H. Niedowniczański, Institute of Nuclear Physics
Polish Academy of Sciences

Doctoral dissertation

Defended by

Jakub ZAREMBA

Computing environment for phenomenology of τ lepton decays.

Thesis Advisors:
prof. dr hab. Zbigniew WĄS,
dr hab. prof IFJ PAN Marcin CHRZĄSZCZ

Krakow July, 2018

Abstract

The thesis discusses computing environment for phenomenology of τ lepton decays. The study is performed with TAUOLA Monte Carlo library and concentrates on decays $\tau^- \to (\pi\pi\pi)^-\nu_\tau$. Importance of connection between the experimental data and model building is investigated. Special emphasis is put on available data distributions as a limiting factor in model development.

τ decays apart from being interesting field on their own are also used as a tool for measurement of hard electroweak and also quantum chromo-dynamic processes. Investigation of how different theoretical models can affect such measurements was done with help of Neural Networks on the example of Higgs charge parity state measurement

Last but least, future needs of experiments and possible paths of tools development are studied and discussed together with recent improvements in TAUOLA library. Most notably new initialization is introduced and option for user-programmed models is added. Context behind introduced modifications and features is given.

Streszczenie

Rozprawa przedstawia środowisko obliczeniowe dla fenomenologi rozpadów leptonów τ. Do badań użyto biblioteki Monte Carlo TAUOLA pozwalającej na generowanie rozpadów leptonów τ. Praca koncentruje się na rozpadach typu $\tau^- \to (\pi\pi\pi)^-\nu_\tau$. Tworzenie modeli teoretycznych powinno być oparte o dane experymentalne, ze względu na konieczność optymalizowania przewidywań oddziaływań silnych pośrednich energii. Zbadano jak jakość danych oraz sposób ich przedstawienia może potencjalnie wpływać na uwydatnienie lub ukrycie właściwości fizycznych opisywanych rozpadów.

Rozpady leptonów τ są także narzędziem do badania innych procesów, jak chociażby pomiary parametrów bozonu Higgsa. Takie pomiary mogą być zależne od użytego modelowania rozpadów lepronów τ. Z pomocą sieci neuronowych podjęto próbę oceny jak silna jest to zależność.

Praca opisuje rownież rozwój narzędzia TAUOLA w kontekście potrzeb przyszłych eksperymentów, a także dyskutuje dalsze ściezki rozwoju. Umotywowanie zmian również zostało opisane.

Acknowledgments

My work in the topic of the τ leptons started with master thesis. Since then, for over six years I have received great support and guidance, but also some lenience and patience when required. For all that I would like to thank my supervisor Prof. Zbigniew Wąs.

My cooperation with dr hab. Marcin Chrząszcz shorter, but also very fruitful. For all his help and support I am very grateful.

Ever since I began studies on AGH University of Science and Technology, one person was with me since day one, through whole journey, up until now doing phd at the IFJ PAN. For all the great time together and friendship I want to thank Basia Wasilewska.

I would also like to thank Amanda Bartkowiak, who gave me lots of fun and made me a better person.

They are too numerous to bring all the names here but to all my friends and colleagues who helped me along the way, I am really grateful.

Last but not least, I would like to thank my parents and sisters, who supported me to reach my goals.

Contents

1 Introduction **1**

2 Theory for non specialist **3**
2.1 Introduction to elementary particles 3
2.1.1 τ leptons 5
2.2 Monte Carlo methods 5

3 Precision and nature of experimental data in τ lepton physics **7**
3.1 Past measurements 7
3.2 Future measurements 9

4 Monte Carlo method implementation and its requirement for the τ decay physics **11**
4.1 Basic components and techniques of MC methods 11
4.1.1 Random number generation 11
4.1.2 Changes of variables and rejection method 13
4.2 TAUOLA implementation of MC methods 16

5 τ lepton physics **18**
5.1 Fermi's point-like interaction 18
5.2 Intermediate energy QCD 19
5.3 Beyond Standard Model hypotheses 19
5.3.1 Lepton Flavour Violation 21
5.4 Higgs parameters measurements 21

6 Hadronic τ decay models comparison **23**
6.1 Implementation of theory into Monte Carlo simulation 23
6.2 Essentials of Resonance Dominance Model 26
6.3 Essentials of CLEO model 27
6.4 Essentials of Resonance Chiral Lagrangian model 29
6.5 Comparison of RDM, CLEO and RχL models 31

7 TAUOLA development **42**
7.1 Fitting model parameters 42
7.1.1 Template morphing 42
7.1.2 Fitting analytical distributions 44
7.1.3 Further reflections on fitting 45
7.2 New initialization 45
7.2.1 User defined currents 46
7.3 Future plans 47

8 Higgs CP state measurement with the help of Machine Learning techniques 49
8.1 Introduction to Machine Learning and NN 49
8.2 Assessing systematic errors associated with different models using NN 50

9 Summary 52

Bibliography 54

List of Figures

2.1 Elementary particles/fields and their basic parameters. Particles are assigned into groups and subgroups based on their properties. . . . 4
2.2 Feynman diagram of τ^- decay. 5

3.1 $e^+e^- \to \tau^+\tau^-$ cross section (black) and researched energy ranges of different experiments (red). Data points are taken from $e^+e^- \to l^+l^-$ at CESR, PEP, TRISTAN, LEP. B-factories operate at 10 GeV τ-charm factories at 2-5GeV. 8

4.1 Initial (X-axis) and following (Y-axis) (pseudo)random number generated using middle square method, for first 100 numbers. Visible line indicate correlations between numbers in the sequence. For "good" RNG such plot should look homogeneous. 12
4.2 Visualization of rejection method. If we plot our PDF and randomly generate (x,y) pairs, those below the $f(x)$ (PDF) curve are accepted and those above rejected. 14
4.3 Block diagram of most basic TAUOLA components. Some are omitted, e.g. of phase-space pre-samplers steering. It is to simplify view and because only expert users will actually look into them. 17

6.1 Invariant mass distributions obtained from CLEO (red), RχL (green) and RDM (blue) models. Number of events in $\pi^0\pi^-$ distribution is doubled because both of possible $\pi^0\pi^-$ combinations are used. . . . 33
6.2 8 Dalitz plots for slices in Q2: 0.36- 0.81, 0.81-1.0, 1.0-1.21, 1.21-1.44, 1.44-1.69, 1.69- 1.96, 1.96-2.25, 2.25-3.24 GeV^2. Each Dalitz plot is distribution for RDM model in s1, s2 variables (GeV^2 units). s1 is taken to be the highest of the two possible values of $M^2_{\pi^0\pi^-}$ in each event. 34
6.3 8 Dalitz plots for slices in Q2: 0.36- 0.81, 0.81-1.0, 1.0-1.21, 1.21-1.44, 1.44-1.69, 1.69- 1.96, 1.96-2.25, 2.25-3.24 GeV^2. Each Dalitz plot is distribution for RχL model in s1, s2 variables (GeV^2 units). s1 is taken to be the highest of the two possible values of $M^2_{\pi^0\pi^-}$ in each event. 35
6.4 8 Dalitz plots for slices in Q2: 0.36- 0.81, 0.81-1.0, 1.0-1.21, 1.21-1.44, 1.44-1.69, 1.69- 1.96, 1.96-2.25, 2.25-3.24 GeV^2. Each Dalitz plot is distribution for CLEO model in s1, s2 variables (GeV^2 units). s1 is taken to be the highest of the two possible values of $M^2_{\pi^0\pi^-}$ in each event. 36
6.5 Ratio of distributions obtained from CLEO model (Fig. 6.4) to the one from RDM model (Fig. 6.2). 37

6.6 Ratio of distributions obtained from RχL model (Fig. 6.3) to the one from RDM model (Fig. 6.2). 38
6.7 Ratio of distributions obtained from RχL model (Fig. 6.3) to the one from CLEO model (Fig. 6.4). 39

List of Tables

3.1 Comparison of data samples, center-of-mass (CM) energies and luminosities of selected experiments researching τ leptons. OPAL and ALEPH operated in the years 1989-2000. CLEO started operating in 1979 and stopped the year 2000 and undergone two major upgrades along the way (we exclude CLEO-c here). Belle experiment ran from 1999 to 2010 and BaBar from 2000 to 2008. Belle II started collecting data in 2018. 8

6.1 Partial decay width of $\tau^- \to \pi^0\pi^0\pi^-\nu_\tau$ obtained form different models and from experimental measurements. Results given in GeV/c^2 units. 32

6.2 Analysis of histogram bins of high ratio for compared models (Fig. 6.5 contains ratio A to B we consider here also ratio B to A). Bins where difference exceeds 50% and 100% are counted, as well as number of events in those bins and how big fraction of all events does those bins constitute to. 40

6.3 Analysis of histogram bins of high ratio for compared models (Fig. 6.6 contains ratio A to B we consider here also ratio B to A). Bins where difference exceeds 50% and 100% are counted, as well as number of events in those bins and how big fraction of all events does those bins constitute to. 41

6.4 Analysis of histogram bins of high ratio for compared models (Fig. 6.6 contains ratio A to B we consider here also ratio B to A). Bins where difference exceeds 50% and 100% are counted, as well as number of events in those bins and how big fraction of all events does those bins constitute to. 41

Chapter 1

Introduction

Elementary particles physics is a field devoted to describe most fundamental building blocks of mater. There are 17 particles we consider elementary - 6 quarks, 6 leptons (together called fermions), 4 gauge bosons and 1 scalar boson. This thesis focuses on the heaviest known lepton which is the τ lepton.

The τ leptons were first discovered at SLAC in a series of experiments [1] between years 1974-1977. The mass of the τ lepton is 1.777 GeV/c^2 and a lifetime of 2.9×10^{-13} s [2]. It is the only known lepton that can decay into hadrons, and does so roughly 65% of the times. The decays are P parity sensitive. There are dozens τ lepton decay modes, most of which are not described as precise as the available data. Therefore models aiming to describe τ lepton decays need to be constantly updated.

One type of tools that are used by the experiment are Monte Carlo (MC) generators, which allow modeling of the physical process in simple and comprehensible way, without need for laborious analytical calculations. Such modeling can help finding optimal experimental setup for measurement of specific expected/possible properties of particles and interactions. In τ lepton physics one of the most prominent MC generators is TAUOLA library [3], which allows for modeling of τ lepton decays. Research described in this thesis was dominantly performed with the help of particular MC generator.

The thesis is organized as follows: Chapter 2 is dedicated to non specialists and gives brief introduction on particle physics, τ leptons and MC generators. It is followed by Chapter 3 discussing experimental measurements relevant to the τ leptons and topics brought later in this thesis. Special emphasis is put on precision of measurements as factor dictating our approach towards MC and theoretical models.

Chapter 4 gives further insight into MC techniques and their implementation, concentrating on TAUOLA MC library. Next Chapter 5 gives brief introduction to electroweak interaction, only on the level necessary for understanding topics discussed in following chapters. It also introduces and describes importance of searches strongly related to the τ physics. Topics of intermediate energy (0.5-2GeV/c^2) quantum chromo-dynamics, rare τ decays, measurements of Higgs properties and Beyond Standard Model hypotheses like Lepton Flavour Violation are touched upon.

Comparison of different models for $\tau^- \to \pi^0\pi^0\pi^-\nu_\tau$ decay and its implications for future progress in the field are discussed in the Chapter 6. This chapter is an important supplement to the publication [A] written by the author of this thesis. Some results from that publication are brought up for the context.

Development of TAUOLA library over past few years is recounted in the Chap-

ter 7. First Section 7.1 discusses different approaches to fitting tested on Resonance Chiral Lagrantian model that later became one of the options available through TAUOLA. It is followed by Section 7.2 describing features of most recent initialization [B]. Special emphasis is set on the newly added framework for plugging user-defined models. Possible future improvements are mentioned in last section of the chapter.

Chapter 8 gives brief introduction into vast topics of Machine Leaning and Neural Networks, then follows with practical application of said techniques. Systematical errors associated with different modeling are estimated when Neural Network technique is used for distinction between even and mixed Higgs CP (C -charge conjugation, P- parity symmetries) state.

Chapter 9 summarizes the thesis.

References to papers I am co-author are given numbered by capital letters.

Chapter 2

Theory for non specialist

Field of elementary particles physics and Monte Carlo methods, stand at the base of the research presented in this thesis. Even people who do have some general knowledge in physics often do not know the details of what the τ lepton is. Hopefully this introductory chapter is enough to answer that for people outside the field of particle physics.

2.1 Introduction to elementary particles

Elementary particles are most basic building blocks of nature. What was considered elementary evolved with time as our knowledge expanded, however the basic definition remains the same: The elementary particle is an object that does not have any internal components. These particles can be grouped based on their properties. Figure 2.1 features all currently known elementary fields with their basic parameters and groups to which they are assigned to. Division into the groups is based on characteristics that will be discussed below. Most basic differentiating factor is the spin. Particles with integer spin are called bosons, while those with half spin are fermions.

There are five elementary bosons, whereas four of them are gauge bosons and one is a scalar boson. Gauge bosons (photons, gluons, W and Z bosons) are particles carrying basic forces of nature: electromagnetism, strong force and weak force. In further discussion we ignore gravity, as no established quantum theory of gravity exist. While most people understand forces as invisible springs connecting two interacting objects, in reality it looks more like mail exchange with help of carrying-pigeons (gauge bosons). The only known scalar boson - Higgs particle [4] serves a whole different purpose. Its importance is commonly simplified to giving other particles their masses[1]. It is not within the scope of this thesis to fully explore this topic, but it should be noted that there is way more to Higgs boson than that. The Higgs boson is closely related with other particles couplings also P, CP parity properties. It is even considered to participate in the so called inflation of the universe e.g. [5] or be responsible for vacuum stability [6]. That is why it is sometimes called God Particle [7].

Fermions the construction blocks of the mater we see daily. There are twelve known fermions divided into two basic categories (six particles each). Quarks are the fermions that interact with environment through all forces. Most notably quarks

[1] Acquisition of masses happens via spontaneous symmetry breaking.

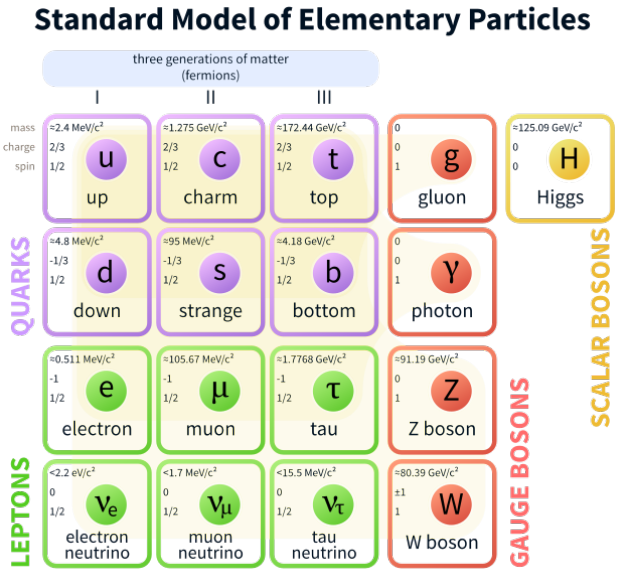

Figure 2.1: Elementary particles/fields and their basic parameters. Particles are assigned into groups and subgroups based on their properties.

interact strongly and are building constituents of all hadrons, like proton and neutron. Quarks can only be seen in bound states called hadrons. Even though they are elementary, they can only exist while interacting with other quark, because of confinement [8]. Leptons have two subgroups based on their interactions. Uncharged leptons are called neutrinos, which can interact only by weak force, while charged leptons can also interact by electromagnetism.

Within fermions we can distinguish three families, which differ by mass and quantum numbers but have the same spin and charge. Particles from heavier families decay into lighter ones (except neutrinos). Therefore, in everyday life we see mater build only with particles from first family.

Apart form described above particles there are also anti-particles. Anti-fermions have opposite charge and quantum numbers, while bosons are their own antiparticles except for W boson (which is charged). W^+ boson is an anti-particle to W^- boson. As far as we know they do not have any additional properties over regular particles. If particle and its antiparticle counterpart meet they annihilate into photons. In addition all known processes of creation both particles and antiparticles are produced in equal amount. That creates the matter - antimatter asymmetry problem. Though the Standard Model[2] [9] provides some mechanisms of CP symmetry breaking, it does not provide any way of breaking lepton nor barion number. Therefore, most of the matter should annihilate based on equal production, leaving almost only photons and no barionic matter. Some of the approaches to tackle this problem will

[2] A theory describing interactions between elementary particles.

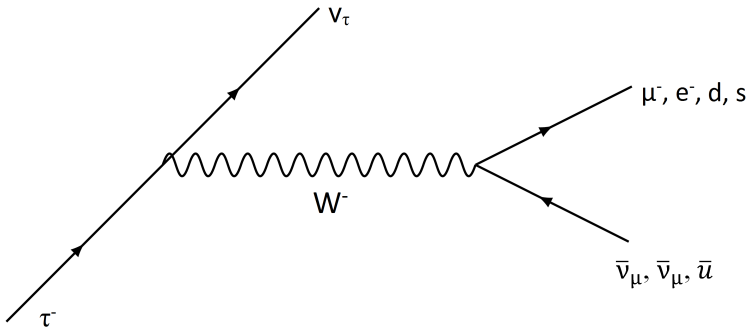

Figure 2.2: Feynman diagram of τ^- decay.

be discussed later.

2.1.1 τ leptons

Mass (1.777 GeV/c^2) of the τ leptons makes them the only leptons with mass higher than the lightest hadron (that is pion $M_{\pi^0} = 0.135\text{GeV/c}^2$) and therefore can decay into them. Hadronic decays cover around 65% of all τ lepton decays [10]. Remaining percentiles are split between decays $\tau \to \mu\nu_\tau\bar{\nu_\mu}$ and $\tau \to e\nu_\tau\bar{\nu_e}$. Taus decay only through weak interaction, by emission of W boson and subsequent creation of ν_τ, which is illustrated in Fig. 2.2. Leptonic decays require quantum electro-dynamics (QED) corrections, while hadronic decays fall into calculation schemes of quantum QCD, of intermediate energy range $(0.5 - 2.0\text{GeV})$).

2.2 Monte Carlo methods

Monte Carlo methods [11, 12] are typically defined as a class of numerical methods based on random number generation. More complicated, but also precise definition is given by J. H. Halton [13]: Representing the solution of a problem as a parameter of a hypothetical population, and using a random sequence of numbers to construct a sample of the population, from which statistical estimates of the parameter can be obtained.

