TECHNISCHE UNIVERSITÄT WIEN

Diploma Thesis

High-Precision Magnetic Field Prediction and Control in Synchrotrons

submitted in satisfaction of the requirements for the degree
Diplom-Ingenieur
of the TU Wien, Faculty of Physics

Diplomarbeit

Hochpräzise Magnetfeldvorhersage und -steuerung in Synchrotronen

ausgefuehrt zum Zwecke der Erlangung des akademischen Grads
Diplom-Ingenieur
eingereicht an der TU Wien, Fakultät für Physik

Anton Lu, M.Sc.Eng. M.Sc.

Matr.Nr.: 12231206

Betreuung: Prof. Dipl.-Ing. Dr.techn. **Michael Benedikt**
Atominstitut
Technische Universität Wien
Stadionallee 2, 1020 Wien, Österreich

Wien, im Date undefined _____

Kurzfassung

Abstract

Acknowledgements

Contents

Acknowledgements **6**

1 Introduction **19**
1.1 Motivation and Problem Statement . . . 19
1.2 High-Level Scope and Contributions . . . 19
1.3 Thesis Outline . . . 19

2 Background and Literature Review **20**
2.1 Accelerator Physics Fundamentals . . . 20
2.1.1 Transverse Beam Dynamics as Quasi-Harmonic Oscillators . . . 20
2.1.2 Nonlinear Transverse Beam Dynamics and Momentum Coupling . . . 21
2.1.3 Principal Magnet Types in Synchrotrons . . . 21
2.1.4 Betatron tune and beta function . . . 22
2.1.5 Dipole Field Errors and Closed Orbit Distortion . . . 23
2.1.6 Magnet errors and resonances . . . 23
2.1.7 Longitudinal Beam Dynamics . . . 24
2.1.8 Beam Injection . . . 25
2.1.9 Beam Extraction . . . 25
2.2 The CERN Accelerator Complex . . . 25
2.3 The CERN Super Proton Synchrotron . . . 26
2.3.1 Principal Machine Design . . . 26
2.3.2 Principal Magnet Families . . . 26
2.3.3 Beam Instrumentation and Diagnostics . . . 27
2.4 Normal Conducting Accelerator Magnets . . . 28
2.4.1 Magnetic Field Generation in Normal Conducting Magnets . . . 28
2.4.2 Physical Mechanisms of Hysteresis . . . 28
2.5 Hysteresis Models . . . 29
2.5.1 Operator-type Models . . . 30
2.5.2 ODE-type Models . . . 32
2.5.3 Machine Learning Approaches . . . 35
2.6 Eddy Current Decay . . . 36
2.6.1 Eddy Currents in Conducting Materials . . . 36
2.6.2 Eddy Currents in Synchrotron Components . . . 37
2.6.3 Beam Impact of Eddy Current Decay in Fast Cycling Synchrotrons . . . 39
2.6.4 Mitigation Strategies . . . 40

3 Accelerator Magnets and Magnetic Measurements **42**
3.1 Magnetic Field in the Magnet Aperture . . . 42
3.1.1 Field Quality Conventions . . . 43
3.1.2 Forbidden Harmonics and Magnet Symmetry . . . 43
3.1.3 Fringe Fields and the Full-Length Measurements . . . 44
3.2 Measurement Techniques . . . 44
3.2.1 Local, Integral, and Average Measurements . . . 45

3.2.2 Hall Probes . . . 45
3.2.3 NMR Probes . . . 47
3.2.4 Stretched Wire . . . 47
3.2.5 Fluxmeters . . . 48
3.2.6 Rotating Coils . . . 50
3.3 Magnetic Measurements at CERN . . . 51
3.3.1 SPS Main Dipole Measurement Bench . . . 51
3.3.2 SPS Main Quadrupole Measurement Bench . . . 51
3.4 Online B-Train Measurement Systems . . . 51
3.4.1 Operating Principle of a B-train . . . 51
3.4.2 B-Train Implementations . . . 52
3.4.3 SPS B-Train . . . 53

4 Beam Operation at CERN and the SPS 54
4.1 Accelerator Operation in Cycles . . . 54
4.2 Main SPS Cycles . . . 56
4.2.1 SPS North Area Fixed Target Proton Cycle (SFTPRO) . . . 56
4.2.2 Remaining Cycle Families . . . 57
4.3 Typical SPS Supercycles . . . 58
4.3.1 Proton run supercycles . . . 58
4.3.2 Ion run supercycles . . . 60
4.4 Accelerator Controls Infrastructure at CERN . . . 62
4.4.1 LSA . . . 63
4.4.2 Timing . . . 65
4.4.3 Middleware . . . 66
4.4.4 Logging . . . 66
4.4.5 UCAP . . . 67
4.4.6 FGCs . . . 68
4.5 Magnetic Field Control in the CERN Injectors . . . 70
4.5.1 Hierarchical Field Control and Field Regulation in the PSB and PS . . . 70
4.5.2 SPS B-Train in Operation . . . 70
4.6 Magnetic Precycle and Degaussing in Operation . . . 71
4.6.1 Dedicated Magnetic Precycle in the SPS (MD1) . . . 71
4.6.2 Quasi-Degaussing of Higher Order Magnets in the SPS . . . 72
4.6.3 Operational Limitations from Hysteresis . . . 72
4.7 Machine Development Time . . . 73

5 Problem Statement and Proposed Approach 74
5.1 Hysteresis as an Operational Problem . . . 74
5.2 Motivating Evidence . . . 77
5.2.1 Slow Extraction Perturbation from Supercycle Changes . . . 77
5.2.2 Orbit Displacement as Result from Dipole Errors . . . 79
5.2.3 Field Response and Hysteresis Quantification in the Main Dipoles . . . 81
5.2.4 Eddy Current Decay in the SPS Main Dipoles . . . 83
5.3 Problem Statement . . . 84
5.4 Proposed Approach . . . 85
5.4.1 Beam-Based Approaches . . . 86
5.4.2 Data-Driven Modeling . . . 86
5.4.3 Phenomenological Modeling Limitations . . . 87

5.4.4 Residual Modeling with Machine Learning . . . 88
5.4.5 Data Requirements and Measurements . . . 90
5.4.6 Feed-forward Compensation in the Control System . . . 91
5.4.7 Compensation Validation Strategy . . . 93

6 Field Measurement Campaign 95
6.1 Dipole Laboratory Measurements . . . 95
6.1.1 Operational Sequences Measurements . . . 96
6.1.2 Hall Probe and Measurement Averaging . . . 101
6.1.3 Power Converter Limitations . . . 105
6.1.4 Limitations and Data Transferability . . . 108
6.2 B-Train measurements . . . 109
6.2.1 Field Drift and Reproducibility . . . 110
6.2.2 Marker-Based Drift Correction . . . 113
6.2.3 Residual Drift on Long Cycles . . . 114
6.2.4 Data Acquisition and Limitations . . . 116

7 Modeling Hysteresis 118
7.1 Modeling Requirements . . . 118
7.2 Dataset and Target . . . 119
7.2.1 Dataset Composition . . . 119
7.2.2 Signal Conditioning and Sampling . . . 119
7.2.3 Prediction Task and Residual Target . . . 120
7.2.4 Training, Validation, and Test Split . . . 122
7.3 Preliminary Studies . . . 122
7.3.1 Gray-Box and Physics-Parametrized Models . . . 122
7.3.2 Black-Box Baselines . . . 124
7.4 Architecture Selection . . . 125
7.4.1 Candidate Architectures . . . 125
7.4.2 Model Architecture . . . 126
7.5 Training . . . 127
7.5.1 Simulated Pretraining Data . . . 128
7.5.2 Pretraining and Fine-Tuning . . . 129
7.6 Results . . . 131
7.6.1 Ablations . . . 131
7.6.2 Controlled Study . . . 132
7.7 Limitations and Considerations . . . 133

8 Field Compensation and Operational Validation 136
8.1 Compensation Scheme Design . . . 136
8.1.1 The Control Problem . . . 136
8.1.2 Settings Management Integration . . . 137
8.1.3 Timing System Constraints . . . 138
8.1.4 Control Loop Implementation . . . 139
8.1.5 Autoregressive Operation . . . 141
8.1.6 Extensibility to Other Magnet Families . . . 142
8.2 Beam Validation . . . 142
8.2.1 Deployment on the Fixed Target Cycle . . . 142
8.2.2 Sensitivity to Field and Tune Perturbations . . . 143

8.2.3 Dipole Field Compensation at the SFT Flat Top . . . 145
8.2.4 Extended Operational Test . . . 149
8.2.5 Autoregressive Operation . . . 150
8.2.6 Scope of Validated Results . . . 151

9 Eddy Current Decay: A Parallel Case Study 154
9.1 Observations from Measurements . . . 154
9.2 Eddy Current Models . . . 157
9.3 Measurements of Eddy Current Decay . . . 159
9.4 Fitting the Eddy Current Model . . . 160
9.5 Compensation Design and Results . . . 165
9.5.1 A Complementary Compensation Scheme . . . 165
9.5.2 Operational Validation . . . 167
9.6 Discussion . . . 170
9.6.1 Orbit Decay Above Transition . . . 170
9.6.2 Feasibility of the Fitted Model . . . 172
9.6.3 Comparison of the Beam-Fitted and B-Train Models . . . 173
9.6.4 Higher Order Magnets . . . 174

10 Operational Simplification of the Precycle 176
10.1 Isolating the precycle contribution . . . 176
10.2 Dipole precycle . . . 177
10.3 Quadrupole precycle . . . 178
10.4 Function and simplification of the precycle . . . 180

11 Discussion 183
11.1 Attainment of the Stated Goals . . . 183
11.2 Measurement Reproducibility and Identifiability . . . 184
11.3 Data Variety and Generalization . . . 185
11.4 Eddy-Current Decay as a Distinct Effect . . . 185
11.5 Scope of the Learned Model . . . 186
11.6 Control-Loop and Deployment Limitations . . . 187
11.7 Planned Feedback in the B-train . . . 188
11.8 Extension to the Quadrupole and Higher-Order Families . . . 188
11.8.1 Measurement-Fidelity Limitations . . . 189
11.8.2 Proportional Hysteresis Excitation in Dipoles and Quadrupoles . . . 190
11.8.3 Seeding Without an Online Measurement . . . 190
11.9 Outlook . . . 192

12 Conclusions 193
12.1 Summary of Contributions . . . 193
12.2 Future Work . . . 195
12.3 Concluding Remarks . . . 195

List of Figures

2.1 Idealized transverse magnetic field lines in the magnet aperture for a dipole, quadrupole, and sextupole. . . . 22
a Dipole . . . 22
b Quadrupole . . . 22
c Sextupole . . . 22
2.2 Idealized major B–H hysteresis loop . . . 29
2.3 Illustration of a hysteron operator $R_{\alpha\beta}$ which activates when the input signal crosses the upper threshold α, and deactives when the signal decreases below β. The units of the illustrated input signal are arbitrary as the hysteron operator is scale-invariant, and the output is a binary state of either $+1$ or -1. . . . 31

4.1 SFTPRO machine cycle . . . 57
a Dipole field and intensity . . . 57
b Higher-order circuits . . . 57
4.2 Standard fixed-target supercycle . . . 59
4.3 SPS supercycle during LHC filling . . . 59
4.4 LHC-filling supercycle with pilot cycle . . . 60
4.5 SPS LHC pilot-beam supercycle . . . 60
4.6 SPS AWAKE supercycle . . . 61
4.7 SPS HiRadMat pilot supercycle . . . 61
4.8 SPS HiRadMat physics supercycle . . . 62
4.9 SPS supercycle with an MD cycle . . . 62
4.10 SPS SHiP supercycle . . . 63
4.11 LSA settings hierarchy . . . 64
4.12 LSA calibration functions . . . 65
a Quadrupole gradient calibration . . . 65
b Dipole field calibration . . . 65
4.13 Controls middleware architecture . . . 67
4.14 UCAP converter example . . . 67
4.15 Dynamic economy mode . . . 69
a Nominal and economy cycle . . . 69
b Economy ramp-down after injection . . . 69
c Economy end phase . . . 69
4.16 B-Train field marker placement . . . 71
a SFT cycle, 14 GeV/c injection . . . 71
b LHC cycle, 26 GeV/c injection . . . 71
4.17 MD1 magnetic precycle . . . 72
a Dipole field and quadrupole current . . . 72
b Sextupole and octupole reference currents . . . 72

5.1 Distribution of SPS run time across beam destinations, 2023–2025 . . . 76

5.2 Supercycle change from SFT only to an LHC filling sequence 78
a Dipole field evolution 78
b Relative change in dipole field 78
5.3 Supercycle change from an LHC filling sequence to SFT only 79
a Dipole field evolution 79
b Relative change in dipole field 79
5.4 SFT flat-top field decay after transition to fixed-target only 79
5.5 SFTPRO spill macrostructure change during supercycle sequence changes 80
a SFT only to LHC filling sequence 80
b LHC filling sequence to SFT only 80
5.6 Measured orbit displacement for field perturbations on a Q20 MD cycle 81
5.7 Measured dipole field of all operational cycles with the B(I) transfer function . . 82
5.8 Measured dipole field of SFTPRO cycles with the B(I) transfer function 82
5.9 Dipole field measurements plotted as the I to B/I transfer function 83
5.10 Main dipole field decay over a 3.6 s idle period after an LHCPILOT ramp-down . 84
5.11 Proposed feed-forward compensation scheme 92

6.1 Measurement of an operational SHIP and LHC filling sequence on the laboratory bench 97
6.2 SHIP cycle integral field measurements drifting at constant current 98
a Overlaid per-cycle integrated field 98
b Zoomed injection and cycle-end insets 98
6.3 Measured and programmed current at the flat top of one SHIP cycle 100
6.4 SHIP cycle integral field averaged over 16 aligned measurements 102
a Field averaged over 16 measurements 102
b Zoomed injection and cycle-end insets 102
6.5 Averaged Hall probe and fluxmeter field of an LHCPILOT cycle 103
6.6 Averaged Hall field and current at the SHIP flat top across cycle occurrences . . 104
6.7 Mean and standard deviation of SHIP flat-top current across 16 measurements . 106
6.8 Drift of the SHIP flat-top current over seven hours 106
6.9 Time-weighted RMS current during an SFTPRO supercycle against the converter limit 107
6.10 Transfer function of the laboratory Hall probes and the B-train 108
6.11 Difference between programmed and measured current over 95 SFTPRO occurrences 110
6.12 B-train flat-bottom field of SFTPRO cycles on two measurement days 111
a Flat injection plateau, one day 111
b Ramping injection plateau, another day 111
6.13 B-train flat-top field over 125 SFTPRO cycles at steady state 112
6.14 Drift-corrected B-train flat-bottom field of SFTPRO cycles 114
6.15 B-train field over the LHC1 filling cycle with injection-plateau drift 115

7.1 Measured field residual by cycle type after calibration subtraction 121
7.2 Neural-network-parametrized Preisach architecture 123
7.3 Adapted TFT encoder-decoder architecture 127
7.4 Training and validation MSE, scratch training against transfer 130
a Training MSE. 130
b Validation MSE. 130
7.5 Fine-tuned field prediction over the validation sequence 133

8.1 Extended LSA settings hierarchy with the BHYS field-correction program 137
8.2 Data flow of the feed-forward compensation loop over one cycle 140
8.3 Sensitivity of the SFT slow extraction to main dipole field and tune perturbations 144
a Main dipole field offset by -6×10^{-4} T–6×10^{-4} T. 144
b Horizontal tune offset by -6×10^{-3}–6×10^{-3}. 144
8.4 Spill BCT changes under controlled main dipole field perturbations 145
8.5 Spill quality factor and reproducibility error across SFT and LHC supercycle changes 146
8.6 Reproducibility error of the SFT spill across operational supercycles 146
8.7 Normalised spill BCT for the compensated and uncompensated cases across operational supercycles 147
8.8 Reproducibility error across the asymmetric 2×SFT plus LHCPILOT supercycle 148
8.9 Reproducibility error over the 35-hour extended operational test 149
8.10 Reproducibility error during the autoregressive compensation test 150
8.11 Single-shot SFT reproducibility with the preceding precycle flattened 152

9.1 Orbit decay across constant-current injection plateaus of various SPS cycles . . . 156
9.2 First-turn radial trajectory decay of different SPS cycles 160
9.3 Programmed dipole current, beam energy, and field derivative across the injection plateaus 161
9.4 Pareto front of the multi-objective eddy current model optimisation 164
9.5 Eddy current correction on the injection plateau of a short 26 GeV/c MD cycle . 167
9.6 Operational eddy current feed-forward on the SFT injection plateau with the precycle removed 168
a Beam response and applied correction 168
b Measured B-Train current and field 168
9.7 Closed-orbit decay for increasing numbers of inserted ZERO cycles, below and above transition 170
a Ion fixed-target cycle 170
b MD cycle at 26 GeV/c 170
9.8 Predicted eddy current field over the 26 GeV/c MD cycle for inserted ZERO cycles 171
9.9 Predicted eddy current field of the beam-fit and B-Train models over an MD supercycle 173
9.10 Eddy current decay in the main quadrupole circuit, from tune and bench gradient 174
a Tune change across the plateau 174
b Bench gradient decay 174

10.1 Tune against momentum offset on the 14 GeV/c MD cycle with and without the dipole precycle 179
a Measured tune against momentum offset with a 3-polynomial fit. 179
b The same measurement with the linear component removed, showing only the nonlinear chromaticity. 179
10.2 Evolution of linear chromaticity along the injection plateau with and without the dipole precycle 179
10.3 Tune along the injection plateau with and without the quadrupole precycle . . . 180
10.4 Main-dipole current of the MD1 precycle and short parallel MD cycle before and after flattening 181
a MD1 precycle before and after flattening 181
b Short parallel MD cycle before and after flattening 181

10.5 Supercycles after flattening the precycle and short parallel MD cycles 182
a Standard physics supercycle with the flattened MD1 precycle 182
b Standard parallel MD supercycle with the flattened MD cycle 182

11.1 Power converter voltage compared with the main dipole field 191

List of Tables

2.1 Principal parameters of the European Organization for Nuclear Research (CERN) Super Proton Synchrotron (SPS). 27

3.1 Allowed normal harmonics b_n for the principal SPS magnet families. All other harmonics are forbidden and vanish in a geometrically perfect magnet. 43
3.2 Classification of magnetic measurement techniques by the field quantity they provide directly. 45

4.1 Main SPS operational cycle families 57

5.1 Estimated MD1 energy consumption, cost, and share of SPS run time 76
5.2 Theoretical radial displacement at injection from a uniform dipole field error . . 80

7.1 Model and training hyperparameters. 129
7.2 RMSE by model and evaluation mode 132

9.1 Eddy current decay time constants for the B-Train and laboratory coil 155
9.2 Injection parameters of the measured SPS cycles 156
9.3 Fitted time constants and couplings of the eddy current model for one to three modes 163

10.1 Change in linear chromaticity when the dipole precycle is deactivated 178
a 450 GeV/c Large Hadron Collider (LHC) cycle ahead of the precycle . . . 178
b 400 GeV/c SPS Fixed Target (SFT) cycle ahead of the precycle 178

Acronyms

AD Antiproton Decelerator

ALPS A Logarithmic Position System

AWAKE Advanced Wakefield Experiment

BBQ Base-Band Tune (Q) Measurement System

BCT Beam Current Transformer

BLM Beam Loss Monitor

BP Basic Period

BPM Beam Position Monitor

CERN European Organization for Nuclear Research

CMW Controls MiddleWare

CNAO National Center for Oncological Hadrontherapy

COSE Constant Optics Slow Extraction

DBS Dynamic Beam Scheduling

DCCT Direct Current Transformer

DMS Direct Multi-Step

ELENA Extra Low Energy Antiproton Ring

FDI Fast Digital Integrator

FEC Front-End Computer

FESA Front-End Software Architecture

FFMM Flexible Software Framework for Magnetic Measurements

FGC Functional Generator / Controller

FODO Focusing-Defocusing

FORC First Order Reversal Curve

GMT General Machine Timing

GRN Gated Residual Network

HiRadMat High Radiation to Materials

HL-LHC High-Luminosity Large Hadron Collider

JA Jiles-Atherton

LEIR Low Energy Ion Ring

LEP Large Electron-Positron Collider

LHC Large Hadron Collider

LINAC Linear Accelerator

LLG Landau-Lifshitz-Gilbert

LS Long Shutdown

LSA LHC Software Architecture

LSS Long Straight Section

LSTM Long Short-Term Memory

MD Machine Development

MLP Multilayer Perceptron

MTE Multi-Turn Extraction

NAFF Numerical Analysis of Fundamental Frequencies

NARX Nonlinear Autoregressive Network with Exogenous Inputs

NMR Nuclear Magnetic Resonance

NXCALS NeXt Controls Accelerator Logging Serviceg

ODE Ordinary Differential Equation

PCB Printed Circuit Board

PLEP Parabolic - Linear - Exponential - Parabolic

PLP Parabolic - Linear - Parabolic

PPM Per-Pulse Modulation

PS Proton Synchrotron

PSB Proton Synchrotron Booster

QD Quadrupole Defocusing

QF Quadrupole Focusing

RDA Remote Device Access

RDP Ramer-Douglas-Peucker

RF Radio Frequency

RL Reinforcement Learning

SFT SPS Fixed Target

SFTION SPS Fixed Target Ion

SFTPRO SPS Fixed Target Proton

SHiP Search for Hidden Particles

SPS Super Proton Synchrotron

TFT Temporal Fusion Transformer

UCAP Unified Control, Acquisition and Processing

WCM Wall Current Monitor

WRT White Rabbit Timing

Chapter 1

Introduction

1.1 Motivation and Problem Statement

1.2 High-Level Scope and Contributions

1.3 Thesis Outline

Chapter 2

Background and Literature Review

This chapter provides an overview of the fundamental accelerator physics principles relevant to synchrotron accelerators, focusing on the impact of magnetic field errors and hysteresis on beam dynamics and stability. The chapter begins with a review of basic transverse and longitudinal beam dynamics, including the role of dipole, quadrupole, and sextupole magnets in shaping the beam trajectory and focusing properties. The following chapters build on this foundation to develop challenges in beam operation at ac CERN and the Super Proton Synchrotron (SPS), the challenges of attaining high-precision magnetic fields, and methods to mitigate these effects, such as using surrogate models for correction.

2.1 Accelerator Physics Fundamentals

Synchrotron accelerators accelerate charged particles to high energies through the combined action of Radio Frequency (RF) electric fields and precisely controlled magnetic fields. The magnetic field configuration in a synchrotron, also known as a lattice, serves two principal functions: establishing and maintaining an equilibrium circular orbit for the particles, and providing transverse focusing to keep the beam within acceptable apertures. Furthermore, additional, more specialized magnets may be used to stabilize the beam. This section outlines the physics principles governing beam dynamics in synchrotrons, emphasizing why small deviations in magnetic field strength, at the level of $\mathcal{O}(10^{-4})$ relative to the main field, can produce unacceptable changes in beam parameters and beam quality.

2.1.1 Transverse Beam Dynamics as Quasi-Harmonic Oscillators

Modern synchrotrons employ a separated-function design in which distinct magnet families provide the guiding bending and focusing forces. Dipolar magnets provide the main bending field to establish a closed orbit along the reference trajectory, satisfying the relation:

$$B\rho = \frac{p}{q} \tag{2.1}$$

where B is the magnetic field strength perpendicular to the beam trajectory, ρ is the bending radius of the ring, p is the particle momentum, and q is the particle charge. The relation $B\rho$ is also known as the magnetic rigidity of the beam.

For particles with small transverse displacements from the design orbit, dedicated quadrupole magnets provide the additional focusing forces required for stable motion around the ring. In the absence of coupling between planes, the horizontal and vertical motions are treated independently, and each follows Hills equation, which in the linear approximation for small displacement x takes the form

$$\frac{\mathrm{d}^2x}{\mathrm{d}s^2} + K(s)x = 0 \tag{2.2}$$

where s is the azimuthal coordinate around the ring and $K(s)$ is the focusing function set by the quadrupole gradient and local bending geometry, with an analogous equation for the vertical plane. This quasi-harmonic transverse motion about the local equilibrium orbit is termed betatron oscillation, whose amplitude envelope is the beta function $\beta(s)$ and whose number of oscillations per turn defines the betatron tune, both introduced in Section 2.1.4.

2.1.2 Nonlinear Transverse Beam Dynamics and Momentum Coupling

Equation 2.2 assumes that the particle momentum equals the design momentum. In a bunched beam, the particles carry a small relative momentum deviation $\delta = \Delta p/p$. Through Equation 2.1, a momentum deviation changes the curvature required to follow the reference orbit and produces an equilibrium horizontal offset in bending regions, given by the dispersion function $D(s)$,

$$x(s) = x_\beta(s) + D(s)\,\delta, \tag{2.3}$$

where $x_\beta(s)$ is the betatron oscillation about the local equilibrium orbit, and $D(s)$ is the periodic solution of the inhomogeneous equation of motion, generated only in regions of non-zero bending radius $\rho(s)$ and fixed by the same focusing lattice that sets the betatron motion.

A particle with momentum deviation δ also experiences a focusing strength $k(s) \propto 1/(B\rho)$ that differs from the design value, so its betatron tune shifts. This momentum dependence of the tune is quantified by the chromaticity ξ,

$$\xi_x = \frac{\mathrm{d}Q_x}{\mathrm{d}\delta}, \quad \xi_y = \frac{\mathrm{d}Q_y}{\mathrm{d}\delta}, \tag{2.4}$$

and in synchrotrons by the normalized chromaticity

$$Q' = \frac{\xi}{Q}, \tag{2.5}$$

which gives the fractional tune shift per unit relative momentum deviation and is comparable across machines with different tunes. For a lattice with only bending and focusing elements, the natural chromaticity is typically non-zero, spreading the betatron tune across the beam momentum distribution. This spread can cause beam loss where it overlaps resonance frequencies, and the sextupole magnets in Section 2.1.3 correct it to first order by providing momentum-dependent focusing. Section **??** discusses the implications of Q' for transverse beam stability.

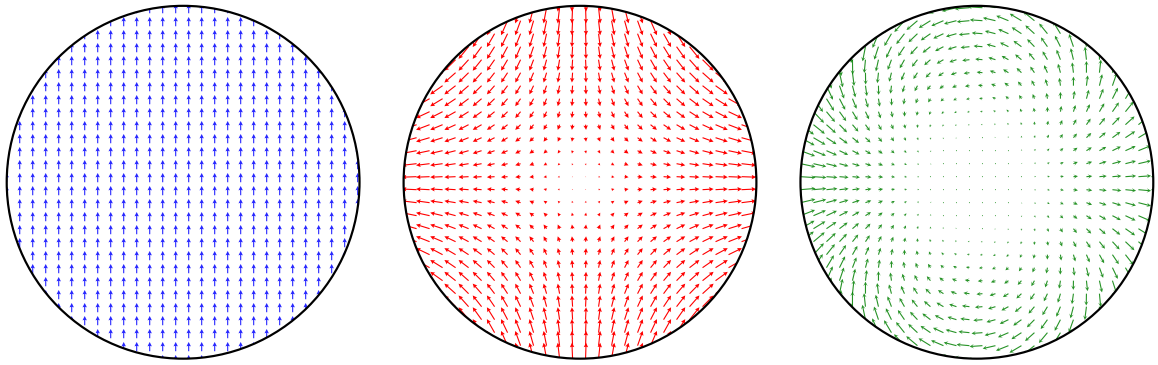
2.1.3 Principal Magnet Types in Synchrotrons

The transverse beam dynamics introduced above are realized by accelerator magnets, each optimized for a specific field component and discussed in detail in Chapter 3. The ideal transverse field in the mid-plane $y = 0$ of a synchrotron magnet is a series expansion,

$$B_y(x,0) = B_y(0,0) + \left.\frac{\partial B_y}{\partial x}\right|_0 x + \frac{1}{2!}\left.\frac{\partial^2 B_y}{\partial x^2}\right|_0 x^2 + \frac{1}{3!}\left.\frac{\partial^3 B_y}{\partial x^3}\right|_0 x^3 + \cdots . \tag{2.6}$$

Factoring out the ideal dipole field on the reference orbit, denoted B_0 and equal to the dipole field B in Equation 2.1, defines the multipole coefficients b_n,

$$B_y(x,0) = B_0 \sum_{n=1}^{\infty} b_n\, x^{n-1}, \tag{2.7}$$



(a) Dipole **(b)** Quadrupole **(c)** Sextupole

Fig. 2.1: Idealized transverse magnetic field lines in the magnet aperture for a dipole, quadrupole, and sextupole.

where b_1, b_2, and b_3 are the dipole, quadrupole, and sextupole terms. Identifying the two expansions, the normalized strengths k_{n-1} used in transverse beam dynamics follow from the on-axis derivatives as

$$k_{n-1} = \frac{1}{B\rho}\,\frac{1}{(n-1)!}\,\frac{\partial^{n-1}B_y}{\partial x^{n-1}}\bigg|_0 = \frac{B_0}{B\rho}b_n \tag{2.8}$$

which removes the dependence on particle momentum through the rigidity $B\rho$ and links the field-measurement coefficients b_n of Chapter 3 to the focusing strengths in the equations of motion. Since every k_{n-1} scales with the main dipole field B_0 and is normalized by the same rigidity $B\rho$, the effective strength seen by the beam changes whenever the main dipole field differs from its design value, even at fixed higher-order excitation.

The three principal magnet types follow from the leading terms of this expansion, and Figure 2.1 shows their idealized aperture field lines. A dipole (b_1) generates a uniform transverse field that bends the reference trajectory, with normalized strength $k_0 = B_0/(B\rho) = 1/\rho$ set by the bending radius. A quadrupole (b_2) produces a field linear in displacement, with normalized strength $k_1 = (B_0/B\rho)\,b_2$ from Equation 2.8, and focuses in one transverse plane and defocuses in the other, so alternating quadrupoles give net strong focusing around the ring. A sextupole (b_3) produces a field quadratic in displacement, with normalized strength $k_2 = (B_0/B\rho)\,b_3$, and corrects chromaticity by relying on non-zero dispersion so off-momentum particles sample an effective quadrupole component. Sextupoles also introduce amplitude-dependent tune shifts and can drive the nonlinear resonances of Section 2.1.6, so their strength and reproducibility affect dynamic aperture and beam lifetime.

2.1.4 Betatron tune and beta function

The beta function $\beta(s)$ is the amplitude envelope of the betatron oscillations, set solely by the magnetic elements of the lattice, which also fix the dispersion function $D(s)$ in Section 2.1.2. As the beam circulates, the transverse oscillations accumulate a phase advance $\mu_{x,y} = \int_0^C \mathrm{d}s\,/\beta_{x,y}(s)$ over one revolution of circumference C, and the betatron tune is the number of oscillations per turn,

$$Q_{x,y} = \frac{\mu_{x,y}}{2\pi}. \tag{2.9}$$

2.1.5 Dipole Field Errors and Closed Orbit Distortion

A dipole magnet with a field error, either a deviation in field magnitude ΔB or misalignment, acts as an unintended deflecting kick. The deflection angle is $\theta = \Delta B\, L/(B\rho)$, where L is the dipole length. With small field errors on individual magnets, the small deflection drives oscillations of the closed orbit around the design orbit. For macroscopic field changes where each magnet experiences the same relative field deviation $\Delta B/B$, the closed-orbit response depends on the beam's longitudinal condition. The momentum compaction factor

$$\alpha_c = \frac{\Delta C/C}{\Delta p/p} = \frac{1}{C}\oint \frac{D(s)}{\rho(s)}\,\mathrm{d}s \tag{2.10}$$

is a measure of how path length changes with momentum, determined by the lattice design through dispersion and bending radius, and independent of beam parameters such as energy or intensity. A particle follows a dispersed closed orbit set by its momentum offset relative to the momentum that the instantaneous field would center on the design orbit. Denoting this offset by $\tilde{\delta} = \Delta p/p - \Delta B/B$, the mean radius responds through the beam rigidity of Equation 2.1 as

$$\frac{\Delta r}{R} = \alpha_c \left(\frac{\Delta p}{p} - \frac{\Delta B}{B}\right) \tag{2.11}$$

where R is the reference radius. Only the difference between the two terms enters. A momentum that tracks the field, $\Delta p/p = \Delta B/B$, gives $\tilde{\delta} = 0$ and no orbit change, and a momentum offset at fixed field gives $\Delta r/R = \alpha_c\, \Delta p/p$. Equation 2.11 reduces to a limiting form depending on how the beam momentum is constrained.

When the RF system is off, as for a coasting beam or during the first turns after injection, the beam momentum is fixed and $\Delta p/p = 0$. The orbit then shifts by

$$\Delta r = -\alpha_c R \frac{\Delta B}{B} \tag{2.12}$$

A field increase raises the bending strength and moves the beam to a smaller radius, so the displacement scales inversely with the beam rigidity of Equation 2.1 at fixed absolute field error. When the RF system holds a fixed frequency, as on an injection plateau or a flat top, the revolution frequency is pinned and the beam momentum adjusts so that the velocity and path-length changes cancel. Using $f = \beta c/2\pi R$ and the slip factor $\eta = \alpha_c - 1/\gamma^2$, the orbit responds as

$$\Delta r = \frac{1}{\gamma^2}\frac{\alpha_c}{\eta} R \frac{\Delta B}{B} \tag{2.13}$$

The remaining case is a beam that stays synchronous during a ramp, where the RF frequency tracks the rising field so that $\Delta p/p = \Delta B/B$. The momentum offset then vanishes, and the closed orbit remains centered, $\Delta r = 0$. This condition is not considered further, since the field errors of interest here act on a beam at fixed frequency or with the RF off.

2.1.6 Magnet errors and resonances

Accelerator magnets always contain small field errors from manufacturing imperfections, misalignments, and finite geometry, and as the beam circulates, these errors drive resonant growth of the betatron oscillation when the tune lies close to a resonance condition,

$$|m_x| Q_x + |m_y| Q_y = p, \tag{2.14}$$

for integers m_x, m_y with $|m_x| + |m_y| = p$ the order of the driving multipole. Plotted against the horizontal and vertical tunes, these conditions form the resonance lines of a tune diagram. Not all lines are equally strong, as the strength depends on the specific multipole error and its distribution around the ring, but the working point (Q_x, Q_y) must still be placed with margin from the low-order lines to avoid beam loss. The lattice design sets the integer part of the tune, and the fractional part is adjusted through the quadrupole strengths.

Resonances are mitigated by moving the fractional tune, by dedicated correction magnets that compensate for driving-field errors, by transverse dampers that suppress coherent oscillations through feedback, and by octupole magnets that introduce amplitude-dependent tune shifts to detune large-amplitude particles for Landau damping.

2.1.7 Longitudinal Beam Dynamics

A synchrotron must also control the longitudinal motion, governed by the interaction between the beam and the accelerating RF cavities that raise the beam energy turn by turn. The RF cavity maintains an oscillating field at frequency f_{RF} that satisfies the synchronism condition

$$f_{\mathrm{RF}} = h \cdot f_{\mathrm{rev}} = h \cdot \frac{v}{2\pi R}, \tag{2.15}$$

so that f_{RF} is an integer multiple h, the harmonic number, of the revolution frequency f_{rev}. More than one RF system may be used, for instance a fundamental at harmonic h and a second harmonic at $2h$ for additional longitudinal focusing. As the beam accelerates, the bending dipole field and the RF frequency increase together to preserve synchronism, so a mismatch between the actual and required RF frequency, whether from an RF frequency error or an unexpected bending field error, makes off-momentum particles drift in RF phase and degrades longitudinal beam quality.

Each particle in a bunched beam carries a slightly different energy and phase relative to the synchronous particle, giving rise to synchrotron oscillations in energy and phase at the synchrotron tune Q_s. Since $Q_s \ll 1$ while $Q_x, Q_y > 1$, the longitudinal and transverse motions occur on well-separated timescales and are usually treated independently, coupling only through chromatic effects.

2.1.7.1 Chromatic Effects and Beam Stability

The chromaticity Q' of Section 2.1.2 and Equation 2.5 has direct consequences for transverse stability. In a bunched beam, particles at the head and tail occupy different longitudinal positions and exchange them over the synchrotron period, and through the chromatic tune spread, this head-tail motion couples to the transverse wake fields excited by the beam in the resistive machine environment. This coupling drives the head-tail instability, in which transverse oscillation amplitudes grow on a timescale much shorter than the synchrotron period, leading to rapid beam loss. Suppressing the head-tail instability requires a chromaticity whose sign follows the slip factor η of Section 2.1.5, which changes sign at the transition energy γ_t. A small positive Q' is therefore maintained above transition ($\eta > 0$) and a small negative Q' below transition ($\eta < 0$). The chromaticity must nonetheless stay within a controlled range, since excessively large values increase the tune spread and push the working point onto the resonance lines of Section 2.1.6.

2.1.8 Beam Injection

The beam is injected into a synchrotron by transferring a bunched beam from an upstream machine through a transfer line and a septum onto the ring orbit, while the circulating beam passes on the other side of the septum blade [53, 58]. Capture without emittance growth or loss requires matching the injected momentum, RF phase, and betatron tune to the receiving ring, with the working point placed clear of the resonance lines of Section 2.1.6 and the incoherent space-charge tune shift at injection energy. At injection flat bottom, the beam rigidity is lowest, so an absolute dipole field error causes the largest relative cycle deviation and produces the largest closed-orbit displacement through Section 2.1.5.

2.1.9 Beam Extraction

Beam extraction transfers the circulating beam, or a fraction of it, into a downstream beamline. Fast extraction removes the whole beam in one or a few turns through a kicker and septum for machine-to-machine transfer, as used for SPS injection into the Large Hadron Collider (LHC). Slow extraction peels particles off over hundreds of milliseconds to seconds by driving them onto a resonance, delivering the quasi-continuous flux required by fixed-target experiments [10, 127]. Chromatic slow extraction is a variant of slow extraction that exploits the momentum dependence of the betatron tune, described below.

2.1.9.1 Chromatic slow extraction

Chromatic slow extraction, employed at the SPS [26, 76] and other fixed-target facilities, keeps the machine optics nominally constant through the spill, with no active tune or emittance manipulation during extraction. Dedicated sextupoles establish the third-integer resonance $3Q_x = p$ of Equation 2.14, forming a triangular separatrix in horizontal phase space whose stable area shrinks as the tune approaches the resonance. The working tune is offset from the resonance, so the chromaticity Q' maps each momentum slice $\delta = \Delta p/p_0$ to a distinct tune, and only a narrow resonant slice becomes unstable and drifts out to the electrostatic septum. As the dipole field evolves over the flat top, the mean radial position of the beam shifts through Equation 2.12, continuously displacing the resonant slice and sustaining a steady spill without active tune manipulation. At the SPS the scheme is operated with Constant Optics Slow Extraction (COSE) [76], co-scaling all machine settings with beam rigidity so that the separatrix orientation stays fixed in normalized phase space through the spill.

The extraction rate is set passively by the interplay of Q', the tune offset, and the radial position, with no turn-by-turn handle to compensate field errors. Errors in the sextupole, quadrupole, and dipole fields perturb the spill directly, through Q', through the tune relative to the resonance, and through the closed orbit at the septum, so the spill quality is a sensitive probe of relative field errors at the $\Delta B/B \sim 10^{-4}$ level and depends directly on the reproducibility of the lattice fields.

2.2 The CERN Accelerator Complex

The European Organization for Nuclear Research (CERN) accelerator complex comprises many accelerators, organized into two injector chains in which the machines are connected in series to reach progressively higher energies. The proton chain runs from Linac4 through the Proton Synchrotron Booster (PSB) and the Proton Synchrotron (PS) to the SPS, and the heavy-ion chain, used predominantly for lead, runs from Linac3 through the Low Energy Ion Ring (LEIR)

and the PS to the SPS. The beam delivered by these chains is used for the LHC and for a broad fixed-target physics programme.

Each synchrotron in a chain accelerates beam both for its own experimental areas and for the next machine downstream, with the initial linac injecting only into the first synchrotron. The SPS is the last synchrotron before the LHC, and is the subject of the following section.

2.3 The CERN Super Proton Synchrotron

The SPS [129] is the second-largest accelerator at CERN, housed in a 6.9 km tunnel under the Franco-Swiss border, and has operated since 1976 as a proton–antiproton collider [41], as the Large Electron-Positron Collider (LEP) lepton injector [31], and now as the injector for the LHC and the fixed-target programme. It accelerates protons and heavy ions such as lead [94], and is being upgraded to serve the High-Luminosity Large Hadron Collider (HL-LHC) and the Search for Hidden Particles (SHiP) experiment [3, 132].

The SPS receives beam from the PS at 14 GeV/c or 26 GeV/c for protons and 17 GeV/(c u) for heavy ions, and accelerates to 450 GeV/c for LHC injection or 400 GeV/c for North Area fixed-target physics. It is a multi-cycling synchrotron, accelerating beams of different intensity, particle type, energy, tune, and RF configuration in interleaved cycles for LHC filling and the fixed-target experiments, which in 2026 include Advanced Wakefield Experiment (AWAKE) [63], High Radiation to Materials (HiRadMat) [65], and the North Area experiments [14].

2.3.1 Principal Machine Design

The SPS is functionally divided into six sextants, each containing two arcs and a Long Straight Section (LSS), in a separated-function Focusing-Defocusing (FODO) lattice. Each FODO cell contains one focusing and one defocusing quadrupole separated by four dipoles over a length of about 60 m, and 108 cells make up the ring, with 16 per arc, 2 per LSS, and the remaining 8 in the straight sections for injection, extraction, and RF systems [129]. In total, the lattice holds 744 bending dipoles of the MBA and MBB designs, powered in two series circuits to the same field, and 216 focusing and 216 defocusing quadrupoles of a common design, powered with opposite polarity for the two planes. It also holds 216 main sextupoles for chromaticity control, as 108 focusing and 108 defocusing magnets of differing design, together with octupoles, skew quadrupoles, and dedicated extraction sextupoles for chromatic slow extraction.

Acceleration is provided by 200 MHz RF cavities in LSS3 at harmonic number 4620, giving a fundamental frequency near 43.3 kHz that varies little across the ramp since the beam is already ultra-relativistic at injection. A complementary 800 MHz system provides bunch rotation, longitudinal emittance control, and improved chromatic slow-extraction spill quality [21]. The machine accelerates proton beams of up to 3×10^{13} protons per batch and lead-ion beams of up to 1.5×10^{11} ions per batch, with injection emittances near 3 µm for protons and 1.5 µm for lead and a momentum spread of order 10^{-4}. Current operation uses integer tune 26 (Q26) for fixed-target slow extraction, injecting protons at 14 GeV/c and lead ions at 17 GeV/(c u), and integer tune 20 (Q20) for LHC-type beams injected at 26 GeV/c, though other integer tunes can be set. Table 2.1 summarizes the principal parameters.

2.3.2 Principal Magnet Families

Although originally specified for 300 GeV/c at a dipole field near 1.8 T, the SPS now reaches the 450 GeV/c required for LHC injection at 2.0 T and 5800 A, close to the practical limit of the normal-conducting design, with the lattice and magnet designs unchanged since the 1970s. All

Tab. 2.1: Principal parameters of the CERN SPS.

Parameter	Value	Unit
Circumference	6911.5	m
Injection momentum (protons)	14 & 26	GeV/c
Maximum extraction momentum	450	GeV/c
Injection momentum (heavy ions)	17	GeV/(c u)
Number of dipole magnets	744	–
Number of quadrupole magnets (F/D)	216/216	–
Number of sextupole magnets (F/D)	108/108	–
Total number of magnets	1317	–
Harmonic number	4620	–
RF frequency	200.2 & 800.8	MHz
Nominal betatron tune (Q_x/Q_y)	20.x / 20.x or 26.13 / 26.18	–

main magnet families operate at room temperature and are water-cooled through their copper coils. The main dipoles are evenly split into the MBA and MBB types, reaching the same integrated field at the same current but differing in yoke geometry and aperture, 153 mm wide by 39 mm high for MBA and 129 mm by 52 mm for MBB to match the differing beam sizes along the magnets, and of length 6.2 m and 6.22 m respectively. Their vacuum chambers are correspondingly rectangular with rounded corners and of different heights [105].

Transverse focusing is provided by two quadrupole families, the focusing Quadrupole Focusing (QF) and the defocusing Quadrupole Defocusing (QD), which share a common design of 88 mm square aperture and 3.05 m length, reaching 1.0 T at the pole tips for a gradient of 22 T/m, and are powered in separate circuits of opposite polarity for the two planes. Chromaticity is controlled by focusing and defocusing sextupole circuits, with dedicated sextupoles for slow extraction, and octupoles and skew quadrupoles correct nonlinearities and transverse coupling. These higher-order magnets are likewise iron-dominated and subject to the same effects as the dipoles and quadrupoles.

2.3.2.1 Magnet Powering and Control

The main magnets are powered in series circuits, the MBA and MBB dipoles forming one global dipole circuit fed by 12 power stations known as SMDs, with the quadrupole and higher-order families powered analogously. Each family is current-regulated to below 10 ppm by the power converters known at CERN as Functional Generator / Controllers (FGCs).

2.3.3 Beam Instrumentation and Diagnostics

The SPS carries a wide range of beam diagnostics, of which the Beam Position MonitorBeam Position Monitors (BPMs), Beam Current Transformers (BCTs), Base-Band Tune (Q) Measurement System (BBQ), and Beam Loss MonitorBeam Loss Monitors (BLMs) are relevant to this thesis [56]. The BPMs measure the transverse beam position around the ring and were upgraded to the A Logarithmic Position System (ALPS) system during Long Shutdown (LS)2 [20, 139], with 108 horizontal pickups beside the focusing quadrupoles and 108 vertical pickups beside the defocusing quadrupoles, resolving single-bunch position to about 290 µm horizontally and 150 µm vertically. Their turn-by-turn and averaged readings feed the closed-orbit feedback [24] and beam-based measurements of the optics and magnetic fields.

The BCTs are toroidal transformers that measure the circulating beam current, with fast AC transformers for bunch-by-bunch intensity and DC transformers for the average intensity [72, 131], and provide the intensity and spill monitoring used through injection, acceleration, and extraction. The measurement is non-interceptive and continuous, providing circulating beam intensity in real time; during slow extraction, the BCTs also records the time profile of the extracted spill.

The betatron tune is measured by the BBQ system, which picks up the small residual transverse oscillations of the beam and derives the fractional tune from their frequency spectrum [55]. From a turn-by-turn position record the dominant spectral line is extracted with the Numerical Analysis of Fundamental Frequencies (NAFF) algorithm [84], giving a precise frequency estimate over a short measurement window, and combined with a controlled momentum change the same measurement yields the chromaticity.

The BLMs detects beam losses through the secondary radiation produced when particles strike the beam pipe, and triggers beam dumps for machine protection when losses exceed preset thresholds [28, 140]. This work does not use other instruments, such as wire scanners and the Wall Current Monitor (WCM), for transverse and longitudinal profiles.

2.4 Normal Conducting Accelerator Magnets

In the SPS the dipole, quadrupole, and sextupole magnets are normal-conducting electromagnets, each a ferromagnetic yoke wound with resistive copper coils and operating at room temperature. The iron yoke gives the field a history-dependent and nonlinear response that underlies magnetic hysteresis and, under changing excitation, eddy-current effects, with the magnet design, tolerances, and field quality treated in Chapter 3.

2.4.1 Magnetic Field Generation in Normal Conducting Magnets

A normal-conducting magnet drives flux through its aperture with a resistive coil wound on the yoke, the excitation current I setting the field intensity $\mathbf{H}$ through Ampère's law,

$$\oint \mathbf{H} \cdot d\mathbf{l} = NI, \tag{2.16}$$

for N coil turns carrying current I. The flux density follows from the constitutive relation

$$\mathbf{B} = \mu_0 \left(\mathbf{H} + \mathbf{M}\right), \tag{2.17}$$

with μ_0 the vacuum permeability and $\mathbf{M}$ the magnetization. In this thesis, magnetic field refers to the flux density $\mathbf{B}$, the quantity acting on the beam, written as the scalar B except where direction matters.

In a linear material $\mathbf{M} = \chi_m \mathbf{H}$ and $\mathbf{B} = \mu_0 \mu_r \mathbf{H}$ with relative permeability $\mu_r = 1 + \chi_m$, but in the iron yoke μ_r depends nonlinearly on $\mathbf{H}$. The magnetization arises from microscopic domains that align with the applied field, and once all are aligned $\mathbf{M}$ reaches its saturation value M_s, beyond which the aperture field no longer grows linearly with current. When the field is reduced, the domains do not fully retrace their configuration, leaving a remanent magnetization set by the prior excitation history, the irreversibility that gives rise to magnetic hysteresis.

2.4.2 Physical Mechanisms of Hysteresis

Magnetic hysteresis makes the relationship between the applied field H, driven by the excitation current, and the flux density B non-single-valued, so the same H corresponds to different B

depending on whether the excitation is increasing or decreasing [19]. Cycling H traces a closed loop in the B–H plane, shown in Figure 2.2 and characterized by the remanent field B_r at $H = 0$ and the coercive field H_c at $B = 0$, with the enclosed area equal to the energy dissipated per cycle.

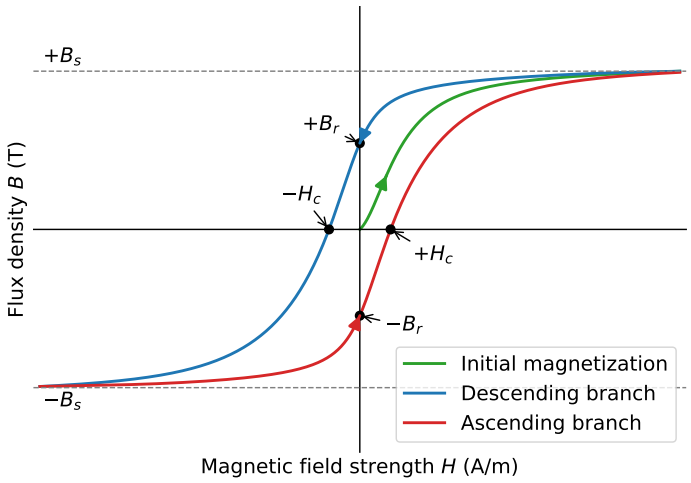


Fig. 2.2: Idealized major B–H hysteresis loop of a ferromagnetic material, showing the initial magnetization curve, saturation flux density $\pm B_s$, remanent flux density $\pm B_r$, and coercive field $\pm H_c$. The arrows indicate the direction of traversal as H is cycled.

Microscopically, the lag arises from domain walls moving through the yoke as H changes but pinned at lattice defects, grain boundaries, and impurities [69]. A wall advances only when the field overcomes a local pinning barrier and stays trapped at new sites on the return excursion, the irreversible motion known as the Barkhausen effect [15], so the magnetization encodes the full prior excitation history and no single-valued $B = f(H)$ exists across different histories. Resolving this from first principles through the micromagnetic Landau-Lifshitz-Gilbert (LLG) equation [57, 83] would demand a nanometre-scale discretization of the whole yoke, intractable for a metre-scale magnet, so the macroscopic B–H response has no first-principles prediction and is instead captured by phenomenological models identified from measurement.

Static hysteresis is a good approximation of rate independence, with the B–H loop set mainly by the sequence of field extrema and depending only weakly on the sweep rate. Rate-dependent contributions also exist, from thermally activated domain-wall relaxation within the hysteresis itself and, dominant in accelerator magnets, from eddy currents induced when the excitation changes rapidly, which add a transient field that decays once the excitation is held constant.

2.5 Hysteresis Models

As established in Section 2.4.2, the macroscopic B–H relationship of a ferromagnetic yoke is nonlinear, history-dependent, and not reducible to a single-valued function. Predictive modeling of this relationship therefore requires models that are capable of representing path-dependent behavior under arbitrary excitation histories. The models reviewed in this section operate at the macroscopic level, treating the yoke as a lumped element characterized by its bulk B–H response, rather than resolving the underlying domain microstructure. Classical phenomenological approaches fall into two broad classes: operator-type models that represent the hysteretic map as a weighted superposition of elementary bistable operators, and differential models that describe

magnetization evolution through ordinary differential equations driven by the applied field. More recently, data-driven machine learning models have been applied to hysteresis modeling, learning the input-to-output mapping directly from measurement data without an explicit physical parametrization.

Although hysteresis models are conventionally formulated in terms of the material quantities H and M, or equivalently H and B, the same mathematical structure applies when the excitation current I is used as the input and the flux density B in the magnet aperture as the output, since I and H are related through Ampère's law in Equation 2.16 and the geometry of the magnetic circuit. In this thesis, models are formulated directly in terms of I and B, as these are the quantities accessible through measurement.

Despite the diversity of available phenomenological models, no single model provides a universal solution that simultaneously achieves high accuracy, broad generalizability, and compact parameterization across arbitrary magnets and excitation histories. Each model class has structural assumptions about the form of the hysteretic response, whether it be a superposition of bistable operators, a low-dimensional differential equation, or a parametric curve. These assumptions typically match well only the class of excitation histories and material behaviors for which the model was designed [19, 97, 99]. Parameter identification is performed on a finite and necessarily incomplete set of measured cycles, so that a model calibrated on major loop data may systematically misrepresent the response under minor loop excitations or multi-reversal histories not represented in the identification dataset. Furthermore, the compactness that makes phenomenological models computationally attractive directly limits their representational capacity, and a model with few free parameters cannot, in general, capture the full complexity of the history-dependent magnetization state of a real iron-yoke magnet. The data-driven approaches reviewed in Section 2.5.3 replace these structural assumptions with flexible function approximators trained directly on measurement data.

2.5.1 Operator-type Models

Operator-type models represent the hysteretic input–output map as a functional acting on the full excitation history, rather than on the instantaneous input value alone. The history dependence is encoded through a collection of elementary bistable operators, each of which switches irreversibly between two output states at threshold values of the input, and the aggregate output is formed as a weighted superposition of the states of all operators. Because the state of each operator depends on the sequence of past inputs rather than only the current value, the model is inherently history-aware without requiring an explicit memory variable. The two most widely used operator-type models in the magnetics literature are the Preisach and Prandtl–Ishlinskii models. The Preisach model assigns a continuous weight distribution over a two-dimensional space of operator thresholds, and is the most general and mathematically well-characterized member of this class. The Prandtl–Ishlinskii model is a special case built from one-dimensional play operators whose symmetric density admits a closed-form analytical inversion [1, 25], which makes it attractive for feedforward compensation in smart actuators. As a strict Preisach subclass, it inherits the congruency property [100], which iron yokes violate.

2.5.1.1 The Preisach Model

The Preisach model, originally introduced to describe magnetic aftereffect in ferromagnetic materials [108] and later given a rigorous mathematical foundation by Mayergoyz [97], is the most widely used classical model for magnetic hysteresis. It decomposes the hysteretic response of a macroscopic material into a superposition of elementary bistable units called hysteron. Each

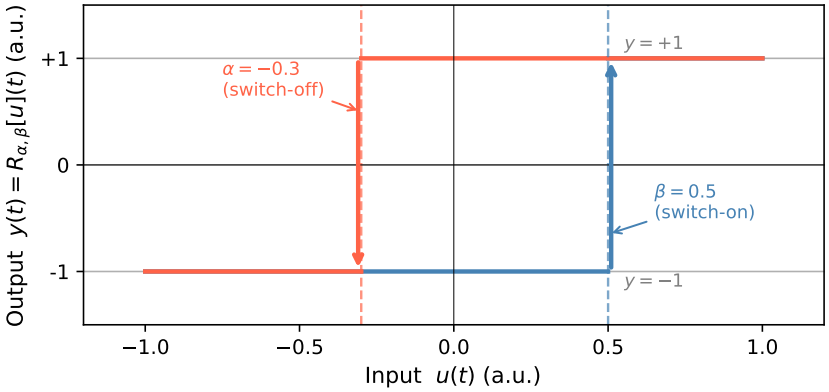


Fig. 2.3: Illustration of a hysteron operator $R_{\alpha\beta}$ which activates when the input signal crosses the upper threshold α, and deactives when the signal decreases below β. The units of the illustrated input signal are arbitrary as the hysteron operator is scale-invariant, and the output is a binary state of either $+1$ or -1.

hysteron represents a microscopic switching event such as the irreversible motion of a domain wall past a pinning site.

Each hysteron is characterized by two threshold values α and β with $\alpha \geq \beta$, which define the input levels at which the unit switches between its two output states. When the input $u(t)$ rises above α, the hysteron switches to the $+1$ state; when $u(t)$ falls below β, it switches to the -1 state; when $u(t)$ lies between β and α, the hysteron retains its current state. Formally, the hysteron output is defined as

$$\hat{R}_{\alpha\beta}[u](t) = \begin{cases} +1 & \text{if } u(t) > \alpha, \\ -1 & \text{if } u(t) < \beta, \\ \hat{R}_{\alpha\beta}[u](t^-) & \text{otherwise,} \end{cases} \tag{2.18}$$

where t^- denotes the previous time instant.

The Preisach plane is the half-plane $\{(\alpha, \beta) : \alpha \geq \beta\}$, and each point in this plane corresponds to a distinct hysteron. The full Preisach model assigns a weight $\mu(\alpha, \beta) \geq 0$ to each point in this plane, known as the Preisach density or weight function, and defines the model output as the weighted integral of all hysteron states,

$$B(t) = \iint_{\alpha \geq \beta} \mu(\alpha, \beta)\, \hat{R}_{\alpha\beta}[u](t)\, \mathrm{d}\alpha\, \mathrm{d}\beta. \tag{2.19}$$

The weight function $\mu(\alpha, \beta)$ encodes the distribution of switching thresholds across the ensemble of hysterons and is the only free function of the model. Physically, a region of high density in the Preisach plane corresponds to a large population of hysterons with similar switching thresholds, reflecting a concentration of pinning sites with similar energy barriers in the material microstructure. The Preisach plane is therefore a material representation: different magnets with different microstructures will generally have different weight functions. For example, soft magnetic materials with low coercivity are expected to have a weight function concentrated near the diagonal $\alpha \approx \beta$, with mostly a linear magnetic response. At any instant, the Preisach plane is partitioned into regions S^+ and S^- holding the hysterons in the $+1$ and -1 states, separated by a staircase interface whose vertices record the sequence of past input extrema. Output follows

as the excess weight of S^+ over S^-, and the interface moves with the input to produce the hysteretic B–I response.

For numerical evaluation and identification, the plane is discretised onto a finite mesh of N points (α_i, β_i), replacing the integral of Eq. (2.19) with the weighted sum

$$B(t) \approx \sum_{i=1}^{N} \mu_i \, \hat{R}_{\alpha_i \beta_i}[u](t), \tag{2.20}$$

with mesh weights $\mu_i = \mu(\alpha_i, \beta_i)$ fitted to data. Classical identification measures First Order Reversal Curves (FORCs), a family of reversal curves from saturation whose mixed second derivative estimates $\mu(\alpha, \beta)$, though the measurement count and the $\mathcal{O}(N^2)$ least-squares fit scale poorly with mesh resolution and lose accuracy for soft materials with complex minor loops [124]. Gradient-based optimization over differentiable implementations of the model [113] addresses this, as discussed in Section 2.5.3.

The Preisach model satisfies two properties that Mayergoyz showed to be both necessary and sufficient to characterise the class of hysteretic systems representable by a Preisach operator [97]. The first is the wiping-out property, also called the return-point memory property. This means that whenever the input reaches a new local maximum, all previous staircase vertices with a β value smaller than that maximum are erased, and similarly for local minima. Consequently, the model memory retains only dominant past extrema, while minor fluctuations that are later exceeded are forgotten. The second is the congruency property, which states that all minor loops traced between the same pair of input reversal values u_1 and u_2 are congruent, meaning they have the same shape and vertical extent regardless of the magnetisation level at which they are initiated. Soft ferromagnetic materials satisfy the wiping-out property to a good approximation, but often violate the congruency property: experimentally measured minor loops in iron yokes are not congruent and depend on the magnetisation level at which reversal occurs [19]. This is a fundamental limitation of the classical Preisach model that cannot be resolved by refining the mesh or improving the identification procedure.

Several extensions of the classical model have been proposed to address the congruency violation. The moving Preisach model introduces a mean-field feedback term that shifts the effective input by a quantity proportional to the current output, thereby making the minor loop shape dependent on the magnetisation state [19]. The nonlinear Preisach model replaces the fixed weight function with one that depends on the current output value, introducing state-dependent weighting at the cost of a more complex identification procedure. Despite these extensions, the classical and extended Preisach models share a common practical limitation, where the weight function is a static object identified from a fixed set of excitation cycles Consequently, it cannot adapt to changes in the magnet's magnetic history or operating conditions without re-identification. Furthermore, the model does not account for rate-dependent hysteresis or rate-dependent effects in the macroscopic magnetic field, such as eddy-current contributions at higher ramp rates, which are discussed separately in Section 2.6.3. More recently, a differentiable implementation of the Preisach model with numerical parametrisations of the weight function has enabled gradient-based identification from arbitrary measurement sequences, which is discussed further in Section 2.5.3.

2.5.2 ODE-type Models

ODE-type hysteresis models describe the magnetic state through a first-order ordinary differential equation in which the rate of change of magnetization with respect to the applied field depends

on the current state and the direction of the field sweep. The general form is given by the Duhem model [45, 136],

$$\frac{\mathrm{d}M}{\mathrm{d}H} = \begin{cases} f^+(M,\, H) & \dot{H} > 0, \\ f^-(M,\, H) & \dot{H} < 0, \end{cases} \tag{2.21}$$

where f^+ and f^- are model-specific functions governing the ascending and descending branches, respectively.

Hysteresis arises from the asymmetry $f^+ \neq f^-$, and the model possesses local memory represented only by the current state (M, H). The future trajectory is then fully determined by state (M, H) and the sign of $\dot{H}$, with no dependence on the prior input history beyond what is encoded in M. This contrasts with operator-type models, whose output depends on a distributed memory structure over the full input history. Special cases of (2.21) differ in the choice of $f^{\pm}$, from the arctangent branches of the Flatley model [50] to the domain-wall formulation of the Jiles–Atherton model [74] treated below.

The main advantages of ODE-type models are computational simplicity, compact parameterisation, and, in some variants, a direct correspondence between parameters and material properties. However, several limitations apply to the general class of models. First, numerical integration of (2.21) over long or complex excitation sequences accumulates truncation and rounding errors, causing the simulated trajectory to drift from the true hysteresis loop. Additionally, no natural reset mechanism exists unless one is explicitly introduced [36, 148]. Second, ODE models of the Duhem form produce congruent minor loops by construction, since the trajectory on any reversal branch is determined solely by the current state. They therefore cannot reproduce the non-congruent minor loop behavior observed in many soft magnetic materials [97]. Third, even with few model parameters, parameter identification is a non-convex optimization problem, and gradient-based methods are not guaranteed to find a global minimum [85]. Finally, the formulation in (2.21) is rate-independent by definition, so this model form cannot capture rate-dependent hysteresis or dynamic effects arising from eddy currents.

2.5.2.1 The Jiles–Atherton Model

The Jiles-Atherton (JA) model [73, 74] derives the magnetization response from a physical picture of domain-wall motion in a medium with pinning sites. The total magnetisation M is decomposed into an irreversible component M_{irr}, associated with domain-wall pinning, and a reversible component M_{rev}, associated with domain-wall bowing without depinning,

$$M = M_{\mathrm{irr}} + M_{\mathrm{rev}} = M_{\mathrm{irr}} + c(M_{\mathrm{an}} - M_{\mathrm{irr}}) \tag{2.22}$$

where $c \in [0, 1]$ is the reversibility coefficient and M_{an} is the anhysteretic magnetisation, representing the equilibrium state in the absence of pinning. Consequently, it also represents the idealised non-hysteretic response of the material, including saturation and nonlinearity but no history dependence. The anhysteretic curve is modelled by a Langevin function,

$$M_{\mathrm{an}}(H_e) = M_s \left[\coth\!\left(\frac{H_e}{a}\right) - \frac{a}{H_e} \right], \tag{2.23}$$

evaluated at the effective field $H_e = H + \alpha M$, which accounts for interdomain coupling through the mean-field parameter α, which accounts for interdomain coupling. The irreversible magnetisation evolves according to

$$\frac{\mathrm{d}M_{\mathrm{irr}}}{\mathrm{d}H} = \frac{M_{\mathrm{an}} - M_{\mathrm{irr}}}{k\,\delta - \alpha(M_{\mathrm{an}} - M_{\mathrm{irr}})}, \tag{2.24}$$

where k is the pinning coefficient controlling coercivity and $\delta = \pm 1$ is the sign of $\dot{H}$, which enforces the Duhem direction dependence of (2.21). Combining (2.22) and (2.24) yields the full ODE for M, which is integrated numerically given the internal state $(M_{\mathrm{irr}}, M_{\mathrm{rev}})$ at the start of the excitation sequence $H(t)$. However, this internal state is not directly observable, as a single measurement of (M, H) does not uniquely determine the decomposition in (2.22). Therefore, initialising the model at an arbitrary point on a minor branch requires additional assumptions about the prior magnetic history [18]. In practice, the model is therefore typically initialised from a known reference state such as saturation or the demagnetised state, which limits its applicability when the excitation history is not fully controlled.

The JA model reproduces the qualitative features of major loops and FORCs, and its parameters can in principle be related to independently measurable material properties. However, several practical difficulties arise. The denominator of (2.24) can change sign near reversal points, producing non-physical negative differential susceptibility and numerical instability in the integration [36, 148]. Parameter identification requires non-convex optimisation over five coupled parameters, and different local minima can yield loops that look similar on the major loop but diverge substantially on minor branches [85]. Like all Duhem models, the JA model produces congruent minor loops and is therefore unable to reproduce non-congruent behaviour [98]. Numerical drift under repeated cycling accumulates over long excitation sequences and limits the practical use of the model as an accurate surrogate without periodic re-initialisation. Inherent minor loop drift is a fundamental limitation of the ODE model structure and cannot be resolved by improving numerical integration or parameter identification.

The five-parameter structure and physical derivation make the JA model the most generic model for magnetic hysteresis within the ODE class. However, the fundamental limitations of drift, congruency, and initialisation ambiguity remain, and are not resolved by the additional physical grounding. Despite these limitations, the JA model remains the most widely used physics-based ODE hysteresis model in the magnetics literature.

2.5.2.2 Other ODE-type Models

Further Duhem special cases differ mainly in how $f^{\pm}$ is specified. Fröhlich's model [54] writes the differential permeability as a rational function of magnetisation and in its original form describes only the anhysteretic curve. A measurement-driven alternative is the Tellinen model [128], which interpolates the interior response between measured upper and lower major-loop branches and so accommodates any loop shape, at the cost of depending on those curves. For soft materials under unipolar excitation, the Coleman–Hodgdon model [38, 68] is common in finite-element codes because it is easy to identify. The Bouc–Wen model [23, 138], originally from structural mechanics, introduces an auxiliary hysteresis variable whose flexible form reproduces a wide range of loop shapes and maps onto a magnetic hysteresis ODE through $u \leftrightarrow H$ and $z \leftrightarrow M$.

All models above are rate-independent, with the output path set by the input trajectory rather than the rate of traversal. Rate-dependent extensions add a term proportional to $\dot{H}$ or $\dot{B}$ to the Duhem ODE of (2.21), widening the loop with sweep rate, as in the viscosity-based model of Zirka and Moroz [148] that augments the Jiles–Atherton framework with a statistical excess-loss term. This macroscopic rate dependence is distinct from the microscopic spin dynamics of the LLG equation. In iron-yoke accelerator magnets, eddy currents drive the dominant dynamic field errors, as treated in Section 2.6.

2.5.3 Machine Learning Approaches

Machine-learning approaches to hysteresis span a range from models that retain an established phenomenological structure to models that learn the input-output dynamics directly from measured sequences. Structural constraints keep the output physically valid and lower the data required for identification, and relaxing them grants flexibility at the cost of larger data requirements and weaker physical guarantees.

Neural parametrisations of the Preisach model of Section 2.5.1.1 retain the operator and its physical guarantees, and replace only part of its identification procedure with trainable networks. A feed-forward network was first used to represent the Preisach output functional, and the method was later extended to identify magnetic hysteresis from noisy measurements [118]. A Preisach recurrent network then combined play operators with a recurrent weighting and represented rate-dependent hysteresis in ARMCO iron from measured loops [61].

Model differentiability makes the derivatives of its output with respect to its parameters available, so gradient descent can identify the parameters. Roussel et al. [113] introduced a differentiable, non-parametric Preisach model in which smooth approximations replace the discontinuous hysteron operators shown in Figure 2.3, and the discretised Preisach density is fitted from sequential measurements. The model was demonstrated on accelerator magnets through direct magnetic measurements and indirect beam-based observations, and it was coupled to a statistical model of the beam response for model-based optimisation; its evaluation covered magnet characterisation and model-based optimisation over controlled excitation scans.

Physics-informed neural networks add governing differential equations, boundary conditions, or constitutive constraints to the training objective, which restricts the learned solution with available physical knowledge [109]. In this way, saturation and hysteresis were incorporated into a one-dimensional time-domain magneto-quasi-static problem, with the governing equations enforced during training [64].

Structure has also been imposed through the evolution of a learned internal state. HystRNN used an ordinary-differential-equation-inspired recurrent update to represent history dependence, evaluated on first-order reversal curves and minor loops outside its training set [33]. Operator learning, which learns mappings between complete input and output functions, has been applied to magnetic-field trajectories using DeepONet, Fourier, and wavelet neural operators [32].

Autoregressive models represent hysteresis through delayed samples of the excitation and response. A neural nonlinear autoregressive model with exogenous inputs, denoted NARX, predicts the next output from these delayed samples without a Preisach-type state. A multilayer NARX network was trained on measured field data from an iron-dominated CERN quadrupole and evaluated on quasi-static and dynamic loops [4], one of the few machine-learning applications to hysteresis in a physical accelerator magnet, with an evaluation based on controlled measurements and individual loops.

General sequence models represent magnetic memory with fewer structural constraints. A learned hidden state is propagated through an ordered sequence in a recurrent neural network, and gated states in a long short-term memory network retain information across longer intervals [47, 67]. In an encoder-decoder model, the encoder forms a latent representation of the input history and the decoder produces the corresponding output sequence [123]. A sequence-to-sequence LSTM encoder-decoder mapped measured flux-density waveforms and operating conditions to complete field waveforms for power magnetic materials [119].

An attention mechanism weights each position in the encoded sequence by its relevance to each prediction, and self-attention is the principal sequence operation of a Transformer [12, 133]. A Transformer encoder-projector-decoder was applied to measured B-H loops across materials, temperatures, frequencies, and excitation waveforms [89]. Across these approaches, the modelling

target is the full hysteresis response, in the form of major and minor loops, first-order reversal curves, or complete field waveforms, and the reported accuracies are referenced to the scale of that response.

2.6 Eddy Current Decay

When particle beams are accelerated in synchrotrons, the electromagnets are ramped by rapidly increasing the excitation to achieve the desired magnetic field for beam steering and focusing. The rapid current change produces a time-varying magnetic flux that induces circulating currents in all conducting materials threaded by that flux. These induced currents, known as eddy currents, generate a secondary magnetic field that opposes the applied flux change according to Faraday's law of induction. Once the current ramp ends and the excitation is held constant, the eddy currents dissipate resistively and the secondary field decays toward zero. The result is a time-decaying perturbation of the magnetic field in the magnet aperture that persists for timescales ranging from milliseconds to several seconds, depending on the geometry and conductivity of the surrounding materials.

This transient field perturbation appears after and during a fast magnet ramp, such as at injection after a previous magnet ramp-down, during acceleration, and right after the ramp-up is complete. During this period, decaying multipole field components, often including strong dipole and quadrupole components, can measurably affect the beam. The dipole component displaces the closed orbit as introduced in Section 2.1.5, while a decaying quadrupole component shifts the betatron tune and chromaticity, which can affect beam stability. Additionally, coupling between the dipole component, or, in other terms, beam energy, implicitly causes tune drifts as the effective focusing strength is normalised by beam rigidity. A qualitatively distinct class of dynamic field perturbations comes from the beam itself. As a charged bunch of particles travels through the resistive vacuum chamber, it excites electromagnetic wakefields that act back on trailing particles, driving transverse and longitudinal instabilities [44]. Eddy-current decay and wakefields are therefore complementary effects, but act on significantly different timescales. For this thesis, we consider only the dynamic effects caused by the power supply history that can be feed-forward controlled.

2.6.1 Eddy Currents in Conducting Materials

A time-varying magnetic flux Φ threading a conducting loop induces an electromotive force, a special case of Faraday's law, which drives a circulating current in the conductor.

$$\mathcal{E} = -\frac{\mathrm{d}\Phi}{\mathrm{d}t} \tag{2.25}$$

Here, the factor N familiar from wound coils is absent because eddy-current loops in a bulk conductor form single closed paths. According to Lenz's law, this induced current flows in a direction that opposes the change in flux, producing a secondary magnetic field that partially cancels the change in the applied field. In a bulk conductor of finite resistivity, the induced currents are distributed continuously throughout the material volume and are referred to as eddy currents.

