European Journal of Control (2000)6:54–62

European Journal of Control

Analysis of Closed-Loop Identification with a Tailor-Made Parameterization*

Edwin T. van Donkelaar† and Paul M.J. Van den Hof

Mechanical Engineering Systems and Control Group, Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands

An analysis is made of a closed-loop identification scheme in which the parameters of the plant model are identified on the basis of input and output measurements of the closed-loop system. The closed-loop transfer function is parametrized in terms of the parameters of the open-loop plant model, and utilizing knowledge of the implemented feedback controller. This is denoted as a tailor-made parametrization as it is tailored to the specific feedback structure at hand. Consistency of the plant estimate is shown to hold under additional conditions, resulting from the requirement that the parameter set should be a connected set. Sufficient conditions for this requirement are formulated, requiring the controller order not to exceed the order of the plant model.

Keywords: Closed-loop identification; Tailor-made parametrization; Prediction error methods

1. Introduction

Many industrial processes operate under feedback control. Due to unstable behaviour of the plant, required safety and/or efficiency of operation, experimental data can only be obtained under so-called closed-loop conditions. Identification methods for dealing with closed-loop experimental data have been developed in the seventies and eighties, see [10] for an overview. These 'classical' methods are typically directed towards solving the consistency problem, considering the situation that plant and disturbance model can be modelled exactly (system is in the model set).

Initiated by an emerging interest in the identification of models that are particularly suitable for model-based (robust) control design, renewed attention has been given lately to the problem of closed-loop identification. There are a number of arguments to prefer closed-loop experiments over open-loop ones in particular situations where one is interested in model-based control design. These arguments comprise aspects of bias and variance, input shaping, and the fact that a controller can linearize the (possibly nonlinear) plant behaviour in a relevant working point, thus enabling approximate linear modelling. Accounts of this area are given in [4,6,12].

Unlike the classical situation, particular attention is now also given to properties of identified approximate models, handling the – more realistic – situation that plant and noise dynamics are not exactly present in the model set considered. In view of this, two questions have been addressed in particular:

- Can a plant model be identified consistently in the situation that the noise characteristics on the data (noise model) are misspecified? and
- Can identification of a reduced order plant model lead to an (asymptotically) identified model that approximates the underlying plant in a well-defined (and known) sense?

The classical 'direct' method of closed-loop identification is not able to deal with these questions; it

*A previous version of this paper was presented at the ECC 97, Brussels, 1–4 July 1997.

†The research of Edwin van Donkelaar, supported by the Dutch Technology Foundation (STW) under contract DWT55.3618.

Correspondence and offprints requests to: Paul Van den Hof, now with the Department of Applied Physics, Delft University of Technology, Lorentzweg 1, 2628 CJ Delft, The Netherlands. Tel.: +31-15-2784509; Fax: +31-15-2784263; Email: p.m.j.vandenhof@tn.tudelft.nl.

Received 20 October 1997; Accepted in revised form 28 June 1999. Recommended by E. Egardt and D. Clarke.

requires both plant and noise model to be modelled exactly (system in the model set) to achieve consistency of the plant model, and the resulting bias expression for approximate models is essentially dependent on the (unknown) noise characteristics.

Recently presented algorithms that comply with the above questions are directly related to the classical prediction error methods known as 'indirect identification', and 'joint input/output method' [8,10]. The two-stage method [11] and the method of coprime factor identification [13] are particular forms of the joint i/o method; identification using the dual-Youla parametrization [5] is a generalized form of the indirect method. For a more extensive description and overview of the several methods and their properties the reader is referred to [9,15]. The several methods have their particular advantages and disadvantages. Whereas the two-stage method requires a high order and accurate estimate of the sensitivity function of the closed-loop system, the coprime factor and dual-Youla method have problems in handling model classes with prespecified model order. As a result the latter two methods often lead to high order models.

An alternative method that was first suggested as an exercise in [8] seems to avoid these problems.

The basic idea is that the closed-loop transfer function from excitation signal r to output signal y (see Fig. 1) is identified using an output predictor

$$\hat{y}(t|t-1;\theta) = \frac{G(q,\theta)}{1+C(q)G(q,\theta)} r(t) \tag{1}$$

using the parameters corresponding to the (open-loop) plant model

$$G(q,\theta) = \frac{b_1 q^{-1} + \cdots + b_{n_b} q^{-n_b}}{1 + a_1 q^{-1} + \cdots + a_{n_a} q^{-n_a}}$$

with $\theta = [a_1 \cdots a_{n_a} b_1 \cdots b_{n_b}]$, and q^{-1} indicating the backwards shift operator.

Using the open-loop plant parameters, and knowledge of the controller C, a prediction error criterion is used to estimate the plant parameters; this requires a nonlinear optimization procedure.

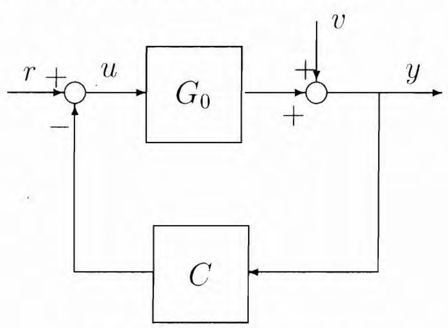

Fig. 1. Closed-loop configuration.