MC methods make no sense when trying to resolve simple problems, but for problem with high variation of initial state or multiple intermediate stages it makes its way. As an example: calculating the odds of getting specific sum of dots when trowing two six sided dices, would be to simple to be worth the effort of writing and using MC simulation, but if you had to trow 10 dices and number of sides of each dice was randomly selected ranging from 4 to 100 then MC simulation might be easier and faster solution. Ultimately it boils down to dimension of the problem - the bigger the dimension, the more useful MC methods are. Weather models, traffic simulations, modeling of diffusion, numerical integration, all of these, use

MC methods.

First published paper on the topic originated in Manhattan Project [11, 12]. In [14] it is said, that E. Fermi already used the method for solving neutron transport equations in nuclear power plants in 1930', but didn't publish anything on it. It was first usage of the method in a modern way, meaning with help of mechanical device - FERMIAC. The very first documented use of randomness was done by G. Compte de Buffon in 1777. Later, his experiment known as Buffon's needle was used by Laplace in 1886 to estimate the value of π.

In particle physics MC generators are invaluable help. They allow for modeling whole experiment from collision to detector response, testing methods of analysis and event selection criteria. They also help with systematic error estimation and even fitting of model parameters (Markow chain MC [15])

Analysis presented in this thesis are done with the help of TAUOLA MC library [3, 16]. It is a MC generator for τ lepton decays with a long history starting in 1991 or even earlier. Recent updates in TAUOLA will be discussed later in this thesis.

Chapter 3

Precision and nature of experimental data in τ lepton physics

Physics is a science strongly connected to empirical observations, therefore, in this chapter precision of the experiments researching τ lepton will be discussed. Some historical overview will be provided, as well as some possible future developments, to complete experimental background for this thesis.

3.1 Past measurements

Since the discovery of the τ lepton [17] experimental environment evolved until establishment of B-factories operating at $\Upsilon(4S)$ peak. It happens so that at those energies cross section for $e^+e^- \to \tau^+\tau^-$ is still relatively high - 0.919nb [18], (cf. Fig. 3.1). Therefore, the B-factories are also superb place to study τ lepton properties. That is not the case for experiments reaching energies high above Z boson peak and those colliding protons, where τ pair production cross section is much lower, while cross section for hadronic processes i.e. background increases.

Comparing experiments associated with LEP facility (OPAL, ALEPH) and operating at the same time CLEO experiment (CLEO II detector [19]) we have two orders of magnitude bigger datasets (of the τ lepton decays) in the latter, see Table 3.1. The advent of B-factories like BaBar and Belle enabled increasing the amount of $\tau\tau$ pairs produced by another two orders of magnitude in following decade. Those experiments concentrating on relatively low energies could achieve big improvement in luminosity.

The sheer amount of data makes measurements at B-factories potentially precise. The experimental data currently reaches precision levels beyond 0.1% in bins of invariant mass distributions coming from decay channels with high branching ratios e.g. $\tau^- \to \pi^-\pi^-\pi^+\nu_\tau$. Such precision already requires QED loop corrections at the level of matrix element calculation. Topic of precision of theoretical predictions will be discussed in greater details in Chapters 5 and 6. Here it should only be stated that experimental precision exceeds that of theoretical predictions by orders of magnitude. Therefore, any models for well established hadronic decays have no sufficient predictive power. Model building has to be performed in a data driven fashion and the result has to be fitted to the experimental data. At the

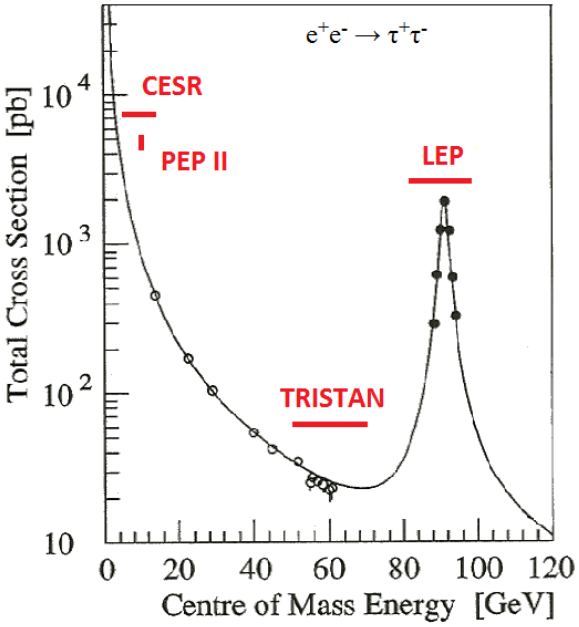

Figure 3.1: $e^+e^- \to \tau^+\tau^-$ cross section (black) and researched energy ranges of different experiments (red). Data points are taken from $e^+e^- \to l^+l^-$ at CESR, PEP, TRISTAN, LEP. B-factories operate at 10 GeV τ-charm factories at 2-5GeV.

Experiment	Accelerator	CM energy	Luminosity [fb^{-1}]	No. τ pairs
OPAL	LEP I/LEP II	90-209 GeV	$\sim$1	200 $\times 10^3$
ALEPH	LEP I/LEP II	90-209 GeV	$\sim$1	300 $\times 10^3$
CLEO	CESR	3.5-12 GeV	$\sim$5	14 $\times 10^6$
BaBar	PEP II	10.6 GeV	513.7	440 $\times 10^6$
Belle	KEKB	10.6 GeV	$>$ 1000	$\sim$1 $\times 10^9$
Belle II	Super KEKB	10.6 GeV	exp. $>$ 40000	$>$ 40 $\times 10^9$

Table 3.1: Comparison of data samples, center-of-mass (CM) energies and luminosities of selected experiments researching τ leptons. OPAL and ALEPH operated in the years 1989-2000. CLEO started operating in 1979 and stopped the year 2000 and undergone two major upgrades along the way (we exclude CLEO-c here). Belle experiment ran from 1999 to 2010 and BaBar from 2000 to 2008. Belle II started collecting data in 2018.

same time, model independent analysis as described in [20] should be the ultimate goal. Until now, the CLEO collaboration was the one closest to achieving this goal and in Ref. [21] declared that such model independent analysis is being performed, while publishing the analysis that has performed a three-dimienional (Dalitz plots in different ranges of Q^2) fit of theoretical model to the $\tau^- \to \pi^0\pi^0\pi^-\nu_\tau$ data. In Ref. [A] it was studied how usage of such distributions instead of 1-dim invariant masses of final state particles can improve tests of theoretical models, as more of underlying properties are disentangled. BaBar collaboration managed to publish only one-dimensional distributions [22]. Ongoing developments in TAUOLA MC, discussed in chapter 7, are in part motivated by the desire to enable/simplify model independent analysis as well as model parameters fitting for τ decays with three scalar particles in final state.

Validation of theoretical models by fitting to the experimental data should be done (if possible) within experimental collaboration even if data is in principle publicly available. Amount of data is not the only thing affecting precision of experimental measurements. Detector design and backgrounds also crucial, and can help or hinder specific subset of measurements. Only experimental collaboration members can really be competent in full understanding systematical errors associated with the detector, triggers, particle reconstruction, background estimation, etc. For example CLEO II detector [19] had very poor separation of charged particles (pions and kaons), while being very good at measuring π^0 mesons. Hence, one-prong 3π channel analysis is the best we have today, while analysis for three-prong channel was never published by collaboration.

3.2 Future measurements

This year, Belle 2 started collecting data. With data sample expected from Belle 2 collaboration we should be able to measure τ lepton decays with branching ratios at the level of 10^{-10}. This gives hope for strong improvement in searches of new physics. Currently upper limits (90% confidence level) on LFV τ decays are at the level of 10^{-8} [10]. This signifies the importance of studying LFV models (and other models predicting New Physics, e.g. second class currents) and puts strong requirement for MC generators to simplify introducing and testing such models into them. This was also addressed in last TAUOLA release [B], and will be discussed in chapter 7, from MC development perspective.

The τ-charm factories of extremely high luminosities could bring even more data than Belle II. Although, the topic is still open, as of now it seems one will be constructed in Novosibirsk [23, 24] There is also similar project considered in China [25]. Before construction begins it is hard to tell exactly how precise will be measurements at super τ-charm factories, but I think it is safe to assume data samples at least one order of magnitude bigger than those we will obtain form Belle II. Also background and reconstruction conditions should be more advantageous.

New data in the field of intermediate energy QCD (which as we discusses is

important to τ lepton decays) may also be obtained from unexpected source of Gravitation Waves measurements from Neutron Stars merger [26]. Already now, those measurements bring some informations about equation of state of neutral matter of those stars. It is possible, that with progress in measurement technology of gravitation waves, we will get some input for models of hadronic τ decays. On the other hand, measurements from medium energy accelerator physics may become useful for the equation of state in Neutron (or quark, strange-quark) Stars. Such possibility certainly is fascinating.

Chapter 4

Monte Carlo method implementation and its requirement for the τ decay physics

In this chapter technical aspects of MC methods will be discussed. We will start with random number generators and their properties. Then typical MC techniques and their implementation will be presented.

4.1 Basic components and techniques of MC methods

4.1.1 Random number generation

Monte Carlo event generator is a stochastic tool using random numbers, therefore its most elementary component (or elementary functionality) is a (pseudo)random number generator (RNG). Here we should put distinction between true random numbers and pseudo-random numbers. True random numbers are those generated in completely unpredictable and unrepeatable way, e.g. from physical experiment. Throwing a good (unloaded) die would be example of obtaining true random number, though limited to integer in range 1-6. Pseudo-random numbers are generated according to strict mathematical formula. They however should have all the statistical properties of true random numbers, hence from now on, we will call them random numbers. A typical example of random number generation is a middle square method [27], invented by J. von Neumann, where you choose a number from specific range (using specific number of digits) e.g. 1-1000, square it and then use middle numbers of the result as your new random number. For example you can start with 234, which gives 54756. Then, you take middle of that - 475, squared it gives 225625, therefore 56 is your next random number. If we continue with that choice, we will eventually arrive at loop: 16, 5, 25, 2, 4, 16. Middle square method is flawed in many aspects, but that makes it great example for what is important in random number generators.

First of all, RNGs require a seed (or seeds). Seed is a number that starts the sequence of random numbers and once set, it determines output. The sequence can be therefore recreated using the same seed. Since the sequence can be recreated it is not truly random (that is why earlier we used term pseudo-random). This

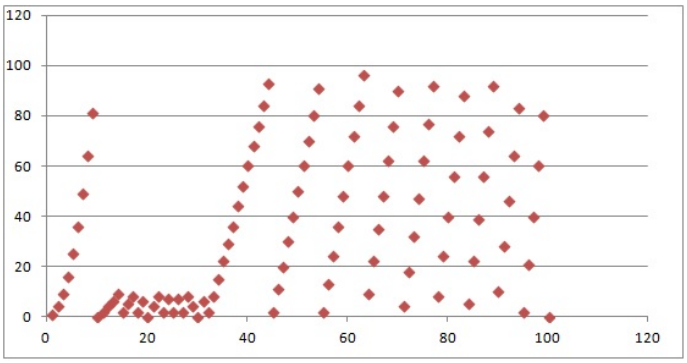

Figure 4.1: Initial (X-axis) and following (Y-axis) (pseudo)random number generated using middle square method, for first 100 numbers. Visible line indicate correlations between numbers in the sequence. For "good" RNG such plot should look homogeneous.

is efficient and useful feature for tests of MC generators. We can e.g. make sure that our computation using RNG gives same result regardless of operating system, user-defined setup, hardware, etc. - comparison can provide installation tests. It is useful during program restructuring, e.g. rewriting into different programing language. Such ability to debug program would not be possible with true random numbers, because true random sequence cannot be reproduced.

Second important lesson from middle square method is that RNG can easily revert to infinite loop or single number. This is an extreme example and most RNGs do not revert into single number or a infinite loop, but they do start repating the sequance of generated numbers eventually. Amount of numbers that can be generated from given method before repeating the sequence is called a cycle (or a period). Long cycle is one of the properties describing good RNG. What defines long cycle depends on the problem you want to resolve, but we usually are talking about length of $\sim 10^{30}$ different numbers. For comparison our example had only 12.

Other crucial characteristic of RNG is a lack of correlations in the sequence of generated numbers. In our example of middle square method correlations between two consecutive numbers can be seen in Fig. 4.1 as clear lines. Same thing may occur for a pair of numbers with any number of others in between and is always a bad thing, as it is definite proof of produced numbers not being random. Similar to Fig. 4.1 plot for good generator should look more like plot of white noise, with no distinct features. This characteristic (lack of correlations) is often called randomness.

Random numbers are usually used in huge quantities. Speed of generation was an issue and is one of the factors defining a good RNG. It limited how complicated mathematical formula could be used. Nowadays computational power available makes the issue of speed less and less important. That being said, we need to acknowledge that RNGs almost always were faster, than obtaining sequence of true random numbers from physical experiment. Speed was at the beginning one of the

main factors behind development of RNGs

While talking about speed, it should be said that parallelization (in terms of using different machines or processor cores) can easily be obtained by setting different RNG seed for each processing core which are then run in parallel. If N generations are performed that way, we can obtain the sample of the required size N-times faster. This, of course, puts on the generator additional requirement of randomness between independently generated series of random numbers, so the results can be summed without introducing new correlations. It is a simple but efficient way for parallelization.

4.1.2 Changes of variables and rejection method

While RNGs generate random numbers from flat distribution[1], for practical applications we usually need random numbers following a particular probability distribution function. There are two main ways of obtaining such number, von Neuman's rejection method (also called elimination method or acceptance-rejection method) and changing the variable by reverting cumulative density function (CDF).

Let us describe the latter method first, starting with some definitions. Assume we have random variable x. Probability distribution of that variable is a function that for any given x is equal to probability of randomly obtaining this x. In continuous space ($x \in [a, b]$) such definition is useless because probability of getting any particular number from infinity of possibilities is 0. Therefore we use probability density functions (PDF), which should be treated as relative likelihood of obtaining specific value. PDF ($f(x)$) is defined such that integral over given range of x is equal to probability of obtaining result from that range. CDF ($F(x)$) is a function which value at point X is a probability of obtaining $x \leq X$ when x has stated before PDF distribution. Hence, $F(x)$ must be continuous[2], non decreasing and fulfill: $F(-\infty) = 0$ and $F(\infty) = 1$. PDF and CDF are connected by formula:

$$F(X) = \begin{cases} 0 & \text{for} X \leq a \\ \int_a^X f(x)dx & \text{for} X \in (a, b) \\ 1 & \text{for} X \geq b \end{cases} \tag{4.1}$$

Reverting CDF method uses changes of variable such that:

$$R \in (0, 1),\ R = F(X)\ \rightarrow\ X = F^{-1}(R). \tag{4.2}$$

Then, X has a PDF distribution of function $f(x)$, while R is generated from RNG, therefore has flat distribution. Such change of variables is correct because:

$$F(X) = P(R \leq F(X)) = P(F^{-1}(R) \leq x) = P(X \leq x), \tag{4.3}$$

[1] At least most of them, but one can find RNG that produces numbers from e.g. Gauss distribution.

[2] For a continuous random variable as considered here.

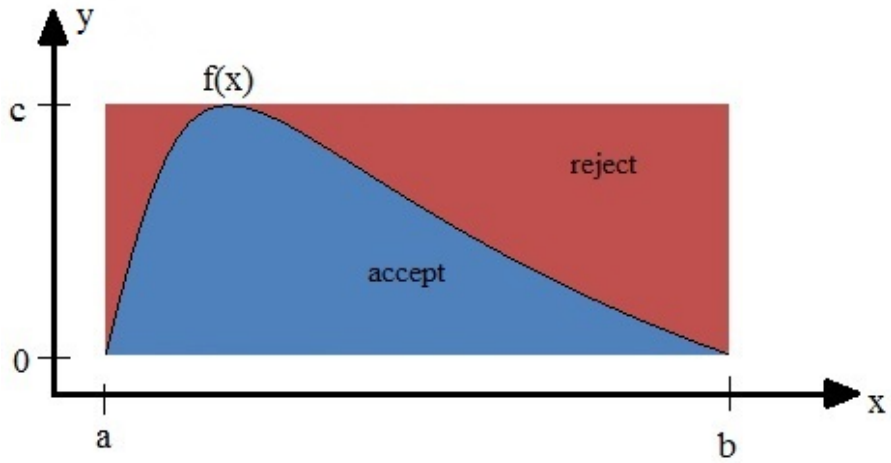

Figure 4.2: Visualization of rejection method. If we plot our PDF and randomly generate (x,y) pairs, those below the $f(x)$ (PDF) curve are accepted and those above rejected.

where $P(A)$ is probability of A occurring.

Changing variable in the presented way is an easy and accurate method. It is also efficient as one random number from RNG gives you one number from given CDF. Downside is, that only simple CDF can be reverted analytically and therefore used in fast manner. Numerical CDF reverting will decrease speed and accuracy.

Rejection method is in its premise similar to physical experiments, where measurement of one variable indicate value of other. An example of such would use of Buffon's needle to estimate value of π. Buffon's needle in its original form was posed as a mathematical question of geometrical probability. "*Suppose we have a floor made of parallel strips of wood, each the same width, and we drop a needle onto the floor. What is the probability that the needle will lie across a line between two strips?*" But knowing actual probability (which can easily be calculated and depends on needle's length and strips width) and performing the experiment numerous times one has simple MC method of estimating π (inversely proportional to the probability). Hence, by measuring where needle was dropped, we can actually measure π.

Let's say we want to generate random number X from given PDF of $f(x)$, where $x \in [a, b]$ and PDF value ranges from 0 to c. We can obtain it by generating a pair of numbers: $r_1 \in [a, b]$, $r_2 \in [0, c]$ and if $f(r_1) > r_2$ then assign $X = r_1$ (accept event), else repeat generation (reject event). This is rejection method in its simplest form. Visualization of such method is presented on Fig. 4.2.

