Chapter 1
Identification for robust control of complex systems: Algorithm and motion application

Tom Oomen and Maarten Steinbuch

Abstract Increasing performance demands in control applications necessitate accurate modeling of complex systems for control. The aim of this chapter is to develop a new system identification algorithm that delivers models that are suitable for subsequent robust control and can be reliably applied to complex systems. To achieve this, an identification algorithm is developed that delivers system model in terms of recently developed coprime factorizations and thereby extends classical iterative procedures to the closed-loop case. These coprime factorizations have important advantages for uncertainty modeling and robust controller synthesis of complex systems. A numerically optimal implementation is presented, which relies on orthonormal polynomials with respect to a data-dependent discrete inner product. Experimental results on a nanometer-accurate positioning system confirm that the algorithm is capable of delivering the required coprime factorizations and the implementation is numerically reliable, which is essential for complex systems as common implementations suffer from severe ill-conditioning.

1.1 Introduction
Increasing performance demands in control applications necessitate accurate models, and hence the development of system identification algorithms for complex systems. Important examples are nanometer-accurate positioning systems, which are
Identification for robust control of complex systems

for instance used in the production of integrated circuits (ICs). Here, increased performance requirements necessitate taking into account high frequent flexible dynamical behavior [19]. Furthermore, for such motion systems, six inputs and six outputs are traditionally used for control in six degrees-of-freedom. However, more actuators and sensors are necessary in the near future to actively compensate for flexible dynamical behavior, as is argued in [28]. Hence, models for controlling next-generation positioning systems are envisaged to be of high order and to have large input and output dimensionality.

Several approaches have been developed to connect system identification and robust controller design, see [23], [17], [32], [13]. Especially, frequency domain approaches, see [31] for a general overview and [7], [26] for robust control related approaches, have led to successful results in traditional motion control, see [41], [8], [36], [28].

Although identification and robust control have been able to enhance robust performance in traditional motion systems, existing techniques are inadequate for complex systems. These complex systems include motion systems with high-order dynamics and large input and output dimensionality. At least three interrelated aspects are highly relevant for identifying complex systems for control:

1. numerical aspects for identifying a nominal model,
2. algorithmic efficiency for identifying a nominal model, and
3. algorithmic aspects related to the model uncertainty structure in robust control design.

Related to aspect 3, model uncertainty structures for robust control have been recently refined to enable a non-conservative and efficient synthesis. First, traditional additive and multiplicative structures have been extended towards coprime-factor based uncertainty structures in, e.g., [21]. These coprime-factor based structures have been further extended towards dual-Youla uncertainty structures, see, e.g., [20], [2], [7], [10]. Such dual-Youla structures are particularly suitable for uncertainty modeling in closed-loop situations. However, a common aspect in these uncertainty structures is that the multivariable situation typically involves a highly structured and frequency weighted uncertainty model, as is confirmed in [41] for the additive case and in [8] for the dual-Youla case. Recently, a new coprime factorization has been proposed in [26] that extends dual-Youla type of uncertainty structures towards the non-conservative use of unstructured uncertainty models. Thus, it automatically provides a suitable multivariable frequency weighting of the uncertainty block. The key advantage is a subsequent non-conservative robust controller synthesis, see, e.g., [11], since the uncertainty structure is \( \mu \)-simple [29] for any input-output system dimension. However, the new coprime factorization in [26] cannot be directly identified using existing techniques, including [18]. In the present chapter, it is aimed to develop a frequency domain algorithm for identifying the coprime factorization in [26].

In view of aspect 1 mentioned above, frequency domain system identification is numerically often ill-conditioned, as is evidenced by the development of approaches to enhance numerical conditioning, including [30], [14], [43]. Since the numerical
1.1 Introduction

Conditioning typically deteriorates for increasing model complexity, the development of numerically reliable algorithms is particularly important for complex systems. A similar reason holds for aspect 2, i.e., the development of efficient algorithms, since the computational time often grows significantly for increasing model complexity.

The main contribution of this chapter is a new algorithm for identifying complex systems for robust control design. The algorithm generalises well-established results for the open-loop case, including [34] to the closed-loop case, and

1. is numerically optimal by exploiting results in orthonormal basis functions with respect to a data-based discrete inner product, see [3] for details;
2. can be efficiently implemented, both in the required number of iterations and the computational load per iteration that can be rendered $O(N)$, where $N$ is the data length;
3. is suitable for efficient robust controller synthesis due to the internal coprime-factor based model structure, enabling a fast and nonconservative robust controller synthesis, exploiting the results in [26].

Although the results presented in this chapter aim at closed-loop identification of a certain coprime factorization, the open-loop case is directly recovered as a special case. As a result, the proposed algorithm also enables the numerically reliable frequency domain identification of open-loop systems. In addition, so-called control-relevant weighting functions are used in this chapter. This can directly be replaced by any user-chosen weighting function. Similarly, the approach is presented for discrete time identification. The presented approach can be directly adapted towards continuous time model identification. Finally, the approach is specifically aimed towards handling complex systems. Of course, the approach can also be applied to relatively low-order systems with smaller input-output dimensionality.

