



LUND UNIVERSITY

Machine Learning for Longitudinal Monitoring and Control of Ultra-Bright Electron Beams

Lundquist, Johan

2026

Document Version:

Publisher's PDF, also known as Version of record

[Link to publication](#)

Citation for published version (APA):

Lundquist, J. (2026). *Machine Learning for Longitudinal Monitoring and Control of Ultra-Bright Electron Beams*. Lund University.

Total number of authors:

1

General rights

Unless other specific re-use rights are stated the following general rights apply:

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Machine Learning for Longitudinal Monitoring and Control of Ultra-Bright Electron Beams

Machine Learning for Longitudinal Monitoring and Control of Ultra-Bright Electron Beams

by Johan Lundquist



LUND
UNIVERSITY

Thesis for the degree of Doctor of Philosophy
Thesis advisors: Dr. Francesca Curbis, Prof. Sverker Werin,
Dr. Erik Mansten
Faculty opponent: Dr. Rasmus Ischebeck

To be presented, with the permission of the Faculty of Science of Lund University, for public criticism at the Department of Physics (Rydbergsalen) on Sunday, the 16th of October 2026 at 13:15.

Organization LUND UNIVERSITY Department of Physics Box 118 SE-221 00 LUND Sweden		Document name DOCTORAL DISSERTATION	
Author(s) Johan Lundquist		Date of disputation 2026-10-16	
		Sponsoring organization	
Title and subtitle Machine Learning for Longitudinal Monitoring and Control of Ultra-Bright Electron Beams			
Abstract In this thesis machine learning applications to modern particle accelerators are developed and tested at a number of facilities. The focus is on improving and advancing longitudinal diagnostics, with a major project focusing on non-destructive predictions of beam distributions in longitudinal phase space. Applications are developed using data collected from MAX IV, FERMI, SwissFEL and FLASH. Beyond this, the use of machine learning for optimization of linear accelerators is implemented with focus on complex tasks such as longitudinal phase space and emittance optimization. Further, we advance the use of longitudinal phase space diagnostics in connection with other studies focused on longitudinal beam dynamics. Projects include investigations into the high-order shaping of the longitudinal phase space and the further development of beam compression schemes, including experimental measurements of arc-like bunch compressors and investigations of their inherent advantages.			
Key words Accelerators, Beam Diagnostics, Machine Learning, Beam Dynamics, Artificial Neural Networks, Free Electron Lasers, Supervised Learning, Transverse Deflecting Structures			
Classification system and/or index terms (if any)			
Supplementary bibliographical information		Language English	
ISSN and key title		ISBN 978-91-6858-030-2 (print) 978-91-6858-031-9 (pdf)	
Recipient's notes	Number of pages 163	Price	
	Security classification		

I, the undersigned, being the copyright owner of the abstract of the above-mentioned dissertation, hereby grant to all reference sources the permission to publish and disseminate the abstract of the above-mentioned dissertation.

Signature 

Date 2026-09-22

Machine Learning for Longitudinal Monitoring and Control of Ultra-Bright Electron Beams

by Johan Lundquist



LUND
UNIVERSITY

Machine Learning for Longitudinal Monitoring and Control of Ultra-Bright Electron Beams

Paper I © 2022 JACoW Publishing

Paper II © 2023 JACoW Publishing

Paper III © 2026 The Authors

Paper IV © 2026 The Authors

Paper V © 2026 JACoW Publishing

Paper VI © 2026 JACoW Publishing

Paper VII © 2024 JACoW Publishing

Paper VIII © 2026 The Authors

For all other material: © 2026 Johan Lundquist

Faculty of Science, Department of Physics

Cover illustration front: Image generated using a large language model, based on the longitudinal phase space of the MAX IV electron beam.

Cover illustration back: MAX IV Meadow (Photo by Emelie Hilner).

Funding information: Work funded by the Synchrotron Radiation Research Division at Lund University

ISBN: 978-91-6858-030-2 (print)

ISBN: 978-91-6858-031-9 (pdf)

Printed in Sweden by Tryckeriet i E-huset, Lund University, Lund 2026

*You cannot touch the world.
You cannot hold it in your hand
for it is made of breath only.*
— Cormac McCarthy, *The Crossing*

Yet I will try the last.
— Macbeth, Act 5, Scene 8

Contents

List of publications	iii
Acknowledgements	vi
Popular summary in English	viii
Populärvetenskaplig sammanfattning på svenska	x
Machine Learning for Longitudinal Monitoring and Control of Ultra-Bright Electron Beams	1
1 Introduction	3
2 Longitudinal Beam Dynamics and Diagnostics	7
2.1 The Longitudinal Phase Space	8
2.2 RF Acceleration and Energy Chirp	9
2.3 Longitudinal Compression	10
2.3.1 Chicane Bunch Compressors	12
2.3.2 Arc-like Compressors	13
2.4 Longitudinal Diagnostics	14
2.4.1 Transverse Deflecting Structures	15
2.4.2 Passive Wakefield Streaker	17
2.4.3 Bunch Compression Monitors	18
2.5 Longitudinal Dynamics in Free Electron Lasers	19
3 Machine Learning for Accelerators	23
3.1 Machine Learning Fundamentals	24
3.1.1 Artificial Neural Networks	24
3.1.2 Convolutional Neural Networks	27
3.1.3 Learning Schemes and Datasets	28
3.2 Machine Learning Applications in Accelerators	29
3.2.1 Virtual Diagnostics	30
3.2.2 Surrogate Models	32
3.2.3 Advantages and Limitations	33
3.3 Bayesian Optimization	33
4 Development of Virtual Diagnostics	35
4.1 MAX IV Virtual Diagnostic	36

4.1.1	Simulated TDS and Dispersive-Screen Reconstruction . . .	37
4.1.2	First TDS Measurements and Full LPS Surrogate Model .	39
4.1.3	MAX IV Virtual Diagnostic for a Multi-Facility Project .	42
4.2	Virtual Diagnostic for FERMI	44
4.3	Virtual Diagnostic for SwissFEL	46
5	Applications Based on Virtual Diagnostics	49
5.1	Non-Invasive Calibration of a Passive Wakefield Streaker using Machine Learning	50
5.2	Bayesian Optimization of Longitudinal Phase Space at MAX IV	54
5.3	Reconstruction of the FLASH FEL Pulse	56
6	Investigations of the Longitudinal Dynamics of Arc-like Bunch Compressors	59
6.1	First Measurements of the First and Second Order Longitudinal Dispersion of a MAX IV Bunch Compressor	60
6.2	Inherent Stabilizing Properties of the Arc-like Bunch Compressors	63
7	Conclusions and Outlook	67
	Scientific publications	81
	Author contributions	81
	Paper I: Virtual Diagnostic for Longitudinal Phase Space Imaging for the MAX IV SXL Project	85
	Paper II: Virtual Diagnostics for Longitudinal Phase Space Imaging .	91
	Paper III: Multi-Facility Virtual Diagnostic for Longitudinal Phase Space Predictions	97
	Paper IV: Non-Destructive Passive Wakefield Streaker Calibration us- ing Machine Learning	109
	Paper V: Bayesian Optimization of Longitudinal Phase Space in the MAX IV Linac	119
	Paper VI: FEL Power Profile Predictions with Image-Based Machine Learning at FLASH	125
	Paper VII: Measurement of First and Second Order Longitudinal Dis- persion in the MAX IV Bunch Compressor	131
	Paper VIII: Local Bunch Compression Stabilization via Nonlinear Lon- gitudinal Dispersion	135

List of publications

This thesis is based on the following publications, referred to by their Roman numerals:

- I **Virtual Diagnostic for Longitudinal Phase Space Imaging for the MAX IV SXL Project**
J. Lundquist, S. Werin and F. Curbis
Proceedings of 40th Free Electron Laser Conference, 2022, pp. 437-440,
Publisher: JACoW

- II **Virtual Diagnostics for Longitudinal Phase Space Imaging**
J. Lundquist, S. Werin and F. Curbis
Proceedings of 14th International Particle Accelerator Conference, 2023,
pp. 4471-4474, Publisher: JACoW

- III **Multi-Facility Virtual Diagnostic for Longitudinal Phase Space Predictions**
J. Lundquist, J. Björklund Svensson, P. Dijkstal, E. Mansten, G. Penco, S. Werin and F. Curbis
Scientific Reports Vol. 16, 12021 (2026)

- IV **Non-Destructive Passive Wakefield Streaker Calibration using Machine Learning**
J. Lundquist, P. Dijkstal, E. Mansten, S. Werin and F. Curbis
Submitted to *Physical Review Accelerators and Beams*, July 2026

- V **Bayesian Optimization of Longitudinal Phase Space in the MAX IV Linac**
J. Lundquist, J. Björklund Svensson, E. Mansten, S. Werin and F. Curbis
Proceedings of 17th International Particle Accelerator Conference, 2026,
Publisher: JACoW

- VI **FEL Power Profile Predictions with Image-Based Machine Learning at FLASH**
J. Lundquist, G. Goetzke, S. Werin and F. Curbis
Proceedings of 17th International Particle Accelerator Conference, 2026,
Publisher: JACoW

VII **Measurement of First and Second Order Longitudinal Dispersion in the MAX IV Bunch Compressor**

J. Lundquist, J. Björklund Svensson, E. Mansten, S. Werin, R. Swärd, S. Thorin, A. Dixon, P. H. Williams and F. Curbis

Proceedings of 40th Free Electron Laser Conference, 2024, pp. 22-24,
Publisher: JACoW

VIII **Local Bunch Compression Stabilization via Nonlinear Longitudinal Dispersion**

J. Björklund Svensson, **J. Lundquist**, E. Mansten, F. Curbis, S. Thorin and S. Werin

Manuscript in Preparation

All papers are reproduced with permission of their respective publishers.

Publications not included in this thesis:

Measurements of Particle Trajectory Caustics in Bunch Compressors

A. Dixon, A. Wolski, T. K. Charles, P. H. Williams, **J. Lundquist** and S. Thorin

Proceedings of 40th Free Electron Laser Conference, 2024, pp. 132-135, Publisher: JACoW

Single-shot Electron Bunch Profile Monitor based on Coherent Transition Radiation Imaging

A. Guisao Betancur, C. Welsch, J. Wolfenden, O. Grimm, E. Mansten, **J. Lundquist** and S. Thorin

Proceedings of 14th International Beam Instrumentation Conference, 2025, pp. 608-611, Publisher: JACoW

Designing a Femtosecond-resolution Bunch Length Monitor using Coherent Transition Radiation Images

A. Guisao-Betancur, J. Wolfenden, E. Mansten, S. Thorin, **J. Lundquist**, O. Grimm, and C. P. Welsch.

Instruments, vol. 9, 4, (2025)

Experimental Generation of Wakefield Accelerator-Suitable Bunch Current Profiles by Arc-like Bunch Compressors

J. Björklund Svensson, F. Curbis, **J. Lundquist**, E. Mansten, S. Thorin and S. Werin

Proceedings of 7th European Advanced Accelerator Conference, *Journal of Physics: Conference Series*, vol. 3266, (2026)

Using Accelerator Toolbox in a Teaching Environment

S. Werin, J. Björklund Svensson, **J. Lundquist** and F. Curbis

Proceedings of 17th International Particle Accelerator Conference, 2026, Publisher: JACoW

Acknowledgements

This thesis is in many ways a culmination of a very long curiosity of mine, a curiosity of physics and of very big machines. My interest and history with accelerators started quite early, when I was only 15 and Kevin Fissum kindly gave me a tour of the old MAXLab. I can mostly recall that day as a mess of wires and components that I could not hope to understand, but I found something very intriguing in the chaos of the machine, so a big thanks to Kevin for that original push in this direction.

The next big step towards this point was my employment with the operations group at MAX IV, and I must thank everyone in that team, both for their support and teaching during my employment, and for all their critical help with the work presented here.

Moving chronologically, I want to thank the supervisors of my bachelor and master thesis works. Starting with the BSc, a big thanks to Åke Andersson and David K Olsson, for taking me on in a project that was a bit over my head at the time. I also want to thank them both for their continued support in my PhD work, David as a fellow student (for a little while) and Åke as a continued supervisor of sorts, and his critical help in teaching me the Wiedemann book.

From my time with the MSc, I want to extend a big thanks to Sverker Werin and Natalia Milas. More on Sverker later, but Natalia deserves some special attention. She is still the one of the most enthusiastic and driven people I have met in the field (or in general), and her energy in supervising and advising me really has pushed me further than I would have thought possible.

This work would not have been possible without my great thesis supervisors, Francesca Curbis, Sverker Werin and Erik Mansten. Francesca, thank you for your continuous support and advice in all aspects of this work, and beyond. Sverker, thank you for your keen eye for communication and editing. Beyond your help in my understanding of physics and accelerators, I truly think you have made me a better speaker and writer. Thank you, Erik, for your enthusiasm and active interest in all my little projects. Even though you showed up as supervisor on paper only after a few years, I believe you have been a guiding hand in all experiments throughout the whole PhD. I would also like to thank Jonas Björklund Svensson and Mihai Pop, who were both office mates of mine at different times. You both gave me some more "unofficial" supervision, and Jonas, you were a great help on all projects with your extensive expertise in analysis and theory.

Staying on the topic of the actual thesis work, I want to thank my closest collaborators at the different facilities I have had the pleasure of working with. Guiseppe Penco at FERMI was the first to join this collaboration and I want to thank him for his great help in experiments, analysis and writing. A big thanks to Philipp Dijkstall at PSI for the massive support in the related projects, as well as for his continuing advice on the world of academia and accelerators. Also, thank you to Gesa Goetzke at FLASH for her enthusiastic support of the work.

More locally, I would like to thank everyone I have had the pleasure of meeting and working with in the accelerator development group at MAX IV. A non-exhaustive list of people that have improved 4.5 years of lunches and evenings out should include: Adam, Henrique, Jonas and Jonas, Murillo, Gus, David, Francis, Neven, Carl, Marco, Magnus, Lennart, Åke, Domenico, Filip, Pedro and Sara.

Work isn't everything, and I want to thank everyone in my personal life that has helped and supported me, both mentally and physically, throughout the years. I am proud to actually have a *list* of close friends to thank, including: Mats, Simon, Adam, Sjövall, Greiff, Alle, Isak, Calle, Axel and Zebbe. I also want to thank the book club which has been going semi-steadily throughout my whole PhD time: Axel and Carmen, Rezvan, Jakob and Anna, and Tova. If nothing else, thanks for forcing me to read War and Peace and act in a very strange movie.

Finally, and most importantly, I want to thank those closest to my heart. My mother, Gerd, has always helped me through everything in life and has always believed in me with such fervor that it is hard not to agree with her, try as I might sometimes. My father, Mikael, has always, and continues to inspired me. My siblings, Tova and Pym, always support me in all I do. Luisa has carried me for so long, in such innumerable ways. I do not say it enough, but I love you all very much.

But this dedication is for others to read:

These are private words addressed to you in public.

— T. S. Elliot

Popular summary in English

Particle accelerators are among the world’s most advanced and complex scientific instruments, pushing the boundaries of human understanding and achievement. They have a multitude of uses, among them novel studies of matter and driving developments of future technologies for medicine, industry, and fundamental physics. To achieve all this, accelerators must produce beams of charged particles with extraordinary precision and control. The tiniest variations and errors in the beam can significantly affect performance, which makes detailed diagnostics to understand and monitor accelerated beams essential. Within the subfield of beam diagnostics, it is especially difficult to monitor the *longitudinal* properties of accelerated beams, since they are often extremely short and move near the speed of light.

In this thesis we investigate the possibilities in introducing machine learning methods to find new ways to monitor and control electron beams, with a focus on longitudinal properties and methods. Machine learning refers to methods in data science in which computer algorithms learn patterns in data by themselves, through iterations and self-updating, rather than explicitly programmed procedures. Once trained, such algorithms can recognize complex patterns and relations in data to make accurate predictions from new measurements. Modern accelerators generate vast amounts of data, often with underlying patterns which are difficult to detect by traditional means. One can train machine learning algorithms to recognize such patterns and reconstruct beam information which is difficult or even impossible to measure without disturbing the beam itself. These *virtual diagnostics* can provide rapid predictions of beam shape and distribution, allowing for new modes of monitoring and stable operation in accelerators and improvements in many areas which rely on such information.

The methods developed in this thesis were first tested and produced for the MAX IV linear accelerator, but later saw wider use at several different particle accelerators. There was an initial development for extracting the longitudinal information, such as the length and energy distribution of electron bunches, from other beam diagnostic sources. Later, more advanced applications were made of the algorithms to areas of beam control and calibration of novel diagnostic systems. Together, these studies demonstrate how machine learning can become a valuable tool for the operation of modern accelerator facilities, assisting in providing new information for operators and optimization systems.

As part of a full characterization of the longitudinal diagnostics and dynamics, this thesis explores the dynamics of controlling and longitudinally compressing

accelerated beams. To achieve the short timescales required for modern accelerator experiments, the machines use specific configurations of magnets to shorten particle bunches, "compressing" them. This can be done in a number of ways, and the MAX IV linear accelerator has chosen a novel configuration, called an *arc-like bunch compressor*. The advantages and dynamics of these rare systems are further explored in this thesis, and their more detailed control of the beam has been characterized to new levels, and their dynamics have been explored to find new advantages of the scheme.

Overall, this thesis serves to introduce new methods to make particle accelerators more precise, efficient and reliable. The work here can hopefully help accelerator facilities achieve the precise longitudinal control and monitoring required to meet the ever rising demands from scientific applications.

Populärvetenskaplig sammanfattning på svenska

Partikelacceleratorer hör till världens mest avancerade och komplexa vetenskapliga instrument och tänjer ständigt gränserna för mänsklig kunskap och teknisk utveckling. De används bland annat för att studera materiens uppbyggnad, utveckla framtidens teknik inom medicin och industri, samt bedriva forskning inom den mest grundläggande fysiken. För att möjliggöra detta måste acceleratorer producera strålar av laddade partiklar med enastående precision och kontroll. Även mycket små variationer eller fel i strålen kan påverka prestandan avsevärt, vilket gör noggranna diagnostiksystem för att förstå och övervaka de accelererade strålarna helt avgörande. Det är särskilt svårt att mäta strålarnas *longitudinella* egenskaper, eftersom de består av extremt korta partikelpaket som färdas nära ljusets hastighet.

I denna avhandling undersöks möjligheterna att använda maskininlärning för att utveckla nya metoder för att övervaka och styra elektronstrålar, med särskilt fokus på deras *longitudinella* egenskaper. Maskininlärning är en gren inom datavetenskap där algoritmer lär sig att identifiera mönster i stora datamängder genom att successivt förbättra sina egna modeller, i stället för att följa i förväg programmerade regler. När algoritmerna väl har tränats kan de upptäcka komplexa samband och göra träffsäkra förutsägelser utifrån ny mätdata. Moderna acceleratorer genererar mycket stora mängder data, där viktiga samband ofta är svåra eller omöjliga att upptäcka med traditionella analysmetoder. Genom att träna maskininlärningsmodeller på sådan data kan man återskapa information om elektronstrålen som annars är svår eller rent av omöjlig att mäta utan att störa den. Dessa så kallade *virtuella diagnostiksystem* kan snabbt uppskatta strålens *longitudinella* form och fördelning och därmed möjliggöra nya arbetssätt samt förbättra många acceleratorsystem som är beroende av sådan information.

Metoderna som utvecklades i denna avhandling togs först fram och testades i linjäracceleratorn vid MAX IV, men kom senare att användas vid flera olika acceleratoranläggningar. De första studierna fokuserade på att återskapa *longitudinell* strålinformation från andra mätningar, medan senare arbeten använde algoritmerna för styrning av elektronstrålar och kalibrering av nya diagnostiksystem. Tillsammans visar dessa studier hur maskininlärning kan bli ett värdefullt verktyg för drift och utveckling av moderna acceleratoranläggningar.

Utöver studierna av maskininlärningsbaserad *longitudinell* diagnostik behandlar avhandlingen även de fysikaliska processerna bakom styrning och *longitudinell* kompression av elektronstrålar. För att uppnå de extremt korta pulser som krävs i dagens acceleratorbaserade experiment används särskilda magnetkon-

figurationer som komprimerar partikelpaketen. Detta kan göras på flera olika sätt, och linjäracceleratorn vid MAX IV använder en ovanlig lösning som kallas en *bågformad paketkompressor*. I avhandlingen undersöks dessa systems egenskaper och dynamik, och resultaten visar att de har flera intressanta befintliga fördelar.

Sammantaget presenterar denna avhandling nya metoder som kan göra framtidens partikelacceleratorer mer precisa, effektiva och tillförlitliga. Förhoppningen är att resultaten ska bidra till utvecklingen av nya acceleratoranläggningar och metoder för att möta de ständigt ökande kraven från framtidens vetenskapliga och tekniska tillämpningar av acceleratorer.

Machine Learning for Longitudinal Monitoring and Control of Ultra-Bright Electron Beams

Chapter 1

Introduction

Particle accelerators enable scientific work in a broad range of areas and applications by producing beams with precisely controlled properties. The rising demands of accelerator applications place correspondingly greater requirements on beam diagnostics and control. Specifically, there is a growing need at modern accelerator facilities for precise measurement and control of the longitudinal properties of accelerated beams. Parameters of the longitudinal phase space (LPS), such as bunch length, energy spread and current profiles, are becoming increasingly critical with the high demands of free-electron lasers (FEL) [1] and emerging advanced accelerator experiments, such as plasma wakefield accelerators [2]. Naturally, new and complex challenges arise when there is a need for high-accuracy diagnostics and control of ultra-fast phenomena, at the femto-second scale and beyond [3].

Beam properties in the LPS are strongly connected to the compression of particle bunches. By understanding and tuning compression systems, improved performance can be reached. With more elaborate compression control follows a need for an online and constant understanding of the beam LPS, which often stands in conflict with the working principles of current diagnostic systems. In this work, we have analyzed a special compressor system, the MAX IV arc-like bunch compressors [4, 5] (as can be seen in Fig. 1.1).

As compressed electron beams now commonly reach the femtosecond-scale and below, demand for novel diagnostic methods for LPS distributions is increasing. The established method of investigating beam LPS, using transverse streaking in a radio-frequency (RF) cavity [6, 7, 8], has enabled detailed insights into longitudinal beam properties, but remains limited in key regards. Issues, such as

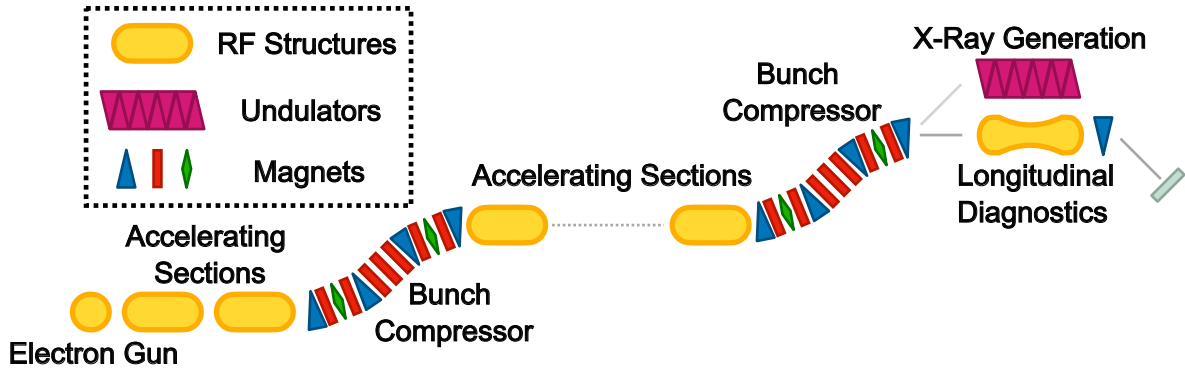


Figure 1.1: A simplified layout of the MAX IV linear accelerator.

the cost to implement these methods, have been mitigated with innovative technologies like passive wakefield streakers [9, 10, 11, 12] and bunch compression monitors [13], but these detector systems are still often invasive, technically complex and time-consuming to operate. With the growing demand for detailed LPS diagnostics, there exists a strong motivation for developments in non-invasive diagnostics and optimization techniques which are fast, robust and generally applicable.

Along with the demand for longitudinal diagnostics and control, the field of machine learning (ML) applied to particle accelerators has steadily been growing. Across multiple fields, ML has proven to be a widely-applicable tool for prediction and optimization in complex and high-dimensional systems. Within accelerator physics, ML can be used to draw on the complex, nonlinear relations between machine states and beam observables to predict beam distributions and properties. ML can complement and potentially surpass more analytic or simulation-based approaches. Surrogate modeling [14], virtual diagnostics [15, 16] and data-driven optimization [17] are areas of rapid development, leading to enhanced machine performance and enabling new diagnostic modes for operators and users of modern machines.

