Initialization of Gradient-based Optimization Algorithms for the Identification of Structured State-Space Models *

O. Prot * G. Mercère **

* University of Limoges, Institut de Recherche XLIM, 123 Avenue A. Thomas, 87060 Limoges Cedex, France (e-mail: olivier.prot@unilim.fr).
** University of Poitiers, Laboratoire d’Automatique et d’Informatique Industrielle, 2 rue P. Brousse, bâtiment B25, B.P. 633, 86022 Poitiers Cedex, France (e-mail: guillaume.mercere@univ-poitiers.fr).

Abstract: Estimating consistent parameters of a structured (grey-box) state-space representation requires a reliable initialization when the vector of parameters is computed by using a gradient-based algorithm. Assuming that a reliable initial fully-parameterized state-space model is available, an algorithm dedicated to non-smooth optimization is introduced in this paper in order to transform this initial model into the structured state-space parameterization of the system to be identified. A specific constraint on the similarity transformation between both system representations is added to avoid singularity. Numerical examples highlight the performance of the developed approach.

Keywords: System identification; Convex optimization; Iterative methods; Parameterization.

1. INTRODUCTION

In the field of system identification, the estimation of the parameters of a model often involves a non-convex optimization (see, e.g., the well-known prediction-error methods (Ljung, 1999)). Usually, such a non-convex optimization problem is solved iteratively by using a gradient-based search. The efficiency and reliability of such an iterative search mainly depend on two conditions: the parameterization of the model and its initialization.

While many solutions are now available concerning the parameterization of the model (McKelvey, 1995; Ribarits, 2002; Verhaegen and Verdult, 2007; Wills and Nimness, 2008), the initialization problem of the gradient-based methods is not totally solved yet. This problem is a significant one because a bad initialization may lead to a local optimum of the cost function and, by extension, to irrelevant estimated parameters. In this paper, the problem of supplying a reliable initial vector of parameters is tackled by

- using a non-iterative method such as, e.g., a subspace-based algorithm (Verhaegen and Verdult, 2007), in order to get an accurate estimation of a (fully-parameterized) state-space form (black-box model),
- transforming this initial black-box model into a structured state-space representation corresponding to parameterization of the system.

This parameterized model is used in a third step as an initial model for the iterative gradient-based algorithm.

This framework has already been considered in the literature, mainly in (Xie and Ljung, 2002), (Parrilo and Ljung, 2003) and, more recently, in (Lyzell et al., 2009). Xie and Ljung (2002) are the first to write down a mathematical solution to this problem. Unfortunately, their solution still resorts to an iterative algorithm which requires good initial values. This problem was reconsidered by Parrilo and Ljung in 2003. More precisely, it is reformulated as a polynomial optimization problem which avoids an iterative algorithm. Despite its mathematical robustness, this solution is limited to small size matrices (around 20-24 unknown parameters as claimed by the authors). Finally, Lyzell et al. introduced in 2009 a numerically reliable approach to solve a specific reformulation of the initial problem. However, this solution is usable only for a small number of model structures.

In this paper, the cost function adopted in (Xie and Ljung, 2002) and (Parrilo and Ljung, 2003) is considered again. Roughly speaking, the underlying optimization problem consists in minimizing the input-output error between a black-box model and a structured state-space model. The optimized variables are the parameters of the structured state-space matrices as well as the similarity transformation matrix T between both linear representations of the same dynamical system. The main difficulty in this optimization problem comes from the non-convexity of the cost function. It is important to notice that our objective is not to find the global optimum, but rather a good local one.

* This work was supported by a “PEPS Automatique” grant from the Institut des Sciences de l’Ingénierie et des Systèmes, Centre National de Recherche Scientifique, France.
in order to initialize the iterative gradient-based algorithm used in the final prediction-error or output-error method.

