Francesco Giordano

UNIVERSITÀ DEGLI STUDI DI NAPOLI
FEDERICO II

itee
Ph D
iNFORMATION tECHNOLOGY
eLECTRICAL eNGINEERING

A method for a first-order approximation estimation of the longitudinal impedance of a synchrotron component, starting from power loss measurements on the device, is proposed. This method also estimates the resonance frequency and the quality factor of the impedance after the execution of several machine runs, without disconnecting the device. After a detailed description of the method, its suitability is demonstrated through a practical case study using power loss measurements of the Large Hadron Collider (LHC) at the the European Organization for Nuclear Research (CERN).

Then, electromagnetic simulations were used to benchmark recent theoretical models and assess their possibility to compute the two beam power loss. It is shown how beam-induced power loss can largely differ from the single beam case when two beams are present in the same component. Simulation studies are shown in the case of a resonant pillbox cavity. This benchmark also allowed simulating cases, for which the lumped impedance assumption of the available analytical formula may not be valid anymore.
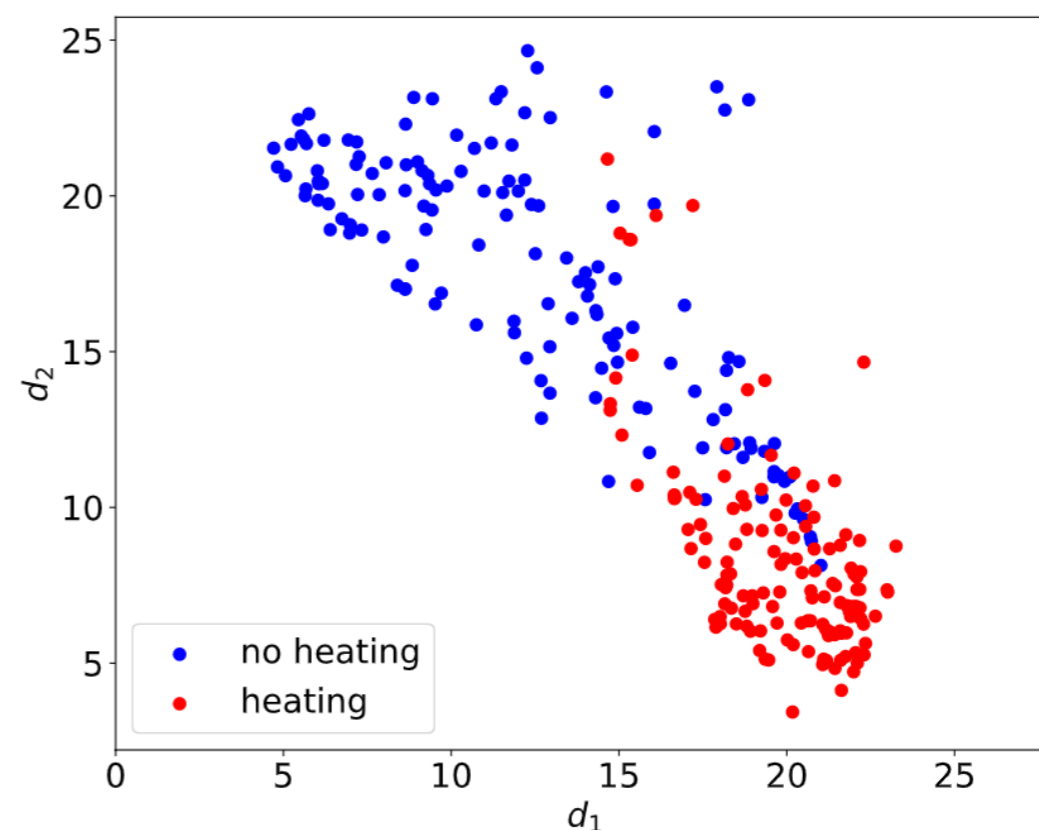
Finally, machine learning models were developed to detect heating from pressure measurements in synchrotron colliders. These results allow to analyse all the pressure measurements in the time avail- able between two consecutive machine runs. Due to the prevalence of noise and the diversity of the behaviours, simple heuristic-based techniques do not achieve high performance. To overcome the limits of simple heuristic-based algorithms several machine learning models have been trained, tested and compared with an heuristic-based approach which is used as base-line. In particular, it is shown for the case of the Large Hadron Collider (LHC) that machine learning models reached better performance both in precision and recall scores with respect to the baseline.

SIMULATION ANALYSIS AND MACHINE LEARNING BASED DETECTION OF BEAM-INDUCED HEATING IN PARTICLE ACCELERATOR AT CERN

FRANCESCO GIORDANO

PH.D. IN INFORMATION TECHNOLOGY AND ELECTRICAL ENGINEERING

Università degli Studi di Napoli Federico II

Ph.D. Thesis
IN
Information Technology and Electrical Engineering

Simulation Analysis and Machine Learning Based Detection of Beam-Induced Heating in Particle Accelerator at CERN

Francesco Giordano

Tutor: Prof. Pasquale Arpaia

Coordinator: Prof. Daniele Riccio

XXXIII Ciclo

Scuola Politecnica e delle Scienze di Base
Dipartimento di Ingegneria Elettrica e Tecnologie dell'Informazione

UNIVERSITY OF NAPLES FEDERICO II

PH.D. THESIS

IN

INFORMATION TECHNOLOGY AND ELECTRICAL ENGINEERING

Simulation Analysis and Machine Learning Based Detection of Beam-Induced Heating in Particle Accelerator at CERN

Supervisors:

Prof. Pasquale Arpaia

Dr. Benoit Salvant

Candidate:

Francesco Giordano

Abstract

A method for a first-order approximation estimation of the longitudinal impedance of a synchrotron component, starting from power loss measurements on the device, is proposed. This method also estimates the resonance frequency and the quality factor of the impedance after the execution of several machine runs, without disconnecting the device. After a detailed description of the method, its suitability is demonstrated through a practical case study using power loss measurements of the Large Hadron Collider (LHC) at the the European Organization for Nuclear Research (CERN).

Then, electromagnetic simulations were used to benchmark recent theoretical models and assess their possibility to compute the two beam power loss. It is shown how beam-induced power loss can largely differ from the single beam case when two beams are present in the same component. Simulation studies are shown in the case of a resonant pillbox cavity. This benchmark also allowed simulating cases, for which the lumped impedance assumption of the available analytical formula may not be valid anymore.

Finally, machine learning models were developed to detect heating from pressure measurements in synchrotron colliders. These results allow to analyse all the pressure measurements in the time available between two consecutive machine runs. Due to the prevalence of noise and the diversity of the behaviours, simple heuristic-based techniques do not achieve high performance. To overcome the limits of simple heuristic-based algorithms, several machine learning models have been trained, tested and compared with an heuristic-based approach which is used as base-line. In particular, it is shown for the case of the Large Hadron Collider (LHC) that machine learning models reached better performance both in precision and recall scores with respect to the baseline.

List of publications

Journal publications:

[1] Arpaia P, De Vito L, Giordano F, Salvant B. *Estimating the impedance of a synchrotron component from beam-induced power loss*, Nuclear Instruments and Methods in Physics Research, 2020.
DOI:10.1016/j.nima.2019.163244.

[2] Arpaia P, Bregliozzi G, Giordano F, Prevete R, Salvant B, et al. *Machine learning for beam dynamics studies at the CERN Large Hadron Collider*, Nuclear Instruments and Methods in Physics Research, 2020. DOI:10.1016/j.nima.2020.164652.

[3] Arpaia P, Giordano F, Prevete R, Salvant B.*Machine Learning-based Heating Detection from Pressure Measurements in the CERN Large Hadron Collider*, Nuclear Instruments and Methods in Physics Research, 2021. DOI: 10.1016/j.nima.2020.164995.

[4] Teofili L, Arpaia P, Giordano F, Lamas I, Rumolo G, Salvant B, Zannini C, Migliorati M. *Wakefunction and Impedance Determination For Two Counter-Moving Particle Beams*, accepted for publication on Physical Review Accelerators and Beams, 2020.

[5] Quartullo D, Arpaia P, Biancacci N, Giordano F, Lamas Garcia I, Mostacci A, Palumbo L, Redaelli S, Salvant B, Teofili L, Migliorati M. *Electromagnetic characterization of the crystal primary collimators for the HL-LHC*, submitted to Nuclear Instruments and Methods in Physics Research, 2020.

Conference proceedings:

[1] Teofili L, Giordano F, Marongiu M, Migliorati M, Lamas I, Nuiry F-X, Romagnoli G, Salvant B. *ANALYSIS ON THE THERMAL RESPONSE TO BEAM IMPEDANCE HEATING OF THE POST LS2 PROTON SYNCHROTRON BEAM DUMP*,10th International Particle Accelerator Conference, Melbourne, Australia, 19 - 24 May 2019, pp.THPTS065.

[2] Teofili L, Giordano F, Calviani M, Carbajo Perez D, Gilardoni S, Mazzacano G, Perillo A. *ANALYSIS ON THE MECHANICAL EFFECTS INDUCED BY BEAM IMPEDANCE HEATING ON THE HL-LHC TARGET DUMP INJECTION SEGMENTED (TDIS) ABSORBER*,September 2018, Journal of Physics Conference Series 1067(6):062011.

[3] Teofili L, Giordano F, Calviani M, Carbajo Perez D, Mazzacano G. *A MULTI-PHYSICS APPROACH TO SIMULATE THE RF-HEATING 3D POWER MAP INDUCED BY THE PROTON BEAM IN A BEAM INTERCEPTING DEVICE*, 9th International Particle Accelerator Conference, Vancouver, Canada, 29 Apr - 4 May 2018, pp.THPAK093.

Technical report:

[1] Iadarola G, Skripka G, Salvant B, Giordano F, et al. *Beam-induced heat loads on the LHC arc beam screens with different beam and machine configurations: experiments and comparison against simulations*, CERN-ACC-2019-0057.

Acknowledgements

I had the opportunity of working in a very exciting science environment at CERN during the last year with very kind people. The team of BE-APB-CEI section has been wonderful and I do believe that I have been a lucky technical student to have been involved in such a special environment.

I am grateful to my supervisor, Dr. Benoit Salvant, who has allowed me to pursue my passion and has always pushed me in doing better. He has always been available ans has always helped me in the most difficult moments.

Il mio più grande ringraziamento va a tutte le persone che mi sono state vicine durante questi tre anni e mi hanno incoraggiato ad andare avanti e fare sempre del mio meglio.

Grazie a i miei genitori per non avermi mai fatto mancare nulla ed a mio fratello Valerio per aver sempre creduto in me.

Grazie a Giulia per essermi sempre stata accanto, per avermi sempre incoraggiato e sopportato. Questa tesi e tutti gli altri obbietivi raggiunti in questi tre anni non sarebbero stati possibili senza di te.

Sono moltissimi gli amici che dovrei ringraziare per il loro supporto durante questo dottorato. Tra tutti il mio grazie più grande va ad Antonio e Nicola.

Contents

List of Abbreviations

CERN	Conseil Européen pour la Recherche Nucléaire
CNN	Convolutional Neural Network
HL	High Luminosity
ML	Machine Learning
NN	Neural Network
LHC	Large Hadron Collider
LR	Logistic Regression
LS1	Long Shutdown 1
LS2	Long Shutdown 2
SPS	Super Proton Synchrotron

A Giulia...

Introduction

The main figure of merit for the high energy particle accelerators performance is luminosity [1]. High luminosity values can be reached by incremental increase of the beam intensity, which means increasing the number of charged particles stored in the accelerator. The strong electromagnetic (EM) field produced by charged particle beams [2] stored in a high-energy, high-intensity accelerator such as the LHC may induce heating of surrounding equipment. Such heating scales quadratically with the bunch intensity [3] and can represent a limit to increasing the performance of the machine: for instance, beam-induced heating was one of the major limitations to reach nominal performance of the machine for the case of LHC during its first Run (2010-2013) [4]. This phenomenon is commonly referred to as *beam–induced radio frequency (RF) heating* [5]. There are various sources of heating in a high-energy accelerator. The major beam-induced phenomena that may induce heating are: direct particle loss on the surrounding accelerator structures [6], electron cloud [7], synchrotron radiation [8], and beam–induced RF heating due to impedance [3].

For each machine component such as collimators, kickers, and magnets, the impedance can be measured with different tests, based on RF measurements (called bench measurements), before installation. Much of the present literature focuses on the global impedance model [9]. Conversely, one of the objective of this thesis focuses on the local impedance information obtained during operation.

In the literature, the heat load and beam induced heating are not used in the estimation of the impedance parameters but they are treated in several application as: the ALBA stripline kicker heatload, the beam induced heating in SOLEIL, the DCCT (DC Current Transformer) and FCT (Fast Current Transformer) heating in PETRA III, the wake field analysis in PEP-II SLAC B-factory, the beam induced RF heating in the LHC and the beam heat load due to geometrical and resistive wall impedance in COLDDIAG [10].

The most common test is the wire measurement method [11]: a wire is stretched inside the component in order to emulate the beam EM field and to assess the S-parameters, which describe the electrical behavior of linear electrical networks when undergoing various steady state stimuli by electrical signals [12]. All these tests require the machine to be shut down for direct access to the component under test or the equipment to be measured before installation or removal from the machine.

Most of the impedances from the accelerator components can be modeled with good approximation as a series of resonator impedances [13]. Three parameters are needed to fully describe an impedance with the resonator model: (i) f_r, the resonance frequency; (ii) Q_r, the quality factor, which characterizes the resonator bandwidth relative to f_r; and (iii) R_s, the shunt impedance, which represents the real part of the impedance at the frequency $f = f_r$. Narrow–band impedances are usually generated by geometric imperfections or unintended cavities, while broad–band impedances are related to the material losses of the machine component. As a result, the impact on the power loss of a narrow– or very narrow–band impedance could be reduced by changing slightly the geometry of the component. This operation can be very difficult if the component is already built. Similarly, for a broad–band impedance, usually the material of the component would have to be changed in order to reduce its impedance, which is also an undesired operation.

Particles injected into the machine are grouped in bunches, spaced by a fixed distance, or analogously at a fixed time interval assuming a constant velocity. The power loss increases with the total number of bunches M injected into the machine [3]. However, the relation between power loss and number of bunches depends on the impedance bandwidth [3]. In particular it is commonly assumed that [14]:

- For narrow–band impedances, $P_{loss} \propto M^2$.

- For broad–band impedances, $P_{loss} \propto M$.

The current literature does not specify the effective boundaries of the impedance bandwidth where the above relations are valid.

This thesis focuses on three main aspects of the beam–induce heating: the impedance estimation from power loss measurements, the impact of two counter-rotating beams on the beam–induced heating and the machine learning detection of the beam–induced

heating from pressure measurements. The first subject of this thesis proposes a method to provide a first-order approximation of Q_r and f_r by measuring the power loss, while the parameters of the beam in the machine are varied. The Q_r and f_r values, modelled as a resonator, are assessed from measurements (or estimations) of the power loss only.

All the above considerations are valid when there is a single beam circulating in a machine component. However, in some devices of circular colliders there is an extra complication due to the presence of two counter-rotating particle beams. In recent circular colliders, the two beams circulate by design in two separate vacuum chambers to avoid beam-beam effects. However, in the collision regions and, sometimes in other components, they have to transit in the same vacuum chamber. In this case, the particles of one beam move in opposite direction with respect to the particles of the other beam. The main studies of the interaction between two beams are: Pellegrini [15] and Wang [16] studied longitudinal and transverse counter-moving two beam instabilities linked to resonant modes for the Large Electron Positron storage ring (LEP) [17]. However their work is focused on the long range effects and does not describe what happens when both counter-moving beams are inside the cavity. Zimmerman [18] discussed the transverse resistive wall wakefield problem for two counter-moving beams. Zannini et al. [3, 19] and Grudiev [20] presented models to compute the RF-heating induced in a vacuum chamber traversed by the counter-moving beams.
Some of these 2-beam components have presented beam induced heating damage or excessive outgassing during operation, and it is important to understand how the two–beam power loss differs from the single beam case inside a resonant cavity. The second subject of this thesis is the better understanding of the two–beam power loss inside a resonant cavity and the limits of its analytical [3] formula with electromagnetic simulations.

The heating can be studied and monitored in different ways. One way of measuring heating is by means of temperature probes. In the LHC, this is commonly done with PT100 [21] probes in the LHC ring, and more recently also with optical fibres

inside the CMS detector [22]. Heating in the LHC can also by monitored by measuring the required cooling power to compensate for the beam–induced power loss [23]. The temperature monitoring system in LHC is focused on known critical components from heating point of view, and large fractions of the accelerator are not equipped with dedicated temperature monitoring. Heating of surrounding equipment can also be detected indirectly with pressure measurements [24]. The idea is based on the outgassing phenomenon [25] produced by the increase in temperature in a high-vacuum environment. The outgassing is directly observed as pressure increase by means of vacuum gauges as for the case of the injection protection device (TDI), during the 2012 LHC run [4].

The LHC vacuum monitoring system has more than 1200 vacuum gauges distributed along the LHC circumference [26]. With respect to the direct temperature monitoring, the vacuum-based indirect measurement is more dense and systematic. However, analysing the pattern of each pressure measurement is not a trivial task. Expert knowledge is needed to identify heating from the behaviour of a pressure measurement. Moreover, due to the large amount of pressure data retrieved after each fill of the LHC, classifying each pressure measurement as indicating heating or not is a tedious and time-consuming task. A fill is an LHC operation cycle, starting with the gradual injection of the two counter-rotating beams and finishing with the beam dump. During the LHC operation, each fill is injected in the LHC roughly 2 hours after the dump of the previous one. In such a short time, all the 1200 pressure measurements cannot be analysed manually in order to prevent heating that could cause damage during the next fill. Therefore, the heating detection process can be improved by automatising the classification of the pressure measurements, providing a significant gain of time with very high performance. Moreover, such a tool can be also used to retrieve old data with heating indication in order to double check that no abnormal heating was missed in the past runs.

Finally this thesis proposes a Machine Learning-based heating detection from pressure measurements for synchrotron colliders with a test case on the CERN Large Hadron Collider.

The structure of the thesis is as follows:

- **Chapter 1: Beam–induced heating**. Introduction to the beam–induce heating

with a particular focus on the power loss due to impedance. Review of the literature and derivation of the formulae used in this thesis.

- **Chapter 2: Machine learning for heating detection**. Introduction to the pressure measurements in the LHC followed by a description of the machine learning techniques applied in this thesis.

- **Chapter 3: Impedance of a synchrotron component from beam-induced power loss**. Methodology of the impedance estimation of a synchrotron component from beam induced heating. Derivation of the formulae adopted in the method. Detailed explanation on how to derive the resonance frequency f_r and the Q factor of an impedance based on power loss measurements.

- **Chapter 4: Power loss simulations**. Methodology applied on power loss simulation studies. Description of the resonant cavity model used for power loss simulation. Single beam power loss simulations set up both for single bunch and multi-bunch scenarios. Two beam power loss simulations with comparison with the single beam case. Investigation of the dependence on the phase shift of the two-beam power loss with comparison between simulation studies and two-beam power loss equation.

- **Chapter 5: Heating detection from pressure measurements**. Methodology of the heating detection from pressure measurements. Description of a simple heuristic algorithm used as base model. Presentation of the preprocessing techniques that helped in improving the models performance. Detailed description of the machine learning models adopted.

- **Chapter 6: Results of impedance of a synchrotron component from beam-induced power loss**. Results of the impedance estimation technique with an example of the implementation of the methodology in the LHC.

- **Chapter 7: Results of power Loss simulations**. Results of the simulation studies for both single beam and two-beam power loss. Analysis of the dependence of the two-beam power loss on the cavity geometry.

- **Chapter 8: Results of heating detection from pressure measurements in the CERN Large Hadron Collider**. Results of the heating detection from pressure measurements. Analysis of the impact of the preprocessing technique on the machine learning models. Comparison between machine learning models and simple heuristic base model for the case of the LHC..

Part I

Background

Chapter 1

Beam-induced heating

1.1 Overview

This chapter focuses on the beam–induced RF heating due to impedance. First, the concept of wake function and beam coupling impedance is introduced focusing in particular on the case of the longitudinal plane. In the longitudinal plane, the relationship between the beam coupling impedance and the energy loss by a bunch of particles is further discussed. After that, the power loss is defined as well as how to compute it numerically for the single beam case. The concept of a filling scheme is then introduced for a circular collider and in particular for the LHC.

Finally the two-beam power loss is introduced with its analytical derivation. Also the assumption that leads to the numeric computation and its expression are presented.

1.2 Wakefields

The wake function describes the electromagnetic interaction of the beam with the surrounding environment. The reference to a "wake" originates from the similarities of the electromagnetic fields generated by a charged particle beam to the wake left by a boat, which is travelling into the see. This reference indeed fits particularly well with the beam wakefields because - due to the ultrarelativistic speed of the particle in the LHC - the fields all lie behind the source, like the waves are always behind the boat, which generates them.
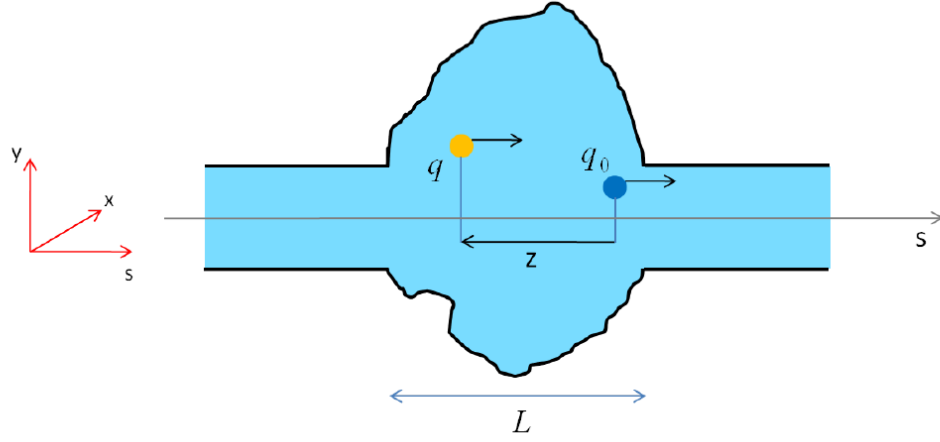
FIGURE 1.1: Source and test charge particles [27]

The electromagnetic (EM) problem is posed by setting the Maxwell's equations with the beam as source term and boundary conditions given by the structure, in which the beam propagates. In order to understand the wakefields, it is common to consider the simple case of a single charge that acts like a source followed by a test charge (Fig: 1.1). In the following, these assumptions are made:

- Wakefields are considered superimposed to the external fields (e.g. dipoles, quadrupoles, RF cavities).

- The particles are moving at the same speed $v = \beta c$, so they keep the same distance between each other (Rigid motion approximation).

The wakefields are related to the momentum variation :

$$\Delta \mathbf{p} = \int_{-\infty}^{+\infty} \mathbf{F}(x_s, y_s, z_s, x_t, y_t, z_t) dt \tag{1.1}$$

where (x_s, y_s, z_s) are the source's coordinates and (x_t, y_t, z_t) are the test particle coordinates. The force acting on the test particle is the Lorentz's force:

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \tag{1.2}$$

E and **B** are obtained by solving the Maxwell's problem. The wakefields is defined as [28]:

$$W(x_s, y_s, z_s, x_t, y_t, z_t, t) = -\frac{1}{q_s q_t} \int_{-\infty}^{+\infty} \mathbf{F} \cdot \mathbf{dz} \tag{1.3}$$

It is possible to observe that this definition is strongly related to the momentum if we consider that $\mathbf{dz} = \mathbf{v}dt$ so [29]:

$$W = -\frac{v}{q_s q_t} \int_{-\infty}^{+\infty} \mathbf{F} \cdot \mathbf{t} \tag{1.4}$$

that is related to the momentum (given the rigid motion approximation):

$$W(x_s, y_s, z_s, x_t, y_t, z_t, t) = -\frac{v}{q_s q_t} \Delta p \tag{1.5}$$

1.2.1 Longitudinal wakefields and impedance

The wakefields can be decomposed on the longitudinal and on the transverse planes. By naming W_l the longitudinal plane and W_t the transverse plane. It is common to distinguish those two components because they lead to different beam dynamics effects. The following will focus on the longitudinal component since only the longitudinal impedance plays a role in the power loss computation [3]. As for the wakefield, also the electromagnetic field of the electromagnetic problem inside the beam pipe can be decomposed as follows:

$$\mathbf{E} = E_l \hat{\mathbf{z}} + E_t \hat{\mathbf{t}}, \qquad \mathbf{B} = B_l \hat{\mathbf{z}} + B_t \hat{\mathbf{t}} \tag{1.6}$$

For the longitudinal case $\mathbf{v} = v\hat{\mathbf{z}}$ implying that the magnetic field does not give any contribution to the longitudinal component of the Lorentz force (Eq. 1.2), so neither to the momentum , nor to the wakefields. In this case it is possible to write:

$$W_l(x_s, y_s, z_s, x_t, y_t, z_t, t) = -\frac{v}{q_s q_t} \Delta p_l \tag{1.7}$$

with

$$\Delta p_l = \int_{-\infty}^{+\infty} q_t E_l(x_s, y_s, z_s, x_t, y_t, z_t) dt. \tag{1.8}$$

The longitudinal coupling impedance is then defined as:

$$Z_l(x_s, y_s, x_t, y_t, w) = \int_{-\infty}^{+\infty} W_l(x_s, y_s, z_s, x_t, y_t, z_t)e^{jwt}dt \tag{1.9}$$

with $t = \frac{s}{v}$ and $s = z_t - z_s$. It is possible to observe with the assumption of rigid motion approximation that this impedance only depends on the relative distance between the particles, s, and not on their absolute position z_s, z_l.

1.2.2   Resonator impedance model

The resonator model is widely used to model impedances for the computation of power loss [30, 31, 32]. The resonator impedance (Z_{res}) formula is

$$Z_{res}(f) = \frac{R_s}{1 + jQ_r(\frac{f}{f_r} - \frac{f_r}{f})}, \tag{1.10}$$

where Q_r is the quality factor, related to the width of the real part of the impedance, f_r is the resonance frequency representing the impedance position, R_s is the shunt impedance, and $j^2 = -1$.

Figures 1.2a and 1.2b show the real and the imaginary parts of the impedance of the resonator model, respectively, for a given value of f_r and different values of the Q–factor.

Narrow and broad band impedance

The beam impedance is defined in the frequency domain, therefore the *bandwidth* concept is widely used. In this manuscript, it is common to distinguish between the *narrow band* and the *broad band* impedance.:

- narrow band : an impedance that covers just few lines of the power spectrum.

- broad band : an impedance that covers all the lines of the power spectrum.