In this thesis it is shown how the application of machine learning in the areas of longitudinal diagnostics and optimization has been explored and developed across multiple facilities and projects. Several questions are relevant in this context.

The first question was *whether the longitudinal phase space (LPS) of an electron beam could be accurately reconstructed from non-invasive accelerator diagnostics*. In this pursuit I initially investigated methods using simulated LPS data and measurements in a dispersive section of the MAX IV linac (**Paper I**), before extending the studies to direct predictions of the full LPS using measurements from the newly commissioned transverse deflecting structure (TDS) (**Paper II**).

With the established capability of virtual diagnostics (VDs) to reconstruct the MAX IV LPS, the next area of interest was *whether the developed methods could be generalized to other facilities*. This led to the application of VD systems at FERMI and SwissFEL, showing that the approach could be expanded to several different facilities (**Paper III**).

The VD developments opened the possibility of the application of ML techniques within other established areas of controls and diagnostics, and the question was raised *whether VD predictions could be incorporated to enhance existing systems and procedures in accelerators*. Such investigations resulted in the development of a novel calibration procedure for the PS system at SwissFEL, using a VD to infer critical longitudinal information using uncalibrated PS signals (**Paper IV**). It also resulted in progress into the application of optimization on LPS parameters using VD signals at MAX IV (**Paper V**) and an application at FLASH, using VD signals to reconstruct the longitudinal properties of FEL radiation (**Paper VI**).

Taken together, the studies in this thesis serve to demonstrate a progression from establishing the feasibility of ML techniques to investigating both their generalization between facilities and their integration into advanced accelerator operations.

The LPS studies described thus far are closely connected with the dynamics of longitudinal compression of particle beams. In connection with the other work on longitudinal diagnostics in this thesis it became natural to investigate *whether there were unexplored advantages of the arc-like compressor scheme* [4]. The decoupled control of second-order compression parameters in arc-like compressors lacked experimental confirmation, and it was of interest to measure the effect of the compressor sextupoles on these parameters at MAX IV. To enable this, I developed a time-of-flight (TOF) setup which allowed for measurements of the bunch compression parameters during sextupole scans (**Paper VII**). In a second project on the arc-like compression scheme, I participated in investigations of an inherent stability in the compression unique to the arc-like compressors (**Paper VIII**). Such inherent stabilizing properties could prove critical in implementing bunch compression for plasma-based accelerators in the future.

The structure of this thesis is as follows. In Chapter 2, I begin with an overview of the relevant longitudinal beam dynamics and diagnostic systems. Here I include a description of the longitudinal phase space and its critical parameters, two longitudinal compression schemes, and a review of established longitudinal diagnostic systems. I also include a description of the FEL lasing process and its effect on the electron bunch LPS. Following this, in Chapter 3, I introduce the basic components of machine learning relevant to accelerator applications, and go on to provide an overview of virtual diagnostics and surrogate models. I also include a brief overview of Bayesian optimization and its relevance for accelerators. The subsequent Chapter 4 describes the development and implementation of virtual diagnostics across multiple facilities, with an emphasis on the stages of development at MAX IV. After this, in Chapter 5, more advanced applications of the VDs into existing accelerator operations are described and shown to improve on multiple processes. Then, in Chapter 6, measurements and investigations of the unique aspects of compression in the MAX IV arc-like bunch compressors are presented. Finally, the results are discussed in a broader accelerator context, and conclusions and future research directions are outlined.

Chapter 2

Longitudinal Beam Dynamics and Diagnostics

Modern electron linear accelerators (linacs) rely on precise control of accelerated beams in six-dimensional phase space. The dynamics in the transverse planes, (x, x', y, y') , governs the focusing and transport of the beam, but it is the longitudinal dynamics in $(z, \frac{\Delta p}{p_0})$, i.e. in longitudinal position and relative energy deviation, that determines the temporal structure, peak current and energy distribution of the beam [18, 19]. These longitudinal properties are critical in linac applications like free-electron lasers (FELs), wherein the lasing process is extremely sensitive to the LPS distribution [20], and other uses of compressed accelerated beams, such as plasma wake-field accelerators [2].

This chapter provides a brief introduction to the longitudinal beam dynamics and concepts critical to the material of the thesis. To start, the longitudinal phase space and its key parameters are introduced and discussed. The principles of longitudinal beam control are then reviewed, with a focus on acceleration in radio-frequency structures, followed by two sections on bunch compression in magnetic structures of two different types. Then some key longitudinal diagnostic techniques are summarized, providing the necessary background for the work on virtual diagnostics and longitudinal optimization in this thesis. Finally, the process of FEL lasing is broadly described with a focus on the connection between this process and the electron beam's longitudinal distribution.

2.1 The Longitudinal Phase Space

The full state of a charged particle beam is described by six-dimensional coordinates, corresponding to the positions and momenta of all particles. When representing the full distribution of a beam it is common to group the coordinates by the transverse phase spaces (x, x') and (y, y') , as well as the longitudinal phase space (LPS) defined by the longitudinal position z and the relative energy deviation δ of the particles. These last two coordinates are defined as

$$z = c(t - t_0) = c\Delta t, \quad \delta = \frac{E - E_0}{E_0} = \frac{\Delta E}{E_0} = \frac{\Delta p}{p_0}, \quad (2.1)$$

where z is the position relative to a reference particle at $z_0 = 0$ m, or $t_0 = 0$ s, and δ is the relative energy deviation, as expressed in either energy E or momentum p , with a reference particle at E_0 or p_0 [18]. Many of the parameters within this phase space are critical for the production of radiation, including the bunch length Δz and energy spread $\Delta\delta$, i.e. the spread of the beam distribution's projection onto either plane, z or δ . This can be defined by different width calculations and within this thesis the root mean square (RMS), full width at half maximum (FWHM) or interquartile range (IQR) are all used, as best applied in each individual case and distribution. There are also cases where the full spread in the δ -projection is of less interest, and the separate energy distributions within smaller slices of the beam along z becomes more relevant. This parameter is referred to as the *slice energy spread*.

The *energy chirp*, h , of the beam describes the internal correlation between the longitudinal position z and the relative energy deviation δ in the bunch. When the beam has an energy chirp, particles at different z -positions will have different mean energies. While h will here be used to refer to the linear part of the total $\frac{\partial\delta}{\partial z}$ correlation, there are often higher orders present as part of this correlation. An example of a beam distribution in the LPS, as measured in the MAX IV linac, can be seen in Fig. 2.1, with the projections onto each axis plotted above and to the right of the image. The relevant FWHM parameters are indicated, as well as the energy chirp based on a third-order approximation of $\frac{\partial\delta}{\partial z}$. The energy chirp plays a central role in longitudinal bunch compression, where the longitudinal phase space is rotated to reduce the bunch length at the expense of increasing the energy spread [18, 19]. In modern linacs, this chirp is primarily generated and controlled through the acceleration process in radio-frequency structures.

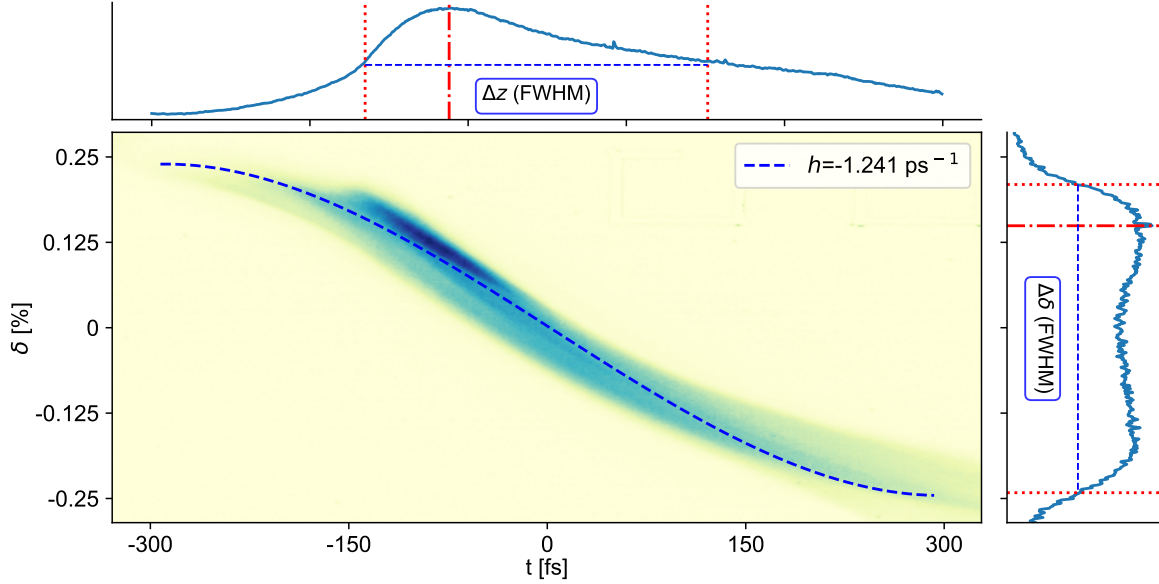


Figure 2.1: Example of a beam's longitudinal phase space measured in the MAX IV linac, with projections plotted and key beam parameters indicated, such as the beam's energy chirp h , bunch length Δz and energy spread $\Delta\delta$.

2.2 RF Acceleration and Energy Chirp

Acceleration of the particles in modern electron linacs is achieved using radio-frequency (RF) electromagnetic fields within resonant accelerating cavities [21, 22, 18, 19]. In such resonant cavities, the electric component of the RF field transfers energy to the beam as it traverses the cavity. The total energy gain of a particle passing through an RF structure can be approximated as

$$\Delta E = eV_{RF} \cos(\phi_0 + kz), \quad (2.2)$$

where e is the elementary charge, V_{RF} is the amplitude of the integrated accelerating field, and ϕ_0 is the RF phase experienced by a reference particle relative to the crest of the accelerating field. The term kz expresses the additional phase experienced by the specific particle at longitudinal position z . Here, $k = 2\pi/\lambda_{RF}$ is the RF wave number, where λ_{RF} is the RF wavelength.

As the RF field varies sinusoidally in time, particles at different longitudinal positions within the bunch experience different accelerating gradients. A correlation is therefore established between the longitudinal coordinate z and the final particle energy, generating an energy chirp h within the beam. Expanding the energy gain around the reference particle of phase ϕ_0 , i.e. performing a Taylor expansion around $z = 0$ m, gives

$$\Delta E(z) \approx eV_{RF} \left[\cos \phi_0 - kz \sin \phi_0 - \frac{(kz)^2}{2} \cos \phi_0 + \mathcal{O}(z^3) \right]. \quad (2.3)$$

The linear term in the expansion generates the first-order energy chirp, while higher-order terms contribute nonlinear curvature in the longitudinal phase space.

Acceleration at the crest of the RF wave, $\phi_0 = 0^\circ$, maximizes the total beam energy while minimizing the induced energy chirp and spread. However, the energy chirp is required in most longitudinal bunch compression schemes, and therefore the beam is more commonly accelerated off-crest. The sign of h can be controlled via accelerating on either the rising or falling slope of the RF sinusoidal field.

The nonlinear curvature expressed in the higher orders of Eq. 2.3 can become problematic during strong compression, where distortions and growing energy spread can appear as a result [19]. Therefore, it is not uncommon to employ additional RF structures of higher frequency, or utilize higher-order magnets to linearize the LPS of the beam, minimizing higher-order terms while controlling the linear part of the energy chirp.

2.3 Longitudinal Compression

The generation of ultrashort electron bunches is a fundamental requirement for many modern accelerator applications. Specifically, FELs and plasma wakefield accelerators utilize effects strongly dependent on the LPS distributions of the electron beams. Typically, electron bunches emerging from guns have durations of several picoseconds, and many applications often require femto- or even attosecond-scale bunch lengths. Collective effects, mainly space charge, limit the ability of generating short bunches at the source. Instead, electron beams are accelerated with a controlled energy chirp, which can then be converted to an effective longitudinal compression of the bunch using magnetic lattices. Several schemes exist for this type of longitudinal compression, including magnetic chicanes, doglegs and arc-like compressors [18]. The work presented in this thesis focuses mainly on compression using chicanes and the more specialized arc-like compression scheme employed at MAX IV.

The effect of a magnetic lattice on the longitudinal beam distribution can be described using the transport matrix formalism [23]. To third order, the longitudinal position of a particle at a final state f , as a function of its energy and position at initial state i , can be expressed as

$$z_f(\delta) = z_i + R_{56}\delta + T_{566}\delta^2 + U_{5666}\delta^3 + \mathcal{O}(\delta^4), \quad (2.4)$$

where z_i represents the initial position and the notation of R_{ij} and T_{ijk} arises from the matrix formulation of beam transport [18, 19, 23]. This formulation expresses the dependency between beam coordinates at different locations in the accelerator, expressed in the $(x, x', y, y', z, \delta)$ 6D space. For example, R_{12} refers to the $\frac{\partial x}{\partial x'}$ dependency and R_{56} then refers to the $\frac{\partial z}{\partial \delta}$ dependency. In principle, R_{56} expresses the difference in path length as dependent on energy. To first order, compression corresponds to a rotation of the beam distribution in the (z, δ) plane, expressed by the transport matrix generated by the compressor optics as $R_{56} \neq 0$ m. The T and U matrices express higher-order dependencies, e.g. $T_{566} = \frac{\partial^2 z}{\partial \delta^2}$ expresses the second-order dependency of the longitudinal coordinate on energy.

The longitudinal transfer elements of R_{56} and T_{566} will appear whenever the transverse effecting elements of R_{16} and T_{166} (or R_{36} and T_{366}) are present. These elements are also known as first- and second-order transverse *dispersion*, and describe the effect of the beam energy on the transverse position [18]. Dispersion appears whenever the beam path is bent, as different relativistic beam momenta p result in different bending angles and radii ρ . The relation to R_{56} is described by

$$R_{56} = - \int \frac{R_{16}}{\rho} ds, \quad (2.5)$$

where s is the coordinate along the beam path [4, 24]. One can also expand to T_{566} as

$$T_{566} = - \int \frac{R_{26}}{2} + \frac{T_{166}}{\rho} ds. \quad (2.6)$$

As R_{56} and T_{566} also describe the dependency between beam coordinates and the beam energy, they are occasionally referred to as elements of *longitudinal dispersion*.

Given Eq. 2.4 one can infer the basic principle of bunch compression: if the beam has an internal correlation between z and δ , i.e. if an energy chirp is present, then the beam's distribution in the LPS can effectively be rotated. With the proper signs on both energy chirp and R_{56} , this rotation can result in a longitudinal compression of the bunch as a whole, bringing both the head and tail closer to the core of the beam. The rotation can even carry on beyond the point of minimum projected bunch length, going into *over-compression*. The R_{56} required for bunch compression can be generated and controlled with specific types of magnetic lattices, described in more detail below.

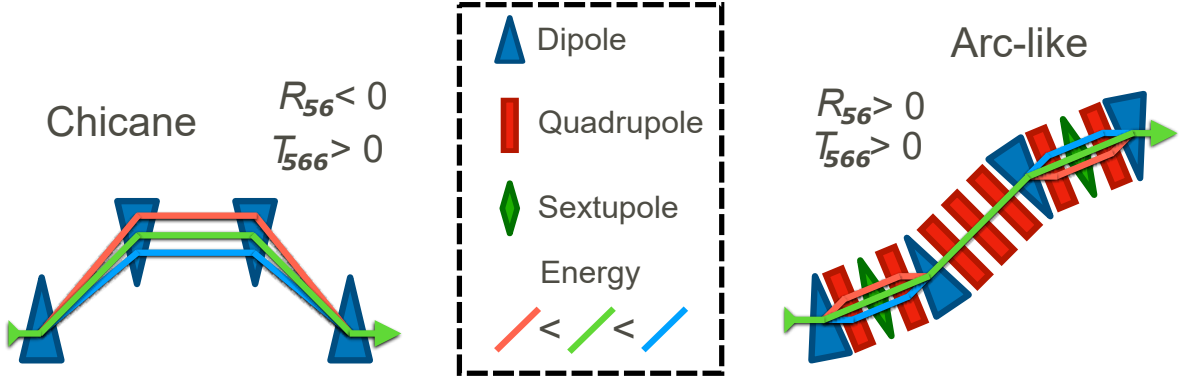


Figure 2.2: Schematics of the two types of bunch compression schemes relevant to this work. On the left, the more well-established magnetic chicane scheme, and on the right, the MAX IV arc-like compression scheme.

2.3.1 Chicane Bunch Compressors

Magnetic chicanes are the most widely used compression scheme for longitudinal bunch compression in modern linear accelerators. The standard magnetic lattice in this scheme consists of four dipoles arranged such that the beam is first deflected from and then returned to the main trajectory of the linac, as illustrated in the left schematic in Fig. 2.2. The dispersion in the dipoles results in an energy dependent path length through the relation in Eq. 2.5. In a chicane the generated R_{56} is negative, i.e. higher-energy particles take a shorter path through the chicane compared to lower-energy particles. This means that a beam arriving with a positive energy chirp, $h > 0$, would experience a rotation of the longitudinal phase space resulting in a shorter bunch in total. This corresponds to the opposite chirp orientation compared to the example shown in Fig. 2.1.

The total R_{56} of a chicane bunch compressor can be tuned and controlled by adjusting bending angles and dipole spacing, although this requires some flexibility in the path laid out by the beam's vacuum pipe, which is not common. It is more common to control the compression by adjusting the energy chirp of the beam arriving at the compressor. One can express the total compression factor C as a combination of the R_{56} of the compressor and the energy chirp h of the incoming beam as [25, 19]

$$C = \frac{1}{1 + hR_{56} + \mathcal{O}(\delta)}. \quad (2.7)$$

The energy chirp can be controlled using the main RF structures used to accelerate the beam, as different phases of the accelerating field will contribute different gradients in energy within the beam. Thus the total compression of the beam can be controlled using the accelerating parameters. It is also common

to use a higher-order RF structure downstream of the final accelerating stage to control for higher-order energy chirp, *linearizing* the LPS before compression.

Chicane bunch compressors are relatively simple and robust, and their effectiveness makes them a standard choice in most linac-based facilities. However, they do have some critical limitations. One detriment is their inability to control higher-order terms such as T_{566} , which can lead to either undesirable curvature in the compressed beam's LPS, or difficulties in controlling desired structures in the LPS reliant on such curvature [19]. This is what drives the need for upstream higher-order RF structures, which are also costly to implement. Further, chicane compressors generate relatively large amounts of coherent synchrotron radiation (CSR) in the strong bending magnets. This radiation can degrade critical beam parameters, such as slice energy spread, when interacting with the beam [26].

2.3.2 Arc-like Compressors

Arc-like compressors, displayed in the right schematic in Fig. 2.2, provide a more advanced alternative to chicane compressors. The arc-like scheme consists of a symmetric lattice of multiple orders of magnets, arranged in an achromatic arc [4, 5]. One key design component is the more gradual and tuneable dispersion within the arc, with quadrupole and sextupole magnets in the dispersive regions allowing for greater control of first- and higher-order longitudinal transport terms. This control alleviates the issues of higher-order energy chirp in the beam, and the need for higher-order RF structures.

The R_{56} element of the total transfer matrix comes out as negative, and as such the requirement on the h of the incoming beam is reversed relative to the chicane scheme; compression requires higher energies in the head and lower energies in the tail, or $h < 0$, as exemplified in Fig. 2.1. This results in an LPS rotation into a shorter bunch. This desired energy chirp can be achieved naturally by accelerating the beam on the falling side of the crest of the RF field. Accelerating in this region also provides some additional stability, as correlated phase and amplitude jitter in the RF field cancels each other [27].

The addition of higher order magnets within the dispersive regions of the arc provides independent control of both R_{56} and T_{566} , as these types of dispersive magnets directly control the contributing R_{16} and T_{166} matrix components, see Equations 2.5 and 2.6. Adjusting R_{56} using the dispersive quadrupoles is rare, as it is simpler to maintain the achromaticity of the transfer and instead adjust h with the accelerating RF phase, but the novel control of T_{566} allows tuning of

the *linearization* in the LPS, minimizing higher-order components of the energy chirp.

2.4 Longitudinal Diagnostics

Diagnostics, systems for measuring the properties of accelerated beams, are critical subsystems for any modern accelerator. In order to maintain and accelerate particle beams, online monitoring and control of the beam is key. For example, one of the principal diagnostic designs in accelerators is the beam position monitor (BPM). There exists a variety of designs of these systems, but in the most common designs they consist of electrodes set in the beampipe walls surrounding the passing beam [28]. The signal from the electrodes can then be collected and analyzed to estimate the transverse position of the beam. This system does not disturb the passing beam in any significant way, and is referred to as *non-destructive* or *non-invasive*. An example of a *destructive*, or *invasive*, diagnostic system is the scintillating screen and beam imaging system. While BPMs may provide the (x, y) position of the beam's center-of-charge, they provide no information of the beam's distribution in the plane. To extract this, one can insert a scintillating crystal into the beam path, allowing the beam to collide with the crystal in order to produce light through scintillation. While many such crystals exist in production, it is most common in electron accelerators to use yttrium aluminum garnet (YAG) crystals. The produced light is then imaged and recorded using a camera setup. This is, in part, the type of system used to produce the data for Fig. 2.1.

Thus far the examples covered have focused on the transverse, (x, y) , dimensions of the beam. The longitudinal position and distribution of the beam in (z, δ) requires additional work to probe. Properties in δ can be investigated by applying the mentioned transverse techniques in dispersive regions of the beam path, where $R_{16} \neq 0$ m (or $R_{36} \neq 0$ m). BPMs in dispersive regions can provide relative measurements of the central momentum p , while dispersive imaging systems can provide the full energy distribution of the beam, allowing for measurements of the beam energy spread, $\Delta\delta$. The longitudinal properties, however, are even more difficult to probe. As the demands of modern FELs and linacs have increased, the requirements on the resolution of longitudinal diagnostics have also grown more drastic. For instance, an expectation of sub-10 fs resolution leaves only a handful of feasible techniques. Here, the three longitudinal diagnostic systems are described, including transverse deflecting structures, passive wakefield streakers and bunch compression monitors.

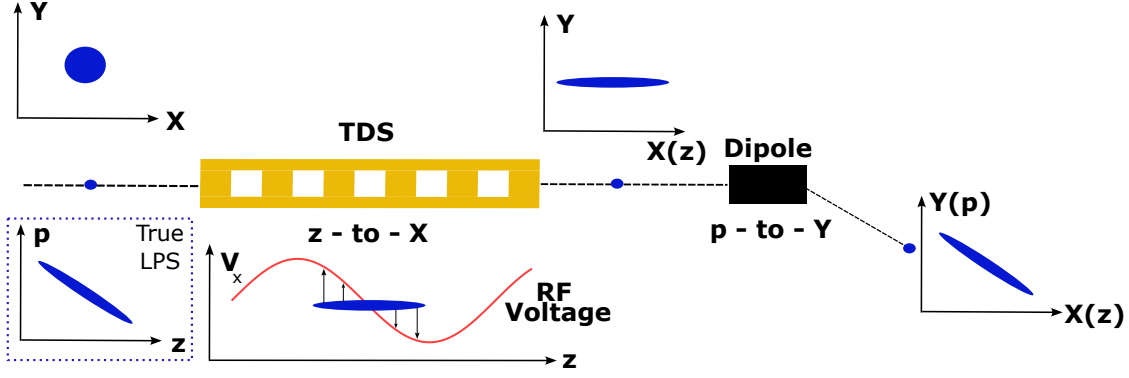


Figure 2.3: A schematic of a transverse deflecting structure (TDS) system, with a simplified image of the RF structure and the passing beam, as well as the RF sinusoidal field and the spectrometer dipole.

2.4.1 Transverse Deflecting Structures

The transverse deflecting structure (TDS), at times referred to as a transverse deflecting cavity (TDC), is one of the most useful diagnostic systems for probing longitudinal beam properties in modern accelerators. TDS systems provide direct access to the temporal structure of electron bunches with femtosecond-scale resolution, making them especially valuable in high-brightness linacs and FEL facilities, where the longitudinal beam properties strongly influence machine performance [20]. Although numerous longitudinal diagnostic systems exist, including the coherent radiation diagnostics and wakefield streaking techniques covered in the coming sections, the TDS has become the reference diagnostic at modern FEL facilities because it can provide direct single-shot measurements of the temporal distribution with femtosecond resolution.