As pointed out by Parrilo and Ljung (2003), the global solution to this optimization problem may be unusable in practice because of a nearly singular transformation matrix $\mathbf{T}$ at the global optimum. This situation must be avoided by taking care of the condition number of this matrix during the identification. In this paper, a new cost function is introduced in order to control the condition number of $\mathbf{T}$. This function is non-smooth because of the constraint on the condition number. To get round this difficulty, specialized algorithms are used to solve such a non-smooth optimization problem. Bundle methods are efficient techniques in the field of non-smooth optimization (Bomans et al., 2006). Hereafter, the spectral bundle method (Aphirian et al., 2008; Noll et al., 2008) is more precisely used because of its nice convergence properties.

The paper is organized as follows. Section 2 is dedicated to the problem statement. Section 3 addresses the problem of optimal $\mathbf{N}$ satisfying $\mathbf{N} = \mathbf{M}$ (Horn and Johnson, 1990). Denoting the inverse operator of the matrices $\mathbf{N}$ and $\mathbf{M}$ (Horn and Johnson, 1990). It is important to point out that Eq. (3) remains valid even in

2. PROBLEM STATEMENT

In many system identification problems, especially when dynamical grey-box models are concerned, the system can be modelled with a linear time-invariant state-space form

$$\begin{align*}
\mathbf{p} \chi(t) &= \mathbf{A}(\theta_0) \chi(t) + \mathbf{B}(\theta_0) \mathbf{u}(t) \\
\mathbf{y}(t) &= \mathbf{C}(\theta_0) \chi(t) + \mathbf{D}(\theta_0) \mathbf{u}(t)
\end{align*}$$

where $\mathbf{p}$ stands for the forward shift operator when discrete-time systems are considered or for the differential operator when continuous-time systems are handled. Hereafter, $\chi(t) \in \mathbb{R}^n$ is the state vector, $\mathbf{u}(t) \in \mathbb{R}^m$ is the input vector and $\mathbf{y}(t) \in \mathbb{R}^{n_y}$ is the system output vector. Furthermore, the matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ and $\mathbf{D}$ are characterized by a specific structure (a priori known by the user) and a set of unknown parameters, denoted herein by $\theta_0$. Notice that $\theta_0$ contains only the parameters to be estimated. The matrices of the system can have parameters which are a priori known (using, e.g., physical insights).

Having access to a set of reliable \footnote{By reliable, it is meant that the input is persistently exciting and that the length of the data set is sufficient. The problem of optimal input design is not considered in this contribution.} I/O data, the vector of parameters $\theta$ can be estimated by minimizing the cost function

$$V_N(\theta) = \frac{1}{N} \sum_{k=1}^{N} \|y(k) - \gamma(k, \theta)\|_2^2$$

with respect to $\theta$ where $\gamma(\cdot; \theta)$ denotes the output of the model satisfying

$$\begin{align*}
\mathbf{p} \chi(t, \theta) &= \mathbf{A}(\theta) \chi(t, \theta) + \mathbf{B}(\theta) \mathbf{u}(t) \\
\mathbf{y}(t, \theta) &= \mathbf{C}(\theta) \chi(t, \theta) + \mathbf{D}(\theta) \mathbf{u}(t)
\end{align*}$$

When the structure of the model leads to a non-convex cost function, an iterative gradient-based search is necessary to optimize $V_N$. Assuming that the numerical problems related to the non-uniqueness of the optimal $\theta$ are circumvented by using, e.g., the DDLC parameterized search

(Wills and Ninness, 2008), the main issue consists in finding an initial vector of parameters $\theta_{init}$ in the domain of attraction of the global optimum of $V_N$ in order to ensure a consistent final estimation. As suggested in (Xie and Ljung, 2002) and (Parrilo and Ljung, 2003), this problem can be solved by following a two-step procedure composed of

- firstly the estimation of a reliable fully-parameterized black-box model $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$ from the available I/O data set using a non-iterative approach, i.e., independent of any initial values.
- then the transformation of $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$ into a representation $(\mathbf{A}(\theta), \mathbf{B}(\theta), \mathbf{C}(\theta), \mathbf{D}(\theta))$.