The spatial distribution of eddy currents is not uniform throughout the conductor. At high rates of change, currents are confined to a surface layer characterised by the skin depth,

$$\delta = \sqrt{\frac{2\rho}{\omega\mu}}, \tag{2.26}$$

where ρ is the electrical resistivity, μ is the magnetic permeability, and ω is the angular frequency of the field variation. A smaller skin depth confines eddy currents to a thin surface layer, reducing the effective loop area and increasing the effective resistance of the eddy current path. This penetration scale is the physical basis for lamination as a mitigation strategy, which is discussed in Section 2.6.2. When the drive slows or stops, the skin depth grows, and currents redistribute through the full conductor cross-section before decaying resistively; this regime is relevant to field settling after a ramp.

The decay of eddy currents after a step change in excitation can be described by analogy with a series Reinforcement Learning (RL) circuit, in which the inductance L represents the magnetic energy stored in the eddy current loop and the resistance R represents the ohmic dissipation [101]. The governing equation for the current I_{eddy} in such a loop is

$$L\frac{\mathrm{d}I_{\mathrm{eddy}}}{\mathrm{d}t} + R\,I_{\mathrm{eddy}} = 0, \tag{2.27}$$

which has the solution

$$I_{\mathrm{eddy}}\,(t) = I_0 \exp\!\left(-\frac{t}{\tau}\right), \quad \tau = \frac{L}{R}, \tag{2.28}$$

where I_0 is the eddy current amplitude at the start of the flat excitation and τ is the decay time constant. The secondary magnetic field contribution in the aperture is proportional to the eddy current,

$$\Delta B_{\mathrm{e}}\,(t) = \alpha\, I_{\mathrm{eddy}}\,(t), \tag{2.29}$$

where α is a geometry-dependent proportionality constant set by the conductor geometry and its position relative to the aperture. The secondary magnetic, or eddy-current-driven, field contribution is proportional to I_{eddy} , so the field perturbation in the aperture decays with the same time constant τ. The amplitude I_0 is set by the particular solution of the inhomogeneous Ordinary Differential Equation (ODE) driven by the excitation ramp. For a constant ramp rate $\frac{\mathrm{d}I}{\mathrm{d}t}$, the steady-state driven solution gives $I_0 \propto \tau\,\frac{\mathrm{d}I}{\mathrm{d}t}$, so a faster preceding ramp produces a larger initial eddy current amplitude. However, this decay picture is valid only when the excitation is held constant. During an arbitrary time-varying excitation, or even small excitation ramps following a larger one, the full inhomogeneous ODE must be solved, and the field perturbation does not take a simple exponential form. Finally, the time constants τ_k are not, in general, fixed material properties when discussing eddy current decay in the iron yoke, but depend on the local magnetic permeability μ of the iron, which varies with the excitation level through the nonlinear B–H curve. During low excitation I, or equivalently H, the time constant is approximately constant. However as the iron approaches saturation, μ decreases and the time constant τ becomes $\tau = \tau(\mu)$ [107]. In practice, a real conductor contains many such loops of differing geometry, each with its own time constant.

2.6.2 Eddy Currents in Synchrotron Components

The preceding section introduced eddy currents as a general phenomenon in conducting materials; however, in a synchrotron, time-varying magnetic flux threads not only the magnet's iron yoke but every conducting structure within and around the magnet assembly. Eddy currents therefore arise in every conducting material threaded by the flux, not only in the iron yoke itself. The relevant conductors in a normal-conducting synchrotron magnet assembly include the laminated iron yoke, the beam vacuum chamber, and other mechanical support structures. Each contributes an eddy current mode with a characteristic amplitude and time constant determined by its geometry, conductivity, and proximity to the beam aperture. [101]

2.6.2.1 Iron Yoke

The iron yoke carries the bulk of the magnetic flux and thus experiences the largest flux variation during a current ramp. Its high relative permeability $\mu_r \gg 1$ yields a small skin depth δ from (2.26), which would confine eddy currents to thin surface layers if the yoke were a solid conductor. In practice, accelerator-magnet iron yokes are constructed from thin laminated sheets, electrically insulated from one another [101]. Each sheet forms an independent eddy-current loop with a cross-sectional area proportional to the sheet thickness squared, increasing the effective resistance R of each loop and reducing both the eddy-current amplitude and the time constant $\tau = L/R$. This increases the total ohmic resistance R of each loop and reduces both the amplitude and the time constant $\tau = L/R$, so eddy currents in the iron decay on timescales of order milliseconds or shorter [126]. As a consequence, although the iron yoke is the most strongly driven conductor, its eddy-current contribution decays rapidly and becomes negligible on longer timescales.

2.6.2.2 Vacuum Chamber

The vacuum chamber runs through the magnet aperture and encloses a large fraction of the flux path, while being physically closer to the aperture centre than the iron yoke. Unlike the iron yoke, the vacuum chamber is not laminated and is typically fabricated from stainless steel, aluminium, or copper, materials with low relative permeability $\mu_r \approx 1$. The low permeability results in a large skin depth, so the full wall thickness participates in eddy current conduction and forms a large effective loop with low resistance. Consequently, the resulting time constants can range from tens of milliseconds to several seconds depending on the chamber material, wall thickness, and cross-sectional geometry [106]. These slow components therefore dominate eddy-current decay on timescales relevant to beam dynamics and are the primary focus of this thesis.

The transverse cross-sectional geometry of the vacuum chamber determines which multipole components the induced eddy currents produce in the aperture. For a circular cross-section, symmetry dictates that a dipole excitation induces eddy currents that generate only a dipole field component inside the chamber [106], as explained in more detail in Chapter 3. When the cross-section deviates from circular symmetry, as in elliptical, rectangular, or flat chambers common in synchrotrons to accommodate the horizontal beam size, the eddy-current distribution loses this symmetry. The asymmetric current distribution then generates higher-order multipole components in addition to the fundamental dipole, with the specific multipoles that arise determined by the symmetry class of the cross-section [86]. For example, an elliptical or rectangular chamber with one axis of symmetry along the dipole field direction generates a sextupole component from a pure dipole excitation [106, 130]. Finally, each multipole mode has its own time constant, so the overall eddy-current contribution from a realistic chamber cross-section is a multi-exponential decay with mode-dependent multipole content.

While individual vacuum chambers can carry different eddy-current modes, in practice, the full ring of a synchrotron is composed of many adjacent vacuum chamber sections joined by bellows and flanges. If these joints are electrically conductive, the eddy-current path is no longer confined to a single chamber section but can extend across longer sections, or even around the full ring, forming a large low-resistance loop with a correspondingly long time constant. In practice, RF shields or insulating flanges are sometimes used to break such longitudinal current paths and limit the effective loop length.

The vacuum chamber material and geometry must simultaneously satisfy competing requirements from field transparency, beam coupling impedance, and grounding continuity. A thicker or more conductive chamber wall increases both the eddy-current time constant and the induced field amplitude, which is undesirable from a field-quality perspective, but a thinner or more resistive wall increases the chamber's beam-coupling impedance [77, 106]. Therefore, the choice

of vacuum chamber material and wall thickness involves a fundamental design tradeoff. Vacuum chambers in modern fast-cycling synchrotrons are therefore commonly fabricated from stainless steel, which offers a compromise between sufficient resistivity to limit eddy current amplitudes and adequate mechanical rigidity for the required aperture geometry. Where lower impedance is required, chambers are sometimes made from aluminium or copper with a thin resistive coating on the inner surface, or from ceramic with a thin metallic layer, so the eddy-current path is geometrically constrained while the beam sees a low-impedance wall. The grounding scheme of the chamber further complicates this tradeoff, as a continuous conducting path along the beam pipe allows eddy currents to propagate over long distances, whereas segmenting the chamber with insulating flanges reduces this effect at the cost of increased impedance at each junction, as was the case in the SPS [78, 79].

2.6.2.3 Other Conducting Structures

While the vacuum chamber and magnet yoke are the dominant sources of eddy currents in synchrotron magnets, other mechanical structures in and around the magnet aperture contribute additional induced current paths that can perturb the field seen by the beam. Even in the magnets, solid steel end plates typically clamp the laminated yokes at each magnet end. Unlike the laminated body, these end plates are bulk conducting structures in which eddy currents can circulate freely, and their time constant can be significantly longer than that of the laminations. [101, 125]

Magnet support girders and steel alignment fixtures that mechanically connect adjacent magnets can form closed conducting loops near the yoke. The contribution of these paths to the total eddy current response is difficult to characterise reproducibly, as it depends on the contact resistance at bolted joints, which varies with assembly and may change over time. [101]

Bus bars and current leads connecting the magnet coils to the power converter run along the magnet length and carry the same time-varying current as the main excitation circuit. Although their distance from the beam axis limits their direct influence on the field in the aperture, inductive coupling between the leads and the surrounding conducting structures can drive additional circulating currents during rapid field ramps. [101]

Finally, in-vacuum instrumentation such as BPMs electrodes introduces local conducting inserts into the vacuum chamber wall at discrete locations around the ring. These inserts locally perturb the otherwise uniform eddy-current distribution in the chamber wall and can introduce asymmetric field multipole contributions that a simple one-dimensional model of the chamber does not capture. [37]

2.6.3 Beam Impact of Eddy Current Decay in Fast Cycling Synchrotrons

2.6.3.1 Beam Impact of Eddy Current Decay

Eddy current decay in the conducting structures of a synchrotron lattice produces time-dependent errors in all principal field components, which in turn perturb the main beam parameters. Since the orbit, momentum, tune, and chromaticity of the beam are coupled through the lattice optics, as established in Section 2.1, a perturbation to any one field component propagates into the others. The resulting beam parameter drift can therefore significantly impact injection efficiency, beam stability during acceleration, and extraction quality, or even cause full beam loss if the field perturbation is large enough to drive the beam outside the dynamic aperture or cause instabilities.

Eddy current-induced field decay in the dipole component causes the actual bending field to fall below the programmed value after a change in excitation, shifting the beam's closed orbit as

described in Section 2.1.5. At injection flat bottom, this manifests as an orbit mismatch between the incoming beam and the receiving ring, since the dipole field has not yet reached its stationary value, and the beam arrives on a displaced orbit relative to the design trajectory.

The decay in quadrupolar field components, whether from the main quadrupole magnets or from asymmetric eddy-current distributions in surrounding conducting structures, causes the effective focusing strength of the lattice to drift over time, producing a corresponding drift in the betatron tune [82]. Additionally, since the beam circulates on a displaced closed orbit due to the dipole field decay, it experiences a different quadrupole gradient through the dispersion function in (2.3), introducing an additional tune shift through chromaticity, as described in Section 2.1.7.1. A simultaneous drift in the sextupolar field component, which eddy currents in vacuum chambers with specific shapes can induce, can further perturb the chromaticity independently of quadrupole decay. The interplay between these effects is further complicated when an RF feedback loop is active, as the loop responds to the momentum deviation caused by the dipole field decay by adjusting the RF frequency, which in turn shifts the closed orbit through the momentum compaction factor and couples the orbit, tune, and chromaticity drifts in a non-trivial way.

In fast-cycling synchrotrons, the beam is typically injected immediately after the preceding ramp-down to minimise machine downtime, so the conducting structures of the lattice have not fully reached their unperturbed remanent state. The residual eddy currents from the preceding cycle therefore set the initial conditions for stationary decay at injection flat bottom, making the beam parameters at injection sensitive to the machine-cycle history [82, 93]. This inter-cycle dependence is particularly pronounced in the quadrupole and sextupole circuits, where decaying field components directly perturb the tune and chromaticity on a timescale comparable to injection and early acceleration.

During a fast field ramp, eddy currents in the lattice's conducting structures cause the actual field to lag behind the programmed excitation, with the lag magnitude determined by the amplitude and time constants of the eddy-current modes [35, 82]. Fast eddy current modes that are negligible during the stationary flat-bottom decay become significant during rapid ramps, and if the resulting field mismatch between the programmed and actual field is not compensated, it can be sufficient to prevent beam acceleration entirely [6]. This intra-cycle effect is not confined to the main magnets, as any conducting structure in the lattice, including vacuum chambers, support elements, and instrumentation, contributes to the total lagging field [101]. Furthermore, eddy-current-induced chromaticity errors during the ramp can drive head-tail instability, as observed during commissioning of rapid-cycling synchrotrons where extra sextupole correction was not applied. [6]

After acceleration and entry into the extraction flat top, a stationary eddy-current decay reappears, with initial conditions set by the preceding ramp, analogous to the inter-cycle effect at injection. An uncompensated field drift at the extraction flat top perturbs the beam orbit and optics at extraction, and both fast and slow extraction schemes require the beam parameters to match the extraction channel geometry within tight tolerances; without proper compensation, performance degrades.

2.6.4 Mitigation Strategies

Mitigation of eddy current effects can be addressed at the hardware, operational, and software levels, and in practice multiple approaches are combined depending on the magnitude of the effect relative to the field tolerance required for beam operation. At the hardware level, lamination of the iron yoke is the primary mitigation, as discussed in Section 2.6.1, and the same principle can be extended to the end plates that clamp the laminated yoke, by replacing bulk steel clamps with laminated or non-conducting composite structures to suppress their slow-decaying

field contribution [101]. The vacuum chamber wall can be slotted or segmented to break the continuous eddy current path along the beam pipe, or insulating flanges can be used to sectorise the conducting path into shorter grounded sections, as implemented in the SPS with one ground connection per half FODO cell of approximately 32 m [78]. As discussed in Section 2.6.2, both approaches reduce the effective loop area and induced field amplitude, but each discontinuity in the chamber wall contributes to the beam-coupling impedance and must be accounted for in the machine impedance model [78, 79].

At the operational level, increasing the dwell time between successive machine cycles allows eddy currents from the preceding ramp to decay further before injection, directly reducing the inter-cycle initial condition sensitivity described in the previous subsections. This comes at the cost of reduced machine duty time, and is therefore most applicable in machines where the cycle repetition rate is not a primary operational constraint. Choosing a slower field ramp rate reduces the amplitude of intra-cycle lagging effects, since the induced eddy-current amplitude scales with the rate of change of excitation, but this also conflicts with the demand for high cycle efficiency in fast-cycling machines.

When eddy-current decay is sufficiently reproducible from cycle to cycle, its effect on beam parameters can be compensated by applying pre-computed feed-forward current offsets to the affected magnet circuits as a function of time within the cycle [6]. This approach has been applied in the CERN SPS, where manually determined offsets to the quadrupole and sextupole circuits during the ramp preserve cycle-to-cycle reproducibility of the beam optics [6]. Its validity is, however, conditional on a controlled magnetic history, since the correction waveforms are derived under specific precycle conditions and lose accuracy when the supercycle structure or cycle sequence changes.

Where greater flexibility or precision is required, a physics-based or data-driven model of eddy-current decay can predict the time evolution of the field error and generate a dynamic compensation waveform that adapts to varying machine histories [35, 82]. Finally, if the eddy current contribution to the field error is small relative to the field tolerance required for beam operation, no active mitigation is applied, and the effect is absorbed into the overall field uncertainty budget, which is typically the case for higher-order conducting structures such as instrumentation inserts and bus bars [101].

Chapter 3

Accelerator Magnets and Magnetic Measurements

The previous chapter introduced the transverse magnetic field in an accelerator magnet as a Taylor expansion of B_y along the horizontal midplane in (2.7). This expression was sufficient to derive the beam-motion equations but was restricted to the on-axis gradient structure of upright magnets. This chapter develops the full two-dimensional multipole description of the field inside the magnet aperture, needed for the magnetic measurement methods described later. The description distinguishes the normal field components that enter beam dynamics from the skew components produced by magnet imperfections, and relates the relative coefficients reported by measurement systems to the absolute integrated strengths used in the equations of motion.

The chapter then surveys the principal techniques used to measure the main magnetic fields and multipole components in practice, and briefly evaluates them against the field accuracy requirements established in Chapter 2. A particular emphasis is placed on integral measurement methods, which directly yield the field quantity relevant for beam dynamics, and the systematic limitations that prevent any single technique from achieving drift-free, absolute, multipole-resolved measurements at the required level during pulsed synchrotron operation.

The chapter concludes by describing the specific laboratory measurement systems and the online magnetic monitoring system used in the CERN synchrotron, including the SPS, which provide the data underpinning the field models developed in Chapters 7 and 8. The measurement accuracy limitations identified here are the primary constraint on the quality of that training data and, consequently, on the compensation performance discussed in Chapter 8.

3.1 Magnetic Field in the Magnet Aperture

The on-axis field expansion of Chapter 2 is a special case of the full two-dimensional field in the magnet aperture. In the current-free interior, Maxwell's equations reduce to Laplace's equation for a scalar potential, whose regular solutions form a multipole series [114]. Introducing the complex transverse coordinate $z = x + iy$ and a reference radius R_{ref}, the field is written as

$$B_y(x,y) + iB_x(x,y) = \sum_{n=1}^{\infty} (B_n + iA_n) \left(\frac{z}{R_{\mathrm{ref}}} \right)^{n-1}, \tag{3.1}$$

where B_n and A_n are the normal and skew harmonic coefficients of order n, with $n = 1$ the dipole, $n = 2$ the quadrupole, and $n = 3$ the sextupole. The normal coefficients B_n describe the field of an upright magnet, while the skew coefficients A_n describe components rotated by $90°/n$ in the transverse plane and arise from magnet tilts, assembly asymmetries, and manufacturing imperfections. Setting $y = 0$ and $A_n = 0$ reduces (3.1) to the midplane Taylor series $B_y(x,0) = \sum_n B_n (x/R_{\mathrm{ref}})^{n-1}$ of (2.7).

3.1.1 Field Quality Conventions

The coefficients B_n and A_n in (3.1) have dimensions of magnetic flux density, and their magnitude depends on the choice of reference radius R_{ref}. For magnet characterisation and comparison across different magnet families, it is convenient to work with dimensionless field quality coefficients, defined by normalising to the principal field component B_{ref},

$$b_n = \frac{B_n}{B_{\mathrm{ref}}}, \qquad a_n = \frac{A_n}{B_{\mathrm{ref}}}, \tag{3.2}$$

where B_{ref} is the main harmonic of the magnet type under consideration: B_1 for a dipole, B_2 for a quadrupole, and so on, following the same convention as in Chapter 2. The coefficients b_n and a_n thus express each multipole component as a fraction of the design field evaluated at R_{ref}.

By convention in accelerator magnet metrology, b_n and a_n are reported in units where $1\,\mathrm{unit} = 10^{-4}$, so that a coefficient of 1 unit corresponds to a field error of 10^{-4} of the reference field at R_{ref}. The normalized strengths k_{n-1} used by accelerator physicists follow from these coefficients as

$$k_{n-1} = \frac{1}{B\rho}\,\frac{B_n}{R_{\mathrm{ref}}^{n-1}} = \frac{B_{\mathrm{ref}}}{B\rho}\,\frac{b_n}{R_{\mathrm{ref}}^{n-1}}, \tag{3.3}$$

with the momentum dependence carried by the rigidity $B\rho$ and separated from the normalised gradient k_{n-1} seen by the beam. This makes explicit that accelerator physicists require the absolute value of B_{ref} and the beam rigidity $B\rho$ to translate the dimensionless b_n into a focusing strength, whereas magnet builders work entirely with the relative coefficients b_n and a_n at a fixed R_{ref}. In practice, beam physicists evaluate the equations of motion using the longitudinal integral $\int B_n\,\mathrm{d}z$, which the normalised strength k_{n-1} in (3.3) represents by implicitly assuming the field is concentrated over an effective hard-edge length. This distinction between relative coefficients reported by measurement systems and the absolute integrated strengths required by beam dynamics drives the accuracy requirements for the measurement techniques discussed in Section 3.2.

3.1.2 Forbidden Harmonics and Magnet Symmetry

The pole geometry of a magnet imposes discrete rotational antisymmetry on the field, so only a subset of harmonic orders in (3.1) can be non-zero in a geometrically perfect magnet [114]. A $2p$-pole magnet is invariant under rotation by π/p combined with a reversal of the field direction, which admits only the orders $n = p(2k-1)$ for $k = 1, 2, 3, \ldots$. All remaining orders are the forbidden harmonics of the magnet, whose coefficients vanish identically when the geometry is perfect. For a dipole ($p = 1$) the allowed orders are $n = 1, 3, 5, \ldots$, for a quadrupole ($p = 2$) they are $n = 2, 6, 10, \ldots$, and for a sextupole ($p = 3$) they are $n = 3, 9, 15, \ldots$, as summarized in Table 3.1.

Tab. 3.1: Allowed normal harmonics b_n for the principal SPS magnet families. All other harmonics are forbidden and vanish in a geometrically perfect magnet.

Magnet type	Allowed normal harmonics	Lowest allowed error harmonic
Dipole	$n = 1, 3, 5, 7, \ldots$	b_3 (sextupole)
Quadrupole	$n = 2, 6, 10, 14, \ldots$	b_6 (duodecapole)
Sextupole	$n = 3, 9, 15, \ldots$	b_9 (18-pole)

The lowest-order field error in a dipole is therefore the sextupole component b_3. Since b_3 is allowed by symmetry, it is generated whenever the magnetisation of the iron yoke deviates from

perfect uniformity, whether through manufacturing tolerances, mechanical asymmetry, or the history-dependent domain configuration of the iron. For the SPS main quadrupoles, the lowest allowed error is the duodecapole b_6, whose beam dynamics effect scales as $(r/R_{\mathrm{ref}})^5$ relative to the main quadrupole gradient and is negligible for beams well within the aperture.

3.1.3 Fringe Fields and the Full-Length Measurements

The multipole expansion of Section 3.1 is strictly valid in the transverse cross-section of an infinitely long magnet. In a real magnet of finite physical length L_{phys}, the field does not terminate abruptly at the iron boundaries but decays continuously over a fringe-field region whose longitudinal extent is set by the aperture size and pole geometry. Each multipole coefficient B_n therefore becomes a function of the longitudinal coordinate z, with a profile $B_n(z)$ that differs from the constant body value in both magnitude and multipole content.

From the perspective of beam dynamics, a particle traverses the full length of the magnet including the fringe regions, and the physically relevant quantity is the longitudinal integral of each multipole component,

$$\int B_n \,\mathrm{d}\ell = \int_{-\infty}^{+\infty} B_n(z)\,\mathrm{d}z. \tag{3.4}$$

The integrated normalised strength entering the equations of motion follows from Eq. (3.3) as

$$\int k_{n-1}\,\mathrm{d}\ell = \frac{\int B_n\,\mathrm{d}\ell}{(B\rho)\,R_{\mathrm{ref}}^{n-1}}, \tag{3.5}$$

which extends the hard-edge definition of k_{n-1} to the full field profile. The integral quantities are often referred to as $B\,\mathrm{d}\ell$, or simply $B\,\mathrm{d}L$, and $k_{n-1}\,\mathrm{d}\ell$ or $k_{n-1}L$ in accelerator physics literature and common usage. The effective magnetic length is then defined as

$$L_{\mathrm{eff}} = \frac{\int B_n\,\mathrm{d}\ell}{B_n(0)}, \tag{3.6}$$

so that $\int k_{n-1}\,\mathrm{d}\ell = k_{n-1}\cdot L_{\mathrm{eff}}$. In the hard-edge approximation, L_{eff} is replaced by L_{phys}, which is exact only when the fringe contribution vanishes. For long magnets such as the SPS dipoles from Section 2.3.2, the fringe region is short relative to the magnet body and the approximation $L_{\mathrm{eff}} \approx L_{\mathrm{phys}}$ is typically acceptable for the main dipole component B_1. The approximate average field,

$$\langle B_n \rangle = \frac{\int B_n\,\mathrm{d}\ell}{L_{\mathrm{phys}}} = B_n(0)\,\frac{L_{\mathrm{eff}}}{L_{\mathrm{phys}}}, \tag{3.7}$$

provides a single-number characterisation equivalent to $\int B\,\mathrm{d}\ell$ once L_{phys} is known.

Near the magnet ends, the pole-face symmetry is broken by the three-dimensional field structure, so additional multipole components appear in the fringe region even in a geometrically perfect magnet, and the integrated sextupole $\int B_3\,\mathrm{d}\ell$ of a dipole is non-zero independently of any magnetisation error [114].

3.2 Measurement Techniques

Magnetic measurements of accelerator magnets serve a range of purposes, including the characterisation of field quality and multipole content during magnet production and subsequent validation, and the determination of magnetic axis position and alignment. In contrast, accelerator physicists are interested in measuring integrated field strengths for beam optics calculations and monitoring

field reproducibility during accelerator operation. Each application places different requirements on the measurement technique. Production characterisation requires multipole resolution across the full aperture; alignment requires knowledge of field centre position and is often performed at a single axial position, with fixed and reproducible excitation. Beam dynamics, by contrast, requires the integrated field along the particle trajectory with non-static excitation, and operational monitoring similarly requires pulsed excitation and calibration stability outside a controlled laboratory environment. No single instrument satisfies all of these requirements simultaneously, and in practice a combination of complementary techniques is used. A further distinction exists between absolute accuracy, which quantifies how closely a measurement reflects the true field, and cycle-to-cycle reproducibility, which quantifies whether the same excitation history yields the same measured signal independent of drift or offset errors. These two concepts of measurement quality impose different, and not always compatible, requirements on the instrumentation. The subsections below summarise the main measurement methods, arranged according to the physical quantity they yield and their suitability for pulsed operation.

3.2.1 Local, Integral, and Average Measurements

Because the beam-dynamics quantity of interest is $\int B_n \, \mathrm{d}\ell$ from Eq. (3.4), magnetic measurement techniques are classified by which field quantity they directly provide. A local measurement yields the field $B_n(x, y, z)$ at a single point or cross-section, and it cannot determine L_{eff} or the fringe-field multipole contributions without an additional axial scan of the full magnet length. An integral measurement yields $\int B_n \, \mathrm{d}\ell$ directly from a single measurement along the z axis, and an average measurement yields $\langle B_n \rangle$ over the coil cross-section and length, related to $\int B_n \, \mathrm{d}\ell$ through Eq. (3.7). The principal instruments belonging to each class are listed in Table 3.2.

Tab. 3.2: Classification of magnetic measurement techniques by the field quantity they provide directly.

Class	Quantity provided	Principal instruments
Local	$B_n(x, y, z)$	Hall probe, NMR probe
Integral	$\int B_n \, \mathrm{d}\ell$ over coil length	Rotating coil, fluxmeter
Integral	$\int B_n \, \mathrm{d}\ell$ over full length	Stretched wire, long fluxmeter

Integral measurements are the primary method for beam-dynamics characterisation, since they directly yield $k_{n-1} L_{\text{eff}}$. Local measurements are easier to implement, more adaptable, and better suited to pulsed operation; they complement integral methods for qualitative field verification, transverse profile mapping, and axial scans used to reconstruct the longitudinal field distribution. The subsections below outline each instrument individually, with particular emphasis on how they are used for magnetic measurements relevant to accelerator beam dynamics.

3.2.2 Hall Probes

A Hall probe exploits the Hall effect, where a current-carrying semiconductor placed in a transverse magnetic field B develops a voltage V_H perpendicular to both the current and field directions according to

$$V_H = R_H \frac{IB}{d}, \tag{3.8}$$

where R_H is the Hall coefficient of the semiconductor material, I is the bias current fed into the semiconductor, and d is the thickness of the sensing element. The output voltage is proportional

to the local flux density B at the probe location, making the Hall probe a local point measurement device. A typical probe assembly consists of a semiconductor chip mounted on a small printed circuit board, with total dimensions on the order of a few millimetres, allowing straightforward installation into magnet apertures, either on coil windings and iron yokes directly or on mechanical supports for fixed-position monitoring. Additionally, a typical Hall probe can be sampled at rates of several kilohertz, or even megahertz for specialised fast probes, making them well suited for pulsed operation and field monitoring during accelerator ramps.

The combination of small physical size, mechanical flexibility, low cost, and the ability to sample the field continuously at rates well above any synchrotron ramp rate makes Hall probes among the most practical field monitoring tools available. A probe installed at a fixed position can track the absolute field through an entire machine cycle, capturing cycle-to-cycle variations, including those arising from hysteresis and eddy-current decay, making it a candidate for the magnetic measurements required for this thesis.

Despite these practical advantages, achieving absolute field accuracy below the $1 \times 10^{-4}\,\mathrm{T}$ level with a Hall probe is intrinsically challenging, due to a combination of calibration, thermal, and mechanical limitations that each contribute errors comparable to the target accuracy and precision. Firstly, Hall probes require external calibration against an independently established reference field, typically provided by Nuclear Magnetic Resonances (NMRs) or a primary field standard [117]. The calibration must be performed individually for each probe, since the Hall coefficient, offset, and temperature sensitivity differ between units even within the same production batch. Additionally, the Hall coefficient R_H depends on the applied field magnitude, introducing a nonlinearity of typically $\leq \pm 0.1\%$ up to 1.5 T for well-characterised devices [5], which consequently require calibration at multiple field points across the measurement range to achieve high accuracy. Temperature variations on the semiconductor can also produce a sensitivity drift of approximately $-150\,\mathrm{ppm/K}$ for GaAs-based sensors [5] such as those discussed in Section 3.3. In controlled laboratory conditions, temperature can be kept approximately constant; however, in an accelerator tunnel, temperature gradients of several kelvin can exist along the magnet string, and the local yoke temperature itself varies with the machine cycle due to resistive heating of the coils, which would make the effective probe temperature difficult to track. A tunnel environment further introduces ionising radiation, which is known to affect the characteristics of semiconductor devices over time, causing gradual calibration shifts or outright sensor destruction. Separate from sensitivity drift, the zero-field offset of the probe output is generally not stable over time, and is particularly problematic near zero field, which is relevant for cycled magnets that leave a small remanent field at the end of the cycle. Beyond intrinsic sensor errors, the measurement is also sensitive to mechanical alignment and positioning, with the severity depending on the local field gradient at the probe location [117].

Because the probe provides only a local field value at a single point, an additional limitation arises when it is used as an absolute field reference. This relies on the assumption that the local field at the probe position represents the average field in the magnet body, which is valid in the uniform central region of a long dipole. However, this assumption breaks down near fringe regions, for higher-order multipole components, and whenever the field distribution across the aperture is non-uniform. The absolute accuracy achievable in this configuration is therefore bounded by both the probe limitations described above and the validity of this representativeness assumption.

3.2.3 NMR Probes

NMR probes provide an absolute reference for the magnetic flux density by exploiting the Larmor precession of nuclear spins in a magnetic field. In a homogeneous field with magnitude B, the nuclear precession frequency f_0 is proportional to the field strength according to

$$f_0 = \frac{\gamma}{2\pi} B, \tag{3.9}$$

where γ is the gyromagnetic ratio of the chosen nucleus, typically the proton for magnet metrology. Modern frequency standards determine f_0 with very high precision, so NMR probes can reach absolute field accuracies of order $1 \times 10^{-7}\,\mathrm{T}$ under well-controlled laboratory conditions.

The main advantage of NMR probes over Hall probes is that they measure frequency and not an electrical voltage proportional to the Hall coefficient. As a consequence, NMR measurements are largely immune to temperature drifts, long-term gain changes, and electronics ageing, and do not require frequent recalibration once the gyromagnetic ratio and frequency reference are fixed. In terms of attainable absolute accuracy under static conditions, NMR therefore provides an upper bound that is typically one to two orders of magnitude better than what can realistically be achieved with Hall probes on the same accelerator magnet measurement bench.

The advantage of the NMR, however, comes at significant trade-offs. To obtain a narrow and well-defined resonance line, the magnetic field must be highly homogeneous across the active sample volume and essentially constant over the full measurement time, which typically ranges from milliseconds to seconds depending on the signal-to-noise ratio and the required precision. NMR are therefore generally incompatible with pulsed operation during an accelerator ramp, since the field variations during the ramp are too rapid to allow a stable resonance signal. Another consequence of the homogeneity requirement is that NMR probes are not suitable for measuring higher-order multipole components, since field gradients across the sample volume broaden the resonance and reduce signal quality. Classical NMR teslameters therefore operate under quasi-static conditions in dedicated reference magnets and are not a universal high-accuracy field-measurement solution for accelerator magnets; rather, they serve as calibration references or highly accurate field monitors at fixed fields.

Recent commercial developments partially address these limitations. Some modern NMR teslameters integrate three-dimensional Hall sensors and advanced locking schemes, using the Hall readings to steer the excitation frequency close to the expected resonance and to track slow field variations [30]. Such hybrid systems extend the range of fields and operating conditions where an NMRs lock can be maintained, and are being adopted in accelerator magnet metrology at the time of writing. However, the measurement bench used in this thesis employs an earlier-generation NMR system without these capabilities, and this work focuses on fast-cycling synchrotron operation, where field dynamics remain incompatible with continuous NMR tracking. For these reasons, Hall probes are used as the primary local field sensor on the pulsed measurement bench, while NMR defines the static reference accuracy that cannot be reached during dynamic operation.

3.2.4 Stretched Wire

A taut conducting wire moved through a static field forms a closed loop with a return conductor, and by Faraday's law the induced voltage integrates to the line integral $\int B_y \,\mathrm{d}l$ of the field along the wire path [114]. A longitudinal scan at fixed transverse position yields the integrated main field and the effective magnetic length of (3.6), and transverse scans locate the magnetic axis and roll from symmetry conditions on the signal. Mechanical uncertainties in wire position, straightness, and sag set the achievable accuracy, reaching relative errors of a few 10^{-5} on the

main component under well-controlled conditions. Controlled mechanical motion restricts the method to static fields, so it provides a static reference for integrated field strength and magnetic axis but cannot follow fast ramps or characterise eddy-current decay.

3.2.5 Fluxmeters

Fluxmeters are integral magnetic measurement devices that infer the longitudinal field integral from the voltage induced in a sensing coil according to Faraday's law. For a coil with N turns and instantaneous flux $\Phi(t)$ through each turn, the induced voltage is

$$V(t) = -N\frac{\mathrm{d}\Phi(t)}{\mathrm{d}t} \tag{3.10}$$

so that an ideal electronic integrator with output signal $U(t)$ realises

$$U(t) = -\int_{t_0}^{t} V(t')\ \mathrm{d}t' = N\big[\Phi(t) - \Phi(t_0)\big] \tag{3.11}$$

The instrument therefore provides the change of magnetic flux linked to the coil between an initial reference time t_0 and the measurement time t. In accelerator magnet metrology, the coil is arranged to span the length of the magnet aperture so that the flux is expressed as a longitudinal integral over the field components of interest [114].

In practice, a fluxmeter system for accelerator magnet characterisation consists not of a single coil but of an array of elements placed at different transverse positions within the aperture. A minimal system for dipole field measurement uses a single element whose calibrated area yields the average field, while multipole characterisation requires at least two elements connected in a compensated arrangement that suppresses the dominant harmonic and isolates the target order. More complete systems can employ five or more at distinct radii, and high-resolution implementations reach twenty or more elements, enabling simultaneous extraction of several harmonic orders from a single excitation cycle [114]. The coils are traditionally hand-wound on mechanical supports, but modern implementations increasingly use Printed Circuit Board (PCB) fluxmeters to achieve higher precision and reproducibility in the coil geometry.

The flux linked to element k, placed at radius $R_{\mathrm{coil},k}$ with N_k turns, is related to the multipole expansion introduced in Section 3.1. Using

$$B_y(x,y,z) + iB_x(x,y,z) = \sum_{n=1}^{\infty}\big[B_n(z) + iA_n(z)\big]\left(\frac{x+iy}{R_{\mathrm{ref}}}\right)^{n-1}, \tag{3.12}$$

and neglecting small variations of B_n across the coil thickness, the contribution of the normal multipole of order n to the flux through element k scales as [114]

$$\Phi_{n,k} \approx N_k\, R_{\mathrm{coil},k}^{n} \int_0^{L_{\mathrm{coil}}} \frac{B_n(z)}{R_{\mathrm{ref}}^{n-1}}\ \mathrm{d}z \tag{3.13}$$

Each element therefore measures a different linear combination of integrated multipole strengths, weighted by its geometry. By constructing and inverting the matrix of geometric weights across all elements, the individual integrated multipoles $\int B_n\ \mathrm{d}z$ can be recovered simultaneously up to the order supported by the number of independent elements [114]. For a pure dipole with

a nearly uniform field over the coil cross-section, the expression simplifies to the total effective area $A_{\mathrm{tot},k} = N_k A_{\mathrm{eff},k}$, and the change of average dipole field follows directly as

$$\Delta B_{\mathrm{avg}} = \frac{\Delta U_k}{A_{\mathrm{tot},k}} \tag{3.14}$$

where ΔU_k denotes the change of the integrator output. The absolute field is then recovered by accumulating changes from a reference time t_0 at which the field is independently known,

$$B_{\mathrm{avg}}(t) = \sum_{t_0}^{t} \frac{\Delta U_k(t')}{A_{\mathrm{tot},k}} + B_{\mathrm{avg}}(t_0) \tag{3.15}$$

Because the induced voltage $V(t)$ is measurable in real time during magnet ramps, fluxmeters are naturally suited to pulsed magnetic measurements. The voltage from the coils is digitised and integrated numerically at sampling rates from the kilohertz regime up to several hundred megahertz, depending on the excitation bandwidth and noise floor. Dedicated Fast Digital Integrators (FDIs) are employed when high sampling rates and long integration windows are required simultaneously, but the underlying principle remains the time integration of Eq. (3.10). This dynamic capability allows fluxmeters to follow complete synchrotron cycles and capture static and dynamic field effects that depend on the full excitation history, which is the primary operational mode for the measurements described in subsequent chapters.

The accuracy of a fluxmeter measurement is ultimately limited by how well the effective area $A_{\mathrm{tot},k}$ of each element is known. In a single-element system, a relative calibration error $\delta A_{\mathrm{tot}}/A_{\mathrm{tot}}$ propagates directly as a relative error of the same magnitude in the inferred field. In a multi-element array, each element must be calibrated independently, since manufacturing tolerances in winding geometry, conductor placement, and substrate thickness produce element-to-element variations in both the effective area and the effective radius $R_{\mathrm{coil},k}$. From Eq. (3.13), a relative error $\delta R_{\mathrm{coil},k}/R_{\mathrm{coil},k}$ in the effective radius of element k introduces a relative error of order $n\,\delta R_{\mathrm{coil},k}/R_{\mathrm{coil},k}$ in the inferred integrated multipole of order n [114]. To reach a precision of order 10^{-4} for a sextupole component with $n = 3$, the effective radius must therefore be known to a few parts in 10^{-5}, which illustrates how geometric tolerances tighten progressively with multipole order. Hand-wound coils carry slightly different geometry in each element from manufacturing imperfections, and this per-element variation is absorbed by calibrating each element individually. A PCB-based set instead reproduces the coil geometry across many channels and extracts the harmonics by inverting the geometric weight matrix over all of them. The inter-element reproducibility of the calibrated areas then sets the floor on multipole accuracy, since the matrix inversion amplifies any systematic discrepancy between elements.

Misalignment of the coil set with respect to the magnetic axis introduces an additional source of systematic error by coupling different multipole orders. For a coil offset by δy in the vertical direction inside a dipole that carries a residual normal quadrupole component B_2, the measured vertical field becomes

$$B_y(0, \delta y, z) \approx B_1(z) + B_2(z)\frac{\delta y}{R_{\mathrm{ref}}} + \dots \tag{3.16}$$

so that the inferred dipole field acquires a systematic bias proportional to δy. If $B_2/B_1 \sim 10^{-2}$, a relative misalignment $\delta y/R_{\mathrm{ref}} \sim 10^{-2}$ already introduces a relative dipole error of order 10^{-4} [114]. Analogous couplings arise at all multipole orders, so that the required transverse alignment tolerance scales inversely with the ratio of the dominant to the target harmonic. Additionally, if the coil is shorter than the magnet, displaced longitudinally, or tilted with respect to the magnetic axis, the measured flux samples only a specific axial window of $B_n(z)$ and applies a

modified weighting to the fringe-field regions, and in practice these effects are minimised by mechanical design and absorbed into the calibrated scale factors.

As Eq. (3.11) shows, the integrator output $U(t)$ is defined only up to an additive constant set by the reference flux $\Phi(t_0)$, so a fluxmeter provides only relative measurements of field changes. Every measurement therefore requires a well-defined integration marker, a point in the cycle at which the integrated field is assigned a known value. In static calibration this reference is established by driving the magnet through a reproducible precycle and anchoring the integrator output to an auxiliary high-accuracy sensor such as an NMR probe in a dedicated reference magnet. In pulsed operation like a synchrotron cycle, the marker is imposed at a reproducible point in the machine cycle, after which the fluxmeter tracks relative changes until the next marker. In addition to the marker requirement, practical measurements are affected by small DC offsets and leakage currents at the integrator input. Denoting the effective offset by V_{off}, the measured signal in Eq. (3.11) becomes

$$U_{\text{meas}}(t) = N\big[\Phi(t) - \Phi(t_0)\big] - V_{\text{off}}(t - t_0) \tag{3.17}$$

where the second term represents a linear drift that is indistinguishable from a genuine slow field variation. Over a machine cycle of duration Δt, the relative error introduced by the offset scales as

$$\frac{\Delta\Phi_{\text{drift}}}{\Delta\Phi_{\text{sig}}} \approx \frac{V_{\text{off}}\,\Delta t}{\int_{t_0}^{t_0+\Delta t} |V(t')|\,\mathrm{d}t'}\,, \tag{3.18}$$

which sets a practical lower bound on accuracy even if the coil geometry is perfectly known. In real integrators V_{off} varies slowly with variables like temperature, and external conditions, so the drift is rarely linear and cannot be removed exactly by subtraction over long integration windows. The large loop areas and high turn counts of multi-element fluxmeter sets also make them sensitive to external magnetic disturbances and mechanical vibrations, which couple additional parasitic contributions into $V(t)$ that are indistinguishable from the magnetic signal. These effects are mitigated by rigid mechanical supports, stable cable routing, and twisted-pair wiring, but can rarely be eliminated completely outside of tightly controlled environments.

In compensated coil arrangements that suppress the dominant harmonic, drift, noise, and external pickup become progressively more limiting relative to the residual output as the target multipole order increases. Rotating-coil systems, discussed in the following subsection, employ the same underlying flux integration principle but use the azimuthal motion of the coil to extract static multipole content with high sensitivity and reduced susceptibility to long-term integration drift.

3.2.6 Rotating Coils

A rotating coil applies the same flux-integration principle as the fluxmeter, and by rotating an induction coil at constant angular velocity through the aperture it maps the azimuthal Fourier harmonics of the induced voltage onto the normal and skew multipole coefficients B_n and A_n, recovering the full harmonic spectrum from a single revolution [9, 114]. Because the coil returns to its starting position each revolution, a linear integrator offset is periodic in the rotation angle and cancels from the extracted coefficients, so the drift accumulates only over one rotation period of order seconds and not over a full machine cycle, which makes the rotating coil the standard technique for production characterisation and acceptance testing of accelerator magnets [9, 114].

The Fourier decomposition requires the field to remain static over one full rotation, and during a synchrotron ramp the field changes within a single rotation period, so the extracted coefficients acquire systematic errors that grow with the rate of field change [114]. A non-rotating single-shot

integration remains possible during a ramp, though it reduces to a fluxmeter-type measurement without the harmonic resolution or drift resistance of the rotating mode. The rotating coil is therefore a static-field harmonic instrument, and it is not among the measurements used in this work.

3.3 Magnetic Measurements at CERN

3.3.1 SPS Main Dipole Measurement Bench

The main dipole bench of the CERN magnetic measurement laboratory is a fluxmeter-based system, and the measurements of this thesis were taken on the spare MBB dipole of the SPS main ring. Setup and calibration of the bench were carried out by the laboratory together with the magnetic measurement teams, and the sensing-coil effective area that scales the integrated flux to the average field was established as part of that calibration. The calibration procedures and the associated accuracy limits follow standard laboratory practice [114], and the resulting absolute accuracy is taken as given for the field measurements of Chapter 6.

3.3.2 SPS Main Quadrupole Measurement Bench

3.4 Online B-Train Measurement Systems

In a synchrotron, the RF frequency must track the beam revolution frequency in real time according to the synchronism condition in Eq. (2.15). As the beam accelerates, the momentum and therefore the required RF frequency evolve continuously over the cycle, and maintaining this synchronism requires knowledge of the instantaneous average bending field $B(t)$. For machines in which the particle velocity changes substantially during the cycle, the RF frequency sweep is large enough that deriving the required frequency from a programmed current function alone is insufficient at the required precision. In those cases, a direct, continuous measurement of the bending field is operationally necessary.

The concept of measuring this field in real time from an induction coil wound on a reference magnet was pioneered at CERN for PS, where the combined-function lattice makes a direct field measurement essential [29, 111]. The measurement system that performs this function is commonly referred to as the B-train, which originally referred to the continuous real-time signal chain that distributed the average bending field estimate to the RF system, the power converters, and the beam instrumentation [27, 29]. The principle has since been adopted at accelerator facilities worldwide, including medical synchrotrons where cycle-to-cycle field reproducibility is a direct requirement for dose precision [48, 52, 96].

3.4.1 Operating Principle of a B-train

For a standard B-Train, a sensing coil is wound on a reference magnet powered in series with the main ring dipole circuit, so the reference magnet receives the same excitation as the ring dipoles. The universally used field measurement instrument for a B-Train is the fluxmeter coil, which provides a continuous measurement of the field integral as described in Section 3.2.5. For most machines using a B-Train, the average dipole field from Eq. (3.7) suffices to describe the beam dynamics and to derive the required RF frequency. However for machines with significant fringe fields or non-trivial design such as sectored magnets, the full integral field $B\,\mathrm{d}\ell$ from Eq. (3.4) is used.

In practical designs, an FDI digitises the coil voltage at high sampling rates at several hundred kilohertz, or even megahertz, and performs the numerical integration in real time. This makes the instantaneous field estimate $B_{\mathrm{avg}}(t)$ available to the RF control system and power converters at rates sufficient to follow the field ramp at ramp rates exceeding 1 T/s while integrating out irregular small amplitude noise in the voltage measurements [27, 29]. As synchotrons operates cycle-by-cycle, the B-Train provides a cycle-by-cycle measurement of the field, and anchores the integration with a field marker that fires when the field passes through a calibrated reference value. This bounds the accumulation of integrator drift to the interval between successive resets. At CERN, all deployed B-trains currently use NMR-based markers for this purpose [2, 27], and marks before injection to avoid large field jumps when beam is present and the integrator is reset to zero. The accuracy and drift limitations that govern the B-train measurement remain those of the fluxmeter integration chain discussed in Section 3.2.5.

A limitation of the B-train beyond the fluxmeter-bound ones is the requirement for near-constant uptime. In practice, the effective voltage offset V_{off} of the integration chain seen in Eq. (3.17) is not stable over time but drifts on timescales ranging from minutes to hours depending on temperature, electronics load, and cable and external conditions [60]. While offline calibration suffices for laboratory measurements, constant, or near-constant uptime of the B-train system means that the offset correction must be determined and maintained in real time in order to avoid accumulated drift that exceeds beam dynamics tolerances. The standard procedure is to create an electronic short-circuit at the integrator input, disconnecting the sensing coil and measuring the residual voltage bias directly, and this offset is then subtracted from subsequent integrations [2]

3.4.2 B-Train Implementations

At CERN, B-train systems are deployed on all synchrotrons in the injector chain, distributing the real-time field estimate to the RF system, the power converters, and the beam instrumentation over optical White Rabbit [120] timing links [2, 27]. The B-train output serves as the primary input to the RF phase and radial position loops, which adjust the RF frequency to maintain synchronism with the circulating beam. In machines regulated on field instead of current, the B-train closes a feedback loop on the main magnet power converter on the dipole field exclusively, which suppresses cycle-to-cycle variability in the bending field such as those caused by hysteresis.

The PS is the original B-train machine, where the combined-function lattice cannot operate on a simulated field function alone and therefore uses field regulation, with a reference magnet identical to the main ring units [29, 111]. At the PSB, LEIR, and Antiproton Decelerator (AD), B-trains likewise serve both RF control and dipole field regulation [2, 60]. In Extra Low Energy Antiproton Ring (ELENA), which decelerates antiprotons to 100 keV at bending fields near the noise floor of standard induction coils, a curved fluxmeter following the dipole arc raises the effective coil area and the signal-to-noise ratio [62, 75].

At the SPS, the B-train serves a different role from the smaller synchrotrons, because the injection energy is sufficiently high that the revolution frequency is nearly constant over the cycle, the B-train output is not required for RF control, and static or simulated fields can be used to set the RF frequency. It is neither used for field-regulation, as the SPS main magnets are regulated in current. This B-train serves as a core part of the data acquisition for this thesis, and is more thoroughly presented in its own Section 3.4.3.

Beyond CERN, B-trains are used at medical synchrotrons where bending-field reproducibility sets dose precision, including National Center for Oncological Hadrontherapy (CNAO) for both RF control and reproducibility monitoring [48, 52]. At MedAustron, the original B-train was superseded by a Hall-probe-based measurement after integrator drift over its long cycles became

the dominant error [95, 96]. Long cycles, demanding absolute accuracy, or infrequent marker resets mark the general boundary of the approach, and make alternative real-time measurements preferable.

3.4.3 SPS B-Train

The SPS B-train measures the average dipole field from two reference magnets of the main dipole types, one of the MBA type and one of the MBB type, powered in series with the main dipole circuit and installed in building 870 of the BA3 area [22, 135]. Both reference magnets contain their nominal vacuum chambers, so the measured integral includes any chamber-wall eddy currents, as discussed in Section 2.6.1.

A single long integral coil spans the length of each reference magnet and feeds one integration channel [134]. The coil output is proportional to the rate of change of the linked flux, and the average field is recovered by integrating that voltage and fixing the integration constant with a low-field NMR marker, following the fluxmeter relation of Section 3.2.5 [134].

The installation carries two parallel measurement chains, an operational chain and a spare chain, which provides redundancy for the measurement [22]. Four NMR probes are installed in the reference magnets, of which one remains functional and is connected to the spare chain alone, so the field marker is present on the spare chain and absent on the operational chain. Acquisition and integration are handled by the FIRESTORM digital system, an FPGA-based flux integrator that distributes the field estimate over a White Rabbit network at a steady rate of 250 kHz with a resolution of 10 nT [60, 134]. The operational use of the marker, the per-cycle integration reset, the data logging, and the offset recalibration are described in Section 4.5.2.

Chapter 4

Beam Operation at CERN and the SPS

[TODO: where to put idle current (155 A) and field (0.625 T) for the main dipoles, or concept of idle current?]

The CERN accelerator complex operates a chain of synchrotrons that each prepare and deliver beam to a variety of experiments and downstream machines. At the center of this chain for the work presented in this thesis is the SPS, which functions as a multi-cycling accelerator serving several different beam destinations within a single repeating sequence of machine cycles. Each cycle accelerates beam to a different energy or for a different purpose, and these cycles are interleaved in a fixed supercycle that repeats continuously during operation.

The accelerator control system that governs this operation is a layered architecture spanning settings management, timing, communication middleware, data logging, postmortem, power converters and hardware controllers. This architecture determines how cycle settings are defined, structured, and propagated from high-level beam physics quantities down to magnet hardware, how operational data is recorded and made available for offline analysis, and how runtime corrections can be computed and applied within the control chain. These aspects of the control system are directly relevant to the work of this thesis, as they set the practical constraints on data acquisition, model design, and the deployment of the compensation strategy developed in the following chapters.

Section 4.1 introduces the cycle and supercycle as the fundamental units of SPS operation. The main SPS cycle families and their role in the machine programme are described in Sections 4.2 and 4.3. The accelerator controls infrastructure is presented in Section 4.4, covering settings management, timing, middleware, logging, and hardware controllers. The field control hierarchy in the injector chain and the operational procedures used to maintain magnetic, and therefore beam, reproducibility are covered in Sections 4.5 and 4.6. The machine development framework within which experimental validation in Chapter 8 is carried out is introduced in Section 4.7.

4.1 Accelerator Operation in Cycles

At CERN, each accelerator in the injector chain operates according to a programmed sequence of settings that is executed repeatedly. One complete execution of this sequence is referred to as a machine cycle. A cycle defines the settings of all accelerator equipment for one period of operation, including the magnets, radio frequency systems, beam instrumentation, and timing devices. It is therefore not a purely magnetic concept but a complete operational programme for the machine, even though the term magnetic cycle often appears in the literature and is used colloquially to refer to the same concept. Within a cycle, the sequence is divided into logical phases called beam processes, which correspond to distinct operational stages such as pre-injection, flat bottom / injection, acceleration ramp, flat top / extraction, and ramp-down. Different settings can in principle be assigned independently to each beam process within the settings management system, described in Section 4.4.1, while the hardware layer is agnostic to this and operates without knowledge of the process boundaries.

The CERN injector chain, introduced in Section 2.2, consists of several synchrotrons that each operate according to their own cycle structure. The duration of a machine cycle is always an integer multiple of the so called Basic Period (BP) of 1.2 s. The number of basic periods per cycle depends on the machine and the operational requirements and beam parameters, and can vary from cycle to cycle within the same machine In the proton injector chain, Linear Accelerator (LINAC)4 and the PSB operate synchronously with cycles of one basic period. PS cycles span one or two basic periods depending on the beam destination, energy and other beam parameters, with the two-period cycles typically used for the 26 GeV/c LHC-type beam. SPS cycles consist of three or more basic periods, with operational cycles reaching up to approximately 50 basic periods for heavy ion operation. The ion injector chain follows the same basic period convention, with LINAC3 and LEIR operating with two to three basic periods per cycle, and the PS and SPS using the same cycle structures as for protons but with different settings and timing.

In addition to real machine cycles, a dedicated placeholder cycle called the ZERO cycle occupies basic period slots in the supercycle where no beam is required, or not desired for operational reasons. This could also be caused by desire to keep the supercycle at a fixed length, or to reduce stress on the machine by inserting a break in the beam delivery. In a ZERO cycle, all magnets are held at their idle current rather than ramped, and most equipment is placed in a standby or ready state instead of cycled. The ZERO cycle is always one basic period long, regardless of the machine in which it appears, and serves as the standard filler slot across the PSB, PS, and SPS supercycles.

In addition to the standard supercycle mechanism with cycles, the control system supports the concept of normal and spare cycles. A normal cycle can be paired with a designated spare cycle of the same length, and the two can be exchanged dynamically with some advance notice, without requiring a full supercycle change. This mechanism allows beam clients to request a change at shorter notice than the supercycle turnaround time would otherwise permit. As an example, a facility such as the AD [13], which receives beam every 2 minutes, may use a nearby spare slot to activate the spare cycle in the PS supercycle rather than waiting for the next supercycle iteration. The normal and spare mechanism normally used in both the SPS and the PS, and the SPS case will be elaborated in Section 4.3.

The ordered sequence of cycles executed by a machine is called the Supercycle [16, 87, 88]. Each position in the supercycle is referred to as a cycle slot, and one supercycle repeats continuously during operation. The composition of the supercycle is determined by the beam requests of all downstream users across the accelerator complex, and is coordinated at the level of the full injector chain. The structure and length of the supercycle is primarily set by the SPS, since it operates the longest individual cycles, and typically ranges between 30 s and 60 s, with occasional extensions up to approximately 75 s depending on the operational programme. Changes to a supercycle composition are always initiated manually by a human operator, and changing which supercycle is played similarly requires operator action. In practice, the supercycle composition is therefore static during operation, and for the SPS, different supercycles exist for different operational programmes, such as fixed-target physics, LHC injection, or machine development, and is described in more detail in Section 4.3. Additionally since 2025, a system known as Dynamic Beam Scheduling (DBS) [70] automates the filling and optimization of PS and PSB supercycle slots to improve machine usage when new supercycles are created, while the SPS supercycle composition or templates remains under manual operator control. Notably, the LHC does not operate on a cycle structure in the same sense, as it is a storage ring that holds beam continuously, and injection time and phase is already timed relative to the SPS supercycle, so the LHC orchestration is not considered in the supercycle context besides the accelerated beam in the SPS

Each cycle slot in the supercycle is associated at runtime with a user identifier, referred to simply as a user. The user is the runtime label under which devices execute their settings and publish their data for that cycle slot. In the settings management system, a cycle is the object that holds the complete configuration of all devices for one slot, and a user is mapped to a cycle before execution. This distinction between the user as a runtime concept and the cycle as a settings concept is used throughout this thesis when referring to logged operational data, and is discussed further in Section 4.4.1.

The ability to hold independent settings per user simultaneously is referred to as Per-Pulse Modulation (PPM), or multiplexing, and devices or parameters that support this are described as PPM-capable or enabled, or simply as multiplexed. A PPM-capable device can apply a different configuration for each user as the supercycle advances, which is the hardware mechanism that makes multi-cycling operation possible. Devices or parameters that are not PPM-capable hold a single global setting regardless of which user is currently playing, and must therefore be set to values compatible with all cycles in the supercycle simultaneously. The distinction between PPM and non-PPM parameters is relevant for the compensation strategy developed in this thesis, as it determines which settings can be varied independently per cycle and which cannot, and is discussed further in Section 4.4.1.

4.2 Main SPS Cycles

The SPS serves several beam destinations, each imposing distinct requirements on beam energy, intensity, emittance, and extraction mode, as discussed in Chapter 2. These requirements translate into different momentum programmes, optics configurations, and cycle lengths, and the resulting cycle families form the building blocks of the SPS supercycles described in Section 4.3. Table 4.1 lists the operational cycle families with their optics, injection and extraction momenta, extraction mode, and length, and the remainder of this section describes only the properties that bear on magnetic history and field reproducibility. A dedicated magnetic precycle MD1 precedes fixed-target proton cycles in the standard supercycle and is described in Section 4.6.1.

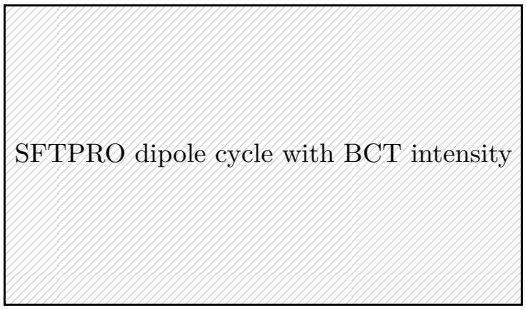
4.2.1 SPS North Area Fixed Target Proton Cycle (SFTPRO)

The SPS Fixed Target Proton (SFTPRO) cycle, commonly referred to as SFT, is the primary fixed-target proton cycle of the SPS and delivers beam to the North Area experiments through the TT20 transfer line. It operates with integer tune 26 (Q26) optics, with fractional working point 26.62 in the horizontal and 26.58 in the vertical plane, and receives two Multi-Turn Extraction (MTE) injections from the PS at $14\,\mathrm{GeV}/c$, at 200 ms and 1400 ms, reaching about 1.3×10^{13} protons per cycle. This injection scheme with Q26 optics at $14\,\mathrm{GeV}/c$ is referred to as an SFT-type beam. The ramp begins at 1500 ms and reaches the flat top at 4460 ms, from where beam is extracted at $400\,\mathrm{GeV}/c$ by chromatic slow extraction, described in Section 2.1.9.1, over a 4800 ms spill before ramp-down. As shown in Figure 4.1, the QF and QD quadrupole circuits follow the momentum programme to maintain the optics, the sextupole circuits are ramped in proportion to momentum with transmission corrections added on top, and the octupole circuits provide Landau damping.

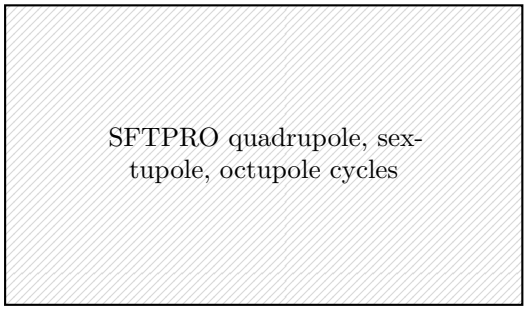
Injection at 0.625 T is the lowest field reached with beam in the SPS, so a fixed absolute field error produces its largest relative error there. The chromatic slow extraction requires reproducible dipole, quadrupole, and sextupole fields at the 1.8 T flat top, and with poor reproducibility the extracted spill rate changes and can trigger operator retuning or an early dump of high-intensity beam to protect the electrostatic septa, as reported by Velotti et al. [135].

Tab. 4.1: Main SPS operational cycle families with their optics, injection and extraction momentum, extraction mode, and length. Ion injection momentum is given per nucleon, depending on particle species, and ion extraction momenta are proton-equivalent. LHC-type cycles use Q20 optics in standard operation and Q26 for special runs such as proton-oxygen in 2025. Machine development cycles are configurable, and the two representative short shapes used in this thesis are listed. The dedicated magnetic precycle MD1 takes no beam and is described in Section 4.6.1.

Cycle	Optics	p_{inj} [GeV/c]	p_{ext} [GeV/c]	Extraction	Length
SFTPRO	Q26	14	400	slow (chromatic)	10.8 s (9 BP)
LHC Pilot/INDIV	Q20	26	450	fast	13.2 s (11 BP)
LHC 4-injection	Q20	26	450	fast	20.4 s (17 BP)
AWAKE	Q20	26	400	fast	7.2 s (6 BP)
HiRadMat Pilot	Q20	26	440	fast	8.4 s (7 BP)
HiRadMat 4-inj	Q20	26	440	fast	**[TODO:** 20.4 s **(17 BP)]**
SHiP (Run 4)	Q26	14	400	slow (chromatic)	7.2 s (6 BP)
SFTION	Q26	13/17	33–300	slow (chromatic)	**[TODO: length]**
MD 14 GeV	Q26	14	200	dump	7.2 s (6 BP)
MD 26 GeV	Q20	26	200	dump	7.2 s (6 BP)



(a) Main dipole field and normal operational beam intensity



(b) Quadrupole, sextupole, and octupole circuits

Fig. 4.1: The SFTPRO machine cycle. Left: main dipole field $B(t)$ as a function of cycle time, overlaid with the beam intensity measured by the BCT. The two-injection pattern is visible at 200 ms and 1400 ms, followed by the acceleration ramp, the 4800 ms slow extraction spill, and the ramp-down. Right: QF or QD quadrupole, sextupole, and octupole circuit functions on the same time axis.

4.2.2 Remaining Cycle Families

Cycle families of Table 4.1 differ mainly in optics, injection momentum, and extraction mode. The fixed-target cycles SFTPRO, SPS Fixed Target Ion (SFTION), and the future SHiP cycle use Q26 optics, while the LHC-type, AWAKE, and HiRadMat cycles use Q20 optics, and the LHC ion cycles use Q26 against the Q20 of the LHC proton cycles. Because the stronger focusing of the Q26 working point requires a larger normalized quadrupole gradient, the Q26 cycles drive the quadrupole circuits to higher excitation currents than the Q20 cycles at the same momentum.

The LHC-type, AWAKE, and HiRadMat proton cycles reach 400 GeV/c–450 GeV/c at extraction and drive the main dipole close to its highest operational field, and are extracted in a

single turn by fast extraction. The ion fixed-target cycles SFTION use the same chromatic slow extraction as SFTPRO, with less stringent spill-quality tolerances, and ion operation is not the main focus of this thesis. Machine development cycles are not fixed by a beam destination and can take an arbitrary shape and length, and the two typical short six-basic-period shapes used in this thesis inject under SPS Fixed Target (SFT)-type or LHC-type conditions and ramp to $200\,\mathrm{GeV}/c$, the same peak field as the MD1 precycle of Section 4.6.1.

In every cycle the sextupole and octupole circuits are driven through a bipolar excursion at the end of the cycle instead of ramping down with the dipole and quadrupole circuits, visible in the higher-order traces of Figure 4.1. This end-of-cycle excursion is described as quasi-degaussing in Section 4.6.2.

4.3 Typical SPS Supercycles

The composition of the SPS supercycle during any given operational period is determined by the active beam programme. As introduced in Section 4.1, each slot in the supercycle is associated with a user identifier at runtime, and each user is mapped to a cycle configuration at the settings level. For standard operational cycles, the mapping between a user and its cycle remains fixed throughout the year, and in this context the two terms are used interchangeably in the following descriptions.

The supercycle composition is static during operation and can only be changed by a manual operator action. During physics operation, supercycle changes follow well-defined patterns driven by the requirements of the active beam programme, and arbitrary cycle insertion in general not done. The exception is beam commissioning and dedicated machine development time, during which cycles can be inserted with greater flexibility, as described in Section 4.7.

4.3.1 Proton run supercycles

For the majority of the operational year, the SPS runs the standard fixed-target physics supercycle, consisting of alternating SFTPRO and MD1 slots:

$$\texttt{SFTPRO1} \rightarrow \texttt{MD1} \rightarrow \texttt{SFTPRO1} \rightarrow \texttt{MD1} \rightarrow \texttt{SFTPRO1} \rightarrow \texttt{MD1} \tag{4.1}$$

This configuration repeats continuously and typically appears in 2 to 3 repeats of the SFTPRO1 $\rightarrow$ MD1 pattern, depending on the requirements of the accelerator complex in supercycle length. The amount of beam extracted to the North Area is unaffected since the same sequence repeats. This sequence is shown in Figure 4.2.

LHC injection interrupts the standard fixed-target configuration for 30 minutes to approximately one hour at a time, repeating every two to twenty hours depending on LHC demand.

The corresponding supercycle takes the form:

$$\texttt{SFTPRO1} \rightarrow \texttt{ZERO} \rightarrow \texttt{LHC1} \rightarrow \texttt{MD1}$$

where the ZERO slot before the LHC cycle provides a timing gap without interrupting the magnetic sequence. Normally, a parallel spare supercycle is prepared where the LHC1 slot is replaced by the shorter LHCINDIV or LHCPILOT, with additional ZERO cycles trailing to reach the same length as the standard LHC supercycle. The supercycles are shown in Figure 4.3 and Figure 4.4, respectively. An additional sequence is used for dedicated LHC pilot beam, which can be a dedicated sequence for LHC testing or for machine development, to avoid the longer LHC1 sequence in Figure 4.3.

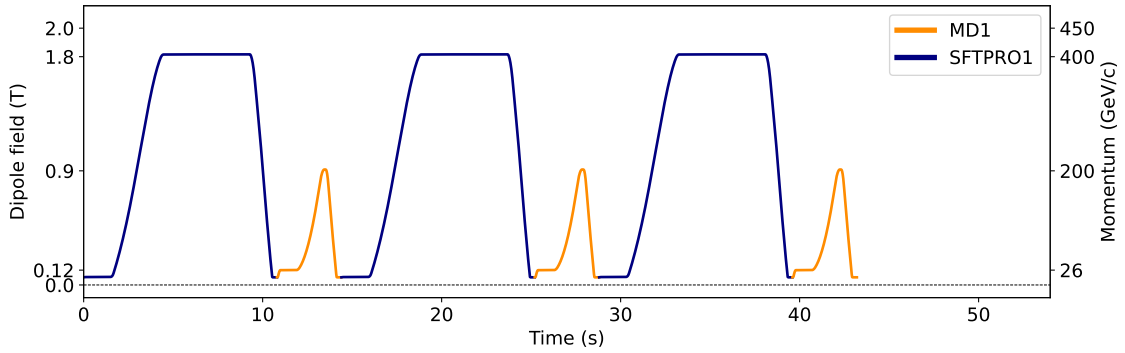


Fig. 4.2: Standard SPS (North Area) fixed-target supercycle, alternating SFTPRO(1) and MD1 slots. Can be reduced to only 2 repeats on discretion of beam operation.

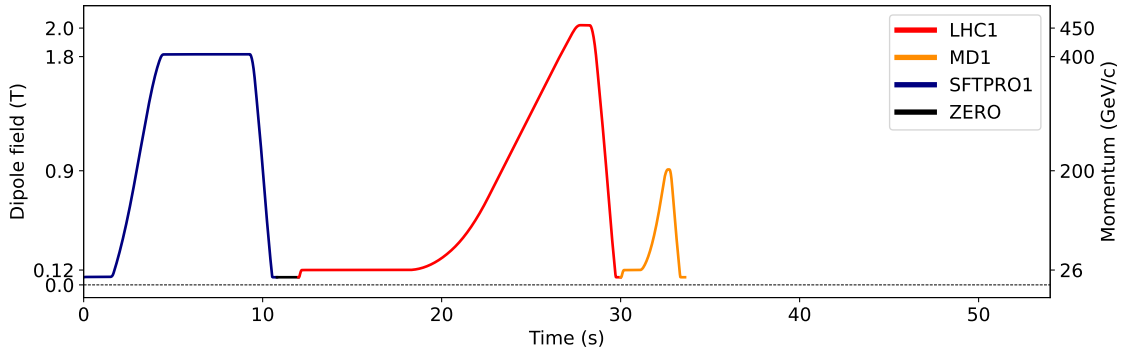


Fig. 4.3: SPS supercycle during LHC filling, with a ZERO slot before the LHC1 slot to provide a timing gap without interrupting the magnetic sequence.

AWAKE runs are typically one to two weeks in duration and use the following structure:

$$\texttt{SFTPRO1} \rightarrow \texttt{AWAKE1} \rightarrow \texttt{MD1} \rightarrow \texttt{SFTPRO1} \rightarrow \texttt{AWAKE1} \rightarrow \texttt{MD1}$$

HiRadMat runs are similarly one to two weeks in duration, and AWAKE and HiRadMat never run simultaneously. Two HiRadMat supercycle variants exist depending on the required cycle length:

$$\texttt{SFTPRO1} \rightarrow (\texttt{ZERO}) \rightarrow \texttt{HIRADMT1} \rightarrow \texttt{MD1} \rightarrow \texttt{SFTPRO1} \rightarrow (\texttt{ZERO}) \rightarrow \texttt{HIRADMT1} \rightarrow \texttt{MD1}$$

$$\texttt{SFTPRO1} \rightarrow (\texttt{ZERO}) \rightarrow \texttt{HIRADMT2} \rightarrow \texttt{MD1}$$

where the ZERO can be inserted at the discretion of the operators. The sequences are visualized in Figures 4.7 and 4.8, respectively. LHC filling is never combined with AWAKE or HiRadMat in the same supercycle. Whenever LHC filling is requested, the supercycle is switched to the LHC1 configuration described above, and the AWAKE or HiRadMat supercycle is paused until LHC filling is complete. Notably for the SFT, the cycle is always present in every supercycle during the physics run, and is always preceded by an MD1.

Machine development cycles can be inserted alongside standard fixed-target operation. Note that Machine Development (MD) cycles are typically 6 basic periods (7.2 s), while MD1 is 3 basic periods (3.6 s), and the two are not magnetically interchangeable. If operators deem the

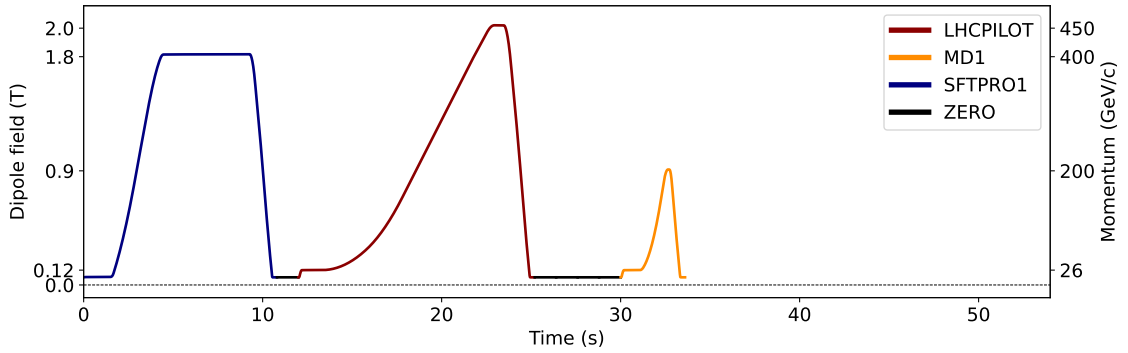


Fig. 4.4: SPS supercycle during LHC filling, but with a shorter, single injection LHCPILOT cycle instead of the standard 4 injection LHC1 cycle. The trailing ZERO cycles are added to reach the same length as the standard LHC supercycle.

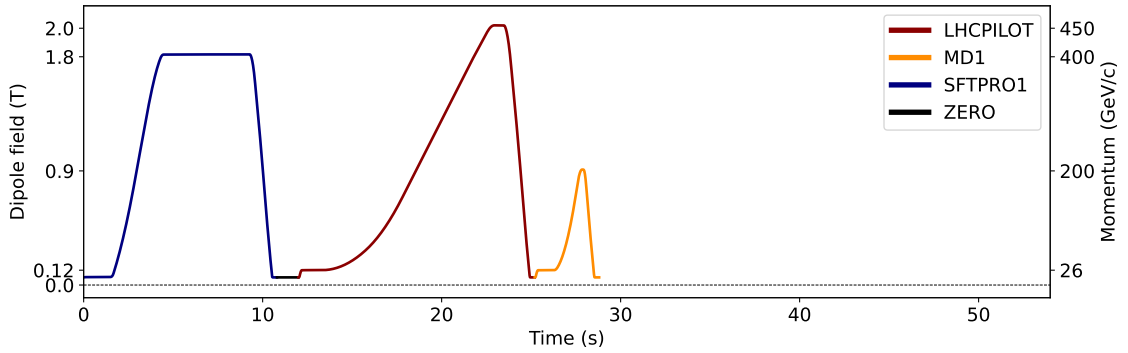


Fig. 4.5: SPS supercycle during LHC pilot beam, which can be a dedicated sequence for LHC testing or for machine development, to avoid the longer LHC1 sequence in Figure 4.3.