A narrow band impedance usually has $Q_r > 10^3$ while a broad band has $Q_r = 0$. For intermediate values of Q_r it there is not a defined nomenclature for the bandwidth. The lines of the spectrum depends on the spacing between two following bunches.
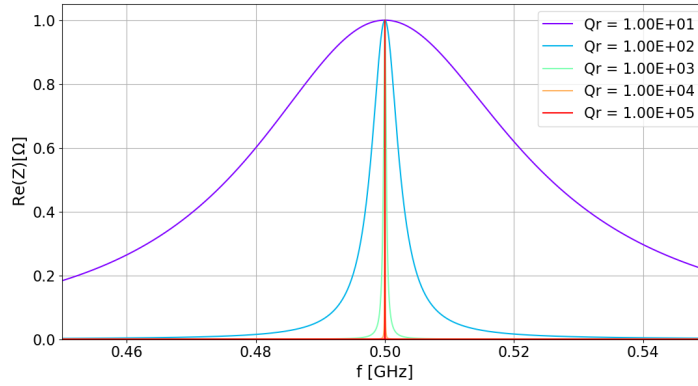
(A) Real part
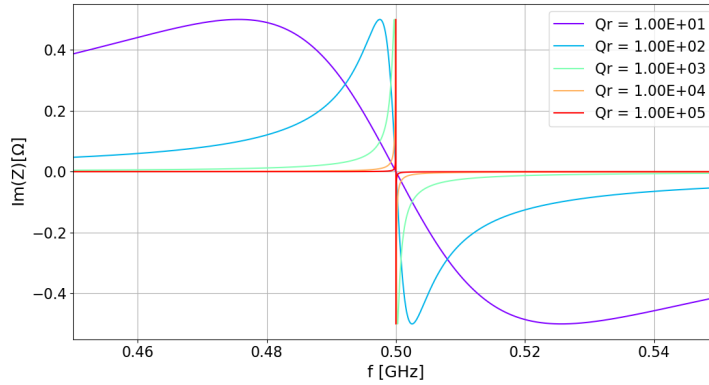
(B) Imaginary part

FIGURE 1.2: Real part (a) and imaginary part (b) of the resonator model impedance for $f_r = 500MHz$ and for different values of Q_r.

1.3 Filling schemes

The *filling scheme* defines the time distribution of the bunches of particles in the machine. Considering the LHC, it has 3564 slots (buckets), in which the bunches could be injected. The LHC can accelerate protons and ions. In this manuscript it is assumed that the particles are protons. The filling scheme is the map of the slots actually occupied by the bunches. It is encoded as 1, if a slot is occupied by a bunch, and 0 otherwise. The filling scheme can be appreciated by observing the longitudinal time

distribution of the beam (Fig. 4.8). The longitudinal time distribution is the longitudinal proton density (i.e. number of protons per unit of time). The longitudinal direction is classically measured either in meters or in seconds in the LHC. The longitudinal time distribution is referred as $\lambda(z)$ or $\lambda(t)$ depending on the considered unit.
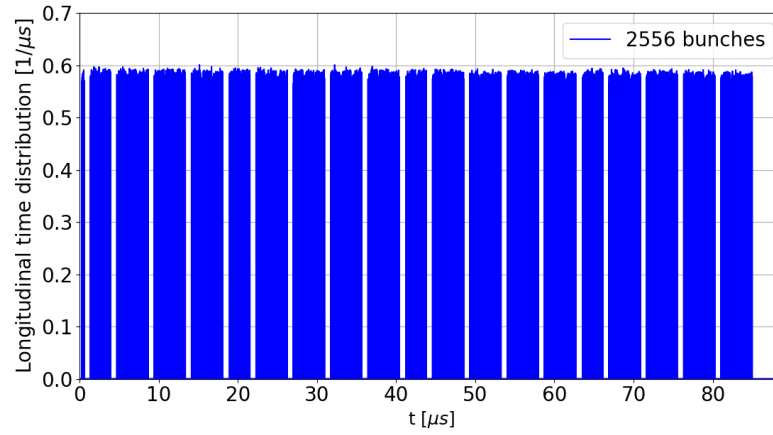

FIGURE 1.3: Longitudinal time distribution of the beam for the fill number 5979 of a LHC run of the 21-07-2017 in the time window from 18:44:26 to 18:48:15 (before reaching collision energy). The horizontal axis can be converted in space as the particles are moving at constant speed.

The bunch longitudinal profile is characterized by two main parameters: *bunch length* and *bunch shape*. The bunch shape is the longitudinal model that best fits its shape. Common bunch shapes that have the same full width half maximum are:

- Gaussian: $s(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(t)^2/2\sigma^2}$.

- q-Gaussian: $s(t) = \frac{32}{5\pi H}(1 - \frac{4t^2}{H^2})^{2.5}$ with $H = 2\sigma\sqrt{\frac{2log(2)}{1-2^{-0.4}}}$.

- Cosine: $s(t) = \frac{1}{\tau_c}[\cos\left(\frac{\pi t}{2\tau_c}\right)]^2$ with $\tau_c = 2.77\sigma$.

- Parabolic: $s(t) = 1 - \left(\frac{t}{\sigma}\right)^2$.

where σ is referred to the Gaussian r.m.s. Assuming that the bunch longitudinal distribution is Gaussian, the bunch length is defined as 4σ of such a Gaussian distribution.

1.4   Power loss

The power lost by the beam in a circular collider depends on whether one or two beams are circulating in the machine. In the following the power loss derivation is presented for both the single beam and two counter-rotating beam cases.

1.4.1   Single beam power loss

It is possible to derive the single beam power loss formula starting from the energy loss under the rigid motion approximation. Considering $v \simeq c$ the energy can be written as $E \simeq mv^2$, so it is possible to formulate the longitudinal wakefield from the Eq. 1.7 as [27]:

$$W_l(z_t - z_s) = -\frac{\Delta E(z_t - z_s)}{q_s q_t} \tag{1.11}$$

and:

$$\Delta E(z_t - z_s) = -W_l(z_t - z_s)q_s q_t \tag{1.12}$$

In order to relate the bunch distribution to the charged particle, the longitudinal particle distribution $\lambda(z)$ is considered. It represents the number of charges per unit length, normalized to the average number of charges in the bunches N_b : the integral of $\lambda(z)$ along a single bunch results to be 1. So, the electrical charge in a bunch slice dz' at a coordinates z' will be $\lambda(z')eN_b dz'$. It is now possible to consider the infinitesimal increment of energy loss as :

$$dE(z_t - z_s) = -W_l(z_t - z_s)\lambda(z_s)eN_b dz_s \lambda(z_t)eN_b dz_t \tag{1.13}$$

which can be rearranged as

$$dE(z_t - z_s) = -W_l(z_t - z_s)\lambda(z_s)\lambda(z_t)e^2 N_b^2 dz_s dz_t \tag{1.14}$$

where $dE(z_t - z_s)$ is the energetic variation given by a source slice of the bunch at z_s acting on a test slice at z_t. The whole bunch energetic variation is given by [33]:

$$\Delta E = -\int_{-\infty}^{+\infty} dz_t \int_{-z_t}^{+\infty} dz_s W_l(z_t - z_s)\lambda(z_s)\lambda(z_t)e^2 N_b^2, \tag{1.15}$$

since only preceding slices can affect a test slice, at ultra-relativistic velocities. Remembering that a convolution integral is equal to the inverse Fourier transform of a product of the Fourier transforms:

$$\Delta E = -\int_{-\infty}^{+\infty} dz_t \int_{-\infty}^{+\infty} \frac{dw}{2\pi} Z_l(w)\Lambda(w)e^{-jw\frac{z_t}{v}}\lambda(z_t)e^2 N_b^2 \tag{1.16}$$

where $\Lambda(w) = F_T[\lambda(z)]$. Repeating the same procedure integrating in dz_t,

$$\begin{aligned}
\Delta E &= -\frac{e^2 N_b^2}{2\pi}\int_{-\infty}^{+\infty} dw Z_l(w)\Lambda(w)\left[\int_{-\infty}^{+\infty} dz_t e^{jw\frac{z_t}{v}}\lambda(z_t)\right]^* \\
&= -\frac{e^2 N_b^2}{2\pi}\int_{-\infty}^{+\infty} dw Z_l(w)\Lambda(w)\Lambda^*(w) \\
&= -\frac{e^2 N_b^2}{2\pi}\int_{-\infty}^{+\infty} dw Z_l(w)|\Lambda(w)|^2
\end{aligned}$$

because the wakefields are real, from the Fourier Transform proprieties results:

$$Z_l^*(w) = Z_l(-w) \tag{1.17}$$

that tells that $Im[Z_l(w)]$ is odd. Now, considering that $\lambda(z)$ is real, $|\Lambda(w)|^2$ is even so their product integrated on a symmetric domain is zero. Observing that the energy loss is real, so the imaginary part of the impedance does not contribute to its calculation. Finally the energy loss is obtained:

$$\Delta E = -\frac{e^2 N_b^2}{2\pi}\int_{-\infty}^{+\infty} dw Re[Z_l(w)]|\Lambda(w)|^2 \tag{1.18}$$

The power loss is derived by dividing ΔE by the time interval in which the energy has been lost from the beam.

Numeric computation of the single beam power loss

In the one turn case, the dE gained or lost from the test particle in the position z_t due to the source slice dz_s is given by:

$$dE(z_t, z_s) = e^2 N_b M \lambda(z_s) W(z_s - z_t) dz_s$$

where $\lambda(z_s)$ is the longitudinal profile of the beam that integrated over z_s and multiplied by $e N_b M$ gives the total charge of the beam.

For the multi turn case, it has to be taken into account that each particle sees also the wake fields of all the previous turns of the beam in the machine, so:

$$dE(z_t, z_s) = e^2 N_b M \lambda(z_s) \sum_{k=-\infty}^{+\infty} W(z_s - z_t + kC) dz_s,$$

where C is the length of the accelerator. The energy seen by the test particle is obtained by integrating the energy over all the preceding source slices from the test particle to $+\infty$ following the causality principle:

$$E(z_t) = e^2 N_b M \int_{-z_t}^{+\infty} dz_s \lambda(z_s) \sum_{k=-\infty}^{+\infty} W(z_s - z_t + kC)$$

Because the integral is zero for $z < z_t$, the lower bound of the integral can be extended to $-\infty$. To obtain the total energy gained or lost by the beam the energy $E(z_t)$ has to be divided by the test particle charge and integrated all over the beam by multiplying for the charge of the test particles:

$$E_{beam} = \frac{1}{e} \int_{-\infty}^{+\infty} dz_t e N_b M \lambda(z_t) E(z_t)$$

$$= e^2 N_b^2 M^2 \int_{-\infty}^{+\infty} dz_t \lambda(z_t) \int_{-\infty}^{+\infty} dz_s \lambda(z_s) \sum_{k=-\infty}^{+\infty} W(z_s - z_t + kC)$$

Now remembering the mathematical identity

$$\sum_{k=-\infty}^{+\infty} W(z_s - z_t + kC) = \sum_{k=-\infty}^{+\infty} W(z_s - z_t) \circledast \delta(z_s - z_t + kC),$$

where $\circledast$ is the convolution operator, results in:

$$E_{beam} = e^2 N_b^2 M^2 \int_{-\infty}^{+\infty} dz_t \lambda(z_t) \int_{-\infty}^{+\infty} dz_s \lambda(z_s) \sum_{k=-\infty}^{+\infty} W(z_s - z_t) \circledast \delta(z_s - z_t + kC)$$

$$= e^2 N_b^2 M^2 \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} dz_s dz_t \lambda(z_s) \lambda(z_t) \sum_{k=-\infty}^{+\infty} W(z_s - z_t) \circledast \delta(z_s - z_t + kC)$$

Because the convolution is also the inverse Fourier Transform of the product of the Fourier transforms, the beam energy E_{beam} becomes:

$$E_{beam} = e^2 N_b^2 M^2 f_0 \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} dz_s df \lambda(z_s) \Lambda(f) e^{j\frac{2\pi f z_t}{c}} \sum_{p=-\infty}^{+\infty} Z(f) \delta(f + p f_0) e^{j\frac{2\pi f(z_s - z_t)}{c}}$$

$$= e^2 N_b^2 M^2 f_0 \int_{-\infty}^{+\infty} df \Lambda(f) \sum_{p=-\infty}^{+\infty} Z(f) \delta(f + p f_0) \int_{-\infty}^{+\infty} dz_s \lambda(z_s) e^{j\frac{2\pi f z_s}{c}}$$

$$= e^2 N_b^2 M^2 f_0 \int_{-\infty}^{+\infty} df \Lambda(f) \Lambda^*(f) \sum_{p=-\infty}^{+\infty} Z(f) \delta(f + p f_0)$$

$$= e^2 N_b^2 M^2 f_0 \sum_{p=-\infty}^{+\infty} \int_{-\infty}^{+\infty} df |\Lambda(f)|^2 Re[Z(f)] \delta(f + p f_0)$$

$$= e^2 N_b^2 M^2 f_0 \sum_{p=-\infty}^{+\infty} |\Lambda(p f_0)|^2 Re[Z(p f_0)]$$

Where in the first step it has been considered that:

$$F[\sum_{k=-\infty}^{+\infty} W(z_s - z_t) \circledast \delta(z_s - z_t + kC)] = f_0 \sum_{p=-\infty}^{+\infty} Z(f) \delta(f + p f_0)$$

with $F[W(z_s - z_t)] = Z(f)$.

The first integral goes away because of the Fourier transformation of $\lambda(z_s)$ while the second one goes away because of the property of the Dirac's delta.

Therefore the power loss per turn is:

$$P_{loss} = E_{beam} f_0$$

$$= e^2 N_b^2 M^2 f_0^2 \sum_{p=-\infty}^{+\infty} |\Lambda(pf_0)|^2 Re[Z(pf_0)]$$

$$= I_{beam}^2 \sum_{p=-\infty}^{+\infty} |\Lambda(pf_0)|^2 Re[Z(pf_0)] \tag{1.19}$$

with $w_b = Mw_0$. From this formula it is very interesting to note that for a broad band impedance, P_{loss} is linear with M while for a really narrow-band impedance, P_{loss} is quadratic with M [34]. The behaviour of the P_{loss} will be described in detail in the following chapters.

It is important note that this formula doesn't clearly explains what will happen with a resonator impedance with quality factors with values intermediate between broad band and narrow band.

One aim of the this thesis is also to better understand the impact of the impedance bandwidth on Eq. 1.19.

1.4.2 Two–beam power loss

The two beams power loss can be derived similarly as the single beam case [19].

Considering two counter rotating beams named beam 1 and beam 2, it is possible to define the following wakefields:

- $W_{11}(s, z)$: effect of charge in beam 1 evaluated on beam 1.

- $W_{22}(s, z)$: effect of charge in beam 2 evaluated on beam 2.

- $W_{21}(s, z)$: effect of charge in beam 2 evaluated on beam 1.

- $W_{12}(s, z)$: effect of charge in beam 1 evaluated on beam 2.

With these definitions it is possible to write the energy loss per turn on beam 1 by rewriting Eq. 1.15 as:

$$\Delta E_1 = e^2 N_b^2 \int_{-\infty}^{+\infty} dz_t \lambda_1(z_t) \int_{-z_t}^{+\infty} dz_s (W_{11}(z_t - z_s)\lambda_1(z_s) - W_{21}(z_t - z_s)\lambda_2(z_s)), \quad (1.20)$$

where the minus sign inside the integral refers to counter rotating beams and plus sign for beams rotating in the same direction. $\lambda_1(z)$ and $\lambda_2(z)$ are the longitudinal distribution of beam 1 and beam 2. Analogously for beam 2 it is possible to write:

$$\Delta E_2 = e^2 N_b^2 \int_{-\infty}^{+\infty} dz_t \lambda_2(z_t) \int_{-z_t}^{+\infty} dz_s (W_{22}(z_t - z_s)\lambda_2(z_s) - W_{12}(z_t - z_s)\lambda_1(z_s)). \quad (1.21)$$

The total energy loss is simply:

$$\Delta E = \Delta E_1 + \Delta E_2.$$

Assuming the same longitudinal distribution, at a certain longitudinal distance s from the interaction point (IP) between the beams,

$$\lambda_1(z) = \lambda(z), \lambda_2(z) = \lambda(z - 2s),$$

by defining the phase shift $\tau_s = 2s/c$ represented in Fig. 1.4 it is possible to rewrite ΔE_1 and ΔE_2 in the Fourier domain as follows:

$$\Delta E_1 = \frac{e^2 N_b^2}{2\pi} \int_{-\infty}^{+\infty} dw |\Lambda(w)|^2 [Z_{11}(w) - Z_{21}(w)e^{-jw\tau_s}] \quad (1.22)$$

$$\Delta E_2 = \frac{e^2 N_b^2}{2\pi} \int_{-\infty}^{+\infty} dw |\Lambda(w)|^2 [Z_{22}(w) - Z_{12}(w)e^{jw\tau_s}]. \quad (1.23)$$

Having $Z_{21}(w) = Z_{12}(w)$ due to the reciprocity theorem [35], the total energy loss by the two beam results to be:

$$\Delta E = \frac{e^2 N_b^2}{2\pi} \int_{-\infty}^{+\infty} dw |\Lambda(w)|^2 [Re[Z_{11}(w) + Z_{22}(w)] - 2Re[Z_{12}(w)cos(w\tau_s)]]. \quad (1.24)$$
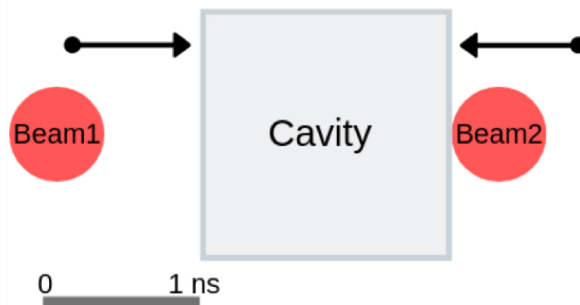
FIGURE 1.4: Two beam approaching the same cavity with a phase shift (τ_s) of $1\,\text{ns}$. In particular beam 1 will enter the cavity $1\,\text{ns}$ after beam 2.

1.4.3   Numeric computation of the two–beam power loss

The two beam power loss formula can be numerically computed given the following approximation [19, 3]:

- Ratio $\frac{L}{s}$ small. In particular the length of the structure L where the impedance is computed has to be at least one order of magnitude smaller than the distance s from the IP. This means that the transit time in the cavity (i.e. the time the bunch takes to pass trough the cavity) has to be at least one order of magnitude smaller than the phase shift τ_s.

- Top-bottom or right-left transverse structure symmetry.

- The wakefield can be approximated as $W(z, x, x_0) = W^0(z) + W^{1d}(z)(x + x_0)$.

Given these approximation the formula can be written as [19]:

$$\Delta W(s) = (2eN_bMf_0)^2 \sum_{p=0}^{\infty} |\Lambda(p\omega_0)|^2 \left\{ \text{Re}\left[Z_{||}^0(p\omega_0)\right] + \right.$$

$$\left. [\Delta y_1(s) + \Delta y_2(s)] \, \text{Re}\left[Z_{||}^1(p\omega_0)\right] \right\} (1 - \cos p\omega_0\tau_s) \,. \tag{1.25}$$

1.5   Pillbox cavity

The pillbox cavity is a simple resonant cavity whose geometry is represented in Fig. 1.5.
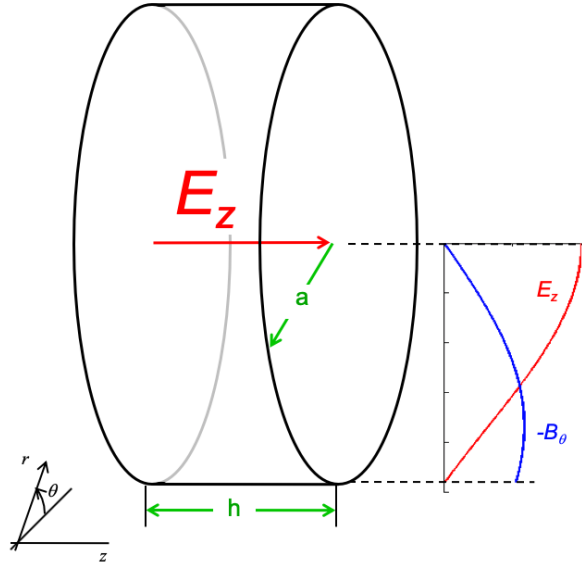
FIGURE 1.5: Pillbox resonant cavity. a is the radius of the cavity and h is its length. E_z is the electric field along the z direction [36].

The modes of a pillbox cavity are expressed in cylindrical coordinates. In particular, TM_{mnp} denotes an EM mode with only electrical component along the direction z while TE_{mnp} denotes an EM mode with only magnetic component along the direction z [37]. The m, n, p notation is explained in the following. The fundamental mode of such a cavity is the TM_{010} with a resonance frequency of $a = 0.383\lambda$ [37].
The Q is given by:

$$Q = \frac{a}{\delta}(1 + \frac{a}{h})^{-1},\tag{1.26}$$

with $\delta = \sqrt{\frac{2}{2\pi f_r \sigma \mu}}$.

In Eq. (1.26), a is the radius of the cavity, h is its length, f_r is the frequency of the electromagnetic field, σ is the conductivity of the material of which the cavity is made and μ is the magnetic permeability.

In Fig. 1.6 is represented the mode chart of a pillbox cavity. The only modes that could couple with the beam are the TM modes due to their component of the electric field in the z direction [38, 37] where the beam is circulating. For a pillbox cavity the electric field along the z direction and the magnetic field along the ϕ direction can be written as [37]:

$$E_z = Acos(\frac{p\pi z}{h})e^{-i2\pi f_r t},,\tag{1.27}$$
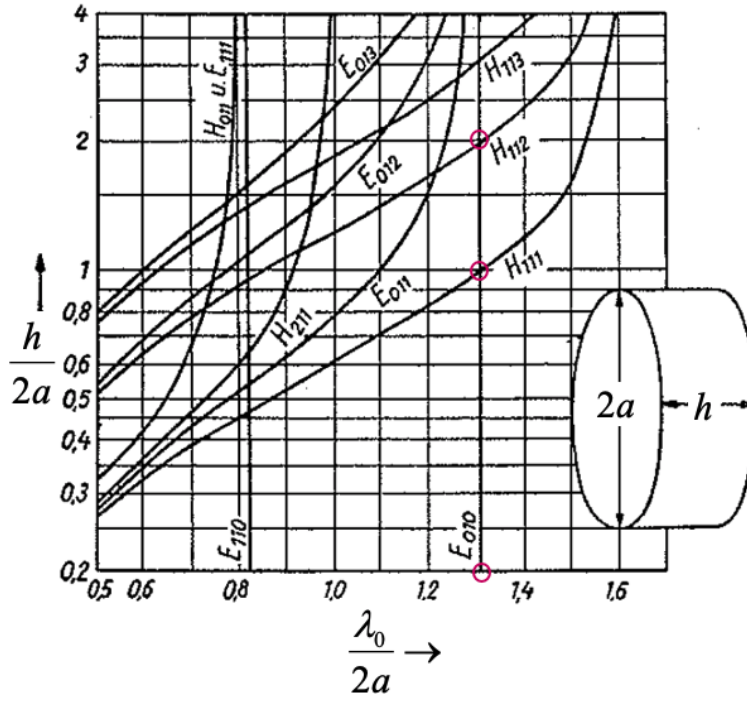
FIGURE 1.6: Pillbox mode chart [36]

$$H_\phi = Bcos(\frac{p\pi z}{h})e^{-i2\pi f_r t}, \tag{1.28}$$

where only the dependence on z and t is explicit. In these equations, A and B are constants with respect to z, h is the length of the cavity and p is an integer value. In the TM_{mnp} notation m is the number of variation of field of the azimuthal variable ϕ. n is the number of nulls in E_z along the radial direction. p is the number of nodes of E_z along the z-axis. In Fig. 1.7 the cases of m, n, p equal to zero and one are reported.

The power loss on the wall of the cavity for the modes TM_{0np} is [38, 37]:

$$P_{wall} = \frac{R_s}{2} \int_{wall} |H_\phi|^2, \tag{1.29}$$

where R_s is the surface resistance of the wall defined as:

$$R_s = \frac{1}{\delta\sigma}.$$
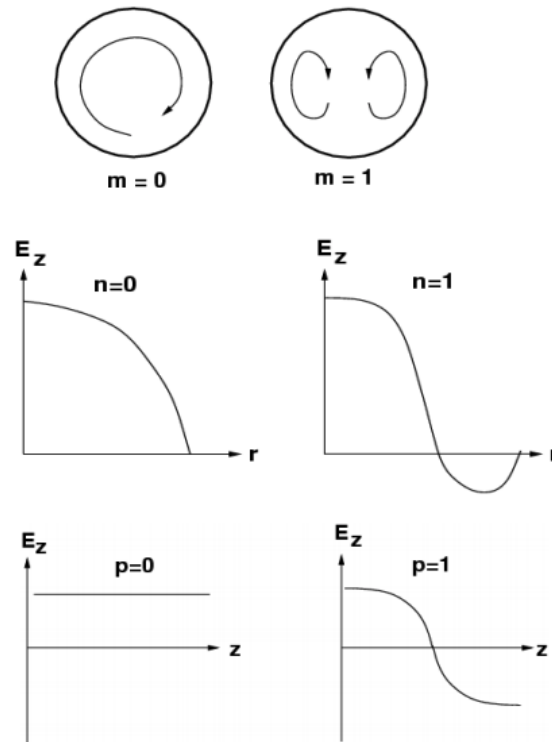
FIGURE 1.7: Representation of the mode indexes m, n, p of a pillbox cavity [38, 37]

1.5.1   CST Studio Suite

CST Studio Suite® is a 3D EM analysis software package for designing, analyzing and optimizing electromagnetic (EM) components and systems [39]. It is widely used in accelerator physics to perform EM simulations. In particular it can compute the impedance of a device given the geometry and can also compute the power lost in the walls of a lossy material.

Chapter 2

Machine learning for heating detection

2.1 Overview

The vacuum monitoring system of the LHC is the largest in the world [26]: it consists of more than 1200 vacuum gauges.

Heating to machine equipment can lead to an anomalous pressure increase, which would be observed in a pressure measurement pattern [40]. A heating example is represented in Fig 2.1: in the top plot the pressure increases from the middle of the LHC fill without any correlation with the beam intensity, beam energy and bunch length. In this specific case, heating was confirmed to be caused by beam induced heating in injection protection collimator (TDI) [5]. Thanks to that observation, TDI was completely redesigned for the current long shutdown LS2 [4].