While the first step can be performed efficiently by using, e.g., a subspace-based algorithm (Verhaegen and Verdult, 2007), the second one is much trickier. Assuming that the system is minimal, this transformation consists in determining the similarity transformation $\mathbf{T}$ as well as the vector $\theta$ satisfying

$$\begin{align*}
\mathbf{T} \mathbf{A} &= \mathbf{A}(\theta) \mathbf{T} \\
\mathbf{C} &= \mathbf{C}(\theta) \mathbf{T}
\end{align*}$$

3. FROM AN INITIAL BLACK-BOX MODEL TO THE FINAL MODEL STRUCTURE

In order to solve the set of bilinear equations (1), a natural idea is to use a least-squares formulation leading to the following cost function

$$F(\theta, \mathbf{T}) = \|T \mathbf{A} - \mathbf{A}(\theta) \mathbf{T}\|^2 + \|T \mathbf{B} - \mathbf{B}(\theta)\|^2$$

$$+ \|\mathbf{C} - \mathbf{C}(\theta) \mathbf{T}\|^2 + \|\mathbf{D} - \mathbf{D}(\theta)\|^2$$

where $\|\cdot\|$ is the Frobenius norm, i.e., for a matrix $\mathbf{M}$, $\|\mathbf{M}\| = \sqrt{\text{tr}(\mathbf{M}^\top \mathbf{M})}$. The value of $F$ gives an idea of the error in the bilinear equations (1). This function is clearly non-convex because of the bilinear terms in $\theta$ and $\mathbf{T}$.

In order to perform the estimation of the vector $\theta$ as well as the similarity transformation matrix $\mathbf{T}$, i.e., in order to find a local optimum of the cost function $F$, a quasi-Newton Broyden-Fletcher-Goldfarb-Shanno (BFGS) method can be used (Nocedal and Wright, 2006). To set up this method, the gradients of the cost function $F$ with respect to $\theta$ and $\mathbf{T}$ must be computed. This gradient computation can be performed as follows. Assuming that the structured state-space matrices depends on $\theta$ linearly, the matrices $d\mathbf{A}$, $d\mathbf{B}$, $d\mathbf{C}$ and $d\mathbf{D}$ can be defined as follows

$$\begin{align*}
\text{vec}(\mathbf{A}(\theta)) &= d\mathbf{A} \theta \\
\text{vec}(\mathbf{B}(\theta)) &= d\mathbf{B} \theta \\
\text{vec}(\mathbf{C}(\theta)) &= d\mathbf{C} \theta \\
\text{vec}(\mathbf{D}(\theta)) &= d\mathbf{D} \theta
\end{align*}$$

where vec(-) is the vectorization operator which transforms a matrix into a vector by stacking its columns on top of one another. The vectorization operator is linear and invertible (Horn and Johnson, 1990). Denoting the inverse operator by vec$^{-1}$, the expression of the gradient of $F$ with respect to $\theta$ is

$$\nabla_{\theta} F(\theta, \mathbf{T}) = -2 (\mathbf{T} \mathbf{A} - \mathbf{A}(\theta) \mathbf{T})^\top \cdot \text{vec}^{-1}(d\mathbf{A})$$

$$-2 (\mathbf{T} \mathbf{B} - \mathbf{B}(\theta))^\top \cdot \text{vec}^{-1}(d\mathbf{B}) - 2 (\mathbf{T} \mathbf{D} - \mathbf{D}(\theta))^\top \cdot \text{vec}^{-1}(d\mathbf{D})$$

$$-2 (\mathbf{C} - \mathbf{C}(\theta) \mathbf{T})^\top \cdot \text{vec}^{-1}(d\mathbf{C})$$

