DOCTORAL DISSERTATION
PREPARED IN THE INSTITUTE OF PHYSICS
OF THE JAGIELLONIAN UNIVERSITY
SUBMITTED TO THE FACULTY OF PHYSICS,
ASTRONOMY AND APPLIED COMPUTER SCIENCE
OF THE JAGIELLONIAN UNIVERSITY

Hyperons @ HADES

Supervised by:
prof. dr hab. Piotr Salabura

Cracow, 2020

Wydział Fizyki, Astronomii i Informatyki Stosowanej
Uniwersytet Jagielloński

Oświadczenie

Ja niżej podpisany Krzysztof Nowakowski (nr indeksu: 1078309), doktorant Wydziału Fizyki Astronomii i Informatyki Stosowanej Uniwersytetu Jagiellońskiego, oświadczam, że przedłożona przeze mnie rozprawa doktorska pt. „Hyperons by HADES" jest oryginalna i przedstawia wyniki badań wykonanych przeze mnie osobiście, pod kierunkiem prof. dr. hab. Piotra Salabury. Pracę napisałem samodzielnie.

Oświadczam, że moja rozprawa doktorska została opracowana zgodnie z Ustawą o prawie autorskim i prawach pokrewnych z dnia 4 lutego 1994 r. (Dziennik Ustaw 1994 nr 24 poz. 83 wraz z późniejszymi zmianami).

Jestem świadom, że niezgodność niniejszego oświadczenia z prawdą ujawniona w dowolnym czasie, niezależnie od skutków prawnych wynikających z ww. ustawy, może spowodować unieważnienie stopnia nabytego na podstawie tej rozprawy.

Kraków, dnia ....

Jakis mądry cytat

Autor *„Zrodlo”*

Ten sam cytat po angielsku

Autor, *“Zrodlo”*
Translation by Tlumacz

Abstract

sOME ABSTRACT

Streszczenie

Jakies streszczenie

Contents

Chapter 1

Introduction

Chapter 2

The HADES detector

Chapter 3

Deta analysis

Chapter 4

Neural networks

4.1 Introduction into artificial neural networks

4.2 The ROC curve and the optimal classifier

One of the most common problem in machine learning is a binary classification, when a data set has to be divided into two subsets, fulfiling serian requirements. A simple example of such a problem is distinction between signal and bacground events in deta collected by experiment. We would like to have a function which takes as agruments set of physical observables (eg. particles' energy, momentum, coordinates of vertexes), represents by $\vec{x}$ and returns sigle number. More formally, a clasyfier can be call any function $h : \vec{x} \to \mathbb{R}$ designed in such a way, that high $h(\vec{x})$values correspond signal events and low $h(\vec{x})$values correspond background event. A threshold value $h(\vec{x})$=c, which is the value separating signal and backgrond events is called a working point, and has to be set by a user. The signal efficiency will be defined as $\epsilon_S = \int d\vec{x}\rho_S(\vec{x})\Theta(h(\vec{x})-c)$ and respectively a background efficiency $\epsilon_B = \int d\vec{x}\rho_B(\vec{x})\Theta(h(\vec{x}) - c)$.

The problems how to represent a clasyfier performence, how to compare different clasyfiers and how to choose proper working point have been discused since many years.

4.3 The data-driven approach

The original paper by Metodiev, Nachman and Thaler [1] the othors show the idea of a data-driven analysis in details. In this chapter I want to introduce main concepts, necessery to understand how the proposed metode helps in week decays recosntruction.

