzed in Section 9.8.6.

## 9.6 Real-time Predictions

Previously, BGNet models were trained for specific purposes and an offline analysis was performed to investigate the beam background decomposition, usually triggered by unexplained increases in sub-detector hit rates. These analyses could only be performed by a few experts with a turnaround time of at least one day. To make BGNet results more accessible and faster, a real-time application was developed by the author. It allows to produce beam background predictions and decompositions based on live EPICS data and publish the data back into EPICS with a latency  $< 1$  s. The data provided by BGNet therefore becomes available in real-time. In this mode of operating BGNet, the models are required to extrapolate and infer the hit rate decomposition from data not seen during training. New BGNet models are trained frequently to adapt the models to the current conditions of SuperKEKB. This aspect of the BGNet performance is evaluated in Section 9.8.3.

Two requirements have had major impact on the design philosophy of the real-time application of BGNet (RT-BGNet). The first is to fulfill the initial premise of providing BGNet predictions in real-time at a frequency of 1 Hz. This matches the standard update frequency in the Belle II EPICS slow control application. The latency of each iteration must be  $< 1$  s, otherwise the time difference between the measured and predicted data grows with time, resulting in a growing time shift that is difficult to reconstruct or data points are skipped resulting in data loss and ambiguities arise for matching measured and predicted values. The second factor in the design of RT-BGNet is the EPICS system itself [69]. It does not guarantee the availability of data or correct transmission of data. While not all caveats of EPICS and its data handling shall be explored here, its use of the User Datagram Protocol (UDP) should serve as an example to illustrate its impact on the design of RT-BGNet. UDP is a connectionless network protocol that sends data packets into the network without being aware of any recipient and without any handling of data loss, best described by a “fire-and-forget” approach. This avoids expensive connections and handshakes, avoids related timeouts and slow devices, which is especially important in large networks with many thousand devices for each data transfer, as is the case for Belle II and SuperKEKB. The requirements for speed and resilience against incomplete data have shaped the design of RT-BGNet, resulting in a highly robust and parallelized application.

RT-BGNet is realized in python3 [114] and relies on the *threading* library that is part of the python standard library. The PCASpy [115] and PyEpics [116] python APIs are used to interface with EPICS. A schematic of the functional parts in the network and their connections is displayed in Figure 9.13. The overall layout of RT-BGNet and its functions is displayed in Figure 9.14. Both figures share a common color scheme in order to relate the common parts shown.

The application starts by establishing channel access to all needed PVs. The list of needed PVs is equal to the list of all input variables for all neural networks  $S_1$  to  $S_8$ . Establishing connections is relatively time consuming and therefore the con-

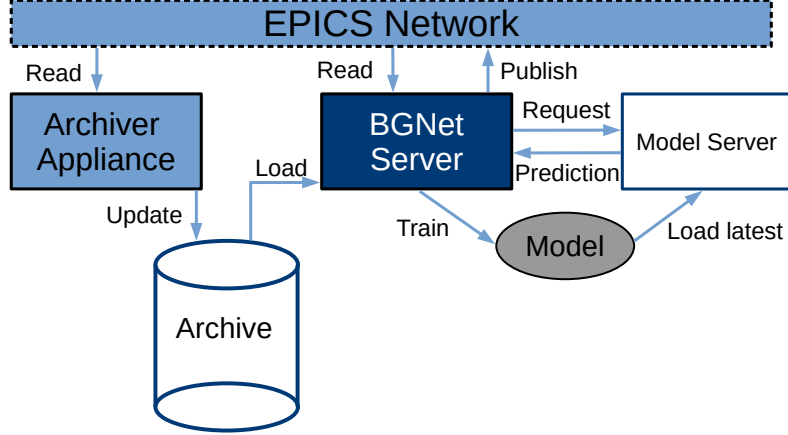


Figure 9.13: Diagram of all functional parts in the network, in which the BGNet server is located. The RT-BGNet application is run on the BGNet server. The BGNet server is a dedicated machine for BGNet related tasks.

nections are kept alive, as long as the application is running. Then, the PV values are read from EPICS, updating the values of all PVs. The update is performed once every second for all PVs at the same time  $t_0$  by pulling the current value. A vector  $x_{t_0}$  is constructed from all values at the current time  $t_0$ . The vector enters a FIFO queue with a fixed length of 7. As long as the length of the queue is smaller than 7, vectors are added to the FIFO queue. The queue is motivated by the delay correction explained in Section 9.3, which needs multiple adjacent vectors. At a length of 7, the FIFO queue holds the vectors  $x_{t_0}, x_{t_0+1}, \dots, x_{t_0+6}$ , i.e. the state of all variables that are input to BGNet for the last seven seconds. Once the queue is full, at some time  $t$ , the current and the previous states of the past 6 s are handed to a separate, concurrent thread that handles the data preprocessing. For the preprocessing, multiple data points are needed to accommodate the delay compensation and compute the delay corrected input vector  $x'(t)$  according to Equation 9.13. At that time the oldest vector in the queue is dropped and the vector of the following time step is added to the FIFO queue and the next input vector is formed for the next prediction.

Each thread handles one type of model: Hit rate and beam loss. The appropriate preprocessing is performed for each model. From the 7 vectors provided by the FIFO queue at time  $t + n$  an input sample  $x_n$  is constructed by the preprocessing. The preprocessing steps are outlined in Section 9.3. The obtained vector describes the (current) state of SuperKEKB and contains all variables needed for the inference of the respective model. The vector is serialized to a string and encoded using Base64 and brought into a JSON compatible form.

A HTTP request is sent to a TensorFlow model server [117] that is locally hosted in a Docker container [118]. The JSON compatible input data is attached to the HTTP request. The request also carries information about the type of model and

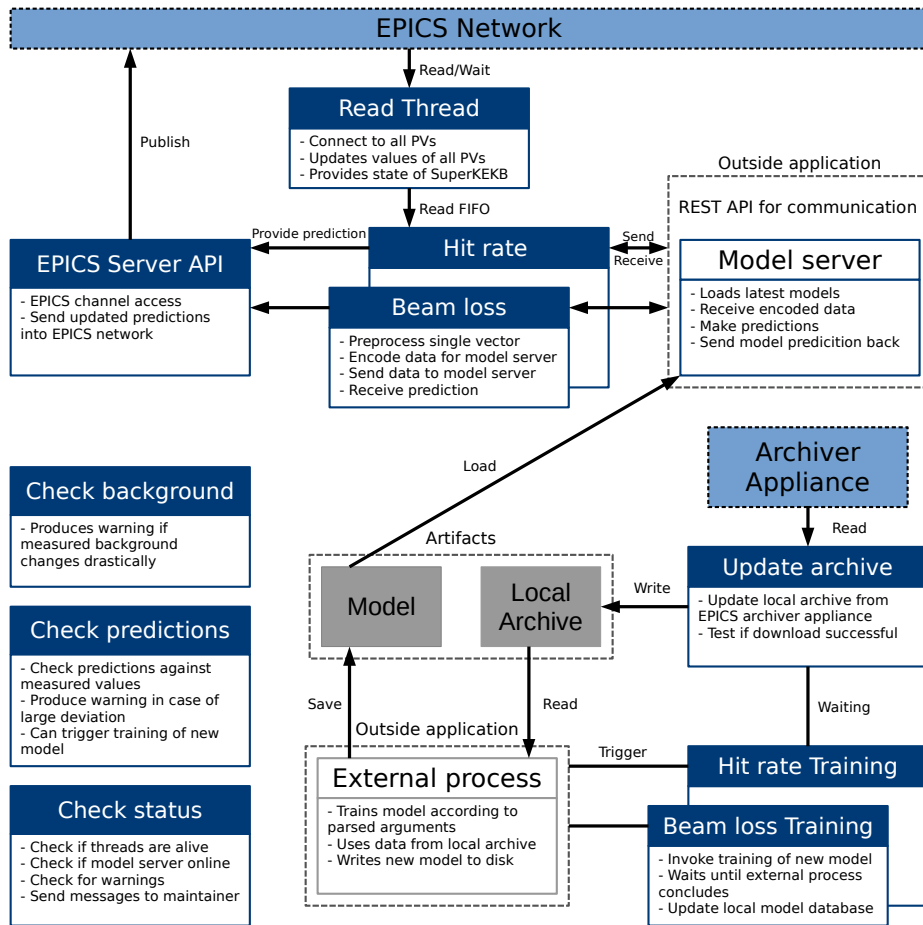


Figure 9.14: Diagram of the functions performed by RT-BGNet. Each box represents a functional part. The EPICS network and the archiver appliance are part of the KEK common infrastructure. The dark blue boxes are tasks performed by RT-BGNet on the BGNet server. Each task is implemented as thread. A thread is an independent sequence of operations that runs in the same context (address space) as all other threads of a process. The threads are operating concurrently and the process is RT-BGNet. The “Outside applications” are running locally on the BGNet server but are executed outside the context of RT-BGNet. They communicate with RT-BGNet via a REST API using Hypertext Transfer Protocol (HTTP) or artifacts written to the local hard drive. Connections between functional parts are drawn in case the communication is required for the operation to continue. For example, the “Update archive” thread has a connection to the “Archiver appliance” as it is required to wait for network to provide the requested data and the “Hit rate Training” and “Beam loss Training” threads need to wait until the updated training data is written to the local archive. If data is passed between functional parts, the direction of data flow is indicated by arrows.

what output is requested. The outputs are requested by calling a model signature, which is a computation graph attached to the model. This computation graph applies the proper scaling to the inputs as discussed in Section 9.3 and performs the inference of the model. The output of the inference can be chosen to be that of an intermediate layer of a BGNet model. To obtain the sensitivity values for instance, the output of the layer before multiplying with the heuristic features is chosen as output. To obtain the hit rates of the individual background components, the output of the layer that performs the multiplication of the sensitivities and the heuristic features, would be chosen. Finally, the obtained rates or sensitivities are scaled back to return physical units to the values. The computation graph for the attribution values also performs calculation of the integrated gradients. The obtained values are sent back from the model server to the RT-BGNet application.

The output is received by the thread that sent the initial request for inference. The thread, and only this thread, has waited for the duration of the inference. The received data does not need to be manipulated anymore and the values are published in EPICS , updating the state and value for the respective prediction in EPICS . Although the thread formally waits for the API to publish the values, this process concludes instantly with respect to the relevant time scale of 1 s. The thread returns to waiting for the read thread to again fill the FIFO queue with 7 entries for the preprocessing in the next iteration. If any issues were encountered during, e.g. due to missing data in EPICS , network issues or SuperKEKB/Belle II being offline, the thread immediately skips any steps left at that point and publishes generic zeros for all outputs of that thread. This concludes the main functionality of RT-BGNet, as BGNet real-time predictions have been served.

The training of new models is important to limit the length of extrapolation a model has to perform, as the performance of models generally degrades the longer the model extrapolates with respect to the latest training data. Detailed studies of extrapolation performance is given in Section 9.8.3.

BGNet models are always trained on local data that is stored in a number of distributed files, simply called *archive*, that was downloaded from the EPICS archiver appliance. The training data is frequently accessed and the bandwidth and time consumption of downloading everything on demand for every training easily justifies keeping a local copy of the relevant data. Thus, the first step of training a model is to update the locally available training data by requesting the latest data from the EPICS archiver appliance. Then the training is started in a separate shell outside the context of the python instance that is running RT-BGNet. The trainings of all 4 models (CDC hit rate, diamond radiation monitor hit rate, HER beam loss and LER beam loss) can be run concurrently on the BGNet server. The time required for each type of model is generally different, yet the different processes do not wait on each other. Once the model is written to disk and no errors are reported by the training process, an entry is made into the BGNet model data base (a CSV table) and is loaded by the model server and the new model is available for predictions. The model data base connects the model name (an integer timestamp of the time

when the model training concluded) with the type of model, the BGNet version it was trained with and the training interval. This is needed when performing offline analyses, as it is performed by the chat bot introduced in the following section. The training of new models immediately starts again once all model trainings from the previous training interval concluded.

Three independent threads are run in the context of RT-BGNet that perform checks and produce warnings.

The first thread checks the measured background and is in first order totally independent of any BGNet predictions. However, drastic changes in measured backgrounds usually warrant further investigation as to the cause. The performance of RT-BGNet should be closely monitored during these events. This functionality was largely superseded by the following thread and it is currently not used anymore.

The second thread checks the ratio of RT-BGNet total hit rate predictions and the measured hit rates for all predicted hit rates. Currently, the predictions are marked “bad” if  $0.75 > \mathcal{O}_{\text{pred}}/\mathcal{O}_{\text{meas}} > 1.25$ . If 80 % of all predicted total hit rates in the last 1200 s were marked bad a warning is produced. It is considered to replace these bad predictions with zeros, essentially masking bad performance. However, currently it is decided that experts can easily judge bad performance by themselves and this was not implemented. Instead, a warning is produced, with no immediate actions being taken, as performance typically recovers with a new model, that was trained on the latest data of SuperKEKB.

The third thread is a collection of small functions that are combined in this description. They perform small checks and functions like:

- Check the connection status to EPICS
- Check the connection status to all EPICS PVs
- Check the status of the BGNet server
- Check the status of all other threads, as they can silently and individually die, while all other threads continue to work, making it difficult to spot issues quickly
- Check the status of the locally hosted model server
- Check for warnings produced by all other threads, and in case of critical issues send messages via e-mail and a chat bot to the author until noticed

Every encountered failure mode discovered since the commissioning of this tool in 2023 lead to the implementation of a counter-measure or an extended logging with expressive warnings to make issues quickly visible. This is especially important to obtain a high availability of the RT-BGNet service in a an environment of continuous 24 h per day operation of Belle II and SuperKEKB.

Since most EPICS variables and related user-interfaces are updated with a frequency of 1 Hz the hit rate decomposition for all supported sensors, the related

sensitivities and attribution values, the beam life prediction and decomposition and attribution values for both rings are available in real-time. A selection of important variables produced by BGNet are also saved by the archiver appliance. This has the added benefit that BGNet integrates into all related EPICS user-interfaces like the web-interface of the EPICS archiver appliance [70] and Control System Studio [119] that are established in the Belle II and SuperKEKB community.

## 9.7 BGNet Chat Bot

A chat bot was developed, that allows to make a request to BGNet and then provides a PDF file with automatically generated plots to the user. This fully automates the BGNet offline analyses with an average generation time of the analysis of less than a minute. The usage of the chat bot requires no setup, technical knowledge or access to the internal network.

The Belle II collaboration uses a common chat platform [120] for internal communication. The BGNet chat bot is available in a BGNet channel called “BGNet Bot”. It can be utilized to generate offline analyses with BGNet. A request to the bot requires specifying several options, like the type of model and the time window of the analysis.

A report is generated, containing the BGNet predictions, attributions and additional control plots, as well as tables with the input variables of the neural networks and the mean predicted sensitivities. For most purposes the generated report contains all information that is expected for a background analysis with BGNet. The generated report is a PDF file that is posted back to the BGNet channel.

Any message sent in the channel is scanned by a python3 application accessing the chat platform via the respective API. A message starting with the special keyword `BOT_REQUEST` will trigger the application to interpret the following message as a JSON compatible dictionary of options that needs to be formatted like:

```
"option1":"value1",
"option2":"value2",
...
"optionN":"valueN",
```

The options available are:

- **target:** Can be `CDC_CUR_AVERAGE` (CDC leakage current) or `VXD_Rad` (diamond radiation monitors) for hit rate reports, `HER_BEAM_LOSS` or `LER_BEAM_LOSS` for beam loss reports and `ECL_LUM_MON_ACC` for luminosity reports.
- **start:** Timestamp that defines the beginning of the time window considered in the report.
- **end:** Timestamp that defines the end of the time window considered in the report.



- **base\_start**: Timestamp that defines the start of the time window considered for the reference samples to calculate the integrated gradients. Default: First valid sample in time window defined by **start** and **end**.
- **base\_end**: Timestamp that defines the end of the time window considered for the reference samples to calculate the integrated gradients. Default: Second valid sample in time window defined by **start** and **end**.
- **test\_start**: Timestamp that defines the start of the time window considered for the test samples to calculate the integrated gradients. Default: First valid sample in time window defined by **start** and **end**.
- **test\_end**: Timestamp that defines the end of the time window considered for the test samples to calculate the integrated gradients. Default: Last valid sample in time window defined by **start** and **end**.
- **tag**: String to name the output file that contains the report. Default: Report.
- **rf**: Defines resample frequency for plotting, for example use **1s** or **5min**. Default: **5min**
- **long**: Boolean to make a report with more attribution control plots. Only available for reports on hit rate models. Default: **False**.
- **model\_index**: Integer to define a sub-model of a background component, if this single sub-model should be used to calculate the feature attribution. **-1**: Full BGNet model, **0**: HER beam-gas, **1**: HER Touschek, **2**: LER beam-gas, **3**: LER Touschek, **4**: HER injection, **5**: LER injection. Valid only for hit rate model. Default: **'-1'**.
- **model\_index**: Integer to define a sub-model of a background component, if this single sub-model should be used to calculate the feature attribution. **-1**: Full BGNet model, **0**: beam-gas, **1**: Touschek. Valid only for beam loss model. Default: **'-1'**.
- **filter\_mode**: Defines the injection gate status for the data used. Can be **None**, **Open**, **Closed**. If **None** is defined no selection is made. Default: **None**
- **ignore\_mdi**: Boolean to specify if data samples should be masked while there is no Belle II high voltage permission. If **True** data will not be masked. Default: **False**.

The chat bot offers to provide templates of requests that can be easily be copied and adapted by writing “help” or “template” into the chat.

The options target, start, end are mandatory. All options are checked for valid input and the user is informed in case the request or the options are malformed. A non-blocking thread is then started to create the report. According to the **start** and

**end** timestamps a model is selected from the local model database. An environment is then set up with a BGNet version that is defined by the BGNet version the model was trained with. The BGNet binary `bgnet_report` is then called performing data loading, preprocessing, inference, plotting and creation of the output file. Multiple reports can be requested at the same time. Typically, it takes about 1 min for a default hit rate BGNet report for a time window of one full day.

An examples of a typical use case, where the chat bot provided quick and valuable information for the operation, will be presented. A new non-linear collimator [121] was introduced recently. It was installed in LER section 5 (D05) in position 1, has vertical (V) jaws and is therefore identified by D05V1. No operational experience existed for that collimator and its effect on the beam needed to be studied in practice. Feedback on the behavior of the individual beam background components after a change of the positions of D05V1 can be provided by BGNet, using the chat bot. The bottom (BTM) jaw D05V1BTM was moved on the 24.11.2025 at 19:06 and BGNet report was requested for this event shortly after. The complete request was as follows:

```
BOT_REQUEST
"target": "VXD_Rad",
"start": "2025-11-24 18:30:00",
"end": "2025-11-24 20:00:00",
"base_start": "2025-11-24 18:40:00",
"base_end": "2025-11-24 18:50:00",
"test_start": "2025-11-24 19:30:00",
"test_end": "2025-11-24 19:50:00",
"tag": "collimator_d05",
"rf": "10s",
```

As a **target** hit rate for BGNet the dose rates of a diamond radiation monitor was chosen and an identifying **tag** was given to the report. An overall time window of the report, around the time of the event, was defined with the **start** and **end** options. The test and baseline samples for the feature attribution were defined to be before and after the event, respectively, using the **base.start**, **base.end** and **test.start**, **test.end** options. The time windows are indicated in the hit rate decomposition plot of the report, that is displayed in Figure 9.15.

The impact on the hit rate and the individual beam background components due to the collimator movement at 19:06 is immediately apparent. The BGNet predictions follows this change in hit rate well and the ratio between the total hit rates stays small during and after the change of the jaw position. The results of the feature attribution are shown in Figure 9.16.

The attribution value identifies that the change in the model prediction between the test and baseline samples is due to the change of the collimator jaw position D05V1BTM. The full systematic name of the collimator jaw position is

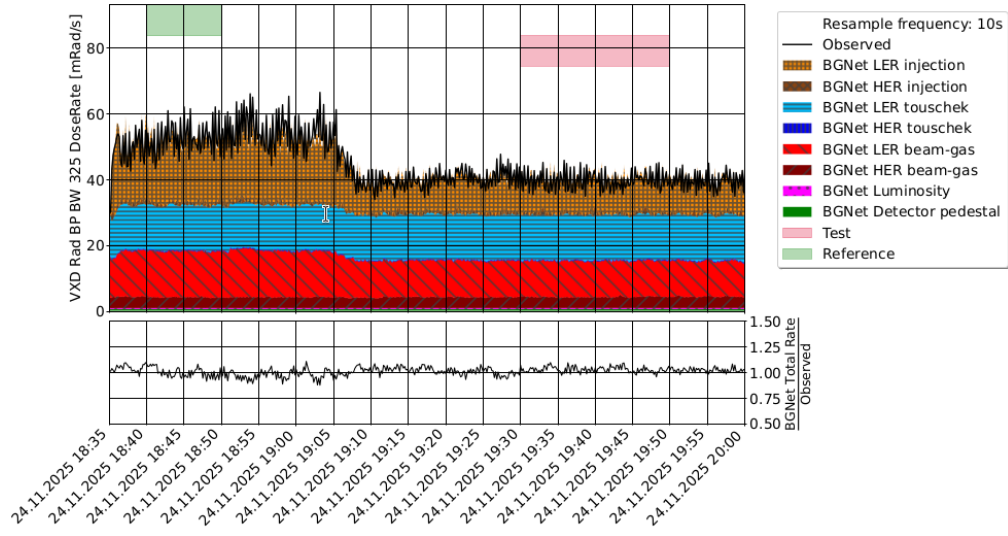


Figure 9.15: In the upper plot, the predicted beam background decomposition is shown as a stacked histogram of all components for the dose rate of a diamond radiation monitor. The total observed, or measured, hit rate is indicated. The chosen time windows for the test and baseline, or reference, data sets are indicated as colored boxes on the top of the plot. The resampling frequency is shown in the legend. The bottom plot shows the ratio of the total measured hit rate and the total predicted hit rate.

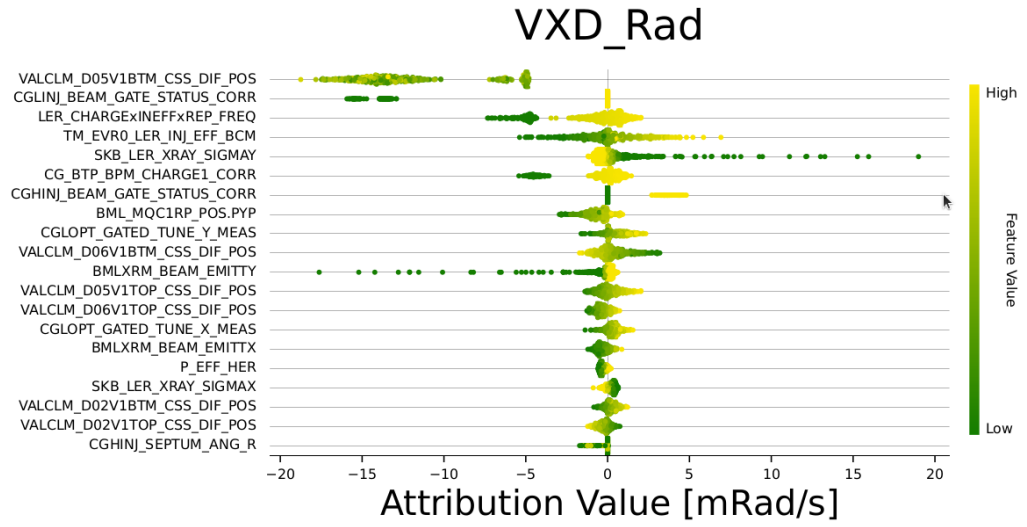


Figure 9.16: Attribution values of input variables for the hit rate model of BGNet. The test and baseline samples are from the time before and after the collimator jaw D05V1BTM was moved, respectively. The attribution values are visualized as point clouds, and the density of points is reflected by the vertical width of the point clouds. Each variable has points equal to the number of samples in the test data set. The color of each point corresponds to the relative feature value compared to all feature values present in all test samples. The input variables are ranked by their mean absolute attribution values.

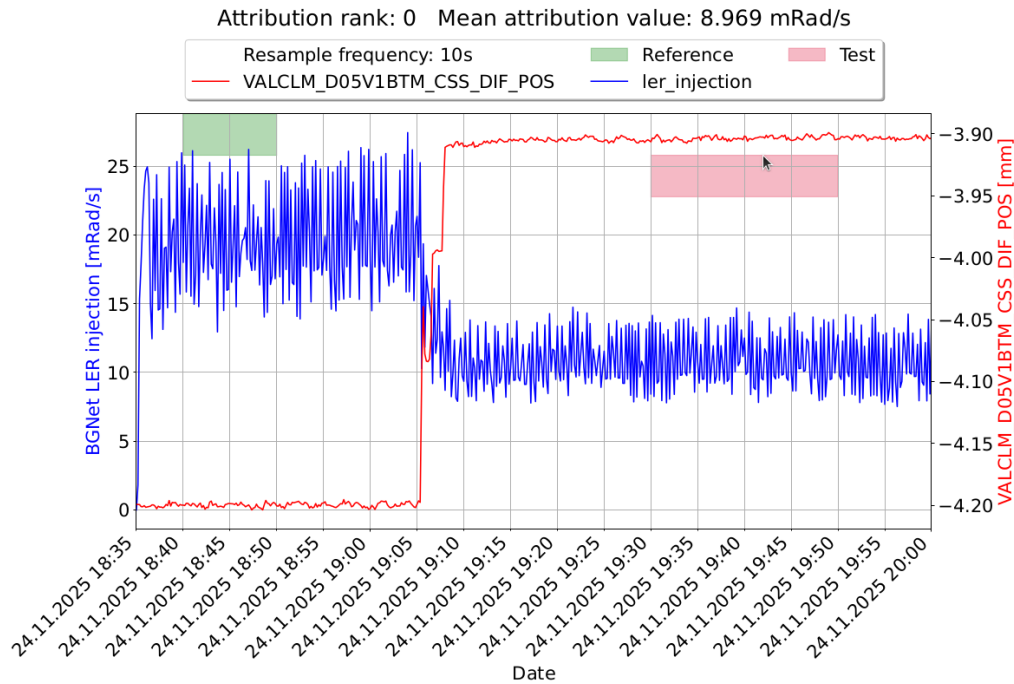


Figure 9.17: The total predicted hit rate and the collimator jaw position D05V1BTM during the selected time window of the report. The test and baseline time windows are indicated as colored boxes on the top of the plot. The closing of the collimator jaw by 0.3 mm coincides in time with the reduction of the predicted hit rate. The mean attribution value is shown. On average, the difference of the total hit rate prediction between the test and baseline samples due to the collimator jaw position is  $12.19 \text{ mrad s}^{-1}$ .