where $\mathbf{N} \cdot \mathbf{M}$ stands for the Frobenius scalar product of matrices $\mathbf{N}$ and $\mathbf{M}$ (Horn and Johnson, 1990). It is important to point out that Eq. (3) remains valid even in
the non-linear case by substituting the matrices \(dA, dB, dC\) and \(dD\) for the derivative of \(A, B, C\) and \(D\) with respect to \(\theta\). The gradient of the cost function \(F\) with respect to \(T\) can be derived quite easily by using the well-known property of the Kronecker product \(\otimes\) (Horn and Johnson, 1990) \((M^\top \otimes N)\) \(\text{vec}(X) = \text{vec}(NXM)\). Thus, the gradient reads
\[
\nabla_T F(\theta, T) = -2(C - C(\theta)) \cdot \text{vec}^{-1}(I_n \otimes C(\theta)) \\
- 2(TB - B(\theta)) \cdot \text{vec}^{-1}(B^\top \otimes I_n) \\
- 2(TA - A(\theta)T) \cdot \text{vec}^{-1}(A^\top \otimes I_n - I_n \otimes A(\theta)).
\]

To summarize, the minimization of \(F\) is used in order to get an approximated solution to the set of bilinear equations (1). Rather than using a global optimization method, it is aimed at focusing on local methods in order to get a good and reliable local optimum. Thanks to this local formulation, it is now possible to handle structured state-space matrices with non-linear dependence on the parameters.

### 3.1 Constraint on the condition number

As claimed by Parrilo and Ljung (2003), the use of the cost function \(F\) can lead to unusable results. Indeed, its minimization can supply a singular similarity transformation matrix \(T\) at the global minimum. In order to get around this difficulty, a constraint on the condition number of \(T\) can be added into the previous optimization problem. To reach this goal, let us consider a scalar \(\alpha > 0\). Then, the constrained optimization problem to be solved becomes

\[
\min_{\theta, T} F(\theta, T) \text{ subject to } \text{cond}(T) \leq \alpha
\]  

(4)

where \(\text{cond}(T)\) is the condition number of \(T\). Now, the constraint \(\text{cond}(T) \leq \alpha\) can be written as

\[
\sigma_{n_x}(T) \geq \frac{\sigma_1(T)}{\alpha}
\]

where \(\sigma_1(T)\) and \(\sigma_{n_x}(T)\) are the highest and the lowest singular values of \(T\) respectively. In order to simplify the problem, the term \(\sigma_1(T)\) can be removed from the constraint. Then, the optimization problem becomes, for a value \(\beta > 0\),

\[
\min_{\theta, T} F(\theta, T) \text{ subject to } \sigma_{n_x}(T) \geq \beta
\]  

(5)

Rather than solving the constrained optimization problem (5), a penalization parameter \(\mu > 0\) is introduced. Thus, without any constraint, the penalized cost function is

\[
F_2(\theta, T) = F(\theta, T) + \mu \sigma_{n_x}(T)^{-2},
\]

or, similarly,

\[
F_2(\theta, T) = F(\theta, T) + \mu \lambda_1 \left( (T^{-1})^\top T^{-1} \right)
\]

where \(\lambda_1(\cdot)\) denotes the maximum eigenvalue function. By minimizing the maximum eigenvalue of \((T^{-1})^\top T^{-1}\), the smallest singular value of \(T\) is maximized. This condition keeps \(T\) away from singularity. Moreover, by tuning the penalization parameter, the condition number of \(T\) can be adjusted easily. In practice, a small value for \(\mu\) is chosen because our objective is to minimize the function \(F\) rather than the condition number of \(T\).

Although \(F_2\) corresponds to an unconstrained optimization problem, its non-smoothness makes its optimization difficult. This characteristic comes from the maximum eigenvalue function \(\lambda_1(\cdot)\) which is convex but non-smooth when the maximum eigenvalue is multiple. To solve this optimization problem, an algorithm able to handle non-smooth functions must be used. In this paper, the developed algorithm resorts to the combination of

- firstly a non-smooth BFGS method as described in (Lewis and Overton, 2008),
- then the spectral bundle algorithm of (Apkarian et al., 2008; Noll et al., 2008) (used herein to improve the numerical results).