In a classical approach to supervized machine learning, a model learns its properties usign sets of labeled data. Of courese providing good training sets is always a problem. To do this someone can use either experimental data, labeled by a user, or simulation. In first case a user uses his external knowledge about the data to describe it. In necond case the user fully rely on simulation. (opisz zagrożenia)

The data-data driven analysis avoids inconveniencees of two mentioned methodes. It requires neither labeling nor simulation. According to Neyman-Pearson lemma [2] the optimal clasyfier for two sets, A and B is a function given by a dencity ratio

$$h_{opt}^{A/B}(\vec{x}) = \frac{\rho_A}{\rho_B} \tag{4.1}$$

or any monotonous function of $\frac{\rho_A}{\rho_B}$. Assuming that both sets A and B contains signal (s) and bacground (b) events and a statistical distribution of s and b is the same in A and B, we can write (4.1) in the following way

$$h_{opt}^{A/B} = \frac{f_1\rho_s + (1-f_1)\rho_b}{f_2\rho_s + (1-f_2)\rho_b} = \frac{f_1\rho_s/\rho_b + 1 - f_1}{f_2\rho_s/\rho_b + 1 - f_2} = \frac{f_1 h_{opt}^{s/b} + 1 - f_1}{f_2 h_{opt}^{s/b} + 1 - f_2}. \tag{4.2}$$

It can be proven that $\partial_{h_{opt}^{s/b}} h_{opt}^{A/B} > 0$, what means that optimal clasyfier for both cases is the same. It is important to underline that the reasoning gives no clue about the working points for both cases.

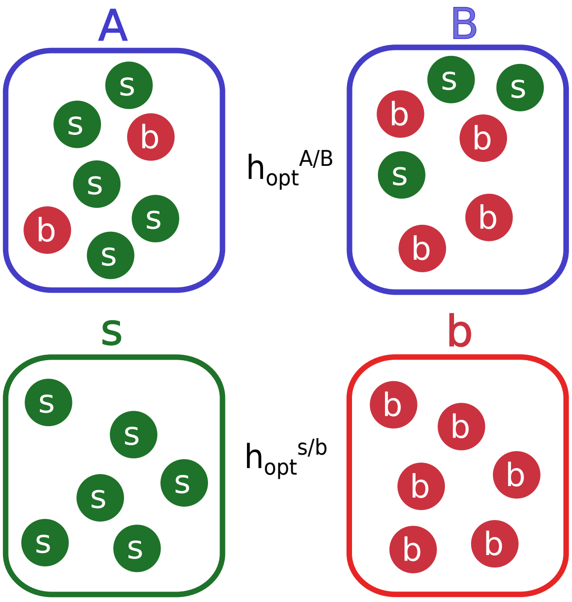

FIGURE 4.1: A data-driven approach visualisation. According to [1] the opitimal clasyfier for sets A and B is equivalent to optimal clasyfier for sets s and b.

4.4 Application for analysis

Chapter 5

Simulations of a new experiment

The HADES collaboration is one of the leading forces of a FAIR Phase-0 project. Within the scope of FAIR project a pp@4.5GeV experiment is going to be pervormed. It gives a great opportunity to measure hyperons' Dalitz decays (see Chapter 1). One of the goals of my work was to carry out a simulation of such an experiment.

5.1 An estimation of cross-sections

In energy range of $1GeV < \sqrt{S} < 6GeV$ an inclusive cross section for $\Lambda(1116)$ and $\Sigma(1193)$ were measured for many different enegys [ref]. Also an inclusive cross section for $\Lambda(1405)$ production was measured for two different energys [ref], and a cross section for $\Lambda(1520)$ in known for one enrgy [ref] in this range. In contrast to the exclusive production cross section , inclusive cross sections for hyperons' production are poorly known.

5.1.1 $\Lambda(1116)$ inclusive cross section

The first step for all estimations is a parametrization of a $\Lambda(1116)$ inclusive prodaction. In a given energy range there are four measured values. More over I made two additional assumptions i) the cross section is equal 0 for threshold enegy, ii) for energy below one pion mass (140MeV) the inclusive and the exclusive cross sections are the same, so for the parametrization I can use cross section measured for $\mathrm{pp} \to \mathrm{pK}^+\Lambda(1116)$ for $\sqrt{S}$ below ??. To the chosen data points I fitted a 3th order polynomial

$$\sigma_{\mathrm{pp}\to\Lambda(1116)X}(\sqrt{S}) = 48 \cdot (\sqrt{S} - 2.55) + 292.6 \cdot (\sqrt{S} - 2.55)^2 - 45.4 \cdot (\sqrt{S} - 2.55)^3. \quad (5.1)$$

Fit result, together with residual plot is show in ??.

