UNBIASED RLS IDENTIFICATION OF ERRORS-IN-VARIABLES MODELS IN THE PRESENCE OF CORRELATED NOISE

Reza Arablouei[1], *Kutluyıl Doğançay*[2], *and Tülay Adalı*[3]

[12]Institute for Telecommunications Research, University of South Australia, Mawson Lakes SA, Australia

[3]Dept. of Computer Science and Electrical Engineering, University of Maryland, Baltimore County, USA

ABSTRACT

We propose an unbiased recursive-least-squares(RLS)-type algorithm for errors-in-variables system identification when the input noise is colored and correlated with the output noise. To derive the proposed algorithm, which we call unbiased RLS (URLS), we formulate an exponentially-weighted least-squares problem that yields an unbiased estimate. Then, we solve the associated normal equations utilizing the dichotomous coordinate-descent iterations. Simulation results show that the estimation performance of the proposed URLS algorithm is similar to that of a previously proposed bias-compensated RLS (BCRLS) algorithm. However, the URLS algorithm has appreciably lower computational complexity as well as improved numerical stability compared with the BCRLS algorithm.

***Index Terms*—**Adaptive estimation, dichotomous coordinate-descent algorithm, errors-in-variables modeling, recursive least-squares, system identification.

1. INTRODUCTION

Errors-in-variables (EIV) models represent linear systems in which both input and output data are corrupted with noise [1]-[4]. Three popular estimation methods that are widely used to identify parameters of the EIV models are the instrumental variables (IV) [5]-[7], total least-squares (TLS) [8]-[10], and least-squares (LS) with bias compensation [11]-[14].

The IV estimation techniques are computationally simple but typically have poor performance compared with the other techniques [15]-[17]. The TLS approach estimates the parameters of an EIV model by fitting the input data to the output data with minimum perturbation. The TLS estimate is unbiased only when the noises on input and output are zero-mean i.i.d. and mutually uncorrelated [9].

The LS methods are biased in the presence of the input noise. However, the bias can be calculated and compensated for by subtracting it from the biased LS estimate. Based on this concept, several bias-compensated LS (BCLS) algorithms, e.g., [18]-[20], and bias-compensated recursive LS (BCRLS) algorithms, e.g., [21], [22], have been devised. The BCLS and BCRLS algorithms proposed in [20] and [22], respectively, have specifically been tailored for the case where there is correlation among the entries of the input noise vector (colored input noise) as well as between the input noise vector and the output noise. This colored-and-correlated-noise case occurs in several signal processing applications related to engineering or econometrics. One example is the identification of the coefficients of an autoregressive predictive model in the context of speech signal analysis where an additive colored noise may contaminate the input data [22].

The BCRLS algorithm of [22], which we will simply refer to as *the BCRLS algorithm*, at each iteration, computes a recursive LS (RLS) estimate of the sought-after parameter vector along with an estimate of the noise-induced bias. It then subtracts the calculated bias from the RLS estimate to yield a bias-eliminated RLS estimate. The BCRLS algorithm calculates the estimation bias using the prior knowledge of the covariance matrix of the input noise vector and the cross-correlation vector between the input noise vector and the output noise.

In this paper, we propose an algorithm, termed unbiased RLS (URLS), for estimating the parameters of an EIV model when the input noise is colored and correlated with the output noise. Unlike the bias-compensation-based algorithms, the URLS algorithm does not explicitly eliminate the estimation bias by subtracting an evaluation of the bias from a biased estimate. Instead, it incorporates the knowledge of the input noise covariance matrix and the input-output noise cross-correlation vector into the recursive estimation process to produce an unbiased estimate. Similar to [22], we assume that the required noise statistical quantities can be estimated or are known *a priori*.

To derive the URLS algorithm, we define an exponentially-weighted LS optimization problem whose solution is an unbiased estimate. We solve the system of linear equations (SLE) associated with the normal equations of the defined problem utilizing the dichotomous coordinate-descent (DCD) iterations [23]. The DCD algorithm is a shift-and-add algorithm that solves an SLE with no multiplication operation. In this algorithm, the number of exercised iterations and the step-size resolution establish a trade-off between accuracy and complexity [24], [25].