MD cycle magnetic programme sufficiently different from MD1, an explicit MD1 slot is retained between the MD and the following SFTPRO slot:

$$\texttt{SFTPRO1} \rightarrow \texttt{MD} \rightarrow \texttt{MD1} \rightarrow \texttt{SFTPRO1} \rightarrow \texttt{MD} \rightarrow \texttt{MD1}$$

MD cycles are not inserted in supercycles containing AWAKE, HiRadMat, or LHC filling slots as those are dedicated physics operation sequences.

The future SHiP cycle, planned for Run 4, will be served by a dedicated supercycle in which multiple SHiP slots follow a single SFTPRO cycle:

$$\texttt{SFTPRO1} \rightarrow \texttt{SHIP} \rightarrow \texttt{SHIP} \rightarrow \texttt{SHIP} \rightarrow \texttt{SHIP} \rightarrow \texttt{SHIP} \rightarrow \texttt{MD1}$$

which can be seen in Figure 4.10. The addition of multiple consecutive SHiP slots will alter the magnetic history pattern seen by all co-running cycles, making the field prediction and compensation methods developed in this thesis increasingly relevant in Run 4.

4.3.2 Ion run supercycles

Ion operation uses a supercycle structure analogous to the proton case, without the AWAKE and HiRadMat slots, which are proton-only in Run 3. Proton and ion cycles can share a supercycle

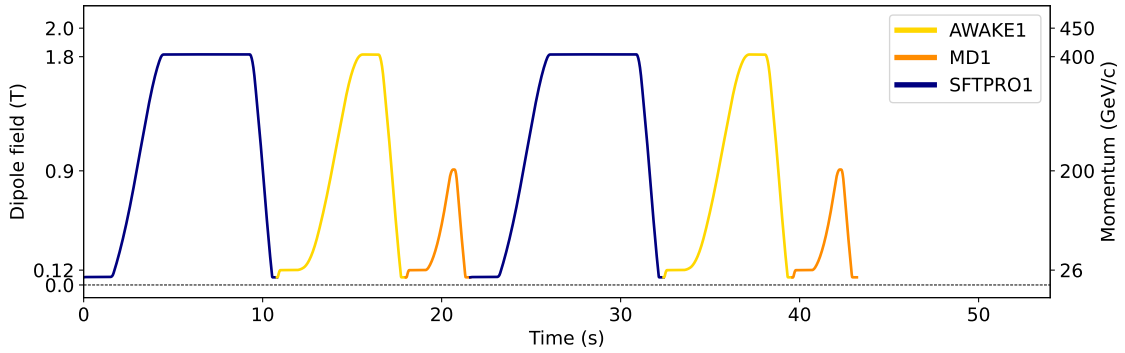


Fig. 4.6: SPS supercycle during standard AWAKE operation, with repeating SFTPRO(1), AWAKE1, and MD1 cycles.

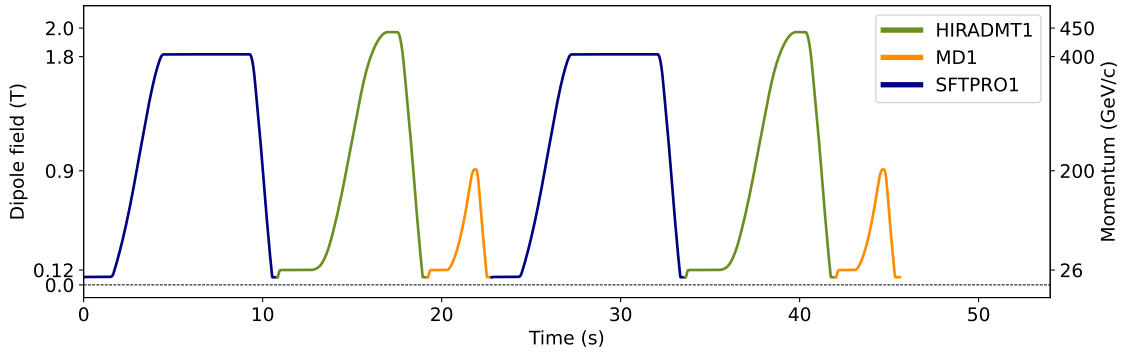


Fig. 4.7: SPS supercycle for HIRADMT1, or HiRadMat pilot cycle. The HIRADMT1 can be preceeded by a ZERO on the discretion of the operators.

during the transition from the fixed-target proton run to the ion run, while SFTPRO and SFTION cannot be scheduled together because of a hardware interlock on the TT20 transfer line. The standard ion fixed-target supercycle alternates SFTION and MD1,

$$\texttt{SFTION1} \rightarrow \texttt{MD1} \rightarrow \texttt{SFTION1} \rightarrow \texttt{MD1}$$

with its length restricting it to two repetitions per supercycle.

LHC ion filling mirrors the proton case, with an LHC ion cycle in place of the proton one and additional MD1 slots padding the length to match timing constraints across the injector complex,

$$\texttt{SFTION1} \rightarrow \texttt{ZERO} \rightarrow \texttt{LHCION1-3} \rightarrow \texttt{MD1}$$

$$\texttt{SFTION1} \rightarrow \texttt{ZERO} \rightarrow \texttt{LHCION1} \rightarrow \texttt{MD1} \rightarrow \texttt{MD1} \rightarrow \texttt{MD1}$$

where LHCION1 through LHCION3 denote single-bunch to multi-bunch ion filling cycles. A proton LHC pilot cycle at Q20 or Q26 is also run during the ion period when the LHC requires it, and a lower-energy SFTION variant reaching 33 GeV/c, below the standard 300 GeV/c, ran in 2025 for low-energy North Area fixed targets following the same structure.

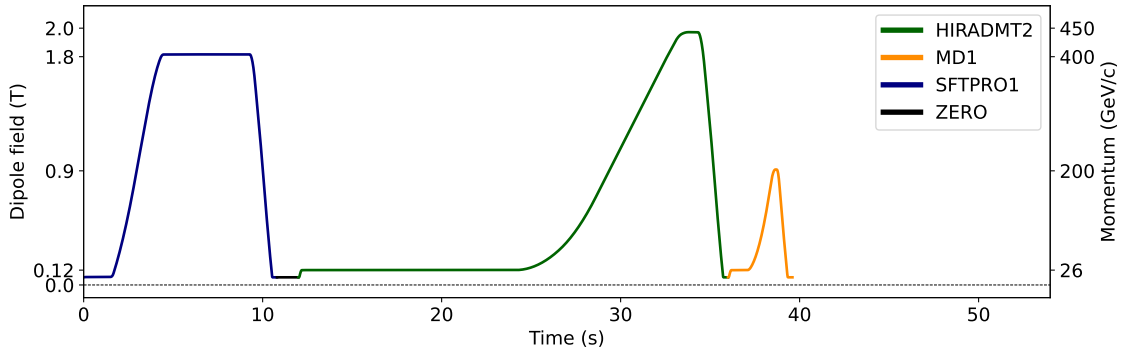


Fig. 4.8: SPS supercycle for HIRADMT2, the standard HiRadMat high-intensity physics cycle. The HIRADMT2 can be preceeded by a ZERO on the discretion of the operators.

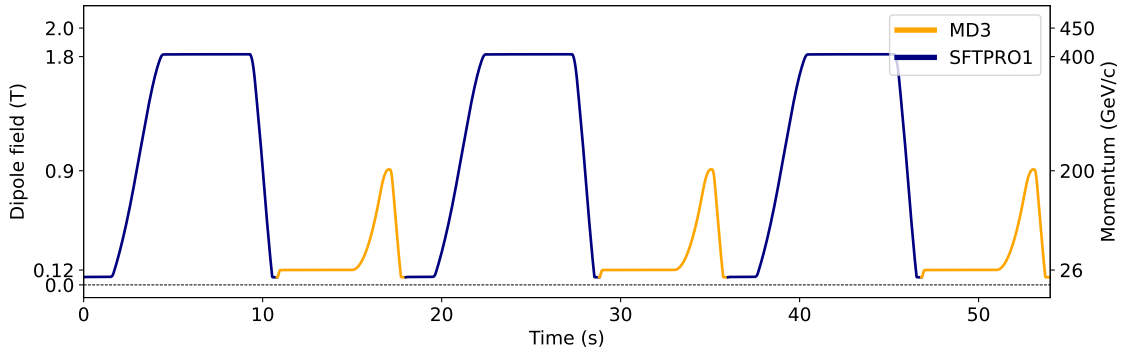


Fig. 4.9: SPS supercycle with a MD cycle inserted alongside standard fixed-target operation. In this case, an MD1 is not inserted between the MD cycle and the following SFTPRO cycle, as the MD cycle is deemed sufficiently similar to MD1.

4.4 Accelerator Controls Infrastructure at CERN

The operation of the CERN accelerator complex relies on a layered controls architecture that separates the definition of high-level physics settings from their execution at the hardware level. It is organized into three tiers, the presentation tier, the business tier, and the front-end tier, connected vertically through a common middleware layer and synchronized by a dedicated timing system. Most of the work in this thesis sits in the business tier, with a targeted extension of the presentation tier to expose the compensation scheme to operators.

Every addressable unit in the hierarchy, whether physical hardware or a purely virtual software entity, exposes its data and settings through the same Device/Property interface [11], where each property can independently be read, written, or subscribed to so that a client receives only the data it requests. A device holds not a single value for a property but a table of values, one for each USER, a mechanism known as PPM [121] that realizes the physics-level cycles and supercycles of Section 4.3 at the hardware level, with the timing system of Section 4.4.2 selecting the active USER. A cycle is a settings object managed in the business tier, while a USER is the runtime identity under which the front-end tier executes [16, 87, 88].

The business tier hosts the components that define, manage, and process settings and operational data, namely the settings management system LHC Software Architecture (LSA), the logging system NeXt Controls Accelerator Logging Serviceg (NXCALS), and the online

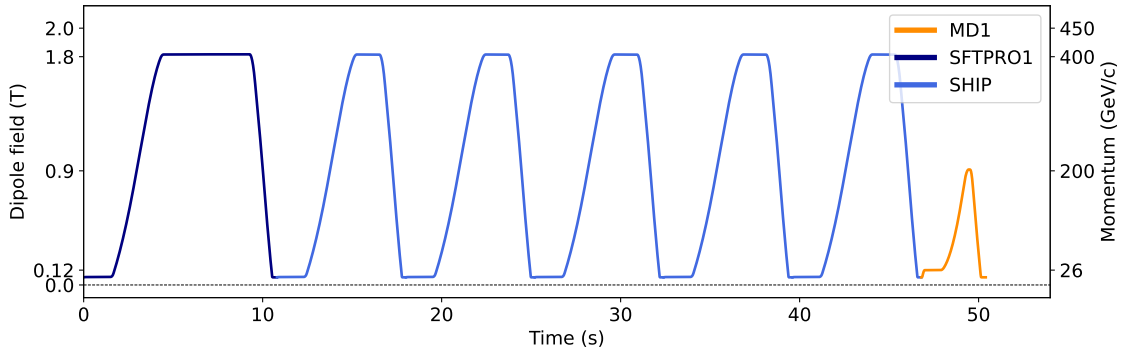


Fig. 4.10: SPS supercycle for the future SHiP experiment, with a single SFTPRO(1) slot followed by multiple SHIP slots and a trailing MD1.

processing framework Unified Control, Acquisition and Processing (UCAP), and communicates with the front-end tier over the Controls MiddleWare (CMW) and Remote Device Access (RDA) middleware [39, 43, 81, 112, 141]. Devices in the front-end tier execute settings on hardware and are not architecturally uniform, with most built on Front-End Software Architecture (FESA) [49], such as the B-Train reference magnet system of Section 3.4.3 and the BPMs, while the main power converters are regulated by the FGC, a dedicated embedded controller interfaced through a gateway front-end computer.

Not all communication follows this vertical structure, since some front-end devices exchange data directly over dedicated real-time networks that bypass the business tier. The field measured by the B-Train is transmitted to the FGC over White Rabbit for closed-loop field regulation in this way, a horizontal path distinct from the timing system that broadcasts synchronization events on the same underlying technology. The following subsections examine these components in the order in which a setting is defined, executed, and logged, beginning with LSA and ending with the FGC, and the front-end internals of FESA are left outside the scope of this thesis.

4.4.1 LSA

The settings of the SPS, and of all CERN accelerators, are managed through LSA, the software system used across the complex to define, structure, and deploy operational settings [43, 112]. It provides the abstraction layer between the accelerator physics models and the low-level control system, and represents the full parameter set of all devices for one operational cycle as a single object, called a Cycle or Context. A context can be cloned and settings copied between contexts, so that operators derive new configurations from existing ones without redefining every parameter.

LSA organises settings in a hierarchy, illustrated in Figure 4.11, in which high-level parameters such as the beam momentum and the working-point tune are converted into dependent lower-level parameters through transformations called make rules, down to leaf nodes that are either virtual or driven to hardware. The momentum converts to the main bending field through the rigidity relation of Equation 2.1, and the tune converts to a normalized quadrupole gradient K through a stored optics model and then to an unnormalized gradient through Equation 2.8, with both the dipole field and the quadrupole gradient converted to a reference current IREF at the final step. For the main magnets this last conversion uses a calibration function instead of an analytic expression, shown in Figure 4.12, since the magnetic response is nonlinear and is characterised experimentally in the laboratory. More elaborate correction models for this step exist at

other CERN accelerators, for instance for magnetic decay and snapback in superconducting magnets [116], and no equivalent model is available for the SPS main magnets.

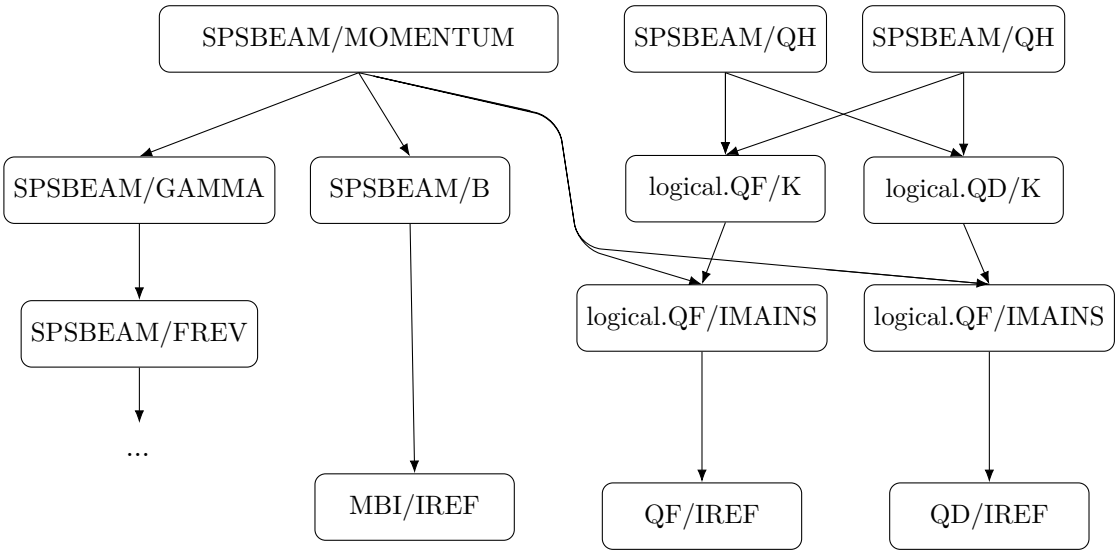


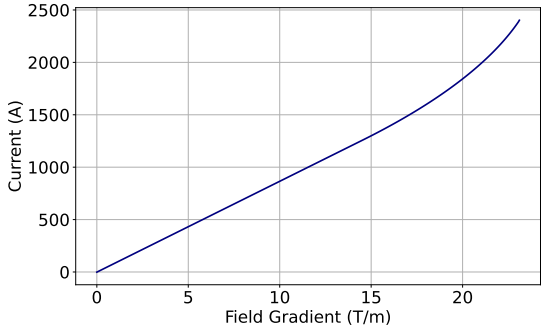
Fig. 4.11: Schematic representation of the LSA settings hierarchy with some example parameter examples, showing how high-level parameters are converted into lower-level parameters through make rules. Here, `SPSBEAM/MOMENTUM` is the design beam momentum p per nucleon, `SPSBEAM/B` the corresponding magnetic field B in the dipoles, and `MBI/IREF` the reference current I required to produce that field in the main bending magnets. The same applies for `SPSBEAM/QH` and `SPSBEAM/QV` which are the horizontal and vertical tunes, respectively, with the `QF/IREF` and `QD/IREF` being the reference currents for the quadrupole families that control the tunes. The intermediate logical parameter `logical.{QF,QD}/K` is the normalized quadrupole gradient. `SPSBEAM/GAMMA` is the relativistic Lorentz factor γ corresponding to the beam momentum, and `SPSBEAM/FREV` is the revolution frequency of the beam in the SPS.

Within each make rule, LSA distinguishes three quantities, the target derived from the physics or optics model, the correction added on top of it, and the resulting value, related by Equation 4.2. The value is propagated to the next make rule in the hierarchy, or driven to hardware where the parameter is a leaf node.

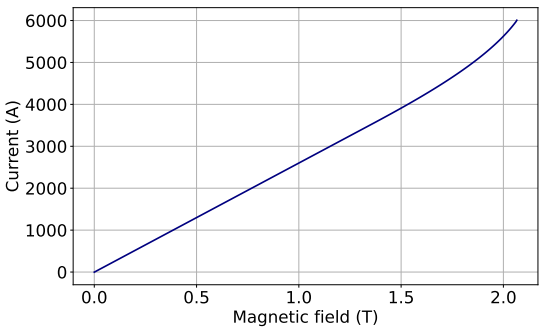
$$\text{Value} = \text{Target} + \text{Correction} \tag{4.2}$$

A parameter can be a scalar, an array, or a discrete function, a two-column array of cycle time in milliseconds from cycle start and the corresponding value, in which form the main dipole reference current and field are stored. Scalar settings are defined either per beam process or globally for the cycle.

A cycle context is divided in time into beam processes, slices such as start of cycle, injection, acceleration, flat top, and beam out, each carrying the same kind of settings over its own window and clonable independently, and the full discrete function of a parameter is the concatenation of its beam processes. The value at a boundary between two beam processes must match on both sides, and where it does not a link rule interpolates linearly between the nearest defined points, which can introduce an unintended transition. Link rules join beam processes in the same way make rules join the hierarchy vertically, and at the beam-out process most equipment is returned



(a) Transfer function absolute quadrupole gradient k_2 to reference current IREF.



(b) Transfer function bending field B to reference current IREF.

Fig. 4.12: Calibration functions used in LSA to convert from a physical quantity to the corresponding reference current IREF for the main quadrupole and dipole power converters. The calibration functions are determined experimentally and stored in the LSA database.

to its armed state by an automated link rule. For the main circuit the momentum and field are defined only from cycle start until extraction, and a link rule programs the ramp-down of the reference current to the idle value and joins it to the start of the next cycle, so that the physical machine state stays continuous across the cycle boundary even though the contexts are discrete.

A modification of a setting, or a SET, is called a trim, applied vertically through the hierarchy to all dependent parameters and, at a leaf, to hardware. Trims are transactional, so a change rejected at any level, for instance by a make-rule constraint or a machine-protection limit on rate of change or absolute value, reverts all affected settings to their last active state. Each trim is stored with a comment, a timestamp, and the full set of parameters it affected, so it can be traced and reverted later, and a full context or a subset of it can be tagged and reverted as a whole.

Trims are normally persistent and remain in the settings history indefinitely, while a transient trim is removed after a grace period of seven days by default. Transient trims serve settings that are corrected repeatedly or automatically, since the LSA database is intended for persistent settings with only occasional deliberate modification and not for cycle-by-cycle corrections.

The settings stack from LSA to hardware is decoupled by the middleware, so settings can be applied directly to hardware without passing through LSA, and an inconsistency between the settings recorded in LSA and the values actually applied is therefore possible despite the transactional guarantees. Once a cycle exists as a settings object, it is associated with runtime execution through the timing and user system described next.

4.4.2 Timing

The timing system assigns the runtime execution context to devices and synchronizes their execution across the complex [16, 87, 88]. A cycle is a settings object in LSA, while the User is the runtime identity under which devices execute their settings and publish their data, and at runtime an LSA cycle is manually mapped to a User that the timing system then refers to in place of the cycle. Each execution is assigned a cycle stamp, its execution time in nanoseconds since epoch, which uniquely identifies that cycle instance.

The number of Users is limited to 24, against the effectively unlimited number of cycles definable in LSA, and SPS User names such as SFTPRO1 through SFTPRO3, LHCPILOT,

LHC1 through LHC4, LHCINDIV, LHCION1 through LHCION4, SFTION1, SFTION2, and MD1 through MD5 change little from year to year. For standard physics the mapping between cycle and User is stable over the run, while during machine development it is remapped frequently, so a name such as MD2 need not denote the same cycle on different days, though the cycle stamp always identifies the executed instance.

Timing-dependent actions such as cycle start, injection, ramp start, flat top, and extraction are triggered at each device by direct analog or digital pulses from local timing receivers, since their precision exceeds what event delivery over the general network guarantees. The timing system also publishes events on the Device/Property schema announcing that a state has occurred or is about to occur, and forewarning events broadcast at 2500 ms and 1000 ms before cycle start which cycle and User are scheduled, with beam parameters in the payload. Because the supercycle is staggered, with the PS receiving beam ahead of SPS injection and the SPS accelerating ahead of extraction, the SPS can know at most 3125 ms before cycle start what will play, and a further event at each supercycle start lists its sequence of cycles and Users. These events let systems not connected to hardware act by subscribing over the middleware.

The SPS historically relied on the General Machine Timing (GMT) system over a 500 kbit/s serial link with hardware-limited advance notice, and from around 2025 the complex began migrating to White Rabbit Timing (WRT), an Ethernet-based system with sub-nanosecond synchronization [102, 120, 122]. For this thesis the migration mainly matters for more reliable announcements of upcoming cycles and Users further ahead of cycle start, which the earlier timing packet sizes limited.

Timing also chains cycles across the injector chain, for instance the two consecutive SPS injections from the PS and the two PS injections from the PSB, and SPS supercycles remain largely static as discussed in Section 4.3. Constraint-based optimization may extend DBS to the SPS in the future, driven by LHC needs [70], which lies outside the scope of this thesis but bears on how its outcomes may later be applied. The commands and data that timing schedules are transported by the middleware described next.

4.4.3 Middleware

The middleware transports settings and data between client applications and the front-end computers that interface with hardware, with settings written as SET operations and measured data read by get or subscribe operations over the Controls Middleware CMW and its Remote Device Access protocol RDA [11, 81]. A multiplexed device holds a distinct value for each User, so an operation is filtered by a selector, normally the timing User and not the LSA cycle, since the middleware has no notion of LSA cycles.

Publication timing depends on the measurement, as some properties are published once at the end of the cycle when complete, such as the SPS main-dipole field from the B-Train after flux integration over the full magnetic cycle, while others are published as soon as acquired, such as BPM orbit values. Figure 4.13 shows the flow of settings and data between clients, the CMW and RDA layer, and the front-end computers, and this stream is archived by NXCALS and processed online by UCAP as described next.

4.4.4 Logging

Operational data is archived by NXCALS, the CERN data logging service that replaced the earlier CALS system to handle the scale of data generated across the complex [141]. Any property published over the middleware can be archived through a logging subscription, each publication recorded with its cycle, User, and cycle stamp, and retrieved by time range and device long after

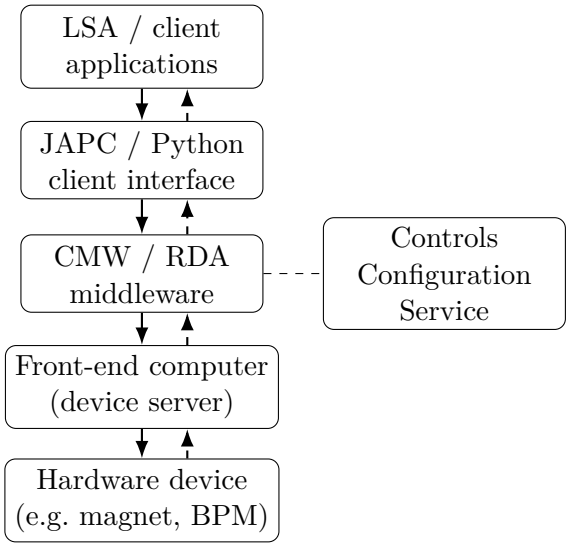


Fig. 4.13: Simplified controls middleware architecture. Solid arrows indicate the set path from client applications down to hardware, while dashed arrows indicate the get and subscribe path returning measured or published data. The Controls Configuration Service provides device registration and addressing metadata used by the middleware.

the cycle was played without a dedicated measurement campaign. NXCALS is the main source of operational data in this thesis, including B-Train field measurements, power converter current signals, and the record of which cycles and Users played and when, all retrievable for a given cycle stamp.

Logging provides the offline data, and real-time compensation instead needs the online processing of the following section.

4.4.5 UCAP

The Unified Controls Acquisition and Processing framework UCAP adds a server-based processing layer on the middleware of Section 4.4.3 [39]. A user registers a Java or Python converter that subscribes to one or more device properties, applies its own processing logic, and republishes the result as a new device property, as illustrated in Figure 4.14.

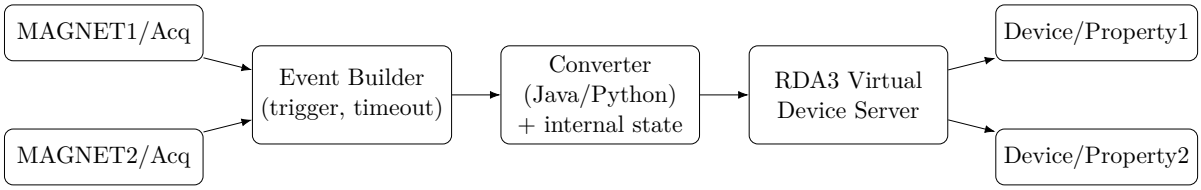


Fig. 4.14: Example of a UCAP converter that subscribes to two input properties, applies a user-defined transformation, and republishes the result as a new property.

UCAP is organized as three stateless stages, Acquisition, Transformation, and Publishing. The Acquisition stage subscribes to the input devices and, through an Event Builder, groups inputs from several sources by a common trigger, cycle, or timestamp, with a configurable timeout so an incomplete set is still forwarded and one source can be buffered until a triggering event arrives. The Transformation stage runs the converter on each grouped event, holding state across invocations and, on supported nodes, running CUDA-accelerated code such as neural-network

inference, while the Publishing stage returns the output to the middleware through a Virtual Device Server.

A converter can also act on machine settings by triggering actions, including trims on the LSA hierarchy, and its outputs can be logged to NXCALS. UCAP is deployed across CERN accelerators, for instance for the online correction of Multi-Turn Extraction splitting in the PS, where the converter output is applied as an LSA trim [51].

4.4.6 FGCs

The power converters that excite the SPS magnet circuits are controlled by Function Generator/Controllers FGCs, dedicated embedded controllers on racks near the converters that translate a requested setpoint into the point-wise reference function the hardware plays, exposed to the control system over RDA [80, 104, 147]. The generation deployed on the SPS main circuits is FGC3, and the relevant setpoints here are current functions, since the SPS main circuits are regulated in current and not in field as discussed below, while other CERN synchrotrons can regulate a field-related function.

4.4.6.1 Function interpolation and regulation period

The FGC runs a regulation loop at a fixed period, 2 ms for the main dipole circuit and 4 ms for the main quadrupole circuit, and interpolates the sparser knot points of the LSA current function, for instance by polynomial interpolation, into a dense reference at that resolution. A required current change is shaped by the Parabolic - Linear - Exponential - Parabolic (PLEP) method, or its simplified Parabolic - Linear - Parabolic (PLP) variant, into a parabolic segment that raises the rate of change, a linear segment at the maximum allowed rate, and a closing parabolic segment back to the target, which respects the converter limits on rate of change and voltage and produces a smooth played current with possible ringing after large steps. The current function stored in LSA is therefore not exactly what the magnets play, and apparent differences between programmed and measured current follow from this shaping and not from a control error.

4.4.6.2 Operation modes of the SPS main circuits

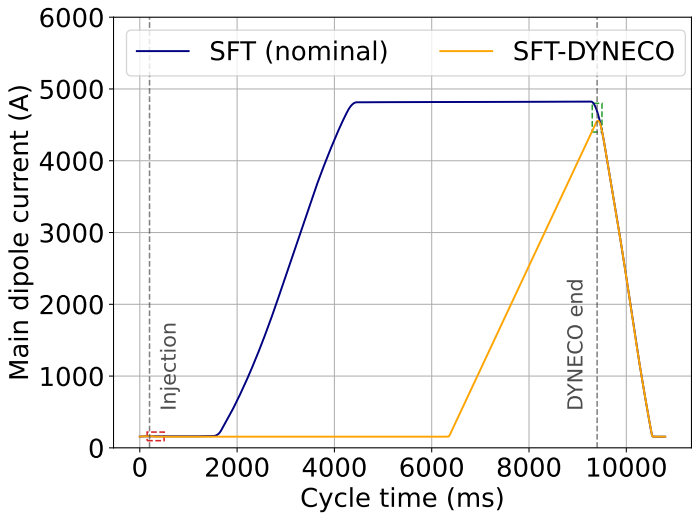
In a cycling machine such as the SPS the FGC arms the full reference function of the coming cycle before it starts, applied as a persistent setting that holds until changed and not as a cycle-to-cycle modulation of successive occurrences of the same cycle. Because the setting is persistent, a change made in LSA carries no guarantee on which cycle occurrence it is first played. An alternative path is the real-time channel, a UDP gateway accepting a current correction ΔI at 50 Hz that the regulation loop applies immediately, originally developed for the coasting-beam LHC.

Because the SPS operates in cycling mode, the current at the end of one cycle matches the start of the next, called the idle current, 155.73 A for the main dipole circuit and 55.5 A for the main quadrupole circuit. The LSA settings enforce this match, since a large current step at the boundary would trip the converter.

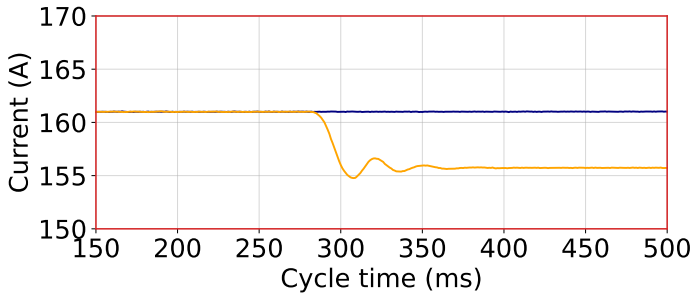
4.4.6.3 Economy modes in the SPS power converters

A further feature of the FGC, and specific to the SPS only, is dynamic economy mode. It is a power regulation behaviour that is native to the power converter and evaluated in real-time, and is not controlled by higher level settings. It is a mode whose purpose is to reduce power

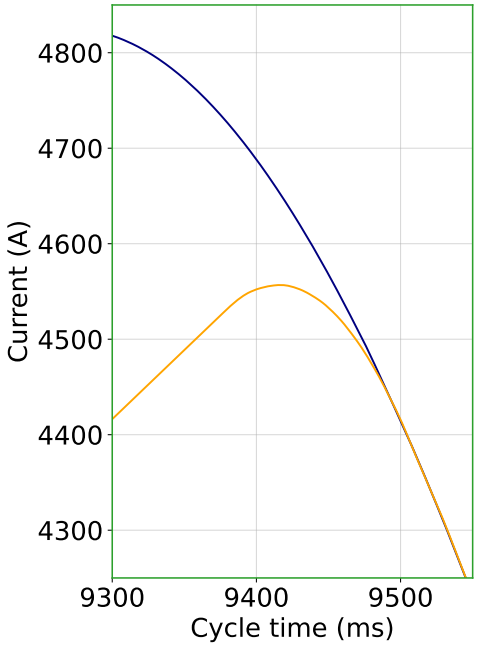
consumption by lowering the excitation current of the main magnetic circuits when no beam is present to be accelerated. Concretely, it can be triggered if if circulating beam is not detected after injection, for instance as a consequence of an injection failure from the PS or a beam dump. If dynamic economy is enabled for the cycle in question, the trigger releases a corresponding asynchronous timing event tied to that cycle time.



(a) Nominal SFT cycle and a cycle put into dynamic economy.



(b) Inset showing start start of cycle and how the power converter ramps down after injection.



(c) Inset showing the dynamic economy end phase, where the economy mode joins the programmed current function at the end time.

Fig. 4.15: Dynamic economy mode of the SPS main dipole circuit on the SFT cycle. The main plot shows the nominal and dynamic economy current functions, as measured. Inset b) shows the start of the cycle, where the power converter ramps down after injection. Inset c) shows the end phase, where the economy mode rejoins the programmed current function at the end time.

The mode has two main phases. First, the power converter is ramped down asynchronously from the current level at the time of triggering to the idle current. Second, at a corresponding setting referred to as the end time, the power converter is ramped back up from idle current to rejoin the programmed current function, which is handled in real time by the FGC rather than through LSA. In most cases this means that the power converter ramps from idle current up to meet the programmed current, which is normally higher than idle current except during the final milliseconds of the cycle, when the programmed current itself returns to idle current. Both the ramp down and the ramp back up follow the same PLP shaping described earlier, using a parabolic segment to move from a fixed current with zero rate of change to the maximum allowed

rate of change in a linear segment, followed by a second parabolic segment that matches the rate of change of the programmed function at the point of reconnection.

As an illustrative example, for the standard SFTPRO beam, the dynamic economy end time is usually around 9.4 s. This means that at cycle time 9400 ms the dynamically regulated current joins the original programme just after ramp down starts. Consequently the main field reaches close to the nominal magnetic cycle without remaining at the 4800 A flat top current for several seconds. An example of the dynamic economy mode is shown in Figure 4.15, for the SFT cycle. The main quadrupole circuit is also subject to dynamic economy, and the same principle applies, though the current levels are different due to the different design of the quadrupole circuit.

A distinct economy mode, referred to as full economy, is triggered directly by the operators rather than by a cycle based timing event. Full economy is equivalent to a standby mode of the power converter, in which the current is held at the idle current for the entire duration of the cycle.

Both dynamic and full economy modes are relevant to this thesis because the excitation of the magnets depends not only on the programmed current function from the high-level control system, but also on runtime changes to that excitation triggered by the FGC. Since the dynamic economy timing event occurs asynchronously after the cycle has already started, the eventual magnetic trajectory played during that cycle is not fully known until the event either occurs or fails to occur. This has direct consequences for reconstructing the magnetic state and the operational history of the machine from logged data, since a given cycle cannot be assumed to have followed its nominal programmed trajectory without additional information on whether dynamic or full economy was active. Dynamic economy can cause the magnet to be excited to a different field level than the nominal field for part of the cycle, while full economy, by holding the current at idle with zero rate of change, avoids the induction of eddy currents that would otherwise arise from a changing current.

4.5 Magnetic Field Control in the CERN Injectors

4.5.1 Hierarchical Field Control and Field Regulation in the PSB and PS

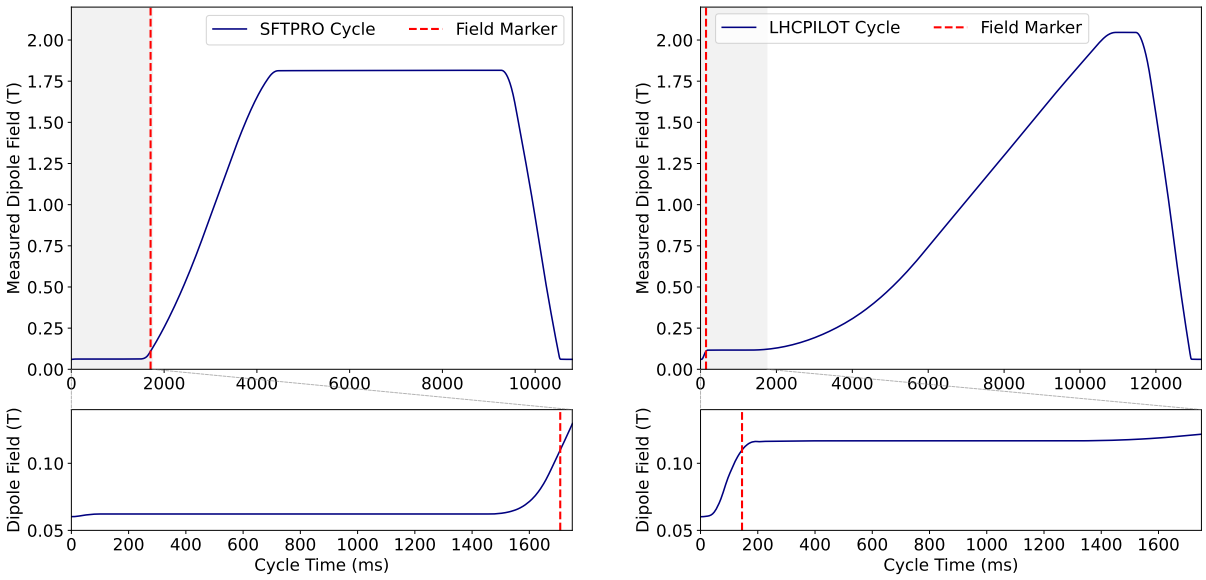
As described in Section 4.4.1, the LSA hierarchy propagates the momentum programme to a required field and then to a current through the calibration functions of Figure 4.12. This current is applied open-loop to the power converter for the quadrupole, sextupole, and higher-order circuits of every injector and for the main dipole of the SPS. For the main dipole of the PS and the PSB the hierarchy instead terminates at the programmed field, and the measured B-Train field of Section 3.4 is compared against it in a closed loop on the FGC over White Rabbit [2]. The current is then the regulator output that drives the measured field onto the programmed field. No equivalent loop exists on any SPS circuit, and every SPS field including the main dipole is realized open-loop.

4.5.2 SPS B-Train in Operation

As established in Section 4.5.1, the SPS B-Train is not part of any control loop and serves only as a measurement and logging system for the main dipole field. It exists as two physically distinct chains detailed in Section 3.4.3, an operational chain whose integrator is reset to a fixed value at the start of every cycle, giving a field trace relative to that arbitrary value, and a spare chain carrying a field marker that provides an absolute measurement.

The spare-chain marker triggers at a fixed field of 1100 G, where the rate of change of field is high enough for a clear NMR peak, and it anchors the integration for that cycle at a point

near injection that depends on the cycle, as shown in Figure 4.16. The marker is detected at the native 250 kHz rate, and its logged application time in cycle time is an estimate.



(a) Magnetic cycle SFT cycle 14 GeV/c injection. **(b)** Magnetic cycle LHC cycle 26 GeV/c injection.

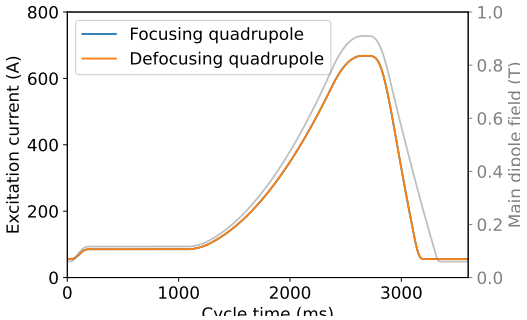
Fig. 4.16: Magnetic cycle of two common SPS cycles, with a vertical dashed line indicating where in the cycle the field marker is applied. For the LHC cycle, the marker is applied prior to injection, while for the SFTPRO cycle it is applied after, due to the lower injection momentum of the 14 GeV/c beam.

Both chains publish to CMW and are logged to NXCALS per cycle instance and USER, with the field estimate decimated to 1 kHz from the native 250 kHz rate and the marker value and time logged alongside. The integration chain has a voltage offset that drifts over time and is recalibrated after a timeout of 3 h and always during a ZERO cycle, the only cycle that holds the magnets at idle current with no excitation. Recalibration thus requires a ZERO cycle in the supercycle, and not every supercycle contains one. In that case the timing cables behind the Front-End Computer (FEC) are shorted to make an ordinary cycle appear as a ZERO cycle and force the recalibration.

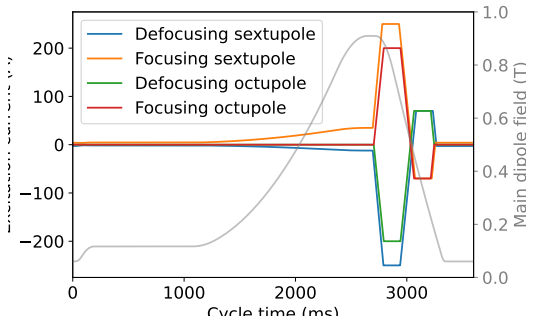
4.6 Magnetic Precycle and Degaussing in Operation

4.6.1 Dedicated Magnetic Precycle in the SPS (MD1)

The SPS maintains a dedicated magnetic precycle, colloquially referred to as MD1, which occupies the `MD1` timing user throughout operation. The cycle is 3 BP long, or 3.6 s, and reaches a peak field of 0.91 T, equivalent to 200 GeV, but does not take beam. Its flat bottom sits at an equivalent momentum of 26 GeV/c, reusing the beam process of the LHC-type cycle of Section 4.2 for the initial ramp up. This value however carries no particular physical significance, since MD1 never injects and the flat bottom itself is very short. The design is an exploitation of an existing cycle template in the construction of MD1, rather than a deliberate matching to any injection condition. The peak field of 200 GeV is shared with the short MD cycles of Section 4.2, both of which reproduce the same ramp up and ramp down beam process as MD1 at the end of the cycle.



(a) Main dipole field $B(t)$ and the two main quadrupole circuits QF and QD reference current $I(t)$



(b) Focusing and defocusing sextupole and octupole reference current $I(t)$

Fig. 4.17: The MD1 magnetic precycle. Left: main dipole field and QF/QD quadrupole reference current. Right: focusing and defocusing sextupole and octupole circuits, shown as reference current only, as the intermediate gradient is not readily accessible through the settings management system like the quarupoles. MD1 does not take beam, and reaches a peak field of 0.91 T, equivalent to 200 GeV/c, over 3 BP (3.6 s).

This is asserted on the basis of operational knowledge of the cycle definitions, as no documented design record could be identified for this thesis. **[TODO: keep this?]**

The main dipole and quadrupole circuits are shown in Figure 4.17 as the derived main field B and quadrupole gradient K, consistent with the units used for the other cycles of Section 4.2. The sextupole and octupole circuits are shown as reference current only, as the intermediate field or gradient of these circuits is not readily accessible through the settings management system The reference current therefore provides the most direct representation of the played cycle for these two circuit families.

4.6.2 Quasi-Degaussing of Higher Order Magnets in the SPS

The sextupole and octupole circuits do not simply ramp down together with the main dipole and quadrupole circuits at the end of an SPS cycle. Instead, they are driven to a maximum (or minimum) reference value, then the maximum value of the opposite polarity, before returning to zero at the end of the cycle. This feature is visible in the sextupole and octupole traces already shown for SFTPRO in Figure 4.1 and for MD1 in Figure 4.17, and is common to all operational cycles. The minimum reference value reached is generally smaller in magnitude than the maximum, so the excursion is not a symmetric bipolar excitation. This end-of-cycle excursion is referred to operationally as quasi-degaussing, in contrast to a classical degaussing sequence, which uses a decaying alternating excitation thaat asymptotically drives the remanent field toward zero.

The fast ramp rate required to execute this excursion within the available cycle time, rather than any requirement arising while beam is present, is understood to be the driving design requirement for the sextupole and octupole power converters.

4.6.3 Operational Limitations from Hysteresis

Hysteresis in the SPS main and higher order magnet circuits means that the cycle preceding a given cycle can influence its behaviour, in addition to its own persistent settings. In regular operation, the MD1 always precedes fixed target beams like SFTPRO and SFTION in the

standard supercycles, as shown in relations 4.1 and 4.3.2 and illustrated for the proton case in Figure 4.2. In general, MD cycles like the one in Section 4.2 carrying an MD1-like ramp are not required to be followed by MD1, for instance as shown in the supercycle in Figure 4.9. However, an explicit MD1 slot is sometimes retained afterward at the discretion of the operator.

Additional MD1 slots are sometimes appended to a supercycle beyond this standard placement. One case is shown in Equation 4.3.2, where MD1 slots are appended to extend the supercycle to match timing constraints shared with the rest of the injector complex, since the PS and PSB supercycle lengths are tied to that of the SPS. A second case arises at the discretion of the operator, who may append one or more MD1 slots ahead of a subsequent cycle, often SFTPRO, if the beam on that cycle seems unstable. In both cases, the appended MD1 slots carry no beam, so the machine time they occupy is not available for physics or MD operation.

MD cycles are, by contrast, not restricted to a fixed position in the supercycle. An MD cycle can in principle follow any preceding cycle, subject to the discretion of the operator and of the physicists requesting the study. The same sequence-dependence nonetheless affects MD cycles, where a cycle configured against one supercycle sequence may behave differently if inserted into a different sequence on another day.

Irrespective of the precycle, When the supercycle sequence changes, for example from the standard fixed-target configuration of Equation 4.1 to the LHC filling configuration of Equation 4.3.1, the magnetic history preceding SFTPRO changes with it. Operators compensate for this by manually trimming the horizontal tune Q_H by a few $\pm 1 \times 10^{-3}$, in order to restore the extracted spill macrostructure to its state prior to the sequence change. This trim is applied manually, after an observed change in spill behaviour, each time the supercycle sequence changes.

These effects were characterised for the SPS by Velotti et al. [135], where changes in supercycle composition were shown to alter the remanent field of the main dipoles and to correlate with measurable changes in the fixed-target closed orbit and in the slow-extraction spill quality. The reproducibility loss was attributed to the remanent field of the main dipoles, with the main quadrupoles identified as a further contributor to the spill, and a slow feed-forward correction of the beam energy derived from the B-Train measurement was proposed. The quantification of these effects and the compensation scheme developed to address them are presented in the following chapters.

4.7 Machine Development Time

Machine development MD is machine time on a CERN synchrotron for studies outside routine physics delivery, available only in a limited number of slots. It comes in two kinds in the SPS, dedicated MD, where one study has the full ring for about 10 h and can play any cycle in any sequence, and parallel MD, where a single short cycle of up to 6 BP is inserted in place of MD1 during physics operation and its beam is usually dumped at low energy.

A parallel MD study controls only its one cycle, with the surrounding supercycle fixed, and control of the preceding sequence is available only in dedicated MD. In either kind only cycles already in the SPS pool are generally available, and a new cycle needs approval from management and equipment experts before commissioning.

Chapter 5

Problem Statement and Proposed Approach

The operational relevance of this problem was established for the SPS by Velotti et al. [135], where the dependence of the fixed-target orbit and slow-extraction spill on the supercycle composition was measured and a slow feed-forward correction driven by the measured B-Train field was proposed. That correction depends on the field being measured for each cycle as it is played, and a model able to predict the realized field ahead of time, for an arbitrary cycle within an arbitrary sequence, was identified as necessary and left for future work. The present thesis develops that predictive model. No equivalent field prediction and compensation scheme is established for the SPS main circuits, and no comparable data-driven solution for hysteresis and eddy current effects in accelerator magnets is reported more broadly, as reviewed in Chapter 2. **[TODO: Fold in other CERN prestudies here if applicable.]**

The hysteresis and eddy current effects described in Section 4.6 are a single underlying phenomenon exposed by different operational triggers, including supercycle sequence changes, power converter economy modes, and the presence or absence of a ramp down between cycles. The dedicated magnetic precycle, MD1, and the analogous end-of-cycle quasi-degaussing excursion used by the higher-order circuits (Section 4.6.2), are the current operational response to this mechanism.

5.1 Hysteresis as an Operational Problem

Hysteresis is an operational problem in any synchrotron operation because the field realized by a cycle depends not only on that cycle's own programmed excitation, but also on the magnetic history left behind by the cycles that preceded it. This dependence arises through the static and dynamic mechanisms described in Section 2.4.2 and Section 2.6.2. Without correction, this dependence means that a cycle played after a different preceding cycle realizes a different field, and therefore different beam parameters, purely as a consequence of sequence, with no change to its own programmed settings.

Three distinct aspects of SPS operation perturb this history. A change in supercycle sequence changes which cycle type precedes a given cycle, and therefore the magnetic history it inherits. Switching between the standard fixed target configuration and an LHC filling configuration is one example.

The same is true, more subtly, of changes to the supercycle's temporal construction, even when the sequence of cycle types itself is unchanged. Examples include inserting or removing a timing gap such as a ZERO slot (Section 4.3), extending a cycle's flat top, or appending additional slots to match timing constraints elsewhere in the injector chain (Section 4.6.3). Each of these changes the elapsed time between the end of a preceding ramp-down and the start of the next cycle, and therefore how much eddy current decay has already occurred by the time that cycle begins, independent of which cycle types are involved. Additionally, power converter economy modes, namely dynamic and full economy, alter what is actually played during a cycle independently of what is programmed at the high-level control system (Section 4.4.6.3). As a result, the history

left behind by a given cycle occurrence is not always predictable from its settings alone. The presence or absence of a ramp-down at the end of a cycle further determines whether eddy current decay occurs at all before the next cycle begins, so that even a cycle with no programmed change like full economy can leave a different history behind depending on how it ends.

Manual tune correction, applied by operators when a sequence changes (Section 4.6.3), is also part of this operational response, intended to restore beam behaviour after the fact at a cost documented below.

Consequences for Beam Parameters

An uncorrected field error is not confined to the circuit that carries it, because the circuits are operationally coupled. A main dipole field error acts as a deflecting kick that displaces the closed orbit following Equation 2.12 of Section 2.1.5, and the same field sets the beam momentum through the rigidity relation of Equation 2.1. Every higher-order circuit whose strength is normalized to that momentum through Equation 2.8 then inherits an apparent error, which propagates to the tune, to the chromaticity, and to RF synchronism as expressed in Equation 2.15, so a dipole hysteresis error becomes a precondition for correctly interpreting every circuit downstream of it. In the higher-order circuits the history-dependent errors act directly, shifting the betatron tune toward the resonance lines of Section 2.1.6, shifting the chromaticity Q' that suppresses the head-tail instability described in Section 2.1.7.1, and degrading the amplitude-dependent detuning used for Landau damping.

These errors differ from static manufacturing and alignment errors in their dependence on magnetic history. A static error is mapped once and absorbed by a fixed correction, while a hysteresis-driven error changes with the preceding sequence and is not addressed by any static scheme. The mechanism acts on every cycle that excites the main and higher-order circuits, and what differs between cycles is its magnitude relative to what each cycle can absorb without correction. SFTPRO and SFTION are singled out in what follows because the chromatic slow extraction of Section 2.1.9.1 places stringent reproducibility requirements on the dipole, quadrupole, and sextupole fields at the flat top, and because COSE co-scales these circuits through independent converters, so a hysteretic offset in any one breaks the fixed ratio and perturbs the rate of extraction. Other cycles such as LHC injection experience the same mechanism without being corrected for it, so the problem is not unique to fixed-target beams even though the case for correcting it first is made on their behalf.

[TODO: actually, the LHC beam avoids this by always prefixing with a ZERO, and a SFTPRO, either full or in dynamic economy.]

The Cost of the Current Operational Responses

The first cost is the MD1 precycle itself. As described in Section 4.6.1, MD1 exists to bring the main dipole and quadrupole circuits to a known state ahead of the proton and ion fixed target beams. It takes no beam, so its cost is pure overhead as machine time that could otherwise be used for physics delivery or machine development, and energy spent exciting the main circuits field described there, for no direct physics purpose.

Figure 5.1 gives an illustrative view of SPS run time by beam destination, during the 2023–2025 run period, counting only the time from start of beam commissioning, until the end of the last physics delivery. On average, MD1 is the second largest single consumer of run time after fixed-target delivery itself. Table 5.1 tracks this cost across the full 2023 to 2025 run period, together with the corresponding energy consumption and estimated cost.

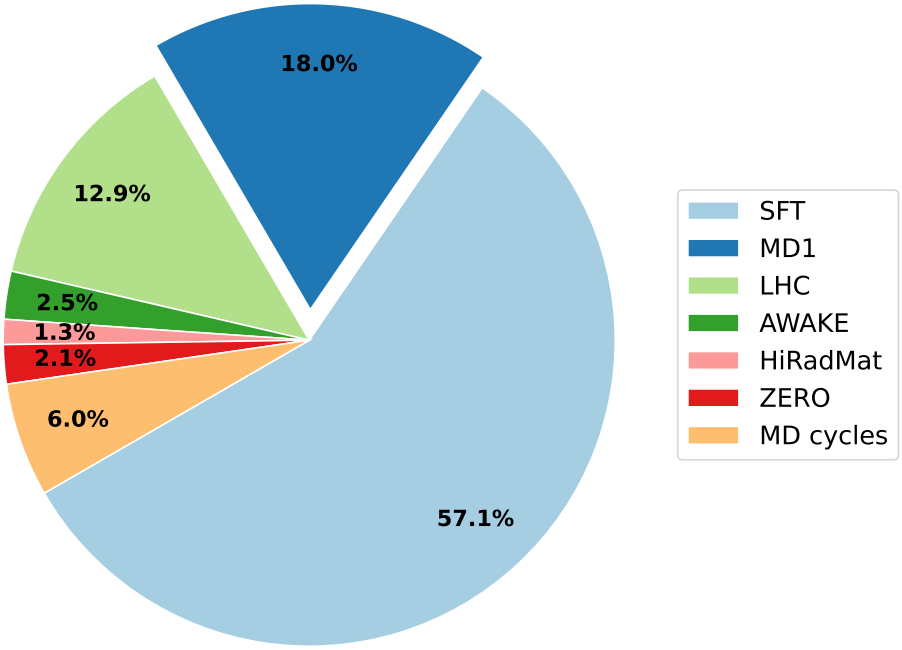


Fig. 5.1: Distribution of the SPS run time across beam destinations 2023–2025. Proton and ion run time like SFTPRO and SFTION, as well as LHC proton and LHCION cycles are grouped together. MD1 is the second largest consumer of run time after fixed target delivery.

Tab. 5.1: Estimated MD1 energy consumption, cost, and share of SPS run time, by year and beam type. The share of run time column counts every scheduled MD1 occurrence, regardless of whether the magnets pulsed during that occurrence. The energy and cost columns count only occurrences in which the magnets actually pulsed. The columns therefore reflect different underlying counts and should not be read as directly proportional to one another.

Year and beam type	Energy [GWh]	Cost [kEUR]	Share of run time [%]
2023 proton	3.96	277	18.3
2023 ion			12.2
2024 proton	4.89	342	18.9
2024 ion			13.0
2025 proton	5.03	352	18.8
2025 ion			11.0

The second cost is the sequence restriction imposed on supercycle construction, together with the manual correction it necessitates. Because the field seen by SFTPRO and SFTION depends on whichever cycle preceded them, changes to the supercycle sequence require operators to manually retune the horizontal tune Q_H, as described in Section 4.6.3.

This procedure carries costs beyond the time spent performing it. It has no defined tolerance against which a correction can be judged sufficient, and it is not recorded in a way that would allow its frequency, or its consistency from one operator to another, to be checked after the fact. It is also inherently reactive, since it depends on an operator noticing that the spill has degraded before any correction is applied, rather than consistently anticipating the change before it occurs. In cases where the supercycle sequence is changed frequently, the operators may even intentionally avoid correcting the tune on every change, because they know that the next change will require a new correction anyway.

This reactive, manual character becomes a harder constraint as scheduling becomes more dynamic. The current approach is tolerable when supercycle sequences are relatively fixed and well rehearsed by operators over a run. It becomes a more significant bottleneck under more flexible or frequently changing scheduling, where there is correspondingly less opportunity for an operator to notice and correct a degraded spill between changes, or when the required change is not consistent from one sequence change to the next.

A smaller, analogous cost is paid by the higher-order circuits, whose end-of-cycle excursion (Section 4.6.2) similarly occupies cycle time and consumes energy without contributing to physics delivery, though at a scale well below that of MD1. As seen in the sextupole and octupole excitation profiles in Section 4.2, the higher-order circuits are never excited to their designed maximum field for physics delivery. The quasi-degaussing excursion, and the fast ramp-rate needed to reach it during beam-out, sets the design requirement for the higher-order circuits instead of physics delivery.

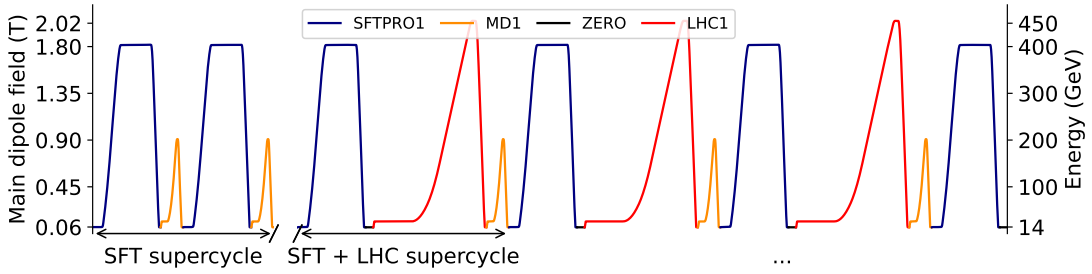
Eliminating these costs is the motivation for this thesis, and drives four related goals. The first is reducing the energy and beam time cost associated with the MD1 precycle and, to a lesser extent, the higher order excursion. The second is increasing SPS duty cycle by recovering the time currently spent on precycling rather than physics or machine development. The third is increasing automation by removing the dependence on manual, intuition based, and undocumented operator correction on sequence change. The fourth is enabling more flexible cycle scheduling, both for the supercycle sequences already in use and for future, more dynamically scheduled beam delivery in the SPS. None of these four goals can be pursued by simply removing MD1 and the manual tune corrections outright, since both currently protect the beam-parameter reproducibility motivated in the preceding section. Achieving them instead requires replacing that protection with an active compensation scheme.

5.2 Motivating Evidence

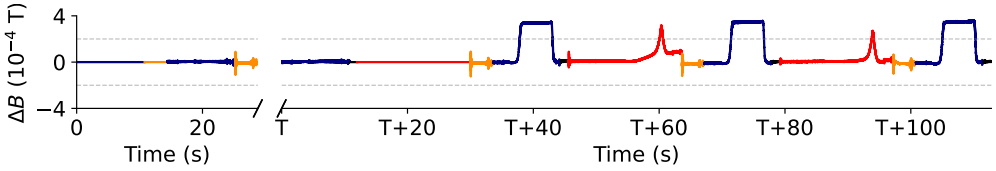
5.2.1 Slow Extraction Perturbation from Supercycle Changes

Figure 5.2 and 5.3 show the evolution of the measured dipole field from the B-train (Section 3.4.3) during supercycle transition, and the difference in measured field compared to an arbitrary reference. The sequences shown are representative of the operational problem described in Section 5.1, switching from a fixed target-only sequence to an LHC filling sequence, and vice versa. The choice of showing field difference with respect to an arbitrary reference rather than absolute field is derived from the reproducibility requirement described in Section 5.1. When introducing the LHC filling into the supercycle as seen in Figure 5.2, the measured field on the flattop of the fixed-target cycle visibly increases by 3.4×10^{-4} T compared to the reference,

immediately after the first LHC cycle is played. When removing the LHC filling from the supercycle, as seen in Figure 5.3, the measured field on the flattop of the fixed-target cycle visibly decreases by 2.5×10^{-4} T compared to the reference, over the course of several cycles. Notably in the change back to SFT only operation, in this case, the measured field does not immediately return to the exact same value as the reference prior to the LHC filling. As can be seen in Figure 5.4, the average measured field at the flat top continues to decay over 15 min after the transition, and stabilizes at a value 3.4×10^{-4} T below the reference, recovering the same field magnitude as before the LHC filling was introduced. Each of these field changes during the slow stabilization is in the order of a 1×10^{-5} T, and as can be observed in the figure, the field does not visually monotonically decay, but oscillates around the measurement limit of the B-train.



(a) Evolution of main dipole field during a supercycle change from SFT only, to an LHC filling sequence.

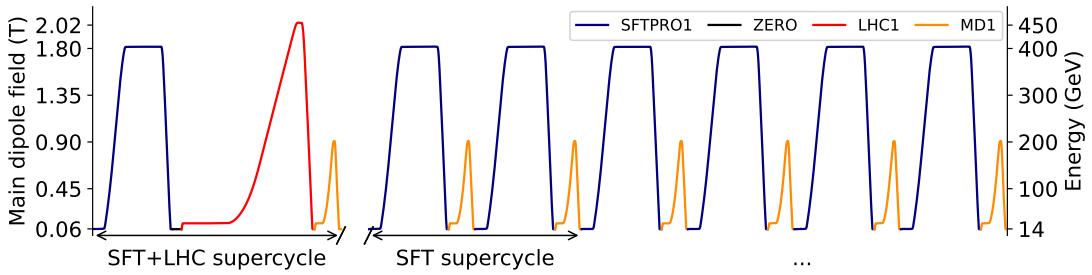


(b) Relative change in measured dipole field during a supercycle change from SFT only, to an LHC filling sequence.

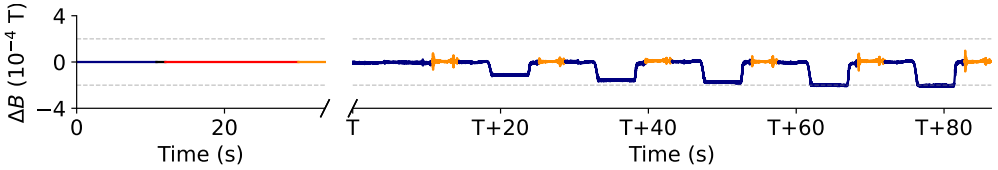
Fig. 5.2: Illustration of a supercycle change from SFT only (Seq. 4.1), to an LHC filling sequence (Seq. 4.3.1), and the corresponding relative change in measured dipole field, where the reference is taken as the first measured field of each cycle.

[TODO: Show an SFTx2+LHCPILOT sequence, where both cycles need to be corrected.]

This field change is not without consequence for the extracted beam. Figure 5.5 compares the BCT trace of an SFTPRO spill recorded before the sequence change to one recorded immediately after, from the same sequence changes in Figure 5.2 and Figure 5.3. During these changes, the spill macrostructure visibly departs from the constant extraction rate operators expect. Following an operator horizontal tune correction, the spill macrostructure returns to a shape visually indistinguishable from the baseline, so only these two representative states are shown rather than every intermediate correction step. This assessment remains primarily qualitative, since the operational figure of merit tracked by operators during sequence changes is the visual linearity of the BCT trace. A percentage-based numerical spill quality factor exists in operation, but does not accurately track the macrostructure of the spill, as is discussed in Chapter 8. Based on estimation by beam dynamics staff, a relative energy error as small as $\Delta p/p \approx 1 \times 10^{-4}$ may significantly perturb the beam extraction, which at the SFT flat top is equivalent to 1.8×10^{-4} T.



(a) Evolution of main dipole field during a supercycle change from an LHC filling sequence, to SFT only.



(b) Relative change in measured dipole field during a supercycle change from an LHC filling sequence, to SFT only.

Fig. 5.3: Illustration of a supercycle change from an LHC filling sequence, to SFT only, and the corresponding relative change in measured dipole field, where the reference is taken as the first measured field of each cycle. The recording follows Figure 5.2a 1.5 h later, after the filling is completed.

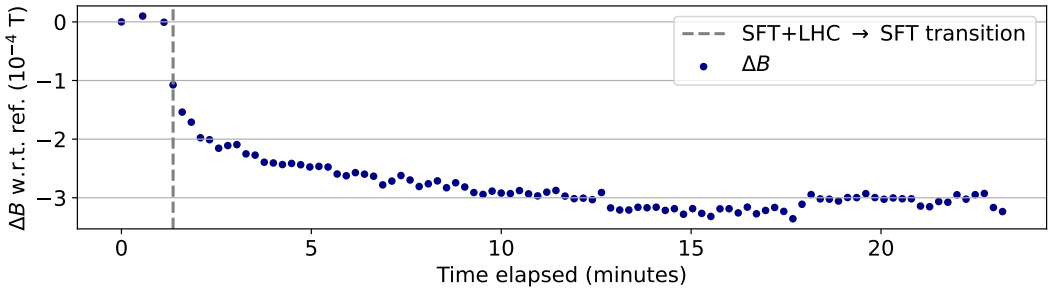
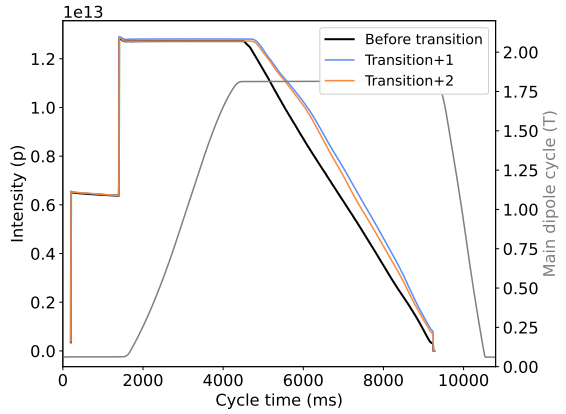


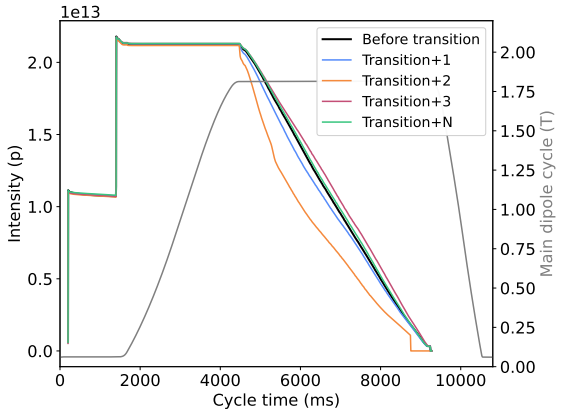
Fig. 5.4: Evolution of the SFT average flat top field after transitioning from an LHC filling sequence to fixed-target only, with respect to a reference taken before transition. The field decays and stabilizes over 15 min after the transition, with a total change of -3.4×10^{-4} T from the reference.

5.2.2 Orbit Displacement as Result from Dipole Errors

The orbit displacement produced by a given field error is largest at injection and decreases with beam energy. On the injection plateau the beam is bunched and held at a fixed RF frequency, so a field error displaces the closed orbit according to Equation 2.13. Table 5.2 lists the resulting displacement at injection for the 14 GeV/c injection, Q26 optics, SFT cycle, and for the LHCPILOT cycle, which injects at 26 GeV/c with Q20 optics, as a comparison to an LHC-type cycle. The ramp and flat top are not tabulated, since the radial loop holds the beam centered during the ramp so that $\Delta r = 0$, and at the flat top the increased rigidity makes the field error manifest as a momentum error and not a radial displacement. A constant hysteresis-driven field



(a) Spill change when transitioning from SFT only to an LHC filling sequence, with a reference measurement just before sequence change in black, and two subsequent measurements in blue and orange. The blue and orange traces show the macrostructure is perturbed significantly immediately after change, with some instability in extraction.



(b) Spill change when transitioning from an LHC filling sequence to SFT only, with a reference in black and four subsequent measurements in blue, orange, red, and green. The blue and orange traces follow the slow flat-top field change of Figure 5.3, the red trace shows an operator tune correction applied before the field has stabilized, and the green trace shows the same correction once the field has stabilized and matches the reference.

Fig. 5.5: Visualized change in the North Area Fixed target beam (SFTPRO) spill macrostructure over the 4.8 s slow extraction, as measured by the BCT, during supercycle sequence changes.

Tab. 5.2: Theoretical absolute radial displacement $|\Delta r|$ in mm at injection from a uniform main dipole field error ΔB, from the fixed-frequency response of Equation 2.13, for the Q26 SFT and Q20 LHCPILOT cycles.

ΔB [T]	SFTPRO (Q26)	LHCPILOT (Q20)
1×10^{-5}	0.55	0.21
2×10^{-5}	1.11	0.43
5×10^{-5}	2.77	1.07
1×10^{-4}	5.53	2.13

offset carried over from a preceding cycle therefore perturbs the beam more strongly at injection. Operationally, an orbit displacement of 1 mm at injection is considered significant, and would be corrected. Therefore, the absolute field error for the most sensitive beam, SFTPRO, at the most sensitive point in its cycle, injection, is used to define the maximum field error tolerance for field reproducibility, at 2×10^{-5} T. In practice, the field error tolerance may be lower, as interplay with RF control loops and energy matching at injection may place the beam at a different radial position. A dipole field error therefore manifests as either a radial error or an energy error, depending on the RF control loop configuration.

The calculated orbit displacements were additionally measured during a parallel MD using a Q20, 26 GeV/c injection cycle. During the tests, all parameters except the main dipole field were held constant, which was shifted by $\pm 5 \times 10^{-5}$ T, without affecting the RF frequency. For the reference measurement, the beam was well centered, and upon applying the field perturbation

the orbit shifted ±2 mm The resulting orbit shift exceeds the theoretically calculated value of 1.1 mm for the equivalent field perturbation for a Q20 LHC-type beam, seen in Table 5.2. This further narrows the in practice maximum error tolerance for dipole field below the previously stated 2×10^{-5} T.

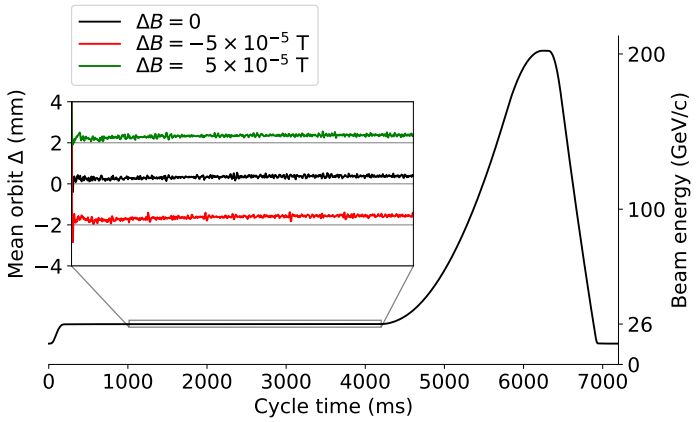


Fig. 5.6: Measurement of mean orbit displacement around the SPS ring for $\Delta B = \pm\, 5 \times 10^{-5}$ T field perturbations applied over the injection plateau of a Q20 optics MD cycle, with beam injected at 1015 ms, and dumped at the end of the injection beam process (start of ramp).

No equivalent quantitative transform exists for the quadrupole circuits, since a quadrupole gradient error produces a tune shift that depends on an integral of the error weighted by the beta function around the entire ring. Consequently, the error manifests mainly as a threshold effect tied to proximity to a resonance line rather than a quantity that scales smoothly with the field error. For this reason, the quadrupole and sextupole consequences described in Section 5.1 are argued mechanistically rather than quantified here.

Beyond the degraded spill quality shown above, hysteresis has been theorized, though not conclusively measured, to have caused beam instabilities and subsequent automated dumps on physics cycles when certain non-standard MD cycles were inserted into a supercycle during parallel MDs. This account is included here as an operational observation reported by operators, not as a measured result, and is not otherwise substantiated in this thesis.

5.2.3 Field Response and Hysteresis Quantification in the Main Dipoles

Figure 5.7 show the $I \rightarrow B$ field response of the main dipole circuit, with B-Train measurements of all operational cycles in the SPS. The field response is shown to be largely linear up to 4000 A, equivalent to 340 GeV/c, well beyond the original SPS design specification of 300 GeV/c given in Section 2.3.1, and below normal operating 400 GeV/c proton beams. Above the 4000 A threshold, a saturation-driven nonlinearity can be observed. The lookup-table current-to-field transfer function from Figure 4.12b is shown in the same figure and demonstrates that while the field response is largely anhysteretic, the magnetization and demagnetization branches are separated by a maximum of 6 mT, corresponding to the hysteresis loop of Section 2.4.2. Figure 5.7 can furthermore be compared to the classical, idealized hysteresis loops in Figure 2.2, where the hysteresis branches are more well-defined and separated in the classical case, while the operational measurements show discrete bands of field response. Consequently, the realized magnetic field

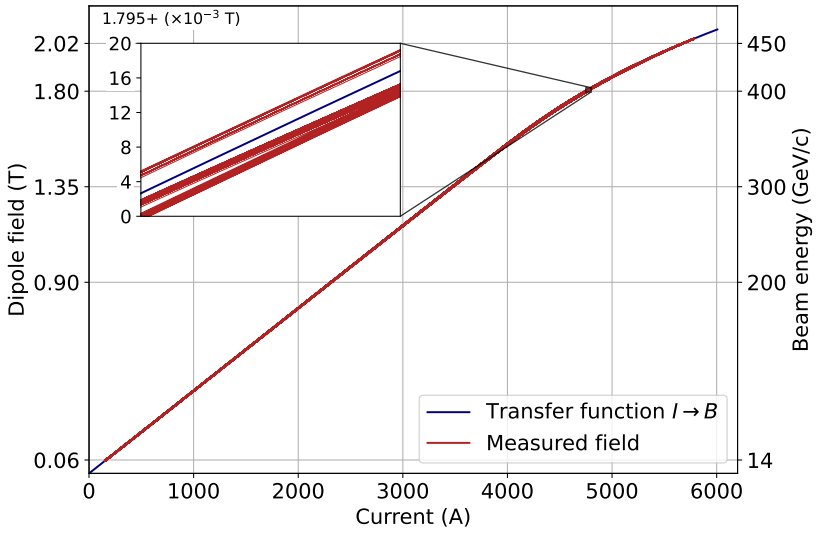


Fig. 5.7: Measured magnetic field of the main dipole circuit with measurements of all operational cycles, including SFTPRO, LHC, HiRadMat, AWAKE, and MD1 cycles, with an anhysteretic I–B calibration function analogous to Figure 4.12b. The inset shows a zoomed view close to the SFTPRO flat top, with excitation between 4750 and 4800 A, where the major hysteresis loops can be distinguished, with about 6 mT difference between the lower (magnetization) and upper (demagnetization) branches, and represent the maximum hysteresis that can be observed operationally.