The TDS is a relatively recent addition to accelerator diagnostics. While early concepts, experimental tests and design studies appeared already during the 1960s [29, 30], the technique saw limited practical use due to technological constraints and the less demanding longitudinal requirements in earlier accelerator facilities. With the emergence of modern, large-scale FELs in the 21st century, the demand for high-resolution longitudinal diagnostics increased significantly, leading to widespread implementation of TDS systems at modern light sources and FEL facilities [7]. Today, TDS systems are routinely used for measurements of bunch lengths, current profiles and full longitudinal phase space distributions.

The layout of a TDS is illustrated in Fig. 2.3. In principle the TDS set up is the same as for RF acceleration, but the cavities are designed for and operated in a different transverse mode for the field (TM₁₁₀) which results in a transverse deflection of the beam, rather than longitudinal acceleration. If the relative phase of this field is tuned such that a passing particle bunch is kicked at the zero-

crossing of the field amplitude on either side of the crest, the bunch center will experience no net transverse deflection (see the lower plot in Fig. 2.3). However, the particles within the bunch will experience transverse kicks linearly proportional to their longitudinal positions z relative to the bunch center, with a sign depending on the slope in field amplitude at the operational zero-crossing. After propagation through downstream drift space, this transverse kick is transformed into a measurable transverse displacement Δx . In this manner, the longitudinal positions z within the bunch are mapped onto a transverse axis, allowing direct reconstruction of the bunch temporal profile. The final transverse displacement of a particle at z can be described by a simplified oscillating RF field as

$$\Delta x(z) = R_{12} \frac{eV_{x,amp}}{E} \sin(\phi_0 + kz), \quad (2.8)$$

where $V_{x,amp}$ is the amplitude of the integrated transverse deflecting field, k the RF wave number, ϕ_0 the RF phase and E the beam energy. At the zero-crossings, and for a short enough bunch in z , this sinusoidal kick is linear, producing transverse displacements directly proportional to the longitudinal positions of beam particles. With a subsequent scintillating screen, the beam's longitudinal distribution can be fully imaged. The achievable temporal resolution of a TDS system depends on several factors, including the RF frequency, cavity voltage, intrinsic beam size, and downstream optics, as explored in detail in the supplemental material of C. Behrens *et al.* [3]. Modern TDS systems are capable of achieving temporal resolutions below 10 fs, making them critical tools for FEL facilities, where structures in these time-scales directly influence the light production [20].

An additional major strength of a TDS system is the ability, when combined with a downstream dispersive section, to directly visualize the full LPS of the beam. If the dispersion is applied in the orthogonal axis to the TDS mapping, in our example y and x respectively, then one transverse axis describes the longitudinal position z and the other has been mapped to the relative energy deviation δ . In Fig 2.3, the sequential mapping of the transverse coordinates to the longitudinal is illustrated, first with an $z \rightarrow x$ mapping using the TDS and then with a $p \rightarrow y$ mapping using the spectrometer dipole. An imaging screen at the end of such a process provides a two-dimensional measurement of the beam distribution in the (z, δ) -space, enabling detailed studies of compression, chirp, and CSR effects on the beam. An example of such a measurement is shown in Fig. 2.1.

Despite their advantages and versatility, TDS systems have inherent drawbacks. For one, it is a fully invasive diagnostic technique, requiring the beam to impact a scintillating screen, destroying it for further use. It is precisely this

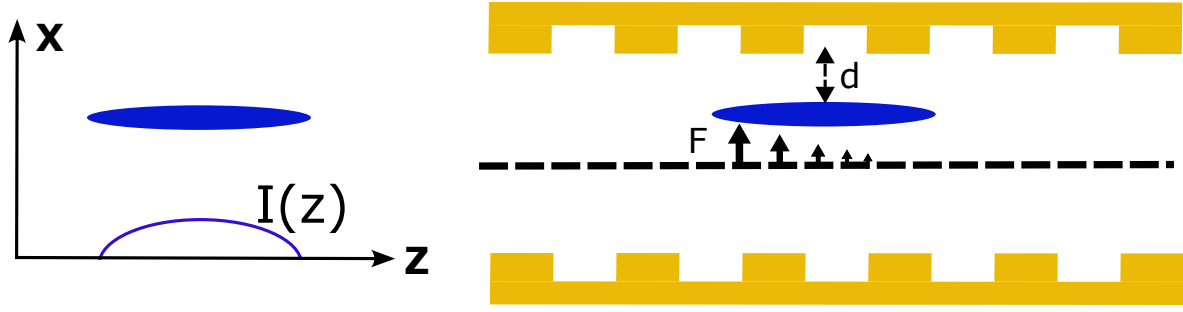


Figure 2.4: Schematic illustration of a passive wakefield streaker using corrugated metallic plates. An electron beam traversing the structure with a transverse offset excites wakefields that impart a longitudinally dependent transverse kick, resulting in a streaked beam profile on a downstream screen.

destructiveness, along with the high complexity of the measurement, which has motivated investigations into machine learning applications for inferring the LPS non-destructively. Secondly, TDS systems are extremely expensive as compared with other diagnostic systems, as they require full, dedicated RF systems. A promising alternative that is actively being explored is the passive wakefield structures described in the coming section [11]. Nevertheless, TDS systems remain one of the most precise, direct, and widely used methods for measuring ultrashort electron bunches in modern accelerator facilities.

2.4.2 Passive Wakefield Streaker

In recent years, considerable effort has been devoted to developing alternatives to RF-based TDS systems. As a compact alternative to conventional transverse deflecting structures (TDS), passive wakefield streakers (PS) have emerged as an intriguing method for obtaining high-resolution longitudinal beam information without requiring an active RF system [9, 10]. A PS system utilizes electromagnetic wakefields generated directly by the beam itself as it traverses a specially designed structure, either a dielectric or corrugated metallic surface which is positioned close to the beam path. If the beam propagates with a transverse offset relative to the symmetry axis of the structure, the generated wakefields impart transverse kicks to the beam of increasing strength along z . Thus the beam is streaked in a manner similar to the forces in an active TDS [11]. A simplified schematic can be seen in Fig. 2.4, with the critical parameters of $I(z)$ and d highlighted.

An analytical wakefield model of PS systems using rectangular corrugated metallic plates has been developed by K. Bane *et al.* [31]. Using this work one can express the experienced transverse displacement on a screen at some distance

from the structure in a similar manner to Eq. 2.8, as

$$\Delta x_{\text{PS}}(z) = R_{12} \frac{L}{E} \int_{-\infty}^z I(z') w(z - z'; d) dz', \quad (2.9)$$

where we now introduce also the current profile I of the beam and its transverse wake function w . The wake function is explored in detail by K. Bane *et al.* and depends strongly on the geometry and material properties of the wakefield structure, but for a given system it is mainly dependent on the distance d from the passing beam to the nearest corrugated plate.

One major advantage of PS over TDS is the compactness of the systems compared to conventional RF-based TDS installations. Since PS systems are also passive, they avoid the need for costly high-power RF systems. These advantages in combination make these systems orders of magnitude simpler to install and operate than an equivalent TDS system. Further, the streaking field is inherently synchronized with the beam, bypassing another operational difficulty with TDS systems.

However, PS diagnostics also introduce new challenges. The induced wakefields depend strongly on beam alignment and transverse beam size, making the measurements sensitive to jitter and machine stability. The wakefield interaction may additionally introduce energy spread growth, especially for strong streaking configurations. Furthermore, the interpretation and reconstruction of temporal profiles often require more involved calibration and modeling than conventional TDS systems. This is because the measured displacement $\Delta x_{\text{PS}}(z)$ depends on both the transverse beam trajectory and the longitudinal beam profile I , as shown in Eq. 2.9. This has led to PS calibration techniques being an active area of research [11, 32]. Despite these limitations, PS systems have become an increasingly important diagnostic tool in modern FEL facilities and ultrafast accelerator experiments.

2.4.3 Bunch Compression Monitors

Many longitudinal diagnostic systems do not necessarily aim to reconstruct the complete LPS distribution of beams. Such diagnostics sacrifice the full information content for robustness and non-invasive operation. One important example is the bunch compression monitor (BCM). BCM systems are generally designed to provide non-invasive, or minimally invasive, estimates of bunch length, allowing continuous observation of compression stability without disturbing the beam.

The operating principles of BCM systems vary between different applications. The most common approaches rely on the radiation emitted by well-compressed bunches. When the bunch length becomes comparable to or shorter than the emitted radiation wavelength, the emitted intensity scales with the square of the number of electrons, enhancing the detected signal. Measurements of radiation produced by the beam can therefore provide indirect information about the bunch length and compression [26]. The radiation required can be produced in a number of ways, but it is common to use the coherent synchrotron radiation (CSR) produced in the bending magnets of bunch compressors. As this radiation is inherently produced in the bunch compression process, the diagnostic can be fully non-invasive and in many FEL facilities, CSR intensity monitors are routinely used as online indicators of compression performance.

Compared to TDS and passive streaking systems, BCM diagnostics generally provide significantly less detailed longitudinal information. While they can often resolve relative compression changes, they do not directly reconstruct the full temporal current profile or LPS distribution of the beam. Furthermore, the relative BCM signals require calibration against more direct diagnostics, such as TDS measurements.

Despite these limitations, BCM systems remain valuable operational diagnostics due to their simplicity and robustness. Further, as they are non-invasive, they are particularly useful for applications such as feedback systems and online monitoring of compression stability. In this thesis they have seen recurrent use as training parameters for machine learning applications designed to reconstruct the LPS from non-invasive diagnostic signals (**Papers III, IV, and VI**).

2.5 Longitudinal Dynamics in Free Electron Lasers

Free-electron lasers (FELs) comprise a family of coherent radiation sources in which relativistic electron beams generate intense synchrotron radiation when traversing periodic magnetic structures known as *undulators* [22, 20]. A fundamental requirement for FEL processes to take place is the interaction between the electron beams and radiation within the undulator. This can be achieved using optical cavities as in oscillator FELs, or in single-pass high-gain FELs. The latter scheme dominates the modern field, and can be subdivided into the self-reliant self-amplified spontaneous emission (SASE) scheme and externally seeded schemes. Since all FELs considered in this thesis operate in the SASE regime, the following discussion focuses on this process.

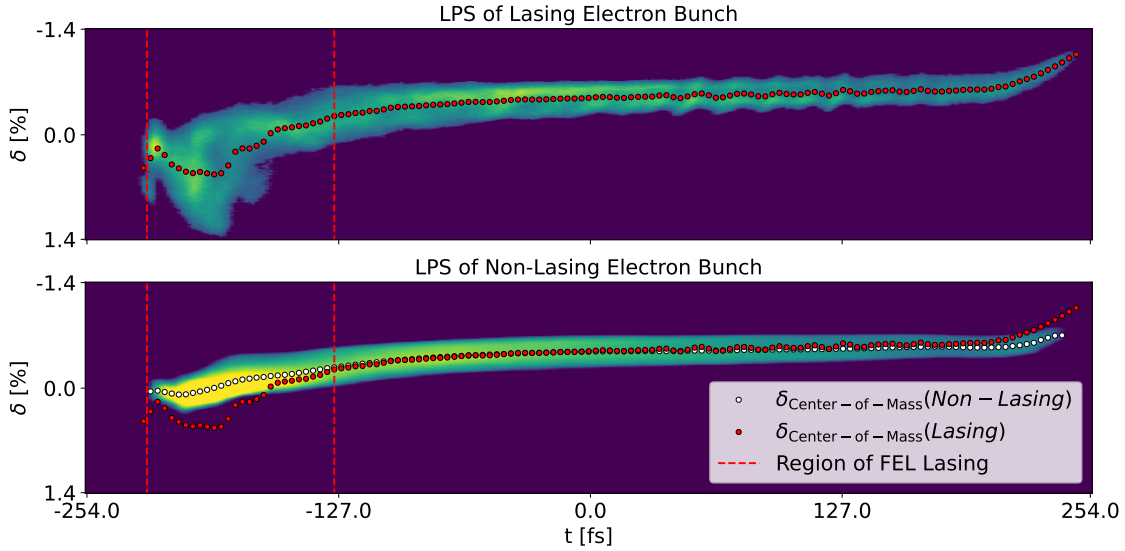


Figure 2.5: Longitudinal phase space measured after active and inactive SASE lasing in the FLASH FEL. The area of FEL radiation production is seen in the region in the head of the bunch, where the energy spread and central energy is affected in a set of slices.

One of the primary motivations for controlling and monitoring the longitudinal phase space distribution of electron beams is the stringent beam requirements for modern FELs. In the undulators of a SASE FEL, the emitted radiation moves through the radiating bunch and interacts with other electrons, producing an energy modulation at the radiation wavelength. Because the electron trajectories in the undulator exhibit longitudinal dispersion ($R_{56} \neq 0$ m), the periodic energy modulation will result in a longitudinal *microbunching* of the beam, which in turn leads to an exponential growth of the radiation power, producing intense, ultrashort photon pulses [1, 20]. This type of self-amplification drives the SASE scheme, though in many alternative FEL schemes the process is instead seeded with an external laser system [20].

The efficiency of the SASE process, or *lasing*, is highly sensitive to the longitudinal beam properties. High peak currents are required to achieve sufficient microbunching (and subsequent gain in radiation power) within the available undulator length, making bunch compression an essential process in FEL facilities. At the same time, sufficiently small slice energy spread is required to achieve resonance between the electrons and the emitted radiation [20]. Distortions in the LPS and excessive energy spread can severely reduce FEL gain and degrade the desired properties of the produced radiation. Consequently, it is of great interest for accelerator facilities driving FELs to maximize peak currents of the electron beams and minimize the degrading processes that appear during beam acceleration, compression, and transport.

An important feature exploited in FEL diagnostics is that the lasing process leaves a measurable imprint on the longitudinal phase space of the electron beam. In the lasing process electrons transfer their kinetic energy to the radiation within the undulators, an energy modulation develops within the bunch, eventually leading to an increased slice energy spread and a reduction in the mean energy in the lasing regions of the bunch. Using LPS diagnostics in the areas downstream of the undulators can allow for indirect longitudinal diagnostics of the photon beam itself through observations of the effects of lasing on the electron beam [33]. The effect is displayed in Fig. 2.5: the region within the red lines in the upper image shows measurably increased energy spread and decreased central energy within the local beam slices under lasing conditions, as compared with the unaffected LPS distribution in the lower image. This example is taken from LPS diagnostics downstream of the undulators in the FLASH FEL.

Given the stringent requirements of FEL facilities for controlling and diagnosing the electron beam longitudinal phase-space distribution, we can see a clear motivation for the investigations in this thesis. The work on compression schemes and virtual LPS diagnostics in this thesis are driven by the need to preserve and optimize beam properties for optimal radiation production and high FEL performance.

Chapter 3

Machine Learning for Accelerators

Machine learning (ML) has been an active field of study for more than seventy years. Defining an exact starting point for a field of study is difficult, but Rosenblatt's 1958 work on the perceptron is commonly regarded as one of the foundational contributions to modern artificial neural networks (ANNs), a key component of the modern ML field [34]. Like most fields, machine learning has also gone through multiple iterations and grown with time, but few other fields have seen such an explosive increase of applications in varied areas. The rapid developments in ML has been driven both by advances in the theoretical understanding of the systems and by the dramatic increase in available computational resources since the middle of the 20th century. Contemporary machine learning has now become a practical tool across a wide range of scientific and industrial applications, as well as everyday life [35].

Traditionally, accelerator physics has relied on analytical modeling and detailed numerical simulations to describe beam dynamics and optimize the performance of modern facilities. Also, accelerators have always generated large amounts of detailed and complex data from diagnostics and machine control systems. Underlying much of this data are highly nonlinear beam dynamics and coupled systems inherent to accelerator physics. This information can be difficult to model, measure and understand by conventional methods. These factors have made accelerator science and controls a particularly suitable field for the development of ML applications [36].

In the current state of accelerator development, ML is being explored for a

broad range of applications, including machine tuning and control, surrogate modeling, anomaly detection and virtual diagnostics [36]. While ML has seen experimental use in accelerators since the 1990’s [37, 38, 39], it has risen to far more prominence in the last ten years [36]. The work in this thesis has mainly focused on the growing need for applications in high-demand linac facilities, such as modern FELs [1]. These facilities have especially strict stability requirements and a need for detailed and extensive beam diagnostics [20]. ML can assist in both areas, with its capability to learn complex correlations directly from operational machine data, enabling both rapid predictions of diagnostic signals and detailed control of machine parameters.

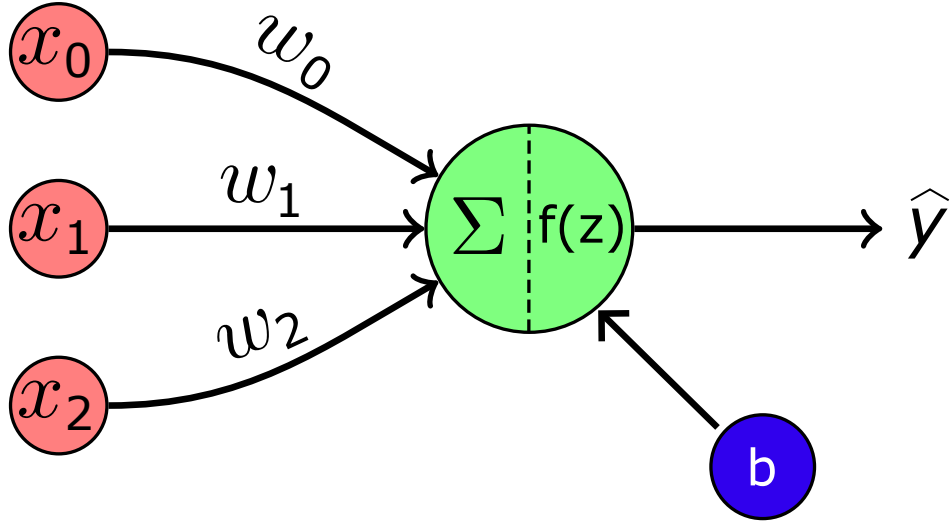
In this work, ML methods are primarily employed in two areas. First, artificial neural networks are used to construct virtual diagnostics and surrogate models capable of predicting longitudinal beam information from non-destructive observables. Second, Bayesian optimization is explored as a method for automated optimization of key accelerator parameters. The following sections introduce the fundamental ML concepts required for these applications, followed by an overview of virtual diagnostics and Bayesian optimization methods in accelerator physics.

3.1 Machine Learning Fundamentals

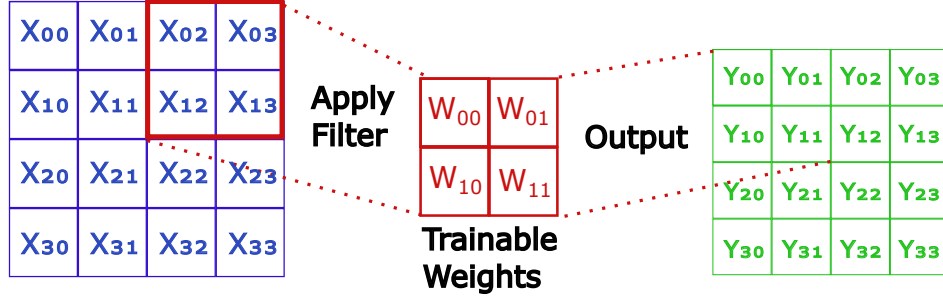
Within the wide range of machine learning systems and applications, a powerful method commonly applied in applications within physics and accelerators is artificial neural networks (ANNs), systems trained to recognize patterns in data using a method called supervised learning. The following sections cover the fundamentals of these ML systems and methods.

3.1.1 Artificial Neural Networks

The fundamental building block which constructs an artificial neural network is the *perceptron*, initially introduced by Rosenblatt in 1958 [34]. In the context of ANNs these are also referred to as *nodes* or *neurons*, and a simplified model is depicted in Fig. 3.1a. Inspired by biological neural systems, the goal of the perceptron is to construct a mathematical system capable of adapting to inputs. This is done through iterative optimization of *weights*, multiplicative and additive factors in the operations applied to the inputs. A single perceptron’s sum z ,



(a) A perceptron, the fundamental building block of an ANN, which combines weighted inputs $w_i x_i$ and a bias b before applying an activation function f and returning an output $\hat{y}$.



(b) A CNN layer, with the input matrix on the left, followed by the convolutional filter of trainable weights w_{ij} in the middle, and the resulting output on the right.

Figure 3.1: The fundamental structure used to construct the ANNs used in the thesis projects: (a) Perceptrons and (b) Convolutional Neural Networks.

given n inputs x_i , can be expressed as

$$z = \sum_{i=0}^n w_i x_i + b, \quad (3.1)$$

where the critical components for iterative optimization of the output can be seen in the multiplicative weights w_i , separate for each input channel, and the bias b . This can also be written in vector form as

$$z = \mathbf{w}^T \mathbf{x} + b. \quad (3.2)$$

For a complete perceptron, it is common that an *activation function*, $\hat{y} = f(z)$, is applied before the output is passed along. A very common activation function is the rectified linear unit (ReLU), returning $\hat{y} = z$ if $z > 0$ and 0 otherwise.

Other commonly used activation functions include simple linear functions and the hyperbolic tangent function $\hat{y} = \tanh(z)$. In supervised learning, it is assumed that the ground truth y of $\hat{y}$ is known, and a *loss function*, $L(y, \hat{y})$, can then be applied and further used to update the weights in the perceptron. Common loss functions include the mean squared error (MSE) or mean absolute error (MAE) for learning tasks focused on regression, and cross-entropy loss functions for tasks focused on classification [35]. Once a loss has been defined for one pass through the node, it is then back-propagated to the trainable weights and biases, which are updated using a gradient-based optimization method. Common optimization algorithms for the task of optimizing weights and biases include stochastic gradient descent [35], and the Adam optimizer [40]. With the perceptron updated, inputs can be reapplied and a new $\hat{y}$ calculated, which can then be used to re-optimize the perceptron once more. This iterative optimization of weights and biases is what enables ANNs to learn patterns and mappings in data.

While a single perceptron can only represent relatively simple relationships between inputs and outputs, the true strength of the ANN structure emerges when many perceptrons are connected together in multiple layers. The output of one layer is used as the input to the next, allowing the network to learn increasingly complex and nonlinear mappings by training not only the weights on given inputs, but also the weights internally connecting the nodes of the network. The first layer receives the external inputs and is referred to as the input layer, while the final layer produces the network prediction and is known as the output layer. Layers between these are commonly referred to as hidden layers [35]. These different layers of an ANN can be trained in the same manner of loss back-propagation as outlined above, but will now be capable of more complex pattern-recognition.

In modern ANN applications, many different types of individual nodes and layers have been developed from the basic perceptron. Nodes and layers can have differing complexity, can be focused on different use-cases, and can be organized into different *architectures*. ANN architectures that consist solely of perceptrons can be called *multi-layer perceptrons* or *densely connected networks*, due to their high amount of weighted connections when compared with other types of nodes [35]. Other modern architectures include recurrent neural networks which have been developed to work with sequential data, e.g. classifying and predicting new values in time-series. There are encoder-decoder networks, such as the modern U-Net architecture [41], which extract key parameters in inputs through data compression. Further examples are graph neural networks for irregular data structures and transformers for language processing. Another

key example is convolutional neural networks, which have become particularly common for accelerator virtual diagnostics owing to their ability to efficiently reconstruct beam images.

3.1.2 Convolutional Neural Networks

Networks consisting solely of perceptron layers, referred to here as multi-layer perceptrons (MLPs), are highly flexible and capable of approximating complex functions, but they can quickly become computationally inefficient when the task involves high-dimensional inputs and/or outputs [35]. Since each node in each layer is connected to every node in the preceding layer, the number of trainable parameters grows rapidly with the dimensions of the data involved and the depth of the architecture, which in turn increases the computational demands on training the network. Further, MLPs will treat every input value independently, and thus possess no inherent information of the spatial arrangement of data. A prime example in accelerator diagnostics where both data size and spatial distribution become key issues is imaging screens, where the resulting images can often contain hundreds of thousands or even millions of individual data points (pixels) and the distribution is critical, making purely MLP-based applications inefficient.

