Identification for Motion Control: Incorporating Constraints and Numerical Considerations

Robbert Voorhoeve∗, Robin de Rozario∗, and Tom Oomen∗

Abstract—Frequency domain identification is a common starting point for model based motion control. The aim of this paper is to tailor parametric identification methods to the specific class of motion systems. The proposed method involves two aspects: 1) incorporating prior knowledge, e.g. it is often known beforehand that the system exhibits rigid-body behavior, and 2) numerical reliability, since the considered class of systems is challenging to identify. The result is a frequency domain algorithm that is particularly suited for the identification of lightly damped motion systems with rigid body behavior, a high model order, and large input-output dimensions. Experimental results on a prototype next-generation motion system clearly demonstrate the advantages of the proposed approach.

I. INTRODUCTION

Increasing accuracy and speed requirements in motion systems lead to increasing requirements on identification techniques. On the one hand, increasing performance requirements necessitate a higher control bandwidth. On the other hand, increasing accelerations are expected to meet the speed requirements, for which it is desirable to have lightweight construction. An immediate consequence is that the resonance phenomena occur at lower frequencies. Combining these two developments implies that resonance phenomena appear in the cross-over region [1]. It is envisaged that active control of these flexible dynamics is required, including the use of additional actuators and sensors, and inferential control [2]. This leads to increasing requirements on the identification approach, since high order, lightly damped flexible dynamics have to be identified with a large input and output dimensionality.

Identification of lightly damped flexible dynamics is challenging from an identification perspective, as is evidenced by the recent dedicated benchmarks [3], [4]. Typically, such systems are identified using a frequency domain approach [5]. In particular, a parametric model is fitted using a nonlinear least-squares type of criterion. The key challenge arises in the numerical aspects of optimizing such a criterion. Many partial solutions have been developed that mitigate the ill-conditioning, including i) frequency scaling [6]; ii) amplitude scaling [7]; and iii) the use of orthonormal polynomials with respect to a continuous inner product [8],[9].

Interestingly, promising approaches have been developed that aim at optimal conditioning of frequency domain identification algorithms through the use of polynomials that are orthonormal with respect to a data-dependent discrete inner product, see [10],[11]. Essentially, these methods do not formulate the normal equations as is typically done using, e.g., monomial basis functions, but tailor the polynomial basis with respect to the problem data at hand. When interpreted as normal equation, these methods essentially achieve $\kappa = 1$. Recently, these data-dependent polynomials have been further extended towards bi-orthonormal ones [12], [13], which enables solving instrumental variable identification problems [14] with $\kappa = 1$.

Besides the development of identification algorithms and their numerical implementation, an important aspect in motion systems is that prior knowledge often is available. For instance, the phase of transfer function elements of collocated actuator sensor pairs is constrained in [15]. Other relevant prior knowledge includes modal model structures [16], [17], [18] and rigid-body degrees of freedom. Arguments for including such prior information include i) it facilitates rigid-body decoupling, is of central importance in overactuated and oversensed systems [19]; ii) it facilitates the design of feedforward controllers [20], especially in the inferential situation [21]; iii) it facilitates the reduction of variance [5]; and iv) it may increase confidence in the identified models for motion control engineers.

Although important developments have been made in identification for model-based motion control, at present the inclusion of prior information in frequency domain system identification is not yet clarified. In particular, the use of orthonormal bases with respect to a data-dependent inner product seems to be numerically advantageous for the identification of motion systems. However, the resulting model may be less clear to interpret and it is not immediate how certain prior information can be included, e.g., through direct embedding in the model structure. This paper aims to constrain the model, parametrized in data-dependent polynomial basis functions, to include rigid-body modes.

The main contribution of the present paper is an approach for numerically optimal frequency domain identification (i.e., $\kappa = 1$) to include constraints on the model parametrization, including rigid-body dynamics. The proposed approach applies to the so-called SK algorithm [11], as well as more recent IV techniques, see [14], [12]. The results are demonstrated on a prototype next-generation motion system.