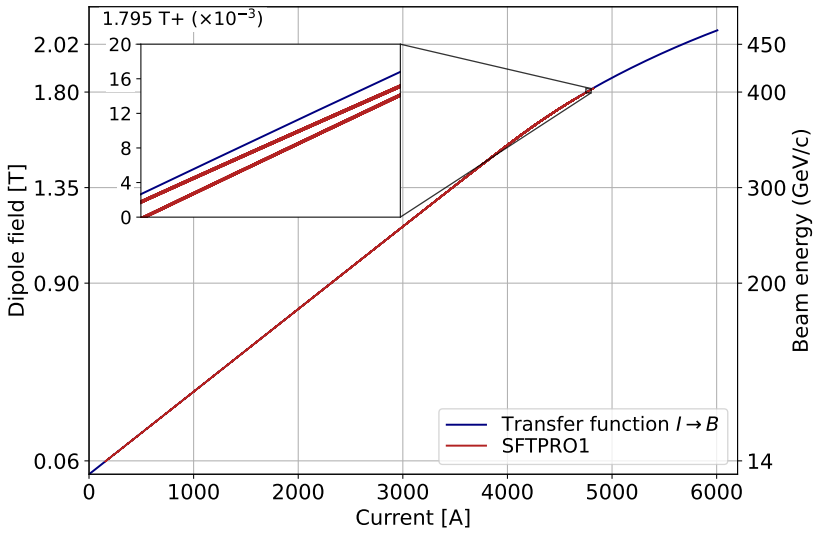


Fig. 5.8: Measured magnetic field of the main dipole circuit with sample measurements of the SFTPRO cycle across many excitations and supercycle transitions, and the anhysteretic I–B calibration analogous to Figure 4.12b. The inset axes stays analogous to Figure 5.7, and shows specifically the upper and lower branches of the SFT cycle under different excitation. It can be seen that cycle-to-cycle hysteresis, that is field reproducibility error, is significantly smaller than the major hysteresis loop, with a maximum difference of 5×10^{-4} T under operational conditions.

response from the main dipoles may differ from the as-programmed magnetic field by the same amount, dependent on the calibration function.

Figure 5.8 instead shows only the SFT cycle measurements with the same static transfer function. It is seen that cycle-to-cycle variation in the magnetic field close to the flat-top remains below 5×10^{-4} T both during magnetization and demagnetization, which is supported by the observations in Section 5.2.1. The SFT cycle is used here as it is the main physics cycle present across different supercycles.

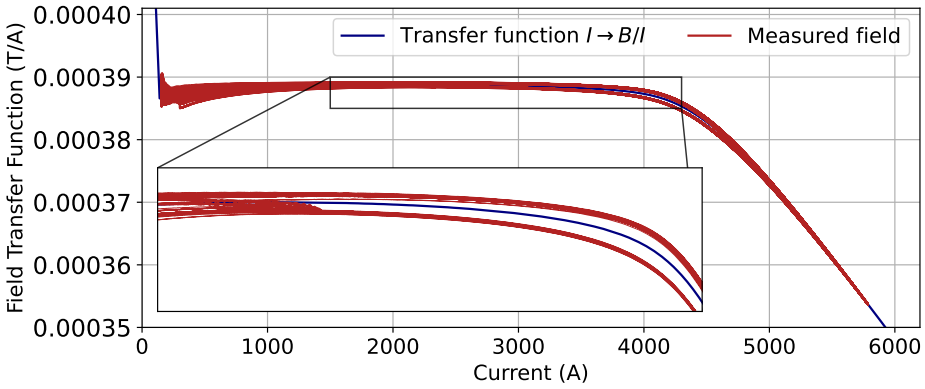


Fig. 5.9: Illustration of the same field measurements of Figure 5.7, plotted as the the $I \rightarrow B/I$ transfer function which is more common practice in magnetic measurements. The inset plot shows a magnification between 1500 A and 4300 A, covering the most of the linear magnetic response, and shows more clearly the separation of magnetization and demagnetization hysteresis branches, while also showing the much smaller cycle-to-cycle changes, also shown in Figure 5.8.

Furthermore, Figure 5.9 shows a different perspective of Figure 5.7, as the transfer function $I \rightarrow B/I$, which shows more clearly the separation between upper and lower magnetization branches, for measurements of all operational cycles.

5.2.4 Eddy Current Decay in the SPS Main Dipoles

Figure 5.10 shows an example of the eddy current decay introduced in Section 2.6.2, measured in the main dipoles by the spare B-Train, where an idle current period of 3.6 s, simulating an MD1 precycle in length, follows the ramp-down of the preceding 450 GeV/c LHCPILOT cycle, prior to the start of the next SFT cycle. In this static current period, the field can be seen decaying almost 2×10^{-4} T over the 3.6 s plateau while the current is kept stable. Using the metrics from Table 5.2, and hypothesizing the absence of the 3.6 s idle period, a field decay at 2×10^{-4} T at the SFT injection, would theoretically decay the closed orbit by over 11 mm, which would be considered operationally unacceptable, and even cause significant beam loss. In reality, the SFT cycle injects at 200 ms after cycle start and not immediately at start cycle, following a small ramp to 14 GeV/c injection at corresponding excitation of 161 A, however the field decay remains significant. Depending on the preceding cycle, the ramp length and rate, the idle period between cycles, and the injection time of next cycle, the field decay is non-negligible, and non-trivial to predict, and compensate for during arbitrary sequences. This puts a strong requirement of keeping predictable and reproducible eddy current decay, in addition to the hysteresis effects, in order to successfully inject beam.

Decaying field during cycle execution during ramp up or at the start of flat top is similarly operationally problematic, however is often corrected for through the static cycle settings, as intra-cycle decay is predictable due to the known excitation profile, and can be compensated for in the cycle design. In contrast, the inter-cycle decay is fully dependent on the preceding cycle, and is not predictable without a model of the magnetic history, and would theoretically require cycle-by-cycle correction of the main field in less rigid cycle sequences than is currently operational.

[TODO: Even fully flat (absence of previous ramp), can be problematic if a ramp-down was expected.]

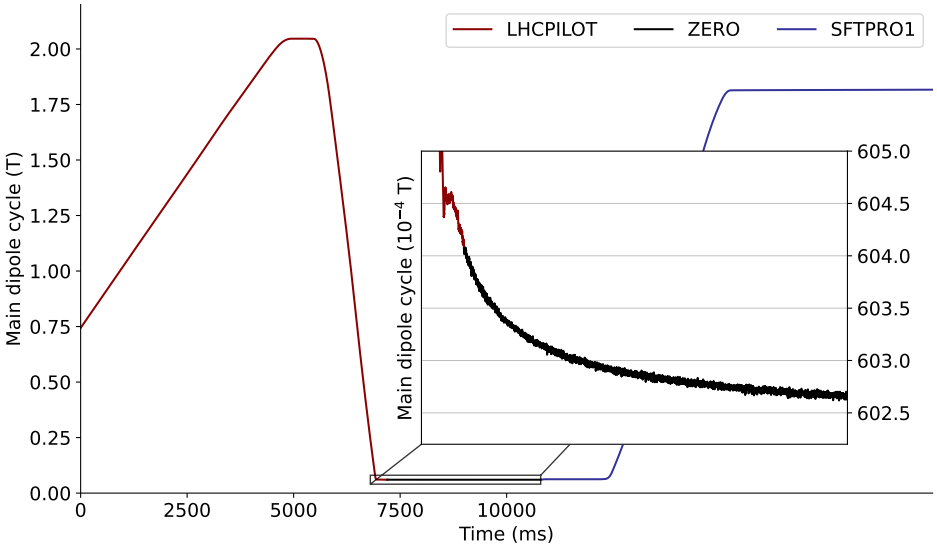


Fig. 5.10: Plot of the main dipole field, as measured by the spare B-Train, where an idle current period of 3.6 s follows the ramp-down of the preceding 450 GeV/c LHCPILOT cycle. A decaying field can be seen immediately after the ringing effect of the power converter that occurs immediately after the ramp down, and the field can be seen decaying over the full 3.6 s plateau where excitation current is kept stable at the 155.63 A idle current. The field can be seen decaying almost 2×10^{-4} T over the decay.

5.3 Problem Statement

The previous sections established that field errors caused by hysteresis and eddy current decay have measurable consequences for beam parameters in the SPS.

Cycle instances in an operational sequence are indexed by j, with cycle type $c(j)$ and programmed excitation $I_{c(j)}(t)$ for a given magnet circuit. The magnetic history $\mathcal{S}(j)$ encodes the cycle types, together with the actually played excitation current and realized field, of the cycles preceding instance j. The realized field is then

$$B_j(t) = \mathcal{F}(I_{c(j)}(t), \mathcal{S}(j)), \tag{5.1}$$

for an unknown, magnet-specific operator $\mathcal{F}$ capturing both the static remanence of Section 2.4.2 and the dynamic eddy current memory of Section 2.6.2. Equation 5.1 restates formally what

Section 5.1 already argued, that the field realized by a cycle is not fixed by its own programmed excitation alone.

We define field reproducibility as a relative property. A magnetic field error only degrades beam performance relative to whatever field the beam has already been tuned against, so a reference field must be fixed for each cycle type before reproducibility with respect to it can be stated. The reference field $B_c^*(t)$ for cycle type c is the field that reproduces itself under repetition of that cycle type, a fixed point of $\mathcal{F}$ restricted to cycle type c. It depends only on $\mathcal{F}$ and $I_c(t)$, and shifts whenever the excitation $I_c(t)$ does.

A tolerance $\varepsilon = \min_t \varepsilon_{\text{beam}}(t)$ is derived from requirements across the cycle for the main dipole circuit in Section 5.2. No equivalent derivation exists for the quadrupole or higher order circuits, and ε for those families is left undefined here, as an open problem.

We define a cycle instance j as reproducible if

$$\left| \mathcal{F}(I_{c(j)}(t), \mathcal{S}(j)) - B^*_{c(j)}(t) \right| < \varepsilon \quad \text{for all } t, \tag{5.2}$$

and the target of this thesis is for Equation 5.2 to hold not for an isolated occurrence but across an entire operational sequence.

Validation of the reproducibility criterion proceeds in two ways. Offline, against held out sequences with field measurements trusted as ground truth, the criterion can be checked directly. Operationally, validation is carried out at beam level, since reproducible beam parameters is the goal established in Section 5.1, with the choice and treatment of the observables used developed in the next section and carried out in Chapter 8.

Solving hysteresis, in the sense used throughout this thesis, means achieving cycle by cycle field reproducibility with respect to B_c^*, across the range of histories encountered in operation, for every magnetic history effect captured by $\mathcal{F}$. Effectively, this means capturing the static remanence of hysteresis and the dynamic decay of eddy currents in $\mathcal{F}$.

5.4 Proposed Approach

We propose to solve the problem stated in Section 5.3 by parametrically modeling the operator $\mathcal{F}$ of Equation 5.1, and by using that model to predict the field $B_j(t)$ realized by a given cycle instance j before it is played. Given a reference field $B_c^*(t)$, the predicted field allows a correction to be computed ahead of time and applied through the existing accelerator control system, so that the compensation is realized as a feed-forward adjustment to the programmed excitation rather than as a new control loop or operational procedure. The compensated quantity is field, or the unnormalized field gradient for higher order circuits, since this is the quantity that is physically realized in a magnet, that underlies the beam parameter effects of Section 5.1, and that is represented, directly or indirectly, in the control system settings that a correction can act on.

A model of $\mathcal{F}$ intended for this purpose must be able to predict field from an arbitrary point within an operational sequence, since a correction is only useful if it can be computed for the sequences actually played in operation, rather than for a restricted class of trajectories beginning from a known reset state. This requires that whatever internal state the model carries forward be initializable from a finite, measurable window of past excitation and field, and not from a canonical seed condition such as full desaturation. This requirement recurs in Sec. **[TODO: ref]** when classical hysteresis representations are evaluated, since it is precisely the property that distinguishes a model usable at arbitrary points in an operational sequence from one usable only from a defined starting condition.

The approach is furthermore designed to be family agnostic. The main dipole, quadrupole, sextupole, and octupole circuits of the SPS all use the same hierarchical settings management system, and the correction mechanism proposed above depends only on that shared structure, not on properties specific to any one family. A predicted field also reaches circuits that carry no absolute field measurement in operation, such as the quadrupole and higher order families, for which only the main dipole is measured by the B-Train, and which a correction driven by the measured field, as proposed by Velotti et al. [135], cannot address. The design goal is therefore a compensation scheme applicable across all four families, even though, as discussed in Section 11.8, the measurements required to realize a trained model for every family were not equally attainable within this work.

5.4.1 Beam-Based Approaches

An alternative to modeling the field directly is to measure the beam's response, such as orbit or tune error, and to translate that error into a corresponding correction of the underlying circuit or directly compute the change in beam parameter. We do not pursue this approach, since it treats a symptom of the field error rather than its cause, and because beam observables are not uniquely determined by the circuit responsible for the error of interest.

The orbit is set primarily by the main dipole field, but it is also shaped by the RF frequency and the associated control loops, so that an orbit deviation does not on its own identify a dipole field error, particularly at injection where energy matching plays a significant role in capture. Orbit is furthermore only observable through BPMs while beam is present and bunched in the machine, which excludes debunched conditions such as those encountered during slow extraction.

The tune is coupled to both quadrupole families, QF and QD, and to the beam momentum and radial displacement, since the field gradient of the quadrupoles is normalized by the beam rigidity of Equation 2.1. A dipole or RF induced energy mismatch can therefore shift the observed tune without any error in the quadrupole field itself. Tune measurement precision is further limited at high beam intensity, and the tune itself is intensity dependent through collective effects such as space charge and wakefields. An absolute tune-based estimate of a quadrupole field error is therefore difficult to establish reliably, and is furthermore beam-specific. Radial displacement additionally alters the horizontal focusing strength, adding to ambiguity between quadrupole error and dipole or RF induced offset.

The same difficulty extends to the higher order circuits. Chromaticity, which reflects the sextupole field, typically requires several measurements across a cycle rather than a single one, and both sextupole and octupole effects can only be meaningfully isolated once the cycle to cycle error of the preceding circuit orders has already been corrected. Every beam based estimate is therefore relative to the beam conditions under which it was obtained, and those conditions, including intensity and optics configuration, are not fixed between measurement sessions.

For these reasons, a beam based method would need to be developed separately for each circuit family and would generally provide only local estimates of field error, valid at specific points in a specific cycle, such as injection or extraction. This does not scale to the requirement of Section 5.3, which is reproducibility of field across arbitrary operational sequences and across all four circuit families. We therefore treat the problem directly at the level of the magnetic field, rather than through its beam level consequences.

5.4.2 Data-Driven Modeling

Any model of the operator $\mathcal{F}$, phenomenological or otherwise, must ultimately be fit to measured data, since $\mathcal{F}$ is not derived from first principles for an assembled production magnet but inferred

from observed excitation-field pairs. Before considering which model class is appropriate, it is useful to recall what the data presented in Section 5.2 already shows about the modeling target.

Figure 5.7 showed that the measured main dipole field and the nominal transfer function are visually indistinguishable across the full excitation range, and that the hysteretic deviation only becomes visible once the current axis is zoomed to a narrow window. This is consistent with the scale argued in Section 5.2, that the field reproducibility tolerance ε is small relative to the total field range. The modeling target is therefore a small deviation superimposed on an otherwise well known relationship, and any candidate model class must be able to represent this small-signal, history-dependent structure at the accuracy this problem requires.

Two sources of data support fitting a model to the field response of the SPS magnets. The first is the operational history logged by the B-train and the accelerator control system, spanning years of cycle-by-cycle excitation and field records across the full range of sequences played in the SPS. This volume is what makes a data-driven approach viable in principle, since it provides many realizations of the operator $\mathcal{F}$ under the histories actually encountered in operation. It is, however, an incomplete sample of the input space relevant to $\mathcal{F}$, since it only contains the sequences that operations happened to play, and not the broader space of excitation histories against which a general model should be validated.

The second source is dedicated measurement in controlled laboratory conditions, namely the CERN magnetic measurement benches as presented in Section 3.3, where a spare magnet from the main ring can be excited with arbitrary current sequences rather than only operational ones. This capability is necessary because a model intended to generalize across arbitrary histories, as required in Section 5.4, cannot be validated, and arguably cannot be reliably fit, using operational data alone. The accuracy and reproducibility of these measurements in turn bound the accuracy any model fit to them can achieve.

Together, these two sources of data motivate treating $\mathcal{F}$ as a data-driven modeling problem, since both a sufficient volume of realizations and a means of sampling arbitrary excitation histories are available to fit and validate a model against.

5.4.3 Phenomenological Modeling Limitations

A full three-dimensional, time-dependent field model is the only representation that makes no simplifying reduction of the magnet, resolving its distributed geometry and material state, three-dimensional saturation and fringe fields, and the eddy current dynamics that depend on the longitudinal structure of the laminations and coil ends, in addition to the static remanence of Section 2.4.2. Two separate obstacles make such a model unusable here. First, constructing it requires explicitly meshing every lamination and insulation layer, a multiscale transient problem that is not identifiable from nominal production data and would demand a dedicated calibration and mesh refinement campaign to approach the required accuracy. Second, even a finely meshed and accurately calibrated model would remain far too expensive to evaluate within the sub-second to tens-of-seconds intervals at which the SPS operates, and so could not serve as a runtime field predictor regardless of its precision. Its most credible role is therefore as an offline source of physical structure or synthetic training data for a reduced model, not as a runtime field oracle. A two-dimensional static model such as ROXIE [115] is a further reduction of this same target, solving only the axially invariant field distribution $B(x, y \mid I)$ for a given geometry. Because it has no time dimension, it cannot represent remanence, and because it assumes invariance along the magnet axis, it cannot capture the end effects and longitudinal eddy current motion that the three-dimensional case above identifies as physically significant.

The ODE models reviewed in Section 2.5.2, such as the Jiles–Atherton or Duhem formulations, reduce the magnetic state to a small number of lumped variables under a fixed-order differential

equation, which fixes in advance a restricted class of admissible trajectories and, being rate-independent, carries no mechanism for eddy current dynamics. An optimizer may fit a major loop or a chosen set of ramps well, while the same fixed-order parameters fail on the minor loops, altered reversal points, varied ramp rates, long dwell times, and changing preceding cycles that make up arbitrary operational histories, which disqualifies these models for the histories this thesis must cover.

The Preisach model and its operator-based relatives (Section 2.5.1.1) represent memory as an ensemble of elementary hysteretic operators, a richer structure than a low-order ODE that approximates a wider range of hysteretic behavior. As established there, it remains rate-independent and rests on congruency and wipe-out assumptions that real iron yokes violate, requiring extensions that weaken its predictive guarantees [46].

A differentiable Preisach model relaxes some of this rigidity and can be fit to measured data by gradient-based optimization. Its rate-independent, operator-superposition structure nonetheless remains, so it carries no mechanism for the rate-dependent eddy current dynamics of Section 2.6.2. It further assumes a canonical, fully demagnetized seed state and offers no natural way to encode an arbitrary initial condition, departing from the initialization requirement established in Section 5.4. Section **[TODO: XXX]** in the next chapter supports this directly, since a differentiable Preisach model fit to the cycle-to-cycle sequences measured for this thesis reproduces arbitrary hysteresis loops only to some extent, and does not reach the cycle-to-cycle accuracy required by the reproducibility criterion of Equation 5.2.

The response this thesis must model is that of an assembled, production magnet, which carries distributed magnetic and dynamic states, from three-dimensional geometry and saturation to eddy current decay and relaxation, that are not generally identifiable from current-field measurements alone and cannot be captured by a single fixed lumped ODE or scalar Preisach operator over arbitrary operational histories. The scale of the required accuracy compounds this difficulty. If the hysteretic deviation lies at the permille level of a unit-range field measurement, as suggested by Figure 5.7, a 1×10^{-4} absolute target requires resolving the hysteretic contribution to roughly ten percent of its own magnitude. Where the cycle-to-cycle differences that matter are themselves below 1×10^{-4}, the model must resolve a small difference between two large, nearly identical field trajectories. This constitutes a stringent residual-modelling problem, distinct in character from the loop-fitting problem addressed by the hysteresis models discussed above and by the machine-learning approaches of Section 2.5.3.

5.4.4 Residual Modeling with Machine Learning

We propose to approach the residual-modeling problem identified above as a multivariate time series forecasting, or function approximation, task using the machine learning approaches introduced in Section 2.5.3. Given the excitation current and realized field over some past window, together with the programmed excitation for a future interval, the model predicts the absolute field over that future interval for the corresponding future cycle. This can be formalized as a function approximator $\hat{\mathcal{F}}_\theta$ parametrized by θ, such that

$$\hat{B}_j(t) = \hat{\mathcal{F}}_\theta(I_{c(j)}(t), \mathcal{S}(j)), \tag{5.3}$$

where $\hat{B}_j(t)$ is the predicted field for cycle instance j, and $\mathcal{S}(j)$ is the measured excitation and field history preceding that instance. The model can be fit to measured data by minimizing a loss function $\mathcal{L}$ over a training set of cycle instances, such as the mean squared error between predicted and measured field. Tracking a reference field $B_c^*(t)$ against this prediction then allows

a correction ΔB to be computed and applied to the control system, as proposed in the opening of this section.

Machine learning is suited to this problem for two reasons. A sufficiently expressive learned function approximator is not restricted, by construction, to a fixed-order state or a prescribed class of admissible trajectories, unlike the models discussed in Section 5.4.3, and its capacity can be matched to the data rather than fixed in advance by a choice of physical state variables. Such models are also naturally fit through mini-batch gradient-based training on large volumes of sequence data, matching the two data sources established in Section 5.4.2, where the classical models were shown to be either structurally inapplicable or difficult to train at all.

A sequence-to-sequence model, of the kind introduced in Section 2.5.3, is proposed, as the measured field response of a cycle instance is a time series, and as the realized field response of a cycle instance depends not only on the program, and prior static state, but also on the dynamic evolution of the excitation and field over time. Hysteresis itself could in principle be captured by modeling only the extrema of the excitation history, like a Preisach model, using a static $I \to B$ mapping to predict the field from the current, but the dynamic effects of eddy currents require a model that also captures time-dependence, which a static mapping cannot. This means that the model must be implicitly, or explicitly time-dependent. Analogously, using cycle type alone to encode history is also insufficient, as the same cycle can be altered dynamically during execution by the power converter, as discussed in Section 4.4.6.3, which further motivates a field-native sequence-to-sequence model. The choice of model architecture, and the reasons for rejecting alternatives, is developed in Chapter 7.

The model is proposed to be architecturally cycle-agnostic, predicting field at arbitrary excitation and at any point within a cycle, rather than being restricted to a fixed set of reference trajectories. This directly satisfies the initialization requirement, since a model that can be queried at an arbitrary point in an excitation sequence can equally be initialized from an arbitrary, measurable window of past excitation and field, rather than from a canonical reset state. The prediction horizon is additionally required to span a full cycle, not only a single point within it. Consequently, the model constitutes a magnetic model of the circuit itself, applied at inference time primarily to operational cycles, rather than a model of any single reference cycle.

Although this work refers to the task as hysteresis compensation, in line with accelerator practice, the quantity being modeled is not limited to intrinsic ferromagnetic hysteresis. The model represents the effective history-dependent response of the assembled magnet as it appears in the measured field signal, which includes hysteresis, rate-dependent eddy current effects, and geometry-dependent nonlinear field contributions together. A single model capturing all of these effects jointly is intended to serve as a digital twin of the magnetic circuit, in the same sense that the B-train is intended to be representative of the field actually realized in the ring.

Lastly, the model is proposed to predict absolute field rather than a correction ΔB directly, because the reference field $B_c^*(t)$ against which a correction is defined is not fixed. It shifts whenever the programmed excitation $I_c(t)$ is retuned, which occurs comparatively rarely for the main dipoles in beam operation but more frequently for the main quadrupoles and sextupoles, through tune and chromaticity correction respectively. A model trained to predict a correction relative to one specific reference would need to be retrained, or at least reinterpreted or conditioned on additional information, each time that reference changes. A model of the absolute field response instead remains valid under any choice of reference, and is therefore reusable across such changes.

Training a model using a pure data-driven approach puts high requirements on the measurements it is trained on, since the model can only be as accurate as the data it is fit to. For a sequence model, it requires long, continuous measurements of both excitation and field, at a high enough sampling rate to capture the fast transients of eddy current decay.

5.4.5 Data Requirements and Measurements

Fitting an accurate model of $\mathcal{F}$ places high requirements on the magnetic measurements it is fitted on. Reproducibility here means that repeating the same magnetic sequence under the same conditions must return the same measured field. We set the measurement reproducibility target at or below $\mathcal{O}(1 \times 10^{-5}\,\mathrm{T})$, a fraction of the field reproducibility tolerance ε of Section 5.3, so that measurement error is not the limiting factor on model accuracy. A full order of magnitude below ε would constrain $\mathcal{F}$ even further, and remains the more conservative bound where measurement instrumentation allows it. ε itself corresponds to the most sensitive point in the most sensitive cycle, so it is already a deliberately conservative bound, and measurements above it in absolute terms may still constrain $\mathcal{F}$ usefully at less sensitive points in a cycle, or for other circuit families. A further fundamental restriction requires the magnetic measurements to be absolute rather than relative. A relative measurement discards the field offset carried over from prior excitation history, the quantity through which hysteresis and eddy current memory are expressed, and that a model of $\mathcal{F}$ is intended to capture.

The quantity to be measured is whichever field-related quantity is used in the control system for a given circuit, with main axis dipole field B for the dipole family, integrated gradient $G\,\mathrm{d}\ell$ for the quadrupole family, and the analogous integrated field derivatives $\int b_3\,\mathrm{d}\ell$ and $\int b_4\,\mathrm{d}\ell$ for the sextupole and octupole families. Standard magnetic measurement campaigns at CERN instead target field quality at a fixed reference radius and the higher order field harmonics defined in Section 3.1.1, quantities that characterize magnet-to-magnet variation and that feed directly into the alignment, placement, and shimming corrections used to absorb that variation elsewhere. Hysteresis and eddy current effects are macroscopic and reproducible across magnets of a given family, so they are not addressed by those corrections, and are the effects this thesis measures. These quantities are output as a time series at the same 1 kHz sampling rate as the B-train, giving laboratory and operational measurements a common format, and able to capture the fast transients of eddy current decay.

The main dipole family can be measured using the existing B-train system, whose spare magnet is equipped with an NMR marker providing an absolute field reference and is continuously logged to NXCALS, as discussed in Sections 3.4.3 and 4.5.2. The main dipole family is additionally measured directly at the CERN magnetic measurement laboratory (Section 3.3), where a spare MBB magnet, equipped with its vacuum chamber, can be excited with arbitrary current sequences rather than only the sequences played operationally. Beyond the bench's standard fluxmeter, additional instrumentation from the techniques of Section 3.2, such as stretched wire systems or Hall probes, can be installed as needed. This makes the data logged by the B-train a suitable source for a proof-of-concept dataset for the dipole family, since a model can be trained directly on it and its predictions validated against the same measurement used operationally. The B-train is nonetheless an incomplete sample of the full input space relevant to $\mathcal{F}$, as established in Section 5.4.2, and offers no control over which sequences are played outside of dedicated measurement sessions. These are, in effect, parasitic measurements, since the sequences played are dictated by operational needs, not by this thesis's modeling requirements, and the B-train's own measurement equipment is similarly outside our control. It also measures the MBA and MBB dipole subfamilies together, averaged into a single signal, whereas a laboratory measurement can only access one magnet at a time. The two measurement systems also differ in calibration, noise, and drift, so the datasets they produce are not directly comparable or interchangeable. The data logged by the B-train therefore complements a dedicated laboratory measurement campaign designed to cover a broader excitation space than operational sequences alone provide.

For the main quadrupole family, and by extension the sextupole and octupole families, dedicated laboratory measurement is intended to proceed analogously. A spare main quadrupole magnet

is available in the measurement laboratory, specifically a QD unit, and since the QF and QD magnets share an identical physical design, differing only by a ninety degree rotation, a QD measurement is representative of the family as a whole. Spare sextupole and octupole magnets are similarly available, and the measurement campaign was designed to be carried out analogously across all four families.

The measurement design is proposed to combine dedicated hysteresis characterization cycles, synthetic combinations of ramps and dwell times intended to excite arbitrary hysteresis states, and varied ramp rates intended to excite eddy current effects, together with pseudo-random excitation sequences constructed to be compatible with design constraints of the power converter and magnet. Pseudo-random sequences are included specifically to cover a broader region of the excitation input-output space for both static and dynamic effects than either fixed reference cycles or narrow parameter sweeps would reach, which is necessary to provide the machine learning approach of the previous section with sufficiently varied training data. Operational sequences and cycles, together with random permutations of operational cycles, are proposed to be measured in addition, to ensure that the hysteresis and eddy current states actually encountered in operation are represented in the training data.

These measurements, combining logged operational data and dedicated laboratory characterization, form the dataset against which the model proposed in the previous section is trained and validated, with the measurement campaign itself detailed in Chapter 6.

5.4.6 Feed-forward Compensation in the Control System

Through the timing system and the events it publishes over the middleware (Section 4.4.2, Section 4.4.3), the evolution of the cycle sequence can be followed in real time. The next cycle and user is known 2.5 s in advance, which is the forewarning period available for prediction and correction.

The model of Section 5.4.4 predicts the field quantity x expected to be realized over the coming cycle, where x denotes the control system representation of the field, B for the main dipole circuit and $G\Delta L$ for the quadrupole circuits. Let $x_c^*(t)$ denote the reference value of this quantity for cycle c, and $\hat{x}_c(t)$ the corresponding model prediction. The correction to be applied is then defined as

$$\Delta_c(t) = x_c^*(t) - \hat{x}_c(t), \tag{5.4}$$

evaluated over the relevant interval of the cycle. The correction is applied as a trim to LSA on the programmed quantity, and propagated through the existing settings hierarchy to the circuit's FGC layer following Figure 4.11, so that the equivalent change in excitation current is set before the next cycle is played. Whether a correction was applied in time can only be confirmed from the measured excitation current of the circuit, which is published to the middleware only once the cycle has completed, and requires experimental validation to confirm that the 2.5 s forewarning period is sufficient. Overlap is in principle possible given the 1.2 s basic period of the complex, since a correction could then be written to a cycle already playing. As the shortest cycle is realistically 3.6 s, and cycles carrying beam are longer still, this is not operationally possible, and the 2.5 s forewarning falls within a single cycle. An overview of the proposed feed-forward compensation scheme is shown in Figure 5.11.

For the main dipoles, the B-Train provides the initial state $\mathcal{S}$ used to seed the model at the start of a cycle, and is also available cycle-by-cycle as a feedback signal once the cycle has completed and data is published to the middleware. Additionally, rather than always seeding from the B-Train, the model prediction is proposed to allowed to be toggled to run in open loop, using its own previous prediction as the initial state for the next cycle, with the B-Train then used only for validation. This is proposed as a way to validate the open loop tracking quality of

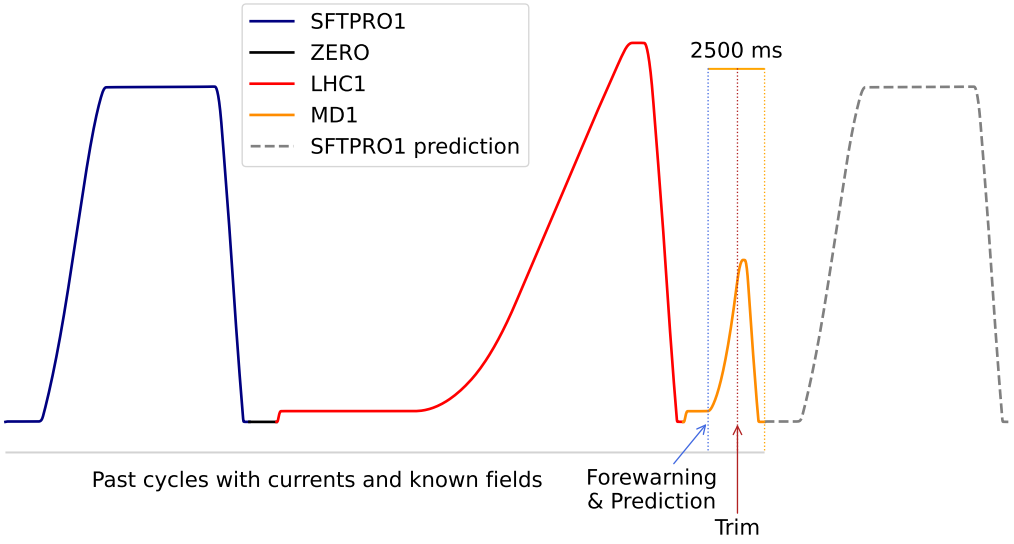


Fig. 5.11: Proposed feed-forward compensation scheme. The model predicts the field expected to be realized over the next cycle, here an SFTPRO cycle, and a correction is computed with respect to the reference and applied to LSA as a trim to the programmed excitation prior to cycle start. This method applies analogously to any cycle, and to any magnet family.

the model, since the higher order magnets have no equivalent online measurement and open loop tracking is the only available proxy for their state.

In cases where the initial state of the magnet is unknown, or whenever the prediction and compensation is in open loop mode, we propose to precycle the magnets to a known states in the SPS before the next cycle is played, so that the model can be initialized to a known state and the open loop predictions continue. This is proposed to be carried out with a cycle that is operationally played and saturates the magnets, and repeating the cycle a number of times to ensure wipe-out of the previous state, where the number of repetitions is to be determined experimentally. Specifically, the precycle must be achieved using an operational cycle, as creation of a dedicated precycle would require approval from management and many commissioning steps, which are out of scope for this thesis. This precycle is distinct from the dedicated MD1 precycle of Section 4.6.1.

A further complication in magnetic state management arises from the power converter economy modes of the FGC in Section 4.4.6.3, which can alter the realized current independently of the programmed function announced at forewarning time. While full economy is announced at the same time as the cycle, dynamic economy is triggered asynchronously, and the realized current, and therefore resulting field response can only be inferred once the cycle has started. This affects any correction at forewarning time, and therefore the field compensation control loop needs to track economy modes, in addition to the cycle that is announced. This is particularly important for open-loop control, where the latent magnetic state must be updated whenever economy modes are triggered. Even for closed-loop control where B-train seeds every initial state, economy mode changes the history of the magnet, and must be tracked to ensure that the next cycle is predicted correctly.

For testing and control room use during sessions with live beam, a frontend client is proposed to be developed that plugs into the control system by subscribing to the relevant middleware timing events and data streams. The application tracks the cycle sequences played, maintains

the internal state of the model across cycles, applies the computed correction to LSA directly, and allows the effect on the field in the following cycle to be observed. It is also intended to visualize the played field and current alongside the predicted field and the tracked reference, and to allow the reference to be retaken when needed. Since the client is agnostic to the underlying prediction model, it can also serve as a monitoring and control surface for the prediction routine itself, including its parameters, and as toggle for the feed-forward compensation scheme.

Beyond the frontend client, the same tracking and correction logic is designed to be run without a human operator in the loop, using the server-based UCAP platform, introduced in Section 4.4.5. UCAP can subscribe to the same middleware events, maintain internal state across cycles as required by the open loop case, and republish the computed correction as an LSA trim directly. This offers a path to headless, production tracking of the cycle sequence and compensation, developed further in Section 8.1.4.

The implementation of the feed-forward compensation scheme under these requirements, and its experimental validation, are developed in Chapter 8.

5.4.7 Compensation Validation Strategy

The field compensation strategy proposed in the previous section, and the trained residual model from Section 5.4.4 is proposed to be validated using operational cycles and beam.

The proposed cycle-to-cycle compensation will be tested in the SPS with the main dipoles as primary subject, during dedicated MDs time outlined in Section 4.7 where full control of the SPS supercycle can be taken. In total, a maximum of 20 to 30 hours of dedicated MD time is available per year, during the 3 year span of this thesis, and the proposed compensation scheme is designed to be tested within that time. In addition to the dedicated MD time, further parallel MDs can be scheduled to validate the compensation scheme, even if the magnetic sequence does not remain under our control. Furthermore, parasitic tests can be carried out during normal operation, as the model prediction and intended feed-forward correction can be observed without actually applying the correction to the control system, in a dry run mode.

The first goal of the validation is to successfully compensate the magnetic field of the SFT beam, which is the only operational cycle that requires active compensation during normal, rigid, operation. As the hysteresis-induced change in flat-top field is large in magnitude in comparison to the lowest beam sensitivity threshold, and the beam perturbation during routine operation is significant, this is the first and most important validation of the proposed compensation scheme. In this case, the compensation needs only to be applied at the flat top of the cycle, where the beam is extracted. Based on BCT measurements we can quantify the improvement in spill quality in the presence of the compensation scheme, and compare it to the uncompensated case. Furthermore, the reproducibility of the field can be evaluated quantitatively by comparing the measured field with and without compensation, with respect to the reference field. We propose that the compensation scheme is considered successful if we can demonstrate stable spill across all currently operational proton sequences, shown in Section 4.3. Ion sequences are not included in this goal due to the short run time of the annual ion runs, and while the extracted spill is perturbed similarly to the proton case, the operational and experimental requirements are less stringent.

The second goal of the validation is to demonstrate that we can successfully compensate the magnetic field of the SFT beam at injection, which is a prerequisite for removing the MD1 precycle, as the precycle is currently used to ensure reproducibility of the magnetic field at injection. In addition to hysteresis, this also includes the effects of eddy currents, which are more pronounced at injection due to the fast down-ramp of the precycle preceding injection. The relevant beam property to compare is the radial beam position across the injection plateau. In

this case, we propose that the compensation scheme is considered successful if we can demonstrate stable beam position, with deviation below 1 mm along the injection plateau, across all currently operational proton sequences, same as for the first goal. The 1 mm threshold is as control room operators would normally re-tune the beam manually if the threshold was exceeded during routine operation. If the MD1 precycle can be removed, this would also allow us to remove it from the supercycle completely, and inject beam immediately after a longer ramp-down, such as an LHC cycle, which puts more emphasis on eddy current effects, and is a more stringent test of the compensation scheme.

The third goal of the validation is to demonstrate that we can successfully compensate the magnetic field of the SFT cycles across supercycles not normally seen during routine operation, such as asymmetric supercycles as in Fig. **[TODO: SFTx2+LHCPILOT SC]** . Furthermore in this case, supercycles can be recomposed rapidly to demonstrate beam stability across different sequences and to demonstrate that the proposed model is able to generalize to unseen sequences. Following this validation, the compensation scheme is proposed to be extended to any SPS cycle that require field compensation besides the SFT beam, which at this stage has not been defined.

Open-loop control with internal state tracking and autoregressive model prediction is proposed to be validated in the same way as closed-loop control, but with the B-train used only for validation and not for continuously seeding the model. This will demonstrate the feasibility of the compensation scheme and prediction model to work with the higher-order magnets. If open loop control is successful, the same measurement campaigns, residual modeling, and validation goals are proposed for the SPS quadrupole and sextupole families. However, the field compensation benefits of the higher-order magnets can only be directly observed after successful compensation of the main dipoles through the optics coupling mentioned in the previous section. Therefore, the primary validation goals proposed in this thesis are focused on the main dipoles, and the higher-order magnets are proposed to be validated in a follow-up study after initial success. The execution of this validation strategy and the operational results with beam are developed in Chapter 8

Chapter 6

Field Measurement Campaign

The field model of the following chapters is fit to measured data, and its accuracy is bounded by the fidelity of that data. Two sources are evaluated in this chapter against the acceptance criteria fixed in Section 5.4.5, namely an absolute field measurement, a reproducibility of order 1×10^{-5} T, a rate of 1 kHz, and coverage of arbitrary excitation sequences.

The first is dedicated measurement on the laboratory dipole bench, presented in Section 6.1, where a spare main dipole is excited with designed waveforms under controlled conditions. The second is the operational B-train, presented in Section 6.2, which measures the field online during operation as the cycles play. Each source is assessed for the measurement fidelity it provides and the excitation histories it can cover.

6.1 Dipole Laboratory Measurements

Following the bench specifications in Section 3.3.1, and goals outlined in the previous chapter, measurement campaigns were launched with the spare MBB dipole in the laboratory. The measurements were divided into three categories: operational sequence setup, absolute measurements with field markers, and arbitrary sequences. As operational sequences are the most important for training and the most relevant to the SPS operation, they were prioritized for initial measurements. They were also used for initial measurements to establish a measurement baseline, and in order to compare measurement reproducibility with the B-train measurements, even in the absence of an absolute field marker. A second reason for proceeding without the field marker was the lack of an existing NMR probe installed in the bench and the lack of a spare probe available for permanent installation. Using an NMR probe would require reinstallation and recalibration for each measurement opportunity in the lab, so initial measurements were conducted without it.

In parallel with setup of the bench, waveforms were designed to measure the operational sequences, as well as the arbitrary sequence. The Flexible Software Framework for Magnetic Measurements (FFMM) framework [7] was used to control the bench and acquire measurements, which supports passing arbitrary waveforms to the power converter, in any constant frequency waveform, in addition to piecewise linear excitations, including trapezoidal shapes. The power converter then plays the waveform, and the FFMM application triggers the measurement synchronously with the waveform or based on a programmed time delay or threshold value. As the programmed current of neither operational cycle nor arbitrary sequences could be designed freely inside of the FFMM framework, each designed measurement had to be exported into a 1 kHz text file, and then imported into the measurement application. The 1 kHz was chosen to be sufficient as it exceeds the regulation period of the SPS ring power converter, and was therefore deemed to be sufficient to program smooth transients.

For the operational sequences, the programmed current for each operational cycle was taken from the LSA database, and then chained together according to the sequences in Section 4.3 in the shape of supercycles. Supercycle transitions were included in the designed waveforms, for

instance, initially playing a SFTPRO-only sequence, transitioning to an LHC filling sequence, and then returning to the original SFTPRO-only sequence for 5 repetitions of each sequence. Since the machine cycles from the SPS always initialize from an idle current of 155.61 A, but the lab power converter must initialize from zero current, the designed waveforms had to start from zero, ramp up to idle current, before playing the operational sequence. Analogously, the end of the design waveforms had to ramp down to zero current following the end of the desired measurement sequence to avoid tripping the power converter. Furthermore, a 10 s idle period was added at the end of the waveform to simplify voltage offset estimation of the fluxmeter in post-processing. While, from a theoretical perspective, the ramp down from idle current to zero and back to idle current between measurements constitutes a minor loop, it was deemed negligible following precycling of the magnet. The connecting ramp-up and ramp-down were designed to be sufficiently slow to avoid eddy-current effects on the first cycle following the ramp, which in practice became a ramp over 2 s, effectively 75 A/s. The ramp was also used as a trigger to start the measurement acquisition in FFMM, where the trigger threshold was set to 100 mV, equivalent to 60 A, after which all signals were acquired.

The acquired signals included the measured current from the 16 bit Direct Current Transformer (DCCT), the raw unintegrated voltage from the fluxmeter center coil, and the voltage measured from the magnet coil. Later iterations included Hall probe measurements, which are discussed in Section 6.1.2. The sampling frequency was set to 100 kHz to improve the signal-to-noise ratio of the fluxmeter output when integrated. The raw voltage from the fluxmeter was integrated in post-processing to obtain the integrated flux, which was then divided by the effective area of the search coil to obtain the average field of the magnet, following Equation 3.7. The effective area of the search coil was known in advance as part of the bench setup.

Prior to every measurement, the magnet was precycled to a reproducible state, following best practices of the laboratory, to ensure that the magnet is in a known state before the measurement is taken. While not strictly necessary for modeling hysteresis, where the goal is to measure the remanent field at different points in the hysteresis loop, for baseline measurements of the machine sequences, the magnets were nonetheless precycled. In practice, this was done using a trapezoidal waveform, with a maximum current of 5800 A, matching the highest current of any dipole cycle in operation, below the power converter maximum current of 6000 A, and repeated 10 times. The ramps were chosen to be as fast as possible to stay below the RMS current limit of the power converter, and dwells were reduced to 0.5 s on the flat top, which, while not allowing the magnet to fully settle, is still sufficient to ensure the magnetic state after the precycles.

The following subsections present the operational-sequence measurements, the introduction of Hall probes and averaging, and the limitations imposed by the power converter. Each is evaluated against the acceptance criteria fixed in Section 5.4.5.

6.1.1 Operational Sequences Measurements

Figure 6.1 shows a measurement of a machine sequence. In this sequence, a standard SFTPRO sequence, followed by an LHC filling transition, and back to a standard SFTPRO sequence is simulated. The original standard sequence was modified to simplify post-processing and follow the power-converter limitations that will be outlined in Section 6.1.3. In this sequence, the standard 10.8 s long SFTPRO with a ramping 4800 ms flat top was replaced with the 7.2 s long SHIP machine development cycle, which has a 1200 ms flat top at constant current, and a similar injection plateau. The LHC filling cycle was similarly replaced with a shorter LHCPILOT cycle, which reaches the same excitation with the same injection. Furthermore, the MD1 precycle was suppressed, and the program was kept at idle current during the 3.6 s precycle duration, to allow measurement of the field response of the fixed-target cycle without a precycle.

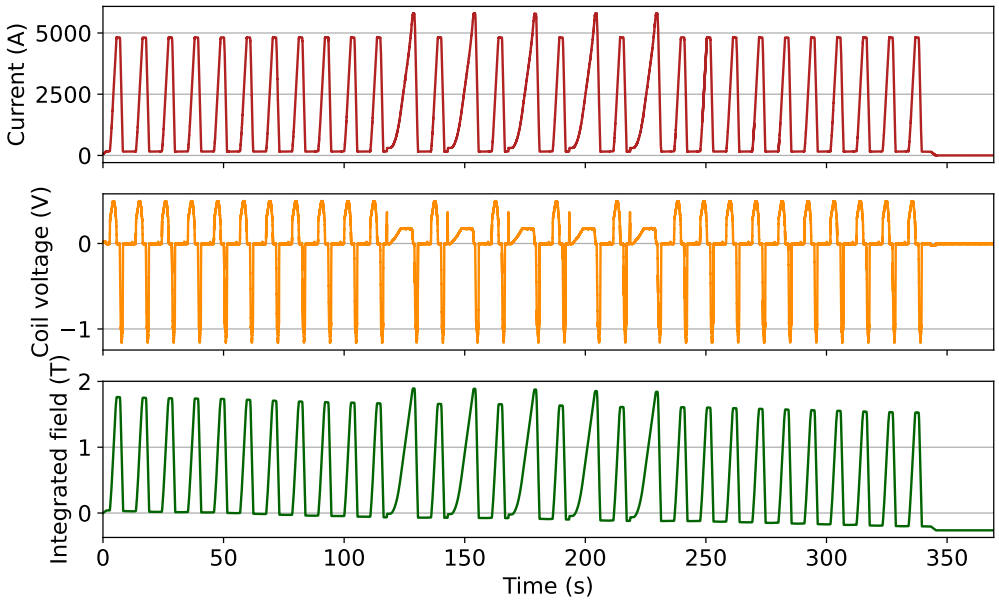
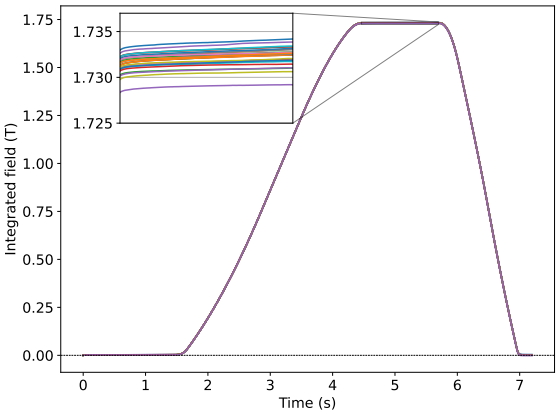


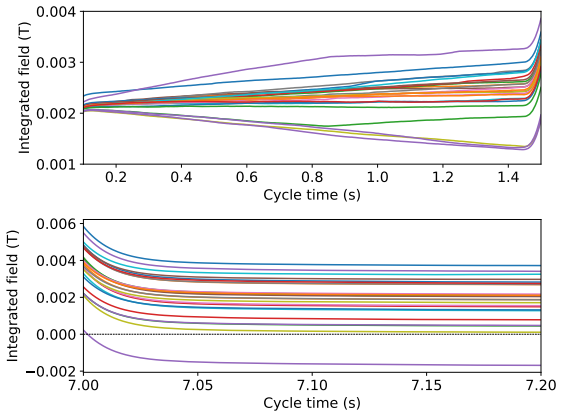
Fig. 6.1: Measurement of an operational sequence, consisting of a modified SPS SFTPRO physics supercycle, followed by an LHC filling supercycle, and then returning to the original sequence. The sequence was modified to replace the SFTPRO cycle with a SHIP cycle, which has an equivalent flat-top field and the same injection plateau. The MD1 cycles between the SFTPRO cycles were also flattened to idle current. The top panel shows the measured current from the DCCT, the middle panel shows the raw, unintegrated voltage from the fluxmeter center coil, and the bottom panel shows the integrated field from the fluxmeter after calibration and voltage-offset correction. The measurement was taken at a sampling frequency of 100 kHz, following the practices outlined in Section 6.1.

The top panel of Figure 6.1 shows the measured current from the DCCT, the middle panel the raw fluxmeter voltage, and the bottom panel the field obtained by integrating the voltage and calibrating by the effective coil area. No field marker was applied, so the integrated field is referenced to zero at the start of the measurement. A single global voltage offset, computed from the average of the final 10 s of dwell time and equivalent to 1 million samples, was subtracted from the raw voltage prior to integration. This correction is insufficient, and the integrated field settles 0.17 T away from its baseline after the current has returned to zero, across the 350 s sequence. The offset is structural: the fluxmeter measures a voltage proportional to the rate of change of flux, and any residual bias in this voltage integrates into a field error that accumulates over the measurement duration.

The accumulated drift is reduced by integrating each cycle individually and applying a separate voltage offset correction per cycle, estimated from the dwell time preceding or following the cycle. A constant voltage offset integrates into a linear ramp, so subtracting a per-cycle offset removes the linear component of the drift. Referencing each cycle to zero at its start additionally provides a per-cycle anchor in the absence of an absolute field marker, from which the reproducibility of the field measurement can be assessed directly.



(a) Average field as calibrated integrated flux for each SHIP cycle, measured in Figure 6.1 and overlaid to show measurement reproducibility. Each cycle is integrated independently and marked from zero at the start of the cycle, with a voltage offset correction applied individually to each cycle. Macroscopically, integration drift is not visible, but the inset plot of the flat top shows a drift-fan of 5 mT.



(b) Zoomed inset of the injection plateau (top) and between ramp-down and cycle end (bottom), of the panel to the left. The top panel shows a drift-fan of 2 mT, whereas the bottom panel shows a range of 6 mT. Both measurements are taken at intervals in the cycle with constant current.

Fig. 6.2: Illustration of SHIP cycle integral field measurements, showing the measurement drifting at constant current, as shown in the inset in the left panel and in the panels to the right.

The integrated field of each SHIP cycle was computed with this per-cycle correction and overlaid in Figure 6.2. At full scale, the cycles coincide, and the measurement appears reproducible. The injection plateau, shown in the top panel of Figure 6.2b, reveals a drift of 2 mT across the 1.2 s plateau at constant excitation, with individual cycle occurrences drifting in opposite directions. This fanning is not consistent with physical field evolution. The current is constant across the plateau, the remanent field is fixed within it, and eddy current decay is excluded by the 3.6 s idle period preceding each SHIP cycle. This idle period greatly exceeds the few tens of milliseconds of eddy decay observed at cycle end in the bottom panel of Figure 6.2b. The residual is therefore attributed to integration drift. The settled traces in the bottom panel span 6 mT at the end of the cycle, as large as the full separation between the upper and lower branches of the major hysteresis loop measured by the B-train in Figure 5.7. This spread cannot be entirely physical, as the cycle-to-cycle variation of the remanent field is far smaller than the major loop, and is therefore dominated by integration drift accumulated by the end of the cycle.

Since a per-cycle offset has already been removed, the remaining drift is nonlinear and originates from a nonstationary voltage offset that a single dead-time estimate cannot track. Some cycles also show a change in the monotonicity of the integrated field within the constant-current plateau, where the field is expected to be constant or, under eddy-current decay, monotonic without abrupt changes in slope. Correcting the nonlinear component would require re-anchoring the integration at intervals within the constant-current cycle, where the coil voltage is near zero. Operational cycles do contain plateaus at injection and extraction, but these are often not at constant current, so they do not provide reliable re-anchoring points; only the dead time surrounding the cycle remains available as an anchor, which removes only the linear component. Cycles with clean constant-current plateaus, or dedicated dead time inserted between successive cycles for offset

re-estimation, could in principle be designed, but such waveforms would no longer represent the operational sequences whose field response the measurement seeks to characterize.

Drift-free integration of the coil signal has been demonstrated for dipole field measurements at CERN by fusing the coil with a co-located second sensor, a Hall probe or the magnet excitation current, through a first-order Kalman filter, reducing the integrated field drift by three orders of magnitude [8]. The fusion anchors the low-noise coil dynamics to the absolute level of the second sensor, so the reproducibility of the fused estimate cannot exceed that of the reference sensor. Neither available reference reaches the required level at the operational field. The Hall probe carries a systematic absolute-level error at high field, shown in Section 6.1.2, which is not random and is therefore not reduced by the filter, and the excitation current is an order less accurate and not reproducible between measurements in the laboratory. The fused estimate would inherit these errors, both above the 1×10^{-5} T reproducibility required to resolve cycle-to-cycle hysteresis, and the method was therefore not adopted here.

An absolute field marker such as an NMR probe would fix the absolute level of the integrated field, but would not mitigate this drift. Under cycle-to-cycle variation of the hysteresis, the remanent field on a plateau manifests as a constant offset in the integrated field, not as a drift, so an absolute anchor corrects the offset while the drift continues to accumulate away from the anchor. The per-cycle referencing to zero already constitutes a best-case anchor, at least as strong as an NMR marker applied once per cycle, and the drift persists in its presence.

The chained operational sequence measured here represents the most favourable case for drift correction, as it contains long constant-current plateaus between cycles from which per-cycle offsets can be estimated, and the SHIP cycle is furthermore the shortest physics cycle in operation and therefore accumulates the least integration drift. The drift across the injection plateau nonetheless remains after per-cycle correction. This 2 mT drift exceeds the cycle-to-cycle hysteresis signal of 5×10^{-4} T and lies two orders of magnitude above the 1×10^{-5} T measurement target. The measurement uncertainty therefore exceeds the signal of interest, and the cycle-to-cycle hysteresis is not resolvable from the fluxmeter measurement even though the major hysteresis loop is clearly visible.

The excitation current recorded by the DCCT at the SHIP flat top is shown in Figure 6.3. The measured current carries a 50 Hz mains ripple of ± 0.7 A and is quantized in steps of 0.1 A from the 16 bit digitization of the DCCT. A 50 Hz notch filter followed by a low-pass filter removes both, shown by the green trace, and the filtered current is adequate as a model input.

The measured current sits approximately 8 A, or SI 0.17 per cent, above the programmed control waveform, a systematic error in the power converter current regulation. This offset could be nulled by recalibrating the programmed waveform to reach the target current, but such a recalibration is of limited value, as the played current is not reproducible between measurements and drifts over time, as examined in Section 6.1.3.

As the field model is trained on measured current paired with measured field, this offset does not by itself introduce an inconsistency, and the involuntary variation of the excitation current adds coverage of the excitation space. It does, however, reduce the ability to compare the laboratory field response one-to-one against the operational sequence played in the ring. The played current can exceed or fall short of the intended value, observed here as an offset of 0.17 % relative to the $\sim$4800 A flattop. The magnet therefore reaches a correspondingly different peak excitation than the programmed cycle and excites more or less hysteresis. The measured field nonetheless corresponds to the actual played current, so the current-field pairing used for training remains physically consistent, and the departure from the intended cycle follows the observed excitation offset. Even after filtering, the current is not perfectly constant at the flat top, drifting by 0.1 A–0.2 A across the plateau, equivalent to a field variation of several times 1×10^{-5} T. As this variation follows the measured current, which is preserved through filtering,

it is reproduced deterministically by a model driven by the current. It does not by itself limit the achievable modelling accuracy, which is bounded by the fidelity of the field measurement and by the calibration of the current. The variation of the played current between measurements, examined in Section 6.1.3, is the more consequential limitation.

Although the 50 Hz ripple is present in the excitation current, it is strongly attenuated in the magnetic field, as the laminated iron yoke shields high-frequency excitation and the field does not follow the current at 50 Hz. The residual ripple that reaches the field introduces a small measurement uncertainty due to the phase of the ripple. The mains ripple is asynchronous with the cycle trigger, so repeated measurements of the same cycle sample the ripple at different phases, and the field at a fixed cycle time takes a range of values that a model driven by the filtered current cannot resolve. This uncertainty is removed by band-limiting the field to the same bandwidth as the filtered current input, so that neither carries the ripple, and is in any case small given the strong attenuation of the ripple in the field.

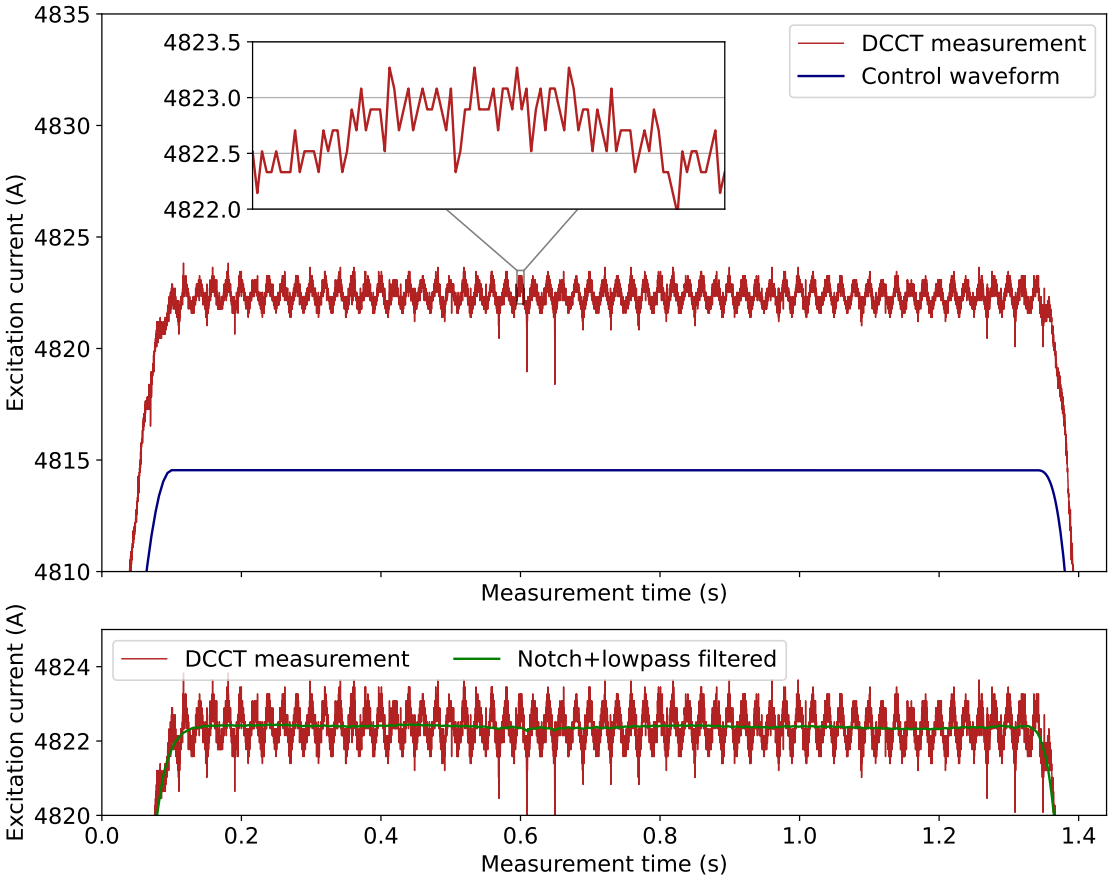


Fig. 6.3: Illustration of the measured current at the 1.2 s flat top of one SHIP cycle, where in this program the current is held constant. The blue trace shows the programmed waveform as sent to the power converter, while the red trace shows the measured current from the DCCT. The 50 Hz ripple can be clearly seen in the measured current, and the quantization of the current can be seen in the inset figure, where the current is shown to be quantized in 0.1 A steps.

The excitation current is therefore an adequate model input, and the limiting factor of the laboratory measurement remains the integration drift of the fluxmeter field. Two routes to reduce

this drift were investigated, averaging over repeated measurements and Hall probes that measure the field directly without integration, and are examined in the following section.

6.1.2 Hall Probe and Measurement Averaging

To measure the field without the integration drift of the fluxmeter, Hall probes were installed on the bench. As described in Section 3.2.2, a Hall probe measures the local field directly and instantaneously, without integrating an induced voltage, and is sampled fast enough to follow the transients of a machine cycle, including the eddy-current decay. Among the techniques of Section 3.2, it is the only direct-field instrument suited to pulsed measurement, as an NMR requires a homogeneous and quasi-static field and is incompatible with a ramping excitation, and the induction coil is the drifting instrument being replaced. The direct measurement removes the integration drift at the cost of a higher instantaneous noise level.

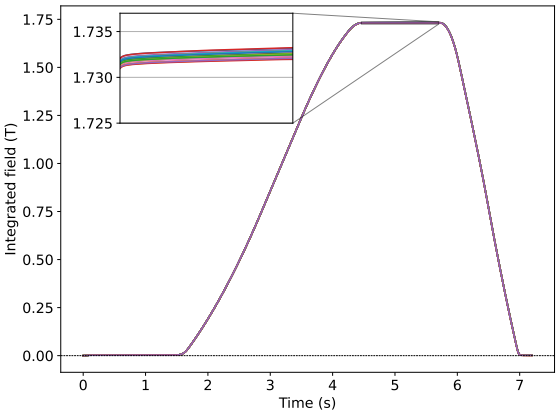
A method to reconstruct the integral field of the dipole from local Hall sensors was developed on the same spare MBB dipole [126]. The integral field is approximated as a linear combination of $k = 4$ longitudinal HE244 Hall sensors, each cross-calibrated against an NMR system and linearized with a third-order polynomial, with the combination weights optimized by least squares against a 7 m integral induction coil calibrated by the single-stretched-wire method. The reconstructed integral field is translated to an average field over the magnetic length, the same quantity obtained from the integrated fluxmeter coil. The filtered reconstruction reached an average-field standard deviation of 1.4×10^{-5} T over a 2 s flat-top window taken after the eddy currents had decayed, comparable to the integral-coil reference and to the B-train.

At the level of a single measurement, the weighted sum of the four calibrated Hall sensors carries significant noise. Over the SHIP flat top, evaluated on a window of 10 000 samples, the weighted-sum average field has a standard deviation of 1.38×10^{-3} T without filtering, reduced to 1.44×10^{-4} T by a third-order Savitzky-Golay filter, a reduction of about one order of magnitude. The window is restricted to 10 000 samples because the Hall probes exhibit a slow thermal drift over the flat top, shown at the end of this section, so the field cannot be assumed stationary over the full plateau and a longer window would inflate the standard deviation with real field change. The filtered value nonetheless remains an order of magnitude above the 1×10^{-5} T target.

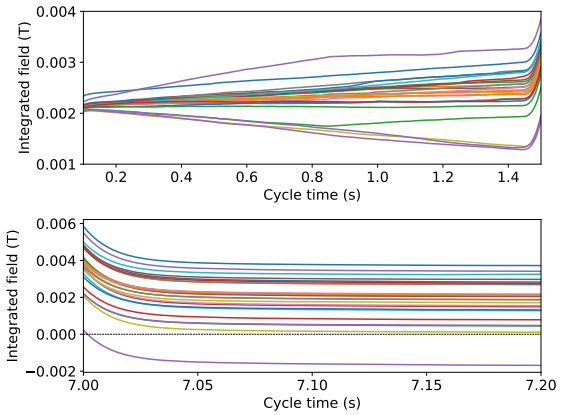
The noise could in principle be reduced by averaging the signal over time within a single measurement. Temporal averaging, however, reduces noise only by lowering the bandwidth of the signal, and the averaging window is bounded by the fastest field transient that the model must reproduce. The eddy-current dynamics of the cycle extend to tens of hertz, so a window long enough to bring the noise to target would attenuate the very transients the measurement is intended to capture, and the flat-top averaging that is admissible does not extend to the ramps where the field changes fast. The 1.4×10^{-5} T reported above was reached by exactly this long temporal averaging over a static flat top, and is not attainable for a dynamic operational cycle. The noise remaining after filtering is moreover not white, as it retains mains harmonics and the slow sensor drift, which a short temporal window does not remove.

Averaging instead over repeated, temporally aligned measurements of the same waveform reduces the noise at every point of the cycle without reducing its bandwidth, and cancels the mains ripple through its differing phase between acquisitions. The measurement shown in Figure 6.1 was repeated 16 times, each measurement precycled as described in Section 6.1, and temporally aligned to the trigger and initial ramp from zero current.

Averaging over repeated measurements not only reduces the Hall probe noise. Applied to the integrated fluxmeter, the same set of measurements provides a diagnostic of the nature of the fluxmeter drift. The drift characterization of Figure 6.2 was repeated with the coil voltage averaged over the 16 measurements, and the result is shown in Figure 6.4. A separate



(a) Average field as calibrated integrated flux of every SHIP cycle as measured from Figure 6.1, where each coil measurement is averaged across 16 measurements of the same waveform. Each cycle is integrated independently and marked from zero at the start of the cycle, with a voltage offset correction applied to each coil measurement prior to averaging. Macroscopically, integration drift is not visible, but the inset plot of the flat top shows a monotonically increasing drift-fan of 1.5 mT.



(b) Zoomed inset of the injection plateau (top) and between ramp-down and cycle end (bottom), of the panel to the left. The top panel shows a monotonically increasing drift-fan of 0.4 mT, whereas the bottom panel shows a range of 1.2 mT. Both measurements are taken at intervals in the cycle with constant current.

Fig. 6.4: Illustration of the SHIP cycle integral field where the coil voltage measurements are averaged over 16 identical and temporally aligned measurements, as compared to Figure 6.2, which shows significantly more integration drift. The inset and panel axes are left intentionally the same as the previous plot for comparison.

voltage offset was computed for each measurement and each cycle before averaging, and the integrated field of each cycle was referenced to zero at its start, as in the single-measurement characterization. Averaging reduces the fanning of the injection plateau from 2 mT to 0.4 mT and the cycle-end spread from 6 mT to 1.2 mT but does not remove it, and a monotonic drift within each cycle survives. Because the offset is subtracted per measurement and per cycle, any zero-mean random component averages toward zero, so the remaining drift is biased and correlated between measurements. The drift is therefore a nonlinear systematic integration drift, and averaging over repeated measurements reduces only the uncorrelated part and cannot remove it. The averaged integrated field consequently still carries a drift-induced variance across cycle occurrences. At the flat top of the cycle, shown in the inset of Figure 6.4a, the averaged integrated field is flat and slightly rising, consistent with the eddy-current catch-up expected after the ramp.

Figure 6.5 shows the Hall measurement of an LHCPILOT cycle at the injection plateau, averaged over the 16 measurements, together with the averaged and unaveraged fluxmeter coil for comparison. Averaging the weighted Hall sum over the 16 measurements gives a standard deviation of 1.4×10^{-4} T over a window of 40 000 samples in the middle of the plateau. A 50 Hz notch, which removes the mains pickup on the Hall bias supply, followed by a 60 Hz low-pass filter, reduces this to 4.7×10^{-6} T. The reduction is set by the low-pass bandwidth, and the 60 Hz cut-off preserves the eddy-current dynamics, so unlike the temporal averaging discussed above, it does not sacrifice bandwidth. Noise is suppressed both across the 16 measurements

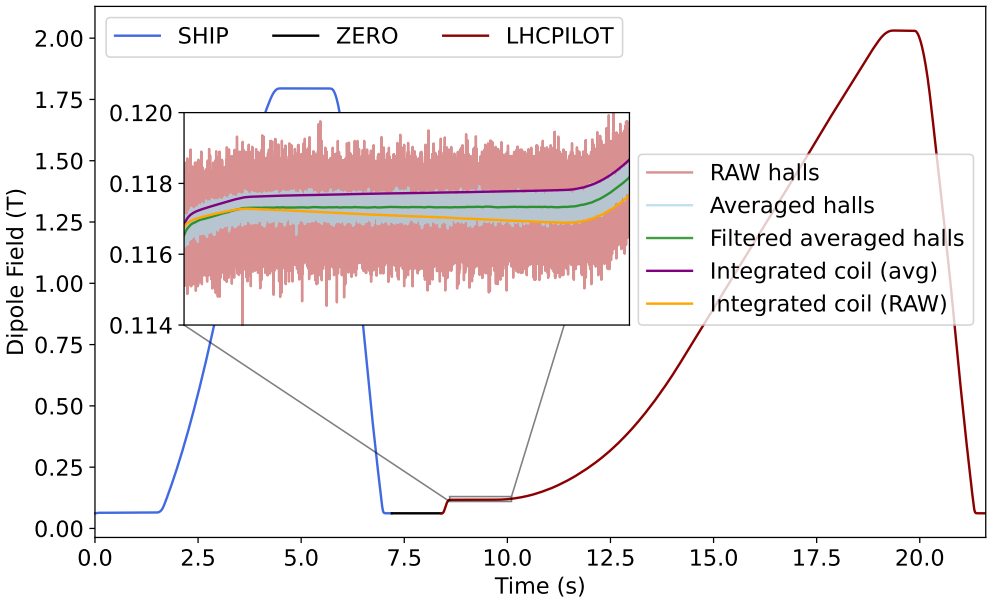


Fig. 6.5: Average field as measured by the Hall probe (faded blue) and fluxmeter coil (purple), averaged over 16 identical and temporally aligned measurements of the same LHCPILOT cycle. The fluxmeter coil was integrated and calibrated to average field, and marked using the Hall probe as an absolute field marker at the start of the cycle. A notch- and low-pass-filtered averaged Hall probe measurement (green) is shown, with a standard deviation of 4.715×10^{-6} T over 40 000 samples in the middle of the injection plateau. Additionally, the non-averaged, non-filtered hall probe measurement (faded red), and non-averaged, integrated fluxmeter coil measurement (orange) are shown for comparison.

and across the four Hall sensors of the weighted sum, and the resulting value lies below the 1×10^{-5} T target. This is a best case, however, obtained at the low field of the injection plateau where the current is low and stable, and it relies on the 16-measurement average without which a single measurement does not reach the target. The window is restricted to 0.4 s of the 1.2 s plateau because the field is not stationary over its full length, so a longer window would inflate the standard deviation with real field change.

Averaging over repeated measurements is only valid if the repeats are identical. Each measurement was precycled identically, as described in Section 6.1, so that every repeat starts from the same magnetic state, and was temporally aligned to the trigger so that corresponding features coincide. Identical precycling removes the variation in initial remanent state that would otherwise add variety to the training data, so the noise reduction from averaging is obtained at the cost of the excitation coverage noted in Section 6.1.1. Reproducibility of the excitation current between repeats is a further condition, as a current that differs from one measurement to the next produces a different field, so the average combines fields of different cycles and does not reduce the noise of a single one. This condition is not met at the flat top, as shown in Section 6.1.3, where the played current is not reproducible between measurements. Averaging also relies on the mains ripple, which appears at a different phase in each acquisition and cancels in the average, while the 50 Hz notch removes any residual. Temporal alignment must finally be precise on the ramps,

as a misalignment between repeats has no effect on a constant-current plateau but distorts the transient on a ramp, where the field changes rapidly, and the average no longer represents a single cycle.

Averaging N measurements reduces the uncorrelated noise by a factor of $\sqrt{N}$, so the 16 measurements are expected to reduce it by a factor of four. At the SHIP flat top, the unfiltered standard deviation of the weighted Hall sum falls from 1.38×10^{-3} T for a single measurement to 4.2×10^{-4} T for the average, a factor of 3.3. This falls short of the ideal factor of four, which indicates that a component of the noise is correlated between measurements and is not reduced by averaging.

Repeated acquisition is furthermore costly in time, as playing the 16 repeats of this comparatively short waveform takes over 1.5 h, excluding setup time and the precycling required before each measurement, so the approach scales poorly to the longer sequences and larger number of waveforms that a full dataset would require.