The parametrization is referred to as a tailor-made parametrization, as it is specifically directed towards (tailored to) the closed-loop configuration at hand, including the use of knowledge of the controller. By specifying the polynomial orders n_a and n_b, the model order can be fixed on beforehand. The method has been used in a recursive version in [7].

In this paper, an analysis will be made of the consistency properties of this method. As the approach simply falls within the framework of prediction error identification methods [8], the available analytical results for these methods can be employed. It will appear that – in comparison with standard open-loop identification methods – particular attention has to be given to the question whether the considered parameter set induced by the parametrization (1) is pathwise connected. This connectedness will be studied in particular, leading to the formulation of sufficient conditions for connectedness and consequently also for consistency of the identified model.

In Section 2 the problem will be specified in more detail. Uniform stability of the resulting model set and connectedness of the parameter set is the topic of Section 3. In Section 4 gradient expressions are presented for the tailor-made optimization. The procedure is illustrated with some examples in Section 5, and finally the procedure is discussed in the context of other closed-loop identification methods in Section 6.

2. Problem Setting

A closed-loop system configuration is considered as sketched in Fig. 1. G_0 is a linear time-invariant discrete-time single input single output plant, C is a known controller, r is an external excitation signal, and u and y are respectively input and output signal of the plant.

It is assumed that measurements of $r(t)$ and $y(t)$ are available. The output noise $v(t)$ is assumed to be generated by filtering a white noise signal $e(t)$ with variance σ^2 using a stable monic filter H_0. The output noise is assumed to be uncorrelated with the excitation signal r. The loop transfer CG_0 is assumed to be strictly proper. The closed-loop system is characterized by

$$y(t) = \underbrace{\frac{G_0(q)}{1+C(q)G_0(q)}}_{R_0(q)} r(t) + \underbrace{\frac{1}{1+C(q)G_0(q)}}_{W_0(q)} H_0(q)e(t), \tag{2}$$

where R_0 denotes the closed-loop transfer function and W_0 the closed-loop noise filter. The sensitivity function is denoted by $S_0 = (1 + CG_0)^{-1}$.

A parametrized plant model $G(q,\theta)$ as in (1) together with knowledge of the controller C can now be used to parametrize a (one-step-ahead) output predictor for the closed-loop system (2), given by [8]

$$\hat{y}(t|t-1;\theta,\eta) = W(q,\eta)^{-1}R(q,\theta)r(t) + [1-W(q,\eta)^{-1}]y(t), \quad (3)$$

where $W(q,\eta)$ is a parametrized (closed-loop) noise model. For a fixed noise model $W(q,\eta)=1$, using a so-called output error structure, this predictor simplifies to

$$\hat{y}(t|t-1;\theta) = \underbrace{\frac{G(q,\theta)}{1+C(q)G(q,\theta)}}_{R(q,\theta)} r(t), \quad (4)$$

and the corresponding closed-loop model set is defined by

$$\mathcal{P} := \left\{ R(q,\theta) = \frac{G(q,\theta)}{1+C(q)G(q,\theta)}, \theta \in \Theta \right\}, \quad (5)$$

where $\Theta \subset \mathbb{R}^{n_a+n_b}$ is the parameter set over which θ varies. The parameter estimate is found by least squares minimization of the prediction error $\varepsilon(t,\theta) = y(t) - R(q,\theta)r(t)$, by solving $\hat{\theta}_N = \arg\min_{\theta\in\Theta} V_N(\theta)$, in which the criterion function is given by $V_N(\theta) = (1/N)\sum_{t=1}^{N}\varepsilon^2(t,\theta)$. The resulting estimate of the plant model will be denoted by $\hat{G}(q) = G(q,\hat{\theta}_N)$. For this identification method the following consistency result holds [8].

Proposition 2.1. Let $\mathcal{P}$ be a uniformly stable model set and let the data generating closed-loop system be stable. If r is a bounded signal and e has bounded fourth moment, then $\theta_N \to \theta^*$ w.p. 1 for $N\to\infty$ with

$$\theta^* = \arg\min_{\theta\in\Theta} \frac{1}{2\pi}\int_{-\pi}^{\pi} |R_0(e^{i\omega}) - R(e^{i\omega},\theta)|^2 \Phi_r(\omega)d\omega. \quad (6)$$

Whenever there exists a $\theta\in\Theta$ such that $G(q,\theta)=G_0(q)$ this choice will be a minimizing argument of the integral expression above, and it is unique provided that $r(t)$ is persistently exciting of sufficiently high order. □

This proposition states that a consistent estimate is obtained with this parametrization under the condition that the model set $\mathcal{P}$ is uniformly stable. This condition is not trivially satisfied in case the tailor-made parametrization given in (5) is used. Therefore, in the next section the conditions under which the model set (5) is guaranteed to be uniformly stable will be investigated.

3. Uniform Stability of the Model Set

Uniform stability of the model set is defined as follows.