The proposed approach has important applications in the identification of mechatronic systems. Indeed, poor identification results are often obtained in the low-frequency region due to a low signal to noise ratio (SNR) in that region as a
result of the closed-loop identification procedure [2]. Hence, utilizing more system knowledge by adding the proposed constraints is useful, e.g., for decoupling, and for feedback and feedforward control design [20].

The outline of the paper is as follows. In section II, the considered identification problem is formulated as well as the problem of incorporating rigid body constraints. In section III, the proposed method of incorporating rigid body constraints is presented. In section IV, the framework for numerically optimal frequency domain identification using data-dependent bases is established. In section V, the proposed identification method is applied to a prototype next-generation motion system, demonstrating the advantages of the proposed approach. And in section VI the conclusions are presented as well as an outlook on future research.

II. PROBLEM FORMULATION

A. Identification problem

In this paper, the parametric identification of linear time invariant models with \( p \) outputs and \( q \) inputs in the frequency domain is considered. The class of models \( \hat{P} \) that is considered is the class of real rational transfer function matrices, i.e., \( \hat{P}(\xi, \theta) \in \mathbb{R}^{p \times q}(\xi) \), where \( \theta \) is the vector of parameters used to parametrize the model and \( \xi \) is an indeterminate variable which depends on the identification domain (e.g. for continuous time \( \xi = s = j\omega \).

As starting point for the parametric identification, a non-parametric estimate of the frequency response matrix (FRM) of the system, \( P_k(\xi_k) \), is used, where \( k \) is used to denote the discrete frequency bins at which the FRM is defined. The identification criterion that is considered is a weighted squared Frobenius norm of the error between the non-parametric FRM and the parametric model, i.e.,

\[
V(\theta) = \sum_{k=1}^{m} \epsilon(\xi_k, \theta)^H \epsilon(\xi_k, \theta),
\]

where

\[
\epsilon(\xi_k, \theta) = W(k) \text{vec}(P_k(\xi_k) - \hat{P}(\xi_k, \theta)).
\]

with weighting matrix \( W(k) \in \mathbb{C}^{m \times p \times q} \). This criterion \( V(\theta) \) includes other regularly used identification criteria such as the sample maximum likelihood criterion [5],

\[
V_{\text{SM}} = \sum_{k=1}^{m} \text{vec}(P_k - \hat{P})^H C^{-1}_{\text{vec}(P_k)} \text{vec}(P_k - \hat{P}),
\]

or the criterion [22],

\[
V_{\text{CRH}} = \sum_{k=1}^{m} \| W_{\text{Schur}} \odot W_{\text{out}}(P_k - \hat{P}) W_{\text{in}} \|^2_F.
\]

Using Kronecker algebra, the matrix \( W(k) \) can be derived from these alternative weighting matrices by

\[
W(k) = C^{-1/2}_{\text{vec}(P_k)}(k) \Lambda W_{\text{Schur}}(k) (W_{\text{in}}(k) \odot W_{\text{out}}(k)),
\]

where \( \Lambda W_{\text{Schur}}(k) = \text{diag}(\text{vec}(W_{\text{Schur}})) \).

What remains is to appropriately parametrize the model \( \hat{P}(\xi) \) as a function of \( \theta \) and to solve the identification problem (1). In general \( \hat{P}(\xi, \theta) \) is not a linear function of \( \theta \) so no closed form solution of (1) exists for this problem, therefore it is often solved using iterative algorithms.

\[
\text{Iteratively reformulate}
\]

\[
V_{\text{SM}} \Rightarrow J_{\text{SM}}(\hat{P}(\xi, \theta)) = 0 \Rightarrow \hat{P}(\xi, \theta) = \hat{P}(\xi, \theta)
\]

\[
V_{\text{CRH}} \Rightarrow J_{\text{CRH}}(\hat{P}(\xi, \theta)) = 0 \Rightarrow \hat{P}(\xi, \theta) = \hat{P}(\xi, \theta)
\]

\[
\text{Iteratively reformulate}
\]

Fig. 1. Graphical depiction of the principles of the SK and IV algorithms.

B. Parametrizations & identification algorithms

In this section the considered class of model parametrizations is introduced and several iterative algorithms which can be used to solve the identification problem (1) are described.

