

# **Comparative Analysis of Simulations and Experimental Outcomes: Slow Extraction Driven by RF Transverse Excitation at the CERN Proton Synchrotron**

THOMAS EDWIN BASS

Department of Physics  
Royal Holloway, University of London



A thesis submitted to the University of London  
for the degree of Masters by Research

November 22, 2023

### **Declaration of Authorship**

I hereby declare that this thesis and the work presented in it are entirely my own. Where I have consulted the published work of others, this is always clearly attributed.

**Thomas Bass**

A handwritten signature in black ink, appearing to read "Thomas Bass".

**Date:** 13th September 2023

This project has received funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement Number 101004730.

## ABSTRACT

Radio-Frequency Knockout (RFKO) is a method of resonant slow beam extraction which is able to produce continuous particle spills over durations longer than can be achieved with fast single-turn or non-resonant multi-turn extraction. By using transverse excitation to gradually drive the circulating particles into resonance, spills can be produced over many thousands of turns, at a low intensity, and with a flat time profile. These spills can be delivered to facilities—such as fixed-target experiments or hadron therapy gantries—which require a continuous, low-intensity particle flux of durations significantly longer than the particle revolution period.

In order to accurately and efficiently simulate this extraction process over a wide range of timescales, new modelling tools and computing platforms must be explored. By utilising optimised computational hardware—such as General-Purpose Graphics Processing Units (GPGPUs), and next-generation simulation software (such as Xsuite), computation times for simulations can be reduced by several orders of magnitude while remaining accurate compared to empirical measurements.

This thesis presents the novel implementation and simulation of RFKO slow extraction at the East Area (EA) experimental facility, located at CERN’s Proton Synchrotron (PS). The results of the experiments outlined seek to confirm both the accuracy of the simulated frequency response and time structure, and the benefits of GPU-accelerated simulation software. These results will also provide an analysis of optimal parameter selections for both the extraction control system—optimising for spill quality—and the simulation system architecture—optimising for performance.

**Keywords** Beam Physics, Computational Physics, Slow Extraction, Radio-Frequency Knockout, Simulation, Xsuite

# **Contents**

## **List of Figures**

## **List of Tables**

# Chapter 1

## Introduction

The CERN Proton Synchrotron (PS), alongside providing intermediary acceleration for hadron and ion beams for the Super Proton Synchrotron (SPS) and eventually the Large Hadron Collider (LHC), provides beams for the many fixed-target experiments located at the East Area (EA) experimental facility. This facility, as illustrated in ??, consists of the proton (IRRAD) and mixed-field (CHARM) irradiation facilities, via a dedicated 24 GeV/c beamline, in addition to a multi-target beamline providing three secondary beams.

Ongoing experiments in these facilities, such as the CHARM High-energy Ions for Micro Electronics Reliability Assurance (CHIMERA) [?], use high-Z ions (Pb) to simulate the harsh radiation environment of cosmic rays that satellites experience. Using low-intensity ( $<10^6$  ions/spill), high-energy ( $>100$  MeV/u) beams, customers such as the European Space Agency can assess the reliability of new materials in these environments. ?? illustrates the two injection chains of the PS, either with Lead ions sourced from the Linear Accelerator 3 (Linac3) and Low Energy Ion Ring (LEIR) chain, or protons from the Linac4 and the Proton Synchrotron Booster (PSB).

### 1.1 Motivation

The CHIMERA experiment, and other experiments located at the East Area and its higher-energy sister site at the SPS North Area, are referred to as fixed-target experiments. These experiments require a semi-continuous flux of particles at lower intensities than those produced by the PS. For example, if the entire beam of the PS was delivered to a material sample in an effectively instantaneous “spill” (a controlled removal of particles from the circular accelerator), any sensitive acquisition equipment would become saturated, and the fixed target itself likely destroyed.

At a beam momentum of  $5.4 \text{ GeV } c^{-1}$  in the PS, the revolution frequency of the beam is around 470 kHz. Therefore, delivering particles over the period of one turn could, at most, provide a spill time on the order of  $2 \mu\text{s}$ . This extraction method, known as *fast* extraction, is commonly used to transfer particles between successive accelerators where the bunched time structure of the

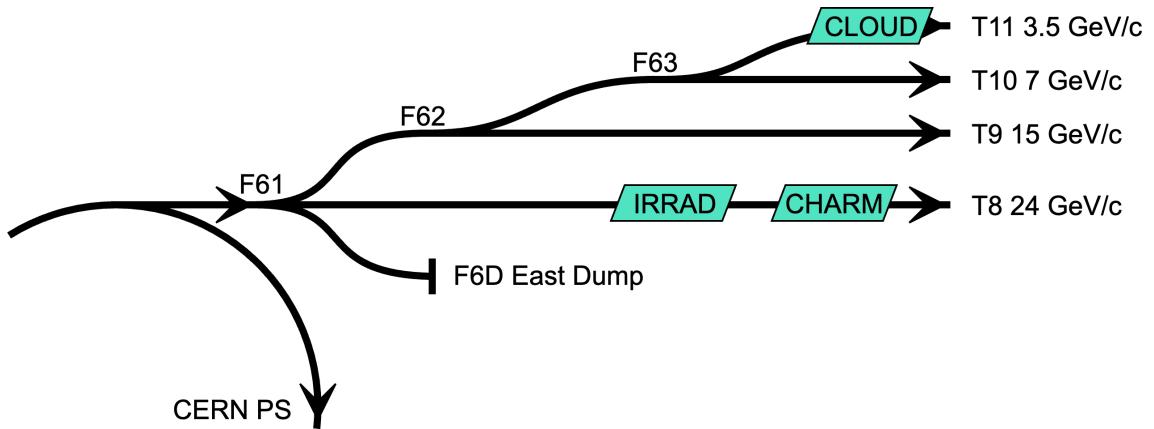


Figure 1.1: Layout diagram of the CERN East Area (EA) beamlines, their proton momenta, and experiments.

beam is preserved. To provide spills on the order of  $\sim 100$  ms required for these experiments, much longer than the revolution period, a “slow” extraction system is employed.

### 1.1.1 Extraction Methods

To move particles between consecutive machines along the primary CERN injector chain, the PSB, PS, and SPS are equipped with fast extraction systems. The principles of these processes will be fully explored in the next section, but a general overview and understanding of their benefits and disadvantages will help in understanding the motivation for the use of *slow* extraction.

As previously mentioned, *fast* extraction provides a spill of particles typically over the duration of one turn. A fast magnet, at a suitable location in the machine, imparts a large deflective angle to the beam. This ejects the beam from the nominal orbit, where it encounters a second magnet—called a septum—to deflect the beam into a transfer line. The opposite effect is then used to inject the bunch into the next machine: a magnet “kicks” the particles onto the circulating beam trajectory.

This method of ***single-turn fast extraction*** (??) is used at two locations in the SPS—Long Straight Sections (LSS) 4 and 6—to fill the two counter-rotating rings in the LHC [?]. The PSB uses the same method, but has the added complexity of “merging” four separate stacked rings to one transfer line [?]. LEIR likewise performs fast extraction [?]. The general layout of these systems is shown in ??, where red diamonds indicate fast extraction locations. Fast extraction systems are also used at the beam dump facilities at various stages of the accelerator chain to discard the circulating beam at the SPS and LHC.

In accelerator-to-accelerator contexts, the synchronous nature of fast extraction allows particles to remain bunched. ***Multi-turn fast extraction*** extends this benefit, and is performed when filling the SPS for its fixed-target experiments (North Area) from the PS. As the SPS (6.9 km)

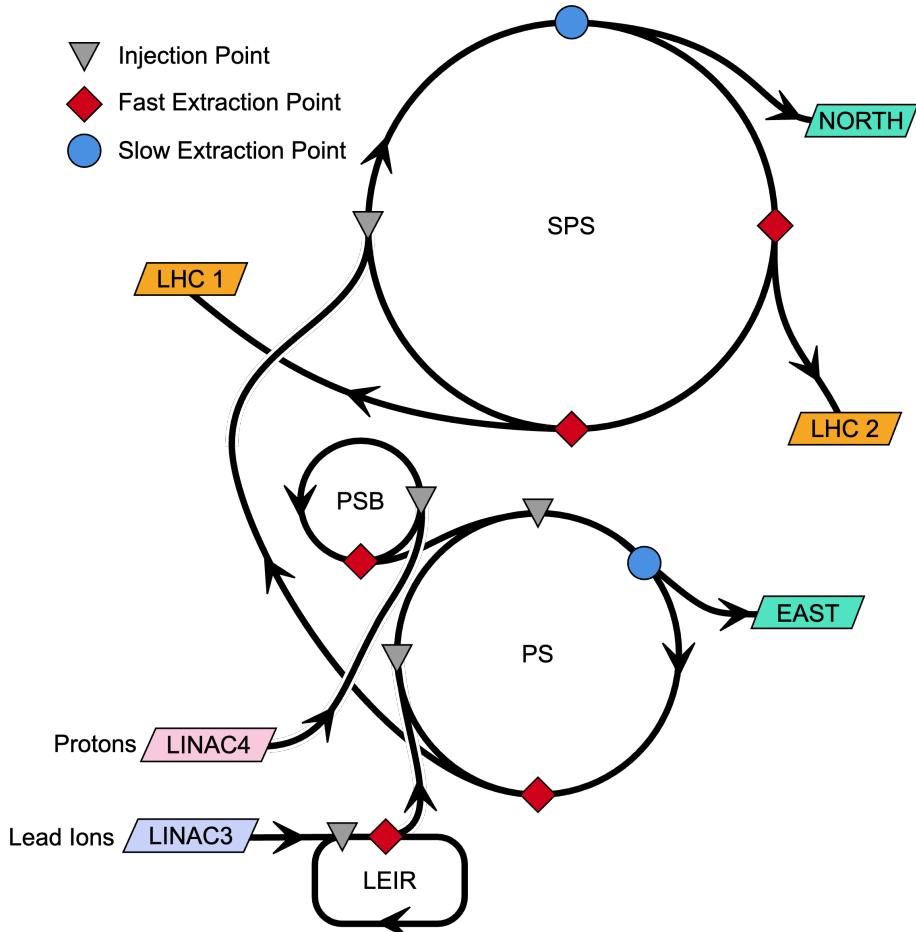


Figure 1.2: Layout diagram of the primary injection chain at CERN. Particles are sourced from either Linac 3 (for Lead ions) or Linac 4 (for Protons), and accelerated through the LEIR or PSB, respectively. After injection and acceleration in the PS, beams are extracted either to the SPS or East Area (EA). Particles at the SPS are either extracted at two LHC extraction locations, or towards the North area. Red diamonds indicate Fast extraction locations, and blue indicates Slow extraction regions. Grey triangles indicate injection regions.

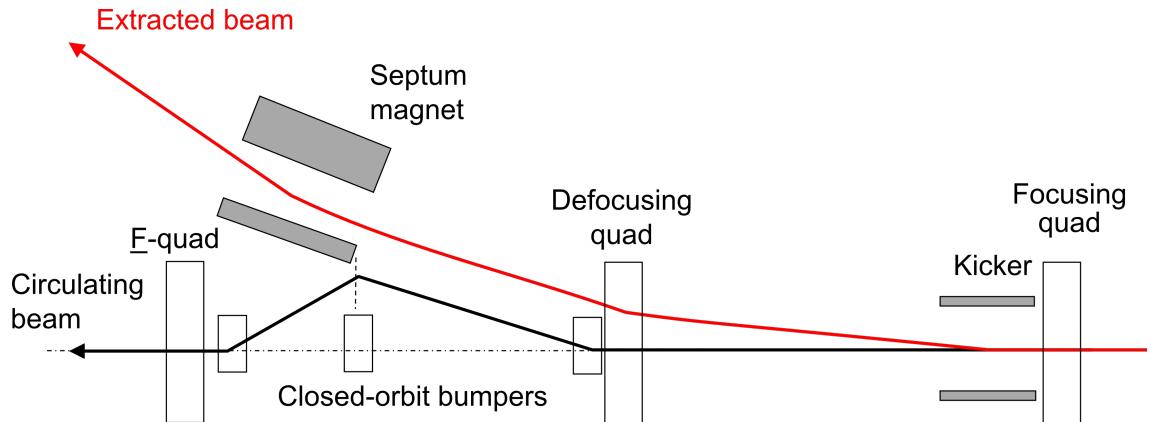


Figure 1.3: Schematic of the general principle of single-turn fast extraction [?].

is approximately 11 times longer in circumference than the PS<sup>1</sup> (628 m), the entire SPS can be filled by 11 successive extractions of the full PS. This method, known as *continuous transfer* (CT), introduces the concept of “shaving” the beam with an electrostatic septum [?]. Previously implemented in extraction from the PS, the beam orbit was bumped similarly to single-turn fast extraction, but was instead forced over the blade of an electrostatic septum (ES), such that a segment of the beam is ‘sliced’ off, kicked towards a magnetic septum (MS), and extracted. The remaining beam is then kicked back onto its orbit. After the next turn, the beam has been “rotated” in phase space by a quarter turn, and the process repeats, until finally the remaining core of the beam is bumped directly to the magnetic septum. This provides a spill to the SPS of five bunches over five beam rotations from a single PS bunch.

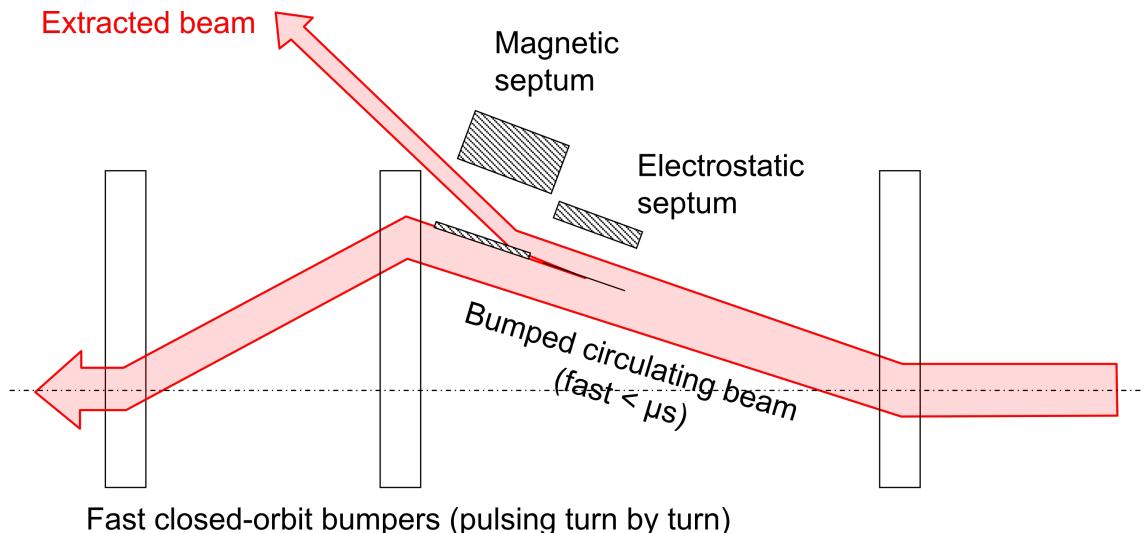


Figure 1.4: Diagram of a general multi-turn fast extraction system [?].

<sup>1</sup>The PS is 100 meters in radius, and therefore  $200\pi$ m

However, this method of increasing the spill time is limited by the width of the septum blades. As the amount of beam scraped off per turn decreases, the effect of the particles lost on the septum blade increases. To maintain a reasonable extraction efficiency, the beam size would have to become exponentially large. For example, if a 50% extraction efficiency is desired over  $10^4$  turns, the beam thickness sliced off by the Electrostatic Septum (ES) must be the same as the thickness of the blade itself. For the SPS, the ES has an effective blade thickness of  $500\text{ }\mu\text{m}$  [?]. The required diameter of the beam, therefore, would be  $2 \cdot 500\mu\text{m} \cdot 10^4 = 10\text{ m}$  [?]. This diameter would be larger than the beam pipes, and clearly physically impossible.

In order to provide spills over a longer duration, over many thousands of turns, *slow extraction* must be used.

### 1.1.2 Introduction to Slow Extraction

**Slow extraction** instead relies on exploiting the resonance of the beam as it oscillates around the orbit of the machine. A full theoretical description of these effects will be detailed in ??, but a brief explanation is offered here: the oscillatory motion of a beam can be controlled such that it returns to the same point along the accelerator's circumference with the same oscillation phase, in order to be in harmonic resonance with the periodic field distribution of the ring. Great care is usually taken to avoid this effect in particle accelerators. However, with the correct configuration, this amplitude growth can be used alongside an ES to separate particles from their orbit, towards an MS, and into the extraction transfer line. Should the effect be slow enough, a small fraction of particles will reach the MS per turn.

The exploitation of resonance to induce extraction is by no means novel; a method akin to slow extraction was first proposed by J. Tuck and L. Teng in 1951 [?], and performed by Le Couteur at the Liverpool cyclotron in 1954 [?]. Many such applications of slow extraction exist, using different effects to drive particles into resonance. This thesis will present some of these methods, particularly focusing on Radio-Frequency Knockout slow extraction.

Many medical synchrotrons, such as the MedAustron Synchrotron, use slow extraction to deliver beam to patients for ion beam cancer therapy [?], producing spills of up to  $30\text{ s}^2$ . In human medical applications, safety is paramount. Therefore, the slow extraction process must be fully understood. Ripples in the beam's intensity have to be carefully controlled in order to provide a steady, predictable irradiation dose to patients.

Small medical synchrotrons are simpler<sup>3</sup>, newer<sup>4</sup>, and have a far higher operational availability requirement than research accelerators such as the PS. At CERN's PS, however, that benefit cannot be relied upon. Machine time is in high demand, and studies are time-shared between machine

---

<sup>2</sup>Limited to 10 s for patient treatment

<sup>3</sup>The PS contains 100 bending magnets along its circumference, as opposed to MedAustron's 16

<sup>4</sup>The PS began operation in 1959, and is CERN's oldest accelerator still in operation. MedAustron was first certified in 2016.

development (MD), fixed-target experimental areas, LHC physics injection, and a wide array of other facilities.

Simulations can, therefore, provide an important tool in understanding slow extraction spills. New methods can be rapidly developed and tested without requiring extensive (and expensive) machine studies, and wide parameter spaces can be explored to construct functional relationships between input parameters and spill output characteristics.

## 1.2 Modelling Slow Extraction

?? will introduce a mathematical model that can be used to calculate particle trajectories in simple machines numerically. However, with even the simplest of accelerators at CERN containing hundreds of components, computational models quickly become necessary. In addition, slow extraction can operate over many more accelerator turns than other processes, such as fast single-turn injection. In the PS, a spill length of 0.5 s would require approximately 230,000 turns. For even a reasonably powerful machine, these computational simulations quickly become time-consuming.

To aid with the design and construction of CERN’s Large Electron-Positron collider (LEP, the predecessor to LHC), the *Methodical Accelerator Design Program* (MAD) [?] was developed. This FORTRAN-based scripting tool allowed physicists to define the arrangement and strengths of magnetic elements around an accelerator (the *lattice*), define beam parameters, and calculate the resulting motion of particles (the *optics*). MAD, and its latest version, MAD-X, quickly became the *de facto* standard within CERN’s beam physics community. However, in the twenty years since the release of MAD-X, the landscape of computational accelerator physics has become awash with a wide array of tools, codes, and standards, each with their own specialisations.

Newer tools are introducing improvements such as “ready-to-go” compatibility with General-Purpose Graphical Processing Unit (GPGPU) acceleration. GPUs are specialised hardware capable of performing highly optimised matrix and vector calculations. This makes them ideal for the matrix calculations required for particle accelerator simulations. Additionally, *embarrassingly parallel*<sup>5</sup> tasks can be run on a large number of smaller processing platforms, rather than a single large platform. These platforms could be processing units within a GPGPU, or a discrete computer host on a batch-processing network.

## 1.3 Outline and Aims

This thesis will present the novel implementation and simulation of a Slow Extraction method at the East Area experimental facility, with the aim to confirm both the accuracy of the simulated frequency response, and the benefits of GPU-accelerated simulation software. To fully understand the effects of the beam extraction method, the challenges in accurate simulation, and the benefits gained, this

---

<sup>5</sup>A problem which requires minimal to no effort to be separated into a large number of independent, concurrently-processed tasks.

thesis will outline the theory required to construct these simulations (??), the implementation of the chosen slow extraction method both in the real-world at the PS, and in simulation (??), and the analysis and results of the simulations (??). Finally, the conclusions of the work will be presented, and future work outlined (??).

## Chapter 2

# Accelerator Physics Theory

To construct accurate simulations of slow extraction methods, one must first understand the fundamental theory behind particle motion in a circular accelerator. This chapter will introduce the theory required to understand the motion of particles as they travel through an accelerator, and how such oscillatory motion can be exploited to perform slow extraction [?].

### 2.1 Transverse Beam Dynamics

#### 2.1.1 Coordinate System

The coordinate systems used to describe the dynamic motion of particles in an accelerator—such as the PS—rely on the periodic nature of a circular accelerator. Using a traditional Cartesian coordinate system would quickly introduce complications when describing circular trajectories, and so a local curvilinear coordinate system is used [?].

The basis vectors for this system,  $(\hat{x}, \hat{y}, \hat{s})$ , rotate in space as the reference particle moves along the accelerator. The longitudinal vector  $\hat{s}$  is tangent to the reference particle's trajectory, and the  $\hat{x}$  direction is parallel to the radius of the accelerator. This system is shown in Figure ??, along with the radius of the circular trajectory  $\rho$ , and the angle from the origin  $s = 0$ .

This coordinate system is known as a *Frenet-Serret* coordinate system, and can be rigorously defined by the unit vector *tangent* to the curve  $\hat{T}$  (pointing in the direction of motion), the unit vector *normal* to the tangent and on the tangential plane  $\hat{N}$ , and the binormal unit vector  $\hat{B}$ , the cross product of  $\hat{T}$  and  $\hat{N}$ . These unit vectors, forming an orthonormal basis spanning  $\mathbb{R}^3$  are equal to the vectors  $\hat{x}, \hat{y}, \hat{s}$  by:

$$\begin{pmatrix} \hat{T} \\ \hat{N} \\ \hat{B} \end{pmatrix} = \begin{pmatrix} \hat{s} \\ -\hat{x} \\ \hat{y} \end{pmatrix} \quad \text{Eq. 2.1}$$

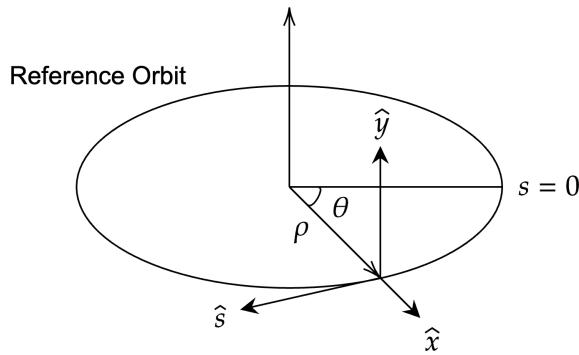


Figure 2.1: The vectors used to define the Frenet-Serret coordinate system.

