NEW INSIGHTS IN CLOSED-LOOP IDENTIFICATION WITH BIAS-ELIMINATED LEAST-SQUARES (BELS)

Marion Gilson*, **Paul Van den Hof**†

* Centre de Recherche en Automatique de Nancy (CRAN)
CNRS UPRESA 7039, Université Henri Poincaré, Nancy 1, BP 239, F-54506 Vandoeuvre-les-Nancy Cedex, France
`fax: +(33) 3 83 91 20 30`
`e-mail: marion.gilson@uhp-nancy.fr`

† Signals, Systems and Control Group, Department of Applied Physics,
Delft University of Technology, Lorentzweg 1, 2628 CJ Delft, The Netherlands,
`fax: +(31) 15 2784263`
`e-mail: p.m.j.vandenhof@tnw.tudelft.nl`

Keywords: System identification; closed-loop identification; prediction error methods; instrumental variables.

Abstract

A bias-correction method for closed-loop identification, introduced in the literature as the bias-eliminated least squares (BELS) method [9], is shown to be equivalent to a basic instrumental variable estimator applied to a predictor for the closed-loop system. This predictor is a function of the plant parameters and the known controller. Corresponding to the related method using a least squares criterion, the method is referred to as the tailor-made IV method for closed-loop identification. The indicated equivalence greatly facilitates the understanding and the analysis of the BELS method.

1 Introduction

Least squares methods based on the bias-correction principle aim at providing unbiased plant parameter estimates, while using linear-in-the-parameters model structures, see e.g. [4, 11]. They retain all merits of the LS method and make it possible to cope with the bias problem in the identification of systems subject to colored disturbances. Recently these kind of methods have also been developed for identification under closed-loop conditions [9, 10]. The proposed method, called the bias-eliminated least-squares (BELS) method, is able to estimate unbiased plant parameters in indirect closed-loop system identification. In [5], it has been shown, based on the work of [6], that the bias-eliminated least-squares estimator proposed in [11] for open-loop system identification is identical to a basic instrumental variable estimator. For the closed-loop identification case, the BELS method is analysed in [8], where a relation is shown with a particular (and rather complex) frequency weighted IV method, applied to the input and output measurement data, gathered under closed-loop conditions.

In this paper, it will be shown that the closed-loop BELS method is equivalent to a so-called tailor-made instrumental variable method, where the predictor for the closed-loop system is used to generate the prediction error and the external reference signals are used as instrumental variables. This connects to the (least squares) tailor-made identification method for closed-loop identification that was recently introduced in the literature, [7, 2]. This equivalence greatly facilitates the understanding and analysis of the BELS method.

2 Preliminaries

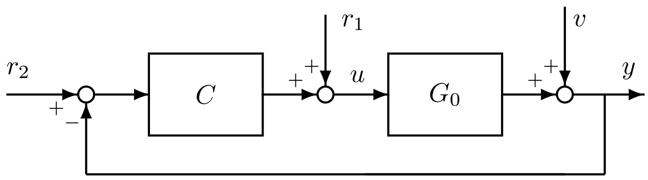

Figure 1: Closed-loop configuration.

Consider a linear SISO closed-loop system shown in figure 1. The process is denoted with $G_0(z)$ and the controller with $C(z)$; $u(t)$ is the process input signal, $y(t)$ the process output signal and $v(t)$ describes the disturbances acting on the loop. The external signals $r_1(t)$, $r_2(t)$ are assumed to be uncorrelated with the noise disturbance $v(t)$. For ease of notation we also introduce the signal $r(t) = r_1(t) + C(q)r_2(t)$. With this notation the closed-loop system can be described as

$$\mathcal{S} : y(t) = \frac{G_0}{1+CG_0} r(t) + \frac{1}{1+CG_0} v(t). \qquad (1)$$

A parametrized process model is considered

$$\mathcal{G} : G(q,\theta) = \frac{B(q^{-1},\theta)}{A(q^{-1},\theta)} = \frac{b_1 q^{-1} + \cdots + b_n q^{-n}}{1 + a_1 q^{-1} + \cdots + a_n q^{-n}} \qquad (2)$$

and the process model parameters are stacked columnwise in the parameter vector

$$\theta = \begin{bmatrix} a_1 & \cdots & a_n & b_1 & \cdots & b_n \end{bmatrix}^T \in \mathbb{R}^{2n}. \qquad (3)$$