To prove the rejection method we need to show that accepted events actually are of desired PDF:

$$P(X < t) = \int_a^t f(r_1) dr_1. \tag{4.4}$$

r_1 and r_2 are generated from flat distributions of PDFs $\frac{1}{b-a}$ and $\frac{1}{c}$ respectively. Therefore:

$$P(X < t) = P(r_1 < t | r_2 < f(r_1)) = \frac{P(r_2 < f(r_1) | r_1 < t) P(r_1 < t)}{P(r_2 < f(r_1))}, \tag{4.5}$$

$$P(r_2 < f(r_1)) = \int_a^b \frac{dr_1}{b-a} \int_0^{f(r_1)} \frac{dr_2}{c} = \frac{1}{(b-a)c}, \tag{4.6}$$

$$P(r_2 < f(r_1) | r_1 < t) P(r_1 < t) = P(r_2 < f(r_1) \cap r_1 < t) = \int_a^t \frac{dr_1}{b-a} \int_0^{f(r_1)} \frac{dr_2}{c} = \frac{1}{(b-a)c} \int_a^t f(r_1) dr_1, \tag{4.7}$$

$$P(X < t) = \frac{\frac{1}{(b-a)c} \int_a^t f(r_1) dr_1}{\frac{1}{(b-a)c}} = \int_a^t f(r_1) dr_1. \tag{4.8}$$

For multidimensional PDFs rejection method works in the same manner. If we have m-dimensional variable, we need to generate $m + 1$ variables and apply: if $f(x_1, x_2, ..., x_m) < x_{m+1} \rightarrow$ accept event, else $\rightarrow$ reject. Multichannel generation, where each channel takes care of distinct structure of enhancements is important to note. We will not go into details here. In case of TAUOLA an example is explained in Section 2 (`PHASE SPACE AND MATRIX ELEMENTS`) of ref [3]. Such methods can always be confirmed with mathematically formal considerations. We will not cover the details as they are lengthy and would require a lot of repetition of published results which are not essential for the present work. We will also not discuss cases, where distribution features extremely sharp peaks like Dirac delta.

4.1.2.1 Issues of numerical stability

Numerical stability can be divided into two main aspects - efficiency of generation and precision of the result. More often than not, those issues can be connected. If required structure of singularities is demanding, density of distribution varies a lot over the phase space, vast spectrum of difficulties may appear, usually related to rounding errors or due to fine distribution of random numbers over some subregion of (0,1) range. That is one of the aspects which needs to be checked whenever precision range improves substantially. Test with semi-analytical results are then useful, but even more effective are the tests with analytic results. First papers on TAUOLA [28] collect such semi-analytical and analytical results for comparison with MC results.

Previous subsection described methods for generating numbers of given PDF. Depending on the choice, some optimization is usually required. In physical problems we usually deal with hard or impossible to revert CDFs. Therefore some variation of rejection method is usually used. Looking at Fig. 4.2 one should notice that for PDFs with very high and thin peaks this method will be extremely inefficient because most of the random points will land above the PDF curve and therefore be rejected. This can be mitigated by defining envelope distribution $g(r_1)$ such that:

$$f(r_1) \leq cg(r_1). \tag{4.9}$$

PDF $g(r_1)$ must be proven to be efficient for such MC use. Then we change variables same way as in first part of this section:

$$R \in (0,1),\ R = G(r_1)\ \rightarrow\ r_1 = G^{-1}(R) \tag{4.10}$$

and

$$G(r_1) = \int_a^{r_1} g(R)dR,\ r_1 \in (a,b). \tag{4.11}$$

Second generated variable is still $r_2[0,c]$ and we reject event if $\frac{f(r_1)}{g(r_1)} > r_2$. Here we should mention that poorly chosen $g(r_1)$ can cause all kinds of problems if e.g. at any point $\frac{f(r_1)}{g(r_1)}$ goes to ∞ or 0.

4.2 TAUOLA implementation of MC methods

MC library TAUOLA has a modular structure allowing for independent tests, optimization or substitution of most important components. Those are RNG, phase space generators, matrix elements and hadronic currents (for semi-leptonic decays only), see Fig. 4.3. Those are treated as black boxes (at least at some stages of work) and no internal structure is enforced, but arguments of routines/functions need to be maintained for communication between blocks.

Hadronic currents are constituents of matrix elements but they contain most of the theoretical assumptions differentiating theoretical models for particular decay. That is why they are coded as independent blocks. Matrix elements together with phase space effectively define PDF of random variables (four-momenta of final state particles) required for generation of τ decay events. Their definitions are given in Chapter 6.

TAUOLA by default uses RANMAR [29, 30] RNG. It can easily be replaced, but as of today no experimental collaboration using TAUOLA has claimed to replace it with more modern generators. RANMAR should still be sufficient in the upcoming years.

τ decay events from TAUOLA are generated using variant of rejection method with weighting events and pre-sampling for maximum weight used to determine whether or not algorithm accepts the event. Usually this method is called Importance Sampling. Section 3 of [28] gives exact technical realization. Multichannel generation is used to control complicated structures of enhancements. It is worth to

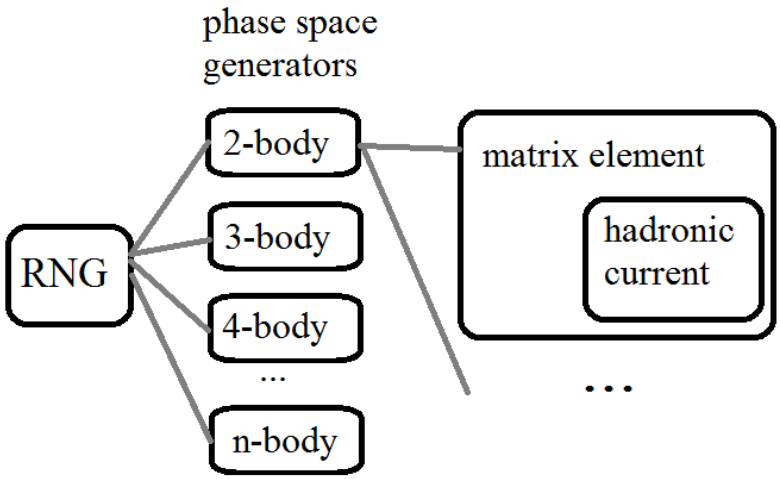

Figure 4.3: Block diagram of most basic TAUOLA components. Some are omitted, e.g. of phase-space pre-samplers steering. It is to simplify view and because only expert users will actually look into them.

note that some of these features of phase space generation are predominantly needed for tests, when in reference analytical calculations narrow width approximations are used.

Chapter 5

τ lepton physics

This chapter gives introduction to the areas of physics directly connected to the τ lepton. Main aspect under consideration is precision predicted by theory, while details irrelevant to my research will not be discussed.

5.1 Fermi's point-like interaction

The τ decays can be described by Feynman's diagram in Fig. 2.2. From there we can use Feynman rules to write down matrix element's general formula, where because of its large mass W boson propagator contributes a factor to a G constant only:

$$\mathcal{M} = \frac{G}{\sqrt{2}} J_\tau^\mu \cdot J_\mu, \tag{5.1}$$

where G is the Fermi constant and $J_\tau^\mu = \bar{u}(N)\gamma^\mu(v + a\gamma_5)u(P)$ is a leptonic current associated with decaying τ lepton. The v and a are handedness defining constants, M is the mass of the τ lepton, P is a four-momentum of decaying τ, q_i's are four-momenta of outgoing particles except ν_τ and N is a ν_τ four-momentum. J_μ is a current associated with second vertex in the diagram Fig. 2.2. J_τ^μ and J_μ are independent[1], which is very useful for semileptonic decays, where J_μ is a hadronic current describing QCD interactions of final state hadrons.

Formulas given in this section have build in two assumptions that slightly limit the precision on MC predictions. First one is lack of loop corrections, which is acceptable for precision down to about $\alpha/\pi \approx 0.2\%$. Second one is reducing W propagator $-i\frac{g^{\mu\nu} - q^\mu q^\nu / M_W^2}{q^2 - M_W^2}$ to $i\frac{g^{\mu\nu}}{M_W^2}$ using $q^2 < m_\tau^2 \ll M_W^2$. It yields error, depending on exchanged virtuality (hence mass of decay product) [31]:

$$\Gamma = \Gamma_{approx.} + \mathcal{O}\left(\frac{3}{5}\frac{M_\tau^2}{M_W^2} - 2\frac{m^2}{M_W^2}\right) \tag{5.2}$$

where m, M_τ, M_W are the mass of decay product, τ lepton and W boson, respectively. Assuming extreme case of decay into some resonance of mass almost equal to τ it cannot exceed ∼0.07% and usually it is below ∼0.04%.

One loop QCD radiative corrections introduce $\sim 10\% = \frac{\alpha_{QCD}}{\pi}$ enhancement to semileptonic decay rates [32]. Those should be introduced in hadronic current as an overall factor, but in general it relies on fits to experimental data anyway.

[1] For leptonic channels such formulation is reaching limits of experimental precision and loop corrections will most likely be necessity for Belle II experiment.

Theoretical constraints are generally not sufficiently precise, as will be discussed in next section.

In the following subsections we will briefly describe areas of physics that could greatly benefit from further exploration of τ lepton decays.

5.2 Intermediate energy QCD

While leptonic τ decays can be relatively well described by weak interaction of W boson and QED, that is not the case for semileptonic decays that fall into low/medium energy (0,1-2GeV) QCD calculation range. No fully reliable analytical QCD methods exist for this energy scale. Perturbative QCD while giving us most precise test of the theory like deep inelastic scattering, LHC measurements and other processes [32], is less suitable for low energies due to bad convergence of perturbative expansion in α_{QCD}. Strong coupling constant is large at those energy scales and therefore makes the calculations non-predictive. Using spontaneous (chiral) symmetry braking one can obtain approximate effective field theories for low energies, such as Chiral Perturbative Theory [33], where planar QCD diagrams are summed to all orders.

χPT has approximate SU(3) symmetry in flavour space where available degrees of freedom are lightest pseudo-scalar mesons: π^+, π^-, K^+, K^-, π^0, K^0, $\bar{K^0}$, η. Its symmetry would be exact if mesons were massless and lightest hadrons were scalars (Goldstone bosons). Because neither of those statements is true, chiral symmetry of χPT is broken. It gives good description of low mass hadrons though, usually limit of applicability is put just below mass of lightest vector particle: $M_\rho = 0.77$GeV. This is not sufficient for τ lepton decays, therefore extrapolations to higher energies were proposed, most notably Resonance Dominance Model [34] and Resonance Chiral Lagrangian [35]. Their implementation into TAUOLA for modeling of $\tau \to \pi\pi\pi\nu_\tau$ decay will be discussed in the next chapter. Here, let us only mention that perturbation expansion can be used also in χPT, but again up to a limit. One can not expect precision of resulting model to be better than 5-10%, may be even 20%, see discussion in [36]. This is by far not sufficient when compared with the precision of experimental data. On the other hand properly analyzed data can provide input for future model builders.

5.3 Beyond Standard Model hypotheses

The Standard Model (SM) [9] describes the strong, weak and electromagnetic interactions via exchange of gauge bosons and is considered a big achievement of modern science. At least last 30 years of experimental research is mainly confirming SM predictions, e.g. recent discovery of Higgs boson [37]. That being said, SM can never be ultimate theory of Particle Physics because it does not include gravity which should become relevant to particle interactions near the Planck scale (10^{19}GeV). Apart from that we do have some experimental issues for SM predictions e.g.: matter/ani-

matter asymmetry [38], neutrino mass [39], dark matter [40]. Theories that try to address the topics are generally called Beyond Standard Model (BSM) or New Physics (NP). In this section I want to briefly describe areas where τ physics is especially relevant to those searches.

Dark matter is a term describing unidentified, invisible constituent of our universe that interacts by gravity and, we suspect, also by weak force. It does not interact by strong and electromagnetic force. A this point all we know about dark matter comes from astronomical observations. If dark matter is made out of elementary particles, they are not present in SM, hence many models predict additional particles that could fit the description. Most scientists lean towards supersymmetric (SUSY) models [41], which predict additional set of particles, similar to that of presented in Fig. 2.1. It introduces additional symmetry between bosons and fermions. According to this theory each elementary particle present in SM has a supersymmetric partner of other type (boson for fermions and fermion for bosons). Naming scheme of such particles goes like this: supersymmetric partners to fermions gain a prefix "s", while partners to bosons get a sufix "ino". Therefore, we get names like squarks, sleptons, Higgsino. Supersymmetry is considered by many as most natural extension of SM. Since τ lepton has relatively large mass those particles could potentially have large Yukawa coupling to τ leptons, hence making them great tool for this type of searches. This tool is as good as our understanding of τ decays. The better modeling we have, the more we can do with measurements.

In current definition of SM neutrinos are assumed to be massless. Recently observed neutrino oscillations, require neutrinos to have a non-vanishing mass or SM to be modified. Oscillation does not allow for direct mass measurement of mass, but allows for probing relation between masses of neutrinos e.g. [42]. Currently upper limit on τ neutrino mass comes mainly from cosmological measurements [43]. Measuring or improving limits on ν_τ mass may also be possible through measuring/setting bounds on Lepton Flavour Violating τ decays. Therefore τ measurements are extremely important for gaining further insight on the topic.

Widely accepted theory of the Big Bang together with standard model predicts equal production of matter and antimatter at the very beginning of our universe. If that was the case, as universe expanded all matter should ultimately annihilate into photons. Therefore scientists search for any signs of inherent asymmetry between matter and antimatter or processes that could lead to it. On atomic scale, recent measurements of anti hydrogen atoms showed exact symmetry in energy spectrum with regular hydrogen [44]. Same experiment tested the idea of anti-gravity from antimatter. Results were leaning towards same gravitational interaction of matter and anti-matter, but uncertainty of the measurement make it inconclusive [45].

On elementary particle scale we look for processes that directly violate barion/lepton number or lepton flavour. τ leptons could in theory provide us with such processes. For instance they are more massive than protons, therefore any experimental sign of taus decaying into barion would be a breakthrough. Other option would be finding Lepton Flavour Violating decay. Those will be discussed in following subsection.

5.3.1 Lepton Flavour Violation

Lepton Flavour Violation (LFV) is one of possible solutions to matter/ani-matter asymmetry. Experiments are constantly searching for LFV processes, but they are yet to be found. In the meantime many models and theories are trying to put predictions on where to find them. Among these theories are (minimal) supersymmetry (SUSY) [41], seesaw models [46], little Higgs scenarios [47] and models with four generations of fermions e.g. [48].

Since we did not observe experimentally SUSY particles yet, the symmetry must be broken, therefore they must be heavier than experimentally available production threshold. With that comes main criticism of the model - in principle the lightest supersymmetric particle can be so heavy that in no foreseeable future we can produce it and test the theory. For example in [49] optimistic prediction was that LEP II could already find evidence of minimal SUSY. Now, 25 years later, we have 2 orders of magnitude higher energies available, and SUSY predictions are again pushed just to the next generation of accelerators [50].

Little Higgs scenarios predict additional particles as well, namely vector like quarks and additional bosons. Those models are more constrained than SUSY and new particles cannot be heavier than roughly 10 TeV [51]

Models with fourth generation of particles, as the name suggests predict another set of fermions. In particular existence of fourth generation leptons strongly affects branching ratios of LFV τ decays [52], where we able to measure them. Such measurements could exclude competing, previously mentioned, theories for NP.

The problem of LFV can be tackled also by an effective field theory approach like in [53]. Such approach allows for estimation of all possible operators of required symmetry with arbitrary couplings. This gives opportunity to test experimental efficiencies in searches of LFV processes.

All mentioned earlier models are important in respect to the τ physics, because they predict mechanism for decays like $\tau^- \to l^-\gamma$ or $\tau^- \to l^-l^-l^+$ which violate lepton flavour. Those types of decays do not require building more powerful accelerators, only collecting more data. Therefore, they provide a simple way of searching for new physics.

5.4 Higgs parameters measurements

In previous sections of this chapter we have briefly discussed areas of physics where τ lepton decays are in the center of the attention themselves, but they can also be great tools for measurement of properties of particles decaying into τ's like Higgs boson. Once discovered, main aspect of Higgs boson measurements is to check weather or not it has properties predicted by Standard Model.

SM predicts Higgs particle to be scalar boson (spin zero particle with symmetrical wave function under CP transformation), therefore deviation from that would be clear indication of NP. The τ leptons is a great tool for measuring Higgs parity because of its large mass thus large coupling to Higgs and P-parity sensitive

decays [54]. Also CP properties of Higgs as well as Z bosons can be studied with the help of the τ lepton final states. The tools like [55] are useful. Not all τ decay channel can be used, also those with high branching ratios are most convenient. The more complicated decay channel, the less precise theoretical modeling. With the help on Machine Learning one can derive methods to exploit more complicated decay channels and study the influence of different decay models ambiguities on the measurement [C]. The multidimensional nature of such measured distributions may offer a window to better models of intermediate energy range of strong interaction results.

Measuring Higgs Yukawa couplings to τ is another important test of the Standard Model. Interestingly first measurements pointed at disagreement [56], but further analysis showed that $H \to \tau\tau$ rates are within SM prediction [57]. BSM hypothesis like charged Higgs are also likely to be tested through usage of τ's [58]. But we will not concentrate on these studies.

Chapter 6

Hadronic τ decay models comparison

As was already mentioned, the τ lepton decays into hadrons are excellent probes of intermediate energy QCD. At the same time we lack good theoretical prediction for that range[1], therefore we rely on models that usually are of limited predictive power. In this chapter we will discuss some models of τ decaying into three pions and τ neutrino. Conclusions of this comparison apply also to other more complicated decay modes. Hadronic τ lepton decays are discussed here in the context of application in MC generator TAUOLA [3, 28].

6.1 Implementation of theory into Monte Carlo simulation

In the following section I will recall descriptions, use notation and naming conventions from [3] and in following couple of sections we also use notation from [A].

When coding theoretical model into MC we are looking for way of calculating differential partial width defined in Eq. 6.1:

$$d\Gamma_X = G^2 \frac{v^2 + a^2}{4M} d\mathrm{Lips}(P; q_i, N) \times |\mathcal{M}|^2, \tag{6.1}$$

where: M is the mass of the τ lepton and q_i's are four-momenta of outgoing particles except ν_τ. It is a product of the flux factor, phase-space and matrix element squared. It is worth noticing that the equations given in this chapter are given for the case of the τ lepton decaying into three scalars and neutrino. Description given here is not original, but it is required for references in later sections e.q Sect. 7.1, so we recall it in a necessary details.