In this work the considered model class of real rational transfer functions is parametrized using polynomial matrix fraction descriptions, where to simplify notation only right matrix fraction descriptions (RMFD’s) are considered, i.e.

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

where \( N(\xi, \theta) \in \mathbb{R}^{p \times q}(\xi) \) and \( D(\xi, \theta) \in \mathbb{R}^{q \times q}(\xi) \). Furthermore, these polynomial matrices are linearly parametrized with respect to the parameters \( \theta \) using a set of basis functions. In this paper general vector polynomials are used as basis functions, which have the following generic form:

\[
\text{vec}
\begin{bmatrix}
\hat{D}(\xi, \theta) \\
\hat{N}(\xi, \theta)
\end{bmatrix}
= \sum_{j=0}^{m} \phi_j(\xi) \theta_j = \Phi(\theta),
\]

where \( \phi_j(\xi) = \xi^j t_{jj} + \cdots + \xi t_{j1} + t_{0j} \),

\[
\tau_j \in \mathbb{R}^{(q+q) \times (q+q)}.
\]

This extra flexibility in the considered class of basis functions is necessary to be able to obtain optimal conditioning of the considered identification algorithms, see [11]. This shows why reducing the freedom in the parametrization by imposing constraints could impair the ability to use the methods described in [11], [12].

Using (6), \( \epsilon(\xi_k, \theta) \) can be reformulated as

\[
\epsilon(\xi_k, \theta) = W(k) \text{vec}(P_k(\xi_k) - N(\xi_k, \theta) D(\xi_k, \theta)^{-1}),
\]

\[
= W(k) \text{vec}
\begin{bmatrix}
P_k(\xi_k) - I_p
D(\xi_k, \theta)^{-1}
\end{bmatrix}
\begin{bmatrix}
D(\xi_k, \theta)
N(\xi_k, \theta)
\end{bmatrix},
\]

or

\[
\tau_j(\xi_k, \theta) = \text{vec}
\begin{bmatrix}
D(\xi_k, \theta)^{-1}
N(\xi_k, \theta)
\end{bmatrix}.
\]

Using (7) & (8) it is clear that

\[
{\begin{bmatrix}
J_{\text{SM}}(\hat{P}(\xi, \theta)) = 0 \Rightarrow \hat{P}(\xi, \theta) = \hat{P}(\xi, \theta)
\end{bmatrix}}
\]

Iteratively reformulate Iteratively reformulate

\[
J(\xi) = 0 \Rightarrow C(\xi') A(\hat{b}(\xi)) = C(\xi') A(\xi)
\]

Iteratively reformulate Iteratively reformulate

\[
J(\xi) = 0 \Rightarrow \hat{b}(\xi) = \hat{b}(\xi)
\]

Iteratively reformulate Iteratively reformulate

\[
\text{Iteratively reformulate}
\]

\[
\text{Iteratively reformulate}
\]

\[
\text{Iteratively reformulate}
\]

In this last equality, \( \text{vec}(ABC) = (C^T \otimes A)\text{vec}(B) \). Using (7) & (8) it is clear that \( D(\xi_k, \theta)^{-1} \) is the only term making (8) nonlinear in \( \theta \). In the well known Santhanam Koerner (SK) algorithm, this non-linearity is eliminated by iteratively substituting this term by \( D(\xi_k, \theta)^{-1} \), which is based on the previous iteration.

\[
\text{Algorithm I (SK-iterations):} \ [12] \text{ Let } \theta^{(0)} \text{ be given. In iteration } i = 1, 2, \ldots, \text{ solve the linear least squares problem}
\]

\[
\theta^{(i)} = \arg\min_{\theta} \sum_{k=1}^{m} \| W_{\text{SK}}(k, \theta^{(i-1)}) \text{vec}
\begin{bmatrix}
D(\xi_k, \theta)
N(\xi_k, \theta)
\end{bmatrix}
\|^2_2,
\]

6210
with
\[ W_{SK}(k, \theta^{(i-1)}) = W(k) \left( D(\xi_k, \theta^{(i-1)}) - T \otimes [P_o(\xi_k) - I_p] \right), \]  
(10)

The second algorithm that is considered in this paper takes a similar approach, but instead iteratively reformulates the gradient of the cost function. This difference between the SK and IV algorithms is also shown in Figure 1.