The real plant G_0 is considered to satisfy

$$G_0(q) = \frac{B_0(q^{-1})}{A_0(q^{-1})},$$

while in these expressions q^{-1} is the delay operator, and the numerator and denominator degree is n_0. The m-th order controller C is assumed to be known and specified by

$$C(q) = \frac{Q(q^{-1})}{P(q^{-1})} = \frac{q_0 + q_1 q^{-1} + \cdots + q_m q^{-m}}{1 + p_1 q^{-1} + \cdots + p_m q^{-m}} \qquad (4)$$

with the pair (P, Q) assumed to be coprime. The closed-loop transfer function (1) can be rewritten in polynomial fraction form

$$y(t) = \frac{B^0_{cl}}{A^0_{cl}} r(t) + \frac{1}{A^0_{cl}} \xi(t) \qquad (5)$$

with $\xi(t) = A_0 P v(t)$. The polynomials B^0_{cl} and A^0_{cl} will generically have orders $n_0 + m$.

For parametrizing the closed-loop transfer function $G_0/(1 + CG_0)$ the following model structure is used

$$B_{cl}(q^{-1}, \Theta) = \beta_1 q^{-1} + \cdots + \beta_{r_B} q^{-r_B} \qquad (6)$$

$$A_{cl}(q^{-1}, \Theta) = 1 + \alpha_1 q^{-1} + \cdots + \alpha_{n+m} q^{-(n+m)} \qquad (7)$$

and the closed-loop parameters are collected in the parameter vector

$$\begin{aligned} \Theta &= \begin{bmatrix} \alpha^T & \beta^T \end{bmatrix}^T \\ &= \begin{bmatrix} \alpha_1 & \cdots & \alpha_{n+m} & \beta_1 & \cdots & \beta_{r_B} \end{bmatrix}^T \in \mathbb{R}^{n+m+r_B}. \end{aligned} \qquad (8)$$

For $r_B \geq (n_0 + m)$ the closed-loop model structure will be flexible enough to exactly represent the reference to output transfer function in the closed-loop system (5).

The bias-eliminated least squares method that is considered in this paper attempts to estimate the process parameters by an indirect closed-loop identification. This means that the closed-loop transfer function (5) is identified, after which process parameters (3) are determined.

The relation between (open-loop) process parameters and closed-loop parameters is determined by the linear equation[1]

$$\Theta = M\theta + \rho \qquad (9)$$

where ρ is a known vector and M is a known full-column rank matrix, given by

$$M = \begin{pmatrix} P_c & Q_c \\ 0 & \bar{P}_c \end{pmatrix} \in \mathbb{R}^{(n+m+r_B)\times 2n} \qquad (10)$$

$$\rho = \begin{pmatrix} p_1 & \cdots & p_m & 0 & \cdots & 0 \end{pmatrix}^T \in \mathbb{R}^{(n+m+r_B)} \qquad (11)$$

P_c, $Q_c \in \mathbb{R}^{(n+m)\times n}$ are Sylvester matrices expanded by $\begin{bmatrix} 1 & p_1 & \cdots & p_m \end{bmatrix}^T$ and $\begin{bmatrix} q_0 & q_1 & \cdots & q_m \end{bmatrix}^T$ respectively, e.g.

$$P_c = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ p_1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ p_m & \vdots & \ddots & 1 \\ 0 & \ddots & \vdots & p_1 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & p_m \end{bmatrix} \qquad (12)$$

$\bar{P}_c \in \mathbb{R}^{r_B \times n}$ is also a Sylvester matrix defined by

$$\bar{P}_c = \begin{bmatrix} P_c \\ 0_{(r_B - n - m)\times n} \end{bmatrix}.$$

[1] Note that in [9, 10] the corresponding equation is written as $\Theta = M\theta - \rho$; the difference in sign is due to the fact that in the mentioned references denominator parameters appear with a negative sign in the parameter vectors.

3 Bias-eliminated least-squares method

The bias-eliminated least-squares method (BELS) for closed-loop identification as discussed in [9, 10] is designed to provide an unbiased estimate for the process model $G(q, \theta)$, while pertaining to simple algorithmic schemes as the linear regression type of estimates. Accurate noise modeling (i.e. finding an noise-shaping filter that models the disturbance signal v) is not considered part of the problem. The method comprises the following main steps:

- Estimate an ARX model for the closed-loop system (5) on the basis of data r and y; this estimate is denoted by $\hat{\Theta}_{ls}$.
- This estimate generally will be biased due to the fact that $\xi(t)$ in (5) will not be white noise; however the bias on $B_{cl}(q^{-1}, \hat{\Theta}_{ls})/A_{cl}(q^{-1}, \hat{\Theta}_{ls})$ can be estimated and subtracted from the closed-loop estimate;
- The corrected closed-loop parameter is converted to an equivalent open-loop process parameter by solving for (9) in a least squares sense.