Convolutional neural networks (CNNs) were developed specifically to address the limitations of MLPs by exploiting the spatial structures present in image data [42]. Instead of the densely connected structure of MLPs, CNNs apply small trainable filters, referred to as convolution kernels, across local regions of a given image. These small filters then move across the input image with a given stride, allowing a relatively small set of trainable parameters to be applied across the entire image. The principle of a single CNN kernel is depicted in Fig. 3.1b. By stacking multiple convolutional layers, which can each consist of many kernels, CNNs can learn increasingly complex image representations, progressing from simple features such as edges and intensity gradients to more abstract structures relevant to the target prediction task [35], i.e. beam structures in images from screens in an accelerator. The CNN is the dominant ANN architecture for accelerator applications involving images, beam profile measurements, and virtual diagnostics [36, 43]. Several of the ML models presented in this thesis make use of CNN architectures to extract key information from beam diagnostics and machine parameters.

3.1.3 Learning Schemes and Datasets

The machine learning algorithms used to train ANNs are commonly divided into three broad categories: *supervised*, *unsupervised*, and *reinforcement learning* [35]. In supervised learning, the desired output y of the ANN must be known during training, often referred to as the *ground truth* or *labels*. This allows for iterative optimization of the network parameters by minimizing the loss function $L(y, \hat{y})$, in each iteration comparing the predicted output $\hat{y}$ with the ground truth. Through this process the weights in an ANN are adapted to learn true mappings between input and output which can then be applied to new data. In contrast, unsupervised learning operates without labeled outputs or ground truth values, instead optimizing a network to find internal underlying structures and correlations within the input data itself. Typical tasks for unsupervised ANNs include data-focused tasks such as clustering of similar observations, detecting anomalous behavior, and reducing the dimensionality of complex datasets to reveal dominant features and correlations present in the data. Reinforcement learning differs from both by training an agent to make decisions through interactions within an external environment, where an objective is set up as a cumulative reward function rather than a loss in reproducing sought outputs. Reinforcement learning therefore differs fundamentally from supervised and unsupervised learning by learning through interaction with an environment rather than from a fixed dataset. While all three paradigms have found applications within the field of accelerators, the work presented here is based almost exclusively on supervised learning, using labeled datasets with known ground truths to optimize networks through loss functions.

A critical step in applying supervised learning is the separation of available data into independent *training*, *validation* and *test* datasets. The training data is used directly in the network optimization, i.e. in the iterative updating of weights and biases inside the network. The validation dataset is used during training only for performance evaluation and hyperparameter selection, and is not used to directly update the network weights. The training data is commonly provided in batches, subsets of the total data, and once all batches have been provided, it is referred to as a full *epoch*. At the end of a training epoch, evaluation on the validation data is usually done to estimate the network’s ability to generalize to unseen data. Finally, the test data is reserved for the final evaluation of the network. This set should consist of data not seen during all training of the model, a separation which is essential to obtain an unbiased estimate of the network’s capability to generalize to new data [35].

The quality and properties of the input and output data used in supervised

learning is key. Prior to training, preprocessing of the data used as inputs and outputs to ANNs is necessary. In the applications developed for this thesis, inputs commonly consisted of vectors of beam information, and outputs often consisted of the full images from scintillating screens. Input channels of this type are commonly normalized to comparable numerical scales, often within the interval $[0,1]$, or standardized to zero mean and unit variance. For the images used, additional preprocessing steps included background subtraction, cropping, resizing, and intensity normalization.

One of the primary challenges in supervised learning is *overfitting*. An overfitted model learns not only the critical mapping present between input and output in the training data, but also the statistical fluctuations and noise specific to that dataset. Such a model may achieve excellent performance on training data while performing poorly in validation and testing. The separation of data into training, validation, and test datasets is therefore essential for detecting and mitigating overfitting, alongside techniques such as regularization, dropout, and early stopping [35].

3.2 Machine Learning Applications in Accelerators

Machine learning has become one of the fastest growing research areas within accelerator physics over the last decade [36]. Modern accelerator facilities generate large volumes of diagnostic and control data while simultaneously exhibiting highly nonlinear and coupled beam dynamics. These characteristics make many traditional accelerator problems well suited to data-driven methods capable of learning complex relationships directly from operational measurements.

Current ML applications in accelerators can be divided into several categories. One of the largest areas is that of *virtual diagnostics*, where ML models are trained to infer beam properties that are difficult to measure continuously. Closely related are *surrogate models*, which approximate accelerator simulations to enable rapid retrieval of physically motivated beam approximations. Early work in the latest wave of ML in accelerators focused on surrogate models for the facilities FACET and LCLS [44]. Machine learning is also increasingly used for automated machine tuning, beam matching and online optimization, often in conjunction with ML-driven methods like Bayesian optimization or reinforcement learning. Other active research areas include anomaly detection for identifying abnormal machine behavior [45], and optimization of accelerator design through fast surrogate-assisted parameter studies [36], as well as the use of language models for summarizing and analyzing facility logbooks and

archives [46, 47].

Many different machine learning architectures have been explored for the varied accelerator applications. Multi-layer perceptrons remain a common choice for problems involving low-dimensional machine parameter inputs and scalar beam quantities as output, due to the MLPs flexibility and computational efficiency. Accelerator diagnostics frequently rely on imaging systems, making CNNs particularly well suited for extracting information from beam images, or in reconstruction tasks. In recent years, encoder-decoder structures, such as the U-Net example, have gained prominence in image-based tasks, owing to their capacity for simultaneously working with compressed, critical information and maintaining high-resolution features [41]. While there is an active drive to implement more of the ever-evolving ML techniques within accelerators, the majority of practical applications remain reliant on supervised learning and the labeled datasets required for such schemes.

Many challenges remain for the continuing integration of ML into accelerator physics and controls. The high-quality labeled datasets required for supervised learning are expensive to acquire in modern facilities, where beamtime is in high demand. The necessary invasive beam diagnostics and expertise in data acquisition also limits the possibilities for setting up such methods. Exacerbating this issue is the fact that accelerators are inherently dynamic systems, with beam conditions and machine states drifting over time, causing the mappings learned by ANNs to gradually become outdated. Maintaining the performance of ML applications remains an active area of research, with approaches including periodic retraining, transfer learning, online adaptation and physics-informed methods [36]. Furthermore, practical deployment requires reliable uncertainty estimation and improved interpretability of model predictions, particularly for applications involving machine protection or automated control [48].

The work presented in this thesis contributes primarily to the areas of virtual diagnostics, surrogate modeling and Bayesian optimization for accelerators. The machine learning methods developed here focus on reconstructing longitudinal beam properties from non-destructive measurements and on applying these models to accelerator control and optimization tasks. The following sections therefore discuss these applications in greater detail.

3.2.1 Virtual Diagnostics

In modern accelerator, many of the most critical beam diagnostic systems are either invasive, destructive, expensive, or difficult to operate continuously. High-

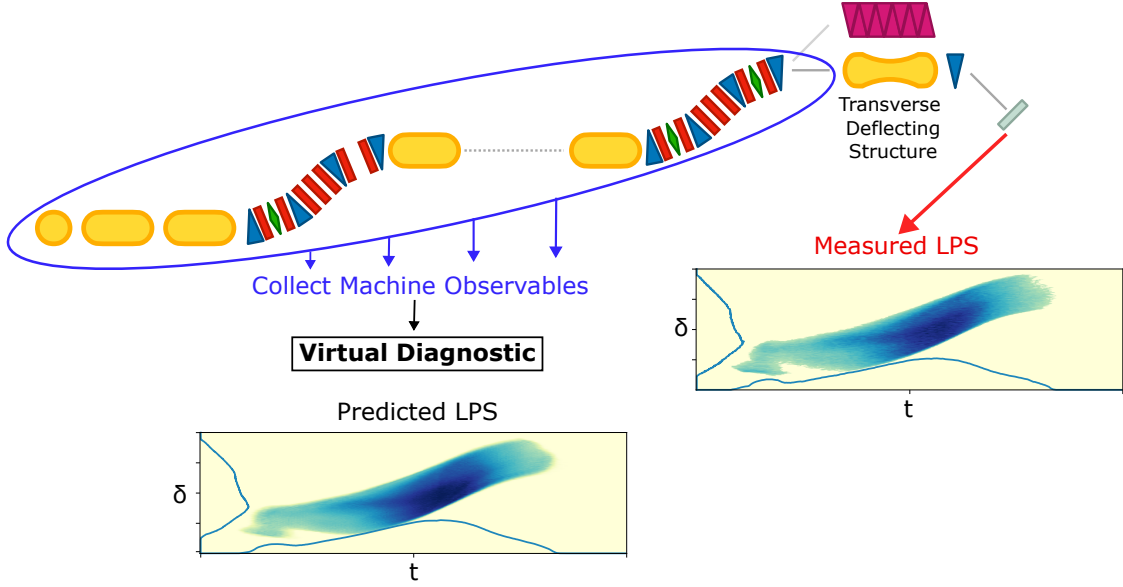


Figure 3.2: Principle workflow of a virtual diagnostic. A real, destructive measurement (example from the MAX IV TDS system on the right) is reconstructed using information from other, non-destructive machine observables (example from the MAX IV VD on the left).

resolution longitudinal diagnostics, such as a transverse deflecting structure (TDS) system, may use scintillating screens to image beam distributions in the LPS destructively (see Sec. 2.4). However, there is also a large amount of non-destructive beam information generated constantly by other beam diagnostics and machine components, which will often include information correlated with the beam’s longitudinal properties, although in nonlinear and obscure ways. The central idea of a virtual diagnostic (VD) is to use machine learning to probe such correlations, training an ANN to construct a mapping between readily available observables and beam properties which are more difficult to measure, thereby providing a fast, ML-based approximation of a physical diagnostic system [44, 36].

The concept of virtual diagnostics emerged during the latest wave of machine learning in accelerator physics and has become one of the field’s most active areas [36, 44, 48]. A VD can be constructed using supervised learning. In this scheme, a training dataset is required including both the invasive measurements to be predicted, and a set of non-destructive *machine observables* recorded simultaneously. The non-destructive beam information, or machine observables, forms the input to the VD, while the invasive diagnostic measurements serve as ground truth for training. Once trained, the VD can infer the invasive diagnostic information using only non-destructive inputs, and can in principle replace or supplement the physical diagnostic during operation. This working principle is illustrated in Fig. 3.2, with the implemented VD workflow and resulting predic-

tion on the left and a measured LPS on the right.

While virtual diagnostics can significantly reduce operational complexity and enable continuous beam monitoring, their performance ultimately depends on the quality and quantity of the training data. A trained VD will also be limited to the parameter space explored during initial data collections, with limited native ability to extrapolate outside provided datasets. Maintaining the stability and reliability of VDs in control room operation, where accelerator conditions drift and machine configurations change over time, remains an active field of study [36, 48].

3.2.2 Surrogate Models

Within machine learning for accelerators, a closely related system to virtual diagnostics are *surrogate models*. While a virtual diagnostic seeks to reproduce the output of a physical measurement system, a surrogate model aims to approximate the underlying behavior of the accelerator itself. This can be done in congruence with physical measurements, calculating underlying properties from the detected signals, but surrogate models are more commonly trained using data generated from detailed accelerator simulations [36]. Modern particle tracking codes can accurately model beam dynamics and the more complex collective effects that drive the development of beam parameters, but such simulations can be demanding, both in terms of time and computational power. Tasks for simulators such as optimization and beam matching may require thousands or millions of evaluations, making the use of high-fidelity simulations with large amounts of particles impractical. By training ANNs or other machine learning models on data from such detailed simulations, a computationally inexpensive surrogate can be constructed. Once trained, such surrogate models can evaluate key beam parameters and distributions orders of magnitude faster than the original simulations [44].

The distinction between surrogate models and virtual diagnostics lies primarily in their intended use. VDs are made to supplement a measurement system by predicting diagnostic outputs, whereas surrogate models replace a physics model by predicting accelerator behavior and beam dynamics [14]. Both approaches rely on supervised learning of correlations from data and often employ similar machine learning architectures, but they serve different roles for accelerator operations.

3.2.3 Advantages and Limitations

Virtual diagnostics can offer several advantages for accelerator operation. By replacing invasive measurements with inferred values from ML applications, continuous monitoring of beam properties otherwise inaccessible becomes possible. VDs can reduce the operational complexity of modern accelerators and provide critical new beam information. A trained VD is also computationally inexpensive to use, which enables real time queries of the inferred values. The computational lightness of these pre-trained ANNs is also what enables the surrogate models focused on replacing far more computationally expensive numerical simulations.

There are, however, important limitations to these applications. Unlike physical measurements of the beam, the predictions of the most basic VDs are not based on real time measurements of the parameters of interest, but on historical correlations learned from collected datasets. Consequently, their performance is inevitably strongly tied to these datasets, their quality and quantity, how well they represent the changing state of the machine [35, 36].

Accelerators are inherently dynamic systems. They will always exhibit drifts in beam parameters and state, due to underlying changes in the machines surrounding environment and in the beam itself. These drifts will gradually alter the mapping learned by models, reducing their accuracy over time. There are strategies which can be implemented to maintain long-term performance, and techniques such as periodic retraining, adaptive learning or transfer learning from simulations remain active areas of research [36].

Finally, ML applications such as virtual diagnostics and surrogate models should at this time mostly be viewed as complementary systems to physical detectors and numerical systems, rather than complete replacements. While VDs can be used to significantly reduce the need for invasive measurements, periodic data collection from physical diagnostics remains necessary to ensure that the learned mappings remain representative of the current machine state.

3.3 Bayesian Optimization

Many scientific and engineering problems can be easily formulated as tasks for optimization algorithms, where a set of controllable parameters is adjusted to maximize or minimize specific objectives or performance metrics. Common examples include optimizing designs in engineering systems, tuning hyperparamet-

ers in machine learning models [49], or maximizing the performance of complex physical experiments. While many conventional optimization methods can perform well when objective functions are inexpensive to evaluate or their gradients are readily available, many real-world problems involve expensive and noisy objectives, which are costly to measure reliably. In such cases, large-scale parameter scans and traditional optimizations rapidly become operationally impractical.

Bayesian optimization (BO) is a ML-assisted optimization technique for problems where evaluations of objectives are expensive and noisy [50, 51, 52]. Rather than sampling the parameter space blindly, BO uses a probabilistic model, commonly based on Gaussian processes, to construct an approximate function of the objectives response to the controlling parameters. A model based on Gaussian processes yields both predictions of the objective, as well an associated uncertainty estimate. BO then uses this model to determine which region in the parameter space to explore further. This decision for queries is controlled by an *acquisition function*, which should be tuned to balance between exploration and exploitation, i.e. between exploring new regions of the response function or exploiting the discovered regions to find the actual optimum. A common acquisition function is the upper confidence bound (UCB), which incorporates both the predicted objective value and the estimated uncertainty from the Gaussian process when selecting subsequent evaluations [53].

In accelerator physics, many tuning tasks can be formulated as optimization problems in which machine parameters are adjusted to maximize or minimize beam properties or facility performance metrics. Examples include minimizing transverse emittance [54, 55], maximizing FEL pulse energy [56], optimizing beam transport or producing desired longitudinal properties (**Paper V**). These objectives are typically evaluated using beam diagnostics, which may be invasive, noisy, and time-consuming to operate. Furthermore, the parameter spaces for accelerators are often high-dimensional and inherently nonlinear. Bayesian optimization is therefore particularly attractive in the field due to its information efficiency, as it can identify high-performing operating points using significantly fewer measurements than conventional scan-based approaches and traditional, gradient-based optimizations. Since early demonstrations of BO’s use in live tuning [51], BO has become one of the most widely used machine learning techniques for online accelerator optimization [36]. Despite its effectiveness, BO performance can degrade as the dimensionality of tasks increase. Another drawback, which is more critical to accelerators, is the underlying assumption that the response function will not change during tuning, while accelerators can often drift over time, which requires some mitigation in the optimization procedure.

Chapter 4

Development of Virtual Diagnostics

Reliable knowledge of the longitudinal properties of accelerated electron beams is essential for the operation and optimization of modern linear accelerators and free-electron lasers. However, direct measurements of the longitudinal phase space (LPS) of high-energy electron beams are typically destructive, time-consuming, or require dedicated diagnostics that cannot be used continuously during machine operation. This motivates the development of *virtual diagnostics* (VDs), where machine learning models use non-invasive accelerator observables to reconstruct beam properties that would otherwise require invasive measurements. We have investigated the feasibility of such ML methods through the development of operational VD systems for LPS reconstruction and validated the method across three accelerator facilities (**Papers I–III**).

The studies of VDs represent a gradual increase in both complexity and scope. An initial step was to establish the feasibility of reconstructing beam LPS from non-invasive machine observables using both a surrogate model trained on *elegant* simulations of the future transverse deflecting structure (TDS) system [57] and a VD trained on measurements from a scintillating screen in a dispersive section of the MAX IV bunch compressors (**Paper I**)¹. This approach was extended to the operational MAX IV TDS, comparing surrogate-model predictions and VD reconstructions using measured TDS data (**Paper II**)¹. Building on

¹The terminology used in this thesis differs slightly from that of **Papers I** and **II**, in which models trained on simulated data were also referred to as "virtual diagnostics". Since then the terminology regarding ML applications in accelerator physics has evolved and this thesis consistently adopts the distinction established in Sec. 3.2.

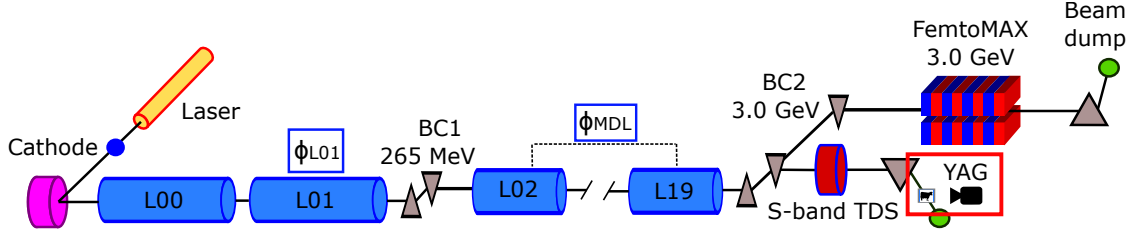


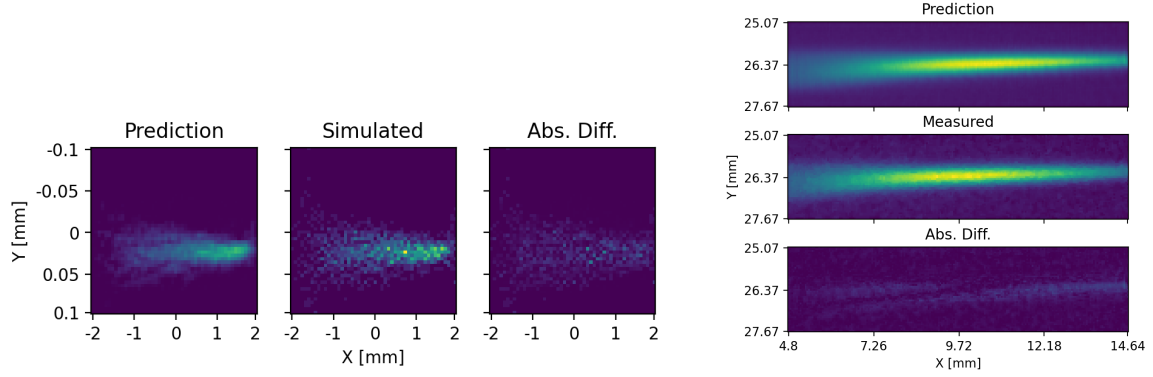
Figure 4.1: Simplified layout of the MAX IV linear accelerator. Highlighted in blue rectangles are the main controlling parameters for longitudinal control of beams, the RF phases before the two stages of bunch compression. Marked by a red rectangle is the point of measurement in the TDS system, where the LPS distribution is imaged on a screen.

the successful implementation at MAX IV, the methodology was subsequently applied to measurements from the FERMI FEL in Italy and SwissFEL at PSI in Switzerland, resulting in a large-scale multi-facility study (**Paper III**).

4.1 MAX IV Virtual Diagnostic

The MAX IV TDS was designed for a ~ 1 fs temporal resolution and is housed in a dedicated beamline within the short pulse facility (SPF) which also accommodates the experimental beamline FemtoMax [58, 59]. A simplified layout of the MAX IV linear accelerator is shown in Fig. 4.1. The inherent invasive nature of TDS measurements has been covered in Sec. 2.4, but here it is exacerbated by the additional operational difficulties of switching between the two separate beamlines leaving the second bunch compressor. The FemtoMAX beamline has stringent and varying requirements on the LPS of the electron beam for their experimental applications. These limitations and requirements are the underlying motivation for investigating the feasibility of a VD for the MAX IV TDS system.

The acceleration in the MAX IV linac before the SPF can be considered as taking place in two major stages, each before a bunch compressor [60]. Electrons leave a photocathode gun and are first accelerated to 265 MeV in the L00 and L01 sections, before the beam enters the first arc-like bunch compressor, BC1. After BC1 follows the main linac where 18 accelerating sections raise the energy to the final 3 GeV. The klystrons supplying the power to the main linac are all fed by a common main drive line (MDL), which enables centralized control of the RF properties across the entire section. The 3 GeV beam then enters the second bunch compressor, BC2, and can then exit into either FemtoMax or the TDS diagnostic beamline [58]. In the TDS beamline, a final spectrometer dipole deflects the beam onto imaging screens, providing the full LPS information. There are three different branches downstream of the dipole, each equipped



(a) Example of the surrogate model's output based on simulations of the TDS system. (b) VD prediction of an image from a dispersive screen in the second bunch compressor of the MAX IV linac.

Figure 4.2: Example of results from **Paper I**, with (a) showing predictions of a surrogate model trained on simulations, and (b) displaying VD results based on a dispersive screen measurement. In both figures the subfigures show the ML prediction, the real or simulated results, and the absolute difference between the two images.

with multiple scintillating screens, allowing for control of the levels of dispersion and streaking strength on the screen during data acquisition. Throughout the development of VDs for this thesis, data collections have mainly consisted of scans of the RF parameters before the two compressors, since these are strongly correlated with the final LPS distribution, affecting energy spread, bunch length and energy chirp.

4.1.1 Simulated TDS and Dispersive-Screen Reconstruction

As the MAX IV TDS system was not yet available during the initial stages of this thesis work, the first studies focused on whether ML-based virtual diagnostics could be deployed on simulated LPS results and existing machine diagnostics. These studies served mainly as an initial proof of principle for later virtual diagnostics based on TDS data.

The results presented in **Paper I** can be divided into two main parts: one based on simulations and another based on machine measurements from diagnostics other than the TDS. The simulations consisted of particle tracking in the *elegant* framework, with a model of the future TDS system included [57, 58, 59]. For the measurements, the TDS could not yet be used for full LPS investigations, so a scintillating screen in the dispersive region of the second bunch compressor, BC2, was used in the data acquisition. If one assumes a completely linearized LPS, the distribution on the dispersive screen should also be linearly dependent on the longitudinal distribution of the beam. If this assumption holds, the

dispersive-screen reconstruction provides a useful experimental test of whether non-invasive machine observables contain sufficient information to infer longitudinally relevant beam properties.

In both cases, simulation and measurement, the same ANN architecture was used as the basis for the surrogate model and virtual diagnostic for predicting images. This ANN consisted of three pure multi-layer perceptron (MLP) layers, and three additional layers of CNN filters, see Sec. 3 for more details. At the center of the network, after a final MLP layer, the data is restructured into matrices for use in the CNN part. The addition of CNN layers makes the network more suitable for particle distribution regression, as it allows the ANN to learn the spatial relations in the outputs, as compared with a pure MLP architecture [42]. More complex architectures, including convolutional encoder-decoder networks, conditional variational autoencoders [61] or generative adversarial networks [35], could provide greater capacity for the ANN, but would introduce additional model complexity that was not required for the image reconstruction task investigated here. Convolutional encoder-decoder structures are utilized in more advanced projects later in this thesis (see Sec. 5.1). The training utilized the Adam optimizer [40] along with a MAE loss function. Throughout the network, the ReLU activation function was applied.

The study of a surrogate model on the simulated data showed that the ANN was capable of reconstructing the TDS images over a range of parameters. The data for this application originated from 1000 6D simulations, spanning randomized RF parameters in L01 and the MDL, with the final beam distributions binned into images for training. In the simulated environment, the TDS was actively streaking, but the downstream dipole used to resolve the energy distribution was deactivated, which simplified the learning task of the ANN. A representative prediction on these images is shown in Fig. 4.2a. On a 100 image test dataset, the model achieved an RMS error of 12.5%, indicating that the surrogate model could reconstruct the images with good fidelity.