The coordinate system  $(\hat{x}, \hat{y}, \hat{s})$  relies on the right-hand chirality (Figure ??) of the vector cross product for the  $\hat{x}$  vector to point radially outwards from the centre of the machine, the  $\hat{y}$  vector points upwards, and the  $\hat{s}$  vector points in the direction of motion.

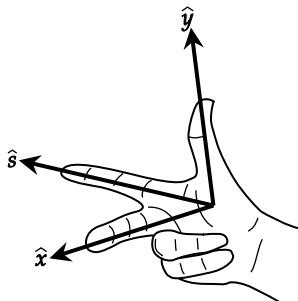


Figure 2.2: The right-hand chirality rule defining the direction of  $\hat{x} = \hat{y} \times \hat{s}$

### 2.1.2 Special Relativity

The equations of motion in this thesis consider particles moving at extremely high velocity, approaching the speed of light  $c$ . At CERN, before injection into the PS, particles are accelerated to 2 GeV, and reach 94.8% of the speed of light. At such speeds, relativistic effects are significant.

The Lorentz factor  $\gamma$  is defined as the derivative of *coordinate* time  $t$  in the relativistic particle's reference frame, with respect to the *proper* or *lab* time  $\tau$  in the observer's reference frame. It is calculated as:

$$\gamma_c = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}} \quad \text{Eq. 2.2}$$

Also defined is the factor  $\beta$

$$\beta_c = \frac{v}{c} \quad \text{Eq. 2.3}$$

and the factor  $\alpha$

$$\alpha_c = \frac{1}{\gamma} \quad \text{Eq. 2.4}$$

These factors will be referred to as  $\alpha_c, \beta_c, \gamma_c$  to distinguish them from other parameters using  $\alpha, \beta, \gamma$  (??).

One important distinction in a relativistic reference frame is the *invariant* mass  $m_0$  versus the relativistic mass  $m$ . The invariant mass, or rest mass of any object is the Newtonian mass measured in its own reference frame. From the *lab* reference frame, this mass is measured as the relativistic mass, defined as:

$$m = \gamma_c m_0 \quad \text{Eq. 2.5}$$

This relativistic mass is conserved just the same as rest mass, and follows the  $p = mv$  formula as in the rest frame. The rest energy  $E_0$  and mass  $m_0$  are likewise defined in the relativistic frame:

$$E = \gamma_c m_0 c^2 \quad \text{Eq. 2.6}$$

$$p = \gamma_c m_0 v \quad \text{Eq. 2.7}$$

For brevity, this report will always assume a relativistic frame, and all mentions of mass  $m$ , momentum  $p$ , or energy  $E$  should be assumed to be relativistic unless otherwise stated.

### 2.1.3 Particle Variables

Following the basis vectors defining the coordinate system, the following variables are used to characterise the orbit of particles, dependent on  $s$ :

#### In the transverse planes:

- $x(s)$  [m] — the particle's horizontal position with respect to the reference orbit.
- $y(s)$  [m] — the particle's vertical position with respect to the reference orbit.
- $x'(s) = \frac{dx}{ds}$  [rad] — the particle's horizontal angle with respect to the reference orbit.
- $y'(s) = \frac{dy}{ds}$  [rad] — the particle's vertical angle with respect to the reference orbit.

#### In the longitudinal plane:

- $z = s - s_0$  — the particle's longitudinal position with respect to the reference particle  $s_0$ .
- $\delta = \Delta p/p_0$  — the particle's momentum difference with respect to the reference particle  $p_0$ .

These variables provide 6 vectors fully describing the position of a particle. However, for the remainder of this thesis, the pair of longitudinal variables  $(z, \delta)$  will be treated separately to the transverse variable pairs  $(x, x')$  and  $(y, y')$ , enabling a simpler understanding of the effects of beam bending and focussing magnets.

### 2.1.4 Design Orbit and Dipoles

An understanding of particle dynamics in an accelerator can be built upon the Lorentz force, describing the force  $\vec{F}$  acting on a charged particle with charge  $q$  as it moves through an electric

field  $\vec{E}$  and magnetic field  $\vec{B}$  with velocity  $\vec{v}$ , given as

$$\vec{F}_{\text{Lorentz}} = q \cdot (\vec{E} + \vec{v} \times \vec{B}) \quad \text{Eq. 2.8}$$

This equation can already show why magnetic fields are typically used to bend particles along a circular trajectory and not electric fields; a particle with higher velocity  $v$  will experience a greater force from a constant magnetic field  $\vec{B}$ , whereas the contribution from the electric field  $\vec{E}$  will not benefit from the increasing velocity of a particle.

The other force experienced by a particle in an accelerator is a consequence of conservation of momentum:

$$\vec{F}_{\text{centrif.}} = \frac{mv^2}{\rho} \quad \text{Eq. 2.9}$$

where mass  $m = m_0\gamma_c$  is the particle's rest mass times the relativistic gamma factor, and  $\rho$  is the radius of the particle's orbit inside the accelerator.

Equating equations ?? and ?? yields

$$B\rho = \frac{p}{q} \quad \text{Eq. 2.10}$$

The quantity  $B\rho$  is known as the *magnetic rigidity* of a beam, and defines the bending radius of a particle for a given magnetic field.

If the relativistic momentum  $p = mv$  of the particle is measured in  $\text{GeV}/c$ , the magnetic rigidity can be expressed in Tesla-meters:

$$B\rho [\text{T m}] = 3.3356 \cdot p [\text{GeV } c^{-1}] \quad \text{Eq. 2.11}$$

Referring again to ?? (with the help of ??), a magnetic field pointed vertically upwards (along the  $\hat{y}$  vector) will cause a positively-charged particle with motion along the  $\hat{s}$  vector to experience a bending force along the  $-\hat{x}$  direction (illustrated in ??). This formulation, however, only considers a single dipole. In reality, the bending magnet structure of accelerators typically consists of multiple dipoles placed around the ring, their power converters connecting in series. The arrangements of these bending magnets—referred to as Main Units (MU) in the context of the PS—will define the reference<sup>1</sup> orbit in the accelerator.

If  $\alpha$  is the bending angle of a single MU, then  $\alpha$  can be related to the magnetic rigidity by

$$\alpha = \frac{ds}{\rho} = \frac{B}{B\rho} ds \quad \text{Eq. 2.12}$$

Integrating this over all magnets in the ring, therefore, must be equal to one full revolution,  $2\pi$ . This provides one of the first key insights which is not obvious at first—that “circular” accelerators are in fact more complex shapes, with bending magnets at each vertex, and a *drift* (accelerator sections without any significant electromagnetic field) or other non-bending multipole magnet at

---

<sup>1</sup>Sometimes referred to as *design* orbit

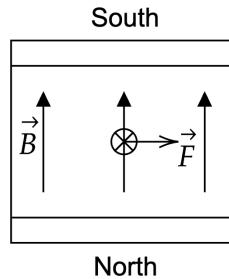


Figure 2.3: A “front-on” view of a dipole magnet, showing its magnetic field pointing down, and the Lorentz force experienced by a positively charged particle travelling into the page.

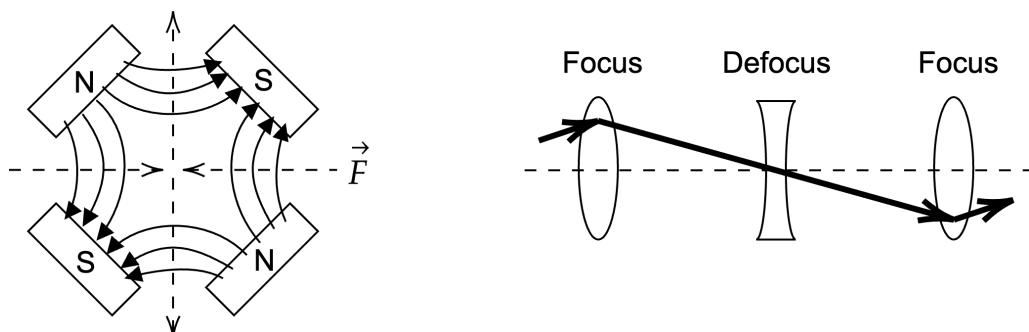


Figure 2.4: Diagram (left) showing the field of a *normal* oriented quadrupole, providing focussing in  $x$  but defocussing in  $y$ . The force experienced by a positively-charged particle moving out of the page plane is shown as dashed arrows. Right-hand diagram illustrates a FODO cell, where the structure of repeating focussing and defocussing quadrupoles can be related to the optical analogy of lenses.

each side: in the case of the LHC, 1,232 dipole magnets are used to keep the beam on its circular path; in the PS—100.

### 2.1.5 Focussing and Multipoles

While single-pair dipole magnets are used to bend the beam along the circular path, double-pair *quadrupole* magnets are used to bound the transverse motion, and to prevent the beam from spreading (increasing in size) due to the intrinsic spread of position and angle. An arrangement of four magnetic poles creates a field stronger at higher  $|x|$  displacement, causing particles to be pulled back towards the design orbit, “focussing” the beam in one transverse plane (while necessarily defocussing in the other). To correct for the defocussing in the other plane, a second quadrupole is placed further along in the beamline, with the magnetic poles arranged in the opposite orientation. This arrangement, known as a FODO lattice, is shown in ?? (right). ?? depicts a *sextupole* magnet, with three pole pairs, which is used to correct for chromaticity by exploiting dispersive effects.

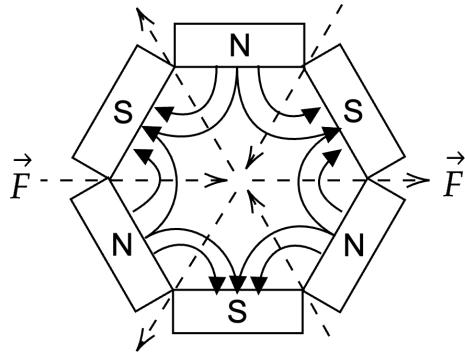


Figure 2.5: A diagram showing the field of a *normal* oriented sextupole. The force experienced by a positively-charged particle moving in to the page plane is shown as dashed arrows.

### 2.1.6 Single-Particle Oscillatory Motion

In a periodic (circular) accelerator, the linear equation for transverse motion takes the form of a differential equation:

$$x'' - x \cdot \left( k - \frac{1}{\rho^2} \right) = 0 \quad \text{Eq. 2.13}$$

where  $k$  is the normalised quadrupole coefficient, defined as

$$k = \frac{q}{p} \cdot \frac{dB_y}{dx} \quad \text{Eq. 2.14}$$

The periodicity of circular accelerators requires that  $k$  is also periodic, and the same gradient is seen after one turn. The  $1/\rho^2$  term represents the effect of the dipole bending field, providing a weak focussing effect. Suppose we enforce the periodicity<sup>2</sup> of  $k$  in  $s$ . In that case, ?? can be rewritten in a compact form

$$x''(s) - \tilde{k}(s)x(s) = 0 \quad \text{Eq. 2.15}$$

where the quadrupole focussing component  $\tilde{k}(s)$  is now a function of  $s$ . With the given periodic boundary condition  $\tilde{k}(s + L) = \tilde{k}(s)$ , ?? is known as Hill's equation. Following Floquet's theorem [?], the general solution to this equation takes the form of

$$x(s) = \sqrt{\varepsilon} \sqrt{\beta(s)} \cos(\varphi(s) - \varphi_0) \quad \text{Eq. 2.16}$$

where  $\varphi(s)$  is the oscillation phase,  $\varphi_0$  is the initial condition, and  $\varepsilon$  is an integration constant.  $\varphi_0$  and  $\varepsilon$  are constants of a trajectory, and so the transverse oscillation of a particle is given by the special  $\beta$  function. The  $\beta$  function must also satisfy the periodic constraint  $\beta(s + L) = \beta(s)$ , and is related to the  $\varphi$  phase function by the following equation:

$$\varphi(s) = \int_0^s \frac{ds'}{\beta(s')} \quad \text{Eq. 2.17}$$

---

<sup>2</sup>i.e. that the quadrupoles do not change between consecutive revolutions

Taking the derivative of Eq. ?? with respect to  $s$  gives an equation for the angle of the trajectory:

$$x'(s) = \frac{-\sqrt{\varepsilon}}{\sqrt{\beta(s)}} \sin(\varphi(s) - \varphi_0) + \alpha(s) \cos(\varphi(s) - \varphi_0) \quad \text{Eq. 2.18}$$

The overall phase advance of transverse oscillations over one revolution of  $2\pi$  is known as the *betatron tune* of the accelerator, and is defined as

$$Q = \frac{1}{2\pi} \oint \frac{ds}{\beta(s)} \quad \text{Eq. 2.19}$$

This quantity, often simply referred to as *tune*, effectively indicates how many complete transverse oscillations a particle makes in one turn. The fractional part of this value is the most significant, and the integer part of a referenced tune may sometimes be omitted.

Particles with momentum different to the reference particle  $p_0$  will experience different focussing effects in quadrupoles. This slight change in focussing will affect the oscillatory motion of off-momentum particles, and result in a spread of tunes throughout a distribution of particles. The variation of this tune, with respect to the variation in momentum, is characterised by the *normalised chromaticity* (where it is divided by  $Q$ ), variously defined as  $Q'$  or  $\eta$ :

$$Q' = \frac{\Delta Q/Q}{\Delta p/p} \quad \text{Eq. 2.20}$$

### 2.1.7 Phase Space

?? and ?? introduced an integration constant  $\varepsilon$ , which is a characteristic parameter of the particle or ensemble of particles. It is known as (transverse) emittance, and is usually considered with particle ensembles in mind. These two equations can be manipulated to provide an expression for  $\varepsilon$ :

$$\varepsilon = \gamma(s)x^2(s) + 2\alpha x(s)x'(s) + \beta(s)x'^2(s) \quad \text{Eq. 2.21}$$

where

$$\alpha(s) = -\frac{1}{2} \frac{d\beta(s)}{ds} \quad \text{Eq. 2.22}$$

and

$$\gamma(s) = \frac{1 + \alpha^2(s)}{\beta(s)} \quad \text{Eq. 2.23}$$

These two values, along with  $\beta$ , form the *Courant-Snyder Parameters* [?], commonly referred to as the *Twiss parameters*<sup>3</sup>.

These parameters can be readily understood in the *phase space* of  $(x, x')$ . Illustrated in ??, ?? describes an ellipse, which covers the area of the particle ensemble in phase space, the area of which is given by  $\pi \cdot \varepsilon$ .

---

<sup>3</sup>Even Richard Q. Twiss, a British astronomer, was himself unsure as to how the parameters became attributed to him [?].

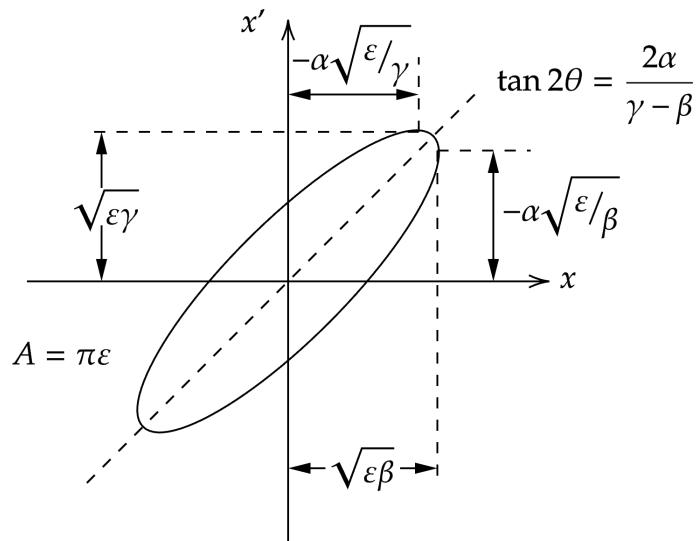


Figure 2.6: A phase-space ellipse in  $(x, x')$  and its relation to the Twiss parameters.

Liouville's theorem states that, under conservative forces and at fixed energy, this emittance is conserved<sup>4</sup>, and no external field can change it. Hence, ?? provides it as an invariant of motion.

### 2.1.8 Dispersion

We now consider the oscillatory motion of *multiple* multi-energetic beams, with reference to the ideal particle. This ideal particle, with reference momentum  $p_0$  and initial conditions all equating to zero, has a path through the field centre of all quadrupoles, and so experiences no quadrupole effect  $k$ . This means that its motion is controlled by the weak dipole focussing  $1/\rho^2$ . A nominal particle, still with momentum  $p_0$  but with initial displacement and/or angle conditions, will perform closed-loop betatron oscillations around this path.

A particle with a momentum offset,  $\Delta p = p - p_0$ , will still satisfy ?? from before, but in the inhomogeneous form

$$x''(s) - x \left( k - \frac{1}{\rho^2} \right) = \frac{1}{\rho} \frac{\Delta p}{p_0} \quad \text{Eq. 2.24}$$

We can use the homogeneous solution previously found to form the complete solution  $x(s) = x_h(s) + x_i(s)$ . This inhomogeneous term  $x_i(s)$ , is normalised by the momentum difference to produce the Dispersion function

$$D(s) = \frac{x_i(s)}{\Delta p/p_0} \quad \text{Eq. 2.25}$$

---

<sup>4</sup>Normalised emittance is conserved during acceleration where energy is not fixed

describing how the additional amplitude of an off-momentum particle depends on the momentum difference. This dispersion function satisfies, for the circumference of the machine:

$$D'' - D \left( x - \frac{1}{\rho^2} \right) = \frac{1}{\rho(s)} \quad \text{Eq. 2.26}$$

### 2.1.9 Matrix Formalism

?? and ?? can be solved by imposing initial conditions  $x_0$  and  $x'_0$  instead of periodic boundary conditions:

$$x(s) = x_0 \cos(\sqrt{K} \cdot s) + \frac{x'_0}{\sqrt{K}} \sin(\sqrt{K} \cdot s) \quad \text{Eq. 2.27}$$

$$x'(s) = -x_0 \sqrt{K} \sin(\sqrt{K} \cdot s) + x'_0 \cos(\sqrt{K} \cdot s) \quad \text{Eq. 2.28}$$

where  $K$  now combines the quadrupole focussing effect  $k$  (?) and the dipole weak focussing effect:

$$K = \frac{1}{\rho^2} - k \quad \text{Eq. 2.29}$$

These two equations can be written as a matrix equation:

$$\begin{pmatrix} x \\ x' \end{pmatrix} = \mathbf{M}_{\text{foc}} \begin{pmatrix} x_0 \\ x'_0 \end{pmatrix} \quad \text{Eq. 2.30}$$

where  $\mathbf{M}_{\text{foc}}$  is the focussing quadrupole transfer matrix [?]

$$\mathbf{M}_{\text{foc}} = \begin{pmatrix} \cos(\sqrt{K} \cdot s) & \frac{1}{\sqrt{K}} \sin(\sqrt{K} \cdot s) \\ -\sqrt{K} \sin(\sqrt{K} \cdot s) & \cos(\sqrt{K} \cdot s) \end{pmatrix} \quad \text{Eq. 2.31}$$

The transfer matrix for a defocussing quadrupole can also be defined as [?]:

$$\mathbf{M}_{\text{defoc}} = \begin{pmatrix} \cosh(\sqrt{K} \cdot s) & \frac{1}{\sqrt{K}} \sinh(\sqrt{K} \cdot s) \\ \sqrt{K} \sinh(\sqrt{K} \cdot s) & \cosh(\sqrt{K} \cdot s) \end{pmatrix} \quad \text{Eq. 2.32}$$

The definition of this transfer matrix  $\mathbf{M}_{\text{foc}}$  can be expanded to define the variety of magnets used in an accelerator. In the simple case, where no magnet is present, we define the transfer matrix for a drift space of length  $L$ :

$$\mathbf{M}_{\text{drift,x}} = \mathbf{M}_{\text{drift,y}} = \begin{pmatrix} 1 & L \\ 0 & 1 \end{pmatrix} \quad \text{Eq. 2.33}$$

For a normal dipole magnet bending only horizontally, where  $k = 0$  and  $K_x = 1/\rho^2$ , the transfer matrix is:

$$\mathbf{M}_{\text{dipole},x} = \begin{pmatrix} \cos\left(\frac{L}{\rho}\right) & \rho \sin\left(\frac{L}{\rho}\right) \\ -\frac{\sin\left(\frac{L}{\rho}\right)}{\rho} & \cos\left(\frac{L}{\rho}\right) \end{pmatrix} \quad \text{Eq. 2.34}$$

$$\mathbf{M}_{\text{dipole},y} = \mathbf{M}_{\text{drift},y} \quad \text{Eq. 2.35}$$

Further matrix equations can be defined for other magnets. If we define the properties of each element in the accelerator (the lattice), we can then define the transfer matrix for the entire accelerator as the product of the transfer matrices of each element. For example, a FODO lattice with an initial bending dipole can be defined as:

$$\mathbf{M}_{\text{total}} = \mathbf{M}_{\text{foc}} \cdot \mathbf{M}_{\text{drift}} \cdot \mathbf{M}_{\text{defoc}} \cdot \mathbf{M}_{\text{drift}} \cdots \quad \text{Eq. 2.36}$$

This is the principle by which *particle tracking codes*, such as MAD-X, operate. An input lattice file defines the structure and properties of the magnets composing the accelerator, and particle variables are then computed with matrix multiplication.

