RECURSIVE BIAS-COMPENSATING ALGORITHM FOR THE IDENTIFICATION OF DYNAMICAL BILINEAR SYSTEMS IN THE ERRORS-IN-VARIABLES FRAMEWORK

T. Larkowski, J. G. Linden, B. Vinsonneau and K. J. Burnham
Control Theory and Applications Centre, Coventry University, Prior Street, Coventry, U.K.
larkowst@coventry.ac.uk

Keywords: Bias compensation, Bilinear systems, Errors-in-variables, Recursive estimation, Regularization, System identification.

Abstract: The paper investigates a recursive approach for the bias compensating least squares (BCLS) technique. The method presented is applied to the problem of on-line identification of single-input single-output bilinear models in the errors-in-variables framework. Within this framework the recursive bilinear BCLS algorithm is realized when a bilinear Frisch scheme (BFS) is iteratively applied for the estimation of the parameters of an exemplary bilinear system, giving rise to the exact recursive BFS (ERBFS) method. Moreover, a further extension of the ERBFS incorporating Tikhonov regularization with variable exponential weighting is considered and this is shown to be beneficial in the initial period of the identification procedure.

1 INTRODUCTION

The errors-in-variables (EIV) framework addresses the identification of dynamical systems where both the input and the output signals are corrupted by the measurement noise (Söderström, 2007). The EIV approaches are found to be of considerable benefit when the underlying physical laws characterizing the system are of a prime interest, as opposed to the prediction of the external signals (Söderström et al., 2002). In this case, the classical methods based on the least squares (LS) principle such as recursive LS (RLS) or the Kalman filter (Ikonen and Najim, 2002) are shown to yield estimates of the system parameters that are asymptotically biased and, therefore, inconsistent (Zheng, 1998; Söderström, 2007).

In the field of modelling for nonlinear systems, the bilinear system (BS) models have been used to advantage in various practical applications, e.g. control plants, biological and chemical phenomena, earth and sun science, nuclear fission, fault diagnosis and supervision, see (Mohler, 1991; Mohler and Khapalov, 2000) or (Ekman, 2005). Due mainly to the fact that BS models are so widely applicable has prompted the need to extend the EIV approaches developed for linear systems to encompass the BS case.

Recently, a technique for off-line compensation of the bias in the case of dynamical BS, i.e. the bilinear bias compensating LS (BBCLS) scheme has been proposed (Larkowski et al., 2007), upon which a bilinear Frisch scheme (BFS) has been constructed (Larkowski et al., 2008). The focus of this paper is the extension of the BBCLS along with the BFS for the purpose of on-line system identification. The proposed approach consists of a recursively performed update and bias compensation procedure for the data covariance matrices, whilst the BFS equations are applied in an iterative manner at each recursion step. Moreover, a further extension of the BFS incorporating the Tikhonov regularization (TR) technique with a variable exponential weighting is considered. It is shown via simulation studies that use of TR can be of considerable benefit in the initial period of the identification procedure.

The paper is organized as follows: in the second section the mathematical representation of the EIV BS together with the assumptions stated and the notation used are introduced. The third section presents a brief review of the BBCLS and the BFS techniques. In section four a recursive implementation of the BBCLS method is proposed. Subsequently, within the BBCLS framework the BFS technique is applied resulting in the exact recursive BFS (ERBFS) algorithm. The section ends with an extension of the ERBFS that incorporates the TR technique. Section five presents the results of a numerical simulation study involving the proposed algorithms, whilst the overall conclusions and the further work are summarized in section six.

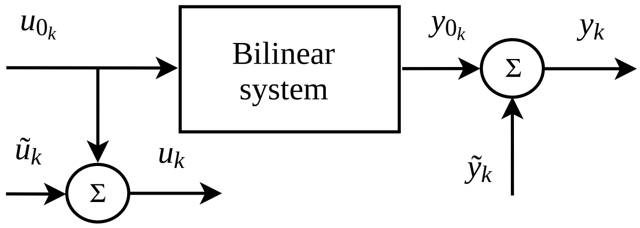

Figure 1: The basic setup for the EIV BS.

2 ASSUMPTIONS AND NOTATION

Consider a discrete time-invariant single-input single-output (SISO) class of BS that can be represented by the following input/output difference equation

$$A(q^{-1})y_{0_k} = B(q^{-1})u_{0_k} + \sum_{i=1}^{p}\sum_{j=1}^{r} \eta_{ij} u_{0_{k-i}} y_{0_{k-j}} \qquad (1)$$

with the polynomials $A(q^{-1})$ and $B(q^{-1})$ given by

$$A(q^{-1}) \triangleq 1 + a_1 q^{-1} + \ldots + a_{n_a} q^{-n_a} \qquad (2a)$$

$$B(q^{-1}) \triangleq b_1 q^{-1} + \ldots + b_{n_b} q^{-n_b} \qquad (2b)$$

where $r \leq n_a$, $p \leq n_b \leq n_a$, q^{-1} is the backward shift operator, defined by $x_k q^{-1} \triangleq x_{k-1}$ and u_{0_k}, y_{0_k} are the noise-free input and output sequences, respectively. A diagrammatic illustration of the typical EIV setup for a SISO BS is depicted in Figure 1.

The BS can be classified into three main categories, see (Pearson, 1999) for more details, namely: $a)$ subdiagonal $\eta_{ij} = 0\ \forall j > i$; $b)$ diagonal $\eta_{ij} = 0\ \forall j \neq i$; and $c)$ superdiagonal $\eta_{ij} = 0\ \forall j < i$. Noting that both the subdiagonal and superdiagonal cases include the diagonal case, reference will be solely made here to the diagonal BS (DBS) case for the remainder of the paper. This is due to the fact that DBS exhibit some crucial properties of interest, see (Rao and Gabr, 1984; Liu, 1992) or (Kotta and Nomm, 2003) for a detailed discussion. At the same time DBS are possibly the most commonly utilized class of BS for the purpose of industrial applications, see (Burnham, 1991; Yu, 1996; Martineau et al., 2004).