Definition 3.1[8]. A parametrized model set $\mathcal{P}$ is uniformly stable if

- Θ is a connected open subset of $\mathbb{R}^{(n_a+n_b)}$,
- $\mu : \Theta \to \mathcal{P}$ is a differentiable mapping, and
- the family of transfer functions $\{R(q,\theta), (\partial/\partial\theta) R(q,\theta)\}$ is uniformly stable. □

In this section it will be made clear that in case a tailor-made parametrization is used, the parameter set Θ is possibly not connected due to the specific parametrization of the closed-loop transfer function $R(q,\theta)$. Also a sufficient condition will be derived for guaranteed connectedness of the parameter set.

Let the strictly proper[1] plant model be parametrized as

$$G(q,\theta) = \frac{B(q,\theta)}{A(q,\theta)} = \frac{b_1q^{-1}+\cdots+b_{n_b}q^{-n_b}}{1+a_1q^{-1}+\cdots+a_{n_a}q^{-n_a}}, \quad (7)$$

where $\theta = [a_1\cdots a_{n_a}b_1\cdots b_{n_b}]^{\mathrm{T}}$. The controller is given by

$$C = \frac{N_c(q)}{D_c(q)} = \frac{n_0+n_1q^{-1}+\cdots+n_{n_N}q^{-n_N}}{1+d_1q^{-1}+\cdots+d_{n_D}q^{-n_D}}, \quad (8)$$

where $N_c(q), D_c(q)$ are coprime polynomials. With this notation the parametrization of the output predictor is given by

$$\hat{y}(t|t-1;\theta) = \frac{D_c(q)B(q,\theta)}{D_c(q)A(q,\theta)+N_c(q)B(q,\theta)} r(t). \quad (9)$$

All closed-loop models $R(q,\theta)$ are stable if the absolute value of the roots of the denominator $D_c(z)A(z,\theta)+N_c(z)B(z,\theta)$ is strictly less than one. Hence, the parameter set corresponding to closed-loop stable models is given by

$$\Theta := \{\theta\in\mathbb{R}^{n_a+n_b} \mid \mathrm{sol}_z\{D_c(z)A(z,\theta) + N_c(z)B(z,\theta)=0\}| < 1\}. \quad (10)$$

The corresponding set of plant models is denoted by

$$\mathcal{G} := \{G(q,\theta),\ \theta\in\Theta\}. \quad (11)$$

[1] For simplicity of notation only the case of a strictly proper plant and a proper controller is regarded. However, the case of a strictly proper controller and a proper plant can be described similarly.

It can be verified that the parameter set for which the polynomial $A(q, \theta)$ is stable, is pathwise connected. As a result, connectedness of the parameter set when using a (standard) numerator–denominator parametrization of the plant in an open-loop setting, will not be a problem. However, in case the tailor-made parametrization (5) is used, with Θ given by (10), Θ may not be pathwise connected as the following simple example shows.

Example 3.2. Given the 7th order controller defined by the continuous-time transfer function

$$\begin{aligned} C(s) =&(0.499s^5 + 0.715s^4 + 2.577s^3 + 3.397s^2 \\ &+ 2.155s + 2.620) \\ &/(s^7 + 1.717s^6 + 5.100s^5 + 8.410s^4 \\ &+ 4.198s^3 + 6.631s^2). \end{aligned}$$

The plant that is to be identified is parametrized by a simple constant $G = \theta$. The parameter space $\Theta \subset \mathbb{R}$ for which the closed-loop system is stable can be simply derived from a root locus plot and is approximately given by

$$\Theta = \{\theta \mid \theta \in (0, 1.27) \cup (2.64, 4.69) \cup (9.98, \infty)\}.$$

This set is a disconnected subset of $\mathbb{R}$. Therefore the corresponding model set $\mathcal{P}$ is not uniformly stable. □

The fact that a parameter set is not connected has not only consequences for the formal proof of consistency as was mentioned before, but also for the non-linear optimization that has to be performed to obtain an estimate. If, for example, a gradient search method is used and an initial estimate is selected in a region of the parameter set that is disconnected from the region where the optimal parameter vector is located, it will be extremely hard if not impossible to reach the optimum.

The denominator of the closed-loop transfer function can be written as a function of the open-loop parameter θ as

$$\begin{aligned} &D_c A(q, \theta) + N_c B(q, \theta) \\ &\quad = 1 + [q^{-1} \ q^{-2} \ \ldots \ q^{-n}]\theta_{cl}, \end{aligned} \tag{12}$$

where the order of the closed-loop polynomial of (12) is given by $n = \max(n_a + n_D, n_b + n_N)$, the closed-loop parameter vector is given by $\theta_{cl} := S\theta + \rho$, the vector $\rho = [d_1 \cdots d_{n_D} \ 0 \cdots 0]^{\mathrm{T}} \in \mathbb{R}^n$ and the matrix $S \in \mathbb{R}^{n \times (n_a + n_b)}$ is constructed from submatrices $P_D \in \mathbb{R}^{(n_a + n_D) \times n_a}$ and $P_N \in \mathbb{R}^{(n_b + n_N) \times n_b}$ defined by

$$P_D = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ d_1 & 1 & & \\ d_2 & d_1 & \ddots & \vdots \\ \vdots & d_2 & \ddots & 1 \\ d_{n_D} & & \ddots & d_1 \\ 0 & \ddots & & d_2 \\ \vdots & \ddots & & \vdots \\ 0 & \cdots & 0 & d_{n_D} \end{bmatrix},$$
$$P_N = \begin{bmatrix} n_0 & 0 & \cdots & 0 \\ n_1 & n_0 & & \\ n_2 & n_1 & \ddots & \vdots \\ \vdots & n_2 & \ddots & n_0 \\ n_{n_N} & & \ddots & n_1 \\ 0 & \ddots & & n_2 \\ \vdots & \ddots & & \vdots \\ 0 & \cdots & 0 & n_{n_N} \end{bmatrix} \tag{13}$$

such that

$$S = \begin{bmatrix} P_D & P_N \\ 0_{(n - n_a - n_D) \times n_a} & 0_{(n - n_b - n_N) \times n_b} \end{bmatrix}. \tag{14}$$

