The Bilinear-Exponential Closed-Orbit Model and its Application to Storage Ring Beam Diagnostics

Bernard Riemann

2016-04

eingereicht als Dissertation
zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften
(Dr. rer. nat.)

Fakultät Physik
Technische Universität Dortmund

Erster Gutachter	Prof. Dr. Thomas Weis
Zweiter Gutachter	Prof. Dr. Andreas Jankowiak
Wissenschaftlicher Mitarbeiter	Dr. Ulf Berges
Vorsitzender der Prüfungskomission	Prof. Dr. Jan Kierfeld
Tag der Prüfung	2016-07-22

Zentrum für Synchrotronstrahlung
Technische Universität Dortmund

Contents

1. Introduction

The field of beam dynamics in closed-orbit accelerators like synchrotrons and synchrotron storage rings is of ongoing interest, as the analysis of particle oscillations (*beam optics*) in such a system enables checking and correction of the components of an accelerator.

In such periodic systems, the ensemble of particles is focused onto one[1] specific trajectory through the ring that is repeated (mapped to itself) on each passage and thereby defined as *closed orbit* [7]. One of the simplest methods to observe this special orbit is given by beam position monitors. In their most basic implementations, which are installed at presumably every storage ring in existence, these devices allow to observe the closed orbit in the low-frequency range up to a few Hz.

While experimental data on closed orbits exists in abundance for many accelerators, its use for beam dynamics studies is sparse in comparison. This is due to the fact that if the beam optics at monitor and actuator (*corrector*) positions are known, closed-orbit perturbations can be computed by elementary means – but if only closed-orbit perturbations are known, beam optics at monitor and corrector positions can only be computed with additional data sources or adding assumptions, by non-trivial means. This problem, where model parameters (beam optics) are to be computed by measurement data (closed orbits), belongs to the class of *inverse problems* [3] and will be described as *response problem* in this work.

The LOCO approach [41] (sec. 6.2.1) has been used successfully to approach the response problem by utilizing an existing detailed simulation of all magnetic fields in the accelerator under study. The accelerator model parameters are modified by fitting simulated closed-orbit perturbations to measured ones. The simulation has to be tailored to an existing storage ring, containing user-defined decisions about what parameters are different between the accelerator model and reality, which corresponds to an arbitrary set of parameters. The model is non-analytical as simulated response matrices are generated by numerical tracking for every optimization step, which results in a non-negligible time for the optimization to run [4], especially for larger storage rings. If the deviations between simulation and reality are too large (e.g. for the DELTA storage ring, chapter 7), LOCO does not converge.

Besides the LOCO approach, methods for measuring beam optics in closed-orbit accelerators extensively use analysis of oscillation data. The methods and experiments can roughly be classified into two groups. Turn-by-Turn analysis [4], working with short transient or coherent excitations, is possible using special monitors at medium- or small-scale storage rings. AC dipole techniques, on the other hand, do use standard monitor hardware, while excitation is performed using special magnets called AC dipoles [5]. This technique, which yielded impressive results, has only been possible at large storage rings like LHC and RHIC [43, 85] or the Fermilab Tevatron [44], where the beam eigenfrequencies (which essentially scale inversely to the ring circumference) are very low in comparison to typical storage rings ($\leq 25\,\mathrm{kHz}$, sec. 6.2.2).

For both TbT analysis and AC dipole techniques, additional dedicated hardware needs to be available at the closed-orbit accelerator under consideration. In addition, the available installation space in existing storage rings is often limited.

[1] In the nonlinear case, more than one closed orbit is possible to exist. In practice, this only occurs deliberately, e.g. in the transverse plane using resonant extraction [1], or in longitudinal direction (α-buckets) [2].

1.1. Scope of this work

The procedure for studying an inverse problem can be divided into the steps [3]

1) parameterization with a minimal set of model parameters,
2) forward modeling and *3) inverse modeling.*

Considering LOCO, it is possible to find a more suitable model and optimization in the context of inverse problems, with an emphasis on step 1). The goal is to reduce the model to its dependencies in the measured closed orbits alone ("Occam's razor" [42]). This leads to a modification of the subsequent steps.

Resulting from the aforementioned situation, the scope of this work can be subsumed as the completion of the following two tasks.

1. *Build an elementary and general model for (small) closed orbit perturbations.*

   When approaching this task, it becomes clear that most of the existing beam optics parameterizations are unsuitable, either due to superfluous assumptions (decoupled approximation) or large 'calculation overhead'. Therefore, one task of this work is the *reduction* of single particle and closed orbit motion into a simplified framework without unnecessary assumptions. Here, removing assumptions does not necessarily imply that the resulting model has small number of free variables – as the model is more general, it is required to be more complex than e.g. decoupled models.

   On the way of extracting the relations between closed orbits and single particle oscillation (forward modeling), one obtains a natural parameterization of the problem based on eigenorbits. Reduction of the parameters to observable quantities by elementary means leads to the expressions for the Bilinear-Exponential model (sec. 3.2).

2. *Find an algorithm to solve the inverse problem.*

   During development of a solution to the problem, parts of existing diagnostics routines [17, 23, 29, 30] will be incorporated and reused. From a conceptual standpoint, only a few key ideas will need to be added to sketch the structure of the solution. Then a step-by-step solution of the inverse problem in form of the COBEA algorithm is shown to be possible (chapter 6).

   To validate the algorithm, its results are compared with results from other successful diagnostic methods, based on experiments done in three storage rings (chapters 7–9). All of the compared methods either need more input data respectively assumptions or additional measurement devices in comparison to the method presented in this work.

In this work, the model building corresponds to the forward modeling, in which a type of "natural parameterization" is derived without further efforts. The inverse modeling part corresponds to the construction of the algorithm.

If the algorithm works as expected, we can consider the response problem as solved.

1.2. Further remarks

In this work, a colloquial style is used instead of a more traditional passive form. The 'we' statements include author and reader and do not imply a third contributing party.

While the author has taken care to introduce variables and operators in a meaningful and sufficient order, the interested reader is referred to an overview of used symbols and brief descriptions of variables, which can be found in sec. 1.3.

1.3. Definitions and syntax

After a short description of mathematical operations, this section describes the terms and definitions used in this work, followed by separated index definitions.

If not stated otherwise in the context,

1. real-valued scalars and vectors are written using lowercase letters, e.g. $x, \vec{g}$,
2. complex-valued scalars and vectors are written using uppercase letters, e.g. $X, \vec{G}$.
3. An index is written lowercase, e.g. m, while its range is written uppercase, e.g. M. This does not hold for the direction index w, which range is also M by definition (see sec. 1.3.3).
4. Matrices $\mathbf{M}$ are written in bold shape. Any multi-dimensional array with dimension ≥ 2 is also written in bold shape. No rules apply for uppercase or lowercase.
5. Operators are written in calligraphic style, e.g. $\mathcal{K}$.

1.3.1. Mathematical operations

conjugation $\bullet^*$
changes the sign of the imaginary part of a complex number, so that for real a, b, $(a + \mathrm{i}b)^* = a - \mathrm{i}b$. Works element-wise on vectors and matrices.

real part $\mathfrak{R}\{\}$
Returns the real part of a complex number. Works element-wise on vectors and matrices. $2\mathfrak{R}\{x\} = x + x^*$.

imaginary part $\mathfrak{I}\{\}$
Returns the imaginary part of a complex number. Works element-wise on vectors and matrices. $2\mathrm{i}\mathfrak{I}\{x\} = x - x^*$.

matrix adjunct $\bullet^\dagger$
Transpose a matrix and replace all entries by their complex conjugates. When used on real matrices, reduces to transpose. Turns column vectors into (conjugate) row vectors. $(\mathbf{A}^\dagger)_{mn} = A^*_{nm}$.

1.3.2. Terms and definitions

BE-d model An elementary model developed in this work that can be used to represent any valid general response matrix without dispersion. (chapter 3)

BE+d model An extension of the BE-d model that allows to treat dispersion and can be used to represent any valid general response matrix of a storage ring. (chapter 3)

BESSY II A 1.7 GeV storage ring built to produce synchrotron radiation for external users, located in Berlin and run by HZB. (chapter 9)

(Betatron) tunes A colloquial term for modal phase advances (see eigenorbit). In a narrow sense, the fractional betatron tune is the modal phase advance per unit circle $\mu_m/(2\pi)$. In a broader sense, the betatron tune also includes the integer windings of the complex eigenorbit around the ring. To give an example, an integer tune of 9 and a fractional tune $\mu_m/(2\pi) = 0.16$ results in a betatron tune of 9.16.

CM mapping Corrector-Monitor mapping, an algorithm developed in [23] and this work to compute all monitor vectors and corrector parameters from a given subset of monitor vector and phase advances. (sec. 5.1)

COBEA An algorithm developed in this work that solves the response problem with the only additional information of accelerator topology (represented by the topology matrix $\mathbf{S}$). (chapter 6)

Corrector parameters A_{km}
Complex scalars that hold perturbation information for a specific corrector k and mode m. They can be related to optical parameters at the corrector location $\tilde{s}_k$ with additional assumptions. (sec. 3.1)

DELTA The Dortmund electron storage ring facility is a 1.5 GeV synchrotron light source located in Dortmund, NRW, Germany. (chapters 7 and 8)

DTFT Discrete-Time Fourier Transform, the continuous band-limited spectrum of a sequence [73].

Eigenorbit $\vec{R}_m(s), X(s)$
A complex-valued orbit which is the solution of the eigenvalue problem for the linear one-turn operator of a segment as defined in this work. The absolute value of all eigenorbit components at start and end positions of a segment is identical, while the complex phases at both positions differ by the modal phase advance μ_m. (sec. 2.3)

FFT Fast Fourier Transform algorithm [73], used to compute the DTFT (respectively DFT) at equidistant frequencies.

General response matrix $\mathbf{r}$ with components $\vec{r}_{jk}$ and coefficients r_{jkw}
A matrix in which each column k holds the orbit perturbations originating from the corrector indexed with k. A general response matrix is not normalized, and the specific kick angles θ_k may be unknown. Note that for an accurate response matrix to be recorded, the additional perturbations to the closed orbit must be small (linearization around an arbitrary closed orbit).

In the cases discussed in this work, a general response matrix can be either a current-response matrix, where the excitation current is known, or an angle-response matrix, for which the kick angle is known and normalized out.

HZB Helmholtz-Zentrum Berlin, a part of the Helmholtz Association (government-funded).

MCS Monitor-Corrector Subspace algorithm (developed in this work), with which an approximate set of BE+d parameters can be obtained from a general response matrix. Used as start-value layer of COBEA. (secs. 5.2, 6.1)

MLS Metrology Light Source, a small storage ring facility located in Berlin, used by Germany's national metrology institute PTB and operated by HZB. (chapter 9)

Modal phase advance μ_m, μ_x
See eigenorbit and betatron tune. (sec. 2.3)

Monitor vectors $\vec{R}_{jm}$ with components R_{jmw}
Complex-valued vectors that represent eigenorbits for mode m at monitor j in direction w.

PCA Principal Component Analysis [31, 32].

Perturbed segment A segment which is voluntarily perturbed during a diagnostic measurement by the accelerator operator. In this definition, any storage ring setup which is not subjected to a measurement process is considered as unperturbed, including rings with undesired magnetic field deviations. (sec. 2.2)

Principal orbits $\langle \mathbf{r} \rangle$ with components $\langle \vec{r}_{jp} \rangle$ and coefficients $\langle r_{jpw} \rangle$
Orbits obtained from a PCA [31, 32] using SVD [28] (sec. A.3.1) on a subset of monitors and correctors ("block") from a general response matrix.

Rectification Jacobian $\mathbf{D}_j(\vec{r}_j^{\,\#})$
Distorted measurements of closed orbits respectively response matrices can be compensated posterior to the optimization procedure if the Jacobian of the rectification (inverse distortion) around the measured unperturbed closed orbit is known. (sec. 3.2.5)

Response problem Given a valid general response matrix, how can optical parameters at monitor and corrector positions (represented by BE+d model parameters) be found with a minimum of additional information? (chapter 4)

SVD Singular Value Decomposition [28]. (sec. A.3.1)

TbT Turn-by-Turn, a term used for data and data sources in which the sampling rate of the respective device is equal to the circulation frequency of particles in the accelerator. (sec. 8.2)

Topology matrix $\mathbf{S}$ with components S_{jk}
The entries $S_{jk} \in \{-1, +1\}$ of this matrix state whether monitor j is located "upstream" ($s_j < s_k$) or "downstream" ($s_j > s_k$) of corrector k.

1.3.3. Definition of indices

If not stated otherwise, indices start from 1. This is the case for all indices but the turn number. Note that the following definitions only hold if the indices are not defined in a differing, local context. This is especially the case for the seldom used indices p, n.

direction index $w \leq M$

The direction index w represents either the horizontal ($w = 1 \equiv \mathrm{x}$) or vertical ($w = 2 \equiv \mathrm{y}$) direction. Note that for the direction index, numbers and direction characters x,y are exactly equivalent. The number of considered directions is given as M and is either 1 or 2.

monitor index $j \leq J$

Each monitor index j represents a monitor in the storage ring. In this work, we assume "twin-view" monitors which can measure the beam position in horizontal and vertical direction simultaneously.

corrector index $k \leq K$

Each monitor index k represents a corrector in the storage ring. If not stated otherwise, it is generally not assumed that the angle or direction of the corrector kick $\vec{\theta}_k$ is known.

mode index $m \leq M$

The mode index distinguishes oscillation eigenmodes. In the M-dimensional (quasi-)harmonic oscillator considered as approximation for betatron motion, the number M of modes and dimensions must be identical.

principal component / partial orbit index $p \leq P$

The orbit in an unperturbed segment in a storage ring, generated by perturbations outside of that segment, can be described completely by 5 principal or partial orbits, which form a complete basis for all "allowed" orbits in a segment. This basis is the relevant quantity, and therefore the directions of specific vectors in that basis are arbitrary. Therefore, the index also denotes the pth principal orbit of a PCA of measured orbits [31, 32].

turn number (sec. 8.2) $0 \leq n \leq N - 1$

In the TbT analysis chapter, n refers to the turn number. The typical case for DELTA is $N = 2048$.

2. Linear particle motion

The general motion of particles in storage rings is a topic that has been discussed extensively in many works, e.g. [6, 7, 8, 9]. The common notation for linearized motion around a reference trajectory (decoupled linear optics) in accelerators is Courant-Snyder parameterization [10]. Although this notation is reasonable for many cases, using it would vastly complicate the expressions in chapter 5.1 and the decomposition of perturbations in chapter 6.[1] On the other hand, using Hamiltonian mechanics [7, 8] would add superfluous complexity to the derivations.

Therefore we introduce a "natural" parameterization for the problem based on the concept of segment eigenorbits, which is closely related to Floquet normal forms [12], but adapts the typical coordinate system of storage rings. It may also be interpreted as a complex analogue to [13] (sec. A.1.2) or a complex-Cartesian analogue to the (coupled) Mais-Ripken parameterization [14, 15] (sec. 2.4.4). Due to its simplicity and elementary interpretation (sec. 2.3.3 and Fig. 2.3), it is beneficial for the derivations performed in this thesis.

2.1. Coordinate system

In any particle accelerator, the particles are guided along a reference trajectory. This special trajectory defines a Frenet-Serret coordinate system [7] with a momentous path length s and transverse coordinates x, y, so that the motion of particles is considered in terms of deviation from the reference orbit. The classical particle state is completely described by the corresponding spatial deviations $x, y, \delta s$ and momentum deviations $p_x, p_y, \delta p_s$.

In the common notation used for particle trajectories in accelerators, the transverse deviations of momentum are replaced by proportional derivatives x', y' with respect to s [7]. Throughout this work, $\bullet'$ is used as an abbreviation for the derivative $\mathrm{d}\bullet / \mathrm{d}s$. In all cases for which experiments were performed in this thesis, we can use the ultra-relativistic and small-angle approximations (no path-length effects), so that $p_\mathrm{x} = x'|\vec{p}|$.

Instead of using the impulse deviation δp_s for the longitudinal coordinate, it is also common to use the relative momentum deviation [7]

$$\delta = \frac{\delta|\vec{p}|}{|\vec{p}|}.$$

The particle state respectively orbit relative to the aforementioned design trajectory can then be represented by a phase space vector

$$\vec{z}(s) = \begin{pmatrix} \vec{r}(s) \\ \vec{r}'(s) \end{pmatrix} = \begin{pmatrix} x & y & \delta s & x' & y' & \delta \end{pmatrix}^\dagger .$$

[1] In [11], which emphasized the strictly decoupled case, the author used chains of angular functions and substitutions extensively to derive expressions for β, ϕ directly; thus turning simple linear equation systems to nonlinear ones.

In consequence, one may define a vector-valued function or *map* between two positions $\tilde{s}$, s, which can be expressed using an $s-$dependent operator $\mathcal{M}(s)$ by

$$\vec{z}(s) = \mathcal{M}(s)\vec{z}(\tilde{s} = 0).$$

For the linear case,[2] maps reduce to translations $\vec{z}_0(s)$ and transfer matrices $\mathbf{M}(s)$ via

$$\mathcal{M}(s)\vec{z}(\tilde{s} = 0) = \vec{z}_0(s) + \mathbf{M}(s)\vec{z}(\tilde{s} = 0).$$

2.2. Segment types

For the following considerations, we will classify partitions of the beam path (*segments*) of an accelerator structure by their role in the measurement process into the three classes

1. known segments,
2. unperturbed segments and
3. perturbed segments.

A known segment is also either an unperturbed segment (true for all discussions in this work) or a perturbed segment.

The following discussion is limited to the start-to-end map of a segment. For a segment of length l, it is simply defined as

$$\mathcal{T} = \mathcal{M}(s = l).$$

2.2.1. Known segments

The designation *known segment* implies that the start-to-end map $\mathcal{T}$ of the segment (implicitly defined by electromagnetic fields) is known, either analytically or by measurements. The simplest example for a known segment is a drift space with known length.

Phase space vectors in known segments

In advance of sec. 3.2.1 we state that the particle orbit in an accelerator is often only known by its transverse position x, y. The start-to-end transfer map $\mathcal{T}$ of any segment can be decomposed into block operators, so that

$$\begin{aligned}\vec{r}_{\mathrm{end}} &= \mathcal{T}_{r\to r}\vec{r}_{\mathrm{start}} + \mathcal{T}_{r'\to r}\vec{r}'_{\mathrm{start}}\\ \vec{r}'_{\mathrm{end}} &= \mathcal{T}_{r\to r'}\vec{r}_{\mathrm{start}} + \mathcal{T}_{r'\to r'}\vec{r}'_{\mathrm{start}}.\end{aligned}$$

From the equation for $\vec{r}_{\mathrm{end}}$, one can obtain the orbit derivative at the entry plane of the known segment

$$\vec{r}'_{\mathrm{start}}(\vec{r}_{\mathrm{start}}, \vec{r}_{\mathrm{end}}) = \mathcal{T}^{-1}_{r'\to r}(\vec{r}_{\mathrm{end}} - \mathcal{T}_{r\to r}\vec{r}_{\mathrm{start}}). \tag{2.1}$$

[2] applications of nonlinear maps are considered in chapter 10.

Drift segments

The simplest possible segment is a drift space with length l_{drift}. For this known segment, by use of the intercept theorem, the block operators resolve to

$$\mathcal{T}_{r'\to r}\vec{r}' = l_{\mathrm{drift}}\vec{r}', \quad \mathcal{T}_{r\to r}\vec{r} = \vec{r},$$
$$\text{so that} \quad \vec{r}'_{\mathrm{start}}(\vec{r}_{\mathrm{start}}, \vec{r}_{\mathrm{end}}) = \frac{\vec{r}_{\mathrm{end}} - \vec{r}_{\mathrm{start}}}{l_{\mathrm{drift}}}. \tag{2.2}$$

For drift spaces, the components of $\vec{r}, \vec{r}'$ are also decoupled from each other, so that x', y' can be obtained at the start position from x, y at the ends without knowledge of δ.

2.2.2. Unperturbed segments

In unperturbed segments, arbitrary but quasi-static forces exist. In this context, quasi-static forces imply that fast, transitory perturbations of forces in the time scale of passage time through the segment do not exclude a segment of being unperturbed.

In consequence, the transfer maps $\mathcal{M}(s)$ in such segments are also assumed to be quasi-static.

Independent orbits in unperturbed segments

For the following, we will consider the setup in Fig. 2.1 with an emphasis on the unperturbed segment with a start-to-end transfer map $\mathcal{T}$. The orbit $\vec{r}(s)$ inside the segment can then be described by a map in the form

$$\vec{r}(s) = \mathcal{M}_{r'\to r}(s)\vec{r}(s=0) + \mathcal{M}_{r'\to r}\vec{r}'(s=0)$$

and thus only depends on the initial conditions in $\vec{z}(s=0)$. This fact will be used in the following to decompose orbits inside a segment into linear combinations of P partial orbits,

$$\vec{r}(s) = \sum_{p=1}^{P} a_p \vec{r}_p(s).$$

For a general, nonlinear map of a given order, P is the number of monomial summands in the components up to that order (see chapter 10). In this work we will only consider the following linear cases of unperturbed segments, in which the maximum value of P is the number of phase space dimensions.

- *Linear unperturbed segments with static fields.*

  If a segment is governed by time-independent fields, it is a time-invariant system, and thus the arrival time expressed by δs cannot have an influence on the particle trajectory. Therefore, $P = 5$ initial conditions remain, corresponding to the number of partial orbits.

- *Linear unperturbed segments with dynamic fields.*

  This case occurs when either radio-frequency cavities are installed, or non-negligible changes along the cross section of the particle chamber occur. Then no further constraints can be stated, and $P = 6$ partial orbits exist.

Obtaining maps of unperturbed segments

If an unperturbed segment is enclosed by two known segments, it is possible to recover its transfer map $\mathcal{T}$ (Fig. 2.1). This fact has been in long-term use for measuring segments of linear accelerators, called 'R matrix from Trajectory Fit' [16, 17] in this context. The idea might also have been put to use sporadically to measure unperturbed segments in rings [18].[3]

As noted in sec. 2.2.1, it is possible to recover the phase space coordinates of a known segment using spatial coordinates from its ends. Thus if the unperturbed segment is enclosed by two known segments, their phase space vectors can in turn be used to recover the start-to-end map $\mathcal{T}$ by the relation

$$\vec{z}_p^{\,\mathrm{end}} = \mathcal{T}\,\vec{z}_p^{\,\mathrm{start}} \quad \text{with } 1 \leq p \leq P$$

with a finite number of perturbation measurements P that matches (error-free measurement) or exceeds (noisy measurement) the number of independent orbits in the unperturbed segment.

To give an example without further constraints (see sec. 2.3.1), $P = 5$ measurements would be needed to obtain the linear transfer map of an unperturbed segment with static fields.

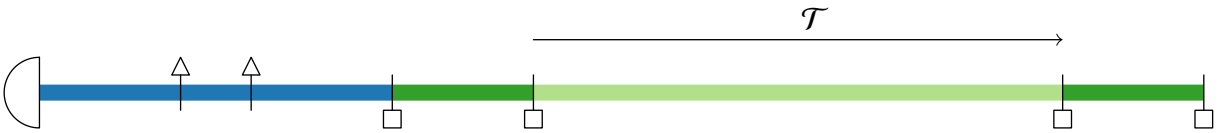

Figure 2.1.: Illustration of the 'R matrix from Trajectory Fit' method, described in [16, 17] for the decoupled case, by which the start-to-end map $\mathcal{T}$ of an unperturbed segment (light green) can be obtained. Monitor positions are denoted by white boxes, the particle source is denoted by a half circle. Perturbations are generated upstream (that is in reverse beam direction) of the segment (perturbed segment, blue), e.g. by dipole correctors (white triangles). The unperturbed segment is enclosed by known segments (dark green).

2.2.3. Perturbed segments

Perturbed segments are the most general segment type. In these segments, arbitrary time-varying forces act on the particle which may change in the course of subsequent measurements.

Again, the nature of the perturbation is assumed to be quasi-static. Thus transient effects on the time-scale of the passage time through to a segment are not considered as perturbations that would classify a segment as perturbed.

[3] This basic idea is expanded in chapter 6 and incorporated into the start-value layer to allow measurement of one-turn maps.

2.3. Linearized motion in rings

Two complementary approaches exist to discuss the motion of particles in accelerators, which we will call device-based approach and Poincaré-based approach. The device-based approach follows the argumentation of

1. assuming a longitudinal piece-wise constant force with a definite transverse force pattern defined by a given magnetic element or multipole decomposition,
2. computing a transfer map between phase space points at entry and exit planes of the magnet element or multipole,
3. building of periodic solutions by different magnet combinations.

While this approach is useful in the design process of accelerators, many assumptions are often posed implicitly, e.g. decoupled, linear motion around a predefined design trajectory in step 1, or specific symmetries of solutions in step 3.

Poincaré theory has been an active topic of research, starting with the divergence of the series expansion for the three-body problem [19]. The Poincaré-based approach in the context of this thesis can be subsumed by the following procedure.

1. Find one or multiple fixpoints of transfer maps (Poincaré sections) for any respectively all longitudinal positions, where the stable fixpoint(s) represent the closed orbit solutions,
2. discuss properties of particle oscillations in Poincaré sections around the stable fixpoint(s),
3. find a suitable parameterization for the problem.

Using this approach, fewer and weaker assumptions are posed on the particle trajectories in comparison to the device-based approach. This is important in the development of diagnostic methods, as the necessity of their use implies that the magnetic fields of the accelerator differ considerably from the desired patterns and values.

For the scope of this thesis, it is thus sufficient to use only the second approach and discuss general properties of such maps. In consequence, magnetic devices are not discussed. Many steps in the following derivation are analogous to the ones found in [15, 20].

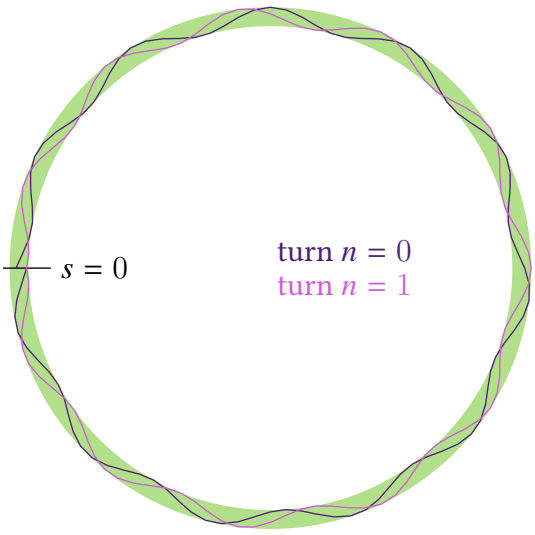

Figure 2.2.: Particle motion in a ring for two consecutive passages or turns. The ring is assumed as unperturbed segment of a fixed length, beginning at $s = 0$. For every turn, the orbit is described by another combination of independent orbits. The change of turns is defined by the longitudinal reference position $s = 0$.

2.3.1. Separation of static and dynamic orbits

Let us assume the storage ring as an unperturbed segment with dynamic fields (sec. 2.2.2) which start and end planes are connected to each other and the start position being referenced as $s = 0$ (ring topology, Fig. 2.2). From sec. 2.1 it is known that the particle state space has an even number of dimensions, which we will assume to be $2M$.

We define a map $\mathcal{T}(s)$ that maps the orbits phase space $\vec{z}(s)$ at any position s to the next turn. If the motion is stable, then at least one orbit $\vec{r}_0(s)$ must exist for which

$$\mathcal{T}(s)\vec{z}_0(s) = \vec{z}_0(s). \tag{2.3}$$

This orbit will be called *closed orbit* in the following and is thus defined by the fixpoints in the Poincaré section plane. Its properties and perturbations will be discussed extensively in chapter 3.

Then, the map $\mathcal{T}(s)$ can be linearized around z_0 as[4]

$$\mathcal{T}(s)\vec{z}(s) = \vec{z}_0(s) + \mathbf{T}(s)(\vec{z}(s) - \vec{z}_0(s)) + O(2).$$

for all positions s. For all following considerations, we will substitute $\vec{z}(s) \leftarrow \vec{z}(s) - \vec{z}(s_0)$ without loss of generality, which means that the unperturbed closed orbit is set as reference orbit. This is a reasonable shift of our coordinate system, as we are only interested in deviations from the unperturbed closed orbit, which are either static, as discussed in sec. 3.1, or transient, as discussed in the following.

2.3.2. Oscillation modes

To study the motion of particles around a closed orbit, we are interested in properties of the one-turn transfer matrix $\mathbf{T}(s)$. We neglect synchrotron radiation effects and state that the phase space density of the resulting system must be preserved due to Liouville's theorem [6, 7]. From this, it follows that the transfer matrix must have full rank and preserves the phase space volume so that $\det \mathbf{T}(s) = 1$.

We can also assume that the particle motion for subsequent turns n is bound and a particle beam can thus be contained in the ring. It can be shown that this is only possible if $\mathbf{T}(s)$ has a valid eigendecomposition (see sec. A.1.1).[5] In this case, one may characterize $\mathbf{T}(s)$ completely by $2M$ eigenvalues λ_m and eigenvectors $\vec{Z}_m(s)$ that solve the eigenproblem

$$\mathbf{T}(s)\vec{Z}_m(s) = \lambda_m \vec{Z}_m(s). \tag{2.4}$$

Note that $\vec{Z}_m(s)$ at a single arbitrary position s characterizes a complete orbit in the unperturbed segment. In consequence, the eigenvalues λ_m are global quantities of that segment.

Then, $\vec{z}$ can be written as a linear combination of all eigenvectors $\vec{Z}_m$. As (2.4) can be multiplied by any complex number, the coefficients of this combination can be omitted or chosen arbitrarily. Here, we use pre-factors $1/2$ to simplify following expressions, so that

$$\begin{aligned} \vec{z}(s) &= \frac{1}{2}\sum_{m}^{2M} \vec{Z}_m(s) \\ \text{and} \quad \mathbf{T}^n(s)z(s) &= \frac{1}{2}\sum_{m=1}^{2M} \vec{Z}_m(s)\lambda_m^n. \end{aligned} \tag{2.5}$$

[4] In other words, $\mathbf{T}$ is the Jacobian of the map $\mathcal{T}$ at $\vec{z}_0$.

[5] An example for a matrix without valid eigendecomposition is the transfer matrix of a drift space. Although assuming a unit determinant, the motion in an accelerator without focusing is unbound.

Further constraints can be applied on eigenvectors and -values. As we assume that the particle motion is bound, the condition $|\lambda_m| \leq 1$ must hold for every eigenvalue. Furthermore, we also assume that the particle motion in an accelerator is governed, in very good approximation, by magnetic fields which do not change the energy of the system. Then, by Liouville's theorem, the phase space volume must be preserved for any transfer, and

$$\det \mathbf{T} = \prod_{m=1}^{2M} \lambda_m \overset{!}{=} 1 \quad \text{leading to} \quad |\lambda_m| = 1 \quad \text{and} \quad \lambda_m = \mathrm{e}^{\mathrm{i}\mu_m},$$

where the *modal phase advances* $\mu_m \in]-\pi, \pi]$ are proportional to the fractional betatron and/or synchrotron tunes.

As $\mathbf{T}^n(s)\vec{z}(s)$ must be a real-valued vector for all turns n, and all eigenvalues λ_m are located on the complex unit circle, it follows from (2.5) that the eigenvectors and eigenvalues must occur in complex-conjugate pairs if $\mathbf{T}$ is nondegenerate. We choose the pairs so that the indices up to M hold one item of each pair, resulting in eigenvalues $\lambda_{m+M} = \lambda_m^*$ and eigenvectors $\vec{Z}_{m+M} = \vec{Z}_m(s)^*$. Then, the summation reduces to

$$\mathbf{T}^n \vec{z}(s) = \frac{1}{2}\left[\sum_{m=1}^{M} \vec{Z}_m(s)\lambda_m^n + \sum_{m=1}^{M} (\vec{Z}_m(s)\lambda_m^n)^*\right] = \frac{1}{2}\left[\sum_{m=1}^{M} \vec{Z}_m(s)\mathrm{e}^{\mathrm{i}n\mu_m} + \sum_{m=1}^{M} \vec{Z}_m(s)^*\mathrm{e}^{-\mathrm{i}n\mu_m}\right].$$

From here on we will always assume that the summation over m extends from 1 to M if not stated otherwise, and obtain a concise expression for subsequent turns in phase space ($\Re$: real part)

$$\vec{z}_n(s) = \mathbf{T}^n \vec{z}(s) = \Re\left\{\sum_m \vec{Z}_m(s)\mathrm{e}^{\mathrm{i}n\mu_m}\right\}. \tag{2.6}$$

In summary, we have formalized the reasonable notion that any motion in a bound, undamped, M-dimensional system with linearized forces can be described as a combination of M oscillators. Contrary to intuition, the usage of harmonic functions $\mathrm{e}^{\mathrm{i}n\mu_m}$ has not implied that these oscillators are harmonic oscillators in the strict sense (see sec. 2.3.4), as they have more degrees of freedom than required. These degrees of freedom are contained in the phase space vectors $\vec{Z}_m(s)$.

Synchrotron radiation damping

It was assumed that synchrotron radiation can be neglected. If this is not the case, one can either consider these damping effects as perturbations, or simply assume that the one-turn transfer matrix is not defective (sec. A.1.1).

Then the above derivation can be generalized to linear damping effects and we can include a damping factor ζ in the respective eigenvalues which are then located either on or inside the complex unit circle, so that

$$\lambda = \mathrm{e}^{\mathrm{i}\mu-\zeta} \text{ and } \lambda^* = \mathrm{e}^{-\mathrm{i}\mu-\zeta} \text{ with } \zeta \geq 0.$$

2.3.3. Eigenorbits

One may split (2.6) into spatial and s-derivative components, so that

$$\vec{r}_n(s) = \Re\left\{\sum_m \vec{R}_m(s)\mathrm{e}^{\mathrm{i}n\mu_m}\right\} \tag{2.7}$$
$$\text{and } \vec{r}'_n(s) = \Re\left\{\sum_m \vec{R}'_m(s)\mathrm{e}^{\mathrm{i}n\mu_m}\right\}.$$

Differentiation of (2.7) with respect to s leads to the second equation, so that (2.7) is equivalent to (2.6).

To clarify the interpretation of the complex trajectory $\vec{R}_m(s)$, we construct a linear one-turn operator[6]

$$\mathcal{N} = \mathcal{T}_{r\to r} + \mathcal{T}_{r'\to r}\frac{\mathrm{d}}{\mathrm{d}s}$$

that maps any orbit for a turn n in the considered segment to the next turn $n+1$,

$$\vec{r}_{n+1}(s) = \mathcal{N}\vec{r}_n(s) \quad \text{with } s \in [0, L[.$$

In consistence with the previous considerations in this section, it can be shown by elementary means that the eigenproblem for $\mathcal{N}$ is then given by

$$\mathrm{e}^{\mathrm{i}\mu_m}\vec{R}_m(s) = \mathcal{N}\vec{R}_m(s),$$

where $\vec{R}_m(s)$ is the *eigenorbit* of the one-turn operator on the segment.[7]

Although being complex-valued quantities, eigenorbits have the same transformation behavior as real orbits and can be mapped using transfer matrices. They may be interpreted as spirals in the complex plane which map to subsequent trajectories for turns n by rotating the spiral about μ_m for each turn and then projecting it into the real trajectory plane (Fig. 2.3).

[6] Expressing the dynamics by one-turn operators should also be possible for the nonlinear case, although it is unlikely to be treated using the eigenvalue problem.

[7] The complete derivation could also be done starting with the introduction of one-turn operators; this approach was discarded to gain more familiarity with existing approaches.

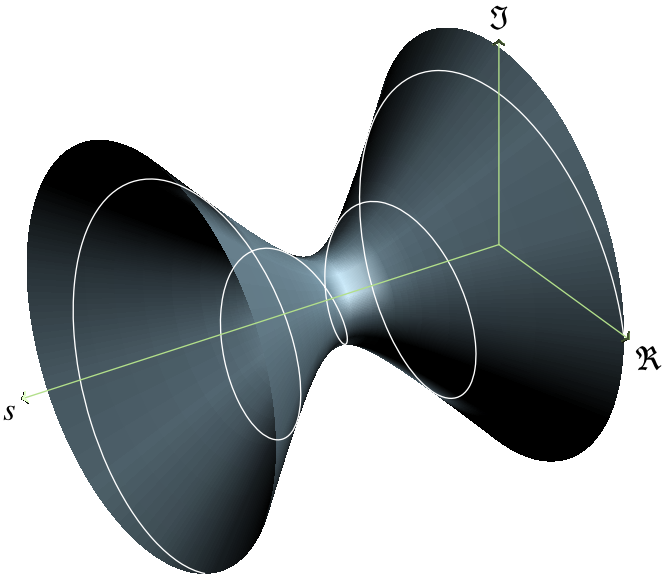

Figure 2.3.: Illustration of an eigenorbit (white line) $\vec{R}_m(s)$ in a given direction w and its beam envelope (surface). The complex angle of the eigenorbit is equivalent to the oscillation phase (respectively betatron phase). To obtain the trajectory for a given turn n according to (2.7), the eigenorbit is rotated by $n\mu_m$ and then projected onto the $\Re - s$ plane.

Relation to Floquet normal forms

Eigenorbits of segments are closely related to Floquet normal forms, which appear as solutions to differential equations of the form [12]

$$\partial_s \vec{r}(s) = \mathbf{A}(s)\vec{r}(s) \text{ with } \mathbf{A}(s) = \mathbf{A}(s+L),$$

and are defined as

$$\vec{r}(s) = \sum_q \vec{g}_q(s) \mathrm{e}^{sB_q} \text{ with } \vec{g}(s) = \vec{g}(s+L).$$

For multiple turns, we would have

$$\vec{r}(s+nL) = \sum_q \vec{g}_q(s) \mathrm{e}^{sB_q} \mathrm{e}^{nLB_q}.$$

For stable motion all B_q must be imaginary. We can then identify $\mu_m = LB_m/\mathrm{i}$ and limit $B_{-m} = -B_m$. Also, we have $\vec{R}_m(s) = \frac{1}{2}\vec{g}_m(s)\mathrm{e}^{sB_q}$ and $\vec{g}_{-m} = \vec{g}_m^*$.

We conclude that the significant difference between eigenorbits and normal forms is that Floquet theory treats all points along the periodic dimensions in an equal fashion, while the segment on which the one-turn operator is defined has a start ($s = 0$) and end plane. This is in accordance with coordinate systems used for storage rings and therefore useful for following calculations.

2.3.4. Decoupled motion

Synchrotrons and storage rings are often designed for "decoupled optics" in which, by definition, each modal oscillation only occurs in a given direction w. When we define the w-th component of an eigenorbit $\vec{R}_m(s)$ as $(R_m(s))_w$, and use an analog definition for $\vec{R}'_m$, the decoupled optics assumption can be formalized as

$$(\vec{R}_m(s))_w = \begin{cases} m = w = 1 & : X(s) \\ m = w = 2 & : Y(s) \\ m \neq w & : 0 \end{cases}, \quad (\vec{R}'_m(s))_w = \begin{cases} m = w = 1 & : X'(s) \\ m = w = 2 & : Y'(s) \\ m \neq w & : 0 \end{cases}. \tag{2.8}$$

By this assumption one can further reduce (2.7), limiting the discussion to a spatial deviation x_n and an angle deviation x'_n in that plane. After substitution of the remaining μ_m by μ_x for clarity, one obtains

$$x_n(s) = \Re\left\{X(s)\ \mathrm{e}^{\mathrm{i}n\mu_\mathrm{x}}\right\}, \tag{2.9}$$

$$x'_n(s) = \Re\left\{X'(s)\mathrm{e}^{\mathrm{i}n\mu_\mathrm{x}}\right\}, \tag{2.10}$$

which means that the phase space motion for any mode m should correspond to a phase space plane spanned up by two axes of the coordinate system.

2.4. Invariants of motion and normalization

We have found expressions for linearized trajectories around a reference orbit by using eigenorbits. In accelerator physics and modeling of beam dynamics, the quantities of interest are the inherent properties of the magnetic lattice, and not properties of the special single-particle motion respectively its initial conditions. Therefore one would like to obtain normalized quantities which some or all of the initial conditions are removed by normalization. This can be accomplished by computing invariants of motion, which are also called Courant-Snyder invariants in this context [10].

2.4.1. Invariants for decoupled motion

We investigate the resulting two-dimensional phase space for all possible phases that can occur and substitute $n\mu_\mathrm{x} \to \phi$. As the exponential term has exactly the same imaginary argument in both x and x', the resulting curve for any given complex amplitudes or *phasors* X_j, X'_j is always a 1:1 Lissajous figure and thus an ellipse.

One condition for a term to be an invariant is that it is constant for all ϕ. Thus one invariant candidate is the area πI_x of the ellipse which is quadratic in phase space coordinates.[8] The ellipse area can be found, either by construction (Fig. 2.4) or by sec. A.1.2, to be a product of phasors

$$I_\mathrm{x} = -\Im\left\{XX'^*\right\} = \frac{\mathrm{i}}{2}\left(X^*X' - XX'^*\right). \tag{2.11}$$

The second condition for an invariant is that it does not change when transfers along the ring are applied. As motion is decoupled, the transfer can be split into M different 2×2 independent transfer matrices, of which each determinant must equal unity. Due to the size of the decoupled matrix, this is exactly equivalent to preservation of area under transformation. This also holds for the full decoupled transfer matrix, where the sub-plane areas are preserved. As I_x also fulfills this second condition, it is indeed an invariant of motion. If we interpret the unit circle as phase space trajectory of a harmonic oscillator with normalized energy 1, we can interpret the invariant I_x as proportional to the energy stored in the oscillation. As such, all invariants must be positive (see sec. 6.1.2).

[8] similar to a circle, which area is proportional to $x^2 + y^2$.

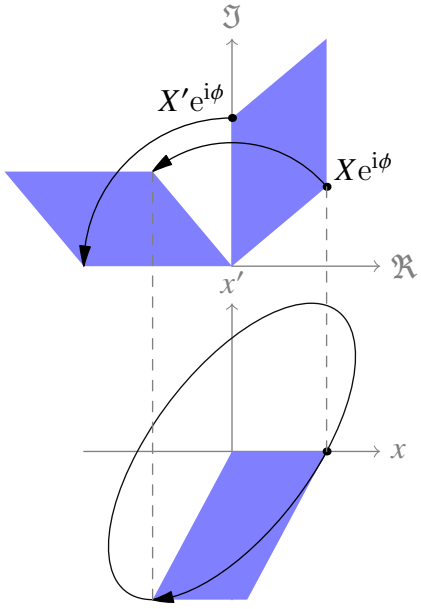

Figure 2.4.: Motion in a decoupled phase space plane and its relation to phasors X, X'. All blue shapes have the same area. An area spanned up by two complex vectors $\vec{a}, \vec{b}$ is given as $|\Im(a^*b)|$; thus the area of the blue shapes is $|\Im(X^*\mathrm{e}^{\mathrm{i}(-\phi+\phi)}X')| = |\Im(X^*X')| = |I_\mathrm{x}|$.

2.4.2. Invariants for general coupled motion

Next, we discuss the general trajectory (including betatron coupling) for a given mode in phase space. For this, we assume that the only oscillation occurs with a phase advance μ_m and the other modes are not excited. Then, we motivate the concepts of general invariants of motion by the decoupled invariants. Nevertheless, the expressions for the invariants will be completely general and not depending on the decoupled assumption.

As discussed in sec. 2.3, $\mathbf{T}$ is invariant on complex scaling of $\vec{Z}$, so the trajectory shape is encoded in the remaining degrees of freedom. All possible $\vec{z}$ for a given amplitude are part of a curve

$$\vec{z} \in \Re\left\{\vec{Z}\mathrm{e}^{\mathrm{i}\phi}\right\} = \Re\vec{Z}\cos\phi - \Im\vec{Z}\sin\phi,$$

with arbitrary ϕ. We conclude that the motion for each mode always happens on a plane spanned up by $\Re\vec{Z}$ and $\Im\vec{Z}$, and the aforementioned curve is again an ellipse on that plane.

In the decoupled case (sec. 2.3.4), we have found one invariant of motion in the (x, x') phase space (2.11), and corresponding expressions exist for the other spatial directions. We are looking for an invariant that can be generalized to the coupled case, but returns the decoupled invariant if the corresponding eigenvectors are used. This is accomplished by a bilinear form

$$I_m = \vec{Z}_m^\dagger \mathbf{\Omega} \vec{Z}_m = \Im\left\{\vec{R}_{jm}^\dagger \vec{R}'_{jm}\right\},$$

where we used a matrix[9]

$$\mathbf{\Omega} = \frac{\mathrm{i}}{2}\begin{pmatrix} & +\mathbf{1} \\ -\mathbf{1} & \end{pmatrix}. \tag{2.12}$$

By definition, the invariants I_m must not change when moving along a segment of the ring, which mapping is described by a segment transfer $\mathbf{M}$. As eigenvectors can be traced like phase space points, the eigenvectors at the new positions are $\mathbf{M}\vec{Z}$, and $(\mathbf{M}\vec{Z})^\dagger$ evaluates to $\vec{Z}^\dagger\mathbf{M}^\dagger$, we have

$$\begin{aligned} \vec{Z}_m^\dagger \mathbf{M}^\dagger \mathbf{\Omega} \mathbf{M} \vec{Z}_m &\stackrel{!}{=} \vec{Z}_m^\dagger \mathbf{\Omega} \vec{Z}_m \qquad \text{for all modes } m, \\ \text{so that} \qquad \mathbf{M}^\dagger \mathbf{\Omega} \mathbf{M} &= \mathbf{\Omega} \qquad \text{must hold.} \end{aligned} \tag{2.13}$$

Any matrix $\mathbf{M}$ that fulfills this condition is called a *symplectic* matrix. It can be shown by elementary means that the product of two symplectic matrices is also a symplectic matrix; this property is used in the following.

For the decoupled case, the expressions reduce to the ones for $I_\mathrm{x}, I_\mathrm{y}$. For this expressions, we have already stated that they are invariants of motion. Thus, any 2×2 matrix that describes motion in a plane[10] is area-preserving and thus symplectic *independently* of the other planes, so that e.g. focusing terms in different planes are not in any way connected when the full decoupled transfer matrix is considered.

To test the coupled case, we just check if any rotated segment

$$\mathbf{M}_\mathrm{rot} = \mathbf{R}^\dagger \mathbf{M} \mathbf{R} \quad \text{with a spatial rotation} \quad \mathbf{R} = \begin{pmatrix} \mathbf{R}_\mathrm{sp} & \\ & \mathbf{R}_\mathrm{sp} \end{pmatrix}$$

is symplectic. $\mathbf{R}$ and $\mathbf{R}^\dagger$ are symplectic matrices, and so the rotated segment transfer $\mathbf{M}_\mathrm{rot}$ must also be symplectic.

[9] This is not the standard symplectic form due to its scaling factor; nevertheless it is valid.

[10] By the assumptions made, this holds for any 2×2 matrix with unit determinant. A reason that symplecticity is hard to illustrate is that an even number of dimensions is required, but it is equivalent to volume preservation for two dimensions.

Therefore, we have shown that any rotated (including "tilted") decoupled transfer element with independent focusing terms in different directions has a symplectic transfer matrix, which includes all rotated quadrupolar-like fields (upright and skew components) and bends (weak focusing) of the coordinate system in arbitrary directions. Then, by the product rule, any combination of decoupled and coupled transfer matrices is also symplectic, and this elementary proof is already sufficient for studies in most storage rings and for the experiments discussed in this thesis, excluding solenoids.[11]

By using Hamiltonian mechanics for electromagnetic forces in a curvilinear coordinate system [7, 20], it has been shown that *any* segment that describes interaction with magnetic fields is symplectic when neglecting synchrotron radiation.[12]

2.4.3. Normalized eigenorbits

After finding the invariants, one may now define normalized[13] eigenorbits via

$$\hat{R}_m(s) = \frac{1}{\sqrt{I_m}} \vec{R}_m(s). \tag{2.14}$$

which only depend on the magnetic lattice and an arbitrary complex phase (betatron phase) offset that has no physical relevance. As shown in sec. A.1.3, the phasor vectors can be traced through the ring like phase space vectors. Therefore, we have found a linear beam optics quantity that is easy to compute and track through the ring.

2.4.4. Relation to common parameterizations

The normalized eigenorbits that have been found need to be related to standard notations of beam optics. We have circumvented such parameterizations until now because all common notations have an essentially polar form. Thus, they often must be subjected to elaborate schemes (e.g. tensor transform for the amplitudes / Betamatrix [6] approach) in derivations. Eigenorbits $\vec{R}$ can be transformed by multiplying their phase space vectors $\vec{Z}$ with the respective transfer matrices (like real orbits), and thus ease derivations enormously even in the decoupled case.

Courant-Snyder parameterization [10] is the first and, so it seems, the only one that is used in the accelerator community for the decoupled case. For general linear (coupled) motion, common approaches either use Edwards-Teng [21, 22] or Mais-Ripken[14] [14, 15, 20] parameters.

The approach of Edwards and Teng transforms an arbitrary phase space ellipsoid into a decoupled frame, which is then described by Courant-Snyder parameters. As this transformation involves locations and momenta of particles, it can not be related in a straightforward way to spatial vectors respectively monitor measurements and is thus not discussed in the following. By contrast, Mais-Ripken parameters can be directly related to spatial motion due to their parameters being constructed independently from eigenorbits for locations and momenta.

[11] A proof in this framework should also be possible but is not investigated here.
[12] As any such transfer map $\mathcal{M}$ is symplectic [7], this also holds for its Jacobian $\mathbf{M}$.
[13] Note that the ˆ symbol is used for general normalization and does not require Euclidean norm to be unity.
[14] These parameters are also sometimes called Ripken-Mais or "Ripken's style" parameters.

Courant-Snyder parameters

For decoupled optics, one can connect the phasors to Courant-Snyder parameters [10] by comparing expressions for successive turns on the phase space ellipse (Fig. 2.5), which leads to the equation

$$\hat{X} = \frac{X}{\sqrt{I_\mathrm{x}}} = \sqrt{\beta}\mathrm{e}^{\mathrm{i}\phi} \quad \text{so that} \quad \beta_\mathrm{x}(s_j) = \hat{R}_{j1} \cdot \hat{R}^*_{j1} = \hat{X}_j \hat{X}^*_j .$$

One can also obtain the betatron phase by the elementary relation

$$\phi_\mathrm{x}(s_j) = \arg X_j . \tag{2.15}$$

Furthermore, the comparison yields

$$\gamma_\mathrm{x}(s_j) = \hat{X}'_j(\hat{X}'_j)^* \quad \text{and} \quad \alpha_\mathrm{x}(s_j) = -\Re\left\{\hat{X}^*_j X'_j\right\} .$$

Mais-Ripken parameters

By their derivation using normalized eigenvectors $\hat{Z}$, the normalized eigenorbits are closely related to the Mais-Ripken parameterization [14, 15, 20], which is the real, polar-like analogue, by

$$\left(\hat{R}_m(s)\right)_\mathrm{x} = \sqrt{\beta_{m\mathrm{x}}(s)}\mathrm{e}^{\mathrm{i}\phi_{m\mathrm{x}}(s)}$$

where the relations also hold when replacing all x with y.

Courant-Snyder parameters can be treated as a special case of Mais-Ripken parameters, which is the approach used for presenting results of our algorithm (sec. 6.1.2).

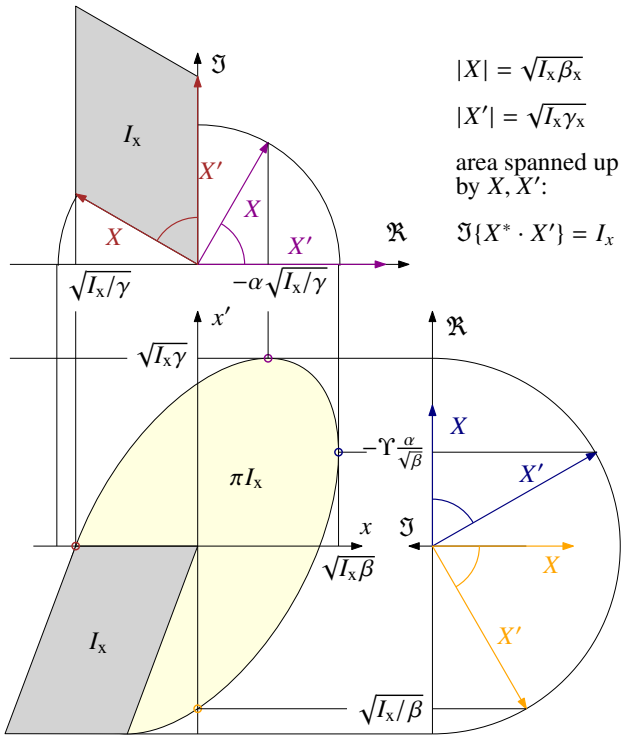

Figure 2.5.: Relations of decoupled phasors X, X' in different phases, the phase space ellipse, and Courant-Snyder parameters.

3. Closed-orbit perturbations using eigenorbits

Building up on the eigenorbit concept, we derive new, generic expressions for orbit perturbations in storage rings, which hold for more general types of perturbations than commonly considered (sec. 3.1). The differences in underlying assumptions can be seen by their reintroduction for comparison purposes in secs. 3.1.3 and 3.1.4.

This approach leads to the definition of the Bilinear-Exponential model (sec. 3.2), which is the basis for the development of all new algorithms and results in this work. The model has also been used in a slightly different form in [23].

For the following derivation, the exact nature of the perturbation inside the perturbed segment is not of interest. The only assumption made is that the perturbations are located in a finite segment of the storage ring.

3.1. Continuous closed-orbit perturbations

We examine the closed orbit and its quasi-static perturbation, which was separated from the oscillation modes and eigenorbits in sec. 2.3.1. We are interested how the fixpoint solution, given by (2.3) as

$$\mathcal{T}(s)\begin{pmatrix}\vec{r}(s)\\ \vec{r}'(s)\end{pmatrix} = \begin{pmatrix}\vec{r}(s)\\ \vec{r}'(s)\end{pmatrix},$$

changes when additional magnetic fields are applied. Still neglecting the unperturbed closed orbit contribution without loss of generality (sec. 2.3.1), $\vec{r}(s)$ represents the closed-orbit perturbation.

3.1.1. Arbitrary orbit perturbations in a storage ring

Again the storage ring is assumed as an unperturbed segment with dynamic fields (see sec. 2.2.2), so that six independent orbits could exist for such a segment.

The closed orbit is linked to the condition that the kinetic energy loss by a particle on this trajectory must equal the energy gain obtained by traveling through the ring and its (radio-frequency) fields; otherwise, the orbit cannot be ring-periodic. This condition links the coordinates δs and δ and thus absorbs one degree of freedom from the orbit, so that $P = 5$ independent orbits remain.

Then one can separate a dispersion orbit $\vec{d}(s)$ that occurs depending on the energy deviation δ from the other orbits by defining[1]

$$\vec{r}_{\text{full}}(s) = \vec{r}(s) + \vec{d}(s)\delta.$$

Until further notice, only the subspace $\vec{r}(s)$ will be considered, which can be decomposed into $P - 1 = 4$ partial orbits.

[1] One could also define this orbit to depend on δs instead of δ due to the above condition.

Next we assume that a local perturbation has occured, which can thus be assigned to a perturbed segment, while the rest of the ring is treated as an unperturbed segment with four partial orbits (see Fig. 3.1). The four partial orbits in this subspace can be chosen arbitrarily as long as they are independent and thus form a complete basis for the subspace. Considering sec. 2.3 (2.7) for the linear case in which monitor vectors were defined, we state that the oscillation's eigenorbits $\Re\hat{\vec{R}}_m(s), \Im\hat{\vec{R}}_m(s)$ for $M = 2$ modes are linearly independent and thus can be used for this purpose. Then for the path from s_+ up to before the reference point $s = 0$, the four initial conditions (e.g. amplitude and phase for each oscillation mode) can be expressed as

$$s \geq s_+ : \qquad \vec{r}(s) = \sum_m \left[a_m \Re\vec{R}_m(s) + a_{m+M} \Im\vec{R}_m(s) \right] = \Re\left\{ \sum_m (a_m - \mathrm{i}a_{m+M})\vec{R}_m(s) \right\}.$$

For the part of the unknown segment in non-negative s direction ($s \geq 0$), the trajectory is described by the next turn of the oscillation ($n = 1$). The unperturbed phase advance μ is included into the expression via

$$s \leq s_- : \qquad \vec{r}(s) = \Re\left\{ \sum_m (a_m - \mathrm{i}a_{m+M})\mathrm{e}^{\mathrm{i}\mu_m}\vec{R}_m(s) \right\}.$$

It is emphasized that μ_m denotes the *unperturbed* modal phase advance and thus is not changed by any perturbation. The perturbed modal phase advances may differ considerably from the unperturbed ones without influencing this equation.

To simplify expressions, one redefines the orbit's initial conditions as M complex quantities

$$A_m = (a_m + \mathrm{i}a_{m+M})\mathrm{e}^{-\mathrm{i}\mu_m/2}$$

Furthermore, a coordinate $\tilde{s}$ is introduced by the condition $s_- < \tilde{s} < s_+$. With these definitions, the perturbed orbit inside the unperturbed segment can be summarized as

$$s \notin]s_-, s_+[: \qquad \vec{r}(s) = \Re\left\{ \sum_m A_m^* \mathrm{e}^{\mathrm{i}\,\mathrm{sign}(\tilde{s}-s)\mu_m/2}\vec{R}_m(s) \right\}, \tag{3.1}$$

$$\text{with} \quad \mathrm{e}^{\mathrm{i}\,\mathrm{sign}(\tilde{s}-s)\mu_m/2} = \begin{cases} \mathrm{e}^{-\mathrm{i}\mu_m/2} & \text{if } s \geq s_+ > \tilde{s} \\ \mathrm{e}^{+\mathrm{i}\mu_m/2} & \text{if } s \leq s_- < \tilde{s} \end{cases}.$$

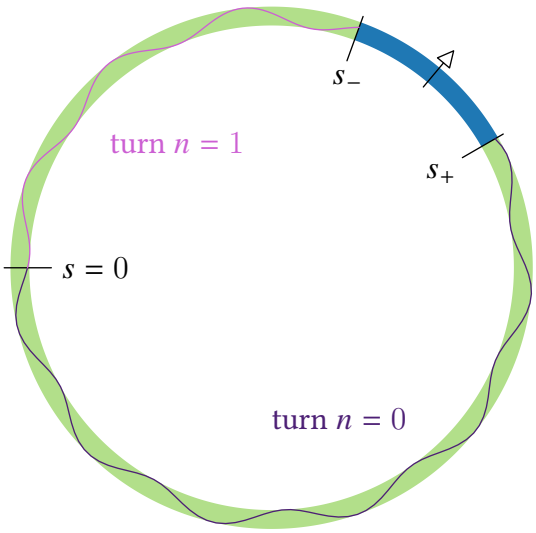

Figure 3.1.: An arbitrary orbit perturbation (oscillating line) in a storage ring. The perturbed segment (blue) between s_- and s_+ is governed by unknown quasi-static fields. In the unperturbed segment (green), the magnetic fields do not change, so its non-dispersive orbit subspace is not influenced by the perturbation and can be expressed by unperturbed eigenorbits. As the unperturbed orbit parts can cross the reference plane $s = 0$, it splits the orbits into different turn parts.