Without loss of generality, the case when all the diagonal terms in the system (1) are present is considered here with their number given as $n_\eta = p^2$ where $r = p$. The following assumptions are introduced:

A1. The DBS is time-invariant, asymptotically stable, strictly stationary, observable and controllable.

A2. The system structure, i.e. n_a, n_b, p, is known *a priori*.

A3. The true input $u_{0_k} \sim \mathcal{N}(0, \sigma_{u_0})$ is white, persistently exciting and of sufficiently high order.

A4. Corrupting input/output noises $\tilde{u}_k \sim \mathcal{N}(0, \sigma_{\tilde{u}})$ and $\tilde{y}_k \sim \mathcal{N}(0, \sigma_{\tilde{y}})$ of unknown variances are additive, white, mutually uncorrelated and uncorrelated with the noise free signals u_{0_k} and y_{0_k}, respectively.

Acknowledging A4, it is implied that the measured input and output can be decomposed into their noise free and noise contributions, i.e.

$$u_k \triangleq u_{0_k} + \tilde{u}_k \qquad (3a)$$

$$y_k \triangleq y_{0_k} + \tilde{y}_k \qquad (3b)$$

where k denotes the discrete time index. The identification problem consists of determining the vector

$$\vartheta^T \triangleq \begin{bmatrix} \theta^T & \sigma_{\tilde{u}} & \sigma_{\tilde{y}} \end{bmatrix} \in \mathcal{R}^{n_\theta+2} \qquad (4)$$

where $\theta \in \mathcal{R}^{n_\theta}$ is the parameter vector with $n_\theta = n_a + n_b + n_\eta$ and $\sigma_{\tilde{u}}$, $\sigma_{\tilde{y}}$ are the input and output noise variances, respectively. The parameter vector is defined as:

$$\theta \triangleq \begin{bmatrix} a \\ b \\ \eta \end{bmatrix} \quad a \triangleq \begin{bmatrix} a_1 \\ \vdots \\ a_{n_a} \end{bmatrix} \quad b \triangleq \begin{bmatrix} b_1 \\ \vdots \\ b_{n_b} \end{bmatrix} \quad \eta \triangleq \begin{bmatrix} \eta_{11} \\ \vdots \\ \eta_{pp} \end{bmatrix} \qquad (5)$$

with $a \in \mathcal{R}^{n_a}$, $b \in \mathcal{R}^{n_b}$, $\eta \in \mathcal{R}^{n_\eta}$ where the extended parameter vector $\bar{\theta}$ is given by

$$\bar{\theta} \triangleq \begin{bmatrix} \bar{a} \\ b \\ \eta \end{bmatrix} \in \mathcal{R}^{n_\theta+1} \qquad \bar{a} \triangleq \begin{bmatrix} 1 \\ a \end{bmatrix} \in \mathcal{R}^{n_a+1} \qquad (6)$$

The regressor vectors for the measured data, noise-free data and noise are defined, respectively, as:

$$\varphi_k \triangleq \begin{bmatrix} \varphi_{y_k} \\ \varphi_{u_k} \\ \varphi_{\rho_k} \end{bmatrix} \qquad \varphi_{0_k} \triangleq \begin{bmatrix} \varphi_{y_{0,k}} \\ \varphi_{u_{0,k}} \\ \varphi_{\rho_{0,k}} \end{bmatrix} \qquad \tilde{\varphi}_k \triangleq \begin{bmatrix} \tilde{\varphi}_{y_k} \\ \tilde{\varphi}_{u_k} \\ \tilde{\varphi}_{\rho_k} \end{bmatrix} \qquad (7)$$

where

$$\varphi_{y_k} \triangleq \begin{bmatrix} -y_{k-1} \\ \vdots \\ -y_{k-n_a} \end{bmatrix} \quad \varphi_{u_k} \triangleq \begin{bmatrix} u_{k-1} \\ \vdots \\ u_{k-n_b} \end{bmatrix} \quad \varphi_{\rho_k} \triangleq \begin{bmatrix} y_{k-1}u_{k-1} \\ \vdots \\ y_{k-p}u_{k-p} \end{bmatrix}$$

$$\varphi_{y_{0,k}} \triangleq \begin{bmatrix} -y_{0_{k-1}} \\ \vdots \\ -y_{0_{k-n_a}} \end{bmatrix} \varphi_{u_{0,k}} \triangleq \begin{bmatrix} u_{0_{k-1}} \\ \vdots \\ u_{0_{k-n_b}} \end{bmatrix} \varphi_{\rho_{0,k}} \triangleq \begin{bmatrix} y_{0_{k-1}} u_{0_{k-1}} \\ \vdots \\ y_{0_{k-p}} u_{0_{k-p}} \end{bmatrix}$$

$$\tilde{\varphi}_{y_k} \triangleq \begin{bmatrix} -\tilde{y}_{k-1} \\ \vdots \\ \tilde{y}_{k-n_a} \end{bmatrix} \quad \tilde{\varphi}_{u_k}^T \triangleq \begin{bmatrix} \tilde{u}_{k-1} \\ \vdots \\ \tilde{u}_{k-n_b} \end{bmatrix} \quad \tilde{\varphi}_{\rho_k}^T \triangleq \begin{bmatrix} \tilde{\rho}_{k-1,k-1} \\ \vdots \\ \tilde{\rho}_{k-p,k-p} \end{bmatrix}$$

with φ_k, φ_{0_k}, $\tilde{\varphi}_k \in \mathcal{R}^{n_\theta}$, φ_{y_k}, $\varphi_{y_{0,k}}$, $\tilde{\varphi}_{y_k} \in \mathcal{R}^{n_a}$, φ_{u_k}, $\varphi_{u_{0,k}}$, $\tilde{\varphi}_{u_k} \in \mathcal{R}^{n_b}$, φ_{ρ_k}, $\varphi_{\rho_{0,k}}$, $\tilde{\varphi}_{\rho_k} \in \mathcal{R}^{n_\eta}$ and $\tilde{\rho}_{k-i,k-j}$ denoting the noise contribution corresponding to the bilinear product terms of the regressor vector φ_{ρ_k}. In agreement with (7), the extended regressor vectors are given by

$$\bar{\varphi}_k \triangleq \begin{bmatrix} -y_k \\ \varphi_k \end{bmatrix} \qquad \bar{\varphi}_{0_k} \triangleq \begin{bmatrix} -y_{0_k} \\ \varphi_{0_k} \end{bmatrix} \qquad \bar{\tilde{\varphi}}_k \triangleq \begin{bmatrix} -\tilde{y}_k \\ \tilde{\varphi}_k \end{bmatrix} \quad (8)$$

where $\bar{\varphi}_k$, $\bar{\varphi}_{0_k}$, $\bar{\tilde{\varphi}}_k \in \mathcal{R}^{n_\theta+1}$.