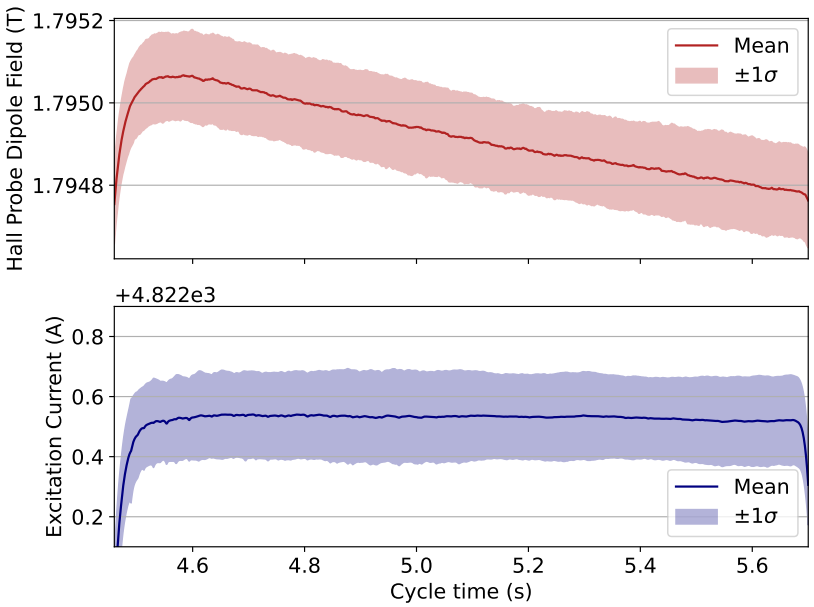


Fig. 6.6: The top plot shows the mean and standard deviation of the already averaged Hall probe measurements over 16 identical and temporally aligned measurements, at the SHIP cycle flat top, across all occurrences in the measurement sequence shown in Figure 6.1. The bottom plot shows the mean and standard deviation of the current measurement over the same cycles. The current measurement shows a near-constant standard deviation of 0.15 A across the plateau, while having a near-constant mean under constant excitation. The averaged hall probe measurements show a similarly constant 1.4×10^{-4} T standard deviation, however show a monotonic decrease over the same flat top, in the opposite direction of any physically motivated drifts like eddy current decay. The drift is also inconsistent with the average field as measured by the fluxmeter seen in the inset of Figure 6.4a.

Figure 6.6 shows the averaged Hall field and the excitation current at the SHIP flat top, taken across all occurrences of the cycle in the measurement sequence. The current is held constant, with a flat mean and a standard deviation of 0.15 A across the plateau, whereas the Hall field drifts downward by approximately 3×10^{-4} T over the 1.2 s plateau. Averaging across the cycle occurrences also averages out the cycle-to-cycle hysteresis, but this does not affect the observation,

as hysteresis appears as a static offset at constant excitation and contributes to the width of the standard-deviation band, which stays constant across the plateau, and not to the monotonic drift of the mean. The drift appears in the mean and not in the band, so it is common to every occurrence of the cycle and is reduced neither by the spatial weighted sum nor by the average over the 16 measurements.

The downward direction of the drift is not consistent with the field in the magnet, as the eddy-current dynamics cause the field to lag during the ramp and to recover upward toward its asymptotic value while the shielding currents decay. At the same flat top, the fluxmeter measurement shown in the inset of Figure 6.4a is flat and slightly rising, in agreement with this expectation and consistent with a stable field across the plateau. This disagreement between the two instruments suggests that the drift originates in the Hall sensor and not in the field, consistent with the temperature-dependent drift of Hall sensors noted in Section 3.2.2.

The drift is on the order of 3×10^{-4} T, about 30 times the 1×10^{-5} T target and comparable to the 5×10^{-4} T cycle-to-cycle hysteresis signal the measurement seeks to resolve. The low-pass filter and averaging do not reduce it, as it is correlated between measurements. Its magnitude is apparent at the flat top only because the current is held constant there, and the drift may also be present during the ramps, where it would be hidden within the changing field. The observed 3×10^{-4} T is in this sense only the portion accumulated over the plateau, and the absolute level reaching the plateau may already be offset by drift earlier in the cycle. Unlike the noise, this drift cannot be mitigated, so the Hall probe does not provide the absolute field to the required accuracy at the operational field, despite its low noise floor.

6.1.3 Power Converter Limitations

The averaging on which the fluxmeter and Hall noise reductions of the previous subsection depend requires the same current to be applied on every repeat, a condition the power converter does not meet. Figure 6.7 shows the mean and standard deviation of the flat-top current measured by the DCCT for each SHIP cycle instance in the chained sequence, computed across the 16 repeated measurements. The mean current differs by up to 0.2 A between cycle instances, and within a single instance it scatters between the 16 measurements with a standard deviation between 0.2 A and 0.35 A. The larger bound is equivalent to 0.0073 % of the flat-top current and corresponds to a field variation of approximately 1.4×10^{-4} T at one standard deviation.

This scatter removes averaging as a means to reduce the measurement noise. The reduction of the uncorrelated fluxmeter and Hall noise relies on repeating the same waveform, so that the common field response is retained while the uncorrelated components are reduced by $\sqrt{N}$. As the played current differs between repetitions, the repeated measurements do not correspond to the same excitation, and the field spread from the current variation enters the average. This spread is above the 1×10^{-5} T target and comparable to the 5×10^{-4} T cycle-to-cycle hysteresis signal, so averaging cannot recover the field of a defined excitation to the required accuracy.

The played current also drifts over longer timescales, as shown in Figure 6.8, where the mean flat-top current measured over 7 h rises from 4819.8 A to a maximum of 4821.1 A near 15:30 and decreases thereafter, a range of about 1.3 A equivalent to a field change of order 5×10^{-4} T. The drift follows the working day and is consistent with thermal drift in the converter and its regulation electronics, with a possible contribution from grid-side load variation; the reversal near 15:30 coincides with the reduction in daytime demand, though the exact cause is not resolved. This drift is the reason the recalibration of the programmed waveform considered in Section 6.1.1 does not provide a stable correction, as the offset between played and programmed current changes over the measurement period and day.

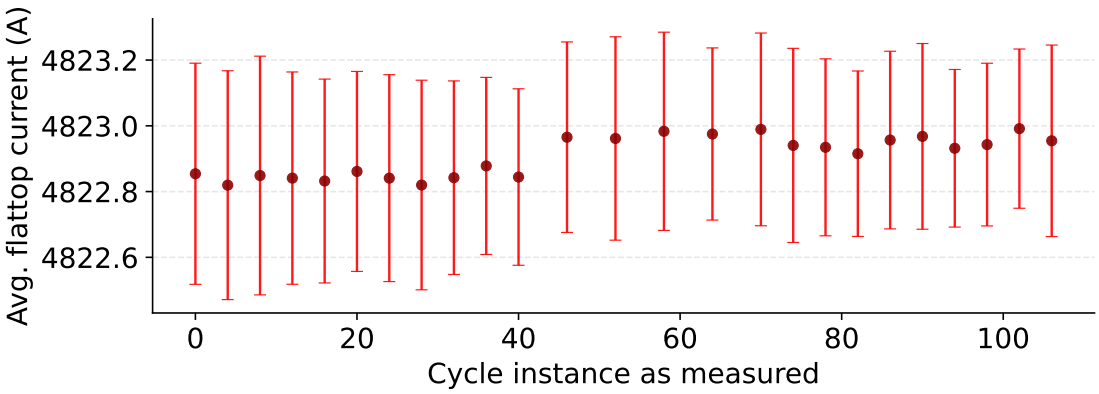


Fig. 6.7: Average and standard deviation of mean current as measured by the DCCT at the 1.2 s flat top of each of the SHIP cycle instances as shown in Figure 6.2, computed across 16 measurements of the same waveform, as described in Section 6.1.2. The measurements show a variation of 0.2 A across cycles, with a standard deviation between 0.2 A and 0.35 A per cycle.

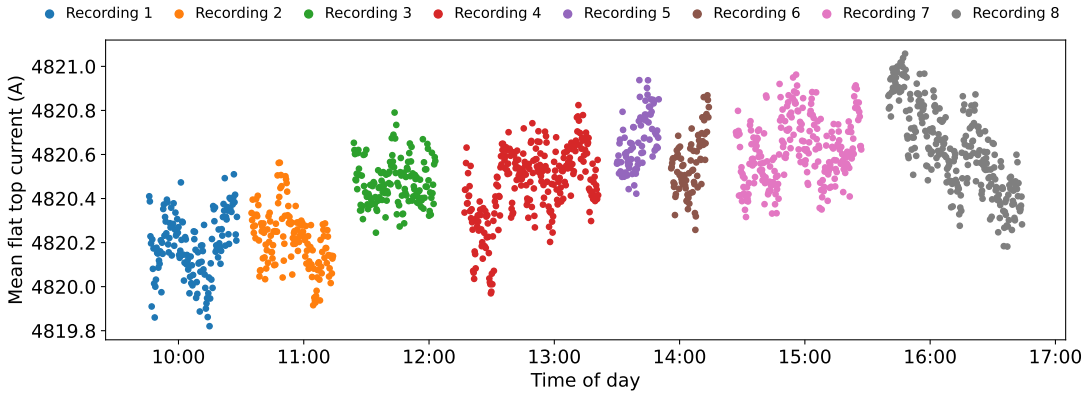


Fig. 6.8: Average current of the 1.2 s flat top of SHIP cycles, as measured by the DCCT, as an extended version of the waveform shown in Figure 6.1, each 1 h long, is repeated over 7 h to capture the drift of the current over time. Each color group shows the average flat top current within its 1 h measurement. The average current can be seen to drift from its lowest at 4819.8 A to its maximum 4821.1 A and then starting to decrease around 15:30.

The current-reproducibility limitations examined above are compounded by a hardware limit that prevents the operationally most significant sequence from being played at all. The magnet bench reaches a peak current of 6000 A and a maximum ramp rate of 3000 A/s. With a second dipole in series disconnected, a downramp of 4000 A/s corresponding to the maximum ramp rate of operational cycles is possible. Therefore, neither the peak current nor the ramp rate of the operational cycles is a limiting factor for powering the bench dipole. The remaining limitation is the RMS current of the power converter, which is limited to 3000 A and trips the converter once exceeded.

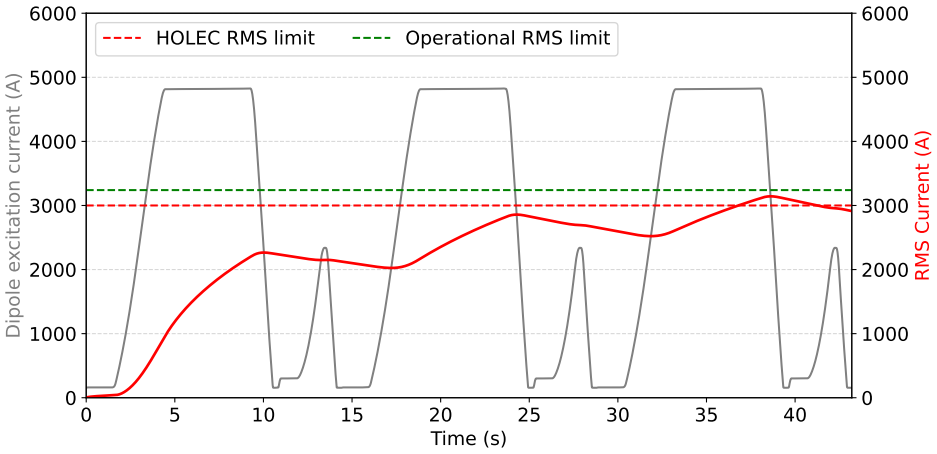


Fig. 6.9: Illustration of the rise of the time-weighted RMS current of the power converter during a standard SFTPRO physics supercycle, overlaid with the programmed current of the sequence. The RMS current can be seen exceeding the 3000 A limit of the power converter during the third SFTPRO cycle at the flat top, which, during measurement, immediately trips the power converter.

$$I_{\mathrm{RMS,static}} = \sqrt{\frac{1}{T}\int_0^T I(t)^2 dt} \tag{6.1}$$

$$I_{\mathrm{RMS,weighted}}(t) = \sqrt{\frac{1}{\tau}\int_0^\infty I(t-t')^2 e^{-t'/\tau}\,\mathrm{d}t'} \tag{6.2}$$

Two definitions of the RMS current are relevant. The bench converter protects on the time-weighted RMS of Equation 6.2, tracked continuously and tripping the converter as soon as it exceeds 3000 A. The SPS converter, on the other hand, protects on the windowed RMS of Equation 6.1, which in operation is evaluated by the operators over a fixed window per supercycle against a soft limit of 3240 A. In operation, however, the RMS limit can be exceeded briefly without tripping the mains. In the lab, the weighting time constant τ was obtained as an effective value of 31.1 s by fitting the decay of the converter RMS readout. Figure 6.9 reconstructs the time-weighted RMS rise from the programmed current of a standard SFTPRO supercycle using this value.

The reconstructed RMS crosses the 3000 A bench limit at the flat top of the third SFTPRO cycle and trips the converter at every measurement of this sequence. The 5 s long SFTPRO flat top sustains the current just over 4800 A, so the RMS accumulates across successive cycles faster than it decays between them. While most operational sequences have a lower RMS current than the bench can handle, the fixed-target production sequences take up most of the SPS physics program, and the sequence changes that motivate this work involve them. In practice, the standard SFTPRO supercycle reaches a windowed RMS of 3241 A, and the combined SFTPRO and SHIP supercycle proposed for SHIP operation in Figure 4.10 reaches 3240 A, both designed to operate at the ceiling of the SPS ring converter. Resolving the change in hysteresis across cycles and across supercycle transitions requires several such supercycles played in succession, yet the bench trips before even a single SFTPRO supercycle completes, so these sequences cannot be measured at all.

For this reason, the operational-sequence measurements of Section 6.1.1 used the SHIP cycle, which shares the extraction energy of SFTPRO on a shorter 7.2 s cycle with a shorter flat top, chained with a 3.6 s idle between cycles to keep the time-weighted RMS below the bench limit. This measured the magnetic response for hysteresis to evaluate the measurement fidelity of the bench, but did not reproduce the operational supercycle one-to-one.

Redesigning the measurements by inserting idle time between the high-current cycles would keep the RMS below the limit but would no longer reproduce the operational sequences. Static staircase measurements hold the current at fixed levels for more reproducible readings, at the cost of much longer acquisition for high-resolution hysteresis. They are infeasible above the RMS limit, as all standard physics cycles reach flat-top currents above 3000 A and up to 5800 A. This therefore also rules out alternative measurement methods such as stretched-wire and rotating coil techniques, even for semi-static measurements like a staircase for the lab dipole.

6.1.4 Limitations and Data Transferability

None of the laboratory measurements pass the pulsed measurement requirements set out in Section 5.4.5. Ultimately, the difficulty of obtaining absolute and reproducible field measurements with the fluxmeter, its only partial mitigation with the Hall probes, and the reproducibility and RMS consequences of the power converter disqualify the laboratory measurements of the dipole for producing data for fitting the field model.

The arbitrary excitation sequences that the laboratory measurements were intended to provide, required in Section 5.4.5 to sample excitation histories beyond those played in operation, were not realized. The operational sequences that serve as the baseline already fail on reproducibility, cannot proceed beyond the RMS limit, and arbitrary high-current sequences are more demanding on both. Intervals of constant current for the fluxmeter drift correction could be inserted into a pseudo-random waveform, but this constrains the excitation toward the operational case and does not remove the reproducibility limit that already disqualifies the baseline.

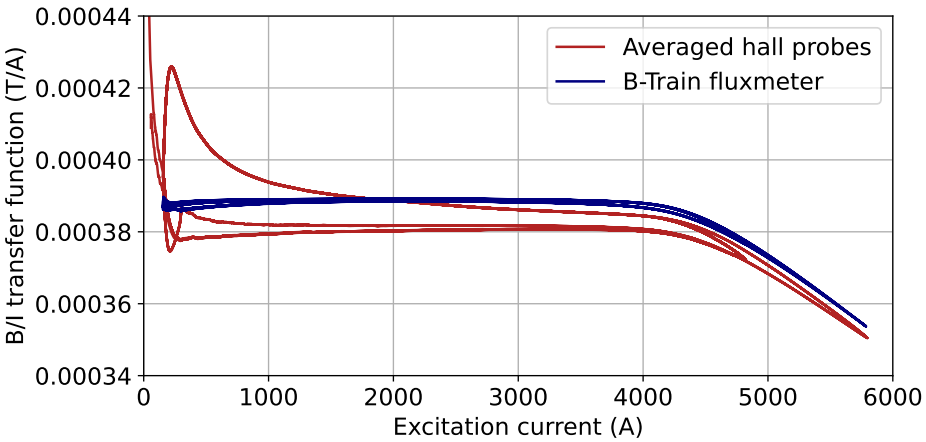


Fig. 6.10: Transfer function $I \to B/I$ of the averaged laboratory Hall probes from this section, and B-Train measurements of a similar supercycle from operation with SHIP and LHCPILOT cycles. The two differ by a constant offset of about 2 % and in the shape of the hysteresis loop, the latter combining the different hysteresis and dynamics of the two systems with a contribution from Hall sensor drift.

Even if these sequences could be measured, the laboratory and the B-train constitute different data distributions, and the shift between them is significant for a machine learning model of the field. The laboratory measures a single spare MBB dipole, whereas the field measured by the B-train is produced by both the MBA and MBB dipoles of the ring, which differ in geometry and therefore in their field response, and a single-magnet measurement does not capture the combined response of the ring. The eddy-current decay is also observed to be faster in the laboratory than in the SPS, settling within tens of milliseconds against several hundred milliseconds, as seen by comparing Figure 6.4b and Figure 5.10. This difference alters the field on the ramps and at the onset of the plateaus, and is developed further in Chapter 9. Figure 6.10 compares the transfer function $I \rightarrow B/I$ of the averaged Hall probes in the laboratory with that of the B-train fluxmeter, showing a constant offset of about 2 % and a different loop shape. Only the offset is removed by recalibration, and the loop-shape difference combines the different hysteresis, the different dynamics, and a contribution from Hall sensor drift, which is itself significant as the Hall probe is the most reliable instrument available on this measurement bench for pulsed measurements. A model trained on laboratory data would therefore not generalize to the B-train measurements from which it is intended to be seeded, so the laboratory measurements could not trivially augment the operational dataset. The B-train measurements described in the following section are used instead.

6.2 B-Train measurements

The B-train system, outlined in Sections 3.4.3 and 4.5.2, measures the dipole field of two reference magnets powered in series with the SPS main dipole circuit, and is intended to represent the field generated by the main dipoles of the ring. It measures the field online during operation, on the cycles as they are actually played.

The excitation current, as driven by the FGC of Section 4.4.6, is designed with higher requirements than the laboratory power converter, and overcomes all the limitations encountered in the laboratory measurements. In Figure 6.11, 95 current measurements of the same cycle, extracted from the logging system, with identical programmed current are shown with the mean and standard deviation of the difference between the played and programmed current. It can be seen that the current is reproducible between measurements, with a standard deviation below 0.02 A on the constant and slowly ramping parts of the cycle, and the main power converters are specified to be ppm-reproducible. Large ripples reaching 1 A appear at the ramp transitions where the second time derivative of the current is high, but these occur where the beam has already been extracted, and are reproducible between measurements. Furthermore, as the ripples are reproducible, they are not expected to be corrected by the model proposed in this thesis. Consequently, neither the measurement of the current nor the reproducibility of the current are limiting factors for using online measurements as model training data.

The main limitation of B-train measurements remains the same as in the laboratory: fluxmeter integration drift. While the B-train has an NMR probe for cycle-by-cycle absolute field marking, it still relies on the integrated fluxmeter for the full cycle, until the next NMR mark. The marking is only performed at the start of each cycle, when the excitation crosses 1100 G. Furthermore, the voltage offset of the B-train system is not re-estimated for each cycle, but rather relies on periodic re-estimation. In practice, this means a recalculation of the offset every few hours, only in the presence of a ZERO cycle in the SPS supercycle, and cannot be triggered by software. The only way to force recalibration is by manually shorting a cable inside the server rack to make every cycle being played appear as a ZERO-cycle and therefore enable the recalibration routine to run. In addition to the tracked voltage offset, there is a second voltage offset in the B-train

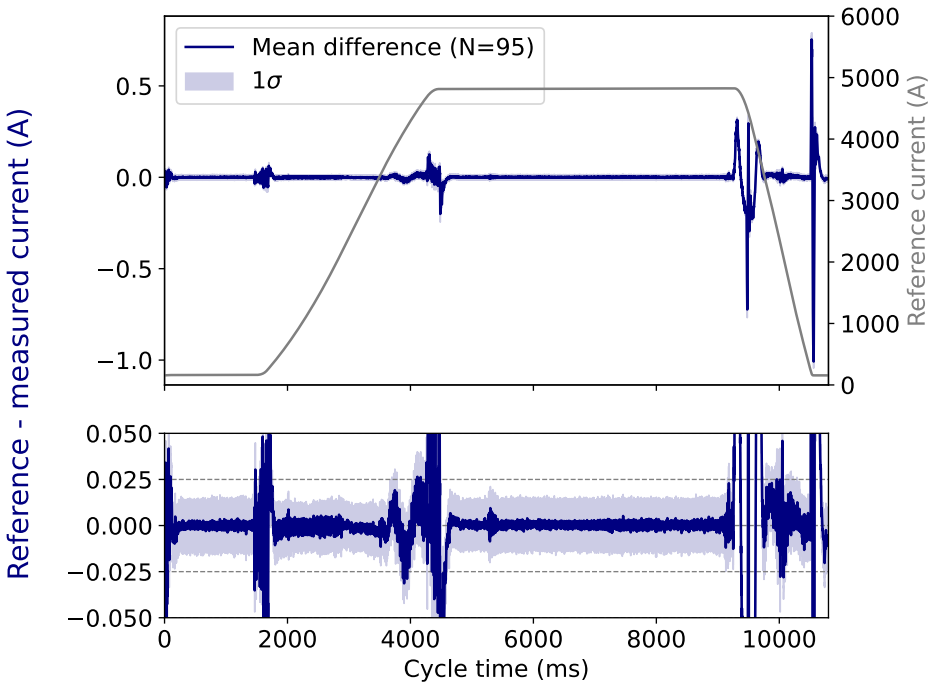


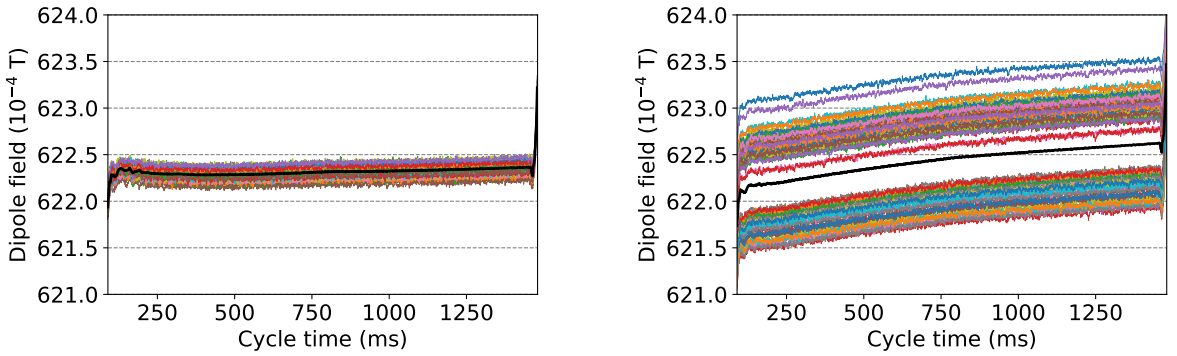
Fig. 6.11: Difference between programmed (reference) and measured main dipole circuit excitation current, as measured by the FGC DCCT, on 95 consecutive SFTPRO cycle occurrences during operation. The plot shows the mean and standard deviation of the difference across the 95 occurrences, with the mean shown in blue and the standard deviation in shaded blue. Large, and on average coherent, ripples can be seen at the start and end of each ramp, with the largest ripples reaching 1 A in amplitude with respect to the reference current, during the beam-out period of the cycle. The lower panel shows the same plot at a smaller scale, where the mean is close to zero, and the standard deviation is below 0.02 A during constant or slow-ramping areas of the cycle.

server rack, which needs manual recalibration, and is normally not done. In consequence, the B-train drifts differently between re-estimations, and the resulting field measurements can differ when taken hours or days apart. The next sections outline the limitations, accuracy bounds and signal processing of the B-Train data for use in the field modeling proposed.

6.2.1 Field Drift and Reproducibility

The laboratory section characterized measurement drift and reproducibility using the SHIP cycle as stand-in for the SFTPRO cycle due to RMS current limitations and need for constant current sections for voltage offset estimation. On operational data, the latter is run continuously over the annual proton physics runs and can be extracted directly from NXCALS. Furthermore, the SFTPRO cycle remains the primary target for measurement, modeling, and correction in this thesis, so its reproducibility is examined here, beginning at the injection plateau. On this cycle, the integration marker is only applied on the ramp after injection, since the 14 GeV/c injection field at around 622 G lies below the marker level, shown in Figure 4.16. The last marker applied prior to the SFTPRO injection is therefore the one applied at the ramp-up of the previous cycle, which in operation is guaranteed to be an MD1 cycle, and therefore to be within 3.6 s of the

injection plateau. Due to operational beam reproducibility requirements, it is assumed that the beam radial position appears approximately stable between cycle occurrences at injection, and it is thus assumed that similarly the injection field is also stable between occurrences, also counting eddy current contributions from the previous cycle. Any field errors as measured at injection are then attributed to the B-train measurement drift and noise, and the spread of the measured field across occurrences is used to quantify the reproducibility of the B-train measurement at injection. The MD1 cycle always precedes the SFTPRO cycle, and the injection field is held constant during operation, so the eddy-current contribution at the injection plateau repeats between occurrences.



(a) Injection flat bottom of the cycle on one measurement day. The field is approximately flat across the plateau, with a drift-fan spanning 3×10^{-5} T and a standard deviation below 5×10^{-6} T over the 125 cycle occurrences.

(b) Injection flat bottom of the cycle on a different measurement day. Each measurement ramps slowly over the plateau, with a larger drift-fan of 1.5×10^{-4} T and a standard deviation of 6.3×10^{-5} T over 99 cycle occurrences.

Fig. 6.12: Two different measurement sets of the same SFTPRO cycles on different days, showing the variation of the flat-bottom field on the 1.4 s injection plateau. The two sets differ in the cycle preceding injection: an MD1 cycle in the first and a 3.6 s idle period during machine development in the second, so the last marker before injection falls within 3.6 s in the first case and a further cycle back in the second. The mean of the measurements is shown in black, and cycle occurrence is plotted individually.

Figure 6.12 shows the field at the injection flat bottom for two sets of SFTPRO cycles recorded on different days, and the second measurement with a 3.6 s idle period instead of an MD1 during controlled machine development. On one day, in Figure 6.12a, the field is flat across the plateau and the occurrences fan over 3×10^{-5} T, with a standard deviation across the 125 injection plateau measurements below 5×10^{-6} T. Contrary to the SHIP cycles measured in the lab, the SFTPRO injection does not have a constant current, but is ramping 0.09 A over the 1260 ms injection window, to compensate for the eddy-current decay of the previous cycle, which will be further examined in Chapter 9. Therefore, the as-measured approximately constant field across injection is expected. This first set is taken as the reproducibility floor of the B-train. The injection plateau is not marked on its own cycle, so its nearest anchor is the marker on the preceding MD1 cycle, already 3.6 s earlier, yet the measured drift over the plateau stays below 5×10^{-6} T. Measured against the acceptance criteria of Section 5.4.5, the standard deviation of 5×10^{-6} T sits below the measurement reproducibility target of $\mathcal{O}(1 \times 10^{-5}\,\mathrm{T})$, while the 3×10^{-5} T peak spread across the 125 occurrences reaches the beam reproducibility tolerance derived in Section 5.2.2. Both figures describe the variability of the measured field between repeated occurrences of the same cycle, the standard deviation as the spread of the distribution

and the peak spread as its worst-case range. The beam tolerance is a peak bound derived from a maximum orbit displacement, so the peak spread is the measure compared against it, and the standard deviation is reported to describe the distribution. At injection, the peak spread reaches the tolerance while the standard deviation stays well below the target, so the two metrics give different verdicts and the peak bound is the binding one. Cycles that are instead marked before injection, such as the LHC filling cycle, are anchored differently and accumulate drift downstream of the marker, as examined later in this section.

On the second measurement set shown in Figure 6.12b, the same cycle fans over 1.5×10^{-4} T with a standard deviation of 6.3×10^{-5} T, and each occurrence follows the programmed current ramp over the plateau in the absence of the preceding MD1 pulse. The programmed excitation of the SFTPRO cycle is identical to the first set, so the difference between the two sets originates in the different preceding cycle and any drift accumulated. Replacing the MD1 with the idle period changes three conditions at the injection plateau at once: the hysteresis branch on which the magnet sits, the eddy-current state decaying into the plateau, and the marker anchoring, since the idle cycle carries no marker and the nearest anchor is pushed a further cycle back in time. The enlarged spread is therefore a combination of these changed preceding conditions and measurement drift, and the two cannot be separated from this figure alone. The drift contribution is revisited in the following subsection once the drift correction is introduced.

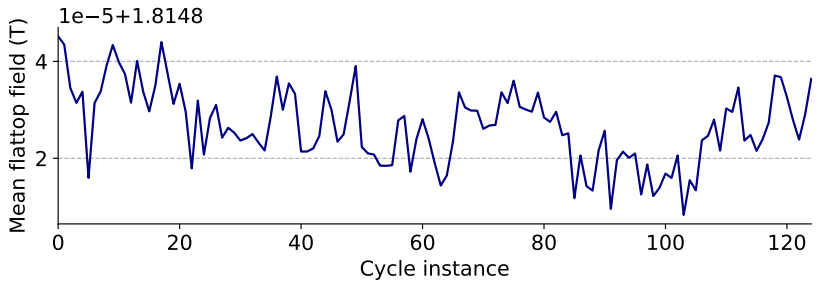


Fig. 6.13: Mean flattop field as measured by the B-Train on 125 SFTPRO cycles after reaching steady state. A variation in the 4×10^{-5} T range can be seen. The programmed field and current do not change over this period, and beam parameters are constant; the variation of the measured field is attributed to accumulated integration drift.

A further study on the drift of the B-train measurements was performed at the flat top of the SFTPRO cycle, which starts 2800 ms after the marker and lasts 4800 ms. For the same measurements as the injection flat bottom is shown in Figure 6.12a, the mean field at the flat top was computed across 4800 samples of the extraction plateau, over 125 consecutive cycles, or 30 min during standard physics. In this case, it was determined that the flat top had reached a steady hysteresis state, with no change in the excitation sequence during the preceding 45 min. The hysteresis-driven relaxation seen in Figure 5.4 settles within 15 min, so the flat-top field is expected to be constant due to the same operational requirements as before, and any variation is attributed to accumulated drift. Figure 6.13 shows this mean flat-top field varying by up to 4×10^{-5} T over the 125 occurrences. Because the injection flat bottom of the same run reproduces to below 5×10^{-6} T, this 4×10^{-5} T spread is close to the best achievable at the flat top. It exceeds both the $\mathcal{O}(1 \times 10^{-5}\,\mathrm{T})$ measurement target and the 3×10^{-5} T beam tolerance, so the flat top is less reproducible than the well-anchored injection plateau. This drift nonetheless remains well below the flat-top field change of 3.5×10^{-4} T produced by the supercycle change in Figure 5.4, which the field model is required to resolve.

6.2.2 Marker-Based Drift Correction

To improve reproducibility of B-Train measurements for the sensitive SFTPRO injection plateau, a per-cycle correction was devised to remove the drift up until the marker. As opposed to the lab, the measured fluxmeter voltage is integrated with FDI internally in the B-train at the native sampling rate, and not made available outside the server rack. Additionally, the NMR marker is applied directly to the integrated field at the native sampling rate, so when the signal is decimated to the 1 kHz logging rate, the marker is effectively applied between two samples. This makes standard voltage-offset-based drift correction impossible, as neither the raw voltage nor a cycle-by-cycle voltage offset is available, and the marker is not applied to every acquisition due to the 1100 G threshold. Furthermore, while a voltage offset is kept in memory and applied prior to integration, it is only re-estimated at most every few hours, and only when a ZERO cycle is played, which in regular operation only occurs during the LHC filling sequence seen in Sequence 4.3.1. Consequently, a different approach was taken by using the field jump at the marker to estimate the drift gap and an effective correction.

The field jump can be estimated by comparing the field change at the marker against the expected field change from the known excitation current, which is reproducible between occurrences. This can be expressed as

$$\Delta B_{\mathrm{drift},k} = (B_k - B_{k-1}) - \big(f(I_k) - f(I_{k-1})\big) \tag{6.3}$$

where $f(I)$ is the known transfer function of the dipole in the linear region, and k is the index of where the marker was applied, which can be found by the time of the field crossing of 1100 G in the measurement, or extracted from logging. By then calculating a linear slope since the time of the previous marker being applied in the preceding cycle, a small linear drift correction can be applied.

Under the linear-drift approximation, this deviation is distributed as a linear ramp over the cycle, which recursively anchors the field to the marker and removes the discontinuity at the marker. The scheme amounts to a first-order Kalman correction, with the marker as the single measurement update per cycle and a linear drift model between markers.

The drift accumulated at the marker combines a linear and a nonlinear component that cannot be separated from the single marker sample. A correction referenced back to the marker of the previous cycle would span the full previous cycle and impose the linear assumption over a region where the drift is partly nonlinear, which can introduce more error than it removes. The correction is therefore restricted to the current cycle, from cycle start to the marker, which covers the fixed-target injection plateau, and leaves the ramp and flat top of the previous cycle uncorrected. This relocates the field discontinuity from the marker to the cycle boundary. The sequence-to-sequence model proposed in Section 5.4.5 operates on long, continuous measurements, so this discrete step is a genuine cost, and it is placed at the cycle boundary where it least affects cycle-by-cycle prediction. The marker gap additionally measures the accumulated drift of each cycle, so cycles and sequences that drift severely can be identified and excluded from the training set, and the model is not exposed to measurements the correction cannot recover. The first cycle of a dataset has no preceding marker and is left uncorrected.

Applying the correction to the measurement set of Figure 6.12b, the large variation in measured injection field goes from the 1.5×10^{-4} T drift-fan to close to the reproducibility floor of the well-anchored set, shown in Figure 6.14. The enlarged spread of that set can therefore be attributed fully to integration drift, and the correction recovers operational reproducibility even in the absence of the preceding MD1 cycle. Different variation in slope is argued to be related to different cycles preceding the 3.6 s idle period, and the clear difference of 8×10^{-5} T in field

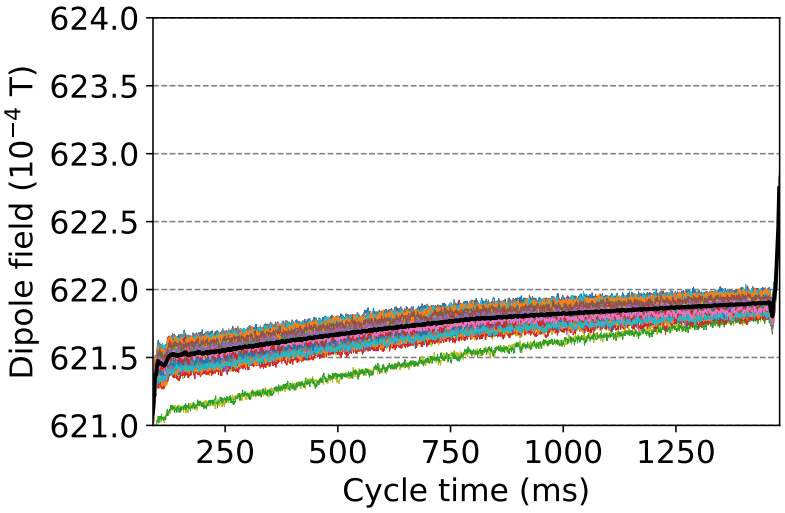


Fig. 6.14: Drift-corrected B-Train measurements, originally shown in Figure 6.12b, where the large variation of the flat-bottom field is collapsed to the same best-case range seen in Figure 6.12a. Each cycle has a different preceding cycle before the 3.6 s idle period, which accounts for the residual variation in slope between occurrences after correction.

at injection compared to a standard cycle is attributed to the different hysteresis branch. For the SFTPRO cycle, the corrected injection field reproduces to a 3×10^{-5} T floor, similar to the best-case measurements, around the beam tolerance and above the $\mathcal{O}(1 \times 10^{-5}\,\mathrm{T})$ measurement target, and an order of magnitude better than the best-case averaged laboratory fluxmeter at the injection plateau. The correction relocates the drift discontinuity from inside the cycle to the cycle boundary, and the field continues to drift away from the marker. The preceding cycle also remains uncorrected, so the recovered reproducibility is local to the corrected cycle and is not a universal removal of drift. The measurement target is therefore not reached, yet the floor sits an order of magnitude below the 5×10^{-4} T cycle-to-cycle hysteresis the model must resolve, so the dominant effect remains measurable. The consequences of this residual floor for model accuracy are developed in Chapter 11.

Limitations remain after correction. The correction removes only the linear component of the drift and leaves a nonlinear residual within the cycle, and it does not disentangle the eddy-current and drift contributions that combine at the plateaus. The correction is also anchored at the marker, so the residual drift grows with distance from it. On cycles whose plateau lies far downstream of the marker, the accumulated drift over the span is large and is neither linear nor reproducible, so a substantial residual remains after correction, which is examined in the following subsection.

6.2.3 Residual Drift on Long Cycles

The integration marker only guarantees an exact absolute field at the marker, and the constructed drift correction only acts up to the marker, so the field is not re-anchored beyond it and drifts freely. On cycles whose measurement extends far beyond the marker, this uncorrected drift accumulates over a long span. The clearest case is the LHC1 cycle, the main proton filling cycle for the LHC. It carries the strictest injection reproducibility requirement, set by the LHC filling

efficiency and the 4 injections over the 12 s, all requiring the same energy matching at injection. Its injection field can therefore not carry cycle-to-cycle hysteresis or eddy current variation in operation, so the spread across occurrences isolates the measurement drift. It is marked before injection, so the full injection plateau lies downstream of the anchor and is never re-anchored, as shown in Figure 6.15.

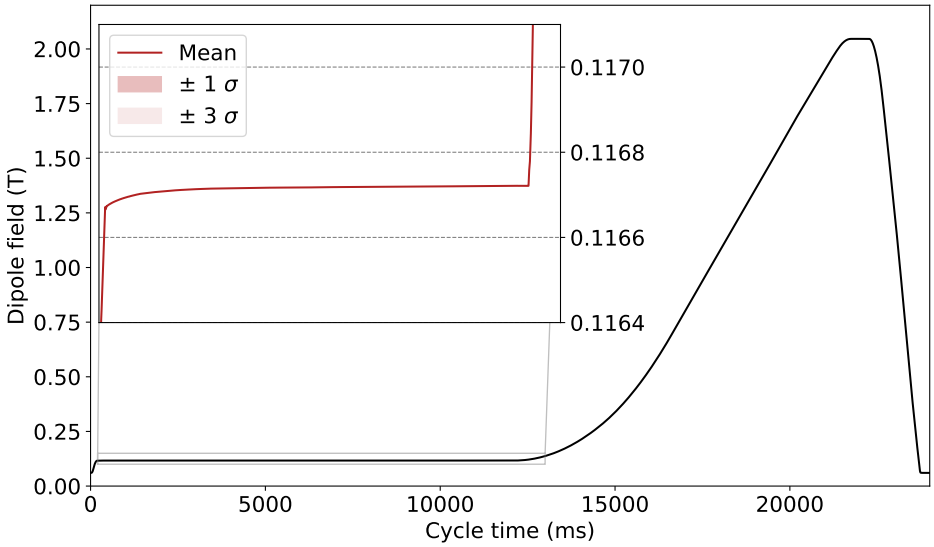


Fig. 6.15: Dipole field over the full LHC1 filling cycle, with the inset showing the injection-plateau statistics across all occurrences during the 2025 proton run ($N = 23562$), including any use by machine development. The mean is shown in dark red, the $\pm 1\sigma$ band in medium shading and the $\pm 3\sigma$ band in light shading. The marker is applied before injection, so the plateau lies downstream of the anchor. The bands widen from the start of the plateau towards its end as the field drifts away from the anchor.

Because the plateau lies entirely downstream of the marker, the correction does not act on it, and the drift accumulates uncorrected, widening the occurrence spread with distance from the anchor. By the end of the 12 s plateau this spread reaches $\pm 1\sigma$ of about 1×10^{-4} T and $\pm 3\sigma$ of a few 1×10^{-4} T. The field contribution of the programmed current, which is shifted vertically but not in trend between occurrences, is removed before these statistics by subtracting the anhysteretic field of the measured current, so the shown spread reflects drift independent of the excitation differences between occurrences. The spread also diverges at different rates between occurrences, so this drift is not a reproducible function of cycle time or of the model inputs. The worst-case spread reaches a substantial fraction of the 5×10^{-4} T cycle-to-cycle hysteresis that the model is required to resolve, so on the longest injections the measurement drift is comparable to the effect being modeled. This spread is well above both the SFTPRO best case and the acceptance criteria of Section 5.4.5, so long injection cycles are not measured to the required reproducibility. The drift therefore limits the reproducibility achievable on these cycles, and usable data for them is restricted to the occurrences with lower drift. It grows further on longer injections still, such as the 48 s LHC Ion cycle filling plateau, which drifts by several 1×10^{-4} T and is excluded from the initial field models. The consequences for the field model, including the effect on training and on the seeded initial state, are developed in Chapter 11.

6.2.4 Data Acquisition and Limitations

The B-train is an operational logging system and its published data is not curated for modeling, so it cannot be extracted in bulk and each dataset is inspected as it is retrieved. The inspection is necessary because the logged field contains failure modes that render individual cycles, and even entire sequences, unusable for modeling. Some cycles publish an all-zero or constant field, others carry no applied marker, so the field is never re-anchored and drift accumulates uncorrected over long sequences, and others carry a marker that is applied incorrectly. The logged data therefore cannot be taken as reliable without inspection, so gathering usable measurements is an active process and data availability is a limiting factor.

During an earlier period of operation, the marker was applied with 1716 G instead of the hardware-configured 1100 G, due to a mismatch between the hardware setting and the LSA value. The measured field then steps 616 G at the marker of each cycle affected and steps back at the marker of the next, which has the correct setting, and because the marker is resolved only once per cycle, this step could not be corrected easily. Logged availability therefore does not guarantee usable data, which limits the mining of historical records for rare sequences.

Measurements of the same cycle can also be misaligned in time between occurrences by less than one sample, arising from the internal downsampling from the native acquisition frequency to the 1 kHz published rate. The misalignment is visible only on the ramps, where the high field rate turns a sub-sample time shift into a measurable difference between occurrences, and is negligible on the flat regions. Correcting it would require aligning each measurement to a fixed per-cycle reference by its field derivative, which is not practical, so the misalignment is accepted and is in any case dominated by the drift.

The main limitation of the B-train data source outside of instrumental error is the limited variety of excitation that can be measured, which arises in two ways. First, individual cycles cannot be freely designed, since the SPS plays only approved cycles, so arbitrary current waveforms cannot be measured on the ring. Second, the arrangement of approved cycles into sequences is constrained in practice, since SPS supercycle construction is manual and coupled to the upstream machines that must populate their own cycle slots, so rapid or arbitrary sequence changes are impractical. Some varied conditions are nonetheless reachable in dedicated sessions, such as SFTPRO played without its preceding precycle, with different preceding cycles, or more asymmetric supercycles. The arbitrary-sequence coverage required in Section 5.4.5 is therefore not achieved, and the usable data is bounded to the operational and MD excitation histories.

Against the acceptance criteria of Section 5.4.5, the B-train provides an absolute measurement through the NMR marker and meets the required rate. It does not reach the $\mathcal{O}(1 \times 10^{-5}\,\mathrm{T})$ reproducibility target, with a reproducibility ranging from a few $1 \times 10^{-5}\,\mathrm{T}$ on well-anchored short cycles to a few $1 \times 10^{-4}\,\mathrm{T}$ on long injections, and it does not achieve arbitrary-sequence coverage. It is nonetheless adopted as the field data source for this work, since it is the only source that measures the operational MBA and MBB average directly and its reproducibility is sufficient to resolve the dominant sequence-change effect of $1 \times 10^{-4}\,\mathrm{T}$ to $3 \times 10^{-4}\,\mathrm{T}$ that the correction targets. The residual drift, the bounded excitation coverage, and the constrained availability of usable data set the limits within which the field model of the following chapters operates.

Before using the data in the field model, we condition it at the dataset level. The field is drift-corrected as in Section 6.2.2, then notch-filtered at 50 Hz and the power-converter harmonics, and low-pass filtered. The mains interference is already small after upstream filtering, and the notch filters remove the residual identified in the measured spectra, which would otherwise alias to a constant offset under the later decimation. The low-pass is a first-order filter at 60 Hz, chosen to remove high-frequency noise while preserving fidelity at fast changes such as the ramps,

where a higher-order filter would introduce ringing. Decimation to the model input rate is left to the data-loading pipeline in the following chapter, so different downsampling rates can be tried, and the reproducibility figures reported above are computed at the native rate and unaffected by it.

Chapter 7

Modeling Hysteresis

7.1 Modeling Requirements

Chapter 5 established the modeling of the operator $\mathcal{F}$ as a data-driven, sequence-to-sequence task, in which the field realized over a cycle is predicted from the past excitation and field together with the excitation of the coming cycle. The model is intended as a generic field model of an SPS magnet, valid at an arbitrary point in an arbitrary excitation sequence, and it is applied at inference to predict one operational cycle at a time, since the aim is feed-forward compensation of individual cycles as proposed in Section 5.4.6. This generality sets the requirements developed here, which constrain the data and the model architecture together.

Because the model is not tied to any single cycle, it is fitted on windows of arbitrary position and length drawn from continuous measurement sequences, so the machine cycle is not used as a unit during training, while prediction in production is carried out one cycle at a time. By construction the model therefore predicts absolute field for use with the proposed feedforward compensation scheme, with the choice of target and its justification given in Section 7.2.3.

The current and field are the only measured quantities available and necessary for magnet characterization, and the model must initialize from an arbitrary magnetization, seeded from a finite measurable window of the past current and field, without a canonical demagnetized or reset state. Its prediction must span the whole cycle as a time series and cannot reduce to a static field value at each setpoint, since eddy currents make the field evolve in time while the current is constant. The model must therefore be time-dependent, whether implicitly or explicitly, since neither a static current-to-field map nor a rate-independent model has a mechanism for these dynamics. Furthermore, the output length must be changeable so that cycles of different duration can be predicted, while a changeable context length is convenient without being essential. The model must condition its prediction on the future current to predict the corresponding field evolution. Finally, the prediction over a window must feed back as the context for the next window, so the model runs autoregressively when no measured field is available to re-seed it, as required for the open-loop operation of Section 5.4.6. This last requirement is met by predicting field from past field, since the output then occupies the same space as the seeding input and no change to the model is needed.

The data requirements that follow from this scope were established in Section 5.4.5, and their intended realization through a dedicated laboratory measurement campaign with arbitrary excitation could not be carried out within this work. Training and validation therefore use the operational B-train field, adopted in Section 6.2.4 as the only available direct measurement of the realized field and the signal that seeds the model in operation, within the reproducibility and excitation-coverage limits established there.

The adopted B-train field serves as the training label, so the modeling problem is supervised. Its noise and drift bound the achievable accuracy, and a self-supervised objective would neither remove that bound nor avoid the field as a decoding target, so supervised learning is adopted as a first approach in this work.

7.2 Dataset and Target

The dataset assembled for the main dipole circuit is drawn from the adopted B-train field, whose provenance and quality are established in Chapter 6. Only the main dipole family is treated here, since it is measured directly by the B-train and is the primary compensation target of this work.

7.2.1 Dataset Composition

The dataset for the main dipole circuit is assembled as collections of uninterrupted measurement sets logged from the main dipole circuit FGC for the current, and B-train for the field measurement. Each set contains at least one supercycle transition, with a start sequence before the transition and an end sequence after it, so that both the field change and its relaxation following the change in excitation are captured. Sets were intentionally selected where the integration drift of Section 6.2.2 is small, so that the field variation attributable to hysteresis and eddy currents is not masked by measurement drift.

More than one set was acquired for each supercycle transition seen in operation, following the transitions that operations actually play among the supercycles of Section 4.3, including all transitions between the SFTPRO, HiRadMat, AWAKE, LHC sequences that occur in physics operations. Machine development sequences and non-standard sequences found are also included, together with ten extended sets measured over varied cycle sequences during dedicated machine development as part of this work. These extended sets contain supercycle versions where the MD1 precycle is kept at idle current or omitted entirely, a configuration whose feasibility is one of the goals motivated in Section 5.1. Ion run SFTION cycles and shorter LHC ion cycles are also present, with the longer ones omitted due to large integration drift over the long duration.

Although these exotic cases are represented, the dataset is skewed toward operational sequences. The SFTPRO cycle is strongly overrepresented, appearing in the large majority of sets, since it is present in every supercycle and many transitions start or end in a sequence where it repeats. Dynamic economy sequences of the same SFT cycle and of some LHC cycles are included, and full economy is excluded because an integrator drifting over a full cycle length yields linear drifting noise without a recoverable signal.

The measured excitation current is logged from the FGCs over the range 155 A to 5800 A, and the measured main field over 0.05 T to 2 T. The programmed current is available from logging for the measured cycles, and it is absent under dynamic economy, where the realized current is altered asynchronously by the power converter economy modes of Section 4.4.6.3. Measured current therefore has to be used for fitting, and since the alignment with the programmed current has been shown to be close in Figure 6.11, this choice introduces no significant discrepancy. Furthermore, this allows the programmed current to be used at inference.

7.2.2 Signal Conditioning and Sampling

Prediction of the dynamic effects, in particular the eddy current decay of Section 2.6.2, requires a sample spacing fine enough to resolve the transient and to avoid the jaggedness that would make a feed-forward correction unusable. A downsampling factor of 20 is selected, reducing the native 1 kHz logging rate of the current and field measurements to 50 Hz, for a sample spacing of 20 ms. Eddy current decay in the B-train relaxes over the few hundred milliseconds reported in Section 6.1.4, a few-hertz bandwidth well within the 25 Hz Nyquist limit of the 50 Hz rate, so the dynamic content is resolved with conservative margin.

Uniform downsampling is used even though the cycles contain long stationary sections such as the injection plateaus. An adaptive scheme keyed to the current would collapse a plateau to its endpoints and erase the eddy current decay that evolves across it, and keying to the field is not

possible at inference, where the field over the coming cycle is unknown. Uniform spacing keeps interior samples on every section and is identical between training and inference.

Preprocessing conditions the full 1 kHz signal independently of the model used downstream. The ramp edges and fast transients of the response must be preserved in the conditioned signal, and a stricter anti-aliasing stage before downsampling, with a lower cutoff and a higher filter order, smooths these edges and adds slow ripple to the field response. A first-order low-pass filter with an empirically set 60 Hz cutoff is therefore applied to the field and current at the native rate, with notch filters at the 50 Hz and 100 Hz mains harmonics. This 60 Hz cutoff sits above the 25 Hz Nyquist limit of the target rate, so the ramp edges are preserved at the cost of mild aliasing, while the notches remove the mains lines before decimation so that they do not fold into the resolved band. Since the content above the few-hertz band of interest is measurement noise and not signal, the aliasing introduced on downsampling remains at the noise floor and does not affect the resolved dynamics. The filtered current is then differentiated to obtain its derivative $\dot{I}$ at the same sampling rate, and the field is partially corrected for integration drift following Section 6.2.2. The downsampling itself is executed at data loading and is a parameter of the training pipeline, and uses the already preprocessed dataset.

7.2.3 Prediction Task and Residual Target

The absolute field spans roughly 2 T, and the reproducibility criterion of Equation 5.2 requires an accuracy of order 2×10^{-5} T, five orders of magnitude smaller. A model predicting the absolute field would therefore have to resolve a signal far below the field range, which is challenging for any function approximator. At full field scale the hysteresis is visually indistinguishable from the current-to-field calibration curve, as seen in Figure 5.7, so the differences relevant to reproducibility are not discernible against the absolute field.

The predicted quantity is therefore taken as the field residual left after subtracting a calibration curve fitted to the measured field, a fixed current-to-field mapping added back to the prediction to recover the full field, in the same way a known polynomial trend is removed from a signal before fitting and restored afterwards. This curve is taken as the midpoint between the magnetization and demagnetization branches of the LHC cycle, which reaches the highest field, so it sits at the center of the hysteresis, analogous to the control-system transfer function of Figure 4.12b and fitted here to the measured field. Written at step i as $r_i = B_i - B_{\mathrm{cal}}(I_i)$, the residual is of order ± 3 mT, a reduction close to a factor of 700 from the 2 T span, at which scale the sub-percent accuracy required by Equation 5.2 is meaningful. Its structure is shown against the current in Figure 7.1, where the injection inset also contrasts a precycled and a non-precycled SFTPRO cycle, the difference the compensation aims to remove.

Even after this subtraction, the residual range is dominated by the major hysteresis loops, whose opening reaches about 6 mT. The cycle-to-cycle field change across a supercycle transition, which is the operationally relevant quantity, is of order 4×10^{-4} T–5×10^{-4} T, eight percent of the 6 mT loop opening, and the 2×10^{-5} T reproducibility tolerance is a few tenths of a percent of that opening. Capturing cycle-to-cycle differences therefore amounts to resolving a small modulation of the major loops, while most of the residual numerical range is taken by loop structure common across the cycles.

No further preprocessing is however applied to isolate this modulation. The calibration curve is a fixed, deterministic function of the current, so subtracting it adds no error to the residual. Any further reduction, whether removing the major loops or rescaling the residual, would need a model or transform that either injects its own error into the small-signal target or distorts a signal spanning the whole loop, so the major loops are left in the residual for the sequence model to represent. Resolution could instead be concentrated on one part of the loop, such as

the magnetization branch where beam is present, but this conflicts with the generic scope of the model, and not every flat top lies on the major loop. Under autoregressive operation it would also leave the demagnetization ramp-down under-resolved, where a misprediction seeds the following cycle and cascades. Subtracting a per-cycle reference B_c^* would isolate the modulation directly, but a reference must then be defined per cycle, which reintroduces the cycle concept the generic model avoids and forces a new reference and retraining for any unseen cycle. The reference is also not fixed, consequently the full residual must be used without further preprocessing.

With the residual fixed as the target, the prediction task maps a past window of measured current and field together with the future current to the residual over a future window of arbitrary length, which at inference spans the coming operational cycle. For a context of length L_c ending at step k and a horizon of length L_h, the residual is predicted as

$$\hat{r}_{k+1:k+L_h} = f_\theta\Big(\{t_i, I_i, \dot{I}_i, r_i\}_{i=k-L_c+1}^{k},\ \{t_i, I_i, \dot{I}_i\}_{i=k+1}^{k+L_h}\Big)\,, \tag{7.1}$$

with the field recovered as $\hat{B}_i = \hat{r}_i + B_{\mathrm{cal}}(I_i)$. Eddy current decay depends on the field rate of change $\dot{B}$, which is unavailable over the future window where the field is the prediction target, so $\dot{I}$, computed from the future current, is used as a proxy for the driving rate. All input channels are scaled to -1–1, and the time input is computed as the relative time between samples and scaled by its maximum, so the model reads the spacing between points and not an absolute timestamp.

Equation 7.1 realizes the approximator $\hat{\mathcal{F}}_\theta$ of Equation 5.3, with the finite context replacing the full magnetic history $\mathcal{S}(j)$ and the window taken at an arbitrary position in place of the cycle index, following the generic scope of the modeling requirements.

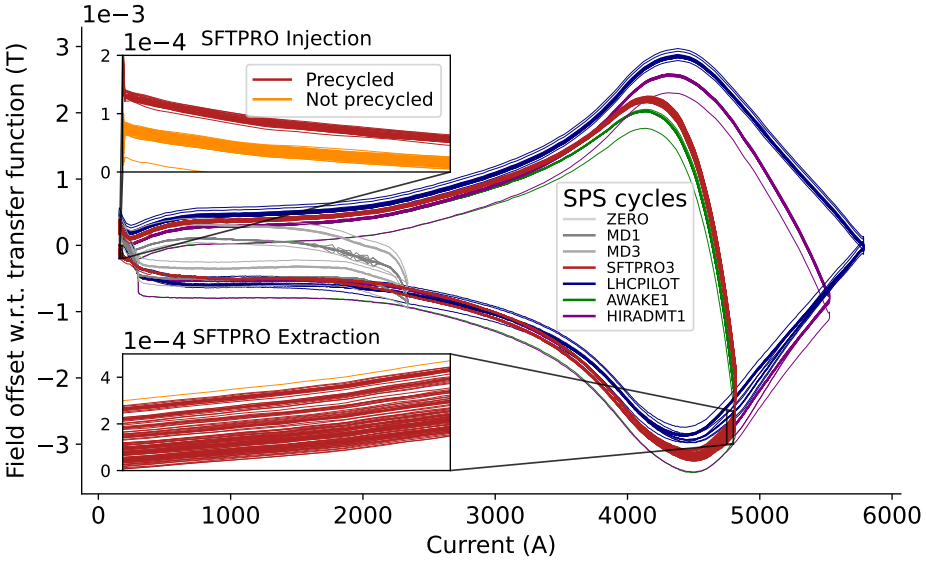


Fig. 7.1: Measured field residual by cycle type after subtracting the fitted calibration curve, shown against the main dipole current, for a custom measured sequence transitioning between different operational supercycles. The insets zoom the SFTPRO injection and extraction points, the injection inset comparing the precycled and non-precycled cases.

7.2.4 Training, Validation, and Test Split

The training set comprises 108 uninterrupted measurement sets, with one further uninterrupted sequence held out for validation and one for test. Any given set, including the validation and test sets must be a single uninterrupted measurement, since the magnetic state is carried continuously through the sequence and separate measurements cannot be stitched together without breaking that continuity. Downsampling by 20 reduces each phase to 50 Hz, and stride-one sampling across the 20 phases retains the full point count, giving roughly 40 million training samples at 50 Hz spacing.

The validation sequence is 578 400 samples long, about ten minutes at the 1 kHz logging rate, and it contains the relaxation of the fixed target cycle across a transition from an LHC filling sequence to a standard physics and AWAKE sequence, and back to a combined SFT and LHC sequence. This sequence was selected for validation because the few fully custom sequences with strongly varied excitation were retained for training, where their coverage of the excitation space is most valuable.

The test sequence pairs a 33 GeV/c SFTION cycle with an LHC cycle, a combination absent from training. That SFTION cycle does appear in training, though only without an adjacent LHC cycle, so the test probes its response to the high excitation the pairing introduces, a magnetic history it did not see in fitting.

7.3 Preliminary Studies

7.3.1 Gray-Box and Physics-Parametrized Models

Physics-parametrized models constrain their output to physically valid magnetic states, carry parameters with physical meaning, and expose a fitted state that can be interpreted and reused, which makes them a first candidate for the operator $\mathcal{F}$. Their general formulation was reviewed in Section 2.5.1.1 and Section 2.5.2, and their structural insufficiency for arbitrary operational histories was established in Section 5.4.3. Models of this class were implemented in both structurally constrained and physics-informed forms, and assessed against the reproducibility target of Equation 5.2.

The differentiable Preisach model [113] from Section 2.5.3 was fitted to the measured residual by gradient descent. Although rate-independent, the model was fitted to the measured residual to absorb its static loop structure, so that a rate-dependent sequence model would then isolate the remaining dynamics, in the same manner the calibration curve already removes the anhysteretic part. The fit reproduced most of the loop structure, and yet the numerical range of the residual was not appreciably reduced, since the loop shape the model can represent is constrained and does not match the measured loop shape, which was also in part caused by the difficulty of the optimization. Subtraction of the anhysteretic response, to make the small-signal problem tractable, also removes the physical magnetization the Preisach structure is built to represent, so the loops fitted on the residual are no longer physical hysteresis loops, and the inductive bias no longer aligns with the data. The model assumes a fixed initial magnetic state shared between training and evaluation, and it carries no mechanism to encode a past context, so it cannot be initialized from the arbitrary measurable window required in Section 7.1, though a known state could be approached by precycling with operational sequences and carrying the final state into the predicted cycle. It supports no mini-batch optimization and is fitted over a whole sequence in a single batch, which prevents fitting on the large volume of logged field data. The mesh is also fixed at training, and its resolution cannot be changed at inference. Furthermore, resolving the field to the reproducibility target over the full 0 T–2 T range would need of order

1×10^5 hysterons to reach an uniform 1×10^{-5} T resolution, whereas the residual is representable with of order 600. Finally, as each hysteron is switched through a sigmoid at every step of the sequence, the forward and backward graph grows with the hysteron count and the sequence length together, and consequently buying resolution by adding fixed hysterons makes training rapidly more expensive.

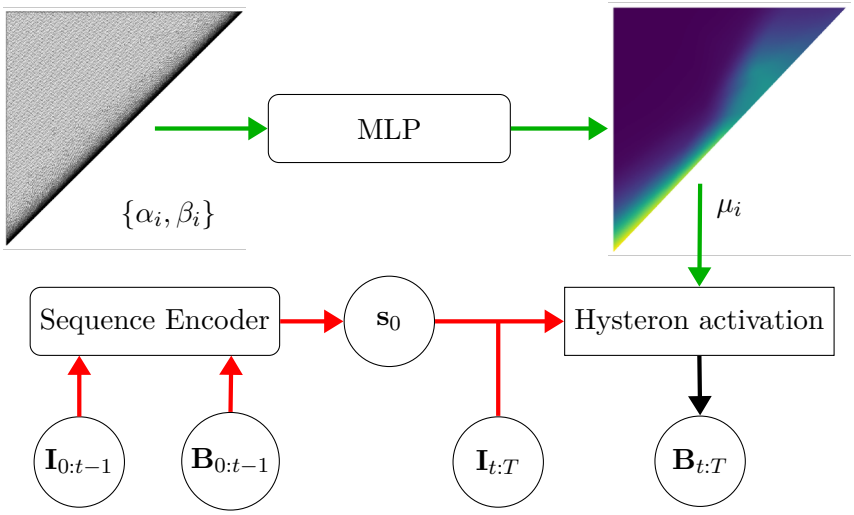


Fig. 7.2: Neural-network-parametrized Preisach architecture.

A neural reparametrization of the Preisach model was explored next as part of this work, aimed at solving the limitations of the fixed-parametrized version above, and the proposed architecture is shown in Figure 7.2. Neural representations of the Preisach density are not new, as [143] realised the Everett function $F(\alpha, \beta)$ with a feed-forward network taking the Preisach-plane coordinates as input, and [61] replaced the density by a recurrent network operating on play-operator outputs to model dynamic hysteresis in accelerator magnet iron. Both treat the initial state as fixed or discard an initial transient, and neither addresses mesh-free resolution or mini-batch training, which are the limitations identified above for [113].

A shallow feed-forward neural network represents the hysteron density as an implicit continuous function, mapping a position on the Preisach plane α_i, β_i to its density weight μ_i, so the mesh points become model inputs and can be sampled at an arbitrary resolution at inference without retraining. During training the mesh coordinates are sampled randomly and perturbed by noise, which smooths the learned density and improves robustness, without adding more mesh points to the forward and backward graph. Because the density is a continuous function of position, mesh coordinates that carry negligible weight can also be discarded after training, and additional hysterons can be added in high-density regions. A sequence encoder produces the initial hysteron state from the past current and field, so the state is sample-specific, and the model supports mini-batch training and initialization from an arbitrary measured history. Physical structure is imposed through forced switching of the hysterons according to the last encoded field, a saturation regularizer that prevents collapse to an all-on or all-off state, an auxiliary loss matching the density-weighted mean of the initial states to the measured field at the end of the encoder window, and soft priors that bias the density toward low coercivity or other material properties. Training was designed in three phases, with the encoder fitted first against a mock density, then the two fitted jointly, and finally in alternation, so that the density network does not corrupt the encoder gradients early in training.

The model reproduced individual loop shapes, produced a physically bounded and interpretable density, and yielded an encoder state that represents magnetic history generically, yet it did not reach the reproducibility target. A reproduced loop describes the response to a single excitation, and the operational target is how that loop is modulated from one cycle to the next by the preceding sequence, a history-dependent memory the model does not capture. Roles of the sequence encoder and the density network are entangled, an identifiability problem that kept the optimization difficult despite the small model, which the three-phase schedule mitigates but does not remove. The encoder is also not constrained to produce a physically valid Preisach memory state, since the forced switching and saturation terms fix only part of the hysteron configuration, so the encoded initial state can be non-physical even where the switching readout keeps the output physical. Constraining the model to the required accuracy needs dedicated first-order-reversal-curve and loop measurements, which standard magnetic measurements at CERN do not produce and which were not gathered here, again limited by data availability. The approach demonstrated the intended capabilities and is left as future work, where a dedicated measurement campaign would support it as a standalone study.

A physics-informed model based on the Bouc-Wen equation was also evaluated [145]. The Bouc-Wen model [138] is a Duhem-type formulation of Section 2.5.2 and similar to the other models in the same family cannot describe the hysteresis of an assembled accelerator magnet. An unrestricted Long Short-Term Memory (LSTM) backbone alone represents the required rate-dependent dynamics and can fit, and even overfit, the measured data, so the rate-independent Bouc-Wen equation was added only as a soft physics-informed loss term [109] to bias the fit toward physical behavior and improve generalization, and not as a constraint on the output. Because the imposed equation does not describe the magnet, this term drives the fit toward a wrong model class, and a prior meant to improve generalization instead prevents it, so no useful accuracy or generalization was reached.

7.3.2 Black-Box Baselines

Three simple data-driven model families are assessed against the requirements of Section 7.1 to establish whether the sequence structure introduced below is necessary. Each predicts the field without that structure and is drawn from a prior approach to hysteresis or time-series prediction.

An autoregressive neural network of the Nonlinear Autoregressive Network with Exogenous Inputs (NARX) type, applied to magnetic measurements by [4], is the previous data-driven approach for this problem, and predicts each field point from a finite window of past field and current fed back through the model. The finite lag carries only limited memory, and under autoregressive feedback on the large volume of sampled data, where long stationary sections and measurement noise dominate, the accumulated error makes the fit unstable.

A plain LSTM improves on the autoregressive network through a gated recurrent state that learns how much past information to retain and carries a persistent, adaptive memory, where the autoregressive network has only a fixed finite lag. A compact network captures the dominant loops of the data with only a few hundred to a few thousand parameters, yet the cycle-to-cycle modulation that is the operational target is not resolved at this size, since a single recurrent state compresses the history. The encoder-decoder LSTM evaluated in Section 7.6.2 recovers part of the modulation by encoding the past window into a separate state carried to the decoder, which motivates the encoder-decoder framing below.

Neural operators have shown recent promise as function approximators, including to hysteresis prediction. The Fourier neural operator application [32] demonstrated high-accuracy prediction on generic hysteresis loops generated from the Preisach model, though it assumes periodic sequences through its Fourier basis, and was benchmarked on data that start and end at the

same value. The operator was applied to the measurements here, with an encoder added to supply the initial magnetic state that it does not natively represent. The measured windows are of arbitrary position and are not periodic, so the Fourier basis assumption does not hold, and the fitted prediction showed oscillatory overshoot at the sharp ramp edges, the Gibbs behavior of a truncated Fourier series, where the fast transients must be resolved most accurately. Restricting the operator to a single cycle, which is closer to periodic, would recover the assumption but conflicts with the model's intended generality.

Across the three families a common gap remains, since none conditions a whole-cycle prediction on the known future current from a generic, non-periodic past window. This motivates the encoder-decoder framing of Section 7.4, where an encoder represents the magnetic history from a past window and a decoder predicts the field conditioned on the future current.

7.4 Architecture Selection

The preliminary studies establish that neither the structured physics models nor the naive data-driven models meet the requirements, so a suitable architecture is selected here from the sequence-modeling literature. Following Section 5.4.4, the task admits two equivalent formulations, as function approximation that maps a past context and the future current to the future field, and as time-series forecasting that predicts the future field with the future current supplied as a known exogenous variable. Both formulations require conditioning on the future current, and the function-approximation view carries no requirement of causality. Together with the initialization from an arbitrary non-periodic history and the variable-length full-cycle horizon of Section 7.1, these properties act as the filter, and narrow the survey to encoder-decoder models that accept future covariates and impose no causal mask.

7.4.1 Candidate Architectures

The closest architectural precedent is the transformer applied to measured hysteresis loops by Li et al. [89], which encodes a full prescribed excitation waveform to predict the field over one periodic cycle, an encoding of the whole excitation that parallels the conditioning on known future current required here. That model supplies no preceding window of measured field and excitation and reduces the data to single-cycle waveforms in periodic steady state, so the initial magnetic state is fixed implicitly by the loop.

The model developed here instead takes a past window of measured current and field together with the programmed current over the coming cycle, and predicts the field realized over that cycle, expressed as $\{I_{\text{past}}, B_{\text{past}}, I_{\text{future}}\} \rightarrow B_{\text{future}}$. The past window identifies the inherited magnetic state, the known future current specifies the excitation trajectory, and no periodic steady state is assumed.

A full-cycle horizon requires long-range handling across the whole window, which a compact recurrent state cannot supply. Attention addresses this by giving each predicted step access to the full encoded sequence.

The task is approached as long-horizon time-series forecasting, for which the architectures evaluated here are the state of the art at the time of this work. Most are transformer-based, with competitive linear and all-Multilayer Perceptron (MLP) alternatives, and target a forecasting setting. A generic transformer is not adopted by default, since [144] report that pure transformer forecasters do not consistently improve on a simple linear predictor. PatchTST [103] segments the input into patches to reach long context lengths, but it processes each channel independently and represents no interaction between the current and the field, and it emits its output through a fixed linear head. iTransformer [91] instead attends across the variables to recover that interaction,

though it embeds the whole lookback per variable and projects to a fixed horizon. Both are encoder-only Direct Multi-Step (DMS) models whose input and output lengths are fixed at training, and neither conditions on a future covariate. TimeXer [137] extends this class with explicit support for exogenous variables and could supply the future current to a forecast, and it remains a fixed-length direct multi-step model that is excluded by the variable-length full-cycle horizon and the arbitrary-history initialization. TSMixer [34], an all-MLP model, extends the NARX-like point-wise predictor across the whole horizon and shares the same fixed-length direct multi-step limitation, so scaling a multilayer perceptron over the horizon does not meet the requirements either. Autoformer [142] decomposes the series into trend and seasonal components through an auto-correlation mechanism that presumes the periodicity the generic sequences here do not have, and Informer [146], which scales attention to long horizons through a sparse approximation, emits a fixed horizon without conditioning on future covariates.

A fixed prediction horizon could still be retained for any of these models through one of two workarounds. Setting the horizon to the longest operational cycle and padding every shorter cycle to that length spends model capacity and computation on the padded region of the many shorter cycles. Shortening the horizon and predicting a full cycle in several chained steps introduces discontinuities at the step boundaries and compounds the prediction due to autoregressive instability.

Conditioning on the known future excitation and decoding a cycle of variable length from an arbitrary history selects an encoder-decoder structure that ingests future covariates, which the encoder-only designs above cannot provide. The Temporal Fusion Transformer (TFT) [90] was the state-of-the-art architecture for interpretable multivariate timeseries forecasting, and meets these requirements, mixing channels, accepting future inputs, and decoding variable lengths, and it combines an LSTM backbone for local structure with attention for the long-range dependencies. The TFT is therefore adopted as the field-prediction model, and its added components are assessed against three encoder-decoder baselines of increasing structure, an encoder-decoder LSTM, the same model with attention, and a transformer variant, with the value of each increment quantified in the controlled study of Section 7.6.

7.4.2 Model Architecture

The past context seeds the encoder and is functionally analogous to the magnet state at the start of the coming cycle. The future window conditions the decoder for the field prediction over that window, mapping the inputs of Equation 7.1 onto the two pathways of the TFT shown in Figure 7.3.

The Gated Residual Networks (GRNs) combine fully connected layers with a gating unit and a residual skip, which lets a component contributing little to the prediction be suppressed by its gate and bypassed through the skip, and variable selection networks weight the input channels by importance, although only a few are present for this task. The field response is a joint function of the current and its rate, and in this architecture the current and field are embedded into a shared representation and processed together by the recurrent backbone and the attention, a channel mixing that a channel-independent model cannot provide. A recurrent LSTM backbone encodes the local sequence structure, and multi-head attention supplies the long-range dependencies across the full cycle. Furthermore, the attention queries are formed over the future horizon with keys and values over the full context and horizon range, so every predicted step is informed by the whole seeded context together with the future excitation. A known-future pathway carries the future current into the decoder.