Expanding on the idea of matrix formalism and the phase-space ellipse, many particle tracking codes operate in *normalised* phase space  $(x_n, x'_n)$ —that is, where the phase space of  $(x, x')$  is linearly transformed in a way which results in the phase-space ellipse becoming circular. The matrix transformation is therefore defined as

$$\begin{pmatrix} x_n \\ x'_n \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{\beta_x}} & 0 \\ \frac{\alpha_x}{\sqrt{\beta_x}} & \sqrt{\beta_x} \end{pmatrix} \begin{pmatrix} x \\ x' \end{pmatrix} \quad \text{Eq. 2.37}$$

## 2.2 Longitudinal Dynamics

Where the transverse planes principally consider bending and focussing effects from magnetic fields, the longitudinal plane considers acceleration effects from RF fields. Returning again to ??, we instead now consider the *electric* component:

$$\vec{F} = q\vec{E} = \frac{dp}{dt} \quad \text{Eq. 2.38}$$

An electric field of magnitude  $E_0$  travelling in the  $s$  direction can be defined as

$$E_s = E_0 \cos(\omega_{RF}t) \quad \text{Eq. 2.39}$$

where  $\omega_{RF}$  is the RF frequency. Therefore, for a synchronous particle with synchronous phase  $\phi_s$ , the rate of energy gain is

$$\frac{dE_s}{ds} = \frac{dp_s}{dt} = eE_0 \sin \phi_s \quad \text{Eq. 2.40}$$

In a synchrotron, where the orbit of particles is of constant radius, the magnetic field  $B$  needs to increase in time (*ramp*) as the particle is accelerated. Additionally, to keep the particle in phase with the RF cavity, the frequency of the RF  $\omega_{RF}$  must follow the particle's energy gain. This synchronous condition, for velocities  $v \ll c$ , is defined as [?]

$$\frac{2\pi\rho}{v_s} = \frac{h}{f_{RF}} \quad \text{Eq. 2.41}$$

where  $v_s$  is the synchronous particle velocity and  $h$  is the RF harmonic, the number of RF cycles per revolution. For East Area beams in the PS,  $h = 8$  acceleration is used. This means that there are eight stable bunches, spread equidistantly around the ring circumference.

To keep a particle with increasing energy bent along a path of constant radius, the magnetic field  $B$  of the bending dipoles must increase as RF acceleration occurs. Recalling ??, and deriving by time, obtains

$$\frac{dp}{dt} = q\rho\dot{B} \quad \text{Eq. 2.42}$$

For one turn in a synchrotron, this results in a momentum gain

$$\Delta p_{rev} = q\rho\dot{B}T_{rev} = \frac{Lq\rho\dot{B}}{v} \quad \text{Eq. 2.43}$$

where  $L$  is the machine length. The energy gain  $\Delta E = v\Delta p$  per revolution is therefore

$$\Delta E_{rev} = Lq\rho\dot{B} = qV \sin \phi_s \quad \text{Eq. 2.44}$$

where  $\phi_s$  is the synchronous phase. This stable phase therefore changes while energy is ramped:

$$\phi_s = \arcsin \left( L\rho \frac{\dot{B}}{V_{RF}} \right) \quad \text{Eq. 2.45}$$

### 2.2.1 Dispersion

Next, akin to ??, we consider small perturbations in momentum with respect to an ideal path. We define two parameters:

$$\alpha_p = \frac{\Delta L/L}{\Delta p/p} \quad \text{Eq. 2.46}$$

where  $\alpha_p$ , the *momentum compaction factor*, is distinct from the relativistic  $\alpha_c$  and Twiss parameter  $\alpha$ .

$$\eta = \frac{\Delta f_{rev}/f_{rev}}{\Delta p/p} \quad \text{Eq. 2.47}$$

where  $\eta$ , the *slip factor*, is the relative change of revolution frequency with momentum.

Using the dispersion function ??, we can relate the momentum compaction factor to its effect on orbit length  $s_0$  [?, 5.1]

$$\alpha_p = \frac{1}{L} \int \frac{D_x(s)}{\rho(s)} ds_0 \quad \text{Eq. 2.48}$$

Using the revolution frequency<sup>5</sup>  $f_{rev} = \beta_c c / 2\pi R$ , we define the relative revolution frequency change  $\Delta f_{rev}$  as

$$\Delta f_{rev} = \frac{d\beta_c}{\beta_c} - \frac{dR}{R} = \frac{d\beta_c}{\beta_c} - \alpha_p \frac{dp}{p} \quad \text{Eq. 2.49}$$

This defines a *transition energy*  $\gamma_{tr}$  such that  $\eta = 0$ , causing no change in revolution frequency  $f_{rev}$  for a small momentum change  $\Delta p$ :

$$\gamma_{tr} = \frac{1}{\sqrt{\alpha_p}} \quad \text{Eq. 2.50}$$

Above this transition energy, where  $\eta < 0$ , the particle has velocity  $v \rightarrow c$ , and so the path-length change dominates the velocity change. This is the standard configuration of accelerators. Below this transition energy, where  $\eta > 0$ , the increase of velocity is significantly larger than the path-length change effect. At this transition energy, both effects compensate, and so  $f_{rev}$  becomes independent of  $\Delta p$ . Here, the *synchrotron* longitudinal oscillations stop, and particles not on the synchronous phase  $\phi_s$  will rapidly lose or gain energy, and become lost due to dispersion  $D$ .

### 2.2.2 Longitudinal Phase Space

These effects can be analysed by considering phase-space coordinates—this time in the longitudinal plane. This space is constructed with variables  $(\phi, \Delta p/p)$ , and shown in ???. A full derivation for the longitudinal motion invariants can be found in [?], where the synchronous phase  $\phi_s$  defines the orbit with zero longitudinal (synchrotron) oscillation. Outside the stable envelope, demarcated by the *separatrix*, particles are lost to instability. The region within the separatrix is known as the *RF bucket*, and the area occupied by stable particles is the *RF emittance*.

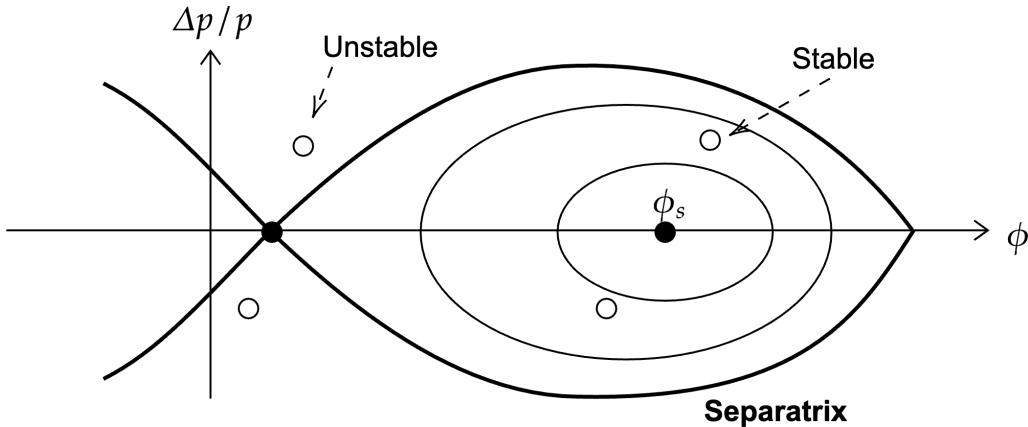


Figure 2.7: Longitudinal phase space, showing particles at the synchronous phase  $\phi_s$ , inside stability, and in unstable motion. The separatrix, bold, separates these two regions.

<sup>5</sup>NB:  $\beta_c$  here refers to the relativistic Lorentz factor  $v/c$

## 2.3 Slow Extraction

Now that the fundamentals of accelerator beam physics are established, this section will introduce the principles behind *slow extraction* (SX). Various methods of performing SX will be covered, but the particular method covered in this thesis—*Radio-Frequency Knockout* (RFKO)—will be described in detail.

### 2.3.1 Sextupoles and $^{1/3}$ -Integer Resonance

?? introduced the concept of dispersion in reference to a multi-energetic beam. These effects are controlled with the use of sextupole magnets. The transverse field of a sextupole magnet has the form [?, Ch 2.4]

$$B_x = -6B_3 xy \quad \text{Eq. 2.51}$$

$$B_y = -3B_3(x^2 - y^2) \quad \text{Eq. 2.52}$$

where

$$B_3 = -\frac{1}{6} \left( \frac{d^2 B_y}{dx^2} \right)_0 \quad \text{Eq. 2.53}$$

In *normalised* coordinates (??), these effect of these fields (in a thin-lens approximation) becomes

$$\begin{aligned} \Delta x_n &= 0 & \Delta x'_n &= S(x^2 - \frac{\beta_y}{\beta_x}y^2) \\ \Delta y_n &= 0 & \Delta y'_n &= -2S\frac{\beta_y}{\beta_x}xy \end{aligned}$$

where  $S$ , the normalised sextupole strength, is

$$S = \frac{1}{2}\beta_x^{2/3}l_s k' \quad \text{Eq. 2.54}$$

where  $l_s$  is the sextupole equivalent length, and  $k'$  is the normalised sextupole field:

$$k' = -6\frac{B_3}{B\rho} \quad \text{Eq. 2.55}$$

For extraction driven in the horizontal transverse plane, the dynamics of the vertical transverse plane are neglected as these terms remain insignificant to the first order [?, Ch 2.2]. Recalling the matrix formalism established in ??, we use a general transfer matrix for  $n$  turns with a betatron tune  $Q$ :

$$\mathbf{M}_n = \begin{pmatrix} \cos(2\pi n Q_x) & \sin(2\pi n Q_x) \\ -\sin(2\pi n Q_x) & \cos(2\pi n Q_x) \end{pmatrix} \quad \text{Eq. 2.56}$$

The tune of a particle is then defined as a third-integer, i.e.

$$Q_x = i \pm 1/3 + \delta Q, \quad i \in \mathbb{Z}, |\delta Q| \ll 1/3 \quad \text{Eq. 2.57}$$

The tune distance,  $\delta Q$ , is therefore the distance of the particle from the resonance. The  $n$ -turn matrix can then be approximated for *modulo-3* turns [?, Ch 2.4]:

$$\begin{pmatrix} x \\ x' \end{pmatrix}_1 = \begin{pmatrix} -1/2 & \pm\sqrt{3}/2 \\ \mp\sqrt{3}/2 & -1/2 \end{pmatrix} \begin{pmatrix} x \\ x' \end{pmatrix}_0 \quad \text{Eq. 2.58}$$

$$\begin{pmatrix} x \\ x' \end{pmatrix}_2 = \begin{pmatrix} -1/2 & \mp\sqrt{3}/2 \\ \pm\sqrt{3}/2 & -1/2 \end{pmatrix} \begin{pmatrix} x \\ x' \end{pmatrix}_0 \quad \text{Eq. 2.59}$$

$$\begin{pmatrix} x \\ x' \end{pmatrix}_3 = \begin{pmatrix} 1 & 6\pi \cdot \delta Q \\ -6\pi \cdot \delta Q & 1 \end{pmatrix} \begin{pmatrix} x \\ x' \end{pmatrix}_0 \quad \text{Eq. 2.60}$$

Therefore, a particle with  $\delta Q = 0$ , i.e. exactly on resonance, will return to its original position after three turns. For such a particle, the change in position  $x$  and divergence  $x'$  from initial conditions  $(x_0, x'_0)$  after three turns is therefore calculated as

$$\Delta x_3 = \varepsilon_Q x'_0 + \frac{3}{2} S x_0 x'_0 \quad \text{Eq. 2.61}$$

$$\Delta x'_3 = -\varepsilon_Q x_0 + \frac{3}{4} S (x_0^2 - x'_0^2) \quad \text{Eq. 2.62}$$

where  $\varepsilon_Q$ , the *modified tune distance*, is defined as  $\varepsilon_Q = 6\pi \cdot \delta Q$ , and  $S$  is the normalised sextupole strength (??).

The values of  $\Delta x_3$  and  $\Delta x'_3$  are known as the *spiral step* and *spiral kick*. The *Hamiltonian* is now used to analyse this motion in phase-space. The Hamiltonian  $\mathbf{H}$  is a time-independent constant of motion, the contours of which describe the *phase-space map*, the motion of a particle in normalised phase-space. Assuming the 3-turn time is small enough to be continuous, the *Kobayashi Hamiltonian* [?] is defined as

$$\mathbf{H} = \frac{\varepsilon_Q}{2} (x^2 + x'^2) + \frac{S}{4} (3x x'^2 - x^3) \quad \text{Eq. 2.63}$$

The first term describes the typical circular phase-space motion familiar from ??, with a sextupole strength  $S = 0$  (i.e. a linear machine). The second term describes the sextupole effect, transforming the elliptical map into a triangular trajectory for increasing values of  $x$  and  $x'$ . The phase-space map is shown in ?. A boundary triangle is easily identified from the map, which defines the trajectory of the stable particle with the greatest possible amplitude, before the motion becomes unstable. Outside this stable region, the particle trajectories do not close on themselves, and the particle will rapidly gain amplitude. The region bounded by this *separatrix* is known as the *acceptance*, the size of which is defined by  $|\varepsilon_Q|/S|$ , and can be found by factorising  $\mathbf{H}$  at a value

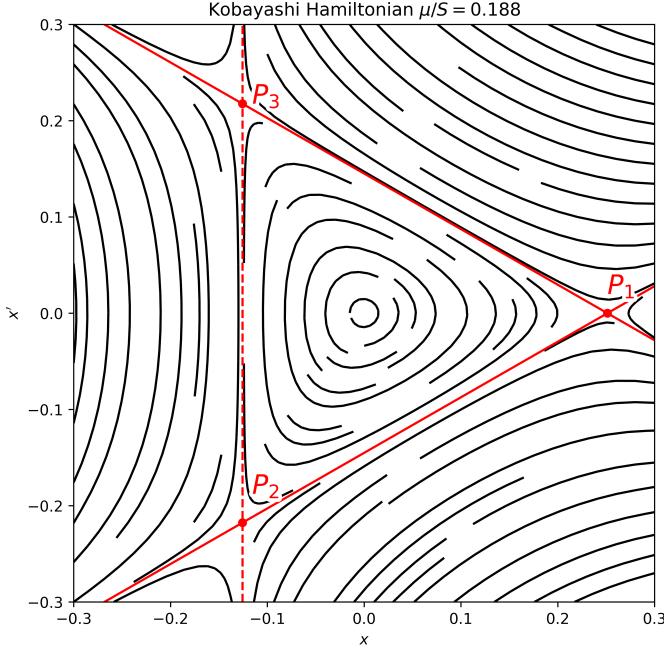


Figure 2.8: Normalised phase-space map of the Kobayashi Hamiltonian, for a ratio  $\varepsilon_Q/S = 0.188$ . The points  $P_1, P_2, P_3$  are shown, with connecting lines.

of  $(2\varepsilon_Q/3)^3/S^2$ , which factorises to three points:

$$P_1 = \left( \frac{4\varepsilon_Q}{3S}, 0 \right) \quad \text{Eq. 2.64}$$

$$P_2 = \left( -\frac{2\varepsilon_Q}{3S}, -\frac{2\varepsilon_Q}{\sqrt{3}S} \right) \quad \text{Eq. 2.65}$$

$$P_3 = \left( -\frac{2\varepsilon_Q}{3S}, \frac{2\varepsilon_Q}{\sqrt{3}S} \right) \quad \text{Eq. 2.66}$$

The  $x$  location of the stable particle,  $h$  can also be defined as

$$h = \frac{2\varepsilon_Q}{3S} = \frac{4\pi}{S} \delta Q \quad \text{Eq. 2.67}$$

Therefore, the lines connecting  $P_1, P_2$  and  $P_1, P_3$  (solid lines in ??), along with the vertical line at  $-h$  (dashed line in ??), define the separatrix. The acceptance, the area bounded by the separatrices, is defined as  $A = 3\sqrt{3}h^2$ .

### 2.3.2 Slow Extraction Principles

The fundamental principle of slow extraction exploits the unstable boundary created by sextupolar fields. After a beam is injected and stable, this *waiting* beam is accelerated towards the  $1/3$ -integer resonance. A particle at resonance will then “lock on” to one of the separatrix arms, and its

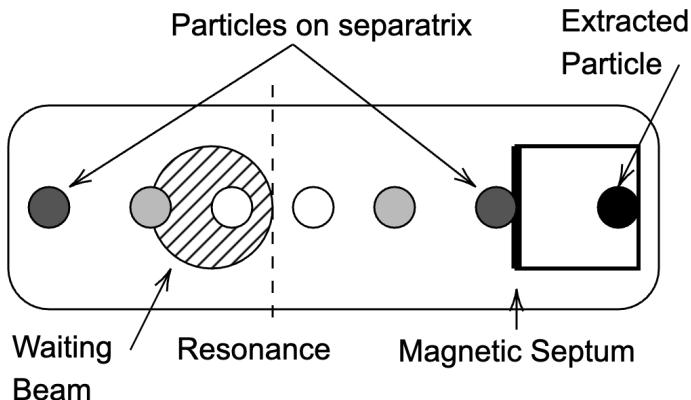


Figure 2.9: A real-space diagram of momentum-driven slow extraction, looking along the  $s$  direction at the magnetic septum location. The waiting beam approaches resonance (dashed), and some particles begin large turn-by-turn oscillations (increasing shaded gradient with turn number) before they jump the septum blade (thick black line) and will be extracted.

amplitude will grow in steps of three turns, rotating around the three separatrices. By carefully controlling the rate at which particles enter resonance, the 3-turn amplitude growth (spiral step) can be adjusted to create a desired flux of particles entering resonance. After a number of turns of amplitude growth (typically hundreds), an electrostatic septum located at some distance from the waiting beam will provide a kick necessary to extract the particles. A real-world diagram of slow extraction is shown in ??.

There are a variety of methods to provide acceleration of the beam towards resonance, which fall into two categories. *Momentum-driven* methods rely on accelerating the beam towards a stationary resonance by controlling the momentum  $\Delta p/p$ . *Amplitude-driven* methods, however, use transverse noise or RF excitation at a third-integer ( $i \pm \frac{1}{3} \cdot f_{rev}$ ) of the beam's revolution frequency to create increasing betatron amplitudes.

Other methods involve moving the resonance itself, by changing the tune of the machine. This can be accomplished by using dipoles, quadrupoles, or sextupoles, but is susceptible to complications arising from the changing transverse optics. Implementations would require active compensation to maintain the beam's position, size, and, ultimately, stability.

These various categories are not mutually exclusive, and some of the methods discussed will use a combination of effects to achieve slow extraction.

### 2.3.3 Steinbach Diagram

The two beam-moving methods of momentum and amplitude-driven slow extraction can be compared using a Steinbach<sup>6</sup> diagram. This plots the beam amplitude  $A_x$  as a function of momentum  $\Delta p/p_0$ . The sextupole instability is shown as a triangle, within which is the unstable region beyond

<sup>6</sup>Attributed to C. Steinbach, CERN

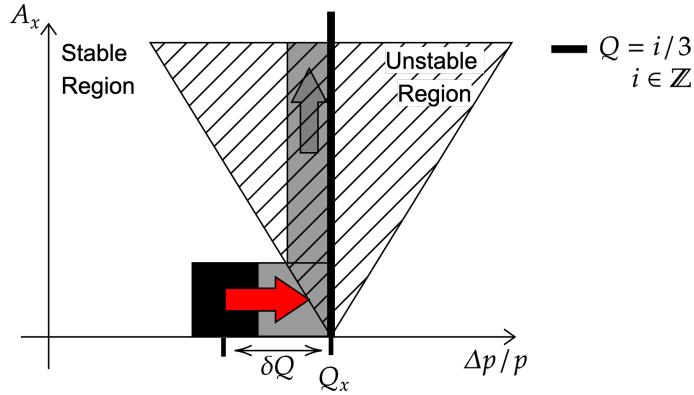


Figure 2.10: Steinbach diagram of momentum-driven slow extraction. Waiting beam in black, resonance driving effect in red, unstable beam in grey.

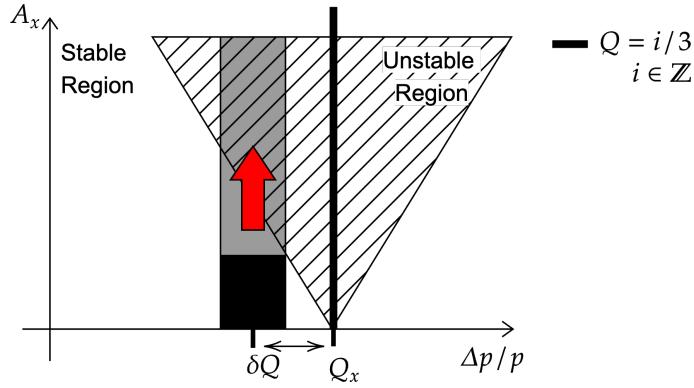


Figure 2.11: Steinbach diagram of amplitude-driven slow extraction. Waiting beam in black, resonance driving effect in red, unstable beam in grey.

the phase-space map separatrices. ?? and ?? illustrate *momentum* and *amplitude* driven extraction, respectively. In order for a particle to be considered stable, its single-particle emittance  $\epsilon$  must be lower or equal to the area of the separatrix triangle. Using ??, this condition is expressed as

$$A^2\pi \leq \frac{48\sqrt{3}\pi}{S^2} \delta Q^2\pi \quad \text{Eq. 2.68}$$

If this condition is plotted on a  $(Q, A)$  graph, the stable triangle shown in ?? and ?? is seen. The width of this unstable triangle, known as the *stopband*, is proportional to  $S$ .

### 2.3.4 Slow Extraction Techniques

A variety of slow extraction implementations can be found across particle accelerators. While all will vary to some degree, most can be categorised by which parameter is used to control extraction, and which equipment or magnet is employed. For example:

- The *National Center for Oncological Hadron Therapy* (CNAO) uses **Betatron Core** slow extraction [?]: a circular magnet produces a longitudinal acceleration, driving particles into 1/3rd-integer resonance.
- The *MedAustron* facility has also used **Constant Optics Slow Extraction** (COSE) [?], simultaneously scaling the magnetic strength of bending dipoles, and sweeping the betatron tune. This method allows for the alignment of separatrices at different momenta, improving the spill quality.
- **Phase Displacement Extraction** (PDE) allows for resonant extraction without the need for ramping magnetic elements. This avoids perturbations or ramp-rate limits which could arise, and has been tested at both the PS and MedAustron [?].
- The *Xi'an Proton Application Facility* (XiPAF) uses **Radio-Frequency Knock Out** (RFKO) extraction [?], an amplitude-driven method that transversely “heats” (excites) the beam, increasing emittance and amplitude into unstable regions.