VALCLM\_D05V1BTM\_CSS\_DIF\_POS. The following pages of the report show the overlaid time series of the total predicted hit rate and the input variables with the largest attribution values. This is shown in Figure 9.17 for the position of the collimator jaw position D05V1BTM. The plot shows the coincidence of the change of the collimator jaw position and the predicted hit rate. Since the measured hit rate behaves similarly to the predicted hit rate, it too coincides with the moving of D05V1BTM. From the hit rate decomposition in Figure 9.15 one can see that the largest change of the hit rate occurs in the LER injection background component. By adapting the option "model\_index": "-1" to "model\_index": "5", the attribution values for the LER injection neural network alone will be shown, instead of the attributions for the total BGNet model. Similarly, the plots comparing the time series of the predicted hit rate and an input variable will show the predicted hit rate due to LER injection background instead of the total predicted hit rate. The results are shown in Figure 9.18.

In summary, the report by the BGNet chat bot provides information about the

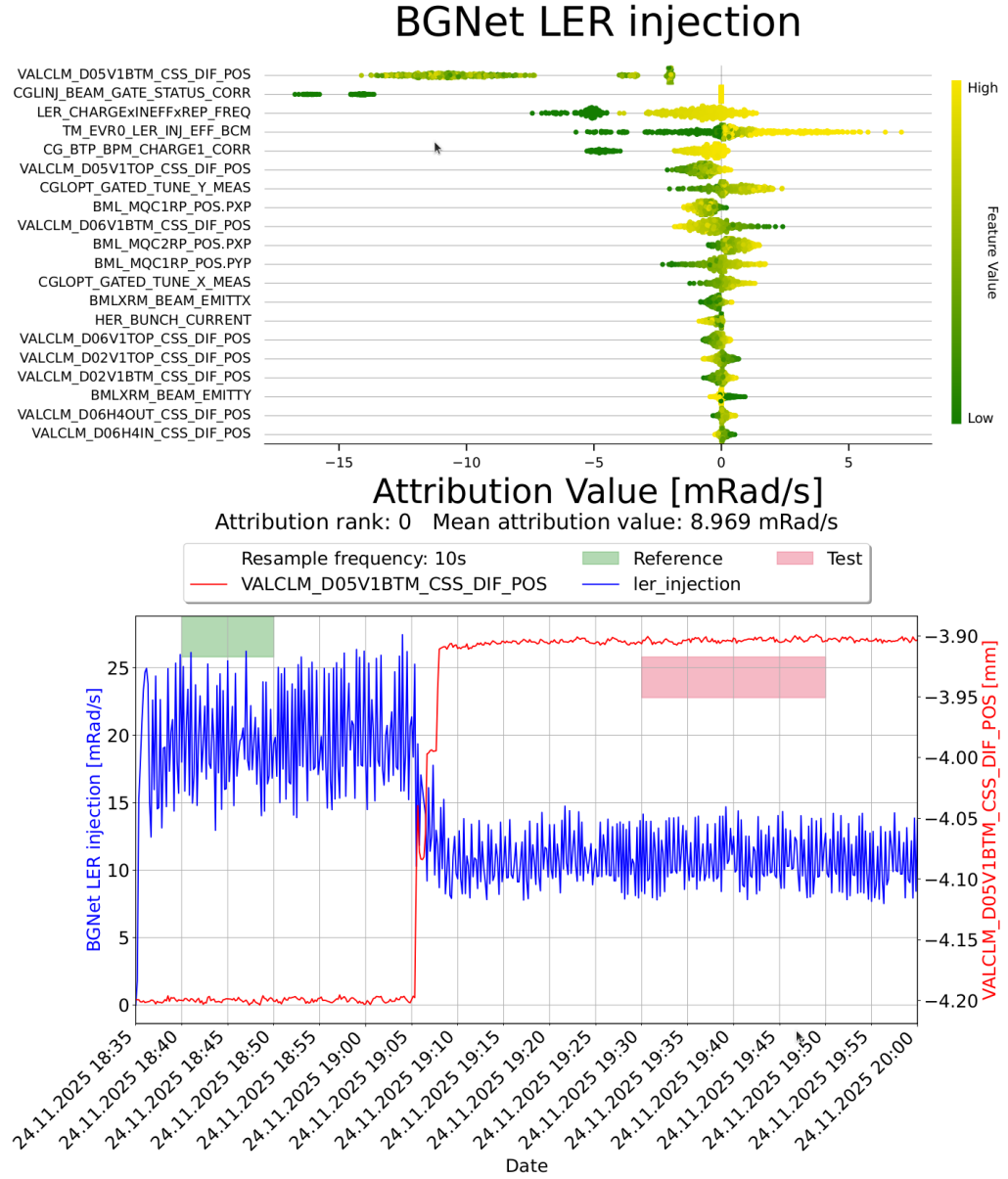


Figure 9.18: Top: Attribution values of input variables for the LER injection network of BGNet. The test and baseline samples are from the time before and after the collimator jaw D05V1BTM was moved, respectively. The attribution values are visualized as point clouds, and the density of points is reflected by the vertical width of the point clouds. Each variable has points equal to the number of samples in the test data set. The color of each point corresponds to the relative feature value compared to all feature values present in all test samples. The input variables are ranked by their mean absolute attribution values. Bottom: The predicted hit rate due to LER injection background and the collimator jaw position D05V1BTM during the selected time window of the report. The closing of the collimator jaw by 0.3 mm coincides in time with the reduction of the predicted hit rate.

measured hit rate and the predicted hit rate with the predicted decomposition. The changes of the magnitude of each beam background component in the hit rate decomposition is shown qualitatively and is then quantified by the attribution values. A ranking of important variables is produced that drove the change of predictions in the individual neural networks and allowed the neural network to accurately reflect the change of the total hit rate. In addition, direct comparison plots are provided between the input variables with large mean absolute attribution values and the neural network predictions. These comparisons allow for an immediate and non-abstract way to inspect how input variables changed and how that change has affected the BGNet prediction, and in extension might have affected the measured hit rate.

The BGNet reports and its information will ultimately always require interpretation by an expert, but so do all tools that are used to understand a complex system, like SuperKEKB. BGNet is not intended as a replacement for any tool or expert but is intended to provide information about the beam background decomposition during real operation. The chat bot is a tool to conveniently and quickly provide information of BGNet to anyone in the collaboration. The creation of the shown report by the chat bot takes less than a minute.

## 9.8 Performance of BGNet

Testing the performance of BGNet is an important step to establish its reliability in the collaboration. Several aspects of the performance of BGNet have been explored during the commissioning in preparation of advertising it to the collaboration. The following performance review was performed by the author for the commissioning. The performance review concentrates on the two main deliverables of BGNet: The prediction of the total hit rate and the decomposition of the predicted hit rate. The performance also needs to be tested separately for scenarios where the data set used for evaluating the performance was part of the training data set (in the following referred to as offline scenario) or where it is not. These two scenarios reflect the performance for offline analyses and for the real-time application, where it is necessary for the models to extrapolate and provide predictions on unseen data. The first part of the performance evaluation of BGNet focuses on the hit rate models as the hit rate decomposition is considered the main deliverable of BGNet. The performance evaluation of the beam loss predictions follows at the end.

The predicted total hit rate can be evaluated by comparing to the measured total hit rate in both offline (see Section 9.8.1 and Section 9.8.2) and real-time scenarios (see Section 9.8.3).

The beam background decomposition is not measured directly but is determined by other methods. The offline performance is compared to the results of fits during background study campaigns (see Section 9.8.2). Additionally, the fast CDC leakage current logger can resolve injection backgrounds and is used to identify the magnitude of storage+luminosity backgrounds, testing the decomposition (see Section 9.8.4). From simulations, the  $\phi$ -profiles of the storage background components are known

[122]. A  $\phi$ -profile for all background components can be obtained from BGNet decompositions by comparing the relative contributions of the components to the dose rate of diamond radiation monitors at different positions in  $\phi$ . The  $\phi$ -profiles obtained from BGNet and the simulation are compared in Section 9.8.5.

The performance of feature attribution is difficult to judge. The prerequisite for correctly identifying the variables driving the changes in background is that the total hit rate and the beam background decomposition are well described at the time the attribution values are computed. Even if this condition is fulfilled, there is no basis for comparison to any measurement or competing method of identifying important variables in a given state of SuperKEKB. However, when broadening the scope of the question one can infer the performance of feature attributions from the extrapolation performance of BGNet. If the model is able to make good predictions for the total hit rate and its decomposition on unseen data, the model must have created a sensible surrogate model of SuperKEKB, that can, to some degree, approximate its physics. Explicitly, the argument would be: The better the extrapolation performance, the better and more physical the approximation of the actual physics processes, the better the feature attribution. The discussed caveats of feature attributions using simplified models in Section 9.4 will always remain, independent of the extrapolation performance. Tests of the applicability of the feature attributions are provided as qualitative examples in Section 9.7 and in Section 9.4.

The second part of the performance study focuses on the beam loss predictions. The performance of the beam loss BGNet models is tested on data seen during the training and during extrapolations. The results are provided in Section 9.8.6.

### 9.8.1 Performance on training data

The performance on training data, or offline performance, of the total hit rate prediction is evaluated by performing the model inference on data seen during training. This is the standard use case when using BGNet for an analysis of the beam background decomposition for historic data. Typically, time windows of the collider operation, where the beam background is very large, very small or is suddenly changing, are analyzed as information on the beam background decomposition is of explicit interest to explain the observed effects. The performance of BGNet on training data is generally better than when the model extrapolates, which should be exploited.

The metric of the performance is chosen to be the ratio between the predicted and measured total hit rate of a sub-detector. The evaluation is performed for two target variables: CDC average layer leakage current and dose rate of a diamond radiation monitor at the backward beam pipe position at  $\phi = 325^\circ$ . A total of 40 different trained models are used in this analysis for each target variable. The training samples overlap between the models but are never equal for any two models. The time window chosen is 1 h long and is arbitrarily positioned within the training data set for each model which generally resembles three weeks of SuperKEKB operation. The ratios are calculated every second and then resampled to 5 min by taking the mean. The results are displayed in Figure 9.19.



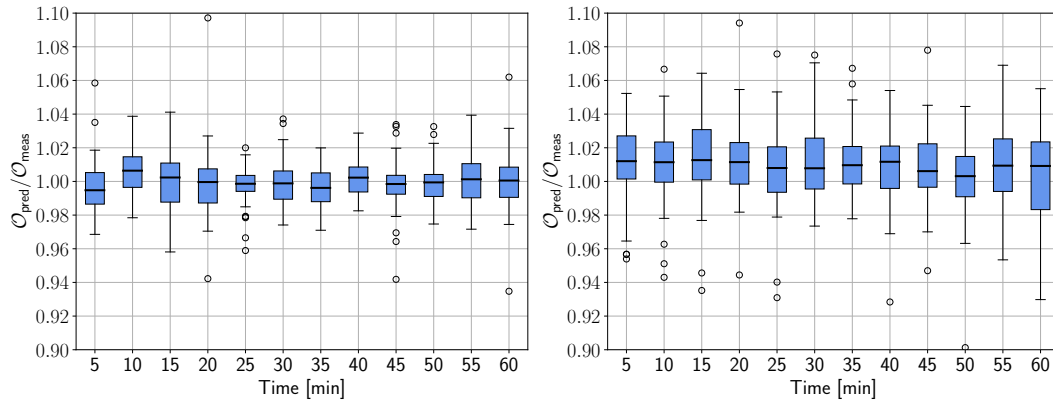


Figure 9.19: Ratio of the predicted hit rate  $\mathcal{O}_{\text{pred}}$  and measured hit rate  $\mathcal{O}_{\text{meas}}$  for 40 models. The predictions were performed on data that was seen during the training. The ratios are calculated every second and then resampled to 5 min, therefore each box represents 40 data points, one for each model, of resampled ratios for the respective time step. The data samples used for the prediction are taken from an arbitrary time window of 1 h within the training data. The box plot follows standard convention where the solid line in the box is the median, the edges of the box extend to the first (Q1) and third quartile (Q3), the whiskers extend the inter quartile range by a factor 1.5 and outliers are explicitly plotted below and above the ends of the whiskers. Left: Ratios for the CDC average layer leakage current. Bottom Left: Ratios for the dose rate of a diamond radiation monitor.

The median of all data points for both target variables is close to unity, fluctuating between 0.99 and 1.01, showcasing the ability for consistent good predictions of the total hit rate. The largest inter quartile ranges over the individual time steps span up to 3% with only few outliers. The model-to-model variations are therefore small. Single models with comparatively worse performance that are under- or over-estimating the hit rate over the whole test window account for most of the outliers. The performance of models on training data is comparable to what is shown here for any sub set of training samples, since the training window of three weeks ensures a similar diverse data set of SuperKEKB states to learn from for each model.

### 9.8.2 Comparison to Background Study Campaign Results

Here, the performance of a BGNet model trained solely on the beam background study campaign data is compared to the series of fits (in the following simply referred to as a fit) that is usually used for the analysis of the beam background decomposition which is described in Section 9.2.1 and [108]. The fit functions are the heuristic formulae, the fit variables are the sensitivities. The fit performed is a  $\chi^2$  fit.

The background study campaign was conducted on 16.04.2024. An overview of the beam background campaign and its stages is displayed in Figure 9.20. The CDC average layer current is used as comparison, as it is a target variable for both BGNet and the study campaign fits. The fits do not take injection background sensitivities into account and are only applied on decay data, thus they are not part of the evaluation, although they are predicted by BGNet. The luminosity sensitivity and detector pedestal are fixed in BGNet and the values are taken from a fit of a previous background study campaign. What is compared are the sensitivities of the storage background components. The fits will produce a single value for the sensitivity during the background study campaign, while BGNet produces a sensitivity prediction every second, meaning for each data point of the 1 Hz time series.

BGNet does not provide an estimate for the uncertainty of any output, as it only predicts the central value. To estimate an uncertainty, the fluctuation of sensitivity predictions during the background study campaign are used. These fluctuations are driven by (small) changes to the inputs of the neural networks during the study day. The fluctuations are quantified by the standard deviation of their distribution over the whole background study campaign. Thus, the errors quoted on the sensitivity predicted by BGNet is the statistical standard deviation of each sensitivity distribution. Since the luminosity sensitivity and detector pedestal are fixed, there is no error given. This method of assigning errors is not optimal, as it implies that the sensitivity is expected to be constant for the duration of the campaign.

However, as discussed below and shown in an example in Section 9.4, an HER beam size blow-up affected the beam background conditions and the sensitivity values. This effect can also be seen in the behavior of the CDC average layer leakage current in Figure 9.20. This also resulted in a poor quality of the fit, while BGNet's flexibility allowed it to cope better with HER beam size blow-up. The errors for the sensitivities obtained by the fit are taken from the covariance matrix obtained from

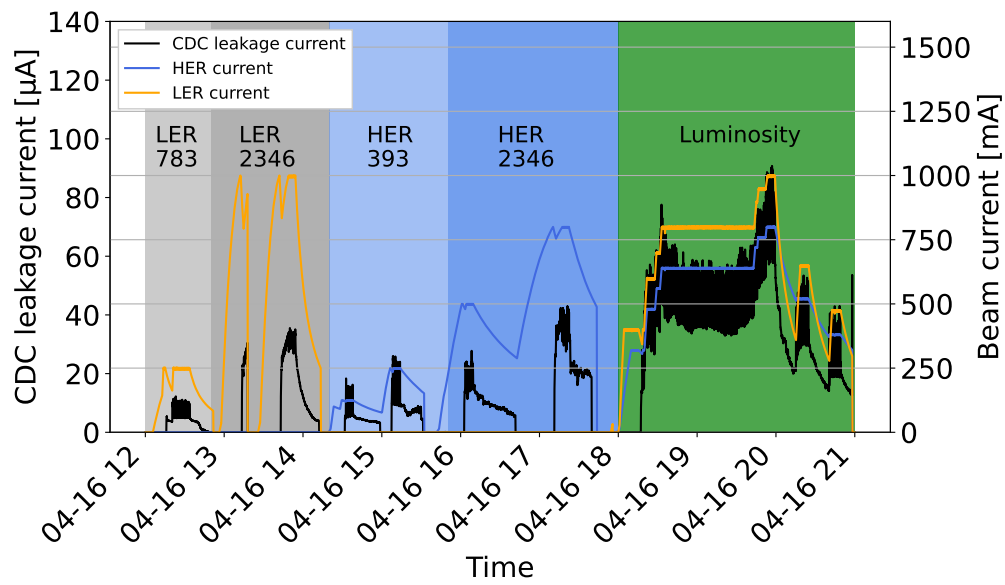


Figure 9.20: Overview of the beam background study campaign conducted on 16.04.2024. The leakage current and the beam currents of both rings are shown. The segments of the study campaign are indicated by the background color and the respective labels. The labels indicate whether a single-beam study is performed, in either the HER or the LER, or a luminosity scan. The number in the label indicates the number of bunches in the respective ring, that needs to be varied in order to differentiate Touschek and beam-gas backgrounds.

| Bkg. components | Fit                              | BGNet                            |
|-----------------|----------------------------------|----------------------------------|
| HER beam-gas    | $(1.28 \pm 0.08) \times 10^{-4}$ | $(4 \pm 2) \times 10^{-4}$       |
| HER Touschek    | $(2.38 \pm 0.02) \times 10^3$    | $(2.0 \pm 0.1) \times 10^{-3}$   |
| LER beam-gas    | $(1.39 \pm 0.01) \times 10^{-4}$ | $(0.92 \pm 0.05) \times 10^{-4}$ |
| LER Touschek    | $(1.28 \pm 0.01) \times 10^3$    | $(1.1 \pm 0.2) \times 10^3$      |
| Luminosity      | $(4.1 \pm 0.2) \times 10^{-4}$   | $4.99 \times 10^{-4}$            |
| Pedestal        | $(1.51 \pm 0.02)$                | 1.35                             |

Table 9.4: Sensitivities obtained during the background study camping on 16.04.2024 using the background fit procedure and BGNet. The errors on the sensitivities are obtained from the covariance matrix of the fit and from the standard deviation of the sensitivity distributions for BGNet.

| Bkg. components | Unit                                         |
|-----------------|----------------------------------------------|
| HER beam-gas    | $\mu\text{A nPa}^{-1} \text{mA}^{-1}$        |
| HER Touschek    | $\mu\text{A}\mu\text{m}^3 \text{mA}^2$       |
| LER beam-gas    | $\mu\text{A nPa}^{-1} \text{mA}^{-1}$        |
| LER Touschek    | $\mu\text{A}\mu\text{m}^3 \text{mA}^2$       |
| Luminosity      | $10^{-30} \mu\text{A}\mu\text{m}^2 \text{s}$ |
| Pedestal        | $\mu\text{A}$                                |

Table 9.5: Units of the sensitivities for the CDC average layer current in units of  $\mu\text{A}$ .

the fit, while the errors on the sensitivities of BGNet are the standard deviations of the sensitivity distributions obtained over the full duration of the background study campaign. The obtained sensitivities are displayed in Table 9.4. The units of the sensitivities are displayed in Table 9.5

The ratio between the BGNet sensitivities and the fit parameters is  $\sim 1$  for all sensitivities except HER beam-gas where the ratio is around 3. The large ratio can be explained by the effects of the HER beam size blow-up.

Further, the ratios between the total measured hit rate versus the BGNet predicted hit rate and the total measured hit rate versus the hit rate obtained by fitting can be compared. The decomposition and the ratios are displayed in Figure 9.21 for a single beam decay of the HER at a number of bunches of 393 and in Figure 9.22 at a number of bunches of 2346. The decomposition and the ratios are displayed in Figure 9.23 for multiple decays during collisions.

The total measured hit rate during the decays with collisions is much better described by the BGNet predictions than by the fit. The fit and the BGNet predictions show a similar poor performance of describing the measured hit rate during the HER decay. During the HER single beam decay, the CDC average layer leakage current has visible artifacts at 16:17 and 16:32, where the hit rate increases, instead of following the smooth beam decay, that is predicted by both the fit and

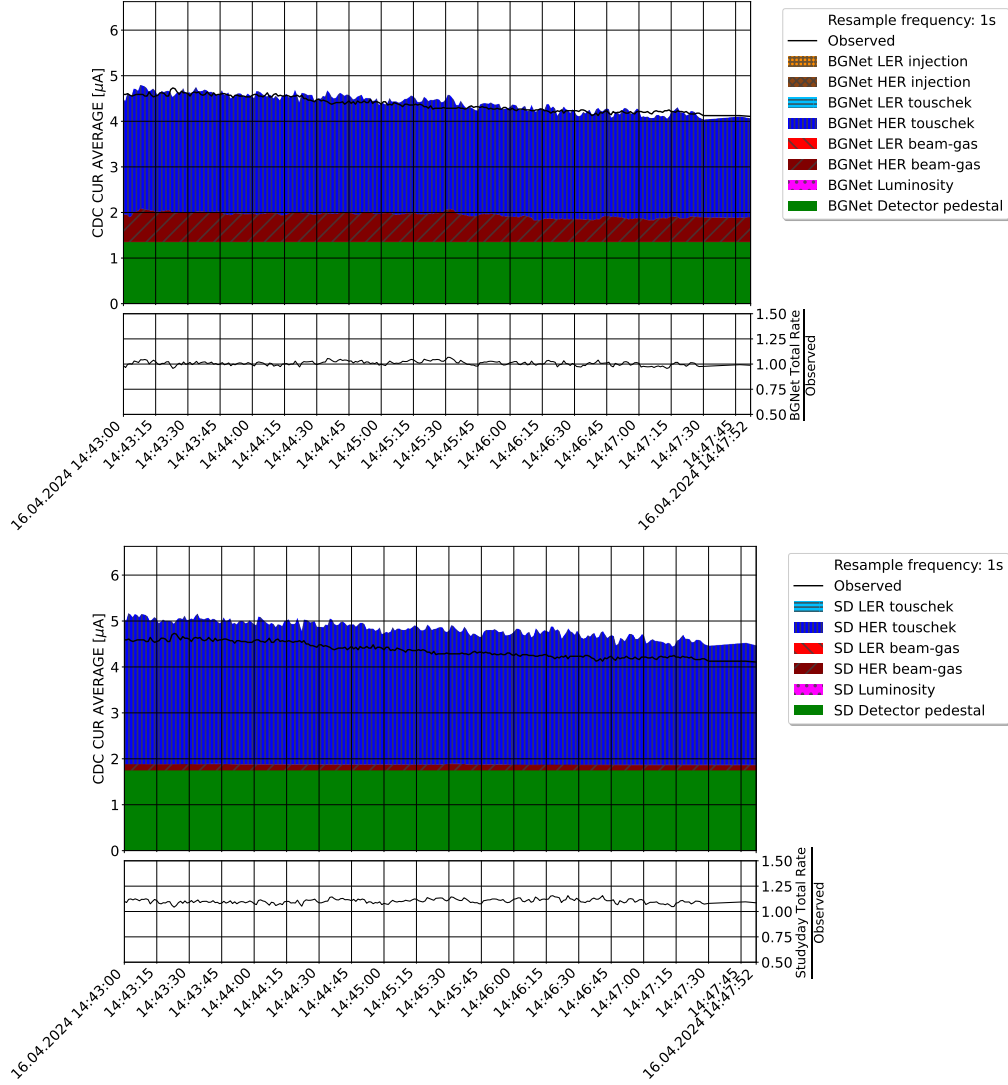


Figure 9.21: Decompositions of the CDC average layer current during a beam back-ground study campaign as a stacked histogram at the time of an HER single beam decay at a number of bunches of 393. The bottom plot shows the ratio between the measured and predicted CDC average layer current. Top: Decomposition obtained from BGNet. Bottom: Decomposition obtained from the fit.