The event generation in MC starts with use of a phase-space parametrization, which for four body decay, is described by formula 6.2:

[1] Especially for decays with three or more hadrons in the final state. Simpler channels, described by single form factor are usually considered as well described, though we use in part empirical models.

$$
\begin{aligned}
&d\mathrm{Lips}(P; q_1, q_2, q_3, q_4) = \\
&\qquad \frac{1}{2^{17}\pi^8} \int_{Q^2_{min}}^{Q^2_{max}} dQ^2 \int_{M^2_{2,min}}^{M^2_{2,max}} dM_2^2 \\
&\qquad\quad \times \int d\Omega_4 \frac{\sqrt{\lambda(M_2^2, Q^2, m_4^2)}}{M_2^2} \\
&\qquad\quad \times \int d\Omega_3 \frac{\sqrt{\lambda(Q^2, m_3^2, M_2^2)}}{Q^2} \\
&\qquad\quad \times \int d\Omega_2 \frac{\sqrt{\lambda(M_2^2, m_2^2, m_1^2)}}{M_2^2}, \\
&Q^2 = (q_1 + q_2 + q_3)^2, \qquad M_2^2 = (q_1 + q_2)^2, \\
&Q_{min} = m_1 + m_2 + m_3, \qquad Q_{max} = M - m_4, \\
&M_{2,min} = m_1 + m_2, \qquad M_{2,max} = Q - m_3.
\end{aligned}
\tag{6.2}
$$

and

$$\lambda(x, y, z) = (x - y - z)^2 - 4yz. \tag{6.3}$$

The formula is exact and depends only on masses of final state particles, but used directly is inefficient for a Monte Carlo algorithm if sharp peaks are present due to resonances in the intermediate states. Therefore, it requires changes of variables, which improve the program efficiency while leaving intact the actual density of the phase space[2]. Phase space, dLips is calculated independently of the matrix element Eq. 6.4. Matrix element (Eq. 6.4) consists of weak and hadronic currents[3]. The Matrix element for τ lepton decaying into ν_τ and hadrons is given by:

$$\mathcal{M} = \frac{G}{\sqrt{2}} \bar{u}(N) \gamma^\mu (v + a\gamma_5) u(P) \cdot J_\mu. \tag{6.4}$$

From Eq. 6.4 we can calculate the formula for matrix element squared and contracted with density matrix:

$$|\mathcal{M}|^2 = G^2 \frac{v^2 + a^2}{2} (\omega + \hat{\omega} + (H_\mu + \hat{H}_\mu) s^\mu) \tag{6.5}$$

[2] These changes could in principle be avoided for the simulations of the physical case, but are indispensable for some tests, see Sect. 4.2.

[3] Recall Section 5.1

where:

$$\begin{aligned}
\omega &= P^{\mu}(\Pi_{\mu} - \gamma_{va}\Pi^{5}_{\mu}),\\
H_{\mu} &= \frac{1}{M}(M^2\delta^{\nu}_{\mu} - P_{\mu}P^{\nu})(\Pi^{5}_{\nu} - \gamma_{va}\Pi_{\nu}),\\
\hat{\omega} &= 2\frac{v^2 - a^2}{v^2 + a^2}m_{\nu}M(J^* \cdot J),\\
\hat{H}^{\mu} &= -2\frac{v^2 - a^2}{v^2 + a^2}m_{\nu}\ \mathrm{Im}\ \varepsilon^{\mu\nu\rho\sigma}J^{*}_{\nu}J_{\rho}P_{\sigma}\\
\Pi_{\mu} &= 2[(J^* \cdot N)J_{\mu} + (J \cdot N)J^{*}_{\mu} - (J^* \cdot J)N_{\mu}],\\
\Pi^{5\mu} &= 2\ \mathrm{Im}\ \varepsilon^{\mu\nu\rho\sigma}J^{*}_{\nu}J_{\rho}N_{\sigma},\\
\gamma_{va} &= -\frac{2va}{v^2 + a^2}
\end{aligned} \tag{6.6}$$

Note: γ_{va}=1 in Standard Model.

For massles ν_{τ} formula 6.5 can be simplified to:

$$|\mathcal{M}|^2 = G^2\frac{v^2 + a^2}{2}(\omega + H_{\mu}s^{\mu}). \tag{6.7}$$

Looking at Eq. 6.4 one can spot that only hadronic current J_{μ} is model dependent and contains all the QCD interactions involved in hadronic part of the decay. It is also main source of systematical error because typical precision of the models is of order of $1/N_c \approx 30\%$ or $1/N_c^2 \approx 10\%$. Experimental data can reach a precision level better than 0.1% in many of the cases.

Equation (6.8) defines[4] hadronic current as used for all three-scalars decay channels of TAUOLA:

$$\begin{aligned}
J^{\mu} \quad = N \quad & \big\{T^{\mu}_{\nu}\big[c_1(p_2 - p_3)^{\nu}F_1 + c_2(p_3 - p_1)^{\nu}F_2 + c_3(p_1 - p_2)^{\nu}F_3\big]\\
& + c_4q^{\mu}F_4 - \frac{i}{4\pi^2F^2}c_5\varepsilon^{\mu}_{\cdot\ \nu\rho\sigma}p_1^{\nu}p_2^{\rho}p_3^{\sigma}F_5\big\},
\end{aligned} \tag{6.8}$$

where $T_{\mu\nu} = g_{\mu\nu} - Q_{\mu}Q_{\nu}/Q^2$ denotes the transverse projector, and $Q^{\mu} = (p_1 + p_2 + p_3)^{\mu}$ is the momentum of the hadronic system, while p_i (i=1,2,3), denotes the four momenta of i-th pion. The same ordering is used for masses (m_i). The $\varepsilon^{\mu}_{\cdot\ \nu\rho\sigma}$ is the Levi-Civita tensor. In equations of this and the following sections we use notation: $s_i = (p_j + p_k)^2$ where i $\neq$ j $\neq$ k $\neq$ i. Constants: c_i (i=1,2,...,5) are Clebsch-Gordan coefficients, defined specifically for particular hadronic current used. Specific form

[4]Five form factors are used instead of four imposed by Lorenz invariance for practical purpose. In principle F_3 can be represented as linear combination of contributions to F_1 and F_2, therefore it is not employed by some models.

factors in (6.8) are describing different types of intermediate states[5]. Form factors 1, 2, 3 are related to pseudo-scalars e.g. a_1. F_4 describes scalar current e.g. $\pi(1300)$ and F_5 the Wess-Zumino term [59] is dedicated to vector intermediate states e.q ω.

Eqs. (6.9), (6.10) and (6.11) describe Breit-Wigner functions that are later used in definition of form factors:

$$P(S,m_1,m_2) \;=\; \frac{\sqrt{(S-(m_1+m_2)^2)(S-(m_1-m_2)^2)}}{\sqrt{S}}, \tag{6.9}$$

$$\Gamma_{L-wave}(S,M,\Gamma,m_1,m_2,L) \;=\; \Gamma\frac{M}{\sqrt{S}}\left(\frac{P(S,m_1,m_2)}{P(M^2,m_1,m_2)}\right)^{2L+1}, \tag{6.10}$$

$$BW(S,M,\Gamma,m_1,m_2,L) \;=\; \frac{M^2}{S-M^2-iM\Gamma_{L-wave}(S,M,\Gamma,m_1,m_2,L)}. \tag{6.11}$$

These are typical building blocks useful for hadronic currents parameterizations, such as of Gounaris-Sakurai parametrization [60] for $\rho \to \pi\pi$. In following sections we will bring mathematical formulas of form factors as coded in TAUOLA for specific models. Exact values of masses and widths used were collected in Ref. [A]. While they are important, their values can be topic of a whole different discussion on reasons why those values were used. In this chapter I want to concentrate on structure of hadronic currents, therefore we will not bring values of masses and widths explicitly. In the following sections I will give technical description of models. For the detailed motivation I will rely on references.

6.2 Essentials of Resonance Dominance Model

Resonance Dominance Model[6] (RDM) in its basic form was described in [34] and later used by BaBar collaboration. It assumes cascade process $\tau \to a_1\nu_\tau \to \rho/\rho'\pi\nu_\tau \to \pi\pi\pi\nu_\tau$ with, Q^2 dependent a_1 width parametrized in Eq. 6.12, as in Ref. [34]:

$$G(x) = \begin{cases} 4.1(x-9m_{\pi_0}^2)^3[1-3.3(x-9m_{\pi_0}^2)]+5.8(x-9m_{\pi_0}^2)^2 & \text{if } x<(M_\rho+m_\pi)^2, \\ x(1.623+\frac{10.38}{x}-\frac{9.32}{x^2}+\frac{0.65}{x^3}) & \text{if } x\geq(M_\rho+m_\pi)^2. \end{cases} \tag{6.12}$$

Above equation comes from dispersive relations [34] and comparison with ARGUS data [61]. Presence of both a_1 and ρ in this decay is well established experimentally. Despite, the RDM model does not perfectly describe the experimental data, but it does describe it well enough to be useful and it is expected not to introduce shaky physical properties to the decay. Form factors in this model have very basic structure of Breit-Wigner distributions:

[5] In case of chain decay, type of the first particle in the chain dictates which form factor is used.

[6] In the literature it is usually called Kuhn-Santamaria (KS) model, because of the names of ref. [34] authors. Here we use alternative name because it is contains idea behind the model.

$$F_1 = \frac{M_{a_1}^2}{Q^2 - M_{a_1}^2 - iM_{a_1}\frac{G(Q^2)}{G(M_{a_1}^2)}} \cdot \begin{cases} BW(s_1, M_\rho, \Gamma_\rho, m_2, m_3, 1) + \frac{\beta_1}{1+\beta_1} BW(s_1, M_{\rho'}, \Gamma_{\rho'}, m_2, m_3, 1) & \text{if } s_1 > (m_\pi + m_\pi)^2, \\ \frac{M_\rho^2}{M_\rho^2 - s_1} + \frac{\beta_1}{1+\beta_1} \frac{M_{\rho'}^2}{M_{\rho'}^2 - s_1} & \text{if } s_1 \leq (m_\pi + m_\pi)^2, \end{cases} \tag{6.13}$$

$F_2(s_2)$ coincide with $F_1(s_1)$ but has an opposite sign when included in hadronic current. Normalization constant $\beta_1 = -0.145$. All other form factors are equal to zero: $F_3 = 0$; $F_4 = 0$; $F_5 = 0$, because this model does not predict any scalar nor vector component in the decay.

6.3 Essentials of CLEO model

Model developed by CLEO collaboration [21] is an variation on RDM model [34], but with some significant improvements. First of all, the Q^2 dependence of the a_1 resonance width is given by new formula (6.14) obtained from dispersive relations fitted to $\pi^0\pi^0\pi^-$ mass spectra performed by the CLEO collaboration [21]. Complicated form is determined by the a_1 decay channels: as a_1 virtuality gets larger thresholds are crossed, allowing for more decay channels to open, therefore changing the Q^2 dependence of the effective width. In the [21] authors suggest that inclusion on KK^* threshold was crucial for improving of the fit to experimental distribution of a_1 in τ lepton decay into three pions. When a_1 virtuality is above this threshold it can decay into KK^*, therefore invariant mass distribution of three pions becomes more steep, see fig. 9 of [21].

$$\begin{aligned} WGA(Q^2) = & \\ & C_{3\pi} \cdot \begin{cases} 0 & \text{if } Q^2 < t_1, \\ 5.809(Q^2 - t_1)^3[1 - 3.0098(Q^2 - t_1) + 4.5792(Q^2 - t_1)^3)] & \text{if } t_1 < Q^2 < t_2, \\ -13.914 + 27.679Q^2 - 13.393Q^4 + 3.1924Q^6 - 0.10487Q^8 & \text{if } Q^2 > t_2, \end{cases} \\ & + C_{3\pi} \cdot \begin{cases} 0 & \text{if } Q^2 < t_1, \\ 6.2845(Q^2 - t_1)^3[1 - 2.9595(Q^2 - t_1) + 4.3355(Q^2 - t_1)^3] & \text{if } t_1 < Q^2 < t_2, \\ -15.411 + 32.088Q^2 - 17.666Q^4 + 4.9355Q^6 - 0.37498Q^8 & \text{if } Q^2 > t_2, \end{cases} \\ & + C_{K^*} \cdot \begin{cases} \frac{\sqrt{(Q^2 - t_3)(Q^2 - (M_{K^*} - m_K)^2)}}{2Q^2} & \text{if } Q^2 > t_3, \\ 0 & \text{if } Q^2 \leq t_3, \end{cases} \end{aligned} \tag{6.14}$$

where:
$t_1 = (3m_\pi)^2$, $t_2 = (2m_{\pi^0} + m_\pi)^2$, $t_3 = (M_{K^*} + m_K)^2$, $C_{3\pi} = 0.2384^2$ and $C_{K^*} = 4.7621^2 C_{3\pi}$ and Q^2 is given in GeV^2 units.

Other part of form factors in CLEO modeling also changed significantly with respect to KS model. New resonances were included to account for additional decay

chains:
$\tau \to a_1\nu_\tau \to \sigma\pi\nu_\tau \to \pi\pi\pi\nu_\tau$
$\tau \to a_1\nu_\tau \to f_2(1270)\pi\nu_\tau \to \pi\pi\pi\nu_\tau$
$\tau \to a_1\nu_\tau \to f_0(1370)\pi\nu_\tau \to \pi\pi\pi\nu_\tau$.
Form factors used in this model are parametrized in Eq. 6.15 and Eq. 6.16

$$
\begin{aligned}
F_1 \;=\; & \left(\frac{M_{a_1}^2}{Q^2 - M_{a_1}^2 - \frac{iM_{a_1}\Gamma_{a_1}}{1.3281\cdot 0.806}\cdot WGA(Q^2)} + \beta\frac{M_{a_1'}^2}{Q^2 - M_{a_1'}^2 - \frac{iM_{a_1'}\Gamma_{a_1'}}{1.3281\cdot 0.806}\cdot WGA(Q^2)}\right)\cdot \\
& \Big(\beta_1\cdot BW(s_1, M_\rho, \Gamma_\rho, m_2, m_3, 1) + \beta_2\cdot BW(s_1, M_{\rho'}, \Gamma_{\rho'}, m_2, m_3, 1) \\
& - \beta_3\cdot\frac{(s_3 - m_3^2) - (s_1 - m_1^2)}{3}\cdot BW(s_2, M_\rho, \Gamma_\rho, m_3, m_1, 1) \\
& - \beta_4\cdot\frac{(s_3 - m_3^2) - (s_1 - m_1^2)}{3}\cdot BW(s_2, M_{\rho'}, \Gamma_{\rho'}, m_3, m_1, 1) \\
& + \beta_5\cdot\frac{(Q^2 + s_3 - m_2^2)(2m_3^2 + 2m_1^2 - s_3)}{18 s_3}\cdot BW(s_3, M_{f_2}, \Gamma_{f_2}, m_1, m_2, 2) \\
& + \beta_6\cdot\frac{2}{3}\cdot BW(s_3, M_\sigma, \Gamma_\sigma, m_1, m_2, 0) \\
& + \beta_7\cdot\frac{2}{3}\cdot BW(s_3, M_{f_0}, \Gamma_{f_0}, m_1, m_2, 0)\Big). \qquad (6.15)
\end{aligned}
$$

The F_2 has the same functional form as F_1. The only difference is interchange for its arguments indices 1 and 2 in eq. (6.8), and that constant c_2 has opposite sign to c_1. The parameter β was introduced by the CLEO collaboration for studies of a_1' influence. Due to insufficient experimental evidence it was set to 0, but remains in the code as an option for future studies. The F_3 takes the form:

$$
\begin{aligned}
F_3 \;=\; & \left(\frac{M_{a_1}^2}{Q^2 - M_{a_1}^2 - \frac{iM_{a_1}\Gamma_{a_1}}{1.3281\cdot 0.806}\cdot WGA(Q^2)} + \beta\frac{M_{a_1'}^2}{Q^2 - M_{a_1'}^2 - \frac{iM_{a_1'}\Gamma_{a_1'}}{1.3281\cdot 0.806}\cdot WGA(Q^2)}\right)\cdot \\
& \Big(\beta_3\cdot\frac{(s_2 - m_2^2) - (s_3 - m_3^2)}{3}\cdot BW(s_1, M_\rho, \Gamma_\rho, m_2, m_3, 1) \\
& + \beta_3\cdot\frac{(s_3 - m_3^2) - (s_1 - m_1^2)}{3}\cdot BW(s_2, M_\rho, \Gamma_\rho, m_3, m_1, 1) \\
& + \beta_4\cdot\frac{(s_2 - m_2^2) - (s_3 - m_3^2)}{3}\cdot BW(s_1, M_{\rho'}, \Gamma_{\rho'}, m_2, m_3, 1) \\
& + \beta_4\cdot\frac{(s_3 - m_3^2) - (s_1 - m_1^2)}{3}\cdot BW(s_2, M_{\rho'}, \Gamma_{\rho'}, m_3, m_1, 1) \\
& - \beta_5\cdot\frac{(s_1 - m_1^2) - (s_2 - m_2^2)}{2}\cdot BW(s_3, M_{f_2}, \Gamma_{f_2}, m_1, m_2, 2)\Big).
\end{aligned}
\qquad (6.16)
$$

In Eq. 6.15 and Eq. 6.16, complex coefficients β_i (i=1,2,...,7) have values: $\beta_1 = 1$, $\beta_2 = 0.12e^{i0.99/\pi}$, $\beta_3 = 0.37e^{-i0.15/\pi}$, $\beta_4 = 0.87e^{i0.53/\pi}$, $\beta_5 = 0.71e^{i0.56/\pi}$, $\beta_6 = 2.1e^{i0.23/\pi}$, $\beta_7 = 0.77e^{-i0.54/\pi}$.

Similarly to KS model F_4 and F_5 are equal to zero, as no scalar nor vector intermediate state is predicted. It is worth noting that scalar contribution ($F_4 \neq 0$),

namely $\tau \to \pi'(1300)\nu_\tau$ subsequently decaying into three pions was studied [62], but experimental evidence was not sufficient to establish existence of such decay chain.