These accelerators and extraction techniques are mentioned for illustrative purposes: it is not an exhaustive list, and most accelerators can employ multiple methods.

The work in this thesis will focus on implementing (and later simulating) *Radio-Frequency Knock Out* (RFKO) for East Area extraction at CERN. A variety of reasons motivate this: RFKO can be performed with both bunched (particles are longitudinally trapped by RF cavities) and coasting (debunched, particles spread along the entire ring circumference) beams. Performing bunched RFKO allows the same beam to be re-accelerated and re-extracted multiple times [?]. Additionally, this method is typically quicker to respond (i.e. turning on and off extraction) than other methods [?, ?, ?].

RFKO can also be used with an amplitude-modulated (AM) excitation, which can be calibrated to provide a constant particle flux over time. However, AM was not used for the experimental application in this thesis, as the extraction process took place over a long enough time period where flux was not a concern.

The principle of RFKO, as illustrated in ??, is to excite particles into the resonant instability region by providing transverse “heating” of the beam—using a radio-frequency exciter to provide a small transverse kick to the beam, increasing its amplitude. A variety of different signals can be used to drive the excitation: theoretically, white noise (a random signal with equal intensity at all frequencies) would be the ideal signal, providing a purely diffusive effect in phase space. However, pre-existing hardware favours the use of a frequency-modulated “chirp” signal, centred at the third-integer fraction of the betatron tune. A range of frequencies around this value is required as there is a range of momenta—and thus a range of tunes—present in the PS East Area beams.

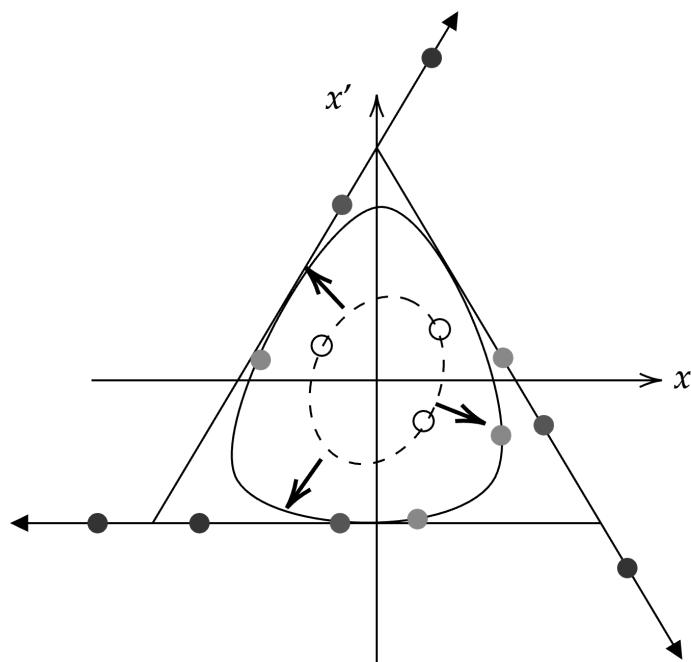


Figure 2.12: A phase-space schematic of RFKO, showing the emittance growth (dashed to solid ellipse) increasing the amplitude of a stable particle (white) into the unstable region (grey), where it moves along the separatrices (arrows) into extraction.

## Chapter 3

# Implementation and Simulation

This chapter presents the novel implementation of RFKO slow extraction at the PS, providing a spill to the East Area fixed-target experiments. As motivated in ??, this chapter will also outline the development of simulation models for the RFKO system using existing particle tracking and accelerator optics software. Significant effort was made to streamline the simulation process, leverage the GPU acceleration available in `Xsuite`, and make full use of the computing platforms available at CERN.

### 3.1 RFKO Slow Extraction at the CERN PS

For the implementation of RFKO extraction to the East Area, a basic EA-type beam was set up and configured as follows. For preliminary studies, a 750 MeV/c/nucleon Lead ( $\text{Pb}^{+54}$ ) ion beam was used, produced by Linac 3 and injected from LEIR. This is the beam configuration used by the CHIMERA experiment, as it provides the correct energy, flux, and beam size for its requirements. After these tests, when an ion beam was no longer available, a proton beam with equivalent magnetic rigidity was used.

This equivalency can be calculated using ??, ??, and ??, producing an equation for the total momentum of an ion beam:

$$pc = E_0 \sqrt{\left(\frac{E_{kin}}{E_0} + 1\right)^2 - 1} \quad \text{Eq. 3.1}$$

where  $E_0$  is the rest mass of the ion, and the relativistic gamma factor  $\gamma_c = \frac{E}{E_0} = \frac{E_0 + E_{kin}}{E_0} = \frac{E_{kin}}{E_0} + 1$ .

A ( $\text{Pb}^{+54}$ ) ion ( $Z = 82$  protons,  $N = 126$  neutrons,  $n_e = 28$  electrons) at kinetic energy 750 MeV/nucleon will have an atomic mass  $A$  of 208 u, and so a kinetic energy of

$$E_{kin} = E_{kin/u} * A = 750 \text{ MeV} \cdot 208 = 156 \text{ GeV} \quad \text{Eq. 3.2}$$

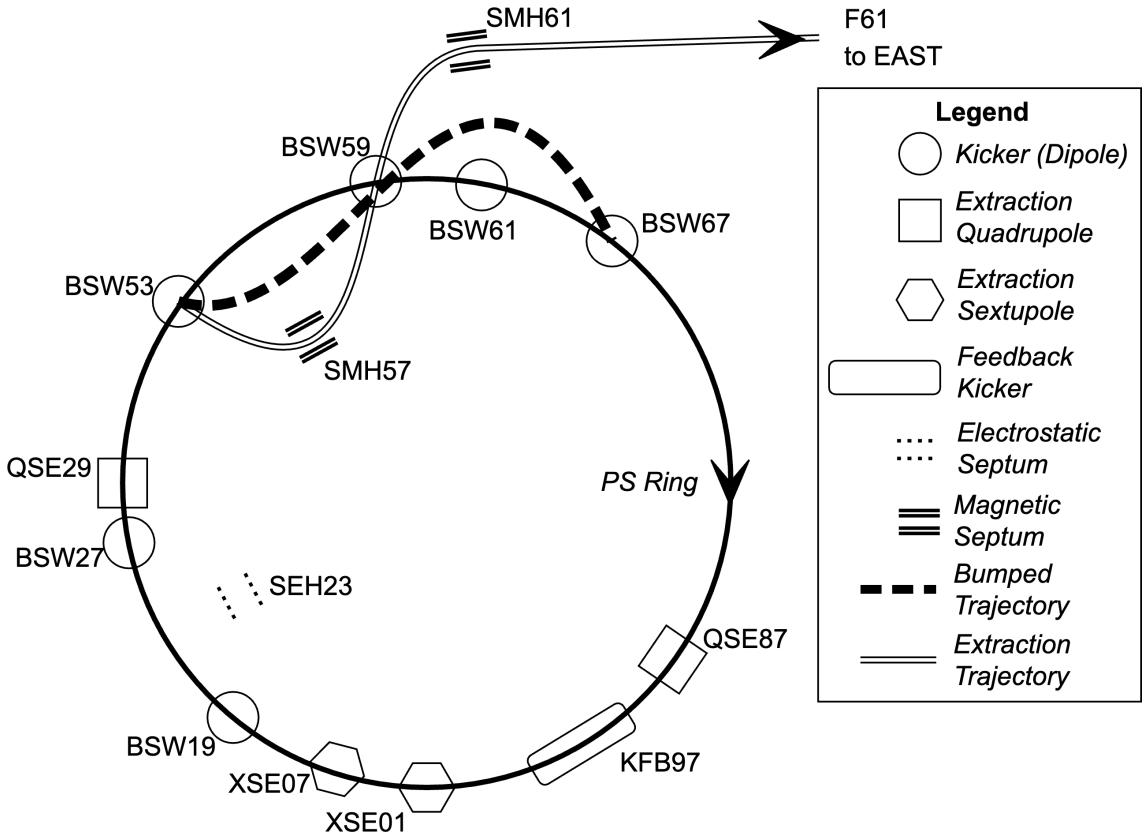


Figure 3.1: A schematic of the PS Ring and East extraction system, as configured for RFKO slow extraction. All important beamline elements are labelled with their CERN layout code. Also shown are the reference orbit (solid line), the bumped orbit (dashed line) created by the kicker magnets, and the extracted trajectory (double line) created by the magnetic septa.

and a rest mass of  $193.74 \text{ GeV } c^{-2}$ . The equivalent proton momentum, using ?? and ??, is then calculated as  $5.397 \text{ GeV } c^{-1}$ .

Therefore, a single-bunch, low intensity ( $30 \times 10^{10}$  particles/bunch)  $5.397 \text{ GeV}^1$  proton beam was chosen, and injected into the PS through a single PSB ring (ring 3). The cycle was configured to use two basic beam-cycle periods of 1.2 s, in order to extend the window available for a spill to 500 ms.

### 3.1.1 Closed Orbit and Extraction Trajectory

As shown in ??, the closed orbit of the PS at extraction differs from that of the nominal orbit. Four *bump* magnets (BSW) create a large horizontal displacement at the two septa SMH57 and SMH61. This brings the closed orbit within range of the septum blades, where resonant particles can then cross the blade and enter the septum window.

<sup>1</sup>a low intensity *beam* of a single, high-intensity bunch

The electrostatic septum, placed before the magnetic septa, is not used in this study, as it is skipped by the beam during ion extraction. It was decided that a proton implementation closer to the ion configuration would be advantageous, so that results can be compared to work carried out by CHIMERA.

### 3.1.2 Cycle Parameters

The dipole magnetic field required for a 5.392 GeV/c proton beam is 0.257 T, which provides the configuration for the PS Main Unit being magnets. In order to control the third-order resonance, and align the separatrices with the septum blade, the extraction sextupole system was enabled. This consists of a high-level parameter PR.XSE/K, which controls the normalised sextupole strength  $k_2$  of the two extraction sextupoles, XSE01 and XSE07 (??).

A high-level parameter QSE is also provided, which controls the normalised strength  $k_1$  for a set of extraction quadrupoles (??). These serve to aid in betatron tune  $Q_x$  control, and to increase the amplitude  $\beta_x$  while decreasing the dispersion  $D_x$  at the septa locations. This large  $\beta_x$  allows the septum to be placed further from the circulating beam, and the small dispersion  $D_x$  aligns the separatrices at different momenta along the septum blade. These two parameters, QSE and XSE, have been pre-configured for nominal East Area slow extraction.

The high-level tune control knobs were also configured for RFKO. These use the Pole Face Winding (PFW) magnets present in every PS Main Unit dipole to control the horizontal and vertical betatron tunes  $Q_{x,y}$  and chromaticities  $Q'_{x,y}$ . Tune control was very important to the set-up of this implementation—the horizontal tune was positioned close to the third-order resonance, at 0.32. The chromaticity was minimised to avoid part of the circulating beam being brought into resonance before the RFKO system was activated, which would cause a low-level initial uncontrolled spill. This tune control is shown in ??, which plots the vertical tune against the horizontal tune. This provides a convenient visualisation of first-, second-, and third-order resonances. These resonances are defined by

$$m \cdot Q_x + n \cdot Q_y = p \quad \text{Eq. 3.3}$$

where  $m$ ,  $n$ , and  $p$  are arbitrary integers, where  $|m| + |n|$  defines the resonance order. Also used is the fractional tune split,  $\Delta$ , defined as

$$\Delta = Q_x - Q_y - p \quad \text{Eq. 3.4}$$

From ??, we find the tune split at injection  $\Delta = -0.1$ , and at extraction  $\Delta = 0.07$ , both of which avoid coupling resonances.

The RF system in the PS (??) is used to accelerate the beam up to 5.392 GeV  $c^{-1}$  during the *ramp* period, and to maintain bunching during *flat-top*. An East-type beam in the PS is bunched at the 8<sup>th</sup> harmonic, or can be disabled to *debunch* the beam (known as a coasting beam). RFKO can be performed with both a bunched and coasting beam, and both scenarios will be examined.

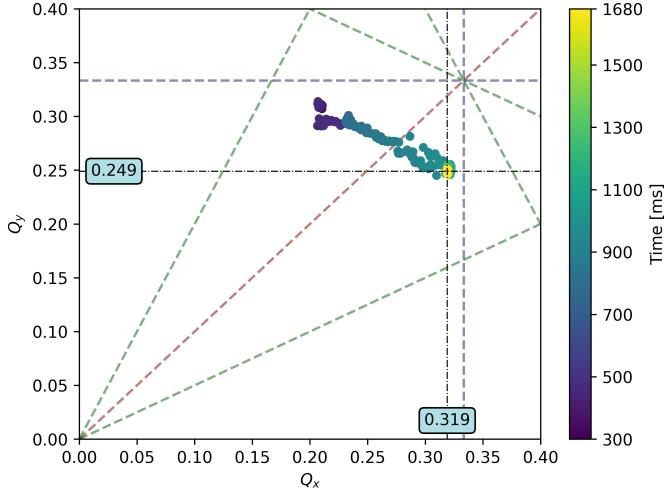


Figure 3.2: A fractional tune map, plotting the evolution of horizontal fractional tune  $Q_x$  against vertical fractional tune  $Q_y$  during RFKO slow extraction. Colour scale indicates time. Dashed lines indicate some of the second-order (red, blue) and third-order (green) resonances. Black lines at  $Q_x = 0.319$ ,  $Q_y = 0.249$  are the final tunes reached during the extraction period. Integer part for all tunes is 6.

### 3.1.3 Transverse Excitation

Once a basic, stable cycle is configured in the PS, the transverse excitation is implemented. This uses a transverse kicker located at section 97 of the PS, KFB97, a  $\sim 1$  m long stripline kicker, with both horizontal and vertical planes [?]. With a peak RF power of 5000 W across the two electrodes, and an impedance of  $100 \Omega$ , the transverse feedback (TFB) kicker can provide a peak voltage at 1000 V and a current of 10 A. The electric and magnetic fields are then defined as:

$$E = \frac{V}{r} \quad \text{Eq. 3.5}$$

$$B = \frac{2I}{2\pi r} \cdot \mu_0 \quad \text{Eq. 3.6}$$

where  $r$  is the electrode aperture 70 mm. The electric and magnetic deflection angles for a proton ( $q = 1$ ) are then [?, 3.1-3.2]

$$\theta_E = \frac{E \cdot L}{p \cdot \beta} \quad \text{Eq. 3.7}$$

$$\theta_B = \frac{B \cdot L}{p} \quad \text{Eq. 3.8}$$

where  $\beta$  is the relativistic beta factor and  $L$  is the electrode length 935 mm. The stripline components  $B$  and  $E$  both contribute to the kick, and so the summation of these two angles provides a maximum deflection of  $10.81 \mu\text{rad}$  for the  $5.397 \text{ GeV } c^{-1}$  proton beam.

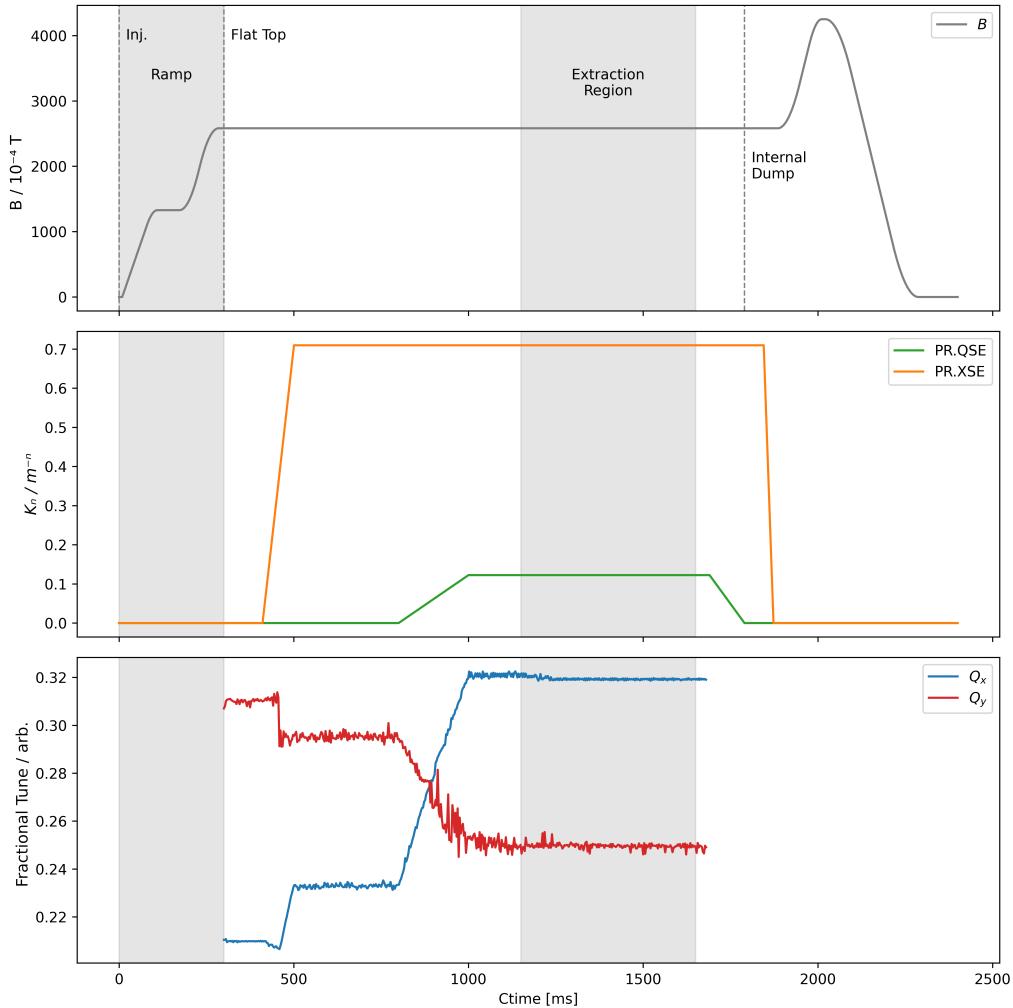
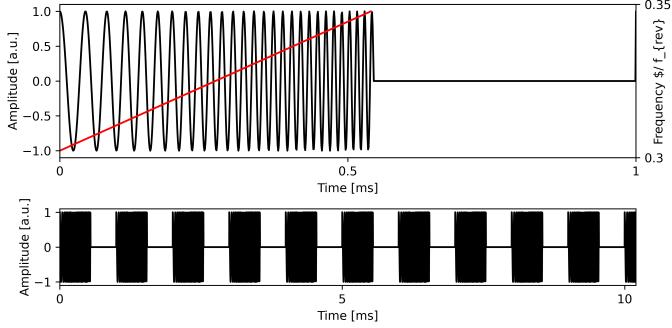


Figure 3.3: The high-level parameter configuration for the implementation of East Area RFKO slow extraction. Top graph shows the bending dipole  $B$  field, points of injection (from PSB) and dumping (internal absorption of the remaining beam), and regions of acceleration (ramp) and stable period for extraction. Second graph shows multipole components  $k_n$  for extraction quadrupole ( $k = 1$ ) and sextupole ( $k = 2$ ). Final graph shows fractional betatron tune measurements for horizontal and vertical plane.

To drive this transverse kicker, two systems were used. Initially, the `Qmeter` control program [?] provided an interface to this kicker. While the program was designed—and used extensively in this study—for tune measurements, it can be exploited to provide a frequency-modulated *chirp* signal crossing the 3<sup>rd</sup>-order resonance to the TFB plates. The program was used to configure the *start* and *stop* frequencies for the frequency-modulated chirp (as unit interval multiples of the revolution frequency  $f_{rev}$ ), chirp length (in  $2^n$  revolutions), and the time after which the chirp would repeat (in milliseconds). This produced the effect of a repeated frequency-modulated chirp windowed by

Parameter	Value
Chirp Start [Hz]	$0.3 \cdot f_{rev}$
Chirp Stop [Hz]	$0.35 \cdot f_{rev}$
Chirp Repetition [ms]	1
Chirp Turns	256

Table 3.1: Initial parameters for RFKO extraction using the *Qmeter* control program.Figure 3.4: An example of the chirp signal generated by the *Qmeter* application: a frequency-modulated chirp windowed by a square signal. Top plot shows the signal (amplitude, black) and frequency modulation (red) in the timeframe of one repetition. Bottom plot shows multiple repetitions over the span of 10 ms.

a square signal (??). Previous studies [?] found suitable parameters to produce a baseline RFKO extraction implementation, tabulated in ??, and found a significant dependency of spill quality on the chirp repetition rate.

To explore chirp repetition rates above the 1 kHz (below 1 ms) limit of *Qmeter*, a new signal generator was configured to control the TFB kicker plates. This function generator was configured as a two-channel source, channel 1 being a sine wave with a frequency modulated by channel 2, a ramp/sawtooth wave. By choosing a central frequency and frequency deviation corresponding to the chirp generated by *Qmeter*, the function generator was able to produce a similar signal to that of the *Qmeter* system, but not windowed by a square function (??). This allowed for chirp repetition rates of 1 kHz and above, and was used to explore the effect of chirp repetition rate on spill quality.

To aid in the research in this paper, and future use-cases of RFKO at the PS, a package was written to simplify the interface with the AFG via the *Standard Commands for Programmable Instruments* (SCPI) protocol. This interface—available as part of the *CERN General Devices* package [?]?—configures the AFG for RFKO extraction, and provides a Python class for modifying the chirp-generation parameters in terms of the revolution frequency  $f_{rev}$ , and the chirp length in milliseconds. This package was used to control the AFG for the results presented in this paper.

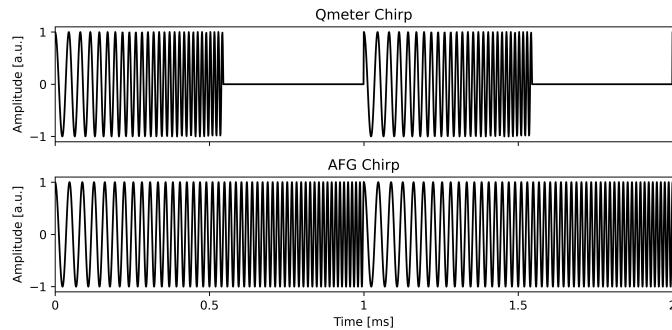


Figure 3.5: An example of the chirp signal generated by *Qmeter* (top) and by the arbitrary function generator (AFG) (bottom).

