2008 American Control Conference
Westin Seattle Hotel, Seattle, Washington, USA
June 11-13, 2008

Errors-in-variables identification through covariance matching: Analysis of a colored measurement noise case

Magnus Mossberg

***Abstract*— The method of covariance and cross-covariance matching is considered for errors-in-variables identification where the measurement noises are colored. The covariance and cross-covariance functions are estimated from the noise-corrupted data and the corresponding theoretical functions, parameterized by the unknown parameters, are matched to the estimated functions. The main contribution of the paper is a step-by-step algorithm for the computation of the covariance matrix of the estimated system parameters.**

I. INTRODUCTION

Consider the errors-in-variables (EIV) [1], [2], [3], [4] system

$$\begin{aligned} A(q^{-1})y_0(t) &= B(q^{-1})u_0(t), \\ A(q^{-1}) &= 1 + a_1q^{-1} + \ldots + a_nq^{-n}, \\ B(q^{-1}) &= b_1q^{-1} + \ldots + b_nq^{-n}, \end{aligned}$$

where q^{-1} is the backward shift operator, and where $u_0(t)$ and $y_0(t)$ denote the noise-free input and output signals, respectively. It is assumed that $u_0(t)$ is a stationary stochastic process with rational spectrum and it is therefore represented as

$$\begin{aligned} C(q^{-1})u_0(t) &= D(q^{-1})e(t), \\ C(q^{-1}) &= 1 + c_1q^{-1} + \ldots + c_mq^{-m}, \\ D(q^{-1}) &= d_1q^{-1} + \ldots + d_mq^{-m}, \end{aligned}$$

where the white noise source $e(t)$ is of zero mean and variance λ_e^2. The measurements

$$u(t) = u_0(t) + \tilde{u}(t), \tag{1}$$
$$y(t) = y_0(t) + \tilde{y}(t) \tag{2}$$

are available for $t = 1, \ldots, N$, where $\tilde{u}(t)$ and $\tilde{y}(t)$ are independent colored noise sequences given by

$$\begin{aligned} G(q^{-1})\tilde{u}(t) &= v(t), \\ G(q^{-1}) &= 1 + g_1q^{-1} + \ldots + g_\alpha q^{-\alpha} \end{aligned} \tag{3}$$

and

$$\begin{aligned} H(q^{-1})\tilde{y}(t) &= w(t), \\ H(q^{-1}) &= 1 + h_1q^{-1} + \ldots + h_\beta q^{-\beta}, \end{aligned} \tag{4}$$

where $v(t)$ and $w(t)$ are independent white noise sources of zero mean and variances λ_v^2 and λ_w^2, respectively, independent of $e(t)$. The case with white measurement noises was considered in [5]. The problem is to estimate the unknown system parameters

$$\boldsymbol{\theta}_0 = \begin{bmatrix} a_1 & \cdots & a_n & b_1 & \cdots & b_n \end{bmatrix}^T$$

from the data $\{u(t), y(t)\}_{t=1}^N$. The parameters

$$\boldsymbol{\psi}_0 = \begin{bmatrix} c_1 & \cdots & c_m & d_1 & \cdots & d_m \end{bmatrix}^T$$

are also unknown, but are not of primary interest.

The solution considered in this paper is to estimate covariance and cross-covariance functions from the noise corrupted data and to match the corresponding theoretical functions, parameterized by the unknown parameters, to the estimated functions. The estimate of the cross-covariance function based on colored noise-corrupted measurements $u(t)$ and $y(t)$ is consistent, provided that the measurement noises (1) and (2) are independent. This makes the method of cross-covariance matching an interesting choice for EIV identification. The main contribution of the paper is the derivation of the covariance matrix of the estimate of $\boldsymbol{\theta}_0$ given by the cross-covariance matching method. The expression is approximative and valid for a large number of data N. A detailed description on how to compute the involving elements is given in the paper.

The outline of the paper is as follows. In the next section, some preliminaries regarding the theoretical expressions for the covariance and cross-covariance functions are given together with definitions of some of their estimates and the asymptotic properties of these estimates. The estimators based on covariance and cross-covariance matching are described in Section III, including discussions on consistency of these estimators. Section IV is devoted to the covariance matrix of the estimate of $\boldsymbol{\theta}_0$, and a step-by-step algorithm for its computation is given in Algorithm 2 in the end of the section. An example in which the theoretical variances are compared with empirical variances from a Monte Carlo simulation is presented in Section V, and conclusions are drawn in Section VI.

II. PRELIMINARIES

Some material on covariance and cross-covariance functions, important for the coming sections of the paper, is presented in this section. First, the covariance and cross-covariance functions for the signals $u_0(t)$ and $y_0(t)$ are given. Represent the system from $u_0(t)$ to $y_0(t)$ as

$$\begin{aligned} \mathbf{x}(t+1) &= \mathbf{A}(\boldsymbol{\theta}_0, \boldsymbol{\psi}_0)\mathbf{x}(t) + \mathbf{B}(\boldsymbol{\psi}_0)e(t), \\ \mathbf{z}_0(t) &= \mathbf{C}\mathbf{x}(t) + \mathbf{D}e(t), \end{aligned}$$

M. Mossberg is with the Department of Electrical Engineering, Karlstad University, Sweden. E-mail: Magnus.Mossberg@kau.se

where

$$\mathbf{z}_0(t) = \begin{bmatrix} u_0(t) & y_0(t) \end{bmatrix}^T.$$

The following result can now be stated.