6.4 Essentials of Resonance Chiral Lagrangian model

Resonance Chiral Lagrangian (RχL) model for $\tau \to \pi\pi\pi\nu_\tau$ has originated in [63], which were inspired by Refs. [35, 64]. The reason behind developing RχL model was to create something theoretically sound, with all desired symmetries, while also improving description of experimental data. RχL is derived from Chiral Perturbation Theory (χPT) and reproduces next-to-leading order chiral behavior with help of effective Lagrangians. In this approach form factors take a form:

$$F_i = F_i^\chi + F_i^R + F_i^{RR}. \tag{6.17}$$

In above equation F_i^χ is a chiral contribution, that means τ lepton decaying directly into final state without intermediate states. F_i^R and F_i^{RR} are respectively single and double resonance contribution. Therefore, we have three types of decay chains included:
$\tau \to \pi\pi\pi\nu_\tau$
$\tau \to X\pi\nu_\tau \to \pi\pi\pi\nu_\tau$
$\tau \to Y\nu_\tau \to X\pi\nu_\tau \to \pi\pi\pi\nu_\tau$
Most recent publicly available version of the model [65] included contribution from a_1 pseudo-vector (Y in above examples of decay chains) and contributions from ρ, ρ', σ (X in decay chains mentioned above). In that model form factors basic (with ρ only) components are parametrized in Eqs. 6.18, 6.19 6.20:

$$F_1^\chi = -\frac{2\sqrt{2}}{2}, \tag{6.18}$$

$$F_1^R = \frac{\sqrt{2}F_V G_V}{3F^2}\Bigg[\frac{3s_1}{s_1 - M_\rho^2 - iM_\rho\Gamma_\rho(s_1, M_\rho, \Gamma_\rho, m_\pi, m_\pi, 0)} - \left(\frac{2G_V}{3F^2} - 1\right)\Bigg(\frac{2Q^2 - 2s_1 - s_3}{s_1 - M_\rho^2 - iM_\rho\Gamma_\rho(s_1, M_\rho, \Gamma_\rho, m_\pi, m_\pi, 0)} + \frac{s_3 - s_1}{s_1 - M_\rho^2 - iM_\rho\Gamma_\rho(s_1, M_\rho, \Gamma_\rho, m_\pi, m_\pi, 0)}\Bigg)\Bigg], \tag{6.19}$$

$$F_1^{RR} = -\frac{4F_V G_V}{3F^2}\frac{q^2}{q^2 - M_{a_1}^2 - iM_{a_1}\Gamma_{a_1}(q^2)}\Bigg[-(\lambda' + \lambda'')\frac{3s_1}{s_1 - M_\rho^2 - iM_\rho\Gamma_\rho(s_1, M_\rho, \Gamma_\rho, m_\pi, m_\pi, 0)} + H\left(\frac{s_1}{q^2}, \frac{m_\pi^2}{q^2}\right)\frac{2Q^2 + s_1 - s_3}{s_1 - M_\rho^2 - iM_\rho\Gamma_\rho(s_1, M_\rho, \Gamma_\rho, m_\pi, m_\pi, 0)} + H\left(\frac{s_2}{q^2}, \frac{m_\pi^2}{q^2}\right)\frac{s_3 - s_1}{s_1 - M_\rho^2 - iM_\rho\Gamma_\rho(s_1, M_\rho, \Gamma_\rho, m_\pi, m_\pi, 0)}\Bigg], \tag{6.20}$$

where

$$H(x,y) = -\lambda_0 y + \lambda' x + \lambda'', \tag{6.21}$$

and

$$4\lambda_0 = \lambda' x + \lambda'', \tag{6.22}$$

$$\lambda'' = -(1 - 2G_V^2/F^2)\lambda', \tag{6.23}$$

$$\lambda' = F^2/(2\sqrt{2}F_A G_V). \tag{6.24}$$

The ρ' contribution is included in similar way to KS model, by replacing in Eq. 6.19 Eq. 6.20:

$$\frac{1}{M_\rho^2 - s^2 - iM_\rho\Gamma_\rho(s^2, M_\rho, \Gamma_\rho, m_\pi, m_\pi, 0)} \rightarrow \frac{1}{1+\beta_{\rho'}}\left(\frac{1}{M_\rho^2 - s^2 - iM_\rho\Gamma_\rho(s^2, M_\rho, \Gamma_\rho, m_\pi, m_\pi, 0)} + \frac{\beta_{\rho'}}{M_{\rho'}^2 - s^2 - iM_{\rho'}\Gamma_{\rho'}(s^2, M_{\rho'}, \Gamma_{\rho'}, m_\pi, m_\pi, 0)}\right). \tag{6.25}$$

Then in Ref. [65] inclusion of σ resonance was introduced by adding new terms[7] into form factors:

$$F_1^R \rightarrow F_1^R + \frac{4F_V G_V}{3F^2}\Big[\alpha_\sigma BW(s_1, M_\sigma, \Gamma_\sigma, m_2, m_3, 0)F_\sigma(q^2, s_1) + \beta_\sigma BW(s_2, M_\sigma, \Gamma_\sigma, m_1, m_3 0)F_\sigma(q^2, s_2)\Big], \tag{6.26}$$

$$F_1^{RR} \rightarrow F_1^{RR} + \frac{4F_V G_V}{3F^2}\frac{q^2}{q^2 - M_{a_1}^2 - iM_{a_1}\Gamma_{a_1}(q^2)}\Big[\gamma_\sigma BW(s_1, M_\sigma, \Gamma_\sigma, m_2, m_3, 0)F_\sigma(q^2, s_1) + \delta_\sigma BW(s_2, M_\sigma, \Gamma_\sigma, m_1, m_3 0)F_\sigma(q^2, s_2)\Big], \tag{6.27}$$

and

$$F_\sigma(x,y) = exp\left(\frac{-\lambda(x, y, m_\pi^2)R_\sigma^2}{8x}\right). \tag{6.28}$$

Most recent variation of this model [66] is not included in any official TAUOLA release, but curious reader can find there inclusion of scalar and tensor contribution to the decay within RχL calculation scheme.

While discussing RχL model we did not include $\Gamma_{a_1}(q^2)$ modeling yet, which as mentioned in previous sections is quite important element of theoretical model

[7] This was because of new experimental data which became available for Ref. [65]

for decays mediated by a_1 resonance. It is described in detail in [36] and I do not want to overextend this already lengthy chapter, nevertheless still missing some important theoretical context. Here I would like to describe its construction at the level of ideas. In opposition to RDM and CLEO model, here strict theoretical considerations are used, based on our knowledge of a_1. The $\Gamma_{a_1}(q^2)$ is constructed as a sum of known, most important contributors to a_1 decay width obtained from dispersive relations. Therefore:

$$\Gamma_{a_1}(q^2) = \Gamma_{a_1\to\pi\pi\pi} + \Gamma_{a_1\to\pi K^+K^-} + \Gamma_{a_1\to\pi K^0K^0}. \tag{6.29}$$

Such constructions greatly complicates mathematical formulation of hadronic current and increases time of calculation, making direct use of the model inefficient for fitting. That was mitigated by calculating table of $\Gamma_{a_1}(q^2)$ values in only some points of phase space and use of polynomial extrapolation [65]. When users run MC simulation 1000 points are generated during initialization for that purpose. This of course increases initialization time compared to other models for given decay.

6.5 Comparison of RDM, CLEO and RχL models

In this section we will compare previously described models. Comparison will be done on three levels starting with decay width of the channel, then all possible combinations of final state pseudoscalars invariant masses will be used. Third method for comparison will be use of Dalitz plots in slices in q^2 that is three-dimensional distributions. We will be comparing MC generated samples of 10M $\tau^- \to \pi^0\pi^0\pi^-\nu_\tau$ decay events. Those results are complementary to three prong channel analysis of Ref. [A].

The most basic property of theoretical model that can be confronted with experiment is predicted partial width of decay channel[8]. In Table 6.1 we collect values obtained from investigated models as well as the one calculated using values from Particle Data Group [2]. One can see drastic difference between older models used by CLEO and BaBar (RDM model) collaborations, and RχL model and experimental value. Fits of RχL parametrization included fitting[9] to the experimental value of decay width [65], while as far as we know previously mentioned collaborations concentrated on recreating shapes of experimentally obtained spectra only.

Use of invariant mass distributions of different combinations of final state particles is usually natural progression from measuring decay widths and as mentioned before, sometimes is considered more important. Not all possible combinations of those are always available, as was the case e.g for early fits of RχL model [67]. Processing the data requires manpower which can be a limiting factor, especially for non mainstream measurements. This is most likely the case with modern analysis of

[8] For MC analysis it can be tuned by simple rescaling of coupling constants, without insight into dynamics. Therefore, it is sometimes ignored.

[9] Note, fit was performed to BaBar data of $\tau^- \to \pi^-\pi^-\pi^+\nu_\tau$ decays and prediction for one prong channel was obtained from isospin rotation.

CLEO	RDM	RχL	experimental
0.15245	0.14689	0.21208	0.20991+/-0.02185

Table 6.1: Partial decay width of $\tau^- \to \pi^0\pi^0\pi^-\nu_\tau$ obtained form different models and from experimental measurements. Results given in GeV/c^2 units.

the τ lepton decays to three pions. Most advanced analysis (in terms dimensionality of distributions used) comes from CLEO collaboration [21] even though later experiments collected a lot more data. Fortunately obtaining distributions from MC is less laborious, therefore Fig. 6.1 collects distributions obtained from CLEO, RDM and RχL models as implemented in TAUOLA MC.

Comparison of spectra reveals clear differences between models, most prominently in the $\pi^0\pi^0$ mass distribution, where peak for RDM model is not as high as the other two. This should be attributed to lack of σ resonance in RDM model. Second big difference is a_1 modeling, displayed in the $\pi^0\pi^0\pi^-$ mass spectra, which is significantly different for RχL model. Here the difference is rooted in RχL being derived strictly from theoretical prediction of (6.29), while a_1 modeling in CLEO and RDM model is obtained from fits of polynomials to the peak shape (see Eqs. 6.12, 6.14). Apart from that differences does not seem outrageously large.

Another step of this analysis is the use of three-dimensional distributions - Dalitz plot slices in q^2. As mentioned before this type of plots was used by CLEO collaboration [21] for fits of their model. Figures 6.2, 6.3, 6.4 contain such distributions obtained from CLEO, RDM and RχL models. While such plots display some differences, it is hard to fully grasp how big those differences are. This is much clearer using ratios of two models results, Figures 6.5, 6.6, 6.7. Already now we can see that three-dimensional distributions reveal a lot more features of models, especially interferences between intermediate states (we will touch on that later). One could ask how big are those differences in measurable sense. A test was performed to count how many bins in dalitz plots differ by more than 50% and 100% relative to other models[10]. Number of events in those bins was also counted. The Table 6.2 collects such analysis for Fig. 6.2 and Fig. 6.4, ratio of which is plotted in Fig. 6.5. Similar comparisons for other two combinations of models are stored in Table 6.3, and Table 6.4. From those test we can conclude that difference between models are substantial and cannot be downplayed as something occurring only in tails of the distributions which usually have no statistical impact. CLEO and RDM models are the most different, almost half of the bins exceed 50% difference and more than 10% of events fall into those bins. It is not as drastic when we compare CLEO with RχL and RχL with RDM models, but still more than 20% of bins display such difference and more that 3% of events is in those bins.

Bigger differences between models visible on three-dimensional than on one-

[10] Here we include both positive and negative difference, so bins with ratios of above 1.5 and below 0.66 are included in the statistics for difference exceeding 50%.

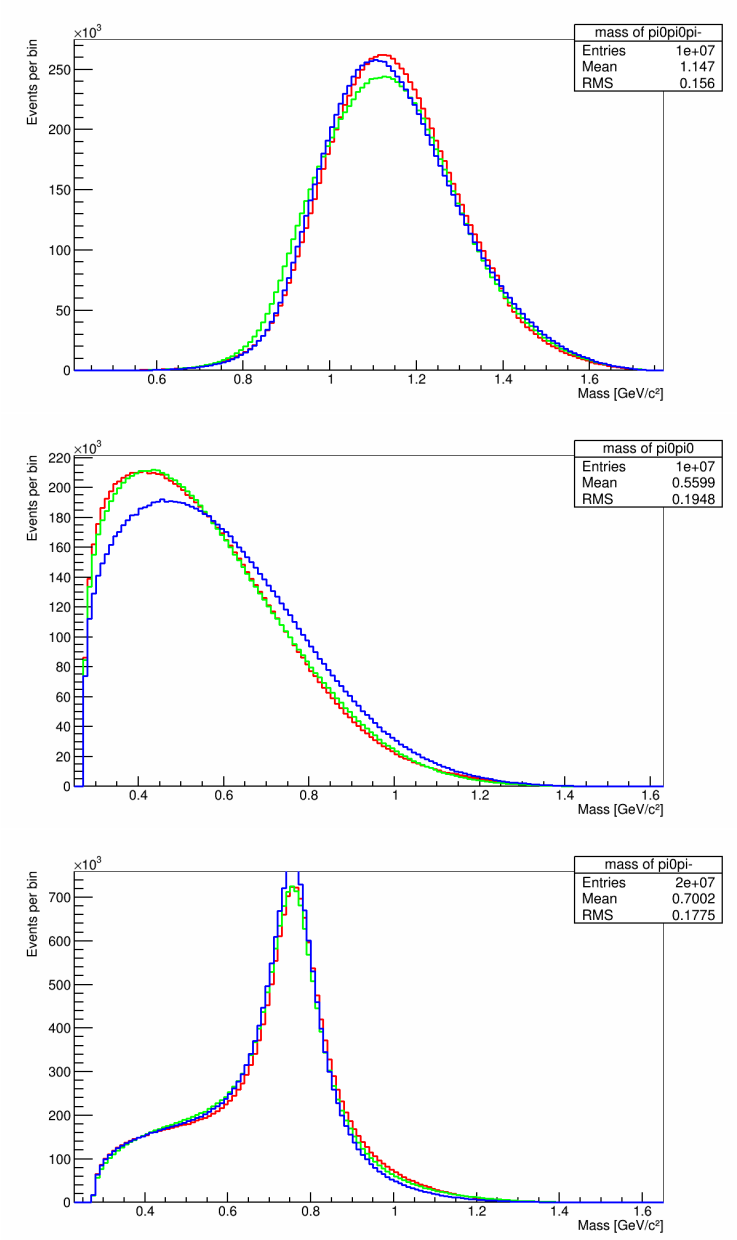

Figure 6.1: Invariant mass distributions obtained from CLEO (red), RχL (green) and RDM (blue) models. Number of events in $\pi^0\pi^-$ distribution is doubled because both of possible $\pi^0\pi^-$ combinations are used.

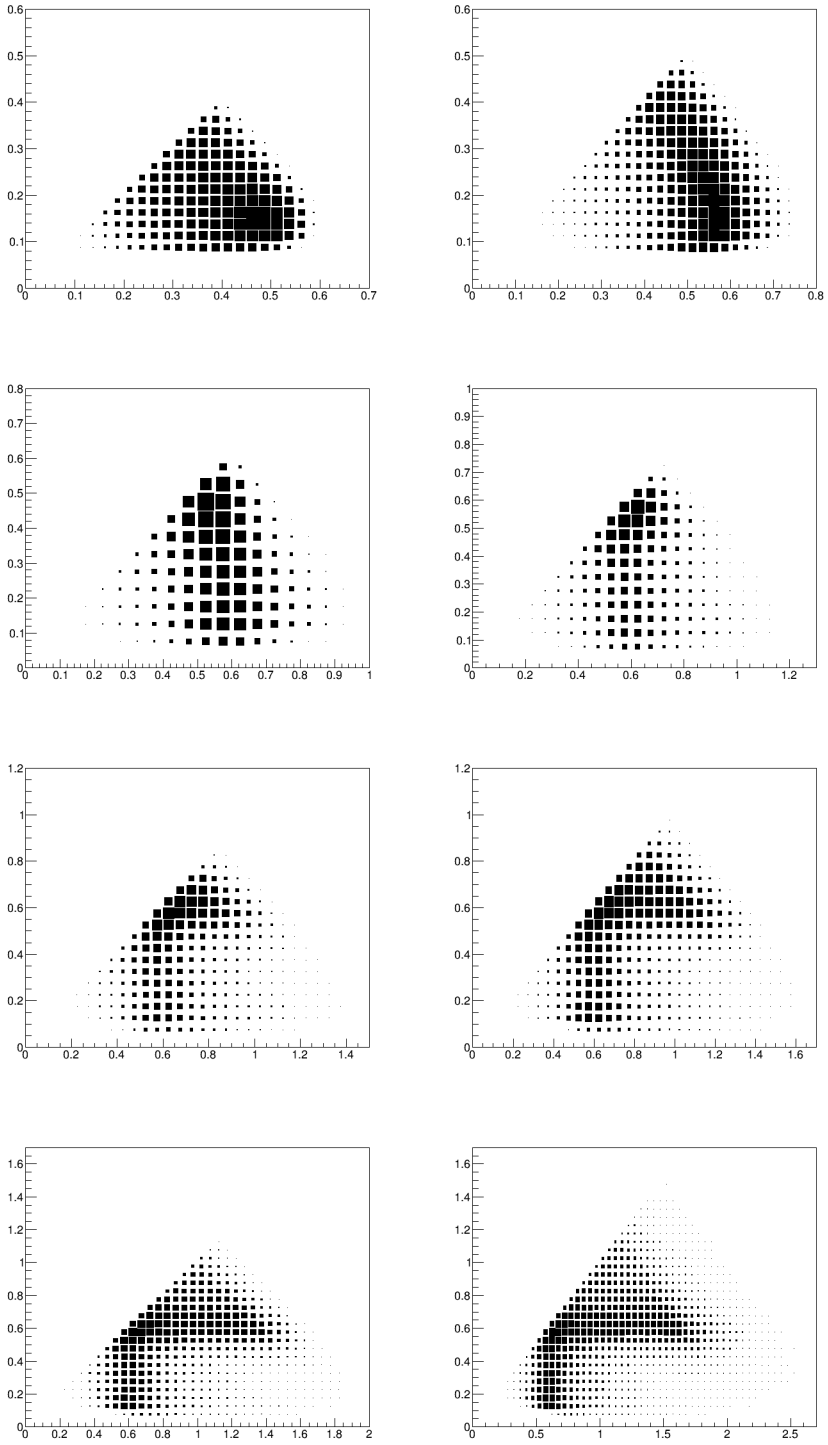

Figure 6.2: 8 Dalitz plots for slices in Q2: 0.36- 0.81, 0.81-1.0, 1.0-1.21, 1.21-1.44, 1.44-1.69, 1.69- 1.96, 1.96-2.25, 2.25-3.24 GeV^2. Each Dalitz plot is distribution for RDM model in s1, s2 variables (GeV^2 units). s1 is taken to be the highest of the two possible values of $M^2_{\pi^0\pi^-}$ in each event.