5.1.2 $\Sigma(1193)$ inclusive cross section

According to PDG [ref] almost all $\Sigma(1193)$s decay into $\Lambda(1116)$. I means that the inclususive $\Lambda(1116)$ signal contains a fraction deriving from Sz decays. However knowing a relation between $\Lambda(1116)$ and $\Sigma(1193)$ it is possible to disantanle both contributions.

A $\Sigma(1193)/\Lambda(1116)$ ratio was measured by COSY and others [ref], see ??. Additionally the COSY collaboration proposed a parametrization of the ratio for eccess energy $\epsilon < 200MeV$. Above this energy ($\epsilon > 200MeV$) a linear parametziation

$$\frac{\Lambda(1116)}{\Sigma(1193)}(\epsilon) = 2.215 - 2.7 \cdot 10^{-5}\epsilon \tag{5.2}$$

describes data quite well ($\chi^2 = 0.89$). In fact for $\epsilon > 200MeV$ the ratio is almost constant and does not depend on energy.

Knowing the $\Lambda(1116)/\Sigma(1193)$ I was able to disantagle a $\Lambda(1116)$ and $\Sigma(1193)$ production. Using determinated ratio and the $\Lambda(1116)$ production parametrization (eq. 5.1) I created following set of equations,

$$P_1(\epsilon) = \frac{L(\epsilon)}{S(\epsilon)} = \frac{L(\sqrt{S} - \Lambda(1116)_{thr})}{S(\sqrt{S} - \Sigma(1193)_{thr})} = P_1(\sqrt{S}), \tag{5.3}$$

$$P_2(\sqrt{S}) = \Lambda(\sqrt{S}) + \Sigma(\sqrt{S}). \tag{5.4}$$

Solving the first equation and shifting an argument by $\Sigma(1385)_{thr}$ I obtained an equation,

$$\Sigma(\sqrt{S}) \cdot P_1(\sqrt{S} + \Sigma(1193)_{thr}) = \Lambda(\sqrt{S} - \Lambda(1116)_{thr} + \Sigma(1193)_{thr}). \tag{5.5}$$

Now, using eq. 5.5 and 5.4 I got a recurrence relation

$$\Lambda(\sqrt{S} - \Lambda(1116)_{thr} + \Sigma(1193)_{thr}) = P_1(\sqrt{S} + \Sigma(1193)_{thr})\left(P_2(\sqrt{S}) - \Lambda(Sqs)\right). \tag{5.6}$$

Assuming that $\Lambda(\Lambda(1116)_{thr}) = 0$ and $\Sigma(\Sigma(1193)_{thr}) = 0$, the above equation can be solved with any given precision. For the purpuse of cross sections estimation a single step was set $\Delta M = \frac{\Sigma(1193)_{thr} - \Lambda(1116)_{thr}}{10}$, obtained decomposition is shown in ref??.

5.1.3 $\Lambda(1520)$ and $\Sigma(1385)$ production cross sections

5.1.4 $\Xi^-(1322)$

5.2 Background channels selection

5.3 Simulations results

TABLE 5.1: List of signal (S) and background (B) channels for simulated benchmark reactions. Channels marked with * have estimated cross-sections, as described in Chapter ??. Each channel containing Δ Dalitz decay is listed below reference channel, used for cross section estimation.