Experimental validation of the promising simulated results was carried out using measurements from a dispersive screen in BC2. Since the TDS was not yet operational, these measurements provided the closest available experimental proxy for longitudinal beam properties. A dataset of 1150 paired measurements of machine observables and dispersive-screen images was acquired by varying the RF settings controlling longitudinal compression, such as the RF phases and the filltime in the SLED systems (the time spent by the RF pulses in the SLED RF pulse compressors [62]). The raw camera images were cropped surrounding the beam center-of-mass. An example from the predictions on the test dataset is shown in Fig. 4.2b, and across this dataset an RMS error of 17.3% was achieved.

This shows a slight degradation in performance as compared with the surrogate model, but it nevertheless proved that readily accessible machine observables provided sufficient information for longitudinal predictions.

For the measured data, a separate simplified network using only MLP layers was also developed to predict the corresponding beam center position on the measuring screen. This allowed a VD utilizing both the image and the center-of-mass ANNs to predict both energy distribution and the central energy of the beam. This can naturally be useful, but for the later VD projects in this thesis a decision was made to focus on distribution predictions, forgoing the center-of-mass information, as the distributions in t and δ are more commonly of interest in a TDS measurement.

The initial results in this study were promising, showing that the machine observables in the MAX IV linac were sufficient for longitudinal reconstructions of the beam, both in simulation and measurement. With this, the proof-of-principle was complete and it clearly motivated the more detailed studies of full LPS VDs to be carried out for **Papers II** and **III**.

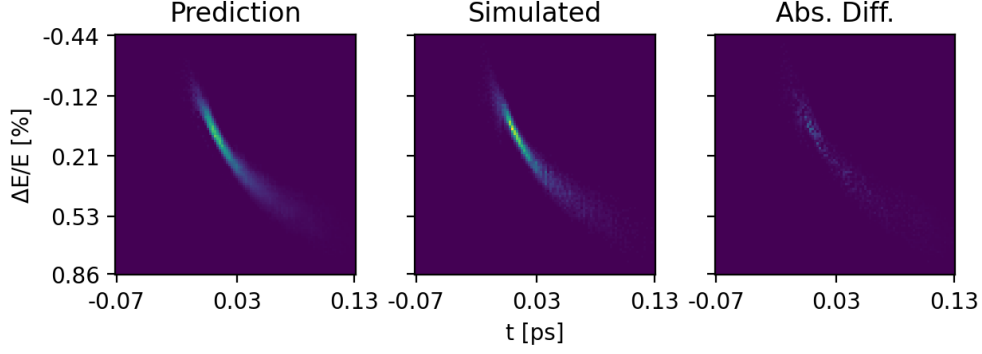
4.1.2 First TDS Measurements and Full LPS Surrogate Model

The initial proof-of-principle established that non-invasive observables contained sufficient information for longitudinal reconstructions. The next step was to test whether the same general framework could be extended to predict full LPS images using a surrogate model on the simulated MAX IV TDS data, and to test the VD on the first measurements from the operational TDS system. In both cases, the aim was now to assess whether the approach could produce quantitatively useful predictions under more realistic conditions.

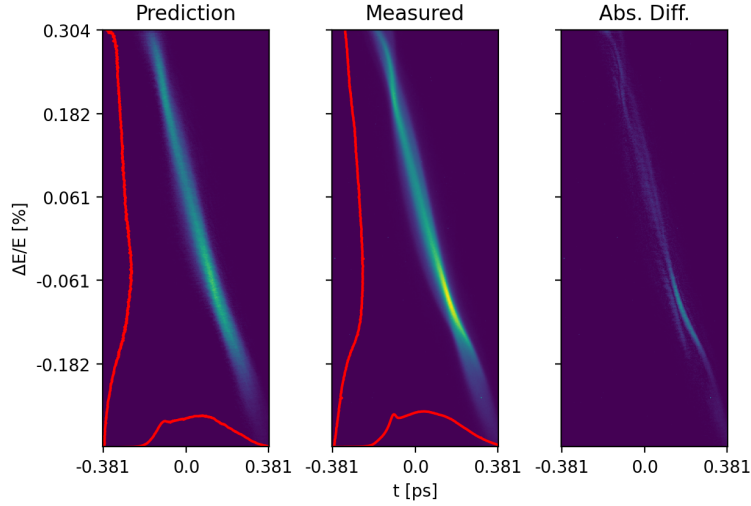
The ANN architecture and training scheme was in principle maintained from the previous study, utilizing a structure combining MLP and CNN layers, and training using Adam with the MAE loss. For a more robust and representative final metric on image predictions, the *coefficient of determination* was introduced, defined as

$$R^2 = 1 - \frac{\sum(y - \hat{y})^2}{\sum(y - \bar{y})^2}, \quad (4.1)$$

wherein y is the ground truth, $\hat{y}$ is the predicted value and $\bar{y}$ is the mean of the ground truth, with sums performed over the full image [44]. This will remain as a recurring metric for image similarity throughout the thesis, since it provides a simple and fast measure of how well a predicted structure reproduces a target distribution.



(a) Example of the surrogate model's output (left) based on simulations of the TDS system, along with the *elegant* LPS distribution (center) and their absolute difference (right).

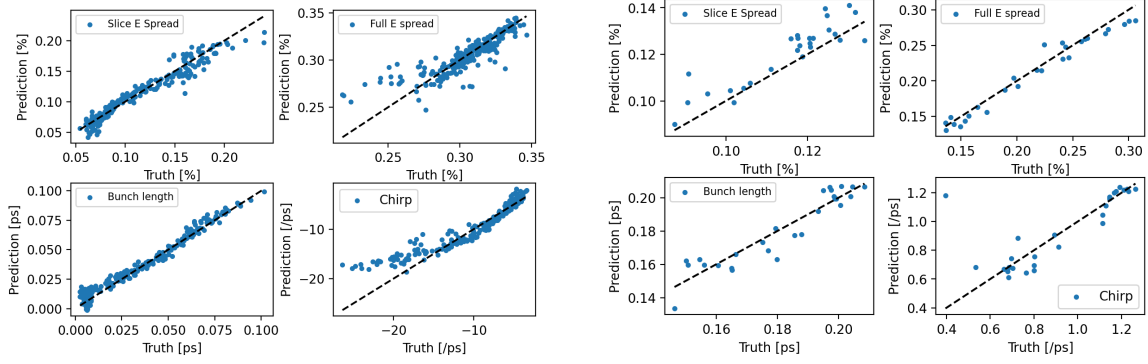


(b) First VD prediction (left) of a MAX IV TDS measurement (center), along with absolute difference (right).

Figure 4.3: Example of results from **Paper II**, including both a surrogate model of *elegant* simulations and first results from the MAX IV TDS VD.

The VDs and surrogate models trained to perform full image reconstructions were computationally heavy, specifically in training of the last MLP layers which represented the full image. Full LPS information may not always be necessary for tuning and control of LPS distributions, but rather key LPS beam parameters may be sufficient. For this reason, a second approach was introduced, where bunch length, slice and full energy spread, and energy chirp was extracted from the full LPS distributions before use in ANN training. With these four values as the new ground truth, a simplified network of two MLP layers could be used for VD and surrogate model applications.

Although experimental data remained limited during commissioning, the expanded simulated dataset allowed the proposed approach to be evaluated under realistic TDS conditions. This was produced using *elegant*, now using the spec-



(a) The TDS surrogate model's beam parameter predictions in correlation plots with the simulated ground truth values. (b) The TDS virtual diagnostic beam parameter predictions in correlation plots with the measured ground truth values.

Figure 4.4: Predictions of beam parameters from the surrogate model and virtual diagnostic, presented in correlation plots with the ground truth. The parameters are the bunch length, slice and full energy spread, and energy chirp.

trometer dipole so the full LPS was imaged. Once more, randomized settings to RF phases and voltages were utilized in L01 and the MDL, resulting in a total set of 2500 simulated 100×100 pixel images of the measurements from the TDS. An example of the image-based surrogate model's predictions can be seen in Fig. 4.3a, whilst correlation plots of the beam parameter predictions on these simulations can be seen in Fig. 4.4a. The image predictions reached a mean R^2 of 89.1% across the 250 image test dataset, which indicates that the dominant features of the distributions were well reconstructed.

The first measured TDS dataset was then used to assess whether the surrogate-model performance could be reproduced on experimental data. To vary compression, the RF phase in the MDL was scanned over a range of $\pm 5^\circ$ around its nominal operating point. For each machine setting, a TDS image was recorded from the TDS dispersive screen, together with the relevant non-invasive machine observables, including readings from 36 BPMs, measurements from two beam current transformers, and relevant RF parameters. Following data collection, the TDS images were processed through background subtraction and cropped to isolate the beam distribution. In total a dataset of 106 paired measurements of machine observables and TDS images was collected. An example from a five image test dataset can be seen in Fig. 4.3b, and the image-based VD reached a mean R^2 -score of 88.1%, a good indication that the model was able to reproduce the main features of the measured LPS, despite the limited training data. Correlation plots for the VD predicting beam parameters can be seen in Fig. 4.4b, showing good levels of correlation between predictions and ground truth. This suggests an expanded application of the beam parameter VD should also be of interest.

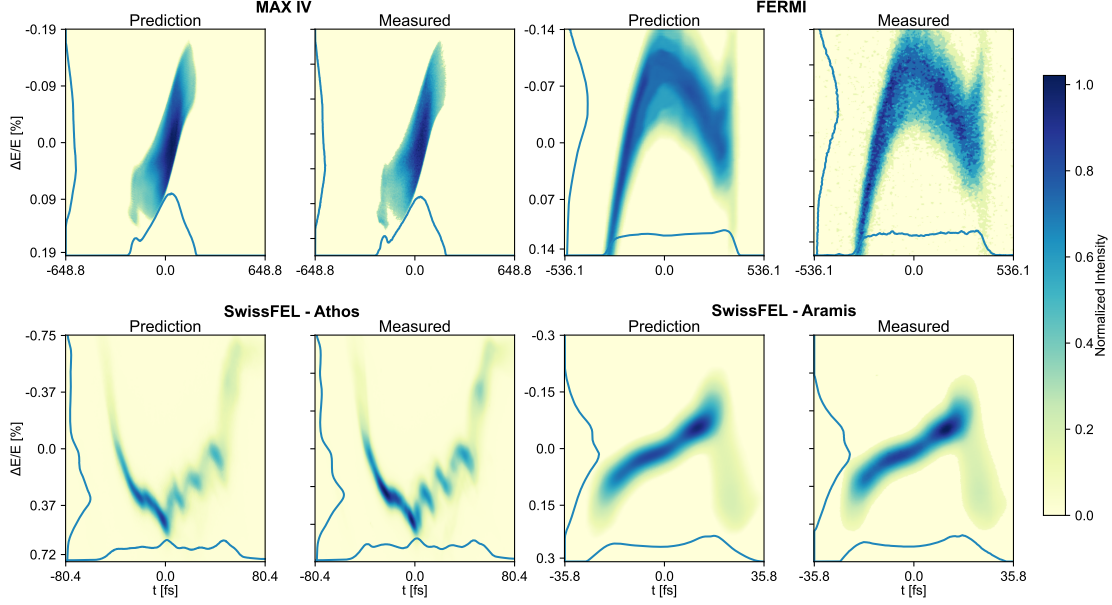


Figure 4.5: From **Paper III**: LPS distributions from MAX IV, FERMI and the two SwissFEL beamlines, as predicted by the LPS virtual diagnostics and measured by TDS systems.

Even with the limited experimental dataset, this represented the first operational virtual diagnostic for LPS predictions at MAX IV. Together, the surrogate model and experimental results demonstrated that the proposed methodology was sufficiently robust for deployment on larger datasets and more realistic operating conditions.

4.1.3 MAX IV Virtual Diagnostic for a Multi-Facility Project

Having established the feasibility of LPS virtual diagnostics, the next challenge was to determine whether they could achieve the robustness and accuracy required for routine operation at MAX IV. In order to further the development towards an operationally viable VD, the data collection from the previous study had to be expanded, not just in quantity, but also in dimension. LPS control at MAX IV involves tuning the energy chirp entering both bunch compression stages, and as such 2D-scans were implemented to reflect this.

A two-dimensional scan of the RF settings controlling the energy chirp upstream of both bunch compressors was used to generate the final datasets. The scan produced a total of 9000 paired measurements of machine observables and TDS images. Although random sampling is often preferable for ML training datasets [49], structured parameter scans were adopted here because they provided uniform coverage of the operating region while maintaining stable beam trans-

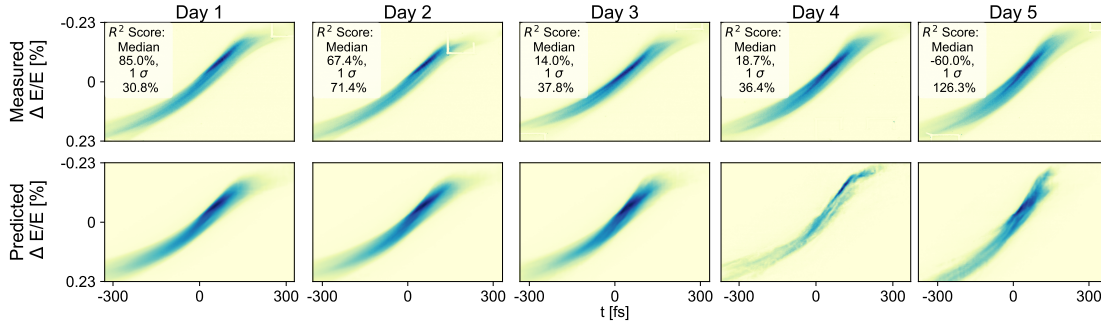


Figure 4.6: Virtual diagnostic predictions and TDS measurements of the MAX IV beam LPS distribution for a VD trained on data from Day 1 and then tested on all subsequent days. Indicated in each image is the median R^2 -score and its 1σ spread.

port during data collection. Training the previously established ANN architecture using this expanded dataset substantially improved the reconstructions, reaching a median R^2 -score of 96.1% for the full LPS predictions (see the top-left figure of Fig. 4.5 for the median example in the test set). The beam parameter approach yielded comparably strong results (Fig. 4.7, MAX IV results marked with blue crosses).

The results demonstrated clearly that the VD could reproduce full LPS measurements with high fidelity across the intended region of operation in the RF parameters. This VD represented a first version that could be suitable for routine use in the MAX IV diagnostic systems and operations. However, as mentioned in Sec. 3.2.3, there is a remaining open research question of how well these systems handle the drifts in the underlying machines [36]. With this in mind, a week-long test of this VD’s stability and reliability was carried out. This consisted of a new training set of 1200 LPS images and machine observables collected on one day, after which test datasets were collected on each subsequent day. The final median LPS images in terms of R^2 -score across the five days of the test are all shown in Fig. 4.6, with TDS measurements in the top row and VD predictions at the bottom. A deterioration in performance of the VD is clearly visible. Nevertheless, the VD maintained satisfactory performance for multiple days without any attempts to mitigate the effects of machine drifts. This suggests that periodic retraining would likely be required in an operational VD, but not with such frequency to be a major hindrance for implementation.

The MAX IV VD developments constituted procedures of data acquisition and network training that appeared suitable for extension to other facilities; ANNs trained on 2D scans of LPS-controlling parameters could be implemented with relatively little effort at multiple facilities. An application utilizing a similar network structure and data analysis pipeline could allow for a new level of integration between generic ML systems and accelerators. In this pursuit, the

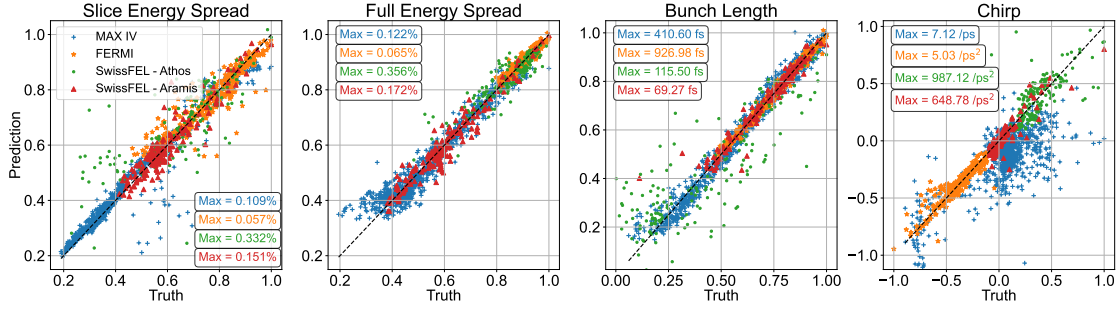


Figure 4.7: Correlation plots for each of the four beam parameters used in the results for **Paper III**. These were slice energy spread, full energy spread, bunch length and the energy chirp. Different colors and marker styles indicate the different facilities. Axes are normalized for readability, using the values displayed in each figure.

final MAX IV VD structure and data management process was expanded to two facilities, FERMI and SwissFEL. The results from all three facilities are reported in **Paper III**, where more details on the procedures and data can be found. The MAX IV LPS VD also saw further use and development for the later project employing Bayesian optimization in conjunction with the VD, outlined in Sec. 5.2, and covered in detail in **Paper V**. For **Paper III**, development continued at the other facilities, and the following sections provide a broad overview of the corresponding applications and results at FERMI and SwissFEL.

4.2 Virtual Diagnostic for FERMI

At the FERMI FEL facility in Trieste, Italy, LPS measurements are routinely used to tune and optimize the performance and X-ray/UV light production of the machine. A simplified layout of the machine can be seen in Fig. 4.8, with components critical for VD development highlighted in red and blue. While non-invasive longitudinal diagnostics such as beam arrival monitors (BAMs) and bunch compression monitors (BCMs) are available at FERMI, the principal diagnostic tool for full LPS information remains the TDS system [63, 64, 65]. This system, with RF structures, spectrometer dipole and beam imaging setup, is installed just before the FEL radiation is produced in the undulators. This means the system is invasive to user operations at FERMI and a virtual diagnostic capable of reconstructing the LPS from non-invasive measurements would therefore provide significant operational benefits by enabling continuous longitudinal beam characterization without disrupting FEL operation. For this purpose, the systems and procedures developed for MAX IV were also applied at FERMI.

Although the overall VD and data acquisition framework remained unchanged, there exists a different longitudinal control scheme at FERMI and adjustments

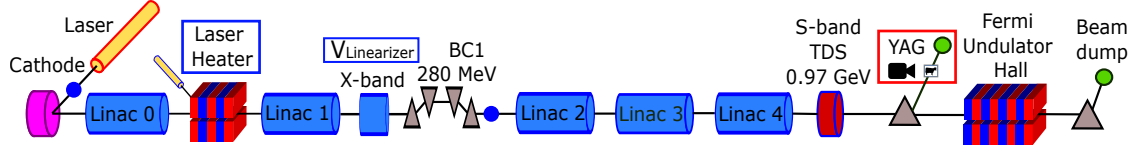


Figure 4.8: Simplified layout of FERMI. Highlighted in blue rectangles are the main controlling parameters for longitudinal control of beams, the RF phase upstream of the bunch compressor and the laser heater settings. Marked by a red rectangle is the YAG screen in the TDS system.

to the scanning procedures were necessary. Unlike MAX IV, FERMI operates with a single chicane bunch compressor², making the higher-order linearizing RF cavity and the laser heater the two dominant parameters governing the longitudinal phase space [66]. These parameters were therefore selected for the training dataset collection. The laser heater uses energy exchange between a laser pulse and the electron beam inside an undulator to induce a controlled increase of slice energy spread. This can be used to mitigate the effects of microbunching instabilities in the subsequent bunch compression stage [67].

Training on a final set of 4000 paired TDS measurements and machine observables produced a VD with reconstruction performance comparable to that of the MAX IV VD. A representative prediction from the FERMI test data is shown in the top-right figure of Fig. 4.5, representing the median test case with $R^2 = 96.4\%$. Similarly good agreement can be seen in the beam parameter approach (Fig. 4.7, orange stars), demonstrating how the VD methodology could be successfully transferred to a second accelerator with a substantially different compression system, LPS distributions and machine observables.

These results completed the initial VD development at FERMI and demonstrated the versatility of the methodology. They formed a substantial part of the results presented in **Paper III**. Further work and collaboration with FERMI included a second data collection at a different final energy, 1.1 GeV instead of the 0.97 GeV used in the original data. Another VD was developed for this dataset which achieved similarly good performance, and although these results were not included in a final published work, they further showed the reliability and versatility of the methods.

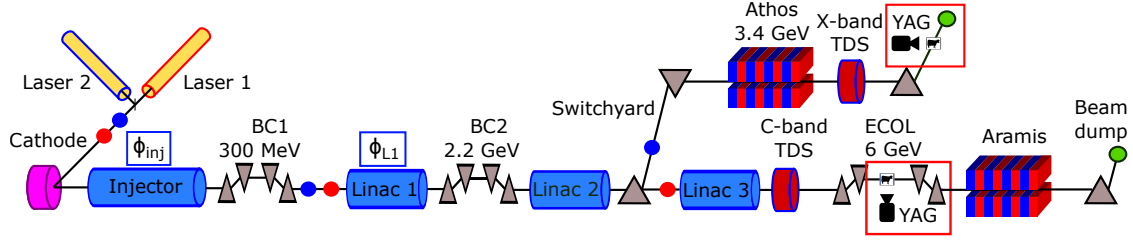


Figure 4.9: Simplified layout of the SwissFEL with its two FEL beamlines, Athos and Aramis. Highlighted in blue rectangles are the main controlling parameters for longitudinal control of beams, the RF phases before two stages of bunch compression. Marked by a red rectangle are the two measurement points in the two TDS systems.

4.3 Virtual Diagnostic for SwissFEL

SwissFEL, located at the Paul Scherrer Institute in Switzerland, was the third accelerator facility included in the multi-facility VD study. As at FERMI, detailed knowledge of the LPS at SwissFEL is essential for bunch compression and FEL performance optimization. SwissFEL also introduces additional challenges, as it provides beam simultaneously to two independent FEL beamlines [68]. A simplified layout of the facility can be seen in Fig. 4.9, with the Athos and Aramis FEL beamlines included. The LPS-controlling parameters are marked in blue and the two TDS systems used in data acquisition are marked in red.

SwissFEL operates with two bunches which are produced by separate laser systems at the gun, and are accelerated and compressed with a 28 ns delay between the bunches, before being separated using a resonant kicker system [68], denoted as the *switchyard* in Fig. 4.9. Since the TDS systems are located downstream of the switchyard, separate LPS measurements are available for both beamlines, making SwissFEL an intriguing facility for extending the VD applications to multiple simultaneous operating modes.

Both TDS systems are located after the bunch separation, and can be operated in parallel during data acquisition, producing two complete datasets with a single parameter scan. The Aramis TDS system is located upstream of the undulators, and is therefore fully invasive, motivating the use of a VD [69]. Although the Athos TDS is located after the X-ray producing undulators [70], it is not routinely used during operation because of radiation safety constraints. Consequently, continuous non-invasive reconstruction of the LPS would provide operational benefits for both beamlines.

As at MAX IV, training datasets were generated by scanning two RF phases, in the injector ϕ_{inj} and linac 1 ϕ_{L1} , marked with blue in Fig. 4.9. These scans

²A second chicane bunch compressor is available at FERMI, but is not used routinely and was not active during the relevant data acquisition.

provided 1600 samples collected from each TDS system, along with machine observables like RF information and measurements from BPMs, BAMs and BCMs. An example of the final LPS VD results for both the Aramis and Athos cases are shown in the lower row of Fig. 4.5, representing the median R^2 -scores of 99.1% and 94.3% respectively. The data from both beamlines was also used in the beam parameter approach and the results are presented in Fig. 4.7, with Aramis shown with red triangles and Athos with green dots.

The SwissFEL study demonstrated that the VD methodology remained effective even in a facility operating two independent FEL beamlines simultaneously. Together with the MAX IV and FERMI VDs, this completed a strong demonstration that a common ML framework and acquisition methodology could be transferred between accelerators with substantially different longitudinal compression schemes, diagnostic layouts and LPS distributions. Critically, individual VDs did not require identical setups and extents of machine observables and data, and the broader methodology proves to be easily transferable and is not restricted to particular accelerator configurations or LPS distributions. At the same time, the differences in the data structures and observables shows that without extensive adaptation, the VDs developed cannot be transferred directly between machines. There also remains a major open question of how to extend these VDs beyond the controlled datasets and parameter spaces they operate under, and further developments are required for the use of LPS predictions in live operating conditions. Still, the multi-facility developments established the generalizability of the overall framework, providing a strong basis for extending the approach to specific longitudinal diagnostic applications and use cases. One such extension followed directly at SwissFEL, where the VD was subsequently developed to support passive wakefield streaker calibration, presented in **Paper IV** and discussed in Sec. 5.1.