## 3.2 Modelling RFKO Slow Extraction

As motivated in ??, the majority of work surrounding the results presented in this thesis was towards developing a simulation framework to accurately model the spill structure of RFKO at the PS. This section will detail the simulation tools used, the implementation of the RFKO extraction simulation, and the methods used to compare the simulation results to machine data.

### 3.2.1 Simulation Tools

Various simulation software packages for accelerator physics are available, each with their own optimisations and application areas amongst the accelerator physics community. For this study, the MAD-X [?] and Xsuite [?] packages were used. MAD-X is a common tool in the area of optics and beam transfer, and CERN provides a common optics repository—acc-models [?—containing MAD-X *sequence* files describing the optics lattices for the CERN accelerator complex. For the PS, many *scenarios* are provided alongside the operational models, including a pre-configured setup for East Area slow extraction.

Xsuite—a relatively recent simulation tool, first made available in 2021—is a Python package that shares many similarities with MAD-X, and can import MAD-X sequences into its own lattice definitions. However, Xsuite has additional capabilities, such as *context* management, which provides the ability to run simulations on different hardware—namely GPUs. This is achieved by using the cupy [?] package, which provides a drop-in replacement for numpy, but with optimisations for accelerating workloads on GPUs. Additionally, Xsuite has an existing Exciter beam element, which can model the time-dependent transverse kicker used in RFKO. There has been a rapid increase in the use of Xsuite within the accelerator physics community at CERN, which—combined with the additional capabilities beyond MAD-X—motivated its use in this study.

### 3.2.2 Lattice Configuration

The *East Area Extraction* scenario provided by `acc-models` was used as the basis for all simulation models, and was modified to include the parameters outlined in ???. This generated lattice file—defining the strengths and location for all 1326 beam elements—was used to produce the optics plots shown in ??, which displays how the kicker magnets create large displacements at the extraction septa, and how dispersion is controlled by the extraction quadrupoles.

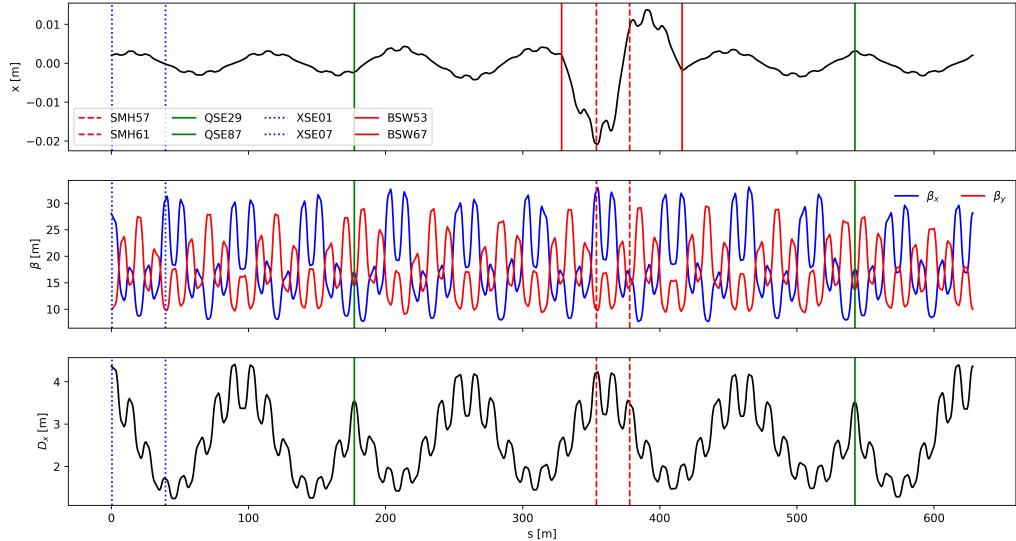


Figure 3.6: Optics plots of the closed orbit for the `acc-models` PS East Area extraction scenario. Top figure shows horizontal closed orbit  $x$ , middle figure shows Twiss beta function  $\beta$ , and bottom figure shows transverse dispersion  $D_x$ . Also labelled are key extraction elements: the magnetic septa (SMH), the extraction quadrupoles (QSE), extraction sextupoles (XSE), and kickers (BSW)

This lattice was then imported to `Xsuite`, where the custom RFKO beam elements could be included. These elements were:

- An `Exciter`, modelling the TFB kicker KFB97.
- A `LimitRect`, modelling the magnetic septum SMH57.
- A `Cavity`, modelling the RF cavities providing acceleration and longitudinal focusing, operating at the 8<sup>th</sup> harmonic.

#### Exciter

Using ?? and ??, a function (??) was written to convert the voltage applied across the TFB kicker to the angle of deflection generated by the electric and magnetic fields. This value was then used to define the strength of the `Exciter` element, which was placed at the location of the TFB kicker in the lattice. Another function (??) was written to generate the chirp signal as generated by `Qmeter`

and the AFG. This function was then used to define the Exciter element’s `samples` parameter, which is used to define the time-dependent kick.

### Septum

The first magnetic septum, SMH57, was modelled as a simple rectangular aperture `LimitRect`. By setting this aperture to coordinates corresponding to the window of the septum, particles which are lost to the aperture can be considered as extracted. In the real machine, particles receive two magnetic septum kicks, from SMH57 and SMH61, but this study only considered the first septum to simplify the model and avoid simulating lost particles for longer than necessary. As only one aperture is modelled, it can be assumed that particles lost to an aperture limit—with `Xsuite` particle state 0 (flag `XT_LOST_ON_APERTURE`)—entered the window of the septum and were extracted.

### RF Cavity

A single RF cavity was added to the model, corresponding to the RF system of the PS. While the real system comprises of many such cavities, modelling a single cavity with the total RF voltage applied provides an accurate representation of the longitudinal dynamics of the beam, given that there are no other longitudinal elements present in the model.

#### 3.2.3 Distribution Generation

To accurately simulate the conditions used in the RFKO studies, the initial particle distribution must be generated using machine measurements. The parameters used included:

- Beam momentum  $p$  [GeV  $c^{-1}$ ]
- Intensity  $I$  [proton/bunch]
- Horizontal RMS emittance  $\varepsilon_x$  [mm mrad]
- Longitudinal RMS emittance  $\varepsilon_s$  [eV s]

These values were obtained during machine studies. The beam momentum and intensity were fixed values, configured by high-level machine parameters. The horizontal emittance  $\varepsilon_x$  was measured using Secondary Emission Monitor (SEM) grids [?], and the longitudinal emittance  $\varepsilon_z$  using tomography of the beam current profile [?]. The measured values are shown in ???. The beam distribution was then generated according to these parameters, and “matched” to the model longitudinal dynamics of the model by `Xsuite`. An example of the generated distribution is shown in ???.

It was decided that the simulation code will generate a new distribution each time it was initiated, rather than loading a pre-generated distribution from a file. This was done to allow for the random distributions to be generated with a different random seed each time, avoiding the same distribution being used for each simulation which could introduce bias in the results. This also allowed for the particle generation parameters to be dynamically adjusted on a per-simulation

Parameter	Value
Beam Momentum $p$ [GeV/c]	5.392
Intensity $N$ [proton/bunch]	$3 \times 10^{11}$
Transverse Emittance $\varepsilon_{x,y}$ [mm mrad]	2.38
RMS Longitudinal Emittance [eVs]	0.276

Table 3.2: Initial parameters for particle generation, corresponding to measurements made during machine studies.

basis, enabling the testing and optimisation of these parameters. This design choice ensured that the simulation code was flexible and could be used for a wide range of studies at different beam energies and intensities.

With a sufficiently large number of particles, these random distributions will be statistically similar. This obviated the need to transfer a large, persistent file to each simulation job; this transfer time was, qualitatively, found to be longer than the time taken to generate a new distribution file of the same size.

### 3.2.4 Simulation Process

Once an initial simulation process was built, one issue became immediately apparent: to simulate the full 0.5 s extraction region (??), the simulation needed to process approximately 235,000 turns. To produce a statistically significant extraction rate, a sufficient number of particles—roughly 500,000—needed to be generated and simulated. A typical desktop computer, for instance, would require over 600 hours to complete such a simulation, at a rate of roughly 13 s/iteration.

To accelerate the simulation speed to a more reasonable rate, two tools were employed. While not unique to Xsuite, they are both simplified by the package. The first method was the simplest: as mentioned previously, Xsuite provides ready-to-go compatibility with CuPy and CUDA GPU acceleration. The second method relies on CERN’s Batch Service, a high-throughput computing service which allows users to submit jobs via the *HTCondor* [?] framework.

To submit the simulation jobs to the batch processing service, the simulation software was *containerised* using a Docker [?] image. This image provides a complete environment for the simulation, including a distribution of Python, the Xsuite package, the CuPy package, and various other dependencies.

This image was made available on CERN’s image registry using a *continuous integration* platform, allowing the image to be automatically re-built and re-published when changes are made to the configuration. It was also distributed via the *CERN Virtual Machine File System* (CVMFS), which HTCondor can use to initialise machines running batch jobs. This container [?] eliminated the need to set up the simulation environment on each machine, ensured that the simulations were run in a consistent environment, and enabled the use of HTCondor’s GPU resources.

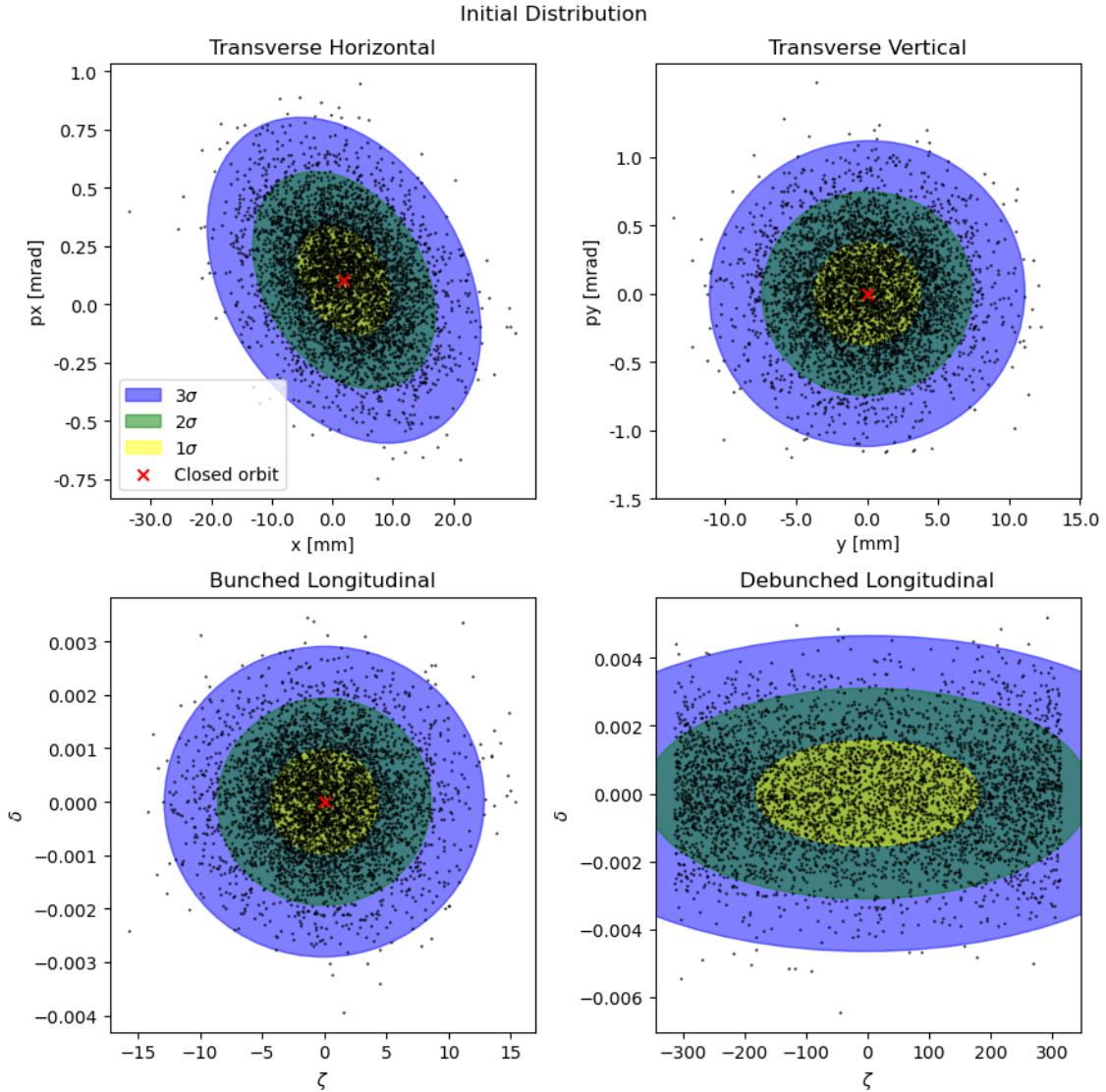


Figure 3.7: Examples of initial distributions generated by Xsuite, for number of particles  $N_p = 3000$ , with parameters as shown in ???. Also plotted on each are the  $1\sigma$ ,  $2\sigma$ ,  $3\sigma$  emittance ellipses, and the closed orbit particle (red). Top plots show transverse phase-space distribution for all simulation types. Bottom plots show longitudinal phase-space distribution for a *bunched* (RF system active) beam (left) and *debunched* (RF inactive) beam (right).

?? shows a flowchart describing the simulation process as successfully implemented. While some efficiency could be made by including shared resources (such as the `ps.seq` lattice file) within the Docker image, developing a general base environment for any `Xsuite` simulation will prove useful in the wider community, and in work outside of the scope of this thesis. Initial test simulations included a `Monitor` object to record the turn-by-turn particle parameters of simulations. This is a very computationally expensive and memory-intensive process, and was removed from the final simulation process, but enabled detailed debugging and analysis of the simulation process. One example, ??, shows the progression of phase-space particle distribution throughout the first 1000 turns of a simulation. Here, the characteristic features of the Kobayashi Hamiltonian phase-space map (??) can be seen, in addition to the transverse heating (emittance growth, ??) and extraction separatrices as discussed in ??.

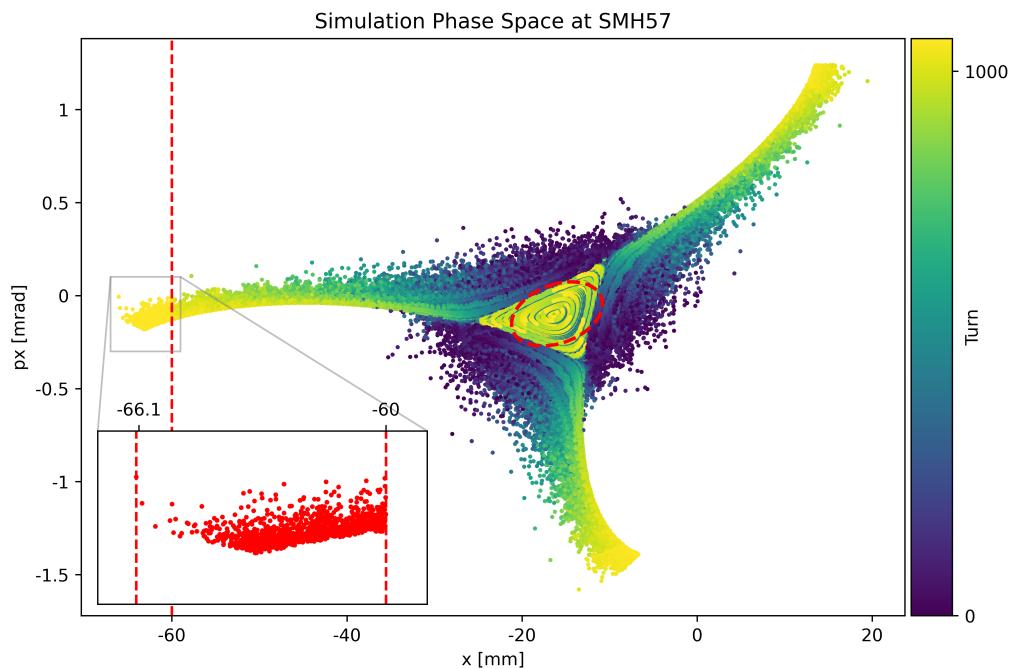


Figure 3.8: Phase-space distribution of the RFKO simulation at the SMH57 septum location, plotting all particles for the first 1000 turns (approx. 2.128 ms in laboratory time). RFKO driven by a transverse excitation frequency chirp across  $f_{rev}/3$  at a *rep-rate* of 0.12 ms. Colour scale indicates turn number. Inset shows extracted particle distribution and a spiral step of  $\Delta X_3 \approx 6.1$  mm.

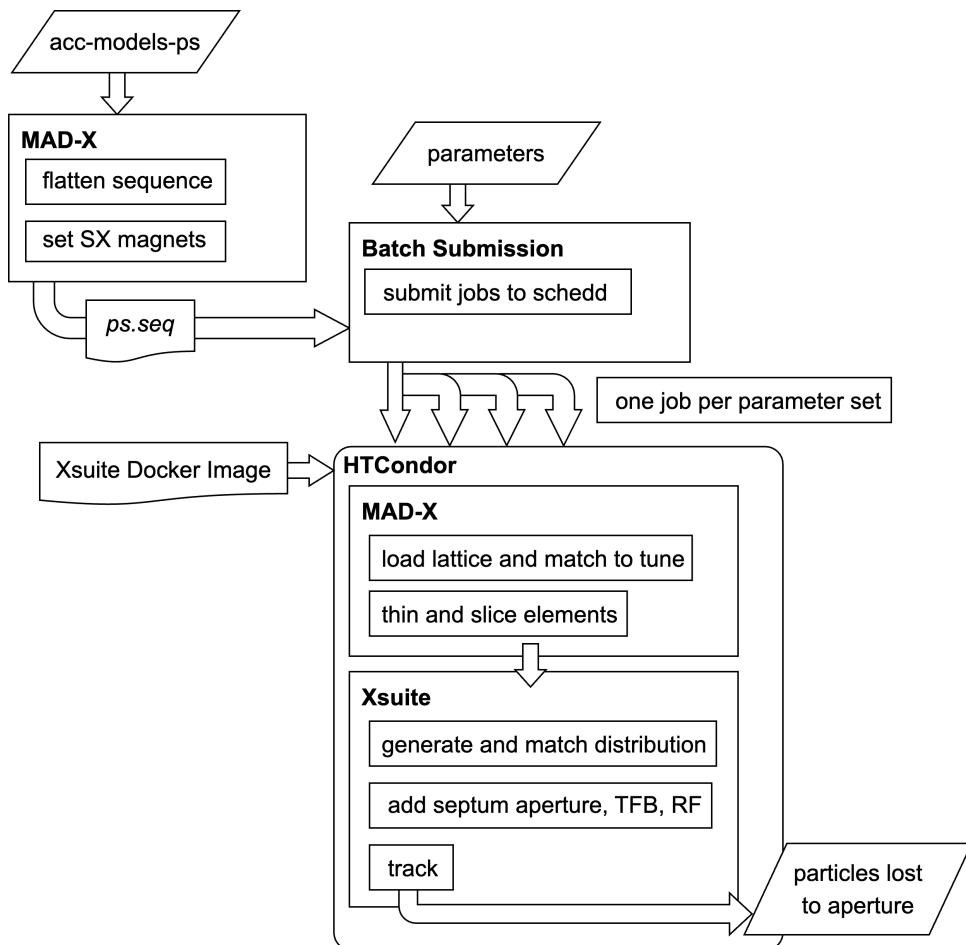


Figure 3.9: A flowchart describing an overview of the simulation process

## Chapter 4

# Analysis and Comparison of Results

This chapter presents the methodology and results of experiments seeking to confirm the successful implementation of RFKO slow extraction at the PS for East Area spills. The accuracy of the simulation models will be confirmed by analysing the spill characteristics, compared between experimental data taken in June 2023 and the results of simulations. This chapter will also seek to demonstrate the benefits of GPU-accelerated particle simulations using `Xsuite`, as compared to typical CPU workflows.

### 4.1 Computation Time

In order to compare the advantages of GPU acceleration provided by `Xsuite`, the simulation process described in ?? was implemented and run using CERN’s batch service. The simulations used single-CPU compute nodes, either using `Xsuite`’s CPU context, or GPU context with additional Nvidia GPUs via `Xsuite`’s CUDA interface. Simulations were run for a range of numbers of particles and turns, and the real-time (*wall-clock*) duration of the simulation was recorded. These results are presented in ?? and ???. The number of particles in a simulation— $N_p$ —effectively controls the computational intensity, as it determines the number of calculations which must be performed per iteration (turn). The number of turns— $N_t$ —controls the number of iterations, and is therefore used to normalise computation time between simulation runs.

These figures show a significantly different dependency of simulation process time on the operational intensity between CPU and GPU workloads. For CPU tasks there is no parallelism available, so for  $N_p > 10^2$ —where performance is not dominated by constant-time *overhead* processes—performance is dependent on  $N_p$ . For GPU tasks, however, there is a significant initial overhead process of allocating and transferring data from the CPU to the GPU. This is apparent from the stratification of performance (time per turn) in ???. The performance, however, is not dependent on  $N_p$  for  $N_p < 10^6$ , where the performance is dominated by the constant-time parallelism available.

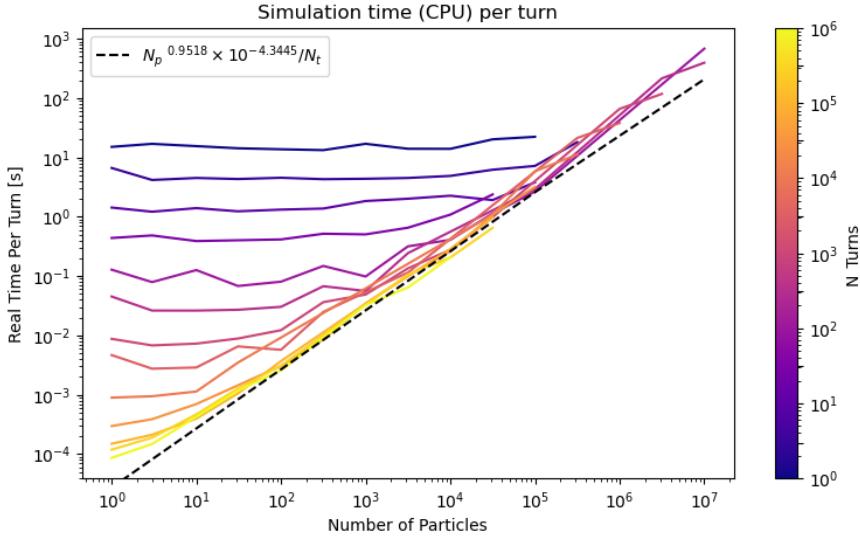


Figure 4.1: Real simulation time (s) per turn for CPU workload per turn as a function of number of particles. Line colour indicates the number of turns simulated.