3 A BRIEF REVIEW OF BBCLS AND BFS

3.1 BBCLS

The BBCLS algorithm for the class of DBS comprises of equations (9), (10) and (11), see (Larkowski et al., 2007). These correspond to the bilinear bias compensation rule, the noise covariance matrix and the noise 'variance' of the bilinear terms, respectively, i.e.

$$\hat{\theta}^{BBCLS} \triangleq \left(\Sigma_{\varphi\varphi} - \Sigma_{\tilde{\varphi}\tilde{\varphi}}\right)^{-1} \Sigma_{\varphi y} \quad (9)$$

$$\Sigma_{\tilde{\varphi}\tilde{\varphi}} \triangleq \begin{bmatrix} \sigma_{\tilde{y}} I_{n_a} & 0 & 0 \\ 0 & \sigma_{\tilde{u}} I_{n_b} & 0 \\ 0 & 0 & \sigma_{\tilde{\rho}} I_{n_\eta} \end{bmatrix} \quad (10)$$

$$\sigma_{\tilde{\rho}} \triangleq \sigma_u \sigma_{\tilde{y}} + \sigma_y \sigma_{\tilde{u}} - \sigma_{\tilde{u}} \sigma_{\tilde{y}} \quad (11)$$

where $\sigma_u \triangleq E[u_k]$ and $\sigma_y \triangleq E[y_k]$ are the expected values of the measured system input and output signals and $(\hat{\cdot})$ denotes an estimate. It is noted, that in the remainder of the paper, the expression Σ_{ab} will be used as general notation for the correlation matrix of vectors a_k and b_k, i.e. $\Sigma_{ab} = E[a_k b_k^T]$. Equation (9) can be alternatively restated as

$$\hat{\theta}^{BBCLS} = \hat{\theta}^{LS} + \Sigma_{\varphi\varphi}^{-1} \Sigma_{\tilde{\varphi}\tilde{\varphi}} \hat{\theta}^{BBCLS} \quad (12)$$

where $\hat{\theta}^{LS}$ denotes the LS estimate. It is implied from the BBCLS algorithm that the knowledge regarding noise variances corrupting input/output of a system together with variances of measured input/output signals is sufficient to obtain unbiased estimates of the true system parameters.

3.2 BFS

The Frisch scheme is a technique that allows the direct estimation of the input/output noise variances together with the parameters of a system (Beghelli et al., 1990; Söderström, 2006). As consequence, the *a priori* knowledge regarding the values of $\sigma_{\tilde{u}}$ and $\sigma_{\tilde{y}}$ is not required leading to a wider practical applicability. The extension of the FS, in the framework of the BBCLS technique, for the class of DBS has been proposed in (Larkowski et al., 2008). Define the partitioned extended data covariance matrix

$$\Sigma_{\bar{\varphi}\bar{\varphi}} \triangleq \begin{bmatrix} \Sigma_{\bar{\varphi}_y\bar{\varphi}_y} & \Sigma_{\varphi_u\bar{\varphi}_y}^T & \Sigma_{\varphi_\rho\bar{\varphi}_y}^T \\ \Sigma_{\varphi_u\bar{\varphi}_y} & \Sigma_{\varphi_u\varphi_u} & \Sigma_{\varphi_\rho\varphi_u}^T \\ \Sigma_{\varphi_\rho\bar{\varphi}_y} & \Sigma_{\varphi_\rho\varphi_u} & \Sigma_{\varphi_\rho\varphi_\rho} \end{bmatrix} \quad (13)$$

where $\Sigma_{\bar{\varphi}_y\bar{\varphi}_y} \in \mathcal{R}^{(n_a+1)\times(n_a+1)}$, $\Sigma_{\varphi_u\bar{\varphi}_y} \in \mathcal{R}^{n_b\times(n_a+1)}$, $\Sigma_{\varphi_u\varphi_u} \in \mathcal{R}^{n_b\times n_b}$, $\Sigma_{\varphi_\rho\bar{\varphi}_y} \in \mathcal{R}^{n_\eta\times(n_a+1)}$, $\Sigma_{\varphi_\rho\varphi_u} \in \mathcal{R}^{n_\eta\times n_b}$ and $\Sigma_{\varphi_\rho\varphi_\rho} \in \mathcal{R}^{n_\eta\times n_\eta}$. The BFS consists of three main phases, i.e. calculation of the maximal admissible value for $\sigma_{\tilde{u}}$, denoted $\sigma_{\tilde{u}}^{max}$ (14), determination of a functional relationship between $\sigma_{\tilde{y}}$ and $\sigma_{\tilde{u}}$ (16a) and the specification of a cost function to find a unique value of $\sigma_{\tilde{u}}$ (18a). The quantity $\sigma_{\tilde{u}}^{max}$ is given by:

$$\sigma_{\tilde{u}}^{max} \triangleq \lambda_{\min}(A_1^*) \quad (14)$$

where $\lambda_{\min}(A_1^*)$ denotes the least eigenvalue of the matrix A_1^*, and $\max(\cdot)$ is the maximum operator. The matrix A_1^* is defined as:

$$A_1^* \triangleq A_1 - B_1 \Sigma_{\bar{\varphi}_y\bar{\varphi}_y}^{-1} B_1^T \quad (15a)$$

with

$$A_1 \triangleq \begin{bmatrix} \Sigma_{\varphi_u\varphi_u} & \Sigma_{\varphi_\rho\varphi_u}^T \\ \Sigma_{\varphi_\rho\varphi_u} & \Sigma_{\varphi_\rho\varphi_\rho} \end{bmatrix} \qquad B_1 \triangleq \begin{bmatrix} \Sigma_{\varphi_u\bar{\varphi}_y} \\ \Sigma_{\varphi_\rho\bar{\varphi}_y} \end{bmatrix} \quad (15b)$$