For both methods, a subgradient of the objective function must be provided. The function \(F_2\) is Clarke-regular (Clarke, 1990) when the similarity transformation matrix is non-singular. Hence, subgradients can be computed by applying the chain rule. See (Lewis, 2003; Apkarian et al., 2008) for more details about the subgradient computation of the maximum eigenvalue function.

The BFGS method is a quasi-Newton method, which is well-known to be very efficient for solving smooth unconstrained optimization problems. In the non-smooth case, it has been observed that specific quasi-Newton methods are also very good in practice, leading rapidly to a good approximation of the optimum (Lemaréchal, 1982). The key point is that a locally Lipschitz function is differentiable almost everywhere. Thus, in (Lewis and Overton, 2008), the authors suggest using an inexact line search in order to keep the iterates away from the non-differentiable points. The “strong” Wolfe condition, which is commonly used in the smooth case, is replaced with the “weak” Wolfe condition in order to ensure the positive definiteness of the inverse Hessian update. The convergence of this inexact line search is studied under general condition in (Lewis and Overton, 2008).

The non-smooth BFGS method is very useful in practice but a proof of convergence is still missing in the framework of non-smooth and non-convex optimization. When the non-smooth BFGS method stops at a non-smooth point, this point is used to initialize the spectral bundle method so that the numerical results are improved. The spectral bundle algorithm is based on the use of a local convex model \(y \mapsto \phi(y, x_k)\) of the objective at the current iterate \(x_k\). This model must satisfy a set of axioms (Noll et al., 2008). When dealing with an eigenvalue optimization, this model can be easily derived by exploiting the intrinsic structure of the maximum eigenvalue function (Helmborg and Kiwiel, 2002; Apkarian et al., 2008, 2009). In this case, the model can be seen as a “non-smooth” Taylor expansion of the objective. In practice, the model \(\phi(\cdot, x_k)\) is not numerically tractable. Hence, a polyhedral “working model” is built iteratively from cutting planes of \(\phi(\cdot, x_k)\). A new candidate point \(y^*\) for the optimization is obtained from this working model. Then,

- if this point leads to a sufficient decrease of the objective, then we set \(x_{k+1} = y^*\) a new local model \(\phi(\cdot, x_{k+1})\) is built,
- else, the working model is improved by updating its bundle of cutting planes.

It is well-known that such a cutting planes algorithm must be stabilized (Bonnans et al., 2006) for instance
by using a proximity control parameter. In the spectral bundle method, this proximity parameter helps us to deal with the non-convexity of the cost function. The proximity parameter is in fact chosen with respect to the distance between the objective and the models.

3.2 Comments and discussion

While the developed approach handles (a modified version of) the objective function used in (Xie and Ljung, 2002) and (Parrilo and Ljung, 2003), several differences and advantages can be highlighted. For instance, in (Xie and Ljung, 2002), because the objective function is biconvex, the authors use alternatively two convex optimizations to find a solution to the minimization of Eq. (2). However, as shown in (Gorski et al., 2007), convergence, even local, cannot be ensured with such an approach. On the contrary, the BFGS method used in our local approach is known to be globally convergent under mild assumption. The use of a local approach (rather than a global one as in (Parrilo and Ljung, 2003)) is also an interesting feature. Indeed, by using such a local approach, it is easier

- to handle a non-linear parametrization of the structured model without changing the optimization algorithm.
- to introduce constraints on the parameters by considering in this case an algorithm dedicated to constrained optimization instead of the BFGS method.