In short

$$\text{CL data} \underset{ARX}{\longrightarrow} \hat{\Theta}_{ls} \underset{Computation}{\longrightarrow} \hat{\Theta}_{corr} \underset{LS}{\longrightarrow} \hat{\theta}_{bels}$$

ARX estimate

An ARX estimate for the closed-loop system is obtained through

$$\hat{\Theta}_{ls}(N) = \hat{R}_{\varphi\varphi}(N)^{-1} \hat{R}_{\varphi y}(N)$$

where

$$\hat{R}_{\varphi\varphi}(N) = \frac{1}{N} \sum_{t=1}^{N} \varphi(t)\varphi^T(t) \qquad (13)$$

$$\hat{R}_{\varphi y}(N) = \frac{1}{N} \sum_{t=1}^{N} \varphi(t) y(t) \qquad (14)$$

$$\varphi(t) = [-y(t-1) \ \cdots \ -y(t-n-m) \ r(t-1) \ \cdots \ r(t-r_B)]^T.$$

Bias correction

The bias correction principle is based on the following reasoning. If the ARX model structure is rich enough to capture all dynamics of the closed-loop system (i.e. if the system is in the model set), then

$$\hat{\Theta}_{ls}(N) = \Theta_0 + \hat{R}^{-1}_{\varphi\varphi}(N)\hat{R}_{\varphi\xi}(N) \tag{15}$$

where Θ_0 is the coefficient vector of the real closed-loop plant, and

$$\hat{R}_{\varphi\xi}(N) = \frac{1}{N}\sum_{t=1}^{N}\varphi(t)\xi(t).$$

Then, under minor regularity conditions on the data, the least squares estimate $\hat{\Theta}_{ls}(N)$ is known to converge for $N \to \infty$ with probability 1 to

$$\Theta^*_{ls} = \Theta_0 + R^{-1}_{\varphi\varphi}R_{\varphi\xi}$$

with

$$\begin{aligned} R_{\varphi\varphi} &= \bar{E}\varphi(t)\varphi^T(t) \\ R_{\varphi\xi} &= \bar{E}\varphi(t)\xi(t), \end{aligned}$$

where the notation $\bar{E} := \lim_{N\to\infty}\frac{1}{N}E$ is adopted from the prediction error framework of [3].
As the noise disturbance ξ is assumed to be uncorrelated with the reference signal r, the bias in the asymptotic estimate is given by

$$\Delta^* := R^{-1}_{\varphi\varphi}R_{\varphi\xi} = R^{-1}_{\varphi\varphi}\begin{bmatrix} I_{n+m} \\ 0 \end{bmatrix} R_{y\xi}.$$

Based on this expression, an estimate for Δ is obtained by considering

$$\hat{\Delta}(N) := \hat{R}^{-1}_{\varphi\varphi}(N)\begin{bmatrix} I_{n+m} \\ 0 \end{bmatrix}\hat{R}_{y\xi}(N).$$

The unknown $\hat{R}_{y\xi}(N)$ in this relation can be obtained by the following reasoning.

As matrix M in (9) has full column rank, there exists a full column rank matrix $H \in \mathbb{R}^{(n+m+r_B)\times(m+r_B-n)}$ that satisfies $H^TM = 0$. Multiplying equation (15) by H^T and using equation (9) for Θ_0 , it follows that

$$H^T\hat{R}^{-1}_{\varphi\varphi}(N)\begin{bmatrix} I_{n+m} \\ 0 \end{bmatrix}\hat{R}_{y\xi}(N) = H^T(\hat{\Theta}_{ls}(N) - \rho). \tag{16}$$

This is a set of $m + r_B - n$ equations with $n + m$ unknowns in $\hat{R}_{y\xi}(N)$. There are two situations to be distinguished

- $m \geq n$ (see [9]). r_B is chosen according to $r_B = n + m$, and equation (16) is an overdetermined set of equations that is solved in least squares sense, leading to

$$\hat{\Delta}(N) = \hat{R}^{-1}_{\varphi\varphi}(N)\begin{bmatrix} I_{n+m} \\ 0 \end{bmatrix}\left[H^T\hat{R}^{-1}_{\varphi\varphi}(N)\begin{bmatrix} I_{n+m} \\ 0 \end{bmatrix}\right]^+ H^T\left[\hat{\Theta}_{ls}(N) - \rho\right] \tag{17}$$

  with $(\cdot)^+$ denoting the matrix pseudo-inverse.