Unfortunately, heating cases are usually not so easy to spot. In general the pressure increase can happen anywhere along the LHC fill and it can last either a few minutes or several hours. Moreover there can be noise during the acquisition process or the cooling system cycle may induce a seasonality on the measurement thus often an expert is required to identify signs of heating. For the LHC, most of the pressure measurements do not indicate heating: among the 1200 vacuum gauge producing data every fill of the LHC, usually only 3 or 4 of them present heating and the gauges that present heating are not necessarily the same for every LHC fill. This makes manual data identification extremely tedious, slow and prone to miss important indications of heating. In this chapter the structure of the pressure measurement data is first described, then the best evaluation metrics are discussed and finally the ML techniques applied in this thesis are presented.
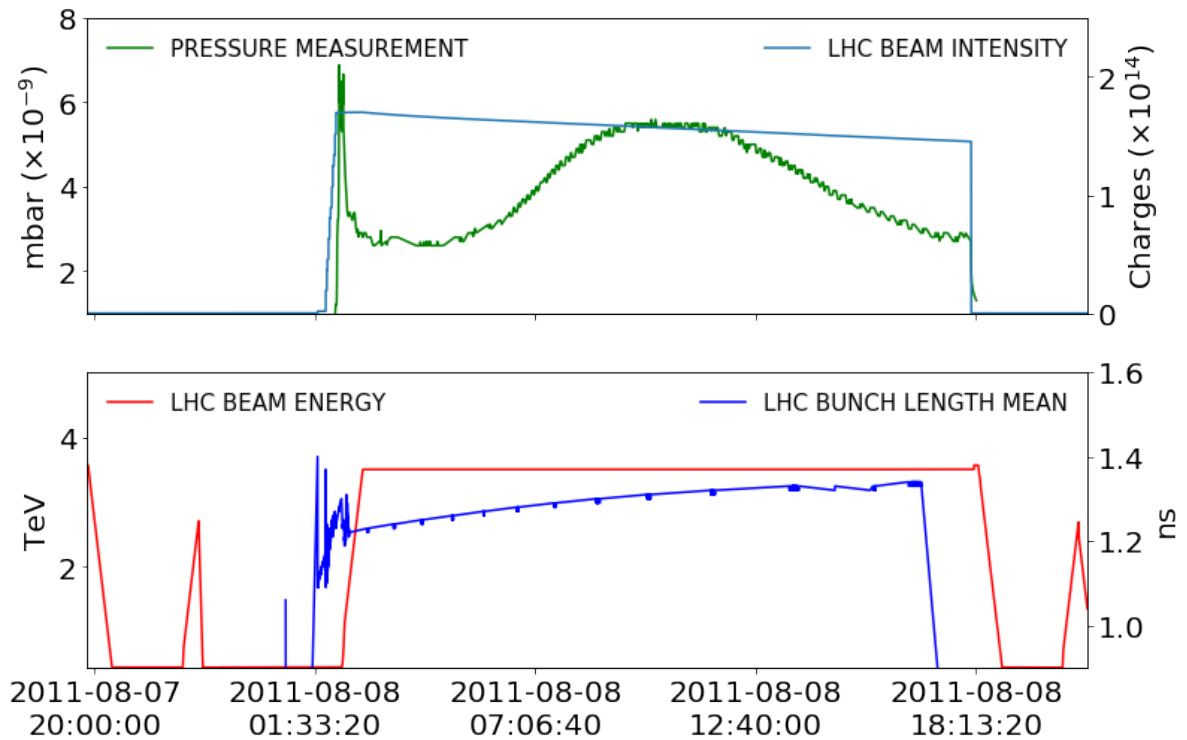
FIGURE 2.1: Example of heating pattern a pressure measurement from a run of the LHC of the 2011. The top plot shows the beam intensity and the pressure measurement, in the bottom are represented the beam energy and the average bunch length. Such a pressure increase is correlated neither to the beam energy nor to variation of the bunch length and has indicated presence of heating in the component. The vacuum gauge is located close to the injection protection collimator (TDI).

2.2   Pressure measurements

Each pressure measurement is stored as a time series with a sampling period that may vary between 1 and 5 minutes. The variability of the sampling period depends on the year of data acquisition: LHC data older than 5 years are downsampled in order to save storage.

The x-axis of the pressure measurements is a time axis. As commonly done for time series [41], for simplicity the *steps* will be adopted as x-axis in the following. Each *step* is an integer value representing the index of the corresponding pressure value in each pressure measurement array. The total number of *steps* is equal to the total values

sampled for each pressure measurement from the acquisition system. Doing so, the time and date information is lost but it is not relevant for the classification purpose.

The total number of *steps* in time of each pressure measurement depends on the sampling period and on the length of the LHC fill. The length of the LHC fill may vary between 5 and 24 hours. The variability of the sampling rate together with the variability of the LHC fill length make the various pressure measurement data different between each other in terms of length.

2.3   Heating monitoring

Heating monitoring is an essential operation during machine runs. The heating has to be kept under control both before and during operation with beam. Neglecting abnormal heating indication may cause damage in components and degradation of machine performance. Heating is a power loss measured in $Watt$. The power loss can be measured either directly or indirectly in different ways. In the LHC, it can be measured directly with temperature probes named PT100 [21] due to the 100 Ω resistors that are used to estimate the temperature and with fiber optics in CMS [22]. Measuring the required cooling power to compensate the beam–induced power loss (e.g. valve opening and coolant temperature for the LHC beam screen [23]) is another direct method. The power loss can also be detected indirectly with the pressure readings [24] giving a more qualitative estimation. The bunch length in the LHC is measured by the beam quality monitor [42].

Heating of surrounding equipment can also be detected indirectly with pressure measurements [24]. The idea is based on the outgassing phenomenon [25] produced by the increase in temperature in a high-vacuum environment. The outgassing is directly observed as pressure increase by means of vacuum gauges as for the case of the injection protection device (TDI), during the 2012 LHC run [4].

With respect to the direct temperature monitoring, the vacuum-based indirect measurement is more dense and systematic. However, analysing the pattern of each pressure measurement is not a trivial task. Expert knowledge is needed to identify heating from a pressure measurement. Moreover, due to the large amount of pressure data retrieved after each fill of the LHC, classifying each pressure measurement

as indicating heating or not is a tedious and time-consuming task. A fill is an LHC operation cycle, starting with the gradual injection of the two counter-rotating beams, their acceleration to collision energy and finishing with the beam dump. During the LHC operation, each fill is injected in the LHC roughly 2 hours after the dump of the previous one. In such a short time, all the 1200 pressure measurements cannot be analysed manually in order to prevent heating that could cause damage during the next fill.

2.3.1 Heating indication from pressure measurements

In the previous section it has been explained how heating can impact pressure measurements through thermally induced outgassing. This section mostly focused in understanding how to visually detect heating from pressure measurements. The main characteristics of a pressure measurement pattern indicating heating are:

- Either sudden or slow pressure increase occurs when the beam energy is constant.

- Pressure increase can happen anytime during the fill.

- There is no defined threshold to determine if a pressure increase is heating.

- Pressure increase can last either few minutes or several hours.

A pressure measurement where the presence of beam induced heating is visible has already been shown in Fig. 2.1. In such case, the pressure increase is smooth and easily detectable by means of simple heuristic based algorithms that computes the sign of the first derivatives.

Unfortunately most of the pressure measurements heating indication are not easily detected by simple heuristic based algorithms. An example is shown in Fig. 2.2. This measurement is more noisy and better represents the most frequent heating pattern observed in the LHC. In this plot there is a first broad increase correlated to the energy ramp that is expected. Following the first broad peak there are several sharp and noisy peaks that can be associated to beam losses or errors of the acquisition system. Finally there is a step increase that stays up until the end of the fill. Only this
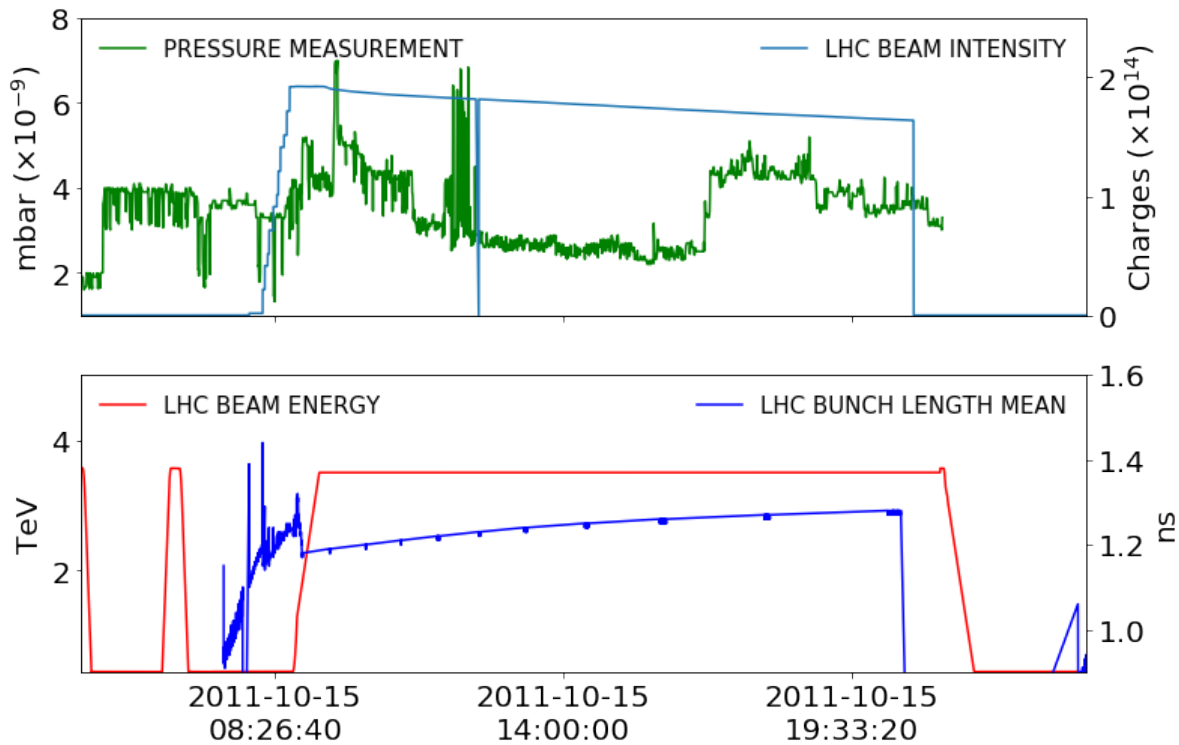
FIGURE 2.2: Example of heating pattern a pressure measurement from a run of the LHC of the 2015. The top plot shows the beam intensity and the pressure measurement, in the bottom are represented the beam energy and the average bunch length. The pressure increase and decrease around 07:00 related neither to the beam energy nor to variation of the bunch length could indicate the presence of heating in the component. The pressure measurement is noisy due to the acquisition system.

last increase could be due to beam-induced heating. In general there is no simple rule to associate a pressure increase to beam-induced heating. Most of the time is expert knowledge based on previous experiences that helps in detecting beam induced heating.

2.4 Machine learning for classification

Machine learning is the science of programming a computer to learn from data. The question is, how can a computer learn something? A good answer to this question is:

A computer is said to learn from Experience E with respect to some task T and some performance measure P, if its performance on T, as measured by P, improves with experience E. Tom Mitchell, 1997. There are several task that machine learning algorithm can tackle. Firstly the algorithms can be either supervised or unsupervised. A supervised algorithm is fed with labeled data (the outcome is known) while an unsupervised algorithm is fed with unlabelled data (the outcome is not known). The unsupervised algorithms are mostly used for clustering or dimensionality reduction of a dataset. The supervised algorithm are mainly divided in two subcategories: regression and classification. Regression consists in predicting an outcome that belongs to a continuous range of values while classification consists in predicting an outcome that belongs to a discrete range of values. This thesis is focused on classification. The goal of a classification is to determine to which class a certain element of the dataset belongs. A class is simply a group or an identifier that has some specified characteristics. A very common example is the spam classifier: given a dataset of e-mails the algorithm should classify them in either spam or not spam. In this context the are are only two classes: spam and not spam. The experience are the data. In general it is desirable to have as much data as possible. The performance is a metric that allows to estimate the goodness of the algorithm. The procedure, for which an algorithm iteratively updates its parameters in order to perform better predictions is called training. The common way to train an algorithm is to define a function named loss function that computes the goodness of the model given the output data of the prediction. The training procedure iteratively updates the parameters aiming to minimise such function.

2.4.1   Dataset

A Dataset is a collection of information used by the ML algorithms for training. It is usually represented as a $m \times n$ table where m are the rows and n are the columns. The rows are also named examples, instances or elements. The column are commonly named features or predictors. Among the features in case of supervised learning, there can be one or more columns that are named labels or targets. An ideal ML model outputs the correct labels given the features as input for each instance in the dataset. It is important when training a ML model to split the dataset in two sets: the

training set and the test set. The training set is used for training the model while the test set is used to evaluate the performance of the model. A model evaluated on the training set can lead to an erroneous interpretation of its performance.

2.4.2 Evaluation metrics

When dealing with ML algorithms it is crucial to choose an evaluation metric properly. Numerous metrics exist in literature for classification problems [41]. The classification of pressure measurements is a binary classification problem: the outcome of the prediction process could belong only to two classes: positive (i.e. the measurement indicates heating) or negative (i.e. the measurement does not indicate heating). Classification metrics classically relies on the definition of True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) [41]:

- TP: an outcome where the model correctly predicts the positive class.

- TN: an outcome where the model correctly predicts the negative class

- FP: an outcome where the model incorrectly predicts the positive class.

- FN: an outcome where the model incorrectly predicts the negative class.

The metrics that will be used in this paper are accuracy, precision and recall defined as follow:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}, \tag{2.1}$$

$$Precision = \frac{TP}{TP + FP}, \tag{2.2}$$

$$Recall = \frac{TP}{TP + FN}. \tag{2.3}$$

The *Accuracy* is the ratio of number of correct responses to the total number of input samples. *Precision* and *Recall* are a trade-off: reaching high precision requires sacrificing recall and vice-versa. High recall is important when positive cases should

not be missed while high precision is needed when the positive outcome of the model should be always correct even if some positive cases are missed.

For the case of the classification of pressure measurements in LHC, high recall is needed during LHC operation in order to maximize the chances of detecting heating issues that could get worse and could lead to damage and long stop of LHC operation, even if this would mean to classify some negative cases as positive (i.e. low precision). High precision is useful to search in the old LHC runs looking for undetected heating case but acknowledging that some of them could be missed (i.e. low recall).

2.4.3   Logistic Regression

The Logistic Regression [43] is a classification algorithm that computes probability. The logistic function is defined as:

$$\sigma(t) = \frac{1}{1 + e^{-t}}. \tag{2.4}$$

The plot of Eq. 2.4 is represented in Fig. 2.3. The logistic function takes as input a real number in the range $]-\infty, +\infty[$ and outputs a real value in the range $]0, 1[$. The output of such function is interpreted as the probability of its input being positive.
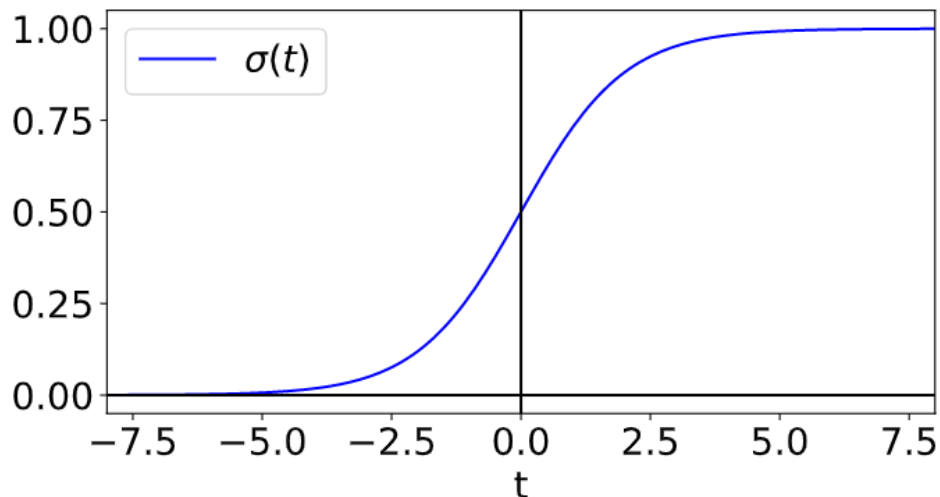
FIGURE 2.3: Logistic function.

The input t is defined as:

$$t = \Theta_0 x_0 + \Theta_1 x_1 + ... + \Theta_n x_n,$$

where Θ_i is the i-th feature parameter of the algorithm and x_i is the i-th feature of the input vector $\mathbf{x}$. The Logistic Regression model predicts 0 if $\sigma(t) < 0.5$ and 1 if $\sigma(t) \geq 0.5$. Training this algorithms means finding the values of Θ that best fit the class label array $\mathbf{y}$. The target is to tune Θ in order to have $\sigma(t) < 0.5$ if $y = 0$ and $\sigma(t) \geq 0.5$ if $y = 1$. This can be done by minimizing the following function also named log loss:

$$J(\mathbf{\Theta}) = -\frac{1}{m} \sum_{j=1}^{m} [y_j log(\hat{y}_j) + (1 - y_j) log(1 - \hat{y}_j)]. \tag{2.5}$$

Where y_j and $\hat{y}_j$ are respectively the class label and the model prediction of the j-th element of the dataset. m is the total number of element in the dataset.

2.4.4   k-Nearest Neighbours

k-Nearest Neighbours (k-NN) [44] is a simple and effective classification algorithm. Its prediction are instance-based rather than model-based. An example of a model-based algorithm is Logistic Regression where the data are used to train the model (i.e. finding the best Θ). In an instance-based model the prediction directly depends on the data used during training.

Fig. 2.4 shows how k-NN performs prediction. In the example there is a two-dimensional dataset (x_1, x_2) where the instances belongs to two classes: class A (yellow) and class B (purple). Given a new instance represented by a red star in the plot and given a fixed value of k, the k-NN counts the k closest elements to the new instance (red star) and predicts the new instance to belong to the most frequent class within the k examples. In the plot can be observed that for $k = 3$ k-NN predicts the new instance to belong to class B while for $k = 6$ predicts class A. This underlines the main difficulty in training k-NN: the choice of k. There are no specified rule on how to set k. In general several values of k are tried and the one that performs better on the chosen metric is taken.
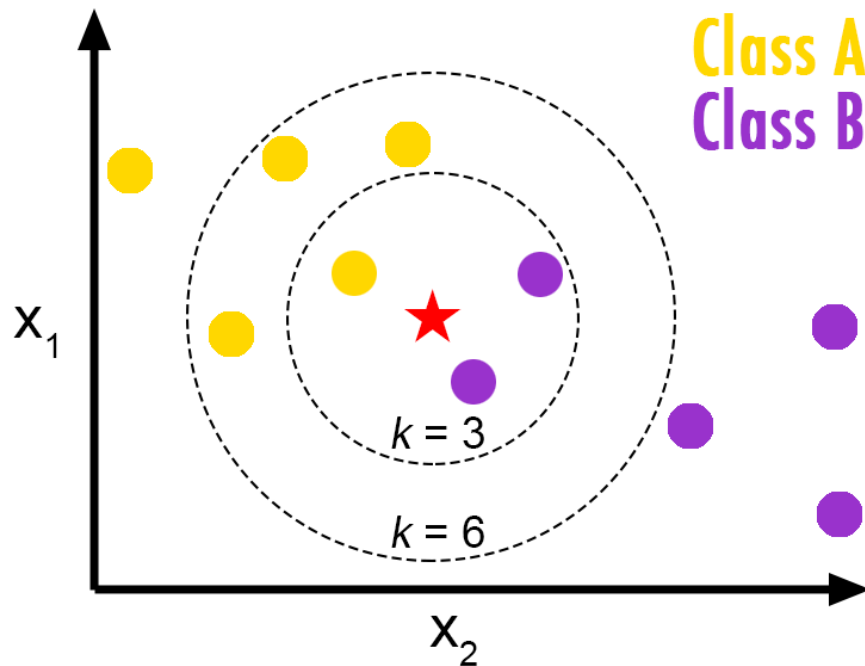
FIGURE 2.4: k-Nearest Neighbours prediction example [45].

2.4.5   Decision Tree

Decision Trees [46, 47] are powerful algorithms that can tackle either classification or regression tasks. As the name suggests, this model builds a tree by subsequent splitting of the dataset according the feature values. In Fig. 2.5 is represented an example of decision tree.

The nodes are the rectangles. In each node there is a feature of the dataset, The root node is the one on the top containing the *Weather* feature. Each node is then split according to the values of the features. According to the value of the node each node can either lead to a prediction (*Yes* if *Cloudy* is chosen) or to a child node (*Humidity* or *Wind* if *Sunny* or *Rainy* are chosen). The most widely used algorithm for training decision tress is the Classification and Regression Tree (CART) [46]. It splits the training set into two subsets according to a feature and a threshold. Naming the feature k and the threshold t_k, the algorithm looks for the couple (k, t_k) that minimises its cost function defined as:
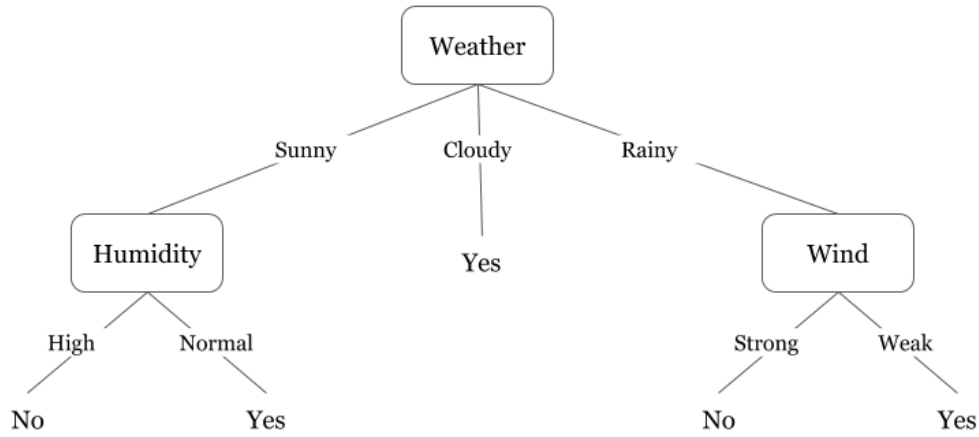
FIGURE 2.5: Decision Tree example.

$$J(k, t_k) = G_{left}\frac{m_{left}}{m} + G_{right}\frac{m_{right}}{m}, \tag{2.6}$$

where m_{left} is the number of instances in the left split and m_{right} is the number of instances in the right split. G is the impunity function defined as:

$$G = 1 - \sum_{k=1}^{n} p_k^2 \tag{2.7}$$

2.4.6 Random Forest

Random Forest [48] is an ensemble of Decision Trees. It builds many Decision Trees (the number can be set by an hyperparameter) and then uses all of them for prediction. The class predicted by the majority of the Decision Trees is the outcome of the Random Forest prediction (Fig. 2.6).
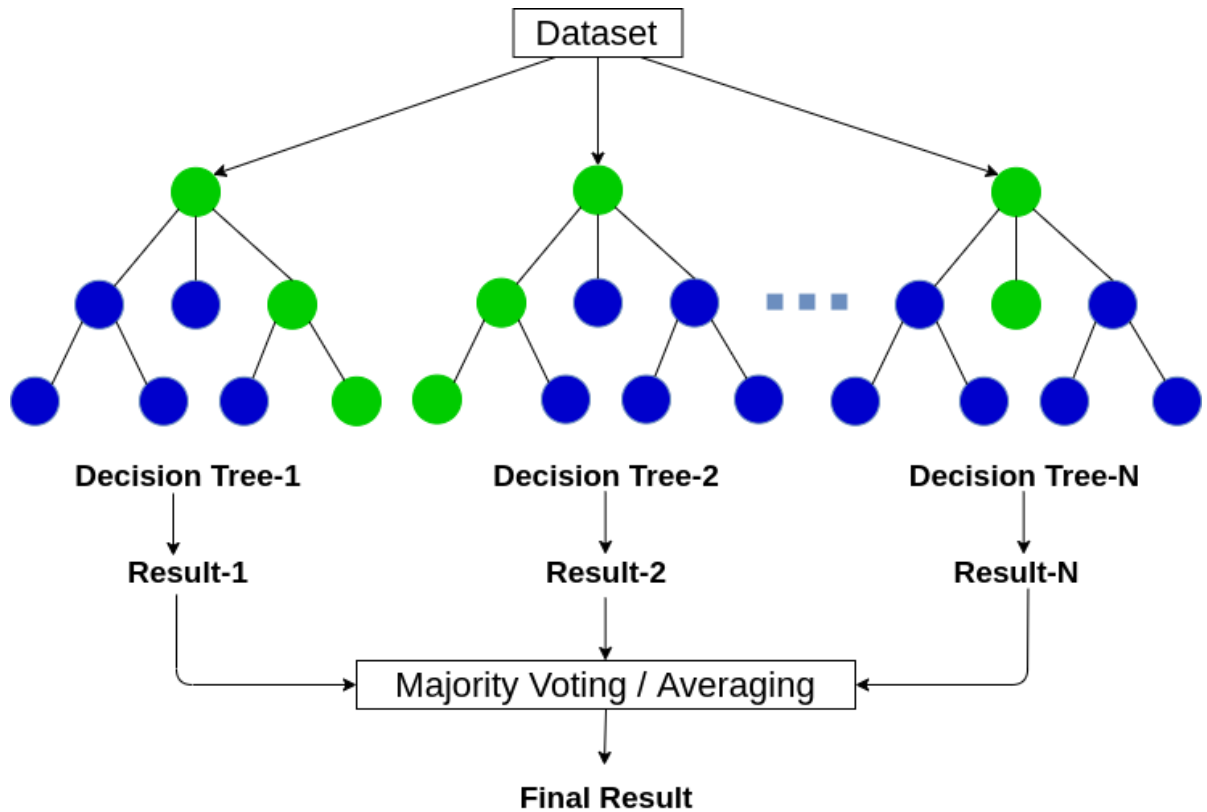
FIGURE 2.6: Random Forest example.

Each Tree is trained on a subset of the training set sampled with replacement. The feature to perform the nodes split is sampled among a random subset of the features. In general Random Forest performs better than Decision Tree and it is less prone to overfitting due to various source of randomness in the algorithm [41].

2.4.7   Neural Network

Neural Networks are scalable, powerful and and versatile algorithms that have become very popular in the last 10 years with the increase of computational power. The elementary component of a neural network is a neuron. A single neuron network is represented in Fig. 2.7.

$x_1...x_i$ are the inputs while $w_1...w_i$ are the weights. Each neuron may also have a bias value b (i.e. a constant) that is not represented in the figure. In its most general
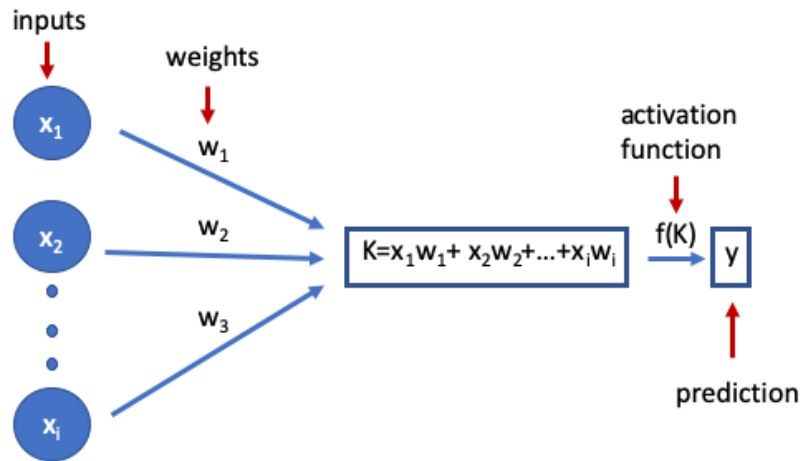
FIGURE 2.7: Single neuron network.

form, a neuron computes:

$$y = f(\sum_{k=1}^{n} x_k w_k + b),\tag{2.8}$$

where f is the activation function. The network just described has only 1 *layer*. In general, networks have more layers and more neurons per layer. A network has two layers if the output of the first set of neurons is fed to another layer. The neural networks are unidirectional, by looking at Fig. 2.7 the inputs come from left and the output goes directly in y or in another neuron if present. The activation function is applied to each neuron to give the network non-linearity. Sometimes the activation function may be omitted (i.e. there could be a neuron without activation function). The two main activation function that one may find in a binary classification problem are: *Rectified Linear Unit (ReLu)* and *Sigmoid*. The sigmoid has already been presented and plotted in Fig. 2.3. The ReLu is represented in Fig. 2.8. It is a simple linear function if the input is positive and zero otherwise. The ReLu is used in the hidden layers, the layers, which output is fed into another layer. The sigmoid is used for the output layer in binary classification task to output a probability. Training a neural
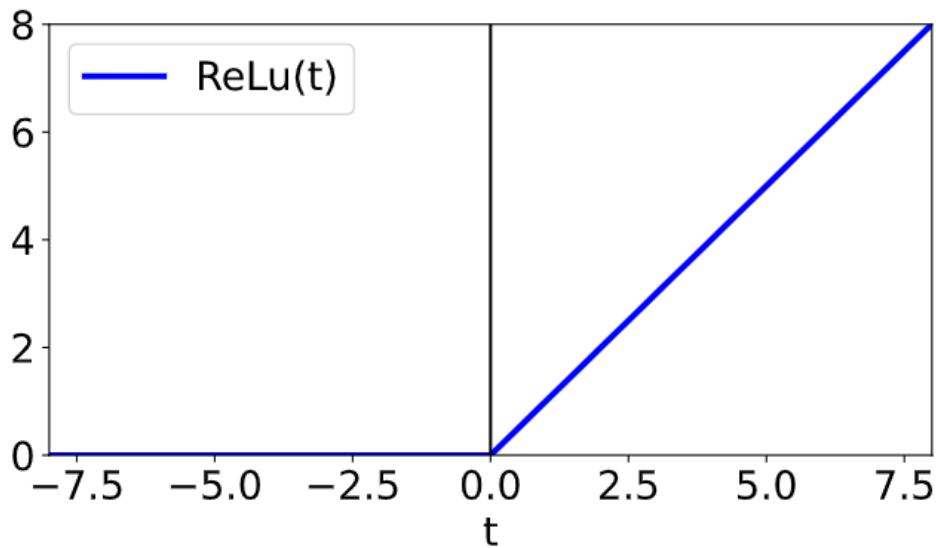
FIGURE 2.8: Rectified Linear Unit (ReLu).

network means finding the weights matrix **W** that minimise the loss function defined for the problem. For binary classification tasks, the loss function adopted is the *binary cross-entropy* previously presented with its synonymous log loss in Eq. 2.5.