Algorithm 2 (IV-iterations): [14] Let \( \theta^{(0)} \) be given. In iteration \( i = 1, 2, \ldots \), solve the linear system of equations for \( \theta^{(i)} \)
\[
\sum_{k=1}^{m} \left[ -\frac{\partial \text{vec}(\hat{P}(\xi_k, \theta^{(i)}))}{\partial \theta^{(i)}} \right] H W(k)^H = W_{SK}(k, \theta^{(i-1)}) \text{vec} \left( \begin{bmatrix} D(\xi_k, \theta^{(i)})^T \nabla(\xi_k, \theta^{(i)}) \end{bmatrix} \right) = 0.
\]
(11)

The third algorithm that is considered in this work is the well known Gauss-Newton algorithm or a variation of this algorithm known as the Levenberg-Marquardt algorithm. These three considered algorithms all iteratively reformulate the nonlinear optimization problem as a linear problem and in fact they are all closely related. See [23] for a discussion on the relation between these algorithms.

C. Incorporating constraints

The considered problem in this paper is the incorporation of constraints for motion systems, in particular rigid-body modes, for the identification problem (1) with parametrization (7). In particular, the goal is to combine these constraints with a data-dependent basis in the form of (7), achieving optimal numerical conditioning of the identification algorithms.

Two steps are taken. A specific MFD structure is developed that allows for the specification of the rigid-body degrees of freedom. Then, this parametrization is implemented in conjunction with data-dependent basis functions. The resulting model inherently contains the rigid-body modes, and hence an optimization problem can be solved without imposing constraints on the basis functions that could impair the ability to use the data-dependent basis functions.

III. ENFORCING RIGID BODY MODES

In this section, the proposed method of enforcing rigid body modes in the identified MIMO model is presented. As described in in section II-B, the model is expressed as a polynomial matrix fraction description where the numerator and denominator matrix polynomials are linear in the parameters. To enable the use of recent numerically reliable implementations ([11],[12]) of the considered identification algorithms, the choice of the exact parametrization, using a set of basis functions \( \Phi \), should remain free and generic.

In this section, an approach is presented that factorizes the rigid body behavior in such a way that it enables the use of generic parametrizations, while at the same time enforcing the desired amount of rigid body degrees of freedom.

Only the continuous time domain (i.e. \( \xi = s = j \omega \)) is considered in this section, to simplify the presentation. The results of this section can directly be extended to other domains, including the discrete time or \( z \)-domain. Also, for the sake of notation, dependencies on the model parameters \( \theta \) have been omitted in this section.

A. Factorizing rigid body dynamics

First, the concept of rigid body modes as is used in this paper is defined. Rigid body modes arise when a linear mechanical system admits movements which do not store potential energy, resulting in a rank deficient stiffness matrix. This rank deficiency is inherited by the A-matrix of its state-space description, and if the rigid body modes are undamped, this results in 2 \( \times \) 2 Jordan blocks in the A-matrix (in Jordan canonical form) with zeros on the main diagonal. In this paper only undamped rigid body modes are considered.

By only considering undamped rigid body modes, a rigid body mode is defined in this paper as a 2 \( \times \) 2 Jordan block with zeros on the main diagonal in the A matrix of the state space model (when transformed to Jordan canonical form). This corresponds to a Jordan chain of two pure integrators in the model per rigid body mode.