Through simulations, we verify that the estimation performance of the URLS algorithm is very close to that of the BCRLS algorithm. However, the URLS algorithm is significantly less complex than the BCRLS algorithm in terms of the required arithmetic operations. The URLS algorithm is also immune to numerical instability caused by the buildup of round-off errors in finite-precision implementations, which can plague the BCRLS algorithm.

2. ALGORITHM DERIVATION

Consider a linear system described by

$$d_k = \mathbf{h}^T \mathbf{x}_k \tag{1}$$

where $\mathbf{h} \in \mathbb{R}^{L\times 1}$ is the column vector of the system parameters, $L \in \mathbb{N}$ is the order of the system, superscript T denotes matrix transposition, and $\mathbf{x}_k \in \mathbb{R}^{L\times 1}$ and $d_k \in \mathbb{R}$ are the input vector and the output at time index $k \in \mathbb{N}$, respectively. Instead of $\mathbf{x}_k$ and d_k, we observe their noisy versions, i.e.,

$$\tilde{\mathbf{x}}_k = \mathbf{x}_k + \mathbf{n}_k \tag{2}$$

and

$$\tilde{d}_k = d_k + \nu_k \tag{3}$$

where $\mathbf{n}_k \in \mathbb{R}^{L\times 1}$ and $\nu_k \in \mathbb{R}$ represent the corresponding input and output noises.

We make the following assumptions regarding the noises and the input data:

A1: The input noise vector, $\mathbf{n}_k$, is i.i.d. with zero mean and covariance matrix of $\widehat{\mathbf{R}}$, i.e.,

$$E[\mathbf{n}_k] = \mathbf{0} \text{ and } E[\mathbf{n}_k\mathbf{n}_k^T] = \widehat{\mathbf{R}}$$

where $\mathbf{0}$ is the $L \times 1$ zero vector.

A2: The output noise, ν_k, is i.i.d. with zero mean and variance of η, i.e.,

$$E[\nu_k] = 0 \text{ and } E[\nu_k^2] = \eta.$$

A3: The input and output noises are mutually correlated with a cross-correlation vector of $\widehat{\mathbf{p}}$, i.e.,

$$E[\mathbf{n}_k \nu_k] = \widehat{\mathbf{p}}.$$

A4: The input noise is statistically independent of the noiseless input.

We also define

$$\mathbf{X}_k = [\mathbf{x}_k, \lambda^{1/2}\mathbf{x}_{k-1}, \dots, \lambda^{(k-1)/2}\mathbf{x}_1],$$
$$\mathbf{N}_k = [\mathbf{n}_k, \lambda^{1/2}\mathbf{n}_{k-1}, \dots, \lambda^{(k-1)/2}\mathbf{n}_1],$$
$$\widetilde{\mathbf{X}}_k = [\tilde{\mathbf{x}}_k, \lambda^{1/2}\tilde{\mathbf{x}}_{k-1}, \dots, \lambda^{(k-1)/2}\tilde{\mathbf{x}}_1],$$
$$\mathbf{v}_k = [\nu_k, \lambda^{1/2}\nu_{k-1}, \dots, \lambda^{(k-1)/2}\nu_1],$$
$$\tilde{\mathbf{d}}_k = [\tilde{d}_k, \lambda^{1/2}\tilde{d}_{k-1}, \dots, \lambda^{(k-1)/2}\tilde{d}_1],$$
$$\mathbf{\Phi}_k = \widetilde{\mathbf{X}}_k\widetilde{\mathbf{X}}_k^T,$$

and

$$\mathbf{z}_k = \widetilde{\mathbf{X}}_k\tilde{\mathbf{d}}_k^T$$

where λ is a forgetting factor that satisfies

$$0 \ll \lambda < 1.$$

By discounting the input noise and its effect, an *unbiased* exponentially-weighted LS estimate of $\mathbf{h}$ at time instant k, denoted by $\mathbf{w}_k \in \mathbb{R}^{L\times 1}$, can be found as

$$\mathbf{w}_k = \arg\min_{\mathbf{w}} \left\|\tilde{\mathbf{d}}_k - \mathbf{w}^T\mathbf{X}_k\right\|^2$$

where $\|\cdot\|$ denotes the Euclidean norm. The vector $\mathbf{w}_k$ can be considered as the weight vector of an adaptive filter and is the solution of the associated normal equations expressed as [26]

$$\mathbf{X}_k\mathbf{X}_k^T\mathbf{w}_k = \mathbf{X}_k\tilde{\mathbf{d}}_k^T. \tag{4}$$