GPUs employ a *Single-Instruction Multiple-Data* (SIMD) architecture to execute a single operation on multiple data points (*vectors*) concurrently. This is in contrast to CPUs, which employ a *Single-Instruction Single-Data* (SISD) architecture, where each operation is performed on a single data point at a time. This means that the GPU can perform  $N_p$  operations in the same time that a CPU could perform a single operation, and so the performance is not dependent on  $N_p$ . The GPU constant-time performance is saturated for  $N_p > 10^6$ , where all available parallelism is being used, and performance is once again dependent on  $N_p$ . Previous research has shown similar performance dependency on  $N_p$  for GPU-accelerated simulations [?].

The parallelism-saturated performance of both CPU and GPU simulations can be formulated as linear relationships in log-log, using the equation

$$t/N_t = N_p^\alpha \times 10^\beta \quad \text{Eq. 4.1}$$

where  $t/N_t$  is the real execution time (seconds) per turn, and  $\alpha$  and  $\beta$  are some constants of the performance of a particular machine. By fitting the linear domain performance to this equation, values of  $\alpha$  and  $\beta$  can be estimated for particular CPUs and GPUs. For these benchmark simulations, various AMD Epyc CPUs were assigned to the machines, and NVIDIA A100 40GB GPUs were used. This specific combination of hardware resulted in  $\alpha_{CPU} = 0.95178$ ,  $\beta_{CPU} = -4.34452$  for CPU-only simulations, and  $\alpha_{GPU} = 0.97671$ ,  $\beta_{GPU} = -6.65887$  for GPU simulations (??, ??).

These optimisation improvements make a clear justification for the use of GPU-accelerated workflows. ?? shows the relative performance increase of GPU acceleration over CPU-only simulations, for the same  $N_p, N_t$ . In the best case, with  $N_p, N_t \approx 5 \times 10^4$ , the GPU-accelerated simulation can be  $\sim 500$  times faster than the CPU-only simulation. At more reasonable  $N_p, N_t$ ,

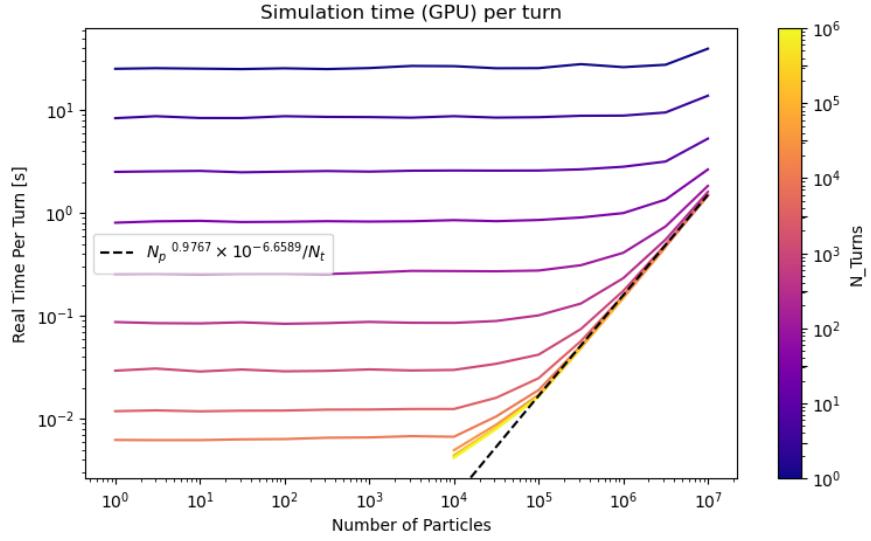


Figure 4.2: Real simulation time (s) per turn for GPU workload per turn as a function of number of particles. Line colour indicates number of turns simulated.

GPU simulations are still several orders of magnitudes faster than a CPU, and only slower at  $N_p, N_t < 10^1$ .

Previous work on optimising simulations slow extraction (at the SPS) achieved a simulation time of 7 hours 30 minutes (27 000 s) for a total of 200 000 turns and 1000 particles [?, Figure 3], using a similar batching process, but without the advantage of GPU acceleration. The simulation process described in ?? achieved a total time of 13 minutes (779.8 s) for the same number of particles and turns running on a GPU platform. This represents a  $\times 34$  speed improvement.

These findings motivated the selection of parameters for later simulations. To simulate the full 500 ms extraction period requires 235 000 turns to be simulated. At this number of turns, the available GPU parallelism becomes saturated around 500 000 particles. By not exceeding this amount, the simulation can be completed in a reasonable amount of time and still yield significant results.

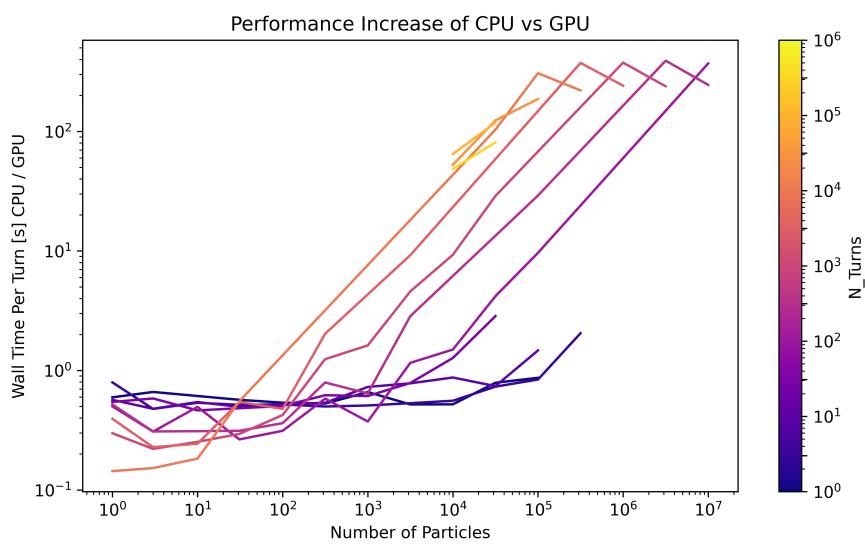


Figure 4.3: Real (*wall time*) simulation time (s) per turn for CPU workload divided by GPU workload for the same  $N_p, N_t$ , showing the relative performance increase of GPU acceleration. All other parameters (e.g. number of compute cores) are held constant.

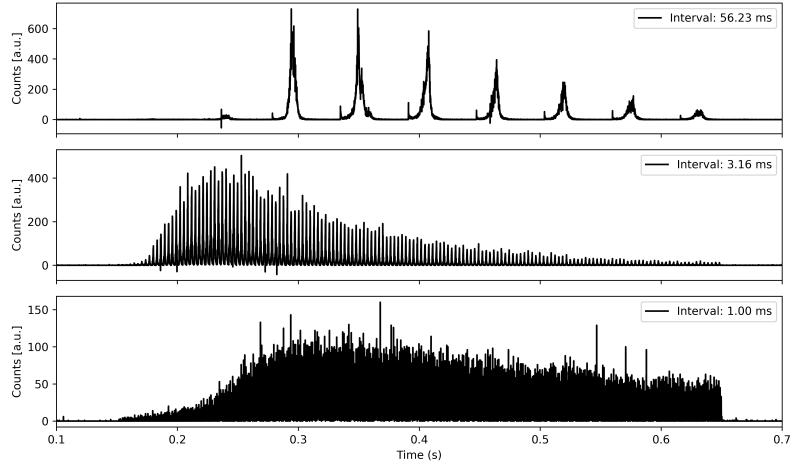


Figure 4.4: Example signals recorded by the BXSCAL scintillator at various *rep-rate* intervals.

## 4.2 Frequency Response

In order to compare the accuracy of RFKO slow extraction simulations to measurements taken from the implementation on the real machine, analytical tools established by previous work [?] were used to measure the frequency response of the RFKO system. By varying the rate at which a chirp was repeated, known as the *interval frequency* or *rep rate*, the resulting spill could be characterised.

### 4.2.1 Analysis Methodology

For the machine implementation, the spill was measured using an  $N_2$  gas scintillator [?] located in the F61 extraction line (see ??). This scintillator is measured by various devices: the interface BXSCAL was selected, which provided time-series measurements of particle counts at  $\sim 18.8$  kHz. The spill was measured for a range of interval frequencies using an automated script interfacing with the *Java API for Parameter Control* (JAPC) software at CERN. Multiple data sets were taken for each interval frequency, for a range of peak voltages applied to the TFB kicker.

Once the scripts completed, the BXSCAL data was retrieved from NXCALCS, CERN's logging database [?]. This time-series scintillator provides a signal proportional to the instantaneous number of particles passing through the F61 extraction line, and can, therefore, be used as the extraction rate spill signal. An example of this data can be found in ??, showing a range of rep-rate intervals. The time-series spill signal was then analysed using a discrete Fourier transform (DFT) to obtain the frequency response of the system.

Once the DFT was computed, the interval rate for each measurement was selected, and normalised by the DFT DC component (the 0 Hz term, equivalent to the average of all samples). A peak-finding algorithm was used to identify this interval rate, allowing for a slight variation in the spill measurement frequency. Over the measurements used, the BXSCAL sample rate varied between

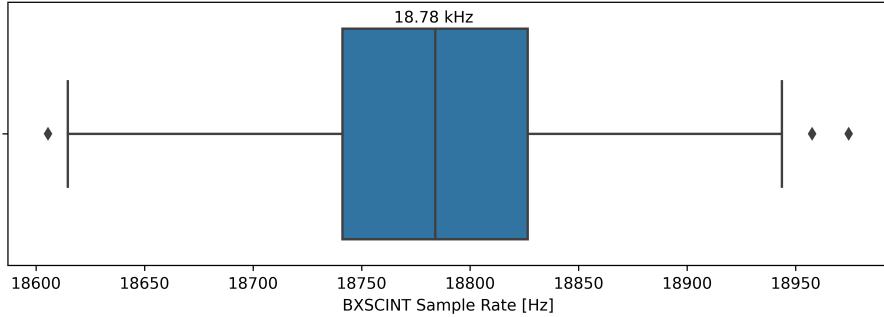


Figure 4.5: Box-plot of the variation in sample rate of the BXSCAL device over the course of the measurement. Box shows the 25<sup>th</sup> to 75<sup>th</sup> percentile, with error bars bounding the 10<sup>th</sup> and 90<sup>th</sup> percentile. Data points outside of this range are considered outliers and shown as diamond points.

18.6 kHz and 18.95 kHz (??). This sample rate variation likely originates from the electrical sampling equipment, or from changes in gas pressure [?]. As the variation was evenly distributed around the mean value of 18.78 kHz, the DFT was largely unaffected. The peak-finding algorithm was used to identify the highest peak within a range corresponding to the inter-quartile range of sample rate variance. An example of these DFTs, with the range and selections of the peak-finding algorithm, can be seen in ??.

To compensate for the “noise” level of the machine data—a result of the measurement device resolution, scintillator “after-glow”, residual current (“dark current”), or readout noise—the noise level was estimated from signal regions before excitation, then subtracted from the measured frequency amplitudes. Error-bar calculations for both machine and simulation data sets were made using the variance between multiple measurements at the same parameters, the variance between all measurements, and the machine noise level.

This selection process was robust but noisy, especially at higher frequencies, where the variation in the BXSCAL sample rate has a higher impact on the accuracy of the DFT. To mitigate this, the peak-finding algorithm was run on multiple repetitions of the same parameters, and the results averaged.

This same analytical method was performed for simulations; the turn number at which a particle was lost to the aperture was recorded, along with its longitudinal  $\zeta$  variable, defined as

$$\zeta = s \frac{\beta}{\beta_0} - \beta c t \quad \text{Eq. 4.2}$$

where  $\beta$  is the relativistic Lorentz factor. This was used to reconstruct the *laboratory* time (in real seconds, not turn number) at which the particle was lost to the aperture—and thus considered as extracted—using the equation

$$t = N t_{rev} - \frac{\zeta}{\beta_0 c} \quad \text{Eq. 4.3}$$

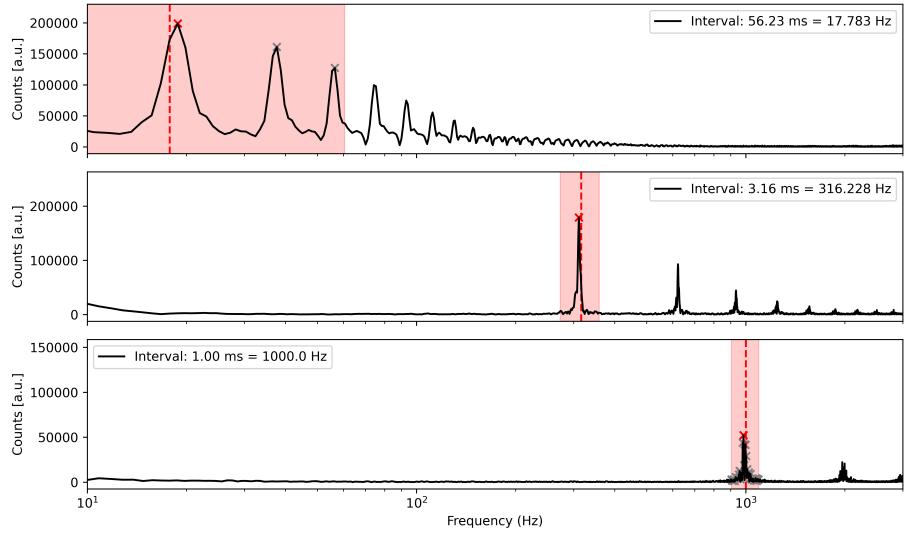


Figure 4.6: Example Discrete Fourier Transforms (DFT) of the spill signals in ???. The *expected* interval frequency is shown as a dashed line, and the range around which the peak-finding algorithm searched is shaded. The accepted peaks are shown in grey marks, and the selected peak is shown in red.

where  $t_{rev} = 1/f_{rev}$ , the revolution time of the reference particle, and  $N$  is the turn at which the particle was lost. This equation can also be found in the source code for the `Exciter` element. The extraction time data for each particle was then used to reconstruct the time-series extraction data (??). The same peak-finding algorithm as used for machine data was then employed to generate frequency response data for simulations.

#### 4.2.2 Results of Frequency Response Analysis

?? presents the results of the analysis method described in ?? applied to data measured from the PS. The graphs characterise the *frequency response* of the spill to different chirp signal *rep-rate* intervals, and indicate how this frequency propagates to the output (spill). The frequency response of the RFKO system is shown for both a bunched beam (upper image) and a coasting beam (lower image). A distinct *low-pass filter* effect is visible in both *log-log* plots, where the frequency response is reduced at higher frequencies. This effect, also observed by previous studies [?, ?], results in higher frequencies being “smoothed out” by the RFKO system.

The frequency response  $H(f)$  of a low-pass filter can be approximated to the first order as

$$H(f) = A \cdot \frac{1}{1 + \left(\frac{f}{f_c}\right)^{2n}} \quad \text{Eq. 4.4}$$

where  $f_c$  is the cutoff frequency,  $n$  is the order of the filter (which determines the steepness of the roll-off), and  $A$  is a gain constant. The frequency response curve at each gain was fitted to this function, and the fitted curve and variance error are shown in red in ?? and ??.

For **bunched beams**, the low-pass filter was found to have a cutoff frequency of  $(1.140 \pm 0.173)$  kHz in machine measurements, and  $(1.490 \pm 0.257)$  kHz in simulation. Both machine measurement and simulation fits resulted in an order of exactly 1, to 7 decimal places. These results show that the frequency characteristics of the bunched-beam RFKO simulations are accurate to the machine implementation.

For **coasting beams**, the fitted low-pass filter has a cutoff frequency of  $(1.120 \pm 0.190)$  kHz in machine measurements, and  $(5.26 \pm 1.03)$  kHz in simulation. While these results have a significantly larger error than the bunched beam results, they are still in agreement with each other. The filter order  $n$ , however, is inconsistent between simulation and measurements: in simulation, the curve is fit to  $n = 10$ , a very steep roll-off; for the machine measurements, the curve is fit to  $n = 1$ .

Both plots also show a *minimum transfer characteristic*, where the frequency response is at its lowest: for bunched beams, this minimum is at 4 kHz, meaning a chirp signal repeating every 0.25 ms will have the least effect on frequencies present in the spill signal. For coasting beams, this minimum is less clear, and appears to shift to higher frequencies with increasing chirp signal voltage.

The low-pass filter effect, and corresponding cutoff frequency and minimum transfer characteristics, can be explained by the time taken for particles to move from a stable phase-space area to the magnetic septum [?]. This time  $T_{tr}$ , in units of 3 turns, is known as the *transit time*, and is given by [?, ?]:

$$T_{tr} = \frac{1}{\sqrt{3}\varepsilon_Q} \ln \left| \frac{n}{n+3} \frac{3}{\frac{\dot{\varepsilon}_Q}{\varepsilon_Q} - \frac{1}{\sqrt{3}} \frac{\dot{\varepsilon}_Q}{\varepsilon_Q^2}} \right| \quad \text{Eq. 4.5}$$

The first fraction within the logarithm, where  $n = |\vec{X}_{MS}|/|\vec{X}_{sep}|$ , accounts for the time taken to move from a fixed point  $O'$  to the magnetic septum, where the distance between the edge of the stable phase-space area and the septum is  $|X_{MS}|$ , and the distance between the edge of the stable phase-space area enclosed by the separatrices and its center is  $|X_{sep}|$ . The second fraction, where  $\varepsilon_Q$  was defined as the modified tune distance in ?? and ??, accounts for time taken to move along the side of a stable triangle to the fixed point  $O'$ . ?? illustrates this transit time in phase-space.

In a machine with a chromatic tune spread, this transit time  $T_{tr}$  will have a spread of values  $\Delta T_{tr}$ , which represents a time interval within which particles which leave the stable phase space simultaneously will arrive at the septum. This results in a *limiting frequency*  $f_{lim} = 1/\Delta T_{tr}$ , above which frequencies in the spill structure are “washed out”. The spread of transit times is therefore approximated to [?, ?]

$$\Delta T_{tr} \approx \frac{dT_{tr}}{d\varepsilon_Q} \Delta \varepsilon_Q \propto \frac{1}{\varepsilon_Q^2} \Delta \varepsilon_Q \quad \text{Eq. 4.6}$$

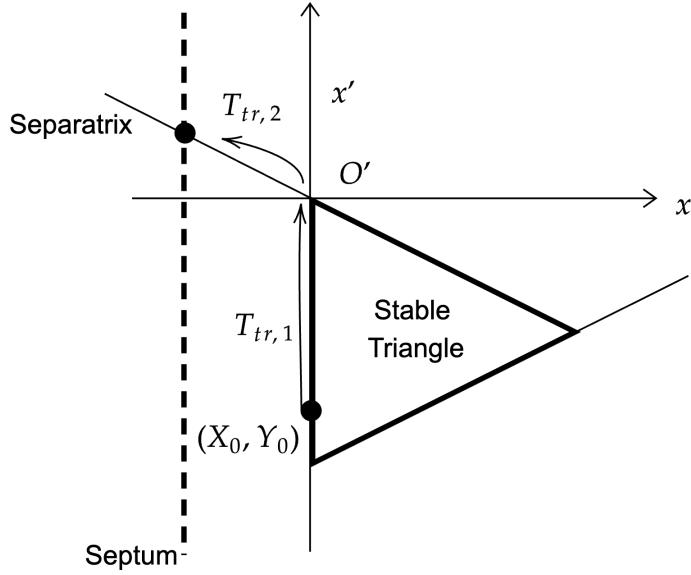


Figure 4.7: A phase-space diagram illustrating the transit time  $T_{tr}$  elapsed between leaving the last stable triangle and arriving at the septum. The particle motion (and transit time) is presented in units of three turns; in real-time, the particle moves between the three separatrix arms. The Kobayashi Hamiltonian forms the separatrices of the stable triangle. Total transit time  $T_{tr} = T_{tr,1} + T_{tr,2}$  depends both on the distance  $(x_0, y_0)$  of the particle from a fixed point  $O'$  ( $T_{tr,1}$ ) and the distance from  $O'$  to the septum blade ( $T_{tr,2}$ ).

where  $\bar{\varepsilon}_Q$  is the mean modified tune distance. The consequence of this transit time spread is a spread in septum arrival time between instantaneously excited particles. Should the chirp repetition period be shorter than this interval, the chirp repetition frequency component will be “washed out” in the spill structure.

The simulated frequency response characteristics are shown in ??, using the same parameter ranges as ???. The bunched beam frequency response shows the same minima characteristic as the real machine, with a minimum at 4 kHz. The coasting beam simulations, however, do not have the same accuracy. The frequency response minimum seen in ?? is still present at 4 kHz, but the typical low-pass filter effect is not seen, with the response returning to a high level as the chirp frequency approaches 10 kHz.

?? and ?? directly compare the frequency response between empirical measurements and simulations, for both bunched and coasting beams. These graphs present the average response curve between different voltages, and the fitted low-pass filter model, with the cutoff frequency  $f_c$  marked. While less pronounced in a *log-log* graph, the similarities in the response frequency for bunched beams are clearly seen in ???. For coasting beams (??), this similarity is poorly observed, motivating the conclusion that—while the low-pass filter is still present—the coasting beam simulations were inaccurate.