Chapter 5

Applications Based on Virtual Diagnostics

As shown in previous chapters, substantial progress has been made in integrating virtual diagnostics (VD) with accelerator operations. Following on these findings it is of interest to analyze how a VD can be integrated into more advanced applications. Aspects of three such cases have been explored in greater detail; VD for calibration, for optimization and in X-ray diagnostics (**Papers IV-VI**). The three projects were also carried out at three different accelerator facilities.

The calibration application consisted of a continuation on the work covered in Sec. 4.3, on VD developments for SwissFEL, by producing a new VD focused on calibration of the passive wakefield streaker in the Aramis beamline [9]. This ML method was compared favorably with a non-ML method for non-invasively reconstructing current profiles (**Paper IV**). A continuation of the work at MAX IV was also carried out, by further applying the LPS VD in a Bayesian optimization framework, as described in Sec. 5.2 (**Paper V**). Finally, an application of the VD methodology to the FLASH FEL at DESY in Germany was explored. The objective was to develop a VD capable of non-invasively predicting FEL power profiles, i.e., the longitudinal distribution of the X-ray pulses produced during the FEL process (**Paper VI**).

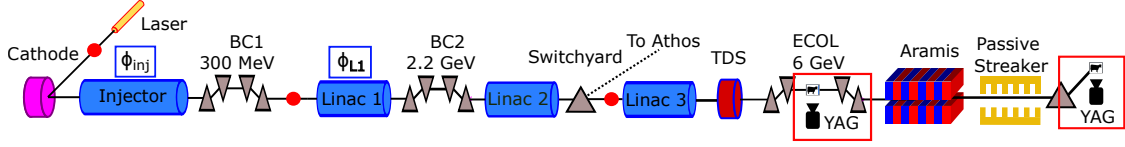


Figure 5.1: Layout of the SwissFEL accelerator and Aramis FEL beamline. The relevant parameters for the data acquisition for **Paper IV** are marked with blue rectangles, and the TDS and PS systems are marked with red.

5.1 Non-Invasive Calibration of a Passive Wakefield Streaker using Machine Learning

While TDS systems have been established as the most prominent and direct LPS diagnostic, the drawbacks of their cost and operational complexity remains recurring issues [29, 71]. Comparatively, the passive wakefield streaker (PS) (see Sec. 2.4.2) is a very cost efficient alternative, with the added benefit of inherent synchronization of the streaking with the passing beam [9]. A full PS system, including a corrugated metal structure for wakefield generation, a spectrometer dipole and a scintillating screen for beam imaging, is available for data acquisition in the Aramis beamline [9]. The following section explores how ML can be leveraged to enable non-invasive time calibration of this system.

The time calibration of a TDS system is relatively simple [71]. As implied by Eq. 2.8, a phase scan of the deflecting RF field is sufficient to solve the underlying equation and find a $t \rightarrow x$ mapping. However, for PS systems there are multiple, interdependent variables that contribute to the final transformation of coordinates. As we can see in Eq. 2.9, $\Delta x(t)$ has a direct dependence on the underlying current profile $I(t)$. The local transverse wake function $w(t)$ will also depend on the transverse trajectory through the PS structure. Now, provided a true current profile $I(t)$, calibration techniques independent of an analytical model of w has recently been shown [32]. The current profile can be measured with an active TDS system, but this calibration procedure is fully destructive, and a non-invasive alternative should be explored. The focus of the study in **Paper IV** was to reconstruct $I(t)$ non-invasively using a VD, for use in the PS time calibration. One can also bypass the reliance on a VD, if non-invasive current profile measurements are available, as in [32] wherein, for the European XFEL facility, the CRISP detector was used [72]. However, no such diagnostic is available at SwissFEL, as for many other facilities. One can also reconstruct the current profiles using the convolution of the measured PS projection $\rho(x)$ and the analytical wakefield function w [11, 73]. This technique was used in comparison with the VD approach.

In principle, two sets of longitudinal information are required to set up the

relevant VD: the PS projections $\rho(x)$ and the current profiles from the Aramis TDS system $I(t)$, which would serve respectively as input and output for the VD. A layout of the SwissFEL machine and the Aramis beamline can be seen in Fig. 5.1, with the TDS and PS systems marked with red boxes. As the TDS measurements are invasive, prohibiting collection of PS data, the two sets of data could not be collected in parallel. Therefore, two identical parameter scans were carried out at different times. The scans were focused on compression, controlled using the injector RF phase (ϕ_{inj}) and Linac 1 RF phase (ϕ_{L1}). The scans resulted in a total of 1685 pairs of profiles¹.

After the full data acquisition and processing, as well as matching of the TDS and PS datasets, the VD development and training followed. Given that only the current profile $I(t)$ is required for PS calibration, the networks and scoring methods were adapted from the image-focused ANNs used in earlier projects, covered in chapter 4. A different metric for comparing network performance was also introduced, namely the correlation coefficient r . Given two different current profiles I_1 and I_2 , their r -score can be defined as

$$r = \frac{\int dt (I_1(t) - \bar{I}_1)(I_2(t) - \bar{I}_2)}{\sqrt{\int dt (I_1(t) - \bar{I}_1)^2} \sqrt{\int dt (I_2(t) - \bar{I}_2)^2}}, \quad (5.1)$$

where $\bar{I}$ represents the mean current. r will be insensitive to the amplitude of the current profiles, but for the ML purposes they are all normalized to an integrated charge of 200 pC in any case. To construct the test dataset, 94 setpoints were chosen from the full scan, and for each point the median sample, in terms of RMS bunch length, was selected for testing. To avoid information leakage, the remaining samples in these 94 setpoints were removed from all sets.

Three different ANN architectures were tested for the task of reconstructing the current profiles $I(t)$, given the PS projected distribution $\rho(x)$ as input. The architectures were a MLP structure, a CNN model and finally an application of the modern U-Net model [41]. The MLP and CNN are both described in Sec. 3.1, while the U-Net can be seen as an extension of the CNN in an encoder-decoder configuration. The U-Net compresses the input into a small, information-dense layer in the center of the network, often referred to as a *bottleneck* or *latent space*. The second half of the network, the decoder, then reconstructs the desired output from this representation. While encoder-decoder networks are commonly used to compress information, the U-Net goes further, incorporating skip connections "across" the central bottleneck to maintain high-resolution features from earlier layers. As a result, the U-Net is generally better for preserving detailed information present in the input data [41].

¹This data acquisition was carried out in conjunction with the collections for **Paper III**.

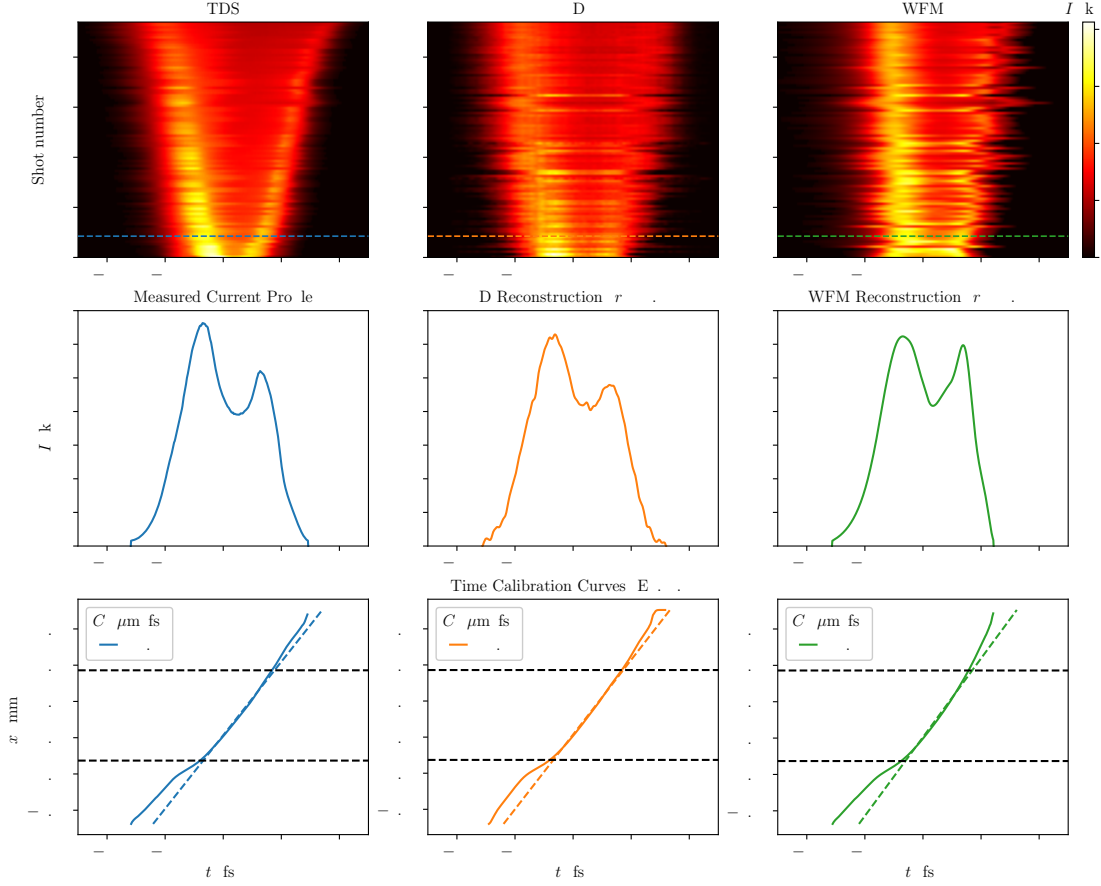


Figure 5.2: Outline of the current profiles and time calibration curves from the results in **Paper IV**. (Top) Waterfall plots of all measured and reconstructed current profiles, with the specific case shown in the middle and bottom rows highlighted. (Middle) Specific current profiles from measurement and the reconstruction methods (VD and WFM). (Bottom) Calibration curves resulting from the current profiles from all three sources (see Eq.5.2).

The U-Net reached a final performance of 95.6%. The U-Net training time remained on the scale of one minute, and it was selected for use in the final results. Further testing was done on using less training data, and it was found that a partition of 30% of the initial 1150 training samples was sufficient for the maximum levels of stable performance. A U-Net trained on this partition of the data was used for the final results.

For a reference method to compare with the VD approach, an established technique of reconstructing the current profiles using a wakefield model was implemented, abbreviated here as *WFM* [74, 73]. For this method, a distance d to the PS structure had to be established, which was done by using the TDS measurement at one compression setting to establish $\Delta x(t)$ and $I(t)$, and then solving for d in Eq. 2.9. Provided the distance d , an analytical model for w and the PS projection $\rho(x)$, $I(t)$ can be iteratively reconstructed [74, 73].

Provided the three different sources of current profiles, they could now all be

applied in a PS calibration technique. The wakefield model-independent method established in [32] uses the normalized cumulative charge distribution, q , based on either $I(t)$ or $\rho(x)$, represented as $F_I(t)$ and $G_\rho(x)$ respectively. Given that they are normalized, the output of both will range between 0 and 1, and as there is no negative current present, both will be monotonic. This implies the functions are invertible, and given this we can construct

$$q = G_\rho(x) \Leftrightarrow x = G_\rho^{-1}(q) \Rightarrow x(t) = G_\rho^{-1}(F_I(t)). \quad (5.2)$$

This can also be expressed implicitly as matching x and t on the condition $G_\rho(x) = F_I(t)$.

The results of the three sources of the current profiles $I(t)$ (TDS measurements, VD and WFM) are shown in Fig. 5.2. Examples of the current profile reconstructions and measurement are shown in the middle row of Fig. 5.2. The current profiles shown in each case were then used with Eq. 5.2 to obtain the $x(t)$ functions shown in the bottom row plots. For a simplified comparison between these different calibration functions, linear fits were performed in the region between the 0.3 and 0.9 charge quantiles, shown with dashed lines in the time calibration figures.

One can compare the results of the time calibrations based on the VD and WFM approaches to the actual TDS-based calibrations. This was done in two ways, both in terms of the linear approximation of $x(t)$, and in terms of the calibrated RMS bunch length as compared with a direct TDS measurement. It was observed that the VD achieved a lower average relative error than the WFM approach, with 9.2% and 16.8% respectively. In the case of the RMS bunch lengths, this difference in performance decreased, but the VD was still marginally more accurate.

Passive wakefield streakers can be considered a recent addition to the field of longitudinal diagnostics [9]. The results in this study represent the first application of ML techniques to PS systems. With these results, future applications for FEL users to extract LPS and FEL pulse information non-invasively could be further studied [11]. The complete details of the study are covered in **Paper IV**.

5.2 Bayesian Optimization of Longitudinal Phase Space at MAX IV

Accelerators require precise and continuous control of both transverse and longitudinal beam parameters. The beam distribution in the longitudinal phase space has a direct and constant influence on delivered parameters at FEL and other light sources [20]. Optimization using ML techniques is an increasingly active field of study within accelerator physics and operations [36]. As part of the work for this thesis, multiple applications of Bayesian optimization (BO) for the MAX IV linac were developed and tested. Some projects focused on transverse beam properties were thoroughly tested and used, such as optimizations of the transverse beam size in the FemtoMAX undulators, and some investigations were also carried out of optimizing the transverse emittance throughout the MAX IV linac. As transverse optimizations are quite well established in the field [36], these studies remained mostly unreported on. The BO developments eventually led to an application to the longitudinal properties of the MAX IV beam, and BO was used in congruence with the other ongoing developments of VDs (see chapter 4) to develop a fully non-invasive longitudinal optimization procedure.

Bayesian optimization was integrated with the MAX IV linac control system using the *XOpt* framework [17]. This provided the core Bayesian optimization components described in Sec. 3.3, including the Gaussian process models and acquisition functions, while the interface to the MAX IV machine was implemented using the control systems relevant to this project. The upper confidence bound method was used as the acquisition function.

The actual task selected was designed not to be overly complicated for this initial study, and as such the bunch length Δz , energy spread $\Delta\delta$ and the sum of the two were selected as beam parameters for optimization. These are critical components of the longitudinal distribution, commonly tuned to optimize the X-ray production in the MAX IV linac. The parameters can be controlled by the off-crest phases in the accelerating RF structures before BC1 and BC2, i.e. in L01 and the main linac, in turn controlled by the SLED filltimes and feedback systems [62]. With the controlling and dependent parameters established, the optimizers need only query the environment, done either through the invasive TDS measurements or through the non-invasive VD.

A VD of similar structure and use to those developed in chapter 4 was brought into operation for the BO application. As in those cases, the VD consisted of combined MLP and CNN layers, trained to predict full images from the TDS,

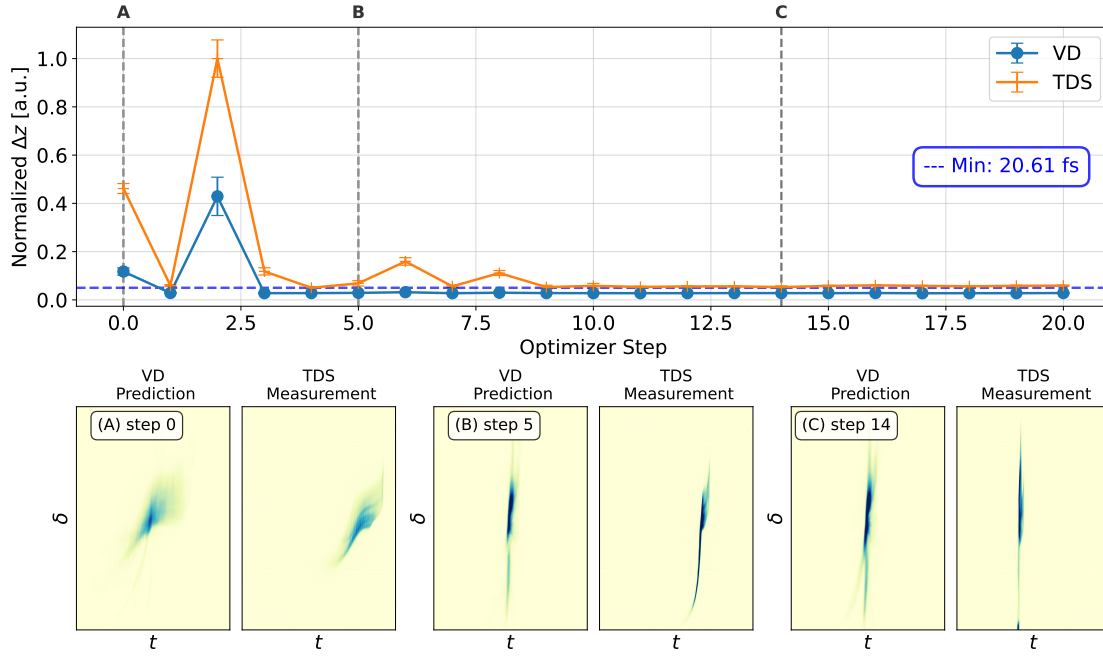


Figure 5.3: Results of an optimization run, optimizing on Δz , based on the VD predictions. Parameter results from TDS measurements in each step are also plotted for comparison. Δz is normalized, as some offsets were present between the VD and TDS results. Indicated in blue is the minimum encountered. In the bottom row, select images and VD predictions are shown for comparison.

from which the parameters to optimize were then calculated. The datasets for training and testing were also collected in the same manner, using 2D scans of the SLED filltimes feeding L01 and the main linac. In total, the dataset consisted of 1400 samples, of which 91 were used for testing. Within the test dataset the VD reached a median R^2 -score of 90%, and a Pearson coefficient of $r = 0.938$ was achieved in terms of correlation in calculated bunch length (see Eq. 5.1). Once the performance was satisfactory, the VD could be implemented in the live control room environment. During operation, the required non-destructive diagnostics were collected, normalized, and passed to the trained network, which provided real-time predictions of the longitudinal phase space. This signal could then be integrated into the BO loop, enabling efficient online optimizations.

Figure 5.3 shows the results of a bunch length Δz optimization based on the VD signal, with the measured and predicted LPS at three points in the campaign shown in the lower row. It is clear that the optimizer is able to increase compression, the goal of the optimization campaign. A minimum value of 20.6 fs was reached during the optimization. As seen in the top figure, for each optimization step, both a VD prediction and TDS measurement was collected, in order to bench-mark the VD predictions. We can see here that the optimizer works well with the VD and that throughout the optimization the VD produces parameters in good agreement with the real TDS measurements. This supports

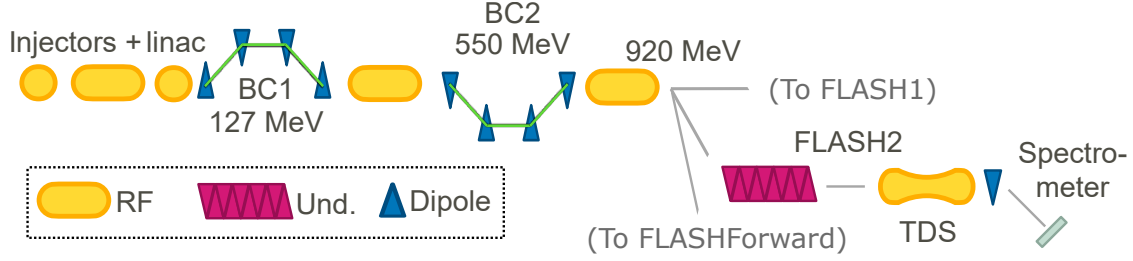


Figure 5.4: An overview of the FLASH linear accelerator, along the FLASH2 FEL beamline and TDS system.

the possibility of fully non-invasive, online optimizations in the future.

The results of this initial study are promising for the future of BO for use in the MAX IV linac, as well as for studies in general combining BO with other ML techniques. The results showed that VD applications could be applied in congruence with Bayesian optimization procedures. Future work in these directions should aim to include more advanced longitudinal beam properties, such as energy chirp or slice emittance, in order to fully utilize the LPS information available. Further details on this study can be found in **Paper V**.

5.3 Reconstruction of the FLASH FEL Pulse

The ML applications covered thus far in the thesis have remained focused on the properties of electron beams in different linear accelerators. An important application of such beams is the generation of high-quality photon beams for user experiments. This is especially true for FELs, where the longitudinal properties of the beam play a critical role in the quality of the FEL radiation [20]. A more direct approach to improve delivery conditions could be to apply ML in applications oriented around the photon beam [33].

At an FEL, the temporal profile and properties of the produced X-ray light is of great interest and a key parameter for time-resolved experiments [20]. In this project, we refer to this profile as the *FEL power profile*, and it remains a challenging diagnostic task to access such information, requiring invasive diagnostics and dedicated time with the machine. As outlined in Sec. 2.5, there are electron beam effects which can be utilized to access such information. Such techniques can be carried without the use of ML, but is then reliant on destructive longitudinal diagnostics. A ML application aiming to alleviate this drawback had been shown to work for the FLASH FEL at DESY, Germany [33]. With this project, we aimed to expand on previous work by applying the image-based VDs developed in earlier studies [16]. The VDs could allow for inference of

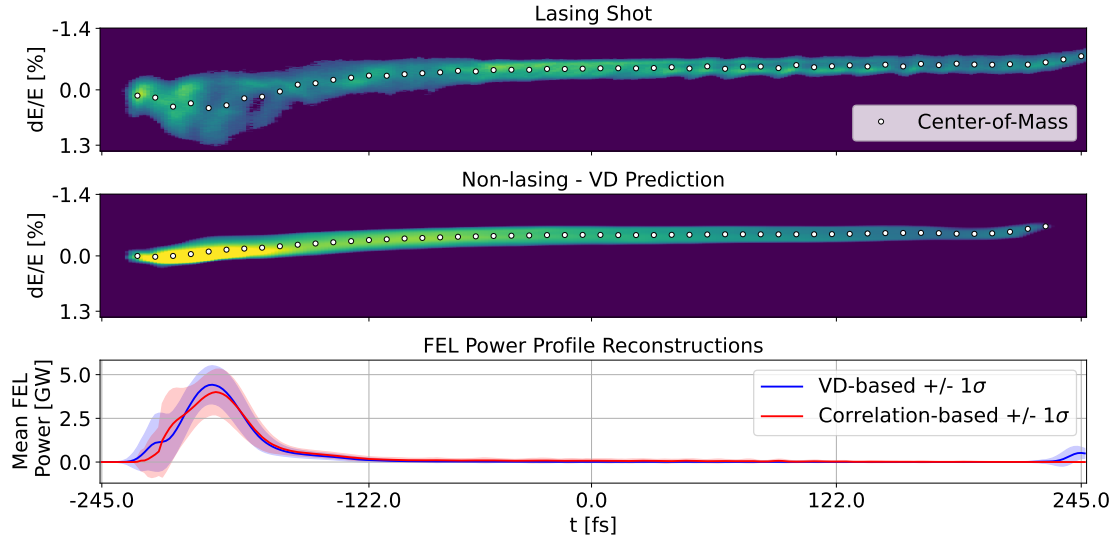


Figure 5.5: An example of the FEL power profile reconstruction. (Top) is the measured lasing electron LPS, (middle) shows the VD prediction of the corresponding non-lasing LPS distribution, where both distributions are shown with the center-of-mass along the y-axis highlighted. (Bottom) shows the resulting mean FEL power profiles from both VD comparisons and the correlation-based approach, with shaded regions representing $\pm 1\sigma$.

destructive signals required for FEL power profile reconstruction.

A simplified layout of the FLASH linac and the FLASH2 FEL beamline can be seen in Fig. 5.4. As FLASH is equipped with a TDS system, and can access the full LPS of the electron beam, one method of acquiring the FEL power profile is through comparisons between beam LPS under *lasing* and *non-lasing* conditions. This refers to beams which are producing FEL radiation and beams which are not [1, 33]. This can provide the temporal characteristics of the FEL pulse, but it relies both on the TDS measurements and on dedicated machine time without active lasing. However, with the previous work in this thesis it has been shown we can reconstruct such measurements. If a VD is constructed to predict the non-lasing LPS distribution, it can be used with non-invasive measurements of lasing bunches to reconstruct the FEL power profile. This was done for the FLASH2 FEL beamline, and the VD approach was bench-marked against a more conventional, non-ML approach.

