THE TITLE3

Von der Fakultät für Mathematik, Informatik und Naturwissenschaften der RWTH Aachen University zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften genehmigte Dissertation

vorgelegt von

THE AUTHOR
PREVIOUS QUALIFICATION

aus HOMETOWN

Berichter: Universitätsprofessor Dr. FOO BAR
Universitätsprofessor Dr. BAZ TEST

Datum der mündlichen Prüfung: XX. Month 2018

Diese Dissertation ist auf den Internetseiten
der Hochschulbibliothek online verfügbar.

THE AUTHOR
THE TITLE3

Dissertation in Physik
Rheinisch-Westfälische Technische Hochschule Aachen
III. Physikalisches Institut A

Berichter: Universitätsprofessor Dr. FOO BAR
Universitätsprofessor Dr. BAZ TEST

Contents

1 Experimental Setup

In this chapter the experimental context of this thesis will be discussed.

1.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is currently the most powerful particle accelerator in the world. Hosted at CERN in Geneva at the Swiss-French border and first put into operation on 10$^{\text{th}}$ September 2008, the LHC is designed for proton and heavy lead ion collisions. The machine has gone through several upgrades between the consecutive data-taking phases (Runs) called Long Shutdowns (LS). During these the proton beam energy has been gradually increased from 3.5 TeV to a recently – on the 5$^{\text{th}}$ July, 2022 to be precise – achieved energy of 6.8 TeV [1] resulting in a total centre-of-mass (CM) proton-proton collision energy of $\sqrt{s} = 13.6\,\text{TeV}$. Similarly, the beam intensity has seen an increase from 1.1×10^{11} protons per bunch (ppb) and ~ 200 bunches to a projected $\sim 1.8 \times 10^{11}$ ppb and ~ 2500 bunches [2, 3]. With a theoretical maximum CM energy of $\sqrt{s} = 14\,\text{TeV}$ and integrated luminosity of $L = 10 \times 10^{34}\,\text{cm}^{-2}\text{s}^{-1}$ it holds the record in these measures among concurring experiments. In order to achieve such luminosities, the beams are kept on a circular trajectory using superconducting NbTi magnets operating at 1.9 K thanks to the superfluid helium bath at about 0.13 MPa [4].
As a result of consecutive accelerator upgrades, today's LHC has an impressive and ever-growingly complex pre-accelerator structure as shown in fig. 1.1; consequently, the proton bunches first go through multiple preparation steps before they get injected into the 27 km tunnel of the LHC where the four main experiments (ALICE, ATLAS, LHCb and CMS) and their interaction points are located.

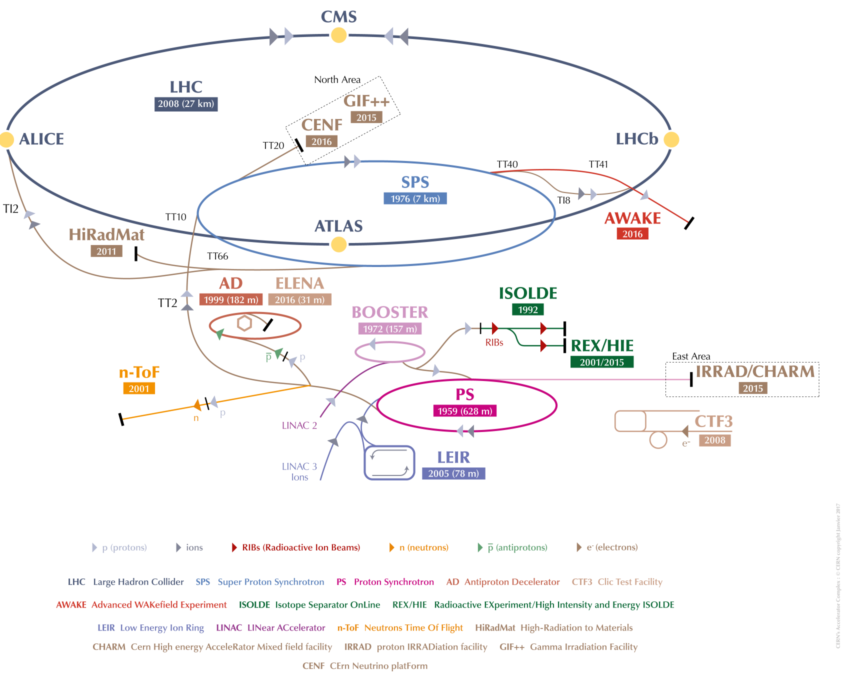

Figure 1.1: The (pre-) accelerator structure of the LHC [5]

1.2 The Compact Muon Solenoid

One of general purpose detectors at the Large Hadron Collider at CERN is the Compact Muon Solenoid (CMS).

