On numerically reliable frequency-domain system identification: new connections and a comparison of methods

Robbert Voorhoeve, * **Tom Oomen,** * **Robbert van Herpen,** *
Maarten Steinbuch *

* *Eindhoven University of Technology, Department of Mechanical*
Engineering, Control Systems Technology group, Eindhoven, 5600MB,
The Netherlands (e-mail: r.j.voorhoeve@tue.nl).

Abstract: Frequency domain identification of complex systems imposes important challenges with respect to numerically reliable algorithms. This is evidenced by the use of different rational and data-dependent basis functions in the literature. The aim of this paper is to compare these different methods and to establish new connections. This leads to two new identification algorithms. The conditioning and convergence properties of the considered methods are investigated on simulated and experimental data. The results reveal interesting convergence differences between (nonlinear) least squares and instrumental variable methods. In addition, the results shed light on the conditioning associated with so-called frequency localising basis functions, vector fitting algorithms, and (bi)-orthonormal basis functions.

Keywords: Frequency domain identification; Mechanical and aerospace estimation; Software for system identification.

1. INTRODUCTION

Frequency domain identification is commonly applied for modelling in various fields, including mechanical systems (Bayard and Chiang, 1998), (Oomen et al., 2013), as well as power systems (Deschrijver et al., 2008). In these fields several identification algorithms have been developed such as the SK (Sanathanan and Koerner, 1963), Vector Fitting (VF) (Gustavsen and Semlyen, 1999) and Instrumental Variable (IV) algorithms (Blom and Van den Hof, 2010). A key challenge that is encountered in the implementation of these algorithms is poor numerical conditioning, leading to inaccurate solutions and unreliable convergence properties of the algorithms.

Several partial solutions to this problem have been developed such as the use of frequency scaling (Pintelon and Kollár, 2005) and the use of various orthonormal basis functions (Ninness and Hjalmarsson, 2001), which have advantageous numerical over the basic monomial basis functions. Also, the use of several rational bases in frequency domain identification has been investigated (Ninness et al., 2000), where a recent development is the use of frequency localising basis functions (Welsh and Goodwin, 2003). Rational bases are also frequently used in the electronics field in the form of the Vector Fitting algorithm (Gustavsen and Semlyen, 1999).

The use of data-dependent polynomial bases has also yielded some important results. It has been shown in Bultheel and van Barel (1995), that orthonormal block-polynomials can be used to achieve optimal conditioning for the iterations of the SK algorithm. Recently, in van Herpen et al. (2014) it has been shown that a similar result can be achieved for the Instrumental Variable algorithm using a set of bi-orthonormal block polynomials.

Although several methods have been developed to improve the numerical conditioning of these algorithms, a thorough comparison between these methods is not yet present.

Therefore, the aim of this paper is to:

1. Establish new connections between these methods.
2. Benchmark the different methods on simulated and experimental data.

The main contributions of this paper are:

C.1 (a) The development of an SK routine with frequency localising basis functions and pole relocation.
   (b) The development of a vector fitting routine using the instrumental variables algorithm.
C.2 A thorough comparison, on simulated and experimental data, of the proposed and existing identification methods with improved numerical conditioning.

In the upcoming section the existing identification algorithms are introduced and the problem is formulated. In section 3, the use of rational basis functions in the identification problem is investigated and two new identification methods are introduced, constituting contribution C.1 of this research. In section 4, the use of data-dependent bases in system identification and the theoretical results that have been achieved with such bases are explained. In section 5, the benchmark results of the various considered methods are presented and discussed, constituting contribution C.2 of this research. Finally, in Section 6 the conclusions of the presented work are given.

2. PROBLEM FORMULATION

2.1 Parametric identification criterion

In this paper, parametric frequency-domain identification based on the quadratic cost criterion (1) is considered,

$$V(\theta) := \sum_{k=1}^{m} \left| W(s_k)(P_o(s_k) - \hat{P}(s_k, \theta)) \right|^2. \qquad (1)$$

The goal is to minimise $V(\theta)$. In (1), $P_o(s_k)$, $k = 1, \ldots, m$, are given frequency response function (FRF) measurements and $\hat{P}(s, \theta)$ denotes the parametric model that is to

be estimated. In the present paper continuous time system identification is considered, but equal results can be obtained for the discrete time case. Using (1), a broad range of system identification criteria can be pursued through an appropriate choice of the weighting function $W(s_k)$, see Pintelon and Schoukens (2012) for an overview. The model structure considered in this paper is,

$$\hat{P}(s,\theta) = \frac{B(s,\theta)}{A(s,\theta)} = \frac{N(s,\theta)}{D(s,\theta)}. \qquad (2)$$

Here $B(s,\theta)$, $A(s,\theta) \in \mathbb{R}[s]$ are the numerator and denominator polynomials respectively. The functions $N(s,\theta)$ and $D(s,\theta)$ are linear in the parameters θ with respect to a set of basis functions, these are not necessarily polynomials. Minimising (1) is a non-convex optimization problem.

