On a Least-Squares-Based Algorithm for Identification of Stochastic Linear Systems

Wei Xing Zheng, Member, IEEE

Abstract—A new form of bias-eliminated least-squares (BELS) algorithm is developed to identify transfer function parameters of a linear time-invariant system, irrespective of noise dynamics. Unlike the BELS estimator previously presented, the main feature with the developed algorithm is that the transfer function parameters are consistently estimated in such a direct way that there is no need to prefilter observed data or to deal with a high-order augmented system. This greatly simplifies implementation of the BELS-based algorithms and reduces numerical efforts, whereas a desirable estimation accuracy can still be achieved. Two simulation examples are presented that clearly illustrate the good performances of the developed algorithm, including its superiority over one type of simple instrumental variable method.

Index Terms—Least-squares methods, parameter estimation, system identification.

I. INTRODUCTION

SYSTEM identification is concerned with fitting parametric models to observed data in order to gain information about a plant in a condensed form. The least-squares (LS) method has since been the dominant algorithm for parameter estimation due to its simplicity in concept and convenience in implementation. However, if the model of disturbances acting on the plant is misspecified, the LS method is unable to give rise to consistent estimates of parameters of the system transfer function [1], [2], [7]. The research on system identification over the last two decades has led to the development of the prediction error (PE) methods and the instrumental variable (IV) methods, which are now considered to be the two most important and commonly used identification techniques in coping with the cases with colored noise (see [5], [8] and the references therein). In particular, the output error (OE) method (i.e., the PE method with an output-error model structure) and the IV methods are more favorably adopted in the presence of inaccurate noise models because they do not need much information on the noise characteristics. However, one primary drawback of the OE method is that it involves intensive computations since its iterative procedure is performed by the Gauss–Newton minimization algorithm. On the other hand, the choice of instruments may affect ultimate identification results substantially in certain situations, whereas simple and efficient methods for selecting appropriate instruments to attain some optimal properties are expected to be developed yet for the IV methods.

Apart from the aforementioned two primary kinds of unbiased estimation approaches, namely, the PE methods and the IV methods, there are other types of bias-correction methods that have received considerable attention [3], [4], [6], [9], [12], [14]. Among them, the least-squares (BELS) method proposed in [3], [14] seems to be more promising from the viewpoint of practical applications. The BELS method is able to estimate transfer function parameters consistently when linear systems are corrupted by colored disturbances. The method is shown to have desirable properties, for instance, high accuracy, fast rate of convergence, and good robustness against noise. One main difference from the bias-correction methods developed in [4], [6], and [12] is that the BELS method is based on noniterative estimation procedures. Another important difference is that the compensated LS (CLS) method [9] and the modified LS method [6] are limited to the particular case where only the plant output is corrupted by white measurement noise, whereas the BELS method is applicable to arbitrary colored disturbances acting on the plant. This type of BELS method is extended to identification of continuous-time systems in the presence of colored noise by using linear integral filters in [13].

It should be mentioned that in recent years, new subspace-based identification algorithms have emerged that identify state-space models from input–output data. The attractive aspect of the subspace-based algorithms lies in its capability in identification of multivariable systems. Nevertheless, this area is still rather immature, especially in the aspect of estimation accuracy [11].

The aim of this paper is to study parameter identification of the system transfer function in the case of colored noise, with an emphasis on addressing two problems that exist with the BELS method presented in [3] and [14]. The first problem concerns prefiltering. The estimation of the bias in the LS estimator is implemented through a prefilter of a proper order to be designed. Unavoidably, prefiltering of observed data contributes, although moderately, an increase in the computational load in application of the BELS method. The second problem, which is closely related to the first one, is that a high-order augmented system resulting from inclusion of the designed prefilter into the identified system has to be handled. This means that on one hand, the parameter estimates of the augmented transfer function are first obtained,
and then extraction of the parameter estimates of the original transfer function of primary interest must be made, and on the other hand, dealing with a high-order system certainly causes added computations. In the following, the prefilter-based BELS method is called the PBELS method for short.

In this paper, a new form of BELS algorithm is proposed that is based on a novel formulation of an auxiliary LS estimator. The proposed algorithm is characterized by its use of no prefiltering and its direct parameter estimation structure, thus being termed the direct BELS (DBELS) algorithm. The procedure of the DBELS algorithm consists of the following steps. First, an LS estimate of the transfer function parameters of the underlying system, which is bound to be biased if the process noise is colored, is calculated. Second, the cross-covariance vector between the plant output and the process noise, which determines the noise-induced bias in the LS estimate, is directly estimated in a novel way that is based on a proper formulation of an LS estimate of another intermediate parameter vector that is introduced. Third, the consistent estimate of the transfer function parameters follows immediately from use of the bias-correction principle [4]. This new DBELS algorithm not only inherits the merits of the previous PBELS method but also, and more importantly, has some significant algorithmic advantages over the latter. This is due to the fact that direct use of unfiltered measurements means that no need to prefilter observed data and that direct identification of the original transfer function means that no parameter extraction needs to be performed. These algorithmic advantages make the DBELS algorithm very appealing in implementation, especially in on-line estimation.

The remainder of this paper is organized as follows. Section II gives a formulation of parameter estimation of transfer functions in the presence of colored disturbances and a brief analysis of biasedness of the ordinary LS estimator. Section III presents the development of the DBELS algorithm. Next, in Section IV, the presented algorithm is analyzed with respect to consistency, robustness against noise, algorithmic aspects and so on, in comparison with the PBELS method, the OE method, the IV methods, and the CLS method. Two examples to estimate transfer function parameters in the case of colored noise are given and evaluated in Section V, whereas the final remarks conclude the paper.

II. PROBLEM FORMULATION

Let \( G(q^{-1}) \) denote the transfer function of a linear time-invariant (LTI), discrete-time system, and be described by

\[
G(q^{-1}) = \frac{B(q^{-1})}{A(q^{-1})} = \frac{b_0 + b_1 q^{-1} + \cdots + b_m q^{-m}}{1 - a_1 q^{-1} - \cdots - a_n q^{-n}}
\]

(1)

where \( q^{-1} \) stands for the backward shift operator. The input/output behavior of the system can be represented by the following output-error type model

\[
y(t) = G(q^{-1}) u(t) + v(t)
\]

(2)

where

- \( u(t) \) plant input;
- \( y(t) \) plant output;
- \( v(t) \) process noise.

In particular, \( v(t) \) may represent observation errors, modeling errors, external disturbances acting on the plant, etc.