The functional relationship between $\sigma_{\tilde{y}}$ and $\sigma_{\tilde{u}}$ is described by

$$\sigma_{\tilde{y}} \triangleq \lambda_{\min}(A_2^*) \quad (16a)$$

where

$$A_2^* \triangleq A_2 - B_2 (\Sigma_{\varphi_u\varphi_u} - \sigma_{\tilde{u}} I_{n_b})^{-1} B_2^T \quad (16b)$$

with

$$A_2 \triangleq \begin{bmatrix} \Sigma_{\bar{\varphi}_y\bar{\varphi}_y} & \Sigma_{\varphi_\rho\bar{\varphi}_y}^T \\ \Sigma_{\varphi_\rho\bar{\varphi}_y} & \Sigma_{\varphi_\rho\varphi_\rho} - \sigma_y \sigma_{\tilde{u}} I_{n_\eta} \end{bmatrix} \quad B_2 \triangleq \begin{bmatrix} \Sigma_{\varphi_u\bar{\varphi}_y}^T \\ \Sigma_{\varphi_\rho\varphi_u} \end{bmatrix} \quad (16c)$$

The cost function utilized is based on the Yule-Walker equations, see (Diversi et al., 2006) for details. The instrumental vector (Söderström and Stoica, 1994) for the measured data is defined as:

$$\bar{\varphi}_k^{IV} \triangleq \bar{\varphi}_{k-n_a-1} \in \mathcal{R}^{n_{\theta+1}\times 1} \quad (17)$$

Using (17) the corresponding cost function is formulated as:

$$J(\hat{\bar{\theta}}) \triangleq \|\Sigma_{\bar{\varphi}^{IV}\bar{\varphi}} \hat{\bar{\theta}}\|_2^2 \quad (18a)$$

such that

$$J(\hat{\bar{\theta}}) = 0 \Leftrightarrow \hat{\bar{\theta}} = \bar{\theta} \quad (18b)$$

Table 1: Summary of the ERBFS algorithm.

Step	Description	Procedure
1	Choose λ_k and j	$0 < \lambda_k < 1,\ j = 2n_a + 1$
2	RLS initialization:	$\hat{\theta}^{LS}_{n_\theta} = 0,\ P_{n_\theta} = 10^3 I_{n_\theta},\ \hat{\sigma}^k_u = 0,\ \hat{\sigma}^k_y = 0$
2.1	RLS loop start	**for** $k = n_\theta + 1 \dots j$
2.2	Data weighting	$\gamma_k = 1/k$
	Computation of:	
2.3	L_k, $\hat{\theta}^{LS}_k$, P_k	$L_k = \frac{P_{k-1}\varphi_k}{\varphi_k^T P_{k-1}\varphi_k + \frac{1-\gamma_k}{\gamma_k}},\ \hat{\theta}^{LS}_k = \hat{\theta}^{LS}_{k-1} + L_k\left(y_k - \varphi_k^T \hat{\theta}^{LS}_{k-1}\right)$
		$P_k = \frac{1}{1-\gamma_k}\left(P_{k-1} - L_k \varphi_k^T P_{k-1}\right)$
2.4	$\hat{\sigma}^k_u$, $\hat{\sigma}^k_y$	$\hat{\sigma}^k_u = \frac{k-1}{k}\hat{\sigma}^{k-1}_u + \frac{1}{k-1}u_k^2,\ \hat{\sigma}^k_y = \frac{k-1}{k}\hat{\sigma}^{k-1}_y + \frac{1}{k-1}y_k^2$
2.5	$\Sigma^k_{\varphi_\rho \bar{\varphi}_y}$	$\Sigma^k_{\bar{\varphi}\bar{\varphi}} = \Sigma^{k-1}_{\bar{\varphi}\bar{\varphi}} + \gamma_k\left(\bar{\varphi}_k \bar{\varphi}_k^T - \Sigma^{k-1}_{\bar{\varphi}\bar{\varphi}}\right)$
2.6	RLS loop end	**end**
3	BBCLS and BFS initialization	$\Sigma^k_{\bar{\varphi}^{IV}\bar{\varphi}} = 0,\ \hat{\sigma}^{max}_{\tilde{u}} = \lambda_{\min}(A^*_{1,j})$
3.1	Recursive BBCLS start	**for** $k = j + 1 \dots N$
3.2	Data weighting	$\gamma_k = 1/k$
3.3	Iterative BFS start	
	Computation of:	
3.3.1	$\Sigma^k_{\bar{\varphi}^{IV}\bar{\varphi}}$	$\Sigma^k_{\bar{\varphi}^{IV}\bar{\varphi}} = \Sigma^{k-1}_{\bar{\varphi}^{IV}\bar{\varphi}} + \gamma_k\left(\bar{\varphi}^{IV}_k \bar{\varphi}^T_k - \Sigma^{k-1}_{\bar{\varphi}^{IV}\bar{\varphi}}\right)$
3.3.2	$\hat{\sigma}^k_{\tilde{u}}$	$\hat{\sigma}^k_{\tilde{u}} = \arg\min_{\hat{\sigma}^k_{\tilde{u}}} J(\hat{\bar{\theta}}_k)$
3.3.3	$\hat{\sigma}^k_u$, $\hat{\sigma}^k_y$	$\hat{\sigma}^k_u = \frac{k-1}{k}\hat{\sigma}^{k-1}_u + \frac{1}{k-1}u_k^2,\ \hat{\sigma}^k_y = \frac{k-1}{k}\hat{\sigma}^{k-1}_y + \frac{1}{k-1}y_k^2$
3.3.4	$\Sigma^k_{\bar{\varphi}\bar{\varphi}}$	$\Sigma^k_{\bar{\varphi}\bar{\varphi}} = \Sigma^{k-1}_{\bar{\varphi}\bar{\varphi}} + \gamma_k\left(\bar{\varphi}_k \bar{\varphi}_k^T - \Sigma^{k-1}_{\bar{\varphi}\bar{\varphi}}\right)$
3.3.5	A^k_2, B^k_2	$A^k_2 = \begin{bmatrix} \Sigma^k_{\bar{\varphi}_y\bar{\varphi}_y} & (\Sigma^k_{\varphi_\rho\bar{\varphi}_y})^T \\ \Sigma^k_{\varphi_\rho\bar{\varphi}_y} & \Sigma^T_{\varphi_\rho\varphi_\rho} - \hat{\sigma}^k_y \hat{\sigma}^k_{\tilde{u}} I_{n_\eta} \end{bmatrix},\ B^k_2 = \begin{bmatrix} (\Sigma^k_{\varphi_u\bar{\varphi}_y})^T \\ \Sigma^k_{\varphi_\rho\varphi_u} \end{bmatrix}$
3.3.6	$A^*_{2,k}$	$A^*_{2,k} = A^k_2 - B^k_2(\Sigma^k_{\varphi_u\varphi_u} - \hat{\sigma}^k_{\tilde{u}} I_{n_b})^{-1}(B^k_2)^T$
3.3.7	$\hat{\sigma}^k_{\tilde{y}}$	$\hat{\sigma}^k_{\tilde{y}} = \lambda_{\min}(A^*_{2,k})$
3.4	Iterative BFS end	
3.5	L_k, $\hat{\theta}^{LS}_k$, P_k	$L_k = \frac{P_{k-1}\varphi_k}{\varphi_k^T P_{k-1}\varphi_k + \frac{1-\gamma_k}{\gamma_k}},\ \hat{\theta}^{LS}_k = \hat{\theta}^{LS}_{k-1} + L_k\left(y_k - \varphi_k^T \hat{\theta}^{LS}_{k-1}\right)$
		$P_k = \frac{1}{1-\gamma_k}\left(P_{k-1} - L_k \varphi_k^T P_{k-1}\right)$
3.6	$\hat{\sigma}^k_{\tilde{\rho}}$	$\hat{\sigma}^k_{\tilde{\rho}} = \hat{\sigma}^k_u \hat{\sigma}^k_{\tilde{y}} + \hat{\sigma}^k_y \hat{\sigma}^k_{\tilde{u}} - \hat{\sigma}^k_{\tilde{u}} \hat{\sigma}^k_{\tilde{y}}$
3.7	$\Sigma^k_{\tilde{\varphi}\tilde{\varphi}}$	$\Sigma^k_{\tilde{\varphi}\tilde{\varphi}} = \begin{bmatrix} \hat{\sigma}^k_{\tilde{y}} I_{n_a} & 0 & 0 \\ 0 & \hat{\sigma}^k_{\tilde{u}} I_{n_b} & 0 \\ 0 & 0 & \hat{\sigma}^k_{\tilde{\rho}} I_{n_\eta} \end{bmatrix}$
3.8	Bias compensation	$\hat{\theta}^{BBCLS}_k = \hat{\theta}^{LS}_k + P_k \Sigma^k_{\tilde{\varphi}\tilde{\varphi}} \hat{\theta}^{BBCLS}_{k-1}$
3.9	Recursive BBCLS end	**end**