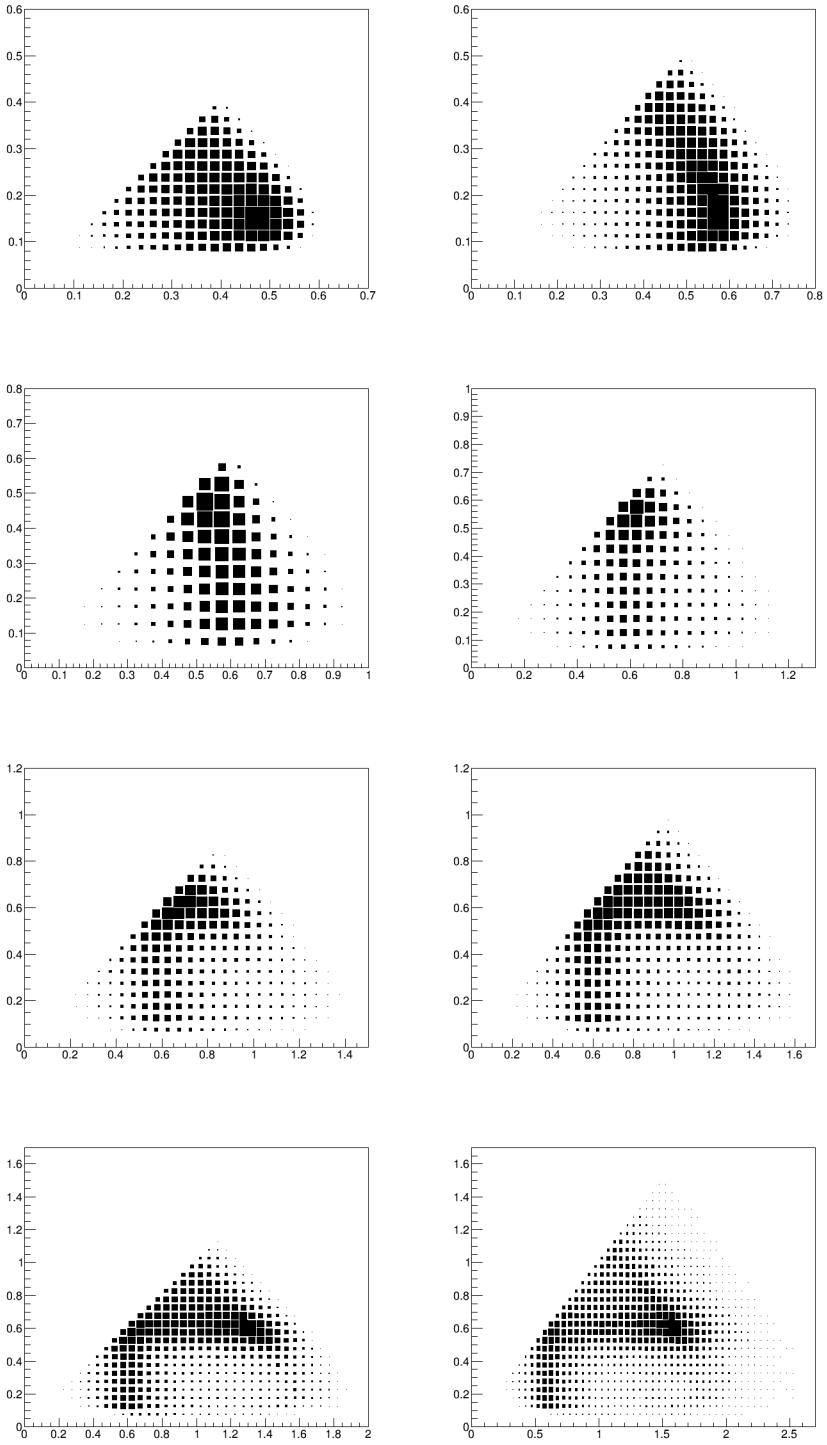

Figure 6.3: 8 Dalitz plots for slices in Q2: 0.36- 0.81, 0.81-1.0, 1.0-1.21, 1.21-1.44, 1.44-1.69, 1.69- 1.96, 1.96-2.25, 2.25-3.24 GeV^2. Each Dalitz plot is distribution for RχL model in s1, s2 variables (GeV^2 units). s1 is taken to be the highest of the two possible values of $M^2_{\pi^0\pi^-}$ in each event.

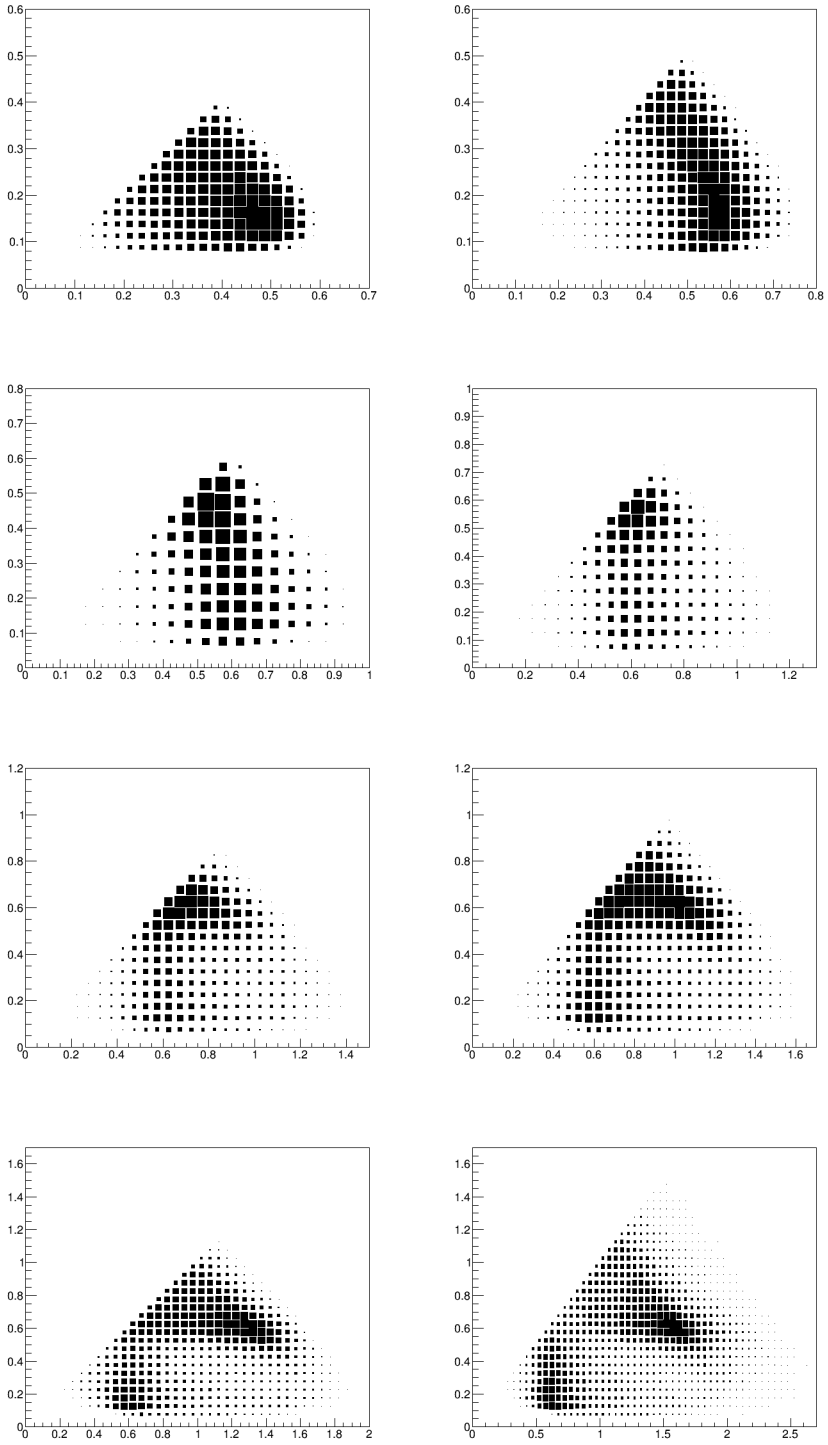

Figure 6.4: 8 Dalitz plots for slices in Q2: 0.36- 0.81, 0.81-1.0, 1.0-1.21, 1.21-1.44, 1.44-1.69, 1.69- 1.96, 1.96-2.25, 2.25-3.24 GeV^2. Each Dalitz plot is distribution for CLEO model in s1, s2 variables (GeV^2 units). s1 is taken to be the highest of the two possible values of $M^2_{\pi^0\pi^-}$ in each event.

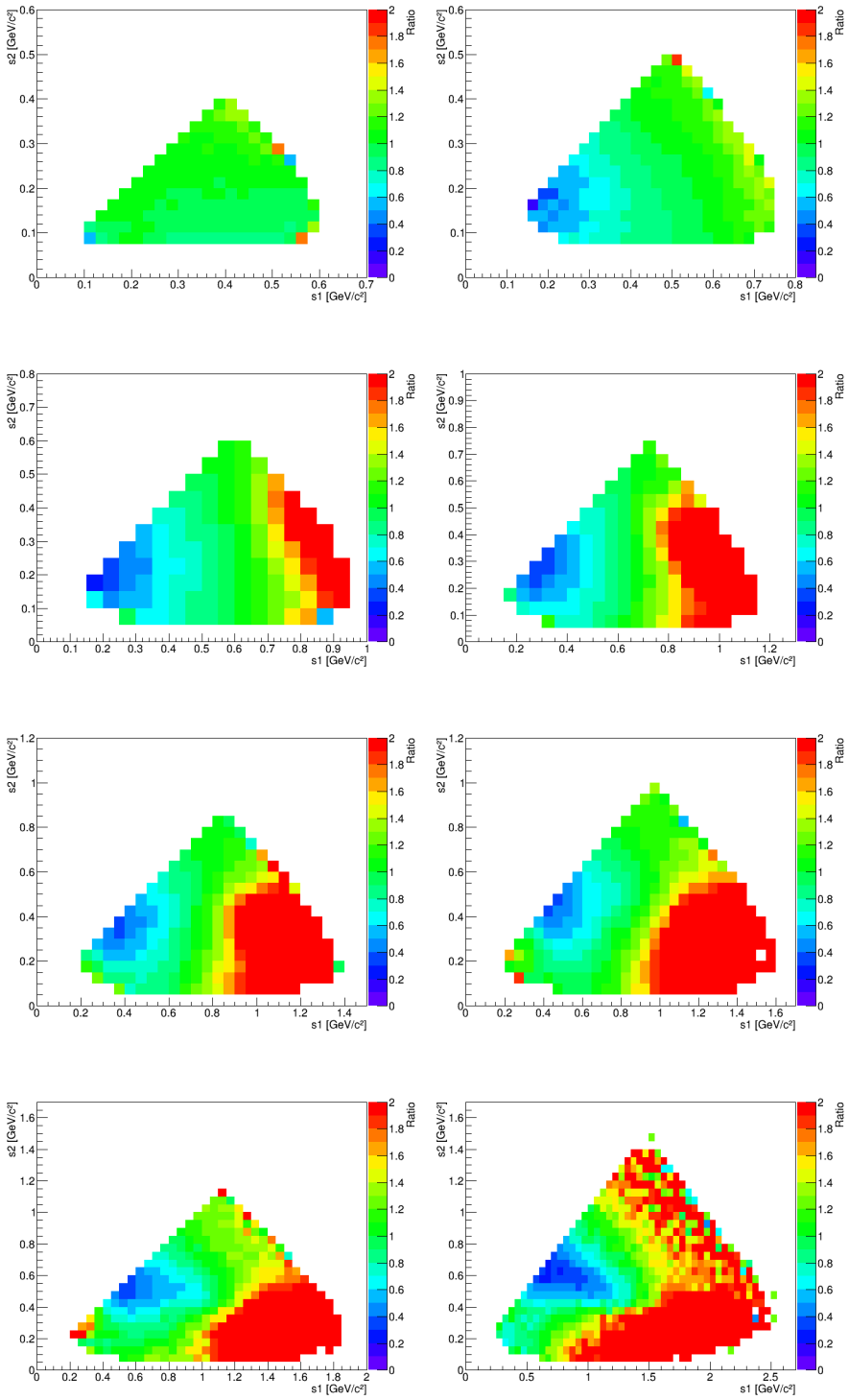

Figure 6.5: Ratio of distributions obtained from CLEO model (Fig. 6.4) to the one from RDM model (Fig. 6.2).

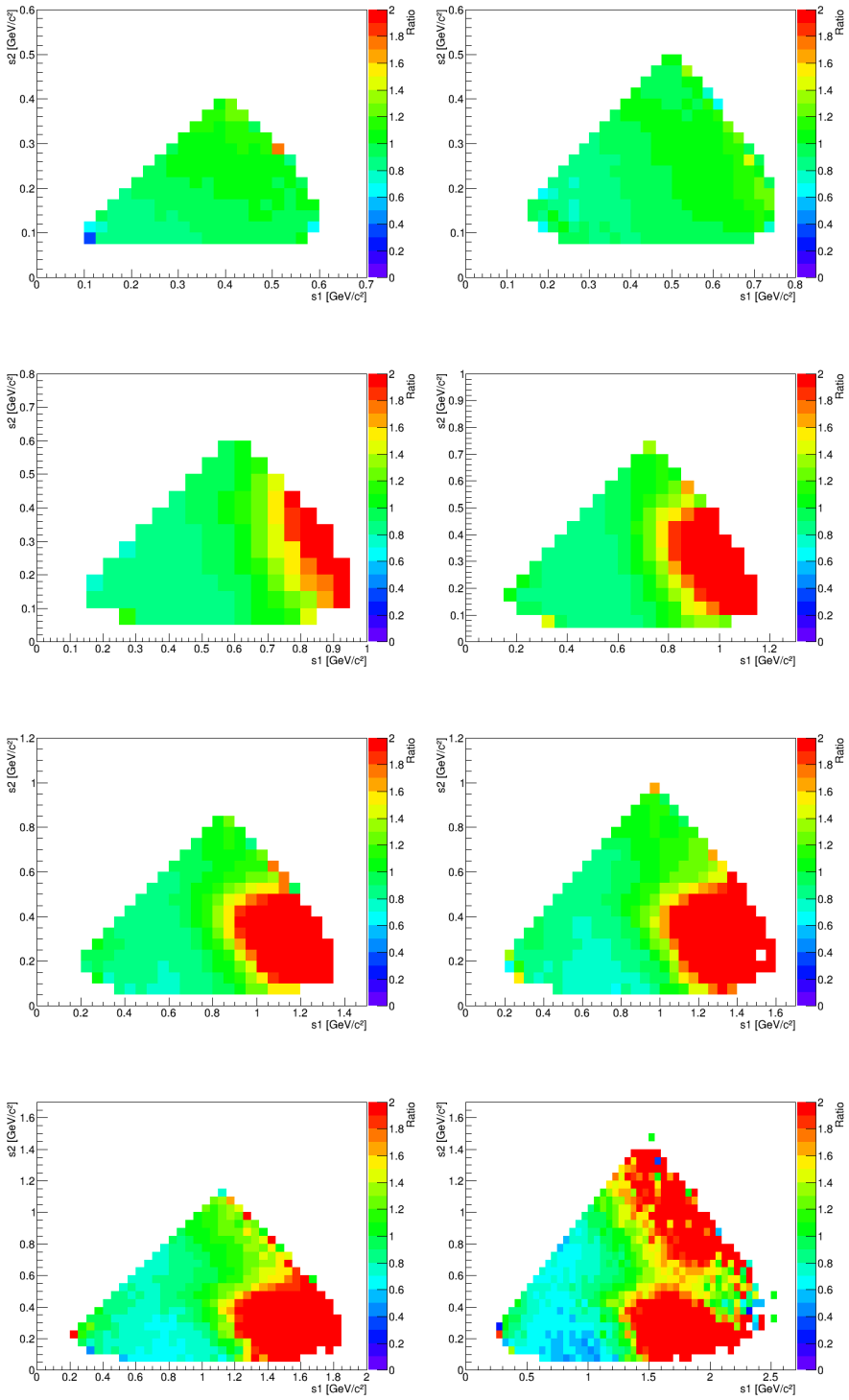

Figure 6.6: Ratio of distributions obtained from RχL model (Fig. 6.3) to the one from RDM model (Fig. 6.2).

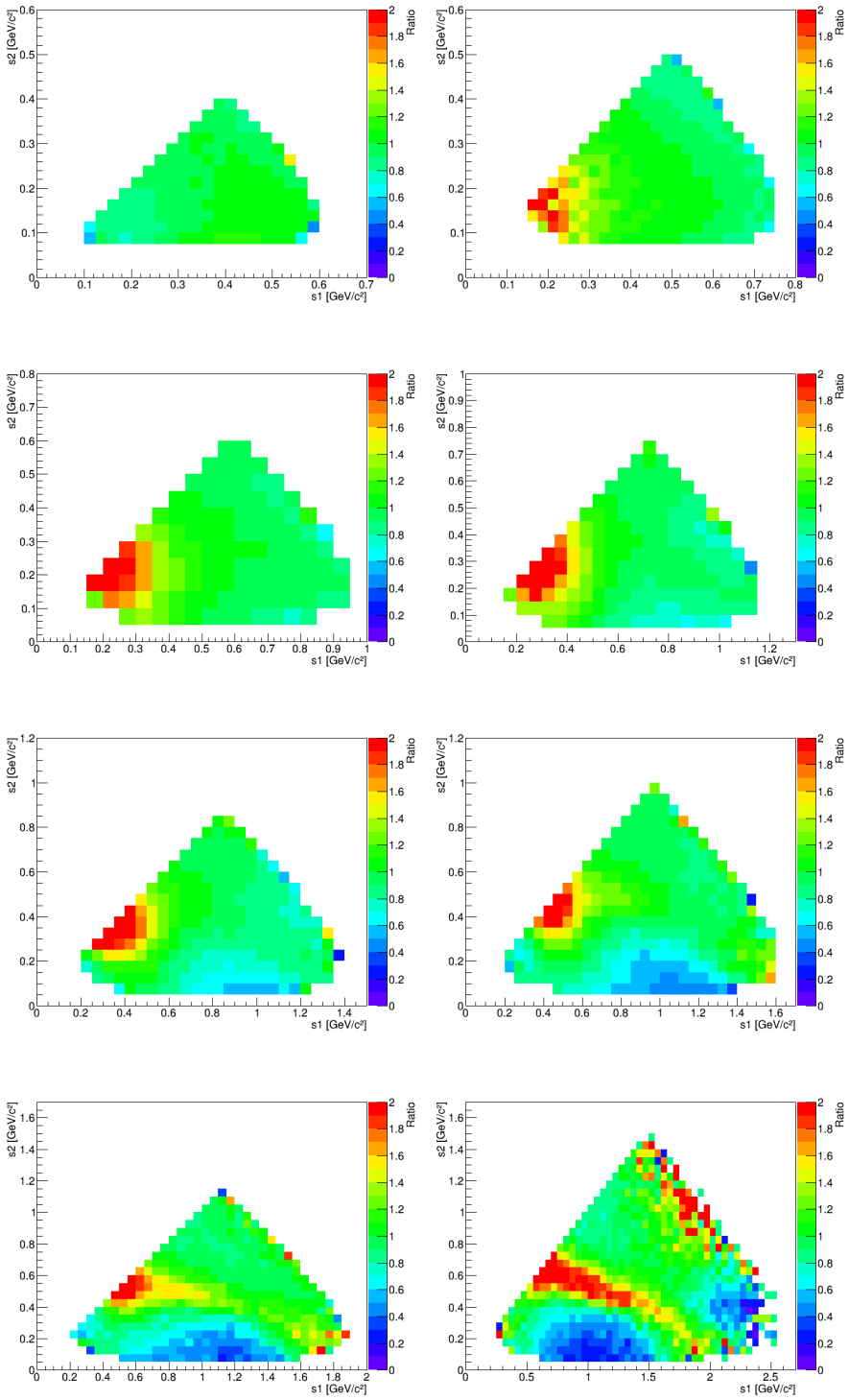

Figure 6.7: Ratio of distributions obtained from RχL model (Fig. 6.3) to the one from CLEO model (Fig. 6.4).

total number of non empty bins for both histograms: 2432
Analysis relative to histogram Fig. 6.4 number of bins with difference over 50%: 1094 number of events in such bins :1345483 number of bins with difference over 100%: 649 number of events in such bins:510036 fraction of events in first and second case: 0.134604, 0.0510248 ==========
Analysis relative to histogram Fig. 6.2 number of bins with difference over 50%: 1094 number of events in such bins :1182125 number of bins with difference over 100%: 649 number of events in such bins:329447 fraction of events in first and second case: 0.118271, 0.0329611

Table 6.2: Analysis of histogram bins of high ratio for compared models (Fig. 6.5 contains ratio A to B we consider here also ratio B to A). Bins where difference exceeds 50% and 100% are counted, as well as number of events in those bins and how big fraction of all events does those bins constitute to.

dimensional distribution should come as no surprise. After all obtaining spectra of lower dimensionality means integrating some of the dimensions and therefore loosing some information. In our case it is particularly important for RχL and CLEO models which display interference between vector and scalar intermediate states. On top of that RDM and RχL models were used to describe those integrated distributions, while CLEO model was fitted to three-dimensional data. It seems we cannot expect significant improvement in modeling without modern experimental analysis in three- (or more) dimensional space. That being said, we can investigate existing models in terms of visible interference depending on type of distribution used. Here we should recall results of Table 5 of [A], which show that CLEO model after integration still allows for decent estimation of impact of it's components. Interference terms contribute 20.5% to total width of the $\tau^- \to \pi^0\pi^0\pi^-\nu_\tau$ decay, while for RχL it is only 4.7%. Same analysis shows that it is respectively 22.4% and 9.5% for three-dimensional distributions. Therefore, use of Dalitz plot slices improves sensitivity in both cases, but also it is much more important for model like RχL. RDM obviously does not have such features because in includes only vector intermediate states.