A single static covariate is used through the static enrichment pathway, the encoder length, which records how many steps of the right-aligned context window are measured data and how

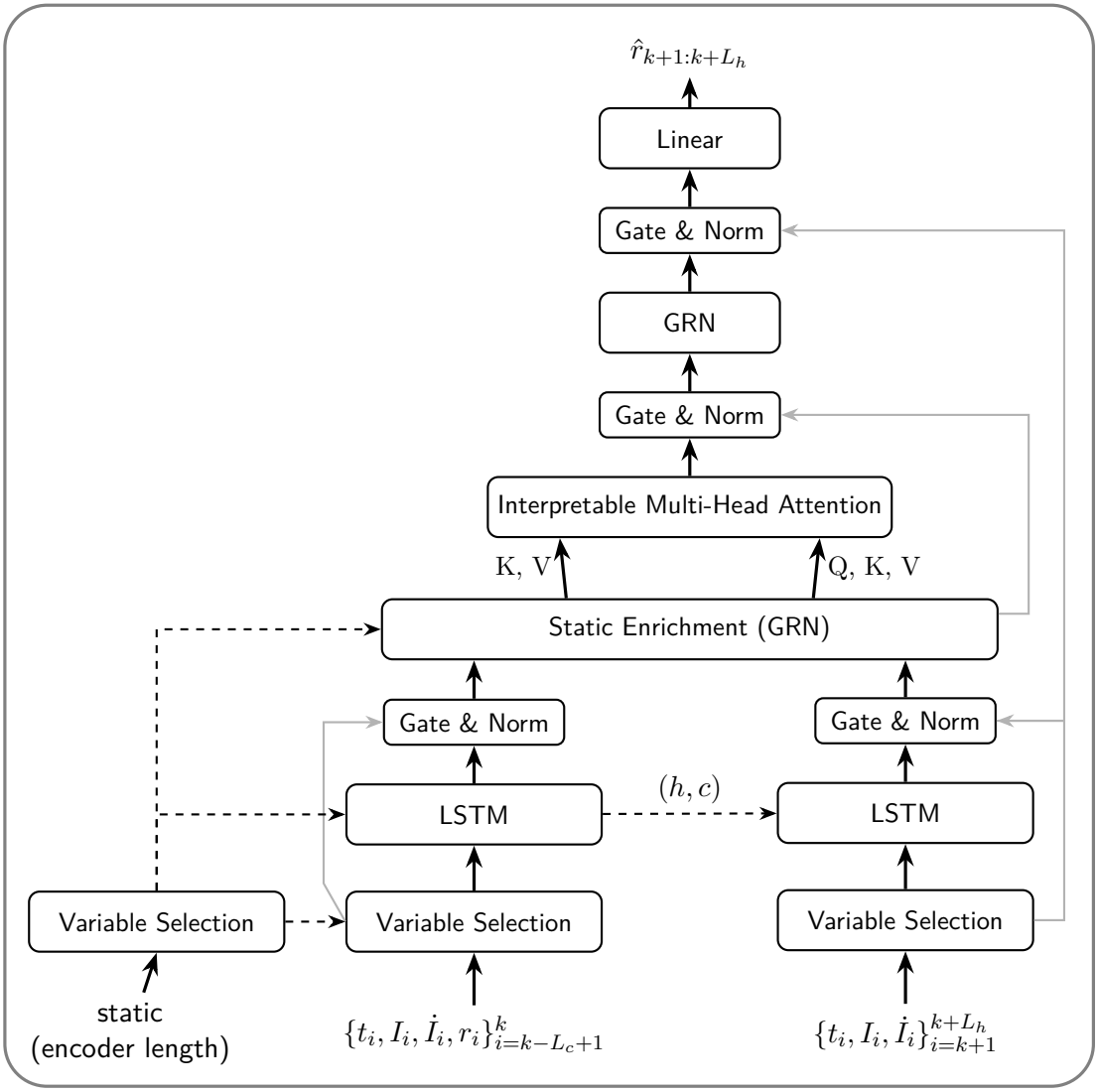


Fig. 7.3: Adapted TFT encoder-decoder architecture, with the encoder length passed as the single static covariate, and attention queries formed over the future horizon drawing keys and values from the full context and horizon range.

many are padding, so a short history is down-weighted accordingly. The decoder length is not passed in the same way, since it would make the prediction at a step depend on the chosen horizon length, an artifact of the query window and not a physical property of the magnet, and the decoder padding is instead excluded through the attention mask and the loss mask. No causal mask is applied over the attention, which admits the full-range keys and values used above.

7.5 Training

Training on the measured field alone is constrained by the scarcity and narrow excitation coverage of the operational data. The logged sequences span only the discrete cycle programs played in operation and carry no arbitrary excitation, and the broader coverage a dedicated campaign would add is absent, so the model is specialized to the operational sequences it was trained on. A model valid at an arbitrary point in an arbitrary sequence requires training data across a broad

excitation and history space, and the long prediction horizon is expected to require a model of sufficient capacity, which in turn requires a large volume of data to fit. Within the available compute budget the model dimension was scanned and set to the value in Table 7.1, large enough for the long horizon. Both needs are met by pretraining on simulated sequences that sample the excitation space densely and are free of measurement drift and noise, from which the structural prior is transferred and then adapted to the measured loop shape by fine-tuning.

7.5.1 Simulated Pretraining Data

To make the synthetic waveforms resemble operation, each simulated current program follows the operational cycle structure of an injection plateau, ramp-up, flat-top, ramp-down, and return to idle, with the amplitude and duration of every phase randomized over operationally plausible ranges so that no two cycles are identical while all remain realistic. The current levels are matched to the operational machine so that the simulated excitation overlaps the operational range, with the idle current at 155.63 A, the injection plateau at 161 A or 302 A and varied by uniform sampling in ± 1 A, and the maximum at 5800 A. A low-energy cycle reaching 2339 A was sampled at 20 percent probability to stand in for the MD1 precycle, while cycle lengths span 3.6 s–24 s and flat-top currents span 500 A–5800 A. Unit geometry in this case is also assumed, so the excitation current I is used directly as the magnetizing field H in the simulation.

The field is simulated with the Jiles-Atherton model of Section 2.5.2, chosen as a physical hysteresis model that reproduces the structural features of history dependence and branch memory at low cost and free of the B-train measurement drift. A single parameter set is used since the model targets a single magnet and takes no hysteresis parameters as input, and parameters are adapted from soft-iron values with the coercivity reduced to produce the small loops of the SPS main dipole field, and each sequence starts from a demagnetized state. By construction the Jiles-Atherton loops widen near coercivity at low field, where the irreversible magnetization change is largest, unlike the measured response that widens near two-thirds of the maximum current as shown in Figure 7.1. This mismatch is expected, since no ODE or operator-based hysteresis model reproduces the response of the assembled magnet to the required accuracy, as established in Section 5.4.3, and such a model would already solve the compensation problem. Therefore, the simulation can only provide a structural prior, not a model of the measured response, and its parameters were selected for representative loops without optimization against the measurements. A more expressive generator such as Preisach could reproduce a wider range of loop shapes, but this offers no advantage here, since only the shape-agnostic structural prior transfers and the shape is corrected during fine-tuning. Preisach would moreover require fitting a density to obtain representative loops, which proved difficult on this data in the study of Section 7.3.1, whereas the Jiles-Atherton form follows from a few hand-set parameters, so it is used here.

To add dynamic behaviour to the simulation, a first-order relaxation term of the form used in Section 2.6.2 is added on top of the static hysteresis, so that the target carries a decaying dynamic component within and between cycles. Its decay time is set to 7 s, longer than a typical plateau, a value chosen so that a pronounced decay evolves across a dwell without settling and carries over between cycles, a slow history-dependent dynamic that a static current-to-field map cannot reproduce. Only the presence of such a carried-over decay transfers as structural prior, and not its exact constant, so the value is set for a clearly visible dynamic and is not fitted to a measured decay time. An anhysteretic curve is computed as the midpoint between the magnetization and demagnetization branches of the precycle, which reaches the maximum current, as for the measured data, and subtracted so that the simulated and measured targets share the same residual form, leaving the hysteresis and eddy residual as the simulated target.

The simulation is generated at the 1 kHz native rate of the measured data, fine enough to integrate the eddy term and to resolve the sharp ramp edges that drive it, which the rate-independent hysteresis itself does not require. The simulation was originally intended to provide excitation waveforms for magnetic measurements at the same rate, but since that campaign did not take place, as explained in Chapter 6, the excitation waveform is used here to pretrain the model instead. To reduce the large and redundant series at the native 1 kHz, it is downsampled with the Ramer-Douglas-Peucker (RDP) [110] algorithm, which preserves the waveform shape while bounding the largest time step to a few hundred milliseconds. The data is generated at several ε tolerances, exposing the model to a range of sample spacings so that it relies on the relative-time input and does not assume a fixed step. The pretraining data comprises over 40 hours of simulated sequences, with one hour each held out for validation and test, respectively.

7.5.2 Pretraining and Fine-Tuning

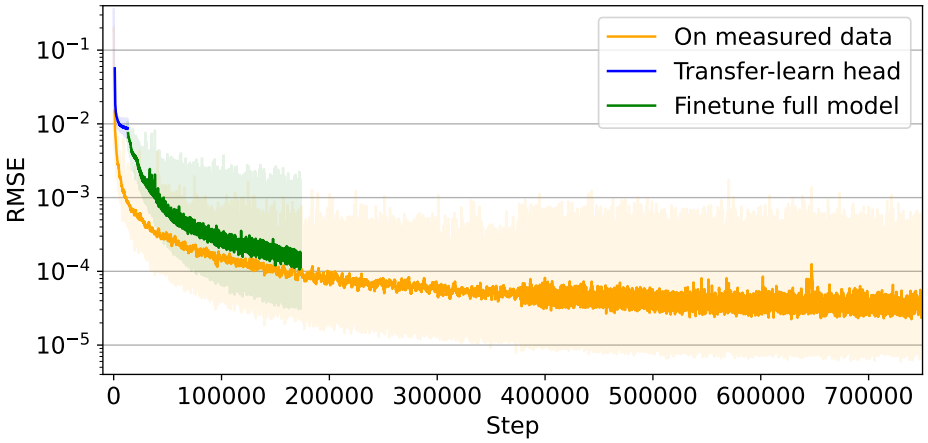
Pretraining establishes that the architecture represents hysteresis on these noise-free and drift-free sequences before the measured problem is approached. Although the architecture accepts arbitrary lengths, the range seen in training bounds where the model is reliable, so the context and horizon lengths are set from the operational cycle durations and the same range is replicated at inference. At the 50 Hz sample rate the lower bound of 180 steps corresponds to 3.6 s, the length of the MD1 precycle, and the upper bound of 1020 steps corresponds to 20.4 s, a three-injection LHC cycle. To promote robustness to arbitrary input and output lengths, sequence length randomization is applied by drawing windowed context and horizon pairs from the sequences with both lengths varied per sample, and the windows packed together with adapted attention and loss masks that exclude the padding. Zero-mean Gaussian noise of standard deviation 0.005 in the normalized input space of -1–1 is added to the inputs at this stage to promote robustness against the measurement noise seen in the measured data, and the residual target is fitted under a mean-squared-error objective with the AdamW optimizer and early stopping on the validation MSE. The model and training hyperparameters are listed in Table 7.1.

Tab. 7.1: Model and training hyperparameters.

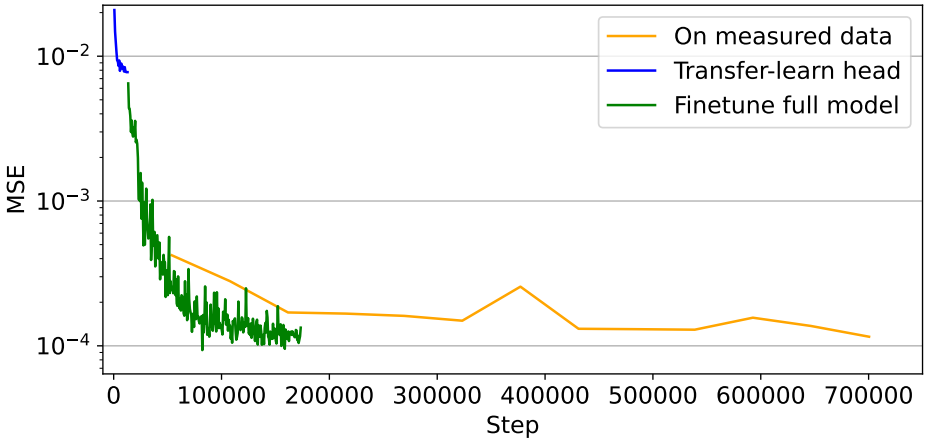
Hyperparameter	Value
Model dimension	256
Attention heads	4
LSTM layers	2
Dropout	0.1
Context length	180 to 1020
Horizon length	180 to 1020
Batch size	64
Optimizer	AdamW
Loss	MSE
Pretraining learning rate	1×10^{-4}
Head-only fine-tuning learning rate	1×10^{-5}
Full fine-tuning learning rate	5×10^{-5}
Weight decay	1×10^{-4}

Fine-tuning transfers the pre-trained model to the measured data in two stages, which retains the shape-agnostic structural prior of history dependence, branch and return-point memory, and rate-dependent dynamics while adapting the loop shape, and avoids the negative transfer of

an unconstrained update that would overwrite the prior. Following standard transfer learning practice, the head layers are adapted first, unlocking only the GRN, gate-and-norm, and linear head layers of Figure 7.3 and freezing the variable selection, LSTM, static enrichment, and attention, and training with a low learning rate until the validation loss plateaus. In the second stage, the whole network is unfrozen and fine-tuned at a higher learning rate until the validation loss plateaus again. In the fine-tuning stage, the windowing, masking, and optimization of the pretraining are reused, with the Gaussian noise removed.



(a) Training MSE.



(b) Validation MSE.

Fig. 7.4: Training and validation MSE for from-scratch training on measured data against the two-stage transfer (head-only, then full fine-tuning).

The benefit of the pretraining is also visible in the training dynamics of Figure 7.4. Head-only fine-tuning converges within a few thousand steps as the output layers adapt to the measured loop shape while the frozen backbone holds the transferred prior, and the full fine-tune then plateaus within about 170 000 steps. Training from scratch on the measured data reaches a similar training loss only after roughly 700 000 steps, and its validation loss stalls while its training loss keeps falling, the widening gap of a fit overfitting the limited measured data, whereas the transfer keeps the two losses together. Fine-tuning moreover draws its windows at a stride of 20, about twenty times fewer positions than the stride-one sampling of Section 7.2.4, and the

twenty phase-shifted copies of each waveform carry nearly the same information, so the reduction is in points and not in equivalent information. The transferred prior therefore shortens training, regularizes the fit, and reaches the measured response from fewer points, with the field accuracy of the final model reported in Table 7.2. A shared prior also opens a path to fine-tuning the model to other magnet families.

7.6 Results

The TFT was selected on the deployment requirements of Section 7.4 and adopted as the operational model, and the controlled study characterizes that selected model against simpler baselines trained from scratch [92]. The study is a component comparison of the chosen model and not an exhaustive search over architectures, so it supports the sufficiency of the model for operation and not its optimality. All models are trained on the same dataset with the same protocol and share the hidden dimension, so the differences in Table 7.2 reflect the architecture and the pretraining under matched conditions.

Each model is evaluated in three modes. Non-cycled evaluation predicts a single window at the maximum horizon from a measured context, giving the standard held-out RMSE. Cycled evaluation predicts one cycle at a time at its own length, which matches the per-cycle unit used operationally. Autoregressive evaluation feeds each predicted cycle back as the context for the next, with no measured field re-seeding after the first, which tests the open-loop error accumulation relevant to autonomous running and to the higher-order magnets that carry no online field measurement.

7.6.1 Ablations

The selected model is compared against a sequence of encoder-decoder baselines of increasing structure, each adding one component over the previous one. The sequence runs from an encoder-decoder LSTM through attention-augmented and transformer variants to the TFT, with a final comparison between the TFT trained from only measured data and the same model after pretraining and transfer.

The encoder-decoder LSTM (EDLSTM) is the simplest baseline. An encoder LSTM compresses the past window into its final hidden and cell state, which initializes a decoder LSTM that produces the field over the horizon from the future current. The encoder-decoder split, in place of a single LSTM over the whole sequence, is what encodes the past context into a state and carries it to the decoder, separating the inherited magnetic state from the prediction of the coming cycle. Its limitation is that single state, since the whole past is compressed into it and the decoder cannot access the past directly.

The attention-augmented LSTM (ATTNLSTM) removes the state bottleneck. The decoder queries the full encoded past and future sequence through cross- and self-attention, so the past is accessible at every decoder step and not only through the compressed state. The attention follows the same query-over-horizon, keys-and-values-over-full-range form as the TFT of Section 7.4.2.

The transformer-LSTM (TFLSTM) stacks several attention blocks on top of an encoder-decoder LSTM, with the LSTMs acting as the embedding layers in place of positional encodings. The past and future features are each passed through a gated residual network into the encoder and decoder LSTM, respectively. The decoder LSTM is initialized from the final encoder state, and each LSTM output is added to its gated residual network output before entering the block stack. Each block is a simplified transformer block, without the position-wise feed-forward sublayer. It applies encoder self-attention, followed by decoder attention over the concatenated encoder and decoder sequences. Each attention operation is followed by a gated residual connection, and the

attention itself is the interpretable multi-head attention introduced in the TFT. A linear layer maps the final decoder representation to the output.

The TFT trained from scratch with only measurements is the reference, using the same recurrent and attention combination with different input handling, through per-signal variable selection, gating at each stage, and the static-covariate pathway, and with a single attention layer in place of the stacked blocks of the transformer-LSTM. The fine-tuned TFT adds the pretraining and two-stage transfer of Section 7.5, isolating the contribution of the transferred prior.

7.6.2 Controlled Study

Tab. 7.2: RMSE calculated between the measured magnetic field from SPS main dipoles and the hysteresis model predictions with different evaluation methods, including autoregressive (AR), on validation and unseen test set.

	validation RMSE (10^{-4} T)			test RMSE (10^{-4} T)		
Model	Non-cycled	Cycled	AR	Non-cycled	Cycled	AR
EDLSTM	0.36	0.42	0.42	1.48	1.56	6.01
ATTNLSTM	0.34	0.43	0.46	2.07	1.36	5.70
TFLSTM	0.27	**0.38**	**0.39**	1.09	1.27	5.28
TFT	0.42	0.52	0.97	1.53	1.31	5.81
TFT (finetuned)	**0.18**	0.41	0.55	**0.33**	**0.42**	**0.46**

On the validation set the five models reach a comparable accuracy below 1×10^{-4} T, since the validation sequence shares the operational programs seen in training, so validation does not separate the architectures, with the TFLSTM marginally the best of the models without transfer. On the unseen test set the models without transfer degrade, most sharply under autoregression, where the error compounds to 5×10^{-4} T–6×10^{-4} T, while the fine-tuned TFT stays at or below 5×10^{-5} T on the test set in every mode.

The TFT without transfer is the largest model yet does not improve on the smaller TFLSTM, so on this data added capacity is not the limiting factor, and the architectural additions give no consistent improvement when trained from scratch. The capacity expected to matter for the long horizon therefore does not dominate here, where the outcome is set by the data variety and the pretraining. The decisive factor is the pretraining, which produces the accuracy on the unseen test sequence and the stability under autoregression that the models without transfer do not reach, with the fine-tuned test error reaching a few 10^{-5} T, near the measurement noise and drift floor.

The test is a single unseen sequence, so the ordering among the models without transfer is not robust, while the order-of-magnitude gap under autoregression and the effect of the pretraining are. A standard held-out RMSE is therefore necessary but not sufficient for a model that runs autoregressively in operation, which the autoregressive evaluation addresses here and the operational test with beam develops in Chapter 8.

Figure 7.5 follows the fine-tuned prediction over the validation sequence in the cycled mode, each cycle seeded from the measured field, as the reference-relative residual $\Delta_{\mathrm{meas}} - \Delta_{\mathrm{pred}}$. This residual is the quantity in which the reproducibility criterion of Equation 5.2 and the feed-forward correction of Equation 5.4 are defined, while the controlled study of Table 7.2 reports the absolute prediction error that ranks the models directly. It reaches an RMSE of 2.71×10^{-5} T over the full sequence and 4.13×10^{-5} T counted only at the fixed-target flat top, close to the measurement

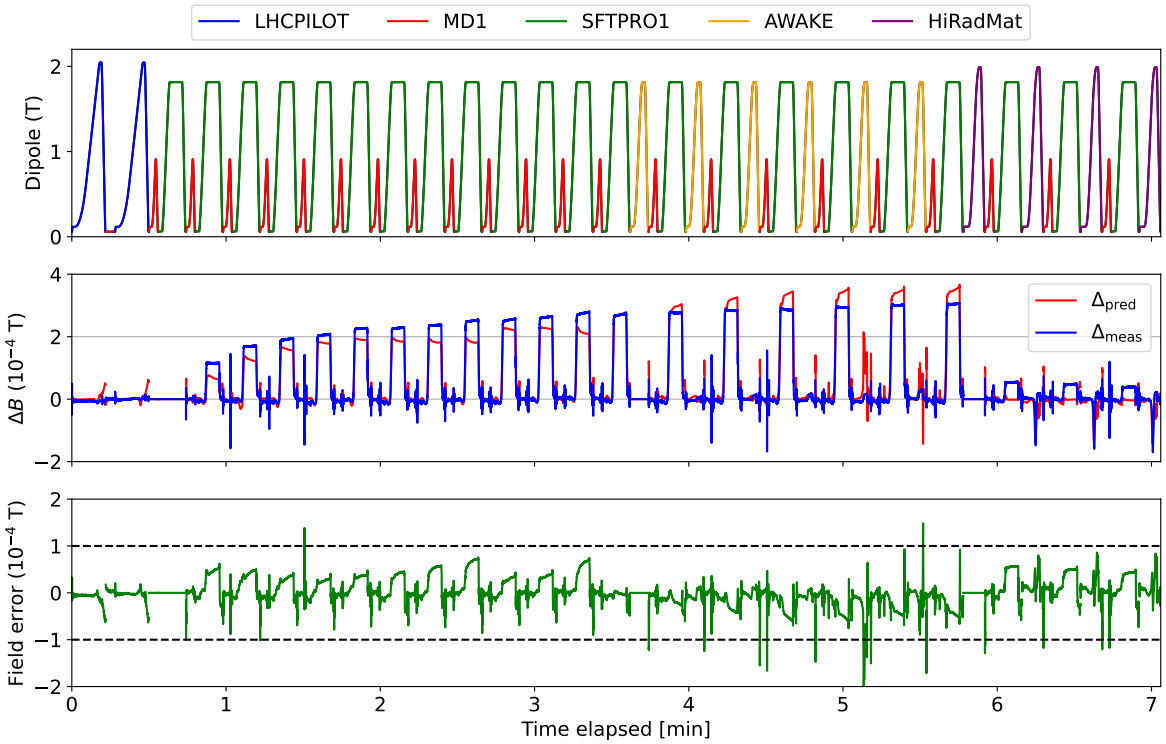


Fig. 7.5: Fine-tuned prediction over the validation sequence. The top panel shows the measured field coloured by cycle type, the middle panel the measured and predicted field relative to the first occurrence of each cycle type, Δ_{meas} and Δ_{pred}, and the bottom panel the reference-relative residual $\Delta_{\text{meas}} - \Delta_{\text{pred}}$ against the 1×10^{-4} T band.

noise and drift floor and above the 2×10^{-5} T reproducibility tolerance that the floor bounds. The residual stays within the 1×10^{-4} T band across the supercycle transitions and the inter-cycle field decay is tracked, while peak excursions at the ramp transitions exceed the band.

The model meets the structural requirements of Section 7.1, initializing from an arbitrary seeded window, predicting a full cycle of variable length, conditioning on the future current, and running autoregressively. The accuracy reported above is established on the held-out operational sequences, including a held-out high-excitation history, and generalization to arbitrary or random excitation remains untested, bounded by the data limitations of Section 7.7. Its sufficiency for compensation is assessed at beam level in the operational validation of Chapter 8.

7.7 Limitations and Considerations

Two bounds limit the accuracy reached in Section 7.6, one set by what the model represents and one by what the measurements constrain. A finite context and horizon, fixed by the operational cycle lengths and the quadratic cost of the attention, constrain the range of magnetic history the prediction draws on. Separately, the accuracy and the variety of the B-train field bound the resolution and the generality of the fitted model. This measurement bound recurs wherever a model is fit to this data, and its fuller treatment continues in the Discussion of Chapter 11.

Taking the first bound, the trained range falls short of the longest operational cycle. At 50 Hz the 1020-sample horizon spans 20.4 s and covers a three-injection LHC cycle, while the longest

operational cycle runs to 24 s, so that cycle is not represented over its full length during training. Because the architecture accepts arbitrary lengths, this reflects the trained range and not a hard limit, though accuracy is expected to degrade on windows longer than those seen in fitting.

Beyond the horizon, the magnetic memory of the circuit extends much further, as the field relaxation after a supercycle change from a higher-energy sequence to standard physics continues over roughly 15 min, as reported in Chapter 5. Representing it directly at 50 Hz would need of order 45 000 samples of context that the quadratic cost of the attention makes infeasible, so the window cannot hold the full memory. This full span is however not necessary, because the accumulated state is read from the seeded measured field at the start of the context, and the relaxation trend is recovered from the decay of the one or two preceding flat tops the window contains, which combined with the regularity of the operational excitation recovers the relaxation on the cycles that follow. History older than the window span is the exception, reaching the prediction only through the seed and not represented directly, so cases depending on it are not guaranteed, however during the operational sequences the model is trained on, longer implicit dependencies has not been observed. A longer context would be able to recover this dependency, at the quadratic cost that fixes the present trade-off.

The second bound follows from the measurements, and it acts first on resolution. The operationally relevant cycle-to-cycle modulation introduced with the residual target is of order 5×10^{-4} T, and the 2×10^{-5} T reproducibility tolerance on it coincides with the measurement noise and drift floor established in Chapter 6. No model, structured or data-driven, recovers accuracy below the information the measurements provide, so this floor bounds the gray-box models of Section 7.3.1 and the sequence models alike.

The variety of the measurements limits generality as much as their accuracy limits resolution. Because the operational data spans only the discrete cycle programs played in operation, with no arbitrary excitation and only small modulations in the programmed current, the controlled study cannot separate the architectures. A recurrent state already captures the regular repeated programs, so the gating and attention increments between the EDLSTM and ATTNLSTM are indistinguishable on this data and their value would appear only on richer excitation, while the single held-out sequence leaves the fine-grained ordering among the models inconclusive. The robust findings reduce to those identified in Section 7.6, the order-of-magnitude error growth under autoregression and the effect of the pretraining.

Within that narrow variety, the few programs present appear in very uneven proportions, which adds to the data scarcity. Because the fixed-target cycle appears in every supercycle its hysteretic states dominate the data, and especially populated with fixed-target idle-like states, in addition to states excited by higher-energy cycles. As measured episodes are continuous measurement sequences, they cannot be reweighted without breaking the state continuity the model depends on. Captured parasitically from operation, the data cannot be rebalanced by resampling, and correcting it needs the dedicated measurement campaign that falls outside this work. While some datasets were measured outside the constraints of regular operation, they are few and overshadowed by the operational sequences in the training data. Furthermore, the model aims to resolve a SFTPRO without a preceding MD1 precycle, which does not occur in operation. Eight sets were measured without the dipole precycle pulsing while retaining its cycle time, but these are few against the full 108 sequences. Furthermore, as seen in Figure 7.1, the precycled and non-precycled injection plateaus reduce to two discrete bands with no measured states in between, so the under-represented non-precycled case is fitted from its few sets, without support from the dominant precycled data.

Below even the selected data sits a residual integration drift in the B-train field, remaining after the low-drift sets are chosen. Because the field is the supervised target, large drift in a set would pull the fit toward an offset uncorrelated with any input and degrade the point-wise

accuracy, while leaving the prediction unbiased on average, so the low-drift sets are selected to keep this offset small. This drift is uncorrelated with any variable known at prediction time and changes sign and magnitude between measurement days, so it cannot be predicted and is intentionally left unmodeled.

At inference the drift of the seed is carried into the predicted cycle, so a prediction seeded under one drift and a reference taken under another are not directly comparable. The feed-forward correction of Equation 5.4 is computed relative to a reference, and a drift common to the seed and the reference cancels in that difference, so the offset does not reach the correction when both are taken under the same conditions. Because training spans only a narrow range of drift, a seed carrying an unfamiliar drift lies outside the training distribution, and the offset carried into its prediction is then unconstrained. On the fixed-target flat top the residual drift reaches up to 4×10^{-5} T, as shown in Figure 6.13, well below the cycle-to-cycle modulation to be resolved, and more selective gathering of low-drift sets narrows the training distribution and improves the statistics.

Chapter 8

Field Compensation and Operational Validation

The feed-forward compensation scheme proposed in Section 5.4.6 predicts the field of the coming cycle with the magnetic field model developed for the main dipoles in the previous chapter and trims the programmed quantity before the cycle is played. This chapter develops that scheme into an operational control loop for the main dipoles and validates it with beam. Section 8.1 sets out the control problem, the constraints the timing system places on it, and how the loop is implemented and deployed, first as an operator-facing client and mirrored to be headless on the server-based UCAP middleware. Section 8.2 reports the operational results on the SPS main dipoles, where the scheme is shown to reproduce the extraction field across operational supercycles.

8.1 Compensation Scheme Design

The compensation scheme is built on the accelerator's existing high-level control system and requires no dedicated commissioned hardware besides a networked computer to serve the model predictions and the control loop. The controls timing system of Section 4.4.2 announces which cycle and user is coming and when, the middleware exposes the programmed and measured quantities of each cycle, which are used for the prediction and correction. The correction can then be delivered as a field correction in the LSA (Section 4.4.1) settings hierarchy for each cycle separately, and the programmed field then propagates through the hierarchy to the excitation current played by the FGC (Section 4.4.6). This compensation scheme is designed to maintain high-level reproducibility, without correcting individual magnets, or correcting field back to an absolute fixed value, including the programmed field in LSA. The scheme is therefore transparent to the operators, who can continue to work with and adjust the same high-level settings and programs as before, without running into a race condition with the compensation loop. Operating within this infrastructure, the scheme reduces to a cycle-by-cycle feed-forward correction that trims the programmed quantity of each cycle before it is played, and the same tracking and correction logic can run either in an operator-facing client or headless on UCAP (Section 4.4.5), so that a result comparable to a dedicated real-time loop is achieved with no additional instrumentation.

8.1.1 The Control Problem

The correction acts at the level of the settings hierarchy, which fixes the temporal granularity of the control problem, since the current program of a cycle is committed once and played open-loop through the cycle. This requires the correction for a cycle to be decided before that cycle is played, and the realized field, although measured in real time and published to the middleware, in practice does not act on the cycle that produced it. The field of the coming cycle is therefore predicted from the magnetic history, and the correction proposed in Equation 5.4 is applied to the programmed quantity ahead of cycle start.

The compensation acts on the magnetic field, and the model developed in the previous chapter predicts that same field, so the loop is formulated entirely in field, with the correction rule of Equation 5.4 taken at $x = B$ for the main dipole circuit. To achieve field reproducibility, the correction is computed as a relative difference between the predicted field of the coming cycle and a reference field for the same cycle. The reference field is chosen as the first prediction of the cycle, so that every later occurrence is corrected back to it and the cycle produces the same field each time it is played. Because the reference and the running prediction come from the same model, any systematic model bias enters B_c^* and $\hat{B}_c$ alike and cancels in the residual $\Delta_c = \hat{B}_0 - \hat{B}_c$, and the bookkeeping that maintains these references across cycles is described in Section 8.1.2. Writing c for successive occurrences of a cycle and $u_c(t)$ for the correction applied to it, the loop realizes the update

$$u_c(t) = u_{c-1}(t) + g\,\Delta_c(t), \tag{8.1}$$

with $\Delta_c(t) = B_c^*(t) - \hat{B}_c(t)$ the predicted residual of Equation 5.4 and $0 < g \leq 1$ a fixed gain. Since the prediction $\hat{B}_c$ is made with the correction u_{c-1} already in place, Δ_c is the residual that remains on top of the previous correction, and Equation 8.1 therefore accumulates the correction across successive occurrences of the cycle.

Because the predictor developed in the previous chapter is imperfect, the gain is deliberately held below 1.0, since applying the full predicted residual at each occurrence would over-correct on the model error and let the field oscillate about the reference. Using a fractional gain closes the residual geometrically over successive occurrences and remains stable for any bounded prediction error, at the cost of the single-step convergence that only a perfect model would afford. The measured field from the B-Train seeds the following prediction and confirms, once the cycle has completed, that the correction reached the circuit in time.

8.1.2 Settings Management Integration

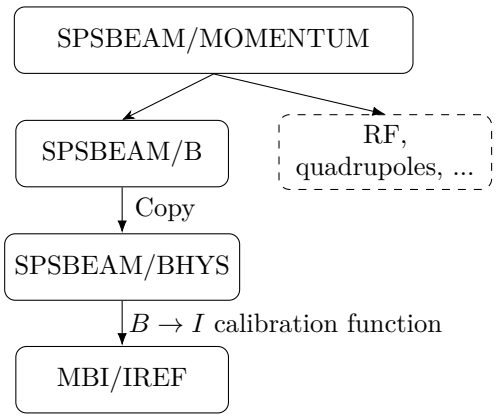


Fig. 8.1: The field correction is delivered to the machine as a trim of the programmed field, which propagates through the LSA settings hierarchy to the excitation current played by the FGC. The trim is keyed by the cycle name, so that each cycle can be corrected independently.

The correction is committed as a trim of the `SPSBEAM/BHYS` parameter, a dedicated field-correction program that sits one level below the programmed dipole field `SPSBEAM/B` in the settings hierarchy and is maintained as a copy of it. The new hierarchy shown in Figure 8.1 extends the standard LSA settings hierarchy shown in Figure 4.11 with the `SPSBEAM/BHYS`

program inserted below `SPSBEAM/B`, an addition that is a routine operation in the control system. Each occurrence writes only the correction settings part of the `SPSBEAM/BHYS` parameter, with the accumulated correction u_c of Equation 8.1 set as a delta on top of that copied field so that the corrected field propagates downward. Placing the correction below `SPSBEAM/B` lets operators and other applications keep acting on the dipole field and the momentum above it while the compensation acts on the correction below, so that the two never contend for the same parameter. The corrected field passes through the nonlinear $B \rightarrow I$ current calibration, seen in Figure 4.12b, to the reference current `MBI/IREF` which is driven to the FGC hardware. Since the compensation trims the field by only a small increment from the investigation in Chapter 5, expected of order 5×10^{-4} T, against at field levels between 1.8 T–2.0 T at extraction field, this increment samples only a locally linear segment of the calibration, so that its nonlinearity introduces no appreciable error and the loop need not model the conversion.

As the control loop corrects field back to a reference B_c^*, a reference is held for each cycle, and seeded on the first occurrence of that cycle, when a prediction has been made. This reference remains valid only while the field that the cycle is meant to produce has not itself been changed, and is discarded when a higher level parameter, like the programmed field or momentum for the main dipole circuit, is trimmed by an operator or another application. Operators and other applications routinely re-tune the machine making small trims in the parameters that sit above the parameter owned by this compensation scheme in the settings hierarchy, and particularly frequent when re-tuning the SFT slow extraction plateau, where changes can occur every few minutes until a satisfactory parameter set is found. Any such change is a deliberate re-tuning of the machine, and it is not a drift that the compensation loop should correct, and therefore the reference for that cycle must be discarded and re-established to avoid a race condition with the operator. Without this reset the loop would treat every deliberate change as a drift to be undone and work against the operator.

The interval of the cycle over which the correction is applied is a parameter of the loop, and can be set to any part of the cycle where beam is present in the machine. Since the model forecasts the field over the whole cycle, and the loop can trim any part of it, so compensation is applied only where it is needed, selected by cycle time, and the correction outside of that interval is left to whatever it was before. The linkrules of LSA then ensures that the correction applied in a given interval connects smoothly to the rest of the cycle.

A trim is committed to the setting only when the correction changes from one occurrence of a cycle to the next by more than a threshold of 2×10^{-5} T, which is set by the accuracy of the model, so that a cycle-to-cycle change smaller than this, carrying no information the model can resolve, leaves the previous correction in place. Within a correction that is committed, point-to-point noise is suppressed by a simple deadband filter that walks along the samples of the correction function and, whenever the next sample lies within a smaller threshold of order 1×10^{-5} T of the previous one, keeps the previous value and moves on to the next. This filter is kept deliberately simple because the shape of the residual for an arbitrary cycle is not known in advance, and a more elaborate filter would risk introducing lag or Gibbs ringing into the committed correction.

8.1.3 Timing System Constraints

The feed-forward correction is upper and lower bounded in timing, in when the information needed to predict a cycle becomes available, and when the correction must be committed for the FGC to play it. As briefly mentioned in Section 4.4.2, the controls timing system announces the next cycle and user through a middleware event 2.5 s before the cycle starts, and this forewarning is the maximum timeframe within which context gathering, prediction, correction computation,

and LSA must all complete. Through empirical studies during machine development, it was found that any correction trim containing points spaced at 50 Hz, completed less than 1.7 s–1.8 s before cycle start, only executes on the following cycle. Therefore while the forewarning is 2.5 s, the effective window for the whole chain is only about 700 ms before cycle start, and the loop must complete within that time. Furthermore, the 700 ms also includes the time for the timing events and the collection of the programmed and measured quantities from the middleware, so that the actual time available for the prediction and correction is even less. Within this budget the model inference dominates, taking of order 100 ms–200 ms **[TODO: measured inference time: mean and spread over N cycles]** . Finally, due to the transactionality of the LSA database, a change is not committed until the trim action returns, and therefore the total trim time includes latency with the LSA server, its database and writers, and the response time of the FGC, and the transaction time can be of order 100 ms, and is not something that can be predicted and reduced in advance.

Besides the timing of the control loop itself, the compensation scheme must take into account the machine mode of the cycle being corrected, since the programmed current in LSA is not always the current that the magnet will play, and is not always known at forewarning. Full economy, discussed in Section 4.4.6.3, is announced together with the cycle at forewarning and is accounted for at prediction time, as the programmed function is replaced by the idle current that the magnet will hold over the whole cycle. Dynamic economy is triggered asynchronously after injection, so that the reference current it produces, and with it the field response, can only be inferred once the cycle has already started, after the correction for the cycle has already been sent. Unlike a normal or full-economy cycle, the dynamic-economy current is not stored as a program in LSA that could be read back, since its shape is generated on the converter itself and is fixed both by an operator-changeable end-time setting in LSA and by the trigger time relative to injection, neither of which is known before the cycle is played, so that the played current must be reconstructed rather than read.

8.1.4 Control Loop Implementation

The previous subsections set out the individual mechanisms of the loop, and this section assembles the implementation into the per-cycle sequence of Figure 8.2, placed on the cycle timeline. The implementation of the loop is a chain of stages that read the timing and data streams of the control system and apply the correction as an LSA trim, and between cycles it carries just two pieces of information, a history of the recent cycles and one reference for each cycle.

The model developed in the previous chapter forecasts the field of a cycle from a context of the recent magnetic history, taking the excitation current, its time derivative, and the field as past covariates, and conditioning the forecast on the programmed current of the coming cycle as a known future covariate. In the loop it is evaluated on the programmed reference current rather than the measured current on which it was trained, since the two were shown in Figure 6.11 not to differ appreciably. Furthermore, the programmed current is free of measurement noise and, unlike the measured current, is known before the cycle is played rather than only afterwards when a cycle has finished playing. The field covariate over the past context is the measured B-Train field, which seeds the prediction on each cycle, and this measurement-seeded mode is closed-loop only in that the model state is corrected by the observation between cycles, the correction itself remaining feed-forward In contrast, the open-loop, autoregressive operation of Section 8.1.5 replaces the seed with the model's own previous prediction. The same preprocessing used in training, a low-pass filter followed by downsampling of inputs to the model rate of 50 Hz, is applied to these inputs at prediction time, and each cycle is predicted in a single shot over the coming cycle only. The resulting field prediction is used directly, at the model's native rate,

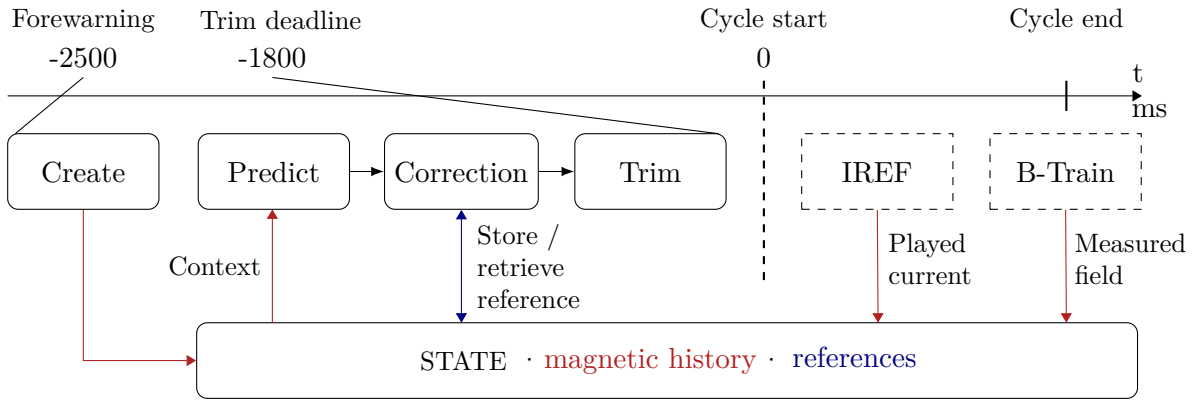


Fig. 8.2: Data flow of the feed-forward compensation loop over one cycle, with time in milliseconds relative to cycle start. Within the forewarning window the processing stages (rounded boxes) create the coming cycle, predict its field, compute the correction, and commit the trim to LSA by the deadline at −1800 ms, in time for the FGC to load the corrected function before the cycle is played. A persistent state carries information between cycles. The created cycle abstraction, the played reference current (IREF, including any economy-mode reconstruction) and the B-Train field measurement write the magnetic history (red), which the prediction reads back as context, while the correction stores and retrieves the reproducibility reference (blue). The state is logical, and in the implementation it is held locally within the individual stages and not as a shared global store.

as $\hat{B}_c$ in the residual of Equation 8.1, which is why the point-to-point noise in the committed correction is suppressed by the deadband of Section 8.1.2.

Assembled on timeline seen in Figure 8.2, most of the loop runs as a forward sequence within the forewarning window. At forewarning, 2500 ms before cycle start, the coming cycle and its machine mode are announced, whereupon the cycle is created from the programmed current announced for that mode, which in full economy is the idle current held over the whole cycle, and otherwise the LSA program for the cycle. For the magnetic history, the last measured field of the same cycle is bootstrapped as context, since at forewarning time the currently played cycle has not yet finished and the last measured field is therefore not available. The current derivative is determined from the programmed current, and the future field is then predicted from the history, and on the first occurrence of a cycle this prediction is saved as the reference and left uncorrected. On every later occurrence the stored reference is retrieved to form the residual of Equation 8.1, which is accumulated and committed as a trim by the deadline at −1800 ms. Within the window the order is therefore linear, running from create through predict, the saving or reuse of the reference, and correct, to trim.

This linear order however breaks due to tracking of the current, since the current the magnet plays is not the one announced at forewarning, as the trim alters it, and a dynamic economy mode, triggered only after cycle start, replaces it altogether with a ramp reconstructed by re-running the FGC's own generation algorithm described in Section 4.4.6.3. Only once the cycle has started is this played current known, at which point it is written back to the history, shown as STATE in the timeline. In addition to the current, only after the cycle ends is the measured field, drift-corrected before publication as described in Section 6.2.2, written back to the context state as ground truth to seed the following predictions. A cycle's record is therefore finalized only after it has played, and because the next cycle is predicted at its own forewarning while the

current one is still playing, each prediction rests on a played current and measured field that settle only once the preceding cycle has run. Should that cycle prove to be an economy cycle, the trim already sent for it is also not played, as the magnet plays the reconstructed ramp or idle current in place of the programmed current and the accumulating correction of Equation 8.1 absorbs the discrepancy in future corrections. A further consequence of measuring the field only after a cycle has played is that, when the following cycle is predicted, the most recent cycle in the history still carries the measured field of its previous occurrence, a limitation of the feed-forward scheme returned to in Chapter 11.

Throughout, the loop follows the UCAP paradigm of stateless stages, each performing a single step, so that the only state persisting between cycles is the history, held in the buffering stage, and the reference, held in the correction stage, with no shared global store.

The same stages run in two deployments that share this logic together with the same data subscriptions, timing events, and trimming mechanism. An operator-facing client was developed that runs the loop in the control room during machine development and test sessions, providing the monitoring surface and manual controls proposed in Section 5.4.6, and a headless UCAP node that runs it autonomously and interfaced through the same operator facing client. In either form the model is run for inference on a local GPU, while serving the predictions online through the Machine Learning Platform [40] at CERN was an alternative that was not implemented.

8.1.5 Autoregressive Operation

In the control loop implemented in the previous section, the predictive model is seeded each cycle by the B-Train field of the main dipoles. The higher-order magnet families have no equivalent online field measurement, so for them the model must carry its own state forward and predict each cycle from its own previous prediction, which defines open-loop, autoregressive operation. The same mode is applied to the main dipoles, as an additional autoregressive prediction mode, with the B-Train held back for validation, so that the open-loop tracking of the model can be assessed against a measured ground truth.

Autoregressive operation is initialized from a known magnetic state, taken from measured history where it is available or established by precycling the magnets to a known state before the sequence begins, so that the carried state starts from a defined point. This state is carried forward and is not rebuilt from the measured history at each cycle, and it must therefore be kept consistent with the current the magnet actually plays. That current is known only after cycle start, where it departs from the program announced at forewarning both through the trim applied to the cycle and, for a dynamic economy cycle, through the reconstructed economy current. In either case the carried state is returned to the value it held before the cycle, and the cycle is predicted again from the played current, so that the state advances from the current the magnet actually played. In closed-loop operation this correction is implicit, since the state is rebuilt each cycle from a history into which the played current and measured field have already been written, and in open loop it must be applied explicitly, because there is no measurement to rebuild from.

The form of the carried state depends on the model architecture: for the TFT model developed for the main dipoles it is the fully decoded field context over the recent cycles, for a recurrent network it is the hidden state, and for other architectures it may be a latent state or a memory token; the re-prediction returns whichever representation the model uses to its pre-cycle value before advancing it with the played current.

Whatever its form, the carried state is only ever the model's own estimate of the played field, never a measured ground truth: open-loop accuracy is therefore bounded by the single-shot quality of the model, and any error in the carried state persists and compounds across the

sequence with no measurement to correct it. For the higher-order magnet families, which have no online field measurement at all, the open-loop performance therefore rests entirely on the accuracy of their field model.

8.1.6 Extensibility to Other Magnet Families

The compensation scheme is not specific to the main dipoles, since its two points of contact with the machine, the timing system that announces the coming cycle and the settings hierarchy through which the correction is trimmed, are common to every magnet circuit. Only the B-train seeding is specific to the main dipoles, but the loop of prediction, correction and open-loop state tracking carries over to the higher-order magnet families without any change to its structure. The quantity that is controlled does change from one family to the next, being the field B for the dipoles and, for the quadrupole and higher-order circuits, their own field quantity, the integrated gradient $G\Delta L$ for the quadrupoles, which is inserted into the settings hierarchy as a dedicated correction program below the programmed quantity and trimmed feed-forward in the same way as for the dipoles, seen in Figure 8.1.

Beyond the controlled quantity, the one further requirement is initialization. Since the higher-order families have no online field measurement to seed the model, they must be compensated in the open-loop, autoregressive mode of Section 8.1.5, which must be started from a known magnetic state. That state can be reached with an operational supercycle configuration that saturates the relevant magnet circuit repeatedly, or with a dedicated precycle, after which the loop runs in open loop from the established state.

This extension is developed at the level of design only in parallel to the dipole deployment, since a field model for the higher-order magnets was not developed in this thesis and the scheme could therefore not be validated on them, so the modeling and validation of the quadrupole and higher-order families is left to follow-up work in Section 11.8 in discussion.

8.2 Beam Validation

The compensation scheme of Section 8.1 applies the magnetic field model of Chapter 7 as a feed-forward correction on the main dipole circuit. This section evaluates that model with beam against the validation goals of Section 5.4.7, on the loss of extracted spill reproducibility across supercycle changes established in Chapter 5. Reproducibility is assessed at beam level on the slow-extracted spill of the fixed-target cycle, the only cycle requiring active compensation during routine operation, recorded by the BCT and logged at 200 Hz.

8.2.1 Deployment on the Fixed Target Cycle

With the control loop assembled in Section 8.1.4, the scheme was deployed on the SFT cycle during dedicated machine development time, following the first validation goal of Section 5.4.7. It is the most frequently played cycle in the SPS and the most prevalent in the training dataset of the model developed in the previous chapter, and it is the only cycle that currently requires field compensation.

Sensitivity to the main dipole field of the fixed-target beams SFTPRO and SFTION arises from the chromatic slow extraction delivered with COSE over the 4.8 s flat top, whose fixed-optics condition a hysteretic dipole offset breaks, as set out in Chapter 5. The fast-extraction cycles are comparatively insensitive to an offset of the same magnitude, so hysteresis is not an operational concern for them, and only the SFT flat top was targeted for correction.

The offset on the SFT flat top is set by the branch of the hysteresis loop that the preceding cycles in the supercycle leave the main dipoles on. This dependence was observed directly only for the change between fixed-target-only and LHC-mixed supercycles, reported in Section 5.2.1. A preceding cycle reaching a higher field than the 400 GeV/c SFTPRO extraction, such as an LHC cycle at 450 GeV/c or the 440 GeV/c HiRadMat cycle of Section 4.2, is expected to shift the following SFT flat-top field. In contrast, in the absence of higher-energy cycles, the SFT cycle flat top slowly relaxes over successive occurrences, as the hysteresis loop settles onto a branch that reproduces the same flat-top field. The AWAKE cycle of Section 4.2 reaches the same 400 GeV/c field as SFTPRO and is expected to leave it unaffected.

Operationally the SFT flat top is disturbed only when the supercycle changes, since a static supercycle settles onto a fixed branch and reproduces its flat-top field. The validation in this chapter reproduced this by alternating between operational supercycles, the fixed-target-only supercycle and mixed supercycles that combine the SFT cycle with LHC, HiRadMat, and AWAKE cycles. The operational supercycles can be found in Section 4.3, where it can also be seen that the SFT cycle is present in every sequence. Constructing bespoke supercycle sequences carries a heavy overhead even in dedicated machine development, so the validation stayed with these operational supercycles and comparably infrequent changes, and did not rapidly switch between sequences or make arbitrary sequence combinations that operation does not use and that are consequently absent from the model training data. In all tests except one, the 24 s standard LHC filling cycle was replaced by the shorter 13.2 s LHCPILOT cycle, as the LHCPILOT excites the same hysteresis and eddy current decay, but allows for faster cycling and more efficient machine time. The operational LHC cycle was only kept in the fixed-target to LHC test of Section 8.2.3, while the extended operational test of Section 8.2.4 followed the normal operational supercycles without substitution, including switches between different LHC filling cycles.

In all tests, the correction was applied only over the flat top of the SFT cycle, between 4460 ms and 9260 ms of cycle time, with a chosen loop gain of 0.9, and only the SFT cycle was trimmed while every other cycle in the supercycle was left unchanged. Throughout the dedicated tests the beam was a single MTE core injection, from a low-intensity core at 5×10^{11} up to 5×10^{12} protons depending on the cycle already set up, the intensity capped by the TED dump in the transfer line to the North Area on which the beam was stopped rather than extracted to a target. Only the extended operational test of Section 8.2.4, which ran during standard physics operation with the beam sent to the target, used the nominal two injections at 1×10^{13}. At the reduced intensity of the dedicated tests the extraction is less stable and the spill less reproducible than in nominal operation, for reasons independent of the field correction, so the reproducibility reported for them is a conservative bound on the performance at full intensity.

8.2.2 Sensitivity to Field and Tune Perturbations

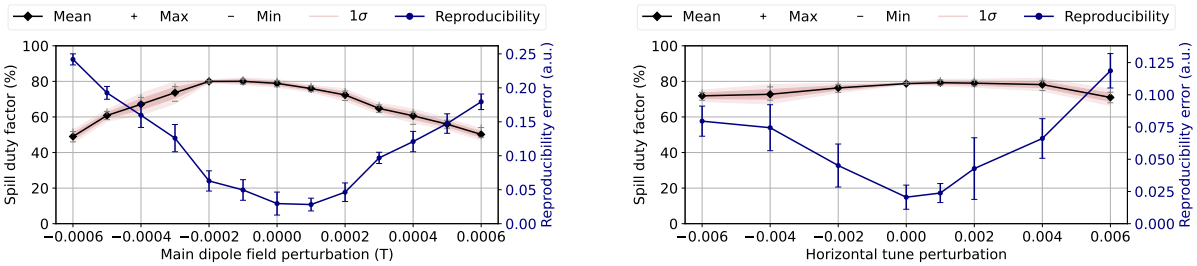
Before actively compensating the beam, the sensitivity of the SFT spill to the main dipole field and to the horizontal tune was characterized under controlled conditions, to separate the effect of each quantity. Each parameter was shifted in isolation by a constant offset held over the extraction plateau, with the spill recorded over 10 cycles at every offset. This isolates the response to a dipole shift and to a tune shift, and it reproduces under controlled conditions the operational spill degradation of Chapter 5.

Two scalar measures of the quality of the recorded spill are used in the results that follow. The first is the spill duty factor, a measure of slow-extraction quality based on the length of the extracted spill [76]. Introduced in this work, the second is the reproducibility error, defined as the root-mean-square difference between a chosen normalized BCT of a spill and that of a chosen

reference spill. The normalization is performed by dividing the BCT by the average intensity over the ramp before the extraction start, where beam losses are negligible.

Normalizing the BCT to the beam intensity before extraction enables the measure to be sensitive to only the shape of the extraction and not to the injected intensity, which varies from cycle to cycle. Being a difference of normalized traces, the reproducibility error is dimensionless and reported in arbitrary units. Such a beam or spill-based measure is necessary because the field itself cannot serve as a reproducibility reference, as the B-train integration drift of Section 6.2.1 exceeds the field differences to be resolved. At the fixed-target flat top the beam is debunched during the slow extraction, so the closed orbit cannot be measured, and the extracted intensity recorded by the BCT is used as the beam observable. The spill depends on all magnet families, including the main dipole field the compensation corrects, so it is not a measure of the dipole circuit alone, and the reproducibility error formed from it is retained as the most direct available beam measure of extraction reproducibility.

The two metrics also differ in their interpretation, as the duty factor is absolute and characterizes a single spill, while the reproducibility error is relative to a chosen reference and is therefore smallest when that reference is reproduced. An absolute reproducibility error would follow from taking the reference as an ideal linearly decaying BCT from the extraction start, representing a perfectly uniform spill, but the compensation targets reproducibility of the operational spill and not its absolute quality. For this perturbation study, the reference is therefore taken as the mean of the unperturbed baseline spills.



(a) Main dipole field offset by -6×10^{-4} T–6×10^{-4} T.

(b) Horizontal tune offset by -6×10^{-3}–6×10^{-3}.

Fig. 8.3: Sensitivity of the SFT slow extraction to the main dipole field and horizontal tune, quantified by the change in spill duty factor and reproducibility error. The reference BCT for the reproducibility error was taken as the average of the baseline spills.

A constant offset of -6×10^{-4} T–6×10^{-4} T was first applied to the main dipole field, beyond the hysteresis-induced range of up to 4×10^{-4} T established in Chapter 5. In Figure 8.3a the duty factor peaks at about 80 %, below the 100 % of a uniform spill, as the low beam intensity of this study and the operational reference not being retuned to the optimum both limit the achievable spill quality. Its maximum lies away from the reference, as expected of an absolute measure, while the reproducibility error is smallest at the reference and grows with the offset, reaching between 0.18 and 0.24 at the ends of the range. The underlying spill BCT in Figure 8.4 shifts earlier or later within the flat top with the sign of the offset, so a constant field offset alone reshapes the extraction, consistent with the behaviour observed in Figure 5.5 on the operational beam. Consistent with the reproducibility error, it can also be seen that the spill on the -6×10^{-4} T side of the reference is more distorted than on the 6×10^{-4} T side, which is reflected in the asymmetry of the reproducibility error.

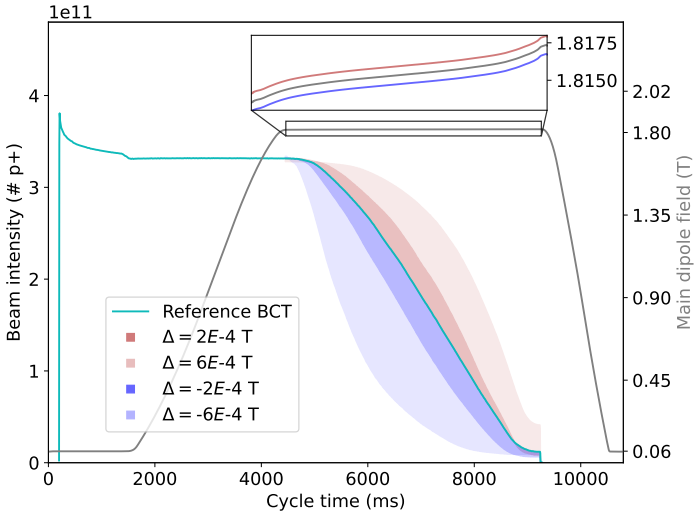


Fig. 8.4: Spill BCT changes as result of the main dipole field perturbations of Figure 8.3a, showing the sensitivity of the SFT slow extraction to the extraction dipole field.

Repeating the study on the horizontal tune, offset by -6×10^{-3}–6×10^{-3} over the same plateau and shown in Figure 8.3b, one can observe that the spill duty factor changes by at most 10 percentage points. In comparison, the reproducibility error grows to between 0.08 and 0.12 at the ends of the range, with the minimum at the reference as expected. The smaller magnitude of the reprodicibility error compared to the field perturbation is expected, as the tune corrections are very small in comparison to the main field perturbations. The compensation developed in this chapter acts on the main dipole field and does not touch tune, whose correction across operational supercycles is evaluated in the following section.

8.2.3 Dipole Field Compensation at the SFT Flat Top

Active compensation was first evaluated across supercycle changes between the LHC filling supercycle and the standard and fixed-target only sequence as the first target of the operational validation, shown in Figure 8.5. In this measurement, the reference was taken as the mean of three fixed-target spills recorded in the initial supercycle prior to the measurement. After each supercycle change the control loop was first left disabled to quantify the uncompensated spill, then enabled to correct it back to the reference. In the uncompensated intervals the reproducibility error rises with the supercycle change and, once compensation is enabled, returns to its baseline level, which demonstrates that both the correction and the control loop that applies it perform as intended.

The two measures respond differently across the changes in Figure 8.5. At the change from the initial supercycle to the SFT-only sequence the duty factor remains unchanged while the reproducibility error increases. When compensation is then disabled with the existing correction left in place and the sequence changes back to LHC filling, the duty factor drops by 15 percentage points while the reproducibility error rises to a level similar to the previous change. The two are only moderately anticorrelated across the study, with a coefficient of -0.56, since the reproducibility error responds to the full spill macrostructure while the duty factor reflects mainly the spill length, which does not always change with the macrostructure or the underlying field error. As the compensation targets field reproducibility and is assessed through the beam, the

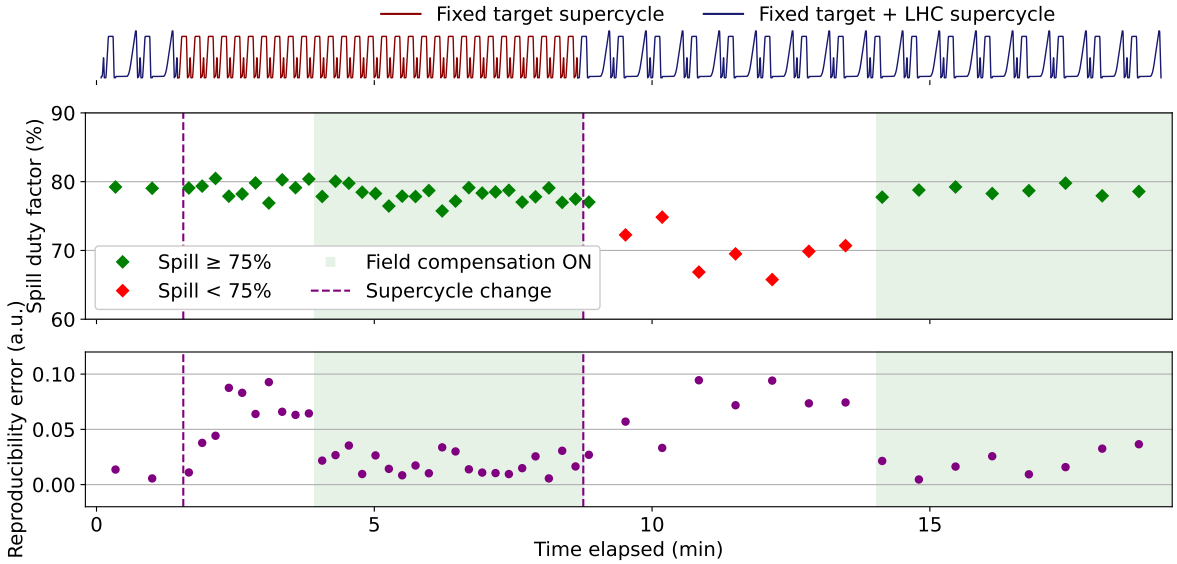


Fig. 8.5: Evolution of spill quality factor and BCT reproducibility error across supercycle changes from SFT with LHC filling, to SFT only, enabling compensation, and back to SFT with LHC filling without compensation, showing the effect of the compensation on the extraction field reproducibility.

reproducibility error is taken as the primary measure for the operational results that follow, and the spill duty factor is dropped altogether.

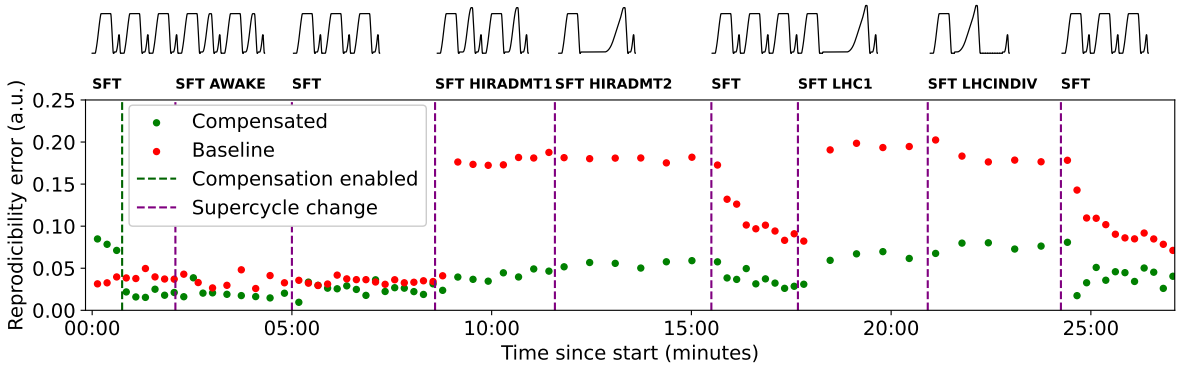


Fig. 8.6: Reproducibility error of the SFT spill for the compensated loop and the uncompensated baseline, played across a succession of operational supercycles. The supercycle composition is labeled along the top and the supercycle changes are marked.

To test the compensation across the full range of representative operational supercycles, the compensation was evaluated across the range of operational supercycles in a dedicated study, seeded from the B-Train on every cycle and run at an injection of a single MTE core of 5×10^{12} protons. Each supercycle was played first without compensation and then with the loop enabled, so that the compensated and uncompensated cases are available for the same sequences. Figure 8.6 shows the reproducibility error of the fixed-target spill through the study for both cases, as the supercycle composition is changed between the fixed-target only sequence and sequences mixing

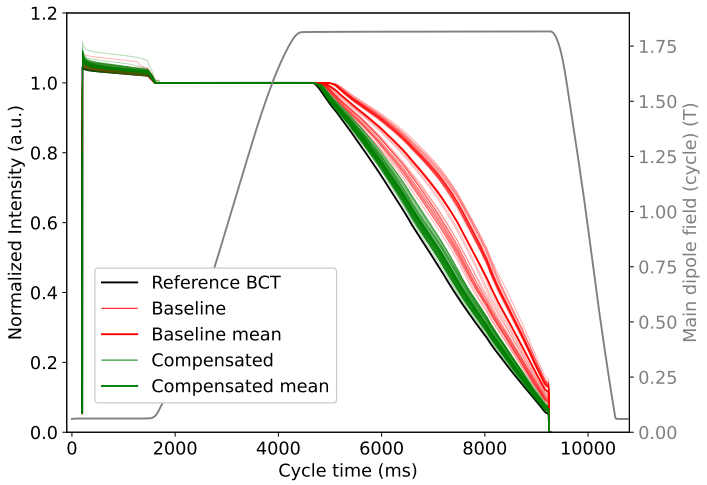


Fig. 8.7: Normalised spill BCT over the fixed target cycle for the compensated and uncompensated cases in Figure 8.6, each shown as individual traces with the mean within each category, against the reference BCT. The programmed main dipole field is overlaid.

the fixed-target cycle with the LHC, HiRadMat, and AWAKE cycles. In this measurement, the reference BCT was taken as the mean of nine fixed-target spills before the baseline study in the same machine development session.

In the uncompensated baseline the reproducibility error follows the supercycle, staying below 0.05 for the fixed-target only sequences and rising to between 0.15 and 0.20 for the sequences that interleave the higher-energy HiRadMat cycle, labelled by its HIRADMT timing user in Figure 8.6, and the LHC cycle, whose hysteresis perturbs the following fixed-target flat top. With the loop enabled, the reproducibility improves immediately from 0.07 to 0.02, and in 400 GeV/c operation performs similarly to the uncompensated baseline, which demonstrates that the compensation loop does not drift or degrade the extraction when the supercycle is static. The compensation maintains reproducibility up to 0.07 during HiRadMat sequences, and up to 0.08 on the LHC sequences. In this measurement, the compensated run began immediately after the baseline, before the field of the preceding sequences had fully relaxed, as shown in Figure 5.4, however the loop still compensated the extraction field close to the baseline. The fixed-target-only baseline, where no supercycle change perturbs the field, does not itself fall to zero but sets the reproducibility floor of the spill of this measurement. On the static fixed-target and AWAKE sequences the compensation reaches this floor, so there the residual is limited by the spill reproducibility and not by the correction.

The measured spill BCT for each fixed-target cycle is shown in Figure 8.7 for the compensated and uncompensated cases against the reference. The uncompensated spills spread away from the reference through the extraction, while the compensated spills form a tight bundle close to it, so the reduction in reproducibility error corresponds to the extraction shape being restored at the level of the individual spill. Operationally in the large control room displays, this spill correction is visually close enough to the reference that operators would not correct the spill, which we consider an operational success of the compensation scheme. The compensation is not always perfect, yet it improves on the uncompensated baseline consistently across the range of operational supercycles, which demonstrates the capability of the control loop and the model to improve field reproducibility in existing operation.

The field reproducibility reached by the loop was in this case measured from the B-Train over the compensated run, by computing the difference between the flat-top field of the reference cycles and each compensated cycle. The per-cycle field error has a median of 1.14×10^{-5} T and an interquartile range of 2.09×10^{-5} T, so the central half of cycles reproduces the field to a scale of order 1×10^{-5} T, that of the model accuracy. **[TODO: ref ch7 table]** The interdecile range of 7.03×10^{-5} T is 3.4 times the interquartile range, against 1.9 for a normal distribution, so the distribution is heavy-tailed and the poorly compensated cycles lie much further from the reference. The compensation reproduces the field close to the reference on the majority of cycles but not on every one, and the field reproducibility tolerance set in Section 5.4.5 is not reached, as anticipated from the model accuracy and the seeding B-Train drift. The higher field errors occur on the sequences with the long HiRadMat and LHC cycles, the same sequences that show the largest reproducibility error, with the performance degradation understood to be in part caused by the longer preceding cycles and inherent field drift in the seeding B-Train.

Because the loop corrects only the main dipole field, the large reduction it achieves in the supercycle-induced reproducibility error is suggestive of the main dipoles being the dominant hysteresis contribution to the operational spill. Had the quadrupole or higher-order circuits contributed appreciably to the perturbation, a correction acting on the dipole field alone could not have brought the spill back this close to the reference. The residual that remains on the higher-energy sequences is consistent with the dipole field-reproducibility limit of the model and the seeding B-Train, so it need not be assigned to an uncorrected higher-order contribution. This measurement does not isolate the individual circuit contributions and cannot exclude a small higher-order effect, yet it points to the main dipoles as the principal source of the hysteresis-driven spill error in current operation.

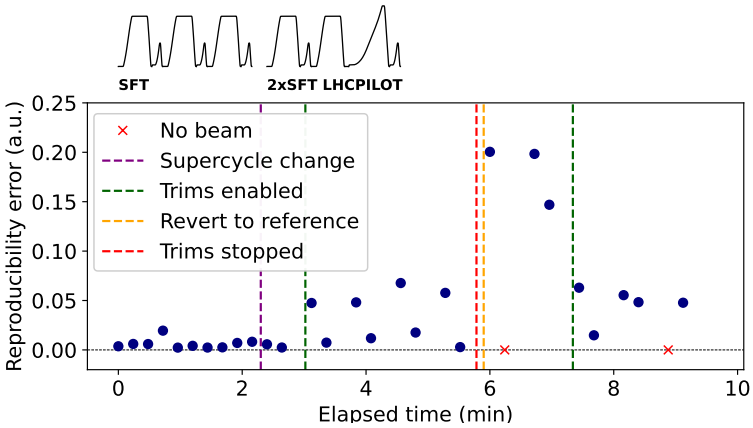


Fig. 8.8: Reproducibility error of the SFT spill across a change from a symmetric SFT supercycle to the asymmetric 2×SFT plus LHCPILOT supercycle. The compensation is enabled, then the trim is reverted to the uncompensated settings and the loop stopped to expose the uncompensated spill, before compensation is re-enabled. Cycles without beam are marked.

A further, smaller study evaluated the compensation on an asymmetric supercycle, a 2×SFT sequence with a single LHCPILOT cycle, shown in Figure 8.8, with the magnetic sequence on the top. The loop was seeded from the B-Train as in the operational study, and the reference BCT was taken as the mean of the preceding ten fixed-target spills. The measurement was made during beam commissioning at a reduced injection of 2.7×10^{12} protons. Such a supercycle is not

used in normal operation and represents a potential asymmetric sequence of the third validation goal. Compared to any other operational sequence, this one is a static-impossible case, since the two SFT cycles occupy different positions in the supercycle and sit on different branches of the hysteresis loop following the LHC cycle, so their flat-top fields differ even when the supercycle does not change. The reproducibility that static operation provides for a symmetric supercycle is therefore unavailable here, and only a per-cycle correction can bring both SFT cycles back to the reference.

The study followed the beam commissioning schedule and setup, so it was kept short and did not include a separate baseline run or a long stability measurement. The loop was instead enabled with the sequence change, reverted to the uncompensated settings to expose the perturbed spill, and re-enabled within the one sequence. Without compensation the two SFT positions show distinct reproducibility errors, of order 0.20 for the cycle following the LHCPILOT and 0.15 for the second, as the field relaxes toward its setpoint over the held SFT cycles. With the loop enabled the reproducibility error is reduced to of order 0.07 on both, and below 0.02 on most of the second occurrence, short of the floor reached on the operational supercycles but well below the uncompensated case. The correction is therefore not perfect on this sequence, yet it restores the majority of the reproducibility that static operation cannot provide.

8.2.4 Extended Operational Test

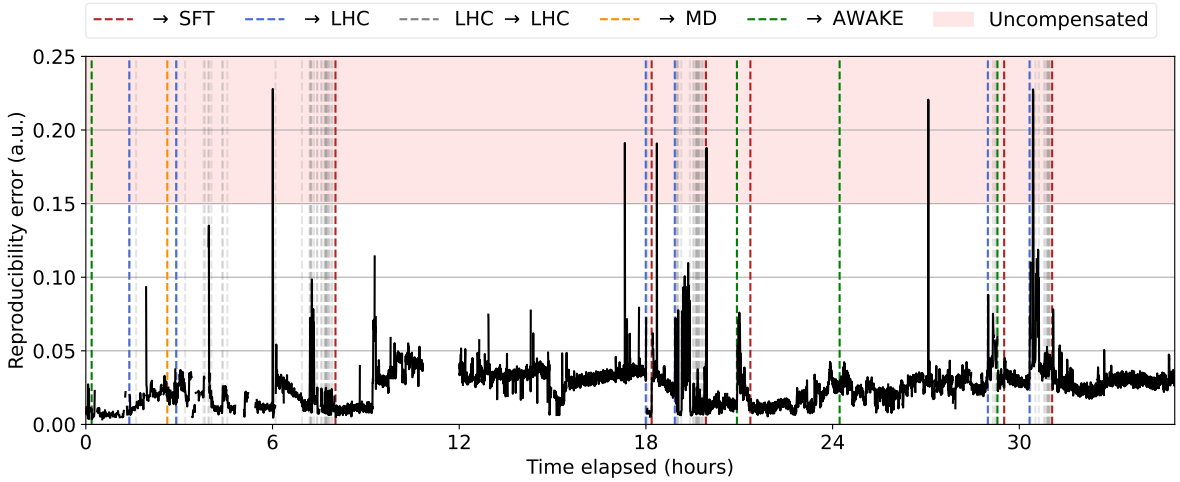


Fig. 8.9: Reproducibility error of the fixed-target spill over a 35 h operational test, with the loop running autonomously through the machine's operational supercycles. The shaded red band marks the uncompensated level reached across supercycle changes in Section 8.2.3, and the magnetic sequence changes are marked by type.