Some technical assumptions are imposed on the plant. We assume that the measurable input signal \( u(t) \) is stationary and persistently exciting of a proper order. We also assume that the disturbance signal \( v(t) \) is stationary but may be with arbitrary dynamic representations. Further, we assume that the polynomial \( A(q^{-1}) \) is stable, the model orders \((n, m)\) are given, and \( u(t) \) is statistically uncorrelated with \( v(t) \). Therefore, our attention is focused on the problem of parameter estimation of a plant in open-loop operation.

The output-error model (2) may be recast in the following linear regression form

\[
y(t) = \phi_t^T \theta + \epsilon(t)
\]

(3)

where the vector \( \theta \) represents the unknown parameters, i.e.,

\[
\theta^T = [a_1^T; b_1^T] = [a_1 \cdots a_n; b_1 \cdots b_m]
\]

(4)

and the regression vector \( \phi_t \) contains delayed input and output signals, i.e.,

\[
\phi_t^T = [y(t-n); u(t-n); u(t-1) \cdots u(t-m)]
\]

(5)

Moreover, the equation error \( \epsilon(t) \) in (3) is given by

\[
\epsilon(t) = A(q^{-1}) v(t).
\]

(6)

For the identification of the unknown parameter vector \( \theta \), we now consider the LS criterion

\[
V(\theta) = E[\epsilon(t)^2]
\]

(7)

where \( E \) denotes mathematical expectation. Minimization of \( V(\theta) \) yields the LS estimate

\[
\hat{\theta}_{LS} = \mathbf{R}_{\theta\theta}^{-1} \mathbf{r}_{\theta\epsilon}
\]

(8)

(see e.g., [2]), where the cross-covariance matrix \( \mathbf{R}_{\theta\theta} \) and the cross-covariance vector \( \mathbf{r}_{\theta\epsilon} \) are defined, respectively, by

\[
\mathbf{R}_{\theta\theta} = E[\phi_t \phi_t^T], \quad \mathbf{r}_{\theta\epsilon} = E[\phi_t \epsilon(t)].
\]

(9)

Next, we analyze the consistency of the LS estimate \( \hat{\theta}_{LS} \). Premultiplying (3) with \( \phi_t \) and then taking mathematical expectation leads to

\[
\mathbf{r}_{\theta\epsilon} = \mathbf{R}_{\theta\theta} \hat{\theta} + \mathbf{r}_{\theta\epsilon}.
\]

(10)

A closer insight into the cross-covariance noise vector \( \mathbf{r}_{\theta\epsilon} = E[\phi_t \epsilon(t)] \) in (10) reveals that

\[
\mathbf{r}_{\theta\epsilon} = \begin{bmatrix} E[y(t) \epsilon(t)] \\ E[u(t) \epsilon(t)] \end{bmatrix} = \begin{bmatrix} \mathbf{r}_{ye} \\ \mathbf{0} \end{bmatrix} = \mathbf{Q} \mathbf{r}_{ye}
\]

(11)

since \( E[u(t) \epsilon(t)] = 0 \) follows from the uncorrelation assumption on \( u(t) \) and \( \epsilon(t) \). Note that in (11), \( \mathbf{Q}^T = [\mathbf{I}_n, \mathbf{0}] \in \mathbb{R}^{n \times (n+m)} \), and \( \mathbf{I}_n \) denotes an identity matrix of order \( n \). Substituting (10) and (11) into (8) gives

\[
\hat{\theta}_{LS} = \hat{\theta} + \mathbf{R}_{\theta\theta}^{-1} \mathbf{Q} \mathbf{r}_{ye}.
\]

(12)

It is seen that \( \hat{\theta}_{LS} \) is an unbiased estimator of \( \theta \) if and only if \( \mathbf{r}_{ye} = \mathbf{0} \), namely, if and only if the equation error \( \epsilon(t) \) is further...
assumed to be white noise. This is obviously a very special case in practice. Otherwise, the colored noise \( \epsilon(t) \) definitely causes a bias in the ordinary LS estimate, and this bias is determined by the cross-covariance noise vector \( \mathbf{r}_{ye} \).

### III. The Direct BELS Algorithm

By using the well-known bias-correction principle \([4]\), the BELS estimator of \( \boldsymbol{\theta} \) can be obtained via the equation

\[
\hat{\boldsymbol{\theta}}_{\text{BELS}} = \hat{\boldsymbol{\theta}}_{\text{LS}} - \mathbf{R}_{\phi \phi}^{-1} \mathbf{r}_{\phi y}
\]

provided that the noise vector \( \mathbf{r}_{ye} \) is known or an estimate of it is available. Hence, to implement the unbiased estimation scheme \((13)\), it is desired to get an estimate of \( \mathbf{r}_{ye} \). Note that in \([3]\) and \([14]\), a somewhat indirect technique for consistent parameter estimation is presented that is based on use of prefiltering (see also Section IV-B). In what follows, a more direct approach to estimating \( \mathbf{r}_{ye} \) is described.

#### A. Estimation of the Cross-Covariance Noise Vector \( \mathbf{r}_{ye} \)

To begin with, we consider identifying the LTI system \((1)\) but with a different order choice for the numerator polynomial. We introduce an auxiliary transfer function defined in \((14)\), shown at the bottom of the page, with

\[
\begin{align*}
\mathbf{b}^T &= [b_{m+1} \cdots b_{m+n}] = \mathbf{0}^T \in \mathbb{R}^n, \\
\mathbf{b}^\top &= \begin{bmatrix} \mathbf{b}^T \end{bmatrix} \in \mathbb{R}^{n+1}, \\
\mathbf{b}^\top &= \begin{bmatrix} \mathbf{b}^T \end{bmatrix} \in \mathbb{R}^{n+1}.
\end{align*}
\]

Equation \((15)\) shows that the introduced \( n \) parameters \( b_{m+1}, \ldots, b_{m+n} \) in the numerator of \( \mathcal{G}(q^{-1}) \) all take zero as the true values. Therefore, the input/output behavior of the LTI system described by \( \mathcal{G}(q^{-1}) \) should be equivalent to that described by \( \mathcal{G}(q^{-1}) \). However, the substantial difference between \( \mathcal{G}(q^{-1}) \) and \( \mathcal{G}(q^{-1}) \) is that the former is with the high model orders \((n, m+n)\), whereas the latter is with \((n, m)\).

For notational convenience, the new variables introduced in the auxiliary model are marked with an overbar.