- $m < n$ (see [10]). By choosing $r_B = m + n$ the number of equations in (16) is not sufficient to uniquely determine $\hat{\Delta}$. In [10], this is solved by applying a dynamic prefilter to the reference signal such that effectively a system with higher numerator degree is estimated. This is equivalent to simply choosing $r_B = 2n$, thus obtaining the situation that (16) is uniquely solvable for $\hat{R}_{y\xi}(N)$. An estimate $\hat{\Delta}(N)$ can then be constructed according to

$$\hat{\Delta}(N) = \hat{R}^{-1}_{\varphi\varphi}(N)\begin{bmatrix} I_{n+m} \\ 0 \end{bmatrix}\left[H^T\hat{R}^{-1}_{\varphi\varphi}(N)\begin{bmatrix} I_{n+m} \\ 0 \end{bmatrix}\right]^{-1} H^T\left[\hat{\Theta}_{ls}(N) - \rho\right]. \tag{18}$$

Combining both situations it appears that r_B can be set to

$$r_B = \max(2n, n + m).$$

The bias elimination can now be performed by constructing the corrected closed-loop parameter vector

$$\hat{\Theta}_{corr}(N) = \hat{\Theta}_{ls}(N) - \hat{\Delta}(N). \tag{19}$$

Finally the plant parameter estimate $\hat{\theta}_{bels}$ is obtained by solving (9) in a least squares sense

$$\hat{\theta}_{bels}(N) = (M^TM)^{-1}M^T(\hat{\Theta}_{corr}(N) - \rho). \tag{20}$$

It has been shown in [9, 10] that this resulting parameter estimate is asymptotically unbiased.

4 Tailor-made IV identification

4.1 Main result

In the first step of the BELS method the closed-loop system is estimated with a general ARX (black-box) model structure. However as we know that the closed-loop system has a particular structure (1) with a known controller C, this structure can also be imposed on the parametrization of the closed-loop.

When defining

$$\bar{B}_{cl}(q^{-1},\theta) = B(q^{-1},\theta)P(q^{-1}) \tag{21}$$

$$\bar{A}_{cl}(q^{-1},\theta) = A(q^{-1},\theta)P(q^{-1}) + B(q^{-1},\theta)Q(q^{-1}) \tag{22}$$

a parametrization of the closed-loop system has been obtained, in terms of the process parameters θ. In the literature this is known as a tailor-made parametrization, and has been applied before in prediction error identification with least squares criteria, see e.g. [7, 2].

Next the main result is formulated.

Proposition 1 *Consider a data generating system according to (1), such that the closed-loop system is asymptotically stable, and consider the BELS estimate $\hat{\theta}_{bels}(N)$ given by (20), with $r_B = \max(2n, n + m)$, and r persistently exciting of sufficiently high order.*

Define the weighted tailor-made IV estimate $\hat{\theta}_{iv,F}(N)$ as the solution to the set of $2n$ equations

$$\frac{1}{N}\sum_{t=1}^{N}\varepsilon(t,\hat{\theta}_{iv,F})\eta(t)=0 \tag{23}$$

$$\varepsilon(t,\theta)=\bar{A}_{cl}(q^{-1},\theta)y(t)-\bar{B}_{cl}(q^{-1},\theta)r(t) \tag{24}$$

$$\eta(t):=F\varphi_r(t),\quad F\in\mathbb{R}^{2n\times r_B} \tag{25}$$

$$\textit{with } \varphi_r(t):=\textit{col}\begin{bmatrix}r(t-1) & \cdots & r(t-r_B)\end{bmatrix} \tag{26}$$

with

$F=I_{2n}$ *in the situation* $m\leq n$, *and*

$F=M^T\hat{R}^T_{\varphi_r\varphi}(\hat{R}_{\varphi_r\varphi}\hat{R}^T_{\varphi_r\varphi})^{-1}$ *in the situation* $m>n$.

Then

$$\hat{\theta}_{bels}(N)=\hat{\theta}_{iv,F}(N).$$

Proof. The line of reasonning of the proof is given in the appendix and a full proof can be found in [1].