Comparisons presented in this section did not include experimental data, therefore should not be treated as definite proof of superiority of specific model. All things considered, the three compared models were made in different times and with different goals in mind and fair comparison would require fit of all of them to same experimental data, optimally three-dimensional one and running the tests again. Until then, this comparison can be our rough estimate of systematical uncertainty when the τ leptons are used for other measurements and some model of the decay needs to be chosen. An example will be discussed in Chapter 8.

total number of non empty bins for both histograms: 2425
Analysis relative to histogram Fig. 6.3 number of bins with difference over 50%: 785 number of events in such bins :614649 number of bins with difference over 100%: 461 number of events in such bins:256047 fraction of events in first and second case: 0.0614842, 0.0256127 ========== Analysis relative to histogram Fig. 6.2 number of bins with difference over 50%: 785 number of events in such bins :318735 number of bins with difference over 100%: 461 number of events in such bins:81371 fraction of events in first and second case: 0.0318861, 0.00814031

Table 6.3: Analysis of histogram bins of high ratio for compared models (Fig. 6.6 contains ratio A to B we consider here also ratio B to A). Bins where difference exceeds 50% and 100% are counted, as well as number of events in those bins and how big fraction of all events does those bins constitute to.

total number of non empty bins for both histograms: 2432
Analysis relative to histogram Fig. 6.3 number of bins with difference over 50%: 526 number of events in such bins :366445 number of bins with difference over 100%: 214 number of events in such bins:71840 fraction of events in first and second case: 0.0366544, 0.00718595 ========== Analysis relative to histogram Fig. 6.4 number of bins with difference over 50%: 526 number of events in such bins :316968 number of bins with difference over 100%: 214 number of events in such bins:62157 fraction of events in first and second case: 0.0317094, 0.00621816

Table 6.4: Analysis of histogram bins of high ratio for compared models (Fig. 6.6 contains ratio A to B we consider here also ratio B to A). Bins where difference exceeds 50% and 100% are counted, as well as number of events in those bins and how big fraction of all events does those bins constitute to.

Chapter 7

TAUOLA development

In this chapter, we will be discussing TAUOLA developments over the past few years. It is worth pointing out that any advancements are always dictated by the needs of the users. In previous chapters some of the motivations behind specific modifications had been laid out. Here, we will connect those motives to specific work done for TAUOLA library. Most important issue is the presence of big differences in the available models and their poor intrinsic precision[1]. Because of that, models need to be constantly updated and adapted to new experimental input. One face of such adaptation are fits of model parameters, other is modification of internal structure[2]. When developing TAUOLA we want to accommodate models improvement and adaptation in an user-friendly way.

7.1 Fitting model parameters

In the Chapter 3 experimental precision was discussed, while Chapter 5 provided prediction on theoretical uncertainties. Later on, in Chapter 6 we have compared different models for the $\tau \to \pi\pi\pi\nu_\tau$ decay. Having that knowledge in mind we can say that fitting theoretical models to experimental data is a must. Therefore, much work was put into technical aspects of different options for fitting [D]. Pros and cons of template morphing approach [69] and the use of analytical functions will be discussed in following subsections.

7.1.1 Template morphing

Template morphing is a standard technique oriented towards models where an analytical approach is impossible or suffers some case specific problems, like lack of unfolded data and efficient way of including backgrounds and cuts. For example, variant of the method was already used in domain of τ physics in [69] for the precision tests of the Standard Model at LEP. With this method theoretical distribution is obtained from MC sample. A pre-generated set of MC events is the template. When change of model parameters occurs weights are calculated for each event. Effectively weight is a relative probability of event with given dynamics occurring,

[1] In this statement I follow the logic of [68]. The theoretical model precision is derived from models assumption, not from comparison with the data. So, if model introduces assumption that is valid up to 30% deviation that is the precision of the model regardless of its perfect agreement with the data.

[2] Like addition of σ resonance to RχL model described in Section 6.4.

that is the ratio: $\frac{|\mathcal{M}_{new}|^2}{|\mathcal{M}_{old}|^2}$ is used. Template weight is not even always required to be from the same model we use for calculating $\mathcal{M}_{old}$, it is enough, that the ratios $\frac{|\mathcal{M}_{new}|^2}{|\mathcal{M}_{old}|^2}$ will be reconstructed properly. That is the case for the spin weight in the case of $\tau^{\pm} \to \pi^{\pm}\pi^0\nu_\tau$ decay models, but in general case $\mathcal{M}_{old}$ and not only $\mathcal{M}_{new}$ must be known. What is important is that we have four-momenta of final state particles and we can calculate matrix element squared for an event. When we apply those weights to events from the template we are morphing the distributions, hence the name of the method.

The use of template morphing [69] for fitting may seem not intuitive at first, because single re-weighting of template gives user only a single point in n-dimensional space of model parameters. From the series of events histograms can be obtained. Such histograms can be then used as a functions to be fitted. The parameters dependence is introduced with accurate weights. Most common method used for fitting by public libraries like MINUIT [70] is a gradient descent [71]. Without going into much details, this method uses linear in respect to model parameter approximation of fitted function. It is obtained by calculating first derivatives over model parameters and using Taylor's expansion, so model parameters are arguments of the function[3]. Minimizing the χ^2 between this function and data points means optimal fit. If prefect fit is possible that means finding zero of the function $y = \chi^2(model\ parameters)$ can be achieved. In physical cases function y resembles parabola floating above x-axis. Therefore, going for values of model parameters predicted from zero point of the y's linear approximation usually will make us go closer to the minimum, but somewhere on the other arm of a parabola. Repeating the procedure we will zig-zag around the minimum, newer reaching it. To avoid this scenario limitations have to be set on maximum change in model parameters per iteration. Usually estimation of second derivative is used for that purpose. If minimum χ^2 lies high above zero zig-zagging will strengthen and limits on step size in parameters space have to be more restrictive.

Direct use of re-wighting algorithm in the fitting library would be highly CPU consuming. Re-weighting of 10M sample may take anywhere between 2 and 7 hours, and fitting with MINUIT can take a few hundreds iterations, sometimes even exceeding thousand. Default methods are too restrictive to provide even rough result fast enough. Therefore we are forced to use nested gradient method. We obtain a linear approximation of function with help of Taylor's theorem numerically. For that at least n+1 points (n is the number of fitted parameters), that means n+1 re-weightings for the sample is required (fortunately those can be parallelized). Constructed function (containing linear terms of Taylor expansion) is then plugged to fitting algorithm of MINUIT through ROOT environment [72]. Values obtained from there can be used for another re-weighting, which produces new points for construction of new linear function in parameters space, which is fitted by MINUIT, and so on. Incorporating second quadratic term estimation can be helpful [D], but

[3] Histograms are linear with respect to filling events and thus their weights.

requires additional n re-weightings for the sample. Nevertheless, we found it beneficial [D] as long as it means using more computing cores and not increasing time needed for single step of calculation. We should also mention that gradient descent convergence is asymptotic, so the biggest drop in χ^2 value happens in early stage of fitting. While nesting gradient as explained before we can control size of the steps in accordance to our knowledge about fitted model. This allows us to optimize for those early stages and opt for only rough estimate of model parameters, which from our experience could take as little as 10-20 iterations [73].

From this description drawbacks of the method should be clear - it is extremely slow (time per iteration), and therefore without incorporating big cluster computers[4] can be used to obtain only rough estimate on parameter values. But, considering systematical uncertainties associated with theory more often than not, rough estimate can be good enough. We should also mention that size of the template should be sufficient to ignore statistical errors of MC sample. The sample is usually 10 times bigger than experimental one, so its statistical error is negligible in comparison to the one of experimental data. This may be detrimental for use of the method with large datasets due to increased CPU consumption.

The upside of this method is its wide range of applications. It can be used to fit data containing background, or with experimental cuts strongly deforming spectra. That being said for unfolded data, if possible, analytical/semi-analytical approach are more convenient for use.

7.1.2 Fitting analytical distributions

Fitting analytical distribution to the experimental data may seem to be the most basic solution for recreating experimental shapes, but sometimes obtaining analytical distributions is not as simple as it may seem. For τ lepton decays theoretical models are coded in a from of hadronic currents as described in Chapter 6. To obtain analytical formula for experimentally available spectra one has to construct hadronic function W_A of [20], which then can be integrated over bins used for experimental histogram. This is three-dimensional integration of convoluted function, so numerical stability needs to be checked at any point. Also stability of numerical calculation of derivatives is essential, otherwise algorithms like MINUIT (using gradients) fail. Some simplifications may be required to save CPU time. One of such simplification is replacing integration of one dimension by using value in the center of the bin. Then, additional normalization is required, but for one-dimensional distributions like invariant mass spectra, it can be done without introducing significant error. Technical details of using this exact approach for RχL model are given in [D].

For three-dimensional distributions like those presented in Chapter 6 integration over all three dimension within each bin cannot be avoided. This significantly increases time of integration by adding on complexity but also by increasing number of bins. The latter can be mitigated, but then we are loosing on resolution.

[4] Even with such computers whole procedure could last a few weeks, therefore have increased risk of failure from software or hardware malfunction.

Additionally instability of integrals can be expected in tails of distributions, where chosen binning covers only small volume of available in the decay phase-space, while being mostly empty. Therefore, distribution obtained from analytical formulation of the problem may be dependent on chosen integration scheme.

All things considered fitting analytical distributions is a "go to" method in most cases, mainly due to superior CPU time efficiency. It is also optimal for numerical stability tests (parameter correlations, multiple similar local minima) before confronting theoretical models with experimental data. Main limitation is the inability to fit distributions with irremovable backgrounds and strong experimental cuts removing big chunk of phase space.

7.1.3 Further reflections on fitting

Brought in previous sections experiences with fitting do not fully explore the topic. Other methods should investigated in the future, but here I would like to concentrate on limitations we experienced, that are not connected to the methods used.

First of all, experimental data available consisted of invariant mass distributions, at first of $\pi^-\pi^+$ and $\pi^-\pi^-\pi^+$ systems, later also $\pi^-\pi^-$ was added. When fitting simultaneously to those distributions each event is used multiple times, introducing strong correlations between histograms and their bins. As a result we were more likely to encounter degeneration of the result and multiple similar local minima. The study [D] actually confirmed that, though other minima had higher χ^2. Fitting three-dimensional distribution similar to those of Chapter 6 would most likely mitigate this issue, because each event is used only once and they monitor more details of the model than its three one-dimensional projections.

Second hindrance to our fits was big number of fitted parameters. Masses and widths of resonances could probably be set to PDG values, as was done in e.g. [21] with CLEO model. Fitting masses and widths is more theoretically sound, but complicates fitting and well established resonances should not deviate from PDG values in significant way.

Third problem was presence of correlations between parameters. We had at least 4 strongly correlated parameters. Resolving such issue is not always straightforward and usually requires fits of the model to a sample generated with its help but with different set of parameters. Discussed later in this chapter modifications to TAUOLA should help with such procedures in the future. Nevertheless, at the time of analysis form [D] this was not viable choice, because it would strongly delay the work.

7.2 New initialization

Recently new initialization for TAUOLA was introduced [B]. It is a much needed response to the needs of modern experiments such as Belle II but also fulfills scientific duty of archiving the progress made by BaBar collaboration. This TAUOLA initialization mirrors the one used by aforementioned collaboration for basic simulations. At the same time historically available options are still included and can

be accessed by internal flags. Such solutions maintains continuity of the code development[5], which is especially important for seasoned users, fluent in FORTRAN, but not necessarily in more modern programing languages. For such users ability to recognize and modify code on the most basic level is very important and complete removal of some options and too drastic changes may be met with a backlash: loss of physics content.

Important aspect of the new initialization is an increased number (up to 200) of available decay channels. This feature was introduced already with [74, 75] and BaBar initialization made a use of this possibility. Not all of those channels were actually used, some of them are placeholders, but it was an important step of development because it simplified addition of new channels and new parameterizations based on theoretical models. Therefore, this initialization introduces some new decay channels, but also improves modeling of some old ones.

Furthermore, some previously merged channels became separated. One example of such is the τ decay into three pions, which at the time generated both one and three-prong decays. User wishing for generation of only one option was forced to modify internal parameters. Now, those are two separate decay channels, therefore can be chosen with program input parameter. Models for the two channels can be studied separately. The partial widths calculated from event samples are obtained and not just the average of the two.

Parallel to introduction of above mentioned features groundwork for future updates has been laid. Whenever possible, the code was split into independent blocks that could be tested and modified separately, see Fig. 4.3. While some parts of the code were already separated, only now structure of the code can be considered optimal for further improvements that will be discussed in Section 7.3. Those changes also allowed for introduction of a framework for user-defined hadronic currents and matrix elements.

7.2.1 User defined currents

Since the TAUOLA is open source any user can always modify the code, but up until last release [B] it required familiarity with FORTRAN source code. Right now, it became possible to code new hadronic current models in any programing language and plug them into TAUOLA through pointers to user-defined functions. With increasing precision and amount of data available more decay channels can be researched and with greater detail. Models predicting e.g. LFV decays, baryonic decays and second class currents are not well established, therefore it is important that many options for those can be easily introduced into MC simulation and tested. The same arrangements can be prepared and be useful for fitting program. In Chapter 5 some of those ideas and their importance were discussed, but even well established decays could benefit from new models, because intrinsic precision of the models is much lower than experimental one for the data used for fits. With the

[5] Note, that TAUOLA project has almost 30 years now.

barrier of required familiarity with the source code lifted, more people can involve themselves into model building.

With the framework for user-defined models we included two examples to help users learn how to use it. First example was quite trivial, but very important as a technical test. We have supplemented C++ functions recreating TAUOLA modeling for $\tau^- \to (\pi\pi)^-\nu_\tau$ decay. With this users got simple template to follow while introducing own models. This example function generates exactly the same events as FORTRAN counterpart used as a basis. Therefore, we have a proof of framework working correctly, showing at the same time that rewriting TAUOLA into C++ can be done channel by channel and easily tested.

Second example can already be used for LFV searches. We have coded higher order operators from LFV model of [76, 53]. Fig. 1 of Ref [B] collects results obtained with this example code. Switching between the options needs to be done by commenting/uncommenting parts of the source code. Such solution may not be elegant, but it is good for learning purposes.

7.3 Future plans

In previous section some historical aspects of TAUOLA were mentioned. One of them is TAUOLA being written in FORTRAN while less and less people are familiar with this programming language. Migration to C++ or other more recent language seems to be unavoidable in modern environment. All the preparations for such translation to new language have been made and hopefully it can be done in a seamless for the users way. It is not set in stone that C++ will be language of choice, but it is most likely scenario. With current block structure of TAUOLA it does not need to be performed at one go and the process can be discussed with users community[6] along the way.

Currently, fitting model parameters for hadronic currents requires a lot of effort. There is no straightforward way of fitting model coded for TAUOLA. The possibility of plugging user-defined currents may enable creation of a tool that could use those currents in the same way for fitting. Previously mentioned possible migration to newer programing language would further facilitate fitting programs. Some effort has been put into this topic but it is still far from finished. Nevertheless, with the help e.g. of ROOT libraries we hope that such tool will be prepared at least for decay channels into three hadrons in the final state. If such a tool will be prepared it can be further improved to allow for fits of distributions obtained with projection operators of [20]. It need to be stressed that all effort put into fitting is necessity due to models being data driven and that is dictated by the precision levels of experimental distributions compared to theoretical predictions.

As time progresses new tools are becoming more prominent. The hot topic right now are the Neural Networks (NN), which development seems to have speed up over

[6] The community of FORTRAN users is quite important and we don't want to force on them too many drastic changes in a short period of time.

the past few years. While at this point NN's don't look too promising[7] as a tool for fitting they are meant for pattern recognition and therefore may be an alternative to projection operators. It is too early to say anything definite, the very idea seems worth exploring. In the next chapter an example of NN usage will be presented for evaluation of systematical errors associated with different models for τ decays when used for Higgs CP parity measurement. This learning experience may be helpful in finding ways of using NNs for other problems.