This chapter is organised as follows. In Sec. 1.2, the role of coprime factorizations for modeling uncertainty in complex systems is established, revealing the advantages of recently developed coprime factor-based uncertainty structures. In addition, the considered robust control framework is outlined in Sec. 1.2. Then, in Sec. 1.3, new algorithms for identifying coprime factorizations are presented. The presented algorithms generalize commonly used SK-iterations to the closed-loop case. Next, a numerically optimal implementation is presented in Sec. 1.4, which is based on orthonormal polynomials that are orthonormal with respect to a certain data-dependent inner product. In Sec. 1.5, the proposed approach is applied to an industrial wafer stage system. Finally, conclusions are provided in Sec. 1.6.

_Notation._ The pair $\{N, D\}$ is an RCF of $P$ implies that $[40] P = ND^{-1}$; $N, D \in \mathcal{RH}_\infty$; and $\exists X, Y \in \mathcal{RH}_\infty$ such that $XN + YD = I$. Throughout, $P$ is assumed to have $n_y$ outputs and $n_u$ inputs. To facilitate the notation, the dimensions of matrices are often omitted if no ambiguity can arise.
1.2 Coprime factor identification for refined uncertainty structures in robust control

1.2.1 Robust control framework

The general $\mathcal{H}_\infty$ norm-based criterion

$$J(P, C) := \|WT(P, C)V\|_\infty$$

is considered, where

$$T(P, C) = \begin{bmatrix} P \\ I \end{bmatrix} (I + CP)^{-1} \begin{bmatrix} C \\ I \end{bmatrix},$$

where $T(P, C) : \begin{bmatrix} r_2 \\ r_1 \end{bmatrix} \mapsto \begin{bmatrix} y \\ u \end{bmatrix}$, see Fig. 1.1. The criterion (1.1) in conjunction with the four-block encompasses many relevant $\mathcal{H}_\infty$-design problems, including the loop-shaping approach in [21] and typical mixed-sensitivity problems. Throughout, it is assumed that $W = \text{diag}(W_y, W_u)$, $V = \text{diag}(V_2, V_1)$, and $W, V, W^{-1}, V^{-1} \in \mathcal{RH}_\infty$.

![Feedback interconnection](image)

The criterion (1.1) is formulated such that it is to be minimized for $P_0$, i.e.,

$$C^0 = \arg\min_C J(P_0, C).$$

The key idea in robust control is to represent the unknown system $P_0$ by a model set $\mathcal{P}$ such that

$$P_0 \in \mathcal{P}.$$  \hspace{1cm} (1.4)

Consequently, the robust performance controller design

$$C^{RP} = \arg\min_C J_{WC}(\mathcal{P}, C),$$

where

$$J_{WC}(\mathcal{P}, C) := \sup_{P \in \mathcal{P}} J(P, C),$$

leads to the performance guarantee.
1.2 Coprime factor identification for refined uncertainty structures in robust control

\[ \mathcal{J}(P_o, C^{RP}) \leq \mathcal{J}_{WC}(\mathcal{P}, C^{RP}). \quad (1.7) \]

A key observation is that the resulting performance guarantee in (1.7) hinges on the shape and size of the model set \( \mathcal{P} \). Typically, this model set is structured as a general LFT, i.e., the model set \( \mathcal{P} \) is constructed by connecting an \( \mathcal{H}_\infty \)-norm-bounded perturbation \( \Delta_u \in \Delta_u \subseteq \mathcal{R}\mathcal{H}_\infty \) to the nominal model as

\[ \mathcal{P} = \{ P | P = F_u(\hat{H}(\hat{P}), \Delta_u), \Delta_u \in \Delta_u \} , \quad (1.8) \]

where \( \hat{H}(\hat{P}) \) contains the nominal model \( \hat{P} \) and the uncertainty structure, as is elaborated on in the next section. Throughout, the model uncertainty set

\[ \Delta_u := \{ \Delta_u | \| \Delta_u \|_\infty \leq \gamma \} \quad (1.9) \]

is considered, where it is assumed that \( \Delta_u \) contains multivariable operators with suitable dimensions.

Two requirements are imposed on the model set in view of high performance robust control:

1. The model set \( \mathcal{P} \) should lead to a small bound (1.7), and
2. The model set should have low complexity to enable an efficient and non-conservative robust control synthesis.

To anticipate on the results in the next sections, a coprime factor-based approach will be pursued to address both aspects, since

- it enables connecting the size of model uncertainty and the control criterion, addressing Requirement 1, and
- it enables the non-conservative use of unstructured model uncertainty, enabling an efficient and non-conservative robust controller synthesis in view of Requirement 2.

### 1.2.2 Identification for robust control approach

To achieve high robust performance, i.e., a small bound in (1.7), a model set is identified that is robust control relevant. To obtain a general framework, the robust-control-relevant identification criterion

\[ \min_{\mathcal{P}} \mathcal{J}_{WC}(\mathcal{P}, C^{\text{exp}}), \quad \text{subject to (1.4)} \]

is considered as in [7], [26]. In (1.10), \( C^{\text{exp}} \) denotes the controller that is present during the identification experiment. Indeed, many systems are open-loop unstable or need to have suitable feedback controller implemented for safety reasons. As such, the criterion in (1.10) explicitly takes the experimental conditions into account. In
addition, note that the controller $C^{\text{exp}}$ is often updated iteratively, which is well-established in iterative identification and control [35], [7], [1]. Note that (1.10) also encompasses the open-loop situation, since in case $\hat{P}$ is stable, then $C^{\text{exp}} = 0$ is an internally stabilizing controller.