Note that since $n = \max(n_a + n_D, n_b + n_N)$ either P_D or P_N is padded with zeros to satisfy the correct row dimension of S.

The closed-loop parameter can vary over a parameter set

$$\Theta_{cl} := \{\theta_{cl} = S\theta + \rho \mid \theta \in \Theta\},$$

where the allowable closed loop parameters are restricted by the affine relation given above. Now, define a parameter set for stable polynomials of order n as follows

$$\Theta_n := \{\theta_n \in \mathbb{R}^n \mid \mathrm{sol}_z\{1 + [z^{-1} \ \cdots \ z^{-n}]\theta_n = 0\}| < 1\}.$$

The parameter set for stable polynomials is connected.[2] From this it can be concluded that the

[2] A justification of this claim is added in the Appendix.

parameter set Θ_n is also connected. In the following theorem a sufficient condition for connectedness of the parameter space Θ is given using the connected set Θ_n as a starting point.

Lemma 3.3. Full row rank of the matrix S (14) with P_D, P_N given in (13), is a sufficient condition for pathwise connectedness of the parameter set Θ given in (10). □

Proof The closed-loop parameter θ_n can vary over the connected set Θ_n. Now define the set

$$\bar{\Theta}_{\rm cl} = \{\bar{\theta}_{\rm cl} | \bar{\theta}_{\rm cl} = \theta_n - \rho, \theta_n \in \Theta_n\}.$$

This set is a shifted version of Θ_n and is therefore also pathwise connected. An open-loop parameter vector $\theta \in \Theta$ and a parameter vector $\bar{\theta}_{\rm cl} \in \bar{\Theta}_{\rm cl}$ are related via $\bar{\theta}_{\rm cl} = S\theta$, $S \in \mathbb{R}^{n \times (n_a+n_b)}$. If S has full row rank it defines a surjective map, hence image$(S) = \bar{\Theta}_{\rm cl}$. In the connected set $\bar{\Theta}_{\rm cl}$ a continuous path can be constructed between two parameter vectors. This path can be mapped into a continuous path in Θ using the inverse mapping of S. Therefore, Θ is also pathwise connected. □

This result implies that the parameter set for which the parametrized transfer function (5) is stable, is only a connected set in specific cases. Therefore, it is not guaranteed that the model set defined in (5) is uniformly stable following the definition of uniform stability in Definition 3.1. The following lemma gives an easy test for guaranteed uniform stability of the model set with a tailor-made parametrization.

Proposition 3.4. Let a model of the plant G_0 be parametrized as in (7) and let the controller be given by (8). A sufficient condition for connectedness of the parameter set Θ for a tailor-made parametrization given in (5), is given by

$$\{n_b \geq n_D \text{ and } n_a \geq n_N\}.$$

Proof. From lemma 3.3 it follows that full row rank of S is a sufficient condition for connectedness. For S having full row rank it is necessary that $n \leq n_a + n_b$. This implies that $n_D \leq n_b$ and $n_N \leq n_a$. By reordering the columns of S a 2×2 upper triangular block matrix can be constructed given by

$$S = \begin{bmatrix} S_1 & S_{12} \\ 0 & S_2 \end{bmatrix}$$

where $S_1 \in \mathbb{R}^{(n_N+n_D)\times(n_N+n_D)}$ is given by

$$S_1 = \left[\begin{array}{cccc|cccc} 1 & 0 & \cdots & 0 & n_0 & 0 & \cdots & 0 \\ d_1 & 1 & & & n_1 & n_0 & & \\ d_2 & d_1 & \ddots & \vdots & n_2 & n_1 & \ddots & \vdots \\ \vdots & d_2 & \ddots & 1 & \vdots & n_2 & \ddots & n_0 \\ d_{n_D} & & \ddots & d_1 & n_{n_N} & & \ddots & n_1 \\ 0 & d_{n_D} & & d_2 & 0 & n_{n_N} & & n_2 \\ \vdots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & d_{n_D} & 0 & \cdots & 0 & n_{n_N} \end{array}\right].$$

For the structure of S_2 we first consider the situation that $n_a + n_D = n_b + n_N$. Then $S_2 \in \mathbb{R}^{(n_a-n_N)\times(n_a-n_N+n_b-n_D)}$ satisfies

$$S_2 = \left[\begin{array}{ccc|ccc} d_{n_D} & \cdots & d_{n_D-n_a+n_N+1} & n_{n_N} & \cdots & n_{n_N-n_b+n_D+1} \\ & \ddots & \vdots & & \ddots & \vdots \\ 0 & & d_{n_D} & 0 & & n_{n_N} \end{array}\right]. \tag{15}$$

