Grey Box Identification for the Free Electron Laser FLASH exploiting Symmetries of the RF-System

Sven Pfeiffer ∗ Christian Schmidt ∗ Gerwald Lichtenberg ∗∗ Herbert Werner ∗∗

∗ DESY Hamburg, Notkestrasse 85, 22607 Hamburg, Germany (e-mail: sven.pfeiffer@desy.de, christian.schmidt@desy.de)
∗∗ Institute of Control Systems, Hamburg University of Technology, 21071 Hamburg, Germany (e-mail: Lichtenberg@tuhh.de, h.werner@tuhh.de)

Abstract:
Appropriate models are essential to achieve high performance control for particle accelerators, which are the basis for free electron lasers. This paper deals with a significant reduction of the model complexity by exploiting physical symmetries of the plant. It is shown how models obeying the inherent SO(2) symmetry of the cavities and the radio frequency subsystem are set up and their parameters are identified from data generated at the FLASH facility. Validation of the symmetric models gives results which are competitive to those of black box models but with a reduced number of parameters.

Keywords: Grey-Box Parameter Identification, Symmetries, Free Electron Lasers

1. INTRODUCTION

At the German Electron Synchrotron (DESY) in Hamburg a challenging European research project is conducted called XFEL - X-ray Free Electron Laser with the goal to supply laser light with a tunable wavelength in the X-ray range (10⁻¹⁰ m), (Brinkmann (2002)). The process uses a linear particle accelerator, which increases the energy of electrons by interaction with electromagnetic radio frequency (RF) fields inside so called cavities. These are high quality resonators, which are currently studied at the free electron laser FLASH at DESY. They are operated at a frequency of 1.3 GHz with field gradients up to 30 MV/m. At this plant, recently its own record was exceeded with a laser wavelength to 4.2 nm, which is in the far ultraviolet range.

In order to reach wavelengths in the X-ray range it is required to control the electro-magnetic fields inside the cavities with extremely high precision: the root mean square (RMS) values of the error between the setpoint and field gradient should be lower than 0.01 % in amplitude and 0.01° in phase. It was experimentally verified that a control scheme consisting of an iterative learning feedforward component and a linear MIMO feedback controller is able to cope with these requirements, (Kirchhoff et al. (2008); Langkowski et al. (2009)). So far, the design of the controller parameters of this scheme relies on a linear black box model which is identified from measurements at FLASH. For this approach, only engineering methods and tools have been used so far.

In this paper, we show how these methods can be enriched by physical insight, i.e. the fact that the accelerators display certain symmetries. Symmetries in linear dynamical systems has been exploited in many contexts, e.g. the behavioral framework, (Fagnani and Willems (1991)) or a more general nonlinear case for symmetries of oscillator networks (Barany and Colbaugh (1999); Barany (2001)), where the focus was on linear system identification. MIMO systems with a similar symmetry have been also investigated before, but in contrast to our case the signs of the cross coupling terms are the same, (Jugo and Arredondo (2005)).

The following questions serve as a guideline to successfully exploit symmetries for controller design:

(1) What are the symmetries of the accelerator system?
(2) How do these symmetries restrict state space models?
(3) How can symmetric model parameters be identified?
(4) Which kind of data preprocessing is required?
(5) Are symmetric models better than black box models?

The guiding questions also give an outline of this paper: The first question will be discussed in Section 2.1, the second in Section 3. Section 4 gives solutions for the remaining questions, especially to the fourth question in Section 4.1 and concludes with the results for experimental data to answer the last question.

2. FREE ELECTRON LASER FLASH

2.1 System Overview

The main components of the system can be found in Fig. 1. They are distributed in physical parts (master oscillator, vector modulator, analog-digital converters (ADC), digital-analog converters (DAC), klystron, cavities, down
sampling unit) and software components (calibration unit, set-point and feedforward tables, controller). The operation of this system can be roughly described as follows. The so-called master oscillator provides a constant RF signal at 1.3 GHz. With the vector modulator one is able to tune amplitude and phase of this signal, before it is amplified with the klystron and distributed in a waveguide system to cavities. These are resonators housing standing electromagnetic waves of the same frequency.

The signals from eight cavities enter a software module inside a Field Programmable Gate Array (FPGA) System. The main advantage of this is parallel high speed data processing and the possibility of reconfiguration without hardware changes. As the measurement equipment is very sensitive to cable movements or different cable lengths, a calibration of the signals is necessary. This is done by rotating the measured field vector of each cavity and by processing their sum. Using only one actuator for a cascaded system is done for economic reasons. Finally the set-points $r_1, r_Q$ are subtracted from the system outputs $y_I, y_Q$ and the resulting error enters the controller. The indices $I, Q$ denotes 'in-phase' and 'quadrature', which is common in accelerator physics.