Next, by evaluating the performance of the feedback interconnection of the controller $C^{\text{exp}}$ and all candidate models $P$ in $\mathcal{P}$ for the general uncertainty description (1.8),

$$\mathcal{J}_{\text{WC}}(\mathcal{P}, C^{\text{exp}}) = \sup_{\Delta \in \Delta_u} \| \mathcal{F}_u(\hat{M}, \Delta_u) \|_{\infty}$$

where $\hat{M}(\hat{H}, C^{\text{exp}}) = \left[ \begin{array}{c|c} \hat{M}_{11} & \hat{M}_{12} \\ \hline \hat{M}_{21} & \hat{M}_{22} \end{array} \right]$. (1.12)

The key point in (1.10) is that in general $\hat{M}_{11} \neq 0$ for many classical uncertainty structures, including additive uncertainty in which case

$$\mathcal{P}^{\text{ADD}} := \{ P | P = \hat{P} + \Delta_u, \Delta_u \in \Delta_u \},$$

where all considered systems have appropriate dimensions.

To ensure boundedness of (1.11), the dual-Youla-Kučera uncertainty structure has been considered in, e.g., [20], [2], [7], [10], [22]. Specifically,

$$\mathcal{P}^{\text{DY}} := \{ P | P = (\hat{N} + D_c \Delta_u) (\hat{D} - N_c \Delta_u)^{-1}, \Delta_u \in \Delta_u \},$$

where the pairs $\{ \hat{N}, \hat{D} \}$ and $\{ N_c, D_c \}$ are any RCF of $\hat{P}$ and $C^{\text{exp}}$, respectively. The model set $\mathcal{P}^{\text{DY}}$ leads to

$$\hat{M}^{\text{DY}} = \begin{bmatrix} D_c + PN_c & \hat{N} \\ \hat{D} & D_c + PN_c \end{bmatrix}^{-1}$$

and

$$\hat{M}^{\text{DY}}(\hat{P}, C^{\text{exp}}) = \begin{bmatrix} 0 & D_c + PN_c \hat{N} \\ \hat{D} & D_c + PN_c \hat{N} \end{bmatrix}^{-1} (C^{\text{exp}}, I)_V \begin{bmatrix} \hat{D} \\ \hat{N} \end{bmatrix}$$

Interestingly, (1.16) leads to

$$\mathcal{J}_{\text{WC}}(\mathcal{P}^{\text{DY}}, C^{\text{exp}}) = \sup_{\Delta_u \in \Delta_u} \| \hat{M}_{22} + \hat{M}_{21} \Delta_u \hat{M}_{12} \|_{\infty}$$

which is an affine function of $\Delta_u$ and hence bounded for all $\Delta_u \in \Delta_u$. Essentially, the dual-Youla-Kučera model uncertainty structure is especially useful from a robust stability perspective, since it excludes candidate models that are not stabilized by $C^{\text{exp}}$. However, it is emphasized that $\hat{M}_{12}$ and $\hat{M}_{21}$ in (1.17) are frequency-dependent and multivariable transfer function matrices. Consequently, the bound in (1.17) is
finite but in general arbitrary. This typically necessitates the use of highly structured perturbation models, e.g., as in [8].

Recently, the result (1.17) has been significantly strengthened in [26], leading to

$$
\mathcal{J}_{WC}(\mathcal{P}^{DY}, \mathcal{C}^{exp}) \leq \mathcal{J}(\hat{P}, \mathcal{C}^{exp}) + \sup_{\Delta_u \in \Delta_u} \|\Delta_u\|_{\infty}.
$$

The result (1.18) has been established through the introduction of a new coprime factorization, which will be described in detail in the next section. The key advantage is that the result (1.18) holds true for unstructured perturbation models and is independent of the input-output dimensions. As a result, the uncertainty structure remains $\mu$-simple [29] for complex systems with high input-output dimensionality. Consequently, this enables an efficient and non-conservative robust controller synthesis, see, e.g., [37], [11].

### 1.2.3 Identifying robust-control-relevant coprime factorizations

The key step in establishing the result (1.18) is a certain robust-control-relevant coprime factorization. In [26], it is shown that

- in the identification of the model set in (1.10), the nominal model $\hat{P}$ can be identified in a separate step, followed by quantifying model uncertainty, i.e., determining $\gamma$ in (1.9).
- the identification of the nominal model in terms of the relevant coprime factorization has a direct connection to classical control-relevant identification techniques, see [35].

To show the latter aspect, recall that typical control-relevant identification approaches involve the criterion

$$
\hat{P} = \arg\min_P \| W (T(P_o, C^{exp}) - T(P, C^{exp})) V \|_{\infty}.
$$

(1.19)

To proceed, let $\{\tilde{N}_e, \tilde{D}_e\}$ be a left coprime factorization (LCF), see [40] for a definition, with co-inner numerator of $C^{exp}V_2 V_1$, i.e., $\{\tilde{N}_e, \tilde{D}_e\}$ is an LCF and satisfies the additional condition that $\tilde{N}_e\tilde{N}_e^* = I$. Given $C^{exp}$, $V_2$, and $V_1$, such a coprime factorization can directly be computed, see [44] for details. Next, algebraic manipulations reveal that (1.19) is equivalent to

$$
\min_{\tilde{N}, \tilde{D}} \| W \left( \begin{bmatrix} N_o \\ D_o \end{bmatrix} - \begin{bmatrix} \tilde{N} \\ \tilde{D} \end{bmatrix} \right) \tilde{N}_e \|_{\infty}
$$

subject to $\tilde{N}, \tilde{D} \in \mathcal{RH}_{\infty}$,

(1.20)

where

$$
\begin{bmatrix} N \\ D \end{bmatrix} = \begin{bmatrix} P \\ I \end{bmatrix} (\tilde{D}_e + \tilde{N}_e V_2^{-1} P)^{-1}
$$