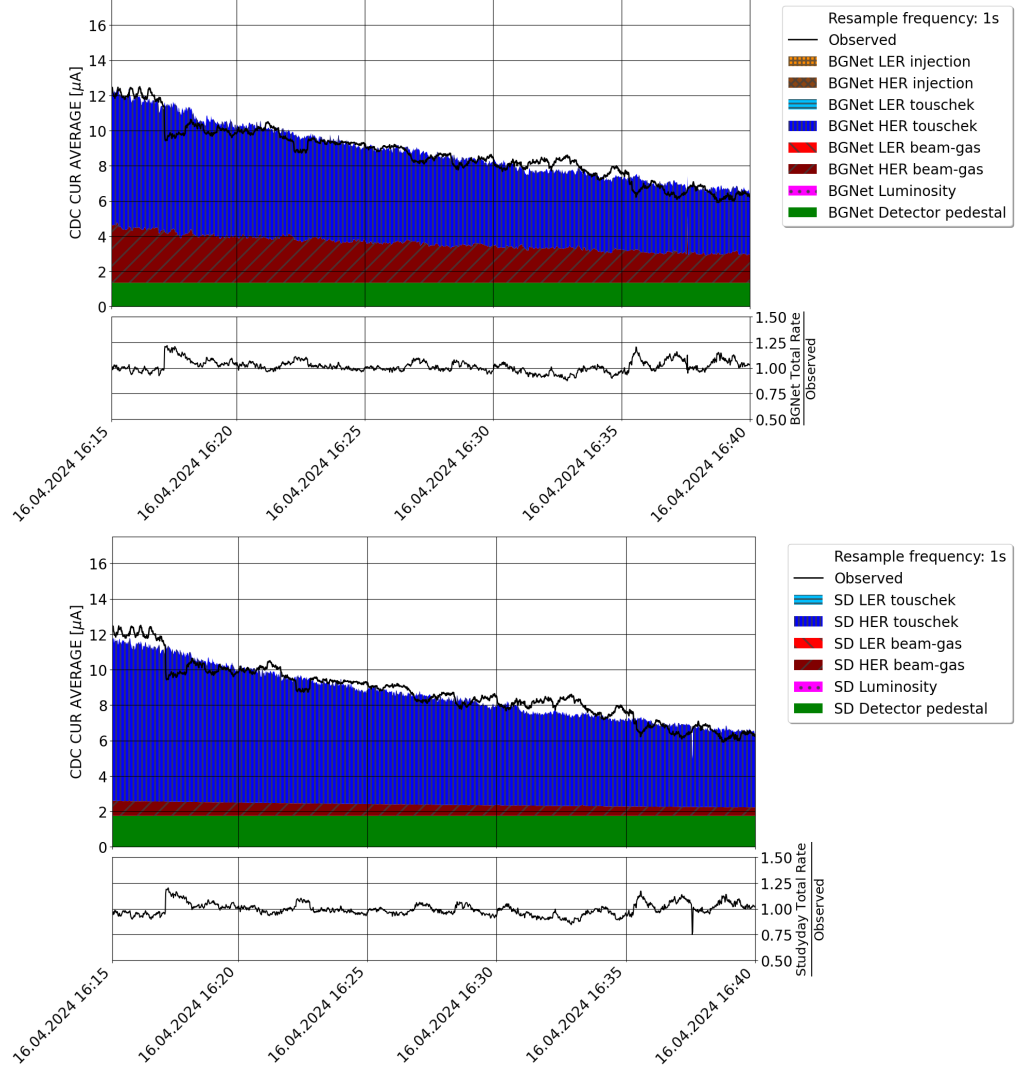


Figure 9.22: Decompositions of the CDC average layer current during a beam background study campaign as a stacked histogram at the time of an HER single beam decay at a number of bunches of 2346. The bottom plot shows the ratio between the measured and predicted CDC average layer current. Top: Decomposition obtained from BGNet. Bottom: Decomposition obtained from the fit.

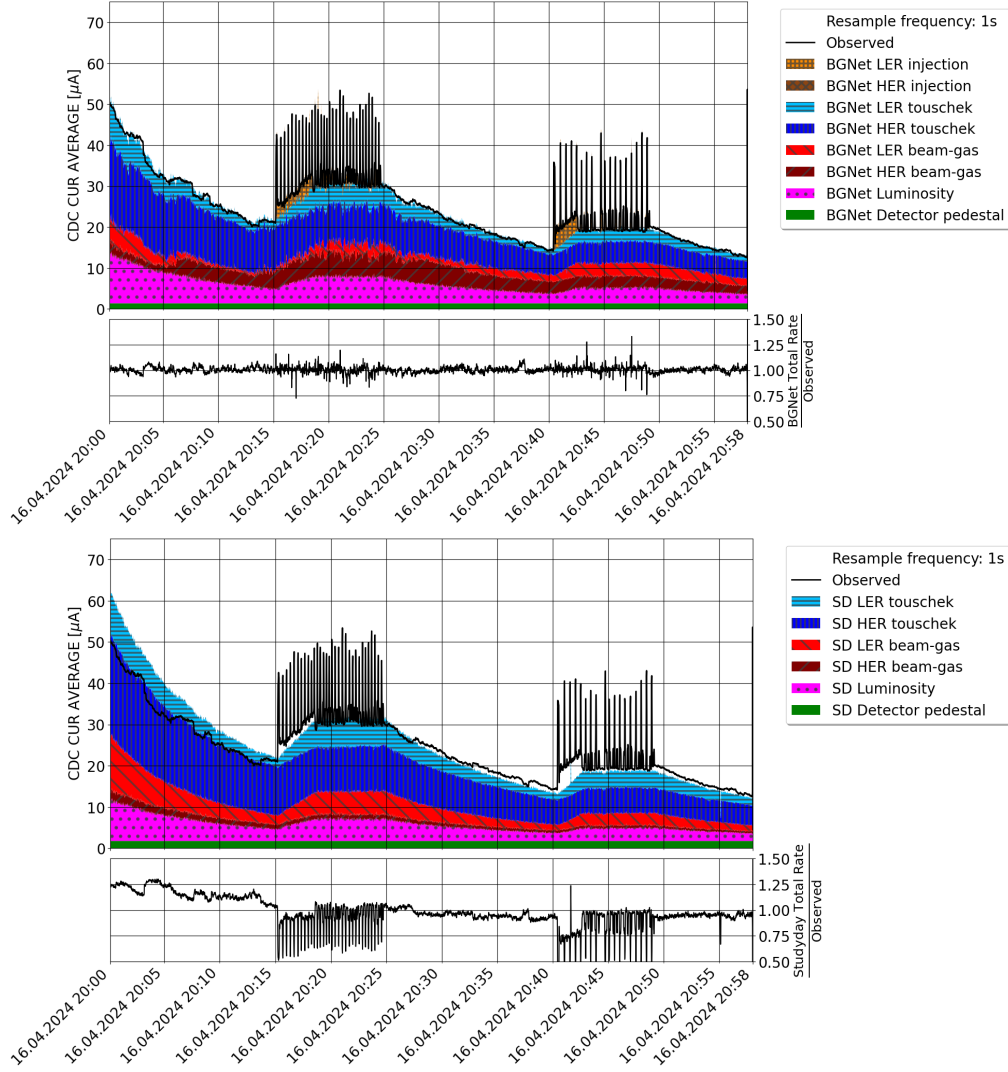


Figure 9.23: Decompositions of the CDC average layer current during a beam background study campaign as a stacked histogram at the time of multiple decays during collisions. The bottom plot shows the ratio between the measured and predicted CDC average layer current for predictions by BGNet and the fit, respectively. Top: Decomposition obtained from BGNet. Bottom: Decomposition obtained from the fit.

BGNet.

The situation is summarized here, but explored in more detail in Section 9.4. The immediate issue was a HER beam size blow-up during the beam decays. The HER beam size is part of the heuristic function for HER Touschek and the dependence of the beam background on the beam size should be well described, however, the descriptive function of the heuristic functions is limited. Here, it becomes clear that the heuristic function multiplied with a sensitivity that is not dependent on the parameters of the heuristic function, and the HER beam size does not deliver a valid description during the transient phases of rapidly fluctuating HER beam sizes.

The beam size blow-up was traced back to issues with the SuperKEKB bunch-by-bunch feedback system [27, 59] during the beam decay. The problematic decays are not regarded in the fit. The smaller fluctuations are explained by the very low hit rates exhibited during this single beam decay, making the noise, or pedestal, of the measurement of the CDC average layer leakage current visible. This can be seen in Figure 9.21, where the pedestal component has a significant contribution. Similar issues can be observed during the first beam decay during collisions, where the hit rate increases despite decreasing currents. However, these moments were correctly described by BGNet, qualitatively following the measured hit rate. The ratio between the measured and predicted total hit rate by BGNet stays in a tight window around unity with small peaks during injections, where the timing of the hit rate increase due to the injections was slightly off. The performance of the fit is worse since it is not flexible enough to account for such behavior of the hit rate. Additionally, the fit can only be applied on beam decay data that was affected by the HER beam size blow-up, while BGNet was trained on all available data points, including the injection data. Since the injections are not accounted for by the fit the ratio between the hit rate predicted by the fit and the measured data during the injections should not be used to evaluate the performance of the fit.

### 9.8.3 Extrapolation with BGNet

The extrapolation performance of BGNet is important to judge whether the surrogate models of the accelerator learned to approximate real physical effects and can predict the total hit rate on unseen data by appropriately scaling the beam background components. The feature attribution methods can also only provide sensible results if this is true. The real-time predictions only make predictions on unseen data. Since the real-time predictions is one of the main deliverables of BGNet, quantifying the extrapolation performance of BGNet is important. The extrapolation performance was an important factor during the commissioning of BGNet.

The metric for the extrapolation performance is chosen to be the ratio between the predicted and measured total hit rate of a sub-detector. The training of the largest BGNet model, the multi-sensor hit rate model for the diamond radiation monitors, takes usually about  $T_{\text{train}} = 40 \text{ min}$  up to  $T_{\text{train}} = 1 \text{ h}$ . In the following, the longest training time of  $T_{\text{train}} = 1 \text{ h}$  is assumed. The first and last timestamps of the training window are  $t_{\text{start}}$  and  $t_{\text{end}}$ . Thus, for the real-time predictions, which



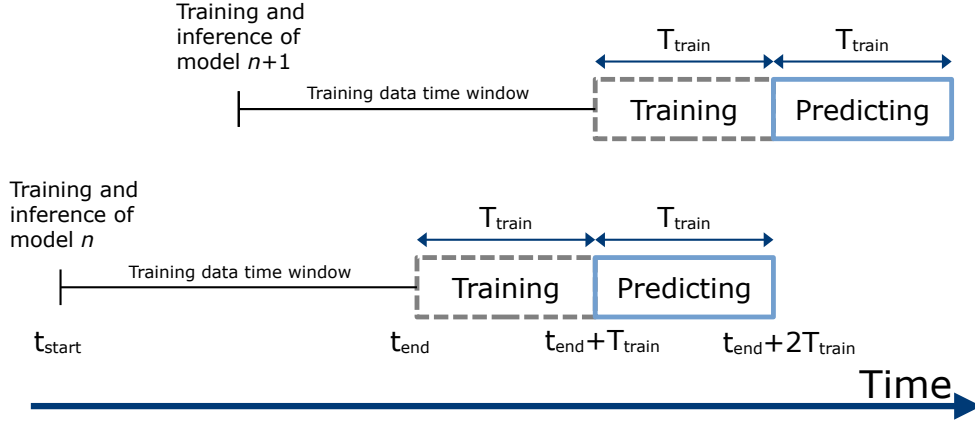


Figure 9.24: Training scheme of new BGNet models for the real-time application of BGNet. The time window for the training data is defined from  $t_{\text{start}}$  to  $t_{\text{end}}$ . The time needed for training a model is  $T_{\text{train}}$ . A model is deployed and used for predictions from  $t_{\text{end}} + T_{\text{train}}$  until  $t_{\text{end}} + 2T_{\text{train}}$ .

immediately use the latest model, the extrapolation is performed from  $t_{\text{end}} + T_{\text{train}}$  until  $t_{\text{end}} + 2T_{\text{train}}$ . A training window typically includes 3 weeks of data. A visualization of the training scheme and the relevant timestamps is given in Figure 9.24.

To observe the full development of the predictions on data beyond the training data set, the extrapolation results from  $t_{\text{end}}$  until  $t_{\text{end}} + T_{\text{train}}$  are analyzed as well. The BGNet models used for this test are models trained during December 2025 for the real-time operation of BGNet to reflect the actual performance of the BGNet predictions delivered in the latest operation period. The evaluation is performed for two target variables: CDC average layer leakage current and dose rate of a diamond radiation monitor at the backward beam pipe position at  $\phi = 325^\circ$ . A total of 40 consecutively trained models are used in this analysis for each target variable. The calculation of the ratio was performed at 1 s and the results were then resampled to 5 min by taking the mean. The results are displayed in Figure 9.25.

During the first hour after the training, the performance of the extrapolated prediction for most models is observed to be stable. The median of the ratio between the predicted and measured total hit rate keeps close to 1 with minor fluctuations. The width of the distribution is visualized as the height of the box, that is the interquartile range (IQR) from the first (Q1) to the third quartile (Q3). The IQR is  $\lesssim 5\%$  for the CDC average layer leakage current, while it is typically a few percent larger for the predictions of the diamond radiation monitor dose rate. There are only a few outliers for the in both cases, that mostly belong to the same two models, showing poor performance for the duration of this analysis. The performance in the second hour of extrapolated predictions show a similar median, that keeps close to 1.

A minimal bias of the median towards higher ratios can be observed for both target variables. A similar trend, and more pronounced, is seen in the interquartile ranges, which tend towards larger ratios. The interquartile ranges have also become larger, as have the number of outliers. A degradation of the prediction performance is expected, the further the SuperKEKB state diverges from the states seen during training. The divergence becomes larger as small changes accumulate over time. However, the majority of all predictions for both target variables lies within a few percent, which is more than sufficient for the purpose of BGNet. This shows that the BGNet models generalize sufficiently well to produce a beam background decomposition that allows extrapolation over several hours and reflecting the changes in background conditions well. A similar argument can be made for the feature attributions produced, as the BGNet models need to have found effective dependencies between the beam background and the input variables.

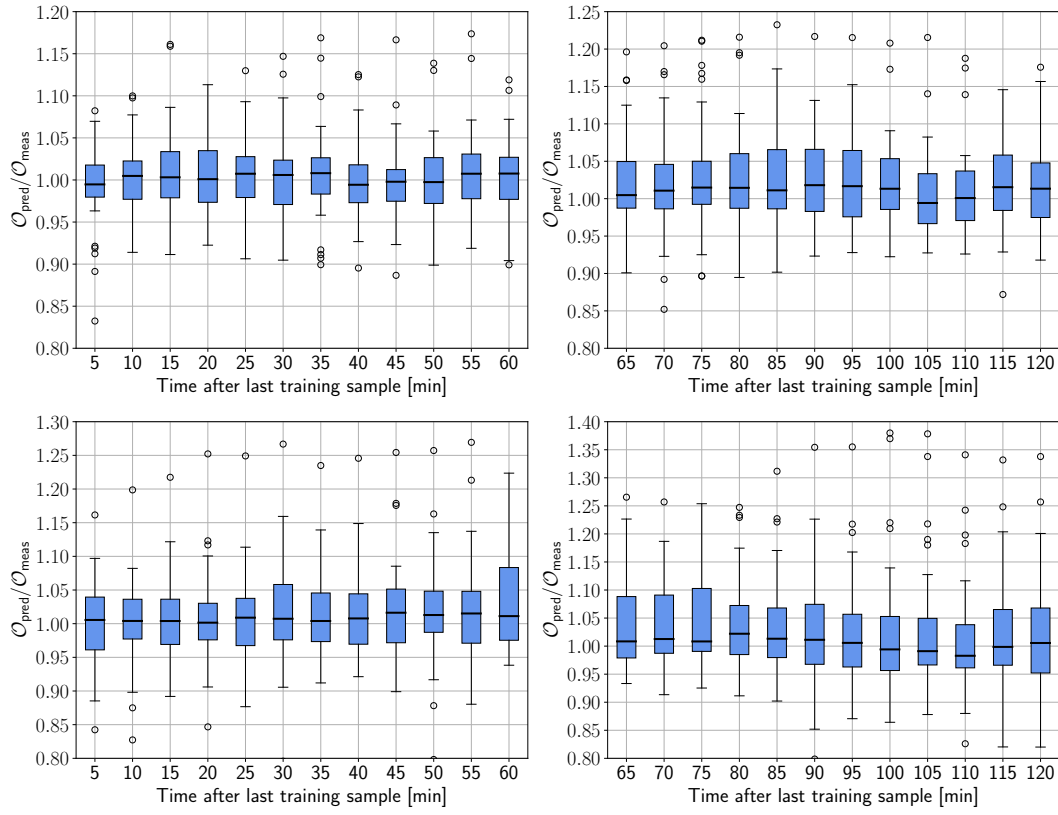


Figure 9.25: Ratio of the predicted hit rate  $\mathcal{O}_{\text{pred}}$  and measured hit rate  $\mathcal{O}_{\text{meas}}$ . The predictions were performed on data that was not seen during the training and the BGNet model is extrapolating. The ratios are calculated every second and then resampled to 5 min, therefore each box represents 40 data points, one for each model, of resampled ratios for the respective time step. The data samples used for the prediction are the ones immediately following the last sample seen during the training. The box plot follows standard convention where the solid line in the box is the median, the edges of the box extend to the first (Q1) and third quartile (Q3), the whiskers extend the inter quartile range by a factor 1.5 and outliers are explicitly plotted below and above the ends of the whiskers. Top Left: Ratios for the first hour after the training for the CDC average layer leakage current. Top Right: Ratios for the second hour after the training for the CDC average layer leakage current. This time window is relevant for the performance of the real-time predictions. Bottom Left: Ratios for the first hour after the training for the dose rate of a diamond radiation monitor. Bottom Right: Ratios for the second hour after the training for the dose rate of a diamond radiation monitor. This time window is relevant for the performance of the real-time predictions.

#### 9.8.4 Differentiation of Injection Backgrounds

The average leakage current of the CDC is monitored and is sampled with a 1 Hz frequency and is available as 1 Hz time series in EPICS . The average leakage current is proportional to the hit rate of the CDC and is used as a target variable in BGNet. Additionally, a fast current logger is installed with a sampling rate of 1 kHz. It measures the leakage current  $I_{\text{Log}}$  of high-voltage sector 3 of layer 4. An annotated screen capture of the attached oscilloscope is displayed in Figure 9.26. The injection spikes are visibly resolved by the fast current logger. The LER injection frequency is 12.5 Hz with alternating 1- and 2-bunch injections. Here, the ratio between the number of 1- and 2-bunch injections is 1:1. During the 2-bunch injection scheme two bunches are injected with a spacing of nanoseconds. The current spikes corresponding to 2-bunch injections are larger. The HER injections have a frequency of 25 Hz and are exclusively 1-bunch injections. A baseline shift is observed during injections into both rings. This occurs, when the injection background is large and the leakage current takes a longer period of time to fall back to the baseline, while the injections are too frequent, not allowing the leakage current to drop to the actual baseline without leakage current due to injections.

Looking at the part with only LER injections, after an injection spike the current returns to a baseline as the additional background from the injection dissipates. The leakage current at the baseline then corresponds to the leakage current due to the storage and luminosity backgrounds and the detector pedestal. A 1 Hz time series of the fast current logger is available in EPICS , corresponding to the minimal value of  $I_{\text{Log}}$  during a given second, here called  $I_{\text{Log}}^{\text{min}}$ . The minimum leakage current  $I_{\text{Log}}^{\text{min}}$  provides the baseline value that can be used to estimate the average layer leakage current, which is the BGNet target variable, due to the sum of all background components except the injection backgrounds. For  $I_{\text{Log}}^{\text{min}}$  to be comparable to the average layer leakage current, it needs to be multiplied by a factor of four, as the sector covered by the fast current logger accounts for one quarter of layer 4. The average layer leakage current and the leakage current of layer 4 need to be the same for such a comparison which is an approximation made here. This comparison also needs to be conducted outside time windows where the baseline shifted.

The comparison of  $4 \cdot I_{\text{Log}}^{\text{min}}$  and the BGNet prediction for the average CDC layer leakage current is displayed in Figure 9.27. A good agreement between  $4 \cdot I_{\text{Log}}^{\text{min}}$  and the sum of all non-injection backgrounds can be observed. The BGNet model, therefore, is able to correctly identify and differentiate injection background from non-injection background despite being trained on only 1 Hz time series data. It learned an effective description of the superposed injection backgrounds at 1 Hz, while the underlying time scales of the injection backgrounds are much smaller.

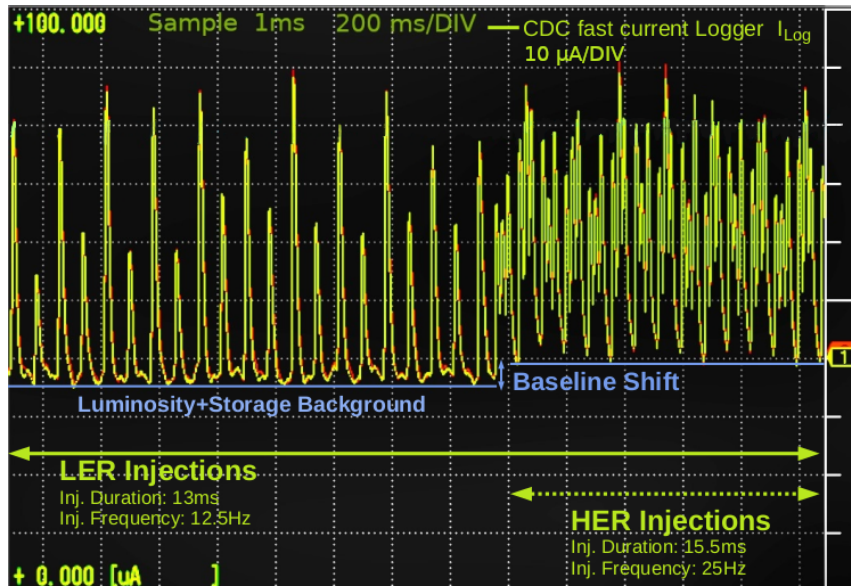


Figure 9.26: Oscilloscope screen capture of the fast CDC current logger on June 12, 2022 at 20:52:39 (JST). The  $y$ -axis shows the leakage current in units of  $10 \mu\text{A}$  per DIV. It was taken during physics data taking during a period with top-up injections. Initially, only the LER received injections with asynchronous HER injections starting about 2 s later [103].

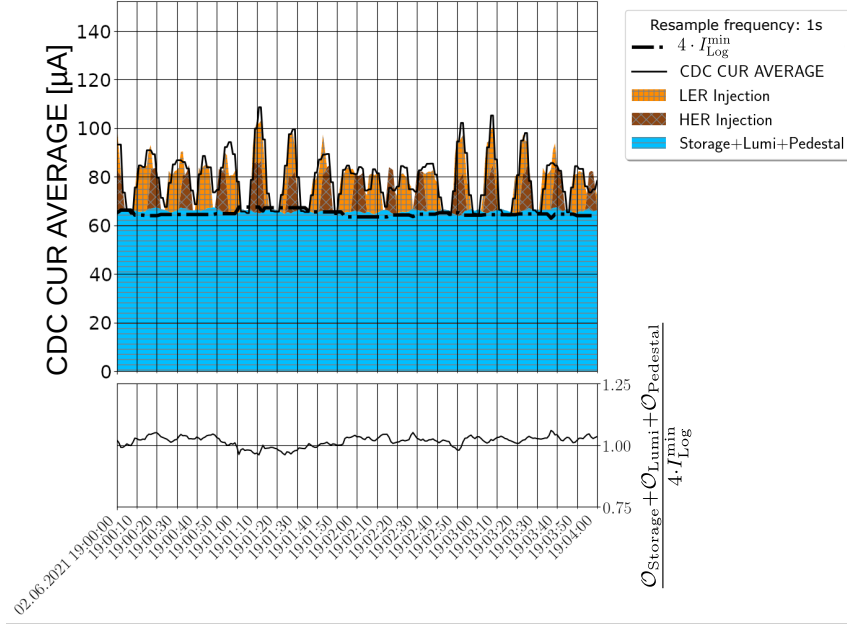


Figure 9.27: BGNet predictions during top-up injections. The non-injection backgrounds are summed up and displayed together. The injection backgrounds for HER and LER are displayed separately. The current  $4 \cdot I_{\text{Log}}^{\text{min}}$  can now be compared to the average CDC layer leakage current due to non-injection background to gauge BGNet’s ability to differentiate injection backgrounds from the other background components [103].

### 9.8.5 $\phi$ -Profiles of the Beam Background

During the design and build phase of Belle II, efforts were made to simulate the beam background that the PXD would be exposed to [122]. The simulation was done on a per module basis, taking its position around the IP into account and with that  $\phi$ -profiles of the storage background components have been established. Thus, the  $\phi$ -profiles describe the distribution of the beam backgrounds in the  $xy$ -plane. If BGNet can correctly separate the storage backgrounds of each component such  $\phi$ -profiles should be approximately reproduced by BGNet. Several dose rates for diamond radiation monitors at different positions in  $\phi$  are predicted by BGNet. They are used to create a  $\phi$ -profile for all background components considered by BGNet. The diamonds used are the forward SVD diamonds at  $\phi = 60^\circ, 120^\circ, 180^\circ, 240^\circ, 300^\circ$ . The diamond at  $\phi = 0^\circ$  is not operational. Since these diamond radiation monitors are positioned at different radii and at different positions along the  $z$ -axis an exact match of the  $\phi$ -profiles is not expected.

The outer modules of the PXD are positioned at  $30^\circ$  intervals, starting at  $\phi = 0^\circ$ . The position  $\phi = 0^\circ$  is defined to be at  $(x, y, z) = (1, 0, 0)$ . A geometry of the PXD, where the numbers of PXD modules are indicated, is displayed in Figure 9.28. The appropriate PXD modules to compare to the available diamond radiation modules have been identified to be 2.3.1, 2.5.1, 2.7.1, 2.9.1, 2.11.1 which are positioned at the same angles in  $\phi$  and are located in the forward direction.

The BGNet model used for obtaining the  $\phi$ -profiles was trained on operation data from December 2025 and the data used for the inference was seen during the training. The hit rate for each diamond is the mean hit rate over 4 min buckets of a 1 Hz time series. The total predicted hit rate needs to agree for with the measured total hit rate of all diamonds in order for the decomposition to be sensible. The  $\phi$ -profile for the total BGNet predictions and the measured hit rate is displayed in Figure 9.29. Additionally, the decompositions of the hit rate over the full time window for all considered diamonds are displayed in Appendix B.4. The  $\phi$ -profiles for the BGNet predictions of all storage and injection background components are shown Figure 9.30 and Figure 9.31.