By the fixpoint condition, any perturbation pattern originating in the perturbed segment will cause an orbit perturbation in the unperturbed segment with distinct A_m values. In conclusion, (3.1) is sufficient to study the linear transfer maps $\mathbf{M}(s)$ and segment maps $\mathbf{T}(s_+ \to s_-)$ for the unperturbed segment, without knowing what and where exactly in $]s_-, s_+[$ the perturbation is.

As our condition with four principal orbits applies only to the unperturbed segment, the start-to-end transfer map of the perturbed segment under consideration may also be nonlinear. Note that when excluding this case, the superposition principle holds for A_m with linear combinations of different perturbations in the region $]s_-, s_+[$.

3.1.2. Scaling invariance of eigenorbits and corrector quantities

One may note that a scaling invariance exists, as the simultaneous scaling by any set of M complex numbers C_m

$$\vec{R}_m(s) \leftrightarrow C_m \vec{R}_m(s) \quad \text{and} \quad A_m \leftrightarrow \frac{1}{C_m^*} A_m = \frac{C_m}{|C_m|^2} A_m$$

does not have any effect on the left-hand side of the considered closed-orbit equations. The two degrees of freedom for each mode m correspond to the invariant I_m of the orbit, and its unknown betatron offset for all modes which does not possess physical relevance.

After having defined normalized orbit vectors, we use the scaling invariance and scale with $C = 1/\sqrt{I_m}$ to define normalized corrector quantities via the normalized closed-orbit equation

$$s \notin]s_-, s_+[: \qquad \vec{r}(s) = \Re\left\{ \sum_m \hat{A}_m^* \mathrm{e}^{\mathrm{i}\,\mathrm{sign}(\tilde{s}-s)\mu_m/2} \hat{R}_m(s) \right\}. \tag{3.2}$$

3.1.3. Arbitrary orbit perturbations for decoupled motion

After reviewing the general full-coupled case of orbit perturbation, one can consider expressions for the special case of decoupled motion.

These expressions are only used to cross-check their coherence with common descriptions in literature and to clarify notions for the introduced variables. The approximations shown in the following are not necessary for the methods discussed in this thesis, which are implemented for general coupled motion ($M = 2$).

Inserting (2.8) into (3.2) while evaluating only in the horizontal direction results in

$$\vec{z}_\mathrm{x}(s) = \begin{pmatrix} x(s) \\ x'(s) \end{pmatrix} = \Re\left\{ \begin{pmatrix} X(s) \\ X'(s) \end{pmatrix} \mathrm{e}^{\mathrm{i}\,\mathrm{sign}(\tilde{s}-s)\mu_\mathrm{x}/2} A_\mathrm{x}^* \right\}. \tag{3.3}$$

All calculations and results for A_x and the horizontal decoupled closed orbit in this section also hold for A_y and the vertical closed orbit when substituting all x indices and components with their y counterparts.

3.1.4. Corrector quantities in the decoupled, dipolar, thin-lens approximation

In addition to the decoupled assumption, it is assumed in the following that the perturbation is located in an infinitely small segment located as $\tilde{s}$. It is then possible to formulate a condition on the corrector quantity $\hat{A}_\mathrm{x}$ by using (3.3) with normalized quantities around $\tilde{s}$. The expression can be reduced to[2]

$$\lim_{\epsilon\to 0}\left[\vec{z}_\mathrm{x}(s_+ = \tilde{s}+\epsilon) - \vec{z}_\mathrm{x}(s_- = \tilde{s}-\epsilon)\right] = 2\sin(\mu_\mathrm{x}/2)\Im\left\{\hat{A}_\mathrm{x}^* \begin{pmatrix}\hat{X}(\tilde{s})\\ \hat{X}'(\tilde{s})\end{pmatrix}\right\} \overset{!}{=} \begin{pmatrix}0\\ \theta_\mathrm{x}\end{pmatrix}.$$

From the first equation of the system one can deduce $\hat{X}$ and A_x are parallel, so that

$$\hat{A}_\mathrm{x} = \hat{X}(\tilde{s})\bar{A} \quad \text{with } \bar{A}\in\mathbb{R}.$$

Then by using (2.11), the braced term in the second equation simplifies to

$$\Im\left\{\hat{A}_\mathrm{x}^*\hat{X}'(\tilde{s})\right\} = \bar{A}\Im\left\{\hat{X}^*(\tilde{s})\hat{X}'(\tilde{s})\right\} = \bar{A}$$

$$\text{and results in the corrector quantity} \quad \hat{A}_\mathrm{x} = \hat{X}(\tilde{s})\frac{\theta_\mathrm{x}}{2\sin(\mu_\mathrm{x}/2)}. \tag{3.4}$$

Therefore we have related the corrector quantities to optical functions by stating the additional (decoupled, dipolar, thin-lens) assumptions commonly used in literature.

Rewriting the closed-orbit perturbation equation in the decoupled, dipolar, thin-lens assumption (3.4)

$$\begin{aligned} x(s) &= \Re\left\{\hat{A}_\mathrm{x}^* \mathrm{e}^{\mathrm{i}\,\mathrm{sign}(\tilde{s}-s)\mu_\mathrm{x}/2}\hat{X}(s)\right\} \\ &= \frac{\theta_\mathrm{x}}{2\sin(\mu_\mathrm{x}/2)}\ \Re\left\{\hat{X}^*(\tilde{s})\ \mathrm{e}^{\mathrm{i}\,\mathrm{sign}(\tilde{s}-s)\mu_\mathrm{x}/2}\ \hat{X}(s)\right\} \\ &= \frac{\theta_\mathrm{x}}{2\sin(\mu_\mathrm{x}/2)}\ \sqrt{\beta(s)\beta(\tilde{s})}\cos\left(\phi(s)-\phi(\tilde{s})-\frac{\mu}{2}\,\mathrm{sign}(s-\tilde{s})\right), \end{aligned}$$

one therefore arrives at the standard closed-orbit equation from literature as found e.g. in [7, 16] with equivalent parameters.

[2] This can be regarded as a system of equations containing $\Re$ to be solved for the unknown complex parameter A_x^* (sec. 5.1.1).

3.2. The Bilinear-Exponential (BE) model

The general treatment of continuous closed orbit perturbations now has to be reduced to handle observable quantities only. It is emphasized that the closed-orbit equation (3.1) and thus the resulting Bilinear-Exponential model does not depend on the exact alignment of magnetic structures in the ring, but only on the topology of elements.

3.2.1. Monitors, correctors, and general response matrix

In the following, we will only assume that the orbit can be observed at J given

$$\text{monitor positions } s_j \text{ for } 1 \leq j \leq J$$

in the ring. In abbreviation, we call the eigenorbits at these positions

$$\text{monitor vectors } \vec{R}_{jm} \equiv \vec{R}_m(s_j) \in \mathbb{C}^M.$$

Furthermore, we assume that the origin positions of K closed-orbit perturbations are fixed on a set of

$$\text{corrector positions } \tilde{s}_k \text{ for } 1 \leq k \leq K,$$

leading to the definition of

$$\text{corrector parameters } A_{km} \equiv A_m(\tilde{s}_k, \theta_k) \in \mathbb{C}.$$

Now that a setup is defined where many perturbations can occur, we can subsume the K perturbed orbits as columns of a general response matrix $\mathbf{r}$ of size $J \times K$, so that each vector element $\vec{r}_{jk}$ denotes the orbit perturbation observed at monitor j and originating from corrector k.

The general response matrix can thus be understood as orbit response matrix with unknown kick angles θ_k, and a connection to common decoupled kick-angle response matrices $\mathbf{x}^{\mathrm{rsp}}, \mathbf{y}^{\mathrm{rsp}}$ exists via

$$\begin{aligned} x_{jk} &= x^{\mathrm{rsp}}_{jk}\theta_k \quad \text{with } k \in \mathbb{K}_{\mathrm{x}}, \\ y_{jk} &= y^{\mathrm{rsp}}_{jk}\theta_k \quad \text{with } k \in \mathbb{K}_{\mathrm{y}}. \end{aligned}$$

Note that for decoupled motion, the corrector indices k are grouped into two subsets $\mathbb{K}_{\mathrm{x}}, \mathbb{K}_{\mathrm{y}}$, and the response matrices in each plane are often only measured for the indices in the respective subset (see also sec. 4.3).

3.2.2. Monitor-corrector topology

In conclusion one can treat the exponential term by defining[3]

$$\text{phase jump coefficients } E_{jkm} = \mathrm{e}^{\mathrm{i}\,\mathrm{sign}(s_j - \tilde{s}_k)\mu_m/2}.$$

Note that the only dependence on j, k of the coefficients is given in the sign term. Thus, there is no dependence on the distance between s_j and s_k in the expression, but only one on the Boolean value of the expression $s_j > \tilde{s}_k$ or the *ordering* of $s_j, \tilde{s}_k$.

This ordering is equivalent to the ordering of the elements j, k along the ring. For example, we could define an ordered list like

$$(\text{mon } 1,\ \text{corr } 7,\ \text{mon } 2,\ \text{mon } 3,\ \text{corr } 4, \dots)$$

[3] Note that the positions of $\tilde{s}$ and $\tilde{s}$ have been switched in the definition, so that E^* occurs in (3.6).

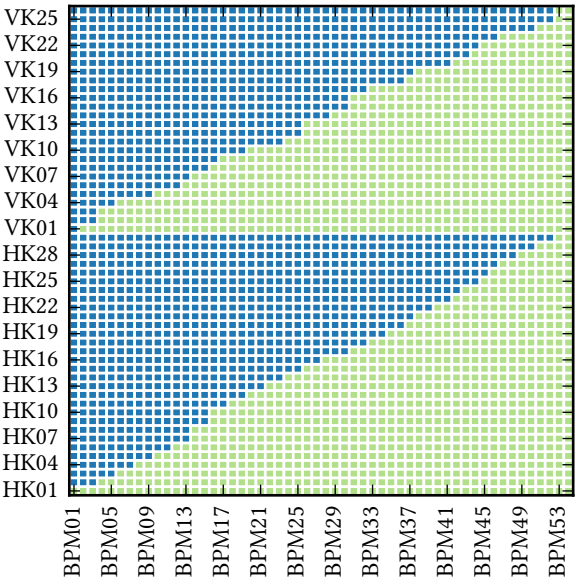

Figure 3.2.: Example topology for DELTA standard correctors (chapter 7). Blue color indicates a negative sign (respectively -1), green color indicates a positive sign (respectively +1).

consisting of $J + K$ identifiers that are ordered towards increasing s coordinates along the ring, which is simply along the beam path. Here, the mon identifiers represent the respective j index of the response matrix, while the corr identifiers represent the respective k index. By comparing the indices n of each monitor and corrector in this list against each other,

$$S_{jk} = \mathrm{sign}(n(\mathrm{mon}\ j) - n(\mathrm{corr}\ k)) \stackrel{!}{=} \mathrm{sign}(s_j - \tilde{s}_k) \in \{-1, +1\}.$$

In practice, the ordered list will contain element names instead of row/column identifiers, and two lists containing the names of all monitors and correctors for each row and column of the response matrix are given. While the matrix S_{jk} is almost always created from an ordered list, we will consider S_{jk} to be given as input to the algorithms and define $\mathbf{S}$ as *topology matrix*. One can then write the phase jump coefficients as

$$E_{jkm} = \mathrm{e}^{\mathrm{i}S_{jk}\mu_m/2}. \tag{3.5}$$

Note that the ordered list and S_{jk} originates from elementary topological properties of the accelerator – the order along the path in which monitors and correctors are installed. No knowledge of magnetic properties or detailed positioning is needed. An example for a topology matrix is shown in Fig. 3.2.

3.2.3. Bilinear-Exponential model without dispersion (BE-d model)

By localizing $\tilde{s}$ between two adjacent monitors, we have ensured that $s_j \notin]s_-, s_+[$ for all j. Therefore one may rewrite (3.1) as

$$\vec{r}_{jk} = \Re\left\{\sum_m \vec{R}_{jm} E^*_{jkm} A^*_{km}\right\}. \tag{3.6}$$

This linear equation (which does not include dispersion) is the starting point for the description of monitor and corrector subroutines in the next chapter.

The scaling invariance discussed in sec. 3.1.2 is also valid for the monitor vectors and corrector parameters. In consequence, an equation with normalized quantities follows as

$$\vec{r}_{jk} = \Re\left\{\sum_m \hat{R}_{jm} E^*_{jkm} \hat{A}^*_{km}\right\}. \tag{3.7}$$

3.2.4. Bilinear-Exponential model with dispersion (BE+d model)

Again integrating the dispersive effects (sec. 3.1.1), one may define

$$\text{dispersion vectors } \vec{d}_j \equiv \vec{d}(s_j) \text{ and dispersion coefficients } b_k.$$

The dispersion coefficients represent energy deviations of the orbit for a given perturbation k, which result from an interplay between path length changes and different energy gains in the cavity or cavities, leading to a new equilibrium. Thus, we arrive at the BE model with dispersion, given by

$$\begin{aligned}\vec{r}_{jk} &= \Re\left\{\sum_m \vec{R}_{jm} E^*_{jkm} A^*_{km}\right\} + b_k \vec{d}_j \\ &= \Re\left\{\sum_m \vec{R}_{jm} \mathrm{e}^{-\mathrm{i}S_{jk}\mu_m/2} A^*_{km}\right\} + b_k \vec{d}_j .\end{aligned} \tag{3.8}$$

This equation will be treated as part of a decomposition problem in chapters 4 and 6.

Note that it is possible to include the dispersion summand into the modal sum by defining an additional mode 0,

$$\vec{r}_{jk} = \Re\left\{\sum_{m=0}^{M} \vec{R}_{jm} \mathrm{e}^{-\mathrm{i}S_{jk}\mu_m/2} A^*_{km}\right\} \text{ with } \lambda_0 \equiv 1, \vec{R}_{j0} \equiv \vec{d}_j, A_{k0} \equiv b_k .$$

3.2.5. Effects of monitor distortions

We investigate how systematic inaccuracies of a monitor effect the applicability of the BE±d models. Approximating the deviations from the unperturbed closed orbit $\vec{r}$ to be small, such monitor distortions at a monitor j can be modeled via a *rectification Jacobian* $\mathbf{D}_j(\vec{r}^{\#}_j)$ for the unperturbed closed orbit as computed by the monitor system (#) at monitor j. Abbreviating the Jacobian as $\mathbf{D}_j$, one obtains

$$\vec{r}_{jk} = \mathbf{D}_j \vec{r}^{\#}_{jk} = \Re\Big\{\sum_m \underbrace{\mathbf{D}_j \vec{R}^{\#}_{jm}}_{\vec{R}_{jm}} E^*_{jkm} A^*_{km}\Big\} + \underbrace{\mathbf{D}_j \vec{d}^{\#}_j}_{\vec{d}_j} b_k .$$

Thus monitor distortions effect the model monitor parameters $\vec{R}_{jm}$ and $\vec{d}_j$ to their # counterparts, but will not prevent the applicability of the BE±d models. This also means that the correction of monitor distortions can be applied to the monitor parameters (see sec. 6.1.2).

Summary

To close this chapter, we summarize properties of the Bilinear-Exponential model that distinguish it from existing approaches to parameterize closed orbits.

1. *Monitor distortions.* Naturally, the values of BE+d monitor parameters change, but the resulting response can still be decomposed into a BE+d model, and rectification can be performed in the model parameters.
2. *Unknown corrector strengths and multipole fields.* Unknown corrector characteristics again only influence the corrector parameters A_{km}. As no assumptions about the exact nature of the perturbation were made but the fact that they only occur in the space between two adjacent monitors, multipolar errors of any kind will only change the A_{km} values (as long as the beam is not lost by the perturbation).

   If, on the other hand, the kick angle θ_k is known, the corrector can be approximated as having no extension in beam direction, not producing additional multipolar perturbations, and motion is decoupled (sec. 3.1.4), additional optical information is available at the correctors via (3.4), as (analogous for y)

$$\hat{X}(\tilde{s}_k) = \sqrt{\beta(\tilde{s}_k)}\mathrm{e}^{\mathrm{i}\phi(\tilde{s}_k)} = 2\frac{\sin(\mu_\mathrm{x}/2)}{\theta_k}\hat{A}_{k\mathrm{x}} \quad \text{for } k \in \mathcal{K}_\mathrm{x}.$$

3. *Inclusion of strong betatron coupling* has been performed without further considerations.

Further benefits of this model become apparent in chapters 4 and 5.

4. The response problem

Having developed the Bilinear-Exponential model with dispersion in chapter 3, we are able to rephrase and clarify the inverse problem mentioned in the introduction (chapter 1). After some considerations about the solvability of the problem under ideal conditions (no measurement errors, sec. 4.1), it is reformulated into an optimization problem (sec. 4.2) which will be the problem statement used by the optimization layer of the algorithm to be developed (sec. 6.1.1).

4.1. Problem statement and solvability

In the following considerations, we include dispersion and use the BE+d model (3.8),

$$r_{jkw} = \sum_m \Re \left\{ R_{jmw} \mathrm{e}^{-\mathrm{i}S_{jk}\mu_m/2} A^*_{km} \right\} + d_{jw} b_k, \tag{4.1}$$

where w denotes the respective vector component.[1]

As the available input data consists only of a general response matrix $\mathbf{r}$ and the topology matrix $\mathbf{S}$ of an accelerator, the inverse problem briefly discussed in the introduction (chapter 1) can be rephrased as shown in the following.

For given general response matrix $\mathbf{r}$ and topology matrix $\mathbf{S}$,
find all parameters of the BE+d model, given as
$R_{jmw}, A_{km}, \mu_m, d_{jw}, b_{kw}$ for each monitor j, corrector k, and direction w,
so that (4.1) is fulfilled.

Taking into account the phase advances μ_m, the number of dimensions for the search space is

$$D = M + 2(KM + JM^2) + K + JM \qquad \text{for the BE+d model and}$$
$$D = M + 2(KM + JM^2) \qquad \text{for the BE-d model,}$$

both with JKM constraints.

4.1.1. Scaling invariants

In (4.1), each response component consists of $M + 1$ summands. For the modal summands starting with $\Re$, we already know from sec. 3.1.2 that the monitor vector and corrector parameters for a given mode m can be scaled by any complex quantity C_m, which results in $2M$ degrees of freedom that do not change the summand. These degrees of freedom are equivalent to the unknown invariant of motion and starting phase of the respective oscillation mode m.

In addition, if all factors in the $\Re$ summands are simultaneously conjugated, the summand will not change. This ambiguity is equivalent to the quadrant problem (secs. 5.1.3 and 6.1.2).

[1] In other words, R_{jmw} respectively d_{jw} denotes the component of $\vec{R}_{jm}$ respectively $\vec{d}_j$ in direction w.

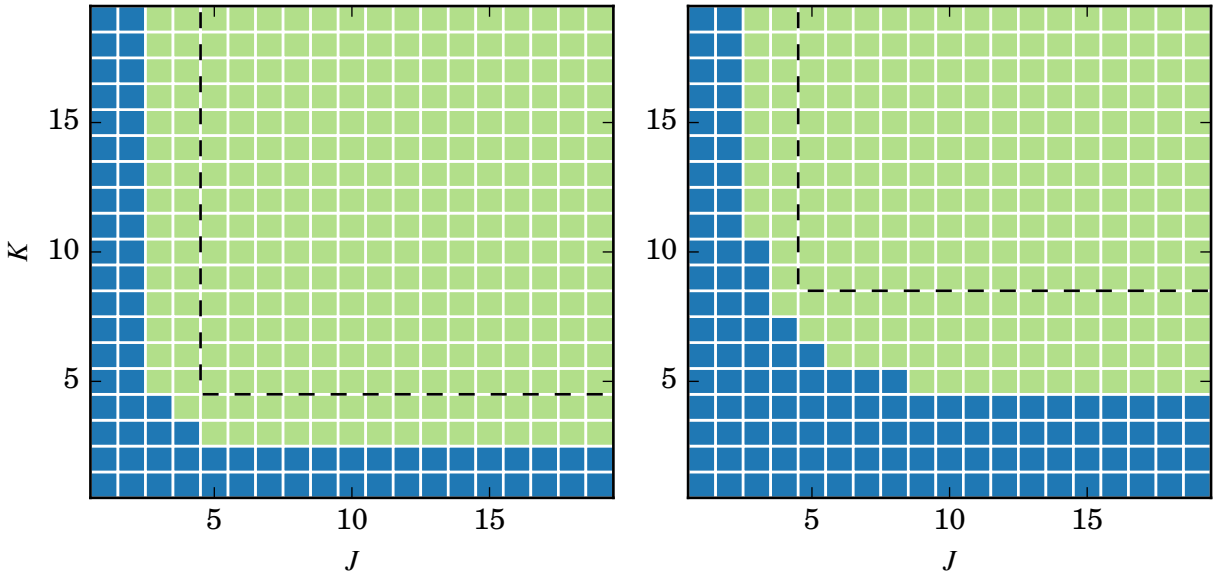

Figure 4.1.: Solvability of the response problem without dispersion (BE-d model) in dependence of the number of monitors J and correctors K for $M = 1$ (left) and $M = 2$ (right). Light green indicates that the problem is solvable. The dashed line separates the region in which solutions can be found by the MCS algorithm (sec. 5.2) from the one for few correctors and monitors for which this is impossible.

Further degrees of freedom are contained in the dispersion term, as any scaling by a real number does not affect the summands. This leads to $D_{\text{inv}} = 2M$ scaling invariants for the BE-d model and $D_{\text{inv}} = 2M + 1$ scaling invariants for the BE+d model, which do not change any summand of the equations.

This means that a solution can be constructed if $D - D_{\text{inv}} \leq JKM$, which is equivalent to

$$(2M+1)(K+JM) \leq (KJ+1)M+1 \qquad \text{for the BE+d model and}$$
$$\text{or} \quad 2(K+JM) \leq KJ+1 \qquad \text{for the BE-d model.}$$

The solvability of a set of equations in dependence of J and K is shown in Figs. 4.1 and 4.2.

4.1.2. BE+d model with fixed μ as bilinear equation system

One may reformulate (4.1) by constructing a vector $\vec{x}$ that holds the components of *all* A_{km} and b_k, and a vector $\vec{y}$ that holds components of all $\vec{R}_{jm}$ and $\vec{d}_j$ (sec. A.3.4). Using a set of JKM different sparse matrices $\mathbf{B}_{jkw}$ which components are either zero, $\mathrm{e}^{\pm \mathrm{i}\mu_m/2}$, or one, the equations can then be rewritten as (sec. A.3.4)

$$r_{jkw} = \Re\left\{\vec{y}^{\dagger}\mathbf{B}_{jkw}\vec{x}\right\}.$$

For a fixed set of matrices $\mathbf{B}_{jkw}(\mu_m)$ and unknown $\vec{x}, \vec{y}$, this is a bilinear equation system [24, 25].[2]

Several strategies exist to solve bilinear systems, which is generally considered a non-trivial task and an active topic of research [24, 25]. At the moment, no directly applicable global one-step solution[3] seems to exist. Even if it existed, we would have to include μ_m into an outer optimization loop (similar to LSSA optimization, sec. 8.2.1).

[2] Usually being defined for real-valued $\vec{x}, \vec{y}, \mathbf{B}$ and a single index $(j,k,w) \to b$, the above system can be redefined by elementary means to fulfill this condition.

[3] without selecting rows or columns like in the MCS algorithm (sec. 5.2)

Quadratic surfaces and solution spaces

The aforementioned system is asymmetric and contains the $\mathfrak{R}$ operator. To obtain more insight into the problem, we reformulate the bilinear system into a quadratic system. We obtain a set of quadratic (hyper)surface equations

$$2r_{jkw} = \begin{pmatrix} \vec{x} \\ \vec{y} \end{pmatrix}^{\dagger} \begin{pmatrix} & \mathbf{B}_{jkw}^{\dagger} \\ \mathbf{B}_{jkw} & \end{pmatrix} \begin{pmatrix} \vec{x} \\ \vec{y} \end{pmatrix}.$$

Each equation of this system (indexed by j, k, w) defines the solution to reside on a surface in the search space. Assuming no measurement noise, the solution is thus the intersection of all JKM surfaces defined by the equations. If noise is included in r_{jkw}, one would like to find the point that is closest to being an intersection by a given measure (see sec. 4.2).

If all hypersurfaces are different from each other, then each intersection of two surfaces reduces the dimension of the resulting solution space by at least one. This is the case if all matrices $\mathbf{B}_{jkw}/r_{jkw}$ differ from each other (see sec. A.3.4) and all $r_{jkw} \neq 0$.

For construction of bilinear equations, we have neglected and fixed the modal phase advances μ_m to their optimal values. If they are subjected to variation and not all S_{jk} are identical, the orientation of the quadric surfaces relative to another will change due to their dependence on $\mathbf{B}_{jkw}$, excluding the case of conjugating the respective summands (and thus changing the sign of μ_m) and the periodicity of μ_m.

While a full proof that a unique solution exists for all thinkable situations could not be given, we conclude that the solution is unique in the bilinear subspace (which only holds M dimensions less than the full problem) and very likely to be unique in the full solution space for non-degenerate oscillation eigenmodes at monitor and corrector positions. Note that in all cases where the MCS algorithm (sec. 5.2) can be applied, a unique solution must exist when neglecting dispersion; this is the case for common setups with $K \geq 4M$ and $J \geq 4$ (see sec. 5.2 and Figs. 4.1, 4.2).

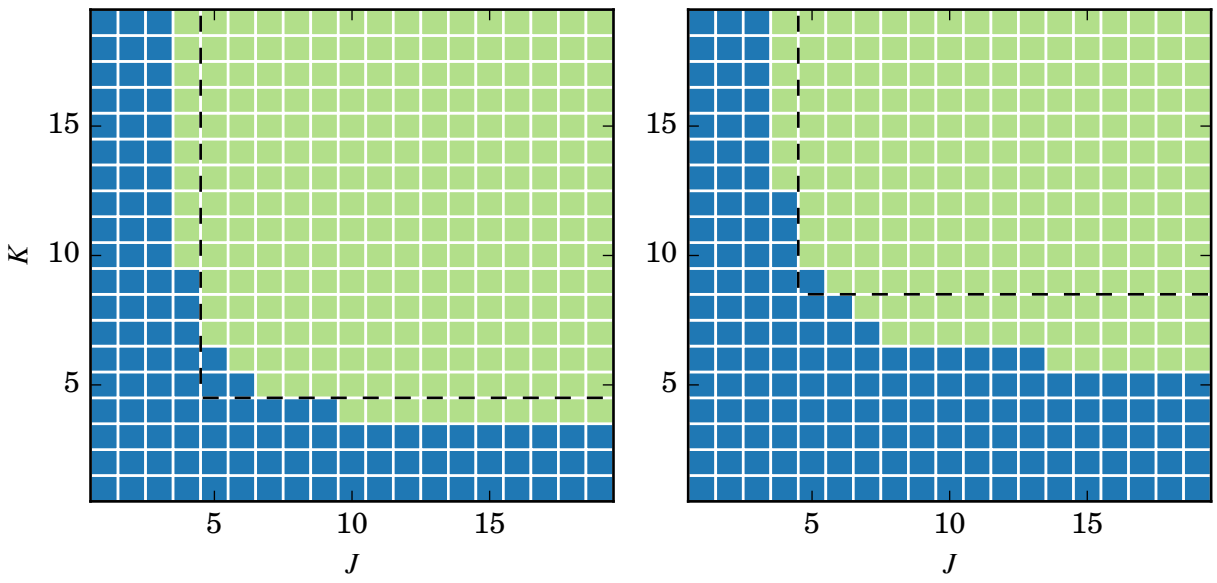

Figure 4.2.: Solvability of the response problem with dispersion (BE+d model) with identical indicators as in Fig. 4.1. Note that an overlap of unsolvable combinations and applicability of the MCS algorithm (sec. 5.2) occurs. For these 3 ($M = 1$) respectively 1 ($M = 2$) cases, the resulting BE+d model would be ambiguous.

4.2. Restatement as an optimization problem

We described the general response problem without any measurement error in the form of a bilinear-exponential equation system. In alternative to this approach, one may express the general response problem as a nonlinear least-squares problem [26] as follows.

For a given general response matrix $\mathbf{r}$ of size $J \times K$ with vector elements $\in \mathbb{R}^M$ and a topology matrix $\mathbf{S}$ of size $J \times K$,

$$\text{find } R_{jmw}, A_{km}, \mu_m, d_{jw}, b_k \text{ for all monitors } j, \text{ correctors } k, \text{ modes } m, \text{directions } w$$
$$\text{so that} \quad \chi^2 = \sum_{jkw} \left[f_{jkw}\left(\ldots, \mathbf{S}\right) - r_{jkw} \right]^2 \text{ is minimal.}$$

In this context, each $f_{jkw}(\ldots, \mathbf{S})$ is a regression function which is given by the right-hand side of (4.1) and thus depends on all variables therein. We can further abbreviate the notation by defining the BE+d vector $\vec{\rho}$ which components contain *all* $R_{jmw}, A_{km}, \mu_m, d_{jw}, b_k$ in an arbitrary but fixed order, and which is located in a D-dimensional (sec. 4.1) *search space*. By this definition one arrives at the following formulation.

$$\text{Find a BE+d vector } \vec{\rho} \quad \text{so that } \chi^2 = \sum_{jkw} \left[f_{jkw}\left(\vec{\rho}, \mathbf{S}\right) - r_{jkw} \right]^2 \text{ is minimal.} \tag{4.2}$$

As stated in sec. 4.1, this problem has D free parameters corresponding to dependent variables respectively number of dimensions of $\vec{\rho}$. Many of such high-dimensional problems are only solvable in reasonable time scales if information about gradients and higher derivatives is provided. As f_{jkw} consists only of linear and exponential parameters in the B±d model, we can find equivalently simple analytical expressions for gradient and second-order derivatives.

4.2.1. Gradient of χ^2 in search space (Jacobian matrix)

We follow the standard procedure of nonlinear regression problems [26]. To find the gradient, we introduce component residuals, we express the squared error in (4.2) as

$$\chi^2 = \sum_{jkw} \chi^2_{jkw} \quad \text{with fit residuals} \chi_{jkw} = f_{jkw}(\rho, \mathbf{S}) - r_{jkw}.$$

The gradient of the objective function in respect to $\vec{\rho}$, respectively its components ρ_n, is then given via chain rule as

$$\frac{\partial}{\partial \rho_n} \chi^2 = \sum_{jkw} \frac{\partial \chi_{jkw}}{\partial \rho_n} \frac{\partial \chi^2}{\partial \chi_{jkw}} = 2 \sum_{jkw} \frac{\partial f_{jkw}}{\partial \rho_n} \chi_{jkw}.$$

The remaining problem consists of finding the gradient of f_{jkw} in the space of dependent variables. By an index mapping $(j, k, w) \to b$, this can be rewritten defining the *Jacobian matrix* of residuals in search space $\mathbf{J}$ by

$$\frac{\partial}{\partial \rho_n} \chi^2 = 2 \sum_b J_{np}(\rho) \chi_b \text{ with } J_{nb}(\vec{\rho}) = \frac{\partial f_b(\vec{\rho})}{\partial \rho_n}$$
$$\text{or} \quad \vec{\nabla}_\rho \ \chi^2 = 2 \ \mathbf{J}(\vec{\rho}) \ \vec{\chi}.$$

The Jacobian matrix can be constructed by analytically calculating all respective gradients as done in the following. As some dependent variables are complex and others real, it is advisable to construct $\vec{\rho}$ using real and imaginary parts of the complex quantities separately. This corresponds to the following treatment of derivatives.

Gradients of linear quantities

With an arbitrary, complex constant a and a dependent variable ρ_m, we obtain two useful relations for the derivatives of linear quantities

$$\frac{\partial}{\partial \Re \rho_m} \sum_n \Re \left\{ a_n^* \rho_n \right\} = \frac{\partial}{\partial \Re \rho_m} \sum_n \Re \left\{ a_n \rho_n^* \right\} = \delta_{mn} \Re a_n,$$
$$\frac{\partial}{\partial \Im \rho_m} \sum_n \Re \left\{ a_n^* \rho_n \right\} = \frac{\partial}{\partial \Im \rho_m} \sum_n \Re \left\{ a_n \rho_n^* \right\} = \delta_{mn} \Im a_n.$$

To prevent index mixing, we replace the indices j, k, m, w corresponding to identical indices in the respective derivative variable by barred counterparts. Then we obtain the Jacobian matrix elements

$$\begin{aligned}
\frac{\partial f_{j\bar{k}w}}{\partial \Re A_{km}} &= \frac{\partial}{\partial \Re A_{km}} \sum_{\bar{m}} \Re \left\{ R_{j\bar{m}w} E^*_{j\bar{k}\bar{m}} A^*_{\bar{k}\bar{m}} \right\}, \\
&= \delta_{\bar{k}k} \Re \left\{ R_{jmw} E^*_{j\bar{k}m} \right\} \\
\frac{\partial f_{j\bar{k}w}}{\partial \Im A_{km}} &= \delta_{\bar{k}k} \Im \left\{ R_{jmw} E^*_{j\bar{k}m} \right\}
\end{aligned}$$

for the corrector parameters, so the corresponding objective function derivatives are

$$\begin{aligned}
\frac{\partial}{\partial \Re A_{km}} \chi^2 &= 2 \sum_{j\bar{k}w} \delta_{\bar{k}k} \chi_{j\bar{k}w} \Re \left\{ R_{jmw} E^*_{j\bar{k}m} \right\} \\
&= 2 \sum_{jw} \chi_{jkw} \Re \left\{ R_{jmw} E^*_{jkm} \right\}, \qquad (4.3) \\
\frac{\partial}{\partial \Im A_{km}} \chi^2 &= 2 \sum_{jw} \chi_{jkw} \Im \left\{ R_{jmw} E^*_{jkm} \right\}. \qquad (4.4)
\end{aligned}$$

With a similar computation, we have monitor vector gradients of f_{jkw} as elements of the Jacobian matrix briefly defined by

$$\left(\frac{\partial}{\partial \Re R_{jmw}} + \mathrm{i} \frac{\partial}{\partial \Im R_{jmw}} \right) f_{\bar{j}k\bar{w}} = \delta_{\bar{j}j} \delta_{\bar{w}w} A_{km} E_{\bar{j}\bar{k}m}, \quad \text{so that} \qquad (4.5)$$
$$\left(\frac{\partial}{\partial \Re R_{jmw}} + \mathrm{i} \frac{\partial}{\partial \Im R_{jmw}} \right) \chi^2 = 2 \sum_k \chi_{jkw} A_{km} E_{jkm}. \qquad (4.6)$$

Likewise we obtain the derivatives regarding dispersion parameters as

$$\frac{\partial}{\partial b_k} f_{j\bar{k}w} = \delta_{\bar{k}k} d_{jw} \qquad \Rightarrow \frac{\partial}{\partial b_k} \chi^2 = 2 \sum_{jw} \chi_{jkw} d_{jw} \qquad (4.7)$$
$$\frac{\partial}{\partial d_{jw}} f_{\bar{j}k\bar{w}} = \delta_{\bar{j}j} \delta_{\bar{w}w} b_k \qquad \Rightarrow \frac{\partial}{\partial d_{jw}} \chi^2 = 2 \sum_k \chi_{jkw} b_k. \qquad (4.8)$$

Phase advance derivatives

As the modal phase advances μ_m only occur in f_{jkw} as exponential dependencies ($E_{jkm} = \mathrm{e}^{\mathrm{i}S_{jk}\mu_m}$), one can compute the respective derivatives by elementary means as[4]

$$\begin{aligned}\frac{\partial f_{jkw}}{\partial \mu_{\bar{m}}} &= \frac{1}{2}\sum_m \left(R_{jmw}\frac{\partial}{\partial \mu_{\bar{m}}} E^*_{jkm} A^*_{km} + R^*_{jmw}\frac{\partial}{\partial \mu_{\bar{m}}} E_{jkm} A_{km} \right) \\ &= \frac{S_{jk}}{4\mathrm{i}} \left(R_{j\bar{m}w} E^*_{jk\bar{m}} A^*_{k\bar{m}} - R^*_{j\bar{m}w} E_{jk\bar{m}} A_{k\bar{m}} \right) \\ &= \frac{S_{jk}}{2} \Im \left\{ R_{j\bar{m}w} E^*_{jk\bar{m}} A^*_{k\bar{m}} \right\}.\end{aligned}$$

The corresponding derivative of the objective function follows as

$$\frac{\partial}{\partial \mu_m}\chi^2 = \sum_{jk} S_{jk} \sum_w \chi_{jkw} \Im \left\{ R_{jmw} E^*_{jkm} A^*_{km} \right\}. \tag{4.9}$$

One concludes that (after a lengthy but straightforward computation) all elements of the Jacobian matrix respectively the complete gradient information for the general response problem have been calculated.

4.2.2. Error computations and approximate Hessian

If we had found an optimal solution to the general response problem, contained in the solution vector $\vec{\rho}^{\,\mathrm{opt}}$, we would like to know how sensitive the solution is to small changes in that vector respectively the dependent variables. This information is contained in the Jacobian matrix for the solution $\mathbf{J}(\vec{\rho}^{\,\mathrm{opt}})$.

In vicinity of the optimal solution $\vec{\rho}^{\,\mathrm{opt}}$, the nonlinear problem is approximately equivalent to a least-squares *linear* problem of the form

$$r_b = \chi_b + \sum_n J_{nb} \cdot (\rho_n - \rho_n^{\mathrm{opt}}) \quad \text{with} \quad \mathbf{J} = \mathbf{J}(\vec{\rho}^{\,\mathrm{opt}}) = \text{const.}$$

For this case and identical noise levels for all inputs, one may compute an estimate of the covariance matrix as [26]

$$\mathbf{C} = \sigma^2 (\mathbf{J}\mathbf{J}^\dagger)^{-1},$$

where σ^2 can be interpreted as the best approximation error given in [27]. The significance respectively reliability of this error depends on an "overdetermination factor", which is the ratio of input data components JKM and the VC dimension [27], being $D - D_{\mathrm{inv}}$ in our case. Including its dependence on the *estimated* in-sample error χ^2, the best approximation error can be estimated as [27]

$$\sigma^2 = \frac{JKM}{JKM - D + D_{\mathrm{inv}}}\chi^2.$$

As a rule of thumb, the estimation becomes inaccurate for overdetermination factors ≤ 10 [27].

[4] For purposes of implementation, one could define variables $c_{jkwm} = R_{jmw} E^*_{jkm} A^*_{km}$ that can be used in the following and for computing the f_{jkw} terms.

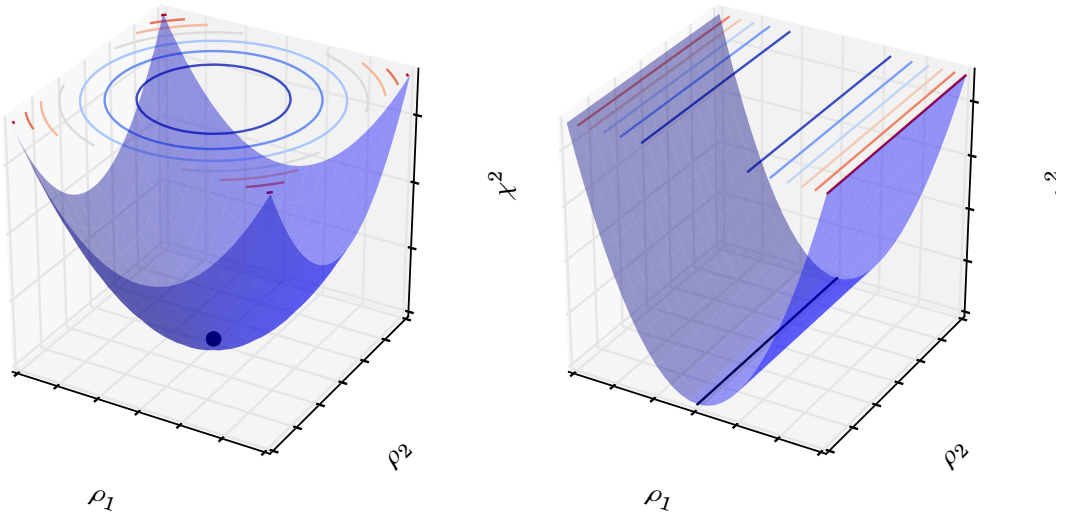

Figure 4.3.: Residual error function χ^2 in vicinity of a valid solution (black points and lines) to a two-dimensional optimization problem with isolines. Left: fully constrained problem. Right: underconstrained problem.

Hessian matrix

As f_{jkw} is an analytical expression, it is also possible to calculate higher-order derivatives. For the second-order derivative, this information is contained in the Hessian matrix $\mathbf{H}$ of the function $\chi^2(\vec{p})$, defined by its components via

$$\begin{aligned} H_{mn} &= \frac{\partial}{\partial p_m}\frac{\partial}{\partial p_n}\chi^2 = \frac{\partial}{\partial p_m} 2\sum_b J_{nb}(\vec{p})\,\chi_b \\ &= 2\sum_b \chi_b \frac{\partial}{\partial p_m} J_{nb}(\vec{p}) + 2\sum_b J_{mb}J_{nb} \end{aligned}$$

In the case of linearization with constant $\mathbf{J}$, this reduces to an approximate Hessian in vicinity of the solution vector

$$\tilde{\mathbf{H}} = 2\mathbf{J}\mathbf{J}^\dagger, \quad \text{so that} \quad \mathbf{C} = 2\sigma^2\tilde{\mathbf{H}}^{-1}.$$

Due to its symmetry, this approximate Hessian must be positive semidefinite, which is useful if its inverse has to be computed. Due to the scaling invariants (sec. 4.1.1) respectively the fact that the expressions f_{jkw} are underconstrained in terms of $\vec{p}$, we cannot expect $\tilde{\mathbf{H}}$ to possess full rank (see Fig. 4.3).

As the computation of an appropriate (pseudo-)inverse is required,[5] it is useful to express the covariance matrix by SVD [28] of $\mathbf{J}$ (sec. A.3.1) as

$$\mathbf{C} = \sigma^2\,\mathrm{pinv}(\mathbf{J}\mathbf{J}^\dagger) = \sigma^2\,\mathrm{pinv}(\mathbf{J}^\dagger)\,\mathrm{pinv}(\mathbf{J}) = \sigma^2\,\mathbf{U}\mathbf{S}_{\mathrm{inv}}\,(\mathbf{U}\mathbf{S}_{\mathrm{inv}})^\dagger.$$

It is possible to use an analytical cutoff criterion (sec. A.3.1) on $\mathbf{S}_{\mathrm{inv}}$. As the number of scaling invariants (sec. 4.1.1) is known for a given model, we set the same number of smallest singular values in the pseudoinverse to zero and invert the remaining singular values.

[5] Another numerical reason is the fact that computing the eigenvalues of $\tilde{\mathbf{H}}$ may be more efficient (especially as $\mathbf{J}$ is sparse, see sec. 6.1.2), but precision is lost due to explicit component multiplication. This is also circumvented by SVD [28].

4.3. Response assumptions

Up to this point, we have implicitly assumed that all general response matrix elements r_{jkw} are known. This is not given for many data sets recorded in storage rings.

In some measurement setups, the orbit perturbation in the direction w which is not under consideration is simply set to zero, so that

$$r_{jkw} = 0 \text{ for } k \notin \mathbb{K}_w.$$

Here, $\mathbb{K}_w$ is the set of corrector indices of the respective direction. This setup assumption is called decoupled measurement assumption in the following, in contrast to the coupled measurement assumption being used when all general response matrix elements are known.

As a result, two fully independent sets of equations

$$\begin{aligned} x_{jk} &= \Re\left\{X_j \mathrm{e}^{-\mathrm{i}S_{jk}\mu_\mathrm{x}/2} A^*_{k\mathrm{x}}\right\} + d_{j\mathrm{x}} b_{k\mathrm{x}} \quad \text{for } k \in \mathbb{K}_\mathrm{x}, \text{ and} \\ y_{jk} &= \Re\left\{Y_j \mathrm{e}^{-\mathrm{i}S_{jk}\mu_\mathrm{y}/2} A^*_{k\mathrm{y}}\right\} + d_{j\mathrm{y}} b_{k\mathrm{y}} \quad \text{for } k \in \mathbb{K}_\mathrm{y} \end{aligned} \tag{4.10}$$

exist for the decoupled measurement assumption. These two sets can each be represented by (4.1), when assuming the number of modes and directions $M = 1$, and then identifying x_{jk} respectively y_{jk} with r_{j11} and X_j respectively Y_j with R_{j11}. In consequence, the resulting optimization problem can be split into two optimization problems.

By inspecting (4.1), we note that only one dispersion orbit $\vec{d}_j$ exists in independence of the oscillation mode m. Thus, when using decoupled response matrices, the relative scaling of the dispersion orbit in horizontal and vertical direction is lost, and different b_{kw} variables are used for each direction w.

Note that this is only an assumption for a limited measurement, while the physics of the problem is always governed by (4.1). Therefore, it is always advisable to use complete response data and the coupled response assumption whenever possible. In consequence, we can only hope to describe the problem properly with (4.10) if the beam optics are decoupled in good approximation, that is $R_{jmw} \approx 0$ for $m \neq w$.

5. Procedures to find approximate BE parameters

Having described the general response problem as a nonlinear regression problem in sec. 4.2, sufficient starting values (respectively BE+d parameters) for the optimization procedure need to be found. Thus we introduce basic algorithms that can be used to extract approximate BE+d parameters from a general response matrix without optimization, gradually removing additional assumptions and input information but topology and increasing abstraction from existing measurement techniques.

Corrector-Monitor mapping (sec. 5.1) uses systems of linear equations to compute all monitor vectors from a subset of monitor information (neglecting dispersion) [23], formalizing and extending existing notions [29, 30, 11]. The missing monitor subset information has to be obtained by other means, e.g. lattice assumptions [29, 30] or multiturn measurements, as shown for a predecessor of CM mapping with additional assumptions in [11].

The Monitor-Corrector Subset algorithm (sec. 5.2) computes this missing information from a small number of $\mathbf{r}$ elements and topology, subsequently utilizing CM mapping to compute B-d parameters. While no additional assumptions are placed, the method relies on the accuracy of monitors in the arbitrarily chosen subset – a property originating from the MCS algorithm using CM mapping as a subroutine.

5.1. Corrector-Monitor (CM) mapping

By use of the BE-d model (sec. 3.2.3), we are able to obtain optical parameters at all monitors if monitor vectors at a sufficient subset of monitors are available (effectively *mapping* known to unknown monitors vectors). This is facilitated by subsequent application of the corrector (sec. 5.1.2) and monitor (sec. 5.1.1) subroutines, which is named as CM mapping (sec. 5.1.3) in this work.

Somewhat outside of the scope of this thesis, we will show a basic example of using CM mapping to process incomplete Turn-by-Turn (TbT) data. This technique called TbT-assisted mapping (sec. 5.1.4) has been developed by the author in previous works [23, 11]. Note that the MCS algorithm (sec. 5.2) which uses CM mapping does not require any TbT data as input.

5.1.1. Monitor subroutine

The basic notion for the development of the monitor subroutine has been used in determining β functions based on lattice model values [29]. Numerical validation of a monitor subroutine precursor had been performed for a simulation of the Fermilab Tevatron [29], while a similar technique was also developed and used for measurements at the KEKB storage rings, again using lattice model data as input [30, 16, 17].

Equation system for all correctors and one monitor

Reviewing (3.6), it is valid to replace the braced term by its complex conjugate. Also, we transpose both sides of the equation. As the left side is real, transpose and matrix adjunct refer to the same operation.

Therefore we can apply the matrix adjunct to the braced term and the left side of the equation, so that

$$\vec{r}^{\dagger}_{jk} = \Re\left\{\sum_m A_{km}E_{jkm}\vec{R}^{\dagger}_{jm}\right\} = \Re\left\{\begin{pmatrix} A_{k1}E_{jk1} & \dots & A_{kM}E_{jkM}\end{pmatrix}\begin{pmatrix}\vec{R}^{\dagger}_{j1}\\ \vdots \\ \vec{R}^{\dagger}_{jM}\end{pmatrix}\right\}$$

As this matrix does not depend on the k index, one can create a linear equation system

$$\begin{pmatrix}\vec{r}^{\dagger}_{j1}\\ \vdots \\ \vec{r}^{\dagger}_{jK}\end{pmatrix} = \Re\left\{\mathbf{G}_j\begin{pmatrix}\vec{R}^{\dagger}_{j1}\\ \vdots \\ \vec{R}^{\dagger}_{jM}\end{pmatrix}\right\} \quad \text{or}$$

$$\begin{pmatrix}\vec{r}_{j1} & \dots & \vec{r}_{jK}\end{pmatrix}^{\dagger} = \Re\left\{\mathbf{G}_j\begin{pmatrix}\vec{R}_{j1} & \dots & \vec{R}_{jM}\end{pmatrix}^{\dagger}\right\} \tag{5.1}$$

which includes all corrector parameters. In the process of simplification we have introduced J matrices $\mathbf{G}_j$ with components defined by

$$\left(\mathbf{G}_j\right)_{km} = A_{km}E_{jkm} = A_{km}\mathrm{e}^{\mathrm{i}S_{jk}\mu_m/2}. \tag{5.2}$$

One concludes that, if not being underconstrained, (5.1) can be used to obtain unknown monitor vectors $\vec{R}_{jm}$ from (all) corrector parameters, corresponding orbit perturbations at the monitors, and knowledge of μ_m.

Decoupled case Note that the monitor equation system is independent for different column vectors of the monitor matrix and the general response matrix, which correspond to different directions. Thus, in the decoupled case, the system can be simplified as e.g. for a horizontal corrector k, the columns corresponding to vertical deviations are zero and do not contribute to the solution.

Equation systems containing $\Re$

Although being a system of linear equations, (5.1) has the slightly unusual form

$$\vec{b} = \Re\left\{\mathbf{C}\vec{x}\right\} \quad \text{or} \quad b_p = \sum_q \Re\left\{C_{pq}x_q\right\}.$$

This is not problematic, as each summand on the right-hand side can be rewritten as

$$\Re\left\{C_{pq}x_q\right\} = \Re C_{pq}\Re x_q - \Im C_{pq}\Im x_q.$$

By this equivalence one may reformulate the original system as

$$\vec{b} = \begin{pmatrix}\Re\mathbf{C} & -\Im\mathbf{C}\end{pmatrix}\begin{pmatrix}\Re\vec{x}\\ \Im\vec{x}\end{pmatrix}$$

which is a standard, real-valued equation system. In comparison with the original system, it has the same number of unknowns, which were contained pairwise in complex numbers of the original form.

Thus, (5.1) is only solvable for all monitor vectors $\vec{R}_{jm}$ if $K \geq 2M$, so that each considered oscillation mode m requires application of $K = 2$ different orbit perturbations to be solved. If the system is overdetermined, the SVD algorithm can be used to find the least-squares result (sec. A.3.1).

$\mathbf{G}_j$ regions and matrix conditions dependent on correctors

As the only dependence of matrix $\mathbf{G}_j$ in (5.2) on j is given by the phase jump coefficients $E_{jkm} = \mathrm{sign}(\tilde{s}_k - s_j)$, it can be stated that many $\mathbf{G}_j$ matrices are identical, and that in fact only $K+1$ different matrices $\mathbf{G}_j$ can occur in the equations. Of these, the two matrices $\mathbf{G}_j$ for $s_j > \tilde{s}_K$ and for $s_j < \tilde{s}_1$ are also identical up to a phase factor, which corresponds to the "switching of turns" at $s = 0$.

To solve (5.1) for monitor vectors, the system matrix generated from $\mathbf{G}_j$ using sec. 5.1.1 must be non-singular. This condition can be formalized for special cases, in which all correctors are located between a given pair of monitors p and $p+1$. Then,

$$S_{jk} = \mathrm{sign}(p + 1/2 - j) = \begin{cases} +1 & \text{for } j \leq p \\ -1 & \text{for } j \geq p+1 \end{cases}$$

Consequently, only two unique $\mathbf{G}_j$ exist in this special case, which are identical up to phase factors and are equally well conditioned. The matrices are singular

$$\text{if } \begin{pmatrix} \Re\mathbf{A} & \Im\mathbf{A} \end{pmatrix} \quad \text{is singular.}$$

This statement is equivalent to the one that different correctors should not be placed at conditions with the same betatron phase in any direction to prevent the resulting orbit perturbations to be identical, in which case the correctors cannot span the set of independent orbits. Thus storage rings are very often designed to prevent this case by choosing the corrector positions so that independent perturbations occur.

In the general case, the above condition is modulated by the phase jump coefficients in the different corrector regions. Nevertheless, there are usually enough correctors in a ring to span the orbit space in each region, and the conditioning can also be checked by numerical means.

5.1.2. Corrector subroutine

Again starting from the BE-d model, we construct a simple corrector subroutine with which unknown corrector parameters can be found from a subset of monitor vectors.

Monitor subsets

In the following, beam oscillation parameters from a subset of input monitors will be mapped to those of output capable monitors. In consequence, the set of all monitors represented by the index set

$$\mathbb{J} = \{j \mid j \in \mathbb{N}^+ \wedge j \leq J\}$$

will be split into subsets ($\mathbb{N}^+$ does not include zero).

The previous expressions in which j occured, e.g. (5.1), have been independent for each j and thus are indifferent to the ordering of elements $\in \mathbb{J}$ as long as the ordering is consistent throughout the calculations. This statement also holds for any subset of $\mathbb{J}$. To keep the ordering consistent, we define a monitor subset by a sequence j_f of length F that contains only unique elements, so that

$$\mathbb{J}_{\mathrm{sub}} = \left\{ j_f \mid f \in \mathbb{N}^+ \wedge f \leq F \wedge j_f \in \mathbb{J} \right\},$$

with f being a second-order index set (see Fig. 5.1). Naturally any equation in chapter 3 including j and the BE-d model (3.6) can be applied to a monitor subset by replacing $j \rightarrow j_f$.

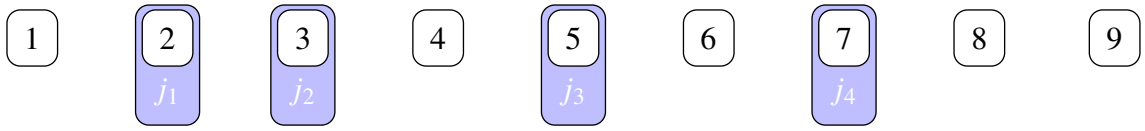

Figure 5.1.: The sequence $j_f \in \mathbb{J}_{sub}$ includes all monitor indices of the subset. The figure shows a subset with $F = 4$.

Equation system for multiple monitors and one corrector

Reviewing (3.6), we can separate A^*_{km} from the other components and use matrix notation. Furthermore, we use the subset notation, so that for any monitor subset $\mathbb{J}_{\mathrm{sub}}$

$$\vec{r}_{j_f k} = \Re\left\{\begin{pmatrix}\vec{R}_{j_f 1}E^*_{j_f k1} & \dots & \vec{R}_{j_f M}E^*_{j_f kM}\end{pmatrix}\begin{pmatrix}A^*_{k1}\\ \vdots \\ A^*_{kM}\end{pmatrix}\right\}.$$

Note that we can rewrite the vector containing corrector parameters as $\begin{pmatrix}A_{k1} & \dots & A_{kM}\end{pmatrix}^\dagger$. It is clear that this vector does not depend on the monitor index j respectively j_f. Then one system containing all orbit perturbations and monitor vectors of the subset can be written as

$$\begin{pmatrix}\vec{r}_{j_1 k}\\ \vdots \\ \vec{r}_{j_F k}\end{pmatrix} = \Re\left\{\begin{pmatrix}\vec{R}_{j_1 1}E^*_{j_1 k1} & \dots & \vec{R}_{j_1 M}E^*_{j_1 kM}\\ \vdots & \ddots & \vdots \\ \vec{R}_{j_F 1}E^*_{j_F k1} & \dots & \vec{R}_{j_F M}E^*_{j_F kM}\end{pmatrix}\begin{pmatrix}A_{k1} & \dots & A_{kM}\end{pmatrix}^\dagger\right\} \text{ for each } k. \tag{5.3}$$

Considering sec. 5.1.1, each equation system (5.3) consists of MF equations with M unknown complex numbers implying $2M$ unknowns. It follows that each system can only be constrained if the monitor subset holds indices of $F \geq 2$ monitors.

Like for the $\mathbf{G}_j$ regions in the monitor subroutine, one may apply singularity conditions to the system matrix of (5.3). Again neglecting phase jumps, the condition reduces to the monitor vectors $\vec{R}_{j_f m}$ being linear dependent.

Decoupled case Again, we can use the decoupled approximation to simplify (5.3). This results in ignoring all rows of the above equation system in which directions and modes are not coupled, so that K different equation systems

$$\begin{pmatrix}x_{1k}\\ \vdots \\ x_{Fk}\end{pmatrix} = \Re\left\{\begin{pmatrix}X_1 E^*_{1k\mathrm{x}}\\ \vdots \\ X_F E^*_{Fk\mathrm{x}}\end{pmatrix}A^*_{k\mathrm{x}}\right\} \text{ for each } k$$

exist. This is also only feasible if the system is not underconstrained, which corresponds to $F \geq 2$.

Matrix conditions dependent on the monitor set

As with the corrector subroutine, there are also conditions for the input monitor vectors to be fulfilled so that the resulting matrix is well-conditioned. In line with the argumentation in sec. 5.1.1 about corrector regions, we state that in the special case of only two regions, the system matrix is singular

$$\text{if } \begin{pmatrix}\Re\mathbf{R} & \Im\mathbf{R}\end{pmatrix} \text{ with } \mathbf{R} = \begin{pmatrix}\vec{R}_{j_1 1} & \dots & \vec{R}_{j_1 M}\\ \vdots & \ddots & \vdots \\ \vec{R}_{j_F 1} & \dots & \vec{R}_{j_F M}\end{pmatrix} \text{ is singular.}$$

For $M = 2$, this is e.g. given if the betatron phase advances between two monitor vectors of different monitors are close to $\pi/2$ for both modes. The matrix condition number is also modulated by the phase jump coefficients, but the problem is unlikely to be well-conditioned if the previous singularity condition is met.

5.1.3. Properties of CM mapping

After having characterized two linear equation systems that allow to obtain corrector parameters from monitor vectors and vice versa, we can summarize that all monitor vectors (for which closed-orbit perturbation information exists) can be obtained by mapping a subset of known monitor vectors to unknown monitor vectors. This is done by applying the corrector subroutine and the monitor subroutine in succession, and this procedure will be called CM mapping in the following.

CM mapping and postprocessing steps

By replacing all $\vec{R}_{j_f m} E_{j_f k m}$ terms in the corrector subroutine equation system with their complex conjugates, the resulting corrector parameters A_{km} will also be conjugated. This in turn also holds for the conjugated monitor vectors $\vec{R}_{jm}$ resulting from application of the monitor subroutine.

As will be shown in sec. 6.1.2, the conjugation of the above terms is exactly what is accomplished by the quadrant correction. Furthermore, the mapping $\vec{R}_{j_f m}$ to $\vec{R}_{jm}$ is linear and the mapping from $R_{j_f m}$ to A_{km} is inverse. These scaling properties are equivalent to scaling invariants (sec. 4.1.1). One may summarize these properties to the informal statement that the CM mapping *commutes* with quadrant correction and invariant normalization.