Result 1. *The covariance function* $\mathbf{R}_{\mathbf{z}_0}(\tau, \boldsymbol{\theta}_0, \boldsymbol{\psi}_0)$ *of* $\mathbf{z}_0(t)$ *is given as*

$$\begin{aligned}\mathbf{R}_{\mathbf{z}_0}(\tau, \boldsymbol{\theta}_0, \boldsymbol{\psi}_0) &= \begin{bmatrix} r_{u_0}(\tau, \boldsymbol{\psi}_0) & r_{u_0 y_0}(\tau, \boldsymbol{\theta}_0, \boldsymbol{\psi}_0) \\ r_{y_0 u_0}(\tau, \boldsymbol{\theta}_0, \boldsymbol{\psi}_0) & r_{y_0}(\tau, \boldsymbol{\theta}_0, \boldsymbol{\psi}_0) \end{bmatrix} \\ &= \mathrm{E}\{\mathbf{z}_0(t+\tau)\mathbf{z}_0^T(t)\} \\ &= \begin{cases} \mathbf{C}\mathbf{P}(\boldsymbol{\theta}_0, \boldsymbol{\psi}_0)\mathbf{C}^T + \lambda_e^2 \mathbf{D}\mathbf{D}^T, & \tau = 0, \\ \mathbf{C}\mathbf{A}^{\tau-1}(\boldsymbol{\theta}_0, \boldsymbol{\psi}_0)\big(\mathbf{A}(\boldsymbol{\theta}_0, \boldsymbol{\psi}_0) \\ \quad \cdot \mathbf{P}(\boldsymbol{\theta}_0, \boldsymbol{\psi}_0)\mathbf{C}^T + \lambda_e^2 \mathbf{B}(\boldsymbol{\psi}_0)\mathbf{D}^T\big), & \tau > 0, \end{cases}\end{aligned} \tag{5}$$

where $\mathbf{P}(\boldsymbol{\theta}_0, \boldsymbol{\psi}_0)$ *is the unique and non-negative definite solution to the Lyapunov equation*

$$\begin{aligned}\mathbf{P}(\boldsymbol{\theta}_0, \boldsymbol{\psi}_0) = \mathbf{A}(\boldsymbol{\theta}_0, \boldsymbol{\psi}_0)\mathbf{P}(\boldsymbol{\theta}_0, \boldsymbol{\psi}_0)\mathbf{A}^T(\boldsymbol{\theta}_0, \boldsymbol{\psi}_0) \\ + \lambda_e^2 \mathbf{B}(\boldsymbol{\psi}_0)\mathbf{B}^T(\boldsymbol{\psi}_0).\end{aligned} \tag{6}$$

Proof: See [6]. ■

An estimate of $\mathbf{R}_{\mathbf{z}_0}(\tau, \boldsymbol{\theta}_0, \boldsymbol{\psi}_0)$ is suggested in the following proposition.

Proposition 1. *If the mean values of* $u(t)$ *and* $y(t)$ *are zero, a possible estimator* $\hat{\mathbf{R}}_{\mathbf{z}}(\tau)$ *of* $\mathbf{R}_{\mathbf{z}_0}(\tau, \boldsymbol{\theta}_0, \boldsymbol{\psi}_0)$ *is*

$$\begin{aligned}\hat{\mathbf{R}}_{\mathbf{z}}(\tau) &= \begin{bmatrix} \hat{r}_u(\tau) & \hat{r}_{uy}(\tau) \\ \hat{r}_{yu}(\tau) & \hat{r}_y(\tau) \end{bmatrix} \\ &= \frac{1}{N-\tau}\sum_{t=1}^{N-\tau} \mathbf{z}(t+\tau)\mathbf{z}^T(t), \quad \tau \geqslant 0,\end{aligned}$$

where

$$\mathbf{z}(t) = \begin{bmatrix} u(t) & y(t) \end{bmatrix}^T.$$

The data are stationary and ergodic under the given assumptions, and

$$\begin{aligned}\mathrm{E}\{y_0(t_1)\tilde{u}(t_2)\} &= 0, \quad \forall t_1, t_2, \\ \mathrm{E}\{\tilde{y}(t_1)u_0(t_2)\} &= 0, \quad \forall t_1, t_2, \\ \mathrm{E}\{\tilde{y}(t_1)\tilde{u}(t_2)\} &= 0, \quad \forall t_1, t_2,\end{aligned}$$

so

$$\begin{aligned}\lim_{N\to\infty} \hat{r}_{yu}(\tau) &= r_{y_0 u_0}(\tau, \boldsymbol{\theta}_0, \boldsymbol{\psi}_0), \quad \tau \geqslant 0, \\ \lim_{N\to\infty} \hat{r}_{uy}(\tau) &= r_{u_0 y_0}(\tau, \boldsymbol{\theta}_0, \boldsymbol{\psi}_0), \quad \tau \geqslant 0.\end{aligned} \tag{7}$$

For the diagonal elements of $\hat{\mathbf{R}}_{\mathbf{z}}(\tau)$, it holds that

$$\begin{aligned}\lim_{N\to\infty} \hat{r}_u(\tau) &= r_{u_0}(\tau, \boldsymbol{\psi}_0) + r_{\tilde{u}}(\tau, \boldsymbol{\gamma}_0), \quad \tau \geqslant 0, \\ \lim_{N\to\infty} \hat{r}_y(\tau) &= r_{y_0}(\tau, \boldsymbol{\theta}_0, \boldsymbol{\psi}_0) + r_{\tilde{y}}(\tau, \boldsymbol{\kappa}_0), \quad \tau \geqslant 0,\end{aligned} \tag{8}$$

where $r_{\tilde{u}}(\tau, \boldsymbol{\gamma}_0)$ and $r_{\tilde{y}}(\tau, \boldsymbol{\kappa}_0)$ are, respectively, the covariance functions of $\tilde{u}(t)$ and $\tilde{y}(t)$. Here,

$$\boldsymbol{\gamma}_0 = \begin{bmatrix} g_1 & \cdots & g_\alpha \end{bmatrix}$$

and

$$\boldsymbol{\kappa}_0 = \begin{bmatrix} h_1 & \cdots & h_\beta \end{bmatrix}$$

contain the filter parameters of (3) and (4), respectively. This information about the asymptotic properties of the estimators of the covariance and cross-covariance functions is needed in the discussion on consistency of the proposed estimators of $\boldsymbol{\psi}_0$ and $\boldsymbol{\theta}_0$ in Section III as well as in the expression for the covariance matrix for the estimate of $\boldsymbol{\theta}_0$ in Section IV.