(1.21)
and $\hat{N}_e = [\hat{N}_{e,2} \hat{N}_{e,1}]$. In addition, the pairs $\{N_o, D_o\}$ and $\{\hat{N}, \hat{D}\}$ are coprime factorizations of $P_o$ and $\hat{P}$, respectively, as is proved in [26, Theorem 2]. It is emphasized that the pairs $\{N_o, D_o\}$ and $\{\hat{N}, \hat{D}\}$ constitute a new robust-control-relevant factorization, and are not equivalent to normalized coprime factors, e.g., as used in [21] and [40].

The important aspect in (1.20) is that $\hat{N}_e$ is co-inner and does not influence the $\mathcal{H}_\infty$-norm. Consequently, it can be removed directly, see (1.22), below. As a result, the four-block control-relevant identification problem (1.19) is recast as a two-block coprime factor identification problem.

Solving the identification problem (1.20) is not immediate and several steps are required to arrive at a suitable identification algorithm. First, notice that (1.20) involves an $\mathcal{H}_\infty$ norm. By employing the frequency domain interpretation of the $\mathcal{H}_\infty$-norm, (1.20) is recast as

$$\min_{\hat{N}, \hat{D}} \max_{\omega_i \in \Omega_{id}} \bar{\sigma} \left( W \left( \begin{bmatrix} N_o \\ D_o \end{bmatrix} - \begin{bmatrix} \hat{N} \\ \hat{D} \end{bmatrix} \right) \right) \quad (1.22)$$

subject to $\hat{N}, \hat{D} \in \mathcal{R} \mathcal{H}_\infty$. (1.23)

Second, $\{N_o, D_o\}$ is unknown. The key idea is that $T(P_o, C_{\text{exp}})$ can be directly identified using frequency response estimation using the results in [31], see also [28, Appendix A], leading to an estimate $\hat{T}(P_o, C_{\text{exp}})$ for $\omega_i \in \Omega_{id}$. A nonparametric estimate of $\{N_o, D_o\}$ is subsequently obtained by

$$\begin{bmatrix} \tilde{N}_o \\ \tilde{D}_o \end{bmatrix} = \hat{T}(P_o, C_{\text{exp}}) V \tilde{N}_e \text{ for } \omega_i \in \Omega_{id}. \quad (1.24)$$

The key step in the actual identification of a nominal model $\hat{P}$ in terms of the required coprime factorization $\{\hat{N}, \hat{D}\}$ requires solving the optimization problem (1.22). This is investigated next.

### 1.3 Generalized SK-iterations for closed-loop coprime factor identification

In this section, a key contribution of this chapter is presented, which is a closed-loop generalization of Sanathanan-Koerner (SK) iterations [34]. First, in Sec. 1.3.1, the model parameterization is introduced. Next, in Sec. 1.3.2, an auxiliary step is taken to address the $\ell_\infty$ norm in (1.22). Then, in Sec. 1.3.3, the generalized SK-iterations for closed-loop coprime factor identification are presented.
1.3 Generalized SK-iterations for closed-loop coprime factor identification

1.3.1 Model parameterization

The model needs to be parameterized in terms of coprime factorizations. Before introducing these, the open-loop model is parameterized as

\[ \hat{P}(\theta) = B(\theta)A(\theta)^{-1}. \]  

(1.25)

Here, \( B \in \mathbb{R}^{n_y \times n_u}[z] \) and \( A \in \mathbb{R}^{n_u \times n_u}[z] \), i.e., polynomial matrices in the complex indeterminate \( z \), see [31] for more details. Hence, \( \hat{P}(\theta) \) in (1.25) is parameterized as a right matrix fraction description (MFD). An analogous parameterization can be obtained for a left MFD. By using a so-called full polynomial form, see, e.g., [38, Chapter 6], the common dynamics between different input-output channels are taken into account. This leads to models with a low McMillan degree. Due to the one-to-one correspondence between MFDs and state-space models, this directly leads to state-space models with a small state dimension. This is in sharp contrast to the use of common-denominator models, e.g., as in [4], that leads to high order state-space models.

In case the MFDs are parameterized linearly in the parameters, e.g., if the full polynomial form is employed, then these can be written as

\[ \text{vec}\left( \begin{bmatrix} B(\theta) \\ A(\theta) \end{bmatrix} \right) = \sum_q \psi_q \theta_q, \] 

(1.26)

is a polynomial vector, which is parameterized such that it is linear in \( \theta = [\theta_1 \ \theta_2 \ \ldots]^T \), and \( \psi_q \in \mathbb{R}^{(n_y + n_u)n_u \times (n_y + n_u)n_u}[z] \).