2.2 Two iterative identification algorithms

In classical literature on parametric system identification, the non-convexity of this optimization problem has been addressed by considering a related convex optimization problem. Observe that $V(\theta)$ in (1) can be rewritten as

$$V(\theta) = \sum_{k=1}^{m} \left| \frac{W(s_k)}{D(s_k,\theta)} \begin{bmatrix} P_o(s_k) & -1 \end{bmatrix} \begin{bmatrix} D(s_k,\theta) \\ N(s_k,\theta) \end{bmatrix} \right|^2. \qquad (3)$$

By multiplying (3) with $D(s_k,\theta)$, $V(\theta)$ is linearised.

Linearised least squares. (Levy, 1959) Approximate the solution θ to (1) by minimising

$$\tilde{V}(\theta) := \sum_{k=1}^{m} \left| W(s_k) \begin{bmatrix} P_o(s_k) & -1 \end{bmatrix} \begin{bmatrix} D(s_k,\theta) \\ N(s_k,\theta) \end{bmatrix} \right|^2. \qquad (4)$$

Although (4) takes the form of a standard linear regression problem that can be solved straightforwardly, a different cost criterion is optimized compared to (1). Compared to (1), the cost function of (4), is weighted with $D(s,\theta)$. For a polynomial parametrisation of $D(s,\theta)$, this typically overemphasises high frequency errors. To improve the estimate, the a priori unknown weighting factor $D(s_k,\theta)^{-1}$, needs to be approximated more adequately. For this purpose, iterative identification algorithms are utilised.

Algorithm 1. (SK-iteration). (Sanathanan and Koerner, 1963) Let $\theta^{\langle 0 \rangle}$ be given. In iteration $i = 1, 2, \ldots$, determine the least-squares solution to

$$\min_{\theta^{\langle i \rangle}} \sum_{k=1}^{m} \left| \frac{W(s_k)}{D(s_k,\theta^{\langle i-1 \rangle})} \begin{bmatrix} P_o(s_k) & -1 \end{bmatrix} \begin{bmatrix} D(s_k,\theta^{\langle i \rangle}) \\ N(s_k,\theta^{\langle i \rangle}) \end{bmatrix} \right|^2. \qquad (5)$$

Typically, Alg. 1 converges and yields a small value of the cost function (1), outperforming the non-iterative solution (4). However, the fixed point of the iteration is generally not a local optimum of $V(\theta)$, as is shown in Whitfield (1987). To address this problem, Gauss-Newton iterations can be used, after the SK algorithm converges, to guarantee that the fixed point of iteration is a local minimum of (1).

An alternative iterative identification algorithm is known in the literature, which is based on an Instrumental Variable (IV) approach. This algorithm is known in time domain identification as the Simplified Refined Instrumental Variable method (SRIV) Young (1976), and has recently been formulated for frequency domain identification (Blom and Van den Hof, 2010). The fixed point of iteration for this algorithm is equal to a local minimum of the cost function, rendering such additional Gauss-Newton iterations superfluous. For this algorithm the first necessary

condition for optimality $\frac{d}{d\theta}V(\theta) = 0$, is considered. For (1), this condition becomes:

$$\sum_{k=1}^{m} \left[\frac{-\partial \hat{P}(s_k,\theta)}{\partial \theta^T} \right]^* W^*(s_k) W(s_k)(P_o(s_k) - \hat{P}(s_k,\theta)) = 0. \qquad (6)$$

This is a non-linear equation in θ which can be solved with an iterative scheme similar to algorithm 1. This leads to the following algorithm.

Algorithm 2. (IV-iterations). (Blom and Van den Hof, 2010) Let $\theta^{\langle 0 \rangle}$ be given. In iteration $i = 1, 2, \ldots$, solve the linear system of equations for $\theta^{\langle i \rangle}$

$$\sum_{k=1}^{m} \left[\frac{-\partial \hat{P}(s_k,\theta)}{\partial \theta^T} \bigg|_{\theta = \theta^{\langle i-1 \rangle}} \right]^* W^*(s_k) \qquad (7)$$

$$\frac{W(s_k)}{D(s_k,\theta^{\langle i-1 \rangle})} \begin{bmatrix} P_o(s_k) & -1 \end{bmatrix} \begin{bmatrix} D(s_k,\theta^{\langle i \rangle}) \\ N(s_k,\theta^{\langle i \rangle}) \end{bmatrix} = 0.$$

Remark 1. This algorithm can be recognised as an instrumental variable method since it replaces the orthogonal projection, $\theta^{\langle i \rangle} = (X^*X)^{-1}X^*y$, of the least squares problem with an oblique projection $\theta^{\langle i \rangle} = (Z^*X)^{-1}Z^*y$, which is the defining characteristic of an IV method.