4 RECURSIVE BBCLS WITH ITERATIVE BFS

In this section a recursive BBCLS (RBBCLS) algorithm is developed, which comprises the recursive update of the data and instrumental covariance matrices, whilst the bias compensation procedure is recursively applied to the data covariance matrix only. Furthermore, an application of the BFS at each iteration step is described with an additional extension incorporating the TR technique. It is to be noted that the RBBCLS method can be interpreted in the framework of the iterative bias compensating LS (BCLS) approaches, see e.g. (Zheng, 1998) and (Zheng, 2000) for more details.

4.1 Recursive BBCLS

The normalized recursive updates of the instrumental and data covariance matrices are given by the following equations, see (Ljung, 1999)

$$\Sigma^{k}_{\bar{\varphi}\bar{\varphi}} = \Sigma^{k-1}_{\bar{\varphi}\bar{\varphi}} + \gamma_k \left(\bar{\varphi}_k \bar{\varphi}_k^T - \Sigma^{k-1}_{\bar{\varphi}\bar{\varphi}}\right) \tag{19a}$$

$$\Sigma^{k}_{\bar{\varphi}^{IV}\bar{\varphi}} = \Sigma^{k-1}_{\bar{\varphi}^{IV}\bar{\varphi}} + \gamma_k \left(\bar{\varphi}_k^{IV} \bar{\varphi}_k^T - \Sigma^{k-1}_{\bar{\varphi}^{IV}\bar{\varphi}}\right) \tag{19b}$$

with

$$\gamma_k = \left(\Sigma_{i=1}^{k}\beta_{k,i}\right)^{-1} = \frac{\gamma_{k-1}}{\lambda_k + \gamma_{k-1}} \tag{20}$$

where the i-th data is weighted at the discrete time k according to the following rule

$$\beta_{k,i} = \lambda_k \beta_{k-1,i} \qquad \text{for} \qquad 0 \le i \le k-1 \tag{21}$$

and $\beta_{k,k} = 1$. It is to be noted that (20) simplifies either to $1/k$ in the case of no adaptivity, i.e. when $\lambda_k = 1$ or to $1-\lambda$ when the exponential forgetting is used, i.e. $\lambda_k = \lambda$ with $0 < \lambda < 1$.