Before proceeding to the general MIMO case, a simplified example is considered. This simplification is twofold, first the single-input single-output case is considered and second this example is parametrized using the standard monomial basis. In this simplified case, a rigid body mode can be enforced by making sure the denominator polynomial of the model,
\[
\hat{P}(s) = \frac{n(s)}{d(s)} = \frac{\sum_{i=0}^{n_0} a_i s^i}{\sum_{i=0}^{n_1} b_i s^i} = \frac{\sum_{i=0}^{n_0} a_i s^i}{\sum_{i=0}^{n_1} b_i s^i} = \frac{n(s)}{d(s)},
\]
(12)
has the parameters \( b_0 \) and \( b_1 \) set equal to zero. Equivalently, the same constraint can be enforced by taking
\[
\hat{P}_{RB}(s) = \frac{n(s)}{s^2} = \frac{\sum_{i=0}^{n_0} a_i s^i}{s^2} = \frac{\sum_{i=0}^{n_0} a_i s^i}{s^2} = \frac{n(s)}{s^2},
\]
(13)
i.e., by factorizing the rigid body part out of the model description. Here it is tacitly assumed that \( n(s) \) and \( s^2 \) are coprime, i.e., no pole-zero cancellations occur between the factorized rigid body part and the remaining model. Using this factorization, it is no longer required that \( n(s) \) and \( d(s) \) are parametrized using the monomial basis, any parametrization can be used, since as long as \( n(s) \) and \( s^2 \) are coprime, the model will always contain the rigid body mode.

This idea of factorizing the rigid body dynamics out of the parametrization can readily be extended to the MIMO case. For a MIMO polynomial matrix fraction description it holds that if the denominator of the MFD can be factorized as:
\[
D(s) = D(s) \mathcal{S}(s) \quad \forall \quad D(s) = \mathcal{S}(s) \hat{D}(s),
\]
(14)
with \( \hat{D}(s) \in \mathbb{R}^{p \times d[\xi]} \) and
\[
\mathcal{S}(s) = \text{diag}(s^2, \ldots, s^2, 1, \ldots, 1),
\]
(15)
then the model will have \( n_{RB} \) rigid body modes as long as \( \mathcal{S}^{-1} \) is coprime with respect to \( N(s) \). This can be seen from the fact that the diagonal elements of \( \mathcal{S}^{-1} \) indeed correspond to Jordan chains of two pure integrators; and if
the coprimeness holds, these Jordan chains cannot be reduced and thus will always remain in the model.

Using this factorization of $D$, two cases can be recognized, the case where the rigid body dynamics are inside the MFD,
\[
\hat{P}(s) = N(s)F(s)^{-1}D(s)^{-1},
\]
and the case where the rigid body dynamics are outside the parametrization, i.e.
\[
\hat{P}(s) = N(s)D(s)^{-1}F(s)^{-1}.
\]
This last expression corresponds to the case where the rigid body modes are only actuated by $n_{RB}$ distinct inputs, and not by the remaining $q - n_{RB}$ inputs, or in other words the rigid body behavior is decoupled from these additional inputs.

In the remainder of this paper, the decoupled case is considered.

**B. Decoupled parametrization**

Using (17) it is possible to rewrite the identification criterion as follows:
\[
\varepsilon(s_k) = W(k) \text{vec}(P_o(s_k) - N(s_k)\hat{D}(s_k)^{-1}F(s_k)^{-1}) = W(k) \text{vec}((P_o(s_k)F(s_k)^{-1} - N(s_k)\hat{D}(s_k)^{-1})F(s_k)^{-1}) = W_F(k) \text{vec}(P_o(s_k)F(s_k)^{-1} - N(s_k)\hat{D}(s_k)^{-1})\]
in this last equality, vec$(ABC) = (C^T \otimes A)\text{vec}(B)$ is used with
\[
W_F(k) = W(k)(F(s_k)^{-1} \otimes I_q).
\]

In the case that there are additional inputs, i.e. $q > n_{RB}$, but the rigid body modes are not decoupled from additional inputs, the factorization (17) can not be used directly. However, a decoupling step can be performed prior to identification. In motion control one can often find rigid-body decoupling matrices $T_u$ and $T_r$ from first principle or based on the non-parametric plant estimate [24]. Using such decoupling matrices one can rewrite the identification problems as follows:
\[
\varepsilon(s_k) = W(k) \text{vec}(P_o(s_k) - N(s_k)\hat{D}(s_k)^{-1}F(s_k)^{-1}) = W(k) \text{vec}(T_r^{-1}P_o(s_k)T_r - N(s_k)\hat{D}(s_k)^{-1}F(s_k)^{-1}) = W_F(k) \text{vec}(P_o(s_k)F(s_k)^{-1} - N(s_k)\hat{D}(s_k)^{-1}).
\]

where
\[
\hat{N}(s_k) = T_rN(s_k), \quad T_rF(s_k)\hat{D}(s_k) = D(s_k), \quad \text{and}
\]
\[
W_F,\text{dec}(k) = W(k)T_r^{-1}F(s_k)^{-1} \otimes T_r^{-1}.
\]

Both in (18) and (20), no constraints are imposed on the parametrization of $\hat{N}(s_k)$ and $\hat{D}(s_k)$. This means the generic parametrization as given by (7) can indeed still be used.