The dedicated studies of the previous Section validated the compensation over controlled runs during limited duration in machine development tests. To validate the control loop and model under normal beam operation over a longer period, the loop was left running as an extended operational test during normal fixed-target physics, with the reference taken at the start of the run and kept during the entire test. It ran for 35 h in its B-Train-seeded mode, applied to the flat top of fixed-target cycle at every occurrence, while the machine ran its own operational and occasional exotic supercycles. Over this run the supercycle changed many times between fixed-target, LHC filling, and non-standard MD, and AWAKE sequences used by operators. Furthermore, the LHC supercycles alternated between different filling schemes and pilot beams,

including the 4 injection LHC1, 1 injection LHCPILOT, and low intensity LHCINDIV cycles, all marked in gray in Figure 8.9.

Of the 6764 fixed-target cycles played over the run, 6089 carried beam. Of these, 98.31 % had a reproducibility error below 0.05 and 99.74 % below 0.1, with only seven above 0.15. This lies well below the uncompensated level of 0.15–0.20 reached across supercycle changes in Section 8.2.3 and shown as the shaded red band. It did not rise into this band at the supercycle changes, and operators adjusted neither the momentum nor the tune to recover the spill at any change during the run which would normally be required without the active compensation. The compensation therefore reproduced the extraction across the run without manual intervention, and remained transparent to the operators, who continued to work with the same high-level settings. That operators further saw no need to intervene even on the cycles where the spill was not perfect indicates the reproducibility reached was operationally acceptable, since in routine running an inadequate spill would prompt a correction.

Full economy was announced 253 times over the run, placing the fixed-target cycle at idle current with no beam. At the time of this test the full-economy state was published only after the correction had been sent, and a later timing update made it available at forewarning as described in Section 8.1.3. The loop therefore tracked it by reconstructing the played current after cycle start, same as for a dynamic-economy cycle and differing only in the shape of the reconstructed waveform, so that the following fixed-target spills remained reproducible. No dynamic economy cycle was triggered during the run.

A small number of isolated spill errors rise above the shaded red band, seen as narrow spikes in Figure 8.9. These are attributed to occasional model mispredictions or to machine faults such as 800 MHz RF trips or beam dumps during the spill, with the loop recovering on the following cycles in each case. No significant beam loss attributable to the compensation loop was identified over the run.

8.2.5 Autoregressive Operation

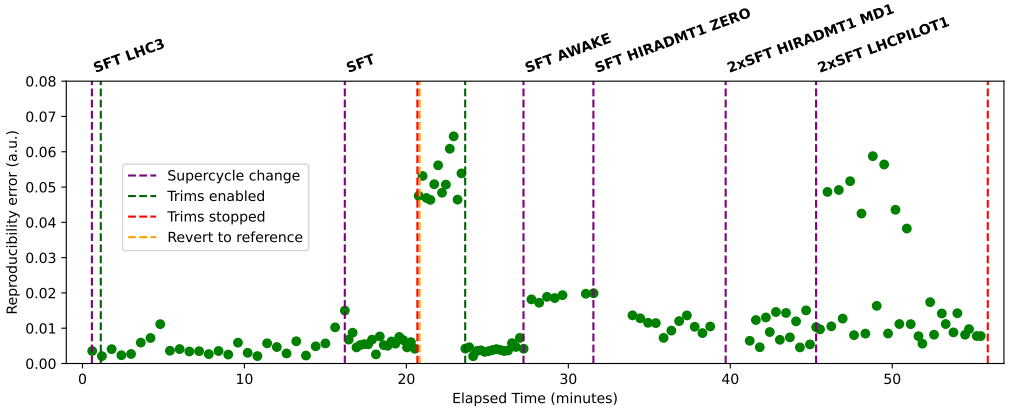


Fig. 8.10: Reproducibility error of the SFT spill during an autoregressive compensation test across a range of unconventional supercycles assembled during the session. The supercycle composition is labeled along the top and the control events are marked. The labels denote these non-standard compositions and differ from the standard supercycles of the same shorthand in Figure 8.6.

All results so far were obtained with the loop seeded from the B-Train on each cycle. The higher-order magnet families have no equivalent online measurement and must be compensated in the open-loop, autoregressive mode of Section 8.1.5, in which the model carries its own predicted

field forward as state. To assess this mode on the main dipoles, the loop was seeded from the B-Train at the first cycle and then run autoregressively, carrying its own predicted field forward, with the B-Train withheld from the following predictions and kept only to evaluate the result.

This run was opportunistic and tested the autoregressive loop and the compensation performance at once, across a range of unconventional supercycles assembled during the session and run at an injection of a single MTE core of 5×10^{12} protons, shown in Figure 8.10. These supercycles are orderings of operational cycles not run in normal operation, so that the individual cycles are represented in the training data while only their ordering is new. No separate uncompensated baseline was recorded, since the same sequences could not be replayed within the available time, so the run is read against the reproducibility levels of Section 8.2.3. Because the reference was taken from a fixed-target cycle in a supercycle that also included an LHC cycle, the reproducibility error on the fixed-target-only sequences is measured against a different reference than in the seeded study of Section 8.2.3, so its absolute level is not directly comparable and reads lower on those sequences.

With the loop enabled the reproducibility error stays below 0.02 on most of the sequences, so the open-loop prediction tracks the field well enough to compensate the extraction on the standard sequences. On the asymmetric 2×SFT sequences the error rises to between 0.05 and 0.06, the weakest performance of the run, consistent with the open-loop accuracy being bounded by the single-shot quality of the model with no measurement to correct the carried state. This sequence was also the hardest for the B-Train-seeded loop in Figure 8.8, and it stays well below the uncompensated level seen there. The autoregressive mode therefore compensates measurably but less well than the B-Train-seeded loop, as expected of open-loop operation.

This run is a limited demonstration, without a baseline and on sequences assembled during the machine development session. Still, it demonstrates the compensation operating in the open-loop mode that the higher-order magnet families require, which have no online field measurement to correct the carried state between cycles and must run from a known initial state established by precycling.

8.2.6 Scope of Validated Results

The preceding sections show the compensation reproducing the fixed-target field at extraction across the operational supercycles. In Section 8.2.3 it reduces the reproducibility error from the uncompensated 0.15–0.20 toward the fixed-target floor, reaching it on most sequences, and reproduces the extraction spill close to the reference to a degree acceptable in operation. The loop then ran autonomously over the extended test of Section 8.2.4 without operator intervention. This meets the first validation goal of Section 5.4.7, the compensation of the fixed-target field across the operational proton sequences, even though the field reproducibility tolerance of Section 5.4.5 is not reached.

The third validation goal, compensation across supercycles not seen in normal operation, was partially met. The asymmetric 2×SFT+LHC supercycle of Section 8.2.3 was compensated with the loop, and further non-standard orderings were exercised in the autoregressive run of Section 8.2.5, though neither was a systematic study with a matched baseline. Current operation plays a small set of discrete supercycles, and the validation spanned these operational sequences, so the range of magnetic conditions it covers is bounded but coincides with the conditions operation produces. Arbitrary supercycle changes do not occur in this mode and would arise only under more dynamic beam scheduling, so the bespoke sequences needed to test them, noted in Section 8.2.1, were not constructed, and the goal was met only for the non-standard sequences that arose in these studies.

The compensation was demonstrated only on the main dipole circuit and only over the fixed-target flat top, the one circuit and cycle phase that require correction in current operation, and the higher-order magnet families and other cycles were left uncompensated as no field model was developed for them. The validation used proton beam, and the fixed-target ion cycle SFTION was not tested, although it undergoes the same chromatic slow extraction and is susceptible to the same dipole hysteresis, so the compensation is expected to apply to it. Ion operation places lower demands on spill quality and more often leaves the spill uncorrected, and ion running is limited to one or two months a year and cannot run alongside the proton beam sent to the North Area targets, so no dedicated session was available. Within this scope the field reproducibility is set by the accuracy of the model and the drift of the seeding B-Train, so the field is reproduced on the majority of cycles but not on every one. The dedicated studies further ran at reduced intensity, at which the extraction is inherently less reproducible, so their reproducibility is a conservative bound on the performance at nominal intensity. The model was trained on recorded operational data, so it is specialized to the sequences that operation plays, and generalizes across orderings of the operational cycles as the autoregressive run of Section 8.2.5 shows, while a sequence that operation rarely plays, such as the flattened precycle, is underrepresented in the training data and correspondingly less well compensated.

The second validation goal, compensation at injection as the prerequisite for removing the MD1 precycle, was not achieved with the fitted model and control loop. Compensating the injection field requires correcting both the static hysteresis offset and the decaying eddy-current transient, both of which the single main-dipole model was intended to predict, and neither held to the required precision at injection. Reproduced to the same order 1×10^{-5} T as at extraction, the static offset corresponds at the lower injection rigidity to 0.55 mm of closed orbit for a coasting beam, as seen in Table 5.2. An interdecile field spread of 7.03×10^{-5} T reached at extraction would therefore give several millimeters of closed orbit distortion at injection, well beyond the 1 mm beam-position tolerance of Section 5.4.7. The eddy-current transient could not be compensated with the same model at injection, and this limitation is taken up in Chapter 9. The injection correction was therefore inconsistent from cycle to cycle, and the partial injection data are noisy and not reported here.

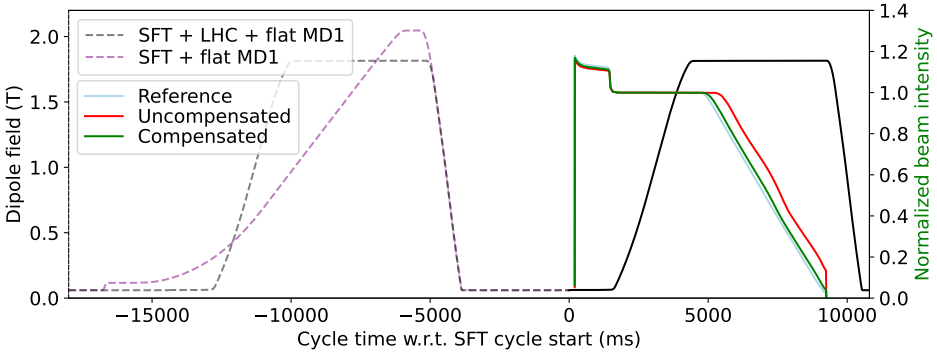


Fig. 8.11: Single-shot reproducibility of the SFT cycle with the preceding precycle flattened in normal physics supercycle and LHC supercycle. BCT traces are shown for the reference measurement, first occurrence after a sequence change, and the following occurrence after compensation is applied.

The configuration in which injection compensation was attempted is the flattened precycle, since the precycle currently fixes the injection field and removing it is the prerequisite for compensating there. Its effect at extraction was checked by flattening the precycle before the

fixed-target cycle to its idle current on the dipoles alone, so that only the dipole contribution is changed. Beam reached extraction in this configuration only because the injection had separately been setup to transmit beam, through measures described in Chapters 9 and 10, so the result that follows is an extraction measurement on top of that setup and not a demonstration of injection compensation. Only a few shots were taken in this brief study, in a fixed-target physics supercycle and in an LHC-filling supercycle. Figure 8.11 shows the fixed-target spill for a reference cycle, the first uncompensated occurrence after a supercycle change, and the following compensated occurrence, with the dipole field of the preceding supercycles and of the fixed-target cycle drawn for context. The uncompensated spill deviates from the reference similar to after a sequence change seen in Figure 8.7, and the compensated occurrence is brought back close to it, so the extraction compensation carries over to the flattened-precycle configuration, as an indicative check and not a systematic validation.

Flattening the precycle was aimed at injection compensation, which nevertheless remained out of reach on two counts, each taken up in the following chapters. The eddy-current transient that the dipole model could not reproduce at injection is investigated in Chapter 9, and the removal of the MD1 precycle it requires, together with the beam stability this exposes, is taken up in Chapter 10.

Chapter 9

Eddy Current Decay: A Parallel Case Study

The feed-forward compensation corrects the realized field of a cycle with a single learned model of the operator $\mathcal{F}$, developed in Chapter 7 to represent the static remanence of hysteresis and the dynamic memory of eddy currents jointly, and fit to the magnetic field measurements of Chapter 6. At extraction this single model reproduced the field on the majority of cycles, as shown in Chapter 8. Compensation at injection, the second validation goal of Section 5.4.7, was not reached. There the correction behaved inconsistently from cycle to cycle, as the single model could not reproduce the injection field to the required precision. This chapter investigates the eddy current decay as the suspected origin, characterising it in the magnetic measurements and as seen by the beam, and finds the decay on the beam to differ from that recorded by the measurements to which the model is fit. A complementary model is developed in response, driven directly by the rate of change of the field and fit to beam observables, and folded into the same feed-forward loop as an additive correction on top of the static hysteresis correction of Equation 5.4.

9.1 Observations from Measurements

To understand why the attempted compensation behaved inconsistently, the eddy current decay is first characterized as it appears in the magnetic field measurements, and then compared against the decay seen by the beam. The eddy current decay can be measured in the B-Train and laboratory pickup coil, described in Sections 3.4.3 and 3.3.1, respectively, following a constant excitation plateau after a ramp, as the induced eddy currents dissipate and the measured field decays to a steady state. The field on such a plateau is a superposition of independently decaying eddy current modes, following the single-loop eddy current model of Section 2.6.2,

$$B_e(t) = \sum_k A_k \, e^{-t/\tau_k}, \tag{9.1}$$

where A_k and τ_k are the amplitude and time constant of the k-th eddy current mode. For fitting the measured field, this stationary decay is truncated to K exponential terms and extended with a constant offset and a linear term,

$$B_e(t) = \sum_{k=1}^{K} c_k \, e^{-t/\tau_k} + d + s\,t, \tag{9.2}$$

where c_k is the fitted amplitude of the k-th mode, d is the settled field, and s is a linear slope that absorbs the slow integration drift discussed in Chapter 6, which is not part of the eddy current physics.

Two independent measurements of the same magnet type were fit in this way. The first is the in-situ B-Train, where three data slices were fit jointly, with each slice from a different acquisition, and therefore different integrator offset and slope. Two flat-bottom decays following

the ramp-down of an SFTPRO cycle, with 3.6 s and 12 s of data, and the flat-top decay of an LHCPILOT with 500 ms of data, each at 1 kHz were taken from logging, and with the initial power converter ringing after each ramp excluded.

The second measurement is the laboratory fluxmeter coil using the same measured dataset shown in Figure 6.1, and using the stationary 3.6 s plateau after the SHiP and LHC cycles in the dataset to fit the decay. Rather than integrating the full cycle with a precomputed voltage offset, resulting in the large drift seen in Figure 6.2b, the field was reconstructed by integrating the coil voltage only starting at the end of the ramp-down, and estimating the voltage offset from the last 500 ms of the plateau. In total, 20 SHiP and 5 LHC slices were used, with 3 s of data at 100 kHz each. For this experiment, the aligned and averaged fluxmeters were used for the integration to reduce impact of the integration drift. The laboratory Hall probe reconstruction were not used here, since the thermal drift identified in Section 6.1.4 remains significant on the plateaus, and since the Hall probes sample the field at discrete longitudinal positions and may not capture a decay localized elsewhere along the magnet.

For each measurement, models with one, two, or three exponential terms were fitted to the data slices using a least-squares objective. Because the time constants are a property of the conducting geometry of the magnet and its eddy current paths, they are taken to be independent of which ramp produced the transient, and are shared across all slices, while the amplitudes c_k, offset d, and slope s are fitted individually to each slice, since these depend on the size and direction of the preceding current step and on the integrator state of each acquisition. Fitting up to three exponential terms avoids assuming a specific number of eddy current modes, and more than three would not be supported by the available data. Slices measured under different conditions were used for each bench so that the shared time constants represent decays of differing magnitude, which also mitigates the non-convexity of fitting a multi-exponential system.

Tab. 9.1: Time constants of the eddy current decay fit to the constant-current field relaxation of Equation 9.2, for the in-situ B-Train and the laboratory pickup coil. The fit cost is weighted per measurement and is comparable only within a measurement, not between them.

Measurement	K	τ_1 [s]	τ_2 [s]	τ_3 [s]	Fit cost
B-Train	1	0.81			1.30×10^{-3}
	2	0.37	1.82		1.10×10^{-3}
	3	0.18	0.46	1.91	1.10×10^{-3}
Laboratory coil	1	0.40			1.13×10^{-5}
	2	0.09	0.44		3.73×10^{-6}
	3	0.05	0.18	0.46	3.64×10^{-6}

The exponential fits and their fit costs are reported in Table 9.1. Going from one to two exponentials reduces the cost substantially in both cases, while the third term reduces it negligibly and its time constants collapse toward one another, indicating that the data do not support a third distinct mode. The two-exponential model is therefore selected. The resulting time constants are $\tau_1 \approx 0.37$ s and $\tau_2 \approx 1.8$ s for the B-Train, and $\tau_1 \approx 0.09$ s and $\tau_2 \approx 0.44$ s for the laboratory coil. The time constant fits with one exponential term differs by a factor two, whereas with the chosen two-exponential fit, the time constants differ by more than a factor four. Even given the uncertainty of the magnetic measurements, these fits suggest that the B-Train system with MBA and MBB magnets connected in series, and the lab bench with one MBB magnet,

show significantly different decay. For consistency, the orbit decays at stationary current was then measured to compare against field decay in the magnetic measurements.

Tab. 9.2: Injection parameters of the SPS cycles measured in this chapter, described per cycle in Section 4.2. The injection time is measured from the start of the cycle. The SFTION and LHCION cycles here accelerate oxygen ions at 13 GeV/c, whereas the same cycles for fully stripped lead ions (Pb^{82+}) normally inject at 17 GeV/c.

Cycle	Injection [ms]	Current [A]	Momentum [GeV/c]
SFTPRO	200	161	14
SFTION / LHCION	725	149	13
MD2 / LHC4 / LHCMD2	1015	302	26

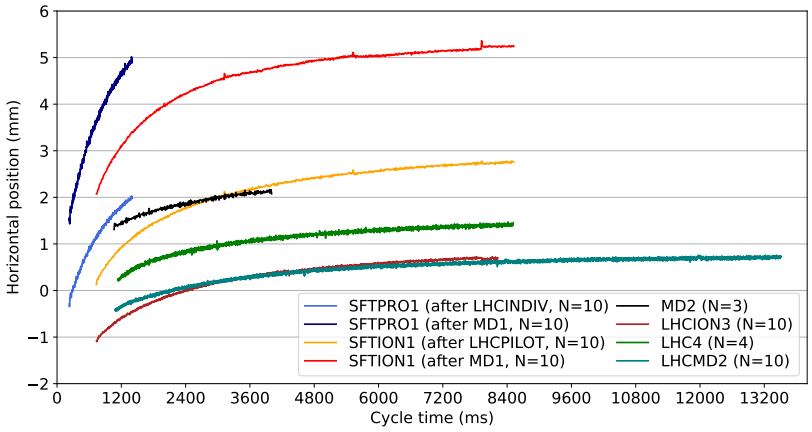


Fig. 9.1: Orbit decay measured from the BPMs across constant current injection plateaus of various SPS cycles, offset by injection time. In each measurement, the initial 10 ms–70 ms are removed to avoid showing the settling of RF loops immediately following injection, and averaged over 3-10 measurements. The injection time, current, and energy of each cycle are given in Table 9.2.

Figure 9.1 shows the beam orbit under controlled cycle measurements where the injection plateau is held at constant current, rather than containing the small ramps used in operation to compensate the eddy current decay. For each cycle, the radial loop is not engaged while the beam is being injected and is enabled only at the start of the ramp, and the RF frequency at injection is computed from the constant programmed field in LSA for the duration of the injection plateau rather than tracking the B-Train field measurement, and the orbit drift follows Equation 2.13. The measurement shows different cycles, with SFTPRO, MD2, LHC4 and LHCMD2 proton cycles, and SFTION and LHCION oxygen ion cycles, following different preceding cycles and therefore different eddy current histories. In this measurement, the LHCMD2 cycle is used for the scrubbing run, where high-intensity proton beam is kept at injection energy for up to 20 seconds, however as all measurements are averaged over 3-10 acquisitions, depending on the cycle, the shortest acquisition limits the maximum plateau duration to 12 s. Furthermore, the initial orbit measurements are removed to avoid showing the settling immediately following injection, when the beam is captured by the RF phase and synchronization loops and by the energy matching of the injected beam, which together can shift the beam by several millimetres at the start of

the plateau and mask the eddy current decay. Therefore, the absolute orbit displacement is not representative of the absolute field error, but only the relative decay of the orbit over the plateau.

The measured orbit can be seen decaying the fastest on the SFTPRO cycle, which injects only 200 ms after cycle start, which in turn is 150 ms–200 ms after the preceding cycle ramp-down ends. The SFTPRO cycle following the 450 GeV/c LHCINDIV cycle can be seen 3.5 mm over the 1200 ms plateau, whereas the same cycle following the operationally standard precycle MD1 drifts 2.2 mm over the same time. Crucially, the SFTION, LHCION, LHC4 and LHCMD2 cycles can be seen drifting over more than 7.5 s, and counting the time before injection until previous ramp-down, this becomes 8.5 s. The longest time constant in Table 9.1 is fitted as 1.82 s, so that at $3\tau = 5.46$ s roughly 95 % of the decay should be complete. However the decays shown in Figure 9.1 suggest that the decays measured on the beam decay for longer than that, making even the longer eddy current decays measured on the B-Train unlikely to explain the beam behaviour. As neither the RF frequency is changing, nor the excitation current of the main dipoles, only a decaying dipolar field can explain the drift in radial position.

The orbit drift measurements suggest that the beam orbit decay does not follow the field decay seen on the magnetic measurements from either the laboratory bench or the B-Train. The two differ because they are observed in different environments. The B-Train measures two spare magnets connected in series with the main ring, and the laboratory bench measures a single magnet in isolation, whereas the beam samples the integrated dipole field around the whole ring. Along that path the beam is deflected not only by the main dipole field but also by the eddy currents induced in the beam pipe and in the other conductive structures of the tunnel, whose current paths are bounded by grounding points spaced approximately every half FODO cell, far longer than the isolated magnets of the measurement benches. The beam therefore integrates a larger and different set of conductors than any magnet-mounted probe can measure. Consequently, the field compensation strategy proposed in Chapter 5, using the fitted field model of Chapter 7 and partially validated in Chapter 8, cannot be expected to fully correct the eddy current decay, since the model is fit to the magnetic measurements and not to the beam.

Although this dipolar field effect cannot be captured in the field measurement of an individual dipole, it can still be compensated through the main bending dipoles by a feed-forward correction of the same form as the hysteresis compensation. The correction is added as a further term to the field correction of Equation 5.4, on top of the static remanence correction predicted by the main field model. Because that field model is fit to the B-Train, its prediction already includes the portion of the eddy current decay visible in the magnet, so this portion must be fitted and tracked so that it can be removed from the field prediction, ensuring that combining the two corrections does not double-correct the magnet-visible decay. The remainder of this chapter develops the model for this term, driven by the rate of change of the field and fit to the beam, since the beam is the only measurement that sees the effect to be corrected.

9.2 Eddy Current Models

The stationary fits of the previous section and Equation 9.1 describe only the free relaxation of the eddy current field, the homogeneous decay of the L–R loop of Section 2.6.1 in Equation 2.27. They yield the decay time constants τ_k, but carry no information about how the eddy current field was accumulated by the excitation preceding the plateau. Operational cycles are rarely stationary, however, and even injection plateaus contain small ramps, so predicting the eddy current field at an arbitrary point in a cycle requires solving the eddy current equation for an arbitrary time-varying excitation.

Under such an excitation the loop is driven through its mutual coupling to the magnet circuit, and by Faraday's law of induction the source term added to Equation 2.27 is proportional to the rate of change of the linked flux, hence to the field rate $\dot{B}$ [101]. The current rate $\dot{I}$, in terms of which the eddy current amplitude of Section 2.6.1 was written, is however not always proportional to $\dot{B}$ due to the saturation of the magnet. This can be seen in the calibration curve in Figure 4.12b, where the nonlinearity starts around 4000 A for the main dipoles. Therefore, $\dot{B}$ is the physically correct driving term. Expressing the eddy current through the field B_e it produces in the aperture following Equation 2.29, and absorbing the loop inductance, resistance, and geometry into the decay time constant $\tau = L/R$ of Equation 2.28 and a dimensionless coupling c, the driven loop obeys the observable first-order linear ordinary differential equation

$$\dot{B}_e(t) = -\frac{1}{\tau}\, B_e(t) - c\, \dot{B}(t), \tag{9.3}$$

where $B(t)$ is the main dipole field produced by the excitation, and the response is characterized by the two constants τ and c.

Equation 9.3 is linear, first order, and time invariant, so it is solved exactly by the integrating factor $e^{t/\tau}$. For a loop at rest in the distant past, the eddy current field is the convolution of the driving rate with a one-sided exponentially decaying kernel,

$$B_e(t) = -c \int_{-\infty}^{t} e^{-(t-t')/\tau}\, \dot{B}(t')\, \mathrm{d}t'. \tag{9.4}$$

The eddy current field at any instant is therefore the entire history of the driving rate, weighted by an exponentially fading memory of duration τ.

A real conductor supports many such loops, and the beam pipe together with the other conductive structures of the ring add further modes of differing geometry, as described in Section 2.6.2. By the linearity of Equation 9.3, the total eddy current field is the superposition of the individual mode contributions,

$$B_e(t) = -\sum_k c_k \int_{-\infty}^{t} e^{-(t-t')/\tau_k}\, \dot{B}(t')\, \mathrm{d}t', \tag{9.5}$$

each mode contributing its own time constant τ_k and coupling c_k. The stationary decay of Equation 9.1 is the special case of Equation 9.5 for a constant excitation, so the two share the same time constants τ_k, while the couplings c_k, which set how strongly each mode accumulates from the driving history, do not appear in a stationary decay and must be determined separately. Each coupling satisfies $c_k \leq 0$, since by Lenz's law the eddy current field opposes the change that induces it, following Section 2.6.1.

For a driving field sampled at a constant spacing Δt, the convolution of Equation 9.5 becomes the Riemann sum

$$B_e(t) = -\sum_k c_k \sum_{t' \leq t} e^{-(t-t')/\tau_k}\, \dot{B}(t')\, \Delta t, \tag{9.6}$$

where the inner sum runs over the sampled times t' up to the current time t, and Δt is the quadrature weight that turns the sampled rate $\dot{B}(t')$ into the field increment contributed over one step. Because the kernel decays, the memory of the system is finite, and the inner sum can be truncated to a history window of a few times the longest time constant τ_k without appreciable loss of accuracy.

Evaluating Equation 9.6 at a single time t costs of order N operations for a history of N samples. Predicting the eddy current field across a whole cycle, as the feed-forward correction

of Section 9.5 requires, evaluates it at every sampled time, which by direct summation costs of $\mathcal{O}(N^2)$ and is computationally unpractical for the cycle histories of order 10^4 to 10^5 samples at 1 kHz, equivalent to 10 s–100 s of history. Because the kernel is a decaying exponential, the convolution over the whole cycle can instead be evaluated in the frequency domain by fast Fourier transform, at a cost of order $\mathcal{O}(N \log N)$.

The time constants τ_k are available from the stationary and orbit decays of Section 9.1, so the remaining unknowns are the couplings c_k, which determine how the eddy current field accumulates from the excitation history. The following sections describes how this accumulated field is measured on the beam through the first-turn trajectory, and converted from a radial displacement to a field quantity, so that the couplings can be fitted.

9.3 Measurements of Eddy Current Decay

As described in Section 9.1, the RF loops shift the orbit by several millimetres at injection independently of the field error, so the closed orbit does not give the field error directly. Instead of using the turn-by-turn orbit acquisitions, the first-turn trajectory from the same BPMs can be used to measure the radial offset of the beam, before the RF kicks for the first time. When the beam is not yet captured by RF, the radial offset is related to the field error by Equation 2.12, which is independent of the transition energy, and therefore of the slip-factor in Equation 2.13.

In practice, the beam is injected in the middle of LSS1, and the RF is installed in LSS3, which results in only one third of the SPS ring, or 36 BPMs out of 108 being used for the first-turn trajectory measurement. The first-turn radial offset is taken as the average of the 36 BPMs in this section. While the number of turns acquired could be extended by not enabling the RF cavities at injection, the BPMs require a bunched beam to accurately read out the radial position, so the first-turn trajectory is the only available measurement.

While the first-turn measurement is free of RF manipulation, the absolute radial offset is still not representative of the eddy current field alone, since it also carries the static hysteresis offset left by the preceding cycle and any energy error of the beam injected from the PS. Only the relative change of the offset between occurrences of the same cycle at different eddy current states is therefore used. This state is varied by changing the time between the measured cycle and its preceding cycle, by inserting additional ZERO cycles to elongate the supercycle in steps of one basic period of 1200 ms as described in Section 4.3. To reduce any impact of variations in injected energy, the cycle preceding the beam-accelerating cycle in the PS is held constant as the supercycle is elongated, though minor residual variation in the injected energy remains a limitation of the measurement. As mitigation of the latter, each point is averaged over approximately ten acquisitions.

Figure 9.2 shows the measured trajectories and their relative change with respect to the first occurrence of each cycle. The left panel shows the absolute first-turn radial offset against the time between the measured cycle and the end of the preceding cycle, for each cycle and preceding history. The absolute offsets differ between cycles because of the hysteresis and energy contributions described above, so only their relative change with elapsed time can be used. The right panel shows the offset relative to the first measured occurrence, $\Delta r = r(n) - r(0)$, converted to the equivalent field B_e through Equation 2.12, giving the accumulated eddy current field as a function of elapsed time, which is the quantity the couplings c_k from Equation 9.5 are fitted against. The measured set comprises four SFTPRO, two SFTION, and one MD cycle, each following a different preceding history. In addition to the trajectory, the field change $\dot{B}$ of the cycle and of the preceding 60 s is acquired from logging, which provides the driving history over which the convolution model is evaluated. As can be seen, the measured curves are not smooth,

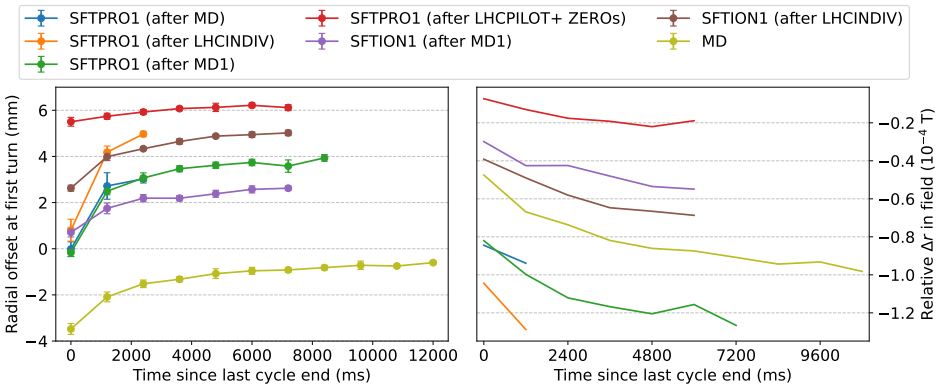


Fig. 9.2: First-turn radial trajectory of different SPS cycles, measured before the RF is enabled, against the time between the measured cycle and the end of its preceding cycle. The left panel shows the absolute radial offset averaged over the BPMs from the injection kicker to the first RF cavities, and the right panel shows the change relative to the first occurrence, $r(n) - r(0)$, converted to the equivalent field B_e through Equation 2.12. Each point is averaged over up to 10 acquisitions, with error bars showing the spread.

which could be attributed to variations in injected energy and position, but the computed B_e can be seen to be decreasing across all cycle measurements, even after 12 seconds which is the longest measurement, with the MD cycle.

Figure 9.3 shows the programmed dipole current and corresponding beam energy across the injection plateau of each cycle, with the injection time marked and the parameters summarized in Table 9.2, as well as the corresponding field changes. The cycles inject at different times and reach their injection plateaus from different excitation changes relative to the 155.63 A idle current. The 26 GeV/c MD2, LHC4 and LHCMD2 cycles ramp up by nearly 150 A to reach their plateau, so that their own ramp induces a large additional eddy current contribution before injection, whereas the SFTPRO and ion cycles change by less than 7 A. The corresponding field ramp rates can be seen in the bottom panel, where the LHC-type cycles reach a peak ramp rate of 0.66 T/s during the 200 ms ramp to injection. In comparison, the maximum ramp down rate of all operational cycles reach −1.82 T/s at its highest when ramping down from flattop fields between 1.8 T–2.0 T. Because the SFTPRO cycle injects earliest, closest to the ramp-down of its preceding cycle, it samples the eddy current field while it is largest, which is why the SFTPRO trajectories show the largest change in Figure 9.2.

9.4 Fitting the Eddy Current Model

The first-turn trajectory data measured in the previous section must be used to fit the coupling parameters c_k of the eddy current model of Equation 9.5. However, the measured data is limited to 39 points when using the four SFTPRO, two SFTION, and one MD cycle. It can also be seen that the measured data does not appear as perfectly smooth curves, as the measurements taken on the beam are subject to uncertainty and variation in the machine. While theoretically, an eddy current model with up to three modes would require only six parameters, the limited and noisy data, together with the non-convex and ill-conditioned nature of the multi-exponential model, require more data to constrain the fit of the time constants τ_k [59, 71]. To enrich the dataset, the orbit decay measurements from Section 9.1 can be used to constrain the time constants τ_k, as the stationary orbit decay follows the same decay physics.

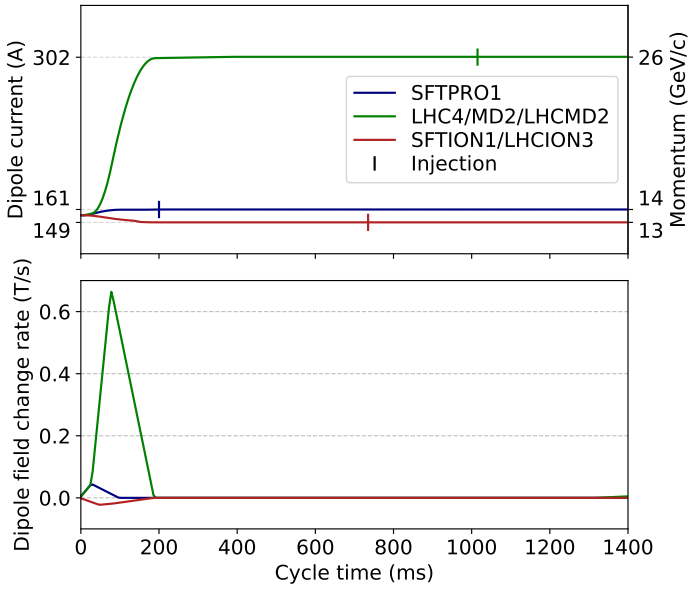


Fig. 9.3: Top panel: programmed dipole current and corresponding beam energy across the injection plateau of the measured cycles, with the injection time of each cycle marked and the idle current at 155.63 A. Bottom panel: the field derivative $\dot{B}$, as programmed in the cycle over the same time range. The injection parameters are given in Table 9.2.

The orbit decay of Figure 9.1 is sampled at the millisecond level, providing several thousand points across the plateau and resolving the shape of the free relaxation, so it constrains the time constants τ_k. Each exponential term amplitude is however not used, because the initial conditions of each mode are not known and do not correspond to the physical couplings. Furthermore, the radial change is proportional to the field change through Equation 2.13, so the time constants are recovered whether the decay is fitted in radial or field units, and the orbit decay is fitted directly in radial units.

The first-turn trajectory of Figure 9.2 is measured once per occurrence, or cycle, at a spacing of 1200 ms, so it is blind to modes faster than $\tau \approx 400$ ms. The orbit decays are on the other hand sampled at up to 1 kHz and are therefore more appropriate for constraining the fast modes. The cycle that injects earliest in the orbit measurements, the SFTPRO cycle at 200 ms, or 350 ms–400 ms after the preceding cycle ramp-down ends, is the main acquisition that can resolve the fastest decay modes. Consequently, decays faster than $\tau \approx 150$ ms can be captured by neither the orbit decay nor the first-turn measurements.

The two measurement sets therefore define two objectives that share the time constants $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_K)$. The orbit objective fits each radial decay to the free relaxation of Equation 9.1 with the shared time constants,

$$\mathcal{L}_{\text{orbit}}(\boldsymbol{\tau}, \mathbf{a}) = \sum_m \left\| r_m - \sum_{k=1}^{K} a_{k,m}\, e^{-t/\tau_k} - o_m \right\|^2 , \tag{9.7}$$

where m indexes the radial decays, r_m is the vector of measured radial samples of decay m, the norm runs over those samples, and the amplitudes $a_{k,m}$ and offsets o_m, collected in $\mathbf{a}$, are fitted to each decay.

The first-turn objective fits the accumulated eddy current field. Writing the per-mode convolution of Equation 9.5 as

$$\varphi_k(t;\dot{B}) = \int_{-\infty}^{t} e^{-(t-t')/\tau_k}\,\dot{B}(t')\,\mathrm{d}t', \tag{9.8}$$

the predicted eddy current field B_e at the injection instant t_{inj} of occurrence n of cycle j is $\hat{B}_{e,j}(n) = -\sum_k c_k\,\varphi_k(t_{\mathrm{inj}};\dot{B}_j^{(n)})$, evaluated over the 60 s of driving history $\dot{B}_j^{(n)}$ preceding that occurrence. Because only the change relative to the first occurrence is used, following Section 9.3, the objective compares the measured change from the reference occurrence against the corresponding predicted change,

$$\mathcal{L}_{\mathrm{ft}}(\boldsymbol{\tau},\mathbf{c}) = \sum_j \sum_n \Big[\big(B_{e,j}(n) - B_{e,j}(0)\big) - \Big(\hat{B}_{e,j}(n) - \hat{B}_{e,j}(0)\Big)\Big]^2, \tag{9.9}$$

where $B_{e,j}(n)$ is the measured field obtained from the radial change through Equation 2.12, and the couplings $c_k \leq 0$, collected in $\mathbf{c}$, are shared across all first-turn data. The absolute eddy current field of each occurrence is computed through the convolution, but only its change relative to the reference enters the objective. The differencing removes the measurement offset left by hysteresis and injection energy, which is not part of the eddy current model, so the couplings fitted from the relative changes still determine the absolute eddy current field.

Introducing the orbit decays constrains the time constants, but it also changes the fit from a single objective for the exponential convolution to a multi-objective optimization. The orbit decays and the first-turn differences are measured in different units and carry different information, so they cannot be pooled into a single residual, and reducing the two to a single weighted objective would fix a weighting between the two measurements and bias the fitted time constants toward one of them. The objective of a sum of exponentials is furthermore non-convex, with many nearly equivalent minima where the time constants are closely spaced, so traditional numerical optimization, whether gradient based or differentiable, converges to whichever minimum lies nearest its starting point. Both properties are addressed by NSGA-II [42], a population based, gradient free, multi-objective evolutionary algorithm. It optimizes the two objectives simultaneously and returns their trade-off as a Pareto front of non-dominated time-constant vectors, so no weighting between the two measurements has to be chosen. Furthermore, its population of candidate solutions searches the bounded time-constant space globally and is not trapped in the nearest local minimum of the non-convex landscape, and it requires only objective values and no gradients. Its elitist non-dominated sorting with a crowding-distance criterion keeps the front well spread, so the returned trade-off solutions are representative. To promote diversity among the time constants, a further separation refinement step is implemented so that candidate solutions in which two time constants lie within a relative tolerance of 5 % are discarded from the population, which prevents two modes from collapsing onto the same value.

Each objective is nonlinear in the time constants but linear in the amplitudes, namely the couplings c_k for the first-turn objective and the mode amplitudes $a_{k,m}$ and offsets for the orbit objective. This separable structure is exploited by variable projection [59]. For any fixed set of time constants, the amplitudes that minimize each objective are the solution of a linear least-squares problem and are obtained in closed form, so they need not be searched over. Both sets of amplitudes carry a fixed sign, so both are recovered by non-negative least squares, the couplings because $c_k \leq 0$ by Lenz's law of Section 2.6.1, and the orbit amplitudes because all measured orbit decays relax in the same direction. Substituting the recovered amplitudes back leaves the reduced objectives $\mathcal{L}_{\mathrm{orbit}}(\boldsymbol{\tau})$ and $\mathcal{L}_{\mathrm{ft}}(\boldsymbol{\tau})$ as functions of the time constants alone, so

NSGA-II searches only the time constants and reconstructs their amplitudes from the linear solve at each candidate.

The only free parameters of the optimization are limited to physically meaningful values. The lower bound is set by the SFTPRO cycle orbit decay with its early injection time, to 150 ms. The upper bound is set to 20 s, supported both by the 60 s of acquired driving history, which spans about three times this bound, and by the orbit decays in Figure 9.1, which show no relaxation on longer timescales.

The number of modes is not known in advance, so models with one, two, and three modes are fitted in turn and the supported order is chosen by the model-order selection below. Decomposing a decay into a sum of exponentials is increasingly ill-conditioned with each added mode, and the number that noisy data can resolve is small, so orders beyond three are not attempted [71]. Each larger model is warm-started from the converged result of the smaller one, keeping the previous time constants and introducing the new mode from a small value. The previous time constants are not frozen, and the enlarged model is re-optimized over all of them jointly, as adding a mode re-partitions the decay across all modes and shifts the optimal value of the existing time constants, and because freezing them would assume the smaller fit had already found their exact values, which the non-convexity does not guarantee. Warm-starting from the smaller model instead makes the larger Pareto front dominate the smaller one, since additional modes cannot fit worse.

The optimization produces a Pareto front of time-constant vectors for each model order, shown in Figure 9.4 and enumerated in Table 9.3. From the front of the selected model order, a single compromise solution is chosen by normalizing the two objectives to the unit interval and taking the vector with the smallest Euclidean distance to the origin of the normalized objective plane.

Tab. 9.3: Time constants τ_k and couplings c_k of the eddy current model fitted for one, two, and three modes. For each model order, the two ends of the Pareto front, minimizing the orbit objective $\mathcal{L}_{\text{orbit}}$ and the first-turn objective $\mathcal{L}_{\text{ft}}$ respectively, and the selected compromise solution are listed, together with the value of each objective. The time constants are given in seconds. The couplings are given for the eddy current field in Gauss and are scaled by 1×10^{-4} to express the field in Tesla.

Model	Selection	$\mathcal{L}_{\text{orbit}}$	$\mathcal{L}_{\text{ft}}$	τ			c		
				1	2	3	1	2	3
1-exp	min orbit	0.0294	0.1072	1.717	-	-	-1.637	-	-
1-exp	min conv	0.2674	0.0282	0.588	-	-	-5.057	-	-
1-exp	compromise	0.1254	0.0522	0.868	-	-	-3.062	-	-
2-exp	min orbit	0.0044	0.0366	0.749	3.406	-	-3.228	-0.216	-
2-exp	min conv	0.0109	0.0095	0.297	3.015	-	-9.832	-0.855	-
2-exp	compromise	0.0081	0.0100	0.302	2.635	-	-9.596	-0.854	-
3-exp	min orbit	0.0042	0.0101	0.391	1.028	3.621	-6.248	-0.000	-0.717
3-exp	min conv	0.0065	0.0093	0.126	0.495	3.369	-89.331	-2.161	-0.792
3-exp	compromise	0.0044	0.0095	0.200	0.709	3.328	-21.899	-0.777	-0.785

The one-mode model fits both objectives poorly, its front lying far from the origin in Figure 9.4, and its single time constant swinging from 1.7 s at the orbit-optimal end of the front to 0.59 s at the first-turn-optimal end, since one mode cannot satisfy both measurements at once. The two-mode model lowers both objectives, though its compromise solution remains above the orbit-optimal end of its own front, at $\mathcal{L}_{\text{orbit}} = 0.0081$ against 0.0044. The three-mode front lies

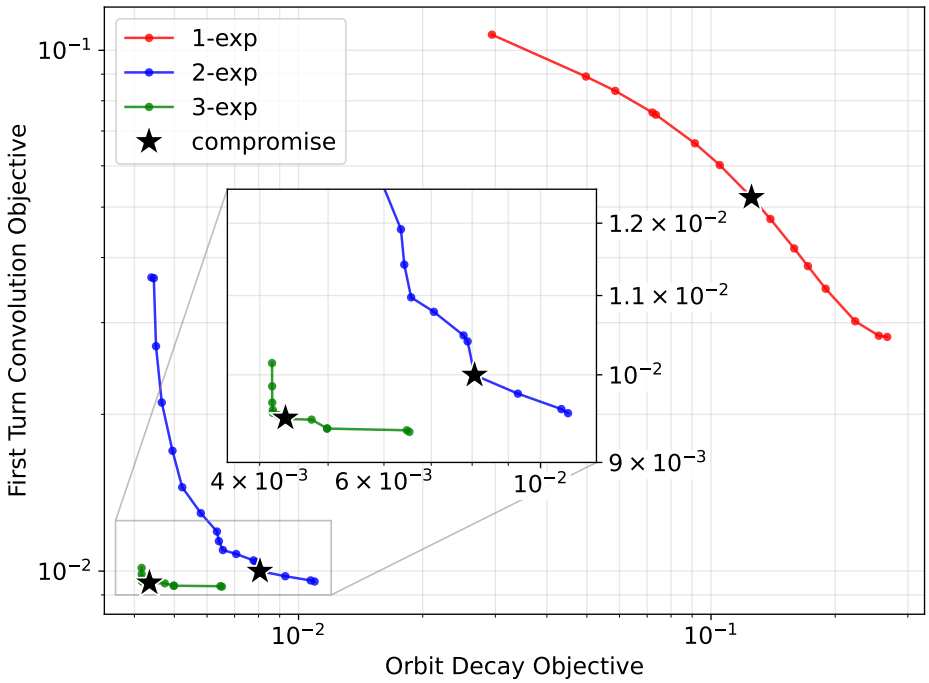


Fig. 9.4: Pareto front of the multi-objective optimization of the eddy current model, with the orbit decay objective on the x-axis and the first-turn trajectory objective on the y-axis. The front is colored by the number of modes in the model, with one mode in red, two modes in blue, and three modes in green. The selected compromise solution is marked with a star.

close to the corner of the objective plane, and its compromise solution nearly reaches the minimum of each objective at once, at $\mathcal{L}_{\text{orbit}} = 0.0044$ against 0.0042 and $\mathcal{L}_{\text{ft}} = 0.0095$ against 0.0093. The orbit objective changes little from two to three modes while the compromise improves, which suggests that the third mode brings the two measurements into agreement rather than lowering one objective alone. The three fitted time constants are separated by more than a factor of three and none reaches the search bounds, so the decomposition is not degenerate, and the three-mode compromise is carried into the next sections.

The selected time constants are $\tau \approx 0.20$, 0.71, and 3.3 s. These values are a compromise fit and not a unique identification of the physical eddy current modes. Across the three-mode front the slow time constant stays near 3.3 s–3.6 s for every selection, whereas the two faster time constants and, more strongly, the couplings vary widely with the chosen trade-off, the leading coupling ranging from -6.2 at the orbit-optimal end to -89.3 at the first-turn-optimal end. The data therefore admit a family of time-constant and coupling sets that fit comparably well, and the reported values are read as one representative solution rather than the true system coefficients. What is robust across this family is the slow mode near 3.3 s, which is longer than any decay recorded in the magnetic measurements of Table 9.1, whose longest fitted time constant is 1.8 s, and consistent with the multi-second orbit drift of Section 9.1 that the magnet-mounted measurements do not reproduce. The couplings, which set the accumulated eddy current field, are moreover calibrated only against the first-turn differences, a proxy sampled once per cycle occurrence, and not against the orbit shape at injection, so how faithfully the fitted model reproduces the accumulated decay is examined in Section 9.6.2.

9.5 Compensation Design and Results

9.5.1 A Complementary Compensation Scheme

The eddy current decay is corrected by the same feed-forward mechanism proposed for hysteresis in Section 5.4.6, with the eddy correction added as a further term to the field correction of Equation 5.4. The two corrections share the same prediction converter, gain, and correction window, so the eddy branch reuses the compensation infrastructure of Chapter 8, described in Section 8.1.4, without a separate event building and control loop.

The eddy field realized over a cycle is predicted from the field derivative history through the exponential convolution model of Section 9.4, using the time constants and coefficients fitted there. The per-cycle prediction $\hat{B}_e(t)$ is evaluated over the full cycle at 1 kHz by the frequency-domain convolution of Section 9.2, whose one-pole kernel reduces the evaluation to an $\mathcal{O}(N \log N)$ operation. The convolution is taken over the same 60 s excitation buffer used by the hysteresis prediction, which guarantees a history long enough to resolve the slowest fitted time constant. At the time of prediction, the programmed field derivative of the preceding cycles held in the buffer are combined with the programmed field derivative of the current cycle up to that time.

The programmed field derivative is required because the field derivative is measured only after a cycle has completed and is not known in advance for the cycle being predicted. It is computed from the programmed excitation current through the chain rule

$$\dot{B} = \frac{\mathrm{d}B}{\mathrm{d}I}\,\frac{\mathrm{d}I}{\mathrm{d}t}, \tag{9.10}$$

where $\frac{\mathrm{d}B}{\mathrm{d}I}$ is obtained by differentiating the transfer function $B(I)$ of Figure 4.12b and $\frac{\mathrm{d}I}{\mathrm{d}t}$ from the programmed current, and the programmed current of cycles in economy mode can likewise be used when published from their respective converters, as outlined in Section 8.1.3. This computation is carried out by a dedicated UCAP converter that subscribes to the programmed current published whenever the main dipole excitation changes, whether by an operator or by the compensation itself, and republishes the field derivative for the programmed cycles. In practice, the implemented converter is restricted to programmed cycles, so the power converter economy modes of Section 4.4.6.3, which alter the realized excitation independently of the program, are not tracked in this deployed converter. Tracking the realized excitation under economy would require an equivalent converter for the economy current and is left as future work, since economy changes the accumulated eddy current history. The start-cycle converter attaches the resulting programmed field derivative to the cycle data passed through the prediction pipeline, in the same way programmed current is attached for the hysteresis prediction.

Analogously for the ML-field model compensation, the correction is defined relative to a reference. The eddy reference $B_{e,\mathrm{ref}}(t)$ is frozen at the first prediction for a given cycle type and left unchanged until a reset is triggered, either manually for that cycle or by an operator change of the reference excitation, at which point the entire eddy state is reset at once. The raw eddy correction is the difference between this reference and the current prediction,

$$\Delta B_{e,\mathrm{raw}}(t) = B_{e,\mathrm{ref}}(t) - \hat{B}_e(t). \tag{9.11}$$

The hysteresis and eddy predictions differ in how they respond to the correction they induce. The main field model of Chapter 7 predicts from the present excitation, as programmed, so a correction feeds into its next prediction and converges on its own. The eddy prediction depends only on the past field derivative history, which a small present correction does not appreciably

change. Therefore, the raw correction of Equation 9.11 stays almost constant from cycle to cycle under unchanged conditions, and re-applying it produces a runaway similar to integrator windup. The applied correction therefore needs to subtract a running reference $R(t)$, set after each trim to the last absolute eddy correction sent to the control system,

$$\Delta B_e(t) = \Delta B_{e,\mathrm{raw}}(t) - R(t), \tag{9.12}$$

so that only the increment still owed is applied, giving the eddy branch the convergence the hysteresis branch has by construction.

The main field model of Chapter 7 is trained on the B-Train and therefore already reproduces the eddy current decay recorded in the magnet. To avoid correcting that contribution twice, the eddy field predicted with the magnet time constants of Table 9.1 needs to be subtracted from the main field prediction before the reference is formed or any correction is computed. The coefficients used are explained later, in Section 9.6.3. This subtraction is evaluated at 1 kHz and downsampled to the 50 Hz field prediction rate, leaving a field prediction free of the first-order eddy current contribution the main model would otherwise duplicate. The beam-fit convolution then supplies the correction for the ring-integrated decay that the magnet measurement does not capture.

The eddy correction is combined with the hysteresis correction before it is applied. For development, the two contributions can be applied together or in isolation,

$$\Delta B(t) = \begin{cases} \Delta B_h(t) & \text{hysteresis only} \\ \Delta B_e(t) & \text{eddy current only} \\ \Delta B_h(t) + \Delta B_e(t) & \text{combined,} \end{cases} \tag{9.13}$$

with the eddy prediction downsampled from 1 kHz to the 50 Hz rate of the hysteresis prediction before the two are summed. The combined correction $\Delta B(t)$ is applied to LSA as a trim to the programmed field over the correction window, through the same accumulation as the hysteresis correction of Equation 8.1, with $\Delta B(t)$ in place of the residual $\Delta_c(t)$. The gain g and correction window are shared with the hysteresis correction, and the gain acts on the total correction or on either contribution alone in the corresponding mode. After the trim, the running reference is updated to the applied correction.

Because the correction is a difference of predictions, its value depends only on the change in predicted eddy field between the reference cycle and the current cycle, and not on the absolute value of the prediction. The compensation is therefore insensitive to the absolute sign of the predicted eddy field discussed in Section 9.6.1, since keeping the macroscopic field reproducible is sufficient for the beam reproducibility that follows from it.

The compensation algorithm computes the eddy field and prepares for correction over the full cycle, following the original goal of correcting the field throughout, set out in Section 5.4.6. The eddy current contribution is nonetheless largest at injection and at the start of the ramp, where the eddy field from the preceding ramp-down has not yet decayed. Once the cycle's own ramp-up dominates the excitation, its contribution is carried by its own cycle, and can be corrected in the cycle program rather than by a dynamic correction. This suggests a case where the eddy field compensation could be decoupled from the hysteresis compensation, which needs to act on the entire cycle, however this case is not explored in this work.

9.5.2 Operational Validation

The eddy current correction was tested operationally on two cycles during dedicated machine development time on different occasions, on a short 26 GeV/c MD cycle above transition and on the SFTPRO injection below transition. Both tests apply the correction as a feed-forward trim to the programmed current through the scheme of the previous section, and in neither test is the main field model correction from the previous two chapters applied, for the reasons given at the end of this section. The correction was applied in a single shot in both tests, and as the applied correction is a deterministic and smooth function of the field-derivative history, it carries no uncertainty of its own.

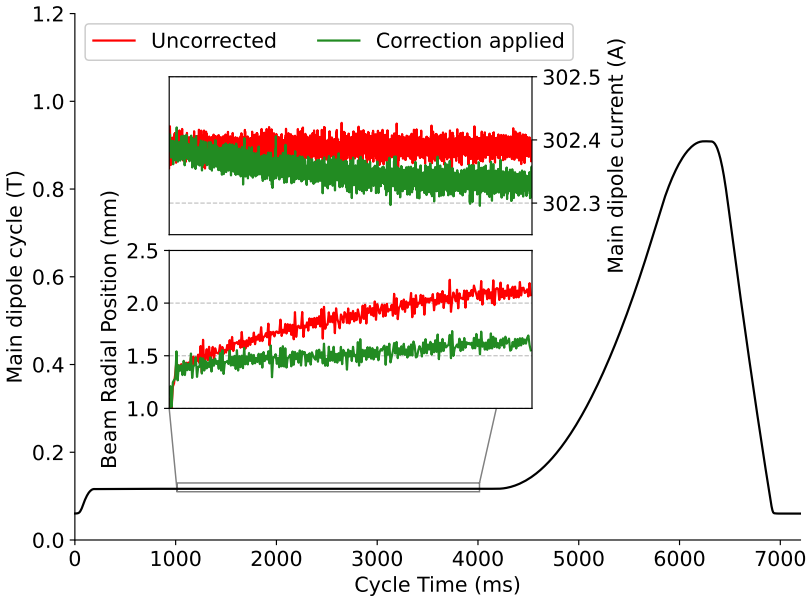
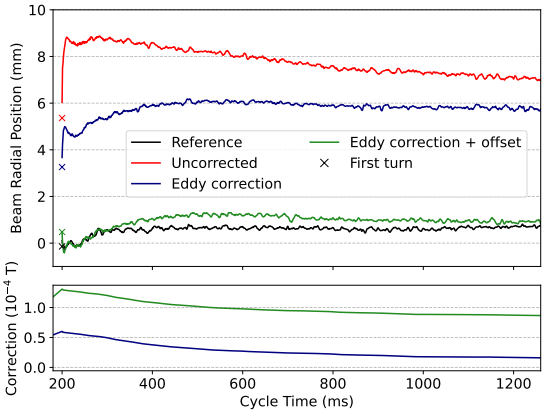


Fig. 9.5: Eddy current correction on the injection plateau of a short 26 GeV/c MD cycle, above transition. The black curve is the full main dipole cycle. The lower inset shows the beam radial position over the injection plateau, which drifts by about 0.9 mm uncorrected (red) and by about 0.4 mm with the correction applied (green). The upper inset shows the corresponding main dipole current, where the applied correction trims the current downward across the plateau. The correction reduces the drift but does not remove it.

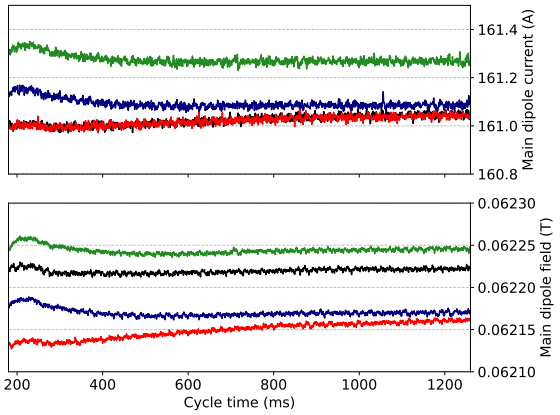
The first test corrects the closed orbit drift on a 26 GeV/c short MD cycle, however at the time of this test the supercycle timing was not under our control, so the correction could not be referenced to a field predicted from a controlled history. The eddy reference was instead set to zero, $B_{e,\mathrm{ref}} = 0$, so that the correction applied drives the eddy field toward a flat plateau, which should flatten the field experienced by the beam. The correction additionally required an empirical sign inversion to reduce rather than enlarge the excursion, which is examined together with the above-transition orbit direction in Section 9.6.1. Figure 9.5 shows the results of the correction, with orbit and current measurements on the injection plateau shown in the insets. The uncorrected orbit drifts by about 0.9 mm across the plateau starting from 1.4 mm offset, after originally being injected at −1.0 mm and kicked by the RF loops. Through the feed-forward correction, the drift is reduced to about 0.4 mm, and the corresponding change in excitation current can be seen in the lower inset. The drift is therefore not fully compensated, since a

residual drift of the same sign remains. The measured field is not shown for this cycle, as on the long injection plateau it is dominated by the integration drift of the B-Train rather than by the eddy current decay, as shown in Chapter 6.

The second test corrects the SFTPRO injection plateau, the most operationally relevant cycle, as outlined in Chapter 5. In this case, the control loop proposed in the previous section is used natively, and the eddy reference is taken from the SFTPRO cycle running in nominal operation with its MD1 precycle, and the proposed correction is applied to the same cycle with the precycle removed. The MD1 removal is achieved by flattening the dipole on the precycle so that the current stays at the idle 155.63 A and does not ramp, contrary to just removing the precycle from the cycle and removing the 3.6 s idle time. This is necessary as the quadrupole and higher order magnet circuits still require precycling, and removing the precycle could add other sources of effects than only dipolar field changes.



(a) Top: beam radial position over the injection plateau for the reference (precycle on), uncorrected (precycle off), eddy correction, and eddy correction + hysteresis offset. Crosses mark the first turn, which is not always the same as the first orbit point. Bottom: the applied field correction for the two corrected cases.



(b) Measurement of excitation current and magnetic field from the B-Train for all conditions in the left plot. Main-dipole current (top) and field (bottom) for the same four conditions. Uncorrected and reference share the same current but differ in field, caused by the hysteresis shift from the removed precycle.

Fig. 9.6: Operational eddy-current feed-forward on the SFT injection plateau with the precycle removed: (a) beam response and applied correction; (b) the corresponding main-dipole current and field. The eddy correction partially flattens the decay, and adding the empirical hysteresis offset brings the orbit close to the reference.

The results of the SFTPRO correction can be seen in Figure 9.6. At the unchanged reference state, the reference orbit, shown in black in Figure 9.6a, is well centered around 0.6 mm after an initial rise from an injection at −0.1 mm. Removing the precycle without any corrections on the dipole field perturbs the beam strongly, and the first turn moves to 5.4 mm and the beam is kicked to 8.8 mm by a combination of the RF loops and injection energy mismatch, after which it decays across the plateau, shown in red. The same orbit can be seen decaying inwards, reaching 7.0 mm at the plateau, as the small current ramp seen in Figure 9.6b at injection is not removed, and the increase in field as programmed moves the beam inward, consistent with Equation 2.13. This small current ramp is part of the programmed cycle at the injection point where the measurement is taken, and appears in the reference as well as in every corrected condition. The initial move of the orbit over the first 80 ms is therefore common to all traces, and since the correction is defined relative to the reference, it is retained rather than removed.

Applying the eddy correction with the three-exponential model, shown in navy, improves the first turn to 3.3 mm and reduces the drift, bringing the orbit further inward, but does not fully flatten it. A residual variation remains, including a bump between 400 ms and 600 ms, where the fast decaying mode of the correction disappears, after which the orbit drifts about 0.5 mm further inward.

The eddy correction applied addresses only the decaying part of the field, so the constant outward shift left by the removed precycle static field remains. Adding an empirical constant offset of $+7 \times 10^{-5}$ T over the injection, compensating the static hysteresis field left by the missing precycle, brings the first turn to 0.47 mm and the orbit to 0.96 mm at the end of injection, with the same bump remaining, shown in green. The applied corrections are shown in the lower panel, the eddy correction decaying from about 0.6×10^{-4} T to 0.15×10^{-4} T, and the eddy correction with the constant offset from about 1.25×10^{-4} T to 0.85×10^{-4} T.

A further point to observe is that the first turn does not correspond to the orbit at later points in the corrected cases, marked by the crosses in Figure 9.6a, since the first-turn trajectory is measured before the RF kick and is not the closed orbit, as introduced in Section 9.3. For the fully corrected case in green, the RF kicks the beam inward by 0.9 mm before it continues to decay outward. Therefore, even when fully centering the beam at first turn, the orbit may still be kicked off-center depending on the RF and energy matching of the beam.

The field measurements in Figure 9.6b are drift corrected following Section 6.2.2. It can be observed that the uncorrected and reference conditions share the same current but differ in field, caused by the static hysteresis shift from the removed precycle. The corrected fields drift upward slightly over the plateau, which is not fully consistent with the nearly constant current, and is attributed to the residual integration drift that remains on the plateau and should not be over-interpreted.

With the corrected field and orbit both close to the reference, the measured field of the eddy correction with offset (green) still lies about 2.5×10^{-5} T away from the reference field at the end of the plateau. This corresponds through the field-to-orbit relation of Table 5.2 to more than 1 mm of radial offset, while the measured orbit differs by only about 0.2 mm. Part of this difference may be residual integration drift, but as the end of the plateau is closest to where the integration anchor is applied, it suggests that the field recorded in the magnet does not map to the orbit the beam realizes, consistent with the observations at the start of this chapter.

The main field model correction was not applied in these tests, and only the static offset was used, for two reasons. The beam is extremely sensitive at injection, where a field error of order 1×10^{-5} T–2×10^{-5} T is already significant, so any noise or transient in the ML model correction over the injection plateau is translated directly to the beam. Furthermore, the next chapter shows a distinct effect of the precycle that the dipole hysteresis compensation cannot address. The static offset was therefore applied alone, to isolate the eddy current correction under this test.

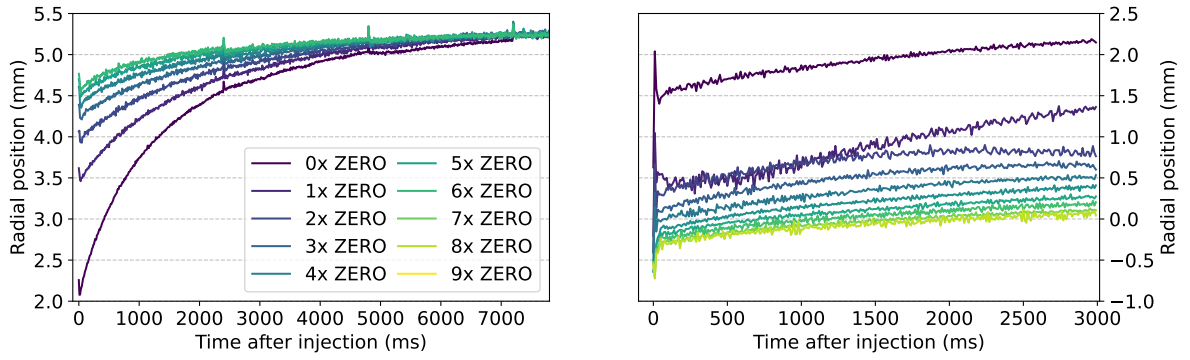
While the fitted model does not fully compensate the beam under the test conditions, it is nevertheless demonstrated that the compensation scheme itself works. The feed-forward path computes the eddy correction from the field-derivative history and applies it to the control system as a bounded trim that moves the beam in the intended direction, on the same infrastructure as the hysteresis compensation. The running-reference scheme of Equation 9.12 converged in a single correction, since the increment still owed on the following unchanged cycle fell below the trim threshold, so no further correction was applied and the loop settled without runaway. On the model side, the eddy correction only partially compensates the decay, reducing the overall drift while leaving a residual shape it does not capture, so it is not ready for deployment on arbitrary cycles. As suggested in the previous section, the fitted coefficients are a compromise

rather than the true description of the beam-observed decay, and the reasons the fit does not constrain the deployed quantity are discussed in Section 9.6.

9.6 Discussion

9.6.1 Orbit Decay Above Transition

The empirical sign inversion required by the MD correction in Section 9.5 points to an inconsistency in the decay direction above transition. In Figure 9.1 it can also be seen that every cycle decays upward even when immediately following a previous cycle, including the three above-transition cycles MD2, LHC4, and LHCMD2. The same cycles measured for the first-turn trajectories in Section 9.3 are re-examined as closed-orbit decays over the injection plateau, for the different numbers of ZERO cycles inserted before each cycle. Throughout each plateau the programmed current is held constant, and the revolution frequency is set to follow the programmed field, which depends only on the current. Therefore any orbit drift after settling of RF loops in the first few milliseconds can only be a deviation of the realized dipole field away from the programmed field.



(a) SFTION1, below transition, for 0 to 7 inserted ZERO cycles. The decay amplitude falls from 3 mm with no inserted cycles to 0.5 mm, and all traces settle at 5.2 mm.

(b) MD cycle at 26 GeV/c, above transition, coloured as in (a). The decay amplitude also falls with the number of inserted cycles, but the settled position differs between them, from 2.1 mm for the cycle with no dwell time, to 0.1 mm for the one with 10x ZEROs.

Fig. 9.7: Closed-orbit decay after injection for different numbers of ZERO cycles inserted before the measured cycle, below transition (SFTION1) and above transition (MD, 26 GeV/c). More inserted cycles lengthen the dwell before injection and reduce the decay amplitude in both cases, and both decay upward.

Figure 9.7a shows the acquired beam orbits of the SFTION1 cycle, below transition. The cycle is injected at 13 GeV/c where the field ramps down to reach the injection plateau, since the injection current lies below the idle current, as seen in Figure 9.3. The orbit decays upward, its amplitude decreases as more ZERO cycles are inserted, and all traces settle at the same position. This is the expected eddy current decay below transition, where the preceding ramp-down leaves the field decreasing over the plateau, and by Equation 2.13 a decreasing field moves the orbit outward. Each inserted ZERO cycle adds 1200 ms of dwell before injection, so less eddy current field remains and the decay is smaller, while the common settled position is the remanent field, which the inserted cycles do not change.

Figure 9.7b on the other hand shows the 26 GeV/c MD cycle, above transition, which does not follow the same behavior. The orbit again decays upward, yet above transition a field relaxing from the preceding ramp-down should move the orbit downward, opposite to what is observed. An upward decay above transition instead requires the field to increase, which the preceding ramp-down cannot produce, since a ramp-down leaves the field decreasing. While it could be argued that the ramp-up to injection of the cycle itself could dominate the previous ramp-down, the measured decay with increased dwell time is not consistent with that either. If it did dominate, inserting more ZERO cycles would decay away the opposing contribution of the preceding ramp-down and make the upward decay larger, while the measured decay seen in the Figure instead becomes smaller. The measured direction and the measured dwell dependence therefore cannot be reconciled within a single eddy current picture. The settled position additionally differs between the inserted-cycle counts, unlike the common settling below transition, and this difference is not explained.

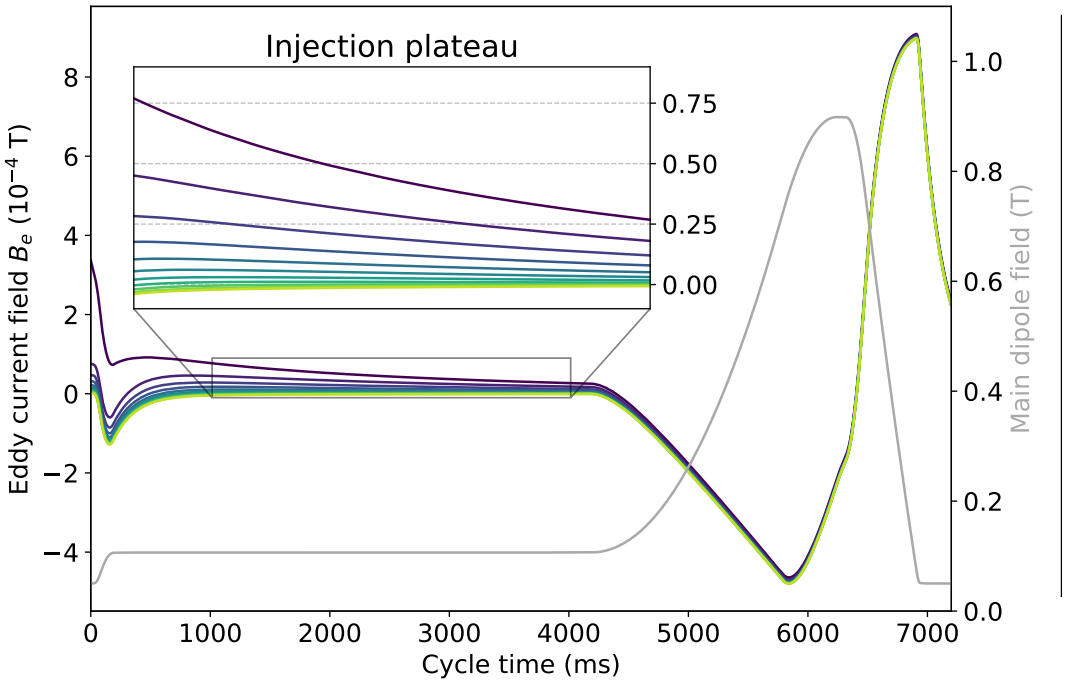


Fig. 9.8: Eddy current field B_e predicted by the fitted model over the 26 GeV/c MD cycle for 0 to 10 inserted ZERO cycles, coloured from dark to light with increasing dwell. The grey curve is the main dipole field on the right axis, shown for context. The inset magnifies the injection plateau, where the predicted field reaches about 7×10^{-5} T with no inserted cycles, stays positive for all but the largest dwell, and crosses negative only beyond seven inserted cycles and then by less than 5×10^{-6} T.

The fitted model also does not resolve the direction. Figure 9.8 shows the eddy current field it predicts at the MD injection plateau for the same range of inserted ZERO cycles. The predicted field is positive, and decreases in magnitude as more dwell time is added. It crosses negative only beyond seven inserted ZERO cycles, and then by less than 5×10^{-6} T. It can be further seen that for all cases except no inserted cycles, the predicted field does indeed go negative at the ramp-up, becoming more negative with more dwell time, which allows previous ramp-downs to decay away. However, as the ramp-up finishes at 200 ms, and beam is injected only 815 ms later,

the fast modes of the ramp-up contribution have already decayed sufficiently to be dominated again by the slower decay modes of the previous ramp-down. The predicted field decreases over the plateau, and the upward orbit above transition requires an increasing field, so the model captures the magnitude of the decay with dwell but not its direction.

The direction of the decay is interpreted through the orbit-to-field relation of Equation 2.13, which changes sign at transition. In Figure 9.1 the three above-transition cycles MD2, LHC4, and LHCMD2 all decay upward in a similar manner, the longest out to more than 12 s, so above transition the field must be increasing over the plateau. Operation leverages the same relation, as the LHC injection cycles above transition carry a small programmed field ramp-down that brings the orbit inward and holds it centered against this upward drift, while the SFTPRO injection below transition carries a small programmed field ramp-up, also seen in Figure 9.6b. The model instead predicts a decreasing field, which by the same relation would move the orbit downward, so it does not reproduce the measured direction. The open question is therefore not only that the model fails, but why, above transition, the field increases over the plateau at all, when the large preceding ramp-down should leave it decreasing, and other identified machine conditions do not support a drift. The MD correction of Section 9.5 required an empirical sign inversion for this reason, and while the behavior remains unexplained a correct prediction cannot be distinguished from an incorrect one, so the model cannot be deployed.