The simulated absolute hit rates due to beam background components, given as occupancy in units of percent, quoted for the PXD [122] cannot be compared to the predicted absolute hit rates for the diamonds, given as dose rate in units of  $\text{mrads}^{-1}$ , by BGNet. To convert the occupancy of the PXD into a dose rate an energy calibration of the average deposited energy per firing pixel is required but not available. Further, the accelerator conditions during the simulation will be different than what was assumed in the simulation, which were performed in the year 2015 before SuperKEKB operations. The simulation is dependent on SAD (Strategic Accelerator Design) simulations [124], which are performed for specific sets of optic and collimator configurations, which are known to have changed considerably since then. Also the beam pipe pressure was increased due to maintenance performed prior to the test time window in December 2025, which required opening the beam pipes. This leads to elevated pressure levels that only return to nominal values after weeks

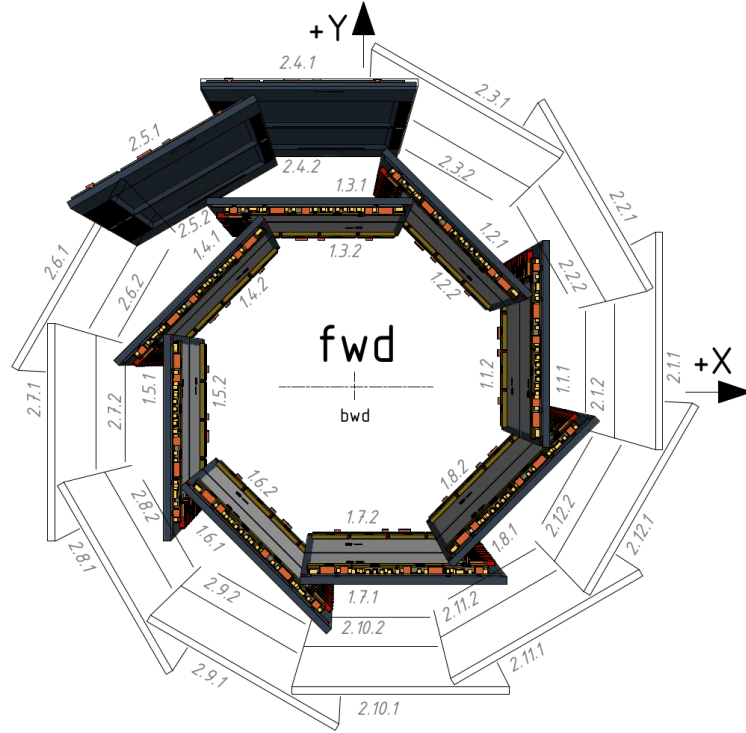


Figure 9.28: A schematic view of the PXD indicating its position in the coordinate system and indicating the module numbers. The colored modules indicate the modules installed in PXD 1 [123].

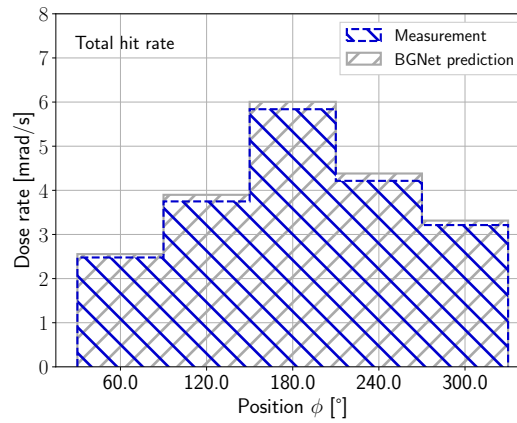


Figure 9.29:  $\phi$ -profile for the total predicted hit rate by BGNet and the total measured hit rate for diamond radiation monitors at  $\phi = 60^\circ, 120^\circ, 180^\circ, 240^\circ, 300^\circ$ .



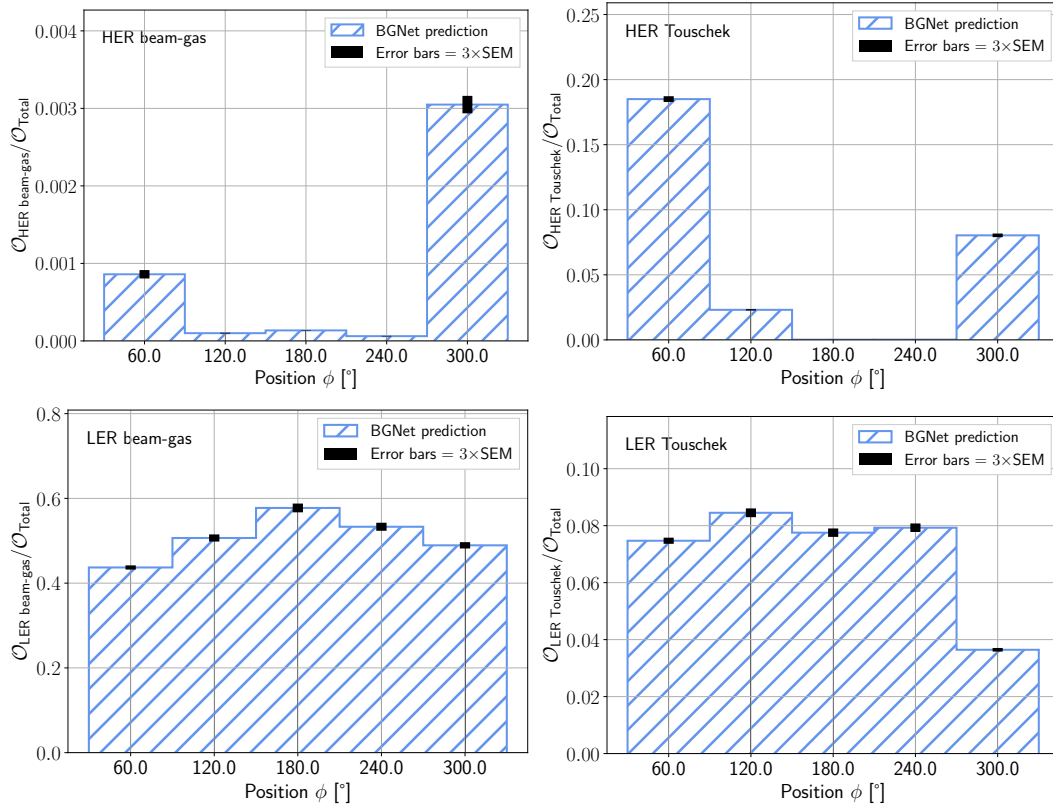


Figure 9.30:  $\phi$ -profiles for the storage backgrounds based on BGNet predictions for diamond radiation monitors at  $\phi = 60^\circ, 120^\circ, 180^\circ, 240^\circ, 300^\circ$ . The hit rate due to each component has been normalized by the total predicted hit rate. The error shown is the standard error of the sample mean, which was increased by a factor 3 to make it visible. Top left: HER beam-gas component. Top right: HER Touschek component. Bottom left: LER beam-gas component. Bottom right: LER Touschek component.

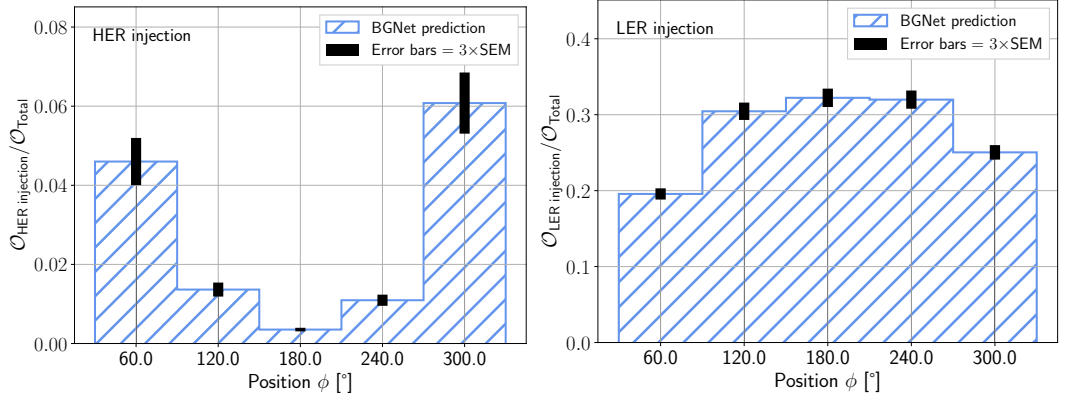


Figure 9.31:  $\phi$ -profiles for the injection backgrounds based on BGNet predictions for diamond radiation monitors at  $\phi = 60^\circ, 120^\circ, 180^\circ, 240^\circ, 300^\circ$ . The hit rate due to each component has been normalized by the total predicted hit rate. The error shown is the standard error of the sample mean, which was increased by a factor 3 to make it visible. Left: HER injection component. Right: LER injection component.

| PXD Module | PXD rel. hit rate [122] | Dia. pos. in $\phi$ | BGNet rel. hit rate |
|------------|-------------------------|---------------------|---------------------|
| 2.3.1      | $(0.6 \pm 0.5)$         | $60^\circ$          | $(0.757 \pm 0.005)$ |
| 2.5.1      | $(0.6 \pm 0.5)$         | $120^\circ$         | $(0.877 \pm 0.007)$ |
| 2.7.1      | $(1.0 \pm 0.8)$         | $180^\circ$         | $(1.000 \pm 0.008)$ |
| 2.9.1      | $(0.8 \pm 0.6)$         | $240^\circ$         | $(0.923 \pm 0.008)$ |
| 2.11.1     | $(0.6 \pm 0.5)$         | $300^\circ$         | $(0.847 \pm 0.007)$ |

Table 9.6: Relative hit rates due to LER beam-gas for PXD modules from simulation [122] and for diamond radiation monitors from BGNet predictions.

and months of operations. So, instead of comparing absolute values, relative values are compared to obtain a qualitative comparison. The relative values are obtained by dividing by largest hit rate in the selection of PXD modules and diamond radiation sensors. The errors for the BGNet predictions are the standard errors of the sample mean. Only the LER beam-gas, HER Touschek and LER Touschek components may be compared, as the HER beam-gas are not available for all PXD modules and have too large errors. The comparison for LER beam-gas is shown in Table 9.6, for LER Touschek in Table 9.7 and for HER Touschek in Table 9.8.

The  $\phi$ -profiles for LER beam-gas and HER Touschek have their highest hit rate at the same positions, comparing the simulation and the BGNet prediction, while around that maximum the hit rate drops to lower values, approaching the minimum of the  $\phi$ -profile on the opposite side of the maximum. The  $\phi$ -profile of LER Touschek shows a difference in the position of the maximum and minimum. The maximum of the BGNet prediction is at the minimum of the PXD simulation and its minimum is at the maximum of the PXD simulation. While the errors on the PXD simulation are

| PXD Module | PXD rel. hit rate [122] | Dia. pos. in $\phi$ | BGNet rel. hit rate |
|------------|-------------------------|---------------------|---------------------|
| 2.3.1      | $(0.9 \pm 0.6)$         | $60^\circ$          | $(0.884 \pm 0.007)$ |
| 2.5.1      | $(0.4 \pm 0.3)$         | $120^\circ$         | $(1.000 \pm 0.009)$ |
| 2.7.1      | $(0.6 \pm 0.4)$         | $180^\circ$         | $(0.917 \pm 0.008)$ |
| 2.9.1      | $(0.4 \pm 0.3)$         | $240^\circ$         | $(0.938 \pm 0.008)$ |
| 2.11.1     | $(1.0 \pm 0.6)$         | $300^\circ$         | $(0.431 \pm 0.004)$ |

Table 9.7: Relative hit rates due to LER Touschek for PXD modules from simulation [122] and for diamond radiation monitors from BGNet predictions.

| PXD Module | PXD rel. hit rate [122] | Dia. pos. in $\phi$ | BGNet rel. hit rate   |
|------------|-------------------------|---------------------|-----------------------|
| 2.3.1      | $(1.0 \pm 0.5)$         | $60^\circ$          | $(1.000 \pm 0.006)$   |
| 2.5.1      | $(0.7 \pm 0.2)$         | $120^\circ$         | $(0.125 \pm 0.001)$   |
| 2.7.1      | $(0.3 \pm 0.1)$         | $180^\circ$         | $(0.0010 \pm 0.0001)$ |
| 2.9.1      | $(0.3 \pm 0.1)$         | $240^\circ$         | $(0.0010 \pm 0.0001)$ |
| 2.11.1     | $(0.7 \pm 0.2)$         | $300^\circ$         | $(0.434 \pm 0.003)$   |

Table 9.8: Relative hit rates due to HER Touschek for PXD modules from simulation [122] and for diamond radiation monitors from BGNet predictions.

large, such that all  $\phi$ -profile would be compatible with an almost flat distribution (which is exhibited by the BGNet predictions apart from the diamond radiation monitor at  $\phi = 300^\circ$ ), the disagreement for LER Touschek could also be due to BGNet. The LER beam-gas dominates the decompositions for all diamonds in this analysis during the relevant time frame and makes up for half or more of the total hit rate. It is possible that due to the difficulties with differentiating the storage backgrounds of one ring, that hit rates actually originating from LER beam-gas contributions leak into the LER Touschek prediction. The LER Touschek component is a much smaller contributor to the total hit rate and the associated  $\phi$ -profile could be dominated by the wrongly assigned LER beam-gas hit rates. This is supported by the fact that LER Touschek  $\phi$ -profile has similar properties compared to HER beam-gas showing a relatively flat distribution, with the exception of the diamond radiation monitor at  $\phi = 300^\circ$ .

The  $\phi$ -profiles for the injection background cannot be compared as there are no simulations available, however, they exhibit a very clear  $\phi$ -dependence. The injection background components can be differentiated well from the other beam background components due to their time-dependent structure. As a result, the issue of ambiguities when differentiating them from other beam background contributions is much smaller and the contributions of other beam background components to the  $\phi$ -profiles of the injection backgrounds are much smaller. The observed  $\phi$ -dependence presents an interesting finding, as there are only few descriptions of the injection background. This could help to devise beam background mitigation strategies for one of the largest beam background components at large currents, that are needed

to reach large instantaneous luminosities.

### 9.8.6 Beam Loss Prediction Performance

Two aspects of the HER and LER beam loss prediction performance are considered in the following: The prediction performance on training data that is relevant for offline analyses and the prediction performance on data not seen during the training that is relevant for the real-time predictions. Only the total prediction of the beam loss will be considered here, as there are no comparisons or tests available for the decomposition of the beam loss.

The performance metric is chosen to be the ratio between the predicted total beam loss and the measured total beam loss. Although the beam loss is not directly measured as it is computed from the measured current as described in Section 9.5, it is referred to as “measured” for the purposes of this section. The beam loss model used by BGNet only considers particle losses due to storage backgrounds. This implies that the measured beam loss is only valid when there are no active injections. As this is a limitation of the used model and not of the beam loss values itself. This would be true for directly measured beam losses as well. The ratio of the measured and predicted total beam loss is therefore only computed while there are no active injections, despite the fact that predictions would be available at all times, predicting the beam loss due to storage backgrounds during active injections.

As before, the performance on training data is expected to be better than on unseen data. This is exploited during offline analyses, where a model, trained on the relevant time window, is available. A total of 40 models are used for the performance analysis of HER and LER beam loss predictions. Each model was trained on three weeks of SuperKEKB operation. The training samples can overlap between different models but no training set is equal to another. For the evaluation of the offline performance, the models are evaluated in a 1 h time window that is positioned arbitrarily within the training data set. The ratios are calculated every second and then resampled to 5 min by taking the mean. The results are displayed in Figure 9.32.

The median of the ratios during the test time window is close to unity with small fluctuations in the range of 1 % to 2 %. The interquartile ranges span 3 % to 5 %. A few outliers in both directions are observed. There are no discernible trends visible for an improvement or worsening of the prediction performance, which is expected, since all samples are part of the training data. Single models with comparatively worse performance that under- or overestimate the total beam loss over the whole test window account for most of the outliers.

The performance evaluation for the real-time predictions of the total beam loss is performed in a similar way as for the hit rate models in Section 9.8.3. For the real-time predictions, the BGNet models need to extrapolate and produce predictions for data samples that were not seen during the training. The metric for the performance is the ratio of the measured and predicted total beam loss. The training time for the beam loss models is around  $T_{\text{train, beam loss}} = 4$  min. However, the training cycle of

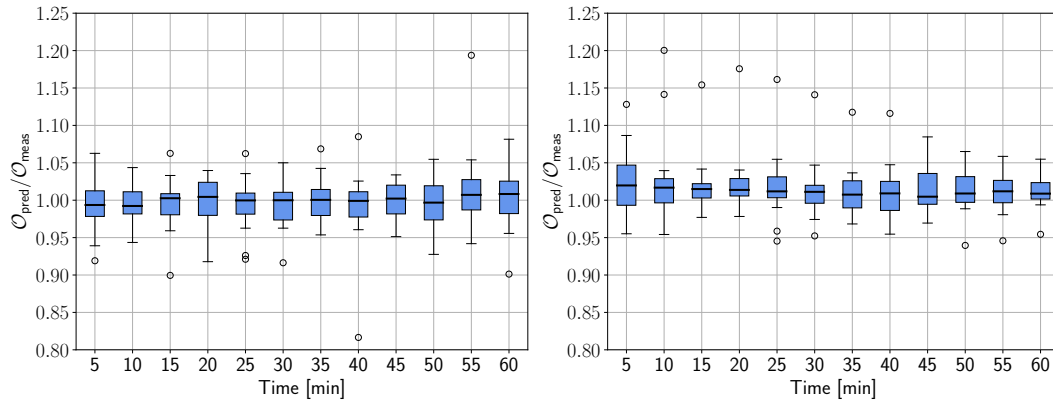


Figure 9.32: Ratio of the predicted beam loss  $\mathcal{O}_{\text{pred}}$  and measured beam loss  $\mathcal{O}_{\text{meas}}$  for 40 different models. The predictions were performed on data that was seen during the training. The ratios are calculated every second and then resampled to 5 min, therefore each box represents 40 data points of resampled ratios for the respective time step. The data samples used for the prediction are taken from an arbitrary time window of 1 h within the training data. The box plot follows standard convention where the solid line in the box is the median, the edges of the box extend to the first (Q1) and third quartile (Q3), the whiskers extend the inter quartile range by a factor 1.5 and outliers are explicitly plotted below and above the ends of the whiskers. Left: Ratios for the HER beam loss. Bottom Left: Ratios for the LER beam loss.

the real-time application of BGNet only begins after all models have been trained, and therefore a new model training will only begin after the slowest model training concluded. The slowest model is the hit rate model with an assumed training time of  $T_{\text{train}} = 1$  h. The training cycle of the hit rate and beam loss models is visualized in Figure 9.33. The training time of the beam loss models is small compared to the training time of the hit rate models. The relevant time window of unseen samples for the real-time predictions therefore starts after the training at  $\approx t_{\text{end}}$  until  $\approx t_{\text{end}} + T_{\text{train}}$ . The extrapolation performance has been tested with samples from the end of the training until 1 hour after the end of the training and a second time window from that extends until 2 hours after the end of the training. The results are displayed in Figure 9.34.

The predictions of the HER beam loss exhibit a small bias towards ratios smaller 1 within the first hour of extrapolation but median is still close to unity. Towards the end of the extrapolation time window the median moves to approximately 97 %. Similarly, the interquartile ranges are tending towards lower values, indicating an asymmetry in the distribution of the ratios. The width of the interquartile range is observed to reach  $\approx 8$  %. There is no strong degradation of the performance of the predicted total HER beam loss observed in the time window relevant for the real-time predictions or even beyond that. Similar observations can be made for the predictions of the LER beam loss. The median is close to unity during the entire evaluation time window with the largest deviations of the median being around 2 % from unity with no discernible bias in either direction. The interquartile ranges reaching widths of up to  $\approx 7$  %. There are only few outliers observed for both the LER beam loss and the HER beam loss. Most of the outliers belong to the same model that has a bad performance during the evaluation time window.

This study shows that BGNet can also be applied on the prediction of the beam loss, and therefore the beam lifetime, with a satisfactory performance that is similar to the performance obtained for the hit rate predictions.

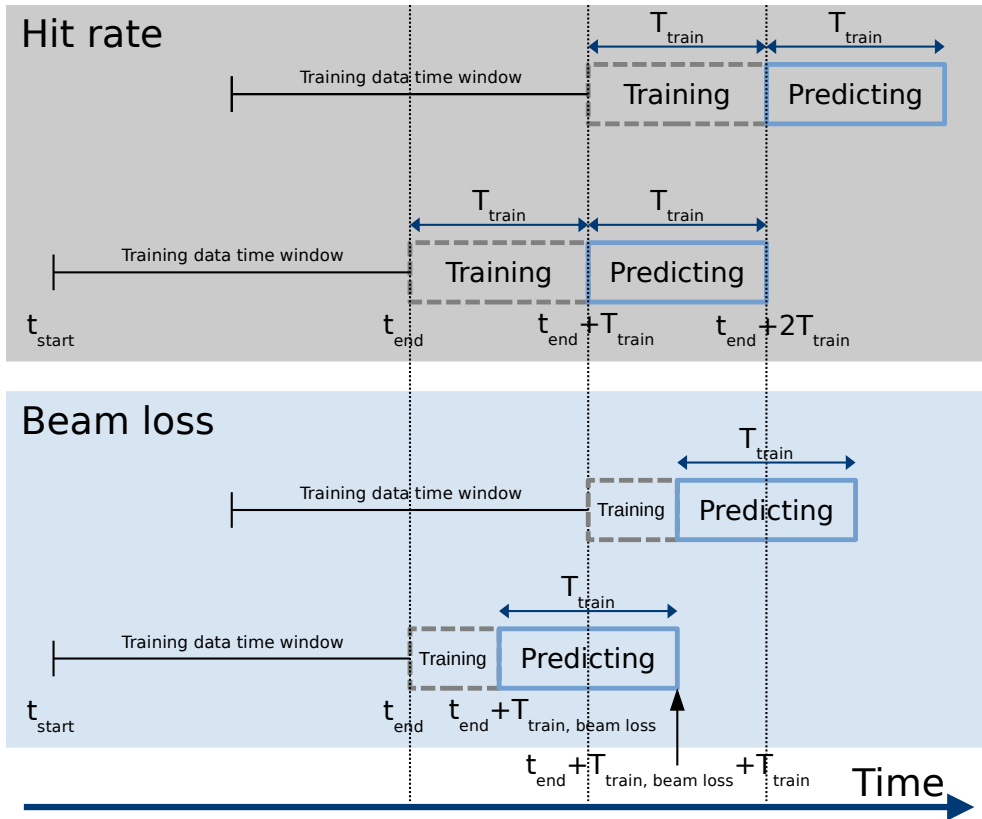


Figure 9.33: Training scheme of the BGNet beam loss models in relation to the hit rate model training for the real-time application of BGNet. The time window for the training data is defined from  $t_{\text{start}}$  to  $t_{\text{end}}$ . The training of all models begins at the same time and the next cycle only starts once all models are trained. The beam loss models are deployed from  $t_{\text{end}} + T_{\text{train, beam loss}}$  until  $t_{\text{end}} + T_{\text{train, beam loss}} + T_{\text{train}}$ , where  $T_{\text{train}}$  refers to the slowest model training, which is a hit rate model.

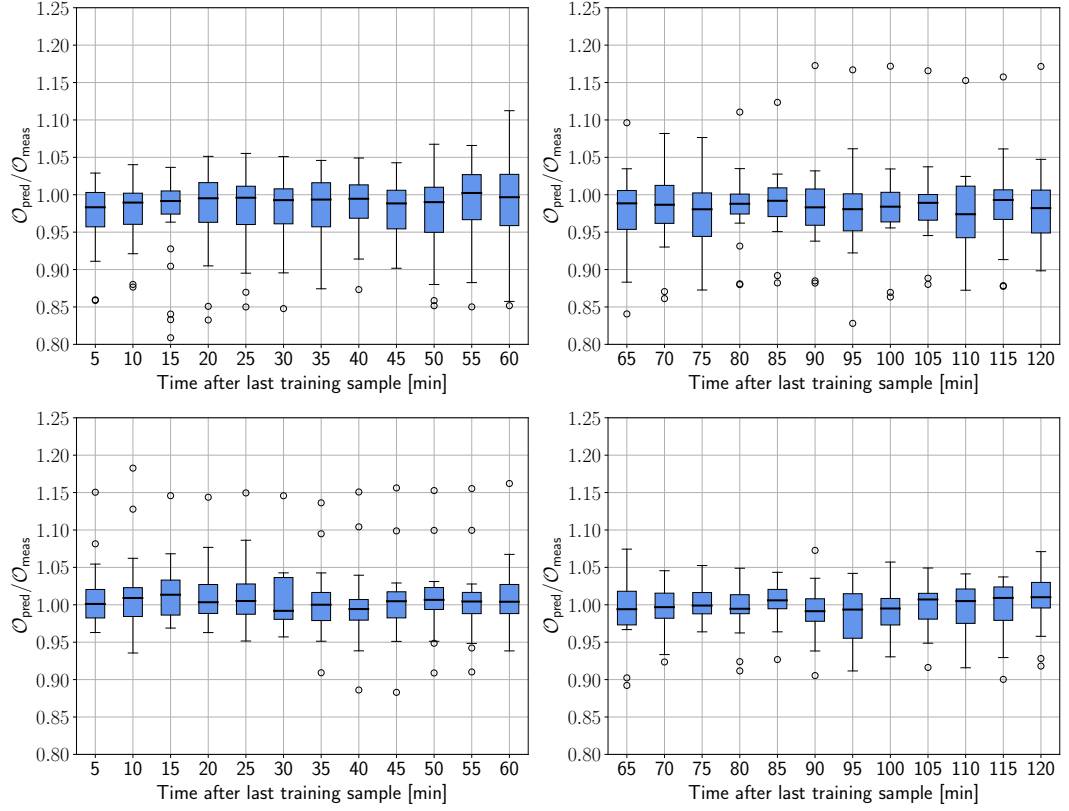


Figure 9.34: Ratio of the predicted beam loss  $\mathcal{O}_{\text{pred}}$  and measured beam loss  $\mathcal{O}_{\text{meas}}$ . The predictions were performed on data that was not seen during the training and the BGNet model is extrapolating. The ratios are calculated every second and then resampled to 5 min, therefore each box represents 40 data points, one for each model, of resampled ratios for the respective time step. The data samples used for the prediction are the ones immediately following the last sample seen during the training. The box plot follows standard convention where the solid line in the box is the median, the edges of the box extend to the first (Q1) and third quartile (Q3), the whiskers extend the inter quartile range by a factor 1.5 and outliers are explicitly plotted below and above the ends of the whiskers. Top Left: Ratios for the first hour after the training for the HER beam loss. This time window is relevant for the performance of the real-time predictions. Top Right: Ratios for the second hour after the training for the HER beam loss. Bottom Left: Ratios for the first hour after the training for the LER beam loss. This time window is relevant for the performance of the real-time predictions. Bottom Right: Ratios for the second hour after the training for the LER beam loss.