Using the above definitions together with

$$\mathbf{X}_k = \widetilde{\mathbf{X}}_k - \mathbf{N}_k$$

and

$$\tilde{\mathbf{d}}_k = \mathbf{h}^T\mathbf{X}_k + \mathbf{v}_k,$$

(4) can be written as

$$(\mathbf{\Phi}_k - \mathbf{X}_k\mathbf{N}_k^T - \mathbf{N}_k\mathbf{X}_k^T - \mathbf{N}_k\mathbf{N}_k^T)\mathbf{w}_k = \mathbf{z}_k - \mathbf{N}_k\mathbf{X}_k^T\mathbf{h} - \mathbf{N}_k\mathbf{v}_k^T. \tag{5}$$

In view of A1-A4, which imply that the noises are stationary and correlation-ergodic, we have

$$\mathbf{X}_k\mathbf{N}_k^T = \sum_{i=1}^{k} \lambda^{k-i}\mathbf{x}_i\mathbf{n}_i^T \approx \mathbf{O}, \tag{6}$$

$$\mathbf{N}_k\mathbf{N}_k^T = \sum_{i=1}^{k} \lambda^{k-i}\mathbf{n}_i\mathbf{n}_i^T \approx a_k\widehat{\mathbf{R}}, \tag{7}$$

and

$$\mathbf{N}_k\mathbf{v}_k^T = \sum_{i=1}^{k} \lambda^{k-i}\mathbf{n}_i\nu_i \approx a_k\widehat{\mathbf{p}} \tag{8}$$

where $\mathbf{O}$ is the $L \times L$ zero matrix and

$$a_k = \sum_{i=1}^{k} \lambda^{k-i} = \frac{1-\lambda^k}{1-\lambda}.$$

Substituting (6)-(8) into (5) alongside defining

$$\mathbf{\Psi}_k = \mathbf{\Phi}_k - a_k\widehat{\mathbf{R}} \tag{9}$$

TABLE I

THE DCD ALGORITHM SOLVING $\boldsymbol{\Psi}_k \mathbf{w}_k = \mathbf{s}_k$

initialize
$\quad m = 1, \alpha = H/2$
$\quad \mathbf{w}_k = \mathbf{0}$
$\quad \mathbf{r}_k = \mathbf{s}_k$
for $j = 1,2,\dots,N$
$\quad p = \arg\max_{i=1,\dots,L}\{\lvert r_k^i \rvert\}$
$\quad$ while $\lvert r_k^p \rvert \le \frac{\alpha}{2}\Psi_k^{p,p}$ and $m \le M$
$\quad\quad m = m + 1, \alpha = \alpha/2$
$\quad$ if $m > M$
$\quad\quad$ algorithm stops
$\quad w_k^p = w_k^p + \mathrm{sign}\{r_k^p\}\alpha$
$\quad \mathbf{r}_k = \mathbf{r}_k - \mathrm{sign}\{r_k^p\}\alpha\boldsymbol{\Psi}_k^{(p)}$

and

$$\mathbf{s}_k = \mathbf{z}_k - a_k \widehat{\mathbf{p}} \tag{10}$$

gives

$$\boldsymbol{\Psi}_k \mathbf{w}_k = \mathbf{s}_k. \tag{11}$$

From the definitions of $\boldsymbol{\Phi}_k$, $\mathbf{z}_k$, and a_k, we have

$$\boldsymbol{\Phi}_k = \lambda \boldsymbol{\Phi}_{k-1} + \tilde{\mathbf{x}}_k \tilde{\mathbf{x}}_k^T,$$