This alleviates the conditions on the storage ring to be met for finding the invariant. If one had to apply invariant normalization before CM mapping (as was the case for the predecessor of CM mapping in [11]), a known segment had to exist between the pair of monitors $\mathbb{J}_{\mathrm{sub}}$ for which the monitor vectors were known. Instead, as the postprocessing can be applied after CM mapping, there only needs to exist a known segment between *any* pair of monitors from $\mathbb{J}$.

Summary

A predecessor of CM mapping has been given in [11] for the standard Courant-Snyder parameterization, where the complications of using this polar-like parameterization became visible in the resulting nonlinear equation systems and their solutions. Using the BE-d model, both underlying subroutines have been generalized and expressed in a concise and directly applicable form as linear system solvers.

Besides the storage ring topology $\mathbf{S}$, the input for CM mapping only consists of $F \geq 2$ known monitor vectors and the modal phase advances μ_m (see also Fig. 5.2), and the origin of this parameters is not specified. Nevertheless, CM mapping reduces the problem of finding all parameters of the B-d model to finding the aforementioned set of inputs.

5.1.4. A digression to TbT-assisted mapping

As an example for the possible direct applications of CM mapping, we will briefly discuss its application to extract information from incomplete TbT monitor data (originating from a monitor subset) of a storage ring.

TbT-assisted mapping [23] is the successor of a method described by the author in [11]. It yields equivalent results in the given case described therein, but drops the condition on a drift space to be enclosed by TbT monitors – instead, general monitors enclosing a known segment are sufficient. In contrast to [11], using the BE-d model allows to write the corrector and monitor subroutines as simple linear equation systems, while also unifying the treatment of decoupled and coupled linear motion. Its application can be summarized by the following steps (see Fig. 5.2).

1. Obtain modal phase advances μ_m and monitor vectors $\vec{R}_{jfm}$ for the set of TbT-capable monitors by TbT analysis (sec. 8.2).

   For TbT-assisted mapping, the set of TbT-capable monitors corresponds to the monitor subset $\mathbb{J}_{\mathrm{sub}}$ introduced in sec. 5.1.2. In consequence, F corresponds to the number of TbT-capable monitors installed in the storage ring, and the algorithm can only be applied for $F \geq 2$.

2. Apply the CM routine to compute all corrector parameters A_{km}, and subsequently, all monitor vectors $\vec{R}_{jm}$ from closed-orbit perturbations $\vec{r}_{jk}$.

Any technique relying solely on CM mapping has a characteristic disadvantage, which can be summarized by the statement that the monitors used for input data need to have a high signal-to-noise ratio, as all other BE-d parameters depend on their accuracy.

Error in the input data will inversely map to the computed corrector parameters and at least linearly to the computed monitor vectors. This problem is can be partially circumvented if the errors are not systematic and input data from a larger number of monitor vectors is available. In this case, the residual error of the overdetermined monitor equation system is minimized by Singular Value Decomposition [28] (sec. A.3.1).

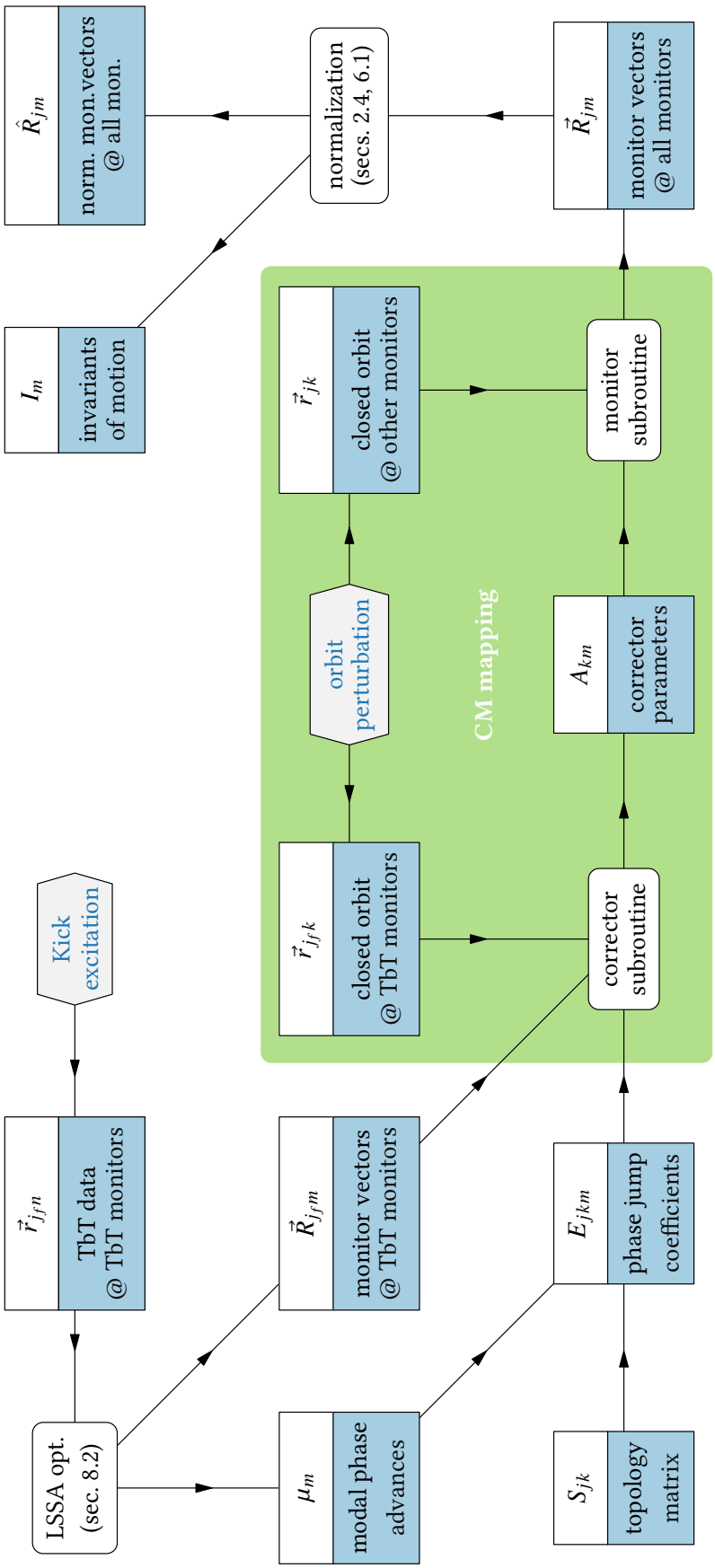

Figure 5.2.: Simplified flow diagram for TbT-assisted mapping (rectangles denote variables, rounded rectangles denote functions, hexagons denote physical processes). Note that all steps outside of the CM mapping also occur when TbT analysis (sec. 8.2) is performed – in this case, the CM mapping part is simply reduced to an identity, as TbT monitors then equal all monitors under consideration.

5.2. Monitor-Corrector Subset (MCS) algorithm

This section describes how approximate BE+d model parameters for a general response matrix can be obtained by performing special computations on four arbitrarily selected monitors $j_1, \ldots, j_4$. This approach works under the following conditions.

1. The signal-to-noise level is sufficiently high.
2. The dispersive effects occurring in the response are small and can be treated as perturbations.

Revisiting CM mapping, we were able to reduce the number of necessary monitor input vectors to 2. In fact, this information is only needed to compute the eigenform $\mathrm{e}^{\mathrm{i}\mu_m}, \vec{Z}_m(s_j)$ of a matrix similar or identical to the one-turn transfer matrix $\mathbf{T}(s_j)$ at a single position s_j. It is thus a large conceptual, but small computational step to compute $\mathbf{T}(s_j)$ using *only* closed-orbit perturbations.

We return to the analysis of unknown segments which are enclosed by known segments introduced in sec. 2.2.2, neglecting dispersive effects. While this technique is often used to compute transfer maps along linear accelerators [16, 17], no assumptions have been made that prevent this technique to be applied to storage rings, as long as the perturbation locations are outside of the segments under consideration.

We introduce the simple but important idea to divide the storage ring into two unknown segments which ends are connected by two known segments (see Fig. 5.3), so that measurement of the unknown segments' transfer maps is possible. Then, these can simply be concatenated to obtain the one-turn transfer map at the start respectively end of any of the segments. The transfer map can then be decomposed into monitor vectors and phase advances and used as input for CM mapping.

5.2.1. Monitor and corrector subsets

To formalize the aforementioned notion, we fix the beginnings respectively ends of the segments to monitor positions s_j for the monitor subset $j_1, \ldots, j_4$ (see Fig. 5.3). The beginning respectively end of the first known segment along s is linked to the monitors j_1 respectively j_2, and that of the second known segment to j_3 respectively j_4. Then, on the particle path between the monitors j_4 and j_1, the s coordinate is reset to $s = 0$.

We need to introduce corrector subsets to ensure that for each measurement of an unperturbed segment, only correctors outside of the unperturbed segment and its enclosing known segments are used. For each measurement cycle A, B (see Fig. 5.3) they are defined as

$$\begin{aligned} \mathbb{K}_\mathrm{A} &= \left\{k \mid \tilde{s}_k \notin [s_{j_1}, s_{j_4}]\right\} = \left\{k \mid S_{j_1 k} > 0 \vee S_{j_4 k} < 0\right\} \\ \text{and } \mathbb{K}_\mathrm{B} &= \left\{k \mid \tilde{s}_k \in]s_{j_2}, s_{j_3}[\right\} = \left\{k \mid S_{j_2 k} < 0 \wedge S_{j_3 k} > 0\right\}. \end{aligned} \tag{5.4}$$

Like in sec. 5.1.2, we introduce sequences $k_g, \bar{k}_{\bar{g}}$ to enumerate all elements in both index sets with G respectively $\bar{G}$ elements, so that

$$\mathbb{K}_\mathrm{A} = \{k_1, k_2, \ldots, k_G\} \text{ and } \mathbb{K}_\mathrm{B} = \left\{\bar{k}_1, \bar{k}_2, \ldots, \bar{k}_{\bar{G}}\right\} \text{ hold.}$$

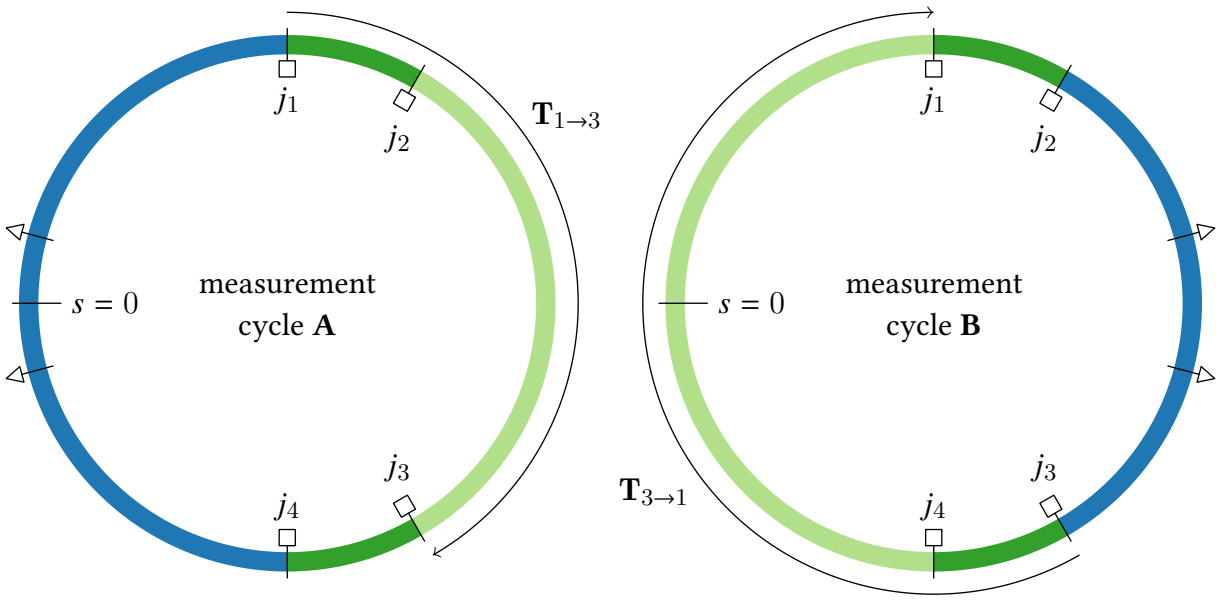

Figure 5.3.: Basic idea for the MCS algorithm (colors and symbols analogous to Fig. 2.1). In each measurement cycle, a transfer map including one of the unknown segments is obtained. Note that in sec. 5.2.2, the known segments are replaced by unperturbed segments, so that no requirements on known segments remain to obtain unnormalized monitor data.

5.2.2. One-Turn transfer maps from closed orbits

In sec. 2.2.1, general expressions have been derived to obtain the phase space at the beginning of a known segment, given here as s_{j_1} and s_{j_3}, so that

$$\begin{aligned} \vec{r}'_{j_1 k} &= \mathbf{T}^{-1}_{r\to r'}(s_{j_1} \to s_{j_2}) \left[\vec{r}_{j_2 k} - \mathbf{T}_{r\to r}(s_{j_1} \to s_{j_2})\, \vec{r}_{j_1 k}\right] \\ \text{and} \quad \vec{r}'_{j_3 k} &= \mathbf{T}^{-1}_{r\to r'}(s_{j_3} \to s_{j_4}) \left[\vec{r}_{j_4 k} - \mathbf{T}_{r\to r}(s_{j_3} \to s_{j_4})\, \vec{r}_{j_3 k}\right] \end{aligned}$$

hold in the linear case for a given perturbation k.

By computing these values, the phase space vectors for each perturbation are known and can be used to obtain the transfer matrix $\mathbf{T}_\mathrm{A}$ by solving the equation system

$$\begin{pmatrix} \vec{r}_{k_1 j_3} & \cdots & \vec{r}_{k_G j_3} \\ \vec{r}'_{k_1 j_3} & \cdots & \vec{r}'_{k_G j_3} \end{pmatrix}^\dagger = \begin{pmatrix} \vec{r}_{k_1 j_1} & \cdots & \vec{r}_{k_G j_1} \\ \vec{r}'_{k_1 j_1} & \cdots & \vec{r}'_{k_G j_1} \end{pmatrix}^\dagger \mathbf{T}^\dagger_\mathrm{A}$$

for all rows of $\mathbf{T}_\mathrm{A}$. To obtain $\mathbf{T}_\mathrm{B}$, the indices of $j_1 \leftrightarrow j_3$ in the above equation system are permuted and the sequence $k_1, \ldots, k_G$ is replaced by its counterpart $\bar{k}_1, \ldots, \bar{k}_{\bar{G}}$ for measurement cycle B. Note that $G, \bar{G} \geq 2M$ is required for the systems not to be underconstrained – this is in coherence with the notion of four independent orbits which will be used in the next subsection. It also implies that $K \geq G + \bar{G} \geq 4M$ is required for the MCS algorithm to be applied to a response matrix.

In conclusion, the straightforward possibility to compute $\mathbf{T}$ at s_1 (or s_3) is then to simply concatenate the partial transfer matrices by

$$\mathbf{T}(s_1) = \mathbf{T}_\mathrm{B}\mathbf{T}_\mathrm{A}.$$

From this one-turn transfer matrix, μ_m and $\vec{Z}_m(s_1)$ can be computed using its eigenform via

$$\mathrm{e}^{\mathrm{i}\mu_m}\vec{Z}_m(s_1) = \mathbf{T}(s_1)\vec{Z}_m(s_1).$$

Returning to the properties of one-turn transfer matrices described in sec. 2.3.2, it is known that the eigenvectors and eigenvalues occur in complex-conjugate pairs, so that eigenvalues corresponding to $\pm\mu$ exist.

One notes that this is the same ambiguity encountered when investing the CM mapping algorithm (sec. 5.1.3). It was shown that the CM mapping algorithm is indifferent to this ambiguity, and that the invariant postprocessing algorithm will correct quadrants and monitor vectors at the end of the computation (sec. 6.1.2). Hence, the ambiguity is not problematic.

From known to unperturbed segments

Up to now we have assumed that the segment transfer matrices between s_1 and s_2 respectively s_3 and s_4 are known a priori. What statements can be made if this is not the case?

Assuming all segments to be linear, there exists a linear unknown map $\mathbf{P}_1$ that relates phase space vectors $\vec{z}$ at the beginning of the segment to the spatial deviations at both ends of it (*composite monitor space*) via

$$\begin{pmatrix}\vec{r}_{j_1k}\\ \vec{r}_{j_2k}\end{pmatrix} = \mathbf{P}_1 \begin{pmatrix}\vec{r}_{j_1k}\\ \vec{r}'_{j_1k}\end{pmatrix} \quad \text{and} \quad \begin{pmatrix}\vec{r}_{j_3k}\\ \vec{r}_{j_4k}\end{pmatrix} = \mathbf{P}_3 \begin{pmatrix}\vec{r}_{j_3k}\\ \vec{r}'_{j_3k}\end{pmatrix}.$$

We now define a similarity transfer matrix $\tilde{\mathbf{T}}(s_1)$ that does not map from and to the particle phase space, but from and to the composite monitor space. In analogy to the case of known segments, this matrix may be computed as a product

$$\tilde{\mathbf{T}}(s_1) = \tilde{\mathbf{T}}_\mathrm{B}\tilde{\mathbf{T}}_\mathrm{A}, \tag{5.5}$$

$$\text{by solving the system } \begin{pmatrix}\vec{r}_{k_1j_3} & \cdots & \vec{r}_{k_Gj_3}\\ \vec{r}_{k_1j_4} & \cdots & \vec{r}_{k_Gj_4}\end{pmatrix}^\dagger = \begin{pmatrix}\vec{r}_{k_1j_1} & \cdots & \vec{r}_{k_Gj_1}\\ \vec{r}_{k_1j_2} & \cdots & \vec{r}_{k_Gj_2}\end{pmatrix}^\dagger \tilde{\mathbf{T}}_\mathrm{A}^\dagger \tag{5.6}$$

as well as the corresponding system for $\tilde{\mathbf{T}}_\mathrm{B}$.

By definition of $\tilde{\mathbf{T}}(s_1)$, we can either relate it to $\mathbf{T}(s_1)$ via similarity transform

$$\tilde{\mathbf{T}}(s_1) = \mathbf{P}_1\mathbf{T}(s_1)\mathbf{P}_1^{-1}$$

or directly argument with the eigenorbit conditions in sec. 2.3, to arrive at the eigenproblem for $\tilde{\mathbf{T}}(s_1)$

$$\mathrm{e}^{\mathrm{i}\mu_m}\begin{pmatrix}\vec{R}_{j_1m}\\ \vec{R}_{j_2m}\end{pmatrix} = \tilde{\mathbf{T}}(s_1)\begin{pmatrix}\vec{R}_{j_1m}\\ \vec{R}_{j_2m}\end{pmatrix}. \tag{5.7}$$

In conclusion we can indeed obtain data that is equivalent to single-particle motion, given by the monitor vectors $\vec{R}_{j_1m}, \vec{R}_{j_2m}$ for and modal phase advances μ_m for all modes m, even if there are only unperturbed and perturbed segments in the ring. A summary of an algorithm using above relations is given in Fig. 5.4.

As one can obtain monitor vectors for two monitors j_1 and j_2 using the algorithm, the minimal condition $F \geq 2$ for correct application of monitor and corrector subroutines (chapter 5.1) is fulfilled. It is then possible to obtain monitor vectors and corrector parameters for all monitors j and correctors k using CM mapping, entirely without any other measurement data than closed orbits and topology.

Data:
orbit perturbations $\vec{r}_{jk}$ for $J \geq 4$ monitors, $K \geq 4M$ correctors, in M spatial dimensions
monitor-corrector topology S_{jk} monitor split indices $j_1 - j_4$
Result:
oscillation-equivalent data $\mu_m, \vec{R}_{jm}$ for M modes and spatial dimensions, at monitors $j \in \{j_1, j_2\}$;
for *each measurement cycle A, B* **do**
  k_g respectively $\bar{k}_{\bar{g}} \leftarrow$ compute the respective subset $\mathbb{K}_{\mathrm{A,B}}$ (5.4) and enumerate it;
  $\tilde{\mathbf{T}}_{\mathrm{A}}$ respectively $\tilde{\mathbf{T}}_{\mathrm{B}} \leftarrow$ compute using orbit perturbations (5.6);
end
$\tilde{\mathbf{T}}(s_1) = \tilde{\mathbf{T}}_{3\to1}\tilde{\mathbf{T}}_{1\to3}$;
for *each eigenvalue* λ_m *of* $\tilde{\mathbf{T}}(s_1)$ **do**
  $\mu_m \leftarrow$ complex angle of λ_m;
  $\begin{pmatrix}\vec{R}_{j_1 m}\\ \vec{R}_{j_2 m}\end{pmatrix} \leftarrow$ eigenvector for λ_m;
end

Figure 5.4.: Simplified algorithm for generating monitor vectors and modal phase advances (equivalent to particle oscillation data) from closed-orbit perturbations. Note that the indices j_1, j_2 have to be chosen arbitrarily. The result can then be used as input data for CM mapping (sec. 5.1).

5.2.3. Principal component eigenvalue problem

While the discussed approach to compute $\tilde{\mathbf{T}}$ works in principle, it is more susceptible to noise than necessary. For this reason a possibility to reduce this effect is discussed.

In each measurement cycle, there exists a large segment that is treated as unperturbed. The monitors in each segment are represented by monitor subsets

$$\mathbb{J}_{\mathrm{A}} = \left\{ j \mid s_j \in [s_1, s_4] \right\} \text{ and } \mathbb{J}_{\mathrm{B}} = \left\{ j \mid s_j \notin [s_2, s_3] \right\}.$$

As we are still neglecting dispersive effects (sec. 3.1.1) it is known that any orbit inside an unperturbed segment can be expressed as a linear combination of $2M$ orbits.

These relations can be exploited using Principal Component Analysis [31, 32] (PCA) which decomposes a set of input vectors from the same space into a set of ordered, orthonormal basis vectors called principal components. The ordering of principal components depends on their relative, overall contribution to the input vectors. PCA is usually performed using Singular Value Decomposition [28, 33, 34] (SVD, sec. A.3.1).

We now take the general response matrix elements for monitor and corrector subsets A and decompose them into principal components $\langle \vec{r} \rangle^{\mathrm{A}}_{jp}$. The first $P = 4$ principal components are then considered as relevant, while the others are treated as noise, so that

$$\begin{aligned} \vec{r}_{jk} &= \sum_p \langle \vec{r} \rangle^{\mathrm{A}}_{jp} a_{pk} + \text{noise} \\ &= \langle \mathbf{r} \rangle^{\mathrm{A}}_{j} \vec{a}^{\mathrm{A}}_{k} + \text{noise} \qquad \text{for } j \in \mathbb{J}_{\mathrm{A}}, \quad k \in \mathbb{K}_{\mathrm{A}}. \end{aligned}$$

$\vec{a}^{\mathrm{A}}_{k}$ is a P-dimensional vector for each segment, while $\langle \mathbf{r} \rangle^{\mathrm{A}}_{j}$ is a $M \times P$ matrix which columns contain the P principal orbits for the monitor set $\mathbb{J}_{\mathrm{A}}$. An analogous decomposition may also be performed for segment B.

The principal orbits can then be used directly as a replacement of the real perturbation orbits at the monitors $j_1 - j_4$ when computing $\tilde{\mathbf{T}}_{1\to3}$, as

$$\begin{pmatrix}\vec{r}_{j_3k}\\ \vec{r}_{j_4k}\end{pmatrix}^\dagger = \begin{pmatrix}\vec{r}_{j_1k}\\ \vec{r}_{j_2k}\end{pmatrix}^\dagger \tilde{\mathbf{T}}^\dagger_{1\to3} \quad \text{for } k \in \mathbb{K}_\mathrm{A}$$
$$\text{implies} \quad \begin{pmatrix}\langle\mathbf{r}\rangle^\mathrm{A}_{j_3}\\ \langle\mathbf{r}\rangle^\mathrm{A}_{j_4}\end{pmatrix}^\dagger = \begin{pmatrix}\langle\mathbf{r}\rangle^\mathrm{A}_{j_1}\\ \langle\mathbf{r}\rangle^\mathrm{A}_{j_2}\end{pmatrix}^\dagger \tilde{\mathbf{T}}^\dagger_{1\to3}. \tag{5.8}$$

Again, the computation for $\tilde{\mathbf{T}}_{3\to1}$ is analogous, as then $j_1, j_2 \leftrightarrow j_3, j_4$ and A $\leftrightarrow$ B are interchanged. In consequence, the solution of the eigenproblem (5.7) benefits from this noise reduction scheme.

Note that, although not necessary, it is possible to further extend the use of PCA cleaning and completely replace the application of the monitor subroutine (sec. A.3.3).

5.2.4. Start values for dispersion coefficients

By using techniques from this section, we can find all BE-d parameters by the following procedure

1. Select two pairs of adjacent monitors j_1–j_4.
2. Use the PCA-enhanced eigenvalue problem (sec. 5.2.3) to solve (5.7) for monitor vectors at this monitor subset, together with modal phase advances.
3. Use CM mapping (chapter 5.1) to get all corrector parameters and all other monitor vectors.

On the other hand, the optimization problem discussed in chapter 4 can use dispersion coefficients as input. Thus we want to obtain starting values for dispersion.

If the dispersion effects on a response matrix are considerably small, approximate values can be obtained by rearranging the BE+d model equation (4.1) to

$$r_{jkw} - \sum_m \Re\left\{R_{jmw}\mathrm{e}^{-\mathrm{i}S_{jk}\mu_m/2}A^*_{km}\right\} = d_{jw}b_k.$$

As all parameters on the left-hand side are known in good approximation, and the right-hand side is an outer product along indices (j, w) with k, one can use outer product decomposition (sec. A.3.2) to obtain estimates of d_{jw} and b_k up to a global scaling factor from measurement data. This also allows to check beforehand if an optimization of dispersion coefficients is viable or superfluous by checking the signal contents in the outer product (sec. A.3.2).

5.2.5. Discrete residual error optimization

With approximate BE+d parameters, we can compute a residual error (sec. 4.2.1)

$$\chi^2 = \sum_{jkw} \chi^2_{jkw} = (r_{jkw} - f_{jkw})^2.$$

This error depends on the monitor subset $j_1, \ldots, j_4$ that was chosen arbitrarily. Therefore, it is reasonable to optimize the chosen monitor subset and select the subset with minimum error. To limit the optimization to a range where the maximum number of correctors can be used, we set $j_2 = j_1 + 1$ and $j_4 = j_3 + 1$, so that only j_1 and j_3 are optimized.

As the number of combinations scales with J^2, we want to limit the iterations and put the combinations with the highest probability of low errors to the beginning. For this, we define $\Delta j = j_2 - j_1$ and set its initial value to $\Delta J = [J/2]$. Then, j_1 runs from 1 to the condition $j_2 = J$.

If the number of permitted steps is large enough, the next iterations start with $\Delta J \in \{J/2 - 1, J/2 + 1, J/2 - 2, J/2 + 2, \ldots\}$. The pattern is shown in Fig. 5.5 for an example value of J.

In all cases presented in this thesis, a small number of only $J/2$ iterations is sufficient to generate proper starting values for the optimization layer.

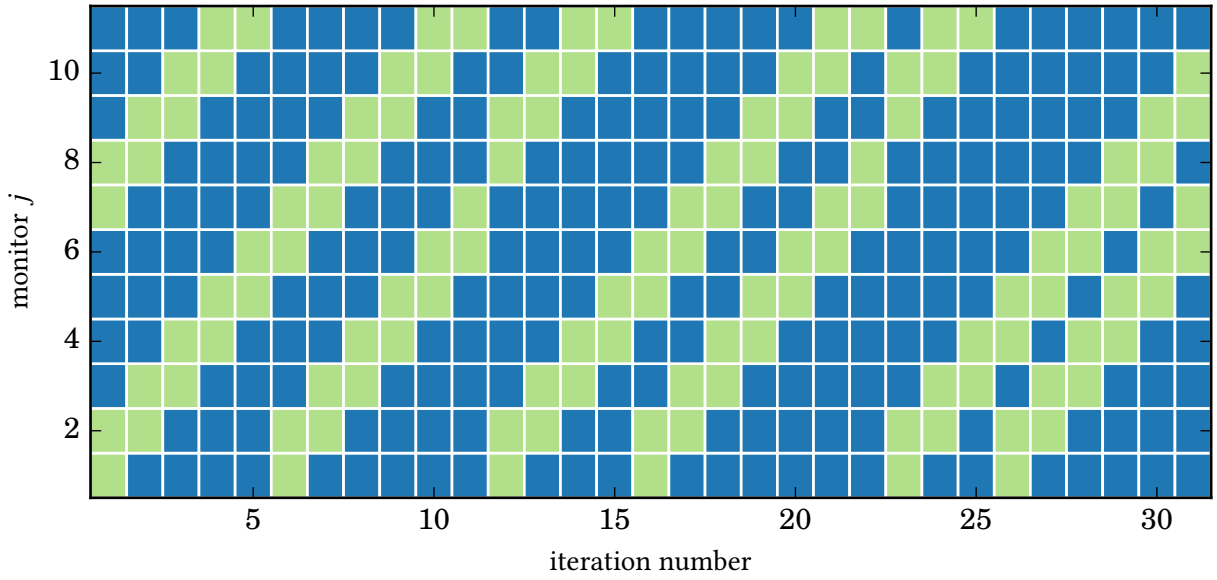

Figure 5.5.: Selected monitor sets (green) for subsequent iteration numbers and the case $J = 11$.

Summary

We found a way to compute approximate BE+d parameters for a given response matrix and topology, neglecting noise and dispersive effects. Nevertheless, these values are useful as starting values for the optimization problem in sec. 4.2. Therefore, we will discuss the implementation of the optimization problem in the next chapter.

6. Closed-Orbit Bilinear-Exponential Analysis (COBEA)

Finding algorithms to compute approximate BE+d parameters (chapter 5) and reformulating the general response problem as a nonlinear regression problem (chapter 4), it is possible to complete the inverse modeling procedure by integrating the aforementioned parts into an algorithm to decompose general response matrices into accurate bilinear-exponential (B+d) parameters (sec. 6.1). The resulting algorithm is termed Closed-Orbit Bilinear-Exponential Analysis (COBEA).

The chapter concludes with a comparison of this algorithm with three techniques (LOCO, AC dipole, Quadrupole Tune Scan) commonly used for beam diagnostics in storage rings (sec. 6.2).

6.1. Architecture of the COBEA algorithm

After finding a sufficient but minimal model for closed-orbit perturbations in storage rings respectively general response matrices (forward modeling, chapters 1 and 3), one may compute a solution to the general response problem by implementing the following steps or 'layers' of the algorithm.

1. *Generate approximate BE+d model parameters for optimization. (Start-value layer)*

   A suitable procedure, the MCS algorithm, has already been discussed in sec. 5.2 and will be used in the following without further inquiry.

2. *Solve the optimization problem (sec. 4.2) using a suitable procedure. (Optimization layer)*

   As gradients can be computed analytically for the given problem, the optimization procedure should be able to use this information efficiently (sec. 6.1.1). For the case of convergence, the inverse problem can be considered as solved.

3. *Compute invariants and errors for the optimal solution. (Postprocessing layer)*

   This allows to normalize the optimal BE+d model parameters using additional monitor drift space information, while also permitting posterior correction of monitor distortion if known (sec. 6.1.2). Mais-Ripken parameters and their errors can be extracted from the resulting BE+d model by elementary means; this step is also integrated into the postprocessing.

6.1.1. Optimization layer

For reasonable noise levels, one may use the solution computed by the MCS algorithm (sec. 5.2) to start an iterative optimization procedure for the general response problem (sec. 4.2). As our optimization function is smooth in the dependent variables, the class of line-search algorithms is appropriate. The discussed algorithm can also use gradient and second-derivative information as inputs, and thus convergence time can benefit from the fact that the gradients and Hessians of the problem under consideration can be computed analytically.

Line-search methods

In the theory of nonlinear optimization, line-search methods are based on multidimensional Taylor approximations around the current optimization vector which are then solved for the minimum argument. We will briefly discuss some possible methods without any claims of completeness. For brevity, step size considerations have been omitted in the following line-search examples.

The most elementary method to use would be gradient descent [26], by stating (see sec. 4.2.1)

$$\vec{\rho}_{n+1} = \vec{\rho}_n - \text{const.} \left[\vec{\nabla}_\rho \chi^2\right]_{\vec{\rho}_n} = \vec{\rho}_n - \text{const.}\ \mathbf{J}(\vec{\rho})\,\vec{\chi}.$$

While this procedure converges to local minima, it completely neglects the second-order effects leading to unnecessarily large number of iterations.

If we also use the Hessian, we can find the next iteration step via Newton's method by solving the equation system

$$\mathbf{H}(\vec{\rho}_n)\delta = \left[\nabla \chi^2\right]_{\vec{\rho}_n} \quad \text{for } \delta = x_{n+1} - x_n. \tag{6.1}$$

Under the assumption that the starting value for the optimization is reasonably close to the optimum solution (sec. 4.2.2), we can approximate $\mathbf{H}$ with $\tilde{\mathbf{H}} \propto \mathbf{J}\mathbf{J}^\dagger$ and could thus also use Newton-Raphson [26] iterations.

Low-memory BFGS (L-BFGS)

While a pursuit of the aforementioned approaches is possible, two complications arise in their implementation.

1. *Singular Hessian.* (6.1) is an underconstrained equation system due to the bilinear nature of the underlying problem and the resulting scaling invariants (sec. 4.1.1). This problem can be circumvented by one of the following options.
   a) Solving the system by using the quasi-inverse of the approximate Hessian, given as the covariance matrix (sec. 6.1.2). The computation of the quasi-inverse has higher time complexity than solving the equation system (6.1) directly.
   b) Adding scaling invariant constraints for normalization to χ^2. This would require an additional discussion about the weighting of this constraint in relation to the other terms in χ^2.
2. *Large number of dependent variables.* For a typical case like the DELTA storage ring, we have $J \approx K \geq 50$, resulting in ≈ 800 dependent variables respectively dimensions of the optimization problem. Hence, set up and solution of equation systems like (6.1) can be a 'bottleneck' of time and/or memory complexity.

To circumvent the aforementioned complications, we will use the low-memory BFGS algorithm (L-BFGS) [35, 36] to optimize the BE+d model. L-BFGS does not use the Hessian matrix directly. Instead, projections of the inverse Hessian computed from function values and gradients are used to update the search direction in each iteration.[1] The number of projections stored is limited and results in low memory complexity. Furthermore, as no direct inversion computation is performed, the rank-deficient Hessian is not problematic for L-BFGS. A flow diagram for the resulting optimization layer is shown in Fig. 6.1.

The problem of a large number of dependent variables is also apparent in the training of deep neural networks, and L-BFGS and its variants are also used in this field [37].

[1] Alternatively, the projections of the Hessian can be given as input to the algorithm.

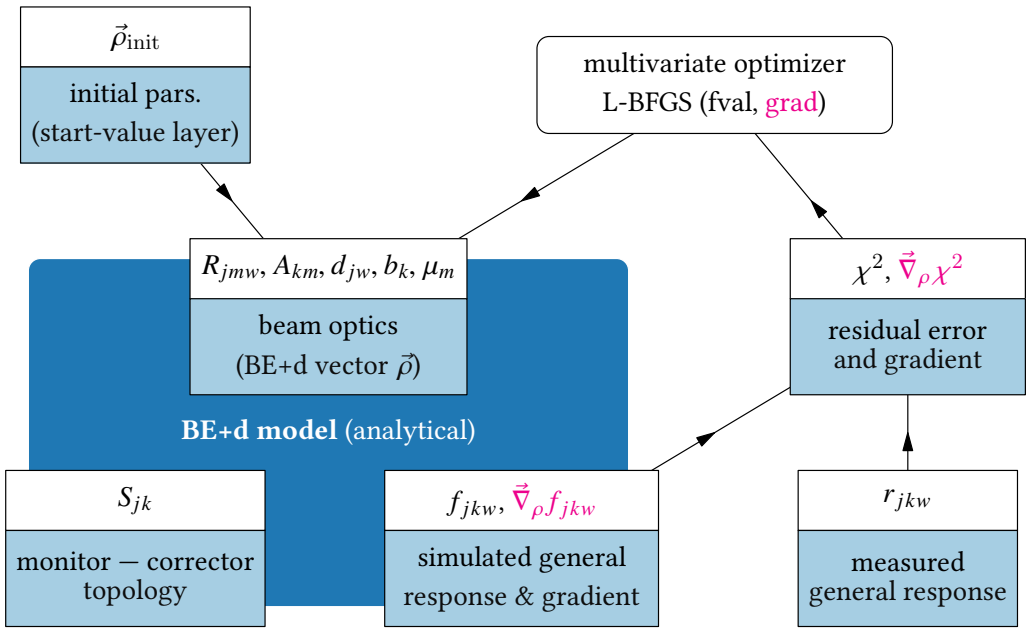

Figure 6.1.: Simplified flow diagram of the optimization layer (identical definitions to Fig. 5.2). The BE+d model is optimized using residual error between measured and simulated response and the respective gradient using L-BFGS.

6.1.2. Postprocessing layer

For the case of convergence, the resulting BE+d parameters hold all optical information regarding the optimal solution.

As many of the underlying principles of postprocessing the data have been discussed in secs. 2.4 and 4.1, we will only briefly analyze them in the order of their appliance. All of the computational steps mentioned in the following are optional and depend on additional assumptions.

Monitor distortion rectification

One may assume that a beam position monitor readout is distorted, so that the reported (transverse) beam position $\vec{r}^{\#}$ and the physical beam position $\vec{r}$ are related by a (nonlinear) operator $\mathcal{D}$ via

$$\vec{r} = \mathcal{D}\vec{r}^{\#}.$$

This relation can be linearized around the unperturbed closed orbit. As we have defined the unperturbed closed orbit as reference orbit in sec. 2.3.1, we can simply write the linearization regarding orbit perturbations as

$$\vec{r} = \mathbf{D}\vec{r}^{\#}.$$

If the map $\mathcal{D}$ and/or the corresponding matrix $\mathbf{D}$ are known, one can

1. run the start-value and optimization layer of COBEA on the distorted beam perturbations $\vec{r}^{\#}_{jk}$,
2. then replace the resulting monitor and dispersion vectors by their rectified equivalents via (sec. 3.2.5)

$$\vec{R}_{jm} = \mathbf{D}\vec{R}^{\#}_{jm}, \quad \vec{d}_j = \mathbf{D}\vec{d}^{\#}_j.$$

This way the rectification of distortion can be carried out or even optimized after a successful run of the optimization layer.

Normalization and tunes

Revisiting the expressions for single-particle oscillation (2.7) and orbit perturbations (sec. 3.1.2), we find both being invariant under each of the following variable transformations (sec. 4.1.1).

1. Switch the sign of μ_m and conjugate R_{jm}, A_{km} for a given mode m,

$$\mu_m \to -\mu_m, \quad \vec{R}_{jm} \to \vec{R}^*_{jm}, \quad A_{km} \to A^*_{km}.$$

2. Scale all monitor vectors and corrector parameters for a given mode m by a complex number C_m using the rule

$$\vec{R}_{jm} \to C_m \vec{R}_{jm}(s) \quad \text{and} \quad A_{km} \to \frac{1}{C^*_m} A_{km} = \frac{C_m}{|C_m|^2} A_{km}.$$

3. Scale the dispersion vectors and coefficients by a real quantity c using the rule

$$\vec{d}_j \to c\vec{d}_j \quad \text{and} \quad b_k \to b_k/c.$$

To resolve the ambiguity for transformation 1, a criterion is needed on how to decide which phase advance $\pm\mu_m$ respectively conjugation is the correct one. For the typical case of $M = 2$, there are 4 regions to choose from, therefore this problem is colloquially known as "quadrant problem".

We return to the definition of invariants of motion, which are proportional to the energy in the oscillation and must be positive (sec. 2.4.1). Notably, the expression(s) used to compute the invariants change sign under transformation 1. In consequence,

> if the computed invariant I_m for a given mode is negative, transformation 1 should be applied.

Thus it is advantageous to choose the correct quadrant directly after computing the invariants, so that μ_m is corrected and $\vec{R}_{jm}$ is corrected and normalized.

To compute the invariant in the first place, the algorithm assumes that a known segment exists between two monitors for which monitor vector data has been computed. Then, given second-order indices p, q so that the monitor index at the beginning respectively end of the known segment are given as j_p, j_q, one can linearize (2.1) in sec. 2.2.1 and obtains[2]

$$\vec{R}'_{j_p m} = \mathbf{T}^{-1}_{r\to r'}(\vec{R}_{j_q m} - \mathbf{T}_{r\to r}\vec{R}_{j_p m}),$$

so that the invariant can be computed by (sec. 2.4.2)

$$I_m = \Im\left\{\vec{R}^\dagger_{j_p m}\vec{R}'_{j_p m}\right\}.$$

After computing the invariants, we can thus first use transformation 1 on any negative invariant. By this transformation, the sign of the invariant also changes. Afterwards, one can use the invariants as scaling factors, applying transformation 2 for each mode m with the scaling factor $C_m = 1/\sqrt{I_m}$. The only transformation which can not be used to remove the respective invariance is transformation 3. To do so, we would need additional information about the energy deviation δ of the closed orbit when different correctors are applied.[3]

A sketch of the resulting algorithm is shown in Fig. 6.2.

[2] Note that, although its components are complex, the eigenorbit $\vec{R}_{jm}$ is treated just like any other orbit.

[3] Such information could be obtained in principle by monitoring the synchronous phase during recording of the general response matrix.

Data:
Monitor vectors $\vec{R}_{jm} \in \mathbb{C}^M$ at J monitors for a given mode m meta-indices p, q for known segment between monitors j_p, j_q transfer matrix blocks $\mathbf{T}_{r\to r}, \mathbf{T}_{r\to r'}$ of known segment
Result:
quadrant-corrected phase advances μ_m for a given mode m normalized, quadrant-corrected monitor vectors $\hat{R}_{jm} \in \mathbb{C}^M$ for mode m optional: invariants of motion I_m for mode m

$\vec{R}'_m \leftarrow \mathbf{T}^{-1}_{r\to r'}(\vec{R}_{j_q m} - \mathbf{T}_{r\to r}\vec{R}_{j_p m})$;
$I_m \leftarrow \Im\left\{(\vec{R}_{j_k m})_\dagger (\vec{R}'_m)_d\right\}$;
if $I_m < 0$ **then**
  $\mu_m \leftarrow 2\pi - \mu_m$;
  $I_m \leftarrow -I_m$;
  for $j \leftarrow 1$ **to** J **do**
    $\vec{R}_{jm} \leftarrow \vec{R}^*_{jm}$;
  end
end
for $j \leftarrow 1$ **to** J **do**
  $\hat{R}_{jm} \leftarrow \vec{R}_{jm}/\sqrt{I_m}$;
end

Figure 6.2.: Simplified algorithm for postprocessing of invariants.

Optical functions From normalized monitor vectors, it is possible to compute Mais-Ripken parameters via the relations derived in sec. 2.4.4. These are given as

$$\phi_{mw} = \arg(\vec{R}_{jmw}) = \arg(\hat{R}_{jmw}) \quad \text{and} \quad \beta_{jmw}(s) = \hat{R}_{jmw}\hat{R}^*_{jmw} = C_m \vec{R}_{jmw}\vec{R}^*_{jmw}.$$

Note that while computing betatron phases is straightforward even if the monitor vectors are not normalized, one can only obtain β values up to a constant factor C_m without the normalization procedure.

This is analogue to the relations for the corrector parameters in the decoupled, dipolar, thin-lens assumption (sec. 3.1.4), where (3.4) results in

$$\hat{X}(\tilde{s}) = \frac{2\sin(\mu_\mathrm{x}/2)}{\theta_\mathrm{x}}\hat{A}_\mathrm{x}$$

$$\text{so that} \quad \phi^k_\mathrm{x} = \arg(\hat{A}_{k\mathrm{x}}) = \arg(A_{k\mathrm{x}}) \quad \text{and} \quad \beta^k_\mathrm{x} = \frac{4\sin^2(\mu_\mathrm{x}/2)}{\theta^2_\mathrm{x}}\hat{A}_{k\mathrm{x}}\hat{A}^*_{k\mathrm{x}}.$$

While the relations for the phase are again invariant under transformation 2, this does not hold for the β values. In addition, the global scaling is modulated by corrector-specific coefficients if the response matrix is not an angle-response matrix and θ_k is unknown. In consequence, the phase information for the corrector parameters is significantly more reliable than the β values.

Error estimation

For a multivariate, scalar-valued problem like (4.2), all linear error estimates regarding variances for an arbitrary variable a can be expressed using the corresponding covariance matrix $\mathbf{C}$ (sec. 4.2.2) via [26]

$$\sigma_a^2 = \vec{A}^\dagger \mathbf{C} \vec{A} = \sum_{pq} A_p C_{pq} A_q \quad \text{with real-valued} \quad \vec{A} = \frac{\partial a}{\partial \vec{\rho}}. \tag{6.2}$$

An estimate for the covariance matrix has been given in sec. 4.2.2 using an SVD of the Jacobian matrix for the solution vector. As (6.2) is a symmetric expression, we can simplify it to

$$\sigma_a^2 = \sigma^2 \; \vec{A}^\dagger \mathbf{U} \mathbf{S}_{\text{inv}} \; \left(\vec{A}^\dagger \mathbf{U} \mathbf{S}_{\text{inv}} \right)^\dagger \tag{6.3}$$

The covariance matrix has indices that correspond to a large set of dependent variables. Therefore, instead of using indices directly, we will use braces, e.g. $\mathbf{C}(\mu_m, \Re D_{km})$, to address components of $\mathbf{C}$ in the following. For computing complicated dependencies on dependent variables via (6.2), it is useful to also adopt this notation for $\vec{A}$ so that

$$\vec{A}(\bullet) = \frac{\partial a}{\partial \bullet}.$$

In the following, we will use the simplifying assumption that the variances of all measured response indices are homogeneous.

Bilinear and modal phase advance errors The above expression becomes very simple if we want to know the errors of the coordinate variables themselves. For any coordinate, the derivative will equal unity for the component representing the coordinate, e.g.

$$\begin{aligned} \vec{A}(d_{jw}) = 1, \text{ all other components zero,} \\ \text{so that} \quad \sigma^2(d_{jw}) = \mathbf{C}(d_{jw}, d_{jw}). \end{aligned}$$

In consequence, the errors for all d_{jw}, b_{kw}, μ_m, just as the errors for real and imaginary parts of R_{jmw} and A_{km}, are all given as diagonal elements of the covariance matrix (replacing d_{jw} with the respective variable).

Mais-Ripken parameter errors When changing to the common polar descriptions of beam optics (sec. 2.4.4), the complexity of expressions slightly increases. However, the Mais-Ripken parameters β, ϕ only depend on the local eigenorbit with identical j, m, w indices. Therefore, (6.2) reduces to

$$\begin{aligned} \sigma_a^2 = \vec{A}(\Re R_{jmw})^2 \mathbf{C}(\Re R_{jmw}, \Re R_{jmw}) + \vec{A}(\Im R_{jmw})^2 \mathbf{C}(\Im R_{jmw}, \Im R_{jmw}) + \\ 2\vec{A}(\Re R_{jmw}) \vec{A}(\Im R_{jmw}) \mathbf{C}(\Re R_{jmw}, \Im R_{jmw}), \end{aligned} \tag{6.4}$$

and the problem of finding two components of the respective $\vec{A}$ remains.

For the phase errors, we drop the discrete distinction criteria for quadrants of complex numbers and just assume $\phi_{jmw} = \arctan(\Im R_{jmw} / \Re R_{jmw})$. Then by using $\partial \arctan y / \partial y = (1 + y^2)^{-1}$ one obtains

$$\begin{pmatrix} \vec{A}^\phi(\Re R_{jmw}) \\ \vec{A}^\phi(\Im R_{jmw}) \end{pmatrix} = \begin{pmatrix} \frac{\partial}{\partial \Re \hat{R}_{jmw}} \\ \frac{\partial}{\partial \Im \hat{R}_{jmw}} \end{pmatrix} \phi_{jmw} = \begin{pmatrix} -\Im R_{jmw} \\ \Re R_{jmw} \end{pmatrix} \frac{1}{\beta_{jmw}}.$$

For the experimental results in this thesis, we will compare phase differences

$$\Delta\phi_{jmw} = \phi_{jmw} - \phi_{(j-1)mw}$$

between two monitor locations s_j, s_{j-1}. Their error is given by

$$\sigma^2(\Delta\phi_{jmw}) = \sigma^2(\phi_{jmw}) + \sigma^2(\phi_{(j-1)mw}).$$

We assume that the optimization result has been normalized by the invariants of motion. As the normalization has not been part of the optimization procedure, we will only compute the errors for Mais-Ripken β parameters *relative to* the normalization. Then

$$\begin{aligned}
\text{from } \beta_{jmw} &= \left(\Re\hat{R}_{jmw}\right)^2 + \left(\Im\hat{R}_{jmw}\right)^2 \text{ one obtains} \\
\vec{A}^{\beta}(\Re R_{jmw}) &= \frac{\partial}{\partial \Re\hat{R}_{jmw}}\beta_{jmw} = 2\Re\hat{R}_{jmw} \text{ and} \\
\vec{A}^{\beta}(\Im R_{jmw}) &= \frac{\partial}{\partial \Im\hat{R}_{jmw}}\beta_{jmw} = 2\Im\hat{R}_{jmw}.
\end{aligned}$$

We conclude that the expressions for Mais-Ripken parameter errors relative to normalization have been calculated.

6.1.3. Aspects of implementation

Following the aforementioned procedures, the author has implemented the COBEA algorithm in Python [38] using the SciPy & NumPy ecosystem [39, 40]. This has the advantage that a many tested subroutines are available for prototyping. Although Python is an interpreted (scripting) language, most computationally expensive operations on NumPy arrays are evaluated by embedded Fortran and C routines.

For the implementation of the residual error and gradient functions, the sparse Jacobian matrix $\mathbf{J}$ is not explicitly constructed. Instead, dense array multiplications using the variables contained in the BE+d vector are used extensively, which leads to a significant acceleration of function evaluations for L-BFGS. In the SciPy `optimize` module, L-BFGS is considered as a special case of the L-BFGS-B algorithm [36].

The Jacobian matrix used for error estimation has been carefully tested to be exactly equivalent to the gradient function evaluations.

Although it is possible for the decoupled response assumption (sec. 4.3) to implement COBEA on two $M = 1$ cases, the prototype implemented for this work uses the equivalent problem

$$r_{jkw} = \Re\left\{R_{jmw}\mathrm{e}^{-\mathrm{i}S_{jk}\mu_m/2}A^*_{km}\right\} + d_{jw}b_{kw},$$

which is almost identical to the coupled problem, the only difference being the substitution $b_k \to b_{kw}$ for two separate dispersion orbits. The gradients are evaluated accordingly, and the resulting optimization procedure is used throughout chapter 9.

The optimization procedure for the coupled case is implemented as described in sec. 4.2.

6.2. Conceptual comparison to existing methods

Observing the particle motion in accelerators is a non-trivial task, and many methods for diagnostics exist to obtain normalized monitor vectors at key positions in a storage ring [9, 16, 17]. Therefore the chapter is concluded by a comparison of COBEA with existing methods and algorithms to solve similar problems. Besides the LOCO approach (sec. 6.2.1), we also discuss two model-independent techniques in secs. 6.2.2 and 6.2.3.

6.2.1. Linear Optics from Closed Orbits (LOCO)

In chapter 1, the LOCO method [41] was briefly introduced as a fitting routine that matches response matrices from an accelerator simulation model to measured ones from a real accelerator. We will now further examine this commonly used and successful method.

One may subsume all magnetic properties that are implicitly (currents) or explicitly (multipole moments around the beam path) varied into a magnetic lattice vector $\vec{\rho}^{\,\mathrm{mag}}$. One can then express the core optimization of LOCO in a form similar to the one in sec. 4.2. For a given accelerator model AM,

$$\text{find } \vec{\rho}^{\,\mathrm{mag}} \quad \text{so that } \chi^2 = \sum_{jkw} \left[g_{jkw}\left(\vec{\rho}^{\,\mathrm{mag}}, \mathrm{AM}\right) - r_{jkw} \right]^2 \text{ is minimal.}$$

The LOCO approach thus essentially consists of the following steps (Fig. 6.3).

1. Construct a detailed accelerator model AM with almost correct values for all magnetic fields in the beam path.
2. Vary magnetic properties $\vec{\rho}^{\,\mathrm{mag}}$ according to an optimization rule.
3. For each set of properties, iteratively excite all correctors and record the orbit responses $g_{jkw}(\vec{\rho}^{\,\mathrm{mag}}, \mathrm{AM})$ in the simulation.
4. Fit g_{jkw} to the measured response r_{jkw} by iterating steps 2 and 3.

If sufficient measurement data is available, LOCO also allows to optimize a variety of effects, e.g. monitor misalignments. Furthermore, LOCO also allows weighting of different errors and uses other extensions known for this type of optimization problems.

In essence, LOCO is a nonlinear optimization procedure for closed-orbit perturbations in ring accelerators and is thus related to the COBEA's optimization layer. The significant difference between them is that the model to be optimized is a complete accelerator with all magnetic fields along the path for LOCO, and the BE+d model for COBEA.

1. For LOCO to give sufficient results respectively present convergence in local minima of χ^2, AM must be a reasonably correct model of the real-world storage ring. This includes all magnetic fields in the beam path, or, if given implicitly, the magnetic properties of all elements (e.g. temperature effects, hysteresis effects, offsets of magnets in the range 10^{-4} m).

   While this is actually possible for many modern storage rings, it poses an obstacle for accelerators in which the magnetic properties are not known everywhere along the beam path to a sufficient degree. If a single element differs significantly, the convergence of LOCO is doubtful.
2. Furthermore, the start parameters $\vec{\rho}^{\,\mathrm{mag}}_{\mathrm{init}}$ must lie in vicinity to their real storage ring values.
3. Depending on the accuracy of the model, a subset of AM properties to be varied via $\vec{\rho}^{\,\mathrm{mag}}$ must be selected. The selection respectively simplified accelerator modeling is done by the user of the algorithm. This is necessary to ensure that the optimization procedure used for LOCO is not underdetermined. In consequence, only effects that have been considered in choosing magnet properties (as variables of the accelerator simulation AM) can be considered in the LOCO result.

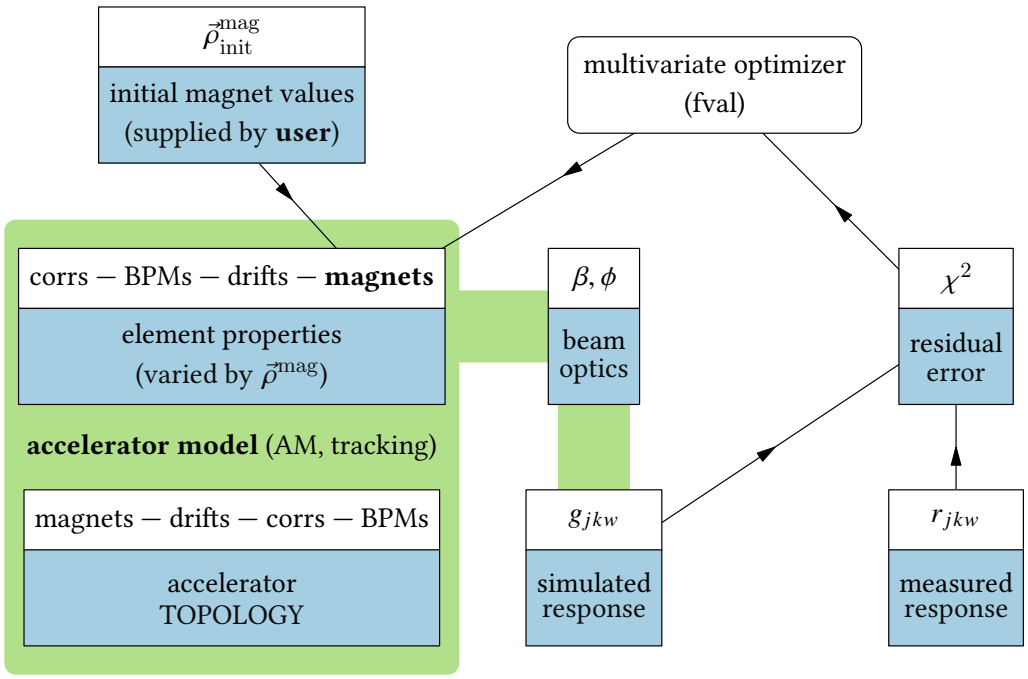

Figure 6.3.: Simplified flow diagram for LOCO (identical definitions to Fig. 5.2). The 'beam optics' step is usually not explicitly computed during optimization. In comparison of LOCO with COBEA (Fig. 6.1), the gradient information is not available in the former, while the complete model but the monitor–corrector topology has been removed in the latter ("Occam's razor" [42]).

4. LOCO relies on an external tracking code to create the response matrices for each iteration step. These tracking codes compute particle trajectories by numerical means (e.g. Runge-Kutta 4 or other tracking algorithms which are symplectic up to a given order), as the dependence of g_{jkw} on $\vec{\rho}^{\,\mathrm{mag}}$ is not analytically known. For this reason, the gradient $\vec{\nabla}_{\rho} g_{jkw}$ of the optimization problem can also not be computed by analytical means.

COBEA has the following properties that allow convergence with a significantly reduced amount of input data.

1. The information of a full magnetic lattice model may be split (Fig. 6.3) into (magnetic and geometric) properties of elements, and the ordering of elements along the beam path (topology). While the detailed properties of elements can often only be obtained by detailed observation or simulation, the topology is a static property that can be obtained without additional instruments (e.g. by visual inspection of a storage ring), as the considered elements are macroscopic.

   Of these properties, COBEA only uses a part of the topology information, namely the relative order of monitors and correctors along the beam path (contained in $\mathbf{S}$, sec. 3.2.2).

2. COBEA does not require any start values given as input. The start values for its optimization layer are generated by COBEA's start-value layer using the input response matrix (secs. 5.2, 6.1).

3. For $K \geq 4M$ and $J \geq 4$, the BE+d model used by COBEA is overdetermined by the response matrix in almost all cases of common storage rings (see Fig. 4.2).

4. As COBEA uses a generic optical model in which the dependence of response matrices to dependent variables is known analytically, its optimizer can utilize gradient information. This property allows COBEA to solve the underlying large-scale optimization problem in a reasonable time scale.

6.2.2. AC dipole (Driven Turn-by-Turn Oscillation)

The AC dipole technique [43] is closely related to TbT analysis (see sec. 8.2) with the main difference that a coherent excitation near the eigenfrequencies is performed by a magnetic device called AC dipole, and not a radio-frequency device. Due to the sinusoidal excitation and the quasi-static low-frequency setup, higher harmonics called resonance driving terms which correspond to nonlinearities in beam dynamics can be observed.

The AC dipole technique can only work on larger storage rings. As there is typically only one AC dipole installed in a storage ring, the information obtained by a static perturbation ($K = 1$, sec. 4.1) is incomplete, and perturbations that are not periodic on the accelerator length need to be excited. Therefore, the AC dipole excitation frequency should be close to the transverse eigenfrequencies (betatron tunes) to obtain accurate information about resonance driving terms [44].