## CHAPTER 10

---

### Conclusion

---

As the instantaneous luminosity delivered by SuperKEKB increases, radiation due beam backgrounds also increases further. This causes a faster degradation of Belle II detector components and increases the difficulty of pattern recognition and reconstruction of signal processes. Excessive beam background levels will pose a major issue that can prevent the experiment to reach the design luminosity.

Monitoring and mitigating beam backgrounds is an important task for both the Belle II and SuperKEKB collaborations. In particular, a new neural network based tool, called BGNet, has been introduced as a diagnostic monitoring tool. BGNet provides predictions of the total hit rate and their decomposition due to beam background for multiple Belle II sub-detectors in real-time. Integrated gradients have been applied as feature attribution method to quantify the impact of individual input variables on the predicted hit rate, in an effort to increase the understanding of the beam background.

The author has participated in the development of BGNet and has conducted extensive performance studies for its commissioning. The performance of BGNet hit rate predictions has been quantified on training data and data not seen during training. BGNet has shown excellent performance in both cases. The beam background decompositions have been compared to beam background decompositions obtained during designated study campaigns and from simulations. During these comparisons BGNet provided compatible results. The quantitative comparisons that can be conducted are limited as simulations and designated study campaigns only provide data for certain configurations of SuperKEKB. BGNet offers predictions during general physics data taking.

The author has developed a real-time application for BGNet that allows for hit rate predictions and beam background decompositions to be available to the collaboration in real-time. This is a distinct advantage over other existing methods

of obtaining beam background decomposition, as they are only available for specific beam background studies. Further, a chat bot for offline analyses has been introduced that delivers detailed reports that summarize the beam background performance for a given time window. The chat bot is accessible for everyone, does not require expert knowledge and reports are generated within less than a minute time. BGNet has also been applied to predict the beam loss and its decomposition. The performance of the beam loss predictions has been found to be excellent as well.

A Belle II upgrade is planned for 2032 and a replacement for the current vertex detector is proposed. The new detector, VTX, is based on a fully pixelated design, improving the performance and tolerance against large beam backgrounds. The OBELIX DMAPS chip that will be used for the VTX is currently in the design stage and based on the TJ-Monopix2 chip. A test beam has been conducted to quantify the effects of irradiation and temperature on the performance of the TJ-Monopix2 chip, define operational temperature ranges and test software/firmware features that are required for the eventual use in the Belle II environment. The VTX is planned to be operated with minimal cooling at room temperature.

The author took part in the measurements and has conducted the test beam analysis. Stable operation with hit-detection efficiencies  $> 99\%$  with a threshold of  $200\ e^-$  at  $20^\circ$  NTC temperature and with a threshold of  $350\ e^-$  at  $40^\circ$  NTC temperature with the DC-coupled front-end at the highest fluence. The highest fluence is the fluence expected after 10 years operation and the design luminosity. The AC-coupled front-end showed larger degradations at high temperatures. The spatial resolution was determined to be  $(9.2 \pm 0.1)\ \mu\text{m}$  along the  $x$ -axis and  $(9.2 \pm 0.1)\ \mu\text{m}$  along the  $y$ -axis for a irradiated TJ-Monopix2 sample with the largest fluence.

A simulation of the TJ-Monopix2 chip has been performed, recreating the conditions during the test beam to confirm that the obtained results are sensible. The simulation results for the cluster properties, the spatial resolution and the hit detection efficiency have been presented. The cluster size and seed charge could be reproduced and are similar between the test beam measurements and the simulation. The spatial resolution for the simulation is smaller with  $(8.3 \pm 0.1)\ \mu\text{m}$  along the  $x$ -axis and  $(8.3 \pm 0.1)\ \mu\text{m}$  compared to what was achieved during the test beam. The simulated hit detection efficiency has been found to be too optimistic at higher thresholds compared to the measurement due to an underestimation of smaller pixel charges of  $\approx 500\ e^-$ . The difference between the simulation to the measurements is explained by the simplifications made during the simulation. The simulation lacks an explicit simulation of the pixel front end, doping concentrations and realistic electric fields. This simulation represents a baseline with a simple setup, making several generalizations and using generic assumptions. It will be tested against more sophisticated and complicated simulation approaches that are developed by the collaboration to gauge the impact of the changes made. The test beam measurement results confirm the feasibility of operating the VTX at room temperatures in Belle II. The DC-coupled front end remains highly efficient at these temperatures and after irradiation. The OBELIX design is planned to conclude in early 2026 and the

project moves towards production and initial measurements of the OBELIX chips towards the end of 2026.



---

## Bibliography

---

- [1] S. L. Glashow and S. Weinberg. “Natural Conservation Laws for Neutral Currents”. In: *Phys. Rev. D* 15 (1977), p. 1958.
- [2] A. Salam. “Weak and Electromagnetic Interactions”. In: *Conf. Proc. C* 680519 (1968), pp. 367–377.
- [3] S. Weinberg. “A Model of Leptons”. In: *Phys. Rev. Lett.* 19 (1967), pp. 1264–1266.
- [4] G. 't Hooft and M. J. G. Veltman. “Regularization and Renormalization of Gauge Fields”. In: *Nucl. Phys. B* 44 (1972), pp. 189–213.
- [5] P. W. Higgs. “Broken symmetries, massless particles and gauge fields”. In: *Phys. Lett.* 12 (1964), pp. 132–133.
- [6] P. W. Higgs. “Broken Symmetries and the Masses of Gauge Bosons”. In: *Phys. Rev. Lett.* 13 (1964), pp. 508–509.
- [7] P. W. Higgs. “Spontaneous Symmetry Breakdown without Massless Bosons”. In: *Phys. Rev.* 145 (1966), pp. 1156–1163.
- [8] G. Aad et al. “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”. In: *Phys. Lett. B* 716 (2012), pp. 1–29. arXiv: 1207.7214 [hep-ex].
- [9] E. Fermi. “Zur Quantelung des idealen einatomigen Gases”. In: *Z Phys* 36 (1926), pp. 902–912.
- [10] S. Bose. “Wärmegleichgewicht im Strahlungsfeld bei Anwesenheit von Materie”. In: *Zeitschrift für Physik A Hadrons and Nuclei* 27 (1924), pp. 384–393.
- [11] S. Navas et al. “Review of particle physics”. In: *Phys. Rev. D* 110.3 (2024), p. 030001.

- [12] S.-I. Tomonaga and J. R. Oppenheimer. “On Infinite Field Reactions in Quantum Field Theory”. In: *Phys. Rev.* 74 (1948), pp. 224–225.
- [13] J. Schwinger. “Quantum Electrodynamics. III. The Electromagnetic Properties of the Electron—Radiative Corrections to Scattering”. In: *Phys. Rev.* 76 (6 1949), pp. 790–817.
- [14] R. P. Feynman. “Space-time approach to nonrelativistic quantum mechanics”. In: *Rev. Mod. Phys.* 20 (1948), pp. 367–387.
- [15] G. Altarelli. “A QCD primer”. In: *AIP Conf. Proc.* 631.1 (2002), p. 70.
- [16] R. Aaij et al. “Observation of a  $J/\psi\Lambda$  Resonance Consistent with a Strange Pentaquark Candidate in  $B\rightarrow J/\psi\Lambda p^-$  Decays”. In: *Phys. Rev. Lett.* 131.3 (2023), p. 031901. arXiv: 2210.10346 [hep-ex].
- [17] B. Andersson et al. “Parton Fragmentation and String Dynamics”. In: *Phys. Rept.* 97 (1983), pp. 31–145.
- [18] S. L. Glashow and others. “Weak Interactions with Lepton-Hadron Symmetry”. In: *Phys. Rev. D* 2 (1970), pp. 1285–1292.
- [19] M. Kobayashi and T. Maskawa. “CP Violation in the Renormalizable Theory of Weak Interaction”. In: *Prog. Theor. Phys.* 49 (1973), pp. 652–657.
- [20] N. Cabibbo. “Unitary Symmetry and Leptonic Decays”. In: *Phys. Rev. Lett.* 10 (1963), pp. 531–533.
- [21] S. W. Herb et al. “Observation of a Dimuon Resonance at 9.5 GeV in 400 GeV Proton-Nucleus Collisions”. In: *Phys. Rev. Lett.* 39 (1977), pp. 252–255.
- [22] S. Abachi et al. “Observation of the top quark”. In: *Phys. Rev. Lett.* 74 (1995), pp. 2632–2637. arXiv: hep-ex/9503003.
- [23] E. Kou et al. “The Belle II Physics Book”. In: *Progress of Theoretical and Experimental Physics* 2019.12 (2019).
- [24] J. Charles et al. “CP violation and the CKM matrix: Assessing the impact of the asymmetric  $B$  factories”. In: *Eur. Phys. J. C* 41.1 (2005), pp. 1–131. arXiv: hep-ph/0406184.
- [25] C. Group. *CKMfitter*. 2025. URL: <http://ckmfitter.in2p3.fr/> (visited on 07/28/2025).
- [26] I. Adachi et al. “Precise measurement of the CP violation parameter  $\sin 2\phi_1$  in  $B^0 \rightarrow (c\bar{c})K^0$  decays”. In: *Phys. Rev. Lett.* 108 (2012), p. 171802. arXiv: 1201.4643 [hep-ex].
- [27] Y. Ohnishi et al. “Accelerator design at SuperKEKB”. In: *PTEP* 2013 (2013), 03A011.
- [28] S. Banerjee et al. “Averages of  $b$ -hadron,  $c$ -hadron, and  $\tau$ -lepton properties as of 2023”. In: (2024). arXiv: 2411.18639 [hep-ex].

- [29] H. Wiedemann. *Particle Accelerator Physics*. ISBN: 978-3-319-18316-9, DOI: 10.1007/978-3-319-18317-6. 2015.
- [30] “Particle Physics Reference Library . Volume 3: Accelerators and Colliders”. In: ed. by S. Myers and H. Schopper. Springer, 2020.
- [31] J. D. Cockcroft and E. T. S. Walton. “Experiments with High Velocity Positive Ions. (I) Further Developments in the Method of Obtaining High Velocity Positive Ions”. In: *Proc. Roy. Soc. Lond. A* 136 (1932), pp. 619–630.
- [32] R. J. Van de Graaff. “A 1,500,000 volt electrostatic generator”. In: *Phys. Rev* 38 (1931), p. 1919.
- [33] R. Wideröe. “Über ein neues Prinzip zur Herstellung hoher Spannungen”. In: *Arch. Elektrotech.* 21.4 (1928), pp. 387–406.
- [34] K. Becks et al. “Commissioning of the DELPHI Pixel Detector”. In: *Nucl. Instrum. Meth. A* 418.1 (1998), pp. 15–21.
- [35] G. Eigen. *Detectors in High-Energy Physics Experiments*. ISBN: 978-3-319-18316-9, DOI: 10.1007/978-3-031-67336-8. 2025.
- [36] H. Kolanoski and N. Wermes. *Particle Detectors: Fundamentals and Applications*. Oxford University Press, 2020.
- [37] H. Kolanoski and N. Wermes. *Teilchendetektoren*. ISBN: 978-3-662-45350-6, DOI: 10.1007/978-3-662-45350-6. 2016.
- [38] H. Bethe. “Theory of the Passage of Fast Corpuscular Rays Through Matter”. In: *Annalen Phys.* 5 (1930), pp. 325–400.
- [39] H. Bethe and J. Ashkin. *Experimental Nuclear Physics*. Ed. by Segré, Ea. New York: Wiley, 1953.
- [40] L. D. Landau. “On the Energy Loss of Fast Particles by Ionisation”. In: *J. Phys. (USSR)* 8 (1944).
- [41] D. Caughey and R. Thomas. “Carrier mobilities in silicon empirically related to doping and field”. In: *Proceedings of the IEEE* 55.12 (1967), pp. 2192–2193.
- [42] W. Shockley. “Currents to conductors induced by a moving point charge”. In: *J. Appl. Phys.* 9.10 (1938), pp. 635–636.
- [43] S. Ramo. “Currents induced by electron motion”. In: *Proc. Ire.* 27 (1939), pp. 584–585.
- [44] M. Moll. “Radiation damage in silicon particle detectors: Microscopic defects and macroscopic properties”. PhD thesis. Hamburg U., 1999.
- [45] G. Kramberger. “Particle Physics Reference Library: Volume 2: Detectors for Particles and Radiation”. In: 2020. Chap. 21, pp. 965–1034.
- [46] F. Hartmann. “Evolution of silicon sensor technology in particle physics”. In: 2nd ed. Springer tracts in modern physics. Springer International Publishing, 2017.

- [47] G. Lindstrom. “Radiation damage in silicon detectors”. In: *Nucl. Instrum. Meth. A* 512 (2003). Ed. by P. Holl et al., pp. 30–43.
- [48] R. Wunstorf. “Systematische Untersuchungen zur Strahlenresistenz von Silizium-Detektoren für die Verwendung in Hochenergiephysik-Experimenten”. PhD thesis. Hamburg: University of Hamburg, 1992.
- [49] G. Lutz and R. Klanner. “Particle Physics Reference Library: Volume 2: Detectors for Particles and Radiation”. In: 2020. Chap. 5, pp. 439–484.
- [50] M. Babeluk. “Development and Testing of Depleted Monolithic Active Pixel Sensors in 180nm CMOS technology for the Belle II VTX Upgrade”. PhD thesis. Technische Universität Wien, 2025.
- [51] C. Bepin Valero. “Characterization of the TJ-Monopix2 Depleted Monolithic Active Pixel Sensor for High-Energy Physics Experiments”. PhD thesis. Rheinische Friedrich-Wilhelms-Universität Bonn, 2024.
- [52] C. Bepin et al. “DMAPS Monopix developments in large and small electrode designs”. In: *Nucl. Instrum. Meth. A* 978 (2020), p. 164460.
- [53] F. Hügging et al. “Recent results and perspectives of the Monopix Depleted Monolithic Active Pixel Sensors (DMAPS)”. In: *Nucl. Instrum. Meth. A* 1070 (2025), p. 170007.
- [54] K. Moustakas. “Design and Development of Depleted Monolithic Active Pixel Sensors with Small Collection Electrode for High-Radiation Applications”. PhD thesis. Rheinische Friedrich-Wilhelms-Universität Bonn, 2021.
- [55] Z. J. Liptak et al. “Measurements of beam backgrounds in SuperKEKB Phase 2”. In: *Nucl. Instrum. Meth. A* 1040 (2022), p. 167168. arXiv: 2112.14537 [physics.ins-det].
- [56] K. Akai, K. Furukawa, and H. Koiso. “SuperKEKB Collider”. In: *Nucl. Instrum. Meth. A* 907 (2018), pp. 188–199. arXiv: 1809.01958 [physics.acc-ph].
- [57] *Schematic of SuperKEKB*. <https://www.kek.jp/en/NewsRoom/Release/20160302163000/>. Accessed: 2022-09-14.
- [58] T. Sugimura et al. “The Linac Upgrade Plan for SuperKEKB”. In: *Conf. Proc. C* 100523 (2010). Ed. by A. Noda et al., THPD007.
- [59] K. Hirosawa et al. “Development of a Longitudinal Feedback System for Coupled Bunch Instabilities Caused by the Accelerating Mode at Superkekb”. In: *8th International Particle Accelerator Conference*. 2017.
- [60] T. Abe et al. “Belle II Technical Design Report”. In: (2010). arXiv: 1011.0352 [physics.ins-det].
- [61] T. Hara, T. Kuhr, and Y. Ushiroda. “Belle II Coordinate System and Guideline of Belle II Numbering Scheme”. In: *Technical Note* (2011).
- [62] L. Andricek. “Operation Status of SuperKEKB”. In: *Belle II VXD open workshop - CERN* (2019).



- [63] H. Bae et al. “A New Scheme for the Injection Veto System of the Belle II Experiment”. In: *2024 IEEE Nuclear Science Symposium (NSS), Medical Imaging Conference (MIC) and Room Temperature Semiconductor Detector Conference (RTSD)*. 2024.
- [64] T. Geßler et al. “The ONSSEN Data Reduction System for the Belle II Pixel Detector”. In: *IEEE Trans. Nucl. Sci.* 62.3 (2015), pp. 1149–1154. arXiv: 1406.4028 [physics.ins-det].
- [65] D. Sun et al. “Belle2Link: A Global Data Readout and Transmission for Belle II Experiment at KEK”. In: *Phys. Procedia* 37 (2012). Ed. by T. Liu, pp. 1933–1939.
- [66] T. Kuhr et al. “The Belle II Core Software”. In: *Computing and Software for Big Science* 3.1 (2018), p. 1.
- [67] S. Bacher et al. “Performance of the diamond-based beam-loss monitor system of Belle II”. In: *Nucl. Instrum. Meth. A* 997 (2021), p. 165157. arXiv: 2102.04800 [physics.ins-det].
- [68] *Experimental Physics and Industrial Control System.*
- [69] S. Park et al. “Environmental Monitoring for Belle II”. In: *21st IEEE Real Time Conference*. 2018. arXiv: 1806.08933 [physics.ins-det].
- [70] *The EPICS Archiver Appliance.*
- [71] Belle II Collaboration. *The Belle II Detector Upgrades Framework Conceptual Design Report*. 2024. arXiv: 2406.19421 [hep-ex].
- [72] V. Bertacchi et al. “Track finding at Belle II”. In: *Comput. Phys. Commun.* 259 (2021), p. 107610. arXiv: 2003.12466 [physics.ins-det].
- [73] M. Babeluk et al. “The OBELIX chip for the Belle II VTX upgrade”. In: *Nucl. Instrum. Meth. A* 1067 (2024), p. 169659.
- [74] S. Bettarini. “The Belle2-VXD Running Experience and the Mechanical Integration of the DMAPS VTX-Upgrade in the new I.R. Design”. In: *'FCC-ee vertex detector R&D workshop'* (2025).
- [75] B. Abelev et al. “Technical Design Report for the Upgrade of the ALICE Inner Tracking System”. In: *J. Phys. G* 41 (2014), p. 087002.
- [76] M. Schwickardi. “Characterisation of CMOS Sensors for the Belle II Vertex Detector Upgrade”. PhD thesis. Georg-August-University Göttingen, 2024.
- [77] Y. Buch et al. “Upgrade of the Belle II vertex detector with depleted monolithic CMOS active pixel sensors”. In: *JINST* 20.12 (2025), p. C12026.
- [78] R. Diener et al. “The DESY II test beam facility”. In: *Nucl. Instrum. Meth. A* 922 (2019), pp. 265–286.
- [79] H. Jansen et al. “Performance of the eudet-type beam telescopes”. In: *EPJ Techniques and Instrumentation* 3.1 (2016).

- [80] J. Baudot et al. “First test results Of MIMOSA-26, a fast CMOS sensor with integrated zero suppression and digitized output”. In: *2009 IEEE Nuclear Science Symposium Conference Record (NSS/MIC)*. 2009, pp. 1169–1173.
- [81] L. Huth et al. “TelePix2: Full scale fast region of interest trigger and timing for the EUDET-style telescopes at the DESY II test beam facility”. In: *Nucl. Instrum. Meth. A* 1080 (2025), p. 170720.
- [82] P. Baesso, D. Cussans, and J. Goldstein. “The AIDA-2020 TLU: a flexible trigger logic unit for test beam facilities”. In: *Journal of Instrumentation* 14.09 (2019), P09019.
- [83] Y. Liu et al. “EUDAQ2—A flexible data acquisition software framework for common test beams”. In: *Journal of Instrumentation* 14.10 (2019), P10033.
- [84] M. Daas et al. “BDAQ53, a versatile pixel detector readout and test system for the ATLAS and CMS HL-LHC upgrades”. In: *Nucl. Instrum. Meth. A* 986 (2021), p. 164721.
- [85] Silizium Labor Bonn. *basil - A data acquisition framework in Python and Verilog*. 1999. URL: <https://github.com/SiLab-Bonn/basil>.
- [86] J. Kröger and L. Huth. “Efficient Analysis of Test-beam Data with the Corryvreckan Framework”. In: *JPS Conf. Proc.* 34 (2021). Proceedings for Vertex2020, 8 pages, 5 figures, p. 010024. arXiv: 2011.09205.
- [87] P. Wieduwilt et al. “Performance of production modules of the Belle II pixel detector in a high-energy particle beam”. In: *Nucl. Instrum. Meth. A* 991 (2021), p. 164978.
- [88] U. Stolzenberg. “Radiation length measurements with high-resolution telescopes”. PhD thesis. Georg-August Univ. Goettingen, 2019.
- [89] B. Schwenker. *Test beam software framework (tbsw)*.
- [90] R. Turchetta. “Spatial resolution of silicon microstrip detectors”. In: *Nucl. Instrum. Meth. A* 335.1 (1993).
- [91] C. Kleinwort. “General broken lines as advanced track fitting method”. In: *Nucl. Instrum. Meth. A* 673 (2012), pp. 107–110.
- [92] V. Blobel and C. Kleinwort. *A New Method for the High-Precision Alignment of Track Detectors*. 2002. arXiv: [hep-ex/0208021](https://arxiv.org/abs/hep-ex/0208021) [[hep-ex](https://arxiv.org/abs/hep-ex/0208021)].
- [93] *Public analysis repository for the 2025 test beam at DESY*. [https://github.com/ybuch/corry\\_config\\_desytb\\_2025](https://github.com/ybuch/corry_config_desytb_2025).
- [94] S. Spannagel and H. Jansen. *GBL Track Resolution Calculator v2.0*. 2016.
- [95] C. J. CLOPPER and E. S. PEARSON. “THE USE OF CONFIDENCE OR FIDUCIAL LIMITS ILLUSTRATED IN THE CASE OF THE BINOMIAL”. In: *Biometrika* 26.4 (1934).
- [96] R. Brun and F. Rademakers. “ROOT: An object oriented data analysis framework”. In: *Nucl. Instrum. Meth. A* 389 (1997), pp. 81–86.

- [97] S. Spannagel et al. “Allpix<sup>2</sup>: A Modular Simulation Framework for Silicon Detectors”. In: *Nucl. Instrum. Meth. A* 901 (2018), pp. 164–172. arXiv: 1806.05813 [physics.ins-det].
- [98] S. Agostinelli et al. “Geant4—a simulation toolkit”. In: *Nucl. Instrum. Meth. A* 506.3 (2003), pp. 250–303.
- [99] G. Kramberger et al. “Effective trapping time of electrons and holes in different silicon materials irradiated with neutrons, protons and pions”. In: *Nucl. Instrum. Meth. A* 481.1 (2002).
- [100] Martín Abadi et al. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. Software available from tensorflow.org. 2015.
- [101] F. Chollet et al. *Keras*. <https://keras.io>. 2015.
- [102] B. Schwenker. *BGNet*. [https://gitlab.desy.de/benjamin.schwenker/bg\\_net](https://gitlab.desy.de/benjamin.schwenker/bg_net). 2022.
- [103] B. Schwenker et al. “A neural network for beam background decomposition in Belle II at SuperKEKB”. In: *Nucl. Instrum. Meth. A* 1049 (2023), p. 168112.
- [104] I. Goodfellow et al. *Deep Learning*. <http://www.deeplearningbook.org>. MIT Press, 2016.
- [105] L. Bottou. “Online Algorithms and Stochastic Approximations”. In: *Online Learning and Neural Networks*. Cambridge University Press, 1998.
- [106] D. P. Kingma et al. *Adam: A Method for Stochastic Optimization*. 2017. arXiv: 1412.6980 [cs.LG].
- [107] P. M. Lewis et al. “First Measurements of Beam Backgrounds at SuperKEKB”. In: *Nucl. Instrum. Meth. A* 914 (2019), pp. 69–144. arXiv: 1802.01366 [physics.ins-det].
- [108] A. Natochii et al. “Measured and projected beam backgrounds in the Belle II experiment at the SuperKEKB collider”. In: *Nucl. Instrum. Meth. A* 1055 (2023), p. 168550.
- [109] A. Wolski. *Beam Dynamics in High Energy Particle Accelerators*. World Scientific, 2023.
- [110] *Realizing High Luminosity at SuperKEKB*. 2021. eprint: <https://journals.jps.jp/doi/pdf/10.7566/JPSCP.33.011007>.
- [111] G. Lemaître et al. “Imbalanced-learn: A Python Toolbox to Tackle the Curse of Imbalanced Datasets in Machine Learning”. In: *Journal of Machine Learning Research* 18.17 (2017), pp. 1–5.
- [112] X. Glorot et al. “Understanding the difficulty of training deep feedforward neural networks”. In: vol. 9. 2010, pp. 249–256.
- [113] M. Sundararajan, A. Taly, and Q. Yan. *Axiomatic Attribution for Deep Networks*. 2017. arXiv: 1703.01365 [cs.LG].