The convergence properties of this Instrumental Variables algorithm are theoretically superior to other algorithms. However it often suffers from prohibitive numerical conditioning problems.

2.3 Problem formulation

The identification methods considered in this paper can all be characterised as one of the two iterative algorithms discussed in the previous section. Both these algorithms are known to suffer from problems regarding numerical conditioning, as is evidenced by the multitude of methods proposed in literature to improve this conditioning (Van Herpen et al., 2014) (Ninness et al., 2000) (Pintelon and Kollár, 2005). The conditioning problems are briefly explained in this section, after which the considered methods to improve this numerical conditioning, are given.

The iterations from the SK algorithm (Alg. 1) can be written as the overdetermined system of equations (Van Herpen et al., 2014):

$$W_1 \Phi \theta^{\langle i \rangle} = b_n. \qquad (8)$$

This system can be solved through QR factorisation followed by backward substitution. The relevant condition number for these iterations therefore is $\kappa(W_1\Phi)$. In (8), Φ is a matrix containing the basis functions. If a monomial polynomial basis is used this becomes a block-Vandermonde matrix which is notoriously ill-conditioned, i.e. $\kappa(W_1\Phi) \gg 1$.

For the IV algorithm (Alg. 2) the set of equations that are solved each iteration can be written as:

$$\Phi^* W_2^* W_1 \Phi \theta^{\langle i \rangle} = \Phi^* W_2^* b_n. \qquad (9)$$

The typical conditioning problems for this set of equations are significantly worse than for (8), since:

$$\kappa(\Phi^* W_2^* W_1 \Phi) \approx \kappa(\Phi^* W_1^* W_1 \Phi) = \kappa(W_1\Phi)^2. \qquad (10)$$

Therefore conditioning is much more critical for the IV algorithm. In this paper, the use of rational basis functions and the use of data-dependent basis functions, to improve numerical conditioning, will be considered in sections 3 and 4 respectively.

3. RATIONAL BASES

In this section, identification using rational basis functions is discussed. Two choices of rational basis functions are considered and with these basis functions two novel identification methods are formulated. This constitutes contribution (C.1) of this paper.

3.1 Rational basis functions

The functions $N(s,\theta)$ and $D(s,\theta)$ are by definition linear in the parameters with respect to a set of basis functions. Although these are usually chosen to be polynomials, this is not necessary in general. In this section rational basis functions are considered, as is done in multiple earlier publications (Ninness et al., 2000) (Welsh and Goodwin, 2003) (Gustavsen and Semlyen, 1999). From these publications, it is known that such rational basis functions can lead to a better conditioning of the identification algorithms considered in section 2.2. When using rational basis functions, $N(s,\theta)$ and $D(s,\theta)$ are given by

$$N(s,\theta) = \sum_{p=1}^{n_n} \theta_{\text{num},p}\, \phi_p(s) = \frac{B(s,\theta)}{E(s)} \qquad (11)$$

$$D(s,\theta) = \sum_{p=1}^{n_d} \theta_{\text{den},p}\, \phi_p(s) = \frac{A(s,\theta)}{E(s)}. \qquad (12)$$

Two choices of rational basis functions ϕ_p, are a partial fractions basis (PF)(Gustavsen and Semlyen, 1999) and frequency localising basis functions (FL) (Welsh and Goodwin, 2003):

$$\phi_{\text{PF},p} = \frac{1}{s+a_p}, \quad \phi_{\text{FL},p} = \frac{|a_p|}{s+a_p} \prod_{l=1}^{p-1}\left(\frac{s}{s+a_l}\right). \qquad (13)$$

Both these sets of basis functions are completely determined by the choice of a set of poles a_p and the common denominator polynomial $E(s)$ is given by:

$$E(s) = \prod_{p=1}^{n}(s+a_p). \qquad (14)$$

With an appropriate choice for the poles of the basis functions, the effect of the a priori unknown weighting introduced in the linearised least squares criterion (4) can be significantly reduced. This is due to the fact that this weighting is given by $D(s,\theta)$, which for the rational case is a bi-proper rational function instead of a (high order) polynomial in the standard case. This significantly reduces the overemphasis on high frequency errors that (4) typically suffers from. In Welsh and Goodwin (2003) the use of such a linearised least squares solution using frequency localising basis functions is proposed for wide band frequency domain system identification. This method is referred to as FLBF-LS in this paper.