At the controller output a feedforward signal $f_I, f_Q$ is added. Finally the input signals $u_I, u_Q$ are converted to an analog signal and applied to the vector modulator such that the controller can change phase and amplitude of the master oscillator. In the following parts of the system, which have a major impact on the system behavior, are described more detailed.

### 2.2 Cavity

The first part is the cavity, the so-called resonator. One can imagine it as a hollow tube consisting of nine cells as presented in Fig. 2. Into this tube electro-magnetic energy in form of microwaves at about 1.3 GHz is injected. As a result of its geometry, the cavities can house different modes of the electromagnetic field applied. The electric component of that field is responsible for the acceleration of the injected electron bunches. It is intended to transfer as much energy as possible from the field to the electrons; an optimal mode is shown in Fig. 2. The arrows indicate the electrical component in the direction of the particle movements of the radio frequency field. The length indicates the field amplitude. All arrows together represent the so-called $\pi$-mode, which is used for the acceleration.

![Fig. 2. Cut through a cavity with electric field vectors](image)

The requirements for controlling the RF fields are determined by the physical processes of laser generation. Roughly, the particles must have the same energy at the end of the acceleration. Therefore an RF control system is used to regulate the fields in a disturbed environment. A typical pulse pattern divided into three stages is shown in Fig. 3, which are filling, flat top and decay.

![Fig. 3. Pulse operation of FLASH](image)

At first the power is increased in the filling phase to its maximum in order to reach a desired level of a field gradient as fast as possible. In the following flat top phase, electron bunches are injected. This is the most important step, in which the gradient has to be kept constant. At the end of the flat top phase power is turned off to avoid overheating and the gradient decays, until the next pulse arrives.

### 2.4 Vector Modulator

Right after the DAC at plant input side the control signal enters the vector modulator. This unit can change the amplitude and phase of the signal produced master oscillator. It controls the real and imaginary part of the complex signal vector. This is done by a multiplication of the real
and imaginary part of the signal with the reference signal and its \((\pi/2)\)-shifted modification, respectively, according to

\[
f(t) = \cos (2\pi f_{RF} \cdot t + \phi_{RF}),
\]

\[
S(t) = u_{I1}(t) \cdot \cos (2\pi f_{RF} \cdot t + \phi_{RF}) + u_{Q1}(t) \cdot \sin (2\pi f_{RF} \cdot t + \phi_{RF}),
\]

where \(f(t)\) denotes a sinusoidal signal coming from the master oscillator and \(u_{I1}(t), u_{Q1}(t)\) are the real and imaginary part of the control signal. The resulting output of the vector modulator is \(S(t)\). The signal \(f(t)\) has frequency \(f_{RF}\) and initial phase shift \(\phi_{RF}\).

2.5 Amplitude and Phase Detection

In a first step the measured 1.3 GHz radio frequency (RF) signal is converted down to an intermediate frequency signal (IF). This is performed by an analog mixer. The underlying mathematical operation is similar to the conversion in the vector modulator described previously. This process is shown in Fig. 4.

![Principle of down-conversion from RF to IF signal](image)

The measured RF-signal is multiplied by a harmonic signal, generated by a local oscillator (LO). This LO-signal contains a harmonic component of about 250 kHz offset from the carrier frequency of 1.3 GHz. After filtering with a low pass filter, the resulting intermediate frequency (IF) signal contains RF-field phase and amplitude information.

The measurement algorithm is visualized in Fig. 5. As the IF-signal is digitized with 1 MHz, 4 sampling instants per period are calculated. Under the assumption that the signal has constant amplitude and phase during measurement, two adjacent samples are phase shifted by 90° with 1μs time difference.

![Sampling method](image)

Two consecutive samples represent the complex field vector by its real and imaginary part. From time step to time step the complex field vector is rotated by 90° and must be roughly merged by different algorithms. There is always a positive real and imaginary part followed by a negative real and imaginary part. This has to be corrected together with the phase offset from different cable lengths to get the output of the system. For detailed information about the used hardware, see Schilcher (1998).