These eigenfrequencies scale inversely with the accelerator circumference. This limits the applicability of AC dipole techniques with regard to storage rings with relatively small circumference. As an example, AC dipole techniques were applied successfully to the (now decommissioned) Fermilab Tevatron, which circumference was $\geq$ 6.4 km [45], and to the CERN SPS with a circumference of $\approx$ 6.9 km [46]. The corresponding maximum of observable frequencies in TbT data [47] are both $\leq$ 25 kHz.

6.2.3. Quadrupole Tune Scan

As dependence of analysis on the accuracy of a specific model is undesirable, some techniques have also been developed to facilitate analysis while reducing reliance on the accurate simulation of the magnetic lattice. A widely used technique of this class (due to its robustness) is the Quadrupole Tune Scan [6, 16, 17], in which the change of modal phase advances μ_m due to a known thin-quadrupole perturbation $\tilde{k}$ is considered, and the Courant-Snyder β functions at the position of the perturbation are derived by the equation

$$\beta_{\mathrm{x}} = 2\frac{\tilde{\mu}_{\mathrm{x}} - \mu_{\mathrm{x}}}{\tilde{k} - k}.$$

The unperturbed and perturbed phase advances μ, $\tilde{\mu}$ can be obtained from any monitor that is TbT capable by methods described in sec. 8.2.

While the technique is applicable to many accelerators, one may only obtain averaged β values over the effective length of the perturbed quadrupole magnets. Furthermore, assumptions on quadrupolar fields and their modeling still apply, so the magnitude of the perturbation must be exactly known. In addition, many storage rings do not possess separate power supplies for each quadrupole, but one for each quadrupole class. Then this technique cannot be applied.

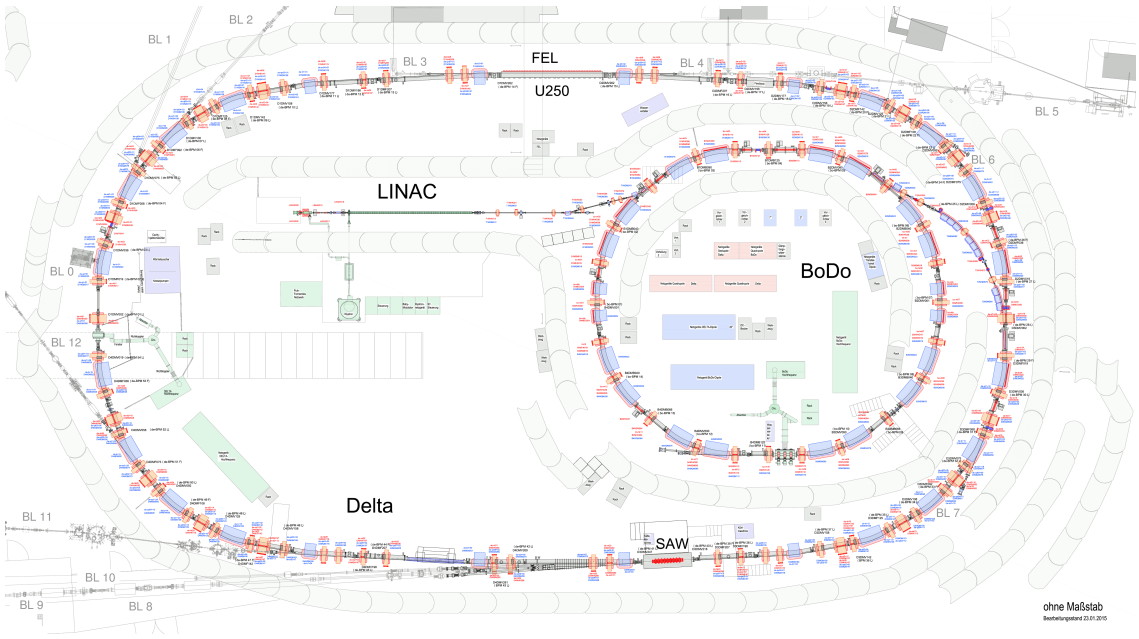

Figure 7.1.: Sketch of the DELTA synchrotron radiation facility [48], which 3 components (LINAC, BoDo, Delta) accelerate electrons up to ≈ 1.5 GeV.

7. Data mining at DELTA using COBEA

As COBEA has been introduced in the last chapter, we continue by showing that COBEA is applicable to real-world storage rings. In the following two chapters, we will evaluate the application of COBEA to the storage ring of the DELTA synchrotron radiation facility (sec. 7.1).

The standard response matrix files from the DELTA control system (sec. 7.2) also include the betatron tune measured by direct excitation in standard operation. This additional information is used to validate COBEAs prediction of modal phase advances respectively betatron tunes from response matrices against measurement data from single-particle oscillations.

7.1. DELTA overview

Parameter	Value	
Lattice structure	Quadrupole triplet (in arcs)	[49, 50]
Reference energy	1.49 GeV	[50]
Circumference	115.1638 m	[51]
RF frequency	499.819 MHz	[50]
harmonic number h	192	[50]
electron beam current	max. 130 mA	[50]

Table 7.1.: DELTA storage ring parameters

DELTA is a facility run by TU Dortmund university for the purposes of providing synchrotron radiation to scientists in the NRW region and beyond (≈ 2/3 of operation time), as well as for accelerator research (≈ 1/3 of operation time). It can be classified as a 3rd generation synchrotron light source [52] (Fig. 7.1), consisting of a 75 MeV linac [53], a synchrotron of 50.4 m circumference that allows full-energy injection [52] and a storage ring which characteristics are shown in Tab. 7.1 [49, 50, 51]. Further information about DELTA, its design and applications can be found in [49, 52, 50].

7.1.1. Monitor hardware and pincushion distortion

A beam position monitor at DELTA's storage ring is based on four pickup electrodes (designed for the ESRF), installed in but electrically isolated from the vacuum chamber, which are each connected to the inner conductor of four coaxial cables [54]. Following A.1.4, knowledge of the signal functions $G_n(x, y)$ for the beam position monitors is needed to accurately measure the transverse beam centroid coordinates at s_j. We introduce electrical coordinates [54]

$$S_\mathrm{x} = \frac{a_2 + a_4 - a_1 - a_3}{a_1 + a_2 + a_3 + a_4}, \quad S_\mathrm{y} = \frac{a_1 + a_2 - a_3 - a_4}{a_1 + a_2 + a_3 + a_4}$$

that only depend on the distribution of voltages (or other, linear measurement quantities, like port variables) on the four pickup electrodes, and thus only on the Green's functions and not on the total charge of the beam distribution.

In standard operation, the beam position is determined using a linear approximation around the electrical center of the monitor, given by

$$x_\mathrm{lin} = C_\mathrm{x} S_\mathrm{x}, \quad y_\mathrm{lin} = C_\mathrm{y} S_\mathrm{y} \quad \text{with} \quad C_\mathrm{x} = 15.9\,\mathrm{mm}, \quad C_\mathrm{y} = 19.6\,\mathrm{mm},$$

where the values $C_{\mathrm{x,y}}$ stem from a simulation in [54]. While this linearization seemingly is sufficient for daily orbit correction tasks, it limits the resolution and applicability of beam diagnostics, as DELTA's standard orbit deviates up to 10mm from the geometrical monitor center at some monitors. For this large deviation, the linearization error is significant (Fig. 7.2).

While the correction of this complications is possible for future measurements, a full compensation of the linearization error via rectification Jacobians $\mathbf{D}_j(\underline{\vec{r}_j}^{\#})$ (sec. 3.2.5) is complicated as they are unknown due to the following facts.

1. The response matrix files recorded by the DELTA control system do not include the reference orbit. This information can possibly be reconstructed by guessing appropriate setup data files which are also present in the control system.

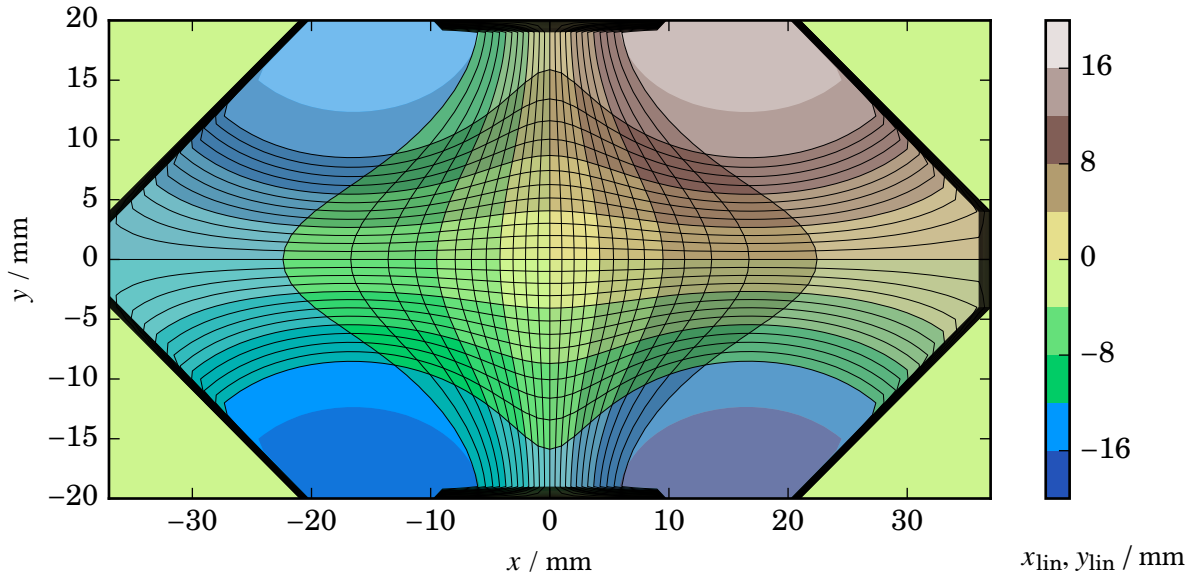

Figure 7.2.: Comparison of computed positions in linear approximation x_lin, y_lin with real positions x, y using data from [54]. The grid lines subdivide distorted squares of 1 mm size in linearized positions.

2. The monitor readout $\vec{r}_j^{\,\mathrm{read}} = \vec{r}_j^{\,\mathrm{lin}} + \Delta\vec{r}_j$ from the EPICS control system is described by a file (`BPM.subst`) with deviation vectors $\Delta\vec{r}_j$. This file is routinely modified according to beam-based calibration measurements. While backups of this file have been made at different times, a one-to-one correspondence for a given date and time cannot be guaranteed.

Note that we use the measured response matrices only in conjunction with *measurement* data from the same system, at the same unperturbed closed orbit. Therefore the measurements to be compared have the same (unknown) rectified Jacobians for any monitor, and

$$\begin{aligned} \mathbf{D}_j(\underline{\vec{r}_j}^{\#})\vec{R}^{\#}_{jm}|_{\mathrm{COBEA}} &= \mathbf{D}_j(\underline{\vec{r}_j}^{\#})\vec{R}^{\#}_{jm}|_{\text{validation method}} \\ \text{implies} \quad \vec{R}^{\#}_{jm}|_{\mathrm{COBEA}} &= \vec{R}^{\#}_{jm}|_{\text{validation method}}. \end{aligned}$$

The equivalence does not hold for any squared errors between both methods as $\det \mathbf{D}_j \neq 1$ is possible. This is not problematic, as one is anyway only interested in measurement errors ($\#$) for fitting and comparison procedures.

Monitors with differing electrode-chamber arrangements

The monitors denoted by `BPM43` and `BPM44` are installed in a special undulator section in which the beam chamber is flattened in vertical direction. The linearization coefficients for these monitors have been obtained by measurement [55] as

$$C_{\mathrm{x}}^{\mathrm{SAW}} = 10.4\,\mathrm{mm}, \quad C_{\mathrm{y}}^{\mathrm{SAW}} = 23.9\,\mathrm{mm}$$

and are assumed to be sufficient if the beam is located near the geometrical center of the monitors. The control system of DELTA's storage ring is set up to assume the standard coefficients everywhere. Thus the positions can be corrected by multiplying with $C_{\mathrm{x}}^{\mathrm{SAW}}/C_{\mathrm{x}}$ and likewise for y.

7.1.2. Monitor readout electronics

To compute positions, signals a_n originating from the monitor electrodes are processed into intermediate S_x, S_y values. For the most monitors at DELTA, this computation is accomplished by analog electronics [56].

To facilitate noise reduction for the static voltage signal, the combination of analog electronics and 8-bit CAN-bus [57] readout system uses a low-pass filter with a characteristic frequency of ≈1.5 Hz. The resulting analog voltages proportional to S_x, S_y are then converted to digital values, with the sampling rate of the voltages being limited to 10 Hz by DELTA's control system. The low-pass filter of the analog electronics thus is the limiting factor ("bottleneck") in fast acquisition of orbit data.

TbT-capable readout electronics on a monitor subset

For 10 monitors at DELTA's storage ring,

BPM 13, 14, 15, 16, 38, 39, 40, 43, 44 and 45,

the position signals are processed by FPGA boards, controlled, read-out and connected to the control system by a Linux server [58]. The fast processing and buffering of signal voltages allows to acquire signal frequencies in the MHz range, and thus TbT data of the beam centroid.

The electronics for BPM 14 are installed parallel to the standard analog electronics using power dividers, for the purpose of obtaining and correcting the betatron tunes in standard operation [55]. After pinging, $N = 2048$ turns can be recorded into the buffer.

The electronics for BPM 38, 39 and 40 were installed for beam diagnostics in dedicated accelerator shifts. These are again only used for TbT analysis and installed parallel to analog electronics, and their low-frequency signals are not used in standard operation.

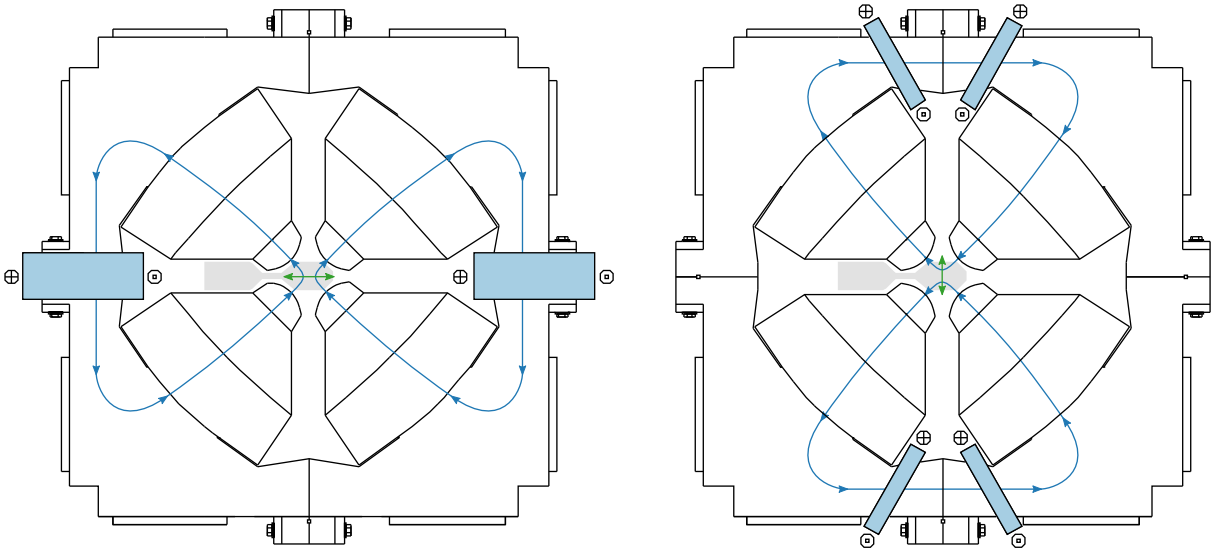

Figure 7.3.: DELTA standard correctors (drawing based on [60]) are installed as additional coils on quadrupole magnets . For this reason, nonlinear magnetic fields are produced.

7.2. Tune validation experiments with standard correctors

At the DELTA storage ring, the standard correctors are not implemented as separate-function magnets, but as additional coil windings installed on the iron yokes of quadrupole magnets. This setup originates from DELTA's original purpose as a test facility without sextupole magnets [49, 59] and is unusual for modern storage rings, as the additional corrector coil current does not only introduce additional dipolar fields, but magnetic multipoles of higher order depending on the specific yoke shape.

Nevertheless, the correctors are sufficient for COBEA analysis as no assumptions about multipolar distortion components of the perturbations have been stated in the derivation of the BE±d model (sec. 3.1.1). Naturally the decoupled, dipolar, thin-lens assumption (sec. 3.1.4) is challenged and one cannot expect the corrector parameters to match with optical functions.

The recording of response matrices is initiated by the accelerator operator. Such recordings are usually done during DELTA's accelerator optimization shifts, or directly before or after user operation. The recording of a standard response matrix takes ≈ 800 s. For the scope of this thesis, 171 response matrices were parsed through COBEA with the following available inputs.

7.2.1. Standard DELTA response matrices as input data

From each closed orbit perturbation measurement recorded by the orbit correction application of DELTA in the past, the following data was available to the author.

1. A coupled angle-response matrix $\mathbf{r}$ using the standard correctors that are used for online orbit correction during operation of the storage ring.
2. a list of identifiers for each corrector respectively column k of the response matrix,
   `VK01, VK02, ..., HK01, HK02, ...` (max. 56 elements)

   For some measurements, single correctors may be missing.
3. a list of identifiers for each monitor respectively row j of the matrix,
   `BPM01, BPM02, ..., BPM54` (54 elements)
4. a list of all monitor and corrector identifiers with the respective elements ordered by increasing s

position along the beam path (“downstream”)

```
HK01, BPM01, VK01, BPM02, BPM03, VK02,  HK02,  VK03,  BPM04, ...
```

with 110 elements that is printed in sec. A.4.2. The list has been created by the author on basis of Fig. 7.1 and has been validated against the `elegant` model [61] of the DELTA storage ring [51, 62]. As is usually the case for storage rings, the topology of DELTA regarding monitors and quadrupole correctors is completely known.

5. The information that
   a) a monitor drift space of length 5.2175 m between BPM14 and BPM15
   b) a monitor drift space of length 0.8678 m between BPM38 and BPM39

   exist [51, 62, 61]. The distances are consistent with the author’s own measurement using a ruler, with a corresponding relative error of 10^{-2}.

While the input described by items 1–3 may change from measurement to measurement, e.g. if some correctors or monitors are nonfunctional, the items 4 and 5 are characteristic for the basic storage ring setup and were compiled by the author. The topology $\mathbf{S}$ is constructed from the three lists included in items 2–4, in analogy to the approach used in sec. 9.1.1.

Validation data

For validation purposes, we use additional information saved into the response matrix files, which is generated by the Q-Pulser application [63] as part of the DELTA control system (Fig. 7.4). The input and validation data for each measurement are extracted from a text file with a strict format, as defined by DELTA’s orbit correction application [60].

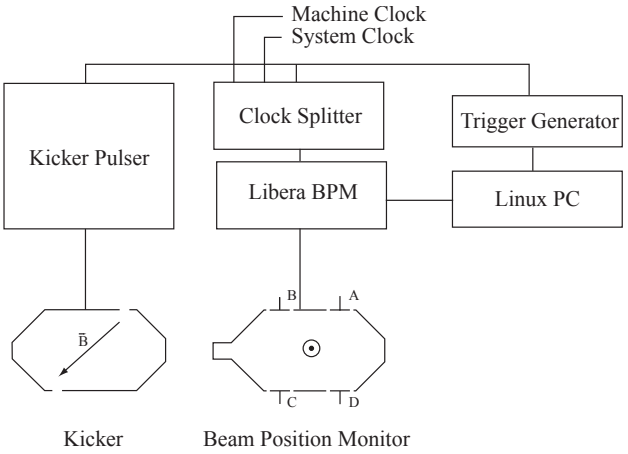

Figure 7.4.: Setup for the Q-Pulser measurement system installed in the DELTA storage ring, taken from [63].

7.2.2. Running COBEA on 171 response matrices

In the time interval between 2006-03 and 2016-01, 171 response matrices were recorded and saved by DELTA’s orbit correction application, initiated by the respective operators. Each of these response matrices (see Tab. 7.2) was analyzed using COBEA, with each run of the optimization layer taking ≈ 15 s on a typical PC.

The 171 analyzed response matrices were sorted by a chronological index (Response ID) starting from RID 1. The indices are related to response matrix file names in sec. A.4.3.

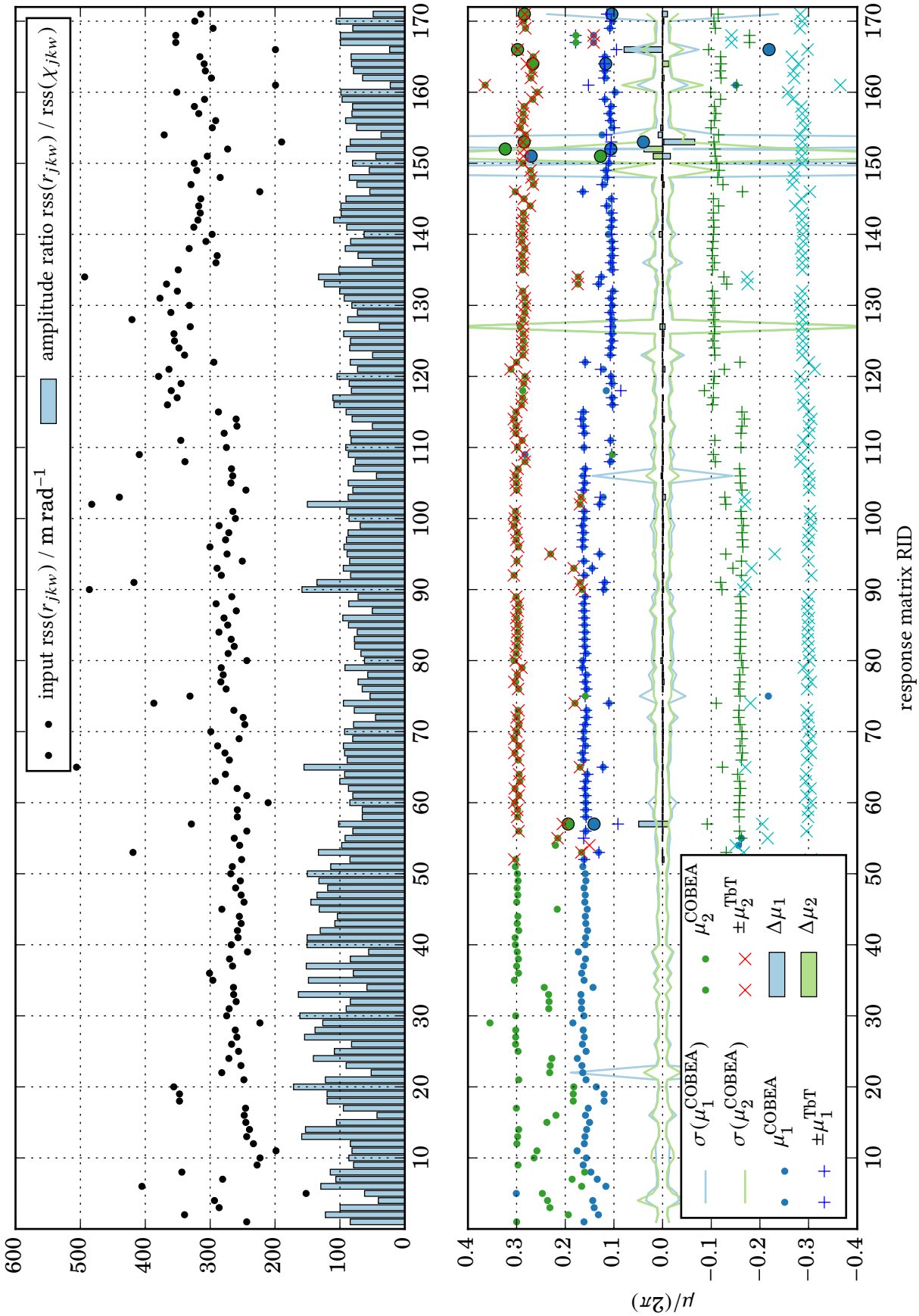

Figure 7.5.: 10-year tune comparison of COBEA and simultaneously recorded TbT tunes using 171 response matrices as input. The labels use the definition $\mathrm{rss}(\bullet) = \sqrt{\sum_{jkw} \bullet^2}$.

Overview of results

In Fig. 7.5, the computed COBEA tunes for each response matrix are shown in comparison to the recorded Q-Pulser tunes. The first 51 matrices cannot be compared with TbT data, as at that time, TbT hardware was either not installed or the tune was not recorded. One can observe that the average input is larger when the modal phase advances μ are lower. This is in consistence with the expected scaling behavior of closed-orbit perturbations with $\propto 1/\sin(\mu_m/2)$ (sec. 3). Also the amplitudes signal-to-noise level for successful measurements is usually larger than 50.

Comparing the results for which TbT tunes are available (Fig. 7.5), we see that from the remaining 120 response matrices,

1. there is a very good agreement between COBEA and TbT measurement for 114 response matrices ($\Delta\mu_m < \pi/50$ for available m from Q-Pulser).
2. Seven response matrices are found for which recorded TbT tunes and computed COBEA tunes are significantly different ($\Delta\mu_m \geq \pi/50$). These matrices are discussed in sec. 7.2.3.
3. Three of the valid response matrices are evaluated by COBEA with (partially) negative modal phase advances. These results are discussed in sec. 7.2.4.
4. A small number of responses with large σ estimations for the μ_m errors exist. These are discussed in sec. 7.2.5.

Quantity	Variable	Value
number of monitors	J	54
max. number of correctors	K	56
max. number of inputs	size($\mathbf{r}$) = JKM	6048
Search space dimensions	D	822

Table 7.2.: Overview of fixed variable values common to RID datasets.

7.2.3. Explanation of deviations in seven response matrices

From inspection of the deviations shown in Fig. 7.5, six responses exist for which the disagreement between Q-Pulser and COBEA tunes $|\Delta\mu_m| \geq \pi/50$. Ordering the response matrices along decreasing deviations, they are indexed by the numbers 166, 153, 57, 152, 151, 164 and 171. In the following, these cases will be evaluated using additional information from DELTA's electronic logbook [64].

RID 166 **response.150317-1_550MeV_below**

This response was recorded in 550 MeV mode. For the response under consideration, there exists a tune inconsistency for mode 1 between Q-Pulser measurement and COBEA estimate. At the time of recording, the stored beam current was significantly lower than average (see Fig. 7.6).

- 550 MeV G. Schmidt, 11:38

  With the current setup, 1.3 mA beam current could be stored. As no horizontal orbit correction was possible, a new response (*RID 166*) has been measured.

- 550 MeV below G. Schmidt, 12:18

  The tune is still located below 9 in the horizontal plane. …

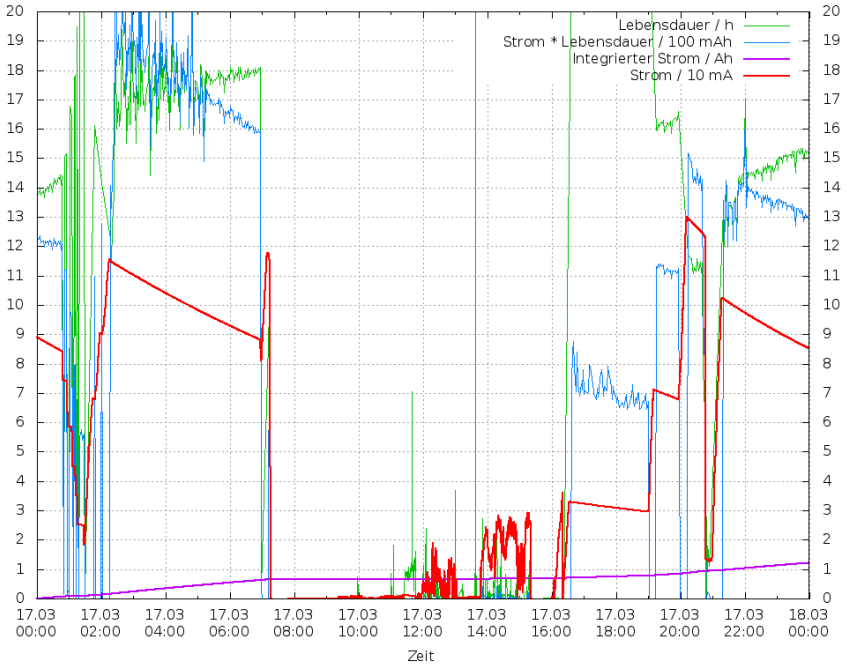

Figure 7.6.: Automatically created current-time plot from DELTA's control system for 2015-03-17.

The signal-to-noise ratio from COBEA is below 40. Taking into account that the response was recorded with less than 1.3 mA beam current, this low value can probably be explained by noisy measurements.

RID 151–154 **response.140924-***

These responses were all recorded on 2014-09-24. The log entries of that day document attempts to repair a beam steering coil (DC1) in the injection region of the storage ring:

- trying to repair DC1 G. Schmidt, 19:21

  …At that location, one could see that the isolation of the supply cable bursted at a constriction. After disassembly, one could measure current flow between magnet and this cable. …

The response matrices recorded were thus compromised by short-circuits of the mentioned beam steering coil (with the possible exception of the last matrix 154, in which no deviations could are observed). This is also underpinned by other observations on orbit problems:

- Orbit problems J. Friedl, 19:56

  The orbit correction application (*running on RID 153*) cannot compensate the occurring deviations.

We can thus assume that the responses indexed with 151 and 153, for which deviations between Q-Pulser and COBEA tunes occur, have not been recorded properly by the control system.

RID 57 **response.081127-2-550MeV**

This matrix has been recorded in the experimental 550 MeV mode using DELTA's previous optical resonator setup [65]. For this day, some problems have also occured with the Q-Pulser system [64].

- Problems with tune measurement J. Friedl, 08:43

  The TbT tune measurement is not working and cannot be restarted.

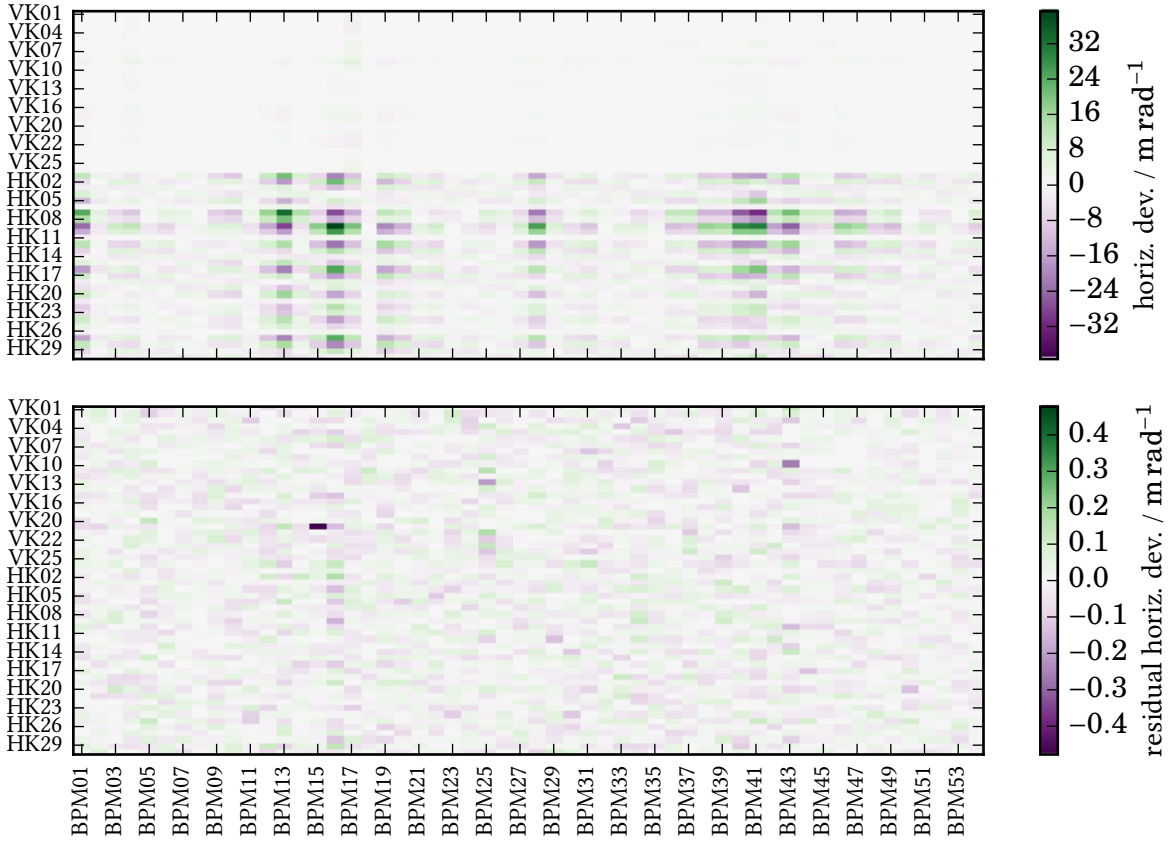

Figure 7.7.: Horizontal orbit deviations (top) and the fit residuals to COBEA's prediction (bottom) for RID 57.

- Tune measurement J. Friedl, 08:51

  The Q-Pulser application can be restarted, but the old window can still not be closed.

The matrix was recorded in a situation where large re-scaling of quadrupole currents had happened.

- 550 MeV tune quadrant H. Huck, 16:50

  I could position the tune into the right quadrant *(RID 57)* by scaling all QF about +2.9% and all QD about +0.5%.

It is therefore possible for this measurement that

1. the tune measurement was compromised by large uncompensated chromaticity,
2. the tune recording was compromised by software errors.

Taking into account the good accuracy of the response matrix fit with its signal-to-noise value > 100 (Fig. 7.5), we can safely conclude that COBEAs tune estimate is correct within its own error margins. A plot of all relevant monitor quantities is shown in Figs. 7.8, 7.9.

RID 164 **response.150205_mitSAW-1**

On this day, new responses for single-bunch mode have been recorded [64].

- new response matrix, single-bunch beam limit P. Ungelenk, 18:21

  Yesterday, the BPMs 54,3,4 and 5 have been calibrated and a new response *(RID 163)* has been recorded. This happened with deactivated SAW.

  Today the SAW has been switched on, which resulted in a new single-bunch current limit of about 11 mA. Therefore a new (and visibly different) response *(RID 165)* has been recorded.

One can state that a new response (RID 165) has been recorded directly after the response in question; for this response, Q-Pulser and COBEA tunes agree very well. It is noteworthy that the tunes predicted for response 165 are in good agreement with the COBEA tune, but not for the Q-Pulser tune, of RID 164.

One may conclude that, while the possibility that COBEAs tune prediction for mode 2 is false exists, the high signal-to-noise value for the vertical deviations (Fig. 7.10) and the coincidence with the RID 165 tune are good indicators that COBEAs tune predictions for this response are correct within the predicted error margins.

RID 171 **response.151103-1**

The deviations between measured tunes for TbT and COBEA are marginal. Nevertheless, problems with the orbit correction have been reported [64] along several other defects for that day.

- Orbit correction works again P. Hartmann, 09:34

  with the right correction file (*other RID*).

- Error management P. Hartmann, 10:35

  - temperature readout works again
  …
  - T2-SSV defective contact → next shutdown
  - automatic plotting into the Elog → in progress
  - orbit correction works with the right setup (*other RID*)

The recorded orbit correction has seemingly never been used. In addition, the amplitude ratio of the

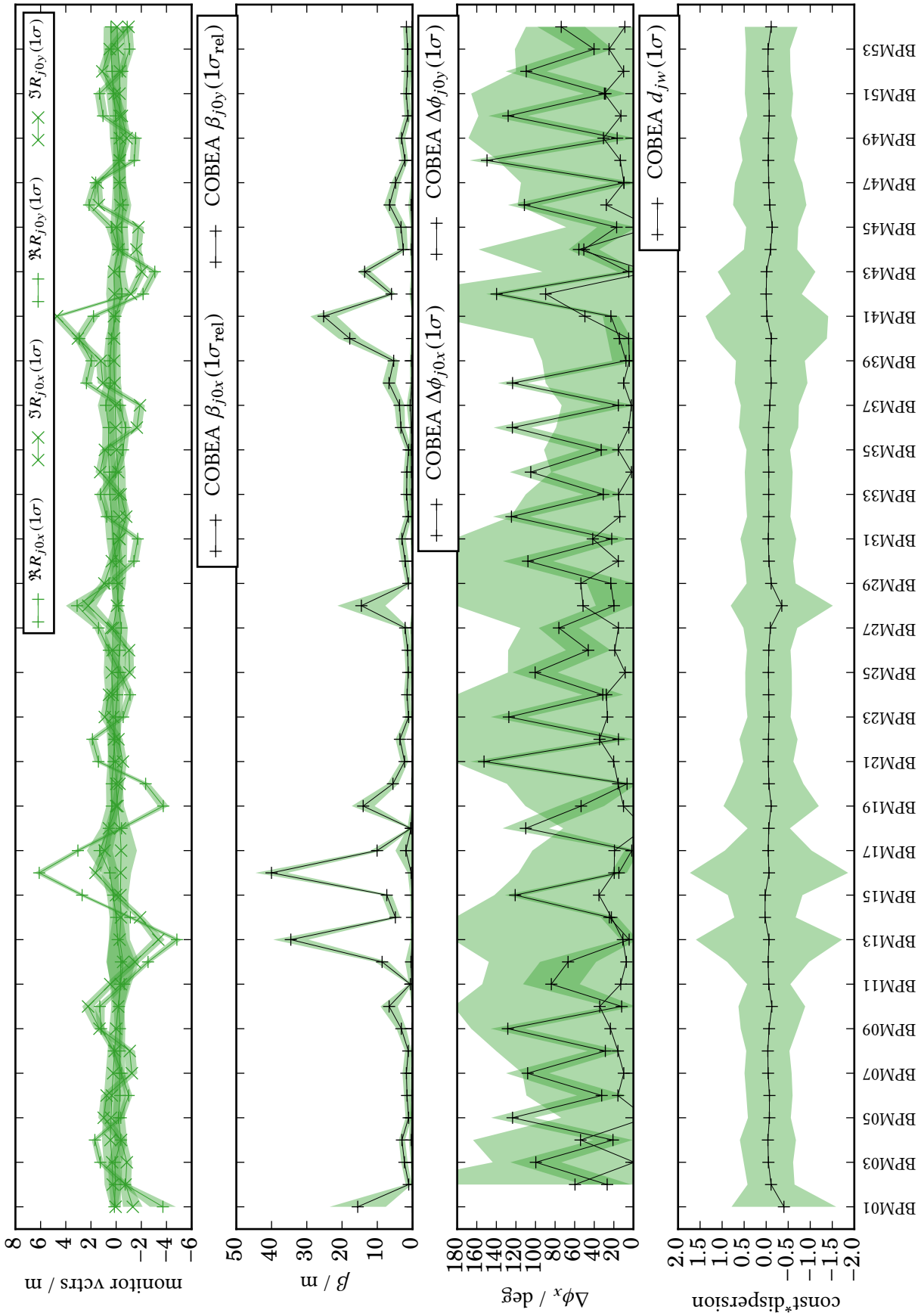

Figure 7.8.: Optical functions of mode 1 predicted by COBEA for RID 57 at monitor positions. The COBEA-predicted tune for this mode is $9.140\,90 \pm 1.17 \times 10^{-2}$, with the TbT tune being 0.091 80.

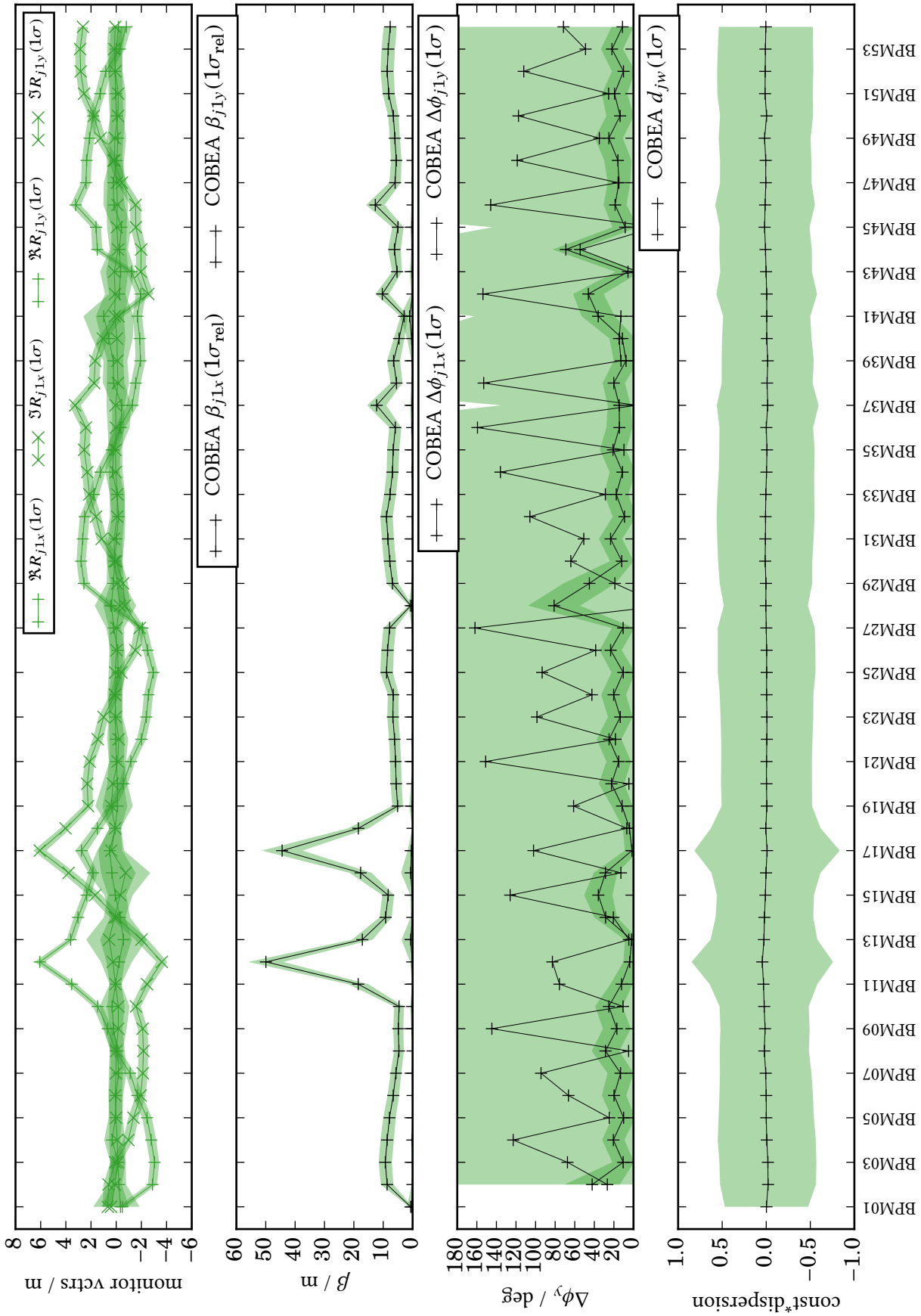

Figure 7.9.: Optical functions of mode 2 predicted by COBEA for RID 57 at monitor positions. The COBEA-predicted tune for this mode is $3.194\,13 \pm 1.05 \times 10^{-2}$, with the TbT tune being 0.205 08.

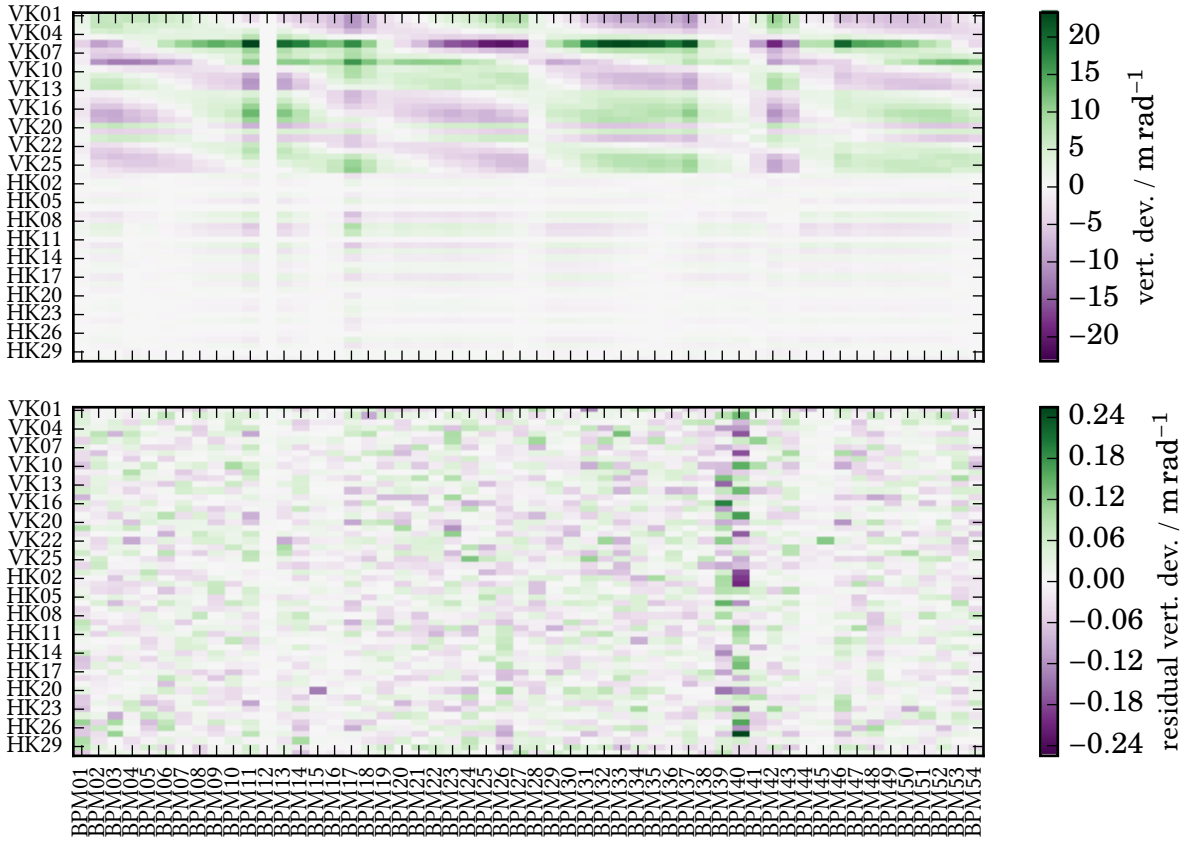

Figure 7.10.: Vertical orbit deviations (top) and the fit residuals to COBEA's prediction (bottom) for RID 164. The COBEA-predicted tune is $9.116\,86 \pm 1.52 \times 10^{-2}$ for mode 1 and $3.266\,71 \pm 1.52 \times 10^{-2}$ for mode 2, with the horizontal respectively vertical TbT tunes 0.120 12 and 0.279 30.

COBEA fit is low. One can conclude that the measurement is either noisy or influenced by malfunctioning devices. Nevertheless, the predicted σ error for both tunes is significantly larger than the existing measurement deviation for this RID.

7.2.4. Results with negative modal phase advances (tunes)

As the Q-Pulser system is only able to record $|\mu_m|$ and the sign information is lost, the actual machine tunes can be either positive or negative. In DELTA's typical operation modes, both phase advances are positive. Therefore we want to check the results with negative μ_m values, which are nevertheless in agreement with the TbT measurement, using additional information from DELTA's electronic logbook [64] again. The response matrices with negative μ_m are indexed by RIDs 54, 55, 75, and 161.

To find a simple criterion to check if μ_1 is really negative in independence of COBEAs estimation, we return to the description of A_x in terms of beam optical functions. It was found that $A_{km} \propto 1/\sin(\mu_m)$, thus a sign change of μ_m for two matrices recorded with otherwise similar beam optics should effect the sign of all A_{km} for that mode. In a weakly coupled setting like that of DELTA, we can then check the quadrant by the condition

$$r^{\mathrm{A}}_{jkw} r^{\mathrm{B}}_{jkw} < 0 \text{ with } w = m \text{ and for most } j, k. \tag{7.1}$$

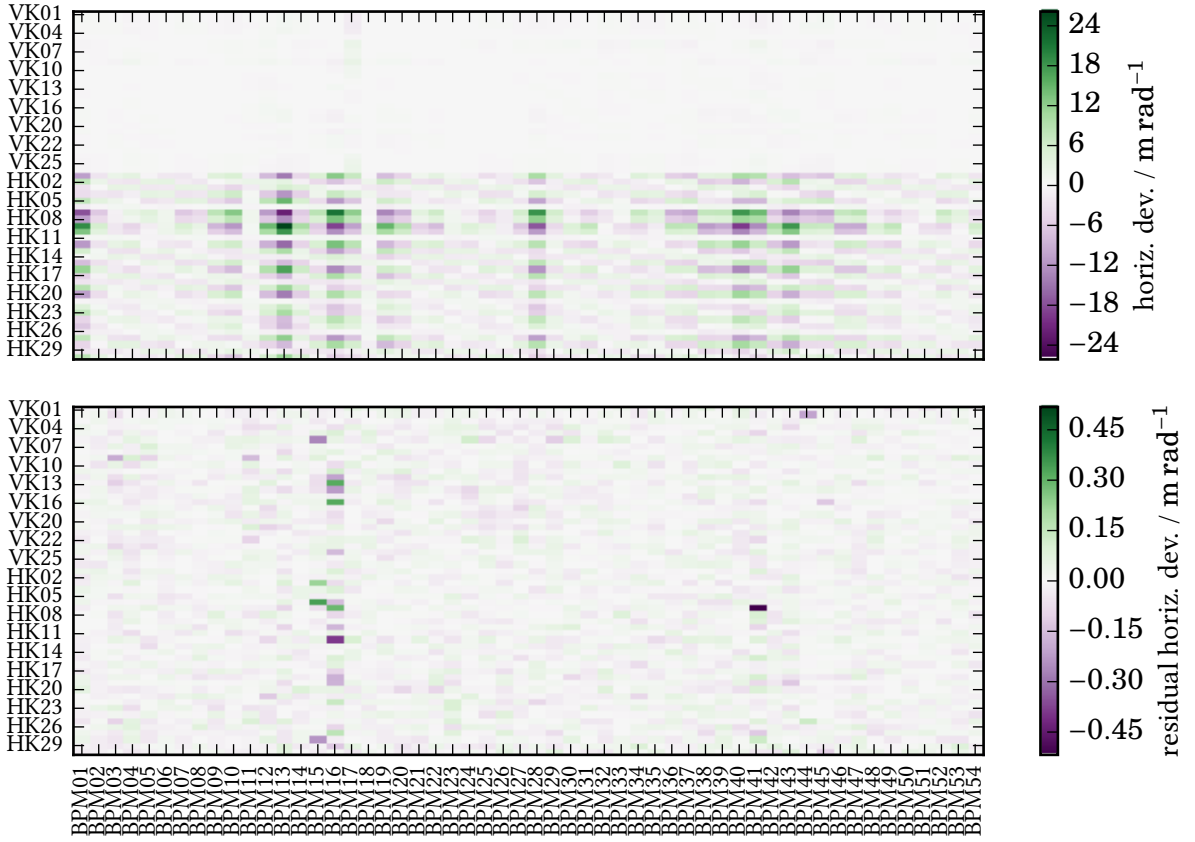

Figure 7.11.: Horizontal orbit deviations (top) and the fit residuals to COBEA's prediction (bottom) for RID 54. The COBEA-predicted tune is $8.843\,99 \pm 1.35 \times 10^{-2}$ for mode 1 and $3.220\,17 \pm 9.72 \times 10^{-3}$ for mode 2, with one TbT tune being $0.150\,39 = 1 - 0.849\,61$, the other being invalid.

RID {54,55} **response.081125-{550MeV, 551_550MeV}**

The signal-to-noise ratio (Fig. 7.5) and the fit errors (Fig. 7.11 for RID 54) indicate a successful optimization using COBEA for both matrices. In addition, the predictions of absolute value $|\mu_1|$ for the mainly horizontal modal phase advance and the full predictions of μ_2 for the mainly vertical modal phase advance are consistent.

For comparison using (7.1), we take RID 57 (see Fig. 7.7) which was also recorded in 550 MeV mode, but two days later with confirmed $\mu_1 > 0$ from COBEA and Q-Pulser. As can be seen in Fig. 7.12, most horizontal components compared between 54 and 57 respectively 55 and 57 have changed sign, while most of these are equal between RID 54 and 55. This indicates that both μ_1 values for RID 54 and 55, which mainly effect the horizontal plane, are indeed negative and COBEAs estimate can be considered as correct.

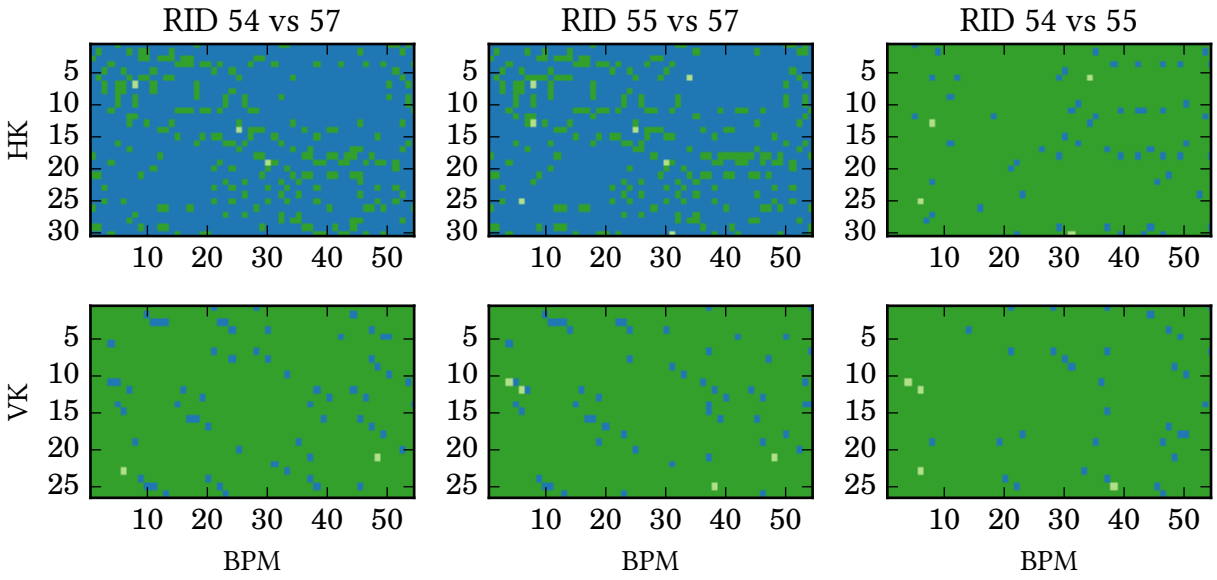

Figure 7.12.: Comparison of sign changes between RIDs 54, 55 and 57 for the horizontal and vertical plane with respective corrector sets. Green indicates values > 0, light green indicates zero, blue indicates values < 0.

RID 75 **response.100415-FEL-1**

The recording of this matrix is documented in DELTA's Elog [64].

- Settings G. Schmidt, 11:46

  …For the orbit correction to work, it was required to record a new response matrix *(RID 75)*. This response deviates significantly from the one used before *(RID 74)*. Due to this and due to the behavior of the Q-Pulser tune under variation of quadrupole current, the horizontal tune is supposedly located below 9. The QF magnets are run with 2% less current in comparison to old settings.

Again checking (7.1) by comparing RID 74 and 75 (Fig. 7.15), we see that most components have changed sign as expected for negative phase advances. Also being supported by the Q-Pulser behavior and the operator, we can consider COBEAs estimate as being correct.

RID 161 **response.150121-alternative-optik-bolsinger**

This response has been recorded in a new beam optics mode for the first time. As the injection efficiency was low, not more than ≈ 1.5 mA could be stored.

As this is the only response for this beam optics mode, we can only compare it (7.1) with the more common operation modes of DELTA for which an abundance of responses exist. Even a comparison with RID 160 (Fig. 7.15), which has been recorded 16 days before using different magnet settings, indicates that the sign of μ_1 is negative and COBEAs prediction is correct.

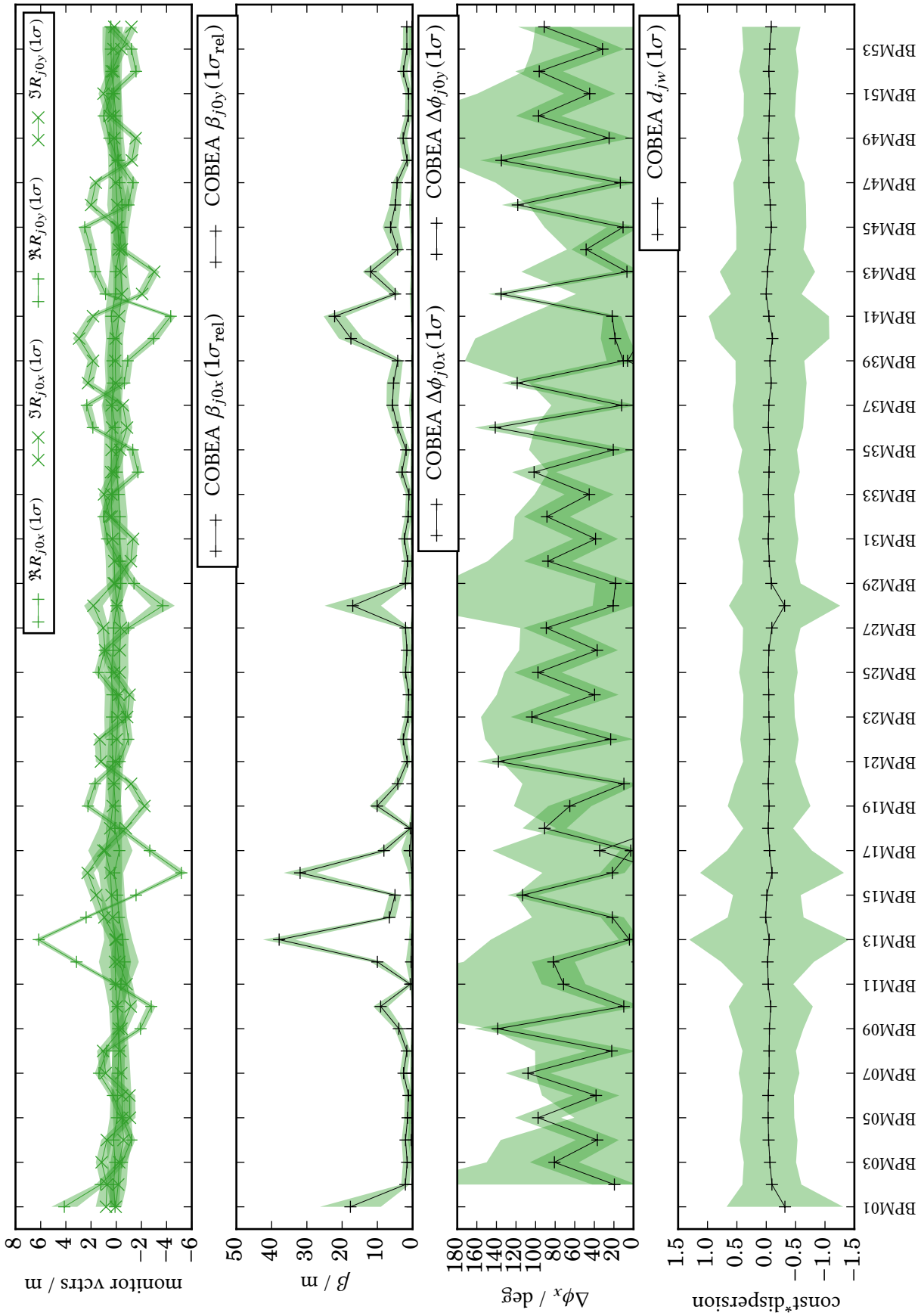

Figure 7.13.: Optical functions of mode 1 predicted by COBEA for RID 55 at monitor positions. The COBEA-predicted tune for this mode is $8.838\,10 \pm 1.56 \times 10^{-2}$, with the TbT tune being $0.162\,11 = 1 - 0.837\,89$.