2 Deep Learning, Normalizing Flows and Invertible Neural Networks

Deep Learning is a subdomain within Machine Learning focussing on the construction and training of deep neural networks. Such networks have been proven to be extremely successful in different, analytically unsolvable problems such as (but not restricted to) image classification tasks, regression tasks, image generation tasks, language translation and reinforcement learning.
In this chapter, a brief introduction to these networks and their properties with a focus on so-called normalizing flows and invertible neural networks is given.

2.1 Foundations of Deep Learning and Neural Networks

2.1.1 The Universal Approximation Theorem

The success of Neural Networks (NNs) lies in their potential of learning (almost) arbitrary models. This property of NNs can be expressed mathematically through the Universal Approximation Theorem which in words states [6]:

> A feed-forward network with linear output and at least one hidden layer with a finite number of nodes can approximate any real continuous function on a given closed and bounded subset to arbitrary precision.

The Universal Approximation Theorem however does not state how the network should be constructed to achieve the desired precision. For this reason, the theorem has been proven for several network architectures; in case of ReLU-activated feed-forward networks the theorem can be written as the following [7]:

Theorem 1 *For any real and continuous function $f : [0,1]^{d_{in}} \to \mathbb{R}^{d_{out}}$ and every $\epsilon > 0$ there is a ReLU-network $\mathcal{N}$ with the same input and output dimension d_{in} and d_{out} and hidden layer width at most w for which*

$$\sup_{x \in [0,1]^{d_{in}}} \|f(x) - f_{\mathcal{N}}(x)\| \leq \epsilon$$

and

$$d_{in} + 1 \leq w_{min}(d_{in}, d_{out}) \leq d_{in} + d_{out}$$

This theorem does not state anything about the exact depth (number of layers) the network needs to have, its speed of convergence and the optimization process it needs to undergo to achieve this arbitrary approximation. On the other hand, it is reassuring to have a mathematical guarantee for convergence for a given feed-forward network structure. For this reason, empiric studies are usually performed to look for a locally optimal solution for a given task.

Bibliography

[1] A. Alici, M. Bomben, I. Dawson, and J. Sonneveld, "The LHC machine and experiments", (2021) . http://cds.cern.ch/record/2773265.

[2] S. Fartoukh, S. Kostoglou, M. Solfaroli Camillocci, G. Arduini, H. Bartosik, C. Bracco, K. Brodzinski, R. Bruce, X. Buffat, M. Calviani, F. Cerutti, I. Efthymiopoulos, B. Goddard, G. Iadarola, N. Karastathis, A. Lechner, E. Metral, N. Mounet, F.-X. Nuiry, P. S. Papadopoulou, Y. Papaphilippou, B. Petersen, T. H. B. Persson, S. Redaelli, G. Rumolo, B. Salvant, G. Sterbini, H. Timko, R. Tomas Garcia, and J. Wenninger, "LHC Configuration and Operational Scenario for Run 3", tech. rep., CERN, Geneva, Nov, 2021. https://cds.cern.ch/record/2790409.

[3] N. Karastathis, M. Barnes, H. Bartosik, K. Brodzinski, X. Buffat, F. Cerutti, S. Fartoukh, B. Goddard, G. Iadarola, S. L. Naour, A. Lechner, J. M. Heredia, A. Mereghetti, E. Metral, D. Missiaen, N. Mounet, F. X. Nuiry, S. Papadopoulou, Y. Papaphilippou, B. Petersen, G. Rumolo, B. Salvant, C. Schwick, M. S. Camillocci, G. Sterbini, H. Timko, R. T. Garcia, J. Uythoven, and J. Wenninger, "LHC Run 3 Configuration Working Group Report", (2019) . http://cds.cern.ch/record/2750302.

[4] O. S. Brüning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock, *LHC Design Report*. CERN Yellow Reports: Monographs. CERN, Geneva, 2004. http://cds.cern.ch/record/782076.

[5] E. Mobs, "The CERN accelerator complex. Complexe des accélérateurs du CERN", (Jul, 2016) . http://cds.cern.ch/record/2197559. General Photo.

[6] M. Erdmann, J. Glombitza, G. Kasieczka, and U. Klemradt, *Deep Learning for Physics Research*. WORLD SCIENTIFIC, 2021. https://www.worldscientific.com/doi/pdf/10.1142/12294. https://www.worldscientific.com/doi/abs/10.1142/12294.

[7] B. Hanin and M. Sellke, "Approximating continuous functions by relu nets of minimal width", 2017. https://arxiv.org/abs/1710.11278.

Selbständigkeitserklärung

Hiermit versichere ich an Eides statt, dass ich diese Arbeit einschließlich evtl. beigefügter Abbildungen, Zeichnungen u.Ä.m. selbstständig angefertigt und keine anderen als die angegebenen Hilfsmittel und Quellen benutzt habe. Alle Stellen, die dem Wortlaut oder dem Sinn nach anderen Werken entnommen sind, habe ich in jedem einzelnen Fall unter genauer Angabe der Quelle deutlich als Entlehnung kenntlich gemacht.

Aachen, den XX. MONTH 2018

THE AUTHOR

Danksagung

Thx.