The system has two inputs \(U_I(z)\) and \(U_Q(z)\) and two outputs \(Y_I(z)\) and \(Y_Q(z)\), which can also be represented as amplitude and phase. The resonant frequency of the cavity differs from the frequency of the master oscillator due to mechanical deformations. This leads to an off-resonant cavity, which can be roughly represented as a phase-error and finally leads to off-diagonal elements in the transfer function. To see this, define a complex vector by a real and imaginary part and a phase offset \(\delta\). All phase offsets can be represented by a rotation matrix

\[
R = \begin{bmatrix}
\cos(\delta) & -\sin(\delta) \\
\sin(\delta) & \cos(\delta)
\end{bmatrix}.
\]

This matrix parametrizes the so-called special orthogonal group \(SO(2)\) of dimension two, (Gilmore (1974)). A linear symmetric-discrete time model has to have the structure

\[
\begin{bmatrix}
Y_I(z) \\
Y_Q(z)
\end{bmatrix} = \begin{bmatrix}
G_{II}(z) & G_{IQ}(z) \\
G_{QI}(z) & G_{QQ}(z)
\end{bmatrix} \begin{bmatrix}
U_I(z) \\
U_Q(z)
\end{bmatrix}.
\]

This fact allows to use dedicated strategies for system identification, which are described in the following.

3. SYMMETRIC MODELS

In this Section, a linear MIMO model with inputs \(u_I, u_Q\) and outputs \(y_I, y_Q\) is presented which fulfills the following requirements:

1. it has a low pass characteristic,
2. it obeys the symmetry discussed in Section 2,
3. it has the ability to model at least one resonance.

3.1 First Order Model

The simplest model which fulfills the first requirement is a first order model

\[
x(k+1) = \lambda x(k) + [b_1 \ b_2] \cdot (u_I(k) \ u_Q(k))^T,
\]

\[
y(k) = [c_1 \ c_2]^T x(k),
\]

with \(b_1, b_2, \lambda \in \mathbb{R}\) and a corresponding transfer function matrix

\[
G(z) = \begin{bmatrix}
G_{II}(z) & G_{IQ}(z) \\
G_{QI}(z) & G_{QQ}(z)
\end{bmatrix} = \begin{bmatrix} c_1 \ c_2 \end{bmatrix}^T \frac{1}{z-\lambda} [b_1 \ b_2],
\]

leading to

\[
\begin{bmatrix}
Y_I(z) \\
Y_Q(z)
\end{bmatrix} = \begin{bmatrix}
c_1 b_1 & c_1 b_2 \\
c_2 b_1 & c_2 b_2
\end{bmatrix} \begin{bmatrix}
U_I(z) \\
U_Q(z)
\end{bmatrix}.
\]

Fulfilling the \(SO(2)\) symmetry demands that the diagonal transfer functions have to be equal

\[
G_{II}(z) = G_{QQ}(z) \forall z \Rightarrow c_1 b_1 = c_2 b_2,
\]

and for the off-diagonal transfer functions

\[
G_{QI}(z) = -G_{IQ}(z) \forall z \Rightarrow c_2 b_1 = -c_1 b_2
\]

holds. From (3) and (4) it follows that

\[
b_2^2 = -b_1^2,
\]

which is possible with model (2) and real valued parameters only for the trivial case \(b_1 = b_2 = 0\). Thus, the
simplest low pass symmetric model has to be at least of second order. Next a general model including also the third requirement is discussed.

### 3.2 Grey-Box Models with SO(2) Symmetry

We consider the discrete time model

\[ x(k+1) = \Phi x(k) + \Gamma u(k), \]
\[ y(k) = C x(k) + D u(k), \]

with sample time \( T \).

In the following, the direct feedthrough \( D \) is always set to zero as this turns out to be always the case when analyzing RF input and output signals of the particle accelerator FLASH. The approach can be extended for time delays by augmenting the state vector.

First consider a diagonal system matrix

\[ \Phi_r = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_1 \end{bmatrix}, \]

and choose the input matrix

\[ \Gamma_r = \begin{bmatrix} b_1 & -b_2 \\ b_2 & b_1 \end{bmatrix}, \]

and the output matrix

\[ C_r = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \]

The resulting discrete time transfer function matrix

\[ G(z) = C_r(zI - \Phi_r)^{-1}\Gamma_r = \begin{bmatrix} b_1 & -b_2 \\ \frac{z - \lambda_1}{b_2} & \frac{z - \lambda_1}{b_1} \end{bmatrix}, \]

obeys – as desired – the SO(2) symmetry.