Let us now focus on the reasons why the term $\sigma_1(T)$ is removed from the constrained optimization (4). Let us assume that a local optimum $(T, \theta)$ can be obtained such that $F_2(\theta, T) < \eta$. Then, we have

$$
\|TA - A(\theta)T\|^2 < \eta, \|TB - B(\theta)\|^2 < \eta,
$$

$$
\|C - C(\theta)T\|^2 < \eta.
$$

These three inequalities can be rewritten as follows

$$
\|M_\theta vec(T) - \beta_\theta\|^2 < 3\eta,
$$

where

$$
M_\theta = \begin{bmatrix} A^T \otimes I_n_{\alpha} - I_n_{\alpha} \otimes A(\theta) \\ B^T \otimes I_n_{\alpha} \\ I_n_{\alpha} \otimes C(\theta) \end{bmatrix}, \beta_\theta = \begin{bmatrix} vec(B(\theta)) \\ 0 \\ vec(C) \end{bmatrix}.
$$

Thus, a vector $\nu$ exists such that $M_\theta vec(T) - \beta_\theta = \nu$ and $\|\nu\| < \sqrt{3n}$. Now, let us assume that the matrix $M_\theta$ is of full rank, i.e., rank($M_\theta$) = $n_\alpha^2$. Then,

$$
T = vec^{-1}(M_\theta^\dagger(\beta_\theta + \nu))
$$

where $M_\theta^\dagger$ denotes the Moore-Penrose pseudo-inverse of matrix $M_\theta$ (Horn and Johnson, 1990). Finally, by computing the matrix 2-norm $\|\cdot\|_2$ of $T$, we get

$$
\sigma_1(T) \leq \|M_\theta^\dagger(\beta_\theta + \nu)\|_2 \leq \|M_\theta^\dagger\|_2 (\|\beta_\theta\| + \sqrt{3n} \eta)
$$

Thus, the greatest singular value of matrix $T$ is bounded.

3.3 Relation with transfer functions

Looking closer at the previous results, it is interesting to see that the cost function $F$ can have an interesting physical meaning. This feature can be outlined and highlighted quite easily when continuous-time systems are addressed. Let us consider $\omega \in [0, \infty]$ and

$$
E(\omega) = C(\omega I - A)^{-1}B - C(\theta)(\omega I - A(\theta))^{-1}B(\theta).
$$

The same experimental condition as in (Xie and Ljung, 2002) are considered, i.e., $m = 1.2$, $M = 5.5$ and $l = 2.3$.

These three inequalities can be rewritten as follows

$$
F(\theta) = \frac{1}{\epsilon} E(\omega) \leq \frac{1}{\epsilon} F(\theta, T).
$$

A straightforward calculation shows that

$$
E(\omega) = (C - C(\theta)T)(\omega I - A)^{-1}B + C(\theta)(\omega I - A(\theta))^{-1}(TB - B(\theta)) + C(\theta)(T(\omega I - A) - (\omega I - A(\theta))T)B.
$$

Thus,

$$
E(\omega) = (C - C(\theta)T)(\omega I - A)^{-1}B + C(\theta)(\omega I - A(\theta))^{-1}(TB - B(\theta)) + C(\theta)(\omega I - A(\theta))^{-1}(TA - A(\theta))T(\omega I - A)^{-1}B.
$$

Finally, with $E_\theta(\omega) = \|\omega I - A(\theta)\|^{-1}$, we get the following inequality

$$
\|E(\omega)\| \leq \sqrt{\epsilon} [E_B(\omega) + E_C(\omega) + E_B(\omega)E_C(\omega)].
$$

This inequality shows that the distance between the transfer functions of both system representations ($A, B, C, D$) and ($A(\theta), B(\theta), C(\theta), D(\theta)$) decreases with $\epsilon$. Moreover, the distance between both transfer functions is smaller for high values of $\omega$.