By analogy with \((8)\), the LS estimate of the parameter vector \( \hat{\boldsymbol{\theta}} \) of the auxiliary transfer function \((14)\) is found to be

\[
\hat{\theta}_{\text{LS}} = \mathbf{R}_{\phi \phi}^{-1} \mathbf{r}_{\phi y}
\]

where \( \mathbf{R}_{\phi \phi} = E[\phi \phi^\top] \), \( \mathbf{r}_{\phi y} = E[\phi y(t)] \), and the regression vector \( \hat{\phi}_t \) is given by

\[
\hat{\phi}_t = [\phi_t^T, \hat{\mathbf{u}}_t^T],
\]

\[
\hat{\mathbf{u}}_t = [u(t - m - 1) \cdots u(t - m - n)].
\]

Similarly to \((12)\), an analysis of \( \hat{\theta}_{\text{LS}} \) reveals that

\[
\hat{\theta}_{\text{LS}} = \hat{\theta} + \mathbf{R}_{\phi \phi}^{-1} \mathbf{r}_{ye}
\]

where \( \mathbf{Q}^\top = [\mathbf{I}_n \mathbf{0}] \in \mathbb{R}^{n \times (2n+1-m)} \). It is very important to note from \((12)\) and \((18)\) that the bias in \( \hat{\theta}_{\text{LS}} \) and the bias in \( \hat{\theta}_{\text{LS}} \) are determined by the exact same noise vector \( \mathbf{r}_{ye} \).

Next, we derive expressions for the LS estimate of the intermediate parameter vector \( \mathbf{b} \) previously introduced and use them to arrive at an estimate for \( \mathbf{r}_{ye} \). By exploiting the particular structure of \( \hat{\phi}_t \) given in \((17)\), we can decompose the covariance matrix \( \mathbf{R}_{\phi \phi} \) and the covariance vector \( \mathbf{r}_{\phi y} \) into the following block form

\[
\begin{align*}
\mathbf{R}_{\phi \phi} &= \begin{bmatrix} \mathbf{R}_{\phi \phi} & \mathbf{R}_{\phi \eta} \\ \mathbf{R}_{\eta \phi} & \mathbf{R}_{\eta \eta} \end{bmatrix}, \\
\mathbf{r}_{\phi y} &= \begin{bmatrix} \mathbf{r}_{\phi y} \\ \mathbf{r}_{\eta y} \end{bmatrix},
\end{align*}
\]

\[\text{(19)}\]

where \( \mathbf{R}_{\phi \phi} \) and \( \mathbf{r}_{\phi y} \) are the same as defined in \((9)\), and \( \mathbf{R}_{\phi \eta} = E[\phi \eta^\top], \mathbf{R}_{\eta \phi} = E[\eta \phi^\top], \mathbf{r}_{\eta y} = E[\eta y(t)] \).

Applying the matrix inversion formula to \( \mathbf{R}_{\phi \phi} \) yields

\[
\mathbf{R}^{-1}_{\phi \phi} = \begin{bmatrix} \mathbf{R}^{-1}_{\phi \phi} + \mathbf{R}^{-1}_{\phi \eta} \mathbf{R}^{-1}_{\eta \phi} \mathbf{R}^{-1}_{\phi \phi} - \mathbf{R}^{-1}_{\phi \eta} \mathbf{R}^{-1}_{\phi \phi} \mathbf{R}^{-1}_{\phi \phi} & -\mathbf{R}^{-1}_{\phi \eta} \mathbf{R}^{-1}_{\phi \phi} \mathbf{R}^{-1}_{\phi \phi} \\
-\mathbf{R}^{-1}_{\phi \eta} \mathbf{R}^{-1}_{\phi \phi} \mathbf{R}^{-1}_{\phi \phi} & \mathbf{R}^{-1}_{\phi \phi} \end{bmatrix}
\]

\[\text{(20)}\]

where

\[
\Delta = \mathbf{R}_{\eta \eta} - \mathbf{R}_{\eta \phi} \mathbf{R}^{-1}_{\phi \phi} \mathbf{R}_{\phi \eta}.
\]

Combining \((16)\) with \((21)\) leads to

\[
\hat{\mathbf{b}}_{\text{LS}} = -\Delta^{-1} \mathbf{R}_{\phi \phi}^{-1} \mathbf{r}_{\phi y} + \Delta^{-1} \mathbf{r}_{\eta y} = -\Delta^{-1} \mathbf{R}_{\phi \phi}^{-1} \hat{\theta}_{\text{LS}} - \mathbf{r}_{\eta y}.
\]

On the other hand, it follows from \((15)\), \((18)\), and \((21)\) that

\[
\hat{\mathbf{b}}_{\text{LS}} = \mathbf{b} - \Delta^{-1} \mathbf{R}_{\phi \phi}^{-1} \mathbf{r}_{\phi y} = -\Delta^{-1} \mathbf{R}_{\phi \phi}^{-1} \mathbf{r}_{\phi y}.
\]

Then, equating \((22)\) and \((23)\) with \( \hat{\mathbf{b}}_{\text{LS}} \) results in

\[
-\Delta^{-1} \mathbf{R}_{\phi \phi}^{-1} \mathbf{r}_{\phi y} = -\Delta^{-1} \mathbf{R}_{\phi \phi}^{-1} \hat{\theta}_{\text{LS}} - \mathbf{r}_{\eta y}.
\]

From \((24)\), it is straightforward to get

\[
\hat{\mathbf{r}}_{ye} = (\mathbf{R}_{\phi \phi}^{-1} \mathbf{Q})^{-1} (\mathbf{R}_{\phi \phi}^{-1} \hat{\theta}_{\text{LS}} - \mathbf{r}_{\eta y}).
\]

The estimate \( \hat{\mathbf{r}}_{ye} \) of \( \mathbf{r}_{ye} \) given by \((25)\) can be used in conjunction with \((13)\) to obtain an unbiased estimate of the parameter vector \( \boldsymbol{\theta} \).