To further reduce the effects of the a priori unknown weighting introduced in this method, Gilson et al. (2013) proposed to use an iterative IV algorithm. However, this IV algorithm is still applied to the linearised problem which is weighted with $D(s_k,\theta)$. The point of convergence of this algorithm is therefore different from the point of convergence of the IV method proposed by Blom and Van den Hof (2010). As a result, this routine does not converge to a minimum of the original cost function (1), as will be seen in the benchmark results. The IV algorithm of Gilson et al. (2013) is referred to as FLBF-IV in the present paper.

A possible reason for solving this weighted problem (multiplied with $D(s_k,\theta)$) is that dividing the problem with $D(s_k,\theta^{\langle i-1\rangle})$ (or multiplying it with $\frac{E(s_k)}{A(s_k,\theta^{\langle i-1\rangle})}$), as is done in the iterative algorithms of section 2.2, leads to:

$$\min_{\theta^{\langle i\rangle}} \left\| \frac{W(s_k)\,E(s_k)}{A(s_k,\theta^{\langle i-1\rangle})}\left(P_0(s_k)\frac{A(s_k,\theta^{\langle i\rangle})}{E(s_k)} - \frac{B(s_k,\theta^{\langle i\rangle})}{E(s_k)}\right)\right\|_2^2. \qquad (15)$$

Which is an optimization problem where the common denominator polynomial $E(s_k)$, is eliminated from the equation. Because of this, the advantageous numerical properties of the rational basis functions are lost. To address this problem, the pole relocation technique, as used in vector fitting (Gustavsen and Semlyen, 1999), can be used. This technique is introduced next, after which it is shown how this technique can be used with the frequency localising basis functions.

3.2 Pole relocation

As mentioned in the previous section, the explicit weighting of the identification problem with $D(s_k,\theta^{\langle i-1\rangle})^{-1}$, results in a loss of the advantageous properties of the rational basis functions. Pole relocation is a technique that implicitly applies this weighting, thus retaining the structure and properties of the rational basis functions. The equivalence between pole relocation and the SK algorithm has been shown in Hendrickx and Dhaene (2006).

The implicit weighting is achieved by iteratively relocating the poles of the rational basis functions. The updated poles are taken equal to the zeros of the denominator function of the previous iteration, i.e.

$$\tilde{z}_p^{\langle i-1\rangle} = \text{zeros}\left\{D(s,\theta^{\langle i-1\rangle})\right\}, \qquad (16)$$

which means $E(s)$ becomes

$$E(s)^{\langle i\rangle} = \prod_{p=1}^{n}(s - \tilde{z}_p^{\langle i-1\rangle}). \qquad (17)$$

Because the polynomial $A(s,\theta^{\langle i-1\rangle})$ is equal to $E(s)^{\langle i\rangle}$, (15) can be rewritten as:

$$\min_{\theta^{\langle i\rangle}} \left\| W(s_k)\left(P_0(s_k)\frac{A(s_k,\theta^{\langle i\rangle})}{E(s_k)^{\langle i\rangle}} - \frac{B(s_k,\theta^{\langle i\rangle})}{E(s_k)^{\langle i\rangle}}\right)\right\|_2^2. \qquad (18)$$

Since the poles of the rational basis functions are also updated such that $a_p^{\langle i\rangle} = \tilde{z}_p^{\langle i-1\rangle}$, the terms $\frac{A(s,\theta^{\langle i\rangle})}{E(s)^{\langle i\rangle}}$ and $\frac{B(s,\theta^{\langle i\rangle})}{E(s)^{\langle i\rangle}}$ are equal to $D(s,\theta^{\langle i\rangle})$ and $N(s,\theta^{\langle i\rangle})$ respectively.

