An approach to identifying linearizable dynamic errors-in-variables systems

Levente Hunyadi and István Vajk

Abstract—We consider an important class of dynamic single-input single-output nonlinear systems where the system model is polynomial in observations but linear in parameters. The investigation is done in the errors-in-variables framework, i.e. both input and output are observed with noise. Assuming white Gaussian measurement noise that is characterized by a magnitude and a covariance structure, we propose a nonlinear extension to the generalized Koopmans–Levin method that can estimate parameters of dynamic nonlinear systems with polynomial nonlinearities given a priori knowledge on the noise covariance structure. In order to estimate noise structure, we apply a covariance matching objective function. Combining the extended Koopmans–Levin and the covariance matching approaches, an identification algorithm to estimate both model and noise parameters is proposed. The feasibility of the approach is demonstrated by Monte-Carlo simulations.

Keywords—dynamic system identification; linearizable nonlinear systems; polynomial eigenvalue problem; covariance matching

I. INTRODUCTION

Errors-in-variables systems where all variables are equally observed with noise are of particular significance in applications where the quantitative description of the internal laws constituting the system is of interest rather than predicting future behavior. Such applications include computer vision, image reconstruction, speech and audio processing, signal processing, modal and spectral analysis, system identification, econometrics and time series analysis. In these applications, the task is to construct a best possible system model based on noisy observations. A general system model takes the form

\[ f(\theta, z_{0,i}) = 0 \]

where the vector \( z_{0,i} \) denotes (noise-free) data for \( \forall i = 1, \ldots, N \), \( N \) being the number of observations, the vector \( \theta \) encapsulates the parameters of interest, and \( f \) represents some constraint between (for dynamic systems, past and present) data. A usual assumption that is satisfied in most applications is that the constraint \( f \) is linear in \( \theta \), i.e.

\[ \theta^T g(z_{0,i}) = 0 \]

where \( g \) is a linearization of \( f \). Given that \( z_{0,i} \) are not directly observable but contaminated with noise, the most common assumption being Gaussian white noise, hence the actual observations \( \tilde{z}_i \) satisfy \( \tilde{z}_i = z_{0,i} + \tilde{z}_i \), the objective is to derive estimates for \( \theta \) given the noisy observations \( \tilde{z}_i \).

For the special case of identifying linear errors-in-variables systems, where the constraint \( f \) is linear not only in \( \theta \) but also in data \( z_{0,i} \), a number of estimation schemes have been proposed. For dynamic single-input single-output (SISO) systems, where the system is described by the linear equation

\[ y_{0,i} = a_1 y_{0,i-1} + a_2 y_{0,i-2} + \ldots + a_m y_{0,i-m} = b_1 u_{0,i-1} + b_2 u_{0,i-2} + \ldots + b_n u_{0,k-m} \]

or more compactly,

\[ A(q^{-1}) y_{0,i} = B(q^{-1}) u_{0,i} \]

where the polynomial variable \( q^{-1} \) denotes the backward shift operator \( q^{-1} = q^{-i-1} \) (in which \( q \) is a generic placeholder for a parameter) but neither the noise-free true input \( u_{0,i} \) nor the true output \( y_{0,i} \) is observable but one is confined to their noise-contaminated variants \( u \) and \( y \), proposed methods include bias-compensating least squares [3], the Frisch scheme [2], instrumental variable [9], higher-order statistics [10], structured total least squares [7], frequency-domain and efficient maximum likelihood methods, see [8] for a comprehensive survey.

This paper deals with a nonlinear extension of the generalized Koopmans–Levin method to estimate model parameters of a dynamic system with given noise structure where the linearization \( g \) in (1) is a polynomial in terms of input and output data (Figure 1), and a subsequent covariance matching objective function to estimate noise covariance structure. The Koopmans–Levin method, proposed in [6], gives a non-iterative quick estimate of the model parameters of a linear system given a priori information on the noise structure. The original method was generalized in [11] to improve estimation accuracy at the cost of increased computational complexity, incorporating as special cases the original Koopmans–Levin method and the maximum likelihood method. In addition, a nonlinear extension to the
original Koopmans method was proposed in [12] for static systems. The Koopmans–Levin method and its generalization are briefly described in Section II. Section III combines and extends the results of [11] and [12] to nonlinear dynamic systems that comprise of polynomial nonlinearities yet are linear in model parameters. The outlined method assumes a preliminarily known noise structure. Section IV extends the estimation method so that no such assumptions are required. In order to demonstrate the feasibility of the method, some simulation results are presented in Section V before the paper concludes with Section VI.