#### B. Summary of the DBELS Algorithm

For practical use, only a finite set of \( N \) input-output data \( \{u(t), y(t)\}, t = 1, \cdots, N \) can be collected on the plant. In this case, the unknown covariances \( \mathbf{R}_{\phi \phi}, \mathbf{r}_{\phi y}, \mathbf{R}_{\eta \eta}, \mathbf{r}_{\eta y} \) in the theoretical equations \((8)\) and \((25)\) may be replaced by its natural estimates

\[
\hat{\mathbf{R}}_{\phi \phi}(N) = \frac{1}{N} \sum_{t=1}^{N} \mathbf{Q} \phi_t \phi_t^\top, \quad \hat{\mathbf{r}}_{\phi y}(N) = \frac{1}{N} \sum_{t=1}^{N} \phi_y(t),
\]

\[
\hat{\mathbf{R}}_{\eta \eta}(N) = \frac{1}{N} \sum_{t=1}^{N} \mathbf{Q} \eta_t \eta_t^\top, \quad \hat{\mathbf{r}}_{\eta y}(N) = \frac{1}{N} \sum_{t=1}^{N} \eta_y(t).
\]

\[\text{(26a, b)}\]

\[
\mathcal{G}(q^{-1}) = \frac{B(q^{-1})}{A(q^{-1})} = \frac{b_0 q^{-1} + \cdots + b_m q^{-m} + b_{m+1} q^{-(m+1)} \cdots + b_{m+n} q^{-(m+n)}}{1 - a_1 q^{-1} - \cdots - a_n q^{-m}}
\]

\[\text{(14)}\]
Hence, the above derivation can be summarized as the following direct BELS algorithm.

**The DBELS Algorithm**

1) Calculate the sample covariance estimates \( \hat{R}_{xx}^e(N), \hat{R}_{yy}^e(N), \hat{R}_{xy}^e(N), \) and \( \hat{r}_{ye}(N) \) in terms of (26).

2) Use the ordinary LS method to get

\[
\hat{\theta}_{LS}(N) = \hat{R}_{x}^{-1}(N)r_{xy}(N). \tag{27}
\]

3) Estimate the noise vector \( r_{ye} \) by virtue of (25), which gives

\[
\hat{r}_{ye}(N) = [\hat{R}_{xy}^e(N)\hat{R}_{xy}^{-1}(N)Q]'^{-1} \cdot [\hat{R}_{xy}^e(N)\hat{\theta}_{LS}(N) - \hat{r}_{xy}(N)]. \tag{28}
\]

4) Perform the bias correction to obtain

\[
\hat{\theta}_{DBELS}(N) = \hat{\theta}_{LS}(N) - \hat{R}_{xx}^{-1}(N)Qr_{ye}(N). \tag{29}
\]

Like the PBELS method, the above DBELS algorithm can be easily transformed into a recursive scheme for on-line estimation because the sample covariance estimates given in (26) may be evaluated recursively.

**IV. ANALYSIS AND COMMENTS**

**A. Consistency of the DBELS Algorithm**

We first show that the DBELS algorithm is strongly consistent. It is evident that the key to establishing the consistency of \( \hat{\theta}_{DBELS}(N) \) is to demonstrate that \( \hat{r}_{ye}(N) \) is consistent. Substitution of (12) into (25) yields

\[
\hat{r}_{ye} = (R_{xx}^{-1}R_{xy}^{-1}Q')^{-1}(R_{xx}^{-1}\theta + R_{xy}^{-1}Qr_{ye} - r_{xy})
= r_{ye} + (R_{xx}^{-1}R_{xy}^{-1}Q')^{-1}(R_{xx}^{-1}\theta - r_{xy}). \tag{30}
\]

On the other hand, using (3) and the uncorrelation assumption on \( u(t) \) and \( e(t) \), we find

\[
\begin{align*}
r_{yu}(k) &= a_1r_{yu}(k-1) + \cdots + a_nr_{yu}(k-n) \\
& \quad + b_1r_{uu}(k-1) + \cdots + b_mr_{uu}(k-m)
\end{align*} \tag{31}
\]

where the covariance functions \( r_{yu}(k) \) and \( r_{uu}(k) \) are defined, respectively, by

\[
r_{yu}(k) = E[y(t)u(t-k)], \quad r_{uu}(k) = E[u(t)u(t-k)]. \tag{32}
\]

Letting \( k = m+1, \cdots, m+n \) in (31) and writing the resultant \( n \) equations in matrix–vector form gives

\[
r_{xy} = R_{x}^{-1}\theta. \tag{33}
\]

Then, combining (30) with (33) leads to

\[
\hat{r}_{ye} = r_{ye}. \tag{34}
\]

Since the sample covariance estimates given by (26) converge to its respective true covariances, letting \( N \to \infty \) in (28) and using (25) and (34), we obtain

\[
\lim_{N \to \infty} \hat{r}_{ye}(N) = (R_{xx}^{-1}R_{xy}^{-1}Q')^{-1}(R_{xx}^{-1}\theta_{LS} - r_{xy}) = r_{ye} \quad \text{w.p.1}. \tag{35}
\]

Equation (35) shows that \( \hat{r}_{ye}(N) \) is an asymptotically consistent estimate of \( r_{ye} \). It then follows immediately from (12), (29), and (35) that

\[
\lim_{N \to \infty} \hat{\theta}_{DBELS}(N) = \lim_{N \to \infty} \hat{\theta}_{LS}(N) - \lim_{N \to \infty} R_{x}^{-1}(N)Q\hat{r}_{ye}(N)
= (\theta + R_{xx}^{-1}Qr_{ye}) - R_{x}^{-1}Qr_{ye} = \theta \quad \text{w.p.1}. \tag{36}
\]

Thus, we have proved the following consistency result.

**Theorem 1:** Let \( \hat{\theta}_{DBELS}(N) \) be the parameter estimate given by the DBELS algorithm under the stated assumptions on the LTI system (2). Then

\[
\lim_{N \to \infty} \hat{\theta}_{DBELS}(N) = \theta \quad \text{w.p.1}. \tag{37}
\]

**B. Comparison Between the DBELS Method and the PBELS Method**

We now look at some major algorithmic differences between the DBELS method and the PBELS method. For this purpose, the PBELS method presented in [3] and [14] is outlined as follows. In the first step of the PBELS algorithm, an \( n \)-th-order all-pole stable filter \( F(q^{-1}) \) is designed and is used to prefilter the input \( u(t) \), which gives \( \hat{u}(t) = F(q^{-1})u(t) \). This prefilter is also inserted into the underlying plant to form an augmented system parametrized by \( \hat{\theta} \in \mathbb{R}^{2n+m} \). The second step is to apply the ordinary LS method to the data set \( \{\hat{u}(t), y(t), t = 1, \cdots, N\} \), which yields

\[
\hat{\theta}_{LS}(N) = \hat{R}_{x}^{-1}(N)r_{xy}(N) \tag{38}
\]