It remains to determine the parametric model \( \{\hat{N}, \hat{D}\} \) in (1.20). The tailor-made parameterization

\[ \begin{bmatrix} \hat{N}(\theta) \\ \hat{D}(\theta) \end{bmatrix} = \begin{bmatrix} B(\theta) \\ A(\theta) \end{bmatrix} (\hat{D}_e A(\theta) + \hat{N}_e z V_2^{-1} B(\theta))^{-1}. \] 

(1.27)

is proposed. This parameterization exploits knowledge of \( C^\text{exp} \) and effectively connects stability of the factors \( \{\hat{N}, \hat{D}\} \) and closed-loop stability of the model, see [26, Theorem 4] for a proof. In addition, the dynamics that are introduced by the experimental controller \( C^\text{exp} \) and weighting filters \( V_2 \) and \( V_1 \) in (1.27) cancel out exactly when constructing \( \hat{P} \), since

\[ \hat{P}(\theta) = \hat{N}(\theta)\hat{D}(\theta)^{-1} = B(\theta)A(\theta)^{-1}, \] 

(1.28)

which indeed corresponds to (1.25).
1.3.2 Frequency Domain Identification Involving $\ell_\infty$-Norms via Lawson’s algorithm

The optimization in (1.22) involves an $\ell_\infty$-type criterion, resulting in an optimization problem that is not guaranteed to be smooth, hence efficient gradient-based optimization techniques cannot be used directly. In addition, the parametrization (1.27) is nonlinear in the parameters $\theta$, resulting in a generally non-convex optimization problem.

To solve the optimization in (1.22), Lawson’s algorithm, see [33] and [5], is employed as follows.

Algorithm 1

Set $\theta^{<o>} = 0$ and $w_i^{<0>} = \frac{1}{n_\omega}$, $n_\omega$ denoting the number of frequencies in $\Omega$. Iterate over $k$ until convergence:

$$\theta^{<k>} = \arg\min_{\theta} \sum_i w_i^{<k>} \|\varepsilon_i(\theta)\|_F^2$$  \hspace{1cm} (1.29)

where $w_i^{<k>} = \frac{\bar{\sigma}(\varepsilon_i(\theta^{<k>})) w_i^{<k-1>}}{\sum_i(\bar{\sigma}(\varepsilon_i(\theta^{<k>})) w_i^{<k-1>})}$, \hspace{1cm} (1.30)

Algorithm 1 iteratively solves the nonlinear least squares problem (1.29), which is defined in detail in Sec. 1.3.3. The iteratively adjusted weighting function $w_i^{<k>}$ is employed to minimize (1.22).

1.3.3 A closed-loop generalization of SK iterations

In this section, a new algorithm is presented to solve the actual computational step (1.29) in Algorithm 1. Observe that the nonlinear least squares problem in (1.29) is equivalent to

$$\sum_i \|W^{<k>} \circ (W(\begin{bmatrix} N_0 \\ D_0 \end{bmatrix} - \begin{bmatrix} \hat{N}(\theta) \\ \hat{D}(\theta) \end{bmatrix})\|_F^2,$$  \hspace{1cm} (1.31)

where the elements of $W^{<k>}_h$ are equal to $\sqrt{w_i^{<k>}}$ and $\circ$ denotes the Hadamard product, which is introduced to separate the weighting of Lawson’s algorithm and the nonlinear least squares problem. Note that the use of the Hadamard product is by no means necessary, but increases flexibility of the algorithm for incorporating user-chosen weighting functions. Rearranging and using the following facts from Kronecker algebra for matrices of suitable dimensions, see, e.g., [6]:

- $\|A\|_F = \|\text{vec}(A)\|_2$
- $\text{vec}(ABC) = (C^T \otimes A)\text{vec}(B)$
- $\text{vec}(A \circ B) = \text{diag}(\text{vec}(A))\text{vec}(B)$,

reveals that (1.31) can be written as
1.3 Generalized SK-iterations for closed-loop coprime factor identification

\[ \sum_i \| W_{lsq,i}^{<k>} (\theta) \text{vec} \left( \begin{bmatrix} B(\theta) \\ A(\theta) \end{bmatrix} \right) \|_2^2, \]  \hspace{1cm} (1.32)

where

\[ W_{lsq,i}^{<k>} (\theta) = \text{diag}(\text{vec}(W_h^{<k>})) \left( \left[ \hat{D}_e V_1^{-1} A(\theta) + \tilde{N}_e V_2^{-1} B(\theta) \right]^{-1} \right)^T \]
\[ \otimes \left[ W \left( \begin{bmatrix} N_o \tilde{N}_e V_2^{-1} & N_o \hat{D}_e V_1^{-1} \\ D_o \tilde{N}_e V_2^{-1} & D_o \hat{D}_e V_1^{-1} \end{bmatrix} - I \right) \right], \]  \hspace{1cm} (1.33)

The key idea in solving the nonlinear least squares problem (1.32) by introducing another iteration index \( <f> \), leading to

\[ \sum_i \| W_{lsq,i}^{<k>} (\theta^{<f-1>}) \text{vec} \left( \begin{bmatrix} B(\theta^{<f>}) \\ A(\theta^{<f>}) \end{bmatrix} \right) \|_2^2. \]  \hspace{1cm} (1.34)

By using the parameterization in (1.26), then (1.32) can be solved by iteratively solving the linear least squares problem

\[ \min_{\theta^{<f>}} \| W_{lsq}^{<k>} (\theta^{<f-1>}) \Psi \theta^{<f>} \|_2^2 \]  \hspace{1cm} (1.35)

subject to an appropriate degree constraint, i.e., ensuring that the resulting denominator satisfies a monotonicity constraint. In (1.35),

\[ W_{lsq}^{<k>} = \text{diag} \left( W_{lsq,1}^{<k>}, W_{lsq,2}^{<k>}, \ldots \right) \]  \hspace{1cm} (1.36)
\[ \Psi = \begin{bmatrix} \psi(\omega_1)^T \\ \psi(\omega_2)^T \\ \vdots \end{bmatrix}^T \]  \hspace{1cm} (1.37)

The iteration (1.35) constitutes a novel generalization of well-known Sanathanan-Koerner iterations, e.g., [42], to the closed-loop case. The fixed point of these iterations is generally not equal to the local minimum of (1.32) but generally very close and achieved after several iterations, as is extensive elaborated on in [42] and confirmed by the experimental results presented below. Therefore, the fixed point of the iteration (1.35) is further refined using a Gauss-Newton optimization step.