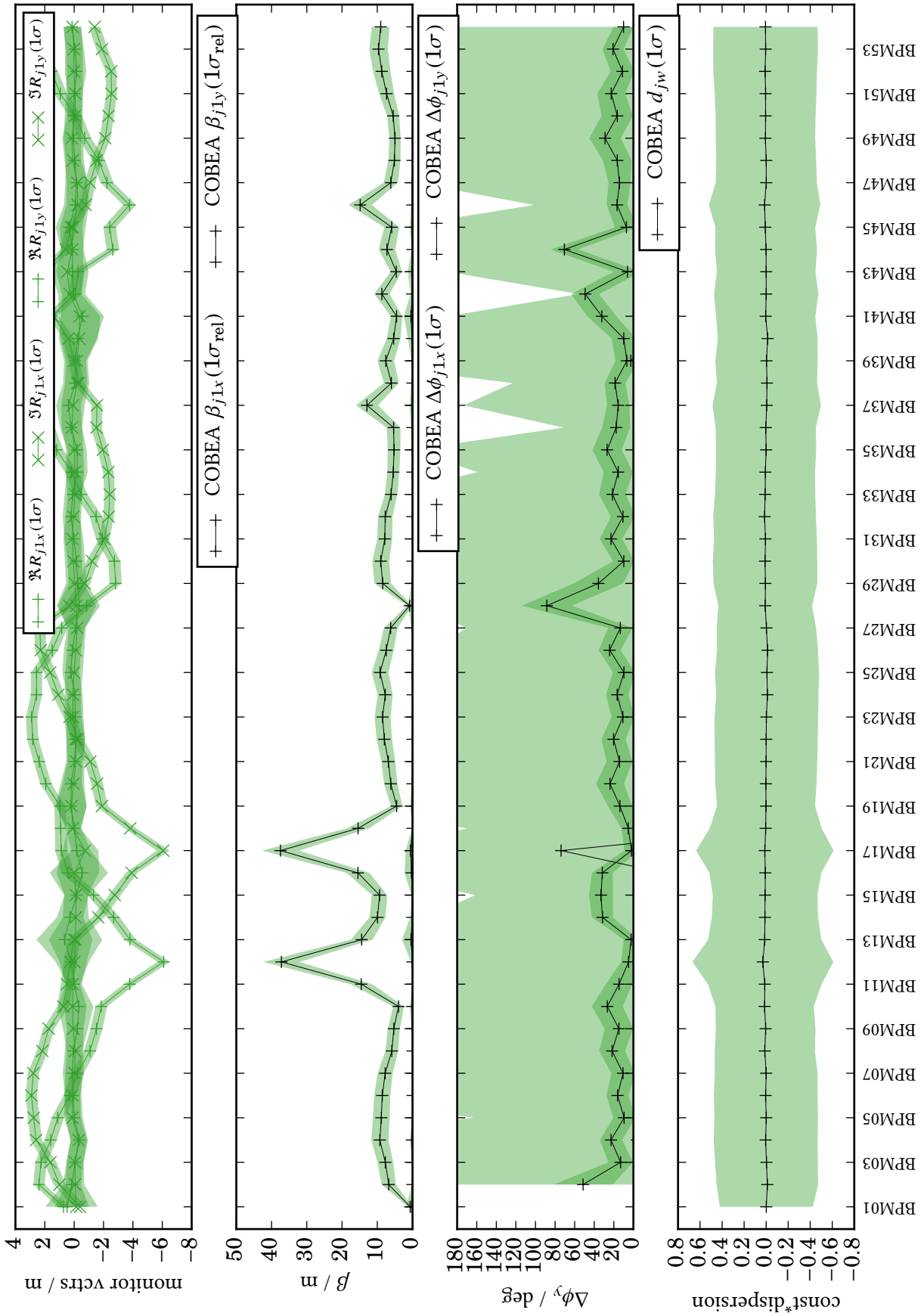

Figure 7.14.: Optical functions of mode 2 predicted by COBEA for RID 55 at monitor positions. The COBEA-predicted tune for this mode is $3.215\,33 \pm 1.10 \times 10^{-2}$, with the TbT tune being 0.215 82.

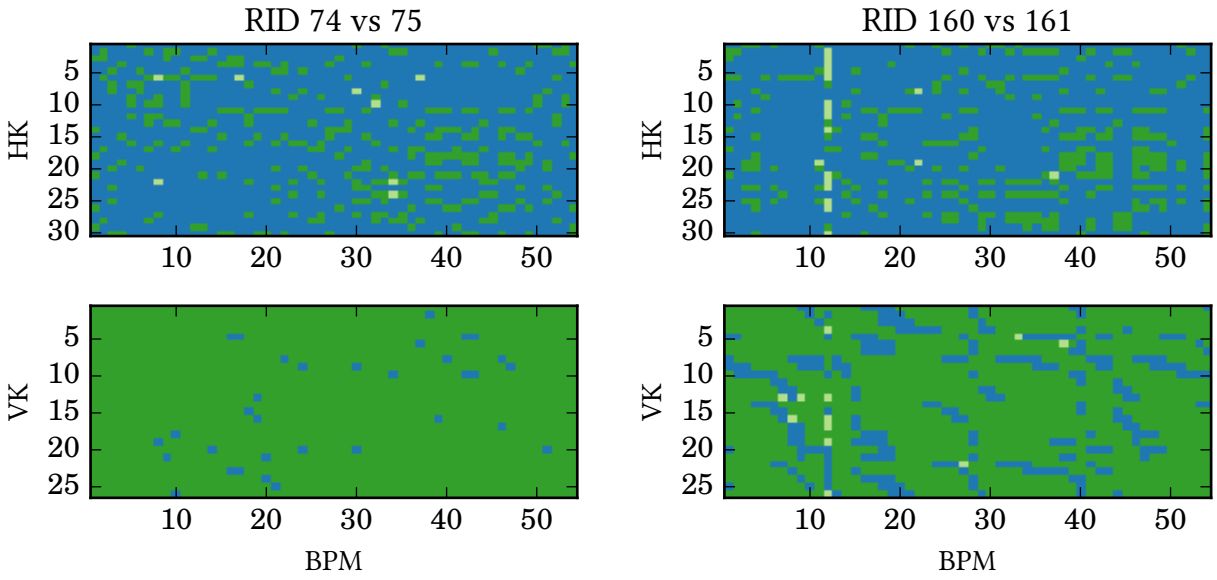

Figure 7.15.: Comparison of sign changes between RIDs 74 and 75 (left) and RIDs 160 and 161 (right) for the horizontal and vertical plane with respective corrector sets. Green indicates values > 0, light green indicates zero, blue indicates values < 0.

7.2.5. Results with large COBEA error estimates

RID 22 **response.070319-2**

This response has been recorded in 550 MeV mode. At that time, TbT monitors had just been installed in the storage ring and were in testing mode. In consequence, the orbit deviations recorded by these monitors are flawed as stated the next day (2007-03-20) in the Elog [64],

- Orbit monitors are flawed — H. Huck, 20.03.2007 08:21

  ...A vertical correction step has been applied. The beam position changed accordingly, which could be observed from the synchrotron light originating from the undulator. The monitored beam positions behaved differently. Especially the measured horizontal positions at BPM13 and BPM15 changed by several millimeters. This cannot be true, to be observed again from the undulator's synchrotron light.

This observation was followed by the recording of a new response matrix (RID 23), for which the σ errors estimated by COBEA are much smaller. It is interesting to see that the tune estimations between RID 22 and 23 do not change significantly, and one can suppose that a correct tune estimate for RID 22 is possible.

RID 127 **response.120817-1-donotuse**

For this RID, problems with two monitors are mentioned in the Elog [64].

- BPM44/45 crashed — P. Hartmann, 15:40

  After opening of the U55 undulator Ulf stated that the BL11 chamber was heating up. The cause for this was that BPMs 44 and 45 showed 0 mm as constant beam position. ...A new response (*RID128*) was recorded.

RID 149 **response.140821-1**

The Elog [64] reports problems with correctors (also called 'steerers').

- Steerers cannot be controlled J. Friedl, 15:35

  ...The orbit correction tried to change the current values (*of correctors*), but the readouts did not change.

Although no origin for the large σ of RID 106 can be found from the Elog, one may suspect that they originate from another unmentioned defect in the storage ring.

We state that for all mentioned response matrices with large σ, the actual tune prediction of COBEA and TbT tunes match nevertheless with sufficient accuracy. This is an indicator that the COBEA procedure is fairly robust.

7.2.6. Summary

171 response matrices recorded at the DELTA storage ring have been evaluated using COBEA. Whenever complementary data from the Q-Pulser system was available (120 responses), the tunes either were consistent (114 cases, 94%) or the deviations could be explained by measurement deficiencies in the respective operation mode of DELTA. The applicability of COBEA is limited by the quality of closed orbit data, which is either decreased by small beam currents (RID 166) or magnetic short-circuits (RIDs 151, 153). The Q-Pulser system is either limited by chromatic effects that modulate the free betatron oscillations with synchrotron sidebands (RID 57) or single-bunch mode with low current (RID 164).

In summary, COBEA has successively analyzed all response matrices for which closed-orbit data of sufficient quality was available. Of the 120 responses which could be validated against Q-Pulser, this was the case for 117 of 120 matrices.

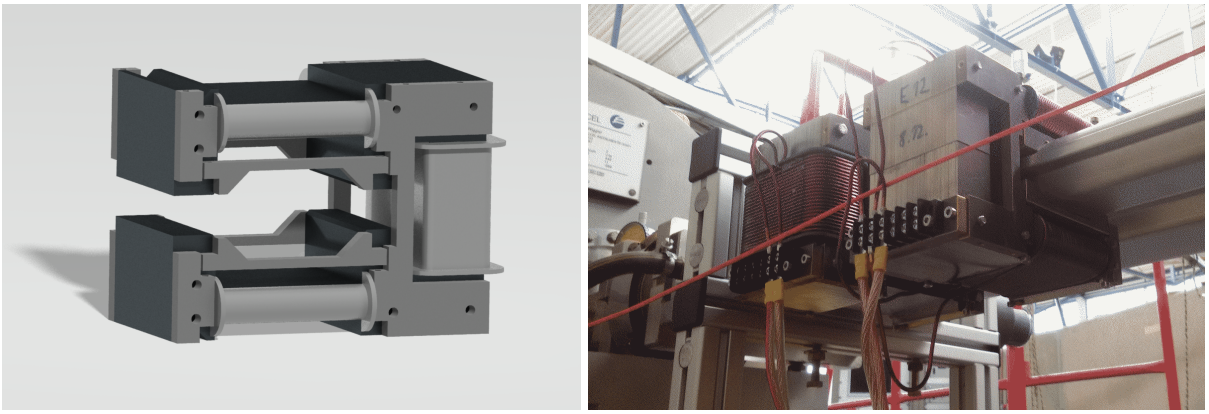

Figure 8.1.: Towalski correctors rendered from CAD [66] and mounted onto the beam chamber.

8. COBEA experiments at the DELTA storage ring

Having used COBEA on already existing data recorded by the DELTA control system, we found the possibility of validating COBEAs tune predictions by the recorded TbT tunes, which is possible using a single TbT-capable monitor. As described in sec. 7.1.2, there are 10 TbT-capable monitors installed in the DELTA storage ring with which it is possible in principle to record monitor vectors. Therefore, the scope of this chapter is to validate COBEA using monitor vectors.

8.1. Towalski correctors

Additional corrector magnets, called Towalski correctors in the following, have been installed in the DELTA storage ring for the purpose of fast orbit feedback [66]. Due to the feedback system still being under commission, these correctors are not used in DELTA's standard operation mode. This has the following advantages.

1. As the correctors are realized as autonomous devices, there is no dependency on other currents in their yokes.
2. As the correctors are not used, their standard current is zero, and thus hysteresis effects are small. In addition, the interference with DELTA's operation is minimal.
3. Due to their design purpose as Fast Orbit Feedback correctors, the Towalski correctors have a characteristic low-pass frequency of $\approx 300\,\mathrm{Hz}$ [66] for applying fields to the beam, only limited by the particle beam chamber. For the time scale of s considered in this thesis, this results in quasi-instantaneous beam excitation.

A picture of an installed Towalski corrector is shown in Fig. 8.1. The detailed topology information is shown in sec. A.4.1.

8.1.1. Applying and recording perturbations

The corrector currents are generated by power supplies [67, 66], which can be remotely controlled by different methods and interfaces. The most direct but slowest method is given by control over a TCP/IP socket. A corresponding Python interface has been implemented in [68]. Although excitation of the beam via these correctors is quasi-instantaneous, we use the TCP/IP socket control, as the recording of orbit perturbations is anyway limited by the low-pass behavior of DELTA's standard monitor readout system (sec. 7.1.2).

The procedure of obtaining a general response matrix can be split into the following steps.

1. For the given corrector k, apply a current I_n from a sequence of length N, which last entry is zero.

   For our measurements, we use the current values $I = (0\,\mathrm{A}, 1\,\mathrm{A}, -1\,\mathrm{A}, 0\,\mathrm{A})$, so that the sum of sequence values vanishes. This has the advantage that the unperturbed closed orbit can be monitored during measurement, hysteresis effects are compensated after each corrector cycle and that nonlinear asymmetric behavior can be checked.

2. Measure and record the closed orbit in intervals of $\Delta t \approx 0.2\,\mathrm{s}$.

   Each measurement is saved into an array $\bar{r}_{jkwp}$ with increasing p. This is done $S = 25$ times so that low-pass behavior of monitors can be analyzed.

3. Increase n by 1 and apply steps 1 and 2 until the end of the current list is reached. p is only reset to zero at the end of the current list, so that SN measurement values exist for any fixed j, k, w.

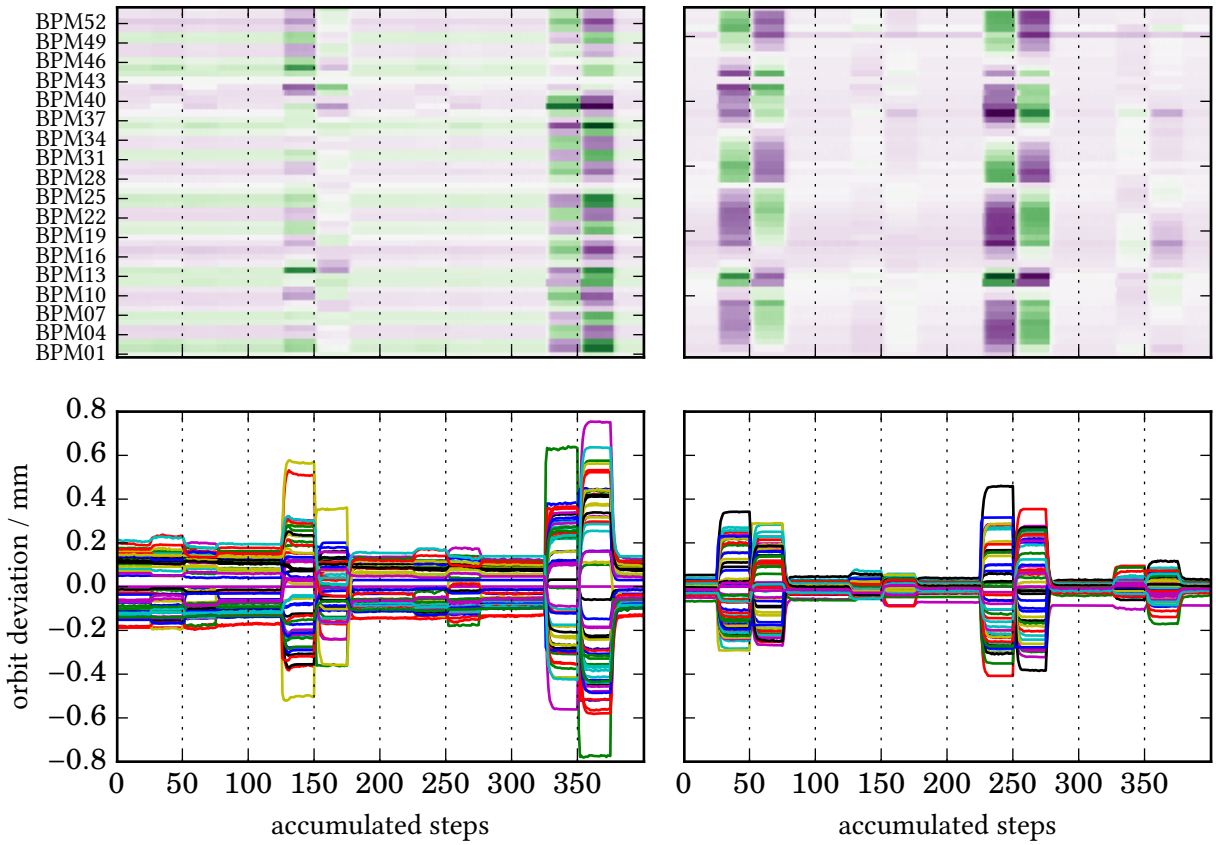

Figure 8.2.: Recording of perturbations for correctors svk01, shk01, svk02, shk02 as projections onto the step–corrector (top) resp. step-deviation (bottom) plane in x (left) and y (right) direction.

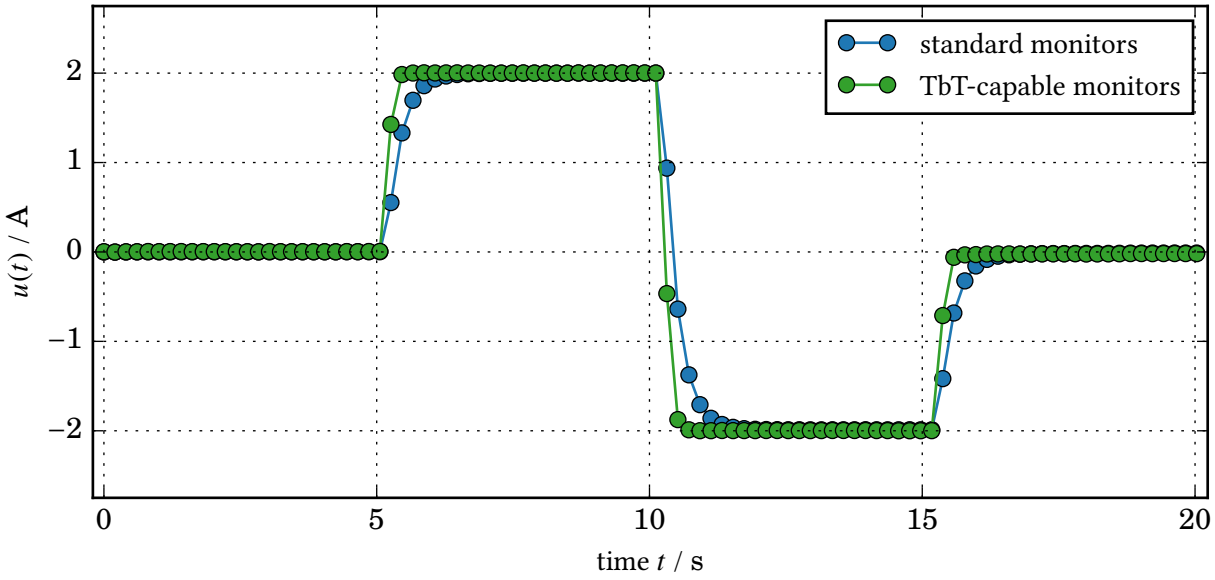

Figure 8.3.: Time pattern $u(t_p)$ resulting from outer product decomposition of measurement data (TID 2), performed separately for standard and TbT-capable monitors.

4. Apply steps 1–3 successively for all K correctors.

The measured closed orbit deviations are shown in Fig. 8.2, where the step counting p is always increasing (accumulated steps) to display perturbations of more than one corrector. The low-pass behavior of the monitors can be observed.

8.1.2. Building of current-response matrices

Many algorithms are suitable to obtain a current-response matrix from the recorded measurement data. In our case we must consider the finite response times due to network latency (TCP/IP), corrector hysteresis and monitor low-pass behavior (sec. 7.1.1), of which the latter is the strongest effect.

In the following, we will consider the storage ring as a linear time-invariant system that is excited by different corrector inputs. We assume that, up to a scaling factor, the impulse response from each corrector k to each monitor j is identical and describe it by a response function $u(t)$. Then we obtain the relation

$$\bar{r}_{jkwp} = u(t_p) r_{jkw} \text{ with } t_p = p\Delta t,$$

which can be used to decompose the measurement data $\bar{r}_{jkw}(t_n)$ into an outer product along indices (j, k, w) with p (sec. A.3.2). From the resulting values $u(t_p)$, the time structure of the response can be extracted.

The time pattern $u(t)$ has the unit of current. For the sample range in which the first non-zero current occurs, it is necessary to check if I_n and the corresponding $u(t_p)$ values have identical sign. If this is not the case, the sign of $u(t)$ and all r_{jkw} must be changed. For the time constants occurring in the low-pass behavior we can then assume that the maximum value is equivalent to the static perturbation excitation

$$\max_p u(t_p) \stackrel{!}{=} \max_n I_n$$

and scale $u(t)$ and all response components r_{jkw} from PCA accordingly, so that their product remains unchanged. According to our assumptions, r_{jkw} is the current-response matrix for the given corrector set.

As the standard and TbT-capable monitor sets at DELTA have different low-pass behavior, we use the PCA separately on both monitor sets.[1] For a typical measurement, this results in the patterns shown in Fig. 8.3, in which the low-pass behavior of the system can be observed.

8.2. TbT data sources and preprocessing

Somewhat a digression on the core topics in this thesis, this chapter describes existing measurement techniques using special, Turn-by-Turn capable beam position monitors and their limits at storage rings. It will be shown how eigenorbits at specific longitudinal coordinates in the ring can be found using turn-by-turn (TbT) capable monitors. The ideas for TbT measurement and analysis have been developed in a variety of works, e.g. [16, 69, 13, 70], and many of them are thus only re-expressed using eigenorbits (sec. 2.3) in the following. A comparison of some existing techniques is given in sec. A.2.

Basic Measurement process

We assume that a kicker device exists in the accelerator that deflects particles from the closed orbit by a short pulse. The duration of this pulse is below the circulation time of the beam and thus a transient effect. For the DELTA storage ring, the kicker device is identical to the one used for the Q-Pulser system [63] (sec. 7.2).

The measurement process begins directly after the pulse has ended with turn $n = 0$ and proceeds a finite time to turn $N - 1$, consequently, $\sum_n$ implies $\sum_{n=0}^{N-1}$ in the following. For the considered turns, particles have a deviation from the closed orbit and behave again like in an unperturbed segment. Using monitor vector components (sec. 3.2.1), one may rewrite the single particle deviation observed at a TbT-capable monitor j at turn n in direction w as (2.7)

$$\tilde{r}_{jnw} = \Re\left\{\sum_m R_{jmw} \mathrm{e}^{\mathrm{i}n\mu_m}\right\}. \tag{8.1}$$

For machine studies, it is often advisable to use small beam excitations, not only due to minimizing the risk of beam loss, but also to stay in the linear regime of the oscillations, so that nonlinear effects are very small. In dependence of the beam charge, this approach can result in a small signal-to-noise ratio for single measurements. To lower the measurement error, it is thus often necessary to average over multiple measurements respectively excitation pulses P.

Due to several problems in DELTA's TbT measurement system, the only assumptions we will use in the beginning of our data analysis is that the excitation of the beam leads to oscillation with the two modal phase advances μ_m, so that the acquired TbT data should have the form

$$\tilde{r}_{jnwp} = \Re\left\{\sum_m C_{jmwp} \mathrm{e}^{\mathrm{i}n\mu_m}\right\}. \tag{8.2}$$

[1] The discussed techniques are not used for DELTA standard response matrices, but may allow to obtain more precise response matrices. With further development of the maximum condition into a fitting of exponential dependencies, the measurement time may also be decreased.

8.2.1. Least-Squares Spectral Analysis (LSSA)

We can reformulate (8.2) as an optimization problem

$$\text{find } \mu_m, C_{jmwp} \text{ for all } j,m,w,p \text{ so that } \quad \chi^2 = \sum_{jnw} \left| \tilde{r}_{jnw} - \Re\left\{ \sum_m C_{jmwp} \mathrm{e}^{\mathrm{i}n\mu_m} \right\} \right|^2 \text{ is minimal.} \tag{8.3}$$

This type of regression problem is linear in all C_{jmwp}, but nonlinear in model phase advances μ_m. Consequently, instead of handling the complete optimization problem using a nonlinear optimizer, one may separate the linear part and treat it as an overdetermined ($JNMP \gg 2JM^2P + M$) equation system.

Solving the problem in this context is known as Vanicek [71] or Lomb method [16] and called Least-Squares Spectral Analysis (LSSA) in the following, based on [71]. LSSA is equivalent to a fit of M cisoids with fixed, arbitrary phase advances μ_m to a sequence in n.

One may reorder the the TbT data $\tilde{r}_{jnwp}$ into a two-dimensional matrix $\mathbf{B}$ in which the indices j,w,p are combined into one axis so that the matrix size is $N \times (JMP)$. Then LSSA reduces to solving an overdetermined equation system containing $\Re$ (see sec. 5.1.1)

$$\mathbf{B} = \Re\{\mathbf{UA}\} \quad \text{with} \quad U_{nm} = \mathrm{e}^{\mathrm{i}n\mu_m} \tag{8.4}$$

and a complex-valued matrix $\mathbf{A}$ of size $M \times (JMP)$ which holds the reordered phasors C_{jmwp}. As the system is overdetermined, one may use SVD (sec. A.3.1) to solve it with a minimal least-square error χ^2, which is also the error to be minimized in (8.3). A sketch of the LSSA optimization step with TbT data is shown in algorithm 8.4. The time complexity is discussed in sec. A.2.

Note that μ_m are only candidates for the correct modal phase advances and thus need to be subjected to optimization, which implies an "outer loop" (similar to the one used in NAFF [72], sec. A.2) which is controlled by a nonlinear bounded optimizer. A way to generate sufficient starting values is discussed in the next section.

Furthermore, by assuming linearity we have implied energy and transverse amplitude-independent constant model phase advances μ_m. Taking these dependencies into account leads to decoherence patterns in the TbT data, which are discussed in sec. A.2.3 in conjunction with extensions of LSSA optimization (sec. A.2.4).

Data:
reshaped TbT data $\mathbf{B} \in \mathbb{R}^{N\times(JD)}$ holding deviations $(\vec{r}_{jn})_d = B_{n(jd)} \in \mathbb{R}$ at J monitors
optional: shape function $u_{nm}(a_1,\ldots,a_P)$. if not given, $u_{nm} \equiv 1$.
for each iteration:
$\tilde{\mu}_m$ candidates for $1 \le m \le M$ optional: shape parameter candidates $a_1,\ldots,a_P$
Result:
Residual squared error χ^2;
shaped phasor data $\mathbf{A} \in \mathbb{C}^{M\times(JD)}$, holding monitor vectors $(\vec{R}_{jm})_d = A_{m(jd)} \in \mathbb{C}$ for M modes
for $m \leftarrow 1$ **to** M, $n \leftarrow 1$ **to** N **do**
  $U_{nm} = u_{nm}(a_1,\ldots,a_P)\mathrm{e}^{\mathrm{i}\tilde{\mu}_m n}$;
end
solve $\mathbf{B} = \Re\{\mathbf{UA}\}$ for $\mathbf{A}$;
$\chi^2 \leftarrow \|\mathbf{B} - \Re\{\mathbf{UA}\}\|^2$;

Figure 8.4.: LSSA optimization step to be encapsulated into a non-linear optimizer for all iteration candidates so that χ^2 is minimized.

8.2.2. Start values for μ by Fast Fourier Transform (FFT)

As the optimization problem (8.3) is nonlinear in modal phase advances μ_m, a procedure to find proper starting values is necessary. To process data from TbT monitors, we introduce the Discrete-Time Fourier transform (DTFT) [73] by an operator $\mathcal{F}$, so that any quantity x_n depending on a turn n has a DTFT spectrum depending on μ, given by

$$\mathcal{F}_\mu x_n = \sum_n^N x_n \mathrm{e}^{-\mathrm{i}\mu n} \text{ with } \mu \in]-\pi, \pi].$$

In consequence, for single particle motion and in the limit of measuring an infinite number of turns, one would obtain a DTFT

$$\lim_{N\to\infty} \mathcal{F}_\mu \tilde{r}_{jnwp} = N \begin{cases} C_{jmwp} & \text{for } \mu = \quad \mu_m \\ C^*_{jmwp} & \text{for } \mu = -\mu_m \\ 0 & \text{otherwise} \end{cases}, \tag{8.5}$$

from which μ_m and C_{jmwp} can be obtained up to an ambiguity in sign respectively conjugation (see sec. 6.1.2) by finding one maximum of $|\mathcal{F}_\mu \tilde{r}_{jnwp}|^2$ and its argument μ_m.[2]

In a realistic measurement, N is usually limited to 10^3–10^5. For finite N the leakage effect [73] occurs and the above spectrum is essentially blurred out by a cardinal sine function which characteristic width scales with $1/N$. Due to this blur, we can use the FFT algorithm and only compute the DTFT at $N/2+1$ equidistant positions (bins) [73]. Then, we use the phase advances corresponding to local maxima of the FFT as starting points for LSSA optimization.

Averaging

The question remains from which of the JMP different turn series this estimate should be taken. Two reasonable approaches exist

1. Use the spectrum of the sum of all turn series $|\mathcal{F}_\mu \sum_{jwp} \tilde{r}_{jnwp}|^2$. This approach is advantageous if the turn series have high coherence and the probability of phasors of different turn series canceling each other is low.
2. Use the sum of the spectrum of all turn series $\sum_{jwp} |\mathcal{F}_\mu r_{jnwp}|^2$. This approach is advantageous if the turn series are mostly incoherent.

For the results discussed in the following, a sufficient level of coherence in the TbT data exists to use the first approach. In Fig. 8.5 it can be observed that two spectral lines have only small variance under increase of averaging turn series, which indicates that these signals have high, although not complete, coherence.

8.2.3. Coherence and Synchronization at DELTA

After completing the LSSA-based optimization on a TbT data set, we can further analyze the relation of phasors in different pulses to each other. If the kicker pulse were perfectly synchronized to the bunch pattern, that is, its repetition time would be an exact multiple of the particle revolution time, and the measurement triggers of the TbT-capable monitors were all triggering at the same turn and bunch position, we would expect that the phasors do not depend on the pulse p,

$$C_{jmwp} = \text{const.} R_{jmw}.$$

[2] As the "power spectrum" $|\mathcal{F}_\mu \bullet|^2 = (\Re\mathcal{F}_\mu\bullet)^2 + (\Im\mathcal{F}_\mu\bullet)^2$ takes less steps to compute than the absolute value, we use it to find the maximum instead.

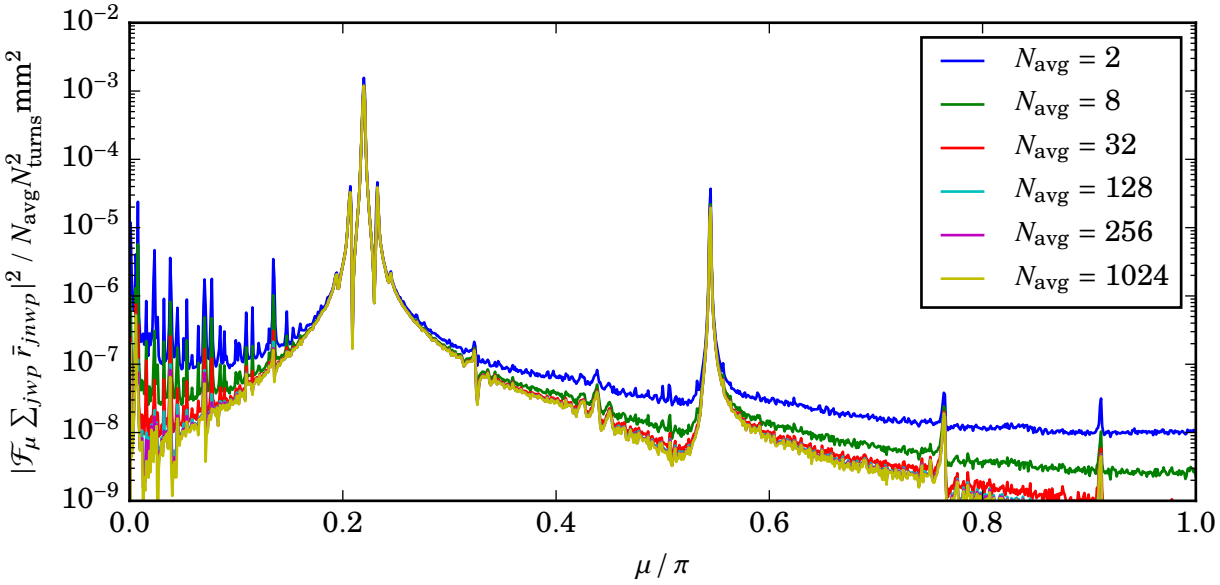

Figure 8.5.: Averaged spectra with N_{avg} different timeseries at a TbT-capable monitor using the unsynchronized kicker and approach 1. The incoherent signal parts reduce as N_{avg} increases.

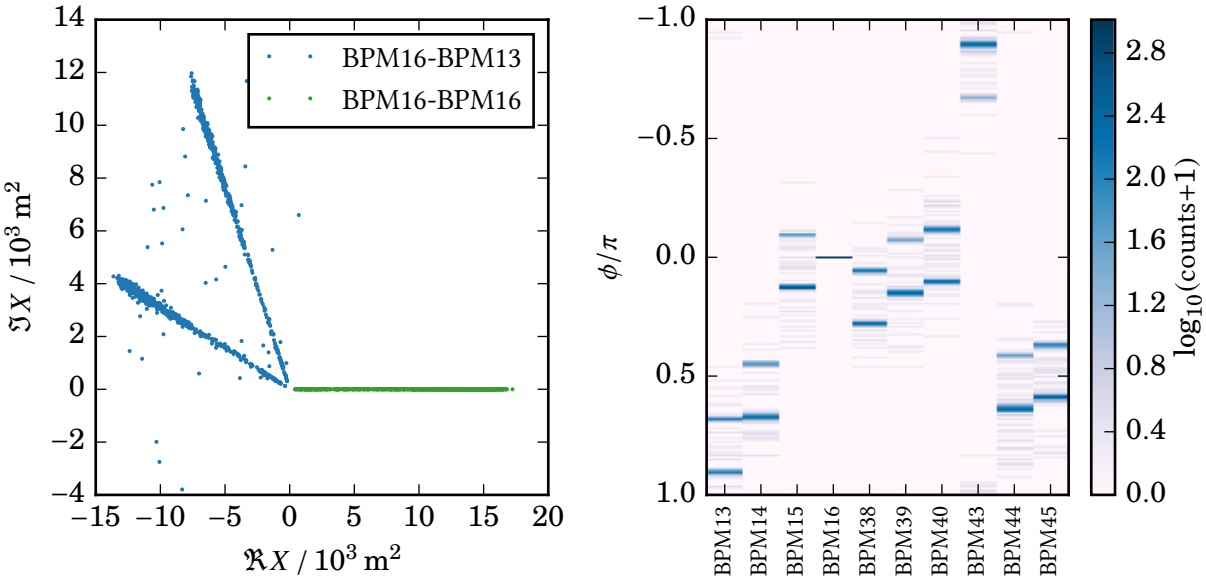

Figure 8.6.: Phasor correlations $D_{mwp}(j_r, j)$ for $m = w = 1$ and reference monitor BPM16. Left: Correlations for two monitors BPM13 and BPM16. Right: Histogram of angle($D_{mwp}(j_r, j)$) for all monitors.

For the used measurement setup, these assumptions are not true. The kicker pulse is not synchronized with the bunch pattern and thus excites different parts of the beam with different amplitudes on every pulse. As the timing of the pulse relative to the monitors is also different, this leads to a phase shift in the observed turn series. Both effects can be taken into account by a complex pulse coefficient c_{mp} so that

$$C_{jmwp} = c_{mp} R_{jmw}. \tag{8.6}$$

To check if the second condition holds nevertheless, we could correlate the phasors from different monitors j at the same pulse p with each other,

$$D_{mwp}(j_1, j_2) = C_{j_2 mwp} C^*_{j_1 mwp}, \text{ so that } \arg D_{mwp}(j_1, j_2) = \arg R_{j_2 mw} R^*_{j_1 mw}$$

should be independent of p. A presentation of all pulses for $w = m = 1$ and can be found in Fig. 8.6, which shows that there are essentially two phases $\arg D_{mwp}$ in which every pulse p for a monitor pair can be grouped into. Therefore the monitors do not all trigger on the same turn at every pulse.

Based on (8.2) and (8.6), a shift of the turn series by one-turn results in a modified pulse coefficient

$$\begin{aligned} \tilde{r}_{j(n-1)wp} &= \Re\left\{\sum_m R_{jmw} c_{mp} \mathrm{e}^{\mathrm{i}(n-1)\mu_m}\right\} \\ &= \Re\left\{\sum_m R_{jmw} c^-_{mp} \mathrm{e}^{\mathrm{i}n\mu_m}\right\} \quad \text{with } c^-_{mp} = c_{mp}\mathrm{e}^{-\mathrm{i}\mu_m}. \end{aligned} \tag{8.7}$$

This discrete shift in the angle of correlation corresponds to the observed behavior in Fig.8.6. To correct for the turn shifts, we use the following procedure

1. Define a reference monitor j_r, which is the monitor with the largest phasor amplitudes.
2. Correlate all C_{jmwp} to the reference monitor by computing $D_{mwp}(j_r, j)$.
3. Cluster the D_{mwp} values for each monitor into two groups separated by the angle μ_m.
4. rotate all C_{jmwp} which are classified into the second cluster by μ_m, computing new phasors $\bar{C}_{jmwp} = C_{jmwp}\mathrm{e}^{\mathrm{i}\mu_m}$.

Neglecting phasors which are shifted by more than one turn, the resulting corrected phasors should now fulfill (8.6),

$$\bar{C}_{jmwp} = \bar{c}_{mp} R_{jmw},$$

which is illustrated in Fig. 8.7. Inspection of this equation shows that, for a given m, $\bar{C}_{jmwp}$ is just an outer product of $\bar{c}_{mp}$ and R_{jmw} along indices p with (j, m, w), and we can thus use outer product decomposition (sec. A.3.2) to obtain both quantities up to constants indexed by m. As the monitor vectors are only defined up to constants indexed by m, we can thus obtain all monitor vectors by this procedure.

An example using measured TbT data is shown in Fig. 8.8. The residual phasors from our SVD fitting procedure are small for a large number of monitor and pulse combinations. Nevertheless, it can be seen that for some pulses, discrete and comparatively large fit errors occur. These errors are likely to stem from trigger timings of monitors which are more than one circulation time apart, and it is reasonable that such timing errors occur with a low probability, assuming a normal distribution of timing errors at all monitors. Neglecting these errors, we are able to reconstruct the input phasors, and thus all monitor vectors, with good accuracy.

Analysis of coherence by pulse coefficients

Outer product decomposition allows to check the coherence level a posteriori by discussion of the $\bar{c}_{mp}$ values. As stated before, the kicker is not synchronized with the circulation frequency of the beam, and therefore different parts are excited by the same kick envelope, which results in different amplitudes. The relative timing of the kicker pulse to the circulation frequency acquisition trigger is proportional to the phase of the observed signal. Therefore, the shape of the filling pattern respectively its convolution with the excitation pulse envelope is encoded in the pulse coefficients.

Summary

Analyzing TbT data, we found that the following problems occur at DELTA.

1. *High noise levels and missing kicker-to-monitor synchronization.*

   We found a way to extract absolute values of monitor vector components with low error levels by combining an extended LSSA optimization procedure with outer product decomposition on the input of a large number of independent measurements. If phase information were available, this procedure can be generalized by elementary means to incorporate phases of monitor vectors.

2. *Imperfect monitor-to-monitor synchronization.*

   This relates to the monitor triggers in relation to each other, and the missing synchronization of the 10 Hz Pulser signal (no phase lock). In consequence, the phase information cannot be used without a detailed investigation that would go beyond the scope of this work.

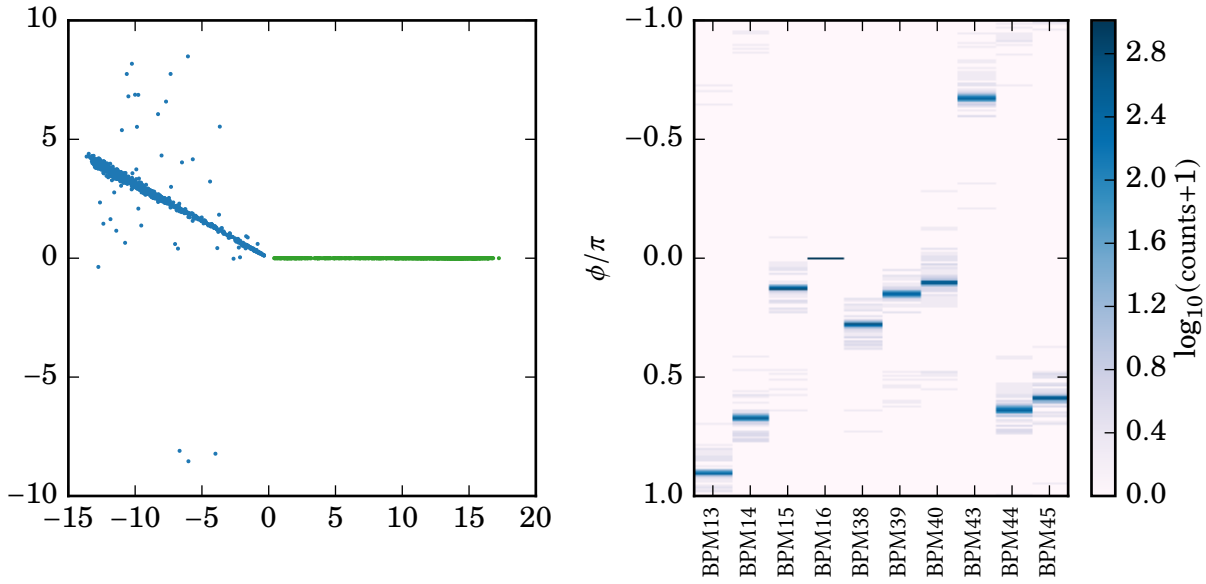

Figure 8.7.: Corrected phasor correlations $\bar{D}_{mwp}(j_r, j)$ for $m = w = 1$ and reference monitor BPM16. Left: Correlations for two monitors BPM13 and BPM16. Right: Histogram of angle$(D_{mwp}(j_r, j))$ for all monitors.

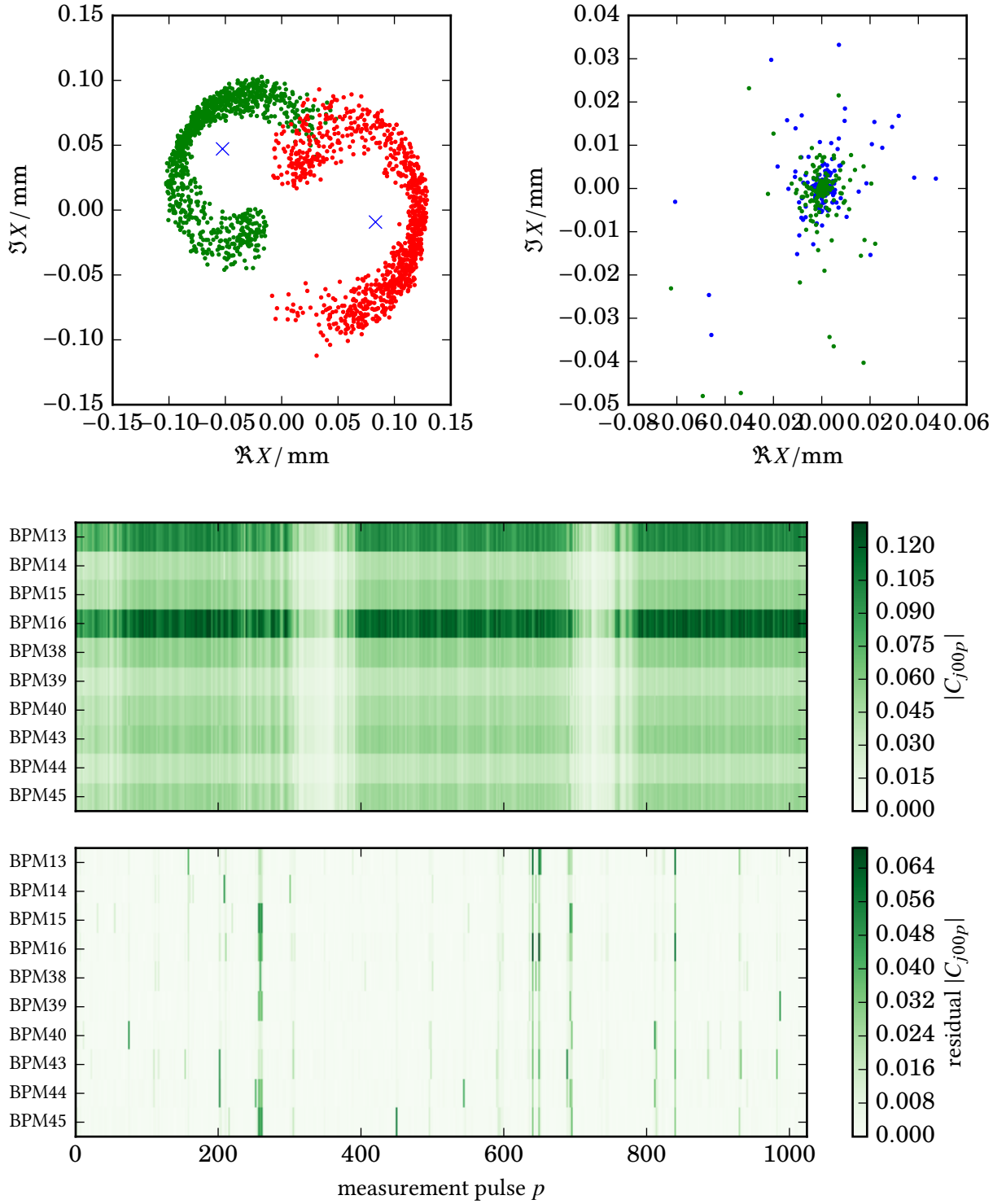

Figure 8.8.: Input phasors (top left), residual phasors (top right) at BPM13 and BPM16 and absolute values with residuals for all monitors (bottom) in mode 1, using all pulses from TbT measurement for TID 4.

8.3. Running COBEA

26 closed orbit measurements with Towalski correctors have been performed in the time between 2015-01 and 2015-06 using the measurement procedure described in sec. 8.1.1. Each measurement was processed to a response matrix with a negligible duration of computation. All obtained response matrices then were processed with COBEA and are identified by a response matrix index called TID in the following.

A typical run of the COBEA optimization layer took about 30 s on a standard PC including convergence output. This fast convergence in comparison to the standard DELTA response matrices relates to the lower number of used correctors. An example of COBEAs fit performance is given in Fig. 8.11. As can be observed for this case, the ratio of maximum values of deviation and residuals is > 100.

For 23 of the 26 matrices, additional TbT data has been recorded, which was processed using the techniques described in sec. 8.2. This allows to compare betatron tunes and monitor vectors at TbT-capable monitors between TbT measurements and COBEA. The measurements for which no TbT data is available can be identified as TID 1, 2, and 8.

8.3.1. Overview of results

A comparison of all computed COBEA tunes in comparison to TbT measurements (sec. 8.2) is shown in Fig. 8.9. It can be observed that the amplitude ratio is always higher than 60, which is a small improvement over the standard response matrix computations in sec. 7.2.

It can be seen that extensive use of TbT data processing (sec. 8.2), which is not possible for the Q-Pulser online feedback allows the TbT predictions to agree very well with those of COBEA. In fact, there is no single measurement for which tune deviations larger than 1.5×10^{-3} could be found.

8.3.2. Comparison of absolute monitor vectors

As TbT data of multiple monitors has been recorded, we are able to compare the computed TbT monitor vectors with those of COBEA up to a global scaling factor. Fig. 8.10 shows the computed $\beta_{j0\mathrm{x}}$ values for all TID responses.

To compare the scaling, we may re-normalize our TbT monitor vectors to the normalized monitor vectors of COBEA by comparing $\sum_j |\vec{R}_{jm}|^2$ for both sets and modes. We can now compute β function values that would be consistent with TbT analysis with those of COBEA and analyze the deviations. An analysis of β values is shown in Figs. 8.12 and 8.13, where it can be seen that small deviations in β values occur.

As a detailed example, a comparison of monitor vectors with extracted TbT information is shown for TID 22 in Figs. 8.14 and 8.15.

Summary

As can be seen in the aforementioned figures, there is a good agreement between COBEA-predicted optical parameters and TbT-predicted optical parameters at a monitor subset with LSSA optimization. Therefore one can consider the application of COBEA to the TID datasets successful.

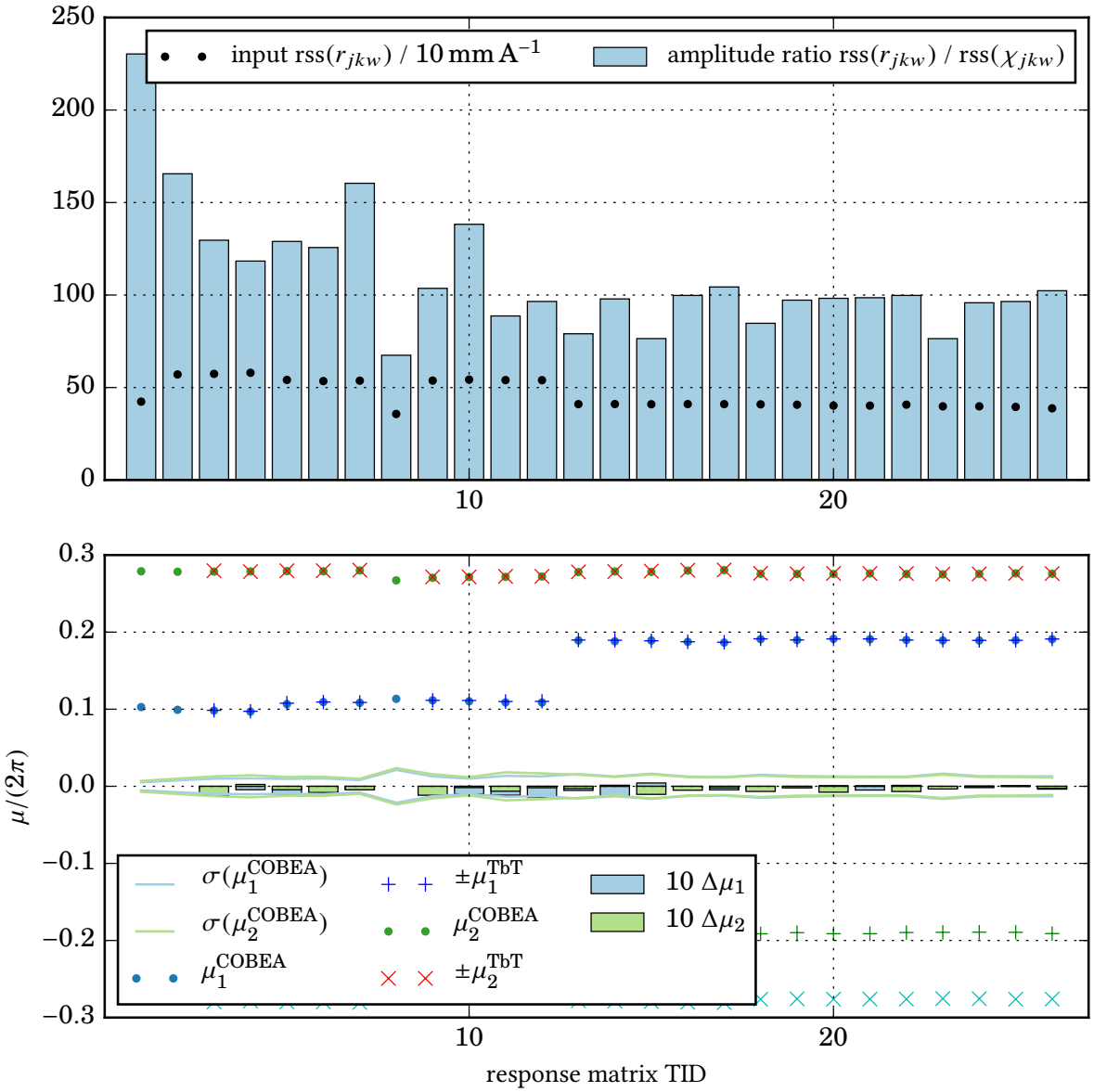

Figure 8.9.: Comparison of COBEA tune results and simultaneously recorded TbT tunes using 26 response matrices which construction is discussed in sec. 8.1. The labels use the definition $\mathrm{rss}(\bullet) = \sqrt{\sum_{jkw} \bullet^2}$. Note that the deviations between COBEA and TbT tunes are scaled by the factor 10.

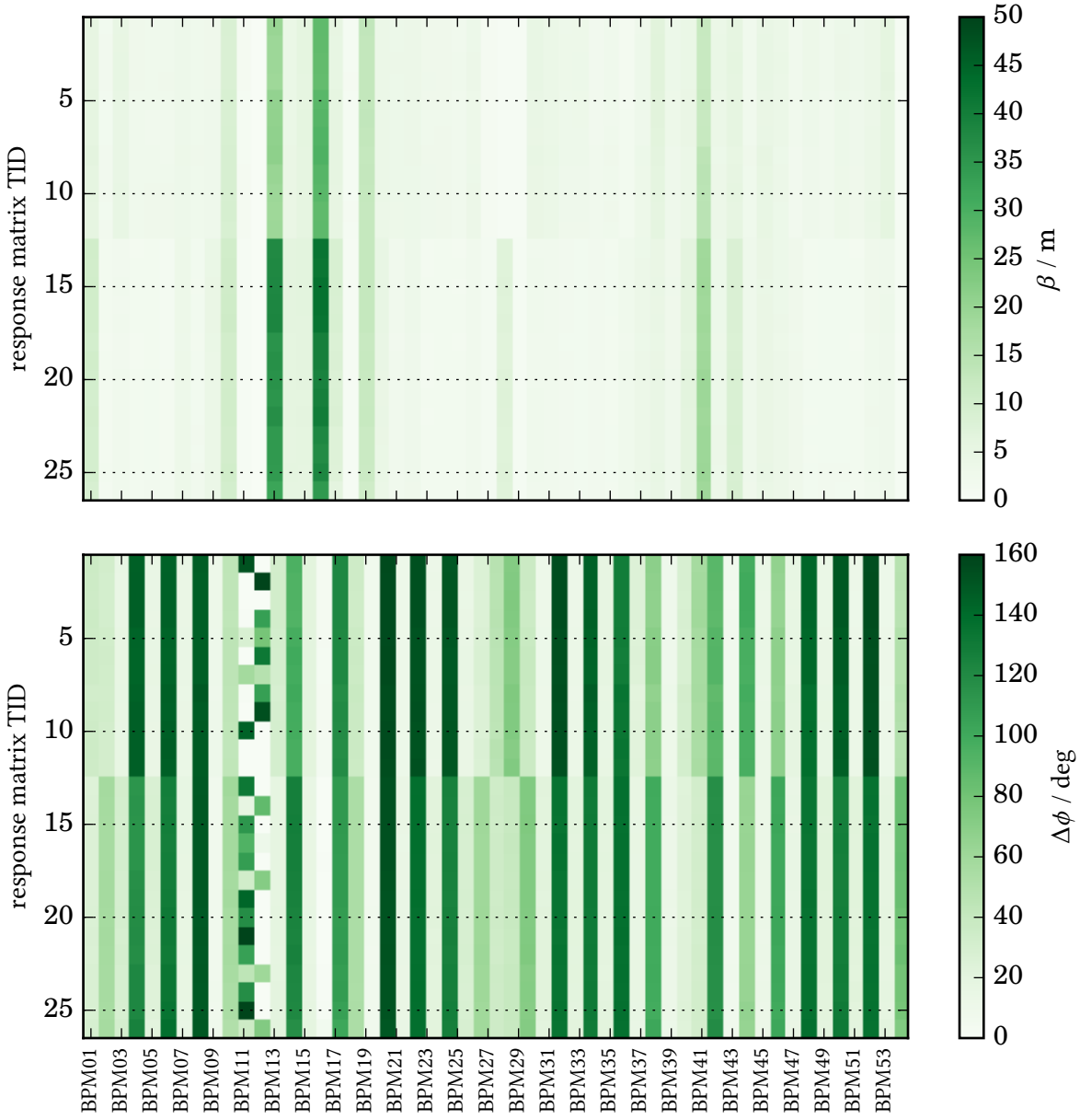

Figure 8.10.: COBEA-predicted β functions and phase advances $\Delta\phi$ per monitor for all measurements with Towalski correctors in the X plane. Five weeks passed between the measurements of TID 12 and TID 13. The behavior in phase advance at BPM11 and BPM12 results from a nonfunctional BPM12, but does not influence the rest of the analysis.