Remarks

- If the order of the process model exceeds the controller order ($n\geq m$), the parameter estimate is a simple tailor-made IV estimate, where the closed-loop prediction error is made orthogonal to delayed versions of the external reference signal. Related estimation algorithms based on a least squares criterion $\sum_t\varepsilon^2(t,\theta)$ have been considered in [7, 2]. The indicated equivalence greatly facilitates the understanding and analysis of the BELS-estimator. Moreover, it also allows the analysis of the estimator under conditions where the real process is not considered to be present in the model set. Note that in the formulation of the main result it is *not* assumed that G_0 has order n.

- In the situation $m<n$, it is suggested by [10] to introduce an auxiliary filter operating on the reference signals in order to increase the number of numerator parameters to identify. Here it is shown, as was also indicated by [8], that this dynamic prefilter is superfluous. The problem can be handled by simply choosing $r_B=2n$, i.e. by deliberately enlarging the number of numerator parameters to estimate.

- When $m>n$, the estimator is obtained by using a linear combination of delayed samples of the reference signal, to act as an instruments in the IV estimator.

4.2 Interpretation of matrix F

In order to interpret the role of matrix F, let us analyse $\varepsilon(t,\hat{\theta}_{bels})$, the equation error of the BELS estimator. As the connection between IV and BELS estimators has been stated, the equation error can be written as

$$\varepsilon(t,\hat{\theta}_{iv,F})=y(t)-\varphi^T(t)(M\hat{\theta}_{iv,F}+\rho). \tag{27}$$

This equation can be simplified by analyzing the constituting expressions. At first, it can be noticed that the controller denominator $P(q^{-1})$ can be used as a prefilter for the output $y(t)$. Then, with

$$\bar{y}(t)=P(q^{-1})y(t) \tag{28}$$

it follows that

$$\bar{y}(t)=y(t)-\varphi^T(t)\rho \tag{29}$$

The second expression $\varphi^T(t)M$ can be rephrased by considering the plant description

$$y(t)=\psi^T(t)\theta_0+A_0(q^{-1})v(t),$$

with

$$\psi(t)=[-y(t-1)\ \cdots\ -y(t-n)\quad u(t-1)\ \cdots\ u(t-n)]^T\in\mathbb{R}^{2n}.$$

In a filtered version this reads:

$$\bar{y}(t)=\bar{\psi}^T(t)\theta_0+P(q^{-1})A_0(q^{-1})v(t) \tag{30}$$

where $\bar{\psi}(t):=P(q^{-1})\psi(t)$. Similarly,

$$y(t)=\varphi^T(t)(M\theta_0+\rho)+P(q^{-1})A_0(q^{-1})v(t)$$

leading to

$$\bar{y}(t)=\varphi^T(t)M\theta_0+P(q^{-1})A_0(q^{-1})v(t)$$

which combined with (30) shows that

$$\bar{\psi}^T(t)=\varphi^T(t)M.$$

Using this expression in (27) leads to

$$\varepsilon(t,\hat{\theta}_{iv,F})=\bar{y}(t)-\bar{\psi}^T(t)\hat{\theta}_{iv,F}. \tag{31}$$

For a further interpretation two cases have to be considered, according to the orders of the controller and the system.

Case $\mathbf{m\leq n}$. If the controller order is smaller than or equal to the system order, it has been stated that F is equal to the identity matrix and the tailor-made IV estimate satisfies

$$\hat{R}_{\varphi_r\varepsilon}=0. \tag{32}$$

By substituting (31) in (32), this yields

$$\hat{R}_{\varphi_r\bar{y}}-\hat{R}_{\varphi_r\bar{\psi}}\hat{\theta}_{iv,F}=0. \tag{33}$$

If the signal $r(t)$ is persistently exciting, of sufficient order, the squared matrix $\hat{R}_{\varphi_r\bar{\psi}}\in\mathbb{R}^{2n\times 2n}$ is invertible and thus the IV estimate is given by

$$\hat{\theta}_{iv}=\hat{R}^{-1}_{\varphi_r\bar{\psi}}\hat{R}_{\varphi_r\bar{y}}. \tag{34}$$

Case $\mathbf{m > n}$. In the case where the controller order is greater than the system order, the vector φ_r is made up of $(n+m)$ components (see the proposition). It follows that the matrix $\hat{R}_{\varphi_r \bar{\psi}} \in \mathbb{R}^{(n+m)\times 2n}$ is not invertible. Thus, the matrix F is added in order to make it invertible, i.e. to make $\hat{R}_{\varphi_r \bar{\psi}}$ regular. In this case, F is equal to $M^T \hat{R}^T_{\varphi_r\varphi}(\hat{R}_{\varphi_r\varphi}\hat{R}^T_{\varphi_r\varphi})^{-1}$ and the tailor-made IV estimate satisfies

$$\hat{R}_{F\varphi_r\varepsilon} = 0. \tag{35}$$

By substituting (31) in (35), it follows

$$\hat{R}_{F\varphi_r\bar{y}} - \hat{R}_{F\varphi_r\bar{\psi}}\hat{\theta}_{iv,F} = 0. \tag{36}$$