where \( \hat{R}_{xx}(N) \in \mathbb{R}^{(2n+m) \times (2n+m)} \) and \( \hat{R}_{xy}(N) \in \mathbb{R}^{2n+m} \) are covariance matrices defined in terms of \( \hat{\phi}_i \) and \( y(t) \) in a similar way to \( \hat{R}_{xx}(N) \) and \( \hat{r}_{xy}(N) \) in (26a), with \( \hat{\phi}_i \in \mathbb{R}^{2n+m} \) being the regression vector formed by \( \hat{y}(t) \) and \( \hat{u}(t) \) in a similar way to \( \phi_i \) in (5). In the third step, the noise vector \( r_{ye} \) is estimated by

\[
\hat{r}_{ye}(N) = [H^T\hat{R}_{xy}^{-1}(N)Q']^{-1}H^T\hat{\theta}_{LS}(N) \tag{39}
\]

where \( H \in \mathbb{R}^{(2n+m) \times n} \) is constructed by use of the known poles of \( F(q^{-1}) \). Then, the BELS estimate of \( \theta \) is found by

\[
\hat{\theta}_{PBELS}(N) = \hat{\theta}_{LS}(N) - \hat{R}_{x}^{-1}(N)Qr_{ye}(N). \tag{40}
\]

Finally, the BELS estimate of \( \theta \), i.e., \( \hat{\theta}_{PBELS}(N) \), is extracted from \( \hat{\theta}_{PBELS}(N) \) by means of the designed prefilter \( F(q^{-1}) \).

It is easy to see that one major difference is that the DBELS algorithm works directly with the measured data \( \{u(t), y(t), t = 1, \cdots, N\} \), whereas the PBELS method needs to employ the prefilter \( F(q^{-1}) \) to get the filtered \( \hat{u}(t) \) from \( u(t) \). Since neither prefilter design nor prefiltering of the sampled data is required in implementation of the proposed algorithm, a saving in the computational load can be achieved.

Another important difference lies in that the DBELS method essentially works with the underlying plant (1) with the model orders \( (n, m) \) directly, whereas the PBELS method has to identify the high-order augmented system with the model orders \( (n, n+m) \). Note that though the auxiliary model with
TABLE I

<table>
<thead>
<tr>
<th>Approach</th>
<th>MAIN PART OF COMPUTATIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DBELS</td>
</tr>
<tr>
<td>Dimensionality of matrices</td>
<td>$(n + m)^2 + (n + m)n + n^2$</td>
</tr>
</tbody>
</table>

$G(q^{-1})$ is introduced in Section III-A, it is only used as an intermediate step to arrive at an estimate of the noise vector $\mathbf{r}_{PE}$ with no need to actually identify it at all in the DBELS scheme. However, in the PBELS method, the $(2n + m)$-dimensional parameter vector $\hat{\theta}$ of the high-order augmented system needs to be estimated [see (38) and (40)]. It is known that the main part of the computations in the parametric algorithms is the updating of a matrix. This leads to a computational complexity that increases with $L^2$, where $L$ is the total number of estimated parameters in the plant model. To make a better illustration, Table I displays the main part of computations in terms of dimensionality of covariance matrices and/or inverse matrices used in the DBELS and the PBELS methods. It is easily seen from the table that with the $n + m$ parameters of the original transfer function being estimated directly, far fewer computations are needed with the DBELS algorithm than with the PBELS method. For example, if it is assumed that $n = m$, then it is expected that the computational load of the DBELS method accounts for roughly 70% that of the PBELS method. A third main difference is that the PBELS method is a certain type of indirect technique for consistent parameter estimation since the BELS estimate of $\theta$ of primary interest has to be recovered from the BELS estimate of the augmented parameter vector $\hat{\theta}_{BELS}(N)$. On the contrary, the DBELS algorithm can produce a direct consistent estimate of $\theta$ without such parameter extraction. Besides, the simple and compact structure of the presented algorithm also enables easier implementation, particularly in on-line identification. These are the primary algorithmic advantages of the DBELS algorithm over the PBELS method.

C. Comparison of the DBELS Method with the OE Method and the IV Methods

We first consider robustness against noise. As shown in Section IV-A, in order to assure consistency, the only assumption imposed on the noise $\psi(t)$ is stationarity. Otherwise, $\psi(t)$ may be represented by any types of noise models or may describe unmodeled colored noise. Therefore, like the PBELS method presented in [3] and [14], the proposed algorithm is consistent, regardless of whether the noise contribution on the observed data can be modeled exactly. This implies that the DBELS method possesses good robustness with respect to noise. The OE method and the simple IV (SIV) method (i.e., the IV method with delayed inputs as instruments) also have a similar robust feature against noise, whereas the optimal IV (OIV) method [8] requires noise modeling in its approximate implementation.

As will be observed from the simulation results presented in Section V as well as several other numerical examples not shown in Section V, the DBELS algorithm can be superior to the SIV method in terms of accuracy and reliability in some situations when the sample length is short and/or the noise level is high. The explanation for this may be that the presented algorithm evaluates the estimation bias explicitly and removes it in a direct manner, thus being an efficient identification algorithm. On the other hand, however, the asymptotic theory regarding the covariance matrix of the SIV estimates, which is valid for large sample lengths, may not necessarily hold for small sample lengths. Note that the similar observations based on the simulation studies are also obtained in [10] for the SIV method.

Furthermore, the computational costs with the OE and the OIV methods are considerably more than that with the DBELS algorithm since the former involve iterative nonlinear optimization procedures, whereas the latter is based on linear regression. Nevertheless, the SIV method is perhaps the simplest algorithm from the numerical point of view.

Finally, we note that, as pointed out in [12], although the IV type methods can be used to perform consistent parameter estimation in the case of colored noise, they have some aspects, e.g., the convergence rate of the algorithm, that are inferior to the bias-correction-based methods because the introduction of an instrumental variable usually means that the resulting covariance matrix is no longer symmetric.

D. Comparison Between the DBELS Method and the CLS Method

It is also interesting to compare the proposed DBELS method with the CLS method since both methods have one common feature in that they are developed based on a similar noniterative bias-compensation principle. First of all, it should be pointed out that the applicability of the CLS method depends entirely on the assumption that the process noise $\psi(t)$ in the output-error model (2) is white noise. This turns out to be the key condition for the CLS method to work. As this white noise condition on $\psi(t)$ is implicitly equivalent to requiring the precise information of the noise contribution on the observed data, the CLS method may not be very robust against noise. On the contrary, the DBELS method can handle much more general cases because $\psi(t)$ in (2) can be any types of colored noise. Therefore, the DBELS algorithm has more extensive practical application domain than the CLS method, which can be considered to be its substantial advantage over the CLS method.