To compute the zeros $\tilde{z}_p^{\langle i-1\rangle}$, a state space representation of the denominator $D(s,\theta)$ is constructed after which the zeros are found by solving the eigenvalue problem:

$$\tilde{z}_p = \text{eig}(\mathcal{A} - \mathcal{B}\mathcal{D}^{-1}\mathcal{C}) \qquad (19)$$

The iterative pole relocation procedure, which is equivalent to the SK algorithm Alg. 1, consists of the following steps. First, construct basis functions from a given set of poles $a_p^{\langle i\rangle}$, then solve

$$\min_{\theta^{\langle i\rangle}} \left\| W(s_k)\left(P_0(s_k)D(s_k,\theta^{\langle i\rangle}) - N(s_k,\theta^{\langle i\rangle})\right)\right\|_2^2. \qquad (20)$$

Next, construct the state space system of the rational denominator function $D(s,\theta^{\langle i\rangle})$ and compute the zeros of this system by solving (19). Finally, take $a_p^{\langle i+1\rangle} = \tilde{z}_p^{\langle i\rangle}$ and iterate until convergence is reached.

This iterative pole relocation procedure, can next be used to implement a new adaptation of the SK algorithm (Alg. 1), utilising the frequency localising basis functions.

3.3 FLBF with pole relocation

Pole relocation can be used to formulate the SK algorithm with frequency localising basis functions, without losing the advantageous numerical properties of these basis functions. To be able to use the pole relocation procedure with the frequency localising basis functions, as proposed in Welsh and Goodwin (2003), two extensions of the existing method need to be made.

First, the basis functions must be able to incorporate complex conjugate pole pairs while maintaining the guarantee of a real-rational transfer function and a real parameter vector θ. Second, a state space representation must be constructed from the basis functions and the parameter vector θ. This state space representation is used to compute the zeros of the rational denominator function.

For complex conjugate poles $a_p^* = a_{p+1}$, the following basis functions are constructed to make sure the transfer function and the parameter vector θ are real-valued:

$$\psi_p(s) = \left(\prod_{l=1}^{p-1} \frac{s}{s + a_l} \right) \frac{|a_p|(s - |a_p|)}{(s + a_p)(s + a_{p+1})}$$

$$\psi_{p+1}(s) = \left(\prod_{l=1}^{p-1} \frac{s}{s + a_l} \right) \frac{|a_p|(s + |a_p|)}{(s + a_p)(s + a_{p+1})} \quad (21)$$

A minimal, real valued, state space realization of $D(s, \theta)$ can be constructed by considering the full state space system as the compound system of the individual basis functions. This method is related to the one used in Deschrijver et al. (2007) to construct the state space representation for orthonormal vector fitting. Using this state space description, the zeros of $D(s)$ can be computed using (19), and an SK-FLBF algorithm using pole relocation is obtained. The development of this method constitutes contribution C.1(a) of this paper.

3.4 Vector fitting with instrumental variables

Recently, it was shown that Vector Fitting is equivalent to the SK algorithm (Hendrickx and Dhaene, 2006), however the SK algorithm is known not to converge to a local minimum of the cost function (1), as is shown in Whitfield (1987). This shortcoming of the current vector fitting algorithms has been acknowledged in various VF related publications (Hendrickx and Dhaene, 2006)(Deschrijver et al., 2007), but has not yet been resolved. The solution proposed in this paper is a new vector fitting procedure that makes use of the IV algorithm (Alg. 2).

To achieve this, the pole relocation technique that is used in vector fitting has to be incorporated into the IV algorithm. This is possible because, in the Instrument Z^* of the IV algorithm, the basis functions are also weighted with $D(s_k, \theta^{\langle i-1 \rangle})^{-1}$. Therefore, relocating the poles implicitly applies the appropriate weighting to both the Instrument Z^* and the problem matrix X of the IV algorithm. The new identification method resulting from the implementation of the oblique IV projection in the vector fitting procedure, will be referred to as IV-VF. The development of this method corresponds to contribution C.1(b) of this paper.

4. DATA-DEPENDENT BASES

The rational basis functions of the previous section have shown good results. A key question that remains is: which choice of basis functions leads to the best possible conditioning of the identification algorithms of section 2.2.

The answer to this is that it depends on the problem data. 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 Bultheel and van Barel (1995) it is shown that choosing the basis functions for the SK iterations to be block-polynomials that are orthonormal with respect to the data-dependent inner product (22), then optimal conditioning of the SK iterations is achieved.

$$\langle \phi_i(\xi), \ \phi_j(\xi) \rangle := \sum_{k=1}^{m} \phi_j^*(\xi_k) w_{1k}^* w_{1k} \phi_i(\xi_k). \quad (22)$$

This result does not hold for IV iteration however, since the IV iterations involve the oblique projection $Z^* X \theta = Z^* b$, instead of the orthonormal projection $X^* X \theta = X^* b$. To achieve optimal conditioning of the IV iterations a separate set of block-polynomials (Van Herpen et al., 2014) needs to be constructed for the instrument and for the variables, transforming (9) to

$$(\Psi^* W_2^* W_1 \Phi) \theta^{\langle i \rangle} = \Psi^* W_2^* b_n. \quad (23)$$