$$\mathbf{z}_k = \lambda \mathbf{z}_{k-1} + \tilde{\mathbf{x}}_k \tilde{d}_k,$$

and

$$a_k = \lambda a_{k-1} + 1.$$

Using these equations together with (9) and (10), we obtain the following recursive update equations:

$$\boldsymbol{\Psi}_k = \lambda \boldsymbol{\Psi}_{k-1} + \tilde{\mathbf{x}}_k \tilde{\mathbf{x}}_k^T - \widehat{\mathbf{R}}$$

and

$$\mathbf{s}_k = \lambda \mathbf{s}_{k-1} + \tilde{\mathbf{x}}_k \tilde{d}_k - \widehat{\mathbf{p}}.$$

In order to find $\mathbf{w}_k$, we solve the SLE of (11) using the DCD iterations [24] given in Table I where a_k^p denotes the pth element of a vector $\mathbf{a}_k$ while $\Psi_k^{p,p}$ and $\boldsymbol{\Psi}_k^{(p)}$ are the (p,p)th entry and the pth column of the matrix $\boldsymbol{\Psi}_k$, respectively. Three design parameters, viz., $N \in \mathbb{N}$, $M \in \mathbb{N}$, and $H \in \mathbb{R}$, govern the accuracy and complexity of the DCD algorithm [23]. The algorithm renders maximum N iterative updates at each run while the entries of the solution vector, $\mathbf{w}_k$, are represented as fixed-point words with M bits within an amplitude range of $[-H, H]$.

We summarize the proposed *unbiased recursive least-squares* (URLS) algorithm in Table II.

3. COMPUTATIONAL COMPLEXITY

Since $\boldsymbol{\Psi}_k$ is symmetric, it is sufficient to update only its upper-triangular part. Moreover, if we select the forgetting factor as $\lambda = 1 - 2^{-c}$ where c is a positive integer, we can replace multiplications by λ with additions and bit-shifts [24].

TABLE II

THE URLS ALGORITHM

initialize
$\quad \boldsymbol{\Psi}_0 = \delta \mathbf{I}$
$\quad \mathbf{s}_0 = \mathbf{0}$
at iteration $k = 1,2,\dots$
$\quad \boldsymbol{\Psi}_k = \lambda \boldsymbol{\Psi}_{k-1} + \tilde{\mathbf{x}}_k \tilde{\mathbf{x}}_k^T - \widehat{\mathbf{R}}$
$\quad \mathbf{s}_k = \lambda \mathbf{s}_{k-1} + \tilde{\mathbf{x}}_k \tilde{d}_k - \widehat{\mathbf{p}}$
$\quad$ solve $\boldsymbol{\Psi}_k \mathbf{w}_k = \mathbf{s}_k$ to find $\mathbf{w}_k$

TABLE III

COMPUTATIONAL COMPLEXITY OF THE BCRLS AND URLS ALGORITHMS IN TERMS OF NUMBER OF REQUIRED ARITHMETIC OPERATIONS PER ITERATION

	$\times$	$+$	$/$
BCRLS	$5L^2 + 5L$	$4L^2 + 2L + 2$	1
URLS	$0.5L^2 + 1.5L$	$1.5L^2 + (4.5 + 2N)L + N + M$	0

In Table III, we present the number of required arithmetic operations per iteration by the BCRLS and URLS algorithms. The ratio of the number of multiplications required by the URLS algorithm at each iteration to that of the BCRLS algorithm for any value of L is bounded as

$$0.1 \le \frac{0.5L^2 + 1.5L}{5L^2 + 5L} \le 0.2.$$

Therefore, the URLS algorithm requires 80 to 90% fewer multiplications per iteration than the BCRLS algorithm. Furthermore, unlike the BCRLS algorithm, the URLS algorithm does not require any division operation that is usually expensive to implement in both hardware and software [27].