On the other hand, the CLS method has a simpler algorithmic structure than the DBELS method since it is purposely built on the white noise assumption on $\psi(t)$ in (2). In the situation when this specific assumption can be satisfied, it is conceivable that the CLS method will be much simpler in computations than the DBELS method. Hence, the CLS method may serve as a better bias-correction-based identification algorithm in terms of its low computational complexity in such particular circumstances with white output noise $\psi(t)$. Another point to mention is that an expression of the covariance matrix of the CLS estimate is derived in [9] for a rigorous analysis of accuracy, whereas the similar work on the DBELS estimate is still under way and has yet to be reported.

E. Existence of the Inverse in (28)

A word is in order regarding the existence of the inverse $[\hat{R}_{\psi\psi}(N)\hat{R}_{\psi\psi}^{-1}(N)Q]^{-1}$ in (28). In certain low-order situations, it can be shown that the matrix $\hat{R}_{\psi\psi}^{-1}(N)Q$
TABLE II

<table>
<thead>
<tr>
<th>Approach</th>
<th>LS</th>
<th>OE</th>
<th>SIV</th>
<th>OIV</th>
<th>PBELS</th>
<th>DBELS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample length</td>
<td>N=50</td>
<td>N=250</td>
<td>N=50</td>
<td>N=250</td>
<td>N=50</td>
<td>N=250</td>
</tr>
<tr>
<td>$a_1$</td>
<td>0.5</td>
<td>0.6260</td>
<td>0.6356</td>
<td>0.5012</td>
<td>0.5081</td>
<td>0.6857</td>
</tr>
<tr>
<td></td>
<td></td>
<td>±0.0836</td>
<td>±0.0863</td>
<td>±0.2094</td>
<td>±0.0867</td>
<td>±0.6649</td>
</tr>
<tr>
<td>$b_1$</td>
<td>1.0</td>
<td>1.0021</td>
<td>1.0026</td>
<td>0.9733</td>
<td>0.9925</td>
<td>1.0998</td>
</tr>
<tr>
<td></td>
<td></td>
<td>±0.1616</td>
<td>±0.0668</td>
<td>±0.2473</td>
<td>±0.1005</td>
<td>±2.2564</td>
</tr>
<tr>
<td></td>
<td></td>
<td>±0.1616</td>
<td>±0.0669</td>
<td>±0.2478</td>
<td>±0.1008</td>
<td>±2.2552</td>
</tr>
<tr>
<td>flops # per run</td>
<td>561</td>
<td>2561</td>
<td>14637</td>
<td>55924</td>
<td>672</td>
<td>3071</td>
</tr>
<tr>
<td>RE</td>
<td>11.27%</td>
<td>12.13%</td>
<td>2.39%</td>
<td>9.8%</td>
<td>18.48%</td>
<td>0.4%</td>
</tr>
<tr>
<td>ACV</td>
<td>20.14%</td>
<td>14.37%</td>
<td>33.67%</td>
<td>13.64%</td>
<td>515.45%</td>
<td>19.45%</td>
</tr>
<tr>
<td>RCL</td>
<td>1.00</td>
<td>1.00</td>
<td>26.49</td>
<td>21.84</td>
<td>1.20</td>
<td>1.20</td>
</tr>
</tbody>
</table>

TABLE III

<table>
<thead>
<tr>
<th>Approach</th>
<th>LS</th>
<th>OE</th>
<th>SIV</th>
<th>OIV</th>
<th>PBELS</th>
<th>DBELS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample length</td>
<td>N=200</td>
<td>N=1000</td>
<td>N=200</td>
<td>N=1000</td>
<td>N=200</td>
<td>N=1000</td>
</tr>
<tr>
<td>$a_1$</td>
<td>0.5</td>
<td>0.6903</td>
<td>0.6923</td>
<td>0.4985</td>
<td>0.5629</td>
<td>0.3627</td>
</tr>
<tr>
<td></td>
<td></td>
<td>±0.0425</td>
<td>±0.0183</td>
<td>±0.2290</td>
<td>±0.0938</td>
<td>±1.9380</td>
</tr>
<tr>
<td></td>
<td></td>
<td>±0.1950</td>
<td>±0.1931</td>
<td>±0.0220</td>
<td>±0.0096</td>
<td>±1.9238</td>
</tr>
<tr>
<td>$b_1$</td>
<td>1.0</td>
<td>0.9887</td>
<td>0.9991</td>
<td>0.9646</td>
<td>0.9899</td>
<td>1.0066</td>
</tr>
<tr>
<td></td>
<td></td>
<td>±0.1643</td>
<td>±0.0673</td>
<td>±0.2623</td>
<td>±0.1148</td>
<td>±0.4133</td>
</tr>
<tr>
<td></td>
<td></td>
<td>±0.1647</td>
<td>±0.0673</td>
<td>±0.2646</td>
<td>±0.1152</td>
<td>±0.4134</td>
</tr>
<tr>
<td>flops # per run</td>
<td>2061</td>
<td>10861</td>
<td>58574</td>
<td>229573</td>
<td>2472</td>
<td>12071</td>
</tr>
<tr>
<td>RE</td>
<td>22.45%</td>
<td>17.81%</td>
<td>33.98%</td>
<td>15.62%</td>
<td>285.71%</td>
<td>22.09%</td>
</tr>
<tr>
<td>ACV</td>
<td>22.45%</td>
<td>17.81%</td>
<td>33.98%</td>
<td>15.62%</td>
<td>285.71%</td>
<td>22.09%</td>
</tr>
<tr>
<td>RCL</td>
<td>1.00</td>
<td>1.60</td>
<td>28.42</td>
<td>22.82</td>
<td>1.20</td>
<td>1.20</td>
</tr>
</tbody>
</table>

is invertible under mild conditions. Although we have not yet been able to generalize this analysis, it has been observed from our extensive numerical studies that the inverse $[\mathbf{R}_N(N)\mathbf{R}_N^\top(N)\mathbf{Q}]^{-1}$ always exists with good matrix condition numbers.

V. IDENTIFICATION EXAMPLES

In this section, the results of some computer simulations are presented to illustrate the performance of the proposed DBELS algorithm and to compare with the other five identification approaches, namely, the ordinary LS method, the OE method, the SIV method, the OIV method, and the PBELS method. For this purpose, two important aspects of these six algorithms, that is, accuracy and computational complexity, are evaluated, with the accuracy being described by bias and variance and the computational complexity being measured approximately by the MATLAB code flops (count of floating-point operations). Note that the OE method and the OIV method are implemented via the standard MATLAB codes oe and iv4, respectively, whereas the coding of the other four algorithms is done in MATLAB.