9.6.2 Feasibility of the Fitted Model

The sign anomaly of the previous section concerns the direction of the predicted field above transition, however the amplitude and shape error of the correction found in the compensation results are a limitation of the fitted model itself. The fitted coefficients in Table 9.3 form a compromise, reproducing the beam-observed decay only in part rather than identifying the physical eddy current modes.

The first-turn objective fixes the couplings that set the accumulated eddy field at the point of injection, while the orbit decays fix the compatibility of the shared time constants. Neither objective constrains the decay shape across the injection plateau, which is the observable that is compensated. Calibration of the couplings uses the first-turn differences alone, and as exemplified in Figure 9.6a, the first turn does not always match the closed orbit at later points on the plateau, so it cannot pin the plateau decay it predicts.

The faster modes are the least constrained of the fitted parameters, since the orbit decays test them only for compatibility of the time constants and not accumulated contribution. A fast mode of large amplitude can then satisfy both objectives while adding a transient over the first few hundred milliseconds that the measured decay does not show, which produces the bump on the corrected orbit of Section 9.5. Modes faster than 200 ms also approach the resolvability limit of Section 9.3 and are not confidently supported by the data. Constraining the model directly on the plateau decay would require an orbit driven by the main dipole field alone, which is not available. The plateau orbit also carries an offset from the RF regulation that cannot be trivially separated from the eddy-driven decay, so no absolute orbit decay can be acquired to fit the convolution model against.

The first-turn offset is not a clean measure of the dipole field error, since effects besides the eddy current field contribute to it. Only 36 of the 108 horizontal-plane BPMs, those between the injection kicker and the RF cavities, give a first-turn trajectory free of the strong RF kick, and these span only two of the six sextants of the ring. The conversion from radial offset to field error of Equation 2.12 is expressed through the momentum compaction, a ring-averaged quantity from Equation 2.10, so it holds only for an offset averaged over the whole ring or corrected by the local dispersion at each BPM. Neither is available here, since the measurement covers two sextants and

the local dispersion is not measured, so converting a partial-ring offset through the ring-averaged momentum compaction carries a systematic error. The offset further depends on the position and energy of the beam injected from the PS, which produce the same dispersive signature as a dipole field error and cannot be separated from it in a single first-turn measurement, reduced but not removed by averaging each point over about ten acquisitions. These errors enter the couplings that set the accumulated eddy field, and because the sum of exponentials is non-convex with many near-equivalent minima, as the spread of the Pareto front shows, they do not only add scatter to otherwise well-determined couplings but shift which solution the fit selects, so the measurement uncertainty and the ill-conditioned fit compound.

9.6.3 Comparison of the Beam-Fitted and B-Train Models

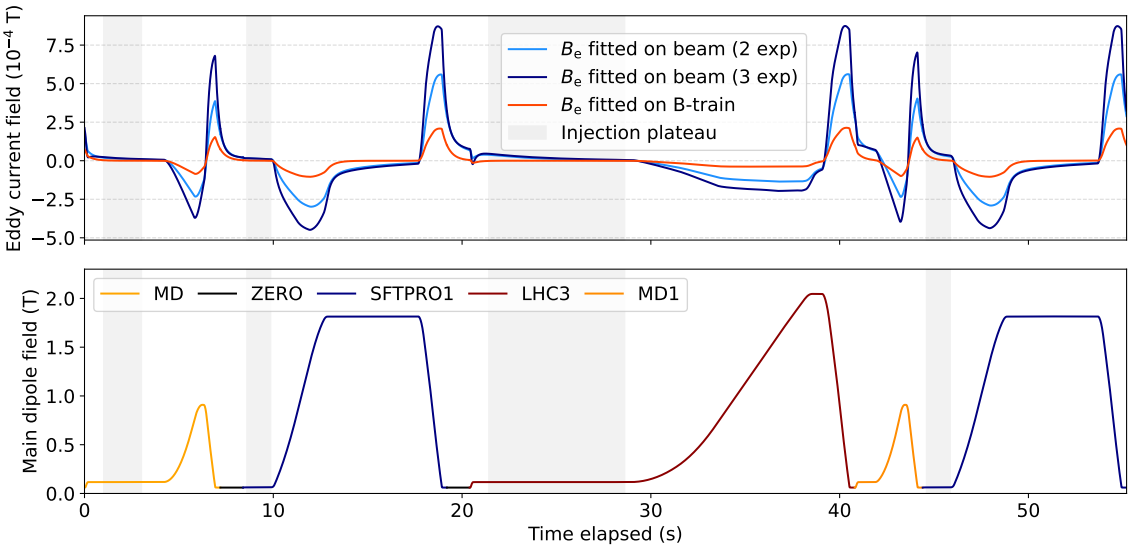


Fig. 9.9: Comparison of the eddy current field B_e predicted by the two-exponential and three-exponential beam-fit models against the B-Train fit of Table 9.1, over a machine development supercycle. The top panel shows B_e with the injection plateaus shaded, and the bottom panel the main dipole field and cycle sequence. The cycle sequence is from an MD session and does not represent an operational sequence. The cycle sequence contains a 26 GeV/c MD cycle, standard SFTPRO and a 3 injection LHC cycle as opposed to a standard 4 injection LHC1 cycle. In this sequence, one SFTPRO cycle is preceded by a ZERO cycle following the MD cycle, whereas the second one is preceded by a standard MD1.

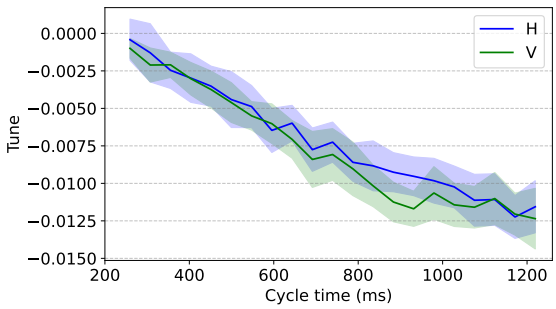
The difference between the beam-observed and the magnet-measured decay can be expressed directly in the model by predicting the eddy field of both fits over the same excitation history. Figure 9.9 compares the eddy current field predicted by the beam-fit model against the field predicted by the B-Train fit of Table 9.1, over a magnetic sequence taken from the SPS during machine development. The B-Train field is produced by the same convolution model as the beam field, but the couplings are obtained by non-negative least squares on the same plateaus used for the B-Train time-constant fit and their field-derivative history, holding the B-Train time constants of Table 9.1 fixed.

Across the injection plateaus, shaded in the figure, the beam-fit field is larger in amplitude and returns to zero more slowly than the B-Train field. Both the two-mode and three-mode beam

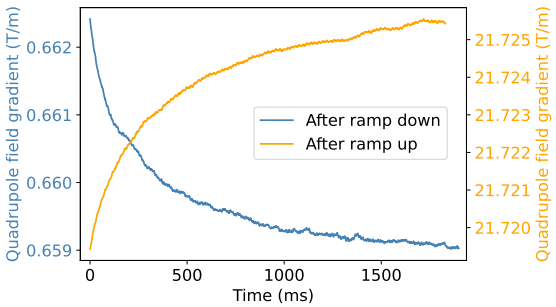
fits show this, so the difference does not depend on the model order. At the ramps the beam-fit field spikes far above the B-Train, most strongly for the three-mode fit, whose large fast-mode coupling amplifies the field derivative. This is the same fast mode identified in Section 9.6.2 as poorly constrained and near the resolvability limit, so the spike is not necessarily physical. The large induced field at the ramps is nonetheless not of concern here, since no beam is present during a ramp-down, and during a ramp-up the eddy field is excited by the cycle's own ramp, which is reproducible and therefore already absorbed by the static cycle settings rather than the inter-cycle correction developed here. The field relevant at injection is that on the plateau, dominated by the decaying tail of the preceding cycle's ramp-down.

The two fits are obtained from physically different measurements. The B-Train records the decay in two spare dipoles powered in series with the main circuit but installed outside the tunnel, so it cannot see the ring conductors of Section 9.1, whereas the beam-fit model is obtained from the field the beam itself experiences around the whole ring. This comparison therefore supports the interpretation that the field the beam experiences decays more slowly, and remains larger, than the field the B-Train records, as first observed from the orbit drift in Section 9.1. Although the fitted coefficients do not uniquely identify the eddy current system, as discussed in Section 9.6.2, this conclusion does not depend on that identification, since both beam fits give a field at injection that is consistently larger and slower to decay than the B-Train field.

9.6.4 Higher Order Magnets



(a) Change in the horizontal (H) and vertical (V) betatron tune across the SFTPRO injection plateau, extracted by NAFF from the strongest spectral line. Each curve is the difference between a flat injection plateau and one carrying the small programmed ramp used in operation, with the MD1 precycle present in both, and each condition averaged over ten acquisitions. The shaded band is the combined standard deviation of the two conditions.



(b) Field gradient of a spare main quadrupole measured on the laboratory bench with a fluxmeter during a trapezoidal excitation, on the plateau following a ramp-down to idle current (blue, left axis) and following a ramp-up to the flat-top current (orange, right axis). The gradient relaxes in opposite directions for the two ramps and continues to decay over more than a second.

Fig. 9.10: Eddy current decay in the main quadrupole circuit, measured on the beam through the tune (a) and in the laboratory bench through the field gradient (b).

The eddy current mechanism is not confined to the main dipoles, and the higher order circuits are expected to show similar behaviour. For the main quadrupoles the beam property that reflects the field is the betatron tune, which shifts with the integrated field gradient. Whether the quadrupole decay follows the main dipole decay depends on where the ring-integrated eddy currents of Section 9.1 are induced.

Figure 9.10b shows the gradient of a spare main quadrupole measured on the laboratory bench across the plateau following a ramp, in the same manner as the dipole bench measurement of Chapter 6. The gradient decays after both a ramp-down and a ramp-up which both share the same length and rate, in opposite directions as expected for an eddy current, and continues to decay over more than a second. The decay is physically attributed to the conductive end plates of the magnet, which support larger eddy current loops than the laminated yoke and therefore decay more slowly.

Figure 9.10a shows the corresponding decay seen by the beam, measured as the change in the horizontal and vertical tune across the SFTPRO injection plateau and extracted by NAFF. The tune was acquired with the small programmed ramp at injection and with a flat injection, both retaining the MD1 precycle, and the difference of the two isolates the drift over the plateau, averaged over ten acquisitions. The beam was centered before both measurements, so the measured tune change reflects the decaying field and not an orbit or energy offset. Both planes drift by about 0.012 over the plateau. As the frequency spectra are chunked over 48 ms over the 1000 ms acquisition, the time resolution of the measurement is limited compared to BPM measurements of the orbit. Furthermore, as the shown quantity is the difference between two acquisitions, the noise of each measurement is combined, which is reflected in the shaded standard deviation band.

Where the quadrupole eddy currents seen by the beam are induced remains an open question. If the conductive structures of the ring carry the dipolar decay observed on the beam, their contribution to a quadrupole field depends on their geometry, since a beam pipe symmetric about the beam axis largely reproduces the multipole symmetry of the field that drives it. A quadrupole eddy component would then be more local to the magnet, such as its end plates, than the ring-integrated dipole decay, which the laboratory bench decay is consistent with, although the present data cannot separate a magnet-local contribution from a ring contribution.

The same convolution model of Section 9.2 applies to the quadrupole, with the measured tune decay converted to a gradient change and corrected through the main quadrupoles by the feed-forward path of Section 9.5. Realizing it requires centering the beam to remove orbit and energy contributions before each measurement, as was done here, and repeating this across every circuit and magnetic state, which makes the higher order case more involved than the dipole. It is left as future work, and no modeling or compensation of the higher order circuits was attempted in this thesis.

Chapter 10

Operational Simplification of the Precycle

The MD1 precycle originally introduced in Section 4.6.1 is always played before proton and ion fixed-target cycles to establish a reproducible magnetic state ahead of injection, and is understood in operation as a reset of the magnetic history of the main and higher-order circuits. For the injection compensation attempted in Chapter 8 and the eddy-current investigation of Chapter 9, the precycle was flattened to its idle current on the dipoles, holding the 3.6 s slot in the sequence while removing the field ramp, so that the eddy-current correction could be evaluated. The flattened precycle was assumed to act on the following cycle only by changing the main dipole field, through a static hysteresis offset and a decaying eddy current contribution.

Through manual adjustment of the main dipole field by matching injection field through B-Train measurements and first turn position, the SFT injection field was recovered to the same values as with the precycle active. Despite having manually compensated the main dipole field, we found that the single fixed-target MTE injection at 1×10^{13} the beam was dumped quickly after entering the ring, within 100 to 200 turns, at every attempt to inject. Beam survived at a reduced intensity of 5×10^{11} protons and was lost again as the intensity was raised. Full intensity was recovered only after the horizontal chromaticity across the injection plateau was lowered empirically through LSA, and the beam reached extraction like the nominal beam once the injection field offset of Chapter 9 was applied, with the resulting fixed-target extraction shown in Figure 8.11.

10.1 Isolating the precycle contribution

The recovery of the beam by a chromaticity trim pointed to the chromatic stability of Section 2.1.7.1 and to a sextupole contribution b_3, despite it being caused by the precycle in the main dipole circuit. To confirm the presence of a sextupole field contribution, an experiment was carried out to measure the chromaticity with the precycle active, and inactive on the main dipole circuit. In addition, a complementary study was carried out on the main quadrupole to assess beam stability in the absence of the precycle on that circuit, to assess the feasibility of compensating the quadrupoles or flattening the precycle in an analogous way to the main dipoles.

To isolate the effects of each circuit, the precycle was deactivated on the main dipole and main quadrupole circuits separately, while leaving the other circuits to precycle normally. In practice, this was done by keeping the circuit at its idle current across the 3.6 s slot in the supercycle, instead of ramping to 200 GeV/c, so that the surrounding cycles were unchanged and only the field of that circuit was affected. The quadrupole precycle spans the separate focusing and defocusing circuits, both held at idle when the quadrupole was deactivated. For this reason the supercycle was not simply suppressed by removing it directly from the supercycle and compressing the cycle sequence. The change in the main dipole field is seen on the closed orbit and on the measured field, treated in Chapters 6 and 9, the sextupole component b_3 of the dipole is measured here through the chromaticity, and the quadrupole gradient is measured through the betatron tune. The absolute value of chromaticity depends on machine settings

and beam conditions that drift over the course of routine operation and seldom match the value programmed in LSA, so it was measured with the precycle active and inactive in the same session, reported as the change between the two. Similarly, the absolute beam tune is intensity dependent, and was also measured in the same way and reported as a difference.

The chromaticity was obtained from the dependence of the tune on the momentum offset $\delta = \Delta p/p$. The momentum offset is a knob in LSA that changes the beam momentum by varying the RF frequency over a time range. For each measurement prior to starting the $\Delta p/p$ sweep, the closed orbit was centered by adjusting the main dipole field, to remove the orbit contribution to the tune. The δ was stepped in increments of 1×10^{-3} over a range of $\pm 4 \times 10^{-3}$, with five tune measurements recorded at each step at an intensity of 3×10^{12} protons. The measurement was performed on a 14 GeV/c MD cycle with a 3 s injection plateau, run behind the precycle and a further high-energy cycle of either 450 GeV/c or 400 GeV/c. This cycle was used as its 14 GeV/c plateau is longer than the 1260 ms plateau of the standard SFT cycle, which also allows the longer chromaticity evolution across the plateau to be measured. The tunes for each δ were extracted with the NAFF algorithm [84] from the horizontal and vertical spectra of the BBQ system, which are logged and retrieved independently from NXCALS. With the precycle inactive the tune varies strongly and nonlinearly with the momentum offset, so a fixed frequency window does not isolate a single tune across the scan, and both tunes appear coupled in each of the two acquisitions. The fifty strongest frequencies of each measurement were therefore grouped into two tunes by k-means clustering [66], the centroid of each cluster taken as the tune, and the larger tune assigned to the horizontal plane and the smaller to the vertical.

For the tune measurement, the same beam was used, but over a standard SFTPRO injection plateau of 1260 ms. The same acquisition was used as for the chromaticity measurement. The tune in each plane was identified from a fixed frequency band, with a continuity constraint between successive slices to prevent switching between peaks. Without the momentum sweep the tune varies little across the plateau, so the k-means clustering used for the chromaticity measurement was not required.

10.2 Dipole precycle

The main dipole field at injection depends on the branch of the hysteresis loop set by the preceding cycle. Deactivating the dipole precycle leaves the magnet on a different branch and reduces the field at injection by 8×10^{-5} T, seen for the SFTPRO plateau in Figure 6.14, as compared to Figure 6.12a with the precycle enabled. In addition to the remanent field, the precycle also introduces the eddy-current decay entering injection. With the decay partially mitigated in Chapter 9, the static field shift in that experiment left the first turn at 3.3 mm and the orbit out at 5.9 mm, recovered by the field offset for hysteresis.

With the beam centered by this offset without the dipole precycle active and confirmed through BPM readings, the chromaticity was measured. The measurement was taken in the case following a 400 GeV/c SFT cycle and a 450 GeV/c LHC cycle, which consitute the standard sequences of physics operation. Deactivating the dipole precycle was found to shift the horizontal chromaticity positive and the vertical negative. Table 10.1 gives the linear chromaticity with the precycle active and inactive for the two preceding high-energy cycles, Figure 10.1a shows the tune against momentum offset from which it is obtained along with a 3-polynomial fit, and Figure 10.1b the same with the linear component removed. The magnitude of the change depends on the preceding cycle while the sign does not, so the chromaticity shift found by disabling the dipole precycle is present regardless of what runs ahead of it.

Tab. 10.1: Change in linear chromaticity when the dipole precycle is deactivated. The measured MD cycle runs behind the precycle, and the precycle behind a high-energy cycle that is either a 450 GeV/c LHC cycle or a 400 GeV/c SFT cycle.

(a) 450 GeV/c LHC cycle ahead of the precycle

	Precycled	Not precycled	Δ
ξ_x	-0.095 ± 0.014	0.066 ± 0.016	0.160 ± 0.021
ξ_y	-0.323 ± 0.011	-0.419 ± 0.015	-0.097 ± 0.018

(b) 400 GeV/c SFT cycle ahead of the precycle

	Precycled	Not precycled	Δ
ξ_x	-0.090 ± 0.027	0.153 ± 0.016	0.243 ± 0.031
ξ_y	-0.266 ± 0.021	-0.420 ± 0.016	-0.155 ± 0.027

The change in chromaticity confirms the presence of a sextupole component change driven by the dipole precycle, which is the first allowed multipole of a dipole magnet listed in Table 3.1, and is suggested to be a hysteresis contribution set by the precycle.

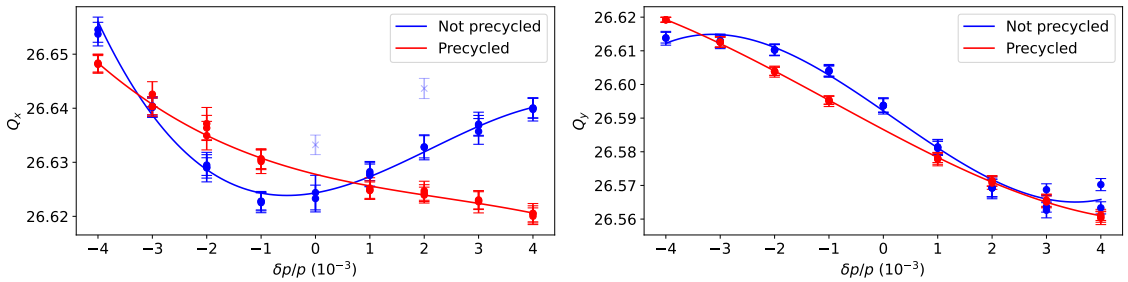
The measurement was taken below transition on the 14 GeV/c MD cycle, where a small negative chromaticity is required for beam stability, opposite to the above-transition case of Section 2.1.7.1. With the precycle active as in standard operation, the horizontal chromaticity is negative, and deactivating it drives the horizontal chromaticity positive and destabilizes the beam at higher intensities, whereas the vertical chromaticity is negative in both cases and does not destabilize the beam. Lowering the horizontal chromaticity in LSA allowed the beam to be injected at full intensity, and the trim is of the same sign as the shift measured here and was found empirically on the operational cycle.

Figure 10.2 shows the linear chromaticity along the injection plateau. A drift of the chromaticity is seen in both cases, but the drift in each measurement stays below 0.05, within the point-to-point scatter of the measurement, and does not change sign. The comparably high uncertainty of the chromaticity measurement in the without precycle case is due to the challenges in tune identification when the tunes are coupled and vary nonlinearly with momentum offset. Nonetheless, the change in chromaticity is clear, and the sign of the change is consistent with operational observation. The cause of the coherent drift of the chromaticity was not investigated as it is out of the scope of this thesis, but it is likely due to unidentified eddy current decay. Furthermore it can be seen that the nonlinear chromaticity, shown with the linear component removed in Figure 10.1b, is similar over the momentum range occupied by the beam and differs only at large momentum offset which operational beam does not reach, so recovering the beam requires only the linear chromaticity trim applied above.

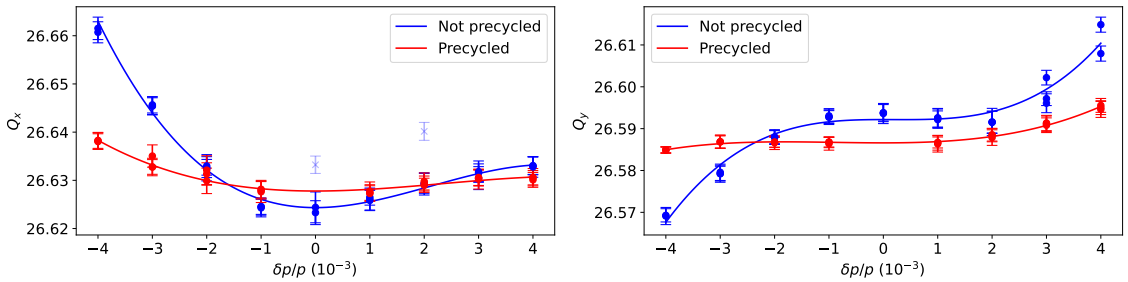
Matching the measured change in an optics model of the Q26 lattice, with a thin sextupole at the centre of each of the 744 main dipoles, gives an integrated strength of $3.096 \times 10^{-5}/\mathrm{m}^2$ per dipole. A change of the b_3 is also confirmed through laboratory measurement of a spare dipole magnet [17].

10.3 Quadrupole precycle

The tune with and without precycle was measured on the SFTPRO injection plateau, whose nominal working point is (26.62, 26.58). Deactivating the quadrupole precycle at this working



(a) Measured tune against momentum offset with a 3-polynomial fit.



(b) The same measurement with the linear component removed, showing only the nonlinear chromaticity.

Fig. 10.1: Horizontal (left) and vertical (right) tune against momentum offset on the 14 GeV/c MD cycle following an LHC cycle, measured with and without the dipole precycle.

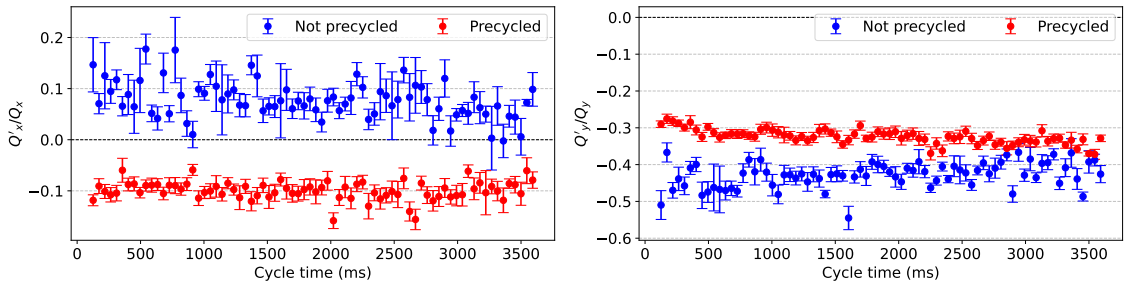


Fig. 10.2: Evolution of the linear chromaticity along the injection plateau of the 14 GeV/c MD cycle following an LHC cycle, with and without the dipole precycle.

point was found to lose the beam, recovered only by trimming the tune in LSA, which pointed to a tune shift toward the half-integer resonance at 26.50 which drive beam instability, as discussed in Section 2.1.6. The working point was therefore raised to (26.70, 26.62) for the measurement, so that the deactivated case stays clear of the resonance, and adds an additional separation of 0.04 between the horizontal and vertical tunes to simplify their identification.

Deactivating the quadrupole precycle shifts both tunes down by 0.08, shown in Figure 10.3, with the same shift in both planes. The focusing and defocusing quadrupoles are the same magnet type and share the same hysteresis, so the integrated gradient of both families is reduced together when the precycle is deactivated, and both tunes shift down. The tune also drifts across the plateau in each condition, from a combination of the space-charge tune shift under beam loss and the eddy-current decay, which are not separated here, as the quantity of interest is the static offset between the two conditions. Matching this shift across the 216 main quadrupoles in

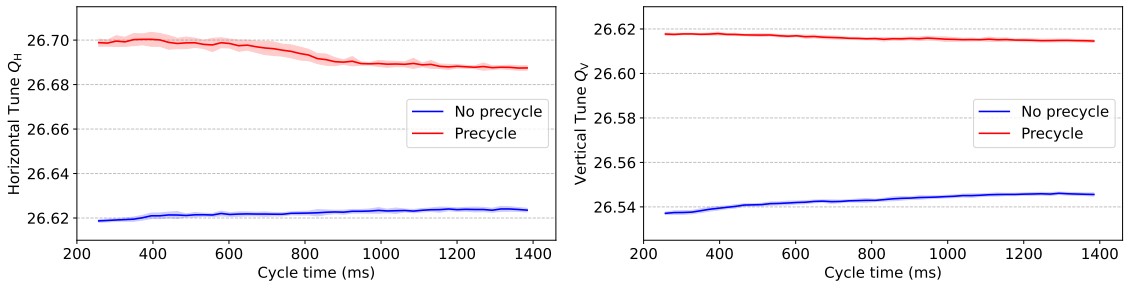


Fig. 10.3: Measured horizontal and vertical tune along the injection plateau with and without the quadrupole precycle. The shaded area shows the 1σ uncertainty of the tune measurement.

the Q26 optics lattice gives a gradient change of $1.05 \times 10^{-3}\,\mathrm{T/m}$, or 48 ppm of the maximum gradient 21.7 T/m. This is comparable to the 40 ppm relative shift of the main dipole field from its own hysteresis, $8 \times 10^{-5}\,\mathrm{T}$ of 2.0 T, so the precycle leaves a similar relative offset on both circuits. The allowed higher multipole of the quadrupole is the duodecapole b_6 of Table 3.1, whose contribution is too small to measure here and negligible for the short cycles that follow the precycle.

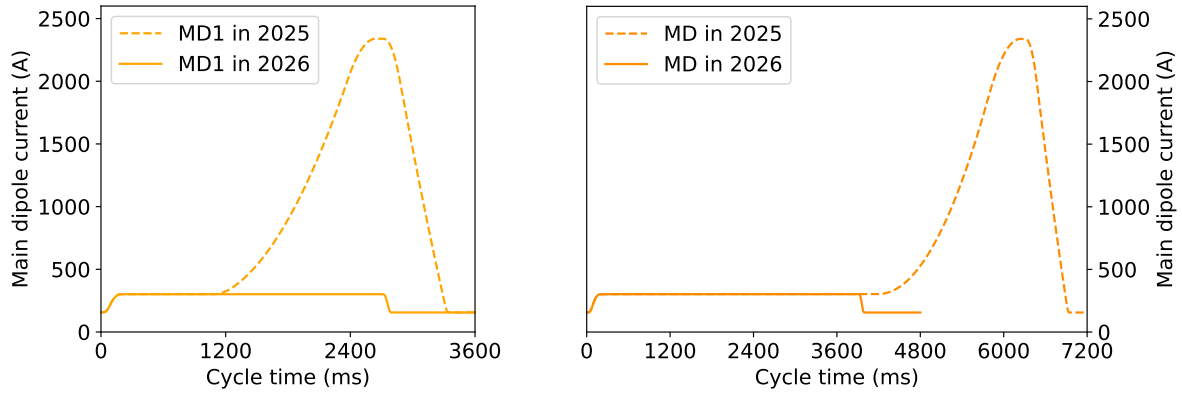
The tune shift is a static and reproducible offset set by the quadrupole hysteresis, and a single trim of the horizontal and vertical tune parameters in LSA corrects it directly. The quadrupole precycle can therefore be flattened in the same way as the dipole, with the working point trimmed once to absorb the offset, as was done empirically to recover the beam above.

10.4 Function and simplification of the precycle

The two deactivation experiments resolve the effect of the precycle at injection, the dipole precycle appearing as a shift of the field and the sextupole b_3, and the quadrupole precycle as a shift of the gradient. Its primary function has therefore been interpreted as method to reach a reproducible main dipole field, sextupole b_3, and quadrupole gradient at injection. Each circuit is placed on a reproducible branch of its hysteresis loop, and the magnetic history is not erased as the operational description of a reset implies. The second function of the dipole and quadrupole precycle is to provide a reproducible eddy-current decay by inserting a 3.6 s slot in the supercycle with a reproducible ramp, rather than allowing arbitrary preceeding cycles to be played. Deactivating the precycle removes this preparation, so the injected beam is perturbed by the dipole field shift, the chromaticity change, and the tune shift at once. This reproducibility is confined to injection, as the SFT flat-top field continues to change from one supercycle to the next even with the precycle active, as shown in Chapter 6 and corrected by the active compensation of Chapter 8.

The injection state is furthermore found to be depending only weakly on the preceding high-energy cycle. Changing the preceding cycle from an SFTPRO to an LHC cycle with a supercycle changes moves the injection field by about 0.4 mm, or $\sim 1.5 \times 10^{-5}\,\mathrm{T}$ equivalently, on the first turn, which is smaller than what can be resolved on the B-Train field measurement. The chromaticity already applied does not need to change as it keeps its sign across the two preceding cycles, as shown in the Section 10.2. Fixed-target proton operation outside of parallel MD is always preceded by a cycle at 400 GeV/c or above, so the injection state is operationally sufficiently reproducible across the sequences used in standard operation, with the short ion

run and arbitrary machine-development cycles the exception, where the MD1 precycle and SFT working point can be restored through a settings change.

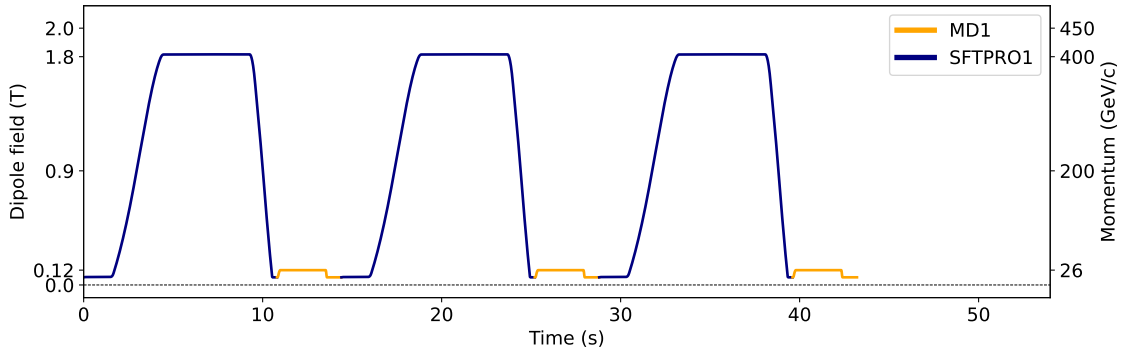


(a) MD1 precycle before and after flattening **(b)** Short parallel MD cycle before and after flattening

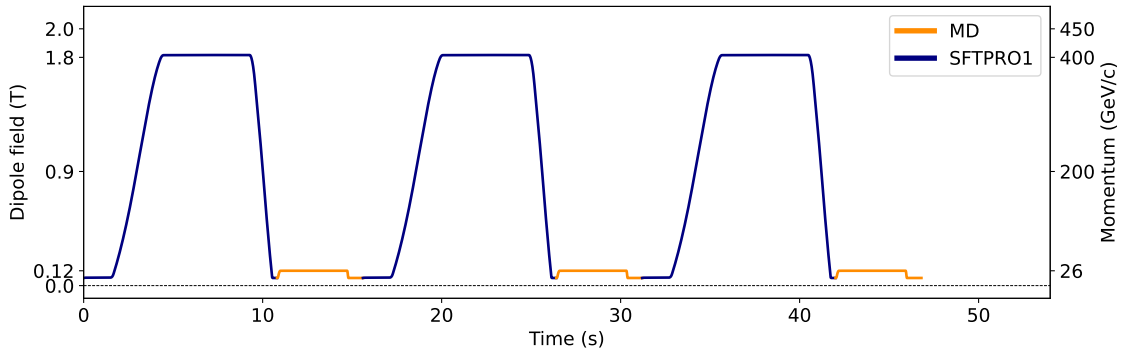
Fig. 10.4: Main-dipole current of the MD1 precycle and the short parallel MD cycle before and after flattening the 200 GeV/c ramp to the 26 GeV/c intermediate field.

As the effects of flattened precycle are reproducible across operational sequences, it can simplified to save energy. It cannot however be removed outright from standard proton operation, as the main dipole circuit has a RMS current limit of 3240 A that the standard fixed-target supercycle already reaches, and running the fixed-target cycles back to back without the 3.6 s slot would exceed it. Outside the scope of this thesis, and carried out by operation, the precycle was flattened to the 26 GeV/c injection plateau on the main dipoles and the equivalent field on the main quadrupoles, and with the SFT injection settings on main B, tune and chromaticity adjusted as described in Sections 10.2 and 10.3. In this configuration, the precycle retains its slot in all operational sequences, while removing the field ramp. The choice to reduce to 26 GeV/c rather than to idle current was made in conjunction with removing the same ramp from the 7.2 s short parallel MD cycle, which also inject beam at 26 GeV/c and is dumped at the end of the plateau. This allows both cycles to follow the same minor hysteresis loops with the ramp to the plateau and back to idle current to maintain injection reproducibility whether in parallel MD or normal operation. The dipole excitation current of both cycles, before and after flattening, is shown in Figure 10.4. It can be seen in the case of the parallel MD cycle that by removing the end of cycle ramp, and shortening the 26 GeV/c plateau from 3015 ms to 2900 ms, the MD cycle is shortened from six basic periods to four, saving 2.4 s in cycle time for every occurrence. The MD cycle is not shortened since the cycle time is necessary for standard operation to stay above RMS current limits. The higher-order circuits in either cycles are unaffected by the flattening, since their quasi-degauss of Section 4.6.2 is applied at the end of every cycle, including the flattened precycle.

Flattening the precycle saves energy while keeping the slot, and was estimated to have saved 4584.8 MW h over the 2025 run when replacing each pulsing precycle with its flattened version for the dipoles and quadrupoles. The supercycles for normal fixed-target physics operation and the same with parallel MD is shown in Figure 10.5, against the prior sequences of Figures 4.2 and 4.9 with the more energy consuming precycle field ramp. Notably in Figure 10.5b it can be seen that as a result of the shortening of the MD cycle, the supercycle length reduced from 54 s to 46.8 s, improving duty cycle by 13.3 % in addition to energy savings, for a tradeoff of 3×115 ms of beam time. While not enabled through active field compensation, this operational



(a) Standard physics supercycle with the flattened MD1 precycle



(b) Standard parallel MD supercycle with the flattened MD cycle

Fig. 10.5: Supercycles after flattening the precycle and the short parallel MD cycles to the 26 GeV/c intermediate field, compared with the sequences of Figures 4.2 and 4.9.

outcome partially meets the energy and duty cycle goals set out in Chapter 5, even though the MD1 precycle could not be suppressed in cycle time, which is inherent from limitation in the RMS current in either power converter or magnetic cooling.

The flattening and operational validation of beam reproducibility was carried out in beam commissioning by operation after the period of this thesis, and follows from the precycle reproducibility findings in Sections 10.2 and 10.3. The consequences of this change in operational working conditions for the active compensation of Chapter 8 are significant, and are taken up in the discussion of Chapter 11.

Chapter 11

Discussion

11.1 Attainment of the Stated Goals

Chapter 5 set four goals for this work, a reduction in the energy and beam-time cost of the MD1 precycle, an increase in the SPS duty cycle, the removal of manual operator correction on sequence change, and support for more flexible cycle scheduling. Field reproducibility within the tolerance ε of Section 5.3, together with the three validation goals of Section 5.4.7, was adopted as the quantitative criterion for these goals. Reproducible fixed-target extraction and its automation were attained for standard operation, while compensation at injection and operation across arbitrary sequences were not, and the energy and duty-cycle gains followed from an improved understanding of the precycle without active field compensation.

Automation of the fixed-target correction was demonstrated in Chapter 8, where the loop ran without operator intervention across the operational proton supercycles and reproduced the extracted spill to a quality operators would not act on. This was attainable because the fixed-target flat top requires only relative reproducibility of the field against a per-cycle reference, and because the correction references the first model prediction of each cycle so that systematic model bias cancels in the difference. Since the dominant cycle-to-cycle error at that point originates in the main dipoles, a dipole-only correction reaches the operational sensitivity of the spill, and the first validation goal is met for standard operation, with reproducibility established at beam level and not against an independent absolute field. Running unattended was made possible by integrating the correction into the existing control system through its shared timing and settings hierarchy, so that no dedicated instrumentation was required and the loop stayed below the operators' own settings, and the same reliance on the shared control structure provides the extension to the other magnet families.

Partial attainment of the energy and duty-cycle goals followed from the precycle reinterpretation of Chapter 10, where flattening the dipole and quadrupole precycle to the injection plateau removed its field pulse together with the associated energy and shortened the parallel machine-development cycle. These gains were adopted in operation after the period of this work, and they reduce the beam-time cost of parallel development, while the precycle slot is retained during normal physics because its removal is prevented by the RMS current limit of the dipole circuit. As the change is operational and not driven by active field compensation, the energy and duty-cycle goals are met in part only.

Compensation at injection, the second validation goal and the precondition for removing the MD1 precycle, was not achieved. At injection the beam is most sensitive to the field, and reproducibility there is degraded by the prediction error of the dipole model, by the drift of the seeding B-train measurement, and by a decaying eddy-current field that the dipole model does not reproduce, as established in Chapter 9.

Flexible scheduling, the fourth goal, is blocked and is the most demanding of the four, since it alone requires reproducible fields on every circuit the beam is sensitive to and at every point in the cycle at the same time. It therefore binds together the absolute injection accuracy, the quadrupole

and higher-order correction of Section 11.8, and a generalized eddy-current correction, each which did not reach operational usability. The control loop was built to apply these corrections jointly, yet neither the dipole field model nor the eddy-current model reached a deployable accuracy at injection, on arbitrary sequences or even on the operational ones, so arbitrary scheduling is not supported.

Generalization to sequences absent from routine operation was shown for individual cases and not established systematically. Asymmetric supercycles and autoregressive operation over novel orderings of operational cycles were compensated in Chapter 8, which supports the third validation goal, although no systematic study across arbitrary sequences was carried out. Validation covered the operational sequences by design, in line with Section 5.4.7, and the overhead of constructing fully arbitrary sequences under the present supercycle system placed them outside the available measurement time.

11.2 Measurement Reproducibility and Identifiability

In this work, the availability and reproducibility of the magnetic measurements remained the biggest limitation on the model's accuracy and downstream applicability. Despite the drawbacks and still significant drift of the B-Train, shown in Chapter 6, it remained the only measurement available for the main dipoles.

In the study shown in Section 6.2.3, on the long LHC filling cycle, the integration drift reaches the 5×10^{-4} T scale of the hysteresis effect being modeled, before the cycle even starts ramping. With no independent reference to separate the two, a supervised fit learns the sum of the true signal and the numerical drift. A fit on such measurements therefore predicts the drift as part of the field, and adding more drifting data trains the model further toward the drift and not the field beneath it. The limit is the drift being unidentifiable from a single measurement and not the quantity of data available. On static plateaus such as the SFTPRO extraction, the field can vary in magnitude from prior accumulated drift, and this appears to the model as a quasi-static effect that cannot be distinguished from true hysteresis. Because of this, the model was deliberately trained on measurements whose drift was empirically determined to be in an acceptable range, and while this constrains the model to mostly clean measurements, it also limits the model's robustness to drift in the inputs.

This limitation has practical implications for the operational deployment of the model, specifically since the drift of the seeding B-train changes magnitude and sign between hours and days. This drift propagates directly as uncertainty into the prediction and correction calculation, and biases the correction applied to the next cycle, a dependence developed in Section 11.6. The same sensitivity also contributes to the gap between the offline validation accuracy of Chapter 7 and the performance observed under live seeding in the operational tests of Chapter 8.

Lowering this limit requires a model that estimates the drift explicitly or a better measurement. A state-space model that represents the drift as a latent state alongside the learned field could separate the two and predict a drift-free field, although that field has no reference against which to be validated. The alternative is to improve the measurement itself, either by reducing the drift or by increasing the frequency of absolute anchoring so that the integration drift is constrained to a smaller range, since drift from a fluxmeter is unavoidable. Neither was pursued here, since the B-train is an operational system outside the control of this work, with the instrumentation and surrounding infrastructure fixed.

11.3 Data Variety and Generalization

The secondary limitation of the work remains the generalization of the model to arbitrary magnetization and sequences, which was blocked by the limited variety of the training data. Compensation of orderings and recombinations of cycles already present in the training data was demonstrated to a limited extent on slightly differing sequences in Chapter 8. Demonstration on arbitrary sequences was neither attempted nor supported by the model and the data. Extrapolation to a cycle shape or an operating point absent from training is the capability that full flexibility requires, and neither data source trains nor tests it.

Operational cycles cover only a few narrow bands of the excitation and history space, fixed by the injection and extraction levels each beam requires. Operators vary the injection current or the extraction field only slightly around these levels, and the measured cycles therefore differ by small modulations on otherwise similar paths. Between the precycled and non-precycled injection states of Chapter 7 no measured data lies, and the model is unconstrained in the gap between such bands.

The flattened precycle and MD cycle of Chapter 10 are an instance of this missing sequence variety. Although the flattened field is at a level that the machine normally operates at, the sequence it forms with the following fixed-target cycle was not measured. The training data contains samples with the pulsing precycle absent but none of the new flattened precycle, so the model cannot be immediately qualified for that configuration. This operating point arose in operation only after the present work, and updated training data was therefore not available within its scope.

Arbitrary and pseudo-random excitation would have added the required variety to the input space, yet it was to come from the dedicated laboratory campaign that the measurement constraints of Chapter 6 made impossible, leaving only the sequences operation happened to play. Older magnetic measurements from the logging system of earlier operation were also intended as a source, but improper integration-marker application and other artifacts in those records left only the current operational period usable. While shorter measurement sessions were possible during dedicated machine development time to schedule more arbitrary or normally unplayed sequences, the amount of data and variety that could be collected within the constraints of beam operation was insufficient to train a model that generalizes to arbitrary sequences.

On this basis, generalization to arbitrary sequences is unsupported and not merely untested, since the histories that would establish it are absent from both data sources. The model is sufficient for the operational sequences present during this work, while any new cycle or operating point requires an update before it can be accurately compensated. Retraining or fine-tuning becomes possible once such a configuration has been exercised long enough to be recorded, and the flattened precycle now present in operation is the immediate candidate for this update.

11.4 Eddy-Current Decay as a Distinct Effect

The single operator of Equation 5.1 was intended to carry the eddy-current decay together with the static hysteresis, and Chapter 9 shows that one model fitted to a magnet measurement cannot hold both. The B-train and the laboratory bench measure the decay of a magnet in a specific measurement bench, while the beam integrates the decay of the whole ring, whose eddy currents flow in the vacuum chamber and the grounded structures that no magnet measurement contains. The decay the beam experiences is therefore distinct from the one the fitted model reproduces, so the beam contribution was modeled separately and the magnet term subtracted from the prediction to avoid counting it twice. The separately fitted beam model is itself limited in accuracy and generality, as shown in Chapter 9, so the separation of the two effects is established

as a mechanism and not yet as a correction deployable on arbitrary cycles. A field feedback on the B-train planned for the main dipoles, discussed in Section 11.7, would not reach this beam-seen decay, since it regulates a magnet measurement that does not sample the conductive structures of the ring.

With the eddy decay handled separately from hysteresis, the correction is needed only at injection and over the early part of the ramp, because the cycle's own ramp sets the dynamic field once it begins and the inter-cycle decay becomes negligible in comparison. Since the intra-cycle eddy currents are reproducible, they are absorbed by the static cycle settings and do not need to be feed-forward corrected. Consequently, the window of correction which in the current implementation is paired to the hysteresis model, should be decoupled in a more realistic implementation.

While the separation of the two effects adds complexity, it also opens the path to a simpler hysteresis model. The proposed model had several limitations imposed on it through the requirement to capture transient effects, such as high-rate sampling for eddy current decay in both context and horizon, in combination with long horizons for hysteresis, and the need to have cycle ramp downs in the context to predict the decay of the following cycle. Without the dynamic effects, the model would no longer need be regularly sampled, and could predict a smaller set of collocation points $B(I)$ along the period of the cycle where beam is present, which is only on the up-ramp of any given cycle. The already known programmed current could then simply be interpolated between those points. Furthermore, magnetic hysteresis is only dependent on the extrema of the excitation, which are the flat top of each cycle and idle current for operational cycles. Since idle current is always the same, the model could be reduced to ingesting averaged past flat top fields only, capturing one point per cycle, and allowing the model to be seeded with a much smaller history and fitting a 15 min relaxation of the SFTPRO cycle in standard physics with only in the order of 60 points rather than 45 000 at 50 Hz. This approach would however require labels of the future field without the eddy current decay, which are by definition captured in pulsed measurements and not trivially separable despite our efforts to do so in Chapter 8. Obtaining static field measurements following operational cycles or any arbitrary sequence would require a dedicated measurement campaign, and it, along with the simplified model is left as future work.

11.5 Scope of the Learned Model

The model was framed in Chapter 5 as a digital twin of the magnetic circuit, a single representation of the effective history-dependent response carrying hysteresis, eddy currents, and geometric nonlinearity together. What the fitted model represents is narrower, since it is a twin of the main field as recorded by the B-train and not of the magnet nor of the field the beam integrates. Furthermore, the premise rested on one model per circuit correcting the main field of that circuit, which Chapter 10 shows to be insufficient for the dipole, since the dipole field also includes a history-dependent b_3 that a main-component model neither represents nor compensates. This concerns the dipole alone, because the lowest systematic higher harmonic allowed by symmetry rises in order with the family, from b_3 in the dipole to b_6 in the quadrupole and b_9 in the sextupole, as listed in Table 3.1. A harmonic contributes less at the beam aperture as its order rises, and its action moves from the chromaticity shift of the dipole b_3 to the weak high-amplitude detuning of the twelve- and eighteen-pole terms, so the history-dependent part of b_6 and b_9 is negligible against the reproducibility tolerance.

The main limitation of the model besides the measurements used to train it, remains the structure it is imposed to learn. As a digital twin of the circuit, the model must reproduce the

full hysteresis loop over any cycle and any horizon to within the reproducibility tolerance, which in Figure 7.1 is shown to be a few tenths of a percent of the 6 mT loop opening, up to eight percent for the largest cycle-to-cycle modulation. The smallest sub-percent accuracy required would be a challenge for any model, while the down-ramp of the cycles are only useful for full field reproduction and not for the beam. Meeting this everywhere is als demanding, and the accuracy degrades on the parts of the loop and the horizons least represented in the data. Therefore, in this case the requirement of the model to be a digital twin is in conflict with the operational necessity of the model, and reducing the model to a specialized predictor on the beam-relevant field and working cycle-by-cycle is a more attainable goal. Keeping only the up-ramp parts of the loops would allow the cycle-to-cycle modulations to gain a magnitude of resolution in the numerical range of the model.

This directly supports and extends the proposed solution of Section 11.4. The suggested model is better cast as a specialized predictor of the beam-relevant field than as a digital twin of the circuit, and this specialized form is the one attainable with the measurements available, besides improvement in the measurement fidelity.

This narrowing of scope recurs across the results and is itself an outcome of the work. The eddy-current decay of Section 11.4 and the sextupole contribution of the precycle established in Chapter 10 were each expected within the single learned model, and in each the resolution came from measurement and diagnosis instead of added modeling capacity, the first through a separate beam-based model and the second through a fixed correction derived from understanding the precycle. Together the measurement campaign and the beam studies that produced these distinctions delineate what a learned model should carry, as much as they support the model fitted here.

11.6 Control-Loop and Deployment Limitations

In the B-train seeded control implementation of Chapter 8, the feed-forward scheme reads the field of a cycle only after that cycle has played, since the realized current and field are published to the control system only once the cycle completes. At the forewarning moment for the next cycle the most recent entry in the history window therefore still contains the field of that cycle's previous occurrence, one occurrence behind the state the next prediction should start from. For the operational sequences tested this stale entry is not a problem, since physics cycles normally do not run back-to-back in the SPS. A supercycle that plays such cycles immediately after each other removes that margin, as planned for the future SHIP operation, where the seed would then be wrong by one occurrence on every cycle. The only available solution is to run the model in autoregressive mode, where the model's own prediction is available for the next step.

The model supports this mode by construction, since it predicts the same field quantity it is seeded on and its own prediction can therefore serve as the next seed, while running without the measured seed lets its error accumulate across cycles with nothing to correct it. Because the model is specialized on the operational sequences, it carries prior information to predict the next cycle from past cycles by cycle identity and not only from the more generic past excitation, so its autoregressive output tends to stay anchored to the learned cycle response and away from a generic point-by-point unrolling. This anchoring is reliable on sequences close to those seen in training, while on an unseen combination of cycles it rests on the pretraining of Chapter 7, where the models without transfer diverge under autoregression and only the pretrained model stays near the seeded level. The operational run of Section 8.2.5 tracked close to the seeded case on novel orderings of operational cycles, as a single trial without a matched baseline. Seeding nonetheless remains the robust choice wherever the B-train is available, since re-grounding each

cycle prevents any accumulation, and the extended operational test was run seeded for this reason. Autoregression is reserved for the cases seeding cannot cover, the back-to-back cycles where the seed would be stale and the higher-order families that have no online measurement, together with badly drifting days on which the drift-free prediction can improve on the seed. Extended unattended operation in that mode across many cycles is not established, bounded by the same data variety of Section 11.3.

11.7 Planned Feedback in the B-train

Field regulation on the B-train was not previously possible, since the FGC did not accept input from the B-train, as described in Section 4.5.2. Hardware support was added in 2025, so a feedback loop can now be closed in field, although such a loop would inherit the integration drift of Chapter 6 and regulate on it, an effect made worse by the long SPS cycles when compared with the shorter PSB and PS cycles that already regulate in field. A fast-locking NMR probe is to be commissioned to track the field on plateaus in place of the single per-cycle marker, constraining the drift between anchors so that field regulation becomes viable.

A further limitation of the current B-train is its integrator offset, which is recalibrated on a ZERO cycle with the magnets at idle, as described in Section 4.5.2. The step is a recalibration and not a zeroing, since a separate hardware offset remains, and it times out after about three hours before it is run again on the next ZERO cycle. For the injectors the dependence is not a rpblem, since unfilled beam slots occur often and serve as ZERO cycles, whereas the SPS is kept filled as much as possible and schedules a ZERO only for deliberate spacing, such as before an LHC filling cycle. When no ZERO is scheduled the recalibration can be forced by manually shorting the front-end timing cables, although this is a manual intervention and cannot be relied upon in operation. A continuous field-control loop cannot rely on an offset corrected only on a scheduled no-excitation slot and valid for a few hours, so the planned scheme must handle the offset without it.

For the main dipole circuit, such a loop could in future correct the measured field directly and reduce the dependence on a predictive model. The model remains necessary where a loop cannot act, since the higher-order families carry no online measurement for a loop to close on, a feed-forward correction anticipates a sequence change before the cycle plays where a reactive loop corrects only after the beam is perturbed, and a loop regulating the main field corrects only the main component and not the history-dependent b_3 of the dipole established in Chapter 10 or the higher-order circuits. Open-loop prediction for the unmeasured families must in any case be validated against the B-train on the main dipole, the only circuit where a prediction can be checked against a measurement, so the two approaches are complementary and the measurement stays central whether or not a loop is in place.

11.8 Extension to the Quadrupole and Higher-Order Families

Compensation of the quadrupole and higher-order circuits formed part of the original hypothesis of Section 5.4 alongside the main dipoles, and the family-agnostic scheme of Section 8.1.6 was designed to accommodate them, with the integrated gradient taking the place of the dipole field in the settings hierarchy. A trained field model was realized only for the main dipoles, since the measurements required for the other families were not attainable within this work, and no compensation of the quadrupole or higher-order circuits is delivered.

Active compensation of these families is nonetheless a precondition for the flexible scheduling set as the fourth goal, since arbitrary sequences require reproducible fields across every circuit

the beam is sensitive to and not only the main dipole field. The sextupole and octupole circuits carry weaker fields and already apply a quasi-degauss at the end of every cycle, as described in Section 4.6.2, so their precycle contribution is expected to be small although this was not quantified, and they were given lower priority and remained blocked by the same measurement obstacles as the quadrupoles.

11.8.1 Measurement-Fidelity Limitations

Unlike the dipole case, the RMS current limit that falsified the laboratory dipole method of Chapter 6 is not the obstacle for the quadrupoles, since their excitation is below the RMS limit. The laboratory method is therefore in principle playable for this family, and the blocks instead arise from measurement fidelity specific to a gradient field. A quadrupole gradient is obtained by subtracting the signals of two horizontally offset coils, which cancels the common dipole-like component only when the coils are geometrically identical. Hand-wound coils of slightly different geometry from the old measurement left that component uncancelled and contaminated the gradient, and a new printed circuit board coil was built to reach the matched geometry its tolerances allow.

Significant drift persisted even with a fast digital integrator operating at high frequency rather than integrating in postprocessing. Disarming the HOLEC power converter removed a large biased component of this drift and large voltage peaks in the measurement, which links it to the converter, since the coil is effectively a large pickup loop and the converter-driven electromagnetic field is integrated into the measurement. Instead of subtracting the signals in postprocessing, the two coils were difference in hardware to cancel the common-mode pickup, which reduced the drift further, but still remained over the scale of the reproducibility target.

Beyond the integration drift, the main problem for the quadrupoles is the absence of an absolute field marker, since an NMR probe requires an approximately homogeneous field to lock and the quadrupole aperture carries a gradient that is zero on axis and grows with radius, so no resonance is obtained where the beam-relevant gradient is measured. A probe placed in the outer yoke region, where the local field is approximately uniform, would sample a field that is not the aperture gradient, and an absolute reference is therefore unavailable. The absolute-measurement requirement of Section 5.4.5 is consequently unmet, so the quadrupole measurements remain relative, which is the main scope limitation for this family where the lab dipole measurements was first disqualified on reproducibility.

Hall probes mounted above the fluxmeter as an alternate anchor did not reach the accuracy needed to recover a reproducible field alongside the drifting fluxmeter, and a Kalman fusion of the local Hall field with the gradient coil, of the kind in [8], is bounded by that same Hall accuracy. A point Hall probe is in any case poorly suited to a gradient, where position and roll errors couple directly into the measured gradient and the sensors of a three-axis package sample different field points, so quadrupoles are conventionally measured with rotating coils or stretched wire.

A reproducibility tolerance for the quadrupoles was also left undefined in Section 5.3, where the dipole tolerance followed from the orbit response of Section 5.2.2 and no equivalent transform exists for a gradient error. A partial estimate follows from the operational retune threshold, since a tune error of 0.01 prompts a retune of the machine and corresponds through the optics to approximately 6 ppm of the maximum gradient, comparable to the 5 ppm that the $1 \times 10^{-5}\,\mathrm{T}$ dipole target represents against its 2 T field. The 0.08 tune shift measured when the quadrupole precycle is removed corresponds for comparison to 48 ppm of the $21.7\,\mathrm{T/m}$ maximum gradient, as reported in Chapter 10. The 6 ppm target translates to a gradient reproducibility of $1.3 \times 10^{-4}\,\mathrm{T/m}$, and measurements of the quadrupole reached only $1 \times 10^{-3}\,\mathrm{T/m}$ down to

upper limits of $1 \times 10^{-4}\,\mathrm{T/m}$, so the measurement fidelity is almost an order of magnitude above the target. Furthermore the measurements remained relative without the absolute reference, so they could not be used for modeling. The main obstacle for successful quadrupole modeling and compensation therefore also remains in measurement fidelity, set by the gradient geometry, the electromagnetic pickup, and the missing absolute reference, and not one of the power converter hardware that limited the dipole. The improvements in measurement fidelity is a precondition for extending the method to the quadrupole and the other higher-order families, left as future work.

11.8.2 Proportional Hysteresis Excitation in Dipoles and Quadrupoles

In one case observed in operation, a supercycle change that normally perturbs the extracted beam produced no visible change, because the preceding cycles excited the quadrupole and dipole circuits in the same proportion and the normalized gradient was suspected to be preserved. A Q26 LHC cycle at $450\,\mathrm{GeV}/c$ was run in an asymmetric sequence with two $380\,\mathrm{GeV}/c$ ion fixed target cycles in operation, a configuration that shifts the spill under normal Q20 LHC operation, analogous to the mechanics seen in Figure 8.8. A momentum change of $0.81\,\mathrm{GeV}/c$ was measured from the flat-top revolution frequency over 20 of each cycle, consistent with the change in dipole field change recorded by the B-train and, while the spill stayed constant across the change.

Because a hysteretic dipole change appears as a momentum error through the action of the radial loop, and because a constant slow-extraction spill requires constant optics at the start of extraction under the COSE scheme of Section 2.1.9.1, the preserved spill implies a quadrupole gradient change in the same proportion as the dipole, even though the gradient is not measured directly in operation. This inference is supported by the precycle study of Chapter 10, where removal of the precycle shifts the quadrupole gradient by 48 ppm against a 40 ppm dipole shift, and where the tune and the dipole field are co-scaled through the control system so that a momentum change propagates to both.

Shared iron of differing geometry between the two families, together with the proportional excitation above, makes a transfer of the learned response plausible. Because the quadrupole stays well below the RMS current limit, it can be driven in the laboratory with the arbitrary and pseudo-random excitation the dipole cannot receive, so a quadrupole model trained on that variety and adapted to the dipole is a route to the generalization coverage the operational dipole data lacks. A dipole model could equally be fine-tuned on limited quadrupole measurements and rescaled to the gradient, since the model learns normalized values, and both directions are gated on raising the quadrupole measurement fidelity to the required level.

This proportionality is a property of the sequences operation happens to play and not a general guarantee, and it also carries a direct consequence for partial compensation. Correcting the hysteresis of one co-scaled circuit while leaving the other uncorrected breaks the preserved ratio, so the extraction optics is no longer constant, and a dipole-only correction can then degrade an optics-sensitive beam where none was degraded before. Compensation of the quadrupoles is for this reason coupled to compensation of the dipoles for the slow-extracted beams, and cannot be deferred without this qualification.

11.8.3 Seeding Without an Online Measurement

The higher-order families carry no online field measurement, so the seeding or periodic re-grounding from the B-train that stabilizes the dipole compensation has no equivalent for them, and a model run open loop eventually diverges from the realized field and must be re-seeded or precycled again to a known state. A signal that encodes the magnetic state is therefore

needed at deployment. Since a hall probe, fluxmeter or an NMR probe cannot be installed in the tunnel due to operational, equipment, and radiation environment constraints, the only parasitic signal available is the power converter voltage. The voltage of the power converter is measured internally by the FGC for the current regulation loop, and is published to the middleware at 1 kHz. Since the powering coils enclose the aperture of the magnet, they link the same flux the beam experiences. By Faraday's law the inductive voltage across the coil follows the rate of change of that flux, and its integral returns the field realized in the aperture, including the hysteretic and eddy-current contributions.

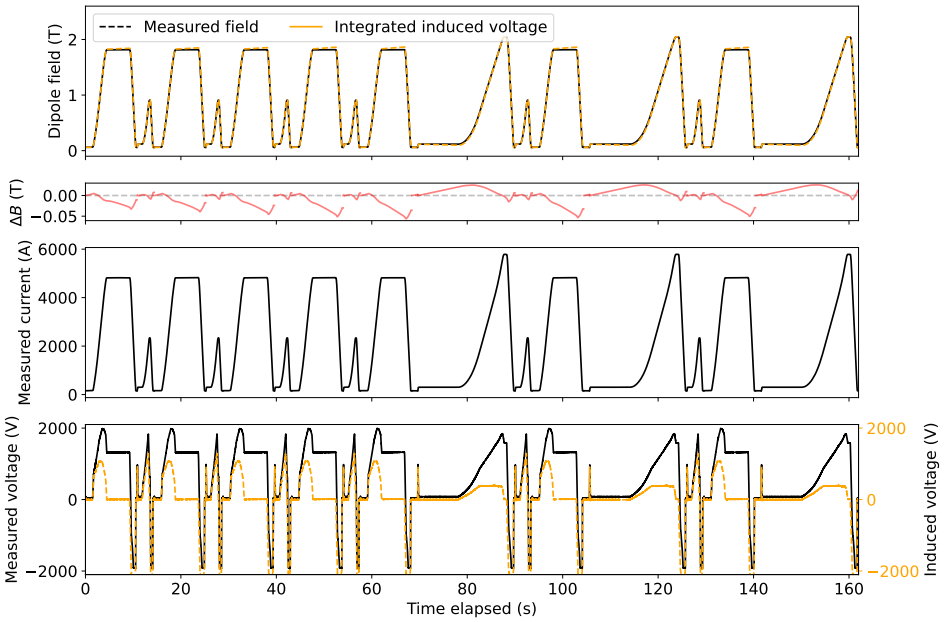


Fig. 11.1: Power converter voltage compared with the main dipole field over a sequence of operational cycles. In the top panel the measured field and the integrated induced voltage calibrated to the field are overlaid, with their difference ΔB in the second panel. The measured excitation current is shown in the third panel, and the measured terminal voltage with the induced voltage, obtained after subtracting the resistive contribution, in the bottom panel.

That published voltage is the full terminal voltage and contains both the resistive and the inductive contribution. Only the inductive part relates to the field, and separating the two is not straightforward, because the coil resistance depends on temperature and magnetoresistance and changes with the excitation and the operational mode even within a cycle, so a constant resistance can only be a first approximation. Figure 11.1 compares the integrated induced voltage with the measured field, where a difference between the two remains after the resistive part is subtracted. With a best-fit constant resistance of 0.272 187 Ω for the one of twelve powering circuits measured, the reconstruction remains about 0.05 T from the measured field, two to three orders of magnitude above the accuracy beam control requires. The constant-resistance reconstruction therefore cannot reproduce the main field, even though the measured circuit is representative of the machine.

Besides the direct field reconstruction, the voltage is nonetheless usable as an additional model input once a constant resistive part is removed, since a model that already knows the operational sequences can learn the hysteretic contribution carried in it. Combined with a precycle of the circuit to a known state from which the prediction can unroll, this gives the higher-order families

a seeding path that does not depend on a field sensor in the magnet. As a seed the voltage however inherits the stale-history limit of Section 11.6, because a cycle's record is complete only once that cycle has played, so the higher-order families carry the autoregressive accumulation and the off-by-one seed together. Neither the voltage input nor the precycle seeding was tested, and both are left as future work.

11.9 Outlook

The validation used proton beam, and the fixed-target ion cycle was therefore not covered. Since it undergoes the same chromatic slow extraction and carries the same dipole hysteresis, the compensation applies to it without modification. Extension to the quadrupole and higher-order families is the larger step, gated on the measurement fidelity of Section 11.8 and on an online signal to seed them, for which the power converter voltage was proposed.

A fast-locking NMR probe is planned for the B-train to constrain the day-to-day integration drift that bounds the absolute accuracy at injection. Covering arbitrary excitation for training is a separate need, and since the dipole cannot be driven with such excitation in the laboratory, the quadrupole transfer of Section 11.8.2 is the main solution for it.

The physics the model represents and the measurement it is trained on can be improved separately and combined in one model. On the physics, the eddy and precycle findings allow a reduced representation that carries only the static hysteresis at a lower rate and over a shorter horizon. On the measurement, the drift could be carried as a latent in a state-space estimator and filtered from the reproducible field by its slow and session-coherent structure. A future model is then simpler in the hysteresis it represents and more structured in how it carries the drift, bounded by the same information limit as the measurement. At the most ambitious end the magnetic state is represented as an explicit latent whose dynamics are learned as a world model or joint-embedding predictive architecture, with the past field as context and the programmed current as the action that advances it. In such a latent the long hysteresis memory is retained beyond the attention window of the present model, and unrolling extends over many cycles. The error accumulation that limits the present autoregressive mode is then reduced. Accuracy stays bounded by the measurement, and whether the structured drift admits this representation is open.

The limitations above point to a model that represents the magnetic history through its excitation extrema and updates its state incrementally, in preference to extending the raw context, since the reduced representation reaches the full magnetic memory at a fraction of the input length and suits autoregressive running. Realizing it depends on the separation of the eddy decay of Section 11.4 and on static field labels from a dedicated measurement campaign, so it stands as a direction for future work and not a result of this thesis.

The beam-based eddy-current correction of Chapter 9 reproduces the decay as a mechanism but does not yet generalize, since the fit is under-constrained by the available beam diagnostics and an increase of the field above transition remains unexplained. A general correction would require orbit diagnostics that resolve the faster decay modes and a reference orbit driven by the main dipole alone, before it can act on the arbitrary cycles that injection compensation and precycle removal need.

Chapter 12

Conclusions

12.1 Summary of Contributions

The magnetic field realized by a cycle in a multi-cycling synchrotron such as the SPS depends on the sequence of cycles that preceded it, through the static hysteresis and the eddy-current decay in each of the main magnetic circuits. A cycle with a different history realizes a different field across the main and higher-order circuits, and the resulting beam parameters, such as the closed orbit, and momentum, tune, chromaticty, are then not reproducible between occurrences. Particularly in the SPS, the slow-extracted spill for fixed-target cycles is sensitive to those parameters, and the resulting variation in the spill in different operational sequences remains a partially mitigated problem. Three operational measures address this in the SPS, two specific to the north-area fixed-target beams, the MD1 precycle that sets a reproducible magnetic state at injection and the manual tune correction operators apply to restore the extraction after a sequence change, and a broader one, the restriction of supercycles to a small set of fixed sequences that keeps those changes infrequent. The precycle alone consumes close to a fifth of the proton run time and several gigawatt-hours of energy each year, as reported in Chapter 5, while the manual correction is reactive and undocumented and the rigid sequencing constrains what can be scheduled. None of the three removes the underlying history dependence, and together they do not scale to the more dynamic scheduling foreseen for the complex. Removing this cost without losing the reproducibility the three measures provide set the four goals of this work, a reduction in precycle energy and beam time, a higher duty cycle, the removal of manual correction, and support for flexible scheduling by allowing arbitrary sequences.

This thesis addressed that problem by predicting and compensating the realized field, an approach posed for every magnet family and implemented for the main dipoles, where usable field measurements were available. A predictive field model and a feed-forward compensation scheme were developed and validated with beam, and together they reproduce the fixed-target extraction to a level sufficient for operation. Coverage reaches the operationally relevant regime of extraction and standard sequences, and a general model for arbitrary excitation histories was not attainable with the available measurements. Alongside the compensation, a dedicated study of the precycle narrowed the cases that require active correction, and six contributions follow from this work.

1. Hysteresis and eddy-current compensation was formulated as a data-driven residual sequence-modeling problem, in which a model predicts the absolute field of a cycle from a seeded window of past excitation and field, and a correction follows as a feed-forward trim against a per-cycle reference. No comparable data-driven formulation of hysteresis and eddy-current compensation for accelerator magnets had been reported before this work.

2. A measurement methodology was established for the main dipole across a laboratory bench campaign and the operational B-train. The bench campaign was disqualified for this problem by the power-converter RMS current limit that excludes the operational cycles

and by an integration drift larger than the hysteresis signal, and the B-train was adopted as the usable source after a marker-gap drift correction and acceptance criteria that recover a reproducibility sufficient to resolve the dominant sequence-change effect. This work identifies the measurement as the limit on attainable accuracy and reproducibility.

3. A residual model of the main dipole field was built on the Temporal Fusion Transformer and trained by pretraining on simulated hysteresis and fine-tuning on measured data, reaching prediction error close to the measurement reproducibility floor. The architecture was chosen through preliminary studies that excluded a differentiable Preisach formulation on accuracy and other state-of-the-art sequence models, and advanced a neural-parametrized Preisach that demonstrated its intended capabilities and is left to future work.

4. A feed-forward compensation scheme was designed and realized within the existing control system through the settings hierarchy, transparent to parallel operator changes. Seeded from the B-train each cycle, it was validated with beam in dedicated machine time across the operational proton supercycles through compensation of the main dipole field for the proton fixed-target slow extraction, including an asymmetric sequence that no static correction can reproduce, and it ran autonomously for 35 h through normal operation with sequence changes and without operator intervention. The open-loop autoregressive mode, in which the model runs on its own prediction with no measured seed, was demonstrated separately on the dipole.

5. The eddy-current decay experienced by the beam at injection was characterized and modeled from beam measurements, and shown to be distinct from the decay recorded by any magnet measurement, since the beam integrates the ring eddy currents in the vacuum chamber and grounded structures that the B-train and the bench do not contain. A correction to the main dipole field, derived from the model, reduced the injection orbit distortion in dedicated beam tests. The fit remains under-constrained, and the correction is not yet deployable on arbitrary cycles.

6. The precycle was understood as the mechanism that sets a reproducible hysteresis branch and a reproducible eddy-current decay at injection of the SFTPRO cycle on the main dipole and quadrupole field, and also sets the sextupole component of the dipole field. These findings informed the flattening of the precycle later adopted in operation to attain the energy savings and duty cycle increase.

Two limits bound how far the work reaches, the reproducibility of the magnetic measurement identified above and the variety of the data available for training. The operational measurements covers only the sequences operation plays, which leave generalization to arbitrary sequences unsupported. Beyond those sequences, the compensation at injection and the extension to the higher-order circuits remain bounded by the measurement and the data, and advance as these improve and not as the model grows.

The proposed and delivered scheme is general in design, predicting the field at any excitation and any point within a cycle and formulated for every main circuit, while only the main dipole at the fixed-target flat top was realized and validated. Of the operational outcomes, the energy savings and increased duty cycle during machine development from the now understood precycle is in routine use. The compensation itself has not yet been adopted for production, partially because the flattened precycle sets an operating point the model was not trained or validated on. Because some operating conditions still require the original precycle, the flattening is a conditional simplification and not a full removal and it delivers those gains without

compensating the field, and the reproducibility it preserves still rests on the precycle. Full flexibility further requires compensation of the quadrupole and higher-order circuits, which the approach immediately transfers to once those circuits can be measured in absolute, and reaching the required reproducibility. Finally, lacking any online field measurement, their compensation must run in the autoregressive mode, seeded once from a known state.

12.2 Future Work

The modeling and measurement directions this work opens are set out in the Outlook of Chapter 11. On the modeling side, the beam-seen eddy-current decay was shown to be a distinct effect that the field measurements, and by extension the fitted field model cannot reproduce and that must be corrected separately, which then allows the learned model to reduce to hysteresis-only, with simpler architecture and shorter prediction horizon. Another direction treats the measurement drift that bounds the accuracy as a latent state in a state-space estimator, as a route to a drift-free field prediction that this work did not explore.

Beyond the dipole, the families that full operational flexibility also requires could not be measured usably within this work. A quadrupole measurement campaign was attempted without reaching a usable result, blocked by the absence of an absolute reference for a gradient field and by electromagnetic pickup from the converter. Extension to the quadrupole and higher-order families therefore waits on an absolute gradient reference and improved fidelity. In addition, quadrupole measurements in the laboratory could achieve the arbitrary sequences we showed to not be possible on the dipoles, which through a transferred model could supply the excitation variety to the dipole model that the operational dipole data lacks.

Planned real-time feedback on the main dipole field, enabled by the new link between the power converter and the B-train, could regulate field directly and reduce the need for the demonstrated feed-forward prediction on the dipole, servoing only the field measured in the B-Train bench. That measurement contains neither the beam-seen eddy decay of Chapter 9 nor any signal from the higher-order families, so the predictive model and the beam-based correction developed here stay necessary even under such feedback.

12.3 Concluding Remarks

This thesis has shown that the history-dependent field of the SPS main dipoles can be measured, modelled and compensated ahead of each cycle. With this compensation, the fixed-target extraction is reproduced under operational beam, without manual intervention.

The same history dependence is present in any fast multi-cycling synchrotron, and in magnetic circuits other than the main dipoles, so the approach should carry over to those cases as well. It adds no hardware, and only needs a control system that announces each upcoming cycle and can trim its settings in advance, plus a magnet measurement as accurate as beam control already requires.

The predictive model that makes this possible is at this stage limited only by the the quality of the field measurement it learns from, and it can replace the dedicated instrumented reference magnet required for feedback, that does not scale to the many circuits of a machine. Beyond the SPS, the pretraining and fine-tuning demonstrated here would allow extension of the model to other normal-conducting magnet families and to other fast-cycling machines when measurements are available.