These two sets of block polynomials are then constructed to be bi-orthonormal with respect to the data-dependent bi-linear form:

$$\langle \psi_i(\xi), \ \phi_j(\xi) \rangle := \sum_{k=1}^{m} \psi_j^*(\xi_k) w_{2k}^* w_{1k} \phi_i(\xi_k). \quad (24)$$

This yields optimally conditioned ($\kappa = 1$) iterations for the instrumental variable algorithm. Efficient algorithms to construct these bases have been developed by Bultheel and van Barel (1995) and van Herpen et al. (2014). In the coming section, the benchmark results will also enable an examination of the numerical robustness of these methods for constructing the data-dependent basis functions.

5. BENCHMARK RESULTS

In this section the performance of the proposed and existing methods is compared for a large set of simulated data as well as an experimental example. This comparison of methods constitutes contribution (C.2) of this paper.

5.1 Simulation data

To be able to thoroughly compare all the proposed and existing methods a dataset is generated consisting of 100 systems of the form:

$$G(s) = \sum_{l=1}^{8} \frac{b_l \omega_l^2}{s^2 + 2\zeta_l \omega_l s + \omega_l^2}, \quad (25)$$

where ω_l spans approximately 9 decades and the parameters b_l, ω_l and ζ_l are generated randomly from a predetermined range. The systems are all evaluated at 1500 logarithmically spaced points that span the entire frequency range of the system. To this sampled data, random noise is added in a multiplicative manner, i.e. $P_o = G(s_k)(1 + \varepsilon(s_k))$, where $\varepsilon(s_k)$ is zero mean Gaussian white noise, such that the signal to noise ratio is 20dB. An example of one of the simulated frequency responses can be seen in Figure 1.

The considered identification methods are :

- the proposed SK FLBF method with pole relocation as presented in section 3.3;
- the Vector Fitting method as developed by Gustavsen and Semlyen (1999), SK VF;
- the SK algorithm with a data-dependent orthonormal polynomial basis, SK OP;

- the SK algorithm with monomial basis, SK Mon;
- the IV FLBF method as proposed by Gilson (2013);
- the proposed IV VF method as presented in section 3.4;
- the IV algorithm with a data-dependent bi-orthonormal polynomial basis, IV BP;
- the IV algorithm with monomial basis, IV Mon.

For the implementation of the Vector fitting methods, a modified version of the freely available `vecfit3.m` routine (`http://www.sintef.no/vectfit`) is used.

For each of the 100 generated frequency response measurements, 25 iterations of all of the considered methods are computed. All of the methods are initialised using the FLBF-LS estimator as described in section 3.1 and as first proposed in Welsh and Goodwin (2003).

Table 1. Converged cost function value of the different methods for the simulation data.

	SK	IV	IV FLBF
$V(\theta^\star)$	$1.4 \cdot 10^4$	$6.9 \cdot 10^3$	$1.3 \cdot 10^4$

In Table 1 the average converged cost function values for the considered methods are depicted. First, the IV based methods converge to a lower value of the cost function than the SK based methods. Second, the IV-FLBF method as proposed by Gilson et al. (2013), on average, converges to a higher value of the cost function than the other IV based methods. This can be explained from the earlier observation that the IV-FLBF method solves a weighted version of the optimal IV criterion (7) , meaning that, the IV-FLBF method does not converge to a local minimum of the cost function (1).

After further inspection of the data, the discrepancy between the converged values of the cost function for the SK and IV methods, is dominated by a small fraction of the analysed systems. For this fraction, the SK algorithms fail to fit one of the eight resonance peaks that are present in the system, leading to a much higher value of the cost function. An example of this is shown in Figure 1.

The fact that for some of these systems the IV algorithm does fit all the resonances peaks and the SK algorithm does not, can be attributed to the known property of SK algorithm (Whitfield, 1987) that the fixed point of the iterations does not coincide with a local minimum of the cost function. Because of this property, the SK algorithm might be more likely to get "stuck". To be able to state this with certainty, however, a more rigorous analysis needs to be performed.

Table 2. Average conditioning of the methods for the simulation data.