The covariance functions $r_{\tilde{u}}(\tau, \boldsymbol{\gamma}_0)$ and $r_{\tilde{y}}(\tau, \boldsymbol{\kappa}_0)$ are found next. To find $r_{\tilde{u}}(\tau, \boldsymbol{\gamma}_0)$, represent $\tilde{u}(t)$ in state space form as

$$\begin{aligned}\mathbf{x}_g(t+1) &= \mathbf{A}_g(\boldsymbol{\gamma}_0)\mathbf{x}_g(t) + \mathbf{B}_g v(t), \\ \tilde{u}(t) &= \mathbf{C}_g \mathbf{x}_g(t).\end{aligned}$$

The covariance function is given by the following result.

Result 2. *The covariance function* $r_{\tilde{u}}(\tau, \boldsymbol{\gamma}_0)$ *of* $\tilde{u}(t)$ *is given as*

$$r_{\tilde{u}}(\tau, \boldsymbol{\gamma}_0) = \mathbf{C}_g \mathbf{A}_g^\tau(\boldsymbol{\gamma}_0)\mathbf{P}_g(\boldsymbol{\gamma}_0)\mathbf{C}_g^T, \quad \tau \geqslant 0,$$

where $\mathbf{P}_g(\boldsymbol{\gamma}_0)$ *is the unique and non-negative definite solution to the Lyapunov equation*

$$\mathbf{P}_g(\boldsymbol{\gamma}_0) = \mathbf{A}_g(\boldsymbol{\gamma}_0)\mathbf{P}_g(\boldsymbol{\gamma}_0)\mathbf{A}_g^T(\boldsymbol{\gamma}_0) + \lambda_v^2 \mathbf{B}_g \mathbf{B}_g^T.$$

Proof: The result follows from Result 1. ■

Alternatively, consider the Yule-Walker equation

$$\begin{aligned}r_{\tilde{u}}(\tau, \boldsymbol{\gamma}_0) + g_1 r_{\tilde{u}}(\tau-1, \boldsymbol{\gamma}_0) + \ldots + g_\alpha r_{\tilde{u}}(\tau-\alpha, \boldsymbol{\gamma}_0) \\ = \begin{cases} 0, & \tau > 0, \\ \lambda_v^2, & \tau = 0 \end{cases}\end{aligned}$$

for $\tau = 0, \ldots, \alpha$, provided that λ_v^2 is known, in order to determine $r_{\tilde{u}}(0, \boldsymbol{\gamma}_0), \ldots, r_{\tilde{u}}(\alpha, \boldsymbol{\gamma}_0)$. The Yule-Walker equation can then be iterated for $\tau > \alpha$ to find $r_{\tilde{u}}(\tau, \boldsymbol{\gamma}_0)$, $\tau > \alpha$.

Analogously, to find $r_{\tilde{y}}(\tau, \boldsymbol{\kappa}_0)$, first represent $\tilde{y}(t)$ in state space form as

$$\begin{aligned}\mathbf{x}_h(t+1) &= \mathbf{A}_h(\boldsymbol{\kappa}_0)\mathbf{x}_h(t) + \mathbf{B}_h w(t), \\ \tilde{y}(t) &= \mathbf{C}_h \mathbf{x}_h(t).\end{aligned}$$

The covariance function is given by the following result.

Result 3. *The covariance function* $r_{\tilde{y}}(\tau, \boldsymbol{\kappa}_0)$ *of* $\tilde{y}(t)$ *is found by solving the Lyapunov equation*

$$\mathbf{P}_h(\boldsymbol{\kappa}_0) = \mathbf{A}_h(\boldsymbol{\kappa}_0)\mathbf{P}_h(\boldsymbol{\kappa}_0)\mathbf{A}_h^T(\boldsymbol{\kappa}_0) + \lambda_w^2 \mathbf{B}_h \mathbf{B}_h^T,$$

and computing

$$r_{\tilde{y}}(\tau, \boldsymbol{\kappa}_0) = \mathbf{C}_h \mathbf{A}_h^\tau(\boldsymbol{\kappa}_0)\mathbf{P}_h(\boldsymbol{\kappa}_0)\mathbf{C}_h^T, \quad \tau \geqslant 0.$$

Proof: The result follows from Result 1. ■

III. Estimation

In this section, estimators of ψ_0 and θ_0 based on covariance and cross-covariance matching are suggested. Two different estimators for ψ_0 and one estimator for θ_0 are given, and consistency of these estimators are discussed. It is assumed that $\hat{\mathbf{R}}_{\mathbf{z}}(\tau)$ from Proposition 1 is available for $\tau = 0, \ldots, \ell$.