The matrix S has full row rank if S_1 and S_2 have full row rank. The first is a Sylvester matrix which has full row rank if and only if the numerator and denominator of the controller are coprime [2]. The second has full row rank if $d_{n_D} \neq 0$ or $n_{n_N} \neq 0$, which is true by definition. For the situations $n_a + n_D > n_b + n_N$ and $n_a + n_D < n_b + n_N$ the analysis follows similarly. □

Remark 3.5. If the controller satisfies $n_N = n_D = n_c$ and the parametrized model satisfies $n_a = n_b = n_s$, then the condition of Proposition 3.4 reduces to $n_s \geq n_c$. In other words: connectedness of the parameter set Θ is guaranteed if the order of the controller does not exceed the model order. □

For identification of a simple model based on experiments with a complex controller, connectedness of the parameter set may be a problem. Note that this is the case in Example 3.2.

4. Gradient Expressions

To obtain a parameter estimate the optimization problem given in the previous section has to be solved. Due to the used parametrization this is a nonlinear optimization problem. To find a solution to this optimization problem, gradient search methods can be used like Newton–Raphson and Gauss–Newton as suggested in [8].

However, the character of the function $V_N(\theta)$ that is optimized as well as the parameter set Θ over which is optimized is highly influenced by the controller. Both the function and the set can be extremely non-convex which can make it difficult to apply gradient search methods successfully because the optimization can get stuck in a local minimum or at the boundary of the parameter set. To alleviate these problems it is essential that a good initial estimate is chosen for the iterative search and a good strategy is applied for the choice of the step size.

To apply gradient search methods the gradient of $V_N(\theta)$ needs to be available and for some methods also the Hessian. In this section these derivatives are derived. For convenience of notation, the situation $n_a = n_b = n_s$ and $n_N = n_D = n_c$ is considered. The more general case can be derived similarly. For brevity of notation we also employ $\hat{y}(t,\theta) := \hat{y}(t\,|\,t-1;\theta)$. The derivatives of the cost function can be expressed as

$$\frac{\partial V(\theta)}{\partial \theta} = -\frac{2}{N}\sum_{t=1}^{N} \varepsilon(t,\theta)\frac{\partial \hat{y}(t,\theta)}{\partial \theta} \in \mathbb{R}^{2n_s},$$

$$\frac{\partial^2 V(\theta)}{\partial \theta^2} = 2\sum_{t=1}^{N} \frac{\partial \hat{y}(t,\theta)}{\partial \theta}\left(\frac{\partial \hat{y}(t,\theta)}{\partial \theta}\right)^{\mathrm{T}} - 2\sum_{t=1}^{N} \varepsilon(t,\theta)\frac{\partial^2 \hat{y}(t,\theta)}{\partial \theta^2} \in \mathbb{R}^{2n_s\times 2n_s}.$$

Hence, these derivatives can be calculated if the first and second derivatives of the output prediction are known. These can be calculated by differentiating (9). Differentiating this expression once yields

$$\frac{\partial}{\partial \theta}\{\hat{y}(t,\theta) + [\hat{y}(t-1,\theta) \ \cdots \ \hat{y}(t-n,\theta)] \times ([P_D \ P_N]\theta + \rho)\} = \frac{\partial}{\partial \theta}\{[r(t-1) \ \cdots \ r(t-n)][0 \ P_D]\theta\}$$

or equivalently

$$\frac{\partial \hat{y}(t,\theta)}{\partial \theta} + \left[\frac{\partial \hat{y}(t-1,\theta)}{\partial \theta} \ \cdots \ \frac{\partial \hat{y}(t-n,\theta)}{\partial \theta}\right] \times ([P_D \ P_N]\theta + \rho) + \begin{bmatrix} P_D^{\mathrm{T}} \\ P_N^{\mathrm{T}} \end{bmatrix}\begin{bmatrix} \hat{y}(t-1,\theta) \\ \vdots \\ \hat{y}(t-n,\theta) \end{bmatrix} = \begin{bmatrix} 0 \\ P_D^{\mathrm{T}} \end{bmatrix}\begin{bmatrix} r(t-1) \\ \vdots \\ r(t-n) \end{bmatrix}.$$

This can be written more concisely as

$$F(q,\theta)\frac{\partial \hat{y}(t,\theta)}{\partial \theta} = M^{\mathrm{T}}\psi(t,\theta) \tag{16}$$

with a filter

$$F(q,\theta) = \left(1 + [q^{-1} \ \cdots \ q^{-n}]([P_D \, P_N]\theta + \rho)\right),$$

a matrix

$$M = \begin{bmatrix} P_D & P_N \\ 0 & P_D \end{bmatrix} \in \mathbb{R}^{2n\times 2n_s}$$

and a regression vector $\psi^{\mathrm{T}}(t,\theta) = [-\hat{y}(t-1,\theta)\cdots-\hat{y}(t-n,\theta)r(t-1)\cdots r(t-n)]$. Equation (16) can also be expressed by

$$\frac{\partial \hat{y}(t,\theta)}{\partial \theta} = M^{\mathrm{T}}\psi_F(t,\theta), \tag{17}$$

where $\psi_F(t,\theta) = F^{-1}(q,\theta)\psi(t,\theta)$ is a filtered version of the regression matrix.