4. SIMULATION EXAMPLES

The numerical example considered in this study has already been used in (Xie and Ljung, 2002) and (Parrilo and Ljung, 2003). The system is composed of a cart with an inverted pendulum (see (Xie and Ljung, 2002) for a picture). The underlying control problem consists in keeping the pendulum upright via the signal $u$. The physical modelling is available, e.g., in (Ogata, 2009), and writes

$$
A(\theta) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}, C(\theta) = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}
$$

$$
B(\theta) = \begin{bmatrix} \theta_3 \\ \theta_4 \\ 0 \end{bmatrix}.
$$

with $x(t) = [y(t) \ y(t) \ \alpha(t) \ \dot{\alpha}(t)]^T$. Herein, $m$ and $M$ are the mass of the pendulum and the cart respectively, $l$ is the length of the pendulum and $g$ the gravitation constant. The same experimental condition as in (Xie and Ljung, 2002) are considered, i.e., $m = 1.2$, $M = 5.5$ and $l = 2.3$. 
To start with, the cost function $F_2$ is initialized with the following black-box model (introduced in (Xie and Ljung, 2002))

$$
\mathbf{A} = \begin{bmatrix}
-18.4268 & 30.1636 & -15.0469 & 3.3101 \\
-3.3101 & -5.0311 & 9.9023 & -1.6512 \\
1.6512 & -9.9023 & 5.0311 & 3.3101 \\
-3.3101 & 15.0469 & -30.1636 & 18.4268
\end{bmatrix}
\quad \mathbf{B} = \begin{bmatrix}
-2.5304 \\
0.8296 \\
-0.8214 \\
2.4391
\end{bmatrix}
\quad \mathbf{C} = 10^{-4} \times \begin{bmatrix}
8.118 & -8.301 & -8.800 & 8.262
\end{bmatrix}
$$

The user-defined parameter $\mu$ is chosen equal to $1e-18$ because of the very bad condition number of $\mathbf{T}$ near the initial point. This bad condition number justifies the use of $F_2$ instead of $F$. The tolerance value for the optimization algorithm stopping test is set to $1e-12$. We obtain the following solution from the non-smooth optimization algorithms

$$
\hat{\theta} = \begin{bmatrix}
-1.287422 \\
5.184889 \\
0.1639943 \\
-0.06322288
\end{bmatrix}
$$

with a final condition number equal to $3.6e3$. Using the same procedure as in (Xie and Ljung, 2002), we get

$$
\hat{M} = 5.4522970966, \quad \hat{l} = 2.090192372, \quad \hat{m} = 1.291004849.
$$

These values are quite similar as the ones found$^2$ by Xie and Ljung. However, it is important to notice that, contrary to the results shown in (Parrilo and Ljung, 2003), the condition number of $\mathbf{T}$ is quite good.

To quantify the estimation quality, let us introduce the error function

$$
e(\theta, \mathbf{T}) = \|\mathbf{T} \mathbf{A} - \mathbf{A}(\theta) \mathbf{T}\| + \|\mathbf{TB} - \mathbf{B} (\theta)\|
\quad + \|\mathbf{C} - \mathbf{C}(\theta) \mathbf{T}\| + \|\mathbf{D} - \mathbf{D}(\theta)\|
$$

i.e., the distance between the estimated structured model and the initial black-box model. In (Xie and Ljung, 2002), the value of this error at the optimum is equal to $4.72e-6$. With our approach, it is $1.14e-5$. Although the error with our algorithm is ten times greater than with the solution obtained by Xie and Ljung, the frequency behavior of both models are quite similar, as shown in Fig. 1.

Fig. 1. Function $\omega \mapsto \sigma_1(\mathbf{E}(\omega))$ for the solution given in (Xie and Ljung, 2002) (dotted line) and for $\hat{\theta}$ (plain line), i.e. the model estimated in this paper. The value of the upper bound for $\hat{\theta}$ (dashed line) is also displayed.

The second experiment consists in initializing the vector of parameters with zeros and $\mathbf{T}$ with the identity matrix.