Proposition 2. *Define the loss function*

$$V(\psi) = \sum_{\tau=j_1}^{j_2} \big(\hat{r}_u(\tau) - r_{u_0}(\tau, \psi)\big)^2,$$

where $0 \leqslant j_1 \leqslant j_2 \leqslant \ell$, *from which an estimate* $\hat{\psi}$ *is obtained as*

$$\hat{\psi} = \arg\min_{\psi} V(\psi). \tag{9}$$

The estimate $\hat{\psi}$ from (9) is not consistent due to the colored measurement noise. Consider

$$\lim_{N\to\infty} V(\psi) = \sum_{\tau=j_1}^{j_2} \{ \lim_{N\to\infty} \hat{r}_u^2(\tau) + r_{u_0}^2(\tau, \psi) - 2\big(r_{u_0}(\tau, \psi_0) + r_{\tilde{u}}(\tau, \gamma_0)\big) r_{u_0}(\tau, \psi)\}, \tag{10}$$

where (8) is used. Note that nothing is said about $\lim_{N\to\infty} \hat{r}_u^2(\tau)$ since this information is not needed. It is seen that (10) is minimized by

$$\big(r_{u_0}(\tau, \psi)\big)_{\min} = r_{u_0}(\tau, \psi_0) + r_{\tilde{u}}(\tau, \gamma_0),$$

i.e., not by $\psi = \psi_0$. This means that

$$\big|\big(r_{u_0}(\tau, \psi)\big)_{\min} - r_{u_0}(\tau, \psi_0)\big| = |r_{\tilde{u}}(\tau, \gamma_0)|$$

in the limiting case. The following proposition is an alternative if α in (3) is known.

Proposition 3. *Define the loss function*

$$S(\psi, \gamma) = \sum_{\tau=j_1}^{j_2} \big(\hat{r}_u(\tau) - r_{u_0}(\tau, \psi) - r_{\tilde{u}}(\tau, \gamma)\big)^2,$$

where $0 \leqslant j_1 \leqslant j_2 \leqslant \ell$, *from which estimates* $\hat{\psi}$ *and* $\hat{\gamma}$ *are obtained as*

$$\{\hat{\psi}, \hat{\gamma}\} = \arg\min_{\psi, \gamma} S(\psi, \gamma). \tag{11}$$

The estimates $\hat{\psi}$ and $\hat{\gamma}$ from (11) are consistent since

$$\begin{aligned}\lim_{N\to\infty} S(\psi, \gamma) = \sum_{\tau=j_1}^{j_2} \{ &\lim_{N\to\infty} \hat{r}_u^2(\tau) + r_{u_0}^2(\tau, \psi) + r_{\tilde{u}}^2(\tau, \gamma) \\ &- 2\big(r_{u_0}(\tau, \psi_0) + r_{\tilde{u}}(\tau, \gamma_0)\big) r_{u_0}(\tau, \psi) \\ &- 2\big(r_{u_0}(\tau, \psi_0) + r_{\tilde{u}}(\tau, \gamma_0)\big) r_{\tilde{u}}(\tau, \gamma) \\ &+ 2 r_{u_0}(\tau, \psi) r_{\tilde{u}}(\tau, \gamma)\}\end{aligned} \tag{12}$$

is minimized by

$$\big(r_{u_0}(\tau, \psi)\big)_{\min} = r_{u_0}(\tau, \psi_0)$$

and

$$\big(r_{\tilde{u}}(\tau, \gamma)\big)_{\min} = r_{\tilde{u}}(\tau, \gamma_0),$$

i.e., by $\psi = \psi_0$ and $\gamma = \gamma_0$. In (12), just as in (10), (8) is used and information about $\lim_{N\to\infty} \hat{r}_u^2(\tau)$ is not needed.

After an estimate $\hat{\psi}$ is obtained, for example as described in Proposition 2 or in Proposition 3, an estimate of θ_0 can be found as described next.

Proposition 4. *Consider the loss function*

$$W(\theta) = \sum_{\tau=0}^{k} \big(\hat{r}_{yu}(\tau) - r_{y_0 u_0}(\tau, \theta, \hat{\psi})\big)^2, \tag{13}$$

where $k \leqslant \ell$, *from which an estimate* $\hat{\theta}$ *is given as*

$$\hat{\theta} = \arg\min_{\theta} W(\theta). \tag{14}$$

It holds that the estimate $\hat{\theta}$ from (14) is consistent, provided that $\hat{\psi}$ is a consistent estimate of ψ_0. Consider

$$\begin{aligned}\lim_{N\to\infty} W(\theta) = \sum_{\tau=0}^{k} \{ &\lim_{N\to\infty} \hat{r}_{yu}^2(\tau) + r_{y_0 u_0}^2(\tau, \theta, \hat{\psi}) \\ &- 2 r_{y_0 u_0}(\tau, \theta_0, \psi_0) r_{y_0 u_0}(\tau, \theta, \hat{\psi})\},\end{aligned} \tag{15}$$

where (7) is used. Here, information about $\lim_{N\to\infty} \hat{r}_{yu}^2(\tau)$ is not needed. It holds that (15) is uniquely minimized by

$$\big(r_{y_0 u_0}(\tau, \theta, \hat{\psi})\big)_{\min} = r_{y_0 u_0}(\tau, \theta_0, \psi_0),$$

i.e., by $\theta = \theta_0$, provided that $\hat{\psi}$ is a consistent estimate of ψ_0.

The estimation method is now summarized in Algorithm 1.

Algorithm 1. *Summary of the estimation method.*

1) *Estimate the covariance function* $\mathbf{R}_{\mathbf{z}_0}(\tau, \theta_0, \psi_0)$ *as described in Proposition 1.*
2) *Compute the estimate* $\hat{\psi}$ *as suggested in Proposition 2 or in Proposition 3.*
3) *Compute the estimate* $\hat{\theta}$ *as described in Proposition 4.*

IV. Covariance matrix

An approximative expression, valid for large N, for the covariance matrix of the estimate of θ_0 described in Proposition 4 is given in this section. The computation of the involving elements is the main contribution of the paper, and a step-by-step algorithm is given in the end of the section.