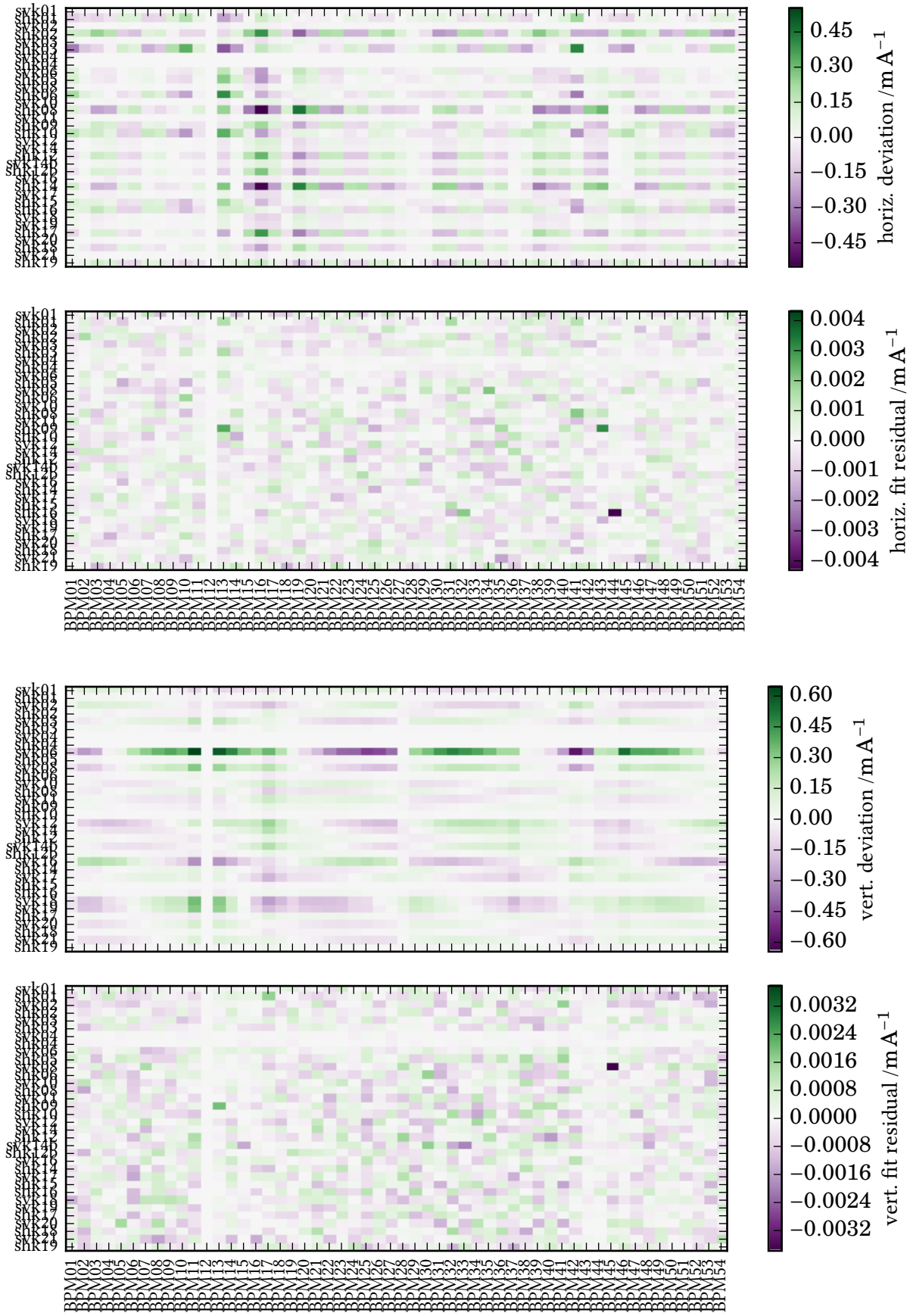

Figure 8.11.: Orbit deviations and fit residuals to COBEA's prediction for TID 12 in the horizontal (2 upper plots) and vertical (2 lower plots) plane.

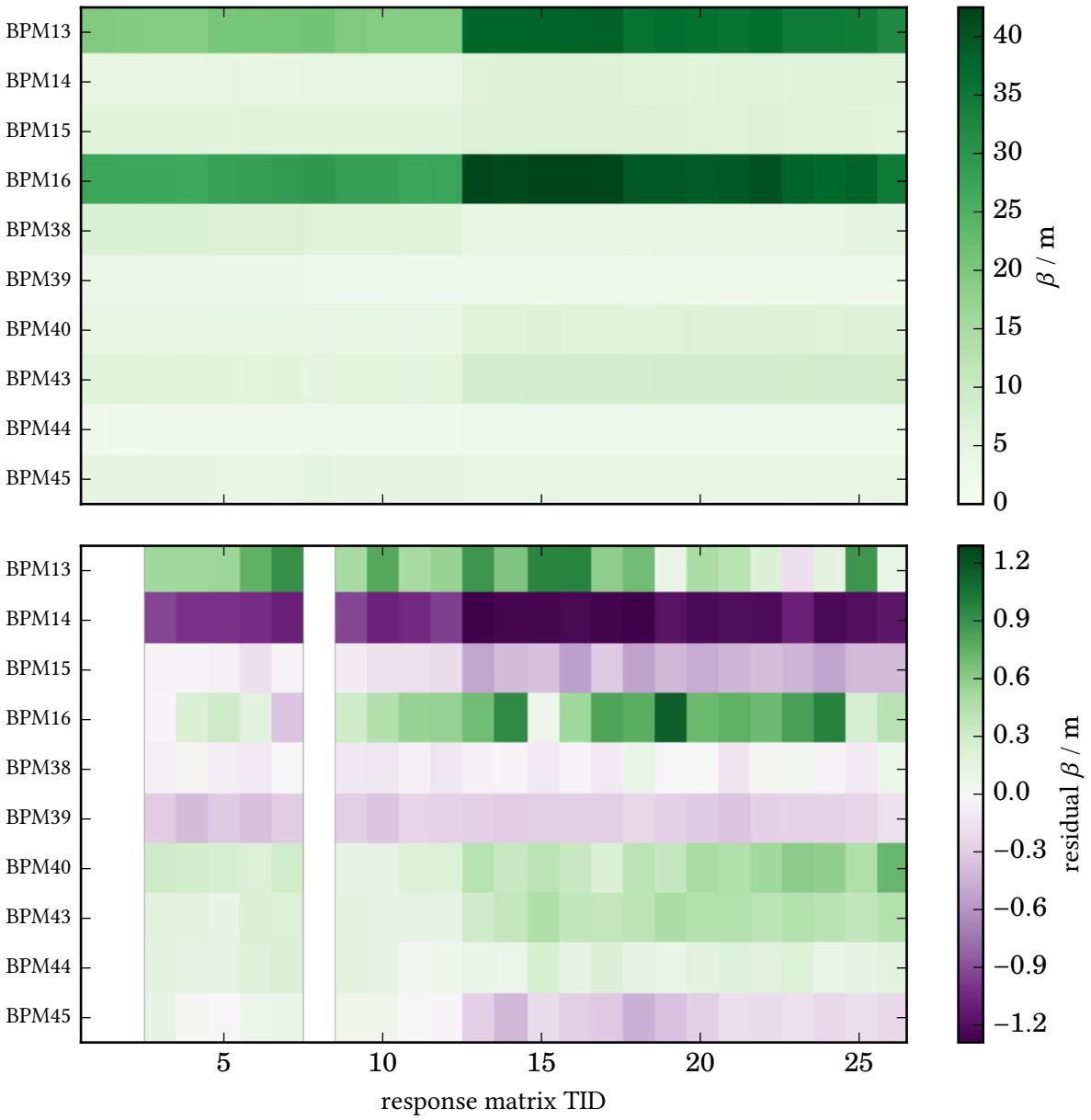

Figure 8.12.: TbT-predicted β functions and their deviation from COBEAs prediction for all measurements for $m = w = 1$ (horizontal plane).

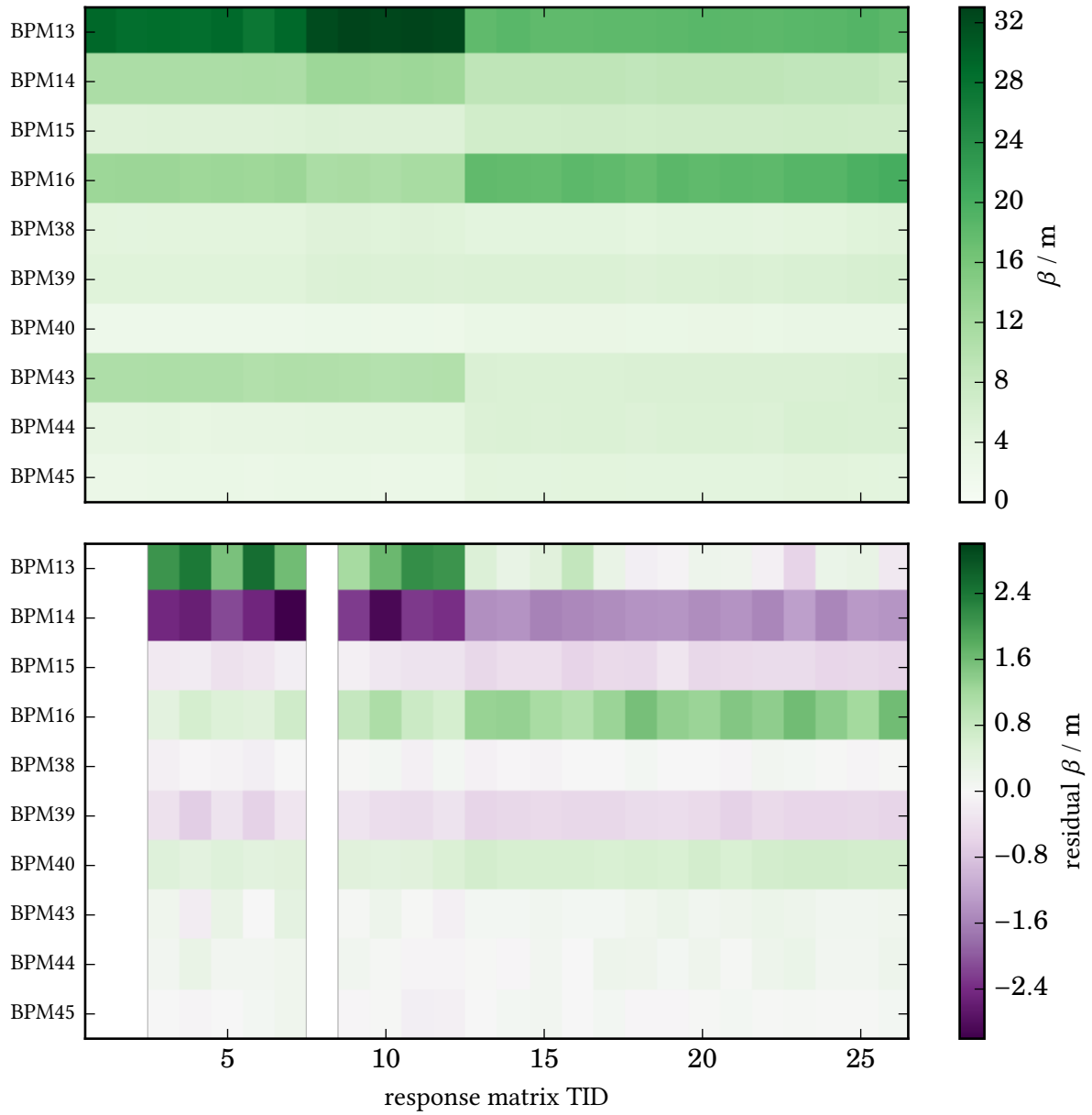

Figure 8.13.: TbT-predicted β functions and their deviation from COBEAs prediction for all measurements for $m = w = 2$ (vertical plane).

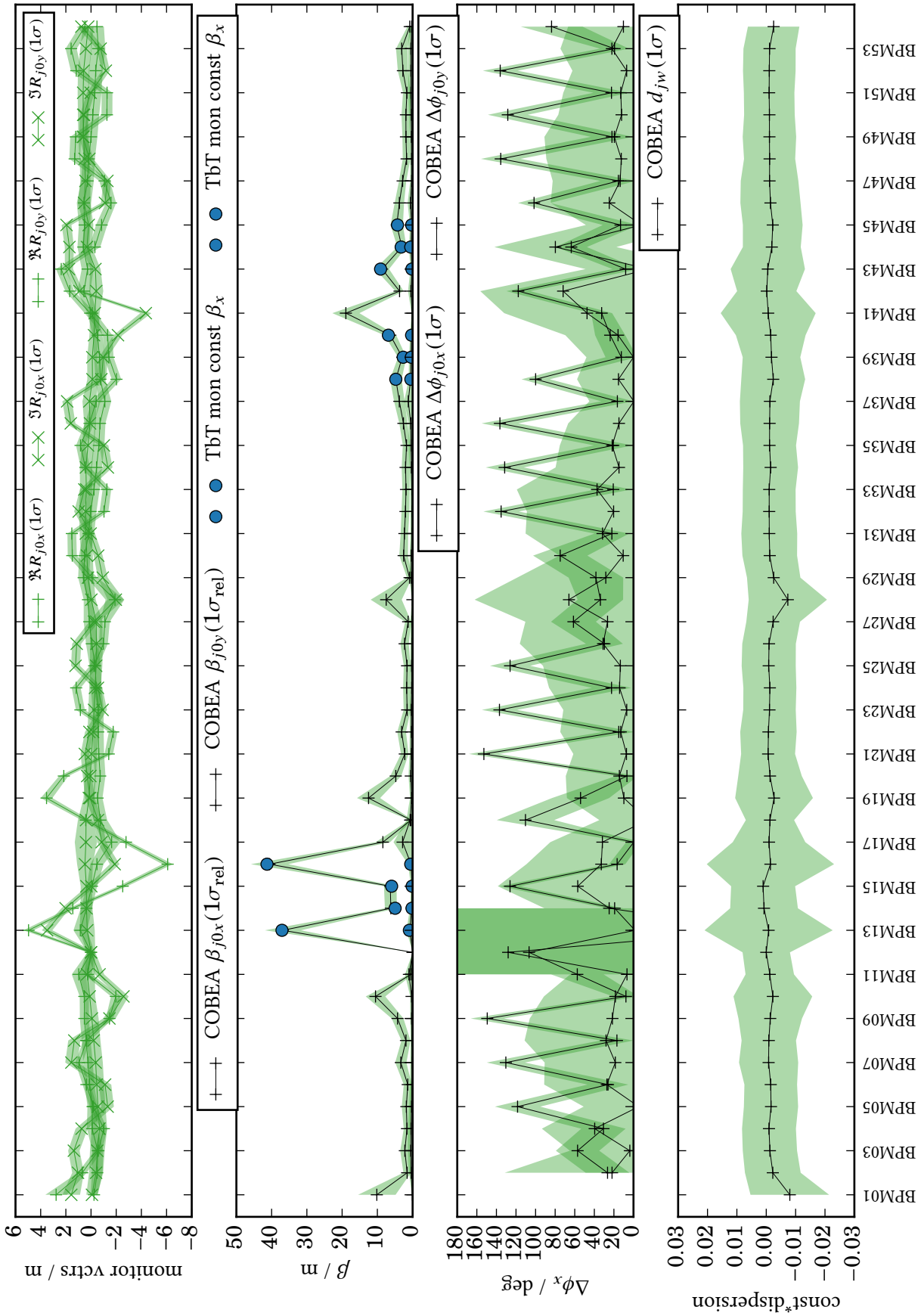

Figure 8.14.: Scaled TbT-predicted Mais-Ripkenβ functions and COBEAs prediction for TID 22 in mode $m = 1$.

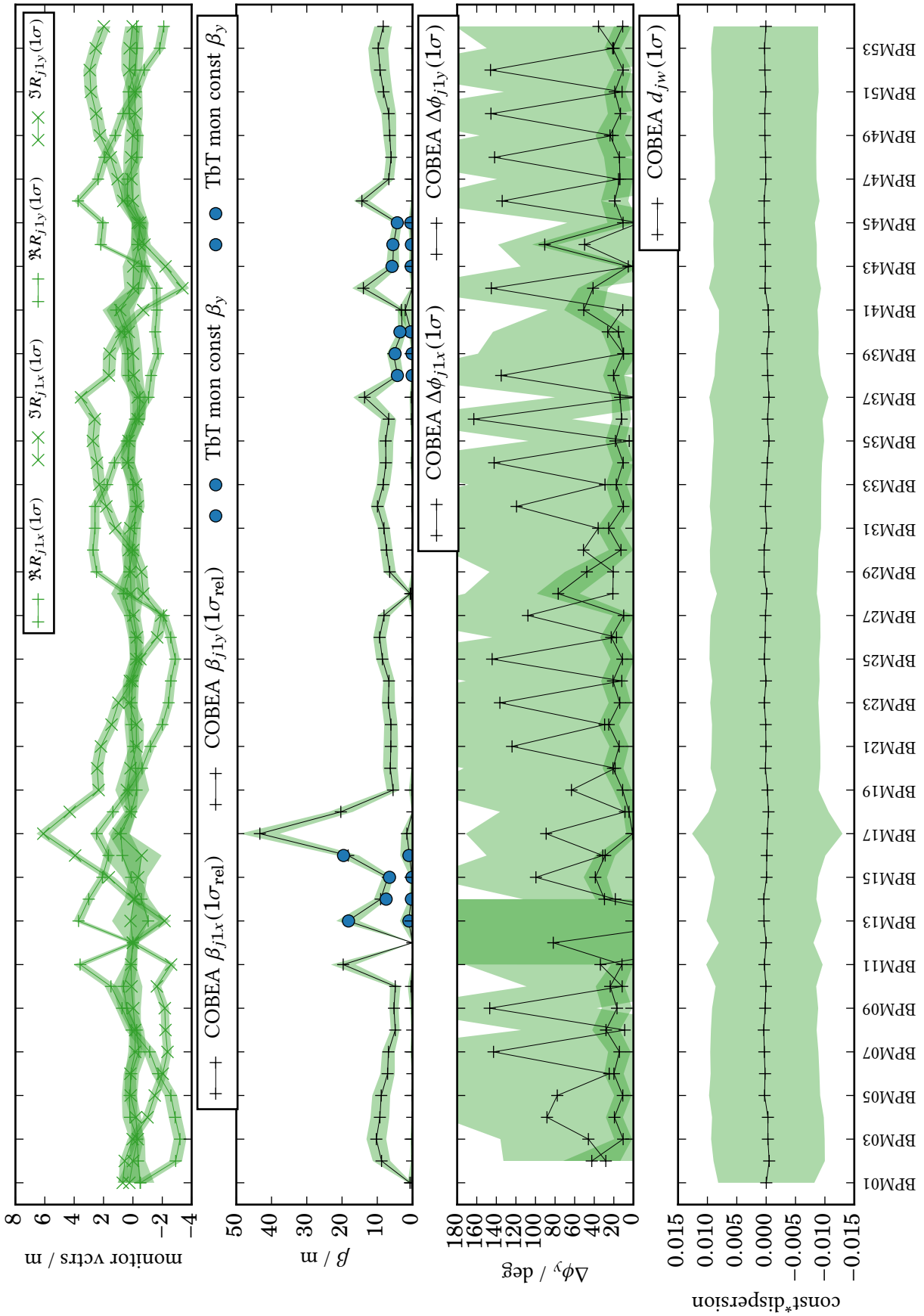

Figure 8.15.: Scaled TbT-predicted Mais-Ripkenβ functions and COBEAs prediction for TID 22 in mode $m = 2$.

9. COBEA evaluation with HZB data

After having validated the COBEA algorithm against TbT data from the DELTA storage ring, we would like to compare its results with those of the related fitting routine LOCO. LOCO generates the most information in comparison to the other validation techniques used in this thesis. This technique does not converge sufficiently for the DELTA storage ring due to large differences [74] between the real storage ring's magnetic fields and its simulation.

Fortunately, response matrix data for two accelerators installed and operated by Helmholtz-Zentrum Berlin (HZB) – the Metrology Light Source (MLS) and BESSY II – has been provided [75]. For both storage rings respectively their simulations, LOCO does converge. It is emphasized that the detailed magnetic fields (or, equivalently, the complete accelerator lattice with all magnet strengths) needed for an accelerator simulation are unknown to the author. Only after preliminary runs and transmission of the results, the author had access to detailed optical information from successful LOCO application at HZB [75].

9.1. Comparison with LOCO at the Metrology Light Source (MLS)

The Metrology Light source is a facility run by the Physikalisch-Technische Bundesanstalt (PTB), Germany's national metrology institute, with a scope on metrology and research in the THz to extreme UV spectral range [76]. Of the storage rings considered in this thesis, MLS has the smallest circumference and is the most recent to be commissioned. Due to its size and steady advances in accelerator construction, we can assume that it is the most accurately set-up storage ring considered in this work. Further information about MLS is shown in Fig. 9.1.

Parameter	Value
Lattice structure	double bend achromat
Injection energy	105 MeV
Maximum energy	630 MeV
Circumference	48 m
Revolution frequency	6.25 MHz
RF frequency	500 MHz
Dipole bending radius	1.53 m
electron beam current	150 mA

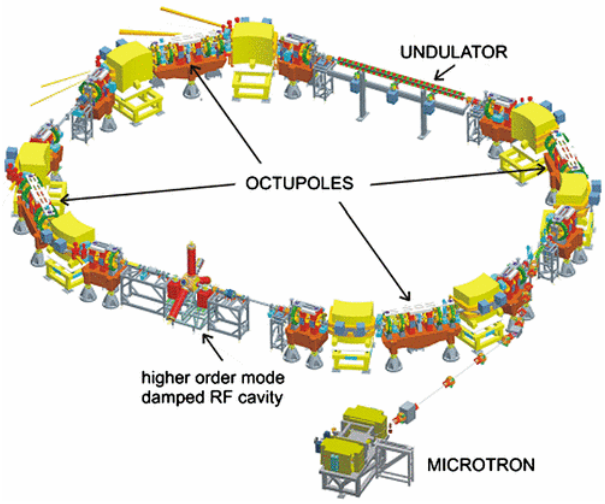

Figure 9.1.: Basic parameters and schematic view of MLS, both compiled from [76].

9.1.1. MLS input data

For the purpose of testing COBEA, the author was kindly provided [75] with input and validation data originating from the MLS control system. The input data is composed of

1. decoupled response matrices $\mathbf{x}$, $\mathbf{y}$ (sec. 4.3) in units $10^3\,\mathrm{m\,rad^{-1}}$.

   The matrices have been produced at HZB by fitting BPM data and magnet-dependent relations between applied current and kick angle for each corrector.

2. two lists of identifiers for each corrector respectively column of the matrix,
   `HS3M2K1RP, HS1M2K1RP, HS3M1L2RP, ...` for $\mathbf{x}$ (12 elements) and
   `VS3M2K1RP, VS2M2K1RP, VS2M1L2RP, ...` for $\mathbf{y}$ (16 elements).

   While the lists were used as given, they could easily be constructed from the first list by taking all elements starting with `HS` respectively `VS` in the respective order (matching the columns).

3. a list of identifiers for each monitor respectively row of both matrices,
   `BPMZ5K1RP, BPMZ6K1RP, BPMZ7K1RP, ...` (28 elements)

   Again, this list can be constructed from the first identifier list by taking all entries starting with `BPM`.

4. two lists of all monitor and corrector identifiers with the respective elements ordered by increasing s position along the beam path ("downstream"),

   ```
   BPMZ5K1RP, HS3M2K1RP, BPMZ6K1RP, BPMZ7K1RP, HS1M2K1RP, BPMZ1L2RP, ... (x-s plane)
   BPMZ5K1RP, VS3M2K1RP, BPMZ6K1RP, VS2M2K1RP, BPMZ7K1RP, BPMZ1L2RP, ... (y-s plane)
   ```

   with 40 respectively 44 elements. Both lists are printed in sec. A.4.4.

5. To obtain the invariants I_x, I_y, the author was supplied with the information that a drift space with the length 2.4068 m exists between the monitors `BPMZ4K3RP` and `BPMZ5K3RP`, and another drift space of length 5.9068 m exists between the monitors `BPMZ4L4RP` and `BPMZ5L4RP`. As the error of computed invariants reduces for longer drift space lengths, the author chose the second drift space information as input for COBEA.

From the items 2–4 of this list, the topology matrix $\mathbf{S}$ (sec. 3.2.2) can be constructed (Fig. 9.2). Note that although two ordered lists occur, the topology matrix is still unique as only relations between correctors and monitors and not between correctors are needed.

Validation data

After a few successful runs of a preliminary COBEA implementation (missing optimization layer) and exchange of results, the author was again kindly provided [75] with a validation data set from a run of Matlab-based LOCO [77] on the previously sent response data with a sufficient model, consisting of

1. Courant-Snyder parameters $\beta_{\mathrm{x,y}}(s_j)$ and $\phi_{\mathrm{x,y}}(s_j)$,
2. Betatron tunes $Q_{\mathrm{x,y}}$.
3. From additional sent files, the author also extracted the dispersion function $d(s_j)$.

9.1.2. Running COBEA on MLS data

Subsequently, COBEA was applied to the input data. COBEA's core routine, consisting of local and global subroutines, took approx. 20 s to converge on a typical modern PC. The optimization layer took 1251 iterations to converge with standard accuracy settings (Fig. 9.2). Note that after approx. 250 L-BFGS

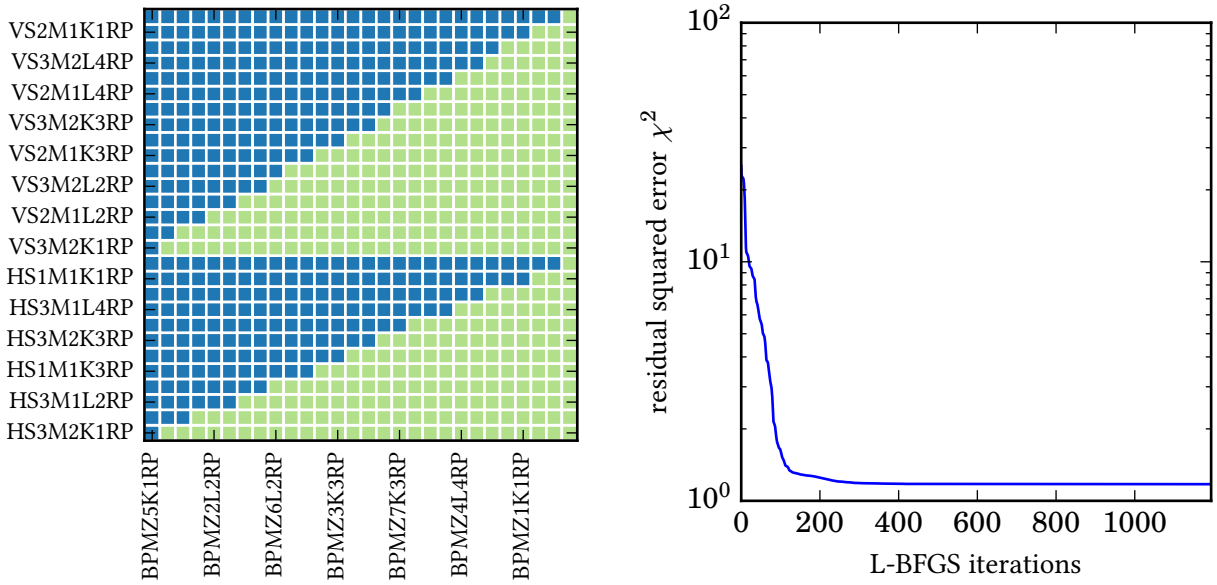

Figure 9.2.: Left: Topology matrix **S** constructed from MLS input data. Blue indicates negative sign, green indicates positive sign. Right: Residual fit error χ^2 for MLS data set in dependence to the number of L-BFGS iterations.

iterations, the residual error does not decrease significantly and an exponential decay to a finite χ^2 limit can be observed. This indicates that it might be possible to apply softer convergence criteria for the COBEA optimization layer and therefore accelerate it further.

Caution should be exercised in the interpretation of COBEA errors for MLS, as the overdetermination factor (sec. 4.2) is low due to the small number of monitors and correctors. Therefore, the error estimate is only a very coarse estimate.

9.1.3. Comparison for x mode

Convergence of the optimization can also be checked by the relative residual fit error of response matrix elements. In accordance with Fig. 9.3, the maximum of fit deviations is 1.7×10^{-2} of that from the response matrix deviations. For the ratio of root mean-squared values, this value is 8.1×10^{-3} (see also Tab. 9.1). The largest fit residuals are located at BPMZ5K1RP (first monitor), BPMZ2K3RP and BPMZ2L4RP.

Starting the comparison of optical parameters (Fig. 9.4) with the betatron phase advances per monitor $\Delta\phi$, one can see that the phase advances predicted by LOCO and COBEA are all in the limit of the error predicted by COBEA and seem almost identical in the plot.

Comparing β functions generated by COBEA and LOCO, one can see that they essentially agree, although COBEAs error margins are a bit too small. It is noteworthy that, while the LOCO values are symmetric around BPMZ4K3RP and BPMZ5K3RP up to small deviations, COBEA has optimized an asymmetric β solution. As no increased fit residual can be found for the two mentioned monitors, this may point to an actual small asymmetry in the magnetic fields of MLS at this time.

For x mode, there is non-negligible dispersion indicated by COBEA and LOCO. When scaling COBEAs dispersion prediction by a global factor to that of the LOCO model, one can find that there is a good agreement of the predicted shapes (Fig. 9.4).

The betatron tunes found by both methods agree very well within the error margin predicted by COBEA.

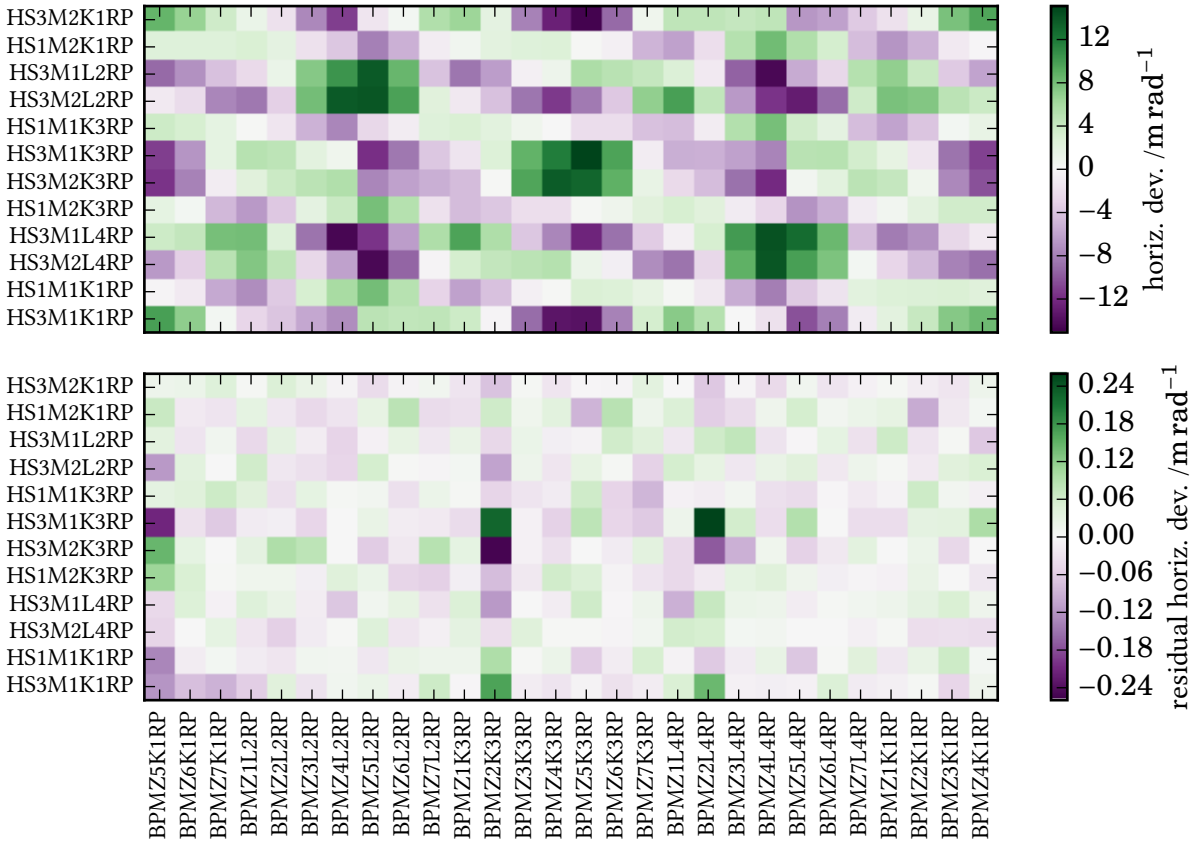

Figure 9.3.: Response matrix (top) and fit residual (bottom) for x mode of MLS.

Quantity	Variable	Value	Error	Unit
number of monitors	J	28		
number of correctors	K	28		
$-x$ mode		12		
$-y$ mode		16		
number of inputs	size($\mathbf{r}$)			
(decoupled resp. assumpt.)	$= JK$	784		
Search space dimensions	D	121+133		
L-BFGS function eval.s	`nf`	1181		
L-BFGS iterations	`it`	1137		
Fit error	$\chi^2(\vec{\rho}^{\,\mathrm{opt}})$	1.174		$(\mathrm{m\,rad^{-1}})^2$
$-x$ mode		0.858		$(\mathrm{m\,rad^{-1}})^2$
$-y$ mode		0.316		$(\mathrm{m\,rad^{-1}})^2$
Betatron tune (COBEA)				
$-x$ mode	$Q_\mathrm{x}^\mathrm{COBEA}$	3.17766	7.21×10^{-3}	
$-y$ mode	$Q_\mathrm{y}^\mathrm{COBEA}$	2.23114	6.28×10^{-3}	
Betatron tune (LOCO)				
$-x$ mode	$Q_\mathrm{x}^\mathrm{LOCO}$	3.17762		
$-y$ mode	$Q_\mathrm{y}^\mathrm{LOCO}$	2.23869		

Table 9.1.: Ring-global COBEA validation results for MLS.

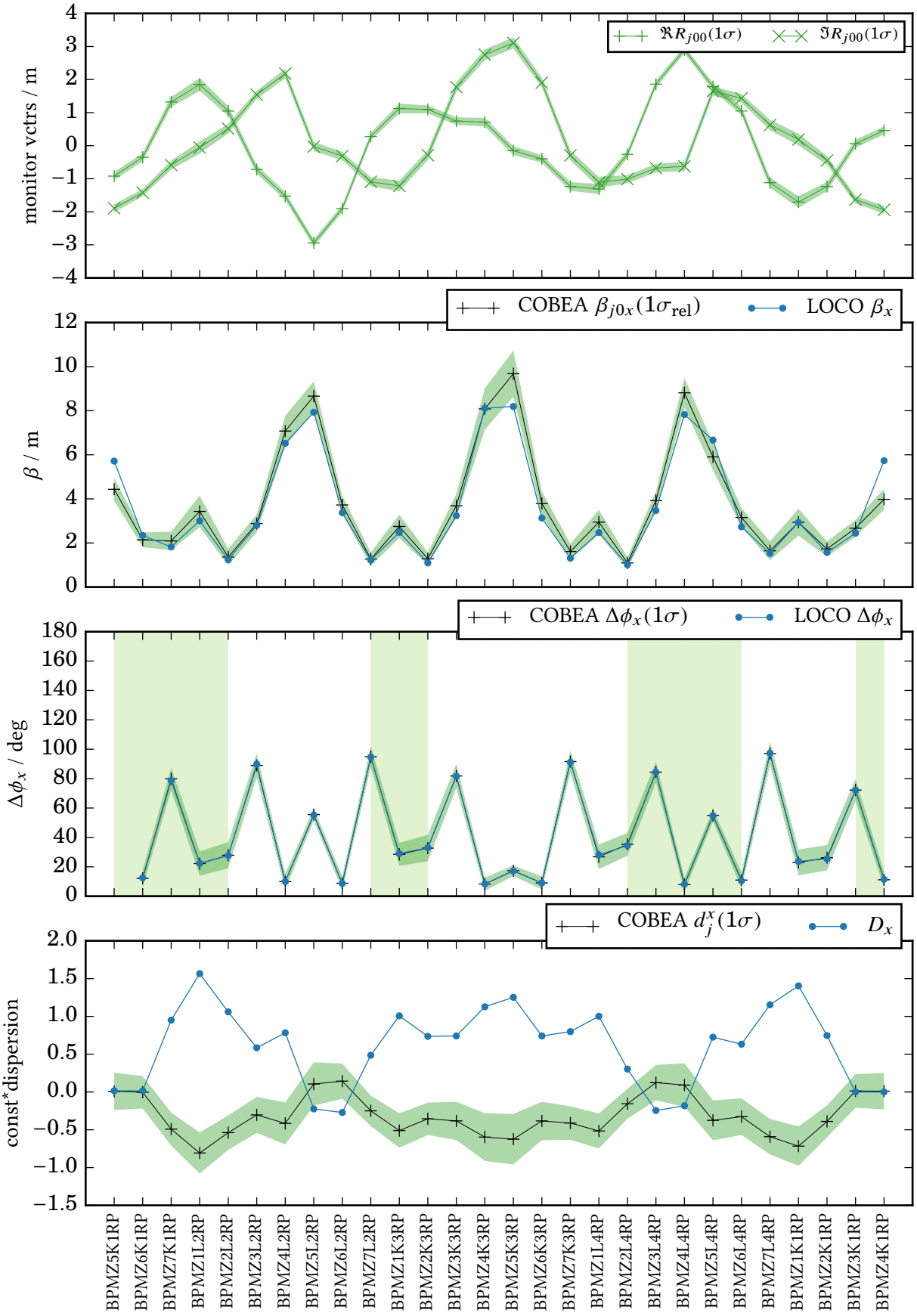

Figure 9.4.: MLS x mode monitor result plot for COBEA (green) and LOCO (blue).

9.1.4. Comparison for y mode

For y mode, an even better agreement of predicted and measured response matrices, measured by the fit residual, has been found (Fig. 9.5). The ratio of maximum values is 1.0×10^{-2}, with the rms ratio being 6.6×10^{-3}. No directly visible patterns remain in the fit residual matrix.

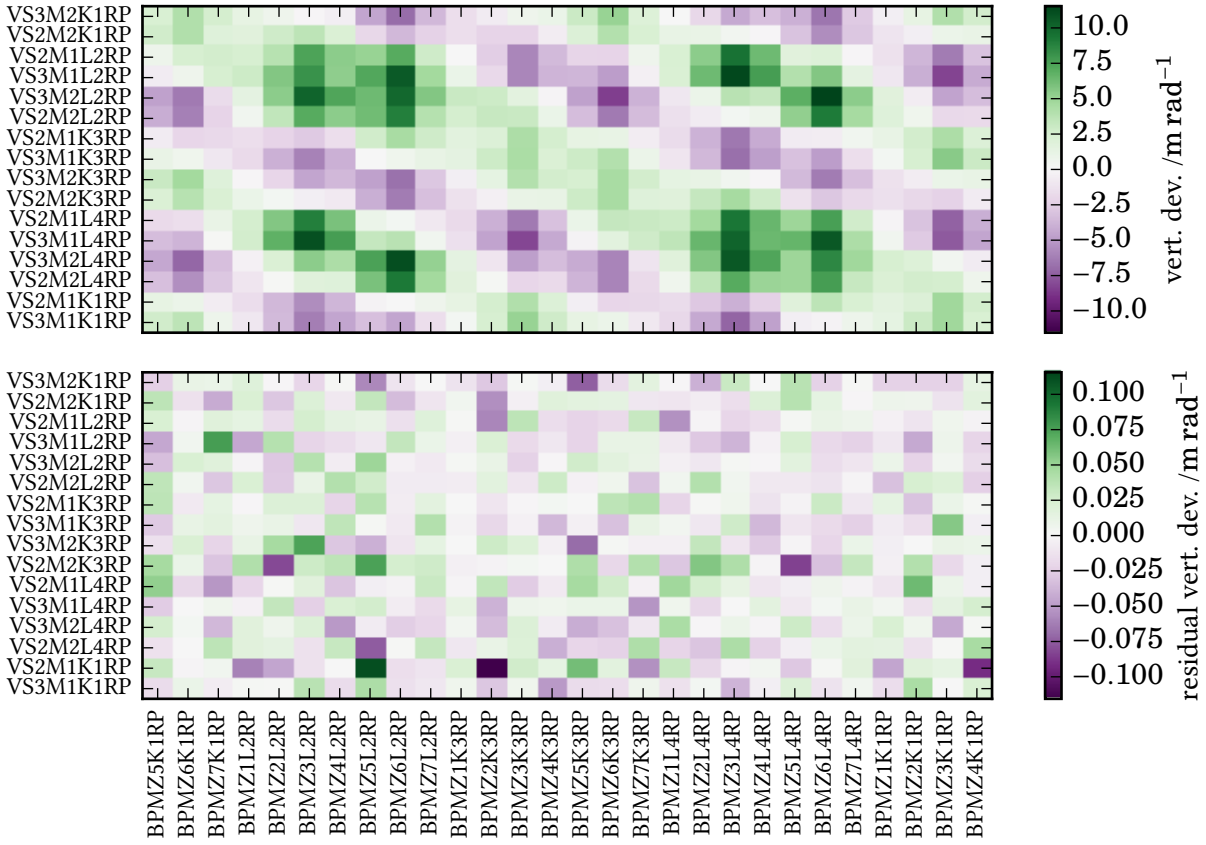

Figure 9.5.: Response matrix (top) and fit residual (bottom) for y mode of MLS.

Like for x mode, the betatron phase advances for y mode from COBEA (Fig. 9.6) agree very well with those of LOCO, which values are inside the error margins predicted by COBEA. Very small deviations can be found around the local maxima of phase advance.

A qualitative comparison of β functions shows a general agreement, although the values predicted by COBEA are larger on average. One explanation for this behavior are an invariant that is measured with significant errors – on the other hand, there is a good agreement for many monitors with smaller β values, and the deviations occur at the local maxima (with regard to positions equipped with monitors).

As to be expected for the vertical plane, no significant dispersion outside of the predicted error margins could be found.

Analyzing the betatron tunes, we find that the tune predicted by COBEA is significantly lower than the one predicted by LOCO. This is in consistence with the previously mentioned larger local maxima of β functions, as β and ϕ have an inverse relation to each other [6] and thus the phase advance is smaller for larger β functions. As these relations are not implied by the used BE+d model, both tune and larger β values indicate that the vertical modal phase advance of MLS is slightly weaker than predicted by LOCO.

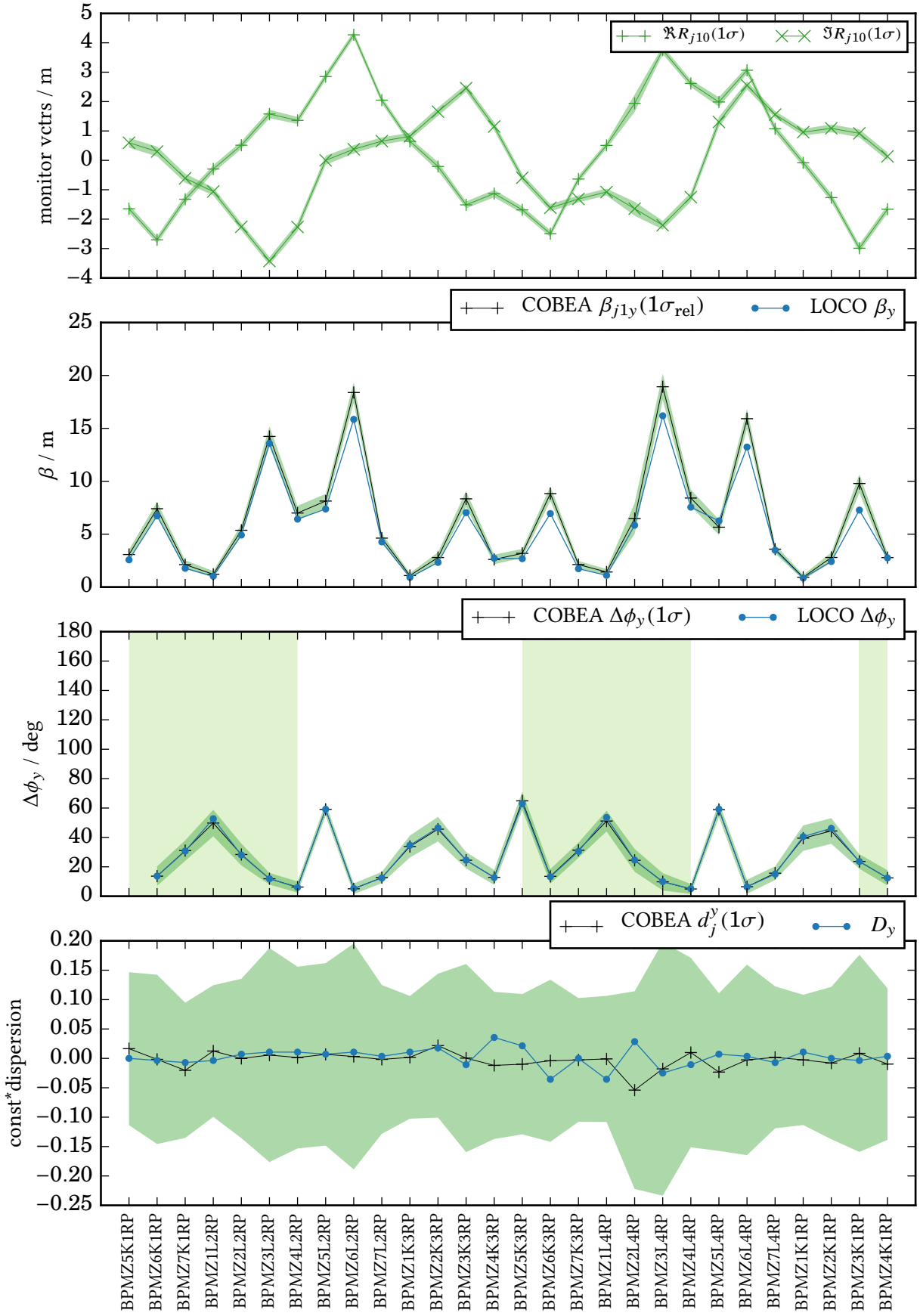

Figure 9.6.: MLS y mode monitor result plot for COBEA (green) and LOCO (blue).

9.2. Comparison with LOCO at BESSY II

BESSY II is classified as a 3rd generation synchrotron radiation source and located in Berlin-Adlershof, Germany. Its > 50 beamlines provide users with radiation from a 1.7 GeV beam that is pre-accelerated by a microtron and a synchrotron with 96m circumference [78]. Further information about the BESSY II storage ring, which will be discussed in the following, is given in Fig. 9.7.

Parameter	Value
Lattice structure	double bend achromat[79]
Beam energy	0.9 GeV −1.9 GeV [78]
Nominal energy	1.72 GeV [78]
Circumference	240 m [79]
Revolution frequency	1.25 MHz[79]
RF frequency	500 MHz [79]
Dipole bending radius	4.355 m [80]
max. beam current	0.3 A [79]
nominal beam current	0.2 A [78]

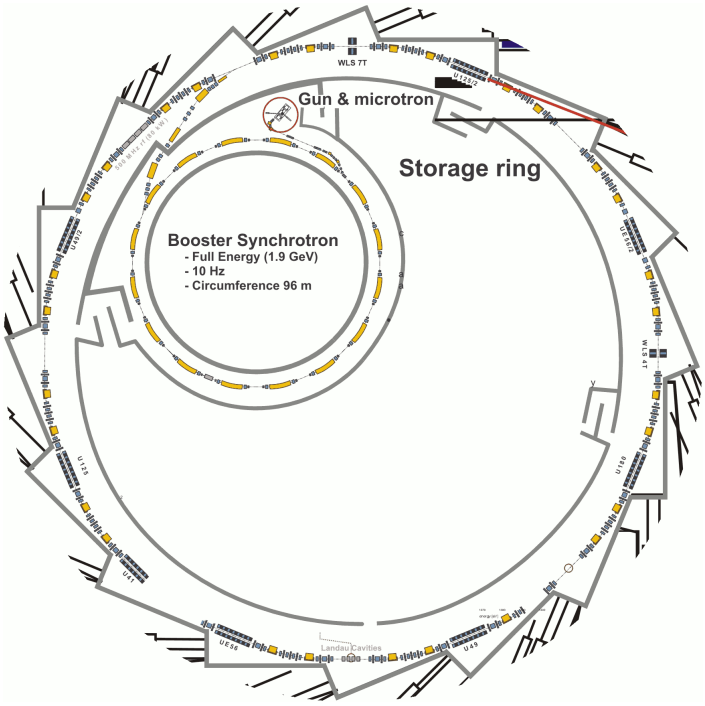

Figure 9.7.: Basic parameters of BESSY II storage ring (left) and sketch of the inner accelerator hall based on [78] (right).

9.2.1. BESSY II input data

Again the author was kindly provided [75] with input and validation data from the BESSY II control system. The input data has the same format as that for MLS (sec. 9.1.1), with the following changes in element topology and drift space information.

2. two lists of identifiers for each corrector respectively column of the matrix,
   `HS4M2D1R, HBM2D1R, HS1MT1R, HBM1T1R, ...` for **x** (80 elements) and
   `VS3M2D1R, VS2M2D1R, VS2M1T1R, VS3M1T1R, ...` for **y** (64 elements).
3. a list of identifiers for each monitor respectively row of both matrices,
   `BPMZ6D1R, BPMZ7D1R, BPMZ1T1R, BPMZ2T1R, ...` (108 elements)
4. two lists of all monitor and corrector identifiers with the respective elements ordered by increasing s position along the beam path,

   ```
   HS4M2D1R, BPMZ6D1R, HBM2D1R,  BPMZ7D1R, HS1MT1R,  BPMZ1T1R, ... (x-s plane)
   VS3M2D1R, BPMZ6D1R, BPMZ7D1R, VS2M2D1R, BPMZ1T1R, VS2M1T1R, ... (y-s plane)
   ```

   with 188 respectively 172 elements. Both lists are printed in sec. A.4.5.
5. To obtain the invariants I_x, I_y, the author was supplied with the information that drift spaces with lengths of 4.722 m exist between the monitors `BPMZ4T`q`R` and `BPMZ5T`q`R`, where $q \in \{1, 2, 7\}$. As only one drift space is required, the author chose $q = 7$ for the following computations.

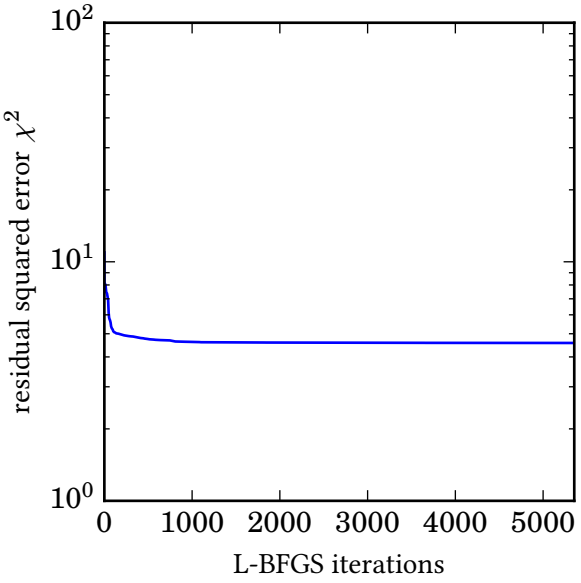

Figure 9.8.: Residual fit error χ^2 for BESSY II data set in dependence to the number of L-BFGS iterations.

The validation data provided also follows the same format as given in 9.1.1.

9.2.2. Running COBEA on BESSY II data

COBEAs optimization layer took ≈ 150 s to converge on an average PC. The convergence data and ring-global results are shown in Fig. 9.8 and Tab. 9.2.

Quantity	Variable	Value	Error	Unit
number of monitors	J	108		
number of correctors	K	144		
$-x$ mode		80		
$-y$ mode		64		
number of inputs	size($\mathbf{r}$)			
(decoupled resp. assumpt.)	$= JK$	15552		
Search space dimensions	D	565+517		
L-BFGS function eval.s	`nf`	5067		
L-BFGS iterations	`it`	4785		
Fit error	$\chi^2(\vec{\rho}^{\,\mathrm{opt}})$	4.57477		$(\mathrm{m\,rad^{-1}})^2$
$-x$ mode		4.20510		$(\mathrm{m\,rad^{-1}})^2$
$-y$ mode		0.36967		$(\mathrm{m\,rad^{-1}})^2$
Betatron tune (COBEA)				
$-x$ mode	$Q_\mathrm{x}^\mathrm{COBEA}$	17.84740	2.87×10^{-3}	
$-y$ mode	$Q_\mathrm{y}^\mathrm{COBEA}$	6.74054	3.55×10^{-3}	
Betatron tune (LOCO)				
$-x$ mode	$Q_\mathrm{x}^\mathrm{LOCO}$	17.84690		
$-y$ mode	$Q_\mathrm{y}^\mathrm{LOCO}$	6.74484		

Table 9.2.: Ring-global COBEA validation results for BESSY II.

9.2.3. Comparison for x mode

Considering the provided response matrix (Fig. 9.9), one can observe the matrix elements with relation to BPMZ4D2R being large in comparison to the other elements. This is due to a fit error that has occured during creation of the input response matrix [75]. It is noteworthy that nevertheless COBEA can fit the corresponding deviations well, and no large fit residual remains. The most reasonable explanation for this behavior is that the mentioned fit error is just a global scaling error of all elements of the monitor, possibly caused by a false effective monitor coefficient C_x. This would also explain why the large deviation does not occur for y mode. The largest fit residuals for x mode occur at elements involving BPMZ2T4R and corrector HS1MT2R.

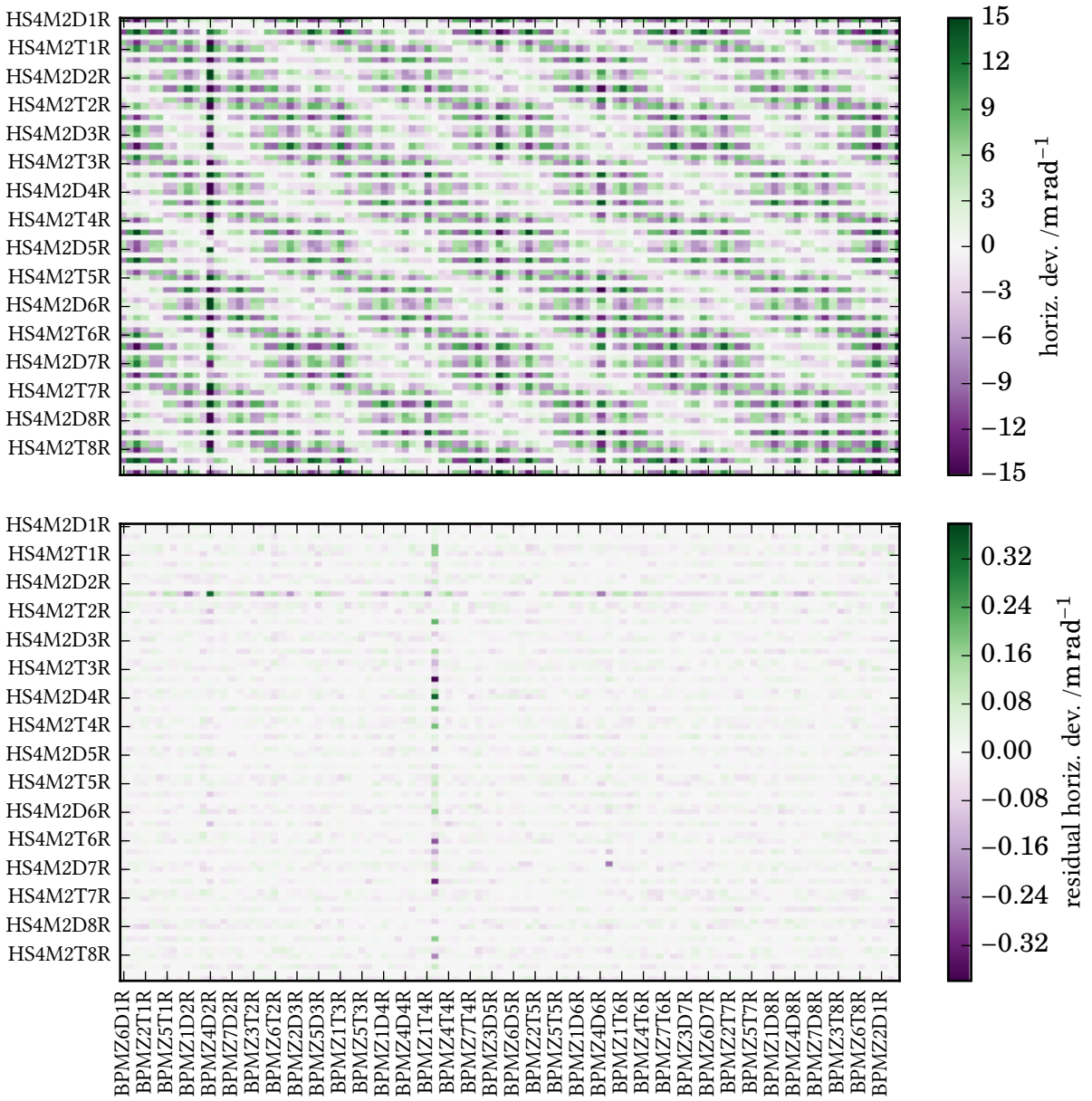

Figure 9.9.: Response matrix (top) and fit residual (bottom) for x mode of BESSY II.

Comparing betatron phase advances per monitor (Fig. 9.10), one can observe that the predictions of COBEA and LOCO qualitatively agree, although differences up to ±10° can be found, concentrating at

the larger phase advances. Considering the low error margins of phase advances, it can be supposed again that a LOCO fit parameter that is identical at a set of many magnetic elements, e.g. absolute identical lattice cells, limits the LOCO solution to those assumed symmetries. Also note that the phase advances for the wrong-fitted monitor BPMZ4D2R show good agreement to the phase advances predicted by LOCO, which reinforces the monitor coefficient hypothesis.

For β functions, there is a rough qualitative agreement, although the deviations between LOCO and COBEA are much larger than in the MLS results (sec. 9.1.3). As the deviations between both predictions have different signs, the effect can not be attributed solely to possible errors in COBEAs computation of the invariant of motion. Instead, the small COBEA error margins for relative β errors may indicate that the actual optics of BESSY II slightly deviate from the ideal assumptions of its lattice model.

As expected for x mode, non-negligible dispersion effects occur in the response matrix data. Although COBEA in its present implementation and without cavity phase data can only fit dispersion factors up to a global scaling factor, there is a good agreement on the dispersion shape and the positions of maximum dispersion. Note that, as the dispersion coefficients have been extracted from the author by hand from additional information [75], the assignment of monitor labels to dispersion values is flawed, which could explain the deviations occurring between the start monitor BPMZ6D1R and monitor BPMZ1D4R (at monitors BPMZ7T1R and BPMZ7T3R).