4. RELATION TO THE BCRLS ALGORITHM

Assuming that the DCD algorithm is adequately accurate, we can consider the filter weight vector of the URLS algorithm to be the exact solution of the LSE of (11), which can be rearranged as

$$\boldsymbol{\Phi}_k \mathbf{w}_k = \mathbf{z}_k + a_k \widehat{\mathbf{R}} \mathbf{w}_k - a_k \widehat{\mathbf{p}}.$$

Multiplying both sides of this equation by $\boldsymbol{\Phi}_k^{-1}$ gives

$$\mathbf{w}_k = \boldsymbol{\Phi}_k^{-1} \mathbf{z}_k + a_k \boldsymbol{\Phi}_k^{-1} \big(\widehat{\mathbf{R}} \mathbf{w}_k - \widehat{\mathbf{p}}\big). \tag{12}$$

Comparing (12) with the update equation of the BCRLS algorithm, i.e.,

$$\mathbf{w}_k = \boldsymbol{\Phi}_k^{-1} \mathbf{z}_k + a_k \boldsymbol{\Phi}_k^{-1} \big(\widehat{\mathbf{R}} \mathbf{w}_{k-1} - \widehat{\mathbf{p}}\big),$$

shows that the update equations of the URLS and BCRLS algorithms are almost identical. The only difference is whether the current or previous filter weight vector is used in the bias correction term. Accordingly, it is expected that the two algorithms exhibit comparable estimation performance. Nonetheless, this seemingly minor dissimilarity in the update terms makes a significant

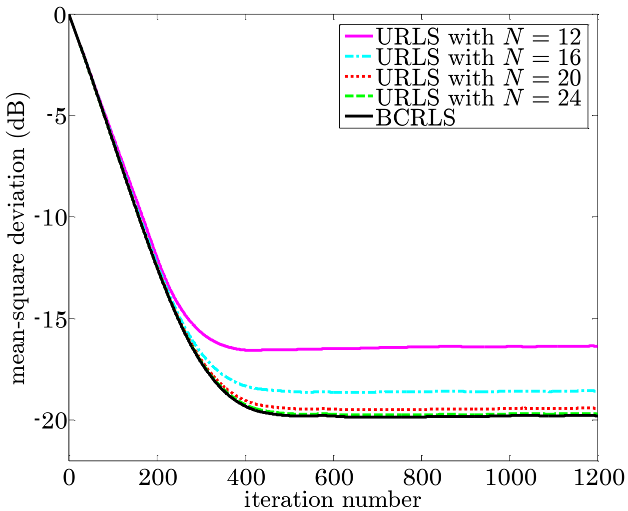

Fig. 1. Learning curves of the BCRLS algorithm and the URLS algorithm with different values of N.

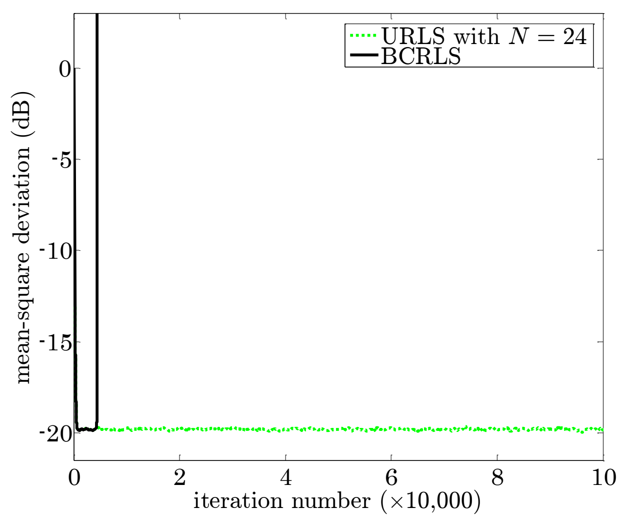

Fig. 2. Learning curves of the BCRLS algorithm and the URLS algorithm with $N = 24$ for 10^5 iterations.

difference in terms of computational complexity and numerical stability by enabling effective utilization of the DCD iterations in the URLS algorithm.