Since $\hat{\theta}$ minimizes $W(\theta)$, it holds that $\dot{W}(\hat{\theta}) = \mathbf{0}$, where $\dot{W}(\theta)$ denotes the first order derivative of $W(\theta)$. By the mean value theorem, $\dot{W}(\hat{\theta}) = \mathbf{0}$ can be written as

$$\mathbf{0} = \dot{W}(\theta_0) + \ddot{W}(\theta_\xi)(\hat{\theta} - \theta_0),$$

where $\ddot{W}(\boldsymbol{\theta})$ denotes the second order derivative of $W(\boldsymbol{\theta})$, and where $\boldsymbol{\theta}_\xi$ is between $\boldsymbol{\theta}_0$ and $\hat{\boldsymbol{\theta}}$. Due to consistency of $\hat{\boldsymbol{\theta}}$,

$$\lim_{N\to\infty} \boldsymbol{\theta}_\xi = \boldsymbol{\theta}_0.$$

Let

$$\bar{\ddot{W}}(\boldsymbol{\theta}) = \lim_{N\to\infty} \ddot{W}(\boldsymbol{\theta})$$

and use the triangle inequality to get

$$\begin{aligned} ||\ddot{W}(\boldsymbol{\theta}_\xi) - \bar{\ddot{W}}(\boldsymbol{\theta}_0)||_2 \leqslant\ & ||\ddot{W}(\boldsymbol{\theta}_\xi) - \ddot{W}(\boldsymbol{\theta}_0)||_2 \\ & + ||\ddot{W}(\boldsymbol{\theta}_0) - \bar{\ddot{W}}(\boldsymbol{\theta}_0)||_2. \end{aligned}$$

Since the right-hand side tends to zero as $N \to \infty$,

$$\lim_{N\to\infty} \ddot{W}(\boldsymbol{\theta}_\xi) = \bar{\ddot{W}}(\boldsymbol{\theta}_0) = \mathbf{H}.$$

This gives the approximation

$$\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0 \approx -\mathbf{H}^{-1}\dot{W}(\boldsymbol{\theta}_0)$$

for large N, provided that $\mathbf{H}^{-1}$ exists. The following result can now be given.

Result 4. *For large N, the covariance matrix of $\hat{\boldsymbol{\theta}}$ is approximately given as*

$$\mathbf{K} = \mathrm{E}\{(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)^T\} \approx \mathbf{H}^{-1}\mathbf{Q}\mathbf{H}^{-1}, \tag{16}$$

provided that $\mathbf{H}^{-1}$ exists, where

$$\begin{aligned} \mathbf{Q} &= \mathrm{E}\{\dot{W}(\boldsymbol{\theta}_0)\dot{W}^T(\boldsymbol{\theta}_0)\}, \\ \mathbf{H} &= \lim_{N\to\infty} \ddot{W}(\boldsymbol{\theta}_0). \end{aligned}$$

Next, the matrices $\mathbf{Q}$ and $\mathbf{H}$ are to be found. Section IV-A is devoted to the computation of $\mathbf{Q}$, whereas Section IV-B describes the computation of $\mathbf{H}$.

A. *Computation of* $\mathbf{Q}$

To find $\mathbf{Q}$, it is first noted that

$$\dot{W}(\boldsymbol{\theta}) = 2\sum_{\tau=0}^{k} \big(r_{y_0u_0}(\tau, \boldsymbol{\theta}, \hat{\boldsymbol{\psi}}) - \hat{r}_{yu}(\tau)\big)\dot{r}_{y_0u_0}(\tau, \boldsymbol{\theta}, \hat{\boldsymbol{\psi}}),$$

where $\dot{r}_{y_0u_0}(\tau, \boldsymbol{\theta}, \hat{\boldsymbol{\psi}})$ denotes the derivative of $r_{y_0u_0}(\tau, \boldsymbol{\theta}, \hat{\boldsymbol{\psi}})$ with respect to $\boldsymbol{\theta}$. The elements of

$$\begin{aligned} &\dot{r}_{y_0u_0}(\tau, \boldsymbol{\theta}, \hat{\boldsymbol{\psi}}) = \\ &\quad = \left[\frac{\partial r_{y_0u_0}(\tau, \boldsymbol{\theta}, \hat{\boldsymbol{\psi}})}{\partial \theta_1} \quad \dots \quad \frac{\partial r_{y_0u_0}(\tau, \boldsymbol{\theta}, \hat{\boldsymbol{\psi}})}{\partial \theta_{2n}}\right]^T, \end{aligned} \tag{17}$$

where θ_i is the ith element of the vector $\boldsymbol{\theta}$, are found by differentiating (5). More exactly, the ith element of (17) is found as element $(2, 1)$ of the matrix

$$\frac{\partial \mathbf{R}_{\mathbf{z}_0}(\tau, \boldsymbol{\theta}, \hat{\boldsymbol{\psi}})}{\partial \theta_i} = \begin{cases} \mathbf{C}\dfrac{\partial \mathbf{P}(\boldsymbol{\theta}, \hat{\boldsymbol{\psi}})}{\partial \theta_i}\mathbf{C}^T, & \tau = 0, \\ \mathbf{C}\dfrac{\partial \mathbf{A}^\tau(\boldsymbol{\theta}, \hat{\boldsymbol{\psi}})}{\partial \theta_i}\mathbf{P}(\boldsymbol{\theta}, \hat{\boldsymbol{\psi}})\mathbf{C}^T & \\ \quad + \mathbf{C}\mathbf{A}^\tau(\boldsymbol{\theta}, \hat{\boldsymbol{\psi}})\dfrac{\partial \mathbf{P}(\boldsymbol{\theta}, \hat{\boldsymbol{\psi}})}{\partial \theta_i}\mathbf{C}^T & \tau > 0, \\ \quad + \lambda_e^2\mathbf{C}\dfrac{\partial \mathbf{A}^{\tau-1}(\boldsymbol{\theta}, \hat{\boldsymbol{\psi}})}{\partial \theta_i}\mathbf{B}(\hat{\boldsymbol{\psi}})\mathbf{D}^T, & \end{cases} \tag{18}$$