Channel	σ (μb)	Type
$\Xi^-(1322)$ production		
$\mathrm{pK^+K^+\Xi^-(1322)}$	$3.6/0.35$	S*
$\mathrm{pp\pi^+\pi^+\pi^-\pi^-}$	600	B†
$\mathrm{p\Lambda(1116)K^0_S\pi^+}$	100	B†
$\mathrm{p\Lambda(1116)K^+\pi^+\pi^-}$	30	B†
$\mathrm{n\Lambda(1116)K^0_S\pi^+\pi^+}$	30	B†
$\mathrm{p\Sigma(1193)K^0_S\pi^+}$	20	B†
$\mathrm{ppK^0_SK^0_S}$	20	B†
background channels the same like for $\Xi^-(1322)$ plus below		
$\mathrm{ppK^+K^-}$	20	B†
Dalitz decays of hyperons		
$\mathrm{pK^+\Lambda(1405)[\Lambda(1116)e^+e^-]}$	69.6, BR = 8.4×10^{-5}	S*
$\mathrm{pK^+\Lambda(1520)[\Lambda(1116)e^+e^-]}$	32.2, BR = 5.3×10^{-6}	S*
$\mathrm{pK^+\Sigma(1385)[\Lambda(1116)e^+e^-]}$	56.24, BR = 1.1×10^{-4}	S*
$\mathrm{pK^+\Lambda(1405)}[X]$	69.6	B†
$\mathrm{pK^+\Lambda(1520)}[X]$	32.2	B†
$\mathrm{pK^+\Sigma(1385)}[X]$	56.24	B†
$\mathrm{pp\pi^+\pi^-\pi^0}$	1840	B
$\mathrm{p\pi^+\pi^-\Delta^+[pe^+e^-]}$	2760, BR = 4.5×10^{-5}	B†
$\mathrm{pn\pi^+\pi^+\pi^-\pi^0}$	300	B†
$\mathrm{p\pi^+\pi^+\pi^-\Delta^0}[n\mathrm{e^+e^-}]$	450, BR = 4.5×10^{-5}	B†
$\mathrm{pp\pi^+\pi^-\pi^0\pi^0}$	300	B†
$\mathrm{p\Lambda(1116)K^+\pi^0}$	43	B†
$\mathrm{K^+\Lambda(1116)\Delta^+[pe^+e^-]}$	64, BR = 4.5×10^{-5}	B
$\mathrm{n\Lambda(1116)K^+\pi^+\pi^0}$	20	B†
$\mathrm{\pi^+K^+\Lambda(1116)\Delta^0[ne^+e^-]}$	30, BR = 4.5×10^{-5}	B
$\mathrm{p\Lambda(1116)K^+\pi^0\pi^0}$	10	B†
$\mathrm{p\Sigma(1193)K^0_S\pi^+}$	18	B†
$\mathrm{p\Lambda(1116)K^+\pi^0\pi^0\pi^0}$	7	B†
Real photon decays of hyperons		
$\mathrm{pK^+\Lambda(1405)[\Lambda(1116)}\gamma]$	69.6, BR = 1.3×10^{-2}	S*
$\mathrm{pK^+\Lambda(1520)[\Lambda(1116)}\gamma]$	32.2, BR = 5.0×10^{-4}	S*
$\mathrm{pK^+\Sigma(1385)[\Lambda(1116)}\gamma]$	56.24, BR = 1.1×10^{-2}	S*
$\mathrm{pp\pi^+\pi^-\pi^0}$	1840	B†
$\mathrm{p\Lambda(1116)K^+}$	54.5	B‡
$\mathrm{p\Lambda(1116)K^+\pi^0}$	35	B†
$\mathrm{p\Lambda(1116)K^+\pi^+\pi^-}$	20	B†
$\mathrm{p\Sigma(1193)K^+}$	23.5	B‡
$\mathrm{p\Sigma(1193)K^+\pi^0}$	20	B‡
$\mathrm{p\Sigma(1193)K^+\pi^+\pi^-}$	2	B‡

Chapter 6

Conclusions

Appendix A

The data-driven approach for a neural network training

The original paper by Metodiev, Nachman and Thaler [1] shows the idea of a data-driven analysis in details. In this chapter I want to introduce main concepts, necessery to understand how the proposed metode helps in week decays recosntruction.