2.4.8   Convolutional Neural Network

A particular kind of Neural Network widely used for image recognition and signal processing are the Convolutional Neural Networks (CNN) [49]. A CNN is a network made of Convolutional Layers. In a convolutional layer the neurons of the input layer are not connected with single elements of the input, but a subset of the input is processed with a filter (also called kernel) that slides over the full input. The size of the filter determine the size of the input that can be processed. In Fig. 2.9 is represented an example of a CNN with a rectangular input and two convolutional layers.

A filter slides over the input and outputs an element that is fed into the first convolutional layer. A single element of the convolutional layer is not connected to the full input but only to a subset of it. All the outputs of a single filter slid over the full input form a *feature map*. The number of the feature maps is equal to the number of filters defined. All the feature maps produced by several filters slid over a single input produce a convolutional layer. When training a CNN the main parameters to
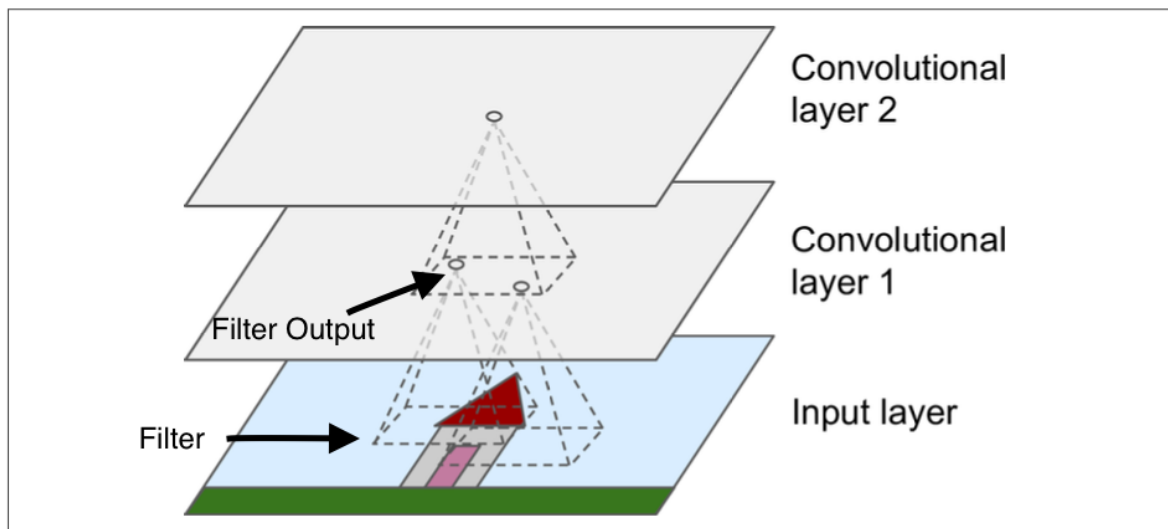
FIGURE 2.9: CNN example [41]

specify are: filter size (kernel size), stride and the number of feature maps (number of filters). It has to be noticed that the filter parameters are not specified and thus are learned trough the training of the network. One has only to specify the filter size and the number of filters. All the filters belonging to the same layer have the same size even if they have different parameters and they produce different feature maps. The stride is the amount of elements of the input that the filter has to skip while sliding over it.

2.4.9 k-means

k-means [50] is an unsupervised clustering algorithm. It groups the input data by a similarity criterion. It has not to be confused with k-NN previously described, which is a supervised learning algorithm. The k-means algorithm takes as input the number of clusters k and the data and works as follows:

[1] It randomly initializes k points in the feature space named centroids.

[2] It computes the distance between each element in the dataset and the centroids.

[3] It labels the closest elements to each centroid with the same class.

[4] It recomputes the centroid position as the average position between the elements of each class.

[5] It repeats from point 2 to 4 until the the position of the centroids does not significantly differ between 2 iterations: their new position is within a predefined convergence distance with respect to their position at the previous iteration.

The most commonly used distance metric is the Euclidean distance defined as:

$$d(\mathbf{p}, \mathbf{q}) = \sqrt{\sum_{i=1}^{n} (p_i - q_i)^2} \tag{2.9}$$

where $\mathbf{p}$ and $\mathbf{q}$ are two elements belonging to the feature space and n is the total number of features. The k-means algorithm can also be used as dimensionality reduction technique proving that it can significantly improve classification accuracy when used as preprocessing step [41].

Part II

Methodology

Chapter 3

Impedance of a synchrotron component from beam-induced power loss

3.1 Overview

For impedances modelled with the resonator formula, the quality factor and the resonance frequency need to be estimated in order to obtain the impedance. In addition, the shunt impedance (R_s) should be evaluated to fully obtain the value of the impedance. However, once the other two parameters have been obtained, a single power loss measurement is needed to assess R_s, by substituting Eq.(1.10) into Eq.(1.19), and by solving for R_s. In this procedure, it is assumed also that different beams could be produced where just one parameter is changed without varying any of the other parameters and that f_r is very close to one of the main lines of the spectrum. In particular, the number of bunches enters in the computation of the quality factor while the bunch length enters in the computation of the resonance frequency. This last assumption is not needed for the computation of the resonance frequency but only for the estimation of the quality factor. Therefore one can first derive the resonance frequency and then evaluate if the quality factor can be computed knowing the frequencies of the main lines of the power spectrum. It should be noted that when the resonant frequency is close to one of the lines of the beam spectrum there are both power loss and beam instabilities can also occur. Nevertheless, a resonance for a single device in a large machine may be a perturbation on top of the impedance

model and a large heat load on one device does not necessarily mean stability issues as was observed for the case of the LHC synchrotron light monitor BSRT and injection protection collimator TDI [4]. In addition, the main heat load spectrum lines are not necessarily resonant lines involved in multibunch beam instabilities [51].

3.2 Computing the resonance frequency

As mentioned in Section 1.2.2, the position of the peak impedance represents the resonance frequency that is being estimated. The impedance does not depend on the beam parameters; therefore, by varying any of these beam parameters, the power loss changes without any variation of the impedance. This is analogous to circuit theory where the impedance is independent from the current. Assuming that it is possible to provide different beams to the machine, with different bunch lengths (but keeping the other beam parameters unchanged), it can be observed that the change of the bunch length modifies the shape of the power spectrum of the longitudinal time distribution of the beam Fig. (3.1). Ideally, the time profile of each bunch is Gaussian [14]. In this case, the power spectrum of the time distribution should also be Gaussian. However, in practice, the bunch length τ is finite and hence, the power spectrum of the bunch distribution will exhibit several lobes, separated by minima, lying at multiples of $1/\tau$, as shown in Fig. (3.2). As a result, a change in the bunch length will affect the position of the first zero of the power spectrum, which will move towards higher values as the bunch length decreases. Figure 3.1 shows the power spectra for different values of the bunch length. It is therefore possible to choose a beam with a given bunch length, such that the first zero lies at a pre-established position.

   As can be seen from Eq. (1.19), the power loss is the product of the power spectrum and the impedance. The result of this product decreases as the resonance frequency approaches the minimum of the power spectrum. However, due to the second lobe of the power spectrum, as the resonance frequency increases to values higher than the frequency of the first minimum of the spectrum, the power loss increases again. Therefore by observing the power loss versus the bunch length, the resonance frequency can be identified as the position of the first relative minimum.

The magnitude of the second lobe of the power spectrum plays an important role. The
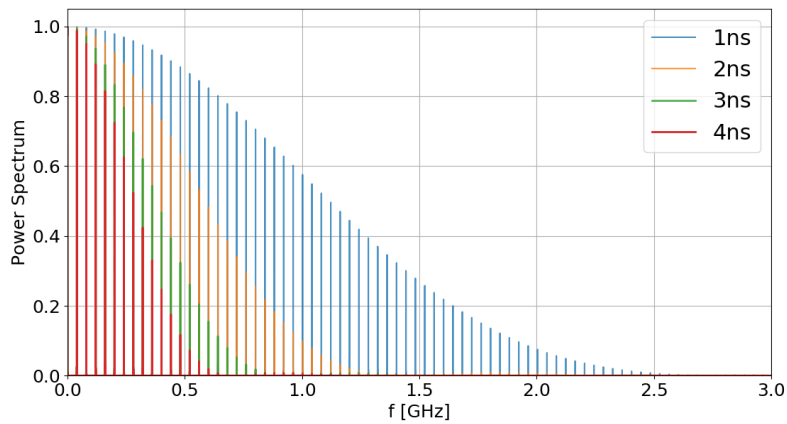
FIGURE 3.1: Power spectra of fills with different bunch lengths. The frequency of the first main line that goes to zero is related to the bunch length. The shape of the bunches is Gaussian.
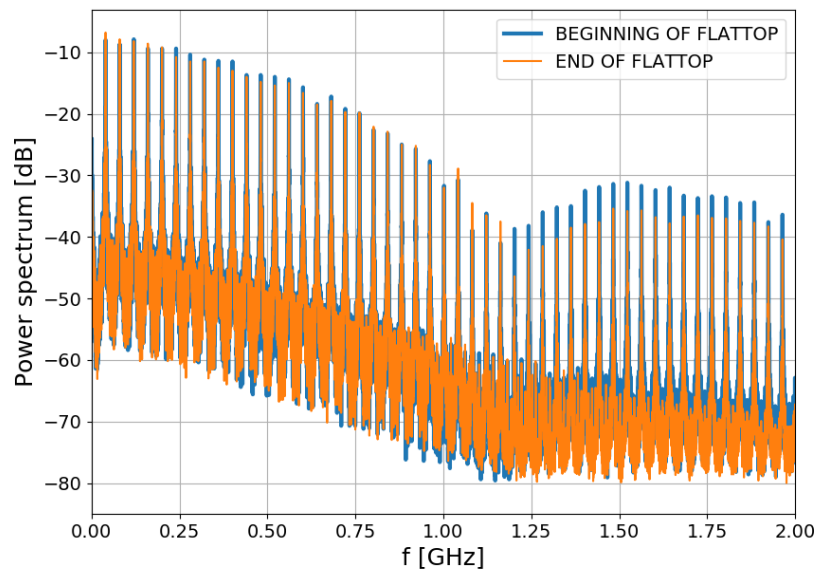
FIGURE 3.2: Measured power spectra of the LHC beam for the fill number 5979 of the (21-07-2017). The figure shows that during the energy flat top the spectrum of the beam is evolving with time. Two lobes of the spectra are visible due to the finite length of the bunches.

accuracy of determining the minimum improves as the amplitude of the second lobe increases. It would be preferable to have parabolic longitudinal bunch charge distributions as shown in Fig. 3.3, where the second lobe is clearly visible. A parabolic shape of the bunch can be seen indeed as the product of a parabola with a time window that results in a parabola for the positive y-axis and zero elsewhere. Here again the presence of the window leads to a sinc function in the Fourier domain and thus again the second lobe. Because the cut in this case is less smooth due to the concavity of the parabola, the second lobe is more visible. Parabolic bunch distributions are present in the LHC after injection oscillations are damped [14]. The distribution can also be tuned by adding some noise, in a procedure called bunch flattening in LHC [52].
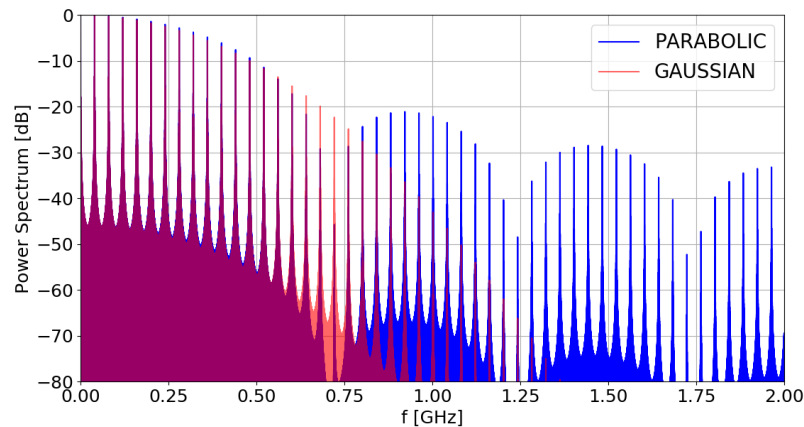
FIGURE 3.3: Power spectrum of a beam with parabolic longitudinal distribution compared with a Gaussian distribution, both with a bunch length of $2\,\mathrm{ns}$. The second lobe of the truncated Gaussian longitudinal distribution is already not visible within $-80\,\mathrm{dB}$.

3.3 Computing the Q factor

Considering a resonator model impedance, with a fixed value of f_r, the relation between beam–induced power loss and the number of bunches in a machine revolution period can be extended to intermediate cases between narrow-band and broad-band
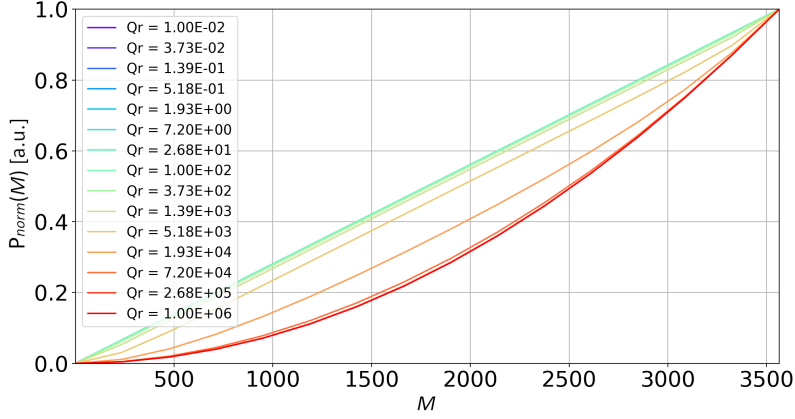
FIGURE 3.4: Normalized power loss vs the number of bunches for various values of Q_r .

impedance, as

$$P_{loss}(M) \propto M^\alpha \tag{3.1}$$

where $\alpha = \alpha(Q_r)$ is a monotonic function of the the quality factor Q_r, where $\alpha = 1$ for $Q_r \to 0$ and $\alpha = 2$ for $Q_r \to +\infty$ [14].

For a fixed value of Q_r, when the trend of the power loss versus the number of bunches is obtained by considering a normalized power loss, $P_{norm}(M)$, as

$$P_{norm}(M) = \frac{P_{loss}(M)}{P_{loss}(M_{max})}, \tag{3.2}$$

where $P_{loss}(M_{max})$ is the power loss corresponding to M_{max} bunches. The normalized power loss is equivalent to Eq. (3.1) as

$$P_{norm}(M) = c \cdot M^\alpha \tag{3.3}$$

where c is a constant that can be determined by observing from Eq. (3.2) that $P_{norm}(M_{max}) = 1$, to be,

$$c = \frac{1}{M_{max}^\alpha}. \tag{3.4}$$

Therefore, Eq.(3.3) becomes

$$P_{norm}(M) = \left(\frac{M}{M_{max}}\right)^{\alpha}, \tag{3.5}$$

where α is assessed by taking the natural logarithm of Eq.(3.5) as follows,

$$\log P_{norm}(M) = \log \left(\frac{M}{M_{max}}\right)^{\alpha} = \alpha \log \left(\frac{M}{M_{max}}\right) \tag{3.6}$$

Finally, Eq.(3.5) is averaged over M, in order to get the best estimation of α over the considered range of the bunch numbers

$$\frac{1}{M_{max}} \sum_{k=1}^{M_{max}} \log P_{norm}(M_k) = \alpha \frac{1}{M_{max}} \sum_{k=1}^{M_{max}} \log \left(\frac{M_k}{M_{max}}\right), \tag{3.7}$$

where k is the number of the considered bunch $M_1, \ldots, M_{Mmax}$.

It can be observed from Eq.(1.19) that α depends on Q_r. By varying Q_r, the bandwidth of the impedance varies. Therefore, the product inside the sum leads to a different number of negligible terms. This means that the value of α giving the best approximation of Eq. (1.19) with Eq. (3.1) also varies. Due to this dependence, it is difficult to obtain a closed form for the $\alpha(Q_r)$ function. For this reason, $\alpha(Q_r)$ was evaluated through numerical computation.

In order to determine the trend of the $\alpha(Q_r)$ function, the value of α has to be computed for several values of Q_r. Before evaluating α, the quantity P_{norm} has to be computed as shown in Fig. (3.4), with M varying from 1 to 3564 that is the total number of slots available in the LHC.

This procedure requires several computations of the power loss with different numbers of bunches for each value of Q_r.

For the numerical computation, the longitudinal distribution of the bunch in the machine is usually assumed to have Gaussian, equally-spaced and identical bunches (as in reference [53]) that are uniformly distributed along the accelerator.

In a more realistic scenario, the bunches are not identical to one another. In addition, they are not equally-spaced and their filling scheme (as defined in Section 1.3) is not periodic.

In this work, the realistic scenario, considering a typical LHC fill patterns (as in Fig. (4.8)) was considered.

Chapter 4

Power loss simulations

4.1 Overview

The power loss depends on numerous parameters. To understand the impact of each parameter on the power loss, their effects have to be decoupled. This is done by studying a cavity with a single resonating mode. Firstly the resonating cavity is defined and its impedance is shown. Then the single beam power loss is presented in this scenario, underlining how this cavity can help in studying the following two-beam power analysis. Finally the two-beam power loss is simulated firstly under the assumptions of Eq.(1.25) and then by observing how the power loss behaves outside these assumptions with the help of the CST solver.

4.2 Pillbox cavity

To perform analysis on the power loss, firstly a cavity with just one resonating mode was identified and modeled. To perform two-beam power loss studies, it is important that the decay time of the wakefield in the cavity is long enough to ensure an EM cross-talk between the bunches of the two beams. By looking at Eq. 1.26 the best trade-off between the conductivity σ and the cavity length h has to be found. A too low value of σ means that the wake field decays too fast, not allowing two bunches of the same beam or even two bunches of two different beams to cross-talk. A cavity with high σ, low Q and f_r around $1\,GHz$ (needed to have just one mode in the cavity) are quite demanding specs, since the required Q and σ lead to a very small value of the cavity length h. This modelling issue can be solved considering that the cavity

has to be linked to a beam pipe, therefore even having an higher value of h does not lead to very high values of Q because some of the power flows into the beam pipe lowering down the Q value. The beam pipes are $50\,\mathrm{cm}$ long per side meaning that the bunch takes $1.6\,\mathrm{ns}$ to pass trough the pipe and reach the cavity where the power loss is dissipated due to the lossy material.

A resonating mode at $1.10\,\mathrm{GHz}$ is obtained with a radius $a = 11.6\,\mathrm{cm}$ and a length $h = 1.0\,\mathrm{cm}$ with a beam pipe with radius $r = 4.0\,\mathrm{cm}$. The resulting model is shown in Fig.. (4.1).
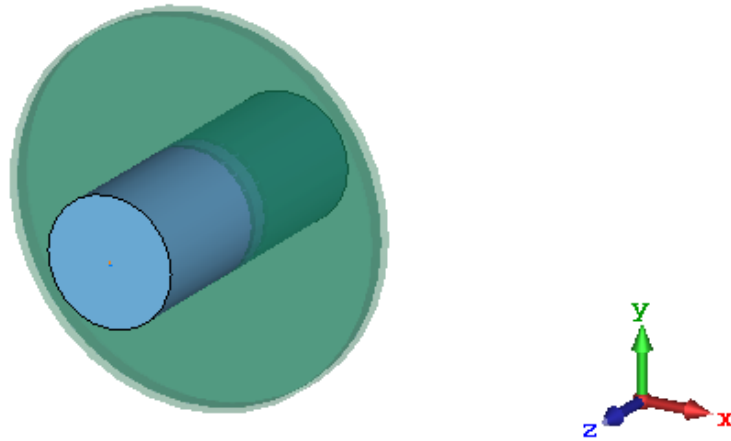
FIGURE 4.1: CST model of a cavity with: a=11.6 cm, h=1.0 cm, σ=10 000 S/m and beam pipe with radius r=4.0 cm.

The cavity has a layer of lossy material with $0.5\,\mathrm{cm}$ of depth and $\sigma = 10^4 S/m$. The cavity was modeled on CST studio suite that allows computing the cavity impedance and wakefields directly. The resulting impedance is shown in Fig. (4.2).

The second mode of the cavity is above $2\,\mathrm{GHz}$. For the beams adopted in simulations, only impedance below $1.5\,\mathrm{GHz}$ may significantly contribute to the power loss. All the impedance computations were performed taking care that the wake potential is decayed within 1% of its initial values as shown in Fig (4.3).

In order to speed-up the simulations that are very time demanding for the two beam power loss case magnetic yz and xz planes were assumed for the simulation to take advantage of the symmetry of the model.
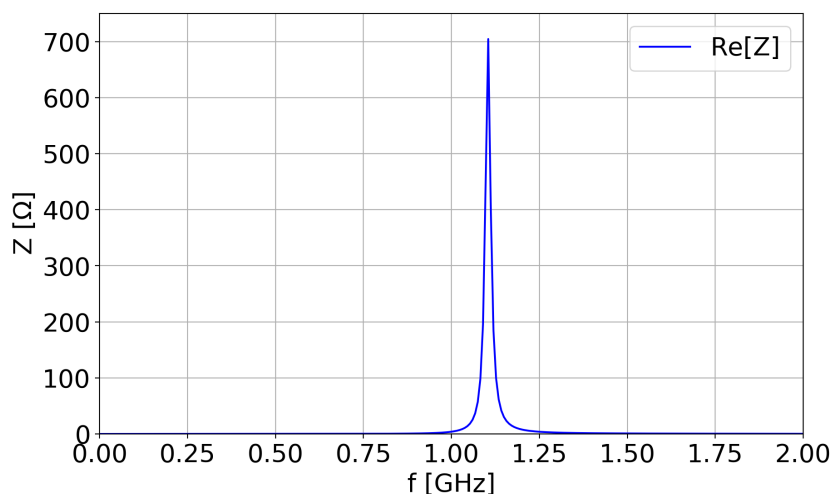
FIGURE 4.2: Real part of the longitudinal impedance of a pillbox cavity with: a=11.6 cm, h=1.0 cm, σ=10 000 S/m and beam pipe with radius r=4.0 cm.
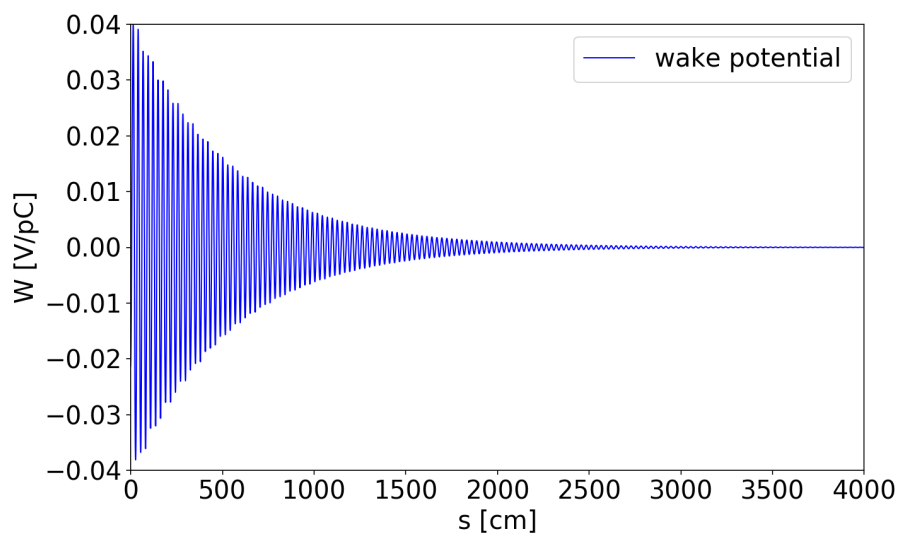
FIGURE 4.3: Wake potential of the impedance of Fig. (4.2).

In Fig (4.4), it is shown that the presence of the symmetry planes do not change the computated wake potential, as expected. On the other hand, the symmetry planes reduce the simulation time significantly by reducing the simulated volume of a factor 4.
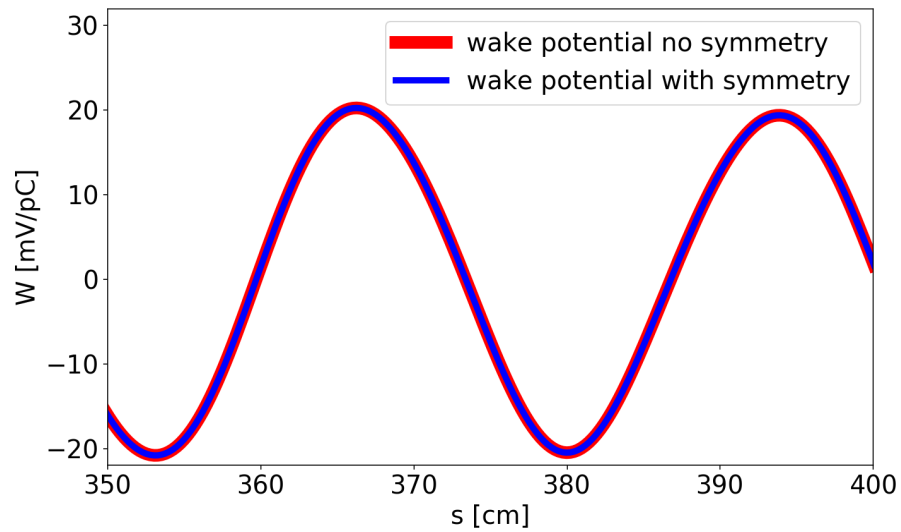
FIGURE 4.4: Wake potential symmetry comparison of the impedance of Fig. (4.2).

4.3   Single beam power loss

The single beam power loss was simulated in two main scenarios:

- Single bunch,

- Multi bunch.