where $\partial \mathbf{P}(\boldsymbol{\theta}, \hat{\boldsymbol{\psi}})/\partial \theta_i$ is given as the unique and non-negative definite solution to the Lyapunov equation

$$\begin{aligned} \frac{\partial \mathbf{P}(\boldsymbol{\theta}, \hat{\boldsymbol{\psi}})}{\partial \theta_i} &= \mathbf{A}(\boldsymbol{\theta}, \hat{\boldsymbol{\psi}})\frac{\partial \mathbf{P}(\boldsymbol{\theta}, \hat{\boldsymbol{\psi}})}{\partial \theta_i}\mathbf{A}^T(\boldsymbol{\theta}, \hat{\boldsymbol{\psi}}) \\ &\quad + \frac{\partial \mathbf{A}(\boldsymbol{\theta}, \hat{\boldsymbol{\psi}})}{\partial \theta_i}\mathbf{P}(\boldsymbol{\theta}, \hat{\boldsymbol{\psi}})\mathbf{A}^T(\boldsymbol{\theta}, \hat{\boldsymbol{\psi}}) \\ &\quad + \mathbf{A}(\boldsymbol{\theta}, \hat{\boldsymbol{\psi}})\mathbf{P}(\boldsymbol{\theta}, \hat{\boldsymbol{\psi}})\frac{\partial \mathbf{A}^T(\boldsymbol{\theta}, \hat{\boldsymbol{\psi}})}{\partial \theta_i}. \end{aligned} \tag{19}$$

Here, $\mathbf{P}(\boldsymbol{\theta}, \hat{\boldsymbol{\psi}})$ is given from (6), and the partial derivatives of the functions of $\mathbf{A}(\boldsymbol{\theta}, \hat{\boldsymbol{\psi}})$ with respect to θ_i are straightforward to find from the chosen state space form.

Assume that

$$\hat{r}_{yu}(\tau) = r_{y_0u_0}(\tau, \boldsymbol{\theta}_0, \hat{\boldsymbol{\psi}}) + \varepsilon(\tau).$$

This means that

$$\mathbf{Q} = 4\sum_{\tau=0}^{k}\sum_{s=0}^{k} \mathrm{E}\{\varepsilon(\tau)\varepsilon(s)\}\dot{r}_{y_0u_0}(\tau, \boldsymbol{\theta}_0, \hat{\boldsymbol{\psi}})\dot{r}_{y_0u_0}^T(s, \boldsymbol{\theta}_0, \hat{\boldsymbol{\psi}}). \tag{20}$$

Compute the element

$$\begin{aligned} \mathrm{E}\{\varepsilon(\tau)\varepsilon(s)\} \approx\ & \mathrm{E}\{\hat{r}_{yu}(\tau)\hat{r}_{yu}(s)\} \\ & - r_{y_0u_0}(\tau, \boldsymbol{\theta}_0, \boldsymbol{\psi}_0)r_{y_0u_0}(s, \boldsymbol{\theta}_0, \boldsymbol{\psi}_0), \end{aligned} \tag{21}$$

where the approximation is motivated by (7), the fact that N is large, and by the assumption that $\lim_{N\to\infty} \hat{\boldsymbol{\psi}} = \boldsymbol{\psi}_0$. Here,

$$\begin{aligned} \mathrm{E}\{\hat{r}_{yu}(\tau)\hat{r}_{yu}(s)\} &= \frac{1}{(N-\tau)(N-s)} \\ &\quad \cdot \sum_{t_1=1}^{N-\tau}\sum_{t_2=1}^{N-s} \mathrm{E}\{y(t_1+\tau)u(t_1)y(t_2+s)u(t_2)\}. \end{aligned} \tag{22}$$

For the four jointly Gaussian variables ζ_1, ζ_2, ζ_3, and ζ_4 of zero mean, it holds that

$$\begin{aligned} \mathrm{E}\{\zeta_1\zeta_2\zeta_3\zeta_4\} = \mathrm{E}\{\zeta_1\zeta_2\}\mathrm{E}\{\zeta_3\zeta_4\} + \mathrm{E}\{\zeta_1\zeta_3\}\mathrm{E}\{\zeta_2\zeta_4\} \\ + \mathrm{E}\{\zeta_1\zeta_4\}\mathrm{E}\{\zeta_2\zeta_3\}. \end{aligned}$$

Hence,

$$\begin{aligned} &\mathrm{E}\{y(t_1+\tau)u(t_1)y(t_2+s)u(t_2)\} \\ &\quad = r_{y_0u_0}(\tau, \boldsymbol{\theta}_0, \boldsymbol{\psi}_0)r_{y_0u_0}(s, \boldsymbol{\theta}_0, \boldsymbol{\psi}_0) \\ &\qquad + \big(r_{y_0}(|t_1+\tau-t_2-s|, \boldsymbol{\theta}_0, \boldsymbol{\psi}_0) \\ &\qquad\quad + r_{\tilde{y}}(|t_1+\tau-t_2-s|, \boldsymbol{\kappa}_0)\big) \\ &\qquad \cdot \big(r_{u_0}(|t_1-t_2|, \boldsymbol{\psi}_0) + r_{\tilde{u}}(|t_1-t_2|, \boldsymbol{\gamma}_0)\big) + f_1f_2, \end{aligned} \tag{23}$$

where

$$f_1 = \begin{cases} r_{y_0u_0}(t_1+\tau-t_2, \boldsymbol{\theta}_0, \boldsymbol{\psi}_0), & t_1+\tau-t_2 \geqslant 0, \\ r_{u_0y_0}(t_2-t_1-\tau, \boldsymbol{\theta}_0, \boldsymbol{\psi}_0), & t_1+\tau-t_2 < 0, \end{cases} \tag{24}$$

$$f_2 = \begin{cases} r_{y_0u_0}(t_2+s-t_1, \boldsymbol{\theta}_0, \boldsymbol{\psi}_0), & t_2+s-t_1 \geqslant 0, \\ r_{u_0y_0}(t_1-t_2-s, \boldsymbol{\theta}_0, \boldsymbol{\psi}_0), & t_2+s-t_1 < 0. \end{cases} \tag{25}$$