Now we take a look at a pair of complex eigenvalues and use a modal state space representation. Similar to the real part one can show that it is necessary to define at least four states for the symmetric model. To see this, define a generic model having one complex pair of eigenvalues of its system matrix

\[ \tilde{\Phi}_c = \begin{bmatrix} \sigma & \omega \\ -\omega & \sigma \end{bmatrix}, \]

and input and output matrices

\[ \tilde{\Gamma}_c = \begin{bmatrix} b_1 & \tilde{b}_2 \\ \tilde{b}_3 & b_1 \end{bmatrix}, \quad \tilde{C}_c = \begin{bmatrix} c_1 & c_2 \\ c_3 & c_4 \end{bmatrix}. \]

The transfer function is

\[ \tilde{G}_c(z) = \tilde{C}_c(zI - \tilde{\Phi}_c)^{-1}\tilde{\Gamma}_c = \begin{bmatrix} D_{11}(z) & D_{1Q}(z) \\ D_{Q1}(z) & D_{QQ}(z) \end{bmatrix} \begin{bmatrix} N(z) \\ N(z) \end{bmatrix}. \]

The matrix \( G_c(z) \) has a common denominator polynomial

\[ N(z) = z^2 - 2\sigma z + \sigma^2 + \omega^2 \]

and four different numerator polynomials

\[ D_{11}(z) = (c_1\tilde{b}_1 + c_2\tilde{b}_3) \cdot (z - \sigma) + (c_1\tilde{b}_3 - c_2\tilde{b}_1) \cdot \omega, \]
\[ D_{QQ}(z) = (c_3\tilde{b}_2 + c_4\tilde{b}_4) \cdot (z - \sigma) + (c_3\tilde{b}_4 - c_4\tilde{b}_2) \cdot \omega, \]
\[ D_{1Q}(z) = (c_1\tilde{b}_2 + c_2\tilde{b}_4) \cdot (z - \sigma) + (c_1\tilde{b}_4 - c_2\tilde{b}_2) \cdot \omega, \]
\[ D_{Q1}(z) = (c_3\tilde{b}_1 + c_4\tilde{b}_3) \cdot (z - \sigma) + (c_3\tilde{b}_3 - c_4\tilde{b}_1) \cdot \omega. \]

The equations \( D_{11}(z) = D_{QQ}(z) \) and \( D_{1Q}(z) = -D_{Q1}(z) \) must hold for all \( z, \sigma \) and \( \omega \) for SO(2) symmetry. This leads to only two non-trivial cases to combine real values of all \( \tilde{b} \) and \( c \) to fulfill symmetry. One for antisymmetric off-diagonals

\[ \tilde{\Gamma}_c = \begin{bmatrix} b_1 & -\tilde{b}_2 \\ \tilde{b}_2 & b_1 \end{bmatrix}, \quad \tilde{C}_c = \begin{bmatrix} c_1 & c_2 \\ c_2 & -c_1 \end{bmatrix}. \]

Numerical experiments show that the results of the grey-box parameter identification, which is described later, could be improved if the state space is extended by another complex conjugated pole pair. Therefore the system has to be extended like in the case for a real eigenvalue. Thus, the number of parameters is not changed, but the order of the system.

The extended system matrix can be written with four states in modal form as

\[ \Phi_e = \begin{bmatrix} \sigma & \omega & 0 & 0 \\ -\omega & \sigma & 0 & 0 \\ 0 & 0 & \sigma & \omega \\ 0 & 0 & -\omega & \sigma \end{bmatrix}. \]

With input matrix

\[ \Gamma_e = \begin{bmatrix} b_{11} & b_{12} & b_{21} & b_{22} \\ -b_{21} & -b_{22} & b_{11} & b_{12} \end{bmatrix}^T \]

and output matrix

\[ C_e = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \]

the whole 6-th order discrete time model with low pass character and one resonance frequency

\[ x(k+1) = \begin{bmatrix} \Phi_e & 0 \\ 0 & \Phi_e \end{bmatrix} x(k) + \begin{bmatrix} \Gamma_e \\ \Gamma_e \end{bmatrix} u(k), \]
\[ y(k) = \begin{bmatrix} C_r & C_e \end{bmatrix} x(k) \]

has two real eigenvalues and two complex-conjugated pairs of eigenvalues and obeys the SO(2) symmetry.