Due to non-convexity of the problem, it is not guaranteed that the proposed iterative algorithm converges to the global minimum of (1.22). However, extensive application of the algorithm for various data has delivered good results. In addition, the convergence is generally fast, since for subsequent iterations good initial estimates are available. Similar observations have been obtained in [5], where related algorithms have been used for open-loop system identification.
1.4 Orthogonal polynomials w.r.t. a data-dependent discrete inner product

In the previous section, a new algorithm is presented that enables the identification of a certain coprime factorization from closed-loop data. Although the algorithm in Sec. 1.3 can be implemented directly, the accuracy of the estimated model and convergence of the algorithm hinges on a numerically reliable implementation. A key contribution of the present section and Sec. 1.5 is to experimentally confirm that standard choices in system identification lead to an extremely ill-conditioned identification problem for the considered class of systems. In addition, the underlying mechanism is analyzed and a solution is provided that is shown to lead to a numerically reliable solution for experimental data.

The essential computational step in the identification problem involves the solution of the linear least-squares problem (1.35) and its subsequent Gauss-Newton refinement. The numerical properties are consequently related to the condition number \( \kappa(.) = \frac{\sigma(.)}{\sigma(\mathbf{A})} \) of the matrix \( \mathbf{W}_{lsq}^{(k-1)} \mathbf{\Psi} \).

A common choice regarding the basis \( \mathbf{\Psi} \) is a monomial basis. In this case, the matrix \( \mathbf{\Psi} \) is a Vandermonde matrix. Even though this leads to a unitary \( \mathbf{\Psi} \) for discrete time identification problems with equidistantly spaced frequencies, and hence \( \kappa(\mathbf{\Psi}) = 1 \), the matrix \( \mathbf{W}_{lsq}^{(k-1)} \mathbf{\Psi} \) generally is ill-conditioned. Consequently, the resulting problem (1.35) is poorly conditioned and cannot be solved accurately. This ill-conditioning is also evidenced by

1. the use of several procedures to enhance the numerical conditioning, e.g., [30], [14], [43]; and
2. by the experimental results in the next section.

The essential observation is that the conditioning in (1.35) hinges on the choice of the polynomial basis in (1.26). The key idea is to select a basis that is orthonormal with respect to the data-based discrete inner product

\[
\langle \psi_m, \psi_l \rangle_{\mathbf{W}_{lsq}} = \sum_i \psi_m(\omega_i)^* \mathbf{W}_{lsq,i}^{(k-1)} W_{lsq,i}^{(k-1)} \psi_l(\omega_i),
\]

(1.38)

for a certain \( m \) and \( l \). If an orthonormal basis with respect to the inner product (1.38) is selected, then \( \kappa(W_{lsq}^{(k-1)} \mathbf{\Psi}) = 1 \). Hence, this would render the computational step (1.35) optimally conditioned. In addition, subsequent Gauss-Newton iterations are then also typically well-conditioned as the estimate \( \theta^{SK} \) is often close to optimality.

The orthonormal basis (1.38) can be computed in a numerically stable and efficient manner. This numerical procedure is beyond the scope of the present chapter and is outlined in [16], which extends the results in [3]. Importantly, the actual implementation can be done in \( \mathcal{O}(N) \), i.e., linear in the data \( N \). The focus of this chapter is on the experimental comparison of these results with respect to the commonly used monomial basis.
1.5 Experimental application

1.5.1 Experimental system

The considered nanometer-accurate positioning system is an industrial wafer stage system is depicted in Fig. 1.3. Wafer stages are part of wafer scanners, see Fig. 1.2 and [41], which are the state-of-the-art machines for the automated production of ICs. During the production process, a photoresist is exposed on a silicon disc, called a wafer. During exposure, the image of the desired IC patterns, which is contained on the reticle, is projected through a lens on the photoresist. The exposed photore sist is then removed by means of a solvent. Subsequent chemical reactions enable etching of these patterns, which is repeated for successive layers. Typically, more than 20 layers are required for each wafer. Each wafer contains more than 200 ICs that are sequentially exposed. During this entire process, the wafer must extremely accurately track a predefined reference trajectory in six motion degrees-of-freedom (DOFs). This precision motion task is performed by the wafer stage, which is the motion system considered in this chapter.

The system is equipped with moving-coil permanent magnet planar motors that enable contactless operation, see [9] for the underlying principle. The motion system consists of two parts: a stator, which is a plate consisting of an ordered array of permanent magnets, and a mover, which constitutes the moving part of the wafer stage.