**IV. DATA-DEPENDENT BASIS FUNCTIONS**

In section II-B it is shown that, when using the considered iterative algorithms, the remaining problem comes down to determining a set of basis functions $\Phi$ with which to parametrize the model and then iteratively solving a linearized problem until a stopping condition is reached. Depending on the choice of basis functions $\Phi$, these linear problems can however become severely ill-conditioned. Finding a generic set of basis functions $\Phi$ for which it is guaranteed that the condition numbers of the problem matrices remain low and bounded seems to be impossible, as this is highly dependent on the data of the identification problem as contained in e.g. $W_{SK}(k)$ (see (9)).

This leads to the concept of data-dependent basis functions. These data-dependent bases can be constructed such that optimal conditioning ($\kappa = 1$) of the iterations is achieved. In [25] it is shown that choosing the basis functions for the SK iterations to be block-polynomials, as in (7), that are orthonormal with respect to the data-dependent inner product (22), then optimal conditioning of the SK iterations is achieved.

\[
\langle \phi_i(\xi), \phi_j(\xi) \rangle := \sum_{k=1}^{m} \phi^H_i(\xi_k)w^H_{1k}w_{1k}\phi_j(\xi_k).
\]

This result does not hold for IV iteration however, since the IV iterations involve a non-symmetric oblique projection $C^H\Phi B = C^Hb$, instead of the orthonormal projection $A^H\Phi B = A^Hb$ as in the SK algorithm. To achieve optimal conditioning of the IV iterations, a second set of block-polynomials [12] is constructed. These two sets of block polynomials are constructed to be bi-orthonormal with respect to the data-dependent bi-linear form:

\[
\langle \psi_i(\xi), \phi_j(\xi) \rangle := \sum_{k=1}^{m} \psi^H_i(\xi_k)w^H_{2k}w_{2k}\phi_j(\xi_k).
\]

Numerically reliable algorithms have been developed and implemented to construct these (bi-)orthonormal basis functions. The major computational step in these algorithms involves solving a structured inverse eigenvalue problem (see e.g. [26]). Especially the construction of the block polynomial basis functions that are orthonormal with respect to (22) is known to be numerically stable since it involves much used unitary zeroing operations.

**V. EXPERIMENTAL EXAMPLE**

In this section, the proposed method is applied to identify a model of a lightweight prototype wafer stage.

**A. The experimental setup**

In Figure 2 the wafer stage setup is shown. It is controlled in six motion degrees of freedom and is equipped with addition actuators and sensors to enable the control of flexible dynamics. In this example only the out-of-plane motions are considered. This means there are three relevant rigid body degrees of freedom (one translation in $z$-direction and two rotations around the $x$ and $y$ axes). One additional actuator and sensor are used, meaning the considered system has a total of four inputs and four outputs, the configuration of these inputs and outputs can be seen in Figure 3, which also contains a bottom view of the considered wafer table where some of these actuators and sensors can be seen.

The fourth sensor is a piezoelectric sensor which is attached to the wafer table and which measures internal
deformations. This sensor is therefore fully decoupled from the rigid body dynamics of the system. Therefore, a decoupled parametrization as described in section III-B can be used without the need for a preliminary decoupling step. In this example however, the rigid body modes are decoupled from the fourth output and not from any of the inputs meaning the RMFD based parametrization (17) cannot be used directly. This problem is easily fixed by either considering an LMDT implementation or by transposing the problem data, switching the roles of the inputs and outputs (as far as the identification problem is concerned).