- [114] G. Van Rossum and F. L. Drake. *Python 3 Reference Manual*. Scotts Valley, CA: CreateSpace, 2009.
- [115] *PCASpy*. <https://github.com/paulscherrerinstitute/pcaspy>.
- [116] *PyEpics: Epics Channel Access for Python*. <https://github.com/pyepics/pyepics>.
- [117] *TensorFlow Serving*. [https://www.tensorflow.org/tfx/tutorials/serving/rest\\_simple](https://www.tensorflow.org/tfx/tutorials/serving/rest_simple).
- [118] D. Merkel. “Docker: lightweight linux containers for consistent development and deployment”. In: *Linux journal* 2014.239 (2014), p. 2.
- [119] *Control System Studio*. <https://www.controlsystemstudio.org/>.
- [120] Mattermost, Inc. *Mattermost*. <https://github.com/mattermost/mattermost>. 2025.
- [121] S. Terui et al. “Collimator challenges at SuperKEKB and their countermeasures using nonlinear collimator”. In: *Phys. Rev. Accel. Beams* 27 (8 2024).
- [122] A. Moll. “Comprehensive study of the background for the Pixel Vertex Detector at Belle II”. 2015.
- [123] B. Spruck et al. “Belle II Pixel Detector Commissioning and Operational Experience”. In: *PoS Vertex2019* (2020), p. 015.
- [124] K. Hirata. “An Introduction to SAD”. In: *2nd ICFA Advanced Beam Dynamics Workshop*. 1988, pp. 62–65.

**CDC** Central drift chamber, gas tracking detector currently used by Belle II. 109, 147, 148, 150, 152, 156, 157, 160

**CMOS** Complementary metal–oxide–semiconductor, a technology used for constructing integrated circuits. The TJ-Monopix2 and OBELIX sensors are manufactured in a 180 nm process. 37, 61, 64

**DAC** Digital-to-Analog Converter, supplies voltages and currents. The TJ-Monopix2 has a 8 bit DACs, allowing to scale the respective output between 0 and 255. A DAC output can be expressed in terms of these 8 bit DAC values, or rather, in terms of DAC units. 75

**DAQ** Data Acquisition System. 73, 74

**DMAPI** Depleted monolithic active pixel sensor. 2, 61, 64, 174

**DUT** Device Under Test, refers to the sensor undergoing testing in the test beam measurements. 74, 77, 85, 87, 90, 103, 104, 106

**EPICS** Experimental physics and industrial control system, slow control system used by Belle II and SuperKEKB. 59, 60, 109, 115–118, 135, 136, 138–140, 160

**FIFO** First In - First Out. 136, 138

**HER** High Energy Ring, electron storage ring of SuperKEKB. 45, 49, 113–116, 123, 126, 127, 129, 132, 138, 150, 152, 156, 160, 166–168, 170

**IP** Interaction point, point of collision of the electron and positron beams. 49, 51, 58, 113, 163

- IR** Interaction region, region around the Belle II interaction point. 61
- LER** Low Energy Ring, positron storage ring of SuperKEKB. 45, 113–116, 126, 132, 138, 142, 145, 146, 160, 161, 166–170, 172, 197, 199, 200
- LINAC** LINear ACcelerator, pre-accelerator for SuperKEKB. 46, 49, 113, 131
- MAPS** Monolithic active pixel sensor. 37, 72
- MIMOSA26** Sensor used by the DURANTA EUDET-type telescope used during the beam test measurements at DESY. 72, 73, 98
- MPV** Most probable value. 100
- NIEL** Non-ionizing energy loss, mode of radiation damage pertaining to structural defects of the silicon bulk material in a silicon pixel detector. 36, 80, 90
- NTC** Negative-temperature-coefficient, thermal resistor for temperature measurements. 98, 99, 174
- OBELIX** Optimized BeLle Ii piXeL sensor, DMAPS pixel sensor developed for the Belle II vertex detector (VTX) upgrade. 2, 29, 37, 38, 61, 64–67, 71, 78, 174, 175
- PV** Process variable, real-time variables in the EPICS slow control system. 59, 60, 135, 136
- QCS** Final focusing quadrupole magnet system. 58, 61, 67
- RF** Radio frequency. Accelerator cavities operate at RF frequencies and have an acceptance, that refers to a longitudinal range of phase space in which particles are accelerated. 16, 17, 49, 113
- TLU** Trigger Logic Unit. 73, 74, 77, 98
- ToT** Time-Over-Threshold, unit for length of digital signal that is proportional to the signal charge collected in a pixel. 40, 75, 79, 99, 100
- VTX** Vertex detector for the Belle II Upgrade. 61, 62, 66, 67, 71, 174
- VXD** Current vertex detector of Belle II. 53, 58, 61

---

## List of Figures

---

|     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      |    |
|-----|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 2.1 | Artistic depiction of the particles in the Standard Model. Measurements are taken from Reference [11]. . . . .                                                                                                                                                                                                                                                                                                                                                                                                                       | 4  |
| 2.2 | Standard Model vertices of fermions coupling with bosons. Each vertex, and associated force, has an associated coupling strength denoted as $g_{\text{QED}}$ , $g_s$ , $g_W$ , $g_Z$ , which can be expressed as dimensionless coupling constants $\alpha$ , $\alpha_s$ , $\alpha_{W/Z}$ . Coupling constant values are at momentum transfers of $q^2 = 1 \text{ GeV}$ [11]. . . . .                                                                                                                                                 | 5  |
| 2.3 | A representation of the unitarity triangle in the complex plane [11]. . .                                                                                                                                                                                                                                                                                                                                                                                                                                                            | 9  |
| 2.4 | The current unitarity triangle fit determined with CKMFITTER [24]. The figure is taken from Reference [25]. . . . .                                                                                                                                                                                                                                                                                                                                                                                                                  | 9  |
| 2.5 | Drawing of two B meson decays with displaced vertices and their distance $\Delta z$ . The event has a tag side and a signal side. The flavor of the tag side B meson is determined by measuring the charge of the electron, allowing to implicitly infer the flavor of the signal side B meson. The apparent spatial separation and the directional bias of the particles of the signal and tag side benefit only the comprehensibility of the drawing and are not due to actual physical effects separating the two decays. . . . . | 10 |
| 2.6 | Averages of $\sin 2\beta = \sin 2\phi_1$ from different experiments. In cases where two errors are quoted for measurements, the first error is the statistical uncertainty and the second error is the systematic uncertainty. In cases, where only error is provided, it is identified as the combined error from statistical and systematic uncertainties [28]. . .                                                                                                                                                                | 12 |

|     |                                                                                                                                                                                                                                                                                                                                                      |    |
|-----|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 3.1 | Schematic of Wideroe drift tubes accelerating a charged particle. The changing voltage $V(t)$ is connected to the accelerating segments in an alternating pattern and at $T = 0$ (top) and $T = \pi$ (bottom), assuming that one period is equal to $2\pi$ . The voltage as a function of the position $x$ along the structure is indicated. . . . . | 15 |
| 3.2 | Sketch of the magnetic field produced by a quadrupole magnet. The direction of the magnetic field is indicated at selected points and the direction of the Lorentz force for an electron traversing the magnet vertically into the plane of the drawing. . . . .                                                                                     | 20 |
| 3.3 | FODO cell consisting of focusing and defocusing quadrupole magnets. The strength of the focusing magnet is typically chosen at half the strength of the defocusing magnet. The spaces in between the quadrupole magnets can be occupied by accelerating cavities. . . . .                                                                            | 21 |
| 3.4 | The betatron functions along a FODO cell. . . . .                                                                                                                                                                                                                                                                                                    | 24 |
| 4.1 | Band structure of insulators, semiconductors and metals. The semiconductor band structure is shown for temperatures of $T = 0$ K and $T > 0$ K. At $T > 0$ K electrons can be lifted from the valence band into the conduction band by thermal excitation. . . . .                                                                                   | 26 |
| 4.2 | Left: Silicon crystal at $T = 0$ K with an implanted boron atom introducing an open bound. Right: At $T > 0$ K an electron from a covalent silicon bound may fill the hole. The hole is mobile. . . . .                                                                                                                                              | 27 |
| 4.3 | Left: Silicon crystal at $T = 0$ K with an implanted phosphorus atom introducing an extra, weakly bound electron. Right: At $T > 0$ K the electron may be lifted into the conduction band, creating a mobile charge. . . . .                                                                                                                         | 28 |
| 4.4 | Band structure for p-type and n-type silicon semiconductors with the introduced acceptor and donor levels. . . . .                                                                                                                                                                                                                                   | 28 |
| 4.5 | Linear stopping power for muons in copper [11]. . . . .                                                                                                                                                                                                                                                                                              | 30 |
| 4.6 | Energy loss distributions for 500 MeV pions in silicon at different thicknesses. The distributions are Landau-shaped and are normalized to unity at the most probable value $\Delta_p/x$ . The width $w$ is the full width at half maximum [11]. . . . .                                                                                             | 31 |
| 4.7 | Energy loss for electrons in copper due to bremsstrahlung and ionization. The electron energy at which the energy loss due to ionization and bremsstrahlung is equal is indicated for different definitions. This critical energy is at $E_e \approx 20$ MeV [11]. . . . .                                                                           | 32 |
| 4.8 | Photon cross-sections for interactions with matter depending on the photon energy in carbon and lead [11]. . . . .                                                                                                                                                                                                                                   | 33 |
| 4.9 | Effects of bulk damage and the effects on the leakage current, depletion voltage and charge collection efficiency (CCE). . . . .                                                                                                                                                                                                                     | 36 |



|      |                                                                                                                                                                                                                                                                                                                                                     |    |
|------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 4.10 | Top: Sensor geometry of a TJ-Monopix2 DMAPS pixel. Bottom: Sensor geometry with indicated boundaries for the depletion region. Modified from [53, 54]. . . . .                                                                                                                                                                                      | 39 |
| 4.11 | Analog input signal with a signal amplitude larger than the discriminator threshold and the corresponding digital output signal with a length equal to the Time-over-Threshold. . . . .                                                                                                                                                             | 41 |
| 4.12 | Schematic of the TJ-Monopix2 front-end circuit. The shape of the signal is indicated throughout the circuit [54]. . . . .                                                                                                                                                                                                                           | 42 |
| 4.13 | S-curve measurement to determine the threshold using injections. . .                                                                                                                                                                                                                                                                                | 42 |
| 5.1  | Map of collimators in SuperKEKB. The arc sections D03, D06, D09, D12 are indicated. They, together with section D01, the Tsukuba straight section and location of Belle II, house most of the collimators installed in SuperKEKB [55]. . . . .                                                                                                      | 46 |
| 5.2  | Schematic depiction of the SuperKEKB collider, as well as the linear accelerator and the positron damping ring [57]. . . . .                                                                                                                                                                                                                        | 47 |
| 5.3  | Schematic illustration of the nano-beam scheme in the horizontal plane. The crossing angle is 83 mrad, the bunch length is $\approx 6$ mm and the horizontal bunch size is $\approx 10 \mu\text{m}$ [27]. . . . .                                                                                                                                   | 49 |
| 5.4  | Schematic depiction of the linear accelerator and the damping ring illustrating the different stages and the respective energy and charge of the electron and positron bunches [56]. . . . .                                                                                                                                                        | 50 |
| 6.1  | The Belle II detector and its coordinate system. The direction of the beams and the sub detectors are indicated. Modified from [61]. . . .                                                                                                                                                                                                          | 52 |
| 6.2  | CAD drawing of the current Belle II PXD with all 20 ladders installed [62]. . . . .                                                                                                                                                                                                                                                                 | 52 |
| 6.3  | CAD drawing of the current Belle II VXD. [62] . . . . .                                                                                                                                                                                                                                                                                             | 53 |
| 6.4  | THE ARICH and TOP sub detectors in an exploded view. . . . .                                                                                                                                                                                                                                                                                        | 55 |
| 6.5  | Cut-view of the ECL and KLM sub systems as well as components of the magnet system. . . . .                                                                                                                                                                                                                                                         | 57 |
| 6.6  | Positions of the diamond radiation monitors at the beam pipe bellows close to the QCS final focusing magnets of SuperKEKB, on the SVD support cone and the beam pipe. The angle indicates the position of the diamond in terms of the angle $\phi$ of the Belle II coordinate system, which is indicated in Figure 6.1. Modified from [67]. . . . . | 59 |
| 7.1  | Comparison of the occupancies of the PXD, SVD, CDC and VTX and the track efficiency for different tracking setups at different background scenarios [71]. . . . .                                                                                                                                                                                   | 62 |

|     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                   |    |
|-----|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 7.2 | Reconstruction efficiency of $\bar{D}^0 \rightarrow K^+\pi^-$ final state as function of the transverse momentum of the soft pion $p_T(\pi_{\text{soft}}^-)$ and the cosine of the polar angle $\theta$ . The current Belle II detector with background SC2 is compared with the upgraded VTX and all background scenarios. The bottom plot shows the the ratio of the efficiency for the Belle II geometry and the VTX geometry for background SC2 [71]. . . . . | 63 |
| 7.3 | Layout of the TJ-Monopix2. The $512 \times 512$ pixels active matrix is divided into four areas with different FEs [54]. . . . .                                                                                                                                                                                                                                                                                                                                  | 64 |
| 7.4 | Top left: 3D cut-view of the 5 layer VTX [71]. Top right: Cross section of the 5 layer VTX [71]. Bottom: Cross section of the 6 layer VTX. . . . .                                                                                                                                                                                                                                                                                                                | 68 |
| 8.1 | Schematic of the test beam setup with the DURANTA EUDET-type telescope, the TJ-Monopix2 DUT and a TelePix2 timing reference plane. The scintillator in front of the telescope produces a trigger signal once a particle passes, which is then distributed by the TLU to the respective DAQ systems. From the run control machine the setup is configured and controlled. . . . .                                                                                  | 73 |
| 8.2 | The plot shows the ToT value in correspondence to the injected charge in DAC units. A fit was performed using the Equation 8.1. The fit parameters are indicated in the top left. The calibration corresponds to the DC-coupled front end of TJ-Monopix2 sensor W8R6. . . . .                                                                                                                                                                                     | 76 |
| 8.3 | Schematic of the test beam analysis flow. . . . .                                                                                                                                                                                                                                                                                                                                                                                                                 | 78 |
| 8.4 | Comparison of the histograms of the cluster size and cluster charge for associated clusters in the DC-coupled front end. It compares an unirradiated sample (W14R12) and an irradiated sample (W8R6). Both measurements were taken at similar thresholds of $\approx 300 e^-$ . . .                                                                                                                                                                               | 80 |
| 8.5 | Comparison of the cluster size for all irradiated samples against the bias voltage. The threshold during these measurements was $\approx 250 e^-$ . Top: Cluster size of the DC-coupled front end. Bottom: Cluster size of the AC-coupled front end. The statistics for these measurements are large and errors are not visible. . . . .                                                                                                                          | 82 |
| 8.6 | Comparison of the calibrated seed charge and cluster charge against the bias voltage for samples W8R6 and W8R4. The threshold during these measurements was $\approx 250 e^-$ . The statistics for these measurements are large and errors are not visible. Top row: Seed charge of the DC-coupled front end. Bottom row: Cluster charge of the DC-coupled front end. . . . .                                                                                     | 83 |

- 8.7 Comparison of the calibrated cluster charge and cluster charge against the bias voltage for samples W8R6 and W8R4. The threshold during these measurements was  $\approx 250 e^-$ . The statistics for these measurements are large and errors are not visible. Top row: Seed charge of the AC-coupled front end. Bottom row: Cluster charge of the AC-coupled front end. . . . . 84
- 8.8 Visualization of a simplified 1D model for the effective pixel pitch of single-pixel clusters in the presence of multi-pixel clusters. The solid lines indicate the physical pixel borders. The dotted lines indicate the boundaries between the inner region of the pixel and the edge region of the pixel. An incidence particle traversing a pixel in the inner region with a length  $a$  of pixel  $n + 2$  produces a response only in pixel  $n + 2$ . If it traverses the pixel in the edge region with length  $b$  between the pixels  $n$  and  $n - 1$ , a response is measured in the two adjacent pixels, producing a multi-pixel cluster. The effective pixel pitch for a single-pixel cluster is equal to the length  $a$ . The pixel pitch is equal to the length of  $a + b$ . . . . . 86
- 8.9 Mean cluster size depending on the reconstructed in-pixel hit position from measurements for the proton irradiated sample W8R6. Pixels are grouped in  $2 \times 2$  cells and mapped onto a common cell. The boundaries between the cells is indicated by solid black lines. This result was obtained using TBSW with cluster shape-based position reconstruction. . . . . 86
- 8.10 Residual distributions for W8R6. The statistical mean and standard deviation are drawn into the plot. Left: Residual distribution along the  $x$ -axis. Right: Residual distribution along the  $y$ -axis. This result was obtained with TBSW. . . . . 87
- 8.11 Residual distributions for single-pixel clusters for W8R6. A fit of a centered box function convoluted with a Gaussian has been applied to the distribution. The statistical mean and standard deviation, as well as the fit parameters, are drawn into the plot. Left: Residual distribution along the  $x$ -axis. Right: Residual distribution along the  $y$ -axis. This result was obtained with TBSW. . . . . 88
- 8.12 Residual distributions for 2-pixel clusters for W8R6. A fit of a centered box function convoluted with a Gaussian has been applied to the distribution. The statistical mean and standard deviation, as well as the fit parameters, are drawn into the plot. Left: Residual distribution along the  $x$ -axis. Right: Residual distribution along the  $y$ -axis. This result was obtained with TBSW. . . . . 88

- 8.13 Hit detection efficiency of all irradiated samples in dependence of the bias voltage for the DC-coupled front end (top left), magnified for high efficiencies (top right), and the AC-coupled front end (bottom). The threshold during these measurements was  $\approx 250 e^-$  except the p-irradiated  $2.5 \times 10^{14} n_{eq}/cm^2$  (W2R17) which was  $\approx 400 e^-$ . . . . . 91
- 8.14 Hit detection efficiency of all irradiated samples in dependence of the threshold for the DC-coupled front end (top left) and the AC-coupled front end (top right). Plots with a smaller scale are provided for the DC-coupled front end (bottom left) and the AC-coupled front end (bottom right). All measurements were taken at maximum bias for the respective front end and chip. . . . . 92
- 8.15 Hit detection efficiency of the AC- and DC-coupled front end of W8R6 that is proton irradiated to a NIEL fluence of  $5 \times 10^{14} n_{eq}/cm^2$ . . . . 93
- 8.16 Hit detection efficiency of all irradiated samples in dependence of the NTC temperature for the DC-coupled front end (left) and the AC-coupled front end (right). All measurements were taken at maximum bias for the respective front end and chip. The threshold during these measurements was between  $170 e^-$  and  $350 e^-$  except the p-irradiated  $2.5 \times 10^{14} n_{eq}/cm^2$  (W2R17) which was between  $350 e^-$  and  $700 e^-$ . . 93
- 8.17 Threshold of the most irradiated samples in dependence of the NTC temperature for the DC-coupled front end (left) and the AC-coupled front end (right). All measurements were taken at maximum bias for the respective front end and chip. . . . . 94
- 8.18 Leakage current of all irradiated samples in dependence of the NTC temperature for the DC-coupled front end (left) and the AC-coupled front end (right). All measurements were taken at maximum bias for the respective front end and chip. . . . . 95
- 8.19 In-pixel efficiency map of the irradiated sample W8R6 with a fluence of  $5 \times 10^{14} n_{eq}/cm^2$  for the AC-coupled front end. Top: Measurement at  $T_{NTC} = 30^\circ C$  with a total efficiency of  $99.60^{+0.02}_{-0.02} \%$  at a threshold of  $239 e^-$ . Bottom: Measurement at  $T_{NTC} = 40^\circ C$  with a total efficiency of  $98.05^{+0.03}_{-0.04} \%$  at a threshold of  $324 e^-$ . . . . . 96
- 8.20 In-pixel efficiency map of the irradiated sample W8R6 with a fluence of  $5 \times 10^{14} n_{eq}/cm^2$  for the DC-coupled front end. Top: Measurement at  $T_{NTC} = 10^\circ C$  with a total efficiency of  $99.912^{+0.003}_{-0.003} \%$  at a threshold of  $180 e^-$ . Bottom: Measurement at  $T_{NTC} = 40^\circ C$  with a total efficiency of  $99.27^{+0.01}_{-0.01} \%$  at a threshold of  $347 e^-$ . . . . . 97

- 8.21 Comparison of cluster properties from simulation and measurements with the p-irradiated sample W8R6 with a fluence of  $5 \times 10^{14} \text{ n}_{\text{eq}}/\text{cm}^2$ . The measurement was taken at a threshold of  $250 e^-$ , a bias voltage of 20 V and at an NTC temperature of  $30^\circ\text{C}$  which is approximately a chip temperature of  $37^\circ\text{C}$ . The mean and MPV are indicated in each plot for the measurement and simulation, respectively. Top: Comparison of the cluster size distribution. Bottom: Comparison of the seed charge distribution. The inverse charge conversion was applied to then apply the charge conversion again to gauge the effects of the conversion of the charge information into integer ToT values. Both distributions have been rebinned to allow for a better visualization of the discrete charge values. A version with finer binning that makes the discrete charge values visible is available in Appendix A.2. . . . . 102
- 8.22 Residual distributions from simulation. The statistical mean and standard deviation are drawn into the plot for the simulation. The measured residuals are indicated for comparison. Left: Residual distribution along the  $x$ -axis. Right: Residual distribution along the  $y$ -axis. This result was obtained with TBSW. . . . . 104
- 8.23 Residual distributions for 1-pixel clusters from simulation. A fit of a centered box function convoluted with a Gaussian has been applied to the distribution from simulation. The statistical mean and standard deviation, as well as the fit parameters, are drawn into the plot. The measured residuals are indicated for comparison. Left: Residual distribution along the  $x$ -axis. Right: Residual distribution along the  $y$ -axis. This result was obtained with TBSW. . . . . 105
- 8.24 Residual distributions for 2-pixel clusters from simulation. A fit of a centered box function convoluted with a Gaussian has been applied to the distribution from simulation. The statistical mean and standard deviation, as well as the fit parameters, are drawn into the plot. The measured residuals are indicated for comparison. Left: Residual distribution along the  $x$ -axis. Right: Residual distribution along the  $y$ -axis. This result was obtained with TBSW. . . . . 105
- 8.25 Hit detection efficiency in dependence of the threshold. The measurement is performed for the DC-coupled front end of TJ-Monopix2 sample W8R6. The parameters of the simulation are chosen with respect to the measurements. The error bars indicate the statistical error of the efficiency. . . . . 107

|      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              |     |
|------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 8.26 | Comparison of the seed charge distribution for clusters associated to reconstructed tracks for simulation and measurement. Both distributions are normalized to an area of 1. This is a visually modified version of Figure 8.21, for details refer to Section 8.4.2. The region of the distribution relevant for effects relating to the threshold is magnified. The maximum threshold applied to TJ-Monopix2 chip W8R6 during measurements and for the simulation are indicated by vertical lines. . . . . | 107 |
| 9.1  | Generic layout of a deep neural network with an input layer, two hidden layers and an output layer with a single node. . . . .                                                                                                                                                                                                                                                                                                                                                                               | 111 |
| 9.2  | Examples for activation functions. The formula of each function is indicated in the respective plot. The tanh and softplus activation functions are used in BGNet. . . . .                                                                                                                                                                                                                                                                                                                                   | 112 |
| 9.3  | Illustration of the storage beam background processes for beam-gas (a) and Touschek (b). . . . .                                                                                                                                                                                                                                                                                                                                                                                                             | 115 |
| 9.4  | Flow diagram of BGNet. The output of the neural network is the total predicted hit rate at time $t$ of a Belle II sub-detector. . . . .                                                                                                                                                                                                                                                                                                                                                                      | 119 |
| 9.5  | Flow diagram for BGNet applied on hit rates with explicit details on the structure of the neural networks and the operations performed. A dense layer is defined as a fully connected layer where each node is connected to every node in the previous layer. The gather operation maps a vector onto a new vector using a list of indices. . . . .                                                                                                                                                          | 122 |
| 9.6  | Flow diagram for BGNet with the multi-sensor scheme. Explicit details on the structure of the neural networks and the operations performed are indicated. A dense layer is defined as a fully connected layer where each node is connected to every node in the previous layer. The gather operation maps a vector onto a new vector using a list of indices. . . . .                                                                                                                                        | 124 |
| 9.7  | The measured and predicted total hit rate and the predicted beam background decomposition during HER single beam decays. The HER single beam decays were conducted during a beam background study campaign on 16.04.2024. . . . .                                                                                                                                                                                                                                                                            | 127 |
| 9.8  | The total measured CDC average layer leakage current and the HER beam current. The measured hit rate of the CDC should decrease as the beam backgrounds should decrease as the HER beam current dissipates. The mean attribution value in units of the hit rate is indicated in the top of the plot. The attribution value refers to the neural network prediction and not the measured hit rate. . . . .                                                                                                    | 128 |

- 9.9 Attribution values of input variables during an HER single beam decay. The attribution values are visualized as point clouds, and the density of points is reflected by the vertical width of the point clouds. Each variable has points equal to the number of samples in the test data set. The color of each point corresponds to the relative feature value compared to all feature values present in all test samples. The input variables are ranked by their mean absolute attribution values. 129
- 9.10 The total measured CDC average layer leakage current and the HER vertical beam size. A large change of the HER vertical beam size can be observed in the middle beam decay, while it is approximately constant during the decays before and after. The mean attribution value in units of the hit rate is indicated in the top of the plot. The attribution value refers to the neural network prediction and not the measured hit rate. . . . . 130
- 9.11 Flow diagram for BGNet applied on beam losses with explicit details on the structure of the neural networks and the operations performed. A dense layer is defined as a fully connected layer where each node is connected to every node in the previous layer. The gather operation maps a vector onto a new vector using a list of indices. . . . . 133
- 9.12 The measured and predicted beam loss of the HER beam current and the predicted beam background decomposition during HER single beam decays. The HER single beam decays were conducted during a beam background study campaign on 16.04.2024 . . . . . 134
- 9.13 Diagram of all functional parts in the network, in which the BGNet server is located. The RT-BGNet application is run on the BGNet server. The BGNet server is a dedicated machine for BGNet related tasks. . . . . 136

