Mgr inż. Maciej Witold Majewski
Katedra Oddziaływań i Detekcji Cząstek
Wydział fizyki i Informatyki Stosowanej

Thesis Abstract

Application of machine learning methods for the analysis of the calibration of the LHCb VELO detector, studies of irradiated silicon pixel sensors and reconstruction of the neutrino interaction in LARTPC detector

This work brings together multiple topics revolving around particle detectors and novel applications of computational intelligence to enhance analysis of data produced by them. As an expert and enthusiast of AI, my main focus of the study is the application of machine learning methods for solving problems in the field of high energy particle detectors. The core of this document relates to the LHCb spectrometer and its high precision silicon Vertex Locator detector both in original strip version and upgraded pixel version. The LHCb is a part of the Large Hadron Collider system of detectors', and its discoveries contributed to the global understanding of fundamental physics. The two data-taking periods, Run 1 (2011-2012) and Run 2 (2015-2018), collected data samples with integrated luminosity of around 9fb -1 . Operating the silicon detectors in such intense radiation field leads to the radiation damage issues. Radiation damage and increased luminosity of LHC are two reasons behind necessity of upgrading the initial design for upcoming LHC Run 3 and 4.

The unique requirements and conditions in which the Velo detector must operate requires an extensive research phase releated not only to hardware studies but also for the new methods of monitoring and analysis of the state of the detector. The latter is presented in this thesis.

The starting chapters, up to Chapter 3 contain the fundamental theoretical and technical introduction to the details of the physics goals, detector composition and analysis methods. The unique analysis of the calibration of the Velo detector in Runs 1 and 2 is described in Chapter 4, and the study of the upgraded detector is contained in Chapter 5.

The first portion of the presented studies investigates two calibration parameters derived from the noise signal picked up by the detector's silicon sensors, pedestal (mean of noise) and threshold (standard deviation of the noise). Those parameters are investigated in views of different dimensionalities.

The analysis of the trend of the pedestals has shown no persisting effects, which confirms that the adaptation of the voltage fed to the detector was in acceptable ranges. The analysis of the threshold parameters revealed the effects of the harmful radiation on the sensors. It inspired the creation of a novel machine learning-based algorithm for the assessment of the calibration itself. I created and implemented an algortihm called ``outlierness'' and successfully introduced it to the monitoring of the Velo detector in 2018. The high dimentionality of the problem (large number of individual sensor readout channels) led to tests of dimentionality reduction algorithms (PCA and autoencoder) and showed that both algorithms could reveal anomalies in the functioning of the detector. The subtle shifts in the relation of the calibration to the noise found in the detector during data taking led us to leverage that effect and use a recurrent neural network to find a method for predicting the need for the calibration of the detector.

The new Velo detector in the upcoming Run 3 exploits pixel matrices for the data taking and increases the number of individual channels more that two order of magnitude.I participated in the

creation of a novel method for finding and tracking clusters of masked (faulty) pixels, using calibration data stream, which will allow for a more informed decision about detector maintenance.

While most of the helpful methods for analysing the calibration of the detector can only be used while the detector is already fully commissioned and functioning, the test data of the sensors can be used for studying the effects of the radiation. The new VeloPix sensor registers a time-over-threshold signal, which can be linked with the energy deposited in the sensor expressed in eV via surrogate function. The change of the function parameters is linked with the amount of total radiation delivered to the sensor. I studied this effect, which led to the development of the basis for breakthrough intelligent method for assessing the fluence of the sensor.

The experience of working with Velo has led to developing the necessary and novel solutions to practical problems presented in Chapter \ref{chap:software}. The contact with the calibration data has led to the creation web-based database service called Storck, which is undergoing the introduction to the LHCb systems during the commissioning of the detector. Along with the database system, an open source framework for efficient creation of the monitoring tools - Titania is presented. Storck and Titania are generic open-source tools that can be used independently in any experiment. I led the development of both projects, along with several developers.

The applications of the machine learning methods for the High Energy Physics detectors usually show high specificity and customisation to the undergoing physics search and design of the detector. While it is usually necessary to develop solutions to particular problems, the problem of tracking and particle identification can be thought of as more universal. The studies in generalisation of machine learning algorithms for particle physics experiments have been largely untouched. The large and robust open dataset created by DeepLearnPhysics has created an opportunity for early search for a more general AI capable of understanding the rules of our universe. The opportunity was leveraged by an ambitious attempt to research the reinforcement learning method's application for particle tracking and identification in the LARTPC dataset. This research is presented in Chapter 8, with a theoretical introduction in Chapter 7.