[7] The talk [77] is somewhat challenging that statement already now.

Chapter 8

Higgs CP state measurement with the help of Machine Learning techniques

In Chapter 5 some areas of physics involving the τ lepton were brought to attention. In this chapter I want to concentrate on those phenomena that can be measured through use of τ's. Electroweak processes involving weak bosons are the typical examples. For years now τ's were used for measuring the properties of Z and W bosons. With the discovery of scalar boson - Higgs, τ's are great tool also for measuring properties of this new particle. This was discussed in Section 5.4. Considerations of optimal measurables [69] available in particular experiment require modeling of the τ's and this introduces systematical error associated with used model. It is hard to asses how much of an impact different models have. In the following sections I will describe briefly what Neural Network is and how it can be used to resolve the issue of systematical errors in Higgs CP measurement [C].

8.1 Introduction to Machine Learning and NN

First of all it should be said that Machine Learning (ML) is very broad term which includes all of many approaches allowing computer to successively improve (learn) performance at certain task, without any additional programing. NN is only one type of ML techniques based on the idea of creating artificial network mirroring that of a brain. Both ideas are relatively old[1] and only recently computers began to have enough computing power to employ ML for scientific tasks. Also simplifications in approaches, without loss of performance were of a great importance for ML practical applications. Speed of the progress increased enormously.

The term NN is also very broad and there are multiple options for creating a NN. The NN techniques develop rapidly and new approaches appear, on a scale of months. It does not seem to be feasible or useful to present snapshot of this development as of today. Things changed a lot since Cracow group started to use the methods in summer 2015. Especially in domain of evaluation of systematic error progress of the last two years was enormous. Software companies seem to

[1] It is hard to pinpoint when the idea pop up, but [78] from 1949 seems one of the earliest examples while 10 years later [79] ML is used on the game of checkers. At that point idea of NN was already established, but due to lack of computing power its use was beyond foreseeable (at the time) future.

play nowadays leading role in this development. That is probably why, CERN organized series of talks devoted to this development let us quote as an example talks: [80, 77]. The practical message is that theoretical predictions, if available with options controlled by weights- matrix element ratios, are useful and even crucial for definition of measures defining distance in the event space.

For the purpose of next section we will use two terms: receiver operating characteristic curve (ROC curve) and area under the curve (AUC) [81]. The ROC [82] can be used to evaluate quality of a classifier (diagnostic test). Classifier is basically an operator that distinguishes between two hypothesis. The AUC is normalized to 1 area under ROC curve and can be interpreted as probability of distinguishing between two measured states using our classifier. Therefore AUC of 0.5 means random classification (we might as well throw a coin to choose the correct answer), the further result is from that value, the better. An AUC of 1 or 0 means perfect classification.

8.2 Assessing systematic errors associated with different models using NN

The classifier (mentioned in previous section) is in our case a tool created with the help of NN, specifically it is a program that recognizes in a dataset Higgs CP characteristic (even vs mixed). NN learning process is based on providing initial algorithm with input of known characteristic. For that purpose 1 million $H \to \tau\tau$ were used, with events subsequent decays $\tau\tau \to a_1^{\pm}\nu_\tau a_1^{\mp}\nu_\tau$ or $\tau\tau \to \rho^{\pm}\nu_\tau a_1^{\mp}\nu_\tau$. First part was generated using Pythia 8.2 [83] with Higgs produced via gluon-gluon fusion, τ decays were simulated using TAUOLA and spin correlations were implemented with TauSpinner [55]. Because experimental measurements cannot always provide most desirable quantities, four types of inputs were were defined:

- the acoplanarity angle (ϕ^*) [54],
- $y = \frac{E_{\pi^\pm} - E_{\pi^0}}{E_{\pi^\pm} + E_{\pi^0}} \left(- \frac{2m_{a_1}^2 - 2m_{a_1}^2 + 2m_{\pi^\pm}^2}{2m_{\rho^0}^2} \right)_{\text{for } \tau \to a_1}$ [54],
- invariant masses of pairs and triplets of final state pions,
- 4-vectors of final state pions.

The goal of the analysis [C] was to check if and how much the classifier is affected by the model used for training depending on available input. The models used for $\tau \to a_1^{\mp}\nu_\tau$ were those presented in Chapter 6 plus one additional variation of CLEO model. Different options for $\tau \to \rho\nu_\tau$ were not considered in the study as they are known to factorize from spin dependence.

While testing how NN responds to training based on different model it is also important to remember that experimental data has specific resolution. Using that knowledge [84, 85], NN was trained also with Gaussian smeared MC samples. It is

only simple estimation of detector effects and further analysis might be required in the future. For now, it should be treated as first order estimation of the experimental effects impact on NN training.

Data collected from above mentioned considerations was collected in a form of AUC values collected in [C]. To summarize those results:

- Smearing the sample used for training has little effect on the AUC score, reducing predictive power of the classifier by about 1%.
- Impact of different models of τ decays used for training than for analyzed sample is bellow 1% in all of the cases, depending on input parameters used it can be even reduced to below 0.1%. This is most important conclusion relying on results of my work on τ hadronic currents.
- Input parameters have the biggest impact on the AUC score. The lowest scores are obtained for usage of ϕ^* and y for input, while best result is obtain when using all of the defined inputs. The difference in the AUC value does not exceed 3%.

While the classifier is quite robust against all considered variations in the training, the AUC score in the best case scenario is 0.604. Note in above consideration we did not use neutrino momentum, because it escapes detection, therefore some information is lost. With possibility to recover neutrino four-momentum the AUC could reach the value of 0.782 [86]. Therefore, any constrains that could be found on neutrino momentum, could greatly improve the measurement of Higg boson CP state using our classifier.

Chapter 9

Summary

The presented thesis has explored vast topic of τ decays in the context of present day experiments like Belle and BaBar and tools used by them. Theoretical background on most interesting aspects has been laid out. Challenges in both modeling and data analysis have been addressed, with special emphasis on available precision level. Based on the needs of experiments, improvements in the MC library TUAOLA were introduced.

New default initialization for TAUOLA was accommodated for progress in the τ decays measurements. That gave users better baseline for further development of theoretical models, while also improved standard modeling for those who use τ decays only as a tool for measurements of other processes like Higgs properties or BSM searches. Historical parametrization were maintained as options available through internal flags.

Most known models for $\tau \to \pi\pi\pi\nu_\tau$ available through TAUOLA library have been compared using data representations of different dimensionality. Interference between vector and scalar intermediate states (when present) has been assessed. Need for model confrontation against three-dimensional experimental data has been voiced. Such need is backed by high precision of experimental data compared to theoretical one. Therefore, models development have to be data driven and fitting to experimental data is a must.

The ability to add user-defined hadronic currents/ matrix elements was also given to the users. While framework for this procedure was prepared with C++ in mind, it gives model-builders ability to program models in a computer language of their choice. At the same time it removes from them the necessity of being familiar with FORTRAN and knowing internal TAUOLA structure in order to introduce new modeling. New decay modes, including the New Physics one, can be added this way. Example functions were included, featuring LFV decays $\tau \to \mu\mu\mu$.

Internal structure of TAUOLA has been modified to provide more flexibility for possible future routes of development. Those were also discussed, notably not only for TAUOLA but for fitting framework as well.

Migration into other - more modern programing language (most likely C++) may be necessary to better facilitate fitting programs. Possible use of NN for data analysis and model development was also discussed.

The topic of systematical errors associated with different models used for simulation of other measurables was brought at the margin of this work. It was addressed in case of Higgs CP state (with τ lepton final states)measurements. ML technique of NN was used to create classifier distinguishing between even and mixed CP. Im-

pact on the classifier was estimated to be insignificant regardless of model, input parameters used for NN's training.

Bibliography

Publications co-authored by Jakub Zaremba

[A] Z. Was and J. Zaremba, Eur. Phys. J. C75 (2015) 566, 1508.06424, [Erratum: Eur. Phys. J.C76,no.8,465(2016)].

[B] M. Chrzaszcz et al., Computer Physics Communications (2018).

[C] E. Barberio et al., Phys. Rev. D96 (2017) 073002, 1706.07983.

[D] T. Przedzinski et al., Computer Science 16(1) (2015).

Remaining publications

[1] M.L. Perl, The Rise of the standard model: Particle physics in the 1960s and 1970s. Proceedings, Conference, Stanford, USA, June 24-27, 1992, pp. 79–100, 1992.

[2] Particle Data Group, C. Patrignani et al., Chin. Phys. C40 (2016) 100001.

[3] S. Jadach et al., Comput. Phys. Commun. 76 (1993) 361.

[4] P.W. Higgs, Phys. Rev. Lett. 13 (1964) 508.

[5] F.L. Bezrukov and M. Shaposhnikov, Phys. Lett. B659 (2008) 703, 0710.3755.

[6] S. Alekhin, A. Djouadi and S. Moch, The top quark and Higgs boson masses and the stability of the electroweak vacuum. Phys. Lett. B 716 (2012) 214, doi:10.1016/j.physletb.2012.08.024 [arXiv:1207.0980 [hep-ph]].

[7] L.M. Lederman and D. Teresi, The God Particle: If the Universe Is the Answer, What Is the Question? (Dell Publishing, 1993).

[8] G. Zweig, An SU(3) model for strong interaction symmetry and its breaking. Version 2, DEVELOPMENTS IN THE QUARK THEORY OF HADRONS. VOL. 1. 1964 - 1978, edited by D. Lichtenberg and S.P. Rosen, pp. 22–101, 1964.

[9] W. Cottingham and D. Greenwood, An Introduction to Nuclear Physics (Cambridge University Press, 2001).

[10] HFAG Collaboration, `www.slac.stanford.edu/xorg/hfag/`.

[11] R.D. Richtmyer, S. Ulam and J. von Neumann, preprint LAMS-551 (1947).

[12] N. Metropolis and S. Ulam, Journal of the American Statistical Association 44 (1949) 335, http://www.tandfonline.com/doi/pdf/10.1080/01621459.1949.10483310, PMID: 18139350.

[13] J.H. Halton, SIAM Review 12 (1970) 1, https://doi.org/10.1137/1012001.

[14] N. Metropolis, 15 (Special Issue, Stanisław Ulam 1909–1984) (1987) 125.

[15] GAMBIT, G.D. Martinez et al., Eur. Phys. J. C77 (2017) 761, 1705.07959.

[16] Z. Was, Nucl. Phys. Proc. Suppl. 98 (2001) 96, arXiv:hep-ph/0011305.

[17] M.L. Perl et al., Phys. Rev. Lett. 35 (1975) 1489, [,193(1975)].

[18] S. Banerjee et al., Phys. Rev. D77 (2008) 054012, 0706.3235.

[19] Y. Kubota et al., Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment 320 (1992) 66 .

[20] J.H. Kühn and E. Mirkes, Zeitschrift für Physik C Particles and Fields 56 (1992) 661.

[21] CLEO Collaboration, D. Asner et al., Phys.Rev. D61 (2000) 012002, hep-ex/9902022.

[22] I.M. Nugent, Precision Measurements of Tau Lepton Decays, PhD thesis, Victoria U., 2009.

[23] Charm-Tau Factory, A.E. Bondar et al., Phys. Atom. Nucl. 76 (2013) 1072, [Yad. Fiz.76,no.9,1132(2013)].

[24] S. Eidelman, Nucl. Part. Phys. Proc. 260 (2015) 238.

[25] G. Huang et al., Chin. Sci. Bull. 62 (2017) 1226.

[26] Virgo, LIGO Scientific, B.P. Abbott et al., (2018), 1805.11581.

[27] J. von Neumann, Various techniques used in connection with random digits, Monte Carlo Method, edited by A. Householder, G. Forsythe and H. Germond, pp. 36–38, National Bureau of Standards Applied Mathematics Series, 12, Washington, D.C.: U.S. Government Printing Office, 1951.

[28] S. Jadach, J.H. Kuhn and Z. Was, Comput. Phys. Commun. 64 (1990) 275.

[29] G. Marsaglia, B. Narasimhan and A. Zaman, Computer Physics Communications 60 (1990) 345, cited By 55.

[30] F. James, Computer Physics Communications 60 (1990) 329, cited By 260.

[31] Q. Ho-Kim and X. Pham, Elementary Particles and Their Interactions: Concepts and Phenomena (Springer, 1998).

[32] R.K. Ellis, W.J. Stirling and B.R. Webber, Camb. Monogr. Part. Phys. Nucl. Phys. Cosmol. 8 (1996) 1.

[33] S. Scherer, Adv. Nucl. Phys. 27 (2003) 277, hep-ph/0210398, [,277(2002)].

[34] J.H. Kühn and A. Santamaria, Zeitschrift für Physik C Particles and Fields 48 (1990) 445.

[35] G. Ecker et al., Nuclear Physics B 321 (1989) 311 .

[36] O. Shekhovtsova et al., Phys. Rev. D86 (2012) 113008, 1203.3955.

[37] ATLAS, G. Aad et al., Phys. Lett. B716 (2012) 1, 1207.7214.

[38] M. Dine and A. Kusenko, Rev. Mod. Phys. 76 (2003) 1, hep-ph/0303065.

[39] M.C. Gonzalez-Garcia and Y. Nir, Rev. Mod. Phys. 75 (2003) 345, hep-ph/0202058.

[40] F.S. Queiroz, Proceedings, 51st Rencontres de Moriond on Electroweak Interactions and Unified Theories: La Thuile, Italy, March 12-19, 2016, pp. 427–436, ARISF, ARISF, 2016, 1605.08788.

[41] I. Aitchison, (2005), arXiv:hep-ph/0505105.

[42] KamLAND, S. Abe et al., Phys. Rev. Lett. 100 (2008) 221803, 0801.4589.

[43] S.A. Thomas, F.B. Abdalla and O. Lahav, Phys. Rev. Lett. 105 (2010) 031301, 0911.5291.

[44] ALPHA, M. Ahmadi et al., Nature 548 (2017) 66.

[45] ALPHA, C. Amole et al., Nature Commun. 4 (2013) 1785.

[46] A. Melfo et al., Phys. Rev. D85 (2012) 055018, arXiv:1108.4416.

[47] M. Schmaltz, Nucl. Phys. Proc. Suppl. 117 (2003) 40, arXiv:hep-ph/0210415.

[48] L. Carpenter, A. Rajaraman and D. Whiteson, (2010), arXiv:1010.1011.

[49] G.L. Kane et al., Phys. Rev. D49 (1994) 6173, hep-ph/9312272.

[50] S. Heinemeyer, International Workshop on Future Linear Collider (LCWS2017) Strasbourg, France, October 23-27, 2017, 2018, 1801.05191.

[51] M. Schmaltz and D. Tucker-Smith, Ann. Rev. Nucl. Part. Sci. 55 (2005) 229, hep-ph/0502182.

[52] A.J. Buras et al., JHEP 09 (2010) 104, 1006.5356.

[53] B. Dassinger et al., JHEP 0710 (2007) 039, arXiv:0707.0988.

[54] G.R. Bower et al., Phys. Lett. B543 (2002) 227, hep-ph/0204292.

[55] T. Przedzinski, E. Richter-Was and Z. Was, Eur. Phys. J. C74 (2014) 3177, 1406.1647.

[56] ATLAS, G. Aad et al., JHEP 09 (2012) 070, 1206.5971.

[57] ATLAS, G. Aad et al., JHEP 04 (2015) 117, 1501.04943.

[58] ATLAS, G. Aad et al., JHEP 03 (2015) 088, 1412.6663.

[59] R. Decker and E. Mirkes, Phys. Rev. D47 (1993) 4012, hep-ph/9301203.

[60] G.J. Gounaris and J.J. Sakurai, Phys. Rev. Lett. 21 (1968) 244.

[61] ARGUS, H. Albrecht et al., Z. Phys. C33 (1986) 7.

[62] CLEO Collaboration, E.I. Shibata, eConf C0209101 (2002) TU05, hep-ex/0210039.

[63] D. Gomez Dumm, A. Pich and J. Portoles, Phys. Rev. D69 (2004) 073002, hep-ph/0312183.

[64] G. Ecker et al., Physics Letters B 223 (1989) 425.

[65] I.M. Nugent et al., Phys. Rev. D88 (2013) 093012, 1310.1053.

[66] J.J. Sanz-Cillero and O. Shekhovtsova, JHEP 12 (2017) 080, 1707.01137.

[67] O. Shekhovtsova et al., AIP Conf. Proc. 1492 (2012) 62, 1208.5420.

[68] K. Popper, Conjectures and refutations : the growth of scientific knowledge (Routledge, London New York, 1989).

[69] SLD Electroweak Group, DELPHI, ALEPH, SLD, SLD Heavy Flavour Group, OPAL, LEP Electroweak Working Group, L3, S. Schael et al., Phys. Rept. 427 (2006) 257, hep-ex/0509008.

[70] F. James and M. Roos, Comput. Phys. Commun. 10 (1975) 343.

[71] R. Fletcher, The Computer Journal 13 (1970) 317.

[72] R. Brun et al., ROOT web page, http://root.cern.ch/, 2001.

[73] J. Zaremba, Model parameter fits to BaBar collaboration tau lepton data, Master's thesis, AGH, 2013.

[74] Z. Was, Nucl. Phys. Proc. Suppl. 98 (2001) 96, hep-ph/0011305, [,96(2000)].

[75] Z. Was and P. Golonka, Nucl. Phys. Proc. Suppl. 144 (2005) 88, hep-ph/0411377, [,88(2004)].

[76] S. Turczyk, Nucl. Phys. Proc. Suppl. 189 (2009) 140, 0812.3830.

[77] M.N. (Facebook), Poincare embeddings for learning hierarchical representations, 2018.

[78] W.S. Mcculloch, Electrical Engineering 68 (1949) 492.

[79] A.L. Samuel, IBM Journal of Research and Development 3 (1959) 210.

[80] D.J.R.D. Mind), Approximate inference and deep generative models, 2018.

[81] S.A.G. Arnaud and L. Dominique, Roc analysis.

[82] C.E. Metz, Seminars in Nuclear Medicine 8 (1978) 283 .

[83] T. Sjostrand et al., Comput. Phys. Commun. 191 (2015) 159, 1410.3012.

[84] ATLAS, G. Aad et al., Eur. Phys. J. C70 (2010) 787, 1004.5293.

[85] ATLAS, G. Aad et al., Eur. Phys. J. C76 (2016) 295, 1512.05955.

[86] R. Jozefowicz, E. Richter-Was and Z. Was, Phys. Rev. D94 (2016) 093001, 1608.02609.