- 9.14 Diagram of the functions performed by RT-BGNet. Each box represents a functional part. The EPICS network and the archiver appliance are part of the KEK common infrastructure. The dark blue boxes are tasks performed by RT-BGNet on the BGNet server. Each task is implemented as thread. A thread is an independent sequence of operations that runs in the same context (address space) as all other threads of a process. The threads are operating concurrently and the process is RT-BGNet. The “Outside applications” are running locally on the BGNet server but are executed outside the context of RT-BGNet. They communicate with RT-BGNet via a REST API using Hypertext Transfer Protocol (HTTP) or artifacts written to the local hard drive. Connections between functional parts are drawn in case the communication is required for the operation to continue. For example, the “Update archive” thread has a connection to the “Archiver appliance” as it is required to wait for network to provide the requested data and the “Hit rate Training” and “Beam loss Training” threads need to wait until the updated training data is written to the local archive. If data is passed between functional parts, the direction of data flow is indicated by arrows. . . . . 137
- 9.15 In the upper plot, the predicted beam background decomposition is shown as a stacked histogram of all components for the dose rate of a diamond radiation monitor. The total observed, or measured, hit rate is indicated. The chosen time windows for the test and baseline, or reference, data sets are indicated as colored boxes on the top of the plot. The resampling frequency is shown in the legend. The bottom plot shows the ratio of the total measured hit rate and the total predicted hit rate. . . . . 143
- 9.16 Attribution values of input variables for the hit rate model of BGNet. The test and baseline samples are from the time before and after the collimator jaw D05V1BTM was moved, respectively. The attribution values are visualized as point clouds, and the density of points is reflected by the vertical width of the point clouds. Each variable has points equal to the number of samples in the test data set. The color of each point corresponds to the relative feature value compared to all feature values present in all test samples. The input variables are ranked by their mean absolute attribution values. . . . . 144



- 9.17 The total predicted hit rate and the collimator jaw position D05V1BTM during the selected time window of the report. The test and baseline time windows are indicated as colored boxes on the top of the plot. The closing of the collimator jaw by 0.3 mm coincides in time with the reduction of the predicted hit rate. The mean attribution value is shown. On average, the difference of the total hit rate prediction between the test and baseline samples due to the collimator jaw position is  $12.19 \text{ mrad s}^{-1}$ . . . . . 145
- 9.18 Top: Attribution values of input variables for the LER injection network of BGNet. The test and baseline samples are from the time before and after the collimator jaw D05V1BTM was moved, respectively. The attribution values are visualized as point clouds, and the density of points is reflected by the vertical width of the point clouds. Each variable has points equal to the number of samples in the test data set. The color of each point corresponds to the relative feature value compared to all feature values present in all test samples. The input variables are ranked by their mean absolute attribution values. Bottom: The predicted hit rate due to LER injection background and the collimator jaw position D05V1BTM during the selected time window of the report. The closing of the collimator jaw by 0.3 mm coincides in time with the reduction of the predicted hit rate. . . . . 146
- 9.19 Ratio of the predicted hit rate  $\mathcal{O}_{\text{pred}}$  and measured hit rate  $\mathcal{O}_{\text{meas}}$  for 40 models. The predictions were performed on data that was seen during the training. The ratios are calculated every second and then resampled to 5 min, therefore each box represents 40 data points, one for each model, of resampled ratios for the respective time step. The data samples used for the prediction are taken from an arbitrary time window of 1 h within the training data. The box plot follows standard convention where the solid line in the box is the median, the edges of the box extend to the first (Q1) and third quartile (Q3), the whiskers extend the inter quartile range by a factor 1.5 and outliers are explicitly plotted below and above the ends of the whiskers. Left: Ratios for the CDC average layer leakage current. Bottom Left: Ratios for the dose rate of a diamond radiation monitor. . . . . 149
- 9.20 Overview of the beam background study campaign conducted on 16.04.2024. The leakage current and the beam currents of both rings are shown. The segments of the study campaign are indicated by the background color and the respective labels. The labels indicate whether a single-beam study is performed, in either the HER or the LER, or a luminosity scan. The number in the label indicates the number of bunches in the respective ring, that needs to be varied in order to differentiate Touschek and beam-gas backgrounds. . . . . 151

- 9.21 Decompositions of the CDC average layer current during a beam background study campaign as a stacked histogram at the time of an HER single beam decay at a number of bunches of 393. The bottom plot shows the ratio between the measured and predicted CDC average layer current. Top: Decomposition obtained from BGNet. Bottom: Decomposition obtained from the fit. . . . . 153
- 9.22 Decompositions of the CDC average layer current during a beam background study campaign as a stacked histogram at the time of an HER single beam decay at a number of bunches of 2346. The bottom plot shows the ratio between the measured and predicted CDC average layer current. Top: Decomposition obtained from BGNet. Bottom: Decomposition obtained from the fit. . . . . 154
- 9.23 Decompositions of the CDC average layer current during a beam background study campaign as a stacked histogram at the time of multiple decays during collisions. The bottom plot shows the ratio between the measured and predicted CDC average layer current for predictions by BGNet and the fit, respectively. Top: Decomposition obtained from BGNet. Bottom: Decomposition obtained from the fit. . . . . 155
- 9.24 Training scheme of new BGNet models for the real-time application of BGNet. The time window for the training data is defined from  $t_{\text{start}}$  to  $t_{\text{end}}$ . The time needed for training a model is  $T_{\text{train}}$ . A model is deployed and used for predictions from  $t_{\text{end}} + T_{\text{train}}$  until  $t_{\text{end}} + 2T_{\text{train}}$ . 157
- 9.25 Ratio of the predicted hit rate  $\mathcal{O}_{\text{pred}}$  and measured hit rate  $\mathcal{O}_{\text{meas}}$ . The predictions were performed on data that was not seen during the training and the BGNet model is extrapolating. The ratios are calculated every second and then resampled to 5 min, therefore each box represents 40 data points, one for each model, of resampled ratios for the respective time step. The data samples used for the prediction are the ones immediately following the last sample seen during the training. The box plot follows standard convention where the solid line in the box is the median, the edges of the box extend to the first (Q1) and third quartile (Q3), the whiskers extend the inter quartile range by a factor 1.5 and outliers are explicitly plotted below and above the ends of the whiskers. Top Left: Ratios for the first hour after the training for the CDC average layer leakage current. Top Right: Ratios for the second hour after the training for the CDC average layer leakage current. This time window is relevant for the performance of the real-time predictions. Bottom Left: Ratios for the first hour after the training for the dose rate of a diamond radiation monitor. Bottom Right: Ratios for the second hour after the training for the dose rate of a diamond radiation monitor. This time window is relevant for the performance of the real-time predictions. . . . . 159

- 9.26 Oscilloscope screen capture of the fast CDC current logger on June 12, 2022 at 20:52:39 (JST). The  $y$ -axis shows the leakage current in units of  $10\mu\text{A}$  per DIV. It was taken during physics data taking during a period with top-up injections. Initially, only the LER received injections with asynchronous HER injections starting about 2 s later [103]. . . . . 161
- 9.27 BGNet predictions during top-up injections. The non-injection backgrounds are summed up and displayed together. The injection backgrounds for HER and LER are displayed separately. The current  $4 \cdot I_{\text{Log}}^{\text{min}}$  can now be compared to the average CDC layer leakage current due to non-injection background to gauge BGNet's ability to differentiate injection backgrounds from the other background components [103]. . . . . 162
- 9.28 A schematic view of the PXD indicating its position in the coordinate system and indicating the module numbers. The colored modules indicate the modules installed in PXD 1 [123]. . . . . 164
- 9.29  $\phi$ -profile for the total predicted hit rate by BGNet and the total measured hit rate for diamond radiation monitors at  $\phi = 60^\circ, 120^\circ, 180^\circ, 240^\circ, 300^\circ$ . 164
- 9.30  $\phi$ -profiles for the storage backgrounds based on BGNet predictions for diamond radiation monitors at  $\phi = 60^\circ, 120^\circ, 180^\circ, 240^\circ, 300^\circ$ . The hit rate due to each component has been normalized by the total predicted hit rate. The error shown is the standard error of the sample mean, which was increased by a factor 3 to make it visible. Top left: HER beam-gas component. Top right: HER Touschek component. Bottom left: LER beam-gas component. Bottom right: LER Touschek component. . . . . 165
- 9.31  $\phi$ -profiles for the injection backgrounds based on BGNet predictions for diamond radiation monitors at  $\phi = 60^\circ, 120^\circ, 180^\circ, 240^\circ, 300^\circ$ . The hit rate due to each component has been normalized by the total predicted hit rate. The error shown is the standard error of the sample mean, which was increased by a factor 3 to make it visible. Left: HER injection component. Right: LER injection component. . . . . 166

- 9.32 Ratio of the predicted beam loss  $\mathcal{O}_{\text{pred}}$  and measured beam loss  $\mathcal{O}_{\text{meas}}$  for 40 different models. The predictions were performed on data that was seen during the training. The ratios are calculated every second and then resampled to 5 min, therefore each box represents 40 data points of resampled ratios for the respective time step. The data samples used for the prediction are taken from an arbitrary time window of 1 h within the training data. The box plot follows standard convention where the solid line in the box is the median, the edges of the box extend to the first (Q1) and third quartile (Q3), the whiskers extend the inter quartile range by a factor 1.5 and outliers are explicitly plotted below and above the ends of the whiskers. Left: Ratios for the HER beam loss. Bottom Left: Ratios for the LER beam loss. . . 169
- 9.33 Training scheme of the BGNet beam loss models in relation to the hit rate model training for the real-time application of BGNet. The time window for the training data is defined from  $t_{\text{start}}$  to  $t_{\text{end}}$ . The training of all models begins at the same time and the next cycle only starts once all models are trained. The beam loss models are deployed from  $t_{\text{end}} + T_{\text{train, beam loss}}$  until  $t_{\text{end}} + T_{\text{train, beam loss}} + T_{\text{train}}$ , where  $T_{\text{train}}$  refers to the slowest model training, which is a hit rate model. 171
- 9.34 Ratio of the predicted beam loss  $\mathcal{O}_{\text{pred}}$  and measured beam loss  $\mathcal{O}_{\text{meas}}$ . The predictions were performed on data that was not seen during the training and the BGNet model is extrapolating. The ratios are calculated every second and then resampled to 5 min, therefore each box represents 40 data points, one for each model, of resampled ratios for the respective time step. The data samples used for the prediction are the ones immediately following the last sample seen during the training. The box plot follows standard convention where the solid line in the box is the median, the edges of the box extend to the first (Q1) and third quartile (Q3), the whiskers extend the inter quartile range by a factor 1.5 and outliers are explicitly plotted below and above the ends of the whiskers. Top Left: Ratios for the first hour after the training for the HER beam loss. This time window is relevant for the performance of the real-time predictions. Top Right: Ratios for the second hour after the training for the HER beam loss. Bottom Left: Ratios for the first hour after the training for the LER beam loss. This time window is relevant for the performance of the real-time predictions. Bottom Right: Ratios for the second hour after the training for the LER beam loss. . . . . 172
- A.1 Comparison of the histograms of the cluster size and cluster charge for associated clusters in the DC-coupled frontend. The irradiated samples W8R4, W2R17 and W2R5 are compared. . . . . 206

|     |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |     |
|-----|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| A.2 | Comparison of cluster properties for all irradiated samples against the bias voltage for the DC-coupled frontend. Top: mean cluster size. Middle: mean cluster charge. Bottom: mean seed charge. The threshold during these measurements was $\approx 250 e^-$ except for the p-irradiated $2.5 \times 10^{14} n_{eq}/cm^2$ (W2R17) which was $\approx 400 e^-$ . . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 207 |
| A.3 | Comparison of cluster properties for all irradiated samples against the bias voltage for the AC-coupled frontend. Top: mean cluster size. Middle: mean cluster charge. Bottom: mean seed charge. The threshold during these measurements was $\approx 250 e^-$ . . . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                        | 208 |
| A.4 | Comparison of the seed charge distribution from simulation and measurements with the p-irradiated sample W8R6 with a fluence of $5 \times 10^{14} n_{eq}/cm^2$ . The discrete charge values for the seed and cluster charge are an effect of the charge-to-digital conversion performed in the pixel frontend. The measurement was taken at a threshold of $250 e^-$ , a bias voltage of 20 V and at an NTC temperature of $30^\circ C$ which is approximately a chip temperature of $37^\circ C$ . The mean and MPV of the distributions are indicated in the plot. The inverse charge conversion was applied to then apply the charge conversion again to gauge the effects of the conversion of the charge information into integer ToT values. Both distributions have the same bin size, the increased linewidth is for better distinction between the histograms. . . . . | 210 |
| B.1 | Hit rate decomposition for the diamond radiation monitor at the SVD forward position at $\phi = 60^\circ$ during the time window used for the $\phi$ -profile analysis of the beam background predictions of BGNet. . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | 217 |
| B.2 | Hit rate decomposition for the diamond radiation monitor at the SVD forward position at $\phi = 120^\circ$ during the time window used for the $\phi$ -profile analysis of the beam background predictions of BGNet. . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | 217 |
| B.3 | Hit rate decomposition for the diamond radiation monitor at the SVD forward position at $\phi = 180^\circ$ during the time window used for the $\phi$ -profile analysis of the beam background predictions of BGNet. . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | 218 |
| B.4 | Hit rate decomposition for the diamond radiation monitor at the SVD forward position at $\phi = 240^\circ$ during the time window used for the $\phi$ -profile analysis of the beam background predictions of BGNet. . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | 218 |
| B.5 | Hit rate decomposition for the diamond radiation monitor at the SVD forward position at $\phi = 300^\circ$ during the time window used for the $\phi$ -profile analysis of the beam background predictions of BGNet. . . .                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                      | 219 |



---

## List of Tables

---

|     |                                                                                                                                                                                                                                                                                                                              |    |
|-----|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----|
| 4.1 | Variables, and their description, of the Bethe-Bloch formula in Equation 4.1. . . . .                                                                                                                                                                                                                                        | 30 |
| 5.1 | Comparison of performance parameters of KEKB and SuperKEKB [56]. (*) marked values are nominal or design values that have not been reached yet. . . . .                                                                                                                                                                      | 47 |
| 5.2 | Parameters of the electron and positron bunches at the moment of injection into the SuperKEKB main ring [56]. . . . .                                                                                                                                                                                                        | 50 |
| 6.1 | Specifications of the Belle II PXD. . . . .                                                                                                                                                                                                                                                                                  | 53 |
| 7.1 | Reconstruction efficiency and purity for the $\bar{D}^0 \rightarrow K^+\pi^-$ final state. The current Belle II detector with background SC2 is compared with the upgraded VTX and all background scenarios [71]. . . . .                                                                                                    | 63 |
| 7.2 | Initial chip requirements for a tracking detector upgrade and the TJ-Monopix2 specifications [71] compared to the final OBELIX specifications [50]. . . . .                                                                                                                                                                  | 66 |
| 7.3 | VTX layer parameters for the 5-layer design [71]. *: the addition of a necessary cooling component will increase the material budget to $0.3\%X_0$ [74] over the previous value in the CDR [71]. †: the material budget of the oVTX layers was increased to $0.8\%X_0$ [74] over the previous value in the CDR [71]. . . . . | 69 |
| 7.4 | VTX outer layer parameters in the alternative 6-layer design [50, 74]. The two radii for each layer reflect the radii of the staggered sub-layers. †: the material budget of the oVTX layers was increased to $0.8\%X_0$ [74] over the previous value in the CDR [71]. . . . .                                               | 69 |
| 8.1 | Summary of the TJ-Monopix2 test beam campaigns. . . . .                                                                                                                                                                                                                                                                      | 72 |
| 8.2 | TJ-Monopix2 samples available for the 2025 test beam campaign. . .                                                                                                                                                                                                                                                           | 74 |

|     |                                                                                                                                                                                                                                                                                           |     |
|-----|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|
| 8.3 | The maximum bias that could be applied to each irradiated sample for the AC-coupled front end (AC FE) and the DC-coupled FE (DC FE). The AC-coupled front end of W2R17 was not operational. . . .                                                                                         | 79  |
| 8.4 | Spatial resolution achieved in measurements of W8R6 for 1-pixel, 2-pixel and all clusters. . . . .                                                                                                                                                                                        | 89  |
| 8.5 | Spatial resolution achieved in simulation for 1-pixel, 2-pixel and all clusters. The simulation is performed with respect to TJ-Monopix2 sample W8R6. . . . .                                                                                                                             | 104 |
| 9.1 | Numbering scheme of background components considered by BGNet.                                                                                                                                                                                                                            | 117 |
| 9.2 | Currently supported target variables. The names of the diamond radiation monitors indicate their positions, compare with Figure 6.6.<br>*: What is here referred to as leakage current of the CDC is in the context of gaseous detectors also referred to as chamber current. . .         | 121 |
| 9.3 | Hyperparameters of BGNet. . . . .                                                                                                                                                                                                                                                         | 121 |
| 9.4 | Sensitivities obtained during the background study camping on 16.04.2024 using the background fit procedure and BGNet. The errors on the sensitivities are obtained from the covariance matrix of the fit and from the standard deviation of the sensitivity distributions for BGNet. . . | 152 |
| 9.5 | Units of the sensitivities for the CDC average layer current in units of $\mu\text{A}$ . . . . .                                                                                                                                                                                          | 152 |
| 9.6 | Relative hit rates due to LER beam-gas for PXD modules from simulation [122] and for diamond radiation monitors from BGNet predictions.                                                                                                                                                   | 166 |
| 9.7 | Relative hit rates due to LER Touschek for PXD modules from simulation [122] and for diamond radiation monitors from BGNet predictions.                                                                                                                                                   | 167 |
| 9.8 | Relative hit rates due to HER Touschek for PXD modules from simulation [122] and for diamond radiation monitors from BGNet predictions.                                                                                                                                                   | 167 |



## APPENDIX A

---

### Appendix: Test Beam Results

---

#### A.1 Cluster Properties

Comparisons for the mean cluster sizes and uncalibrated mean cluster and mean seed charges are displayed in Figure A.2 for the DC-coupled front end and Figure A.3 for the AC-coupled front end. The distributions of the cluster size and uncalibrated cluster charge are shown in Figure A.1.

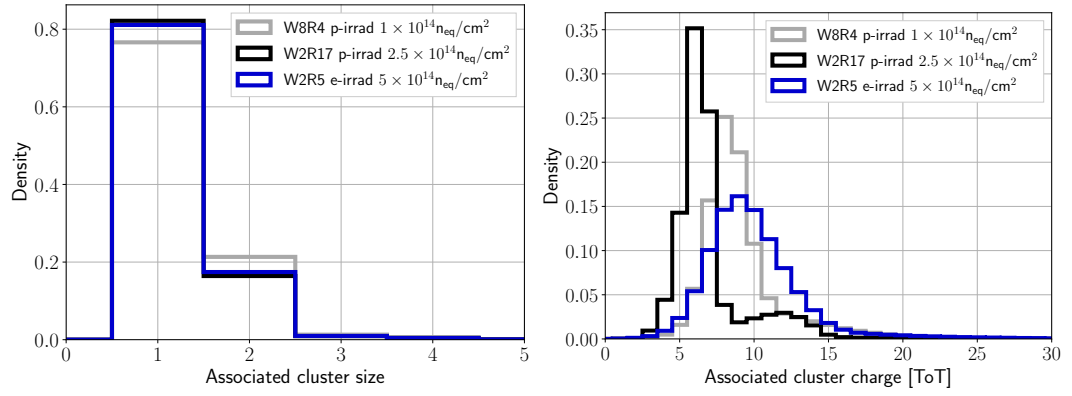


Figure A.1: Comparison of the histograms of the cluster size and cluster charge for associated clusters in the DC-coupled frontend. The irradiated samples W8R4, W2R17 and W2R5 are compared.

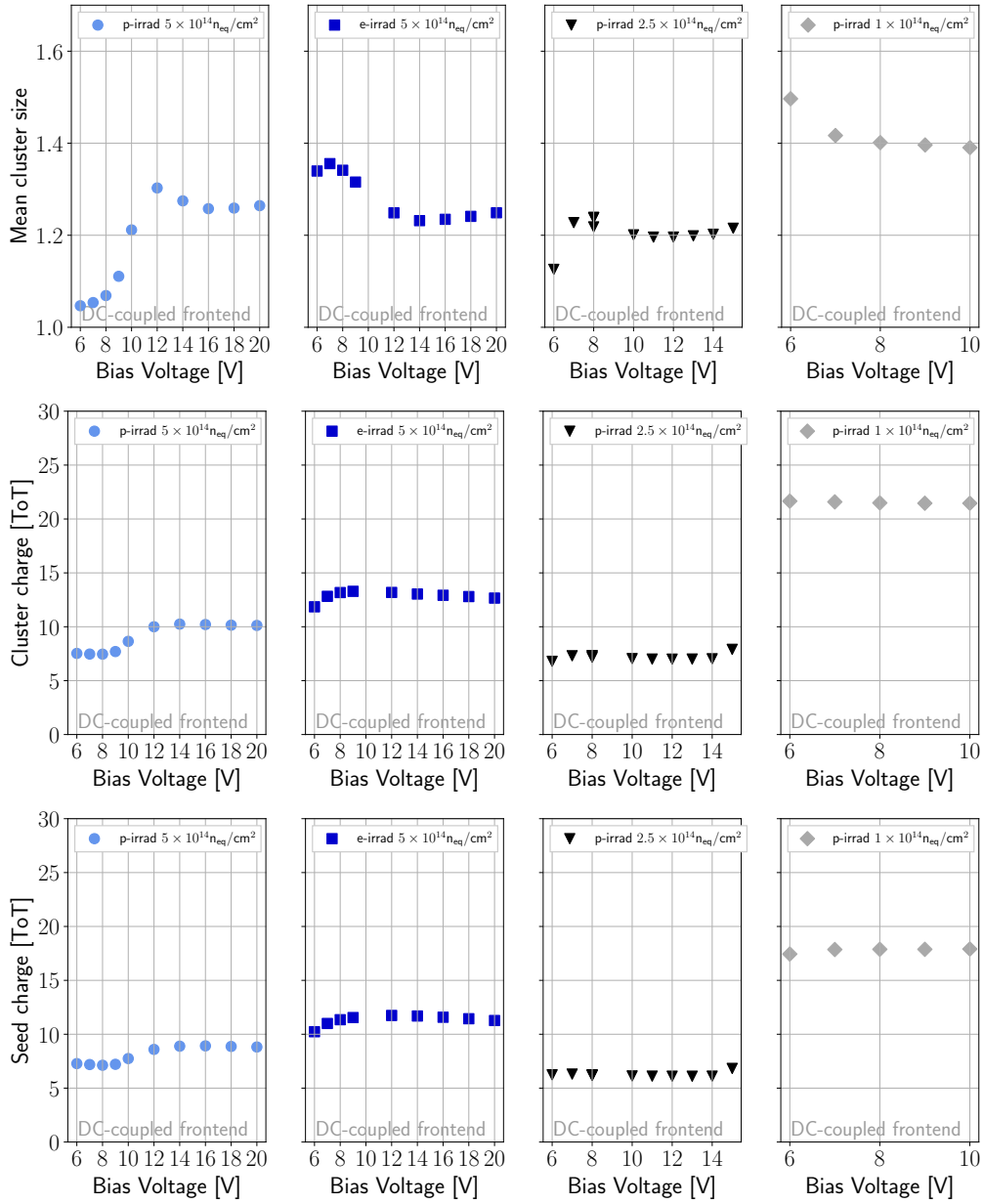


Figure A.2: Comparison of cluster properties for all irradiated samples against the bias voltage for the DC-coupled frontend. Top: mean cluster size. Middle: mean cluster charge. Bottom: mean seed charge. The threshold during these measurements was  $\approx 250 e^-$  except for the p-irradiated  $2.5 \times 10^{14} n_{eq}/cm^2$  (W2R17) which was  $\approx 400 e^-$ .

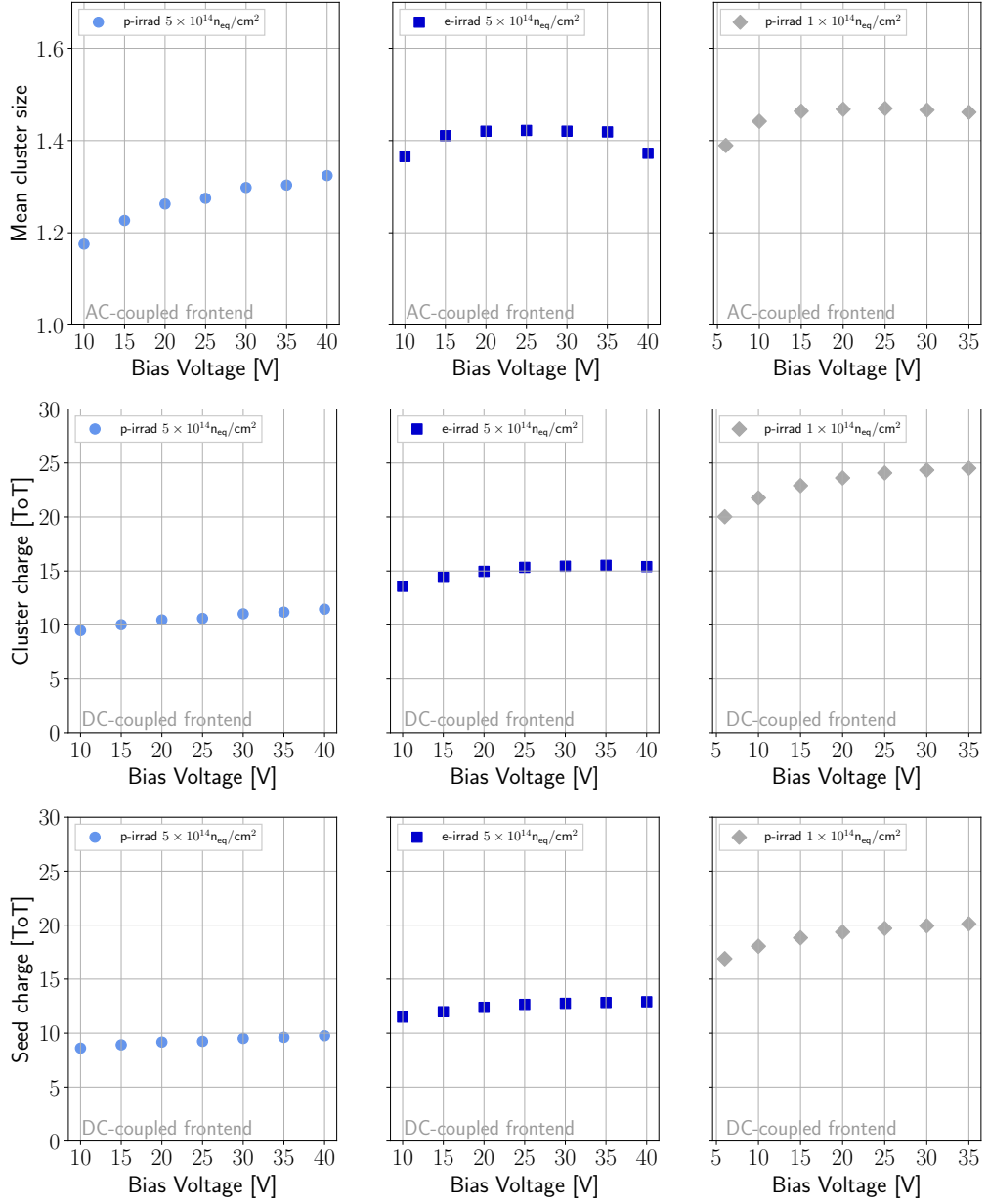


Figure A.3: Comparison of cluster properties for all irradiated samples against the bias voltage for the AC-coupled frontend. Top: mean cluster size. Middle: mean cluster charge. Bottom: mean seed charge. The threshold during these measurements was  $\approx 250 e^-$ .