	κ_{SK}	κ_{IV}
Orth	1.00	1.00
FLBF	$1.9 \cdot 10^7$	$4.6 \cdot 10^9$
VF	$5.7 \cdot 10^{12}$	$2.0 \cdot 10^{27}$
Mon	$9.0 \cdot 10^{39}$	$1.2 \cdot 10^{149}$

In Table 2 the average numerical conditioning properties for seven of the considered methods are depicted. The results depicted in Table 2 show a couple of clear results. First, the conditioning when using Monomial basis functions is prohibitively high for the identification of the considered class of systems. Second, the conditioning for IV based methods is indeed significantly worse than for SK based methods, as expected from the analysis in section
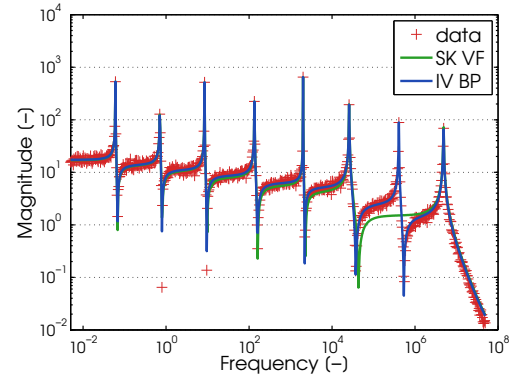

Fig. 1. Simulated FRF, with an IV fit which captures all the resonances and an SK fit which misses one.

2.3. Third, the use of rational basis functions does significantly improve conditioning over the use of a monomial polynomial basis. Fourth, the frequency localising basis functions lead to a much better conditioned problem than the partial fractions basis, which is traditionally used in Vector Fitting.

In addition, the (bi-)orthonormal data-dependent bases indeed lead to an optimally conditioned identification problem. This underlines the results of Bultheel and van Barel (1995) and Van Herpen et al. (2014). It also shows that the recently developed methods for the construction of data-dependent (bi-)orthonormal bases are practically viable for the optimally conditioned identification of complex systems.

5.2 Experimental data

For further validation of the considered identification methods, an experimental example is considered. The experimental set-up is that of the Active Vibration Isolation System which is depicted in Figure 2.

A 15th order parametric model is fitted to the identified FRF data. Fifty iterations for each of the identification methods are computed. The converged value of the cost function (1) is again computed and the results are shown in Table 3. The FRF data and the resulting fits for three methods (SK, IV and IV-FLBF) are shown in Figure 2 and the average conditioning for the considered methods is shown in Table 4.

Again, the methods based on the IV algorithm converge to slightly a lower value of the cost function than those based on the SK algorithm. The FLBF IV algorithm converges to a higher value of the cost function than the other IV methods.

Table 3. Converged cost function values of the different methods for the AVIS data.

	SK	IV	IV FLBF
$V(\theta^\star)$	$5.6 \cdot 10^1$	$2.8 \cdot 10^1$	$4.5 \cdot 10^1$

Table 4. Average conditioning of the methods for the AVIS data.

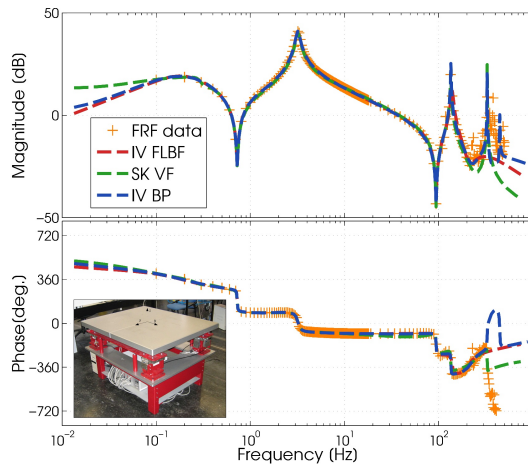|	κ_{SK}	κ_{IV}
Orth	1.00	1.00
FLBF	$4.8 \cdot 10^4$	$1.1 \cdot 10^{10}$
VF	$9.2 \cdot 10^3$	$3.6 \cdot 10^8$
Mon	$2.8 \cdot 10^{22}$	$1.0 \cdot 10^{72}$

Fig. 2. Bode diagram of AVIS FRF measurements and various resulting parametric fits.

The conditioning results for the AVIS data as depicted in Table 4 show a slightly different result than the simulation data. For this identification problem, the conditioning of the methods using frequency localising basis functions is worse than the conditioning of the vector fitting methods that use a partial fractions basis. This difference between the simulated and experimental data might indicate that the frequency localising basis functions are especially suited for systems with the structure (25), as is used in the simulations.

The results in Figure 2 show that an excellent parametric fit is obtained with each of the considered methods up to about 3000 Hz after which the FRF data either becomes too noisy or has a resolution that is too low to accurately determine additional modes. Qualitatively there is not much separating the parametric fits but considering the converged cost function values, it is concluded that the IV algorithm yields the most accurate fit.