4.3.1   Single bunch

The single bunch case is studied to observe the duration of the power loss in the cavity. This is needed to be sure that the power loss lasts long enough to couple two bunches entering the cavity with a phase shift τ_s. The bunch parameters are:

- $N_b = 2.3 \cdot 10^{11}$ protons per bunch.

- bunch length $\tau = 1.2\,\text{ns}$.

The bunch enters the simulation domain $25\,\text{ns}$ after the beginning of the simulation as represented in Fig. (4.5).

The case of two bunches spaced by $25\,\text{ns}$ is shown in Fig. (4.6). Taking into account
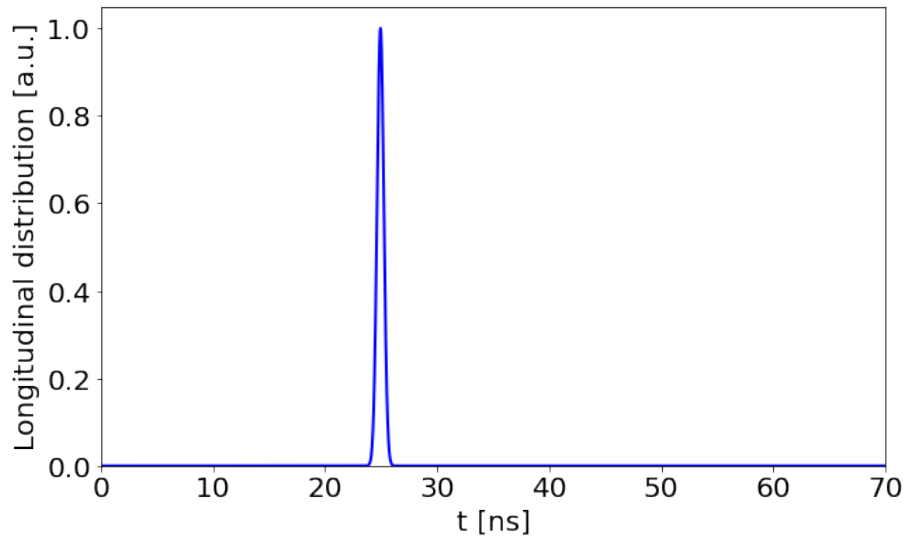
FIGURE 4.5: Longitudinal profile of a single bunch entering the simulation domain 25 ns after the beginning of the simulation.
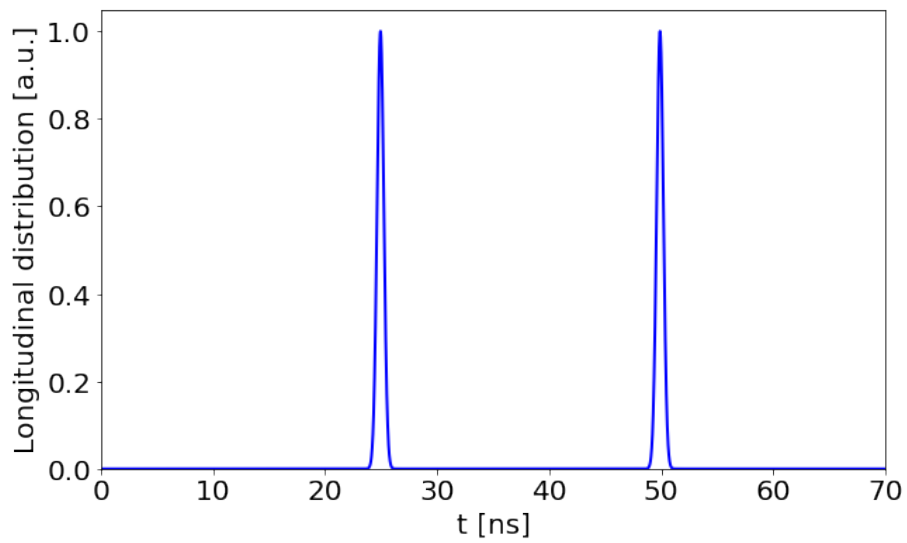
FIGURE 4.6: Longitudinal profile of two bunches entering the simulation domain respectively 25 ns and 50 ns after the beginning of the simulation.

the length of the beam pipe of $50\,\mathrm{cm}$ the bunches in both cases will enter the cavity with an additional delay of $1.67\,\mathrm{ns}$.

4.3.2   Multi bunch

The single beam power loss was also computed from CST and compared with Eq. (1.19) in a simple case scenario. In Eq. (1.19), the impedance is computed with CST and the single beam spectrum of one bunch represented in Fig (4.7) was used. To compute the power loss, the fill 5979 of the Large Hadron Collider (LHC) in 2017 was chosen. Its longitudinal distribution is plotted in Fig. (4.8).
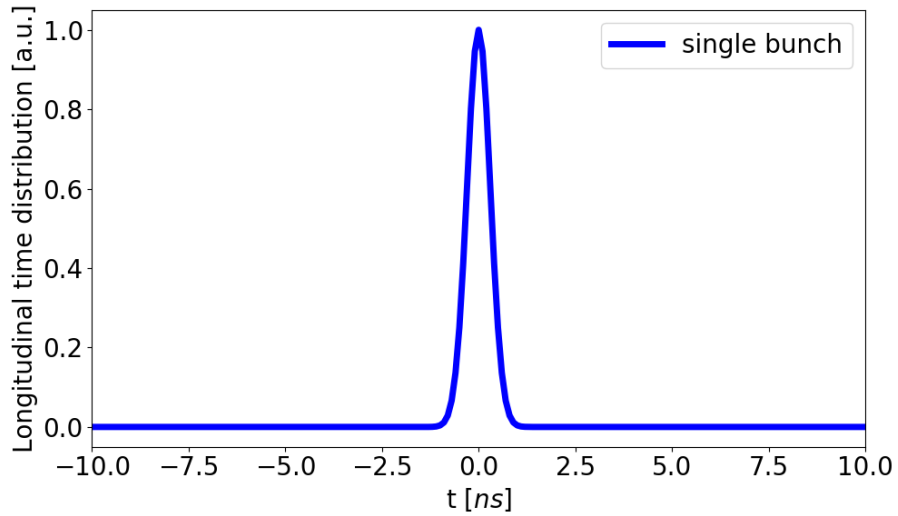
FIGURE 4.7: Longitudinal time distribution of a single-bunch beam.

Such filling scheme has been used directly as input beam distribution in CST to compute the power loss. CST indeed provides as output the instantaneous power loss array $p(nt_0)$ along the duration of the fill T, where n is the length of the array and t_0 is the time between two consecutive points. Thus the power loss for the cavity results to be:

$$P_{loss} = \frac{1}{T} \sum_{n=0}^{\frac{T}{t_0}} p(nt_0). \tag{4.1}$$

The computation of the power loss versus time of a multi-bunch beam like the fill 5979 is very time-demanding. Looking at the longitudinal distribution of the considered fill in Fig. (4.8), it is possible to observe that it is a good approximation to consider just a fifth of the longitudinal distribution to compute the power loss due to
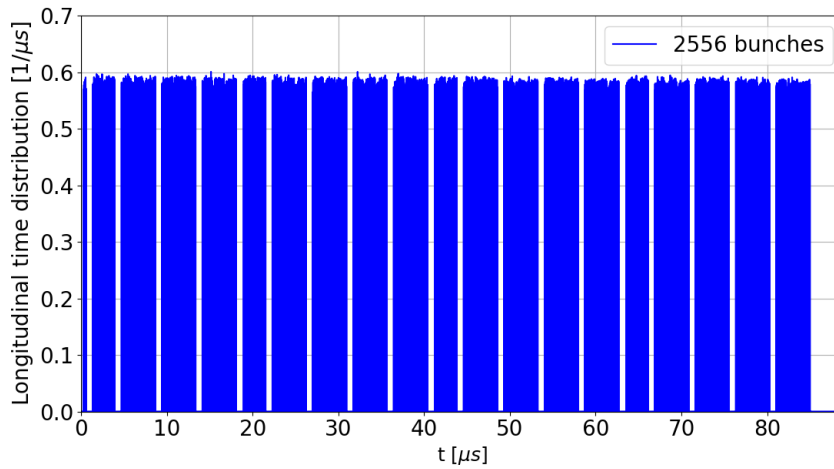
FIGURE 4.8: Longitudinal time distribution of the beam for LHC fill number 5979 on 21-07-2017.

the quasi-periodicity of the filling scheme. It has to be underlined that such method is a good approximation and not the exact computation of the power loss because the first batch of 12 bunches that is needed for machine protection checks does not repeat itself in the filling scheme, contrary to the following batches. An estimate of the error produced by this approximation is given in the part of the thesis dedicated to the *Results*.

4.4 Two beam power loss analysis

CST also allows computing the two-beam power loss with good precision [54]. The simplified two-beam power loss equation Eq. (1.25) has been compared with the CST results in the range of its validity and the the limits of such equation have been explored using CST. To compute the power loss with Eq. (1.25), the impedance computed by CST for the previously defined pillbox cavity was used. Also for this case in CST, the impedance is computed with a single beam made of a single bunch. The procedure to perform two-beam power loss computations with CST is the same as for the single-beam case described in the previous paragraph: the only difference is that a second beam that circulates in the opposite direction to the first beam has to be

defined. Computing simulations with two full beams is extremely time-demanding for CST. It was therefore decided to first perform single bunch simulations for the two-beam case. An analysis with two full beams was performed Eq. (1.25) and the comparison with the formula is given in the results of the thesis. If not otherwise specified, the following two-beam analysis are performed considering two counter-rotating single-bunch beams. In all the simulations the two counter-rotating beams have the same longitudinal distribution (i.e. same filling scheme).

4.4.1 Phase shift

The main parameter in the analysis of two-beam simulations is the phase shift τ_s between the two beams. To simulate the phase shift with CST, the two beams were chosen to enter the simulation domain at the same time, while the length of one of the two beam pipes was adjusted: a longer pipe on one side of the cavity causes the bunch to arrive later inside the cavity. This simulation setup is the one that mostly resemble the real occurrences. In particular the phase shift is:

$$\tau_s = \frac{|l_1 - l_2|}{c},$$

where c is the speed of light, l_1 and l_2 are the length of the beam pipes connected to the cavity. The two beams were assumed to circulate at the center of the cavity, meaning that there is no transverse displacement between the two beams:

$$\Delta y_1 = \Delta y_2 = 0$$

in Eq. (1.25). The first benchmark between Eq. 1.25 and CST is to observe the power loss as function of the phase shift. This is done for the cavity presented in the previous section. With the same cavity parameters, a comparison between the two-beam power loss and the single beam case is also presented. Together with the power loss analysis, an impedance modes analysis was performed to understand which mode is primarily contributing to the power loss. Then the analyses are focused on the case of $\tau_s = 0$, for which the center of the cavity is assumed to be the collision point of the two beams.

Cavity radius

The impact of the radius of the cavity was analysed by computing power losses both with CST and Eq. (1.25) for several configurations. In particular, the cavity radius was swept from $13\,\mathrm{cm}$ to $15\,\mathrm{cm}$ with a step of $0.5\,\mathrm{cm}$, with a fixed length $h = 1cm$. For each of these values the power loss as function of the phase shift was analysed.

Cavity length

The length of the cavity plays a major role in the analysis, in particular in the assumptions to derive Eq. (1.25) as shown in [3]. The limits of these assumptions are revealed in disagreements between Eq. (1.25) and CST computations. Possible explanations of this disagreement are also presented in Chapter 7. The power loss as function of the phase shift τ_s was simulated fixing the radius to $r = 14.5\,\mathrm{cm}$ for cavity length values $h = 1\,\mathrm{cm}$, $2\,\mathrm{cm}$ and $10\,\mathrm{cm}$. Then the case of $h = 30\,\mathrm{cm}$ is presented and discussed. Finally, a bigger sweep from $h = 1cm$ to $h = 100cm$ is analysed for the case of $\tau_s = 0$, focusing also on how this plot changes for radii values of $20\,\mathrm{cm}$ and $30\,\mathrm{cm}$.

4.4.2 Multi–bunch analysis

Considering the LHC fill 5979, a two-beam multi-bunch analysis of the power loss versus the phase shift τ_s was performed with the cavity parameters presented in the previous section. Since the multi-bunch case is very time demanding in CST (a full simulation of a single parameter can take up to 1 day), only a fifth of the beam was simulated. The phase shift was varied from $-0.2\,\mathrm{ns}$ to $1\,\mathrm{ns}$ taking six point in this interval. Negative values of the phase shift means that the shifted beam (i.e. the beam incoming from shorter or longer beam pipe) enters in the cavity before the other beam while positive values means that the shifted beam arrives later. A comparison with the single beam power loss is also presented.

4.4.3   Horizontal offset between beams

When defining the two counter rotating beams, it is possible to specify a transverse offset with respect to the center of the structure, to reproduce for instance the situation of accelerator components with common chambers away from the interaction point in the LHC: there, both beams are offset transversely by design to avoid parasitic crossings between the two beams and reduce long range beam-beam effects [55]. Given the circular symmetry of the pillbox cavity, defining an offset only along the x-axis is the same as defining it only along the y-axis. The offset was applied on the x-axis as shown in Fig. (4.9).
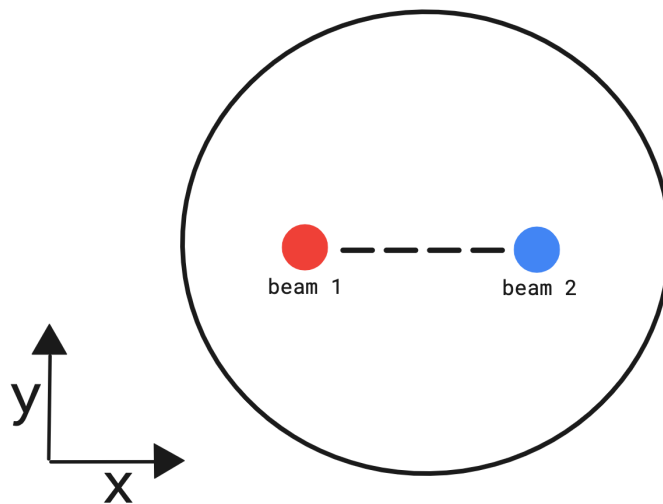
FIGURE 4.9: Horizontal offset of two beams inside the cavity.

The target of this analysis is to understand the impact of the offset on the power loss. In this specific case, two multi-bunch beams were used for the computation in CST using the filling scheme 5979 of Fig. (4.8). By defining the distance between beam 1 and beam 2 as $2x_0$ the power loss was computed for the following value of x_0: $1\,\mathrm{mm}$, $2\,\mathrm{mm}$, $3\,\mathrm{mm}$ and $4\,\mathrm{mm}$.

Chapter 5

Heating detection from pressure measurements

5.1 Overview

In general, machine learning techniques are data-driven methods. In order to detect heating from pressure measurements, a labelled dataset is needed to let the algorithm learn from the measurements indicating heating. Labeling a dataset consists in associating a value, the *label*, to each dataset element representing a particular class [56]. A labelled dataset on LHC heating issues observed by pressure measurement does not exist in the literature. Therefore, it was needed to build it specifically. Then, to this aim, a simple heuristic-based method was developed as baseline and several ML techniques were exploited and compared.

5.2 Labeling the dataset

The dataset is labeled by analysing each record and associating an appropriate label according to the particular class. Each pressure measurement has to be plotted and a binary value belonging to $\{0, 1\}$ has to be assigned, where 1 represents heating cases and 0 the pressure measurements where no obvious heating can be hinted from the data. Heating measurements are very rare with respect to the not-heating measurements, therefore, the construction of a statistically representative dataset turns out to be particularly tedious and time-consuming. On the other hand, it is possible to

download a large amount of pressure measurements unlabeled data from the CERN logging database.

The *k-means* [57] algorithm was exploited to find heating cases from the unlabeled data. The k-means algorithm clusters the data according to their mutual distances. Usually, the Euclidean distance metric is used. The underlying assumption is that heating cases should be close to each other. In particular, the Euclidean distance between two heating cases should be smaller than the distance between a heating and a not-heating case, so that the algorithm should be able to group them. By running the k-means algorithm multiple times for different values of k, the closest pressure measurements to the cluster centers were analysed after every run. In particular, the goal was to look for the value of k, for which at least one of the closest pressure measurements to one of the cluster centers is a heating example. The silhouette score [58] as a function of k for the unlabeled dataset was computed in order to help with this task and in order to minimise the k values, for which the cluster centers should be investigated, . The silhouette score is the average of the silhouette coefficient [58]. The silhouette coefficient can vary between -1 and $+1$, a value close to $+1$ indicates that a sample is situated inside its cluster. A value near 0 indicates that the sample is very close to the decision boundary between two neighboring clusters. A negative value indicates that the sample might have been assigned to the wrong cluster.

The silhouette score plot is represented in Fig 5.1. The points with higher silhouette score were investigated. In particular, one heating example is the closest example in the dataset to one of the cluster centers for $k = 12$. By plotting all the elements belonging to that cluster, 175 heating examples were found. The dataset was completed by choosing 175 not-heating samples. Since finding samples without heating indication is easier, they were chosen in order to have a dataset that would be as representative as possible of all different shapes of the pressure measurements. At the end of this process, a labeled and balanced dataset composed of 350 elements was obtained.

5.3 Preprocessing data

To perform ML classification each pressure measurement was interpolated on a fixed amount of *steps*. The fixed amount of *steps* was chosen to be sufficiently big to avoid
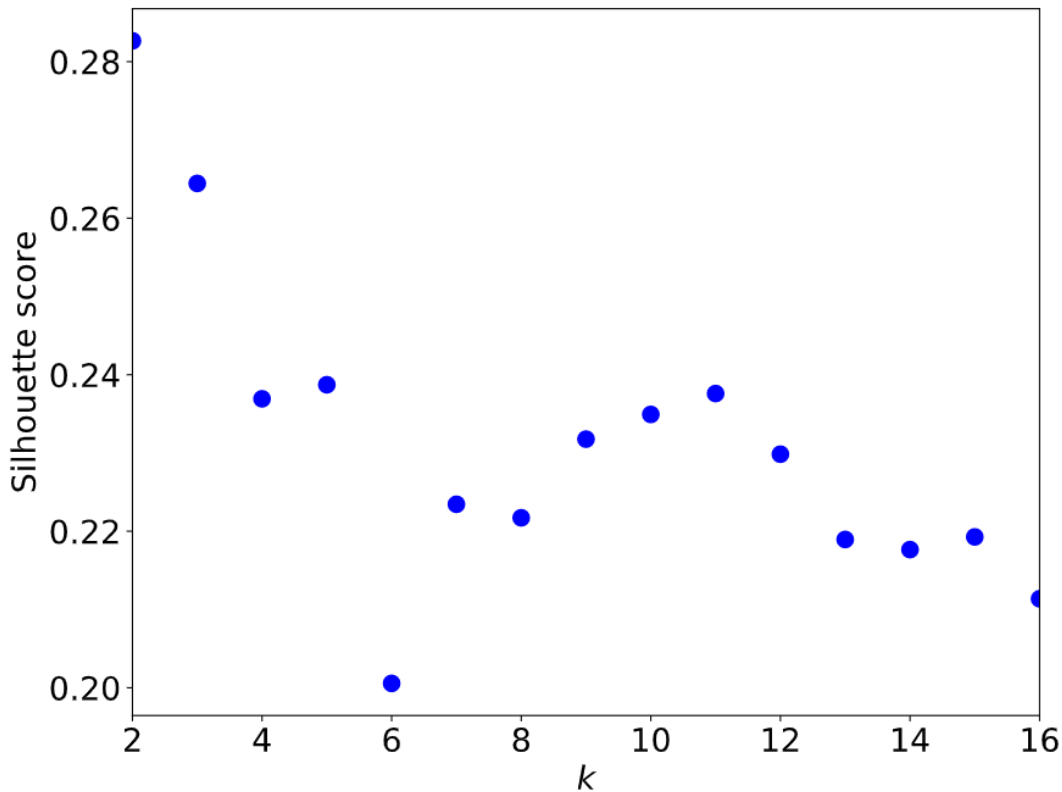
FIGURE 5.1: Silhouette score plot as function of the number of cluster k.

losing information during the interpolation (i.e 10% bigger than the longest pressure measurement sampled). By defining the total number of measurements $m = 350$ and the total *steps* as $n = 3000$, all the pressure measurements were stored row-wise in a matrix $X \in R^{m \times n}$. The labels were stored as a vector $T \in 0, 1^{m \times 1}$, where the n-th entry value of T is the label corresponding to the signal in the n-th row of X.

Taking into account the curse of dimensionality problem [59], this dataset could not be suitable for ML algorithms without further preprocessing. Indeed, there are too many features (i.e., the 3000 time *steps* of each pressure measurement during a fill) with respect to the amount of data (i.e, the 350 labeled examples). To avoid or alleviate this problem, several signal processing techniques were applied to reduce the dimensionality of the feature space and to alleviate the noise coming from the acquisition system while preserving the information useful for classification.

After several experimental investigations, the following preprocessing pipeline was

defined:

[1] Max-pool filter [60].

[2] Median filter [61].

[3] Savitzky–Golay filter [62].

[4] Standardization.

[5] k-means transformation.

As it will be discussed in Section 5.4, the ML algorithms performance is evaluated both when all the pipeline steps are applied, and when some of them are skipped, with the aim to evaluate their impact on the performance of the ML algorithms .

Max pooling is done by applying a max filter (i.e. filter that samples the max value in an array of values) to non-overlapping subregions of the initial representation of the signal. The median filter [61] cleans the signals from Gaussian noise produced by the measurement system. The Savitzky–Golay filter smoothens discontinuities if present. The standardization consists in subtracting from each pressure measurement its mean and dividing by its standard deviation. The k-means transformation goal is to further reduce the input size by transforming each signal into the distance from its closest cluster.

All these filters have hyperparameters that need to be tuned. A common technique in ML is to train the model with various combinations of the hyperparameters and pick the values of the hyperparameters that result in the best performance score. This technique is called *hyperparameter search*, and *Grid Search* [63] is a standard way to perform it.

Because each ML model has also its own hyperparameters to be tuned, some of these preprocessing hyperparameters were deliberately fixed to reduce the computational time. For the others hyperparameters, Grid-Search has been applied with the goal to maximise the accuracy score.

In particular for the max-pool, a filter size equal to 15 was applied, thereby reducing the number of *steps* per pressure measurement to 220. The median filter size was set to 11. The Savitzky–Golay filter order was set to 2. The filter size was set to 9. The

size of the filters was chosen by taking the best trade-off between reducing the signal size and noise while preserving heating indication.

The number of clusters k of the k-means algorithms used for the preprocessing was tuned with a Grid-Search. The set of values for the hyperparameter k given to the Grid-Search is $(5, 10, 30, 50, 100)$. The best value of k resulting from the Grid-Search depends on the algorithms that is following the k-means transformation. As shortly discussed, k-means transformation is not applied to all the algorithms as pre-processing step. When applied, a comparison between the performance of the algorithms with and without the k-means is provided. The value of the best performing k in terms of accuracy when k-means is applied is given in the following sections together with the description of which preprocessing step was applied to each algorithm. The impact of the filters and the k-means preprocessing on the ML algorithms performance is discussed in the results.

5.4   Algorithms

A simple heuristic-based algorithm and several ML algorithms were applied to the classification problem discussed so far, and compared. The selected ML algorithms are: Logistic Regression (LR) [43], k-Nearest Neighbours (kNN) [44], Random Forest (RF) [64] and 1-D Convolutional Neural Network (CNN) [49].

The impact of the preprocessing steps was analyzed by comparing the algorithms in the following scenarios:

- Comparison of all the models when only the max-pooling and standardization is applied.

- Comparison of LR, kNN and RF when max-pooling, standardization and k-means preprocessing is applied (no median and no Savitzky–Golay filtering).

- Comparison of all the models when the full preprocessing pipeline (all the steps presented in sec. 5.3) is applied to each model.

When the k-means transformation is applied the hyperparameter k is tuned with a Grid-Search.

The full preprocessing slightly differs for the heuristic-based algorithm and the CNN. In particular for the heuristic-based algorithm, the full preprocessing consists in all the preprocessing steps without the k-means transformation. For the CNN the full preprocessing steps are:

[1] Max-pool filter.

[2] Median filter.

[3] Standardization.

The motivations for this choice are discussed in sections 5.4.1 and 5.4.5.

5.4.1 Heuristic-based algorithm

A simple heuristic-based algorithm was developed and used as baseline to identify the heating patterns. The need of developing such an algorithm comes from the absence in the literature of approaches for this specific task.

The understanding of the pressure measurements patterns given by the numerous examples retrieved with the k-means algorithm used to label the dataset helped in developing a heuristic-based solution.

The idea behind is: if there is an increase in pressure that lasts at least a fixed amount of *steps* and is not related to the the energy ramp at the beginning of the fill (that usually induces peaks in the pressure measurement observed in Fig. 2.1) it has to be considered as abnormal heating.

The k-means preprocessing step is not performed to preserve the similarity between the raw signals and the preprocessed ones. At the end of the preprocessing procedure, each pressure measurement is a standardized signal of $N = 220$ *steps*. For each of them the following steps are performed:

[1] The *steps* axis is divided in K bins.

[2] The average of the pressure is evaluated for each bin.

[3] If the pressure average of a bin is higher than the previous one (neglecting the beam ramp peak as mentioned above) the measurement is marked as heating.

[4] The value of K that reaches the best performance is tuned.

The heuristic-based model results are discussed in Chapter 8.

5.4.2 Logistic Regression

For Logistic Regression (LR) the number of clusters k for the k-means preprocessing step (when applied) was tuned with a Grid Search. The value k resulting the best accuracy score is 30. To alleviate overfitting [65], l2 regularization [66] was applied.

5.4.3 k-Nearest Neighbours

The main parameter of the k-Nearest Neighbours (kNN) algorithm is the number of neighbours k_n (different from the number of clusters k for the k-means algorithm). With the help of a Grid Search both k (k-means) and k_n (kNN) were tuned. k_n was varied in the range $[3, 7]$ while k was varied in the set $(5, 10, 30, 50, 100)$. The parameters that resulted in the highest accuracy score are $k_n = 3$ and $k = 10$.

5.4.4 Random Forest

Random Forest (RF) is a powerful algorithm that can be used either for regression or classification task. The maximum number of leaf nodes and the number of estimators are the RF specific hyperparameters, which were varied. Their values were chosen with the help of a Grid Search . The first was searched in the interval $[3, 10]$ while the second in the set $(50, 100, 150)$. The number of clusters k of the k-means preprocessing (when applied) was varied in the set $(5, 10, 30, 50, 100)$ as for the previously described algorithm. The best performing combination of parameters resulted to be:

- Max leaf nodes = 4.

- Number of estimators = 100.

- k=30.

5.4.5   Convolutional Neural Network

The last approach adopted is a 1D-CNN. The preprocessing adopted for this approach slightly differs from the previously presented algorithms as mentioned above. In particular, the full preprocessing pipeline is composed of three steps only: Max-pool filter, Median filter and Standardization. The reason of such choice is that CNN networks are usually applied on input signals, which are poorly preprocessed insofar as convolutional layers already act as a sort of feature extractor (see, for example, reference [67]).

Layer	Parameters
Conv1D	filters: 10, kernel size=50, stride=1, activation= ReLu
MaxPooling	kernel size=15, strides=15
Conv1D	filters: 10, kernel size=50, stride=1, activation= ReLu
MaxPooling	kernel size=15, strides=15
Dense	units=1, activation=Sigmoid

TABLE 5.1: CNN layers.