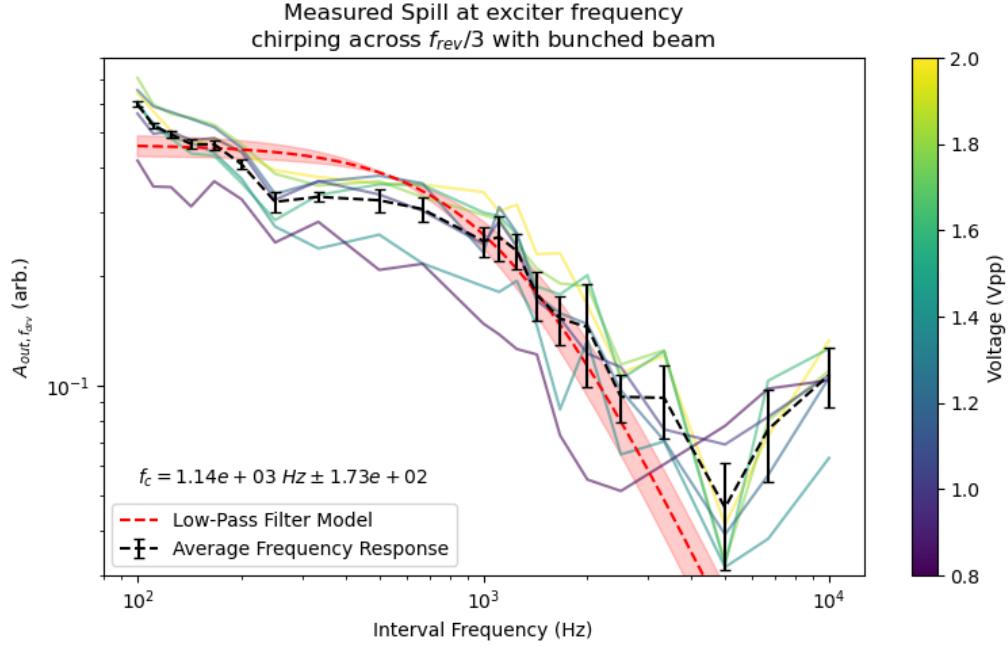
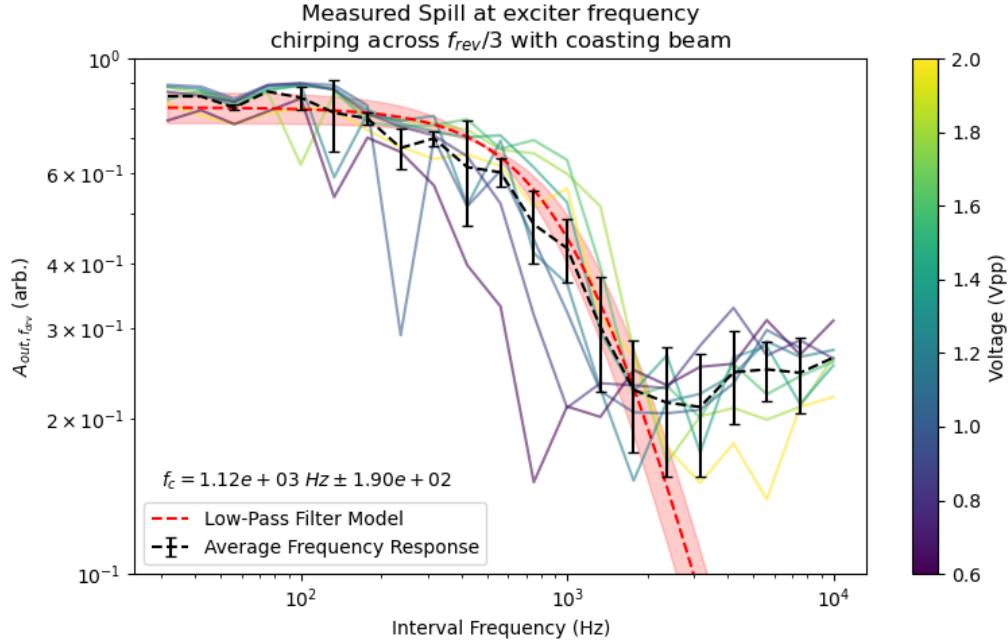
(a) Frequency response of machine measurements with a **bunched** beam.(b) Frequency response of machine measurements with a **coasting** beam.

Figure 4.8: Frequency response of the RFKO system as implemented in the PS, both with a bunched beam (upper image) and coasting beam (lower image).

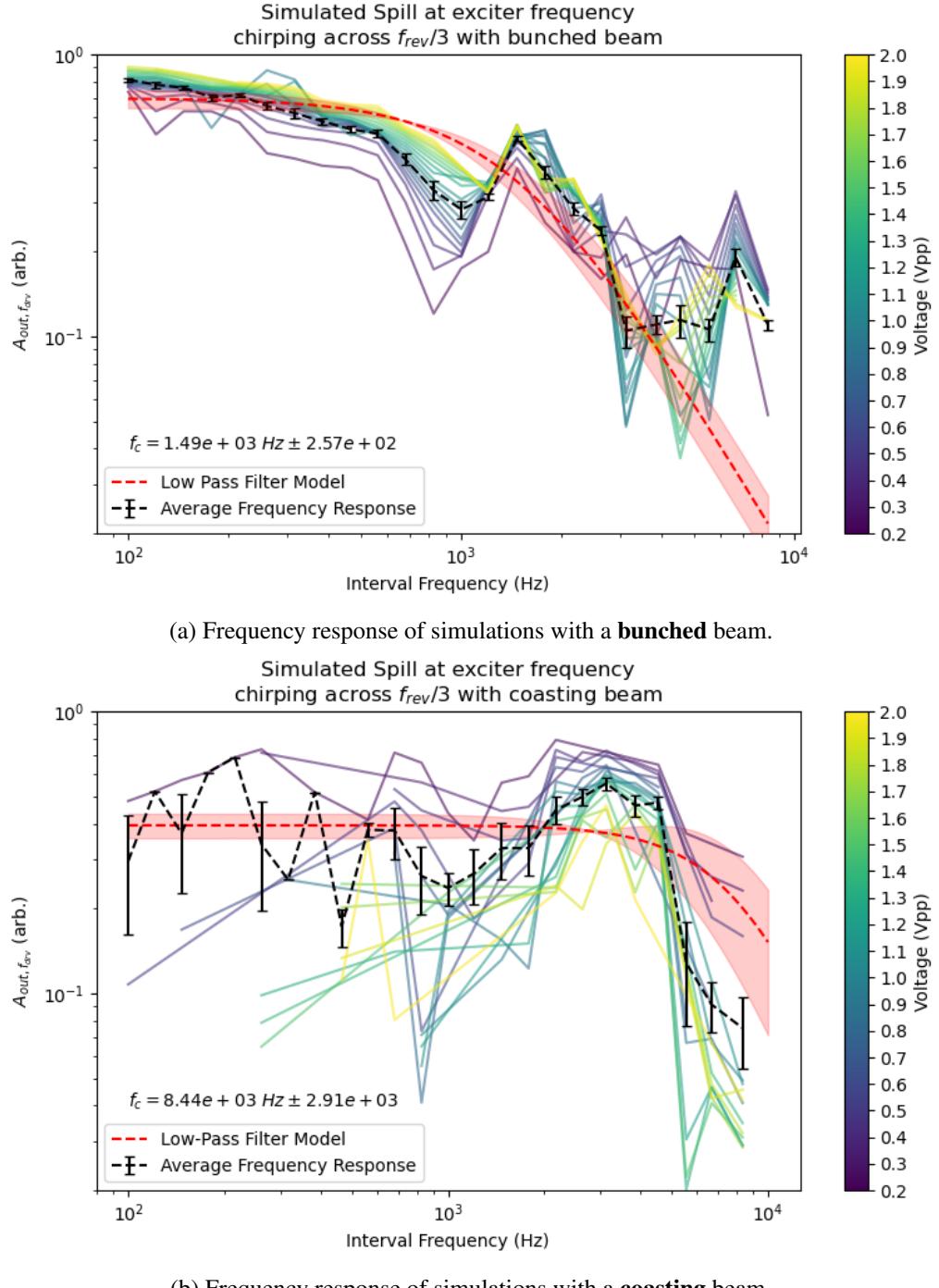


Figure 4.9: Frequency response of the RFKO system as simulated in Xsuite, both with a bunched beam (upper image) and coasting beam (lower image).

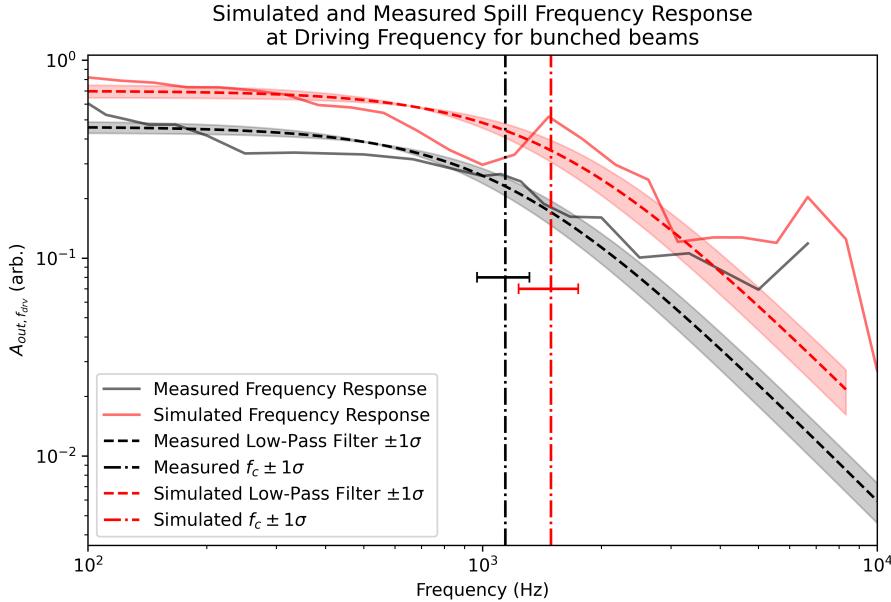


Figure 4.10: Graph comparing the averaged frequency response curves for bunched beams between machine data (black) and simulations (red). Shaded is the  $1\sigma$  error, dashed is the fitted low-pass filter model, and vertical dashed is the calculated cutoff frequency  $f_c$  and the  $1\sigma$  error.

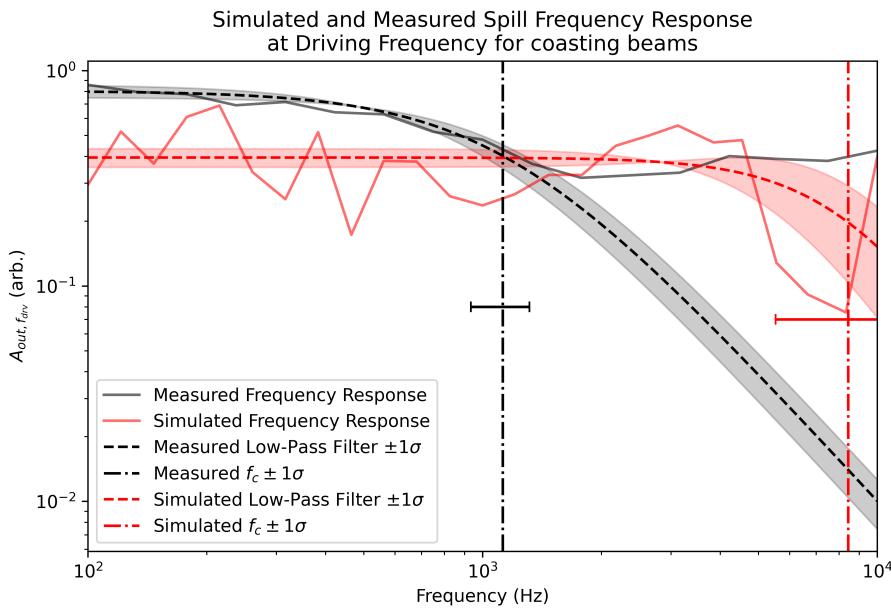


Figure 4.11: Graph comparing the averaged frequency response curves for coasting beams between machine data (black) and simulations (red). Shaded is the  $1\sigma$  error, dashed is the fitted low-pass filter model, and vertical dashed is the calculated cutoff frequency  $f_c$  and the  $1\sigma$  error.

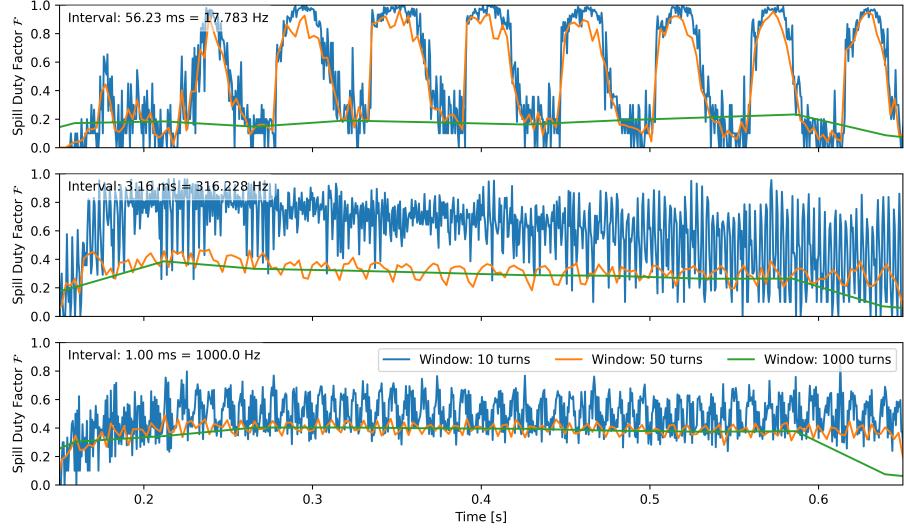


Figure 4.12: Example rolling spill duty factor ( $\mathcal{F}$ ) of the spill signals in ???. Calculations made with 10, 50, and 1000 sample windows.

### 4.3 Spill Duty Factor

Another analytical tool for quantifying slow extraction spill quality established by previous work is the *spill duty factor*. This quantity is defined as [?]

$$\mathcal{F} = \frac{\langle x \rangle^2}{\langle x^2 \rangle} \quad \text{Eq. 4.7}$$

where  $\langle x \rangle$  denotes the average of an arbitrary variable  $x$ . By calculating this quantity within a rolling time window, a rolling average of spill quality can be calculated. ?? presents an example of these calculations for the spill signals in ???: a flatter spill will produce a duty factor closer to  $\mathcal{F} = 1$ , while a highly pulsed spill approaches  $\mathcal{F} = 0$ .

?? shows measurements of spill duty factor for both bunched (top) and coasting (bottom) beam RFKO spills, over a range of TFB signal voltages and chirp repetition frequencies, using a 10 kHz calculation window. Windows between 1 kHz and 20 kHz were tested, and all produced similar heatmaps. These are presented in a heatmap, where the lightness of the cell represents a higher spill duty factor, and thus a flatter spill. Qualitative analysis identifies regions where the spill quality is optimised (high  $\mathcal{F}$ ), making for optimal parameter selection. For bunched beams, the area of optimal spill quality is large, with a maximum duty factor  $\mathcal{F} = 0.8$ . For coasting beams, the region of optimal spill quality is much smaller, with a maximum duty factor  $\mathcal{F} = 0.6$ . Previous work on optimising slow extraction at the PS supports this finding of increased spill duty factor with RF cavities enabled, compared to beams with RF cavities disabled [?].

The mechanism behind the spill duty factor improvement with bunching can be understood by examining how the stable phase space behaves with respect to synchrotron oscillation. As covered in ??, particles moving within longitudinal RF buckets will undergo longitudinal oscillations around the synchronous particle, known as synchrotron oscillations<sup>1</sup>. This can be calculated as

$$Q_s = \sqrt{\frac{hqV_{rf}|\eta| \cos \phi_s}{2\pi\beta_0^2 E_{tot}}} \quad \text{Eq. 4.8}$$

where  $h$  is the RF harmonic,  $V_{rf}$  the RF voltage,  $q$  the particle charge,  $\eta$  the slip factor (??),  $\phi_s$  the synchronous phase, and  $E_{tot}$  the total energy. It can also be found by simulation; for the configuration outlined in preceding sections, the synchrotron tune was found to be  $Q_s = 0.0052$ . In a Steinbach diagram (??), this has the effect of shifting the particle horizontally, in addition to the vertical push from the exciter. This has two consequences: the particles remain at the edge of the unstable region for a shorter amount of time, as the synchrotron motion changes the particle's tune enough to move it into resonance, reducing the effect of the excitation structures. Additionally, it can reduce the effect of magnet ripples changing the shape of the unstable region.

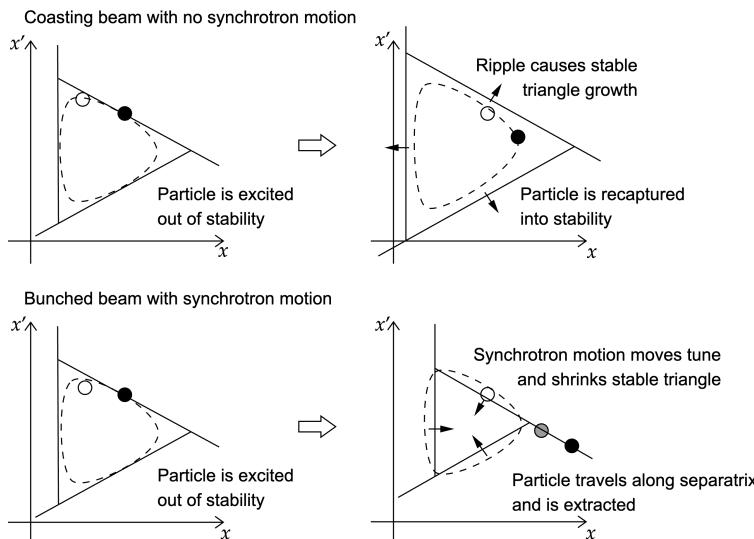


Figure 4.13: Phase space of a particle approaching the separatrix in a coasting beam (top) and a bunched beam (bottom), both with a small sextupole ripple component.

To see this, we will consider the two cases of a particle approaching resonance in a coasting beam and in a bunched beam, visualised in phase space in ???. The upper figure is the case for a coasting beam with no synchrotron motion: particles experience no change in tune, and so a small ripple in the sextupole field could increase the stable region area and “recapture” the particle back into the stable region, preventing the particle’s extraction. The lower figure describes the case for a bunched beam with slow synchrotron motion: the particles receive a kick (RFKO) *and* periodic tune change (RF cavity), which reduces the stable phase-space area. This reduces the chance of a

<sup>1</sup>As opposed to transverse betatron oscillations.

particle being recaptured by a sextupole ripple, pushes the particle deeper into resonance quicker, and “smoothes” out the spill signal. There is, however, a limit to this ripple mitigation, where the synchrotron oscillation is too fast for particles to remain in the unstable region. [?, ?].

This synchrotron motion also effects the extraction transit time. The distance of a particle from the separatrix while in the unstable region grows with each synchrotron period. This reduces the transit time. Therefore, there exists an optimum synchrotron tune, where particles are more likely to remain within the unstable region, but the transit time remains long enough to smoothen the spill “pulses” between successive excitations (??). Both of these effects, and the limiting frequency obtained from transit time spread, contribute to the optimal selection of parameters for an RFKO driving signal.

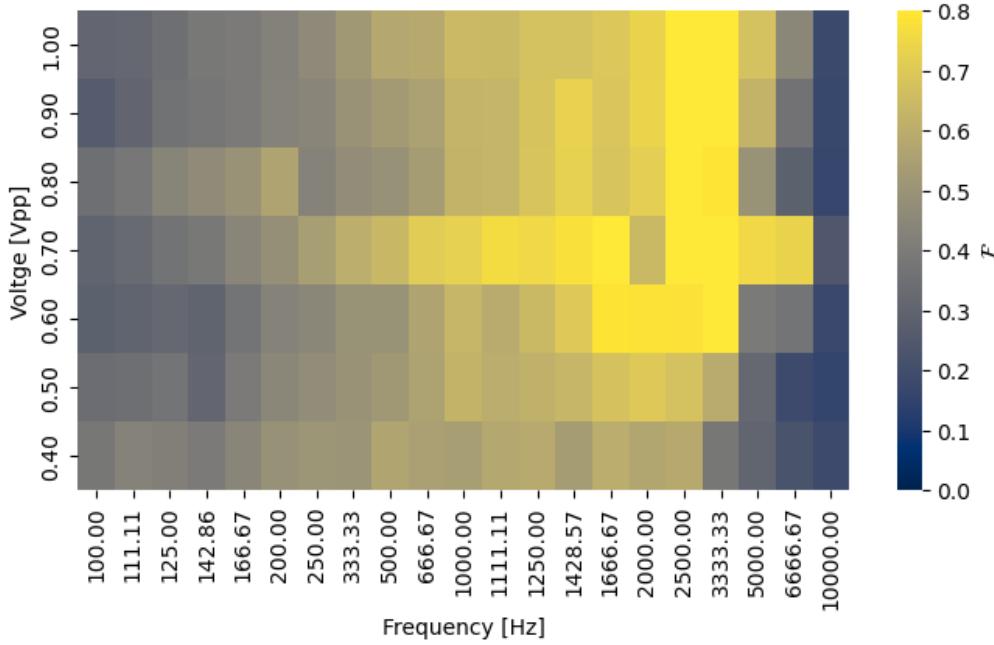
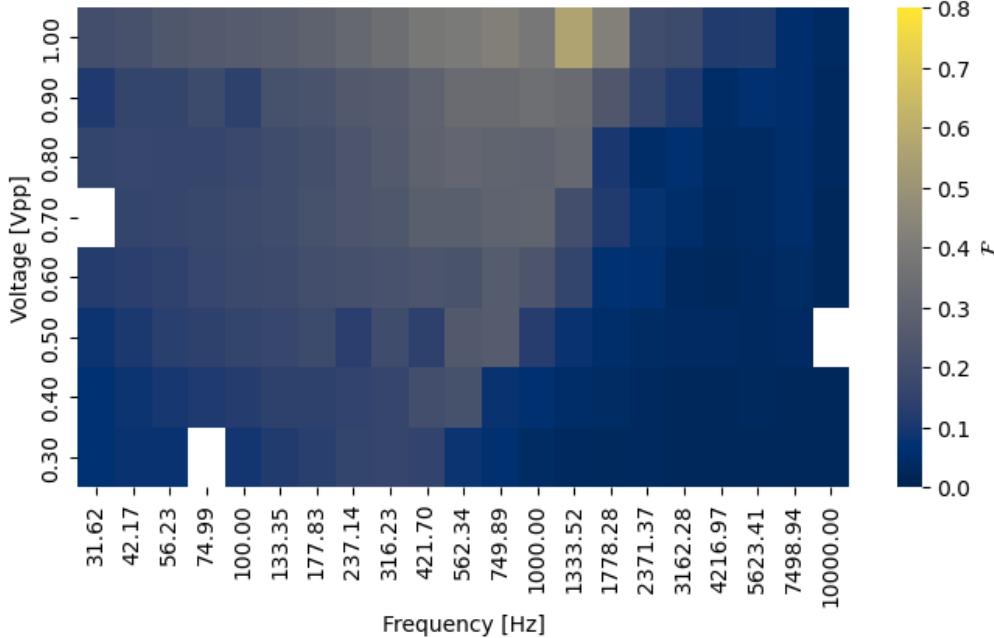
(a) Spill duty factor heatmap for **bunched** beam as measured in the machine.(b) Spill duty factor heatmap for **coasting** beam as measured in the machine.