The second derivative of the output prediction can be calculated by differentiation of (16), which yields

$$\frac{\partial^2 \hat{y}(t,\theta)}{\partial \theta^2}F(q,\theta) + \frac{\partial F(q,\theta)}{\partial \theta}\left(\frac{\partial \hat{y}(t,\theta)}{\partial \theta}\right)^{\mathrm{T}} = -\left[\frac{\partial \hat{y}(t-1,\theta)}{\partial \theta}\cdots\frac{\partial \hat{y}(t-n,\theta)}{\partial \theta}\right][P_D \, P_N],$$

or

$$\frac{\partial^2 \hat{y}(t,\theta)}{\partial \theta^2} = -2F^{-1}(q,\theta) \times \left[\frac{\partial \hat{y}(t-1,\theta)}{\partial \theta}\cdots\frac{\partial \hat{y}(t-n,\theta)}{\partial \theta}\right][P_D \, P_N].$$

These compact expressions are similar to expressions obtained for nonlinear optimization with a standard input–output parametrization and can be fruitfully used in nonlinear optimization routines for closed-loop identification with a tailor-made parametrization.

Alternative expressions for the predictor gradient, employing the sensitivity function of the parametrized closed-loop system, are presented in [3].

5. Simulation examples

In this section two simulation examples are given. One in the case where $G_0 \in \mathcal{G}$ and the parameter set is not connected and the other where $G_0 \notin \mathcal{G}$ with a connected parameter set. In the first example the tailor-made parametrization induces an optimization problem which is difficult to solve while in the second example it is demonstrated that closed-loop identification with this parametrization can be very powerful.

5.1. Simulation 1

In Fig. 2 the three separate branches of the cost function $V_N(\theta)$ for the system from Example 3.2 is depicted for a system $G_0 = 3.5$ and $G(\theta) = \theta$. The output disturbance $v(t)$ in Fig. 1 is white noise with variance $\sigma = 0.1$. The excitation signal $r(t)$ is white noise with variance 1. Note that the parameter regions $(-\infty, 0]$, $[1.27, 2.64]$ and $[4.69, 9.98]$ induce an unstable closed-loop system. The criterion function has several local minima that are located at the boundary of the stability area if the iterative search for the optimal parameter vector is confined to those parameters θ for which the closed-loop system is stable. This makes it difficult to find the optimum with gradient search methods.

The global optimum will generally only be found if an initial estimate is selected from the middle of the three branches of the criterion function. If the number of data points go to infinity, these local minima are not at the boundary of the stability area because in that case the value of the criterion function goes to infinity if the parameter approaches the closed-loop instability area.

5.2. Simulation 2

A simulation is made with a 5th order system, which is given by the transfer function

$$\begin{aligned} G_0(z) = {} & [10^{-5}] \times \{(5.278z^{-1} + 126.7z^{-2} \\ & + 299.3z^{-3} + 110.8z^{-4} + 4.042z^{-5}) \\ & /(1 - 4.391z^{-1} + 7.879z^{-2} \\ & - 7.247z^{-3} + 3.430z^{-4} - 0.6703z^{-5})\} \end{aligned}$$

being an integrator with two resonant modes. The controller used in the simulation is a PI-controller which stabilizes the system, and is given by

$$C(q) = 10^{-2} \cdot \frac{1 - 0.9q^{-1}}{1 - q^{-1}}.$$

The excitation signal $r(t)$ is Gaussian white noise with a standard deviation of $\sigma_r = 1$ and the output noise $v(t)$ is Gaussian white noise with a standard deviation of $\sigma_v = 1.5$. The data length is $N = 500$. The open-loop and closed-loop transfer functions are given in Fig. 3.

For this system a 3rd order model is estimated with a tailor-made parametrization. For the nonlinear optimization a Gauss–Newton method is applied where the initial estimate is obtained with the use of direct identification with an ARX(3, 3, 1) model structure. The speed of the nonlinear optimization routine is improved by using the explicit gradient expression derived in Section 4. The estimated model is given in Fig. 3. Also the initial model is given. From this it can be seen that the estimation with a tailor-made parametrization gives a good fit for the integrator and the first resonant mode, despite the bad signal-to-noise ratio and the bad initial estimate.

6. Discussion

As shown in the previous section, a condition on controller and model order can act as a sufficient condition under which the parameter set in a tailor-made parametrization is connected, leading to a uniformly stable model set, as required for consistency of the identification. This allows the user to verify *a priori* whether consistency problems might occur. As the condition is only sufficient, no definite conclusion about lack of consistency can be drawn if it is not satisfied. Moreover, one can argue whether a possible non-connectedness of the parameter set is a serious problem from a practical point of view. When using gradient-type methods for the nonlinear optimization of the cost function, the choice of a bad initial estimate can lead to a (bad) local minimum, irrespective of the fact whether the global parameter set is connected or not.