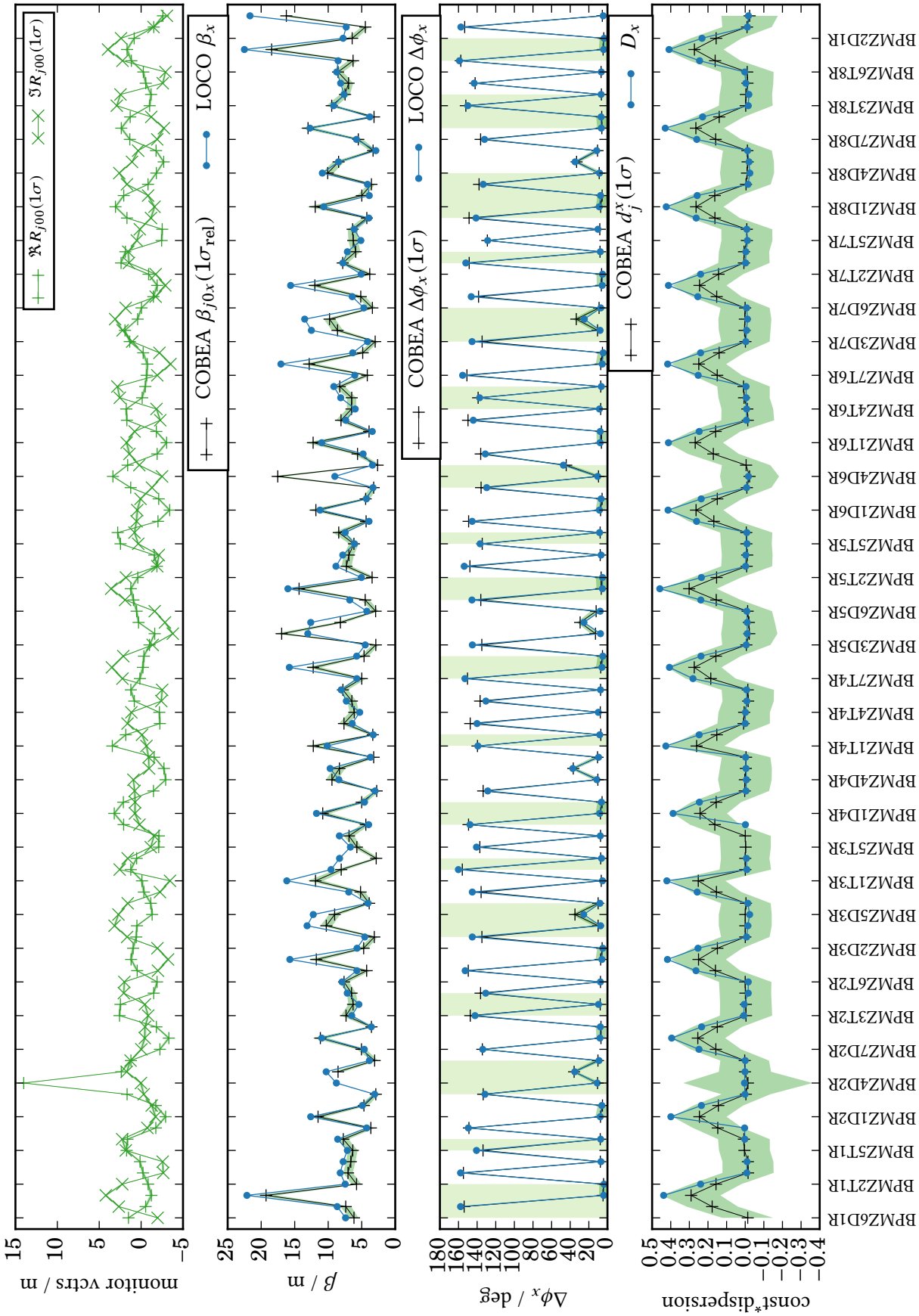

Figure 9.10.: BESSY II x mode monitor result plot for COBEA (green) and LOCO (blue). Corresponding to the BPMZ4D2R input response error (sec. 9.2.3), the β_x value for this monitor has been removed.

9.2.4. Comparison for y mode

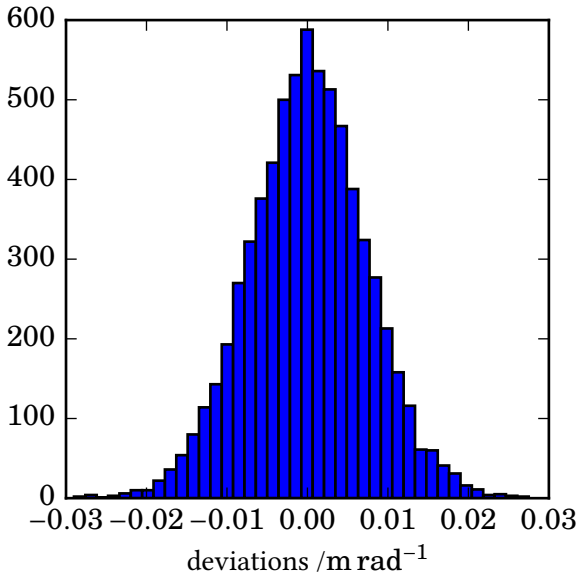

Figure 9.11.: Histogram of the fit residual matrix components for y mode of BESSY II. Up to statistical deviations, the distribution has a Gaussian shape.

The fit residuals for y mode of BESSY II (Fig. 9.12) are small compared to those for x mode, as was the case for MLS. Here, the ratio of global maxima is 4.4×10^{-3} with the rms ratio being 2.3×10^{-3}. Again there are no visible patterns in the residual fit matrix, and a histogram of all elements shows a Gaussian distribution (Fig. 9.11). Thus the elementary necessary conditions for the hypothesis that COBEA has actually fit the complete y mode response matrix signal up to Gaussian noise are met.

In the result BE+d model and the corresponding optical functions (Fig. 9.13), there is a good agreement between phase advances predicted by LOCO and COBEA. As the BE+d model has been used in the vertical plane, where almost no dispersive effects occur in the response matrix, at least one superfluous degree of freedom remains. This could be an explanation for the seemingly arbitrary ambiguity expressed by the high σ levels but low response residuals for monitor BPMZ2T4R. On the other hand, the same monitor has large residuals for the optimization of the X mode.

The largest deviations become visible at the maximum phase advances. As the fit error is very low and the average phase advance error per length should be approximately constant, this is in consistence with small deviations of real and simulated magnetic fields.

Comparing β functions, one can see that there is a systematic deviation of COBEA with regard to LOCO as the predicted values of the former are larger on average. While this could point to an invariant scaling error (due to inaccurate monitor vectors in the monitor drift space), the overall small error of the monitor vectors contradicts this assumption. Like for MLS (sec. 9.1.4), the larger β values should correspond to a lower betatron tune instead.

This hypothesis is also reinforced by a comparison between the predicted betatron tunes (Tab. 9.2), with their difference being slightly larger than 1σ according to COBEAs prediction. Considering the low fit error of COBEA, we may assume that the lower tune is closer to reality.

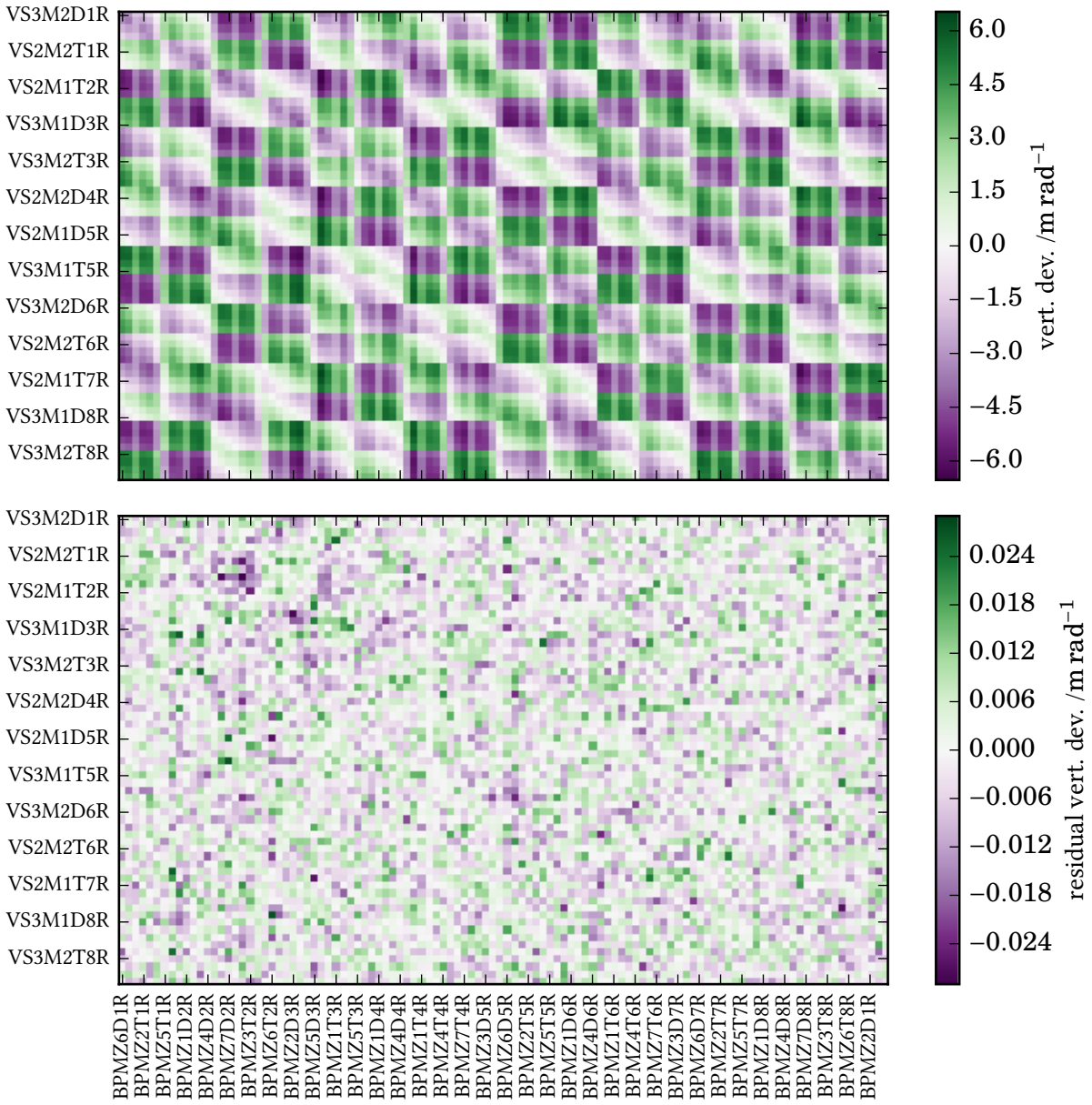

Figure 9.12.: Response matrix (top) and fit residual (bottom) for y mode of BESSY II.

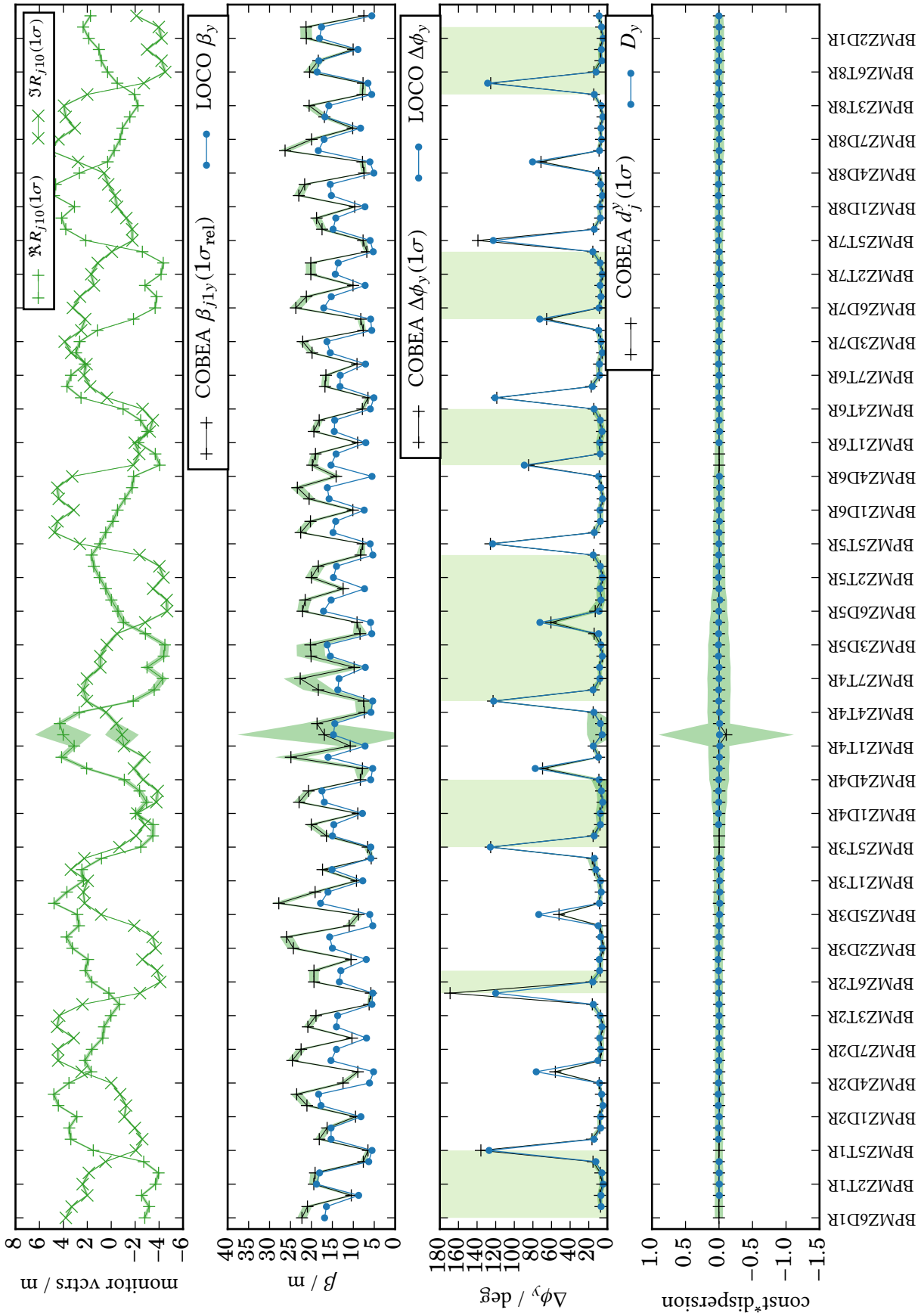

Figure 9.13.: BESSY II y mode monitor result plot for COBEA (green) and LOCO (blue).

10. Conclusion and Outlook

Measurement of closed orbit perturbations respectively response matrices is a technique used for the majority of accelerator storage rings in existence. While these matrices are often used for orbit correction tasks and thus exist in abundance, their direct relation to beam optical parameters had not been fully exploited.

In this thesis, the underlying inverse problem of extracting optical parameters from a response matrix without superfluous or special assumptions and/or additional measurement data has been treated. We have found a linear, transverse-coupled description model, the bilinear-exponential model with dispersion (BE+d model) for closed-orbit perturbations in storage rings. This model uses comparatively few and common assumptions and is the cornerstone for all other theoretical results described in this thesis.

Based on this derivation of a model for optical parameters (forward modeling), a procedure called closed orbit bilinear-exponential analysis (COBEA) has been implemented to decompose measured response matrices into BE+d model parameters (inverse modeling) using an optimization procedure. The algorithm to obtain the start values for this nonlinear problem incorporates and extends ideas used for existing measurement procedures in linear accelerators and storage rings. The optimization layer uses analytical gradient information in search space, which significantly accelerates the convergence time of COBEA. This is facilitated by using the closed-form expressions of the BE+d model and would be impossible by use of numerical particle tracking.

COBEA has been applied to response matrices from the storage rings of DELTA (chapters 7 & 8), MLS (sec. 9.1) and BESSY II (sec. 9.2). In each case, additional data from complementary diagnostic methods (either Turn-by-Turn data or LOCO) was available so that different parts of COBEA results could be validated against their results. Taking into account the discussed and found error sources of these existing procedures and the measurement of response matrices, it could be shown that COBEA converges successfully.

Thus all tasks in the scope of this thesis (sec. 1.1) have been completed and the inverse problem of decomposing general response matrices into beam optics information can be considered as solved under the given assumptions. Due to the low number and weakness of these underlying assumptions and the availability of the necessary input data, COBEA should be applicable to a large number of existing storage rings.

This chapter closes with a short discussion about possible extensions of the COBEA method.

Orbit correction with COBEA-cleaned response matrices

In operating accelerators, unintentional closed orbit perturbations $\vec{r}^{\,\mathrm{unint}}$ may be generated by imperfectly manufactured and/or installed magnetic elements. These closed-orbit perturbations can depend on magnet currents, temperature and time in general, on the time scale of minutes.

Based on measurement of a current-response matrix $\mathbf{r}$, several techniques exist to counteract these undesirable closed-orbit perturbations by shifting the closed orbit to its original position. If no model is included in the correction process, this involves solving the equation system

$$-\vec{r}^{\,\mathrm{unint}} = \mathbf{r}\vec{I},$$

sometimes including additional boundary conditions based on the accelerator setup (like for limited corrector currents available at DELTA [60]).

Any orbit correction by virtue of an accelerator model results in a "cleaning" of the response matrix, which means its effective replacement by a response matrix from the optimized simulation, e.g.

$$-\vec{r}^{\,\mathrm{unint}} = \mathbf{r}^{\mathrm{sim}}(\mathrm{LAT})\vec{I}.$$

For LOCO (sec. 6.2.1), this is done implicitly by computations in the accelerator simulation.

This cleaning procedure is also possible using a Bilinear-Exponential model which is optimized by COBEA,

$$-\vec{r}^{\,\mathrm{unint}} = \mathbf{f}\vec{I},$$

with the coefficients f_{jkw} being generated from optimized BE+d parameters (sec. 4.2).

As the response matrix is overdetermined in relation to BE+d parameters, using COBEA-cleaned response matrices results in a lower noise level of orbit correction. As we only replace the matrix used by the respective orbit correction, no further changes to existing orbit correction applications that would use directly measured response matrices otherwise are necessary.

Online analysis and subspace iteration using COBEA

The approach used in this work for decomposing measurement data is essentially an offline analysis, as it uses response matrices or dedicated machine shifts to obtain BE+d coefficients using COBEA. As the solution of the underlying system by the MCS algorithm (sec. 5.2) is also feasible with a relatively low number of correctors ($K \geq 4M$), it is also possible to obtain a reduced BE+d model with an identical number of monitor vectors J. Depending on K and additional assumptions, dispersion can be included or neglected in the analysis.

A special implementation of COBEA tailored to the possibilities and requirements of the DELTA storage ring is pursued in [68].

For example, it is possible to benefit from the fact that the TbT tune is known with high accuracy in standard operation, so that the μ_m parameters do not need to be optimized in principle. As the resulting optimization problem is now not bilinear-exponential, but only bilinear, subspace iteration schemes that map monitor to corrector parameters and vice versa might be used to replace the optimization layer for DELTA. For the dispersion free case, this would be equivalent to using CM mapping (chapter 5.1) subsequently for a large number of iterations and tracking the residual error until a convergence criterion is met.

Magnetic modeling

Closing the gap between known monitor vector at monitors and full optical information, one can recognize that small deviations around the beam path can be linearized, as has been one during the definition of eigenorbits $\vec{R}_m(s)$.

All linear magnetic fields can be described by magnets up to quadrupolar order. In fact, as the closed orbit is defined by dipolar fields, only the quadrupolar / focusing fields remain, with different focusing in different directions. The eigenorbit solutions are free of dispersive effects by definition. Thus, eigenorbits are governed by the homogeneous Hill differential equation [81]

$$\frac{\mathrm{d}^2}{\mathrm{d}s^2}\vec{R}_m(s) + \mathbf{K}(s)\vec{R}_m(s) = 0.$$

In an accelerator, we may assume the real-valued focusing term $\mathbf{K}(s)$ to be piece-wise constant along s, e.g. inside a given magnet $s \in [s_l^{\mathrm{start}}, s_l^{\mathrm{end}}]$ indexed by l, $\mathbf{K}(s) = \mathbf{K}_l$. Then, the phase-space transfer matrices between start and end positions of the magnet

$$\begin{pmatrix} \vec{R}_m(s_l^{\mathrm{end}}) \\ \vec{R}'_m(s_l^{\mathrm{end}}) \end{pmatrix} = \mathbf{M}_l \begin{pmatrix} \vec{R}_m(s_l^{\mathrm{start}}) \\ \vec{R}'_m(s_l^{\mathrm{start}}) \end{pmatrix}$$

can be found by elementary means, and are essentially transfer matrices of extended quadrupole magnets with arbitrary rotation.

Let us assume that the eigenorbit is known at positions s_j by COBEA, and a known segment exists between monitors f and $f+1$. Then, the magnet elements between them must connect the linear in a steady and differentiable fashion. Finding the linear magnet properties $\mathbf{K}_l$ around the beam for a given set of monitor vectors from COBEA then reduces to finding a least squares solution of ($\mathbf{D}_j$: distortion rectification of monitor vector)

$$\vec{R}_{jm} = \begin{pmatrix} \mathbf{D}_j & 0 \end{pmatrix} \prod_{l}^{l \in [s_f, s_j]} \mathbf{M}_l(\mathbf{K}_l)\, \mathbf{P}_f \begin{pmatrix} \vec{R}_{fm} \\ \vec{R}_{(f+1)m} \end{pmatrix} \text{ for all } j.$$

This is essentially an optimization procedure that has been applied in a similar fashion for linear accelerators and unperturbed storage ring segments in [82] and by the code RESOLVE [83]. By using COBEA eigenorbits as inputs, one can use this approach to model a storage ring.

One can observe that the optimization only depends on the magnetic properties between the known segment and the monitor j considered. One can thus start the optimization in a low parameter space for a few magnets between the first monitors after f and after convergence add more and more monitors and respective magnets, essentially doing a dispersion-free linear fit of magnet elements in a linac.

As the matrices $\mathbf{M}_l$ are known analytically in linear approximation, it is possible to compute the gradient of this optimization problem. Naturally, the additional information where elements start and end is required for this approach, and uniqueness of the solution cannot be guaranteed in general, as it depends on the degrees of freedom $\mathbf{K}_l$ in relation to the monitor vectors and their position towards elements l.

Possible inclusion of nonlinear motion

In this work, the discussion has been limited to a linearization of forces around an existing closed orbit, and it could be shown that the Bilinear-Exponential model can explain the measured closed orbit deviations with a relative error $\leq 2 \times 10^{-2}$, depending on the setup. If a non-linear extension of COBEA is desired, e.g. to quantify anharmonic effects it is thus mandatory to measure with larger excitations respectively corrector currents than those used in this work, leaving the linear regime.

For analysis of the resulting nonlinear response matrices, it is possible to use a modified start-value layer (sec. 6.1), where the linear transfer matrices $\mathbf{T}_{\mathrm{A}}, \mathbf{T}_{\mathrm{B}}$ are replaced by nonlinear transfer maps $\mathcal{T}_{\mathrm{A}}, \mathcal{T}_{\mathrm{B}}$. For parameterization, we will discuss a polynomial approach[1] in the following.

Polynomial mappings $\mathcal{T}_\bullet$ can be derived from measurement data as solutions of linear equation systems. It is also possible in principle to use PCA cleaning techniques (sec. 5.2.3), although the cleaned orbits should be used explicitly, not implicitly using principal orbits, as the superposition principle does not hold.

It is possible to obtain finite symplectic expressions that preserve invariants of motion for a given order. For the quadratic order $P = 2$, such expressions can be found in e.g. [86].

[1] The use of Lie algebra [84] and resonance driving terms [85] might be equally or even better suited for this type of problem.

A problem occurs when connecting both maps. Due to the polynomial parameterization, a product of two matrices with order P will have an order of P^2 (possibly with some coefficients missing). Therefore we either have to neglect coefficients with higher order than P or use symplectic expressions for the order P^2. In the following, we neglect the coefficients so that all maps have the same order.

Then it is feasible in principle to build a model, starting from a known or unperturbed segment $[s_{j_1}, s_{j_2}]$, that is built using the parameters of the symplectic generator function for that order. This nonlinear model then replaces the BE+d model in the optimization layer (sec. 6.1.1). From its parameters, one could then derive the desired quantities by simple computations, or at least by tracking. All linear parameters and thus all eigenorbits could then also be reconstructed by simple computations on the parameters of this nonlinear model.

General Floquet-periodic systems

While the only physical system found by the author to be described by the BE±d model is closed-orbit perturbations in storage rings, our considerations essentially apply to any linear (by Jacobian) system with periodic boundary conditions in which local translations respectively "kicks" occur. In such systems, Floquet normal forms will describe the oscillation solutions, and solutions described in sec. 3.1 will occur as system-periodic static solutions. In any system in which the solution is monitored along the dimension of periodicity, and the translations occur at regions between the monitors, the BE±d model and COBEA are applicable.

Bibliography

If not stated otherwise, all internet sources and weblinks in this bibliography have been last accessed on 2016-04-07. Non-interactive HTTP hyperlinks are aligned to the right, also emphasizing Digital Object Identifier (DOI) and arXiv links.

[1] Y. Zou *et al.*, "Resonant slow extraction in synchrotrons by using anti-symmetric sextupole fields" (2016-01) arXiv.org/abs/1601.03113v1

[2] M. Ries, "Nonlinear Momentum Compaction and Coherent Synchrotron Radiation at the Metrology Light Source", Ph.D. Dissertation (Humboldt University of Berlin, 2013-10) nbn-resolving.de/urn:nbn:de:kobv:11-100217924

[3] A. Tarantola, *Inverse Problem Theory and Methods for Model Parameter Estimation*, 1st ed. (SIAM: Society for Industrial and Applied Mathematics, 2005), ISBN 978-0898715729 www.ipgp.fr/~tarantola/Files/Professional/SIAM

[4] A. Franchi, "Error analysis of linear optics measurements via turn-by-turn beam position data in circular accelerators" (2016-03) arXiv.org/abs/1603.00281v2

[5] S. Peggs, "Nonlinear Diagnostics using AC Dipoles", Proc. PAC 1999, TUP42, New York, USA (1999) accelconf.web.cern.ch/accelconf/p99/PAPERS/TUP42.PDF

[6] K. Wille, *The Physics of Particle Accelerators: An introduction* (Oxford University Press, 2000), ISBN 0198505507

[7] H. Wiedemann, *Particle Accelerator Physics*, 3rd ed. (Springer, 2007), ISBN 978-3-540-49043-2

[8] É. Forest, *Beam dynamics: a new attitude and framework* (Harwood Academic, 1998), ISBN 978-9057025747

[9] Lectures given at the CERN Accelerator School (CAS) course on "Beam Diagnostics", Dourdan, France (2008-06) cas.web.cern.ch/cas/France-2008/Lectures/Dourdan-lectures.htm

[10] E.D. Courant and H.S. Snyder, "Theory of the alternating-gradient synchrotron", Annals of Physics **3**(1), pp. 1–48 (1958-01) dx.DOI.org/10.1016/0003-4916(58)90012-5

[11] B. Riemann, P. Grete and T. Weis, "Model-independent and fast determination of optical functions in storage rings via multiturn and closed-orbit data", Phys. Rev. ST Accel. Beams **14**, 062802 (2011-06) dx.DOI.org/10.1103/PhysRevSTAB.14.062802

[12] C. Chicone, *Ordinary Differential Equations with Applications* (Springer, 1999), ISBN 0387985352

[13] Y. Luo, "Linear coupling parametrization in the action-angle frame", Phys. Rev. ST Accel. Beams **7**, 124001 (Dec 2004) dx.DOI.org/10.1103/PhysRevSTAB.7.124001

[14] H. Mais and G. Ripken, "Theory of coupled synchro-betatron oscillations", DESY Internal Report M-82-05 (1982-02) cds.cern.ch/record/137006

[15] F. Willeke and G. Ripken, "Methods of Beam Optics", sec. 3.4.2, DESY 88-114 (1988-08) cds.cern.ch/record/194174

[16] M.G. Minty and F. Zimmermann, *Measurement and Control of Charged Particle Beams* (Springer, 2003), ISBN 3-540-44187-5

[17] M.G. Minty and F. Zimmermann, "Beam Techniques – Beam Control and Manipulation", USPAS Lecture series script (1999-06) uspas.fnal.gov/materials/99UChicgo/beamtech.pdf

[18] K. Wille, personal communication (TU Dortmund University, 2015-05)

[19] R. Hedrich, *Die Entdeckung der Komplexität: Skizzen einer strukturwissenschaftlichen Revolution* (in German), (H. Deutsch, 1994), ISBN 9783817113835

[20] V.A. Lebedev and S.A. Bogacz, "Betatron motion with coupling of horizontal and vertical degrees of freedom", JINST **5**, P10010 (2010-10) arXiv.org/abs/1207.5526

[21] L.C. Teng, "Concerning n-dimensional coupled motions", FNAL Report FN-229 (1971-05) lss.fnal.gov/archive/fn/FN-0229.pdf

[22] D.A. Edwards and L.C. Teng, "Parametrization of Linear Coupled Motion in Periodic Systems", IEEE Trans. Nucl. Sci. **20** (3), pp. 885–888 (1973-06) dx.DOI.org/10.1109/TNS.1973.4327279

[23] B. Riemann *et al.*, "Two General Orbit Theorems for Efficient Measurements of Beam Optics", Proc. IPAC 2015, MOPWA035, Richmond, USA (2015-06) inspirehep.net/record/1417228

[24] C.R. Johnson, H. Šmigoc, and D. Yang, "Solution Theory for Systems of Bilinear Equations" (2013-02) arXiv.org/abs/1303.4988

[25] C. Cohen and S. Tomasi, "Systems of Bilinear Equations", Stanford University Technical Report CS-TR-97-1588 (1997) i.stanford.edu/pub/cstr/reports/cs/tr/97/1588/CS-TR-97-1588.pdf

[26] V. Blobel and E. Lohrmann, "Statistische und numerische Methoden der Datenanalyse" (in German), DESY library online version (2012-05) www-library.desy.de/elbook.html

[27] Y.S. Abu-Mostafa, "Learning From Data" Machine Learning online course (Caltech, 2012-04) work.caltech.edu/telecourse.html
Y.S. Abu-Mostafa, M. Magdon-Ismail and H.-T. Lin, *Learning From Data* (AMLBook, 2012-03), ISBN 978-1600490064

[28] G. Golub and S. Kahan, "Calculating the Singular Values and Pseudo-Inverse of a Matrix", SIAM J. Numer. Anal. B **2** (2), pp. 205–224 (1964-07) dx.DOI.org/10.1137/0702016

[29] M. Harrison and S. Peggs, "Global Beta Measurement from Two Perturbed Closed Orbits", Proc. PAC 1987, pp. 1105–1107, Washington, USA (1987) inspirehep.net/record/245622

[30] H. Koiso *et al.*, "Correction of optical functions in the KEKB rings", In: *International Workshop on Performance Improvement of Electron-Positron Collider Particle Factories*, Tsukuba, Japan, pp. 253–255 (1999-09) www-acc.kek.jp/kekb/Commissioning_in_PDF/cor.pdf cds.cern.ch/record/871173

[31] K. Pearson, "On Lines and Planes of Closest Fit to Systems of Points in Space", Philos. Mag. **2** (11), pp. 559–572 (1901) dx.DOI.org/10.1080/14786440109462720

[32] H. Hotelling, "Analysis of a complex of statistical variables into principal components", J. Educ. Psychol. **24** (6), pp. 417–441 (1933-09) dx.DOI.org/10.1037/h0071325

[33] T.N.E. Greville, "Some Applications of the Pseudoinverse of a Matrix", SIAM Review **2** (1), pp. 15–22 (1960–01) www.jstor.org/stable/2028054

[34] A. Bjerhammar, "Rectangular reciprocal matrices, with special reference to geodetic calculations", Bull. Geod. **20** (1), pp. 188–220 (1951). Excerpted from "Applications of Calculus of Matrices on Method of Least Squares, with Special Reference In Geodetic Calculations", Trans. Roy. Inst. Tech. Stockholm **48** (1951) dx.DOI.org/10.1007/BF02526278

[35] D.C. Liu and J. Nocedal, "On the Limited Memory Method for Large Scale Optimization", Math. Prog. B **45** (3), pp. 503–528 (1989-08) dx.DOI.org/10.1007/BF01589116

[36] C. Zhu, R.H. Byrd and J. Nocedal, "Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization", ACM Trans. Math. Software **23** (4), pp. 550–560. (1997) dx.DOI.org/10.1145/279232.279236

[37] Q. Le, J. Ngiam *et al.*, "On optimization methods for deep learning", Proc. ICML 2011, pp. 265–272 (2011), ISBN 978-1-4503-0619-5 www.icml-2011.org/papers.php

[38] G. van Rossum and J. de Boer, "Interactively Testing Remote Servers Using the Python Programming Language", CWI Quarterly Amsterdam **4** (4), pp. 283–303 (1991-12) oai.cwi.nl/oai/asset/18204/18204A.pdf

[39] E. Jones, E. Oliphant, P. Peterson *et al.*, "SciPy: Open Source Scientific Tools for Python" (2001) www.scipy.org

[40] S. van der Walt, S.C. Colbert and G. Varoquaux. "The NumPy Array: A Structure for Efficient Numerical Computation", Comput. Sci. Eng. **13**, pp. 22–30 (2011) dx.DOI.org/10.1109/MCSE.2011.37

[41] J. Safranek, "Experimental determination of storage ring optics using orbit response measurements", Nucl. Instr. Meth. Phys. Res. A **388** (1–2), pp. 27–36 (1997-03) dx.DOI.org/10.1016/S0168-9002(97)00309-4

[42] W.M. Thorburn, "The Myth of Occam's Razor", Mind **27** (3), pp. 345–353 (1918) en.wikisource.org/wiki/The_Myth_of_Occam's_Razor dx.DOI.org/10.1093/mind/XXVII.3.345

[43] R. Tomás, "Direct Measurement of Resonance Driving Terms in the Super Proton Synchrotron (SPS) of CERN using Beam Position Monitors", Ph.D. thesis (University of Valencia, 2003-01) cds.cern.ch/record/615164

[44] R. Miyamoto *et al.*, "Nonlinear dynamics studies in the Fermilab Tevatron using an AC dipole", Proc. PAC 2009, TH2PBC02, Vancouver, Canada inspirehep.net/record/822703

[45] R.R. Wilson, "The Tevatron", Fermilab Report TM-763 (1978-02) lss.fnal.gov/archive/test-tm/0000/fermilab-tm-0763.pdf

[46] G. Wendt *et al.*, "SPS Beam Steering for LHC Extraction", Proc. IPAC 2014, WEPRO068 (2014-09) cds.cern.ch/record/1756240

[47] C.E. Shannon, "Communication in the Presence of Noise", Proc. IEEE **86** (2), pp. 447–457 (1998-02). Reprinted from the Proc. of the IRE **37** (1), pp. 10–21 (1949-01)

[48] T. Schulte-Eickhoff, personal communication (TU Dortmund University, 2015-05)

[49] D. Schirmer and K. Wille, "DELTA Optics", Proc. PAC 1991, pp. 2859–2861 (1991-05) cds.cern.ch/record/902705

[50] P. Ungelenk, "Generation and Detection Schemes for Laser-Induced Coherent Terahertz Radiation at the Electron Storage Ring DELTA", Ph.D. Dissertation (TU Dortmund University, 2015-12) dx.DOI.org/10.17877/DE290R-16480

[51] S. Hilbrich and R. Molo, personal communication and DELTA lattice model `del14_03.lte` for elegant (TU Dortmund University, 2015-11)
S. Hilbrich, "Studies of the DELTA Lattice in View of a Future Short-Pulse Facility Based on Echo-Enabled Harmonic Generation", master's thesis (TU Dortmund University, 2015-03)

[52] K. Wille, "Initial Experience with DELTA", Proc. EPAC 1996, pp. 95–100 (1996-06) cds.cern.ch/record/859751

[53] A. Jankowiak, H. Bendfeldt-Hoppe *et al.*, "Reconstruction of the 75 MeV Linac of the DELTA Synchrotron Radiation Facility", Proc. EPAC 2000, pp. 634–636 (2000-06) cds.cern.ch/record/504992

[54] A. Jankowiak, "Strahldiagnose und Closed-Orbit-Charakterisierung mit HF-Strahllagemonitoren am Beispiel der Synchrotronstrahlungsquelle DELTA" (in German), Ph.D. Dissertation (Dortmund University, 1999-11) dx.DOI.org/10.17877/DE290R-8508

[55] P. Hartmann, personal communication (TU Dortmund University, 2015-11)

[56] Bergoz MX-BPM, as described in: *Multiplexed Beam Position Monitor User's Manual*, Rev. 1.5.4 (Bergoz Instrumentation) www.bergoz.com

[57] Controller Area Network (CAN) is described within the ISO standard 11898. www.iso.org

[58] Libera Electron BPM, as described in: *Libera Electron User Manual*, version 1.46, 6th edition (Instrumentation Technologies, 2007) www.i-tech.si
At DELTA, these BPMs are controlled and read out by DIAMOND Libera Epics Driver 1.46.1, which can be found at controls.diamond.ac.uk/downloads/other/libera

[59] D. Schirmer, personal communication (TU Dortmund University, 2016-02)

[60] M. Grewe, "SVD-basierte Orbitkorrektur am Speicherring Delta" (in German), Ph.D. Dissertation (Dortmund University, 2005-01) dx.DOI.org/10.17877/DE290R-8313

[61] M. Borland, "elegant: A Flexible SDDS-Compliant Code for Accelerator Simulation", Advanced Photon Source LS-287 (2000-09) aps.anl.gov/Science/Publications/lsnotes/ls287.pdf

[62] R. Molo, Ph.D. Dissertation in preparation (TU Dortmund University, 2016)

[63] P. Hartmann *et al.*, "Kicker-based tune measurement for DELTA", Proc. DIPAC 2007, WEPB21, Venice, Italy (2007) epaper.kek.jp/d07/papers/wepb21.pdf

[64] Electronic Logbook ("Schichtbuch" in German) of the DELTA facility (viewed in 2016-01).
The entries in German language have been translated by the author.

[65] H. Huck, "Optimierung und Charakterisierung des Free Electron Lasers am Speicherring DELTA" (in German), Ph.D. Dissertation (Dortmund University, 2009-02)
dx.DOI.org/10.17877/DE290R-8254

[66] P. Towalski, Ph.D. Dissertation in preparation (TU Dortmund University, 2016)

[67] "DELTA storage ring corrector magnet power supplies. Bilt system & BE548D 10A fast current source", User manual (Itest, 2013-03)
"Mainframes BN100, BN101, BN103 and BN105", User manual (Itest, 2012)
"BE548 Main SCPI commands", User manual (Itest, 2008-01) www.bilt-system.com

[68] S. Kötter, master's thesis in preparation (TU Dortmund University, 2016)

[69] C.-x. Wang, "Untangling mixed modes in model-independent analysis of beam dynamics in circular accelerators", Phys. Rev. ST Accel. Beams **7**, 114001 (2004-11)
dx.DOI.org/10.1103/PhysRevSTAB.7.114001

[70] C.S. Edmonds *et al.*, "Reconstruction of lattice parameters and beam momentum distribution from turn-by-turn beam position monitor readings in circular accelerators", Phys. Rev. ST Accel. Beams **17**, 054401 (2014-05) dx.DOI.org/10.1103/PhysRevSTAB.17.054401

[71] P. Vaníček, "Approximate spectral analysis by least-squares fit", Ap&SS **4** (4), pp. 387–391 (1969-08)
dx.DOI.org/10.1007/BF00651344

[72] J. Laskar, C. Froeschlé and A. Celletti, "The measure of chaos by the numerical analysis of the fundamental frequencies. Application to the standard mapping", Physica D **56** (2–3), pp. 253–269 (1992-05)
dx.DOI.org/10.1016/0167-2789(92)90028-L

[73] S.W. Smith, *The Scientist and Engineer's Guide to Digital Signal Processing* (California Technical Pub, 1997), ISBN 978-0966017632 www.dspguide.com

[74] O. Kopitetzki, "Iterativer Algorithmus zur strahlbasierten Vermessung und Korrektur von Magnetfehlaufstellungen am Speicherring Delta" (in German), Ph.D. Dissertation (TU Dortmund University, 2009-09) dx.DOI.org/10.17877/DE290R-476

[75] D. Engel, A. Jankowiak, R. Müller, M. Ries and J. Feikes, personal communication (Helmholtz-Zentrum Berlin and TU Dortmund University, 2015-06)

[76] J. Feikes, M. von Hartrott, M. Ries, P. Schmid and G. Wüstefeld, "Metrology Light Source: The first electron storage ring optimized for generating coherent THz radiation", Phys. Rev. ST Accel. Beams **14**, 030705 (2011-03) dx.DOI.org/10.1103/PhysRevSTAB.14.030705

[77] G. Portmann, J. Safranek and X. Huang, "Matlab Based LOCO", SLAC-PUB-14640 (2007) inspirehep.net/record/773188

[78] Website "Electron storage ring BESSY II- description and function" with sublinks to Additional Information. www.helmholtz-berlin.de/quellen/bessy/elektronenspeicherring

[79] A. Jankowiak, J. Knobloch *et al.* (ed.), "BESSY VSR, Variable pulse length Storage Ring, Upgrade of BESSY II", Technical Report (2015-06) dx.DOI.org/10.5442/R0001

[80] T. Becker, D. Krämer *et al.*, "Prototype Developement of the BESSY II Storage Ring Magnetic Elements", Proc. PAC 1995, Dallas, USA (1995) accelconf.web.cern.ch/AccelConf/p95/ARTICLES/FAP/FAP02.PDF

[81] G.W. Hill, "On the Part of the Motion of Lunar Perigee Which is a Function of the Mean Motions of the Sun and Moon", Acta Math. **8** (1), pp. 1–36 (1886-12) dx.DOI.org/10.1007/BF02417081

[82] T.O. Raubenheimer and R.D. Ruth, "A dispersion-free trajectory correction technique for linear colliders", Nucl. Instr. Meth. Phys. Res. A **302** (2), pp. 27–36 (1991–04) dx.DOI.org/10.1016/0168-9002(91)90403-D

[83] B.C. Yunn, R. Li and S. Simrock, "RESOLVE at CEBAF", Proc. PAC 1995, Dallas, Texas (1996) inspirehep.net/record/410357

[84] A.J. Dragt, "An overview of Lie methods for accelerator physics", Proc. PAC 2013, THAP1 (2013-10) accelconf.web.cern.ch/accelconf/pac2013/papers/thap1.pdf

[85] R. Bartolini, "Resonance Driving Term Experiments: An Overview", Proc. ICAP 2006, MOM1MP03, Chamonix, France (2006) accelconf.web.cern.ch/AccelConf/ICAP06/PAPERS/MOM1MP03.PDF

[86] J. Bahrdt and G. Wüstefeld, "A Taylor-Expanded Generating Function for Particle Motion in Arbitrary Magnetic Fields", Proc. EPAC 1992, pp. 670–672 accelconf.web.cern.ch/AccelConf/e92/PDF/EPAC1992_0670.PDF

[87] S.G. Johnson, "A useful basis for defective matrices: Generalized eigenvectors and the Jordan form", MIT Linear algebra course 18.06 (2009-04) web.mit.edu/18.06/www/Spring09/generalized-eigenvectors.pdf

[88] J.D. Jackson, *Classical Electrodynamics*, 3rd edition (Wiley, 1999)

[89] M.G. Minty *et al.*, "Emittance growth due to decoherence and wakefields", WAB09, Proc. PAC95 (1995) epaper.kek.jp/p95/articles/wab/wab09.pdf

[90] M.A. Furman, "Decoherence". In: Alex Chao (Ed.): *Handbook of Accelerator Physics and Engineering*, 2nd edition (World Scientific, 2013), ISBN 978-9814417174

Acknowledgement (in German)

Eine Dissertation entsteht niemals ohne eine Vielfalt externer Impulse und Unterstützung. Deshalb möchte ich mich bei jeder Person, welche in einen Kausalzusammenhang mit der Entstehung und Veröffentlichung dieser Arbeit gebracht werden kann, bedanken. Besonderer Dank für ihre Nachsicht gilt allen Personen, die ich in folgender Auflistung vergessen habe. Ausserdem sollte nicht zuviel in die Reihenfolge der Danksagungen und Ähnliches hineininterpretiert werden.

Ich beginne mit meiner Frau Stella, die mich in jeder Lebenssituation unterstützt hat und ohne die diese Arbeit nicht existieren würde. Weiterhin bedanke ich mich bei meinen Eltern Geraldine und Bernd und Schwiegereltern Ina und Thomas, die mich während meiner nicht kurzen Promotionsphase begleitet haben.

Ich bedanke mich bei allen Mitgliedern des Zentrums für Synchrotronstrahlung für das notwendige "blue skies research"-Klima. Mit meinen (teilweise ehemaligen) Bürokollegen Stephan Kötter, Alessandro Ferrarotto, Helge Rast und Holger Huck habe ich unzählige hilfreiche Diskussionen über meine Arbeit geführt. Benjamin Isbarn und Malte Sommer haben mir ebenfalls seit geraumer Zeit sinnvolle Impulse vermittelt, und ich hoffe, dies mit den weiterhin am Zentrum angestellten Personen fortsetzen zu können.

Peter Hartmann hat mich als mein ehemaliger Betreuer auf das Thema Strahldynamik in Beschleunigern aufmerksam gemacht. Patrick Grete hat mich motiviert, meine Gedanken auch einmal zu veröffentlichen. Weiterhin geht mein Dank an Detlev Schirmer, Gerald Schmidt und alle weiteren Leser der Entwürfe dieser Arbeit für ihre sehr hilfreichen Korrekturanmerkungen.

Patryk Towalski und Gerrit Schünemann gebührt ebenfalls Dank, nicht zuletzt, weil ich Teile des von ihnen designten Fast Orbit Feedbacks zweckentfremden durfte. Vielen Dank auch an Robert Molo und Svenja Hilbrich, mit denen ich Diskussionen über Strahloptik geführt habe, an alle weiteren Mitglieder der Beschleunigerphysik-Gruppe (Fin Bahnsen, Yvonne Bernau, Max Bolsinger, Marcel Bursy, Fabian Goetz, Markus Höner, Maryam Huck, Marc Jebramcik, Daniel Krieg, Nils Lockmann, Carsten Mai, Arne Meyer auf der Heide, Raffael Niemczyk, Andre Nowaczyk, Andreas Schick, Mateusz Suski, Gholamreza Shayeganrad, Peter Ungelenk und Dennis Zimmermann), alle technischen Mitarbeiter (Wolfgang Brembt, Günter Dahlmann, Thomas Dybiona, Andreas Erpelding, Bernhard Hippert, Vadim Kniss, Peter Kortmann, Hans-Peter Ruhl, Tanja Schulte-Eickhoff), an Jochem Friedl, Petra Lindemann und Monika Voits-Besli, sowie an Shaukat Khan und Klaus Wille.

Auch wenn diese Arbeit thematisch nicht mit der DOHRO-Kollaboration zusammenhängt, möchte ich mich bei allen Beteiligten der Theoretischen Elektrotechnik der Uni Rostock (Gruppe Ursula van Rienen), sowie des Helmholtz-Zentrums Berlin (Gruppen Andreas Jankowiak, Jens Knobloch) für das freundliche und kooperative Arbeitsklima bedanken.

Ich möchte mich abschliessend bei der Prüfungskommission bedanken. Im Besonderen gilt mein Dank dem Erstgutachter Thomas Weis, welcher mich während der gesamten Promotionsphase unter anderem über eine durchgehende Anstellung und das damit verbundene ausdauernde Vertrauen in meine Fähigkeiten unterstützt hat. Weiterhin danke ich dem Zweitgutachter Andreas Jankowiak (HZB), welcher zur Prüfung anreist, Ulf Berges als Vertreter der wissenschaftlichen Mitarbeiter und Jan Kierfeld als Leiter der Prüfungskommission dafür, dass sie ihre knappe Zeit für die Begutachtung und Verteidigung meiner Arbeit zur Verfügung stellen.

A. Appendix

A.1. Beam Optics

A.1.1. Defective one-turn transfer matrices allow no bound motion

Matrices are called defective if they are not diagonalizable, that is one or more of its eigenvalues are identical, and its eigenvectors do not form a full basis. We use computations that can be found in [87] for the following derivation.

When considering a defective matrix $\mathbf{T}$, a multiple eigenvalue λ and a corresponding eigenvector $\vec{v}$, so that $(\mathbf{T} - \lambda\mathbf{1})\vec{v} = 0$, the characteristic polynomial has a multiple root λ. One can define a generalized eigenvector $\vec{v}^{(2)}$ by

$$(\mathbf{T} - \lambda\mathbf{1})\vec{v}^{(2)} = \vec{v} \quad \text{and} \quad \vec{v}^{(2)} \perp \vec{v}. \tag{A.1}$$

This can be done iteratively for generalized eigenvectors by replacing $\vec{v}^{(2)} \to \vec{v}^{(3)}, \vec{v} \to \vec{v}^{(2)}$ and so forth until the multiplicity of the respective eigenvalue has been reached and all (standard and generalized) eigenvectors form an orthonormal basis.

If $\vec{x}^\dagger\vec{v}^{(2)}$ does not vanish exactly and by using (A.1), iterative multiplication of any vector $\vec{x}$ with $\mathbf{T}$ results in a term

$$\mathbf{T}^n\vec{v}^{(2)} = \lambda^n\vec{v}^{(2)} + n\lambda^{n-1}\vec{v}. \tag{A.2}$$

Thus if a generalized eigenvector exists so that the matrix is defective, and $|\lambda| \geq 1$ holds for the multiple eigenvalue, any input vector which is not perpendicular to $\vec{v}^{(2)}$ will increase in amplitude for each turn n, and the motion is not bound.

For our considerations, we have assumed (Liouville's theorem) that $\det \mathbf{T} = \prod_m \lambda_m = \prod_m |\lambda_m| = 1$. Thus, if any $|\lambda_m| > 1$, at least one other must be $|\lambda_n| < 1$. There are two possibilities for the multiple eigenvalue λ.

1. $|\lambda| \geq 1$. Then by (A.2), the motion is unbound.
2. $|\lambda| < 1$. Then at least one other eigenvalue $\tilde{\lambda}$ and corresponding eigenvector $\vec{u}$ exists so that

   $$\mathbf{T}^n\vec{u} = \tilde{\lambda}^n\vec{u} \quad \text{with} \quad |\tilde{\lambda}| > 1,$$

   which also results in unbound motion.

Therefore no multiple eigenvalue and thus no defective matrix allows bound motion.

A.1.2. Real-valued block matrix expressions for eigenorbits

By sec. 5.1.1, we can transform the eigenorbit equations (2.6) respectively (2.7) into real-valued equations

$$\begin{aligned}\begin{pmatrix} x \\ x' \end{pmatrix} = \Re\left\{\begin{pmatrix} X \\ X' \end{pmatrix} \mathrm{e}^{\mathrm{i}\phi}\right\} &= \mathbf{L}\begin{pmatrix} \cos\phi \\ \sin\phi \end{pmatrix} \text{ with } \mathbf{L} = \begin{pmatrix} \Re X & -\Im X \\ \Re X' & -\Im X' \end{pmatrix} \\ &= \mathbf{F}\begin{pmatrix} \cos\phi \\ -\sin\phi \end{pmatrix} \text{ with } \mathbf{F} = \begin{pmatrix} \Re X & \Im X \\ \Re X' & \Im X' \end{pmatrix}\end{aligned}$$

Therefore, $\mathbf{L}$ respectively $\mathbf{F}$ transforms a circle with positive respectively negative rotation into an ellipse. Here, connections to Fourier components and a Cartesian parameterization found in [13] can be constructed by the matrix $\mathbf{F}$. In similar fashion, we can construct relations between normalized eigenorbits and matrices $\mathbf{P}$ in [13] by

$$\mathbf{P} = \begin{pmatrix} \Re\hat{X} & \Im\hat{X} \\ \Re\hat{X}' & \Im\hat{X}' \end{pmatrix}.$$

Obtaining the decoupled invariant analytically

By above considerations, we can conclude that the area π of the unit circle is then scaled by

$$I_x = \det \mathbf{L} = \Re X \Im X' - \Im X \Re X' = \Im(X^* X').$$

This is a signed quantity. If the sign is negative, the direction of rotation is opposite for the unit circle and the ellipse regarding phase advance.

A.1.3. Tracking phasor (eigen)vectors

$\vec{z}_n, \vec{Z}_m$ from sec. 2.3 are only related by complex scalars, and thus linear operations. Therefore,

$$\vec{z}_n \to \mathbf{M}\vec{z}_n \quad \text{is equivalent to} \quad \vec{Z}_m \to \mathbf{M}\vec{Z}_m.$$

As this may seem counterintuitive due to $\vec{Z}_m$ being complex quantities, we give a detailed calculation based on $\vec{z}_n, \underline{\vec{z}_n} = \mathbf{M}\vec{z}_n$ in the following. Following (2.5) one may write

$$\vec{z}_n = \sum_m^{2M} \frac{1}{2}\vec{Z}_m \lambda_m^n, \qquad \underline{\vec{z}_n} = \sum_m^{2M} \frac{1}{2}\underline{\vec{Z}_m} \lambda_m^n,$$

Left-multiplying the first equation with $\mathbf{M}$, one obtains

$$\mathbf{M}\vec{z} = \sum_m^{2M} \frac{1}{2}\mathbf{M}\vec{Z}_m \lambda_m^n \quad \stackrel{!}{=} \quad \underline{\vec{z}_n} = \sum_m^{2M} \frac{1}{2}\underline{\vec{Z}_m} \lambda_m^n,$$

By comparison of coefficients, this yields the expected result $\underline{\vec{Z}_m} = \mathbf{M}\vec{Z}_m$, thus phasor vectors can be tracked trough linear maps exactly like phase space vectors. This also holds for their normalized counterparts, the complex Mais-Ripken vectors, and also for the standard Mais-Ripken vectors.

As the relation also holds for components of $\vec{z}, \vec{Z}_m$ related by linear operations, one can, with the same argument, also evaluate that the intercept theorem

$$\vec{r}_n'(s_j) = \frac{\vec{r}_n(s_{j+1}) - \vec{r}_n(s_j)}{L} \quad \text{is equivalent to} \quad \vec{R}_{j,m}' = \frac{\vec{R}_{j+1,m} - \vec{R}_{j,m}}{L}.$$

It should be noted that the aforementioned relations do not hold for Courant-Snyder, Mais-Ripken, or Edwards-Teng parameters, as these are polar-like coordinates for the optical relations and are thus not used the derivation of optical relations in this work.

A.1.4. Beam position monitors and their signals

The standard devices for observation of beam motion are transverse, capacitive beam position monitors, which can be realized by a variety of possible constructions. For all experiments presented in this thesis, the monitor hardware is realized by an arrangement of four electrodes, embedded in but electrically isolated from the particle beam chamber.

We assume in the following that, regarding time constants of the monitor measurement process and superposition of charges, the accelerator structure is a linear time-invariant system. We denote the port signal on the n-th electrode by a_n, so that its relation to a passing particle distribution respectively current density $j(x, y, t)$ through the monitor plane can be described using a Green's function

$$a_n(t) = \int G_n(x, y, t - t')j(x, y, t') \,\mathrm{d}x \,\mathrm{d}y \,\mathrm{d}t'.$$

The particles in the storage ring are assumed to possess high energies with correspondingly high Lorentz factors $\gamma \gg 1$ and are thus ultrarelativistic ($v \approx c$). The distribution of electromagnetic fields, which couple the pick-up signals to the beam charge distribution, is contracted in the laboratory frame with $1/\gamma$ in flight direction; thus one may approximate these fields to be located in a infinesimally thin transverse plane around the particle, so that $G(x, y, t - t') = G(x, y)\delta(t - t')$. With this approximation of the distribution in time, the above relation simplifies to

$$a_n(t) = \int G_n(x, y)j(x, y, t) \,\mathrm{d}x \,\mathrm{d}y.$$

To obtain $G_n(x, y)$, one could choose j to be independent of t, which corresponds to a constant charge respectively current distribution $j(x, y)$. We also assume that static transverse magnetic fields do not influence the electrode.[1] Thus, ignoring changes in the cross-section of the chamber, one can state that the linear port signals obtained at the electrodes are proportional to the static voltages induced on the electrodes by the transverse charge distribution $\rho(x, y)$.

For the following, we assume that in the region which is occupied by the charge distribution, the Green's function can be linearized around the centroid of the beam at x_c, y_c to sufficient accuracy, so that

$$G_n(x, y) \approx G_n(x_c, y_c) + (x - x_c) \left[\partial_x G(x, y)\right]_{x=x_c} + (y - y_c) \left[\partial_x G(x, y)\right]_{y=y_c}.$$

Then, the integral term in

$$a_n(t) \approx G_n(x_c, y_c) \int j(x, y, t) \,\mathrm{d}x \,\mathrm{d}y = G_n(x_c, y_c)I(x_c, y_c)$$

evaluates to the static (or average) current of the distribution.

This is a transverse point-charge aproximation for charge distributions inside the beam chamber. If the particles were assumed to travel in free space, the linearization of G_n would be equivalent to a far-field approximation after which all terms higher than monopole are neglected [88]. Thus, if we assume that the vacuum chamber properties did not significantly perturb the Green's functions, this would be equivalent to approximate the transverse spatial distribution of the beam being much smaller than the distance between beam centroid and each pickup electrode.

[1] This assumption has been made implicity when approximating the Green's function as δ distribution in t,which allows no finite time derivatives.

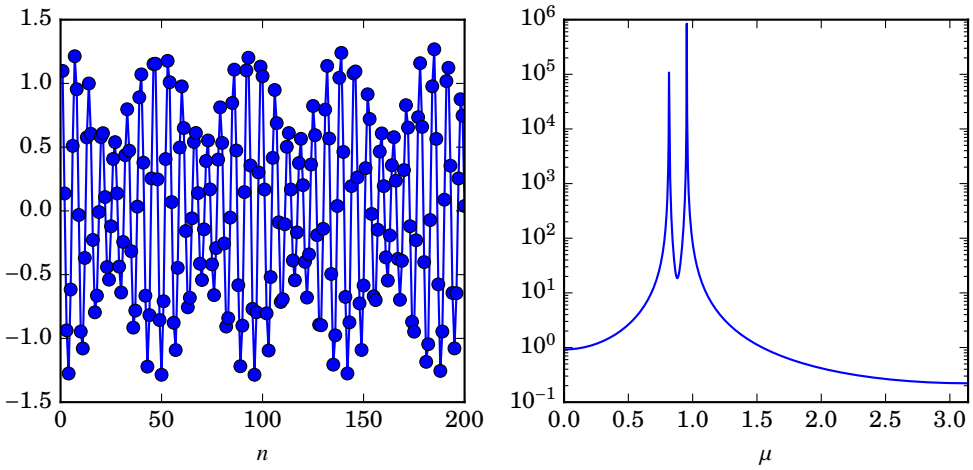

Figure A.1.: Cutting of a sequence with $N = 2048$ consisting of 2 cisoids (left) and its DTFT (right).

A.2. Turn-by-Turn data

An important property of DTFTs that will be used in the following is that they comply with the convolution theorem, so that

$$\mathcal{F}_\mu(x_n * y_n) = (\mathcal{F}_\mu x_n)(\mathcal{F}_\mu y_n).$$

A.2.1. Numerical Analysis of Fundamental Frequencies (NAFF)

When the finite sequence only consists of a few number of cisoids with distinct phase advances, the local maxima of the DTFT are still a good approximation for the respective μ_m. The NAFF algorithm [72] uses this approximation implicitly to obtain the phase advances of cisoids occuring in the sequence.

The following steps can be performed in a loop to obtain all $\mu_m, \vec{R}_{jm}$ using a scalar, real-valued DTFT-derived quantity called $|\vec{F}(\mu)|$ in the following. In this context, the DTFT amplitudes are given as $\vec{F}$.

1. *Compute the maximum argument* $\mu_{\max}$ *of* $|\vec{F}(\mu)|$.

   $\mu_{\max}$ is the DTFT estimate for the phase advance of the strongest (remaining) cisoid in the sequence. To obtain proper start values, a preliminary Fast Fourier Transform (FFT) of the data is performed and the bin number of the maximum amplitude is used to compute the starting value of the optimization. Then, a (bounded) scalar (maximum) optimizer is applied to $|F(\mu)|^2$.[2] In standard NAFF, the accuracy of both estimates is enhanced by using an appropriate window function.