If $r(t)$ is persistently exciting of sufficient order, the matrix $\hat{R}_{F\varphi_r\bar{\psi}} \in \mathbb{R}^{2n\times 2n}$ is invertible and the IV estimate can be written as

$$\hat{\theta}_{iv} = \hat{R}^{-1}_{F\varphi_r\bar{\psi}}\hat{R}_{F\varphi_r\bar{y}}. \tag{37}$$

The matrix $\hat{R}_{F\varphi_r\bar{\psi}}$ can be regarded as the product of two matrices F and $\hat{R}_{\varphi_r\bar{\psi}}$. $F \in \mathbb{R}^{2n\times(n+m)}$ and $\hat{R}_{\varphi_r\bar{\psi}} \in \mathbb{R}^{(n+m)\times 2n}$ have both rank $2n$. Thus, the product $F\hat{R}_{\varphi_r\bar{\psi}}$, or equivalently $\hat{R}_{F\varphi_r\bar{\psi}}$, is squared (dimensions $2n \times 2n$) and has rank $2n$.

5 Relation with other work

Recently it was claimed in [8] that the BELS method of [9, 10] when applied to closed-loop data, is equivalent to a particular frequency weighted optimal IV estimator. In this analysis use is made of earlier results from [6], in the open-loop context.

However, the estimator analysed in [8] is different from the BELS estimator of [9, 10] as will be indicated next.

In [8] the following situation is considered.

$$r_B = n + m \tag{38}$$

M is given by equation (10).

The matrix, perpendicular to M is chosen as

$$H_1^T M = 0 \quad \text{with } H_1 \in \mathbb{R}^{(2n+2m)\times(n+m)}, \tag{39}$$

where in the situation $m > n$, H_1 has a smaller column dimension than the matrix H used before in this paper. As a result, this leads to an estimator different from the one considered in the paper [9].

In order to apply the reasoning of [6] in the closed-loop context, a linear regression has to be found between the closed-loop regression vector $\varphi(t)$ and the open-loop one $\psi(t)$. As a result, the filter $P(q^{-1})$ is used to change the affine map ($\Theta_0 = M\theta_0 + \rho$) into a linear one ($\bar{\psi}^T(t) = M^T\varphi(t)$). Then, the prefiltered open-loop system relation are defined by

$$P(q^{-1})y(t) = P(q^{-1})\psi^T(t)\theta_0 + P(q^{-1})A_0 e(t). \tag{40}$$

or equivalently,

$$\bar{y}(t) = \bar{\psi}^T(t)\theta_0 + \xi(t). \tag{41}$$

The definition of $\bar{y}(t)$ and $\bar{\psi}(t)$ assume that a delay is operating on the loop, due to a hold. Then, a weighted optimal IV estimator operating on the plant input and output signals is denoted by

$$\hat{\theta}_{iv}(N) = (\hat{G}^T W_{iv}\hat{G})^{-1}\hat{G}^T W_{iv}\hat{p} \tag{42}$$

where

$$\hat{G} = \frac{1}{N}\sum_{t=1}^{N} z(t)\bar{\psi}^T(t) \tag{43}$$

$$\hat{p} = \frac{1}{N}\sum_{t=1}^{N} z(t)\bar{y}^T(t) \tag{44}$$

$$z(t) = \begin{bmatrix} r(t-1) & \cdots & r(t-n-m) \end{bmatrix}^T \in \mathbb{R}^{n+m} \tag{45}$$

$$W_{iv} \ \text{ is an optimal weight} \tag{46}$$

and it is stated in [8] that

$$\hat{\theta}_{iv}(N) \mid_{W_{iv}=S^T S^{-1}} = \hat{\theta}_{bels}(N)$$

with $\hat{S} = \begin{bmatrix} 0 & I_{n+m} \end{bmatrix} \hat{R}_{\varphi\varphi}(N) M (M^T M)^{-1}$, assuming that it is non singular.