The identification method discussed here is closely related to the so-called indirect method for the closed-loop identification. In this latter approach first the

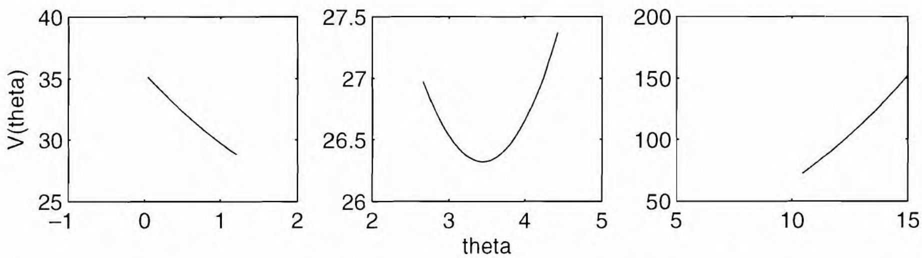

Fig. 2. Criterion function for the identification problem in Example 4.1 with $G_0 = 3.5$ and number of data $N = 100$.

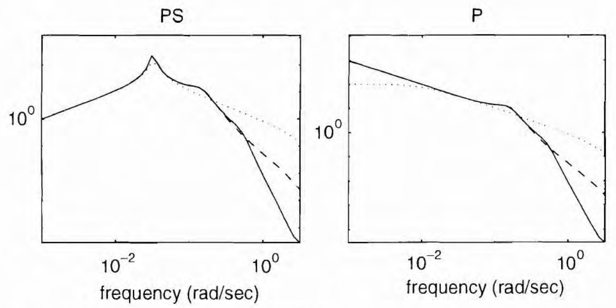

Fig. 3. Amplitude plot closed-loop transfer from $r(t)$ to $y(t)$ (left) and open-loop transfer (right): plant (solid), estimation with tailor-made parametrization (dashed) and direct identification (dotted).

closed-loop transfer function $R(q)$ is identified with a standard numerator–denominator parametrization. Next, a plant model is calculated using knowledge of the controller with $\hat{G}(q) = R(q, \hat{\theta})(1 - R(q, \hat{\theta}) \times C(q))^{-1}$. Estimation of a plant model with a prespecified model order is not a trivial task here. In comparison to the indirect method, the tailor-made approach has the advantage that the model set of plant models can simply be constructed to contain plant models of prespecified order. The same advantage also holds in comparison to the identification using a dual-Youla parametrization [5,14]. There the plant model order is hard to control, but at the benefit of the ability to parametrize all models that are stabilized by the given controller. In this respect the dual-Youla parametrization can be seen as the 'tailor-made' parametrization that is guaranteed to deliver a connected parameter set, however, at the cost of losing control over the plant model order in the model set.

The specific approximative properties of the presented identification method can be obtained from (6). This expression can be further specified as

$$\theta^* = \arg\min_{\theta \in \Theta} \frac{1}{2\pi} \int_{-\pi}^{\pi} |S_0(e^{i\omega})[G_0(e^{i\omega}) - G(e^{i\omega}, \theta)] \times S(e^{i\omega}, \theta)|^2 \Phi_r(\omega)\, d\omega, \tag{18}$$

where $S(q, \theta) = (1 + C(q)G(q, \theta))^{-1}$ is the parametrized sensitivity function. From this it can be seen that the additive plant model error is weighted by both the sensitivity function and the estimated sensitivity function which puts an emphasis on the cross-over region of the closed-loop system. This implies that in the case of approximate modelling ($G_0 \notin \mathcal{G}$) the undermodelling error is particularly small in this frequency region. This has been considered an attractive property in case the identified model is used in model-based control design, as pointed out in [4,12]. In many control-relevant identification schemes this type of weighting is pursued but can only be approximated by the use of specific filtering strategies; by using a tailor-made parametrization this weighting is inherently present. Moreover, by designing the reference signal r, the bias expression (18) can directly be tuned to the designer's needs.

Attention has been restricted to the situation of a fixed noise model in an output error structure $W(q, \theta) = 1$ (see (3)). However, noise models can be identified as well. Actually the choice of particular noise models provides a link with other identification schemes. Utilizing knowledge of the closed-loop structure, a possible parametrization could be

$$R(q, \theta) = \frac{G(q, \theta)}{1 + C(q)G(q, \theta)},$$
$$W(q, \theta) = \frac{H(q, \theta)}{1 + C(q)G(q, \theta)}.$$

This parametrization leads to a prediction

$$\hat{y}(t|t-1; \theta) = y(t) - H(q, \theta)^{-1} \times \{y(t) - G(q, \theta)[r(t) - C(q)y(t)]\}.$$

Using $u(t) = (r(t) - C(q)y(t))$ it follows directly that this predictor is the same as the predictor in a direct identification method on the basis of measured data u and y. Using a particularly structured noise model, a tailor-made parametrization method can thus become equivalent to direct identification.[3] Note that in this situation a consistent identification of G_0 is only possible if the noise model can be exactly modelled within the chosen model set (system is in the model set). If only G_0 can be modelled exactly within the model set, inconsistency will occur due to the fact that plant model and noise model have common parameters. If $R(q, \theta)$ and $W(q, \theta)$ are parametrized independently, the consistency result given in Proposition 2.1 still holds in case $G_0 \in \mathcal{G}$.