As the lasing process will vary randomly from shot to shot in the machine, and it is sensitive to the compression settings [20], parameter scans were not required. Rather, the machine was allowed to drift and jitter randomly, while images were acquired from the X-band TDS system located after the undulators in the FLASH2 FEL beamline. Lasing could be controlled using a horizontal kick to the beam trajectory through the undulators, effectively separating the electron bunch from its generated radiation and deprecating the SASE gain.

The VD for this task was based on the same ANN architecture as used in earlier studies, see chapter 4. It used a total of 29 non-invasive diagnostic inputs to reconstruct the TDS measurements, and could manage to do so to a median R^2 -score of 84.4%. The critical point in the VD setup was the difference between training and testing data, namely that the VD was trained on non-lasing data and then tested for inference from inputs collected during lasing. In theory, this allowed the VD to predict non-lasing LPS distributions matched to the collected lasing data.

Once non-lasing LPS distributions were made available using the VD, the reconstruction of the FEL power profile is relatively straight forward. An example of two LPS distributions, in this case a measured lasing distribution and a non-lasing VD prediction for the same beam, can be seen in Fig. 5.5. The effect of lasing is visible in two major ways in the head of the lasing bunch (left in the image): a decrease in the central energy of longitudinal slices and an increase in the energy spread in these same slices. The central energy in a selection of slices is shown with white dots. Through a comparison in these slice central energies between lasing and non-lasing distributions, and an application of the TDS time calibration, the FEL power profile is reconstructed. The VD approach is compared with a non-ML technique where a bunch profile of maximum similarity to the lasing case is chosen statistically from the non-lasing dataset. This is referred to as *correlation-based* in the results.

In the bottom plot of Fig. 5.5, the mean FEL power profiles from both techniques are shown, calculated from a full test dataset of measurements under lasing conditions. These mean power profiles are both shown with shaded regions representing $\pm 1\sigma$ across the full dataset. We can see a high similarity in the results from the two techniques.

This work was a promising step closer to fully non-invasive FEL diagnostics. It shows further that applications of VD to areas of FEL radiation diagnostics is worth pursuing, extending their uses beyond electron-beam focused projects. Continuous estimates of FEL power profiles could be particularly valuable for monitoring variations in the pulse properties and reporting to users, and additionally to implement in feedback processes and machine optimization. As with other VD applications, further work is required to establish reliable and operational systems based on these developments. Further details are expanded on in **Paper VI**.

Chapter 6

Investigations of the Longitudinal Dynamics of Arc-like Bunch Compressors

The earlier chapters of this thesis has remained focused on ML applications for a variety of diagnostic systems and accelerator controls. The projects have all involved the longitudinal phase space distributions of beams, where the effects of bunch compression plays a critical role. For a deeper understanding and enhanced control of beams, it is of interest to further investigate the underlying dynamics and parameters of such controls. Specifically, the arc-like bunch compression scheme of the MAX IV linac is rare and warrants further investigation [4, 27].

Two aspects of these systems have been analyzed further. In Sec. 6.1, we present studies aiming to investigate the transfer elements R_{56} and T_{566} through an arc-like compressor, both critical for the bunch compression process. This was done using a series of time-of-flight measurements of beams in the first compressor of the MAX IV linac, BC1 (**Paper VII**). The second project on the MAX IV bunch compression scheme investigated an inherent feature of the system that serves to stabilize the compression process (**Paper VIII**). The two projects and their implications are outlined in the following sections.

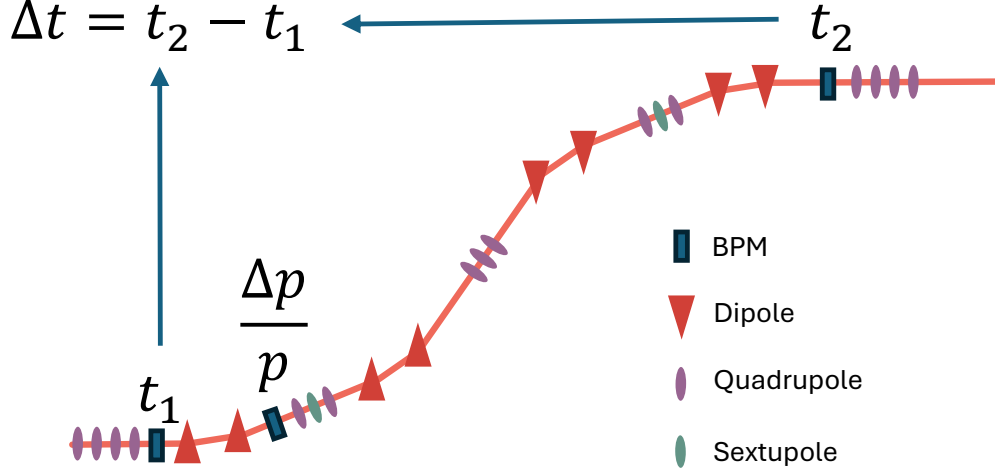


Figure 6.1: Simplified figure of the MAX IV bunch compressor, with the relevant BPMs for time-of-flight Δt and energy calculations indicated.

6.1 First Measurements of the First and Second Order Longitudinal Dispersion of a MAX IV Bunch Compressor

The transfer elements, or elements of the *longitudinal dispersion*, of the bunch compression process were presented in Sec. 2.3. One of the biggest advantages of the arc-like compression scheme, covered in Sec. 2.3.2, is the ability to directly control the nonlinear transfer term T_{566} . This allows for control of the linearization of the longitudinal phase space (LPS) distribution of the beam [4]. The direct linearization control is particularly valuable for FEL applications, where a well-linearized beam improves the conditions for efficient FEL operation. Control of T_{566} can also allow for more specific beam shaping for more advanced application, such as producing ultrashort current spikes or delivering specific current profiles optimal for plasma wakefield acceleration [75]. As the arc-like compression scheme is not widely adopted, the direct T_{566} control has remained without experimental confirmation. Whether the effect of the dispersive elements, as explored in theory and simulation, could be experimentally confirmed was the focus of the study presented in **Paper VII**.

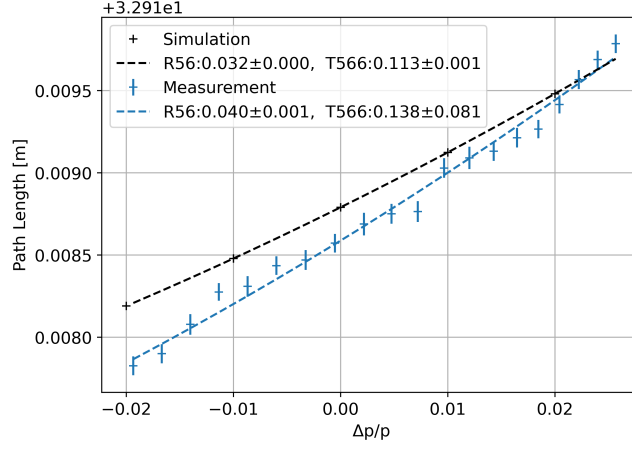
The experimental configuration to measure R_{56} and T_{566} develops naturally from the theory of bunch compression. As the basis of the compression process is the difference in transfer time, or path length, as a function of different particle energies (Eq. 2.4), one should aim to detect the time-of-flight (TOF) during transfer through the bunch compressor of beams of different energies. It is,

however, by no means trivial to perform high resolution TOF measurements of ultrarelativistic electron beams. Multiple different signal pickups were tested in the development of this study, including the use of horn antennas to detect the timing of the passing beam. In the end, it was found most reliable to base the TOF detection on the existing stripline beam position monitors (BPMs) at the entrance and exit of the bunch compressor, indicated in Fig. 6.1. From each of these, one of the four pickups was disconnected from the usual BPM analysis systems, and fed into a high-frequency oscilloscope. This resulted in well-resolved waveforms from which the exact time of the beam's passage could be extracted, and through comparing the time at entry and exit of the bunch compressor, the TOF of the transfer was retrieved.

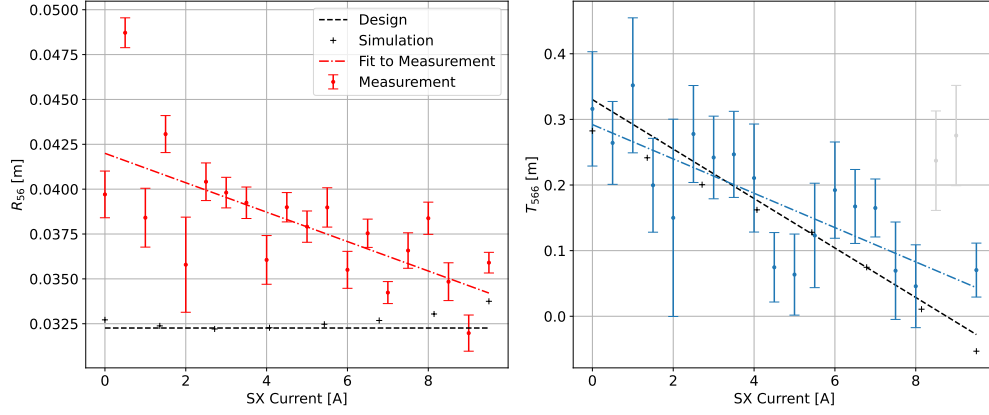
TOF collections were repeated for a number of SLED cavity filltimes, as to establish the critical beam energy-TOF dependency and extract the R_{56} and T_{566} elements [62]. The energy of each passing beam was calculated using readings from the dispersive BPMs inside the bunch compressor, also indicated in Fig. 6.1. The procedure was then repeated for a number of field strengths in the dispersive sextupoles at the beginning and end of BC1, to finally establish the effect of these higher-order magnets on the R_{56} and T_{566} transfer elements. The full procedure was also simulated in *elegant* [57], and design parameters were calculated using the theoretical transfer matrices for the compressor [18].

An example of the TOF results for a scan of incoming beam energies can be seen in Fig. 6.2a. Here the dispersive BPM reading has been recalculated to $\delta = \frac{\Delta p}{p}$ (x-axis) and the TOF has been recalculated to a relative path length (y-axis). Both measured and simulated results are shown, each also along with a polynomial fit to the data. These fits will then describe the transfer elements, as expressed in Eq. 2.4, and the resulting R_{56} and T_{566} are displayed in the labels.

In Fig. 6.2b, we can see the R_{56} and T_{566} for each value in a scan sextupole strengths. Here the simulated results and the measurements are plotted along with the expected dependencies from the matrix model. Both simulation and design suggests there should be no dependency between sextupole strengths and the R_{56} of the transfer, but there is a slight gradient present in the measured data. This may have resulted from a systematically misaligned trajectory through the compressor, or from the settings of the dispersive quadrupoles deviating from their design values. In the T_{566} data we see the expected gradient in both simulation and measurement, even if the measured data suffers from large errors. One may consider the error bars an effect of the limited energy range of the TOF scans.



(a) Example of the TOF measurements for one sextupole setting, along with the corresponding simulation results.



(b) R_{56} and T_{566} for all different sextupole settings, as based on simulations and measurements, plotted along with the theoretical design values.

Figure 6.2: Results from the TOF studies of BC1 at MAX IV. (a) shows resulting TOFs from an energy scan for a single sextupole setting, while (b) shows the full results for all sextupole settings, i.e. the calculated R_{56} and T_{566} of the transfer.

This study included the first measurements of the arc-like design's control over nonlinear compression terms. This is a crucial parameter for many high-level linac applications, and can be used to support the employment of the compression scheme in future facilities. Further, it lays the foundations for more advanced studies of the dynamics in the arc-like compressors, such as the work covered in the next section. This project was carried out on the first bunch compressor in the MAX IV linac, and an expansion to the second compressor, BC2, was also attempted [76]. The full details of the TOF investigations of BC1 are covered in **Paper VII**.

6.2 Inherent Stabilizing Properties of the Arc-like Bunch Compressors

Modern FEL facilities have increasingly extreme demands on the longitudinal properties of electron beams, including bunch lengths nearing attosecond-scales and very stable LPS distributions. This is difficult to achieve, especially as the relative bunch length jitter will naturally scale with the amount of compression applied. A large source of shot-to-shot compression jitter is the jitter in the RF parameters used in accelerating structures. Since the RF acceleration provides the energy chirp required for compression to take effect, RF phase jitter is typically a dominant source of compression jitter further downstream in the accelerator. Previous studies have already found a critical advantage in stability enabled by the specific RF working point when compression is done using the arc-like scheme [27]. There may however be further stabilizing properties made accessible by the decoupled control of the nonlinear compression term T_{566} . An indication of this was observed during other work with the T_{566} -controlling dispersive sextupoles, during a study focused on investigate current profile shaping for plasma wakefield acceleration [75].

Since the arc-like bunch compressors allow for independent control of the first nonlinear compression factor, T_{566} , it can be of interest to expand Eq. 2.7 to a higher order. This was done by Dohlus and Limberg [25] and found to be

$$C = \frac{1}{1 + h(R_{56} + 2T_{566}\delta + \mathcal{O}(\delta^2))}. \quad (6.1)$$

The theoretical stabilizing property of arc-like compressors can be found in the direct dependence of C on both h and δ (see Sec. 2.1). Due to the sinusoidal shape of the accelerating RF fields, h and δ will have a negative correlation in the case of jittering in the RF phase. With this negative correlation, a direct cancellation of jitter in these parameters can be achieved if T_{566} is tuned properly and has the proper sign, such that $(R_{56} + 2T_{566}\delta)$ compensates for changes in h .

We can also express this explicitly using the relative compression sensitivity, defined as

$$\frac{C'}{C_0} = 2(C_0 - 1)\frac{T_{566}}{R_{56}}\delta' - R_{56}C_0h', \quad (6.2)$$

where $\frac{C'}{C_0} = 0$ can be achieved if

$$T_{566} = -\frac{R_{56}}{2} \frac{h'}{h} \frac{1}{\delta'}. \quad (6.3)$$

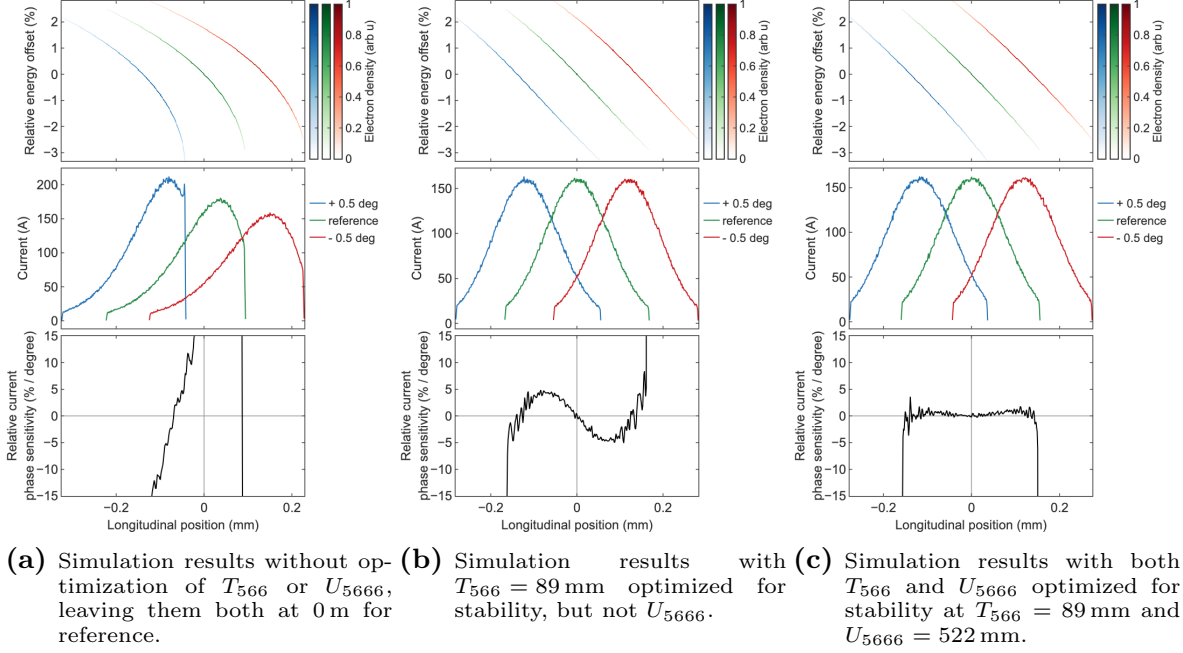


Figure 6.3: LPS, current profiles and phase sensitivity for three beams accelerated at different phases, spanning $\pm 0.5^\circ$, for three different settings of T_{566} and U_{5666} .

Here we see a critical requirement emerge on the signs of R_{56} and T_{566} . With the negative h - δ correlation during RF phase jitter, the stabilizing dynamic can only take effect if $\text{sign}(R_{56}) = \text{sign}(T_{566})$. This is naturally the case for arc-like compressors, but not for chicanes, where $\text{sign}(R_{56}) \neq \text{sign}(T_{566})$. For other jitter sources, such as jitter in the RF field amplitude, the requirement is different and chicanes are applicable. However, as T_{566} is not directly tuneable in chicanes, the specific values required for stabilization may not be achievable.

To investigate whether this theoretical stabilizing mechanism persists in more detailed models and could be observed experimentally, both simulations and experimental studies were carried out. Initial work to explore the parameter ranges in R_{56} , T_{566} and U_{5666} achievable with the dispersive elements of the MAX IV compressors was carried out using *elegant* simulations [57], first established in the work for **Paper VII**. The more detailed simulation work to investigate the stabilizing effect was carried out in custom particle tracking simulations, with parameters chosen to approximately represent the beam conditions going into BC1 in the MAX IV linac. The critical results of the simulations can be seen in Fig. 6.3. For each subplot higher-order compression terms are varied, whilst the first-order compression is maintained at $R_{56} = 32$ mm.

Figure 6.3a shows the case for a non-optimized T_{566} of 0 m. Without the stabilizing effect of the higher-order component, we can see the current profiles

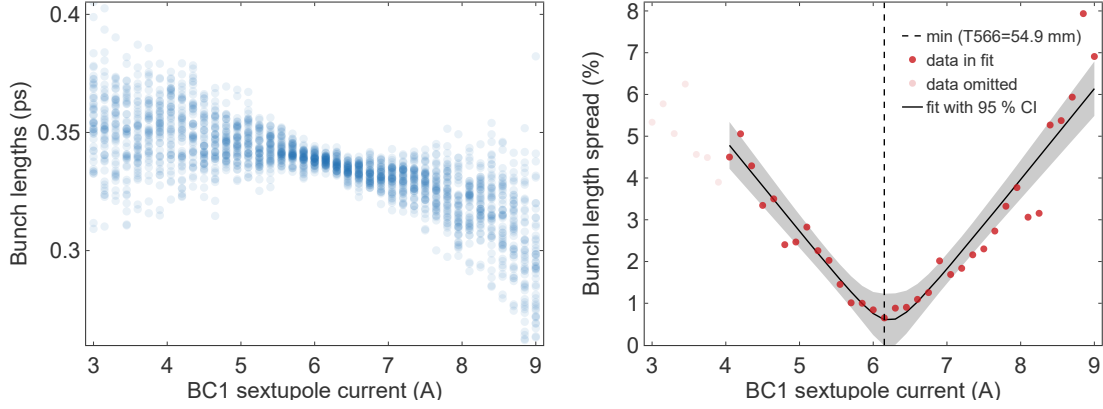


Figure 6.4: Results of TDS measurements of bunch length during scans of the BC1 sextupole field strengths. (Left) All calculated bunch lengths throughout the scan. (Right) Relative bunch length jitter in each scan point, calculated using the robust Q_n scale estimator, along with a fit and its 95% confidence interval (CI).

(Fig. 6.3a, middle figure) clearly deform as accelerating phase changes. In Fig. 6.3b the T_{566} has been optimized both for linearizing the LPS, as can be seen in the top figure, and for canceling the effect of jitter in the accelerating phase. This optimum was calculated to $T_{566} = 89$ mm. In the lower two figures we can see the stability has clearly increased, although some higher order effect is visible in the extracted phase sensitivity plot (bottom figure). This higher order effect is shown to be mitigated when U_{5666} optimization is introduced, as in Fig. 6.3c, where the optimum for stability was calculated to $U_{5666} = 522$ mm. These results indicate that the stabilizing mechanism associated with T_{566} persists in particle tracking and that control of still higher-order longitudinal transport could provide additional stability. It should be noted that direct control to the third order is not possible in the present MAX IV bunch compressors, but with slight modifications to the lattice and the introduction of octupoles it could be enabled. This has been briefly investigated in [4].

The simulation results confirmed the effect as present in detailed tracking, and revealed additional effects to investigate, such as the effect of jitter in the amplitude of the accelerating RF fields, and the future potential of mitigating the effects of U_{5666} on the compression stability.

Based on the simulation work, an experimental plan for investigating the effect in the MAX IV linac could be established. This consisted of large sets of LPS measurements, measured using the transverse deflecting structure (TDS) system, during scans of the field strength in the dispersive sextupoles in the bunch compressors. Since these magnetic elements allow direct control of the T_{566} , the compression stability should also vary during such scans, and optimal points for stable compression could be found. Scans of the BC1 dispersive sextupoles were performed, and for each setting a total of 50 TDS measurements were col-

lected. From these, bunch lengths and bunch length jitter could be calculated. Here, inter-quartile range (IQR) was used as a measure of bunch length, since it is more independent of the specific current profile shape as compared with FWHM or RMS measurements. The results of such scans and measurements are shown in Fig. 6.4. The left figure shows the bunch length of each measurement throughout the scan, where one can observe a minimum in bunch length variance at an approximate sextupole current of 6 A. This is clarified in the right figure, where the relative bunch length jitter for each sextupole setting is shown, calculated using the robust Q_n scale estimator [77]. As such, a reduction in relative bunch-length jitter is observed for particular sextupole settings, consistent with the predicted stabilizing mechanism.

The study established the stabilizing effect of T_{566} and found it clearly observable in simulation and experiment. Leveraging the new effect, bunches were compressed to 350 fs with a relative spread of well below 0.5%. This was without any extensive work on maximizing compression, and far shorter bunches with similar levels of stability should be achievable. This effect could be a critical advantage to leverage for facilities focused on high-quality FEL radiation production or plasma wakefield acceleration. The full details can be found in **Paper VIII**.

Chapter 7

Conclusions and Outlook

This thesis investigates the longitudinal properties and control of electron beams in linear accelerators and free-electron lasers (FELs), with a particular focus on the application of machine learning (ML) methods to these areas. The work combines studies of longitudinal beam dynamics and compression with the development of machine learning-based virtual diagnostics (VDs) for non-invasive beam characterization. The central questions posed in this thesis concern how longitudinal beam properties can be measured and controlled without reliance on invasive techniques, to what degree machine learning can provide reliable and applicable new methods in accelerator environments, and how longitudinal control can be furthered in novel ways with the arc-like compression scheme.

The work presented here answers these questions in several respects. Starting with the early developments for MAX IV, we have demonstrated how ML models can be developed into robust, facility-independent tools capable of enhancing and supporting existing diagnostics. By applying supervised learning with the large amounts of non-invasive diagnostics available in accelerators, the virtual diagnostics can provide fast estimates of LPS beam distributions without resorting to invasive techniques. With the introduction of LPS VD techniques to the multiple facilities involved in this thesis, the project has served to demonstrate the capability of ML techniques for generalization within the field of accelerators. Importantly, the application of LPS VD techniques at multiple facilities has demonstrated that the underlying methodology is not restricted to a single accelerator or experimental configuration. While the individual models remain dependent on the available diagnostics and operating conditions of each machine, the approach itself has proven sufficiently general to be transferred between facilities and adapted to different accelerator systems. This addresses

and answers the central question of whether ML methods could be generalized between facilities.

This work helps to forward the field and support interdisciplinary work to advance operational design and development for facilities, while also identifying many areas to be investigated and improved. The integration of machine learning techniques which are reliant on large datasets is an ever present challenge for systems supported by supervised learning. Wider adoption of machine learning in accelerator facilities will require more standardized approaches to data acquisition, storage, and management. The adoption of standardized data policies for machine learning would pave the way for much wider application of the techniques presented here. This could be paralleled with developments of the ML models themselves, where more focus could be applied to the stability and reliability of the methods [48]. These aspects are especially important if virtual diagnostics are to progress from experimental tools towards routinely trusted components of accelerator operation.

An important continuation on the multi-facility work is therefore to investigate how the dependence on large, facility-specific datasets can be reduced. ML techniques such as transfer learning could allow information obtained at one machine to be reused with minor adaptations when applying a VD to another. This could be combined with physics-informed ML, where datasets or ML architectures and functions are adapted to reflect accelerator dynamics, potentially constraining predictions to physically realistic and relevant results. Similarly, decreasing the dimensionality of the learning problems for VDs could increase their appeal and lower the data volume they require. This could be done with encoder-based approaches, or in the limitation to relevant beam parameters seen in this work. Another multi-facility study comparing such approaches as these would provide a critical next step towards determining how general and versatile virtual diagnostics can become.