With a tolerance value equal to $1e-12$, the BFGS method leads to the following local minimum for the objective function $F$

$$
\hat{\theta}^* = \begin{bmatrix}
0.42634 \\
0.80347 \\
0.16368 \\
0.01675
\end{bmatrix}
$$

The value of the cost function at this point is $F^* = 2.36e - 11$ and the condition number of matrix $\mathbf{T}$ is equal to $1.181e5$. Figure 2 displays the function $\sigma_1(\mathbf{E}(\omega))$ with respect to $\omega$ for the estimate obtained in (Xie and Ljung, 2002) (dotted line) and for $\hat{\theta}^*$ (plain line), the third curve being the bound obtained in Subsection 3.3 (dashed line). These curves show that our estimated model is able to depict the system behavior on a large range of frequencies. Unfortunately, while the frequency plots of both solutions are similar and while the error values are of the same magnitude, the computed parameters are not convenient. For instance, as a ratio of masses and positive constant values times $-1$, the first parameter $\theta_1$ must be a negative number, which is not the case for the local minimum obtained with our algorithm. When no reliable initial models are available, this problem can be solved by introducing a priori information into the optimized cost function, i.e., via the introduction of lower and upper bounds on the parameters. As shown beforehand, this type of constrained cost function can be minimized by using an algorithm dedicated to constrained optimization.

Finally, in order to test the influence of the constraint on the condition number of $\mathbf{T}$, the same numerical tolerance and the same initial point are used with the non-smooth function $F_2$, $\mu$ is chosen equal to $1e-8$ and both algorithms described in Section 3.1 are used. We get the following estimated parameters

$$
\hat{\theta} = \begin{bmatrix}
0.01393 \\
-0.01860 \\
0.16369 \\
0.07840
\end{bmatrix}
$$

Furthermore, for this particular point, $\hat{F} = 2.4e-4$, $e(\hat{\theta}) = 0.01$ and the condition number of the similarity transformation matrix is $1.2e2$. This solution has clearly a greater error than the one obtained previously but, as expected, the condition number is much smaller. Figure 3 displays the function $\sigma_1(\mathbf{E}(\omega))$ with $\hat{\theta}$ and the solution of Xie and Ljung (2002). While the error $e(\hat{\theta})$ is quite large, both frequency responses are very close (except for a pole next to $\omega = 0.1 \text{ rad/s}$). Notice also the gap between the

$^2$ As a reminder, $\hat{M} = 5.56635, \hat{l} = 2.0521$ and $\hat{m} = 1.0730$ in (Xie and Ljung, 2002).
curves and the computed bound. These observations show that the control of the condition number of $T$ can improve and constraint the input-output curve fitting capability of the identified model. As far as the non-physical pole at $\omega = 0.1 \text{ rad/s}$ is concerned, it is interesting to notice that such a mis-estimation also arises when $\mu = 0$. Indeed, this problem is only related to the non-convexity of the cost function. In order to solve the problem of possible non-physical parameter estimates, a modification of the cost functions is under study where constraints are added in order to take into account physical knowledge such as, e.g., $M > 0$ or $m > 0$.

Fig. 3. Function $\omega \mapsto \sigma_1(E(\omega))$ for the solution given in (Xie and Ljung, 2002) (dotted line) and for $\theta$ (plain line). The value of the upper bound for $\theta$ (dashed line) is also displayed.

5. CONCLUSION

In this paper, the problem of initializing a gradient-based identification algorithm is considered. Assuming that an accurate initial fully-parameterized state-space form is available, an algorithm is provided to transform this initial (black-box) model into a structured state-space representation corresponding to the grey-box model to be identified. This technique is valid for both discrete-time and continuous-time system identification. The theoretical developments as well as the simulation examples have shown that this approach can

- deal with MIMO systems with a big amount of parameters without decreasing the condition number of the similarity transformation,
- result in a structured model able to reproduce the system behavior even if the available prior information is poor.

REFERENCES