7. Conclusions

In this paper identification of a model from closed-loop data with a tailor-made parametrization is discussed. Special attention is paid to the possible occurrence of a non-connected parameter set which is induced by the structure of the parametrization. Sufficient conditions are derived for the model order in

[3] Note, however, that the tailor-made approach will require knowledge of the controller in contrast to the direct method. If the controller is linear, time-invariant and finite-dimensional, it can simply be identified from knowledge of r, u and y.

terms of the controller complexity such that the parameter set is connected. These conditions indicate that the parameter set may not be a connected set in case a low complexity model is identified from data obtained with a high complexity controller.

Additionally it is shown that for a specific parametrization of the noise model, the method reduces to closed-loop identification with the classical direct method.

Acknowledgements

The authors wish to thank Carsten Scherer and the anonymous referees for fruitful discussions and comments.

References

1. Åström KJ, Wittenmark B. *Computer controlled systems – theory and design*. Prentice-Hall Inc., Englewood Cliffs NJ 1990.
2. Chen C. *Linear system theory and design*. Saunders College Publishing USA 1984.
3. De Bruyne F, Anderson BDO, Gevers M, Linard N. On closed-loop identification with a tailor-made parametrization. *Proc. 1998 American Control Conf.*, Philadelphia, PA, pp. 3177–3182, 1998.
4. Gevers M. Towards a joint design of identification and control? *Essays on control: perspectives in the theory and its applications*. Birkhäuser, Boston, 1993, pp. 111–151.
5. Hansen FR, Franklin GF, Kosut RL. Closed-loop identification via the fractional representation: experiment design. *In: Proc amer control conf*, Pittsburgh, PA, 1989, pp. 1422–1427.
6. Hjalmarsson H, Gevers M, De Bruyne F. For model-based control design, closed-loop identification gives better performance. *Automatica*, 1996, **32**(12): 1659–1673.
7. Landau, ID, Karimi A. An output error recursive algorithm for unbiased identification in closed loop. *Automatica*, 1997, Vol. 33, No. 5, pp. 933–938.
8. Ljung L. *System identification: theory for the user*. Prentice-Hall, Englewood Cliffs NJ, 1987.
9. Ljung L. Identification in closed-loop: some aspects on direct and indirect approaches. *In: Prepr. 11th IFAC symp system identification*, Vol. 1, Kitakyushu, Japan, 1997, pp. 141–146.
10. Söderström T, Stoïca P. *System identification*. Prentice-Hall, Hemel Hempstead UK, 1989.
11. Van den Hof PMJ, Schrama RJP. An indirect method for transfer function estimation from closed loop data. *Automatica*, 1993; **29**(6): 1523–1527.
12. Van den Hof PMJ, Schrama RJP. Identification and control – closed loop issues. *Automatica* 1995; **31**(12): 1751–1770.
13. Van den Hof PMJ, Schrama RJP, de Callafon RA, Bosgra OH. Identification of normalised coprime plant factors from closed-loop experimental data. *Europ. J. Control*, 1995; **1**(1): 62–74.
14. Van den Hof PMJ, de Callafon RA. Multivariable closed-loop identification: from indirect identification to dual-Youla parametrization. *In: Proc. 35th IEEE conf. dec. control*, Kobe, Japan, 1996, pp. 1397–1402.
15. Van den Hof PMJ. Closed-loop issues in system identification. In: *Prepr. 11th IFAC symp system identification*, 1997, vol. 4, Kitakyushu, Japan, pp. 1651–1664.

8. Appendix

Lemma 8.1. *The parameter set* $\Theta \subset \mathbb{R}^n$ *with elements* $\theta = [p_1 \cdots p_n]^T$, $\{p_i\}_{i=1,\ldots,n} \in \mathbb{R}$ *for which all polynomials*

$$p(z) = z^n + \left[\, z^{n-1}\, z^{n-2} \cdots 1 \,\right] \theta$$

have stable roots, is a pathwise connected subset of $\mathbb{R}^n$.

Proof. First the polynomial $p(z)$ is reparametrized as a product of first and second order polynomials

$$p(z) = \begin{cases} \prod_{k=1}^{n/2} (z^2 + a_k z + b_k), \forall k & n \text{ even}, \\ (z+c) \prod_{k=1}^{(n-1)/2} (z^2 + a_k z + b_k), \forall k & n \text{ odd}. \end{cases} \tag{19}$$

Stability of the full polynomial is guaranteed if stability of the second order polynomials and the first order polynomial is guaranteed, which is guaranteed if and only if $b_k < 1$, $a_k < 1 + b_k$, $-a_k < 1 + b_k$, $\forall k$ and $-1 < c < 1$, see e.g. [1]. This stability area for the quadratic terms describes a triangular area in the a_k, b_k-plane which is not only pathwise connected but also convex. The stability area for the first order term is also convex. The polynomial coefficients of the original polynomial, $\{p_i\}_{i=1,\ldots,n}$, are continuous and continuously differentiable functions in the parameters $\{a_k, b_k\}_{i=1,\ldots,n}$. Therefore, from pathwise connectedness of the set of admissible coefficients $\{a_k, b_k\}_{i=1,\ldots,n}$, pathwise connectedness of the set of admissible parameters $\{p_i\}_{i=1,\ldots,n}$ can be concluded.