### 4. GREY BOX PARAMETER IDENTIFICATION

For the FLASH system, parameter identification is divided into two steps, where in the first step the location of the low frequency pole which gives the low pass characteristic is identified together with its input matrix, using pseudo
random binary (PRB) signals, and in the second step, the frequency and damping of one or more complex conjugated pole pairs are determined by an adaptive excitation scheme that starts with a frequency scan of the system. For both steps data preconditioning is done in the same way which is not standard and will be described in the following.

4.1 Data Preconditioning and Detrending

At a given set-point (amplitude and phase), a linear state space model has to be identified. Assume a stable linear system with a persistent zero mean excitation signal during the flat top phase. If the excitation starts, then the output should also start at the signal without excitation. If the excitation stops, the output should end after some time at the same signal without excitation. But with amplitude and phase drifts, i.e. \( y_1 \) and \( y_0 \) drifts, it is not the case for this system. The obtained output will be detrended such that this behavior is fulfilled. It is assumed here, that the plant follows the excitation signal very fast. If the excitation signal stops, i.e. zero excitation, the steady state is reached. For the output signal it is necessary to use a nonlinear detrend.

Fig. 6. Detrending signals \( u_1 \) and \( y_1 \)

Fig. 6 illustrates the procedure for preconditioning and detrending the datasets:

1. Apply the feedforward signal without added excitation signal (FF wo. Exc.) to the machine.
2. Machine is operated in pulsed mode, such that a series of 20-40 output signals (I wo. Exc.) belonging to input pulses without excitation can be taken.
3. Add an excitation signal to the feedforward signal (FF with Exc.).
4. Take a series of 60-80 output signals (I with Exc.) with excitation (Fig. 7).
5. A maximum likelihood estimate will be computed from all output data without and with excitation by the MATLAB function \( \text{nlmle} \) to get two signals, one without and one with added excitation.
6. Calculate the difference signal (\( \Delta I \)) between the maximum likelihood estimates with and without excitation computed in Step 5.
7. Linearly detrend the signal \( \Delta I \) of Step 6 resulting in \( \Delta I_{\text{Detrend}} \).

This procedure is necessary to generate the preconditioned set of signals used by the following parameter estimation procedure.

4.2 Parameter Identification

With the system identification toolbox of MATLAB, the structure of a linear grey-box model can be defined by the function \( \text{idgrey} \) and used to estimate a grey-box model by the \( \text{pem} \) function. The first goal is estimating the real pole (5) for low pass behavior and its corresponding input matrix (6), thus the model has the form

\[
x(k+1) = \Phi x(k) + \Gamma u(k), \quad y(k) = C x(k).
\]

(12)

In this first step, all parameters are unknown and we set the pole and the input matrix as previously described, but these parameters have to be initialized reasonably. The diagonal elements should dominate, so initialize \( b_1 \) to a higher value, e.g. 0.1, \( b_2 \) as zero or close to zero and the real pole at around 200 Hz. The estimation of the parameters of the simple grey-box model with two real eigenvalues, e.g., leads for some measured data to values

\[
\lambda_1 = 0.9996, \quad b_1 = 0.0028, \quad b_2 = 0.0014.
\]

The next step is to use high frequency excitation signals which allow to excite resonances that are difficult to find by PRBS. The bandwidth of the excitation signal should be limited to the region of interest. This can be done by using chirp sine signals with a linear increasing frequency. By a rough chirp sine signal with a spectrum shown in Fig. 8 the region of possible resonances will be first scanned. The analysis of the output spectrum detects small peaks which are verified by chirp sine signals with a smaller bandwidth as also shown in Fig. 8.

After filtering and detrending these datasets, one can identify at least one complex conjugated pole pair by augmenting the states and matrices from (12) to the structure (11). If there are additional resonant peaks, the model is further extended by state equations of the form (11) with (8)–(10). The old ones will be fixed and the identification of a second complex pole starts at the
Section 5. CONCLUSION

This paper shows a method for system identification of pulsed particle accelerators with bounded excitation time by exploiting symmetries of the RF-system. The excitation for a linear system identification can be done only during the flattop phase. As the high frequency characteristic is only roughly estimated by black box identification, a grey box model is used to cope with these limitations. The system identification procedure is split into two parts, one for the low frequency characteristic and one for the high frequencies. Due to jitter on the sampling of the real and imaginary part of the field gradient it is useful to assume a SO(2) symmetry. By analyzing the resulting grey box model one can say, that the static part is also SO(2) symmetric, whereas the transient part is not. But nevertheless, the grey box model is especially for the high frequencies more accurate as the black box model. The resulting symmetric model is already successfully used during operation of the FLASH plant for fixed order linear controller design.

REFERENCES