In a classical approach to supervized machine learning, a model learns its properties usign sets of labeled data.

Acknowledgements

During the years of my doctoral studies I had the luck to meet a number of great people, without whom this Thesis would never have been created. I would like to thank every one of them for their support, inspiration and friendship.

I would like to express my deepest gratitude to prof. Paweł Moskal for the opportunity to work in his research group, for his supervision over the preparation of this Thesis, and — above all — for his contagious passion for science.

The second person without whom this work would not be possible is dr Eryk Czerwiński. I am greatly indebted to Eryk for his guidance in all my research for the past six years and, perhaps even more importantly, for introducing me to virtually all aspects of academic life.

I would like to extend my special thanks to prof. Antonio Di Domenico, for giving me the possibility to work on the fascinating subject of direct symmetry tests with neutral kaons at KLOE and KLOE-2, and for his careful supervision on my analysis.

I am grateful to prof. Bogusław Kamys for the opportunity to work in the Department of Nuclear Physics of the Jagiellonian University and to prof. Lucjan Jarczyk for all his comments on my work, always motivating me to improve my research and presentation skills.

The time of my work in the Kraków subgroup of KLOE was exceptional thanks my Colleagues Daria Kisielewska, Krzysztof Kacprzak, dr Wojciech Krzemień and dr Michał Silarski. Thank you for the great and inspiring atmosphere and countless help you gave me during these years.

I would like to thank prof. Filippo Ceradini, dr Erika De Lucia, dr Antonio De Santis, dr Paolo Gauzzi and dr Enrico Graziani for sharing their expertise in working with the KLOE data and for their suggestions which helped me overcome several dead ends in my work. I am also indebted to prof. Wojciech Wiślicki for his helpful advice on statistics.

Moreover, my frequent visits in the Laboratories of Frascati would not be the same without the great people I met there and their hospitality, especially dr Elena Perez del Rio, dr Marcin Berłowski, dr Paolo Fermani, dr Gianfranco Morello and all my Colleagues from KLOE.

No less do I owe to my Colleagues from J-PET with whom I have worked on the second part of this Thesis. I would like to especially thank dr Magdalena Skurzok, Monika Pawlik-Niedźwiecka, dr

Grzegorz Korcyl, Szymon Niedźwiecki and dr Sushil Sharma for providing me with building blocks for the calibration and analysis of J-PET data based on their great efforts.

Great thanks also to my officemate Krzysztof Nowakowski, for motivating me when it was time to write, and for saving my sanity with conversations about everything but science when it was time to take a break.

As so many others things in my life, this Thesis would never come to life without the continuous support of my wife Kasia, whom I would like to thank for her patience and unwavering belief that I will finish writing some day. I also owe great thanks to my son Andrzej, for bringing a completely new quality to my life in my last PhD student years, and for his unstoppable will to help me in writing of this text, even if by typing randomly on daddy's keyboard.

Na koniec dziękuję Wam, Mamo i Tato, za wiarę we mnie i za całą pomoc w dotarciu do tego momentu.

This work was supported by the Polish National Science Centre through
Projects No. 2014/14/E/ST2/00262 and 2016/21/N/ST2/01727.

Bibliography

[1] E. M. Metodiev, B. Nachman, and J. Thaler, "Classification without labels: learning from mixed samples in high energy physics," *Journal of High Energy Physics* **2017** no. 10, (Oct, 2017) . http://dx.doi.org/10.1007/JHEP10(2017)174.

[2] J. Neyman and E. S. Pearson, "On the problem of the most efficient tests of statistical hypotheses," *Philosophical Transactions of the Royal Society of London* **231** (Feb, 1933) . https://doi.org/10.1098/rsta.1933.0009.