The CNN model is summarised in table. 5.1. The first layer is a convolutional 1D layer with 10 filters, $kernel size = 50$x1, $stride = 1$, and ReLu activation function. The kernel size is the width of the filter. This layer uses each filter to perform a convolutional operation with the input. The stride represents how many *steps*, from left to right, each filter has to move when sliding over the input signal. The ReLu is the Rectified Linear Unit defined as $max\{0, x\}$ where x is the input value. After the first convolutional layer there is a max pooling layer, whose role is to reduce the size of the input of factor 15 by sampling the max value in each window of size 15. As last layer, a full-connected layer with only one output unit is applied, with sigmoid activation function to compute the output of the output unit. The sigmoid function is defined as

$$h_\theta(a) = \frac{1}{1 + e^{-a}},$$

where a is the activation of the output unit. The loss function used for training is the binary cross-entropy [68]. The model training was limited to 300 epochs for computational reasons.

For this network a Randomized Search [69] was used to tune the number of filters and layers to scan a wider range of hyperparameters. In this context a layer is intended to be the pair Conv1D - MaxPooling.

5.4.6 Training

To evaluate the models described above a variant of Cross-Validation [70] named Stratified Shuffle Split (SSS) [71] was applied. SSS is a variant of Shuffle &Split. Shuffle & Split consists in generating a user-defined number of independent train/test dataset splits. Samples are first shuffled and then split into a pair of train and test sets. Each train and test set is sampled with replacement. SSS differs from Shuffle & Split by returning stratified splits: it creates splits by preserving the same percentage for each target class as in the complete set [71]. SSS was performed with 10 splits with a test size of 33%. The heuristic-based algorithm does not need any training but in order to compare it with the ML models it was tested on the same test splits of the other models to provide an average of each score together with the confidence interval.

Part III

Results

Chapter 6

Results of impedance of a synchrotron component from beam-induced power loss

The methods for estimating the resonance frequency f_r and the Q–factor Q_r described in Chapter 3 have been applied to characterize an impedance inside the LHC. The considered fill is the 5979.

It is assumed that f_r is very close to the position of one of the main lines of the power spectrum. Before going further, it is important to emphasize that the following results are obtained numerically and some of the bunch lengths chosen for the computation cannot be reached by the LHC RF cavities: since it is a 400 MHz system the bunch length cannot exceed 5 ns. This practical constraint limits the minimum resonance frequency that can be estimated with the method to 550 MHz for the LHC. Those values have been chosen to make the example as clear as possible, in particular the power spectrum main lines of the shortest bunches in the range of interest are not visible with a linear scale.

The impedance chosen as test case for the method has $f_r = 400.78\,\mathrm{MHz}$ and $Q_r = 5 \times 10^4$. As previously mentioned, the value of the shunt impedance R_s can always be obtained by using Eq. (1.19) once f_r and Q_r are known, thus it will not be discussed in this section. The power loss computed for the fill 5979 with this impedance is named measured power loss.

First, the f_r value is to be computed. As explained in Section 3.2, it is sufficient to observe the power loss produced by the desired impedance as a function of the

bunch length. By computing the power loss produced by the impedance with several bunch lengths, from $1\,\mathrm{ns}$ to $8\,\mathrm{ns}$, the plot in Fig. 6.1 is obtained. At a bunch length of $7\,\mathrm{ns}$, there is the first relative minimum; therefore, to assess the value of the resonant frequency f_r of the impedance, it is necessary to observe the power spectrum of the $7\,\mathrm{ns}$ bunch length fill. The plot is shown in Fig. 6.2 where it is possible to observe that the first minimum of the main lines is around $400\,\mathrm{MHz}$, which is in good agreement with the assumed $f_r = 400.78\,\mathrm{MHz}$.

With the value of f_r evaluated, it is possible to apply the method explained in Section 3.3. In this case, the plot of Fig. 6.3 is obtained, which shows the value of Q_r corresponding to each value of α. To evaluate Q_r from this plot, the value of α must be estimated, and it can be obtained by evaluating the power loss as a function of the number of bunches. The result of simulations of the power loss, obtained by varying only the number of bunches for different α values, is shown in Fig. 6.4. The value of α that best approximates the measured power loss results seems to be 1.75.

In particular, by computing the mean square error (MSE) between the functions $f(\alpha) = \propto M^{\alpha}$ and the measured power loss for α in the range (1,2) (Fig. 6.5) the value of α that minimise the MSE is found to be 1.74. The evaluation of Q_r from this value of α in Fig. 6.3 leads to $Q_r \approx 5 \times 10^4$, as expected.
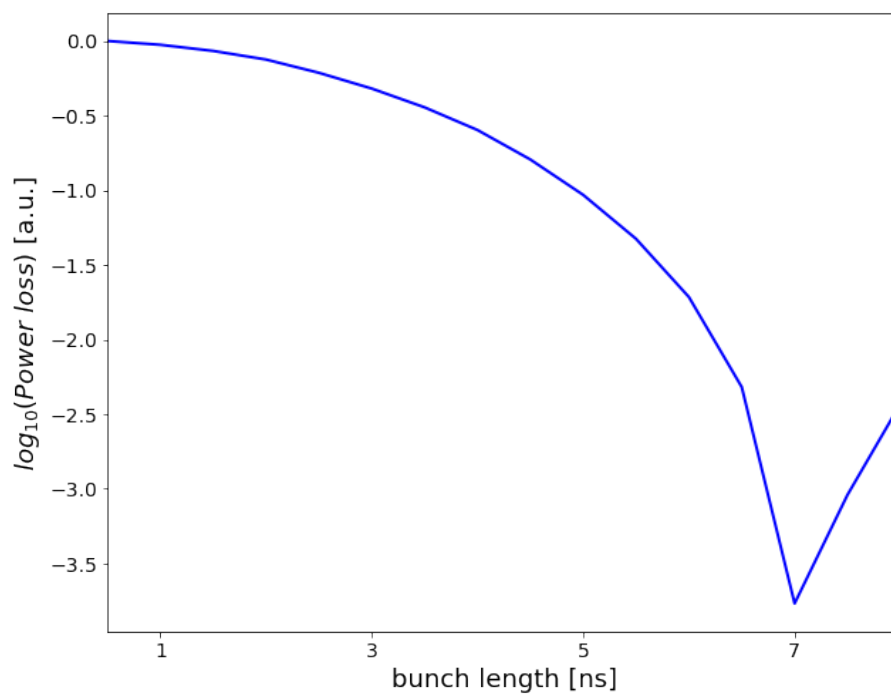
FIGURE 6.1: Logarithm of the power loss as a function of the bunch length with all the other beam parameters kept constant
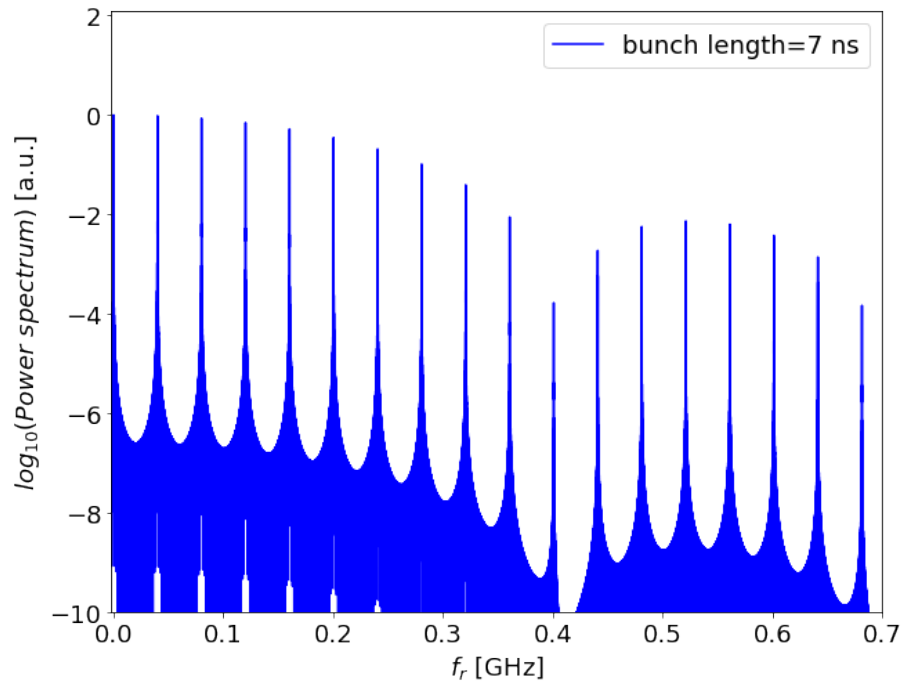
FIGURE 6.2: Logarithm of the power spectrum of a beam with a bunch length of 7 ns. It is important to observe that the first minimum is around $400\,\text{MHz}$.
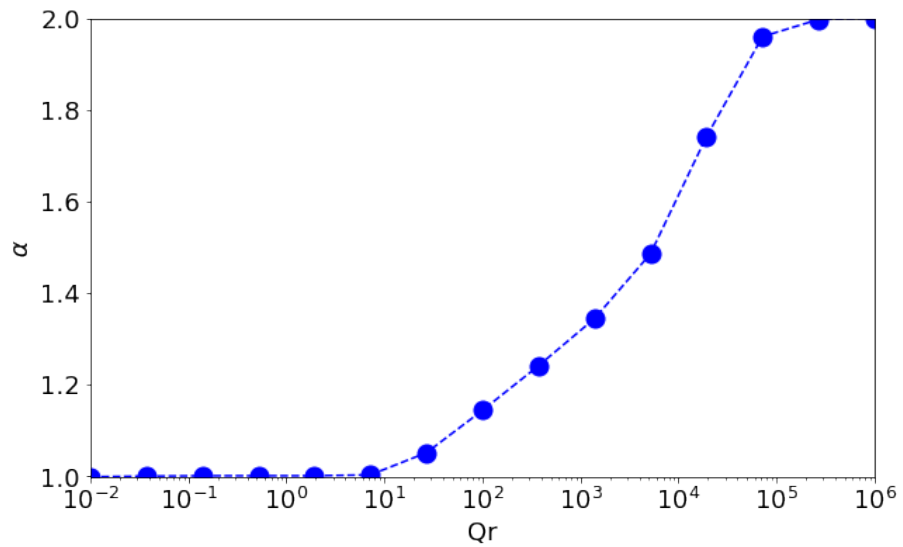
FIGURE 6.3: Assuming $P_{loss}(M) \propto M^{\alpha}$, α is plotted as a function of Q_r for $f_r = 400\,\text{MHz}$.
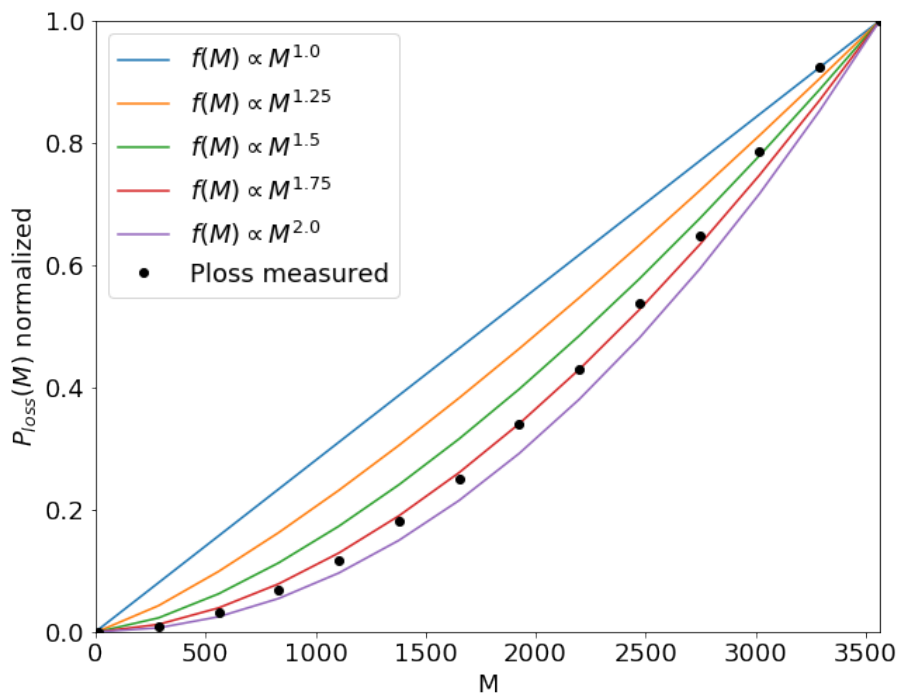
FIGURE 6.4: Power loss normalized to its maximum as a function of the number of bunches M. The black dots correspond to various measurements of the power loss produced by the impedance discussed in this Section by varying only the number of bunches. The lines are the functions $\propto M^\alpha$. The one that best fits that best fits the measurements is
$$f(M) \propto M^{1.75}.$$
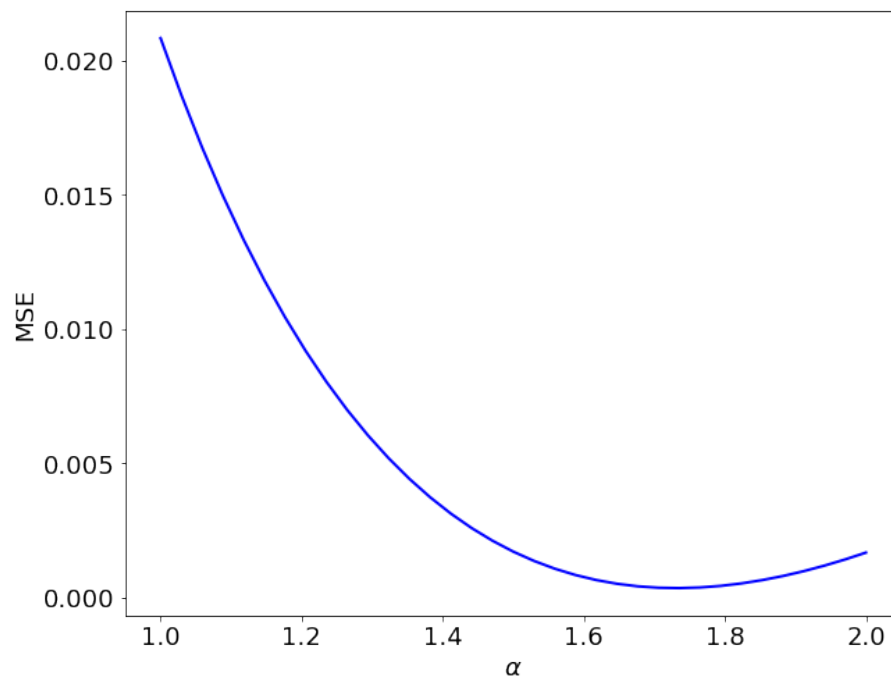
FIGURE 6.5: MSE between the functions $f(\alpha) \propto M^\alpha$ and the measured power loss for α in the range (1,2). The value of α that minimise the MSE is found to be 1.74.

Chapter 7

Results of power loss simulations

7.1 Single beam results

In Fig. (7.1), the power loss of a single beam with one bunch passing trough the cavity of Fig. (4.1) is represented as a function of simulation time. In particular, the single
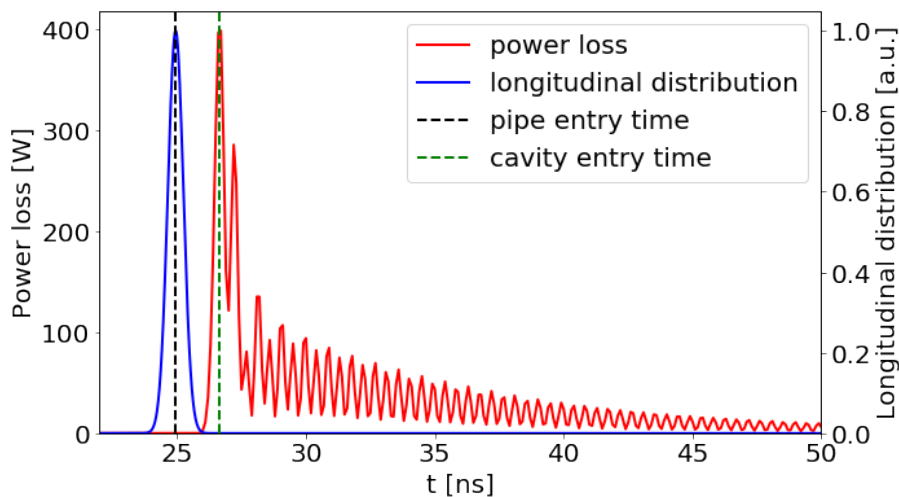
FIGURE 7.1: Instantaneous power loss dissipated inside the cavity as function of time (in red) generated by a single bunch longitudinal distribution passing through the cavity (in blue). The entry time in the beam pipe and the entry time in the cavity are also represented respectively as a vertical black dashed line and as a vertical dashed green line. The reference is taken at the maximum of the longitudinal distribution.

bunch longitudinal profile is plotted together with the instantaneous power loss dissipated in the cavity as a function of simulation time. In the plot are also represented

the entry time in the beam pipe and the entry time in the cavity, the reference for these lines being taken at the maximum of the longitudinal distribution. The peak of the power loss corresponds to the entry time of the bunch in the cavity, $1.67\,\text{ns}$ after the entry in the pipe due to its $50\,\text{cm}$ length . As already introduced in Chapter 4, the bunch parameters are:

- $N_b = 2.3 \cdot 10^{11}$

- bunch length = $1.2\,\text{ns}$

Due to the conductivity of the cavity and its resonant mode, the instantaneous power loss remains significant even $30\,\text{ns}$ after the bunch has passed trough the cavity. The instantaneous power loss reaches its maximum $1.67\,\text{ns}$ after the bunch maximum has entered the cavity and then it decays with an oscillation of $2f_r$ where f_r is resonant frequency of the mode TM_{010} (i.e. the only mode activated in this cavity). In Fig. (7.2)
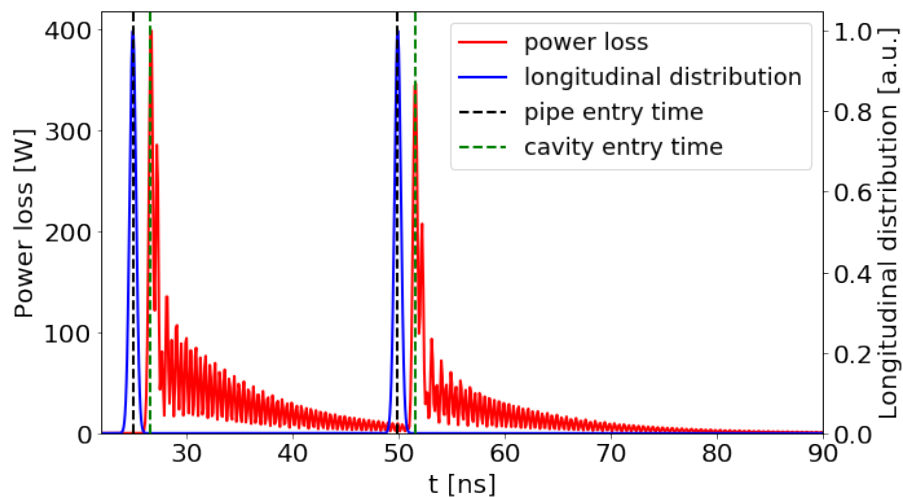
FIGURE 7.2: Instantaneous power loss dissipated inside the cavity as function of simulated time (in red) generated by two bunches passing through the cavity (longitudinal distribution in blue). The entry time in the beam pipe and the entry time in the cavity are also represented for each bunch respectively as a vertical black dashed line and as a vertical dashed green lines. The reference is taken at the maximum of the longitudinal distribution.

the same plot is shown with a beam with two bunches. Even if the two bunches are identical, the instantaneous power loss experienced when the second bunch enters is

overall lower. In fact, the power loss of the first bunch is not decayed completely yet and the coupling of the EM fields does not allow the second bunch power loss to reach the same value. This is due to the phase of the mode TM_{010} when the second bunch enters the cavity. If the phase of the mode is in the same direction as the phase of the mode imposed by the new bunch, the two fields will add up constructively. If the phase is opposite, the field of the first bunch will be subtracted to the field imposed by the new bunch, resulting in a lower instantaneous power loss, as is observed in Fig. (7.2).

The instantaneous power loss of fill 5979 computed by CST is represented in Fig. (7.3). The bunches of operational fills are all different due to the production process in the injector chain (i.e. different bunch length and different N_b). Also in this case, every time a bunch enters an empty cavity (i.e. the EM fields of the previous bunches are completely decayed) the instantaneous power loss reaches higher values. The first bunch of every batch has indeed higher instantaneous power loss.
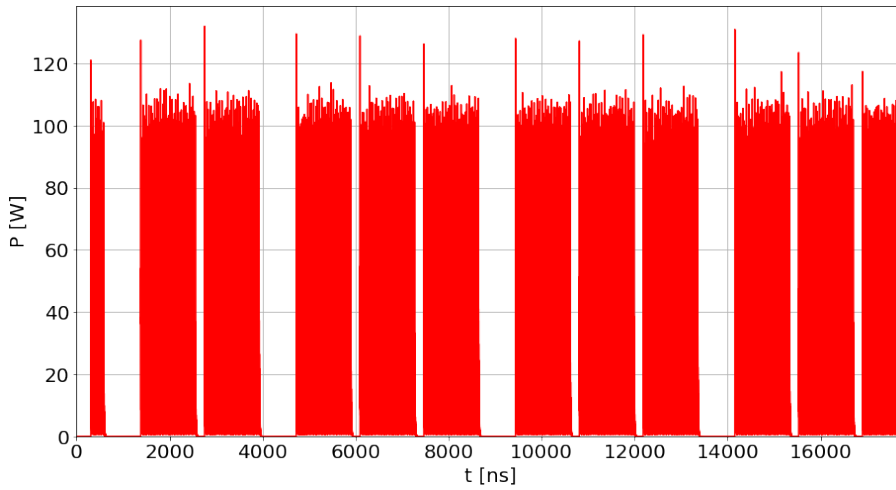
FIGURE 7.3: CST computation of the power loss for the first fifth of the LHC fill number 5979 on 21-07-2017.

Only the first fifth of the fill was computed with CST to reduce the very long computational time of the solver. By applying Eq. (4.1), the total power loss results to be:

$$P_{CST} = 8.05 \, \text{W}.$$

The same power loss was computed with the following parameters in Eq. (1.19):

- $\overline{N_b} = 1.1 \cdot 10^{11}$,

- $M = 2556$,

- $|\Lambda|^2$ represented in Fig. (7.4),

- $Re[Z]$ represented in Fig. (4.2).
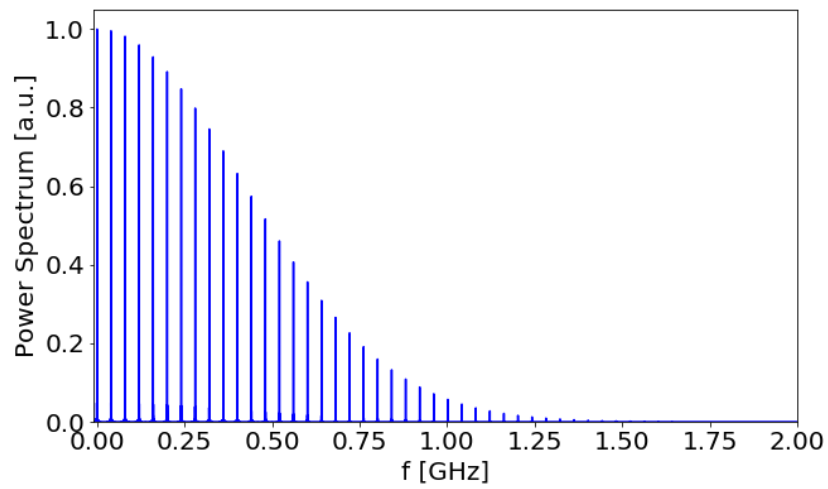
FIGURE 7.4: Power spectrum of fill number 5979 of a LHC run of the 21-07-2017.

The total power results to be:

$$P_{formula} = 8.11\,\text{W}.$$

Resulting in a ratio:

$$\frac{P_{CST}}{P_{formula}} = 0.99\%,$$

thus a very good agreement between CST and the analytical formula. This benchmark gives confidence to use CST for power loss computations.

7.2 Two beams results

The two beam formula Eq. (1.25) has a strong dependence on the phase shift τ_s. In particular the power loss is zero when $p\omega_0\tau_s = 0$, thus for $\tau_s = 0$.

In Fig. (7.5), there is a comparison between Eq. (1.25) and CST for the cavity described in Section 4.2 and a good agreement is found between the two approaches.
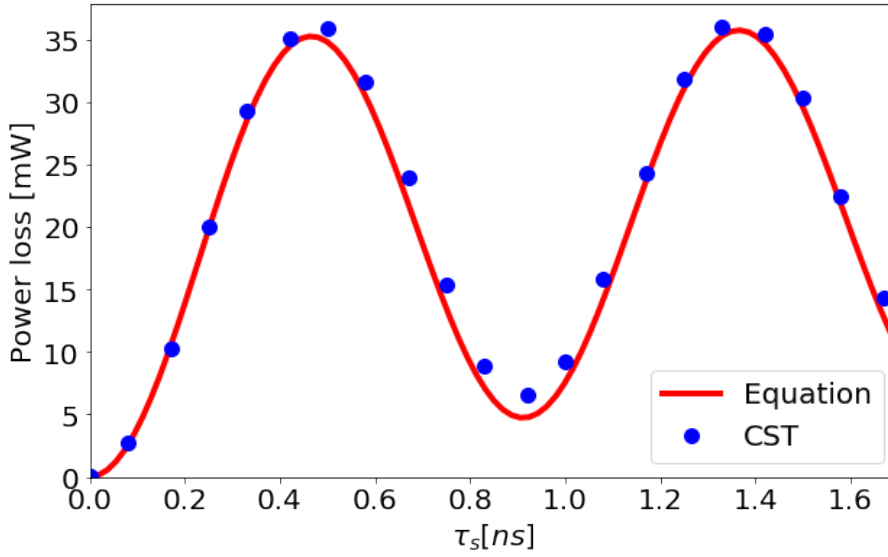
FIGURE 7.5: Two-beam power loss versus phase shift for the cavity introduced in Section 4.2. The analytical results are represented as a red line, while CST simulations are represented as blue dots.