5. SIMULATIONS

Consider an errors-in-variables system identification problem where the system parameter vector has $L = 8$ arbitrary entries. The noiseless input vector, $\mathbf{x}_k$, is zero-mean i.i.d. multivariate Gaussian with covariance matrix

$$\mathbf{R} = \mathbf{Q}_L \text{diag}\{\mathbf{e}_L\} \mathbf{Q}_L^T$$

where $\mathbf{Q}_L \in \mathbb{R}^{L \times L}$ is a randomly-generated unitary matrix and the elements of $\mathbf{e}_L \in \mathbb{R}^{L \times 1}$ are drawn from a uniform distribution in the interval $[0.2,1]$. The augmented noise vector, $[\mathbf{n}_k^T, \nu_k]^T$, is also zero-mean i.i.d. multivariate Gaussian with covariance matrix

$$\begin{bmatrix} \widehat{\mathbf{R}} & \widehat{\mathbf{p}} \\ \widehat{\mathbf{p}}^T & \eta \end{bmatrix} = \xi^{-1} \mathbf{Q}_{L+1} \text{diag}\{\mathbf{e}_{L+1}\} \mathbf{Q}_{L+1}^T.$$

We set $c = 7$ ($\lambda \approx 0.9922$), $M = 10$, $H = 1$, and $\xi = 10$.

In Fig. 1, we compare the estimation performance of the BCRLS algorithm and the URLS algorithm with different values of N. The performance measure is the mean-square deviation (MSD), i.e., $E[\|\mathbf{h} - \mathbf{w}_k\|^2]$, which is evaluated by taking the ensemble-average over 10^3 independent runs. Fig. 1 shows that the larger N is, the closer the performance of the URLS algorithm to that of the BCRLS algorithm becomes.

In Fig. 2, we plot the time-evolution of the MSD of the BCRLS algorithm and the URLS algorithm with $N = 24$ for 10^5 iterations. All the variables are represented in double-precision floating-point format. Fig. 2 shows that, even with such high-precision arithmetic, the BCRLS algorithm diverges after about 4300 iterations. However, the URLS algorithm maintains its numerical stability throughout the simulation. In addition, we carried out several numerical experiments in different scenarios for very large numbers of iterations, e.g., 10^7. In agreement with the findings of [24] pertaining to the numerical stability of the DCD-RLS algorithm, during the experiments, we did not observe any numerical instability in the URLS algorithm.

In the scenario simulated here, i.e., for $L = 8$, the BCRLS algorithm requires 274 multiplications per iteration while the URLS algorithm requires 16% of this amount, which is only 44 multiplications per iteration.

6. CONCLUSION

We proposed an unbiased recursive least-squares (URLS) algorithm for identifying the parameter vector of errors-in-variables models when the input noise is colored and correlated with the output noise. To derive the URLS algorithm, we employed the dichotomous coordinate-descent (DCD) iterations to solve the normal equations associated with a least-squares problem that is specifically formulated to yield an unbiased estimate. Simulation results confirmed that the URLS algorithm performs as well as a previously proposed bias-compensated recursive least-squares (BCRLS) algorithm. However, the URLS algorithm executes considerably fewer computations than the BCRLS algorithm. Moreover, unlike the BCRLS algorithm, the URLS algorithm is virtually immune to numerical instability incurred by the accumulation of round-off errors in finite-precision implementations.

REFERENCES

[1] M. Deistler, “Linear dynamic errors-in-variables models,” *J. Appl. Probability*, vol. 23A, pp. 23–39, 1986.

[2] W. A. Fuller, *Measurement error models*, New York: Wiley, 1987.

[3] S. Van Huffel and P. Lemmerling, Eds., *Total Least Squares Techniques and Errors-in-Variables Modelling: Analysis, Algorithms and Applications*, Dordrecht, The Netherlands: Kluwer Academic Publishers, 2002.

[4] T. Söderström, "Errors-in-variables methods in system identification," *Automatica*, vol. 43, pp. 939-958, 2007.

[5] A. Wald, "The fitting of straight lines if both variables are subject to error," *Ann. Math. Stat.*, vol. 11, pp. 284-300, Sep. 1940.