In the proof of this result the property is used that both weighted IV and BELS estimators can be written as

$$\hat{\theta}_{bels}(N) = \hat{\theta}_{iv}(N) = (\hat{K} - \hat{C}\hat{D}^{-1}\hat{F})\hat{p} \tag{47}$$

where

$$\hat{C} = (M^T M)^{-1} M^T \hat{R}^{-1}_{\varphi\varphi}(N) \begin{pmatrix} I_{n+m} \\ 0 \end{pmatrix} \in \mathbb{R}^{2n\times(n+m)} \tag{48}$$

$$\hat{K} = (M^T M)^{-1} M^T \hat{R}^{-1}_{\varphi\varphi}(N) \begin{pmatrix} 0 \\ I_{n+m} \end{pmatrix} \in \mathbb{R}^{2n\times(n+m)} \tag{49}$$

$$\hat{D} = H_1^T \hat{R}^{-1}_{\varphi\varphi}(N) \begin{pmatrix} I_{n+m} \\ 0 \end{pmatrix} \in \mathbb{R}^{(n+m)\times(n+m)} \tag{50}$$

$$\hat{F} = H_1^T \hat{R}^{-1}_{\varphi\varphi}(N) \begin{pmatrix} 0 \\ I_{n+m} \end{pmatrix} \in \mathbb{R}^{(n+m)\times(n+m)}. \tag{51}$$

However, equation (47) only holds true if matrix $\hat{D}$ is invertible (and thus square). This implies that matrix H_1 has dimensions $(2n+2m)\times(n+m)$. In the method of [9], as considered in this paper, the matrix H has either dimensions $(2n+2m)\times 2m$ (situation $m \geq n$) or $(3n+m)\times(n+m)$ (situation $m < n$). As these dimensions do not match with the dimension of H_1, the method developed in [8] gives consistent estimates but can not be associated with the BELS estimator considered here, neither in the $m \geq n$ case, nor in the $m < n$ case.

6 Conclusions

It has been shown that a bias-eliminated least-squares (BELS) estimator for closed-loop identification is equivalent to an instrumental variable estimator, where the predictor considered reflects the closed-loop system, and where external reference signals act as instrumental variables. This requires a tailor-made parametrization of the closed-loop system, as has been used in the literature before in a least squares setting. The relation between BELS and IV greatly facilitates the understanding and analysis of the former method.

References

[1] M. Gilson, P.M.J. Van den Hof. "On the relation between a bias-eliminated least-squares (BELS) and an IV estimator in closed-loop identification", *Automatica*, **37**, n° 10 (2001).

[2] I.D. Landau, A. Karimi. "An output error recursive algorithm for unbiased identification in closed loop", *Automatica*, **33**, n° 5, pp. 933–938, (1997).

[3] L. Ljung. "System identification, theory for the user", *Prentice Hall, Englewood Cliffs, NJ*, (1987).

[4] S. Sagara, K. Wada. "On-line modified least-squares parameter estimation of linear discrete dynamics systems", *Int. J. Control*, **25**, pp. 329–343, (1977).

[5] T. Söderström, W.X. Zheng and P. Stoica. "Comments on "On a least-squares-based algorithm for identification of stochastic linear systems"", *IEEE Trans. Signal Processing*, **47**, n° 5, pp. 1395–1396, (1999).

[6] P. Stoica, T. Söderström and V. Šimonytė. "Study of a bias-free least squares parameter estimator", *IEE Proc.-Control Theory Appl.*, **142**, n° 1, pp. 1–6, (1995).

[7] E.T. van Donkelaar, P.M.J. Van den Hof. "Analysis of closed-loop identification with a tailor-made parametrization", *Proc. 4th European Control Conference, Brussels, Belgium. Also in European J. Control*, **6**, *pp. 54–62, 2000*, (1997).

[8] Z. Zhang, C. Wen and Y.C. Soh. "Indirect closed-loop identification by optimal instrumental variable method", *Automatica*, **33**, n° 11, pp. 2269–2271, (1997).

[9] W.X. Zheng, C. Feng "A bias-correction method for indirect identification of closed-loop systems", *Automatica*, **31**, n° 7, pp. 1019–1024, (1995).

[10] W.X. Zheng "Identification of closed-loop systems with low-order controllers", *Automatica*, **32**, n° 12, pp. 1753–1757, (1996).

[11] W.X. Zheng "On a least-square based algorithm for identification of stochastic linear systems", *IEEE Trans. Signal Processing*, **46**, n° 6, pp. 1631–1638, (1998).