Figure 4.14: Spill duty factor ( $\mathcal{F}$ ) heatmaps for the RFKO system as measured in the machine, calculated with 10 kHz windows, both with a bunched beam (upper image) and coasting beam (lower image), as a function of chirp repetition frequency (log-scale) and voltage (linear). Colour scale represents  $\mathcal{F}$ .

# Chapter 5

## Conclusion

This thesis presents a number of developments towards the operational implementation of Radio Frequency Knockout (RFKO) slow extraction at CERN’s Proton Synchrotron (PS) for the East experimental area (EA). Within the goal of producing low intensity, high charge, high energy, and long duration spills for fixed-target experiments, the implementation of RFKO presented shows promising spill frequency structure and duty factor performance. The results of machine studies presented in ?? and ?? use both time-domain and frequency-domain analysis to show that the system is capable of producing spills with a flat frequency structure at durations significantly longer than 100 000 turns, and with spill duty factors (measured ) of up to 80 % for bunched beams, and 60 % for coasting beams.

This thesis also demonstrates significant agreement between the frequency response of this RFKO implementation and the characteristics observed in previous studies [?,?]. This agreement is also well-observed between the real machine and simulations with a bunched beam, where both datasets place the optimum repetition frequency (minimising the presence of this frequency in the spill) for an RFKO-driving chirp signal between 2 kHz and 4 kHz. This finding motivates the development of a dedicated signal generator for the PS TFB system for use with RFKO. This agreement is less well-observed for coasting beams, where the simulations show lower frequency response minima than machine observations.

The implementation of RFKO simulations in Xsuite presented in ?? also shows significant performance improvements over previous work simulating Slow Extraction at the SPS with MAD-X [?]. By exploiting GPU-accelerated computing provided by Xsuite’s compatibility with CuPy, CERN’s batch-processing service, and containerised environments, the simulations presented in this thesis are up to 400 times more performant in real-world simulation time than CPU-based methods.

## 5.1 Outlook

There are several opportunities for improvements presented in this thesis, both for potential future studies, and for the operational implementation of RFKO at the PS:

- As mentioned in ??, only one of the magnetic septa—SMH57—was modelled for extraction. The accuracy of the simulation model would benefit from future development to simulate particles through both the septa and into the F61 line. This would require combining the PS optics models with the *F61* optics models provided by *acc-models-tls* [?].
- The setup of the AFG used to generate chirp signals with repetition rates below 1 ms (??) was only a temporary solution, and is not suitable for long-term operational use, as it operates entirely independently from the LHC Software Architecture [?] (LSA) database. A dedicated signal generator, with an interface to CERN’s Technical Network and LSA, would be a significant improvement.
- The extraction procedure outlined in this thesis only used frequency-modulated chirp signals to sweep across the range of particle tunes. Other work has used signals such as coloured noise [?]. These signals, and others such as dual-band frequencies, could be explored in future studies, using the same equipment and methods as presented in this thesis.
- In addition to frequency-modulation control, amplitude-modulation control has been studied in the past, and can offer improvements in macro-scale spill structure [?]. This technique could be used to mitigate the exponential decay seen in all spill structures presented in this thesis, and extend spill times to durations significantly longer than 0.5 s.
- While simulations for bunched beams show good agreement with the machine implementation, the results of coasting (debunched) beams are less clear. The complete dynamics of the coasting beams in RFKO would require further characterisation to understand the differences between the simulations and the machine fully.

# Chapter 6

## Acknowledgements

This work, and my research over the past 12 months at CERN, would not have been possible without the tutelage and support of my studentship supervisor, Yann Dutheil. I cannot thank him enough for his patience and guidance in turning a Bachelors student—without any formal background in Accelerator Physics—into a Master’s candidate. I would also like to thank Matthew Fraser for sharing his endless knowledge and his assistance with my research and machine operation.

I am also deeply grateful to my section colleagues—Elliott, Pablo, Francisco, Vittorio, and Rebecca R.—for their friendship and for welcoming me so warmly to ABT-BTP, my section and group leaders—Chiara and Wolfgang, colleagues across the campus—Emily, Rebecca T., and the ABP, RF, BI, and OP-PS groups—and from among the wider community for their discussions, advice, and knowledge throughout my time at CERN. I would also like to thank my academic supervisor, Stephen, for his encouragement to pursue this project.

Finally, I would like to thank my family, friends, and loved ones for their support from many hundreds of miles away.

# Bibliography

- [1] M. A. Fraser. Fast extraction: single and multi-turn. volume Vol 5, page Extraction and Transfer. CERN, 2018. DOI: 10.23730/CYRSP-2018-005.285.
- [2] M. Fraser, P. A. Arrutia Sota, K. Biłko, N. Charitonidis, S. Danzeca, M. Delrieux, M. Du-raffourg, N. Emriskova, L. S. Esposito, R. García Alía, A. Guerrero, O. Hans, G. Imesch, E. Johnson, G. Lerner, I. Ortega Ruiz, G. Pezzullo, D. Prelipcean, F. Ravotti, F. Roncarolo, and A. Waets. Feasibility of Slow-Extracted High-Energy Ions From the CERN Proton Synchrotron for CHARM. *JACoW IPAC*, 2022:1703–1706, 2022. DOI: 10.18429/JACoW-IPAC2022-WEPOST012.
- [3] K. D. Metzmacher and L. Sermeus. The PSB Ejection and Recombination Kicker Systems for LHC. Technical report, CERN, Geneva, 2000. URL: <http://cds.cern.ch/record/2061508>.
- [4] S. Ghithan, G. Roy, and S. Schuh. Design study of beam transport lines for BioLEIR facility at CERN. *JINST*, 12(09):P09019, 2017. DOI: 10.1088/1748-0221/12/09/P09019.
- [5] J. Barranco García and S. Gilardoni. Simulation and optimization of beam losses during continuous transfer extraction at the cern proton synchrotron. *Phys. Rev. ST Accel. Beams*, 14:030101, Mar 2011. DOI: 10.1103/PhysRevSTAB.14.030101.
- [6] B. Goddard, B. Balhan, J. Borburgh, L. Esposito, M. A. Fraser, L. Jorat, V. Kain, C. Lolliot, L. S. Stoel, P. van Trappen, F. M. Velotti, D. Barna, and D. Veres. Reduction of 400 GeV/c slow extraction beam loss with a wire diffuser at the cern super proton synchrotron. *Phys. Rev. Accel. Beams*, 23:023501, Feb 2020. DOI: 10.1103/PhysRevAccelBeams.23.023501.
- [7] K. Brunner. Improving the slow extraction efficiency of the CERN Super Proton Synchrotron, 2018. Presented 2018. URL: <https://cds.cern.ch/record/2650734>.
- [8] K. J. le Couteur. The regenerative deflector for synchro-cyclotrons. *Proceedings of the Physical Society. Section B*, 64(12):1073, dec 1951. DOI: 10.1088/0370-1301/64/12/308.
- [9] K. J. L. Couteur. The extraction of the beam from the liverpool synchrocyclotron. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 232(1189):236–241, 1955. URL: <http://www.jstor.org/stable/99763>.

- [10] P. A. Arrutia Sota, P. Burrows, A. De Franco, M. Fraser, B. Goddard, V. Kain, F. Kuehteubl, M. Pivi, D. Prokopovich, and F. Velotti. Implementation of a Tune Sweep Slow Extraction with Constant Optics at MedAustron. *JACoW IPAC*, 2022:1715–1717, 2022. DOI: 10.18429/JACoW-IPAC2022-WEPOST015.
- [11] F. C. Iselin. The MAD Program (Methodical Accelerator Design): Reference Manual. 5 1985. URL: <https://inspirehep.net/literature/214504>.
- [12] H. Wiedemann. *Particle Accelerator Physics*. Springer International Publishing, 2015. DOI: 10.1007/978-3-319-18317-6.
- [13] L. Nevay, J. Snuverink, A. Abramov, L. Deacon, H. Garcia-Morales, H. Lefebvre, S. Gibson, R. Kwee-Hinzmann, W. Shields, S. Walker, and S. Boogert. Bdsim: An accelerator tracking code with particle-matter interactions. *Computer Physics Communications*, 252:107200, 02 2020. DOI: 10.1016/j.cpc.2020.107200.
- [14] J. Rossbach and P. Schmüser. Basic course on accelerator optics. 1994. DOI: 10.5170/CERN-1994-001.17.
- [15] E. Courant and H. Snyder. Theory of the alternating-gradient synchrotron. *Annals of Physics*, 3(1):1–48, 1958. DOI: 10.1016/0003-4916(58)90012-5.
- [16] R. D. Ruth. An Introduction to Particle Accelerators . *Physics Today*, 55(8):52–52, 08 2002. DOI: 10.1063/1.1510283.
- [17] W. Hillert. Transverse linear beam dynamics, 2021. DOI: 10.48550/ARXIV.2107.02614.
- [18] F. Tecker. Longitudinal Beam Dynamics - Recap. *CERN Yellow Rep. School Proc.* arXiv: 2004.11908, DOI: 10.23730/CYRSP-2018-005.31.
- [19] F. Tecker. Longitudinal Beam Dynamics in Circular Accelerators. In *CAS - CERN Accelerator School 2019: Introduction to Accelerator Physics*, 11 2020. arXiv: 2011.02932.
- [20] L. Badano, M. Benedikt, P. J. Bryant, M. Crescenti, P. Holy, A. T. Maier, M. Pullia, S. Rossi, and P. Knaus. *Proton-Ion Medical Machine Study (PIMMS)*, 1. 2000. URL: <https://cds.cern.ch/record/385378>.
- [21] Y. Kobayashi. Theory of the resonant beam ejection from synchrotrons. *Nuclear Instruments and Methods*, 83(1):77–87, 1970. DOI: 10.1016/0029-554X(70)90537-9.
- [22] L. Falbo, E. Bressi, S. Foglio, and C. Priano. Betatron Core Slow Extraction at CNAO. In *Proc. 9th International Particle Accelerator Conference (IPAC'18), Vancouver, BC, Canada, April 29-May 4, 2018*, number 9 in International Particle Accelerator Conference, pages 1237–1239, Geneva, Switzerland, June 2018. JACoW Publishing. DOI: 10.18429/JACoW-IPAC2018-TUZGBF3.
- [23] P. A. Arrutia Sota, P. N. Burrows, M. A. Fraser, and F. M. Velotti. Millisecond burst extractions from synchrotrons using rf phase displacement acceleration. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 1039:167007, 2022. DOI: 10.1016/j.nima.2022.167007.

- [24] H. Yao, X. Guan, G. Li, X. Wang, Q. Zhang, and S.-x. Zheng. RF-Knockout Slow Extraction Design for XiPAF Synchrotron. In *57th ICFA Advanced Beam Dynamics Workshop on High-Intensity and High-Brightness Hadron Beams*, page MOPR005, 2016. DOI: 10.18429/JACoW-HB2016-MOPR005.
- [25] G. Russo. Radio frequency knockout (rfko) extraction: Hardware analysis and beam optics simulation and optimization, Mar 2019. URL: <http://webthesis.biblio.polito.it/id/eprint/10253>.
- [26] T. Furukawa, K. Noda, S. Shibuya, M. Muramatsu, M. Kanazawa, E. Takada, and S. Yamada. Progress of rf-knockout extraction for ion therapy. 2002. URL: <https://cds.cern.ch/record/585180>.
- [27] T. Furukawa and K. Noda. Fast beam cut-off method in rf-knockout extraction for spot-scanning. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 489(1):59–67, 2002. DOI: 10.1016/S0168-9002(02)00895-1.
- [28] K. Noda, T. Furukawa, S. Shibuya, M. Muramatsu, T. Uesugi, M. Kanazawa, M. Torikoshi, E. Takada, and S. Yamada. Source of spill ripple in the rf-ko slow-extraction method with fm and am. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 492(1):241–252, 2002. DOI: 10.1016/S0168-9002(02)01318-9.
- [29] G. Sterbini, A. Blas, and S. Gilardoni. Beam-based Performance of the CERN PS Transverse Feedback. page MOPAB10, 2015. URL: <https://cds.cern.ch/record/2158994>.
- [30] C. Bovet, R. Gouiran, I. Gumowski, and K. H. Reich. *A selection of formulae and data useful for the design of A.G. synchrotrons; rev. version*. CERN, Geneva, 1970. URL: <https://cds.cern.ch/record/104153>.
- [31] M. Gasior and R. Jones. High Sensitivity Tune Measurement by Direct Diode Detection. 2005. URL: <https://cds.cern.ch/record/895142>.
- [32] T. Bass. Benchmarking simulations of slow extraction driven by rf transverse excitation at the cern proton synchrotron. In *Proc. IPAC’23*, number 14 in IPAC’23 - 14th International Particle Accelerator Conference, pages 2458–2461. JACoW Publishing, Geneva, Switzerland, 05 2023. DOI: 10.18429/jacow-ipac2023-tupm116.
- [33] CERN General Devices, a Python module providing a high-level API for controls integration. URL: <https://gitlab.cern.ch/abt-optics-and-code-repository/commissioning-tools/cern-general-devices>.
- [34] Xsuite, Python packages for the simulation of the beam dynamics in particle accelerators. URL: <https://xsuite.web.cern.ch>.
- [35] CERN Optics Repository, a common MAD-X optics repository for the CERN accelerator complex. URL: <https://acc-models.web.cern.ch/acc-models/>.

- [36] R. Okuta, Y. Unno, D. Nishino, S. Hido, and C. Loomis. Cupy: A numpy-compatible library for nvidia gpu calculations. In *Proceedings of Workshop on Machine Learning Systems (LearningSys) in The Thirty-first Annual Conference on Neural Information Processing Systems (NIPS)*, 2017. URL: [http://learningsys.org/nips17/assets/papers/paper\\_16.pdf](http://learningsys.org/nips17/assets/papers/paper_16.pdf).
- [37] M. Martini and H. O. Schönauer. Emittance measurements in the CERN PS complex. *Part. Accel.*, 58:137–151, 1997. URL: <https://cds.cern.ch/record/324553>.
- [38] S. Hancock and S. A. J-L. A pedestrian guide to online phase space tomography in the CERN complex. Technical report, CERN, Geneva, 2001. URL: <http://cds.cern.ch/record/960231>.
- [39] HTCondor, a high-throughput computing software framework. URL: <https://research.cs.wisc.edu/htcondor>.
- [40] D. Merkel. Docker: lightweight linux containers for consistent development and deployment. *Linux journal*, 2014(239):2, 2014.
- [41] Xsuite-Batch Docker Image. URL: <https://gitlab.cern.ch/tbass/xsuite-docker>.
- [42] S. E. Hegglin. Simulating Collective Effects on GPUs, 2016. Presented 10 Feb 2016. URL: <https://cds.cern.ch/record/2239398>.
- [43] P. Schicho. Optimizing Simulation Times of SPS Slow Extraction using MAD-X. 2015. URL: <http://cds.cern.ch/record/2039579>.
- [44] M. Pari, F. M. Velotti, M. A. Fraser, V. Kain, and O. Michels. Characterization of the slow extraction frequency response. *Phys. Rev. Accel. Beams*, 24(8):083501, 2021. DOI: 10.1103/PhysRevAccelBeams.24.083501.
- [45] P. Actis, T. Dorenbos, and C. Johnson. An optically coupled differential beam time structure monitor for slow extraction. Technical report, CERN, Geneva, 1976. URL: <https://cds.cern.ch/record/2831368>.
- [46] J. Wozniak and C. Roderick. NXCALS - Architecture and Challenges of the Next CERN Accelerator Logging Service. page WEPHA163, 2020. DOI: 10.18429/JACoW-ICALEPCS2019-WEPEHA163.
- [47] M. Plum, E. Bravin, J. Bosser, and R. Maccaferri. N2 and xe gas scintillation cross-section, spectrum, and lifetime measurements from 50mev to 25gev at the cern ps and booster. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 492(1):74–90, 2002. DOI: 10.1016/S0168-9002(02)01287-1.
- [48] K. Okamura, Y. Arakaki, T. Kimura, S. Murasugi, R. Muto, Y. Shirakabe, M. Tomizawa, and E. Yanaoka. A consideration on the transfer function between rq field and slow extraction spill in the main ring of j-parc. *Proceedings of the 10th Int. Particle Accelerator Conf., IPAC2019:Australia*, 2019. DOI: 10.18429/JACoW-IPAC2019-WEPM008.
- [49] S. Sorge, P. Forck, and R. Singh. Measurements and simulations of the spill quality of slowly extracted beams from the sis-18 synchrotron. *Journal of Physics: Conference Series*, 1067(5):052003, sep 2018. DOI: 10.1088/1742-6596/1067/5/052003.

- [50] M. Pullia. Transit time for third order resonance extraction. Technical report, CERN, Geneva, 1996. URL: <https://cds.cern.ch/record/313678>.
- [51] P. J. Bryant. Resonant extraction. 2018. DOI: 10.23730/CYRSP-2018-005.315.
- [52] R. Singh, P. Forck, and S. Sorge. Reducing fluctuations in slow-extraction beam spill using transit-time-dependent tune modulation. *Phys. Rev. Appl.*, 13:044076, Apr 2020. DOI: 10.1103/PhysRevApplied.13.044076.
- [53] R. Cappi and C. Steinbach. Low frequency duty factor improvement for the CERN PS slow extraction using rf phase displacement techniques. *IEEE Trans. Nucl. Sci.*, 28:2806–2808, 1981. DOI: 10.1109/TNS.1981.4331919.
- [54] S. Sorge, P. Forck, and R. Singh. Spill ripple mitigation by bunched beam extraction with high frequency synchrotron motion. *Phys. Rev. Accel. Beams*, 26:014402, Jan 2023. DOI: 10.1103/PhysRevAccelBeams.26.014402.
- [55] C. Roderick and R. Billen. The LSA Database to Drive the Accelerator Settings. Technical report, CERN, Geneva, 2009. URL: <https://cds.cern.ch/record/1215575>.
- [56] V. Nagaslaev, J. Amundson, J. Johnstone, C. Park, and S. Werkema. Third integer resonance slow extraction using rfko at high space charge. 2011. DOI: 10.48550/arXiv.1209.5967.

## Appendix A

### Code Listings

Listing A.1: Python code to calculate angle kick produced by the TFB system (??).

```
1 from scipy.constants import physical_constants as const
2 import numpy as np
3
4 p = 5.392 # GeV/c, proton momentum for the cycle
5 Brho = p*3.3356 # Tm, magnetic rigidity
6
7 def kick_angle(voltage):
8     """
9         Calculates the angular kick to a proton beam of rigidity 'Brho' [Tm]
10        from a pair of TFB electrodes with voltage 'voltage' [V].
11    """
12
13    c = const["speed of light in vacuum"][0] # [m/s] speed of light
14    E0 = const["proton mass energy equivalent in MeV"][0]*1e6 # eV
15    mu0 = const["vacuum mag. permeability"][0] # H/m, vacuum permeability
16
17    T = (np.sqrt(p**2+E0/1e9**2)-E0/1e9)*1e9 # eV, proton KE
18    P = 5e3 # W, TFB peak power / electrode
19    Z = 100 # Ohm, TFB impedance / electrode
20
21    L = 935e-3 # m, TFB length
22    r = 70e-3 # m, TFB separation
23
24    E = T + E0 # Total Energy
25    gamma = E / E0 # Normalized energy (lorentz)
26    beta = np.sqrt(1-gamma**-2) # normalised velocity (lorentz)
27    cp = p*1e9 # eV, particle momentum
28
29    # Electric Field
30    Vp = np.sqrt(P * Z * 2) # peak voltage
```

---

```

30     V = Vp * voltage
31     Efield = V / r # adjusted for gain
32
33     # Magnetic Field
34     I = np.sqrt((V**2/Z)/Z**2) # current
35     Hfield = (2*I) / (2*np.pi*r) # adjusted for gain
36     Bfield = Hfield * mu0
37
38     # Angle
39     theta_M = c/1e9*Bfield*L / (cp/10**9)
40     theta_E = Efield/1e9*L / (cp/10**9 * beta)
41     theta = theta_E + theta_M
42     return theta

```

Listing A.2: Python code to generate a single chirp signal as produced by the Arbitrary Function Generator (??).

```

1 import numpy as np
2 import scipy.signal as sig
3
4 frev = 470000
5
6 def generate_afg_chirp(
7     chirp_length_time = 0.001,
8     turn_freq = frev,
9     chirp_middle = frev*0.325,
10    chirp_dev = frev*0.025,
11    sampling_freq = 1e9):
12    """
13        Generates a frequency-modulated single chirp signal, centered on ‘
14        chirp_middle’ [Hz], with a deviation of ‘chirp_dev’ [Hz], and a length
15        of ‘chirp_length_time’ [s].
16    """
17
18    chirp_time_samples = np.arange(0, chirp_length_time, 1/sampling_freq)
19    chirp_signal_samples = sig.chirp(chirp_time_samples,          # time array
20                                    chirp_middle-chirp_dev,   # start
21                                    frequency                 # end time
22                                    chirp_length_time,         # end
23                                    chirp_middle+chirp_dev)   # end
24
25    chirp_lasts_for_turns = chirp_length_time * turn_freq
26
27    return chirp_time_samples, chirp_signal_samples, chirp_lasts_for_turns

```

---

Listing A.3: Python code used to calculate the laboratory time of particle extraction and reconstruct the time-series spill array (??).

```
1 t = (particles.at_turn*t_rev) - particles.zeta/rel_beta/sp.constants.c
2
3 max_time = t.max()
4 sample_rate = 18000
5 total_samples = int(max_time * sample_rate) + 1
6
7 time_series_data = np.zeros(total_samples)
8 arrival_indices = np.floor(t * sample_rate).astype(int)
9 for idx in arrival_indices:
10     time_series_data[idx] += 1
```