B. Computation of $\mathbf{H}$

To find the Hessian $\mathbf{H}$, compute

$$\ddot{W}(\boldsymbol{\theta}) = 2\sum_{\tau=0}^{k} \dot{r}_{y_0u_0}(\tau,\boldsymbol{\theta},\hat{\boldsymbol{\psi}})\dot{r}_{y_0u_0}^T(\tau,\boldsymbol{\theta},\hat{\boldsymbol{\psi}}) + 2\sum_{\tau=0}^{k}\big(r_{y_0u_0}(\tau,\boldsymbol{\theta},\hat{\boldsymbol{\psi}}) - \hat{r}_{yu}(\tau)\big)\ddot{r}_{y_0u_0}(\tau,\boldsymbol{\theta},\hat{\boldsymbol{\psi}})$$

to get

$$\mathbf{H} \approx 2\sum_{\tau=0}^{k} \dot{r}_{y_0u_0}(\tau,\boldsymbol{\psi}_0,\boldsymbol{\theta}_0)\dot{r}_{y_0u_0}^T(\tau,\boldsymbol{\psi}_0,\boldsymbol{\theta}_0), \qquad (26)$$

where the approximation is motivated by (7), the fact that N is large, and by the assumption that $\lim_{N\to\infty}\hat{\boldsymbol{\psi}} = \boldsymbol{\psi}_0$.

The computation of the covariance matrix is summarized in Algorithm 2.

Algorithm 2. *The computation of the covariance matrix* $\mathbf{K}$.

1) *Compute* $\mathbf{R}_{\mathbf{z}_0}(\tau,\boldsymbol{\theta}_0,\boldsymbol{\psi}_0)$ *as described in Result 1.*
2) *Compute* $r_{\tilde{u}}(\tau,\boldsymbol{\gamma}_0)$ *and* $r_{\tilde{y}}(\tau,\boldsymbol{\kappa}_0)$ *as described in Results 2 and 3, respectively.*
3) *Compute* $\mathrm{E}\{\hat{r}_{yu}(\tau)\hat{r}_{yu}(s)\}$ *in (22) using (23)–(25) and the results from Steps 1 and 2.*
4) *Compute* $\mathrm{E}\{\varepsilon(\tau)\varepsilon(s)\}$ *in (21) using the results from Steps 1 and 3.*
5) *Compute* $\dot{r}_{y_0u_0}(\tau,\boldsymbol{\theta}_0,\hat{\boldsymbol{\psi}})$ *in (17) through (18) and (19).*
6) *Compute* $\mathbf{Q}$ *in (20) using the results from Steps 4 and 5.*
7) *Compute* $\mathbf{H}$ *in (26) using the results from Step 5.*
8) *Compute* $\mathbf{K}$ *in (16) using the results from Steps 6 and 7.*

V. Example

The system defined by

$$\mathbf{A}(\boldsymbol{\theta}_0,\boldsymbol{\psi}_0)=\begin{bmatrix}-a_1 & 1 & b_1 & 0\\ -a_2 & 0 & b_2 & 0\\ 0 & 0 & -c_1 & 1\\ 0 & 0 & -c_2 & 0\end{bmatrix}=\begin{bmatrix}1 & 1 & 1 & 0\\ -0.5 & 0 & -0.8 & 0\\ 0 & 0 & 1 & 1\\ 0 & 0 & -0.5 & 0\end{bmatrix},$$

$$\mathbf{B}(\boldsymbol{\psi}_0) = \begin{bmatrix}0\\0\\d_1\\d_2\end{bmatrix} = \begin{bmatrix}0\\0\\1\\-0.3\end{bmatrix}, \mathbf{C} = \begin{bmatrix}0 & 0 & 1 & 0\\ 1 & 0 & 0 & 0\end{bmatrix}, \mathbf{D} = \begin{bmatrix}0\\0\end{bmatrix},$$

and $\lambda_e^2 = 1$, where the disturbance dynamics are described by

$$\mathbf{A}_g(\boldsymbol{\gamma}_0) = -g_1 = 0.8, \quad \mathbf{B}_g = 1, \quad \mathbf{C}_g = 1, \quad \lambda_v^2 = 1,$$

and

$$\mathbf{A}_h(\boldsymbol{\kappa}_0) = -h_1 = 0.8, \quad \mathbf{B}_h = 1, \quad \mathbf{C}_h = 1, \quad \lambda_w^2 = 1$$

is considered in an example. The aim is to compare the theoretical variances from $\mathbf{K}$ in (16) with empirical variances from a Monte Carlo simulation with 100 realizations of $N = 10\,000$ data points. Here, $\hat{\boldsymbol{\psi}}$ is taken as $\boldsymbol{\psi}_0$ in the loss function $W(\boldsymbol{\theta})$ in (13) in the estimation of $\boldsymbol{\theta}_0$ in each realization in order to make a fair comparison between the empirical variances and the variances from $\mathbf{K}$.