2. *Compute the amplitudes* $\vec{F}(\mu_{\max})$ *of the signal.*

   If the optimization procedure can be manipulated, this step can be integrated into the first one as the components of $\vec{F}(\mu_{\max})$ are usually computed during the optimization.

3. *Synthesize the corresponding cisoid and substract it from the sequence.*

   This cisoid is simply given as $\vec{F}(\mu_{\max})\mathrm{e}^{\mathrm{i}\mu_{\max}n}/N$.

Performing this three steps M times in a loop leads to finding the phase advances of the M strongest resonances in the signal, corresponding to the μ_m phase advances if the signal-to-noise ratio is sufficient.

[2] In noisy systems where μ can only be obtained to a limited accuracy, it can be avantageous to extend the input timesequence with a sufficient number of zeros (zero-padding) and use the maximum argument of the resulting interpolated spectrum as an estimate of $\mu_{\max}$ instead.

Algorithm (taken out!) is based on NAFF, but uses multiple sequences in n from all monitors j and directions d simultaneuosly. To find the resonance phase advances μ_m using a scalar optimizer, we use a total DTFT power spectrum function

$$F^2(\mu) = \sum_{j=1}^{J} (\mathcal{F}\vec{r}_{jn})^\dagger (\mathcal{F}\vec{r}_{jn}) = \sum_{j=1}^{J} \left\{ \left(\sum_n \vec{r}_{jn} \cos(\mu n) \right)^2 + \left(\sum_n \vec{r}_{jn} \sin(\mu n) \right)^2 \right\}. \tag{A.3}$$

A.2.2. Interpretation of LSSA in comparison to NAFF

Neglecting details about the solution of the equation system, we state that the pseudoinverse of a matrix $\mathbf{M}$ is defined by

$$\operatorname{pinv} \mathbf{M} = (\mathbf{M}^\dagger \mathbf{M})^{-1} \mathbf{M}^\dagger \quad \text{if } \mathbf{M}^\dagger \mathbf{M} \text{ is invertible.}$$

In fact, the pseudoinverse of the LSSA system matrix can be reformulated so that the first multiplicative step would correspond to

$$\mathbf{U}^\dagger \mathbf{B} \quad \text{or, in components, } \sum_n \mathrm{e}^{-\mathrm{i}n\tilde{\mu}_m} B_{n(jd)} = \sum_n \mathrm{e}^{-\mathrm{i}n\tilde{\mu}_m} (\vec{r}_{jn})_d,$$

which is a computation that is carried out row-wise as a step of the NAFF algorithm. Thus, if the cisoids are orthogonal, NAFF and LSSA will yield the same result as then $\mathbf{M}^\dagger \mathbf{M}$ is a diagonal matrix and the substraction of the first resonance does not influence the optimization of the next. In any other case, LSSA will yield more exact results than NAFF in the form in which it is introduced here.

A comparison of NAFF and LSSA is shown in Figs. A.2 and A.3 for a timesequence in which the phase advances of two cisoids are relatively close to each other and thus orthogonality is strongly broken. In this case, NAFF will produce systematic errors.

Time complexity of LSSA

Using the linearity of the $2JM^2P$ free parameters encoded in all C_{jmwp} by LSSA reduces the time complexity of the optimization. Although no commonly accepted minimum complexity of nonlinear opti-

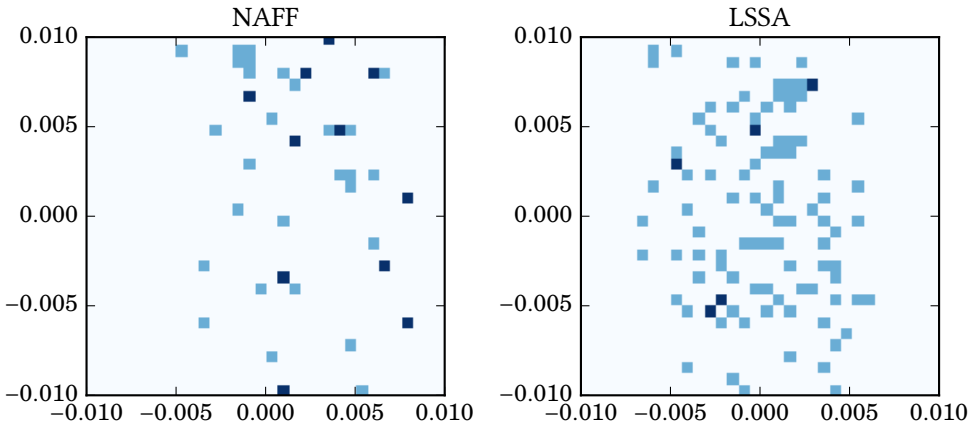

Figure A.2.: Comparison of NAFF and LSSA algorithms for a signal-to-noise level of 4 with 3 channels and 128 different random timesequences. Shown are M = 2-dimensional histograms of μ_m estimates for these timesequences. The axes are $N\Delta\mu/2\pi$, which is the difference between correct and measured μ expressed in units of bins.

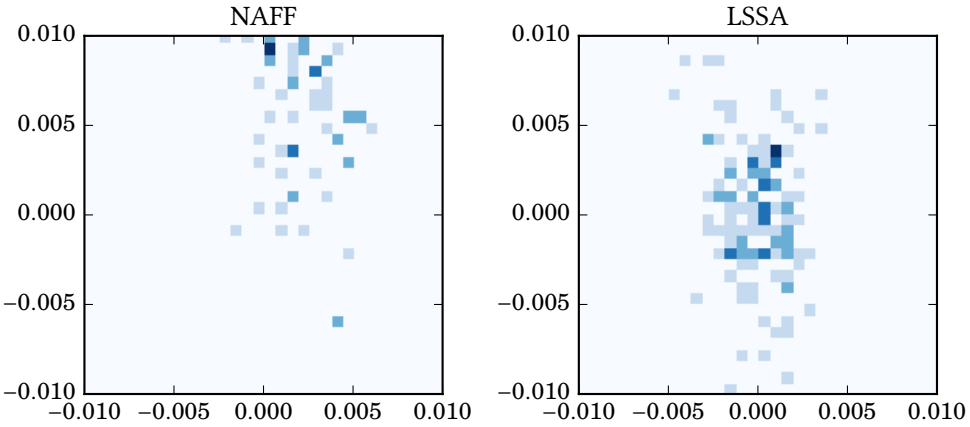

Figure A.3.: Comparison of NAFF and LSSA algorithms for a signal-to-noise level of 8 with the setup taken from Fig. A.2. Note that a systematic error in NAFF remains.

mization in dependence of dimensions can be found, it is certainly greater than $O(C^1)$. Computing the Moore-Penrose pseudoinverse in our context (with full-column rank) only depends on M, N, and the remaining matrix multiplication has the complexity $O(NJD)$, so that the complexity in JD is linear, while the complexity in M remains unchanged.

The time needed to run the algorithm is about one magnitude greater than that of NAFF, due to the small-scale inversion of a $2M \times 2M$ matrix, and performance of the implementation could be accelerated. Hoewever, the LSSA implementation used is sufficient so that phasors can be computed in reasonable time (which for the experiments in this thesis is ≈ 5 seconds).

A.2.3. Decoherence / Filamentation of the beam centroid

So far, our considerations regarding oscillation data were focused on single particle data. For small transverse dimensions of the particle beam, monitor measurements can be approximated as measuring the motion of the transverse beam centroid (sec. A.1.4). Due to nonlinear effects (see also chapter 10), the phase advances μ_m are energy and amplitude-dependent.

There are several effects that are distinguished in common literature based on the first-order expansion of μ. If we only consider the longitudinal component of $\vec{z}$ corresponding to the energy deviation, the resulting expansion is

$$\tilde{\mu}_m = \xi_m \cdot \delta + \mu_m,$$

where $\delta = \Delta E/E$ is the relative energy deviation, a component of $\vec{z}$. In analogy to photons, this is called chromatic abberation, and ξ_m is called chromaticity of the mode.

We now want to examine how the motion of the beam centroid is affected by chromaticity. A recent review of this is given in [70], where the authors also conclude that with knowledge of chromaticity and TbT data,[3] it is possible in principle to compute the energy distribution $\Phi(\delta)$ of the integrated bunch train.

Here, we are primarily interested in the perturbative nature of this effect, that is, if and how we can extend single-particle considerations to centroids. For this, we use the energy distribution $\Phi(\delta, n)$ and

[3] and under the assumption that other nonlinear effects are neglible to chromaticity, which is a valid assumption for EMMA

the linear energy dependence of $\tilde{\mu}$ to obtain the spatial motion of the centroid as

$$\vec{r}_{jn}^{\,\mathrm{ctr}} = \int \vec{r}_{jn}(\delta)\Phi(\delta)\,\mathrm{d}\delta = \int \Re\left\{\sum_m \vec{R}_{jm}\int \Phi(\delta)\mathrm{e}^{\mathrm{i}n(\mu_m+\xi_m\delta)}\,\mathrm{d}\delta\right\}$$

$$\text{or}\quad \vec{r}_{jn}^{\,\mathrm{ctr}} = \Re\left\{\sum_m \vec{R}_m u_{nm}\mathrm{e}^{\mathrm{i}n\mu_m}\right\}\quad \text{with modal envelopes}\quad u_{nm} = \int \Phi(\delta,n)\mathrm{e}^{\mathrm{i}n\xi_m\delta}\,\mathrm{d}\delta. \tag{A.4}$$

Therefore, the centroid motion can be interpreted as a single-particle motion that is modulated by a decoherence term u_{nm} for a given mode m at turn n. Without further assumptions, we can also state that for an infinite timesequence,

$$(Fu_{nm})(\mu = \xi_m\delta) \propto \Phi(\delta).$$

Thus, the modifications in the DTFT directly depend on the shape of the energy distribution. In the following, we will discuss two reasonable and common approximations for this distribution which are connected to corresponding shapes of the modal envelopes u_{nm}.[4]

Static Gaussian energy distribution

We assume that the distribtion is normally distributed with a characteristic relative energy width σ and thus has the form

$$\Phi(\delta) = \frac{1}{\sqrt{2\pi\sigma^2}}\mathrm{e}^{-\delta^2/2\sigma^2}.$$

As u_{nm} is essentially an inverse DTFT of the distribution, this results in Gaussians being transformed to Gaussians, so that

$$u_{nm}(\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}}\int \mathrm{e}^{-(\mu/\xi_m\sigma)^2/2}\cos(n\mu)\,\mathrm{d}\mu = \frac{1}{\sqrt{2\sigma^2}}\mathrm{e}^{-(n\xi_m\sigma)^2/2}.$$

Oscillating Gaussian energy distribution

So far, we have neglected the possibility that the beam centroid can also be excited by the kick along the s, s' direction, oscillating with its synchrotron phase advance μ_0. The derivation of these more general expressions, also including a term for non-linear betatron motion, has already been done in [89, 90].

DTFT decoherence artefacts

We will shortly consider the leakage effect with decoherence. This can be done by multiplying the rectangular sequence into the decoherence term u_{nm}, so that

$$u_{nm}^{\mathrm{win}} = u_{nm}\sqcap_n^N,$$

from which follows that

$$\mathcal{F}z_n^{\mathrm{meas}} = (\mathcal{F}\sqcap_n^N) * (\mathcal{F}\vec{r}_{jn}^{\,\mathrm{ctr}}) = \sum_m (\mathcal{F}u_{nm}^{\mathrm{win}}) * (\vec{R}_{jm}\mathcal{F}\mathrm{e}^{\mathrm{i}\mu_m n}).$$

The DTFT of u_{nm}^{win} is also given by the convolution theorem as

$$\mathcal{F}u_{nm}^{\mathrm{win}} = (\mathcal{F}u_{nm}) * (\mathcal{F}\sqcap_n^N). \tag{A.5}$$

[4] The inverse approach of fitting polynomial expression directly for u_{nm} will not work out for finite timesequences, as the assumption of a given distribution implicitly allows an extrapolation of the timesequence which is finite. Fitting a set of spectral sidebands is possible, but neglects the continuity of the distribution.

If the convolution by the leakage DTFT is initially ignored, it would follow that the phase advance μ_m obtained as the DFT maximum corresponds to the maximum of the energy distribution. Then, reintroducing the leakage DTFT convolution will shift μ_m into the direction of the mean energy distribution. The shift to the mean will be large if the characteristic width of $\mathcal{F}u_{nm}$ is small in relation to $\mathcal{F}\sqcap_n^N$ and vice versa.

In summary, the DTFT maximum thus corresponds to an energy between the maximum argument of the energy distribution $\delta_{\max}$ and the mean of the energy distribution $\langle\delta\rangle = 0$. It is thus reasonable to assume that the μ_m estimate will only slightly be effected by decoherence in the common measurement situation.[5]

DTFT amplitude errors Any window function like the rectangular sequence or the decoherence window also modifies the maximum amplitude of the DTFT and the respective resonance. This scaling is global, which means identical for all monitors j and directions d. In consequence, the relative scaling of the monitor vectors is not influenced.

Furthermore, as the obtained monitor vectors will be normalized using the invariant of motion (2.14), the normalized monitor vectors are independent of the initial amplitude scaling, as this scaling has been absorbed by the invariant.

A.2.4. Inclusion of decoherence by extended LSSA optimization

Having characterized the decoherence effect, our goal is do include this effects when decomposing the timesequence. In sec. 8.2 fitting of the timesequences by LSSA has been performed, which as a least-squares optimization procedure is not influnced by DTFT artefacts. Thus the approach of extending LSSA to include decoherence effects is reasonable.

For any distribution with P unknown parameters a_p, we can generalize the optimization problem in 8.2.1 as

$$\begin{aligned} &\text{find } \mu_m, \vec{R}_{jm} \text{for all } j, m \text{ and } a_1, \ldots, a_P \text{ so that} \\ \chi^2 = \sum_{j,n} \left| \vec{r}_{jn} - \Re\left\{ \sum_m \vec{R}_{jm} U_{nm}(\mu_m, a_1, \ldots, a_P) \right\} \right|^2 &\text{ is minimal.} \end{aligned} \tag{A.6}$$

Here, the elements of $\mathbf{U}$ are the only quantities containing the non-linear optimization parameters and can be defined as

$$U_{nm}(\mu_m, a_1, \ldots, a_P) = u_{nm}(a_1, \ldots, a_P)\mathrm{e}^{\mathrm{i}n\mu_m}.$$

In consequence, the optimization goal will be to minimize the error between the cisoids with the given envelope and the input data. Then, the LSSA equation system (8.4) can be used for optimization with the increased set of non-linear parameters by using the respective matrix $\mathbf{U}$.

Using the oscillating Gaussian energy distribution, which is characterised by synchrotron phase advance μ_0 and the scaled variance $\xi_m\sigma$, we therefore have

$$2M + 1 \text{ nonlinear variables } \mu_0, \mu_1, \ldots, \mu_M, \xi_1\sigma, \ldots, \xi_M\sigma$$

to use in the outer optimization loop.

[5] If the distribution has only one maximum, the phase advance μ_m will occur in the oscillation of many particles in the beam.

A.3. Basic data processing and COBEA computations

A.3.1. Singular Value Decomposition

A singular value decomposition of a matrix $\mathbf{M}$ can be described by [28]

$$\mathbf{M} = \mathbf{U}\mathbf{S}\mathbf{V}^{\dagger},$$

where $\mathbf{U}, \mathbf{V}$ are unitary matrices, and $\mathbf{S}$ is a diagonal matrix with non-negative entries (singular values), which are arranged in decreasing order. Every rectangular matrix with real or complex entries can be decomposed into this form. Further conditions exist to make $\mathbf{U}$ and $\mathbf{V}$ unique [28].

Usage for overdetermined systems

For overdetermined systems of the form $\mathbf{B} = \mathbf{A}\mathbf{X}$, it is very unlikely that an $\mathbf{X}$ exists to solve the system exactly. Instead, a useful approach is to find $\mathbf{X}$ so that the error function[6]

$$\chi^2(\mathbf{X}) = \|\mathbf{B} - \mathbf{A}\mathbf{X}\|^2 \text{ is minimal.}$$

As it is possible that more than one $\mathbf{X}$ minimizes χ^2, we state the additional condition that $\mathbf{X}_{\text{min}}$ has the minimal $\|\mathbf{X}\|$ in this set [28].

This is accomplished by the pseudoinverse of $\mathbf{A}$ via [34, 33]

$$\mathbf{X}_{\text{min}} = \text{pinv}(\mathbf{A})\ \mathbf{B},$$

and it can be computed in a numerically robust way by SVD as [28]

$$\text{pinv}(\mathbf{A}) = \mathbf{V}\mathbf{S}_{\text{inv}}\mathbf{U}^{\dagger},$$

where $\mathbf{S}_{\text{inv}}$ is a diagonal matrix in which each non-zero singular value (in $\mathbf{S}$) is replaced by its reciprocal. For consideration of numerical errors, very small singular values (relative to the largest) are also replaced by zeros in the inverse, depending on a cutoff criterion.

A.3.2. Outer product decomposition

In this thesis, there are many occasions where an input array is decomposed into an outer product. The simplest form of an outer product is the product of one-dimensional arrays

$$A_{pq} = b_p c_q^* \qquad \text{or, using vector notation,} \qquad \mathbf{A} = \vec{b} \otimes \vec{c} = \vec{b}\vec{c}^{\dagger}. \tag{A.7}$$

While more basic techniques exist to solve this problem (also known as rank-one problem), a robust and convenient way is Singular Value Decomposition (SVD). As it is also widely available and also used for the solution of other problems in this thesis, SVD is the technique used in the following. $\mathbf{A}$ can be expressed using its SVD as

$$A_{pq} = \sum_{n=1}^{N} u_{pn}\sigma_n v_{nq}^*.$$

As an outer product has rank one and the singular values σ_n are ordered decreasingly, we can assume that only one non-zero singular value σ_1 does exist, so that

$$A_{pq} = u_{p1}\sigma_1 v_{1q}^*,$$

Then we can, up to a scaling factor, identify $u_{p1} \propto b_p$ and $v_{1q} \propto c_q$. We choose to absorb the scaling factor equally in both terms, so that

$$b_p = \sqrt{\sigma}_1 u_{p1} \text{ and } c_q = \sqrt{\sigma}_1 v_{1q}.$$

[6] $\|\cdot\|$ indicates the Frobenius norm, with is the root of the sum of absolute element squares of a matrix or vector.

Multiple indices

In almost all occasions discussed in this thesis, the arrays for decomposition have more than two indices. Let us assume a problem of the form

$$\bar{A}_{dfrmj} = \bar{b}_{drm}\bar{c}^*_{fj},$$

which we can describe as an outer product along indices (d, r, m) with (f, j). We can define the index mappings

$$(d, r, m) \leftrightarrow p \text{ and } (f, j) \leftrightarrow q,$$

so that we can create a matrix $\mathbf{A}$ from the original multi-dimensional input array $\bar{\mathbf{A}}$. The decomposition problem is again reduced to the form (A.7). After obtaining b_p, c_q, the arrays $\bar{b}_{drm}, \bar{c}_{fj}$ can be created by inverting the above index mapping.

Checking the signal contents

By assuming the problem to be rank one, we assume all signals belonging to other singular values but σ_1 to originate from noise. One can thus check if enough signal is present for a meaningful decomposition of the the input array by the condition $\sigma_1/\sigma_2 \gg 1$.

A.3.3. Replacing the monitor subroutine

Using PCA, it is also possible to represent the monitor vectors as complex linear combinations of principal components. For cycle A, this is simply described as

$$\vec{R}_{jm} = \sum_n \langle \vec{r} \rangle^{\mathrm{A}}_{jn} c^{\mathrm{A}}_{nm} = \langle \mathbf{r} \rangle^{\mathrm{A}}_{j} \vec{c}^{\,\mathrm{A}}_{m} \quad \text{for } j \in \mathbb{J}_{\mathrm{A}}. \tag{A.8}$$

For cycle B, the unperturbed segment contains $s = 0$, so a switching of turns occurs when crossing $s = 0$. Then, the monitor vectors in cycle B are given by

$$\vec{R}_{jm} = \langle \mathbf{r} \rangle^{\mathrm{B}}_{j} \tilde{c}^{\,\mathrm{B}}_{m} \begin{cases} 1 & \text{for } s_j \geq s_3 \\ \mathrm{e}^{-\mathrm{i}\mu_m} & \text{for } s_j < s_3 \end{cases} \text{ and } j \in \mathbb{J}_{\mathrm{B}}.$$

At least four monitors j exist at which $\vec{R}_{jm}$ can be expressed using both A and B; these are represented by the monitor split indices $j_1 - j_4$. We use this property by reformulating these representations as equality constraints, so that

$$\begin{aligned} \vec{R}_{jm} = \quad & \langle \mathbf{r} \rangle^{\mathrm{A}}_{j} \vec{c}^{\,\mathrm{A}}_{m} \stackrel{!}{=} \mathrm{e}^{-\mathrm{i}\mu_m} \langle \mathbf{r} \rangle^{\mathrm{B}}_{j} \tilde{c}^{\,\mathrm{B}}_{m} \quad \text{for } j \in \{j_1, j_2\} \\ \text{and} \quad \vec{R}_{jm} = \quad & \langle \mathbf{r} \rangle^{\mathrm{A}}_{j} \vec{c}^{\,\mathrm{A}}_{m} \stackrel{!}{=} \langle \mathbf{r} \rangle^{\mathrm{B}}_{j} \tilde{c}^{\,\mathrm{B}}_{m} \quad \text{for } j \in \{j_3, j_4\}\,. \end{aligned}$$

These conditions at the start and end regions of unperturbed segments can be subsumed into a generalized eigenvalue problem

$$\mathrm{e}^{\mathrm{i}\mu_m} \begin{pmatrix} \langle \mathbf{r} \rangle^{\mathrm{A}}_{j_1} & 0 \\ \langle \mathbf{r} \rangle^{\mathrm{A}}_{j_2} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \vec{c}^{\,\mathrm{A}}_{m} \\ \tilde{c}^{\,\mathrm{B}}_{m} \end{pmatrix} = \begin{pmatrix} 0 & \langle \mathbf{r} \rangle^{\mathrm{B}}_{j_1} \\ 0 & \langle \mathbf{r} \rangle^{\mathrm{B}}_{j_2} \\ \langle \mathbf{r} \rangle^{\mathrm{A}}_{j_3} & -\langle \mathbf{r} \rangle^{\mathrm{B}}_{j_3} \\ \langle \mathbf{r} \rangle^{\mathrm{A}}_{j_4} & -\langle \mathbf{r} \rangle^{\mathrm{B}}_{j_4} \end{pmatrix} \begin{pmatrix} \vec{c}^{\,\mathrm{A}}_{m} \\ \tilde{c}^{\,\mathrm{B}}_{m} \end{pmatrix}.$$

Note that the eigenvalues are constrained onto the complex unit circle. With included noise, the computed eigenvalue may deviate slightly from this circle, and the monitor vectors computed from this conditions may increase or decrease systematically at the change from segment A to B.

Data:
orbit perturbations $\vec{r}_{jk}$ for $J \geq 4$ monitors, $K \geq 4M$ correctors, in M spatial dimensions
element topology (ordered list)
monitor subset $j_1 - j_4$
Result:
$\mu_m, \vec{R}_{jm}, \vec{D}_{km}$ for all given j, k, m;
residual decomposition error σ^2;
for $SEG \in \{A, B\}$ **do**
  $\mathcal{J}_{\mathrm{SEG}}, \mathcal{K}_{\mathrm{SEG}} \leftarrow$ element topology;
  $\langle \mathbf{r} \rangle_j^{\mathrm{SEG}} \leftarrow$ first 4 principal components of segment orbits $\langle \vec{r} \rangle_{jk}$ for $j \in \mathcal{J}_{\mathrm{SEG}}, k \in \mathcal{K}_{\mathrm{SEG}}$
end
$\mu_m, \vec{c}_m^{\,\mathrm{A}}, \vec{c}_m^{\,\mathrm{B}} \leftarrow$ set up and solve shift-ring PCA eigenvalue problem (eq. γ) for M eigenvalues;
for $SEG \in \{A, B\}$ **do**
  $\vec{c}_m^{\,\mathrm{A}} \leftarrow \vec{c}_m^{\,\mathrm{A}} \mathrm{e}^{\mathrm{i}\mu_m/2} \quad \forall m$;
  $\vec{R}_{jm} \leftarrow \langle \mathbf{r} \rangle_j^{\mathrm{SEG}} \vec{c}_m^{\mathrm{SEG}}$ for $j \in \mathcal{J}_{\mathrm{SEG}}$;
end
$\sigma^2 \leftarrow 0$;
for $k \leftarrow 1$ **to** K **do**
  for $j \leftarrow 1$ **to** J,$m \leftarrow 1$ **to** M **do**
    $E_{jkm} \leftarrow$ element topology;
  end
  $D_{km} \leftarrow$ set up and solve Corrector system using $\vec{R}_{jm}, E_{jkm}$ for all j, m;
  for $j \leftarrow 1$ **to** J,$m \leftarrow 1$ **to** M **do**
    $\sigma^2 \leftarrow \sigma^2 + \Re\left\{\vec{R}_{jm} E_{jkm} D_{km}^*\right\}$;
  end
end

Figure A.4.: PCA-enhanced global algorithm with complete replacement of monitor equation systems.

Therefore, we consider to modify the problem by replacing $\tilde{c}_m^{\,\mathrm{B}}$ with $\vec{c}_m^{\,\mathrm{B}}$ by the relation

$$\tilde{c}_m^{\,\mathrm{B}} = \mathrm{e}^{\mathrm{i}\mu/2} \vec{c}_m^{\,\mathrm{B}},$$

which in turn leads to

$$\begin{aligned} \vec{R}_{jm} &= \langle \mathbf{r} \rangle_j^{\mathrm{B}} \vec{c}_m^{\,\mathrm{B}} \mathrm{e}^{\mathrm{i}\mu_m \,\mathrm{sign}(s_j + 1/2 - s_3)} \quad \text{for } j \in \mathbb{J}_{\mathrm{B}} \\ &= \langle \mathbf{r} \rangle_j^{\mathrm{B}} \vec{c}_m^{\,\mathrm{B}} \begin{cases} \mathrm{e}^{\mathrm{i}\mu_m/2} & \text{for } s_j \geq s_3 \\ \mathrm{e}^{-\mathrm{i}\mu_m/2} & \text{for } s_j < s_3 \end{cases} \text{ and } j \in \mathbb{J}_{\mathrm{B}}. \end{aligned} \tag{A.9}$$

Then, one obtains a generalized eigenvalue problem

$$\mathrm{e}^{\mathrm{i}\mu_m/2} \begin{pmatrix} \langle \mathbf{r} \rangle_{j_1}^{\mathrm{A}} & 0 \\ \langle \mathbf{r} \rangle_{j_2}^{\mathrm{A}} & 0 \\ 0 & \langle \mathbf{r} \rangle_{j_3}^{\mathrm{B}} \\ 0 & \langle \mathbf{r} \rangle_{j_4}^{\mathrm{B}} \end{pmatrix} \begin{pmatrix} \vec{c}_m^{\,\mathrm{A}} \\ \vec{c}_m^{\,\mathrm{B}} \end{pmatrix} = \begin{pmatrix} 0 & \langle \mathbf{r} \rangle_{j_1}^{\mathrm{B}} \\ 0 & \langle \mathbf{r} \rangle_{j_2}^{\mathrm{B}} \\ \langle \mathbf{r} \rangle_{j_3}^{\mathrm{A}} & 0 \\ \langle \mathbf{r} \rangle_{j_4}^{\mathrm{A}} & 0 \end{pmatrix} \begin{pmatrix} \vec{c}_m^{\,\mathrm{A}} \\ \vec{c}_m^{\,\mathrm{B}} \end{pmatrix}. \tag{A.10}$$

In this modified generalized eigenvalue problem, the "unfitting" of the system's solution is balanced out between all split monitors, and threfore no systematic increase or decrease along s in monitor vectors is expected.

We conclude by stating that we can solve the one-turn transfer matrix problem directly without the intermediate step of computing segment transfer matrices. With PCA, the complete analysis routine simplifies to the following procedure.

- Obtain monitor and corrector subsets $\mathbb{J}_{\mathrm{A,B}}, \mathbb{K}_{\mathrm{A,B}}$ from split indices $j_1 - j_4$ and topology,
- perform PCA on the two "blocks" of the general response matrix corresponding to these sets,
- solve the modified generalized eigenvalue problem (A.10),
- reconstruct all monitor vectors $\vec{R}_{jm}$ by (A.8) and (A.9) and
- use the corrector equation system to solve for all D_{km}.
- If a known segment is given, use the invariant postprocessing algorithm to obtain $\hat{R}_{jm}$.

A sketch of this algorithm is given as algorithm A.4.

A.3.4. Construction of the bilinear system

One may display (4.1) in the following form for $M = 2$.

$$\begin{aligned} r_{jkw} &= \sum_m \Re\left\{R_{jmw} \mathrm{e}^{-\mathrm{i}S_{jk}\mu_m/2} A^*_{km}\right\} + d_{jw} b_k \\ &= \Re\left\{\vec{y}^\dagger_{jw} \mathbf{E}_{jk} \vec{x}_k\right\} \quad \text{with } \vec{x}_k = \begin{pmatrix} A_{k1} \\ A_{k2} \\ b_k \end{pmatrix}, \vec{y}_{jw} = \begin{pmatrix} R_{j1w} \\ R_{j2w} \\ d_{jw} \end{pmatrix} \text{ and } \mathbf{E}_{jk} = \begin{pmatrix} \mathrm{e}^{\mathrm{i}S_{jk}\mu_1/2} & & \\ & \mathrm{e}^{\mathrm{i}S_{jk}\mu_2/2} & \\ & & 1 \end{pmatrix}. \end{aligned}$$

This system can be expanded with zeros to

$$r_{jkw} = \Re\left\{\vec{y}^\dagger \mathbf{B}_{jkw} \vec{x}\right\} \quad \text{with } \vec{x} = \begin{pmatrix} \vec{x}_1 \\ \vdots \\ \vec{x}_K \end{pmatrix}, \vec{y} = \begin{pmatrix} \vec{y}_{11} \\ \vdots \\ \vec{y}_{J1} \\ \vec{y}_{12} \\ \vdots \\ \vec{y}_{J2} \end{pmatrix},$$

and $\mathbf{B}_{jkw}$ just containing one $\mathbf{E}_{jk}$ block at the position which corresponds to the rows for $\vec{y}_{jw}$ and the columns for $\vec{x}_k$ elements.

A.4. Accelerator properties

A.4.1. Topology for Towalski correctors

Towalski correctors (sec. 8.1) have been installed in the DELTA storage ring for the purpose of fast orbit feedback. The criteria for their positioning along the ring and further design studies and general construction information can be found in [66].

They are identified by labels with prefixes `shk` for horizontal and `svk` for vertical correctors. The following list shows the labels identifying correctors in relation to the monitors with the prefix `BPM` for increasing s position, starting from $s = 0$.

```
BPM01, svk01, shk01, BPM02, BPM03, svk02, shk02, BPM04, BPM05, BPM06, BPM07, BPM08,
BPM09, svk03, shk03, BPM10, BPM11, svk04, shk04, svk06, shk05, BPM12, BPM13, svk08,
shk06, BPM14, BPM15, BPM16, BPM17, BPM18, BPM19, svk10, shk08, BPM20, BPM21, svk11,
shk09, BPM22, BPM23, BPM24, BPM25, BPM26, BPM27, BPM28, BPM29, BPM30, BPM31, BPM32,
BPM33, BPM34, BPM35, BPM36, shk10, svk12, BPM37, BPM38, svk14, shk12, svk14b, shk12b,
BPM39, BPM40, BPM41, BPM42, svk16, shk14, BPM43, svk17, shk15, BPM44, BPM45, shk16,
svk18, BPM46, svk19, shk17, BPM47, BPM48, svk20, shk18, BPM49, BPM50, svk21, shk19,
BPM51, BPM52, BPM53, BPM54
```

The topology matrix constructed from this list is shown in Fig. A.5. The correctors are controlled by six power supplies that can be interfaced using TCP/IP communication [67, 68, 66].

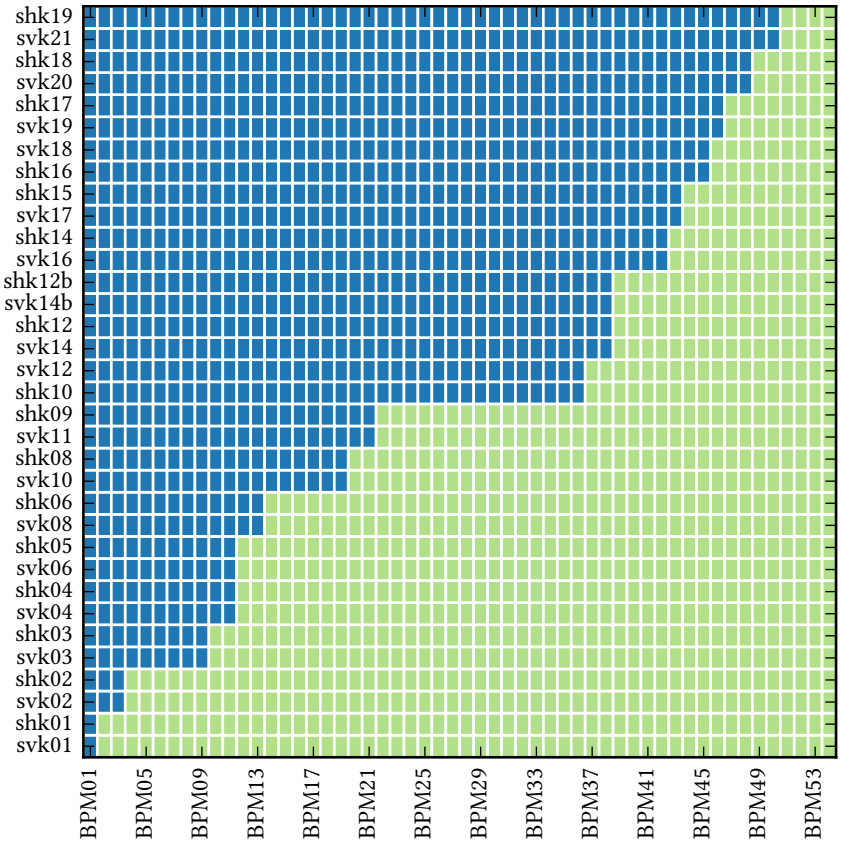

Figure A.5.: Topology matrix for Towalski correctors. Blue indicates $S_{jk} = -1$, green indicates $S_{jk} = +1$.

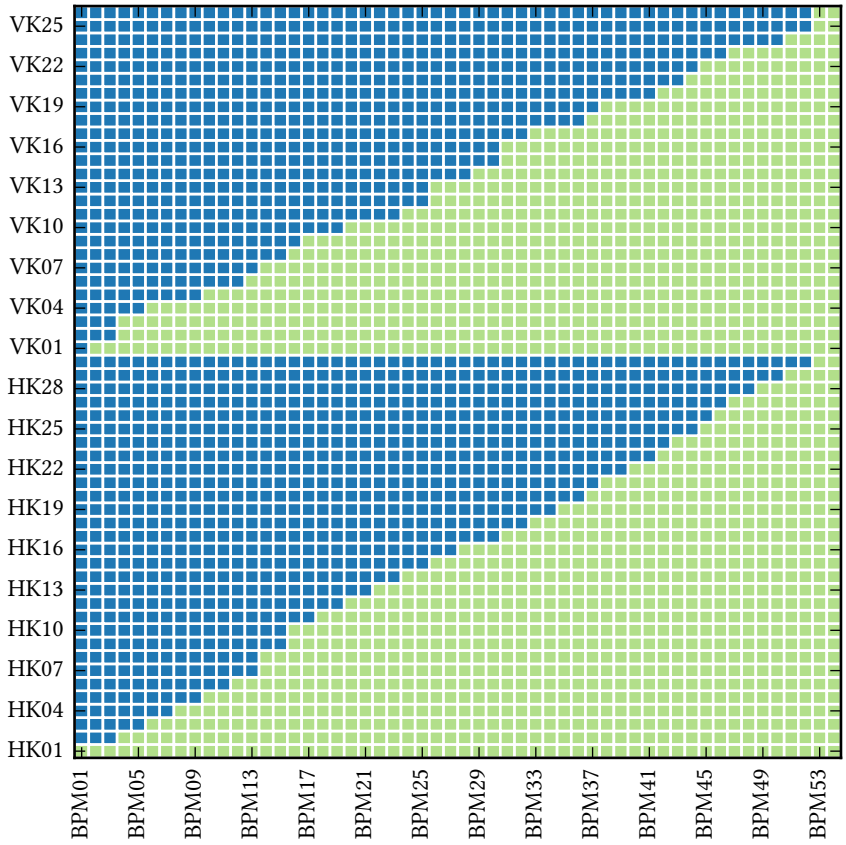

Figure A.6.: Topology matrix for DELTA standard correctors. The rows are sorted into blocks for horizontal respectively vertical correctors. Blue indicates $S_{jk} = -1$, green indicates $S_{jk} = +1$.

A.4.2. Topology for DELTA standard correctors

DELTA standard correctors are realized as additional coil windings on quadrupole magnets that can be controlled separately. Although this is often sufficient for beam steering, the change of corrector currents induces changes in higher-order multipole components of the magnet, as the magnet yoke is designed for quadrupole fields only.

They are identified by labels with prefixes HK for horizontal and VK for vertical correctors. The following list shows the labels identifying correctors in relation to the monitors with the prefix BPM for increasing s position, starting from $s = 0$.

```
HK01, BPM01, VK01, BPM02, BPM03, VK02, HK02, VK03, BPM04, BPM05, HK03, VK04,
BPM06, BPM07, HK04, BPM08, BPM09, VK05, HK05, BPM10, BPM11, HK06, BPM12, VK06,
BPM13, HK07, HK08, VK07, BPM14, BPM15, VK08, HK09, HK10, BPM16, VK09, BPM17,
HK11, BPM18, BPM19, HK12, VK10, BPM20, BPM21, HK13, BPM22, BPM23, VK11, HK14,
BPM24, BPM25, VK12, HK15, VK13, BPM26, BPM27, HK16, BPM28, VK14, BPM29, BPM30,
VK15, HK17, VK16, BPM31, BPM32, HK18, VK17, BPM33, BPM34, HK19, BPM35, BPM36,
VK18, HK20, BPM37, HK21, VK19, BPM38, BPM39, HK22, BPM40, BPM41, VK20, HK23,
BPM42, HK24, BPM43, VK21, BPM44, HK25, VK22, BPM45, HK26, BPM46, HK27, VK23,
BPM47, BPM48, HK28, BPM49, BPM50, VK24, HK29, BPM51, BPM52, VK25, HK30, VK26,
BPM53, BPM54
```

The topology matrix constructed from this list is shown in Fig. A.6. Note that only the relative ordering of HK, VK to BPM (and not HK to HK, VK to VK, HK to VK) is relevant.

A.4.3. DELTA standard responses (RID)

The following tables show the mapping between Response ID (RID) numbers and standard response matrix file names as stored in the DELTA control system. Further information about the recorded response matrices can be found in DELTA's electronic logbook [64].

RID	file name
1	response.060331-1
2	response.060331-2
3	response.060427-550MeV-1
4	response.060515-1
5	response.060921-NeueOptik-1
6	response.060925-1-Nullorbit
7	response.060928-1
8	response.061013-test
9	response.061016-1
10	response.061128-NeueOptik7
11	response.061128-NeueOptik8
12	response.061201-1
13	response.070108-1
14	response.070109-1
15	response.070110-1
16	response.070110-2-550MeV-2ns
17	response.070205-1
18	response.070206-550MeV-1.bak
19	response.070206-550MeV-1
20	response.070207-1-550MeV
21	response.070319-1
22	response.070319-2
23	response.070320-1
24	response.070320-2
25	response.070502-1
26	response.070504-1
27	response.070525-1
28	response.070625-1
29	response.070731-1
30	response.070801-1
31	response.070802-550MeV-linear-1
32	response.070802-551
33	response.070802-FEL-null
34	response.070803-1
35	response.070807-1
36	response.070814-1
37	response.070827-1
38	response.070903-1
39	response.071008-1
40	response.071030-1
41	response.071116-1
42	response.071130-1
43	response.080128-1

RID	file name
44	response.080225-1_ohneSAW
45	response.080227-550MeV-1
46	response.080505-1
47	response.080529-1
48	response.080602-1
49	response.080806-1
50	response.081013-1
51	response.081027-1
52	response.081124-1_ohne_SAW
53	response.081124-2-AP
54	response.081125-550MeV
55	response.081125-551_550MeV
56	response.081127-1
57	response.081127-2-550MeV
58	response.081128-1
59	response.090216-1-nullorbit
60	response.090216-1-nullorbit2
61	response.090216-2-nullorbit
62	response.090216-3-nullorbit
63	response.090217-nullorbit-1
64	response.090217-nullorbit-2
65	response.090422-flat-SAW-an-DC-aus
66	response.090511-1
67	response.090812-1
68	response.090821-1
69	response.090914-1
70	response.090915-1
71	response.091022-1
72	response.100202-1
73	response.100202-2
74	response.100325-1
75	response.100415-FEL-1
76	response.100429-1
77	response.100527-1
78	response.100608-1
79	response.100608-2
80	response.100610-1
81	response.100614-1
82	response.100621-1
83	response.100707-1
84	response.100707-2
85	response.100708-1
86	response.100720-1

RID	file name
87	response.100830-1
88	response.100831-1
89	response.101018-1
90	response.101110-nullorbit
91	response.101129-nullorbit
92	response.101130-1
93	response.101130-neueoptik
94	response.101201-1
95	response.101201-neueoptik-1
96	response.101203-1
97	response.110131-1
98	response.110201-1
99	response.110221-1
100	response.110304-1
101	response.110323-1
102	response.110324-Nullorbit
103	response.110506-1-BBC
104	response.110506-1
105	response.110509-1
106	response.110511-1
107	response.110705-1
108	response.110909-2
109	response.111108-1
110	response.111202-1-SAW
111	response.111205-1
112	response.120109-1
113	response.120109-2
114	response.120116-1
115	response.120117-1
116	response.120118-1
117	response.120120-1
118	response.120215-1
119	response.120216-1
120	response.120523-1
121	response.120612-1
122	response.120613-1
123	response.120814-1
124	response.120814-2
125	response.120815-1
126	response.120815-2
127	response.120817-1-donotuse
128	response.120817-2
129	response.120831-1

RID	file name
130	response.120831-2-ohneSAW
131	response.120831-Nullorbit-0.6Sext
132	response.120910-1
133	response.120913-1-Nullorbit
134	response.120914-1
135	response.130204-1-ohneSAW
136	response.130604-1-SB-9mA
137	response.130604-2-SB-16mA-ohneSAW
138	response.130710-1
139	response.131104-1
140	response.131128_SAW_Nullorbit-1
141	response.140106-1
142	response.140106-Nullorbit
143	response.140116-1
144	response.140314-1
145	response.140318-1
146	response.140328-1
147	response.140505-SAW-AN-1
148	response.140818-1
149	response.140821-1
150	response.140923-1
151	response.140924-1
152	response.140924-2
153	response.140924-3
154	response.140924-4
155	response.140930-1
156	response.141007-1
157	response.141007-2
158	response.141110-1
159	response.141208-1
160	response.150105-1
161	response.150121-alternative-optik-bolsinger
162	response.150202-1
163	response.150204-1
164	response.150205_mitSAW-1
165	response.150205_mitSAW-2
166	response.150317-1_550MeV_below
167	response.150318-550MeV-1
168	response.150318-550MeV-standart
169	response.150519-1_550MeV_AP15
170	response.151013-1
171	response.151103-1

A.4.4. MLS input data

MLS correctors are identified by labels with prefixes HS for horizontal and VS for vertical correctors. This data has been generously provided by [75]. The respective monitor and corrector lists, which were also sent, can be reconstructed by taking all respective elements of the list, without rearranging them.

The following list shows the labels identifying horizontal correctors in relation to the monitors with the prefix BPMZ for increasing s position, starting from $s = 0$.

```
BPMZ5K1RP, HS3M2K1RP, BPMZ6K1RP, BPMZ7K1RP, HS1M2K1RP, BPMZ1L2RP, BPMZ2L2RP,
BPMZ3L2RP, HS3M1L2RP, BPMZ4L2RP, BPMZ5L2RP, HS3M2L2RP, BPMZ6L2RP, BPMZ7L2RP,
BPMZ1K3RP, HS1M1K3RP, BPMZ2K3RP, BPMZ3K3RP, HS3M1K3RP, BPMZ4K3RP, BPMZ5K3RP,
HS3M2K3RP, BPMZ6K3RP, BPMZ7K3RP, HS1M2K3RP, BPMZ1L4RP, BPMZ2L4RP, BPMZ3L4RP,
HS3M1L4RP, BPMZ4L4RP, BPMZ5L4RP, HS3M2L4RP, BPMZ6L4RP, BPMZ7L4RP, BPMZ1K1RP,
HS1M1K1RP, BPMZ2K1RP, BPMZ3K1RP, HS3M1K1RP, BPMZ4K1RP
```

The following list shows the labels identifying vertical correctors in relation to the monitors for increasing s position, starting from $s = 0$.

```
BPMZ5K1RP, VS3M2K1RP, BPMZ6K1RP, VS2M2K1RP, BPMZ7K1RP, BPMZ1L2RP, VS2M1L2RP,
BPMZ2L2RP, BPMZ3L2RP, VS3M1L2RP, BPMZ4L2RP, BPMZ5L2RP, VS3M2L2RP, BPMZ6L2RP,
VS2M2L2RP, BPMZ7L2RP, BPMZ1K3RP, VS2M1K3RP, BPMZ2K3RP, BPMZ3K3RP, VS3M1K3RP,
BPMZ4K3RP, BPMZ5K3RP, VS3M2K3RP, BPMZ6K3RP, VS2M2K3RP, BPMZ7K3RP, BPMZ1L4RP,
VS2M1L4RP, BPMZ2L4RP, BPMZ3L4RP, VS3M1L4RP, BPMZ4L4RP, BPMZ5L4RP, VS3M2L4RP,
BPMZ6L4RP, VS2M2L4RP, BPMZ7L4RP, BPMZ1K1RP, VS2M1K1RP, BPMZ2K1RP, BPMZ3K1RP,
VS3M1K1RP, BPMZ4K1RP
```

The full topology matrix for MLS constructed from these lists is shown in Fig. 9.2.

A.4.5. BESSY II input data

BESSY II correctors are identified by labels with prefixes HS, HBM for horizontal and VS for vertical correctors. This data has been generously provided by [75]. The respective monitor and corrector lists, which were also sent, can be reconstructed by taking all respective elements of the list, without rearranging them.

The following list shows the labels identifying horizontal correctors in relation to the monitors with the prefix BPMZ for increasing s position, starting from $s = 0$.

```
HS4M2D1R, BPMZ6D1R, HBM2D1R, BPMZ7D1R, HS1MT1R, BPMZ1T1R, BPMZ2T1R, HBM1T1R,
BPMZ3T1R, HS4M1T1R, BPMZ4T1R, BPMZ5T1R, HS4M2T1R, BPMZ6T1R, HBM2T1R, BPMZ7T1R,
HS1MD2R, BPMZ1D2R, BPMZ2D2R, HBM1D2R, BPMZ3D2R, HS4M1D2R, BPMZ4D2R, BPMZ5D2R,
HS4M2D2R, BPMZ6D2R, HBM2D2R, BPMZ7D2R, HS1MT2R, BPMZ1T2R, BPMZ2T2R, HBM1T2R,
BPMZ3T2R, HS4M1T2R, BPMZ4T2R, BPMZ5T2R, HS4M2T2R, BPMZ6T2R, HBM2T2R, BPMZ7T2R,
HS1MD3R, BPMZ1D3R, BPMZ2D3R, HBM1D3R, BPMZ3D3R, HS4M1D3R, BPMZ4D3R, BPMZ5D3R,
HS4M2D3R, BPMZ6D3R, HBM2D3R, BPMZ7D3R, HS1MT3R, BPMZ1T3R, HBM1T3R, BPMZ3T3R,
HS4M1T3R, BPMZ4T3R, BPMZ5T3R, HS4M2T3R, BPMZ6T3R, HBM2T3R, BPMZ7T3R, HS1MD4R,
BPMZ1D4R, BPMZ2D4R, HBM1D4R, BPMZ3D4R, HS4M1D4R, BPMZ4D4R, BPMZ5D4R, HS4M2D4R,
BPMZ6D4R, HBM2D4R, HS1MT4R, BPMZ1T4R, BPMZ2T4R, HBM1T4R, BPMZ3T4R, HS4M1T4R,
BPMZ4T4R, BPMZ5T4R, HS4M2T4R, BPMZ6T4R, HBM2T4R, BPMZ7T4R, HS1MD5R, BPMZ1D5R,
BPMZ2D5R, HBM1D5R, BPMZ3D5R, HS4M1D5R, BPMZ4D5R, BPMZ5D5R, HS4M2D5R, BPMZ6D5R,
HBM2D5R, BPMZ7D5R, HS1MT5R, BPMZ1T5R, BPMZ2T5R, HBM1T5R, BPMZ3T5R, HS4M1T5R,
BPMZ4T5R, BPMZ5T5R, HS4M2T5R, BPMZ6T5R, HBM2T5R, BPMZ7T5R, HS1MD6R, BPMZ1D6R,
BPMZ2D6R, HBM1D6R, BPMZ3D6R, HS4M1D6R, BPMZ4D6R, HS4M2D6R, BPMZ6D6R, HBM2D6R,
BPMZ7D6R, HS1MT6R, BPMZ1T6R, BPMZ2T6R, HBM1T6R, BPMZ3T6R, HS4M1T6R, BPMZ4T6R,
BPMZ5T6R, HS4M2T6R, BPMZ6T6R, HBM2T6R, BPMZ7T6R, HS1MD7R, BPMZ1D7R, BPMZ2D7R,
HBM1D7R, BPMZ3D7R, HS4M1D7R, BPMZ4D7R, BPMZ5D7R, HS4M2D7R, BPMZ6D7R, HBM2D7R,
BPMZ7D7R, HS1MT7R, BPMZ1T7R, BPMZ2T7R, HBM1T7R, BPMZ3T7R, HS4M1T7R, BPMZ4T7R,
BPMZ5T7R, HS4M2T7R, BPMZ6T7R, HBM2T7R, BPMZ7T7R, HS1MD8R, BPMZ1D8R, BPMZ2D8R,
HBM1D8R, BPMZ3D8R, HS4M1D8R, BPMZ4D8R, BPMZ5D8R, HS4M2D8R, BPMZ6D8R, HBM2D8R,
```

```
BPMZ7D8R, HS1MT8R,  BPMZ1T8R, BPMZ2T8R, HBM1T8R,  BPMZ3T8R, HS4M1T8R, BPMZ4T8R,
BPMZ5T8R, HS4M2T8R, BPMZ6T8R, HBM2T8R,  BPMZ7T8R, HS1MD1R,  BPMZ1D1R, BPMZ2D1R,
HBM1D1R,  BPMZ3D1R, HS4M1D1R, BPMZ4D1R
```

The following list shows the labels identifying vertical correctors in relation to the monitors for increasing s position, starting from $s = 0$.

```
VS3M2D1R, BPMZ6D1R, BPMZ7D1R, VS2M2D1R, BPMZ1T1R, VS2M1T1R, BPMZ2T1R, BPMZ3T1R,
VS3M1T1R, BPMZ4T1R, BPMZ5T1R, VS3M2T1R, BPMZ6T1R, BPMZ7T1R, VS2M2T1R, BPMZ1D2R,
VS2M1D2R, BPMZ2D2R, BPMZ3D2R, VS3M1D2R, BPMZ4D2R, BPMZ5D2R, VS3M2D2R, BPMZ6D2R,
BPMZ7D2R, VS2M2D2R, BPMZ1T2R, VS2M1T2R, BPMZ2T2R, BPMZ3T2R, VS3M1T2R, BPMZ4T2R,
BPMZ5T2R, VS3M2T2R, BPMZ6T2R, BPMZ7T2R, VS2M2T2R, BPMZ1D3R, VS2M1D3R, BPMZ2D3R,
BPMZ3D3R, VS3M1D3R, BPMZ4D3R, BPMZ5D3R, VS3M2D3R, BPMZ6D3R, BPMZ7D3R, VS2M2D3R,
BPMZ1T3R, VS2M1T3R, BPMZ3T3R, VS3M1T3R, BPMZ4T3R, BPMZ5T3R, VS3M2T3R, BPMZ6T3R,
```

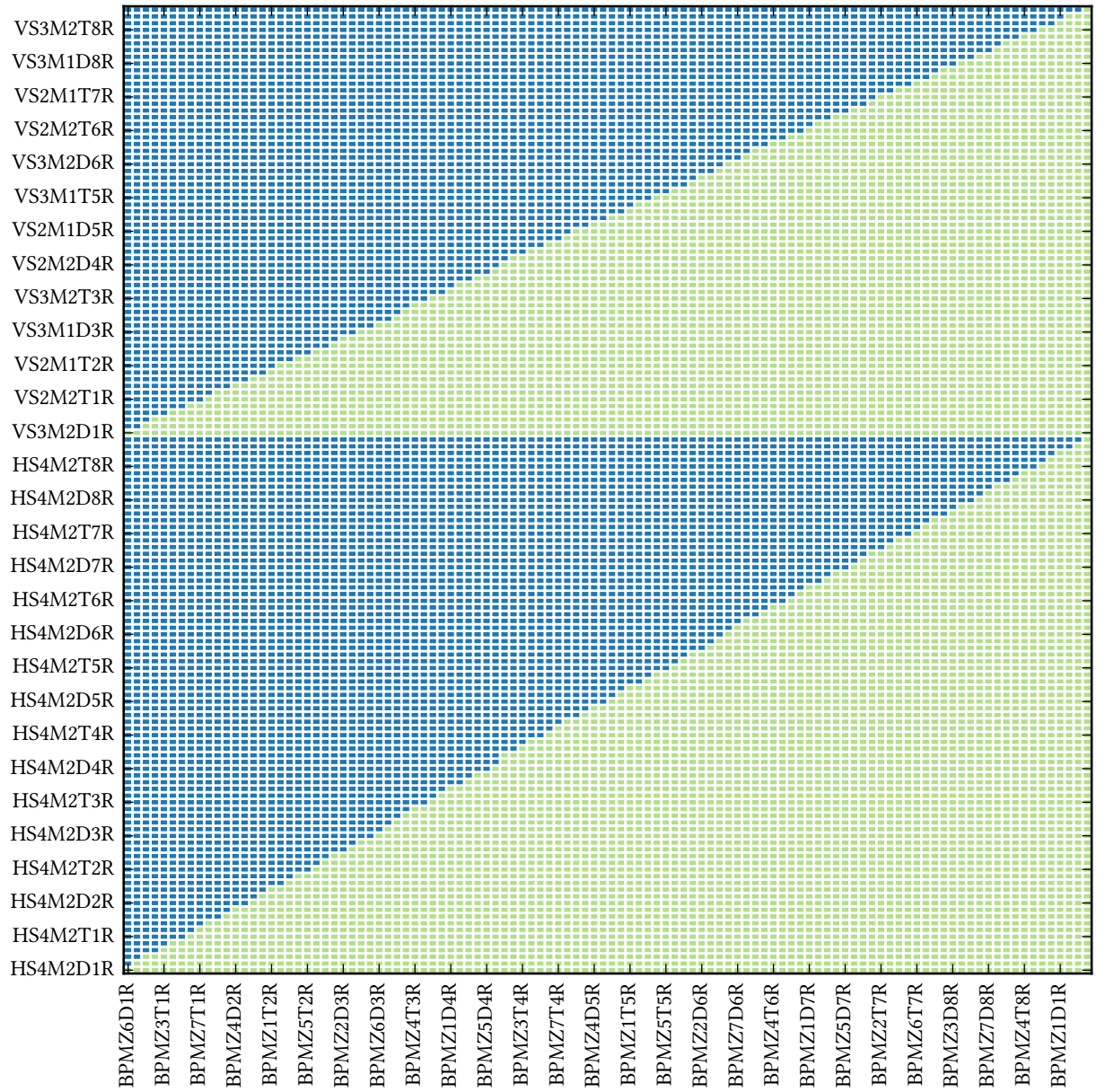

Figure A.7.: Topology matrix for BESSY II. The rows are sorted into blocks for horizontal respectively vertical correctors. Blue indicates $S_{jk} = -1$, green indicates $S_{jk} = +1$.

```
BPMZ7T3R, VS2M2T3R, BPMZ1D4R, VS2M1D4R, BPMZ2D4R, BPMZ3D4R, VS3M1D4R, BPMZ4D4R,
BPMZ5D4R, VS3M2D4R, BPMZ6D4R, VS2M2D4R, BPMZ1T4R, VS2M1T4R, BPMZ2T4R, BPMZ3T4R,
VS3M1T4R, BPMZ4T4R, BPMZ5T4R, VS3M2T4R, BPMZ6T4R, BPMZ7T4R, VS2M2T4R, BPMZ1D5R,
VS2M1D5R, BPMZ2D5R, BPMZ3D5R, VS3M1D5R, BPMZ4D5R, BPMZ5D5R, VS3M2D5R, BPMZ6D5R,
BPMZ7D5R, VS2M2D5R, BPMZ1T5R, VS2M1T5R, BPMZ2T5R, BPMZ3T5R, VS3M1T5R, BPMZ4T5R,
BPMZ5T5R, VS3M2T5R, BPMZ6T5R, BPMZ7T5R, VS2M2T5R, BPMZ1D6R, VS2M1D6R, BPMZ2D6R,
BPMZ3D6R, VS3M1D6R, BPMZ4D6R, VS3M2D6R, BPMZ6D6R, BPMZ7D6R, VS2M2D6R, BPMZ1T6R,
VS2M1T6R, BPMZ2T6R, BPMZ3T6R, VS3M1T6R, BPMZ4T6R, BPMZ5T6R, VS3M2T6R, BPMZ6T6R,
BPMZ7T6R, VS2M2T6R, BPMZ1D7R, VS2M1D7R, BPMZ2D7R, BPMZ3D7R, VS3M1D7R, BPMZ4D7R,
BPMZ5D7R, VS3M2D7R, BPMZ6D7R, BPMZ7D7R, VS2M2D7R, BPMZ1T7R, VS2M1T7R, BPMZ2T7R,
BPMZ3T7R, VS3M1T7R, BPMZ4T7R, BPMZ5T7R, VS3M2T7R, BPMZ6T7R, BPMZ7T7R, VS2M2T7R,
BPMZ1D8R, VS2M1D8R, BPMZ2D8R, BPMZ3D8R, VS3M1D8R, BPMZ4D8R, BPMZ5D8R, VS3M2D8R,
BPMZ6D8R, BPMZ7D8R, VS2M2D8R, BPMZ1T8R, VS2M1T8R, BPMZ2T8R, BPMZ3T8R, VS3M1T8R,
BPMZ4T8R, BPMZ5T8R, VS3M2T8R, BPMZ6T8R, BPMZ7T8R, VS2M2T8R, BPMZ1D1R, VS2M1D1R,
BPMZ2D1R, BPMZ3D1R, VS3M1D1R, BPMZ4D1R
```

The topology matrix constructed from these lists is shown in Fig. A.7.