Assuming that the input/output noise variances, i.e. $\sigma_{\tilde{u}}$ and $\sigma_{\tilde{y}}$ are known *a priori* or can be estimated, allows application of the BBCLS algorithm to the recursively computed estimate of the parameter vector, i.e.

$$\hat{\theta}_k^{BBCLS} = \hat{\theta}_k^{LS} + P_k \Sigma_{\tilde{\varphi}\tilde{\varphi}} \hat{\theta}_{k-1}^{BBCLS} \tag{22a}$$

$$\Sigma_{\tilde{\varphi}\tilde{\varphi}} = \begin{bmatrix} \sigma_{\tilde{y}} I_{n_a} & 0 & 0 \\ 0 & \sigma_{\tilde{u}} I_{n_b} & 0 \\ 0 & 0 & \hat{\sigma}^k_{\tilde{\rho}} I_{n_\eta} \end{bmatrix} \tag{22b}$$

$$\hat{\sigma}^k_{\tilde{\rho}} = \hat{\sigma}^k_u \sigma_{\tilde{y}} + \hat{\sigma}^k_y \sigma_{\tilde{u}} - \sigma_{\tilde{u}}\sigma_{\tilde{y}} \tag{22c}$$

$$\hat{\sigma}^k_u = \frac{k-1}{k}\hat{\sigma}^{k-1}_u + \frac{1}{k-1}u_k^2 \tag{22d}$$

$$\hat{\sigma}^k_y = \frac{k-1}{k}\hat{\sigma}^{k-1}_y + \frac{1}{k-1}y_k^2 \tag{22e}$$

$$L_k = \frac{P_{k-1}\varphi_k}{\varphi_k^T P_{k-1}\varphi_k + \frac{1-\gamma_k}{\gamma_k}} \tag{22f}$$

$$\hat{\theta}_k^{LS} = \hat{\theta}_{k-1}^{LS} + L_k\left(y_k - \varphi_k^T \hat{\theta}_{k-1}^{LS}\right) \tag{22g}$$

$$P_k = \frac{1}{1-\gamma_k}\left(P_{k-1} - L_k \varphi_k^T P_{k-1}\right) \tag{22h}$$

It is remarked that whilst the input/output noise variances are postulated to be known, the noise 'variance' corresponding to the bilinear terms (22c) is required to be recursively approximated at each time step. This involves the recursive estimation of the variances of the input/output signals, i.e. (22d) and (22e) (Young, 1984). Note that due to assumptions A1 and A3, see (Pearson, 1999) for more details, the mean values of the input/output signals do not explicitly appear in expressions (22d) and (22e) since they are both null.

4.2 Iterative BFS

Utilization of the recursively evaluated data and instrumental covariance matrices from the RBBCLS algorithm allows the application of the BFS at each recursion. Thus it is possible not only to estimate the input/output noise variances but also to conduct the noise compensation procedure given by (22a), see (Linden et al., 2007) for more details regarding the linear case. This results in the ERBFS algorithm which is summarized in Table 1. It is to be noted that ERBFS is rather expensive from the computational point of view. However, with reference to Table 1, if the time allowed for the calculation of $\hat{\sigma}^k_{\tilde{u}}$ at step 3.3.2 is bounded, the algorithm at least satisfies the principles of a recursive estimation scheme (Ljung and Söderström, 1987; Ljung, 1999).

4.3 Regularized BFS

In the case when *a priori* knowledge regarding the value of the input noise variance is available, or can be approximately anticipated, a regularization technique may be utilized. The regularization method considered here is that of TR which forces the estimate towards a pre-specified value $\hat{\sigma}^*_{\tilde{u}}$ controlled by the parameter ω (Hansen, 2001). The incorporation of TR into the cost function given by (18a) results in the following regularized cost function

$$J(\breve{\theta}, \omega, \hat{\sigma}^*_{\tilde{u}}) \triangleq \omega \|\Sigma_{\bar{\varphi}^{IV}\bar{\varphi}}\breve{\theta}\|_2^2 + (1-\omega)\|\hat{\sigma}^*_{\tilde{u}} - \hat{\sigma}_{\tilde{u}}\|_2^2 \tag{23}$$

Note that (23) reduces to (18a) for $\omega = 1$. Furthermore, it may be beneficial to consider a variable controlling parameter, i.e. ω_k, such that the impact of the regularization is significant at the beginning of the identification procedure but gradually diminishes as a function of the incoming data stream. It is proposed to realize the concept as follows

$$\omega_k = e^{-\frac{\varsigma}{k}} \tag{24}$$

where ς is a user defined parameter describing the rate at which the impact of the regularization diminishes. This choice allows the potential bias introduced by regularization (Hansen, 2001) to be alleviated as time progresses. With reference to Table 1, the introduction of regularization requires the additional setting of the parameter ς at step 3, the subsequent implementation of equation (24) between steps 3.3.1 and 3.3.2 and replacement of the cost function $J_k(\breve{\theta}_k)$ from 3.3.2 by $J(\breve{\theta}_k, \omega_k, \hat{\sigma}^*_{\tilde{u}})$.

5 SIMULATION STUDIES

This section provides a numerical evaluation and comparison of the proposed ERBFS algorithm with the RLS and the off-line BFS. The SISO DBS system used in the simulations with $n_a = 2$, $n_b = n_\eta = 1$ is simulated for $N = 5000$. It is described by the following difference equation

$$y_{0_k} = 1.2y_{0_{k-2}} - 0.9y_{0_{k-1}} + 0.6u_{0_{k-1}} + 0.1y_{0_{k-1}}u_{0_{k-1}} \tag{25}$$

The input is generated by

$$u_{0_k} \sim \mathcal{N}(0, 0.5) \tag{26}$$

The variances of the input and output noises are selected as $\sigma_{\tilde{u}} = 0.05$ and $\sigma_{\tilde{y}} = 0.16$, respectively, in order to yield an approximately equal signal-to-noise ratio (SNR) on both input and output, i.e. $\text{SNR}_u \approx \text{SNR}_y \approx 10$[dB]. In the case of the ERBFS and the regularized ERBFS (RERBFS) the minimization procedure from step 3.3.2 (see Table 1) is restricted to a maximum of 10 iterations. The parameter λ_k is set to unity, i.e. no adaptivity, for all evaluated algorithms.