Each of the CST simulated points is derived by the instantaneous power loss with Eq. 4.1. The power loss as a function of simulated time computed by CST for the specific cases of τ_s=0 ns and τ_s=0.4 ns is reported in Fig. (7.6). The first observation is that the power loss is zero for τ_s=0 ns. The two bunches are approaching the cavity at the same time and both of them are trying to activate the TM_{010} mode. Because the bunches are identical but counter-rotating they are activating the mode with the same H_ϕ^{010} field constant along z (see Eq. 1.28 for $p = 0$) but with opposite sign. In particular:

$$H_\phi^{010} = H_{\phi 1}^{010} + H_{\phi 2}^{010} = H_{\phi 1}^{010} - H_{\phi 1}^{010} = 0,$$

where $H_{\phi 1}^{010}$ refers to the magnetic field along the ϕ direction activated by the first bunch. Analogously $H_{\phi 2}^{010}$ refers to the magnetic field along the ϕ direction activated
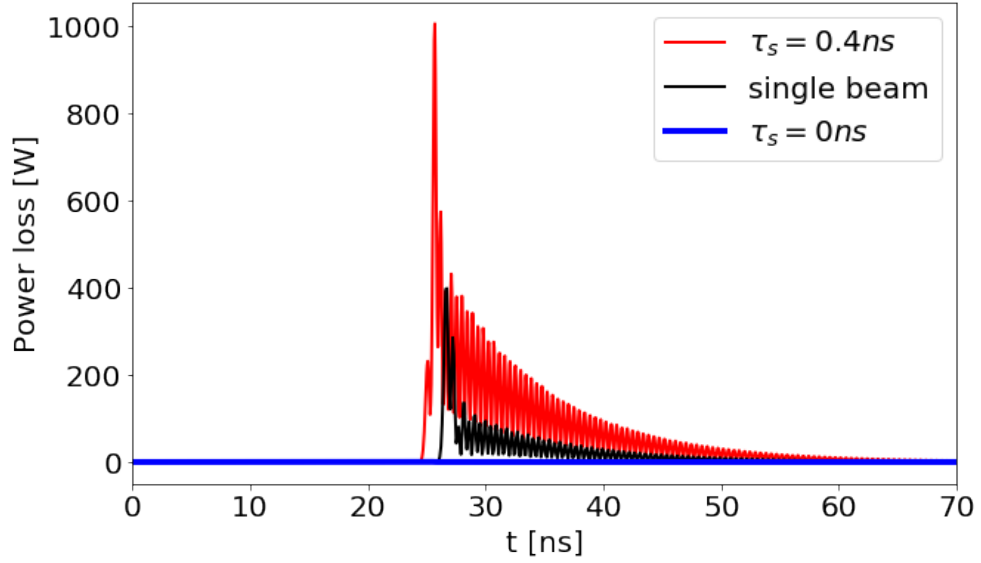
FIGURE 7.6: Instantaneous power loss for the case of $\tau_s = 0$ (in blue) and $\tau_s = 0.4ns$ (in red) of Fig. (7.5). Also the single beam single bunch case is shown for comparison in black.

by the second bunch. The two fields are canceling out thus the power loss is zero. When the phase shift between the two beams differs from 0, the fields do not cancel anymore. The power loss is observed to have an oscillating behaviour with the phase shift (see Fig. 7.5). The period of the oscillation (time between two consecutive minimum) is observed to be $\frac{1}{f_r}$ where f_r is the resonance frequency of the mode TM_{010} in the cavity. This correlation with f_r will be presented in the next section for different radii of the cavity (i.e. different f_r). The fact that the frequency of the oscillation of the power loss with phase shift is f_r can be explained by the fact that H_ϕ has a period of $\frac{1}{f_r}$. Indeed, for a phase shift of $\frac{2}{f_r}$ the field $H_{\phi 1}^{010}$ induced by the first bunch is at its maximum amplitude with opposite sign and is therefore perfectly in phase with the field $H_{\phi 2}^{010}$ induced by a second bunch entering the cavity, leading to a maximum in Fig. 7.5.

In Fig. 7.7, the two-beam and single beam cases are compared for the same impedance (i.e. same cavity). While the single beam case does not depend on the phase shift τ_s, the two-beam case presents a strong dependence with phase shift, varying from zero to almost four times the single beam power loss. When the fields from the two bunches fully add up constructively, one could expect that the power loss would be 4
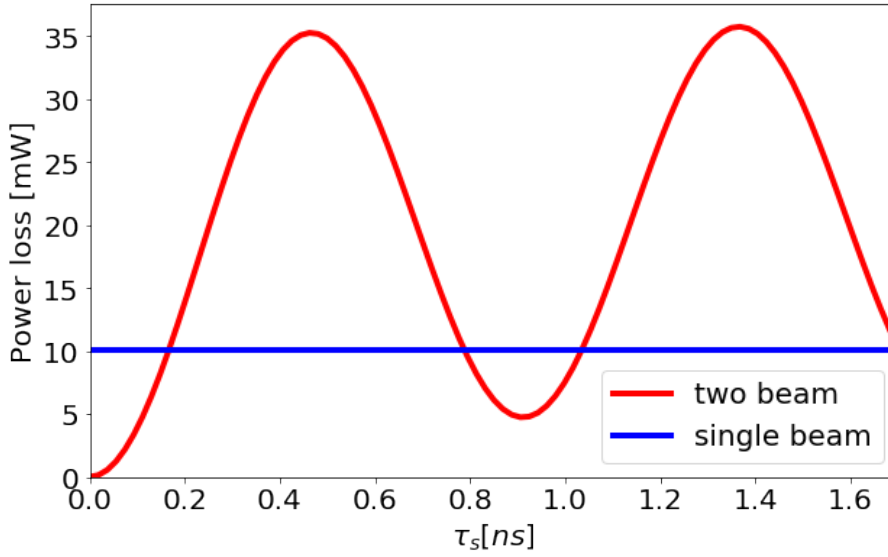
FIGURE 7.7: Comparison between single beam and two beam power loss.

times the power loss caused by a single bunch. It can be noted here that the first maximum in power loss (at $\tau_s = 0.45ns$) is slightly lower than four times the single beam power loss. Since the field of the first bunch decays exponentially until the arrival of the second bunch, when the second bunch enters the cavity its field is in phase with the field left by the first bunch but the amplitude of the latter is smaller. The large fields from the first bunch caused single beam power loss in the cavity before the arrival of the second bunch that could not add up constructively with the fields of the second bunch, and therefore the dissipated power of two bunches entering with a non-zero phase delay cannot reach four times the power loss of a single bunch. The second minimum (i.e. the minimum at $\tau_s = 0.9ns$) is not zero because the first bunch has already dissipated power for 0.9 ns before the arrival of the second bunch. Moreover, the fields left by the first bunch are decayed and do not cancel out with the field of the second bunch even if they have the exact opposite phase. For this reason, it is expected that the exact cancellation of fields and power loss only occurs when there is no phase delay between the bunches.

It is important to note that these observations are valid only if the structure is symmetric with respect to the plane orthogonal to the direction of the counter-rotating beams, and if the counter-rotating beams share the same longitudinal distribution.

For standard operation of a collider, the phase shift between the two counter-rotating beams at a given accelerator component depends only on its longitudinal distance from the nearest interaction point. The knowledge of this distance to the interaction point allows predicting the power loss of a resonating cavity with a single TM_{010} mode. As a consequence, the positions that minimize the power loss to components can be favoured for installation, depeding on the frequency of their resonating modes.

7.2.1 Cavity radius

The cavity radius a is directly related to the first mode of the pillbox cavity, the TM_{010} mode. Considering two beams with only one bunch each, CST and Eq. (1.25) were compared sweeping the cavity radius from $13\,\mathrm{cm}$ to $15\,\mathrm{cm}$ with a step of $0.5\,\mathrm{cm}$ while the power loss versus the phase delay τ_s was computed (see Fig. 7.8). The agreement is consistent for all radii.
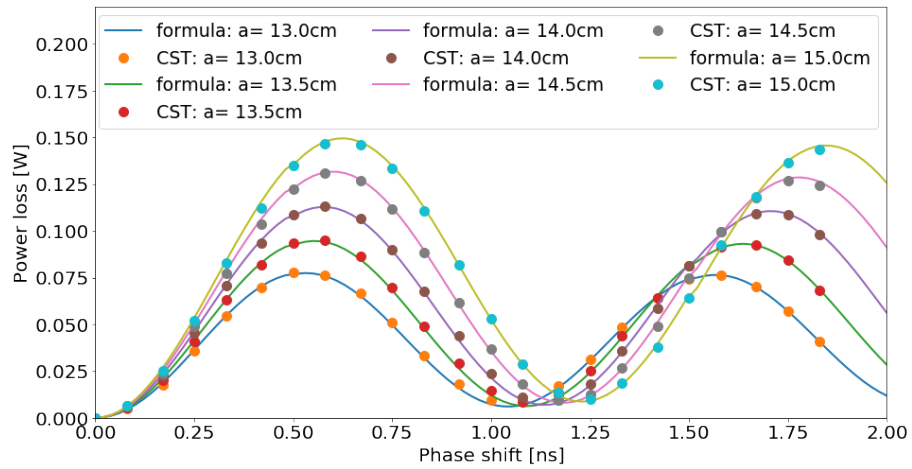
FIGURE 7.8: Comparison between CST and Eq. (1.25) (named *formula* in the plot) for different cavity radii. Two single-bunched beams have been used.

By varying the cavity radius, the transit time of the bunches inside the cavity does not vary. Thus the only impact on the power loss computation is given by the activation of other mode frequencies under the beam spectrum. In Fig. 7.9, it is shown

that all the impedances for the different radii have only one main mode (TM_{010}) under the beam spectrum, as designed.
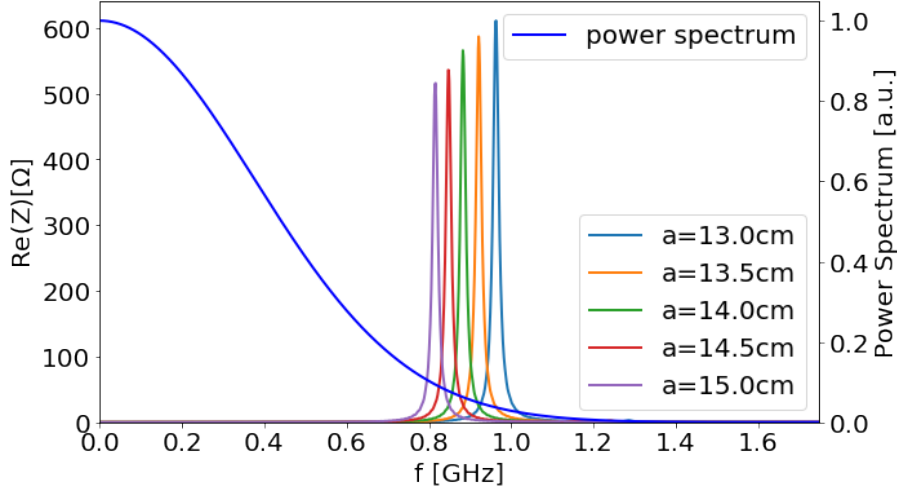
FIGURE 7.9: Impedance for different radii overlapped with the power spectrum.

As expected with the reasoning developed for the case with one radius, the period of the power loss undulation with τ_s is found equal to the inverse of the resonance frequency f_r of the TM_{010} mode for each radius. This can be observed in Fig. 7.10.

7.2.2 Cavity length

The length h of the cavity impacts directly on the transit time of the bunches inside the cavity: the bunches take more time to pass through a longer cavity. In Fig. (7.8) the power loss versus τ_s is shown for various cavity lengths h. In this example, the radius of the cavity was set to $a = 14.5\,\mathrm{cm}$. For $h = 10\,\mathrm{cm}$, one can observe that the agreement between CST and Eq. (1.25) is not as good as for shorter lengths. Indeed, the assumption used in the analytical formula that $\frac{L}{s}$ is small (see Section 1.4.3) is not valid anymore.

If one increases the cavity length further to $h = 30\,\mathrm{cm}$, the agreement between CST and Eq. (1.25) is lost as it can be observed in Fig. (7.12).
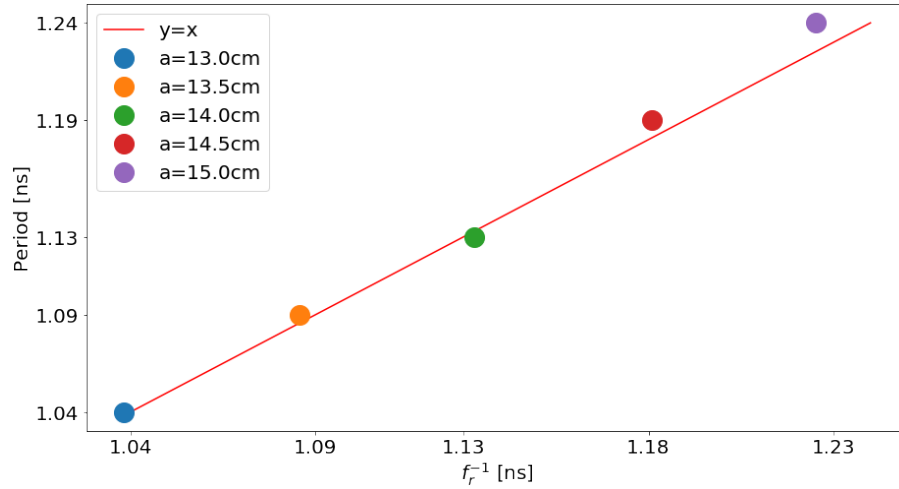
FIGURE 7.10: Period of the power loss undulation with τ_s versus the inverse of the resonance frequency f_r of the TM_{010} mode for each radius. Also the function y=x is plotted for comparison.
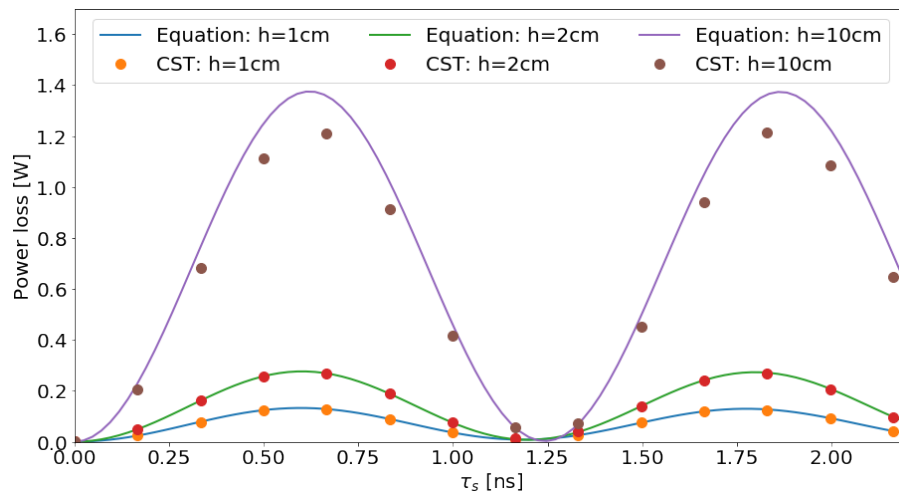
FIGURE 7.11: Power loss versus the phase shift τ_s is shown for cavity lengths h. The agreement between CST and Eq. (1.25) starts to break for h=10 cm.

More importantly, for $\tau_s = 0\,\text{ns}$, one can observe that the simulated power loss is not zero anymore. To analyse this feature further, h is varied from $2\,\text{cm}$ to $100\,\text{cm}$ with step of $2\,\text{cm}$, keeping $\tau_s = 0$ in Fig. (7.13).

This shows that the power loss for $\tau_s = 0$ strongly depends on h, and that the
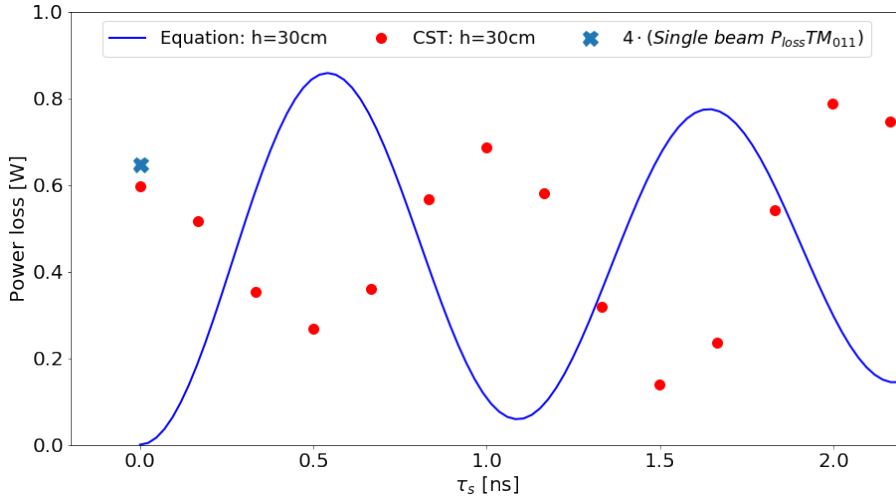
FIGURE 7.12: Power loss versus phase shift τ_s for h=30 cm.
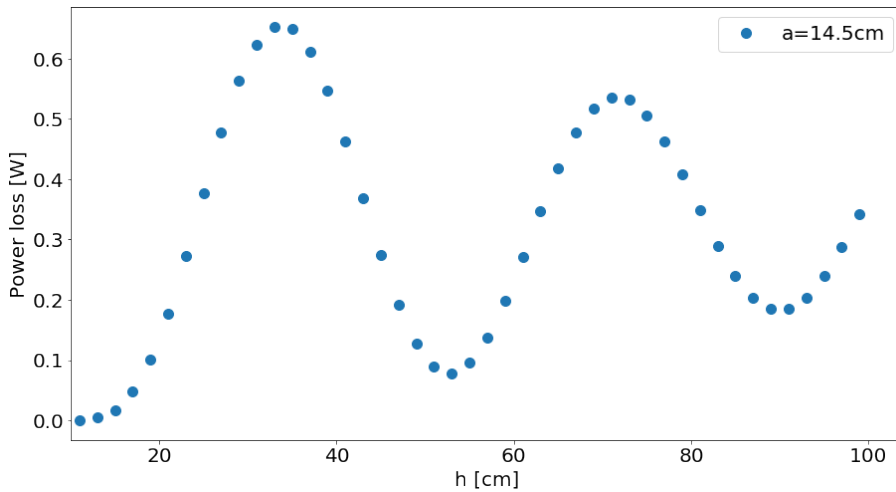
FIGURE 7.13: Power loss in two-beam case when h is varied from $2\,\text{cm}$ to $100\,\text{cm}$ at step of $2\,\text{cm}$ keeping the phase shift $\tau_s = 0$.

situation described in the beginning of Section 1.25 where fields compensate for $\tau_s = 0$ is no longer valid when the length of the cavity increases.

In Fig. (7.14), the power spectrum and the modes in the cavity are plotted for some of the lengths considered in Fig. (7.13). When the length h is increased, the mode TM_{011} is shown to contribute more to the power loss due to its larger shunt impedance and in particular to its lower resonance frequency . In particular for $h = 30cm$ the TM_{011} appears at $980\,\text{Mhz}$. In fact, one can see for $\tau_s = 0$ in Fig. (7.12),
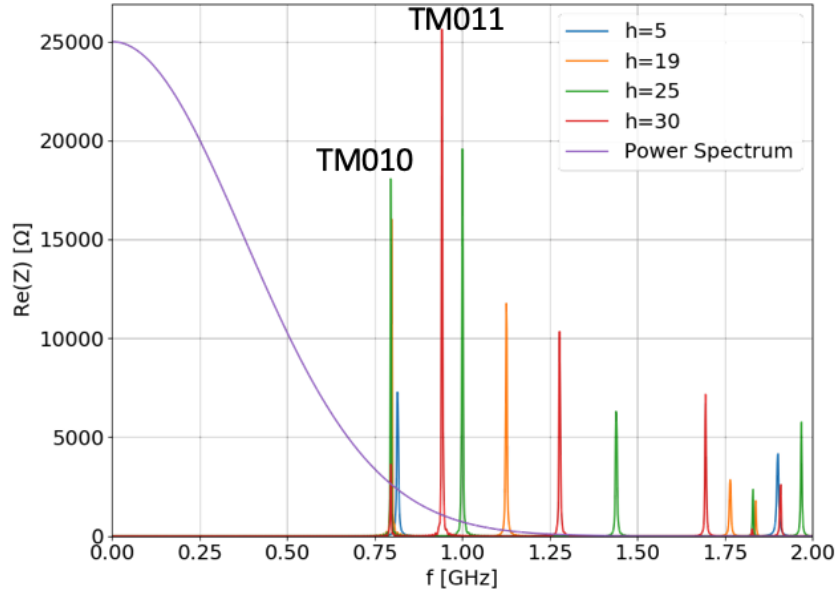
FIGURE 7.14: Modes in the resonant cavity for $a =14.5\,\text{cm}$ at various h. Also the power spectrum is represented in arbitrary unit.When the length h is increased, the mode TM_{011} is shown to contribute more to the power loss due to its larger shunt impedance and in particular to its lower resonance frequency.

that the power loss with the two bunches is comparable with four times the single beam power loss computed with only the TM_{011}. It could be concluded from these observations that the non-zero power loss for longer cavities is due to mode TM_{011}. Indeed, contrary to the case of the TM_{010} mode that has no zero crossing of the field along the longitudinal direction, the TM_{011} mode has a zero crossing of the field. Therefore, when the two bunches enter the cavity at the same time they can excite constructively the TM_{011} with the same sign and direction. This creates an initial condition in the cavity of:

$$H_\phi^{011} = H_{phi1}^{011} + H_{phi2}^{011} = 2H_{phi1}^{011}$$

where H_{phi1}^{011} and H_{phi1}^{011} are the electric fields induced by beam 1 and beam 2 respectively on the z-axis for the TM_{011} mode. Since at $\tau_s = 0$, the field components of the TM_{010} excited by the two bunches arriving in phase cancels out, the TM_{011} is the only

mode contributing to the power loss at this phase shift. The power loss in the cavity scales with $|H_{phi}|^2$, explaining the factor four.

7.2.3 Beams horizontal offset

Considering two beam with the filling scheme 5979 represented in Fig. (4.8), the power loss versus the horizontal offset was simulated. Computing the power loss with CST the results are plotted in Fig. (7.15).
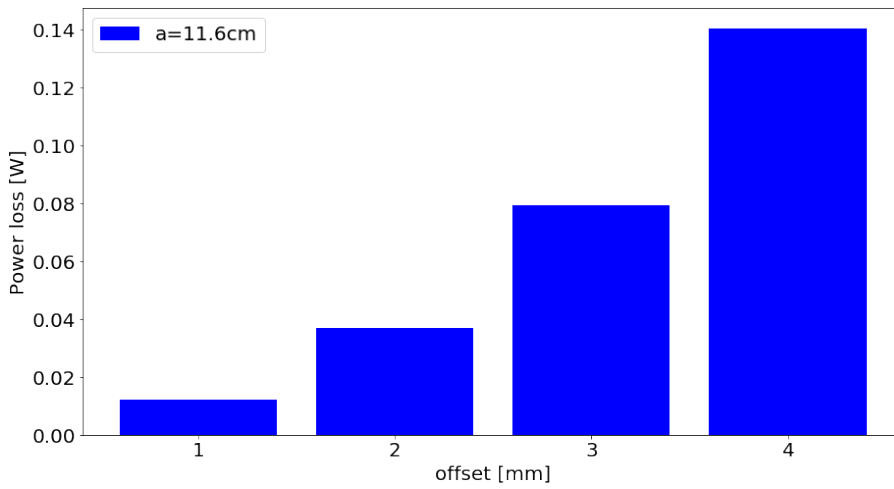
FIGURE 7.15: CST computation of the power loss for two beams with filling scheme 5979 versus the horizontal offset. The two beams are in phase (i.e. $\tau_s = 0$.

Considering that the total power loss of the considered beam for the single beam case is $8.11\,\mathrm{W}$ the impact of the offset can be considered negligible for the range of considered offsets.

7.2.4 Multi-bunch phase shift

The comparison between CST and Eq. (1.25) for two multi-bunch beams is represented in Fig. (7.16). The two beams have the filling scheme of fill 5979 previously presented (Fig. 4.8).

The cavity parameters were presented in Section 4.2. The agreement is therefore found to be good also for the multi-bunch case. The negative values of the phase-shift
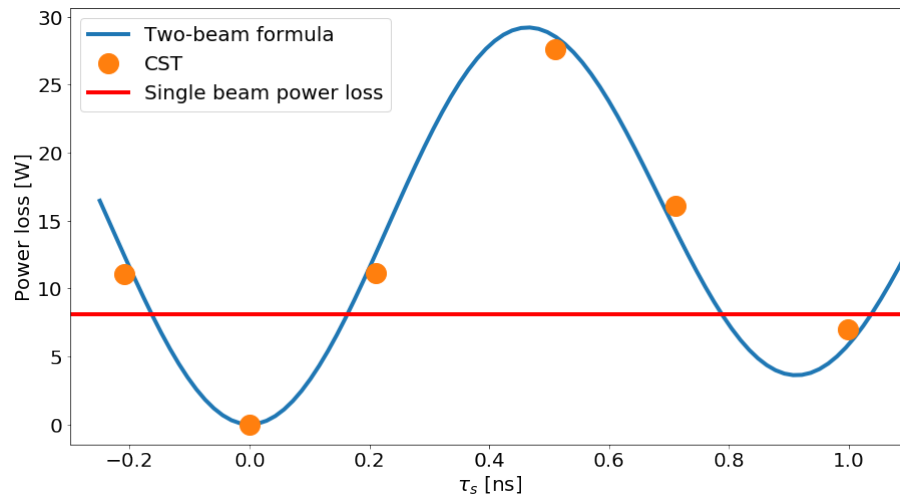
FIGURE 7.16: Comparison between CST and Eq. (1.25). Two multi-bunch beams have been used with filling scheme 5979.

τ_s indicate that the shifted beam is arriving earlier in the cavity, while positive values indicates that it is arriving later in the cavity. Also in this case the two-beam power loss can reach values up to almost four times the single beam power loss.

Chapter 8

Results of heating detection from pressure measurements in the CERN Large Hadron Collider

The result of the filtering flow is summarised in Fig. 8.1 from top to bottom. The raw signal (Fig. 8.1a) is given by the monitoring system and then it is processed by the max filter. The output of the max filter (Fig. 8.1b) completely preserves the broad pressure increase at the end of the signal by reducing the *steps* from 3000 to 220. It is interesting to observe that the max filter preserves also the peaks at the beginning of the fill. After the max filter, the signal is processed by a median filter (Fig. 8.1c), which aim is to reduce the possibly still present noise. The median filter destroys sharp peaks as the one at the beginning of the fill. According to the experts those peaks are not relevant heating indication since thermal processes in this context are very slow compared to the timescale of these sharp peaks. The length of the median filter is chosen to suppress only the peaks that are too short and sharp to indicate heating. Finally, the Savitzky–Golay output represented in Fig. 8.1d smoothens the signal.