![Fig. 1.2 Schematic illustration of a wafer scanner. Light is projected through a sophisticated lens or mirror system. The image of the IC is contained on the reticle stage (top stage), and then projected onto a silicon wafer. This wafer is accurately positioned in six degrees-of-freedom by the wafer stage (bottom stage).](image-url)
Four actuators are connected to the mover to provide the necessary force, each consisting of three coils, which are powered by a three-phase power source. By means of an appropriate position-dependent commutation of the coils, each actuator delivers a well-defined and reproducible force in two directions. As a result, eight independent forces are available.

Laser interferometers in conjunction with a mirror block, which are connected to the metrology frame and the wafer stage, respectively, enable a high accuracy position measurement in all six motion DOFs, i.e., three translations and three rotations. Specifically, subnanometer measurement accuracy is available for the translational DOFs. Throughout, all signals and systems operate in discrete time with a sampling frequency of 2.5 kHz.

It is emphasized that the presented approach in this chapter is aimed to deal with a large number of actuators and sensors, generally more than the number of motion DOFs. However, to facilitate a clear exposition, a two input two output subsystem is considered in this chapter. The other DOFs are controlled by low performance PID controllers. The translational $x$ and $y$ DOFs in the horizontal plane are considered in this chapter. Hence, the input $u$ and output $y$ to the system are given by

$$u = \begin{bmatrix} u^x \\ u^y \end{bmatrix}, \quad y = \begin{bmatrix} y^x \\ y^y \end{bmatrix},$$

(1.39)

whereas the system is partitioned as

[Fig. 1.3 Experimental wafer stage system, where 1: metrology frame, 2: mover, 3: airmount.]
In the experimental setup, an initial controller $C_{\text{exp}}$ is implemented for the considered DOFs that achieves a closed-loop bandwidth of 40 Hz. The control goal is to achieve a closed-loop bandwidth of approximately 90 Hz. This is reflected by a suitable choice regarding the weighting filters $W$ and $V$ in (1.1) using the loop-shaping design rules in [21], [39].

1.5.2 Coprime factor identification results

The identification algorithm in Sec. 1.3 is applied to the experimental wafer stage data. First, closed-loop frequency response function $\tilde{T}(P_0, C_{\text{exp}})$ is identified using the approach in [31], see also [28] for details.

Next, (1.24) is used to compute the coprime factor frequency response function $\{\tilde{N}_o, \tilde{D}_o\}$ in (1.24). The resulting frequency response function $\{\tilde{N}_o, \tilde{D}_o\}$ is depicted in Fig. 1.6.

Subsequently, Algorithm 1 is invoked to solve the identification problem (1.22). The model order is selected using the results in [15]. The McMillan degree of the optimal model equals 8. Specifically, the control-relevant identification criterion (1.22)
Identification for robust control of complex systems

decreases significantly for orders up to 8, whereas increasing the model order beyond 8 does not significantly improve the fit in terms of the criterion (1.22).

The resulting identified coprime factors \( \{ \hat{N}, \hat{D} \} \) are depicted in Fig. 1.6. In addition, the open-loop frequency response function \( P_o \) and model \( \hat{P} = \hat{N}\hat{D}^{-1} = BA^{-1} \) are compared in Fig. 1.5, which facilitates the interpretation of the model in terms of physical system properties. Note that due to the specific parameterization (1.25) in terms of a matrix fraction description, a minimal state space model of \( \hat{P} \) has state dimension 8. From Fig. 1.5, the following observations are made. Below 200 Hz, the system is decoupled, revealing a rigid-body behavior in the diagonal elements, corresponding to translations in the \( x \)-direction and \( y \)-direction, respectively. The first resonance phenomena appear at 208 Hz and 214 Hz in all elements of \( \hat{P}_o(\omega) \). Since these flexible dynamics are not aligned with the motion DOFs, the interaction between the \( x \)-direction and \( y \)-direction is high, i.e., the four elements of \( \hat{P}_o(\omega) \) have an approximately equal gain beyond 200 Hz.

Analysis of the identified model in Fig. 1.5 that has minimal state dimension 8 reveals that it is of low order for two reasons. First, the model only represents a limited number of resonance phenomena of \( \hat{P}_o \) that are observed in Fig. 1.5. This is a direct consequence of the control-relevant identification criterion in (1.19). The coprime factor domain in Fig. 1.6 directly connects to control-relevance in terms of (1.19). From visual inspection it is confirmed that the model is accurate in regions where the amplitude of the corresponding coprime factors is large, which are control-relevant in view of (1.19).

Second, the model exploits the inherently multivariable parameterization in (1.25) that has a one-to-one correspondence with state space realizations. Indeed, observe that four states in the model correspond to the two rigid-body modes in both the \( x \)-direction and the \( y \)-direction. The other four states correspond to resonance phenomena. Since these resonance phenomena correspond to flexible dynamical behavior, these correspond to complex pole pairs, see, e.g., [12]. Specifically, around 208 Hz and 214 Hz two closely spaced resonance phenomena are present. Interestingly, these resonance phenomena correspond to inherently multivariable behavior, since both these resonances appear in all four transfer functions in Fig. 1.5, yet only require two states each. This is significant improvement over the models as delivered by the procedure in, e.g., [41], that does not take into account common dynamics between the different DOFs.