6. CONCLUSIONS

Two new identification methods are developed, which constitutes contribution C.1 of this research. These methods show that pole relocation is a versatile and effective technique to formulate the SK and IV algorithms, using rational basis functions.

Overall conclusions that can be drawn from the benchmarking results of section 5, which constitutes contribution C.2 of this research, are as follows. First, it is clear that the IV algorithm has favourable convergence properties over the SK algorithm. However the conditioning of the IV iterations is in general much worse that for the SK iterations. Second, the considered rational basis functions lead to much better conditioning than the monomial polynomial basis.

It can also be concluded that the most rigorous solution for improving conditioning is the use of the data-dependent (bi-)orthonormal polynomials. This method does require more computational effort, but it has been shown that these bases can be reliably computed. With increasing model order, the conditioning of the identification algorithms will become even more critical. Therefore, the use of a data-dependent basis might prove essential for the formulation of a numerically reliable identification algorithm for increasingly complex systems.

REFERENCES

Bayard, D.S. and Chiang, R.Y. (1998). Identification, uncertainty characterization and robust control synthesis applied to large flexible structures control. *International Journal of Robust and Nonlinear Control*, 8(2), 97–112.

Blom, R. and Van den Hof, P. (2010). Multivariable frequency domain identification using iv-based linear regression. In *Proc. of the 49th Conference on Decision and Control*, 1148–1153.

Bultheel, A. and van Barel, M. (1995). Vector orthogonal polynomials and least squares approximation. *SIAM Journ. Matrix Analysis and Applications*, 16(3), 863–885.

Deschrijver, D., Haegeman, B., and Dhaene, T. (2007). Orthonormal vector fitting: A robust macromodeling tool for rational approximation of frequency domain responses. *IEEE Transactions on Advanced Packaging*, 30(2), 216–225.

Deschrijver, D., Mrozowski, M., Dhaene, T., and De Zutter, D. (2008). Macromodeling of multiport systems using a fast implementation of the vector fitting method. *IEEE Microwave and Wireless Components Letters*, 18(6), 383–385.

Gilson, M., Welsh, J.S., and Garnier, H. (2013). Frequency-domain instrumental variable based method for wide band system identification. In *Proc. of the 2013 American Control Conference*, 1663–1668.

Gustavsen, B. and Semlyen, A. (1999). Rational approximation of frequency domain responses by vector fitting. *IEEE Transactions on Power Delivery*, 14(3), 1052–1061.

Hendrickx, W. and Dhaene, T. (2006). A discussion of "rational approximation of frequency domain responses by vector fitting". *IEEE Transactions on Power Systems*, 21(1), 441–443.

van Herpen, R., Oomen, T., and Steinbuch, M. (2014). Optimally conditioned instrumental variable approach for frequency-domain system identification. *Automatica, accepted*.

Levy, E.C. (1959). Complex-curve fitting. *IRE Transactions on Automatic Control*, 37–43.

Ninness, B., Gibson, S., and Weller, S.R. (2000). Practical aspects of using orthonormal system parametrisations in estimation problems. In *Proc. of the 12th IFAC Symp. Syst. Ident.*

Ninness, B. and Hjalmarsson, H. (2001). Model structure and numerical properties of normal equations. *IEEE Transactions on Circuits and Systems - I: Fundamental Theory and Applications*, 48(4), 425–437.

Oomen, T., van Herpen, R., Quist, S., van de Wal, M., Bosgra, O., and Steinbuch, M. (2013). Connecting system identification and robust control for next-generation motion control of a wafer stage. *IEEE Trans. Contr. Syst. Techn.*

Pintelon, R. and Kollár, I. (2005). On the frequency scaling in continuous-time modeling. *IEEE Transactions on Instrumentation and Measurement*, 54(1), 318–321.

Pintelon, R. and Schoukens, J. (2012). *System Identification: A Frequency Domain Approach*. Wiley-IEEE press, Hoboken, NJ, USA, second edition.

Sanathanan, C.K. and Koerner, J. (1963). Tranfer function synthesis as a ratio of two complex polynomials. *IEEE Transactions on Automatic Control*, 8(1), 56–58.

Welsh, J.S. and Goodwin, G.C. (2003). Frequency localising basis functions for wide-band identification. In *Proc. of the 2003 European Control Conference*.

Whitfield, A.H. (1987). Asymptotic behaviour of transfer function sythesis methods. *International Journal of Control*, 45(3), 1083–1092.

Young, P. (1976). Some observations on instrumental variable methods of time-series analysis. *International Journal of Control*, 23(5), 593–612.