[6] J. D. Sargan, "The estimation of economic relationships using instrumental variables," *Econometrica*, vol. 26, pp. 393-415, Jul. 1958.

[7] J. Pearl, *Causality: Models, Reasoning, and Inference*, Cambridge Univ. Press, 2000.

[8] G. H. Golub and C. F. Van Loan, "An analysis of the total least squares problem," *SIAM J. Numerical Anal.*, vol. 17, no. 6, pp. 883–893, Dec. 1980.

[9] S. Van Huffel and J. Vandewalle, *The Total Least Squares Problem: Computational Aspects and Analysis*, Philadelphia, PA: SIAM, 1991.

[10] I. Markovsky and S. Van Huffel, "Overview of total least-squares methods," *Signal Process.*, vol. 87, pp. 2283-2302, 2007.

[11] P. N. James, P. Souter, and D. C. Dixon, "Suboptimal estimation of the parameters of discrete systems in the presence of correlated noise," *Electron. Lett.*, vol. 8, pp. 411-412, 1972.

[12] S. Sagara and K. Wada, "On-line modified least-squares parameter estimation on linear discrete dynamic systems," *Int. J. Control*, vol. 25, no. 3, pp. 329-343, 1977.

[13] P. Stoica and T. Söderström, "Bias correction in least-squares identification," *Int. J. Control*, vol. 35, pp. 449-457, 1982.

[14] K. Wada, M. Eguchi, and S. Sagara, "Estimation of pulse transfer function via bias-compensated least-squares method in the presence of input and output noise," *Syst. Sci.*, vol. 16, no. 3, pp. 57–70, 1990.

[15] T. Söderström and P. Stoica, *System Identification*, Cambridge, U.K.: Prentice Hall, 1989.

[16] S. Van Huffel and J. Vandewalle, "Comparison of total least squares and instrumental variable methods for parameter estimation of transfer function models," *Int. J. Control*, vol. 50, pp. 1039–1056, 1989.

[17] T. Söderström and K. Mahata, "On instrumental variable and total least squares approaches for identification of noisy systems," *Int. J. Control*, vol. 75, pp. 381–389, 2002.

[18] T. Söderström, "Extending the Frisch scheme for errors-in-variables identification to correlated output noise," *Int. J. Adaptive Control Signal Process.*, vol. 22, pp. 55-73, 2008.

[19] R. Diversi, "A bias-compensated identification approach for noisy FIR models," *IEEE Signal Process. Lett.*, vol. 15, pp. 325-328, 2008.

[20] R. Diversi, "Bias-eliminating least-squares identification of errors-in-variables models with mutually correlated noise," *Int. J. Adapt. Control Signal Process.*, 2012; DOI: 10.1002/acs.2365.

[21] K. Ikeda, Y. Mogami, and T. Shimomura, "Bias compensation of recursive least squares estimate in closed loop environment," *SICE J. Control, Measurement, System Integration*, vol. 1, no. 5, pp. 400–405, Sep. 2008.

[22] A. Bertrand, M. Moonen, and A. H. Sayed, "Diffusion bias-compensated RLS estimation over adaptive networks," *IEEE Trans. Signal Process.*, vol. 59, pp. 5212–5224, Nov. 2011.

[23] Y. V. Zakharov and T. C. Tozer, "Multiplication-free iterative algorithm for LS problem," *Electron. Lett.*, vol. 40, no. 9, pp. 567–569, Apr. 2004.

[24] Y. Zakharov, G. White, and J. Liu, "Low complexity RLS algorithms using dichotomous coordinate descent iterations," *IEEE Trans. Signal Process.*, vol. 56, pp. 3150–3161, Jul. 2008.

[25] J. Liu, Y. V. Zakharov, and B. Weaver, "Architecture and FPGA design of dichotomous coordinate descent algorithms," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 56, pp. 2425–2438, Nov. 2009.

[26] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed., Baltimore, MD: Johns Hopkins Univ. Press, 1996.

[27] K. Doğançay, "Complexity considerations for transform-domain adaptive filters," *Signal Process.*, vol. 83, no. 6, pp. 1177–1192, Jun. 2003.