Beyond demonstrating the capability of these systems to supplement destructive diagnostics, this thesis has further demonstrated their use in a number of integral systems of large accelerators. By integrating VD-based methods into the areas of passive wakefield streaker calibration, beam optimization and FEL X-ray pulse diagnostics, the work furthers the potential uses of ML techniques beyond basic signal inference, and addresses the central research question of integrating virtual diagnostics in routine accelerator operations. These developments show how the value of a VD is not limited to reproducing existing diagnostics: once integrated into calibration and optimization procedures, these systems can enable fast measurements and operations that would be difficult and disruptive to perform continuously otherwise. These VD applications, and

further developments in such directions, are required to find the specific areas of accelerator physics, instrumentation and operations where these new techniques can make a tangible impact and produce novel solutions, rather than treating ML diagnostic developments as isolated predictors.

Several questions related to the integration of VD techniques remain outside the scope of this work. In particular, the long-term performance of the developed VD applications under machine drifts and changing operational conditions has not been systematically investigated. This would require longer and large-scale studies, with integration of ML techniques to alleviate the potential degradation in performance. This could include periodic model adaption and retraining, implementing uncertainty estimates in predictions and results through ensembles of ANNs, and thresholds for identifying unreliable VD inferences. A first step in this direction is the type of long-term data collections done as part of **Paper III**, but more work and analysis in dedicated studies is required.

In conjunction with the studies of VD for LPS control and diagnostics, the thesis work also investigated the more general longitudinal control of electron beams with studies of bunch compression. The work on the arc-like compressors at MAX IV has served to present these systems as an emerging, powerful alternative to the established chicane solutions. Investigations of time-of-flight data showed experimental confirmation of direct T_{566} control, and a new inherent stabilizing property of the scheme was revealed and confirmed in simulation and experiments. The first project furthers the full characterization and understanding of the arc-like scheme, whilst the second proves that new and fundamental possibilities in bunch compression still warrants exploration and further study. The results of these studies serves to answer the question of advantages in the arc-like compression scheme: higher-order decoupled control of compression has been experimentally confirmed, and the arc-like scheme has been shown to provide intrinsic stabilization of compressed beams under accelerator perturbations. Both properties explored show the arc-like scheme as particularly attractive for next-generation FELs, where stringent beam quality requirements demand increasingly sophisticated control of the compression process. Arc-like compressors also show advantageous properties for plasma-wakefield accelerators [75].

Future studies should be carried out to more fully characterize the arc-like compressors. More in-depth work could include investigations of even higher-order effects and control, moving on to investigate decoupled U_{5666} controls using octupoles. Also, the collective effects within the compressors has not been in focus for this work and would require dedicated studies. Future work should also be carried out in congruence with users of the compressed beams, i.e., the FEL and

plasma wakefield accelerator communities, to incorporate the scheme fully with the demands of high-end user facilities. The projects on bunch compression have in total furthered understanding of the dynamics and control in the arc-like scheme. They can be used to support further research projects and design decisions of future systems and facilities.

In conclusion, the work presented in this thesis demonstrates both new methods for applying ML to the observation and monitoring of longitudinal electron-beam properties and new possibilities for controlling them. The results presented answer the principal questions addressed in this thesis, while identifying key limitations of the methods which must be overcome before more general adaption is possible for the betterment of future FELs and advanced accelerator facilities.

References

- [1] John M. J. Madey. Stimulated emission of bremsstrahlung in a periodic magnetic field. *Journal of Applied Physics*, 42(5):1906–1913, 04 1971.
- [2] Pisin Chen, J. M. Dawson, Robert W. Huff, and T. Katsouleas. Acceleration of electrons by the interaction of a bunched electron beam with a plasma. *Phys. Rev. Lett.*, 54:693–696, Feb 1985.
- [3] C. Behrens, F.-J. Decker, Y. Ding, V. A. Dolgashev, J. Frisch, Z. Huang, P. Krejcik, H. Loos, A. Lutman, T. J. Maxwell, J. Turner, J. Wang, M.-H. Wang, J. Welch, and J. Wu. Few-femtosecond time-resolved measurements of x-ray free-electron lasers. *Nature Communications*, 5(1):3762, 4 2014.
- [4] Peter H. Williams, Gustavo Pérez-Segurana, Ian R. Bailey, Sara Thorin, Bill Kyle, and Jonas Björklund Svensson. Arclike variable bunch compressors. *Phys. Rev. Accel. Beams*, 23:100701, Oct 2020.
- [5] Adam Dixon, Peter Williams, Sara Thorin, Andrzej Wolski, Alexander Brynes, Tessa Charles, and Ian Bailey. Arc and chicane bunch compression schemes for hard and soft x-ray free electron laser facilities: A comparison, 2026.
- [6] W. K. H. Panofsky and W. A. Wenzel. Some considerations concerning the transverse deflection of charged particles in radio-frequency fields. *Review of Scientific Instruments*, 27(11):967–967, 11 1956.
- [7] R. Akre, L. Bentson, P. Emma, and P. Krejcik. A transverse rf deflecting structure for bunch length and phase space diagnostics. In *PACS2001. Proceedings of the 2001 Particle Accelerator Conference (Cat. No.01CH37268)*, volume 3, pages 2353–2355 vol.3, 2001.
- [8] M. Huening. Bunch Length Measurements using Transverse Deflecting Systems. In *Proc. LINAC’18*, number 29 in Linear Accelerator Conference, pages 972–976. JACoW Publishing, Geneva, Switzerland, 1 2019.

- [9] S. Bettoni, P. Craievich, A. A. Lutman, and M. Pedrozzi. Temporal profile measurements of relativistic electron bunch based on wakefield generation. *Phys. Rev. Accel. Beams*, 19:021304, Feb 2016.
- [10] Jimin Seok, Moses Chung, Heung-Sik Kang, Chang-Ki Min, and Donghyun Na. Use of a corrugated beam pipe as a passive deflector for bunch length measurements. *Phys. Rev. Accel. Beams*, 21:022801, Feb 2018.
- [11] Philipp Dijkstal, Alexander Malyzhenkov, Paolo Craievich, Eugenio Ferrari, Romain Ganter, Sven Reiche, Thomas Schietinger, Pavle Juranic, and Eduard Prat. Self-synchronized and cost-effective time-resolved measurements at x-ray free-electron lasers with femtosecond resolution. *Phys. Rev. Res.*, 4:013017, Jan 2022.
- [12] Philipp Dijkstal, Weilun Qin, and Sergey Tomin. Longitudinal phase space diagnostics with a nonmovable corrugated passive wakefield streaker. *Phys. Rev. Accel. Beams*, 27:050702, May 2024.
- [13] Nils Maris Lockmann, Christopher Gerth, Bernhard Schmidt, and Stephan Wesch. Noninvasive thz spectroscopy for bunch current profile reconstructions at mhz repetition rates. *Phys. Rev. Accel. Beams*, 23:112801, Nov 2020.
- [14] Ryan Roussel, Juan Pablo Gonzalez-Aguilera, Eric Wisniewski, Alexander Ody, Wanming Liu, John Power, Young-Kee Kim, and Auralee Edelen. Efficient six-dimensional phase space reconstructions from experimental measurements using generative machine learning. *Phys. Rev. Accel. Beams*, 27:094601, Sep 2024.
- [15] S. Bettoni, G. L. Orlandi, F. Salomone, R. Boiger, R. Ischebeck, R. Xue, and A. Mostacci. Machine learning based longitudinal virtual diagnostics at swissfel. *Review of Scientific Instruments*, 95(1):015110, 01 2024.
- [16] J. Lundquist, J. Björklund Svensson, P. Dijkstal, E. Mansten, G. Penco, S. Werin, and F. Curbis. Multi-facility virtual diagnostic for longitudinal phase space predictions. *Scientific Reports*, 16(1):12021, Apr 2026.
- [17] R. Roussel, A. Edelen, A. Bartnik, and C. Mayes. Xopt: A simplified framework for optimization of accelerator problems using advanced algorithms. In *Proc. IPAC'23*, number 14 in IPAC'23 - 14th International Particle Accelerator Conference, pages 4796–4799. JACoW Publishing, Geneva, Switzerland, 05 2023.
- [18] Helmut Wiedemann. *Particle Accelerator Physics*. Springer, Cham, 4 edition, 2015.

- [19] Thomas P. Wangler. *RF Linear Accelerators*. Wiley-VCH, 2 edition, 2008.
- [20] Zhirong Huang and Kwang-Je Kim. Review of x-ray free-electron laser theory. *Phys. Rev. ST Accel. Beams*, 10:034801, Mar 2007.
- [21] Rolf Wideröe. Über ein neues prinzip zur herstellung hoher spannungen. *Archiv für Elektrotechnik*, 21(4):387–406, Jul 1928.
- [22] John David Jackson. *Classical electrodynamics*. Wiley, New York, NY, 3rd ed. edition, 1999.
- [23] Karl L. Brown. A First and Second Order Matrix Theory for the Design of Beam Transport Systems and Charged Particle Spectrometers. *Adv. Part. Phys.*, 1:71–134, 1968.
- [24] David Robin, Etienne Forest, Claudio Pellegrini, and Ali Amiry. Quasi-isochronous storage rings. *Phys. Rev. E*, 48:2149–2156, Sep 1993.
- [25] M. Dohlus and T. Limberg. Bunch Compression Stability Dependence on RF Parameters. *eConf*, C0508213:TUOB004, 2005.
- [26] G. Stupakov and S. Heifets. Beam instability and microbunching due to coherent synchrotron radiation. *Phys. Rev. ST Accel. Beams*, 5:054402, May 2002.
- [27] Erik Mansten, Robin Svärd, Sara Thorin, Mikael Eriksson, and Pedro Fernandes Tavares. Cancellation of klystron-induced energy and arrival-time variations in linear accelerators with arc-type bunch compressors. *Phys. Rev. Accel. Beams*, 27:040401, Apr 2024.
- [28] Robert E. Shafer. Beam position monitoring. *AIP Conf. Proc.*, 249:601–636, 1992.
- [29] G. A. Loew and O. H. Altenmueller. Design And Applications Of R.F. Deflecting Structures At Slac. In *5th International Conference on High-Energy Accelerators*, pages 551–555, 1965.
- [30] O. H. Altenmueller, R. R. Larsen, and G. A. Loew. Investigations of traveling-wave separators for the stanford two-mile linear accelerator. *Review of Scientific Instruments*, 35(4):438–442, 04 1964.
- [31] K. Bane, G. Stupakov, and I. Zagorodnov. Analytical formulas for short bunch wakes in a flat dechirper. *Physical Review Accelerators and Beams*, 19(8):084401, 8 2016.

- [32] Sergey Tomin and Igor Zagorodnov. A method for passive streaker LPS reconstruction, 2026.
- [33] Till Korten, Vladimir Rybnikov, Peter Steinbach, and Najmeh Mirian. Virtual pulse reconstruction diagnostic for single-shot measurement of free electron laser radiation power. *Phys. Rev. Accel. Beams*, 28:030703, Mar 2025.
- [34] Frank Rosenblatt. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65(6):386–408, 1958.
- [35] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. <http://www.deeplearningbook.org>.
- [36] Auralee Edelen and Xiaobiao Huang. Machine learning for design and control of particle accelerators: A look backward and forward. *Annual Review of Nuclear and Particle Science*, 74(Volume 74, 2024):557–581, 2024.
- [37] J.A. Howell, C.W. Barnes, S.K. Brown, G.W. Flake, R.D. Jones, Y.C. Lee, S. Qian, and R.M. Wright. Control of a negative-ion accelerator source using neural networks. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 293(1):517–522, 1990.
- [38] W.C. Mead, P.S. Bowling, S.K. Brown, R.D. Jones, C.W. Barnes, H.E. Gibson, J.R. Goulding, and Y.C. Lee. Optimization and control of a small-angle negative ion source using an on-line adaptive controller based on the connectionist normalized local spline neural network. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 72(2):271–289, 1992.
- [39] E. Bozoki, A. Friedman, and S. Ramamoorthy. First results with a non-linear digital, orbit feedback system at the nsls. In *Proceedings Particle Accelerator Conference*, volume 4, pages 2696–2698 vol.4, 1995.
- [40] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *International Conference on Learning Representations (ICLR)*, 2015.
- [41] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation, 2015.

- [42] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.
- [43] A.L. Edelen, S. Biedron, J.P. Edelen, and S.V. Milton. First Steps Toward Incorporating Image Based Diagnostics into Particle Accelerator Control Systems Using Convolutional Neural Networks. In *Proc. of North American Particle Accelerator Conference (NAPAC’16), Chicago, IL, USA, October 9-14, 2016*, number 3 in North American Particle Accelerator Conference, pages 390–393, Geneva, Switzerland, Jan. 2017. JACoW. <https://doi.org/10.18429/JACoW-NAPAC2016-TUPOA51>.
- [44] C. Emma, A. Edelen, M. J. Hogan, B. O’Shea, G. White, and V. Yakimenko. Machine learning-based longitudinal phase space prediction of particle accelerators. *Phys. Rev. Accel. Beams*, 21:112802, Nov 2018.
- [45] Yusuke Suetsugu. Machine-learning-based pressure-anomaly detection system for superkekb accelerator. *Phys. Rev. Accel. Beams*, 27:063201, Jun 2024.
- [46] L. Stuhlmann et al. Intelligent retrieval systems for accelerator physics: Matching retrieval approaches to use cases. In *Proc. IBIC 2025*.
- [47] Qing Dai, Rasmus Ischebeck, Maruisz Sapinski, and Adam Grycner. Application of large language models for the extraction of information from particle accelerator technical documentation, 2025.
- [48] Owen Convery, Lewis Smith, Yarin Gal, and Adi Hanuka. Uncertainty quantification for virtual diagnostic of particle accelerators. *Phys. Rev. Accel. Beams*, 24:074602, Jul 2021.
- [49] James Bergstra and Yoshua Bengio. Random search for hyper-parameter optimization. *J. Mach. Learn. Res.*, 13(null):281–305, February 2012.
- [50] Bobak Shahriari, Kevin Swersky, Ziyu Wang, Ryan P. Adams, and Nando de Freitas. Taking the human out of the loop: A review of bayesian optimization. *Proceedings of the IEEE*, 104(1):148–175, 2016.
- [51] M.W. McIntire, T.M. Cope, S. Ermon, and D.F. Ratner. Bayesian Optimization of FEL Performance at LCLS. In *Proc. of International Particle Accelerator Conference (IPAC’16), Busan, Korea, May 8-13, 2016*, number 7 in International Particle Accelerator Conference, pages 2972–2975, Geneva, Switzerland, June 2016. JACoW. doi:10.18429/JACoW-IPAC2016-WEPOW055.

- [52] Ryan Roussel, Auralee L. Edelen, Tobias Boltz, Dylan Kennedy, Zhe Zhang, Fuhao Ji, Xiaobiao Huang, Daniel Ratner, Andrea Santamaria Garcia, Chenran Xu, Jan Kaiser, Angel Ferran Pousa, Annika Eichler, Janis O. Lübsen, Natalie M. Isenberg, Yuan Gao, Nikita Kuklev, Jose Martinez, Brahim Mustapha, Verena Kain, Christopher Mayes, Weijian Lin, Simone Maria Liuzzo, Jason St. John, Matthew J. V. Streeter, Remi Lehe, and Willie Neiswanger. Bayesian optimization algorithms for accelerator physics. *Phys. Rev. Accel. Beams*, 27:084801, Aug 2024.
- [53] Niranjana Srinivas, Andreas Krause, Sham Kakade, and Matthias Seeger. Gaussian process optimization in the bandit setting: no regret and experimental design. In *Proceedings of the 27th International Conference on International Conference on Machine Learning*, ICML'10, page 1015–1022, Madison, WI, USA, 2010. Omnipress.
- [54] Ryan Roussel, Adi Hanuka, and Auralee Edelen. Multiobjective bayesian optimization for online accelerator tuning. *Phys. Rev. Accel. Beams*, 24:062801, Jun 2021.
- [55] Sara Ayoub Miskovich, Willie Neiswanger, William Collocho, Claudio Emma, Jacqueline Garrahan, Timothy Maxwell, Christopher Mayes, Stefano Ermon, Auralee Edelen, and Daniel Ratner. Multipoint-BAX: a new approach for efficiently tuning particle accelerator emittance via virtual objectives. *Mach. Learn. Sci. Tech.*, 5(1):015004, 2024.
- [56] J. Duris, D. Kennedy, A. Hanuka, J. Shtalenkova, A. Edelen, P. Baxevanis, A. Egger, T. Cope, M. McIntire, S. Ermon, and D. Ratner. Bayesian optimization of a free-electron laser. *Phys. Rev. Lett.*, 124:124801, Mar 2020.
- [57] M Borland. Elegant: A flexible sdds-compliant code for accelerator simulation. Technical report, Argonne National Lab., IL (US), 08 2000.
- [58] D. Olsson, A. Bjermo, L. Christiansson, J. Lundh, D. Lundström, E. Mansten, M. Nilsson, E. Paju, L.K. Roslund, and K. Åhnberg.
- [59] N. Blaskovic Kraljevic, L. Isaksson, E. Mansten, and S. Thorin. Beam Dynamics for the MAX IV Transverse Deflecting Cavity Beamline. In *Proc. LINAC'22*.
- [60] MAX IV Laboratory. Detailed design report. <https://www.maxiv.lu.se/beamlines-accelerators/accelerators/accelerator-documentation-2/>, 8 2010.

- [61] Chen Zhang, Riccardo Barbano, and Bangti Jin. Conditional variational autoencoder for learned image reconstruction, 2021.
- [62] Z. D. Farkas, H. A. Hoag, G. A. Loew, and Perry B. Wilson. SLED: A Method of Doubling SLAC’s Energy. In *9th International Conference on High-Energy Accelerators*, pages 576–583, 1974.
- [63] P. Craievich, D. Alesini, S. Di Mitri, M. Ferianis, M. Petronio, and M. Veronese. A transverse RF deflecting cavity for the FERMI@elettra project. In *Proc. DIPAC’07*, pages 168–170. JACoW Publishing, Geneva, Switzerland.
- [64] Paolo Craievich, Marco Petronio, Sandra Biedron, Davide Castronovo, Massimo Dal Forno, S. Di Mitri, Nicolas Faure, Daniele La Civita, Giuseppe Penco, Luca Rumiz, Luca Sturari, Roberto Vescovo, and Defa Wang. Implementation of radio-frequency deflecting devices for comprehensive high-energy electron beam diagnosis. *IEEE Transactions on Nuclear Science*, 62:1–11, 02 2015.
- [65] C.J. Bocchetta, Alessandro Abrami, Enrico Allaria, Ilyas Andrian, D. Bacescu, Laura Badano, Luca Banchi, D. Bulfone, C. Bontoiu, R. Bracco, F. Cargnello, K. Casarin, Massimo Cornacchia, Paolo Craievich, Daniele Cocco, Miltcho Danailov, G. D’Auria, Alexander Demidovich, G. Ninno, and Dino Zangrando. *FERMI@Elettra Conceptual Design Report*. 01 2007.
- [66] S. Spampinati, E. Allaria, L. Badano, S. Bassanese, S. Biedron, D. Castronovo, P. Craievich, M. B. Danailov, A. Demidovich, G. De Ninno, S. Di Mitri, B. Diviacco, M. Dal Forno, E. Ferrari, W. M. Fawley, L. Fröhlich, G. Gaio, L. Giannessi, G. Penco, C. Serpico, C. Spezzani, M. Trovò, M. Veronese, S. V. Milton, and M. Svandrlik. Laser heater commissioning at an externally seeded free-electron laser. *Phys. Rev. ST Accel. Beams*, 17:120705, Dec 2014.
- [67] Z. Huang, M. Borland, P. Emma, J. Wu, C. Limborg, G. Stupakov, and J. Welch. Suppression of microbunching instability in the linac coherent light source. *Phys. Rev. ST Accel. Beams*, 7:074401, Jul 2004.
- [68] Martin Paraliiev, Arturo Alarcon, Vladimir Arsov, Simona Bettoni, Roger Biffiger, Marco Boll, Hans Braun, Alessandro Citterio, Paolo Craievich, Andreas Josef Dax, Philipp Dijkstal, Sladana Dordevic, Eugenio Ferrari, Franziska Frei, Romain Ganter, Zheqiao Geng, Christopher Gough, Nicole Hiller, Martin Huppert, Rasmus Ischebeck, Pavle Juranic, Mario Jurcevic,

- Babak Kalantari, Roger Kalt, Boris Keil, Christoph Kittel, Waldemar Koprek, Daniel Llorente, Florian Löhl, Alexander Malyzhenkov, Fabio Marcellini, Goran Marinkovic, Gian Luca Orlandi, Cigdem Ozkan Loch, Marco Pedrozzi, Eduard Prat, Sven Reiche, Colette Rosenberg, Thomas Schietinger, Serguei Sidorov, Alexandre Trisorio, Carlo Vicario, Didier Voulot, Guanglei Wang, and Riccardo Zennaro. Swissfel double bunch operation. *Phys. Rev. Accel. Beams*, 25:120701, Dec 2022.
- [69] P. Craievich, R. Ischebeck, F. Loehl, G. L. Orlandi, and E. Prat. Transverse Deflecting Structures for Bunch Length and Slice Emittance Measurements on SwissFEL. In *Proc. FEL'13*, pages 236–241. JACoW Publishing, Geneva, Switzerland, 2013.
- [70] P. Craievich et al. Sub-Femtosecond Time-Resolved Measurements Based on a Variable Polarization X-Band Transverse Deflecting Structures for SwissFEL. In *Proc. FEL'17*, pages 499–502. JACoW Publishing, Geneva, Switzerland, 2017.
- [71] P. González Caminal, F. Christie, R. D’Arcy, S. M. Jaster-Merz, R. Assmann, F. Burkart, B. Conrad, H. Dinter, M. Foese, J. Herrmann, M. Hoffmann, M. Hüning, R. Jonas, O. Krebs, S. Lederer, B. Marchetti, D. Marx, J. Mueller, J. Osterhoff, I. Peperkorn, M. Reukauff, H. Schlarb, S. Schreiber, G. Tews, T. Vinatier, M. Vogt, A. de Z. Wagner, S. Wesch, P. Craievich, M. Bopp, H.-H. Braun, A. Citterio, R. Fortunati, R. Ganter, T. Kleeb, F. Marcellini, M. Pedrozzi, E. Prat, S. Reiche, K. Rolli, R. Sieber, A. Grudiev, W. L. Millar, N. Catalan-Lasheras, G. McMonagle, S. Pitman, V. del Pozo Romano, K. T. Szypula, and W. Wuensch. Beam-based commissioning of a novel x -band transverse deflection structure with variable polarization. *Phys. Rev. Accel. Beams*, 27:032801, Mar 2024.
- [72] Nils Maris Lockmann, Christopher Gerth, Bernhard Schmidt, and Stephan Wesch. Noninvasive thz spectroscopy for bunch current profile reconstructions at MHz repetition rates. *Phys. Rev. Accel. Beams*, 23:112801, Nov 2020.
- [73] Philipp Dijkstal. *Temporal FEL Pulse Shaping and Diagnostics at SwissFEL*. PhD thesis, ETH Zurich, 2022.
- [74] Philipp Dijkstal, Alexander Malyzhenkov, Paolo Craievich, Eugenio Ferrari, Romain Ganter, Sven Reiche, Thomas Schietinger, Pavle Juranic, and Eduard Prat. Self-synchronized and cost-effective time-resolved measurements at x-ray free-electron lasers with femtosecond resolution. *Phys. Rev. Res.*, 4:013017, Jan 2022.

- [75] Jonas Björklund Svensson, Francesca Curbis, J. Lundquist, Erik Mansten, S. Thorin, and Sverker Werin. Experimental generation of wakefield accelerator-suitable bunch current profiles by arc-like bunch compressors. *Journal of Physics: Conference Series*, 3266:012005, 06 2026.
- [76] Tolonen, Sofia. Longitudinal dispersion in the second bunch compressor of the MAX IV linear accelerator, 2025. Student Paper.
- [77] Peter J. Rousseeuw and Christophe Croux. Alternatives to the median absolute deviation. *Journal of the American Statistical Association*, 88(424):1273–1283, 1993.