The theoretical and empirical variances for the estimates $\hat{a}_1$, $\hat{a}_2$, $\hat{b}_1$, and $\hat{b}_2$ as functions of the maximum lag k considered in the loss function $W(\boldsymbol{\theta})$ are shown in Fig. 1. It is seen that the empirical variances are well described by the theoretical variances and that the validity of the theoretical expressions is confirmed for this example.

VI. Conclusions

The EIV identification problem with colored measurement noises was studied. The solution considered in this paper was to estimate covariance and cross-covariance functions from the noise corrupted data and to match the corresponding theoretical functions, parameterized by the unknown parameters, to the estimated functions. A step-by-step algorithm for the computation of the elements of an expression for the covariance matrix of the estimated system parameters, valid for a large amount of data points, was given. The theoretical variances were verified by empirical variances from a Monte Carlo simulation in an example.

References

[1] B. D. O. Anderson and M. Deistler, "Identification of dynamic errors-in-variables models," *Journal of Time Series Analysis*, vol. 5, pp. 1–13, 1984.

[2] V. Solo, "Identifiability of time series models with errors in variables," *Journal of Applied Probability*, vol. 23A, pp. 63–71, 1986.

[3] R. Guidorzi, R. Diversi, and U. Soverini, "Optimal errors-in-variables filtering," *Automatica*, vol. 39, pp. 281–289, 2003.

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

[5] M. Mossberg, "Analysis of a covariance matching method for discrete-time errors-in-variables identification," in *Proc. 14th IEEE Statistical Signal Processing Workshop*, Madison, WI, August 26–29 2007, pp. 759–763.

[6] T. Söderström, *Discrete-Time Stochastic Systems*, 2nd ed. London, U.K.: Springer-Verlag, 2002.

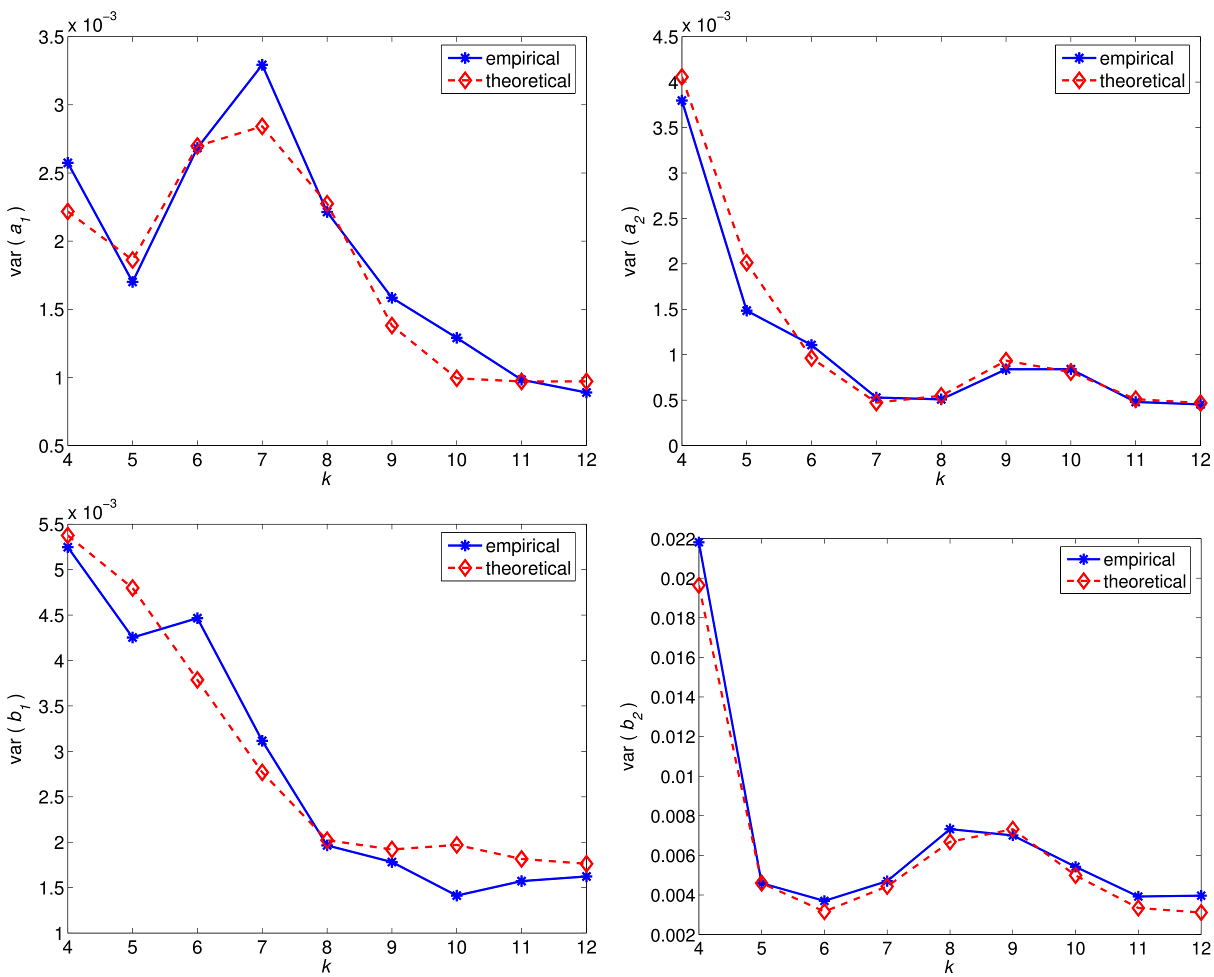

Fig. 1. The theoretical and empirical variances for the estimates $\hat{a}_1$ (upper left), $\hat{a}_2$ (upper right), $\hat{b}_1$ (lower left), and $\hat{b}_2$ (lower right) as functions of the maximum lag k.