B. Identification procedure

The parametric identification procedure for this setup starts from a non-parametric FRM estimate with 5 · 10^3 frequency points obtained from a closed-loop multisine identification experiment, see [5] for details. Modeling for this example will be done in continuous time, therefore the data is first compensated for delay, as delay cannot be accurately modeled by a rational function in s-domain. Subsequently, the non-parametric FRM is transposed so the rigid body behavior becomes decoupled from the new, virtual fourth input. Next, the FRM is right-multiplied by $I_{ex}(s_k) = \text{diag}(s_k^2, s_k^2, s_k^2, 1)$.

Some additional weighting (by a factor 10) is applied in the middle frequency range (from 15 to 160 Hz). Also, the weight placed on the high frequency range (upward of 2 kHz) is reduced by a factor 100 and frequency points upward of 4 kHz are omitted. This additional weighting reflects the region in which the most accurate model is desired for subsequent use in model based control or observer design.

Finally, models of McMillan degree 82 are identified using 75 iterations of the SK algorithm, described in section II-B, superseded by 150 iterations of the Levenberg-Marquardt algorithm. This is done using the numerically reliable implementation with data-dependent orthonormal basis functions, as described in section IV, as well as two standard implementations, one using a monomial basis and one using a monomial basis scaled such that columns of the problem matrix are normalized. Note that the identified models are inherently MIMO systems, i.e., the full 4 × 4 system is of McMillan degree 82, see (16), (17). This order can be specified lower, especially when a control relevant model is desired, see [2].

C. Results

In Figure 4 the fit results of the described identification procedure, using the data-dependent orthonormal basis functions, are shown. It can be seen in this figure that the identification procedure indeed achieves a good quality fit. Convergence of the cost function value has been achieved and the fitted model accurately describes the rigid-body dynamics as well as the most relevant flexible dynamics of the system. The rigid body dynamics are clearly present in the first three outputs and are entirely absent in the fourth output, corresponding to the piezoelectric deformation sensor. Up to a frequency of about 2 kHz the models overlap well with the measured frequency response matrix, at even higher frequencies the model mismatches become more apparent. This is to be expected since the amount of flexible modes that are present at these high frequencies becomes very high and also the quality of the data decreases significantly in this region. The choice to accept a poorer model fit in this high frequency region is also reflected by the additional weighting that is used, as described section V-B.

As can be seen in Figure 5, the condition number during iterations for the implementation with data-based orthonormal polynomials is indeed equal to one. The maximum conditioning number during iterations for this implementation was in fact $\kappa_{\text{max}} \approx 1.0008$. This shows that this orthonormal basis can indeed be constructed in a numerically reliable fashion for MIMO system identification with a relatively high order and with a large amount of data points.

For the implementation with unscaled monomial basis
functions, the conditioning number is high leading to a significantly reduced performance of the algorithm in terms of achieved minimum cost function value and convergence.

For the implementation using the scaled monomial basis, the condition number during iterations can be seen to increase from $O(10^5)$ to $O(10^{10})$. During the first iterations, where this implementation is still relatively well conditioned, the performance of the SK algorithm is identical to the optimally conditioned case. This should be the case since a change of basis functions does not change the algorithm. After the 20th iteration however, the cost function values for these two implementations diverge, showing that the conditioning indeed impacts the performance of the algorithm.

VI. CONCLUSIONS AND OUTLOOK

In this research, the identification of motion systems containing rigid body degrees of freedom is considered. The main goal is to enforce the desired rigid body behavior while using a specific data-dependent basis to achieve optimal numerical conditioning of the iterative identification algorithms. As shown in section III it is often possible to factorize the rigid body behavior out of the model description. This is an effective way to enforce the presence of rigid body modes in the model while still being free to generically parametrize the model. The performance of this procedure is shown experimentally on a lightweight prototype wafer stage.

From the experimental results it can be concluded that the data dependent orthonormal basis functions can indeed be constructed in a numerically reliable way for practically relevant MIMO system identification.

Ongoing work includes establishing generic necessary and sufficient conditions for enforcing rigid body modes. Also extending the numerically reliable identification framework, is a topic of ongoing research, as well as improving the current implementations.

REFERENCES