Considering the results presented in Figure 2, where the ERBFS is compared with its off-line counterpart and the RLS, the following observations are made:

a) RLS yields estimates that are asymptotically biased for the case of all system parameters.

b) The estimate of the vector ϑ obtained via the off-line BFS is quite close to its true value.

c) The EBFS achieved virtually identical estimates as the off-line BFS at the last recursion step, i.e. $k = N$.

d) Estimates given by EBFS converge to their off-line counterparts obtained by the BFS algorithm over the succesive recursions.

e) There is a clear correlation between the quality of the estimated variances of the input/output signals and the quality of the estimated input/output noise variances which, subsequently, has an impact on the estimates of the system parameters.

f) The ERBFS encountered some difficulties in the estimation of the input noise variance in the initial part of the identification procedure, i.e. up to about first 1000 samples which is indicated by the relatively highly scattered values of $\hat{\sigma}_{\tilde{u}}$.

In fact observation (f) can be regarded as a premise for considering regularization of the input noise variance such that the uncertainty in the initial part of the identification is alleviated, leading to improved accuracy of the estimates. In the second experiment the ERBFS is compared with the RERBFS, where the parameters are set as follows: $\varsigma = N/\omega_0$ where $\omega_0 = 100$ and $\sigma^*_{\tilde{u}} = 1.5\sigma_{\tilde{u}}$. For completeness, the results obtained by the BFS are also included. Consideration of the results in Figure 3 leads to the following observations:

g) Although the guess of the regularization parameter $\sigma^*_{\tilde{u}}$ was rather 'rough', a substantial improvement w.r.t. the input noise variance is observed in the initial period of the recursion.

h) The impact of applying TR is also evident in the case of the estimated output noise variance and the system parameters leading to the faster convergence.

i) Due to the utilization of the exponential controlling variable weighting ω_k the results obtained by the RERBFS and ERBFS at $k = N$ are practically indistinguishable. As a consequence, any potential induced bias due to the use of regularization is kept to a minimum, for the case considered.

6 CONCLUSIONS

A new recursive technique, i.e. the RBBCLS method, for identification of the class of SISO DBS has been developed and evaluated. Within the RBBCLS framework the ERBFS algorithm has been formulated in which the Frisch equations are evaluated at each recursion. The further extension incorporating the variable regularization via the TR method, giving rise to the RERBFS, was considered and shown to be beneficial in the initial period of the identification procedure. The methods proposed have been demonstrated when applied to a SISO DBS EIV identification problem. Comparisons made with the standard RLS technique illustrates the superiority and relatively high noise robustness of the proposed algorithms.

The further work will address two outstanding issues. Firstly, the extension and subsequent recursive implementation of the BCLS method to a wider class of the polynomial nonlinear EIV systems. Secondly, the alleviation of the computational burden via the application of a fully recursive BFS based on gradient approaches and/or on other fast methods via linearization of the Frisch equations.

REFERENCES

Beghelli, S., Guidorzi, R. P., and Soverini, U. (1990). The Frisch scheme in dynamic system identification. *Automatica*, 26(1):171–176.

Burnham, K. J. (1991). *Self-tuning Control for Bilinear Systems*. PhD thesis, Coventry Polytechnic.

Diversi, R., Guidorzi, R., and Soverini, U. (2006). Yule-Walker equations in the Frisch scheme solution of errors-in-variables identification problems. In *Proc. of the 17th Int. Symposium on Mathematical Theory of Networks and Systems*, Kyoto, Japan.

Ekman, M. (2005). *Modeling and Control of Bilinear Systems: Applications to the Activated Sludge Process*. PhD thesis, Uppsala University.

Hansen, P. C. (2001). Regularization tools: A matlab package for analysis and solution of discrete ill-posed problems. Technical report, Department of Mathematical Modelling, Technical University of Denmark.

Ikonen, E. and Najim, K. (2002). *Advanced Process Identification and Control*. Marcel Dekker, Inc., USA.

Kotta, U. and Nomm, S.and Zinober, A. (2003). On state space realizability of bilinear systems described by higher order difference equations. In *Proc. of 42nd IEEE Conf. on Decision and Control*, volume 6, pages 5685–5690.

Larkowski, T., Linden, J. G., Vinsonneau, B., and Burnham, K. J. (2008). A novel errors-in-variables approach for bilinear models: the bilinear Frisch scheme. Internal report no. CTAC/TL-1/2008, Control Theory and Applications Centre, Coventry University, Coventry.

Larkowski, T., Vinsonneau, B., and Burnham, K. J. (2007). Bilinear model identification in the errors-in-variables framework via the bias-compensating least squares. In *IAR and ACD Int. Conf.*, Grenoble, France.

Linden, G. J., Vinsonneau, B., and Burnham, K. J. (2007). Fast algorithms for recursive Frisch scheme system identification. In *IAR and ACD Int. Conf.*, Grenoble, France.

Liu, J. (1992). On stationarity and asymptotic inference of bilinear time series models. *Statistica Sinica*, 2:479–494.

Ljung, L. (1999). *System Identification - Theory for the User*. Prentice Hall PTR, New Jersey, USA, 2nd edition edition.

Ljung, L. and Söderström, T. (1987). *Theory and practice of recursive identification*. MIT Press, Cambridge, UK.

Martineau, S., Burnham, K. J., Haas, O. C. L., Andrews, G., and Heeley, A. (2004). Four-term bilinear pid controller applied to an industrial furnace. *Control Engineering Practice*, 12(4):457–464.

Mohler, R. R. (1991). *Nonlinear Systems: Applications to Bilinear Control*, volume 2. Prentice Hall, Englewood Cliffs, NJ.

Mohler, R. R. and Khapalov, A. Y. (2000). Bilinear control and application to flexible a.c. transmission systems. *Journal of Optimization Theory and Applications*, 105(3):621–637.

Pearson, R. K. (1999). *Discrete-Time Dynamic Models*. Oxford University Press, New York, USA.

Rao, T. S. and Gabr, M. M. (1984). *An Introduction to Bispectral Analysis and Bilinear Time Series Models*. Springer-Verlag, Berlin, Germany.

Söderström, T. (2006). Statistical analysis of the Frisch scheme for identifying errors-in-variables systems. Technical report, Upsala Univercity, Department of Information Technology, Upsala, Sweden.