The k-means algorithm was applied to transform the pressure measurements into their distances from the cluster center. When applied after the full preprocessing pipeline, it is possible to observe that the k-means transformation is able to split the heating cases from the not-heating ones quite well with already $k = 2$, as shown in Fig. 8.2. This can be explained by the fact that, after the filtering and the normalization, each signal shares the same scale and most of the outliers are removed. When
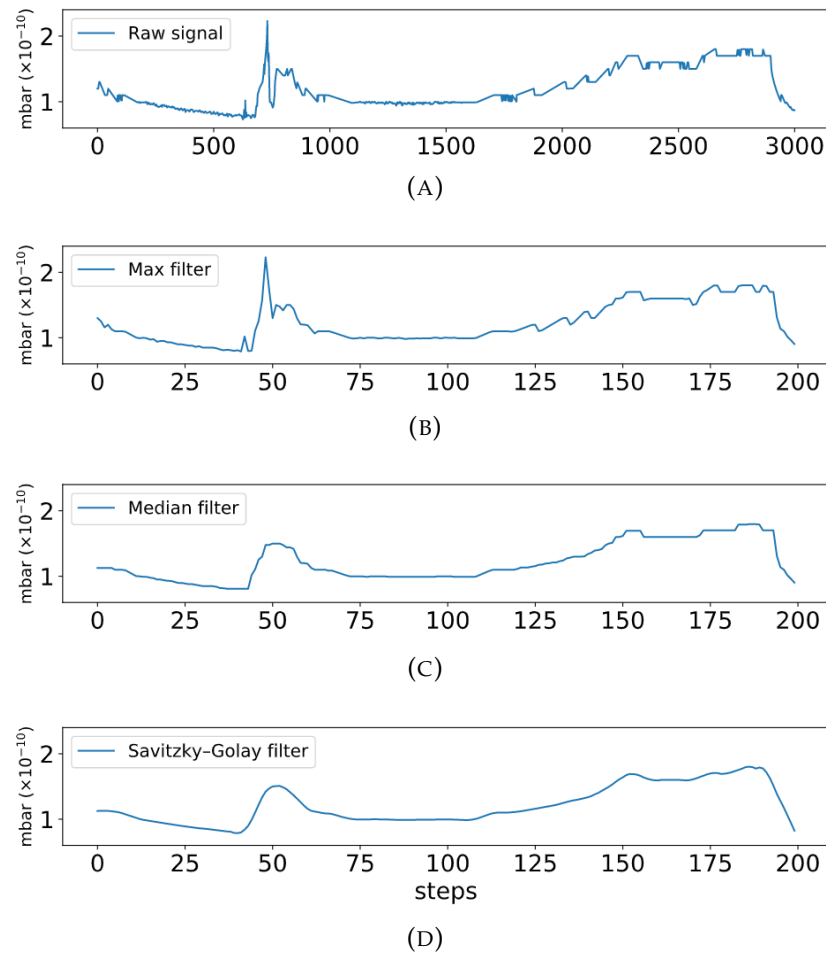
FIGURE 8.1: Filtering of the input signal: (8.1a) raw input signal, (8.1b) output of the max filter, (8.1c) output of the median filter and (8.1d) output of the Savitzky-Golay filter [62].

the distance from each cluster is computed, all the heating measurements that usually present an increasing pressure pattern towards the end of the signal turn out to have similar distance from the clusters centers and thus they are grouped together. The same happens for the not-heating cases but with different distances with respect to the heating ones. The difference in distances from the cluster centers results in the split of the two classes into two different areas of the plot.

The models were compared in three scenarios by applying different steps of the preprocessing techniques. The results are summarised in table where the average performance and its standard deviation is reported Table 8.1 shows the performance
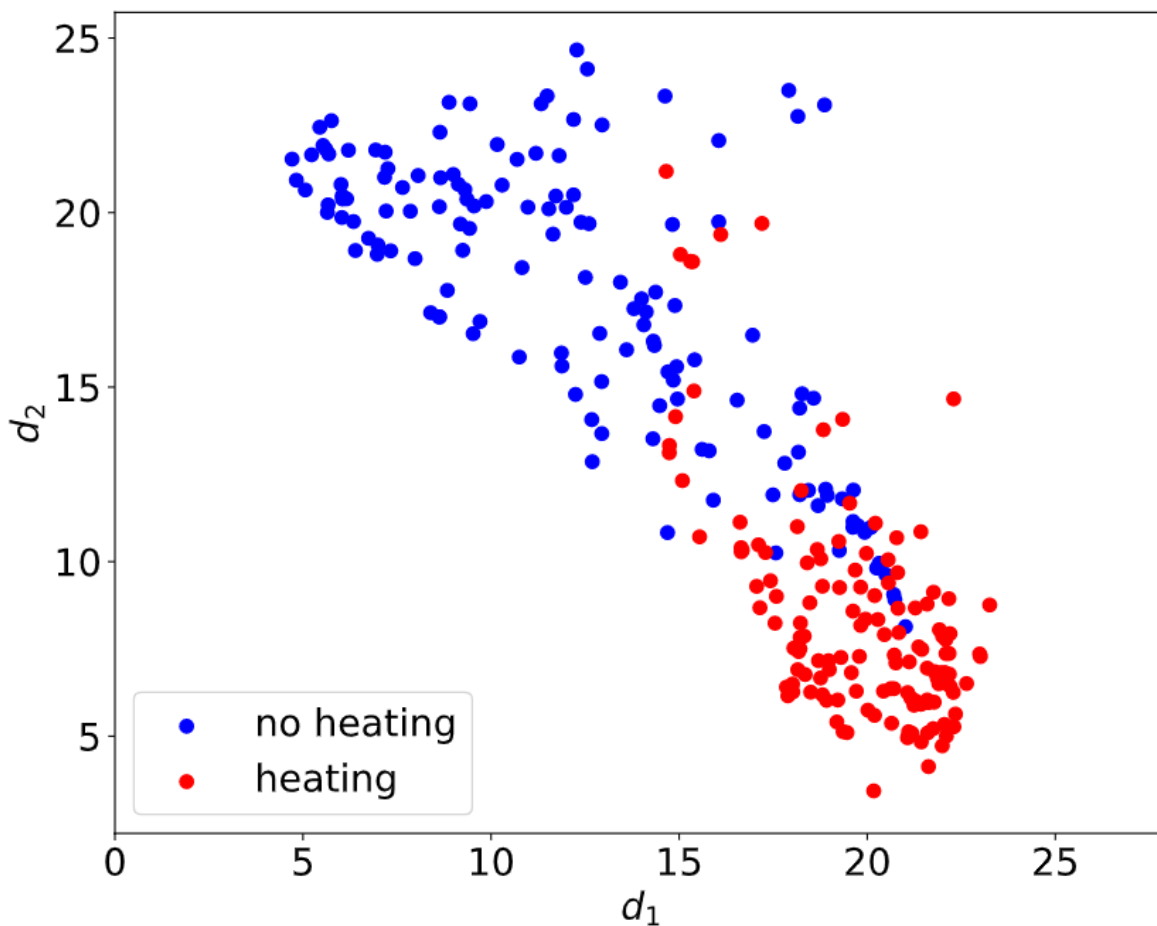
FIGURE 8.2: k-means distances computed with two clusters (the x- and y-axis represent the euclidean distance of each pressure measurement from the first and the second cluster center, respectively.

of the algorithms when only the max-pooling to reduce the dimensionality and the standardization is applied. From these first results it can be observed that Logistic Regression and 1-D CNN already performs very well with minimal preprocessing. The best performing model in this case is the 1-D CNN. It is also important to underline that all the models outperform the baseline model (i.e. the analytical solution built for comparison) both in accuracy and precision.

Table 8.2 shows the results when also the k-means transformation is applied as

TABLE 8.1: Results summary with max pooling and standardization.

Model	Accuracy	Precision	Recall
Heuristic-based Algorithm	0.78±0.03	0.72±0.02	0.93±0.03
Logistic Regression	0.94±0.02	0.94±0.03	0.93±0.03
kNN	0.93±0.02	0.91±0.02	0.95±0.03
Random Forest	0.90±0.03	0.88±0.04	0.92±0.04
1-D CNN	0.95±0.02	0.96±0.03	0.95±0.03

TABLE 8.2: Results summary with max pooling, k-means transformation and standardization.

Model	Accuracy	Precision	Recall
Logistic Regression	0.93±0.02	0.93±0.03	0.94±0.03
kNN	0.96±0.02	0.96±0.03	0.96±0.03
Random Forest	0.94±0.02	0.96±0.03	0.92±0.03

preprocessing step. Random Forest scores improve significantly by adding this preprocessing step. Also kNN benefits from the k-means transformation. Logistic Regression gains little improvement in recall but performs slightly worse in accuracy and precision.

Table 8.3 summaries the best performance achieved by each model. Those scores are achieved when the full preprocessing pipeline is applied to Logistic Regression, kNN and Random Forest. The heuristic-based algorithm uses the full filtering preprocessing without the k-means transformation. The 1-D CNN follows the preprocessing described in Section 5.4.5.

In this case the heuristic-based model reaches very high recall (0.99) but with low precision with respect to the other models. It is important to observe that all the ML models outperform the heuristic-based model both in accuracy and precision, and

TABLE 8.3: Results summary with full preprocessing.

Model	Accuracy	Precision	Recall
Heuristic-based Algorithm	0.77±0.02	0.69±0.02	0.99±0.01
Logistic Regression	0.96±0.02	0.96±0.02	0.95±0.03
kNN	0.96±0.01	0.97±0.02	0.95±0.03
Random Forest	0.95±0.02	0.97±0.02	0.92±0.03
1-D CNN	0.96±0.02	0.96±0.02	0.95±0.02

that using all the preprocessing steps allows gaining performance with respect to results in table 8.1. It is also interesting to see that the heuristic based model does not gain much from the preprocessing, except from the recall point of view.

The 1D-CNN model has a very good precision/recall trade-off: its recall is comparable to the other models with the advantage of lower standard deviation. Moreover it does not need an elaborate preproccessing of the inputs to reach very good scores. The 1D-CNN was trained on 300 epochs for each of the train and test split of SSS. The train and test cross-entropy loss during training is shown in Fig. 8.3. The loss is averaged on the split folds used during training and the standard deviation is also reported as error bar. It is reasonable to assume that the performance of the 1D-CNN would improve by increasing the training epochs that were limited to 300 for computational reason.
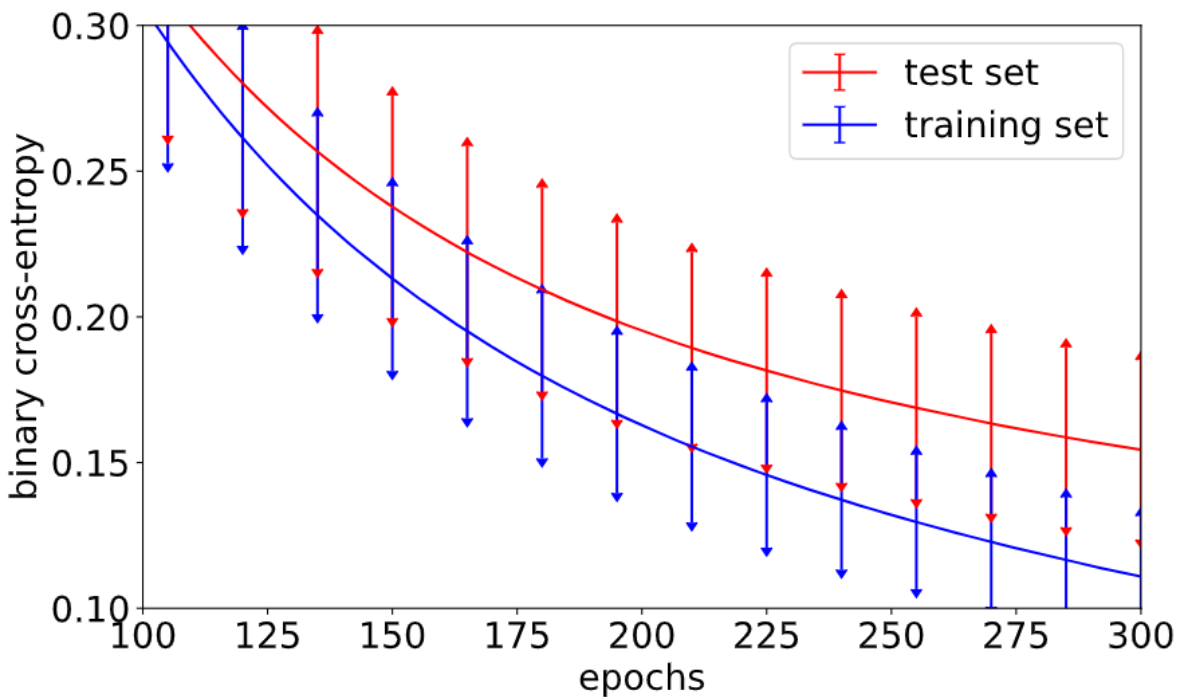
FIGURE 8.3: Binary cross-entropy loss for training and test sets. The loss is averaged on the split folds used during training. The standard deviation is also reported as error bar.

As discussed in Chapter 5, the best model to be chosen for the implementation depends on the particular application of the detection tool. When it is crucial to not

miss heating cases, the models with higher recalls should be preferred, even if a trade-off with precision is needed. When the historical data has to be searched to look for undetected heating cases then the model, which exhibits higher precision has to be preferred.

Conclusion

A method for a first-order estimation of the resonance frequency and the quality factor of the impedance of an accelerator component was proposed. A practical example of the application of the method on an LHC component was reported. Pending specific practical issues that need to be solved for each machine, the results show that the method is suitable to estimate the resonance frequency and the impedance Q-factor.

One of the crucial advantages of this method is the ability to predict impedance characteristics without requiring components be disconnected from the machine: it uses power loss measurements, acquired during several machine runs with various specific filling schemes.

The method requires only two assumptions: 1) the bunch length of a given beam can be changed in a wide range without varying any other beam parameter, and 2) the resonance frequency f_r lies approximately over one of the main lines of the beam spectrum. The latter assumption can be waived if the bunch spacing can be also changed within a wide range. Nevertheless, further work will be dedicated to remove this latter assumption that the resonant frequency f_r needs to be close to one of the main line of the beam spectrum.

The two-beam power loss was studied in a resonant cavity. It was shown that the phase shift between the beams entering the cavity plays a major role on the power loss. Indeed, the power loss of two beams in a cavity can vary from zero to four times the power loss of a single beam. For the specific case of a given mode TM_{010}, it is possible to predict the location in the accelerator with respect to the interaction point that minimizes the power loss in presence of two beams. As shown with an example, the studies performed on two single-bunch beams also apply to two multi-bunch beams.

Various machine learning techniques have been presented to detect abnormal

heating from the many available pressure measurements in the CERN LHC. In particular, these ML models have been evaluated experimentally to investigate their ability to automatically detect abnormal heating. It is important to note that all the selected ML models outperform the heuristic-based model properly built for the task as baseline. In particular, Logistic Regression, kNN and 1D-CNN exhibit the most promising performance. It has also been shown how customized preprocessing techniques could improve the performance of all the ML models for this classification task. Having such automatic tool to detect heating is expected to be of significant help for the next LHC run. The tool has already been in use internally at CERN to study old heating effects and it is planned to be used during operation in the next LHC Run that will start in 2021.

Appendix A

Computing power loss with CST

CST Studio is widely used in accelerator physics to compute the beam impedance of an accelerator component. Following an update of the software in 2018, it is now possible to compute the power loss directly with CST. To properly compute the power loss, it is important to set the *charge* parameter to the total beam charge in Fig. (A.1) in the beam definition.

The total beam charge is computed as:

$$q_{beam} = N_b q M. \tag{A.1}$$

As already defined beforehand, N_b are the particle per bunch, M is the number of bunch in the beam and q is the elementary charge of the proton ($1,602 \cdot 10^{-19} C$). It is then possible to run a power loss simulation by following these steps:

[1] Define a structure with a lossy material.

[2] Click on *Setup Solver*.

[3] Open the *Specials...* menu.

[4] Navigate into *Material*.

[5] Tick *Time power metal losses (1D)*.

[6] Run the simulation.

For simplicity steps 3, 4 and 5 are shown in Fig. (A.2).

FIGURE A.1: CST beam definition settings.
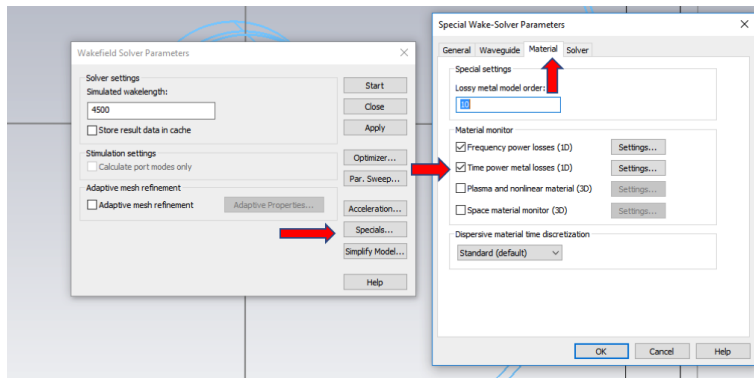
FIGURE A.2: Setting to run a power loss simulation with CST.

Bibliography

[1]   Michael Benedikt and Frank Zimmermann. "Proton colliders at the energy frontier". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 907 (2018), pp. 200–208.

[2]   JD Lawson. *The Physics of Charged-Particle Beams (The International Series of Monographs on Physics)*. 1988.

[3]   Carlo Zannini, G Rumolo, and G Iadarola. "Power loss calculation in separated and common beam chambers of the LHC". In: *5th International Particle Accelerator Conference, Dresden, Germany, 15 - 20 Jun 2014, pp.1711*. 2014.

[4]   Benoit Salvant et al. *Update on beam induced RF heating in the LHC*. Tech. rep. 2013.

[5]   B Salvant et al. *Beam induced RF heating*. Tech. rep. CERN, 2012.

[6]   Andrea Mostacci. "Beam-Wall interaction in the LHC liner". PhD thesis. Rome U., 2001.

[7]   Giovanni Rumolo et al. "Electron cloud effects on beam evolution in a circular accelerator". In: *Physical Review Special Topics-Accelerators and Beams* 6.8 (2003), p. 081002.

[8]   Arsenij Aleksandrovič Sokolov and Igor Michajlovic Ternov. "Synchrotron radiation". In: *Akademia Nauk SSSR, Moskovskoie Obshchestvo Ispytatelei prirody. Sektsia Fiziki. Sinkhrotron Radiation, Nauka Eds., Moscow, 1966 (Russian title: Sinkhrotronnoie izluchenie), 228 pp.* (1966).

[9]   Alexandre Samir Lasheen. "Beam Measurements of the Longitudinal impedance of the CERN Super Proton Synchrotron". PhD thesis. 2017.

[10]   *Workshop on Simulation of Power Dissipation and Heating from Wake Losses, Diamond Light Source, Oxfordshire, UK*. 30 Jan 2013.

[11]   T Kroyer, E Gaxiola, and Friedhelm Caspers. *Longitudinal and transverse wire measurements for the evaluation of impedance reduction measures on the MKE extraction kickers*. Tech. rep. CERN, 2007.

[12]   Dean A Frickey. "Conversions between S, Z, Y, H, ABCD, and T parameters which are valid for complex source and load impedances". In: *IEEE Transactions on microwave theory and techniques* 42.2 (1994), pp. 205–211.

[13]   G Bassi et al. "Analysis of coupled-bunch instabilities for the NSLS-II storage ring with a 500 MHz 7-cell PETRA-III cavity". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 810 (2016), pp. 151–163.

[14]   Elias Métral. "RF HEATING FROM WAKE LOSSES IN DIAGNOSTICS STRUCTURES". In: *2nd International Beam Instrumentation Conference, Oxford, UK, 16 - 19 Sep 2013, pp.THBL1*. 2013.

[15]   C Pellegrini. "Longitudinal Coupled Bunch Instability for Two Counterrotating Beams". In: *Report No. CERN-LEP-TH-86-17* (1986). URL: http://cds.cern.ch/record/170396.

[16]   J.M. Wang. "Transverse Two-Beam Instability". In: *Report No. CERN/LEP-TH/87-65* (1987). URL: https://cds.cern.ch/record/183705.

[17]   Stephen Myers. *The LEP Collider, from design to approval and commissioning*. John Adams' Lecture. Delivered at CERN, 26 Nov 1990. Geneva: CERN, 1991. DOI: 10.5170/CERN-1991-008. URL: http://cds.cern.ch/record/226776.

[18]   F. Zimmermann. "Two-Beam Resistive-Wall Wake Field". In: *Proceedings of PAC07*. 3. Albuquerque, New Mexico, USA, 2007, pp. 4237–4239. ISBN: 1424409179.

[19]   C Zannini and G Rumolo. *2-beam impedance: review of theory*. Tech. rep. 2019. URL: https://indico.cern.ch/event/809558/.

[20]   A Grudiev. "LHC Project Note 413 Simulation and Reduction Of Longitudinal And Transverse Impedances Of A Collimation Device With Two Beams In One Vacuum Chamber". In: *Report No. LHC-Project-note-413* (2008).

[21]   Lorena Vega et al. "Thermal analysis of the LHC injection kicker magnets". In: *Journal of Physics: Conference Series*. Vol. 874. IOP Publishing. 2017, p. 012100.

[22]   Andrea Saccomanno et al. "Long-term temperature monitoring in CMS using fiber optic sensors". In: *IEEE Sensors journal* 12.12 (2012), pp. 3392–3398.

[23]   C Darve et al. *VLHC beam-screen cooling*. 2001.

[24]   Giovanni Rumolo et al. *Electron cloud observation in the LHC*. Tech. rep. 2011.

[25]   Oswald Gröbner. *Dynamic outgassing*. Tech. rep. CERN, 1999.

[26]   JM Jimenez. "LHC: The world's largest vacuum systems being operated at CERN". In: *Vacuum* 84.1 (2009), pp. 2–7.

[27]   Carlo Zannini. "Electromagnetic Simulation of CERN Accelerator Components and Experimental Applications". PhD thesis. ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE, 2013. URL: https://infoscience.epfl.ch/record/187002/files/EPFL_TH5737.pdf.

[28]   Mauro Migliorati. *Wake Fields and Instabilities*. JUAS, 2016.

[29]   V.G.Vaccaro L.Palumbo and M.Zobov. *WAKE FIELDS AND IMPEDANCE*.

[30]   Sjoerd Rienstra. "Impedance models in time domain, including the extended Helmholtz resonator model". In: *12th AIAA/CEAS Aeroacoustics Conference (27th AIAA Aeroacoustics Conference)*. 2006, p. 2686.

[31]   M Bei et al. "The potential of an ultimate storage ring for future light sources". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 622.3 (2010), pp. 518–535.

[32]   Kheifets Semyon and Zotter Bruno. *Impedances and Wakes in High Energy Particle Accelerators*. World Scientific, 1998.

[33]   G.V.Stupakov. "Wake and Impedance". In: *SLAC-PUB-8683* (2000).

[34]   Francesco Giordano. *Impact of filling scheme on beam induced RF heating in CERN LHC and HL-LHC*. Tech. rep. CERN, 2017.

[35]   S Heifets and B Zotter. "On the symmetry of the impedance [particle accelerator]". In: *Proceedings of the 1999 Particle Accelerator Conference (Cat. No.99CH36366)*. Vol. 2. IEEE, 1999, pp. 1118–1119. ISBN: 0-7803-5573-3. DOI: `10.1109/PAC.1999.795467`. URL: `https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=795467`.

[36]   F Caspers and M Wendt. *RF Engineering JUAS 2019*. `https://indico.cern.ch/event/779575/`.

[37]   Thomas P Wangler. *RF Linear accelerators*. John Wiley & Sons, 2008.

[38]   S De Silva. *Fundamentals of RF Cavities*. Tech. rep. 2016. URL: `https://casa.jlab.org/publications/viewgraphs/USPAS2016/L_10_Fundam_RF_Cav.pdf`.

[39]   *CST Studio Suite*. `https://www.3ds.com/products-services/simulia/products/cst-studio-suite/`.

[40]   G Bregliozzi et al. "Vacuum pressure observations during 2011 proton run". In: (2012).

[41]   Aurélien Géron. *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. O'Reilly Media, 2019.

[42]   Giulia Papotti et al. *Longitudinal beam measurements at the LHC: the LHC beam quality monitor*. Tech. rep. 2011.

[43]   Scott Menard. *Applied logistic regression analysis*. Vol. 106. Sage, 2002.

[44]   Yihua Liao and V Rao Vemuri. "Use of k-nearest neighbor classifier for intrusion detection". In: *Computers & security* 21.5 (2002), pp. 439–448.

[45]   *Classification Series 5 – K-Nearest Neighbors (knn)*. `https://dslytics.wordpress.com/2017/11/16/classification-series-5-k-nearest-neighbors-knn/`.

[46]   Wei-Yin Loh. "Classification and regression trees". In: *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 1.1 (2011), pp. 14–23.

[47]   Yoav Freund and Llew Mason. "The alternating decision tree learning algorithm". In: *icml*. Vol. 99. 1999, pp. 124–133.

[48] Tin Kam Ho. "Random decision forests". In: *Proceedings of 3rd international conference on document analysis and recognition*. Vol. 1. IEEE. 1995, pp. 278–282.

[49] Steve Lawrence et al. "Face recognition: A convolutional neural-network approach". In: *IEEE transactions on neural networks* 8.1 (1997), pp. 98–113.

[50] Khaled Alsabti, Sanjay Ranka, and Vineet Singh. "An efficient k-means clustering algorithm". In: (1997).

[51] E Métral. *Overview of single-beam coherent instabilities in circular accelerators*. Tech. rep. 2005.

[52] T Argyropoulos et al. *Beam induced heating reduction by bunch flattening*. Tech. rep. 2014.

[53] M Furman, H Lee, and B Zotter. "Energy loss of bunched beams in SSC RF cavities". In: *Conf. Proc.* 1987, p. 1049.

[54] Lorenzo Teofili. *A Thermomechanical and Electromagnetic Approach For The Design Of High-Intensity Accelerator Components*. Tech. rep. 2019.

[55] J Barranco García et al. "Beam-beam studies for FCC-hh". In: *Proceedings of the IPAC* 17 (2017), pp. 14–19.

[56] Ian H Witten and Eibe Frank. "Data mining: practical machine learning tools and techniques with Java implementations". In: *Acm Sigmod Record* 31.1 (2002), pp. 76–77.

[57] Tapas Kanungo et al. "An efficient k-means clustering algorithm: Analysis and implementation". In: *IEEE transactions on pattern analysis and machine intelligence* 24.7 (2002), pp. 881–892.

[58] Lovisa Lovmar et al. "Silhouette scores for assessment of SNP genotype clusters". In: *BMC genomics* 6.1 (2005), p. 35.

[59] Christopher M Bishop et al. *Neural networks for pattern recognition*. Oxford university press, 1995.

[60] Naila Murray and Florent Perronnin. "Generalized max pooling". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2014, pp. 2473–2480.

[61]   David RK Brownrigg. "The weighted median filter". In: *Communications of the ACM* 27.8 (1984), pp. 807–818.

[62]   Ronald W Schafer. "What is a Savitzky-Golay filter?[lecture notes]". In: *IEEE Signal processing magazine* 28.4 (2011), pp. 111–117.

[63]   Steven M LaValle, Michael S Branicky, and Stephen R Lindemann. "On the relationship between classical grid search and probabilistic roadmaps". In: *The International Journal of Robotics Research* 23.7-8 (2004), pp. 673–692.

[64]   Mahesh Pal. "Random forest classifier for remote sensing classification". In: *International journal of remote sensing* 26.1 (2005), pp. 217–222.

[65]   Douglas M Hawkins. "The problem of overfitting". In: *Journal of chemical information and computer sciences* 44.1 (2004), pp. 1–12.

[66]   Florentina Bunea et al. "Honest variable selection in linear and logistic regression models via l1 and l1+ l2 penalization". In: *Electronic Journal of Statistics* 2 (2008), pp. 1153–1194.

[67]   Xiaojiao Mao, Chunhua Shen, and Yu-Bin Yang. "Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections". In: *Advances in neural information processing systems*. 2016, pp. 2802–2810.

[68]   Zhilu Zhang and Mert Sabuncu. "Generalized cross entropy loss for training deep neural networks with noisy labels". In: *Advances in neural information processing systems*. 2018, pp. 8778–8788.

[69]   James Bergstra and Yoshua Bengio. "Random search for hyper-parameter optimization". In: *Journal of machine learning research* 13.Feb (2012), pp. 281–305.

[70]   Michael W Browne. "Cross-validation methods". In: *Journal of mathematical psychology* 44.1 (2000), pp. 108–132.

[71]   F. Pedregosa et al. "Scikit-learn: Machine Learning in Python". In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830.

List of Figures

List of Tables