Appendix

Proof of Proposition 1. The equations that constitute $\hat{\Theta}_{bels}$ are collected in the following set of equations:

$$\hat{\Theta}_{bels} = M\hat{\theta}_{bels} + \rho \quad (52)$$

$$\hat{\theta}_{bels} = (M^TM)^{-1}M^T(\hat{\Theta}_{corr} - \rho) \quad (53)$$

$$\hat{\Theta}_{corr} = \hat{R}_{\varphi\varphi}^{-1}\hat{R}_{\varphi y} - \hat{\Delta} \quad (54)$$

$$\hat{\Delta} = \hat{R}_{\varphi\varphi}^{-1}\begin{bmatrix} I_{n+m} \\ 0 \end{bmatrix}\hat{R}_{y\xi} \quad (55)$$

$$H^T\hat{R}_{\varphi\varphi}^{-1}\begin{bmatrix} I_{n+m} \\ 0 \end{bmatrix}\hat{R}_{y\xi} = H^T(\hat{\Theta}_{ls} - \rho) \quad (56)$$

$$\hat{\Theta}_{ls} = \hat{R}_{\varphi\varphi}^{-1}\hat{R}_{\varphi y}, \quad (57)$$

where (56) can only be solved exactly if $r_B = 2n$. If $r_B > 2n$ this equation has to be solved in a least square sense. To this end we denote

$$\mu := H^T\hat{R}_{\varphi\varphi}^{-1}\begin{bmatrix} I_{n+m} \\ 0 \end{bmatrix}\hat{R}_{y\xi} - H^T(\hat{R}_{\varphi\varphi}^{-1}\hat{R}_{\varphi y} - \rho)$$

which by use of (55), (56) and (57) can be shown to satisfy

$$\mu = H^T(\hat{\Delta} - \hat{R}_{\varphi\varphi}^{-1}\hat{R}_{\varphi y} + \rho).$$

Minimization of $\mu^T\mu$ over Δ now has to be performed, under the additional constraint (55), which can be reformulated as $[0 \;\; I_{r_B}]\hat{R}_{\varphi\varphi}\hat{\Delta} = 0$, or equivalently

$$\hat{R}_{\varphi_r\varphi}\hat{\Delta} = 0. \quad (58)$$

Minimizing $\mu^T\mu$ under the constraint (58), is achieved by solving the corresponding Lagrangian equation, leading to

$$\begin{pmatrix} HH^T & \hat{R}_{\varphi_r\varphi}^T \\ \hat{R}_{\varphi_r\varphi} & 0 \end{pmatrix}\begin{pmatrix} \hat{\Delta} \\ \lambda \end{pmatrix} = \begin{pmatrix} HH^T(\hat{R}_{\varphi\varphi}^{-1}\hat{R}_{\varphi y} - \rho) \\ 0 \end{pmatrix} \quad (59)$$

where $\lambda \in \mathbb{R}^{r_B}$ is the Lagrange multiplier.

The left matrix is square with dimensions $(2r_B + m + n) \times (2r_B + m + n)$. However in its current form it can not be simply inverted. Therefore additional information has to be added by directly combining (59) by the remaining equations (52)-(54) of the general solution for $\hat{\Theta}_{bels}$. Premultiplication of (52) with the nonsingular matrix $[M \;\; H]^T$ then leads to

$$\begin{pmatrix} M^T & M^T & 0 \\ H^T & 0 & 0 \\ 0 & HH^T & \hat{R}_{\varphi_r\varphi}^T \\ 0 & \hat{R}_{\varphi_r\varphi} & 0 \end{pmatrix}\begin{pmatrix} \hat{\Theta}_{bels} \\ \hat{\Delta} \\ \lambda \end{pmatrix} = \begin{pmatrix} M^T\hat{R}_{\varphi\varphi}^{-1}\hat{R}_{\varphi y} \\ H^T\rho \\ HH^T(\hat{R}_{\varphi\varphi}^{-1}\hat{R}_{\varphi y} - \rho) \\ 0 \end{pmatrix}.$$

Then, the solution of this global system leads to

$$M^T\hat{R}_{\varphi_r\varphi}^T(\hat{R}_{\varphi_r\varphi}\hat{R}_{\varphi_r\varphi}^T)^{-1}\hat{R}_{\varphi_r\varepsilon(t,\hat{\theta}_{bels})} = 0$$

or, equivalently

$$\hat{R}_{F\varphi_r\varepsilon(t,\hat{\theta}_{bels})} = 0.$$

This proofs the result for $m \geq n$, corresponding to $r_B = n+m$.

In the situation $m \leq n$, corresponding to $r_B = 2n$, the matrix F is square and invertible under persistency of excitation conditions on r. As a result the set of equations can equivalently be characterized by $F = I_{r_B}$.