## **A.2 Simulated Cluster Properties**

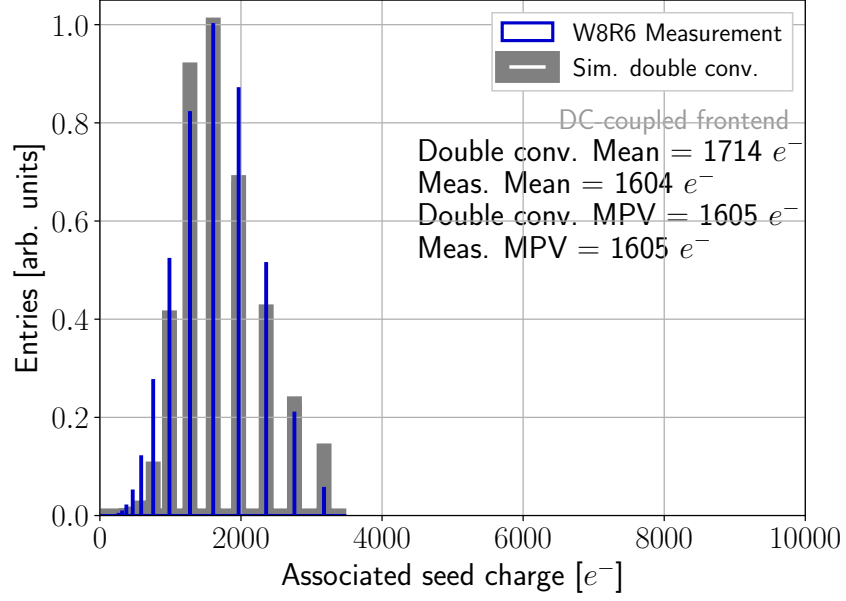


Figure A.4: Comparison of the seed charge distribution from simulation and measurements with the p-irradiated sample W8R6 with a fluence of  $5 \times 10^{14} n_{eq}/cm^2$ . The discrete charge values for the seed and cluster charge are an effect of the charge-to-digital conversion performed in the pixel frontend. The measurement was taken at a threshold of  $250 e^-$ , a bias voltage of 20 V and at an NTC temperature of  $30^\circ C$  which is approximately a chip temperature of  $37^\circ C$ . The mean and MPV of the distributions are indicated in the plot. The inverse charge conversion was applied to then apply the charge conversion again to gauge the effects of the conversion of the charge information into integer ToT values. Both distributions have the same bin size, the increased linewidth is for better distinction between the histograms.

## B.1 Input Variables for Hit Rate Models

```
# Featurs for HER beam gas background model
'sens_her_beamgas': [
    "SKB_HER_P_avg",
    "BMHXRM_BEAM_EMITTX",
    "BMHXRM_BEAM_EMITTY",
    "VAHCLM_D01V1TOP_CSS_DIF_POS", # Collimator variables
    "VAHCLM_D01V1BTM_CSS_DIF_POS",
    "VAHCLM_D12V4BTM_CSS_DIF_POS",
    "VAHCLM_D12V3TOP_CSS_DIF_POS",
    "VAHCLM_D09V1BTM_CSS_DIF_POS",
    "VAHCLM_D09V3_CSS_DIF_POS",
],

# Featurs for HER touschek background model
'sens_her_touschek': [
    "BMHXRM_BEAM_EMITTX",
    "BMHXRM_BEAM_EMITTY",
    "VAHCLM_D09H3_CSS_DIF_POS", # Collimator variables
    "VAHCLM_D09H4_CSS_DIF_POS",
    "VAHCLM_D12H2_CSS_DIF_POS",
    "VAHCLM_D12H3_CSS_DIF_POS",
    "VAHCLM_D01H5OUT_CSS_DIF_POS",
    "VAHCLM_D01H5IN_CSS_DIF_POS",
```

```

],

# Featurs for LER beam gas background model
'sens_ler_beamgas': [
    "SKB_LER_P_avg",
    "CGLOPT_IP_BETA_X",
    "CGLOPT_IP_BETA_Y",
    "BMLXRM_BEAM_EMITTX",
    "BMLXRM_BEAM_EMITTY",
    "BML_MQC2RP_POS.PYP",
    "BML_MQC1RP_POS.PYP",
    "VALCLM_D02V1TOP_CSS_DIF_POS", # Collimator variables
    "VALCLM_D02V1BTM_CSS_DIF_POS",
    "VALCLM_D05V1TOP_CSS_DIF_POS",
    "VALCLM_D05V1BTM_CSS_DIF_POS",
    "VALCLM_D06V1TOP_CSS_DIF_POS",
    "VALCLM_D06V1BTM_CSS_DIF_POS",
    "VALCLM_D03V4TOP_CSS_DIF_POS",
    "VALCLM_D03V4BTM_CSS_DIF_POS",
],

# Featurs for LER touschek background model
'sens_ler_touschek': [
    "BMLXRM_BEAM_EMITTX",
    "BMLXRM_BEAM_EMITTY",
    "BML_MQC2RP_POS.PXP",
    "BML_MQC1RP_POS.PXP",
    "VALCLM_D02H1OUT_CSS_DIF_POS", # Collimator variables
    "VALCLM_D02H1IN_CSS_DIF_POS",
    "VALCLM_D02H4OUT_CSS_DIF_POS",
    "VALCLM_D02H4IN_CSS_DIF_POS",
    "VALCLM_D06H4OUT_CSS_DIF_POS",
    "VALCLM_D06H4IN_CSS_DIF_POS",
    "VALCLM_D06H3OUT_CSS_DIF_POS",
    "VALCLM_D06H3IN_CSS_DIF_POS",
],

# Features for HER injection background
'sens_her_injection': [
    "HER_CHARGExINEFFxREP_FREQ",
    "CG_BTE_BPM_CHARGE1_CORR",
    "TM_EVRO_HER_INJ_EFF_BCM",
    "CGHINJ_BEAM_REPETITION_REQ",

```



```

"CGHOPT_IP_BETA_X",
"CGHOPT_IP_BETA_Y",
"BMHXRM_BEAM_EMITTX",
"BMHXRM_BEAM_EMITTY",
"CGeBT_DPPO",
"CGHOPT_GATED_TUNE_X_MEAS",
"CGHOPT_GATED_TUNE_Y_MEAS",
"SKB_HER_XRAY_SIGMAX",
"SKB_HER_XRAY_SIGMAY",
"BUNCH_CURRENT_RATIO",
"LER_BUNCH_CURRENT",
"HER_BUNCH_CURRENT",
"CGHINJ_INJECTION_YPOS", # Injection related variables
"CGHINJ_INJECTION_YANG",
"CGHINJ_SEPTUM_ANG_R",
"CGHINJ_SEPTUM_POS_R",
"CGHINJ_KICKER_HEIGHT_R",
"CGHINJ_KICKER_JUMP_R",
"VAHCLM_D01V1TOP_CSS_DIF_POS", # Collimator variables
"VAHCLM_D01V1BTM_CSS_DIF_POS",
"VAHCLM_D12V3TOP_CSS_DIF_POS",
"VAHCLM_D12V4BTM_CSS_DIF_POS",
"VAHCLM_D09V1BTM_CSS_DIF_POS",
"VAHCLM_D09V3_CSS_DIF_POS",
"VAHCLM_D09H3_CSS_DIF_POS",
"VAHCLM_D09H4_CSS_DIF_POS",
"VAHCLM_D12H2_CSS_DIF_POS",
"VAHCLM_D12H3_CSS_DIF_POS",
"VAHCLM_D01H5OUT_CSS_DIF_POS",
"VAHCLM_D01H5IN_CSS_DIF_POS",
],

# Features for LER injection background
'sens_ler_injection': [
    "LER_CHARGExINEFFxREP_FREQ",
    "CG_BTP_BPM_CHARGE1_CORR",
    "TM_EVRO_LER_INJ_EFF_BCM",
    "CGLINJ_BEAM_REPETITION_REQ",
    "BMLXRM_BEAM_EMITTY",
    "BMLXRM_BEAM_EMITTX",
    "CGLOPT_GATED_TUNE_X_MEAS",
    "CGLOPT_GATED_TUNE_Y_MEAS",
    "BUNCH_CURRENT_RATIO",

```

```

"LER_BUNCH_CURRENT",
"HER_BUNCH_CURRENT",
"VALCLM_D02V1TOP_CSS_DIF_POS", # Collimator variables
"VALCLM_D02V1BTM_CSS_DIF_POS",
"VALCLM_D05V1TOP_CSS_DIF_POS",
"VALCLM_D05V1BTM_CSS_DIF_POS",
"VALCLM_D06V1TOP_CSS_DIF_POS",
"VALCLM_D06V1BTM_CSS_DIF_POS",
"VALCLM_D03V4TOP_CSS_DIF_POS",
"VALCLM_D03V4BTM_CSS_DIF_POS",
"VALCLM_D02H4OUT_CSS_DIF_POS",
"VALCLM_D02H4IN_CSS_DIF_POS",
"VALCLM_D06H4OUT_CSS_DIF_POS",
"VALCLM_D06H4IN_CSS_DIF_POS",
"BML_MQC2RP_POS.PYP",
"BML_MQC2RP_POS.PXP",
"BML_MQC1RP_POS.PYP",
"BML_MQC1RP_POS.PXP",
],

```

## B.2 Input Variables for Beam Loss Models

```

# Featurs for HER beam gas background model
'sens_her_beamgas': [
    "SKB_HER_P_avg",
    "BMHXRМ_BEAM_EMITTX",
    "BMHXRМ_BEAM_EMITTY",
    "VAHCLM_D01V1TOP_CSS_DIF_POS", # Collimator variables
    "VAHCLM_D01V1BTM_CSS_DIF_POS",
    "VAHCLM_D12V4BTM_CSS_DIF_POS",
    "VAHCLM_D12V3TOP_CSS_DIF_POS",
    "VAHCLM_D09V1BTM_CSS_DIF_POS",
    "VAHCLM_D09V3_CSS_DIF_POS",
],

# Featurs for HER touschek background model
'sens_her_touschek': [
    "BMHXRМ_BEAM_EMITTX",
    "BMHXRМ_BEAM_EMITTY",
    "VAHCLM_D09H3_CSS_DIF_POS", # Collimator variables
    "VAHCLM_D09H4_CSS_DIF_POS",
    "VAHCLM_D12H2_CSS_DIF_POS",

```

```

    "VAHCLM_D12H3_CSS_DIF_POS",
    "VAHCLM_D01H5OUT_CSS_DIF_POS",
    "VAHCLM_D01H5IN_CSS_DIF_POS",
    "CGHOPT_GATED_TUNE_X_MEAS",
    "CGHOPT_GATED_TUNE_Y_MEAS",
],

# Features for LER beam gas background model
'sens_ler_beamgas': [
    "SKB_LER_P_avg",
    "CGLOPT_IP_BETA_X",
    "CGLOPT_IP_BETA_Y",
    "BMLXRM_BEAM_EMITTX",
    "BMLXRM_BEAM_EMITTY",
    "BML_MQC2RP_POS.PYP",
    "BML_MQC1RP_POS.PYP",
    "VALCLM_D02V1TOP_CSS_DIF_POS", # Collimator variables
    "VALCLM_D02V1BTM_CSS_DIF_POS",
    "VALCLM_D05V1TOP_CSS_DIF_POS",
    "VALCLM_D05V1BTM_CSS_DIF_POS",
    "VALCLM_D06V1TOP_CSS_DIF_POS",
    "VALCLM_D06V1BTM_CSS_DIF_POS",
    "VALCLM_D06V2TOP_CSS_DIF_POS",
    "VALCLM_D06V2BTM_CSS_DIF_POS",
],

# Features for LER touschek background model
'sens_ler_touschek': [
    "BMLXRM_BEAM_EMITTX",
    "BMLXRM_BEAM_EMITTY",
    "BML_MQC2RP_POS.PXP",
    "BML_MQC1RP_POS.PXP",
    "VALCLM_D02H1OUT_CSS_DIF_POS", # Collimator variables
    "VALCLM_D02H1IN_CSS_DIF_POS",
    "VALCLM_D02H4OUT_CSS_DIF_POS",
    "VALCLM_D02H4IN_CSS_DIF_POS",
    "VALCLM_D06H4OUT_CSS_DIF_POS",
    "VALCLM_D06H4IN_CSS_DIF_POS",
    "VALCLM_D06H3OUT_CSS_DIF_POS",
    "VALCLM_D06H3IN_CSS_DIF_POS",
    "CGLOPT_GATED_TUNE_X_MEAS",
    "CGLOPT_GATED_TUNE_Y_MEAS",
],

```

### B.3 Requirements for Training Data

- Require MDI high voltage permission: Belle II is powered and in defined state.
- Require non-constant values for target hit rate: Missing data is artificially filled with previous value by the EPICS archiver appliance and is not usable as training data.
- Require non-constant values for beam currents: Missing data is artificially filled with previous value by the EPICS archiver appliance and is not usable as training data.
- Require both beam currents above threshold of 10 mA: Rings are filled and produce beam background larger than the measurement noise.
- Require target variable over threshold. Currently only for CDC average layer leakage current of 4.2  $\mu$ A: Unpredictable, noisy behavior below threshold.
- Require standard optics settings: Defocused optics are used for debugging, special scans and beam scrubbing, that are not suitable as training data.

### B.4 Decompositions of Diamond Radiation Monitors for $\phi$ -Profile Analysis

The time series for the predictions of all diamonds considered for the  $\phi$ -profile analysis performed in Section 9.8.5 are shown in Figure B.1 for  $\phi = 60^\circ$ , Figure B.2 for  $\phi = 120^\circ$ , Figure B.3 for  $\phi = 180^\circ$ , Figure B.4 for  $\phi = 240^\circ$  and Figure B.5 for  $\phi = 300^\circ$ .

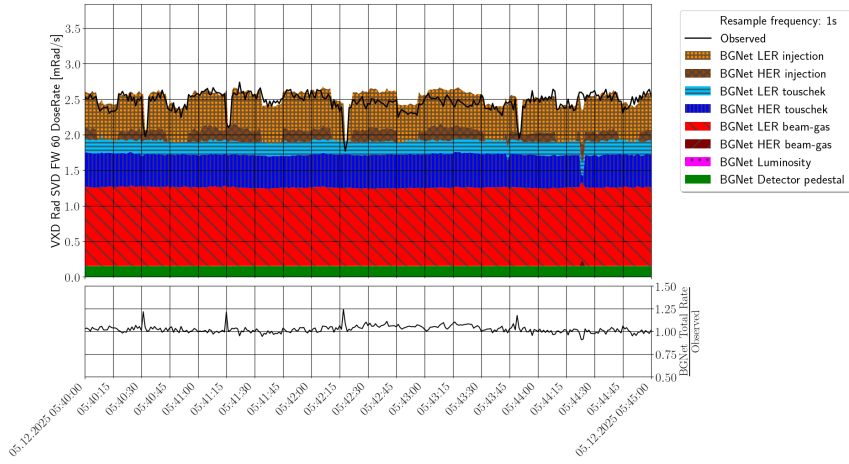


Figure B.1: Hit rate decomposition for the diamond radiation monitor at the SVD forward position at  $\phi = 60^\circ$  during the time window used for the  $\phi$ -profile analysis of the beam background predictions of BGNet.

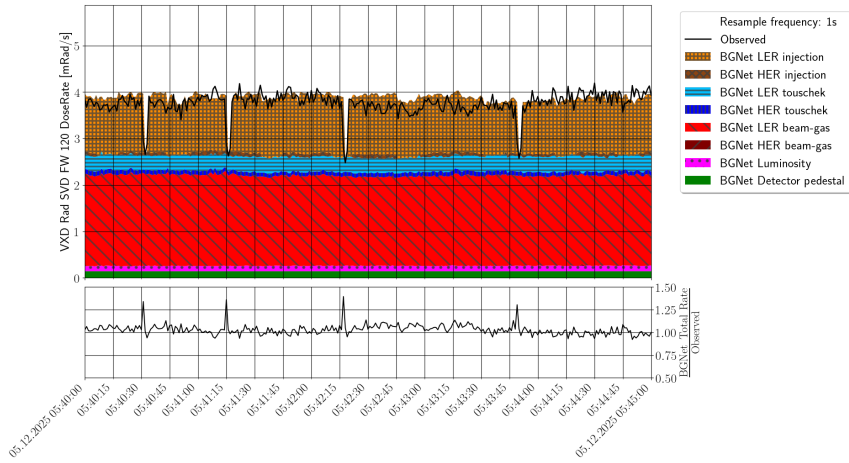


Figure B.2: Hit rate decomposition for the diamond radiation monitor at the SVD forward position at  $\phi = 120^\circ$  during the time window used for the  $\phi$ -profile analysis of the beam background predictions of BGNet.

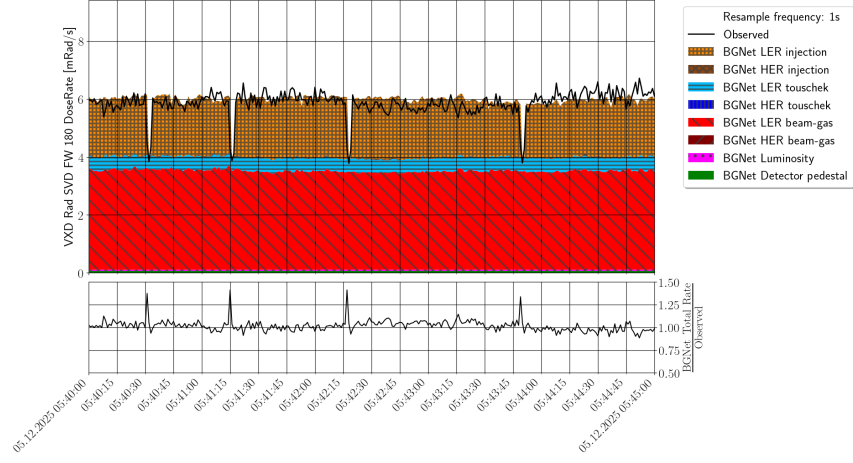


Figure B.3: Hit rate decomposition for the diamond radiation monitor at the SVD forward position at  $\phi = 180^\circ$  during the time window used for the  $\phi$ -profile analysis of the beam background predictions of BGNet.

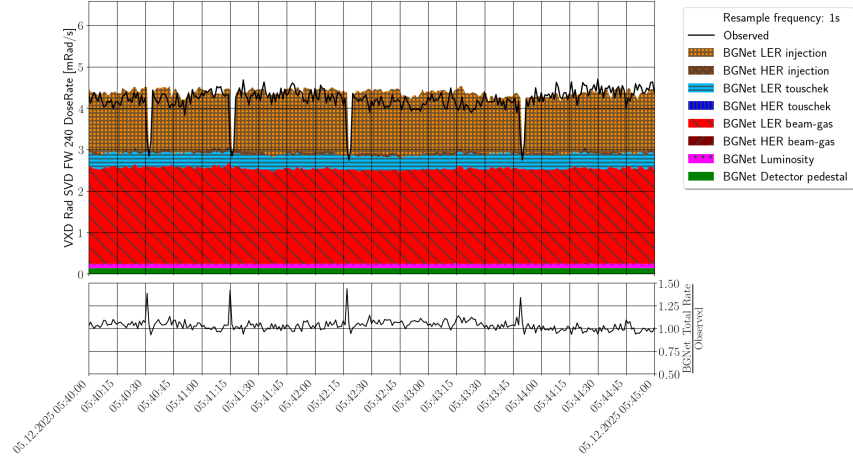


Figure B.4: Hit rate decomposition for the diamond radiation monitor at the SVD forward position at  $\phi = 240^\circ$  during the time window used for the  $\phi$ -profile analysis of the beam background predictions of BGNet.

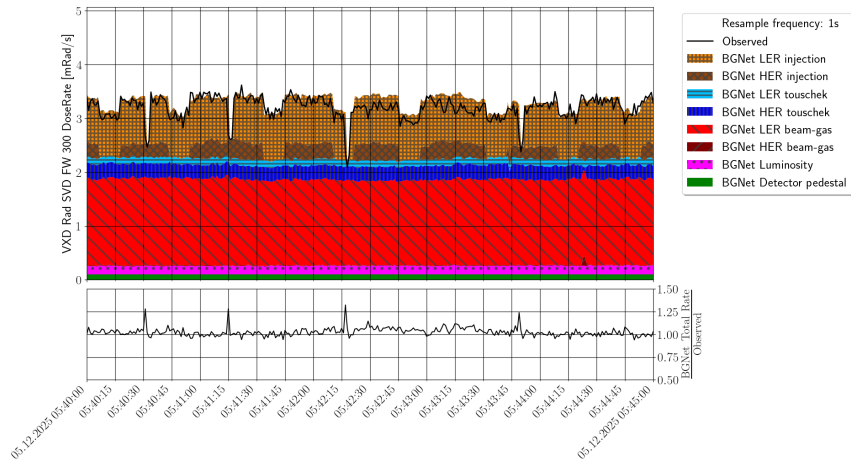


Figure B.5: Hit rate decomposition for the diamond radiation monitor at the SVD forward position at  $\phi = 300^\circ$  during the time window used for the  $\phi$ -profile analysis of the beam background predictions of BGNet.





---

## Acknowledgements

---

I want to thank Prof. Dr. Ariane Frey for fostering such a friendly and happy group and making me a part of it. You have allowed me to grow into the scientist that I have become over the course of so many years. I feel especially grateful for all the travel opportunities and allowing me to explore new places and make new friends along the way.

I am deeply grateful to PD Dr. Benjamin Schwenker for his supervision and mentorship from start to finish. You have lead me back when my work went astray (more than once), taken the time to explain concepts (again) when I was confused (again) and inspired me to be a better physicist. Yet, you have always given me the freedom to tinker, experiment and find out for myself. I have always enjoyed working with you and there have only been very few problems that we couldn't solve in a brainstorming session and a healthy amount of Haribo.

I want to thank the test beam crew and in particular Maximilian Babeluk, Bernhard Pilsl, Alice Gabrielli and Ludovico Massaccesi. Somewhere between mental collapse, sleep deprivation, malnutrition and one too many ouzos at Samarkand we have become friends. You made the long hours in the test beam hut an unforgettable experience and a highlight of my PhD.

My duckiest gratitude goes to Marike Schwickardi, who has been the best desk neighbor one could ever wish for. You have provided emotional support, a target to throw the office ball at and a lot of distraction (from both stress and work). If we had met anywhere but an otherwise empty office at university, we would have not been allowed to sit together in the same room (for good reason). But we did — and you have been there for me during the good and the bad. I enjoyed every bit of our adventure and I hope there are more adventures waiting.

I want to thank Leon Brauns for his friendship since the literal first day I entered the university building. Our weekly game night has been going on for over 4 years now and has become an institution in my life and, especially during the final phase of my dissertation, it was a pillar of my social life.

I also want to thank André Wilhahn for his friendship and being there for so many years. You have been a part of my entire journey at the university and we made many memories together, which I will forever cherish. We have started together and it seems like we will also finish this endeavor together.

Many thanks to my friend and Gym-Buddy Matthias Drescher. You have motivated me time and again to get to the gym and made every training session so much more enjoyable.

Thanks to all the kind people at the institute, particularly those who laughed with me about my jokes. I am lucky to count so many of you among my friends.

Lastly, I want to thank all those who have been there for me and have no clue what happened in this dissertation but were all the more unwavering in their support. Many thanks to Adrian for being my best friend, no matter where life has brought us. Special thanks to Joana for your understanding and love and motivation to continue writing. Last but not least, I want to thank my family — Mama, Papa and Oma — for their support during my studies and beyond.

I would not have gotten this far without you all.

*The test-beam measurements leading to results in this thesis have been performed at the Test Beam Facility at DESY Hamburg (Germany), a member of the Helmholtz Association (HGF)*

---

## Declaration on the Use of ChatGPT and Comparable Tools in the Context of Examinations

---

In this paper, I have used ChatGPT or another AI as follows:

- ☒ not at all
- ☐ during brainstorming
- ☐ when creating the outline
- ☐ to write individual passages, all together to the extent of ...
- ☐ for the development of software source texts
- ☐ for optimizing or restructuring software source texts
- ☐ for proofreading or optimizing
- ☐ further, namely: ...

I hereby declare that I have stated all uses completely. Missing or incorrect information will be considered as an attempt to cheat.