Finally, the optimization algorithm in Sec. 1.3 has an additional advantage from an efficiency perspective. In particular, by recasting the control-relevant identification problem (1.19) as the equivalent coprime factor identification problem (1.20), redundant data is effectively removed, leading to 50% reduction of the complexity for the considered two-input two-output wafer stage application. This can directly be understood when comparing Fig. 1.4 and Fig. 1.6 that involve \( 4 \times 4 \) and \( 4 \times 2 \) transfer function matrices, respectively. Hence, the coprime factors in Fig. 1.6 involve only half the amount of data when compared to Fig. 1.4.
1.5.3 Numerical conditioning

The numerically optimal procedure in Sec. 1.5.3 is essential to obtain the accurate models in Sec. 1.5.2. To show this, it has first been attempted to solve the least squares problem (1.35) using monomial basis functions. The resulting condition numbers corresponding to (1.35) and subsequent Gauss-Newton iterations are depicted in Fig. 1.7 and are in between $10^{20}$ and $10^{30}$, leading to a breakdown of the algorithm even if implemented using a $QR$-factorization. Extensive computations have revealed that the commonly used monomial basis leads to an ill-conditioned least-squares problem (1.35) and subsequent Gauss-Newton iteration for the considered class of motion systems.

To improve the conditioning associated with (1.35), basis functions that are orthonormal with respect to the data-based discrete inner product (1.38) are employed. As a result of the developed algorithm, the condition number corresponding to the least-squares problem equals the optimal value of one. In addition, the condition number during the Gauss-Newton iterations, where the basis is fixed, remains close to optimal as is observed in Fig. 1.7. This leads to a converging algorithm and an accurate computation of the optimal model.

Summarizing, these results underline the importance of numerical aspects in system identification of complex systems such as next-generation motion systems.

1.5.4 Illustration of robust-control-relevance

In Sec. 1.2, it is claimed that the identified coprime factorization in Fig. 1.6 is essential for modeling uncertainty in view of robust control, i.e., for obtaining the result (1.18). To show that this from a control design perspective, the model is extended with uncertainty. In particular, the unstructured perturbation model (1.9) is considered. Herein, $\gamma$ is determined using the model validation procedure presented in [24]. The resulting model set $\mathcal{P}$ as defined in (1.8) is depicted in Fig. 1.5. In addition, the model set corresponding to $\frac{\gamma}{2}$ is depicted, i.e., only 50 % uncertainty. It is observed in Fig. 1.5, that the uncertainty is automatically shaped. First, it has a frequency dependent shape, since the uncertainty is the smallest around the target bandwidth of 90 Hz and the resonance phenomena around 200 Hz, which are known to be essential for closed-loop stability. On the other hand, at low frequencies the uncertainty is very high, since the controller will have integral action. In addition, at high frequencies, the control-relevant aspect addresses the fact that the controller has roll-off, hence the high frequent resonance phenomena are not relevant for stability and performance.

Second, the specific coprime factorizations inherently scale the input-output channels of the model uncertainty. Indeed, observe that (1.9) is a multivariable and unstructured perturbation block. Still, from visual inspection the model uncertainty is scaled well with respect to the two inputs and outputs. This aspect is further explained in [25].
Fig. 1.5 Bode magnitude diagram: nonparametric estimate (dot), nominal model $\hat{P}$ (solid blue), model set $\mathcal{P}$ using $\gamma$ (cyan), and $\mathcal{P}$ for only 50% uncertainty, i.e., using $\hat{\gamma}$ (yellow).
Fig. 1.6 Coprime factorization: nonparametric $N_o$, $D_o$ (blue dots), eighth order parametric model coprime factorizations $\hat{N}$, $\hat{D}$ (dashed red).
Fig. 1.7 Condition number corresponding to linear least-squares and Gauss-Newton iterations. Top: using basis functions that are orthonormal with respect to a data-based discrete inner product (solid blue). Bottom: using standard monomial basis functions (dashed red). In the latter case, the iterations break down after the first Gauss-Newton iteration.
Finally, a subsequent robust control design in [28] confirms that the model set $\mathcal{P}$ is adequate for generating a robust controller that significantly enhances the control performance.

1.6 Conclusions

In this chapter, a new algorithm is presented for identifying complex system. The proposed algorithm has three important advantages. First, a numerical optimal implementation is presented. Second, the new implementation enables an efficient implementation, as is shown in, e.g., [16]. Third, the algorithm enables the identification of a certain coprime factorization. This enables the use of unstructured uncertainty, which is essential for synthesizing robust controllers for complex systems.

The algorithm is demonstrated on an industrial wafer stage system. Experimental results confirm that common algorithm implementations based on monomial basis functions are generally ill-conditioned, prohibiting the identification of accurate models. In addition, it is shown that the presented approach is numerically reliable and provides good identification results.

Present research focuses on extending the algorithm in this section towards rendering Gauss-Newton iterations superfluous, see, e.g., [16] for early results in this direction. In addition, the coprime factor framework is being extended to deal with non-measurable performance outputs in [27]. Interestingly, this also leads to an enhanced complexity of the identification problem, as it leads to an additional number of inputs and outputs that have to be addressed.

Acknowledgements

The authors are grateful for the contributions of Okko Bosgra that have led to the results in this chapter. The authors want to thank Robbert van Herpen for his important contributions and suggestions. In addition, Marc van de Wal and Sander Quist are gratefully acknowledged for their help in performing the experiments and sharing their expertise on wafer stage motion control. Finally, the authors would like to thank Philips Innovation Services, Eindhoven, for providing access to the experimental facilities.

This work is supported by the Innovational Research Incentives Scheme under the VENI grant "Precision Motion: Beyond the Nanometer" (no. 13073) awarded by NWO (The Netherlands Organisation for Scientific Research) and STW (Dutch Science Foundation).