To get an overall description of the performances, we introduce the following three performance criteria. The first is the relative error (RE) defined as

$$RE = \frac{\|\mathbf{m}(\hat{\theta}) - \mathbf{\theta}\|}{\|\mathbf{\theta}\|}$$

where $\mathbf{m}(\hat{\theta}) = [m(\hat{\theta}_1) \cdots m(\hat{\theta}_{n+m})]$ represents the sample mean of an estimator $\hat{\theta}$ and $\| \cdot \|$ indicates the Euclidean norm on vectors. The second is the averaged coefficient of variation (ACV), which is defined as

$$ACV = \frac{1}{n+m} \sum_{i=1}^{n+m} \frac{\sigma(\theta_i)}{|m(\theta_i)|}$$

where $\sigma(\theta_i)$ denotes the standard deviation of $\hat{\theta}_i$ from its true value $\theta_i$ for $i = 1, \cdots, n+m$. The third is the ratio of the computational load (RCL) of one algorithm with respect to that of the ordinary LS method. In the following, the signal-to-noise ratio (SNR) at the plant output is calculated as

$$SNR = 10 \log_{10} \frac{\sum_{t=1}^{N} |G(q^{-1})u(t)|^2}{\sum_{t=1}^{N} |v(t)|^2} \text{ (db)}$$

and the simulation results are all based on 500 independent runs.

Example 1: The transfer function of an LTI system to be identified is given by

$$G(q^{-1}) = \frac{1.0}{1 - 0.5q^{-1}}$$

Thus, we have $a_1 = 0.5$ and $b_1 = 1.0$. The input $u(t)$ is generated as a zero-mean unit-variance white noise sequence. The colored process noise $v(t)$ is described by

$$v(t) = \frac{1 + 0.5q^{-1}}{1 - 0.5q^{-1}} u(t)$$

where $u(t)$ is white noise of zero mean with variance $\sigma_u^2$ and uncorrelated with $u(t)$. Then, the equation error $e(t)$ is
This example is from [8]. In the first case considered, $\sigma_n^2 = 1$ is selected, and the corresponding SNR takes the value of $-2.41$ dB. By setting the sample length $N = 50$ and $N = 250$, respectively, the system is identified off line by using the ordinary LS method, the OE method, the SIV method, the OIV method, the PBELS method with the prefilter $F(q^{-1}) = 1/(1 - 0.5q^{-1})$, and the DBELS algorithm. The arithmetic means and standard deviations of the estimated parameters $a_1$ and $b_1$ from both the mean and true values, the averaged number of $\text{flOps}$ per run, RE, ACV, and RCL are summarized in Table II. In the second case, a larger noise variance $\sigma_n^2$ is chosen to examine the robustness against noise, with the resulting SNR $9.40$ dB. For the cases of $a_1$ and $a_2$, the comparative estimation performances obtained by these six different algorithms are listed in Table III, whose entries are defined similarly to those in Table II.

As expected, the ordinary LS method produces a substantial bias due to the presence of the colored noise. According to the RCL values, the amount of computations associated with the OE method and the OIV method is over 20 times and 14 times, respectively, that of the ordinary LS method, although the former two methods give rise to unbiased estimates. On the other hand, in spite of its least numerical requirements, the SIV method fails to work properly in short sample length problems with high noise, as shown in the cases of $a_1$ in Table II and of $a_2$ in Table III. On the contrary, the DBELS algorithm performs very well in these two cases, thus demonstrating a good robustness in low SNR environments. In fact, it is seen from Tables II and III that the proposed algorithm is more accurate in terms of lower ACV than the SIV method, the PBELS method, and even the OE method. In agreement with the analysis given in Section IV-B, the computational costs with the DBELS algorithm are reduced by about 30% from that of the PBELS method because the latter requires prefiltering of the observed data and handling of a high-order augmented system. It seems that for the simulation example considered here, the proposed algorithm is much more cost effective in terms of accuracy and computations than the other four consistent estimation algorithms.

### Table IV

<table>
<thead>
<tr>
<th>Approach</th>
<th>LS</th>
<th>OE</th>
<th>SIV</th>
<th>OIV</th>
<th>PBELS</th>
<th>DBELS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample length</td>
<td>$N=100$</td>
<td>$N=500$</td>
<td>$N=100$</td>
<td>$N=500$</td>
<td>$N=100$</td>
<td>$N=500$</td>
</tr>
<tr>
<td>$a_1$</td>
<td>$1.5$</td>
<td>$1.2053$</td>
<td>$1.2200$</td>
<td>$1.4967$</td>
<td>$1.4995$</td>
<td>$1.7056$</td>
</tr>
<tr>
<td>$b_1$</td>
<td>$1.0$</td>
<td>$0.9935$</td>
<td>$0.9989$</td>
<td>$0.9865$</td>
<td>$0.9909$</td>
<td>$1.0227$</td>
</tr>
<tr>
<td>$a_2$</td>
<td>$-0.7$</td>
<td>$-0.4321$</td>
<td>$-0.4412$</td>
<td>$-0.6985$</td>
<td>$-0.6999$</td>
<td>$-0.9103$</td>
</tr>
<tr>
<td>$b_2$</td>
<td>$0.5$</td>
<td>$0.7884$</td>
<td>$0.7789$</td>
<td>$0.5208$</td>
<td>$0.5046$</td>
<td>$0.2892$</td>
</tr>
<tr>
<td>$\text{flOps} # \text{per run}$</td>
<td>$3066$</td>
<td>$14266$</td>
<td>$61925$</td>
<td>$62349$</td>
<td>$4085$</td>
<td>$19284$</td>
</tr>
<tr>
<td>RE</td>
<td>$24.62%$</td>
<td>$23.51%$</td>
<td>$1.25%$</td>
<td>$0.27%$</td>
<td>$18.14%$</td>
<td>$2.23%$</td>
</tr>
<tr>
<td>ACV</td>
<td>$35.93%$</td>
<td>$30.82%$</td>
<td>$13.87%$</td>
<td>$6.31%$</td>
<td>$278.65%$</td>
<td>$26.81%$</td>
</tr>
<tr>
<td>RCL</td>
<td>$1.00$</td>
<td>$1.00$</td>
<td>$20.20$</td>
<td>$18.50$</td>
<td>$1.33$</td>
<td>$1.35$</td>
</tr>
</tbody>
</table>