Söderström, T. (2007). Errors-in-variables methods in system identification. In *Automatica*, volume 43, pages 939–958.

Söderström, T., Soverini, U., and Mahata, K. (2002). Perspectives on errors-in-variables estimation for dynamic systems. In *Signal Processing*, volume 82(8), pages 1139–1154.

Söderström, T. and Stoica, P. (1994). *System Identification*. Prentice Hall Int., New Jersey, USA.

Young, P. (1984). *Recursive Estimation and Time-Series Analysis*. Springer-Verlag, Berlin, Germany.

Yu, D. (1996). *Fault diagnosis for industrial systems with emphasis on bilinear systems*. PhD thesis, Coventry University.

Zheng, W. X. (1998). Transfer function estimation from noisy input and output data. *Int. Journal of Adaptive Control and Signal Processing*, 12:365–380.

Zheng, W. X. (2000). Parametric identification of linear noisy input-output systems. *Cybernetics and Systems*, 31(7):803–816.

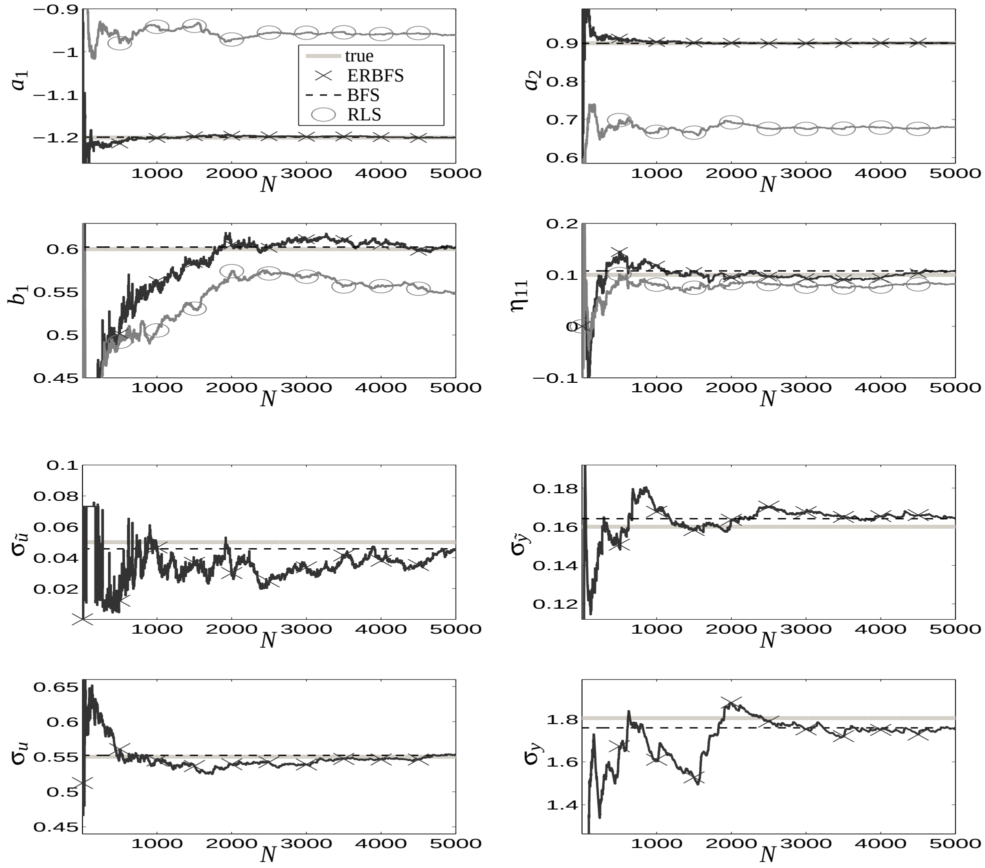

Figure 2: The results of the identification procedure using ERBFS, BFS and RLS algorithms.

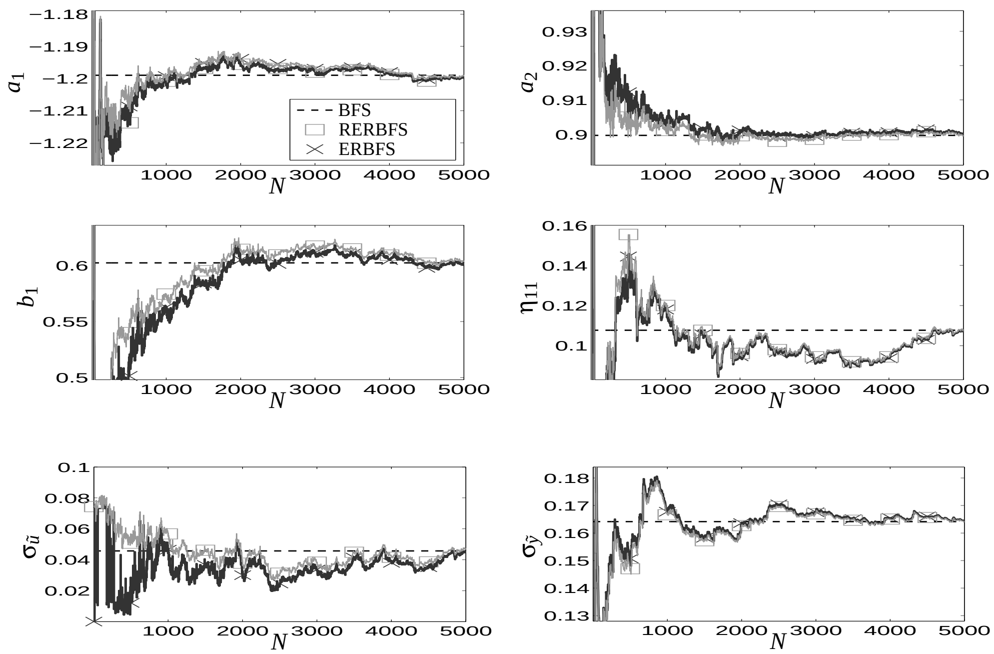

Figure 3: The results of the identification procedure using BFS, RERBFS and ERBFS algorithms.