### Table V

<table>
<thead>
<tr>
<th>Approach</th>
<th>LS</th>
<th>OE</th>
<th>SIV</th>
<th>OIV</th>
<th>PBELS</th>
<th>DBELS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample length</td>
<td>$N=100$</td>
<td>$N=500$</td>
<td>$N=100$</td>
<td>$N=500$</td>
<td>$N=100$</td>
<td>$N=500$</td>
</tr>
<tr>
<td>$a_1$</td>
<td>$1.5$</td>
<td>$0.9861$</td>
<td>$0.9977$</td>
<td>$1.0694$</td>
<td>$1.0491$</td>
<td>$2.0347$</td>
</tr>
<tr>
<td>$b_1$</td>
<td>$1.0$</td>
<td>$0.9599$</td>
<td>$0.9955$</td>
<td>$0.9927$</td>
<td>$1.0198$</td>
<td>$1.0101$</td>
</tr>
<tr>
<td>$a_2$</td>
<td>$-0.7$</td>
<td>$-0.1852$</td>
<td>$-0.1845$</td>
<td>$-0.6908$</td>
<td>$-0.7004$</td>
<td>$-1.2362$</td>
</tr>
<tr>
<td>$b_2$</td>
<td>$0.5$</td>
<td>$1.1320$</td>
<td>$1.1019$</td>
<td>$0.5298$</td>
<td>$0.5678$</td>
<td>$0.6441$</td>
</tr>
<tr>
<td>$\text{flOps} # \text{per run}$</td>
<td>$11466$</td>
<td>$50266$</td>
<td>$263624$</td>
<td>$107119$</td>
<td>$15484$</td>
<td>$76983$</td>
</tr>
<tr>
<td>RE</td>
<td>$25.26%$</td>
<td>$49.83%$</td>
<td>$1.58%$</td>
<td>$0.53%$</td>
<td>$43.74%$</td>
<td>$2.17%$</td>
</tr>
<tr>
<td>ACV</td>
<td>$105.52%$</td>
<td>$101.58%$</td>
<td>$19.04%$</td>
<td>$7.07%$</td>
<td>$1959.25%$</td>
<td>$29.36%$</td>
</tr>
<tr>
<td>RCL</td>
<td>$1.40$</td>
<td>$1.00$</td>
<td>$22.99$</td>
<td>$19.50$</td>
<td>$1.35$</td>
<td>$1.36$</td>
</tr>
</tbody>
</table>
Example 2: The underlying LTI system to be identified is described by the transfer function

\[ G(q^{-1}) = \frac{10q^{-1} + 0.5q^{-2}}{1 - 1.5q^{-1} + 0.7q^{-2}} \]

with the equation error \( e(t) \) being given by

\[ e(t) = (1 - 1.0q^{-1} + 0.2q^{-2})u(t), \]

Here, the input \( u(t) \) is a zero-mean unit-variance white noise sequence and uncorrelated with \( w(t) \), which is white noise of zero mean with variance \( \sigma_w^2 \). This is a standard example in system identification (see, e.g., [5] and [8]). Two cases of the choice of \( \sigma_w^2 \) are considered. One is \( \sigma_w^2 = 1 \), leading to SNR \( \approx 10.92 \) dB. Table IV shows the simulation results by employing these six identification algorithms when \( N = 100 \) and \( N = 500 \). The other choice is \( \sigma_w^2 = 5 \), yielding SNR \( \approx 3.93 \) dB. Table V displays the results by applying these six different algorithms with \( N = 400 \) and \( N = 2000 \). Note that for this second-order system, the PBELS method employs a prefilter designed as \( F(q^{-1}) = 1/(1 - 1.7q^{-1} + 0.72q^{-2}) \).

As shown in Tables IV and V, the ordinary LS estimates are seriously biased and with high variance, especially in the presence of high noise. In spite of its good accuracy, the implementation of the OE method and the OIV method is most expensive computationally. Again, not only does the SIV method perform worst in terms of estimation accuracy among the five unbiased identification algorithms, but in addition, the cases of \( N = 100 \) in Table IV and of \( N = 400 \) in Table V illustrate that it is unable to give consistent estimates when the sample length is relatively short and/or the SNR is relatively low. The two BELS-based algorithms both yield good results in all the situations considered. Further, it is observed that by comparison with the PBELS method, an over 20% reduction in the computational load is achieved by use of the DBELS algorithm, which has no effect on the consistency results. Although, in this example, the price paid for the computational reduction is a fairly moderate increase in the ACV values, the accuracy of the DBELS estimates is still very acceptable from the practical point of view.

VI. Final Remarks

A direct BELS algorithm has been proposed and analyzed for transfer function identification of LTI systems subject to colored disturbances. A significant feature with the presented algorithm is that a direct consistent estimate of the transfer function parameters can be obtained, without prefiltering observed data and without handling a high-order augmented system. It has been demonstrated that except for its ability to accommodate the situations with arbitrarily colored noise, the DBELS algorithm has a simpler algorithmic structure and requires fewer computations than the PBELS method, without sacrificing too much of the estimation accuracy. Owing to these prominent algorithmic advantages, the proposed direct algorithm can be used as a replacement for the implementation of the BELS-based methods in many situations. Furthermore, our simulation results have indicated that compared with the SIV method, the DBELS algorithm can be more accurate in terms of relatively low variance and can work more efficiently in the cases of short sample lengths and/or low SNR values.

REFERENCES


Wei Xing Zheng (M'93) was born in Nanjing, China. He received the B.Sc. degree in applied mathematics and the M.Sc. and Ph.D. degrees in electrical engineering in January 1982, July 1984, and February 1989, respectively, all from the Southeast University, Nanjing.

From 1984 to 1991, he was with the Institute of Automation at the Southeast University, first as a Lecturer and later as an Associate Professor. From 1991 to 1994, he was a Research Fellow with the Department of Electrical and Electronic Engineering and the Interdisciplinary Research Centre for Process Systems Engineering, Imperial College of Science, Technology, and Medicine, University of London, London, U.K.; in the Department of Mathematics, University of Western Australia, Perth, Australia; and in the Australian Telecommunications Research Institute, Curtin University of Technology, Perth, Australia. He joined the Department of Mathematics, University of Western Sydney, Nepean, Sydney, Australia, as a Lecturer in 1994 and has been a Senior Lecturer in the same department since 1996. His research interests are in the areas of systems and controls, signal processing, and operations research. He coauthored the book Linear Multivariable Systems: Theory and Design (Nanjing, China: SEU, 1991).

Dr. Zheng has received several science prizes, including the Chinese National Natural Science Prize awarded by the Chinese Government in 1991.