II. THE GENERALIZED KOOPMANS–LEVIN METHOD

Consider the linear SISO errors-in-variables system \( G(q^{-1}) \) described by the autoregressive moving average (ARMA) difference equation

\[
A(q^{-1})y_{0,i} = B(q^{-1})u_{0,i}
\]

with

\[
A(q^{-1}) = a_0 + a_1 q^{-1} + \cdots + a_m q^{-m}
\]

\[
B(q^{-1}) = b_0 + b_1 q^{-1} + \cdots + b_m q^{-m}
\]

Given the aforementioned system description, we may introduce the model parameter vector \( \theta \) and the extended regressor vector \( z \)

\[
\begin{align*}
\theta^\top &= \begin{bmatrix} a_0 & a_1 & \ldots & a_m & -b_0 & -b_1 & \ldots & -b_m \end{bmatrix} \\
z_i^\top &= \begin{bmatrix} y_{i} & y_{i-1} & u_{i} & \ldots & y_{i-m} & u_{i-m} & \ldots & u_{i-1} \end{bmatrix}
\end{align*}
\]

such that \( \theta^\top z_{0,i} = 0 \) \( \forall i = m + 1 \ldots N \) with \( m \) being the order of the model and \( z_i = z_{i-1} \) is the noise contribution following a normal distribution such that \( y_i \sim N(0, \sigma_y^2) \) and \( u_i \sim N(0, \sigma_u^2) \). Notice that the system is linear in components \( y_{0,i} \) and \( u_{0,i} \) as well as in model parameters \( a_k \) and \( b_k \) with \( 0 \leq k \leq m \). Furthermore, introduce the observation sample and noise covariance matrices as

\[
D = \mathbb{E}(zz^\top) \approx \frac{1}{N} \sum_{i=1}^{N} z_i z_i^\top
\]

\[
\mu C = \mathbb{E}(z z^\top) \approx \frac{1}{N} \sum_{i=1}^{N} z_i z_i^\top
\]

with \( \mu \) denoting noise magnitude and \( C \) representing a normalized noise covariance matrix, or noise (covariance) structure, which is a diagonal matrix due to our assumptions on noise. We assume that the input signal \( u_0 \) is an excitation of sufficient order to make identification possible.

The essence of the Koopmans–Levin method is that the (full-rank) sample covariance matrix \( D \) comprising of (noisy) observations can be decomposed into a (rank-deficient) noise-free component \( D_0 \) and a noise component \( C \):

\[
\theta^\top D \theta = \theta^\top D_0 \theta + \theta^\top \mu C \theta = \theta^\top \mu C \theta
\]

minimizing the objective function

\[
J = \frac{1}{2} \theta^\top D \theta
\]

with \( ||\theta|| = 1 \), which can be effectively tackled by solving the eigenvector decomposition problem

\[
(D - \mu C) \theta = 0
\]

or

\[
\det(D_0) = \det(D - \mu C) = 0
\]

so that the model parameter vector is found by solving a generalized eigenvector problem on the matrix pair \( (D, C) \). The problem may alternately be formulated using matrix notation where

\[
\begin{align*}
Z &= \begin{bmatrix} y_1 & \cdots & y_m & u_1 & \cdots & u_m \\
y_2 & \cdots & y_{m+1} & u_2 & \cdots & u_{m+1} \\
\vdots & & \vdots & & \vdots & \vdots 
\end{bmatrix} \\
D &= Z^\top Z
\end{align*}
\]

\[
\mu C = \text{diag} \left( \sigma_y^2, \ldots, \sigma_y^2, \sigma_u^2, \ldots, \sigma_u^2 \right)
\]

with \( Z \) being an \( (N - m + 1) \times 2m \) matrix and \( C = \mu C_\rho \otimes I_m \) denoting the noise structure such that the noise covariance matrix is known up to a multiplication by a scalar \( \mu \) representing the noise magnitude, i.e.

\[
\mu C_\rho = \mu \begin{bmatrix} \sin^2 \rho & 0 \\
0 & \cos^2 \rho \end{bmatrix} = \begin{bmatrix} \sigma_y^2 & 0 \\
0 & \sigma_u^2 \end{bmatrix}
\]

in which we assume that the noise structure matrix \( C_\rho \) (uniquely determined by the noise direction \( \rho \) that reflects the relative distribution of input vs. output noise) is preliminarily known.

The variance of the estimates thus obtained, however, is rather large. One way [11] to improve the robustness of the parameter estimation approach outlined above is by instead of (3) minimizing the objective function

\[
J = \frac{1}{2(q-m)} \text{trace} \left( G_q^\top Z_q (G_q^\top C_q G_q)^{-1} Z_q G_q \right)
\]

where \( Z_q \) is an \( (N - q + 1) \times 2q \) matrix obtained by augmenting \( Z \) with \( q-m \) columns of additional observations for both \( y \) and \( u \); \( G_q \) is a \( 2q \times (q-m) \) matrix of model parameters such that \( Z_{0,q} G_q = 0 \); and \( C_q = (\mu C_\rho) \otimes I_q \) is a diagonal covariance structure matrix of size \( 2q \), and \( m+1 \leq q \leq N \); i.e.

\[
G_0 = \begin{bmatrix}
\circ_0 \\
\circ_1 \\
\vdots \\
\circ_m
\end{bmatrix}_{q,q-m}
\]

\[
G_q = \begin{bmatrix}
G_y & 0 \\
-G_u & 0
\end{bmatrix}_{2q,q-m}
\]
In short, both the model parameter vector $\theta$ and the original observation matrix $Z$ have been extended from size $m$ to $q$.

The above problem can be reformulated as

$$ J = \frac{1}{2(q-m)} \text{trace} \left( (G_q^\top C_q G_q)^{-1} G_q^\top D_q G_q \right) $$

which can be gradually approximated with the iteration scheme

$$ \frac{\text{trace} \left( (G_q^\top (\theta_k) C_q G_q(\theta_k))^{-1} G_q^\top (\theta) D_q G_q(\theta) \right)}{\text{trace} \left( (G_q^\top (\theta_k) C_q G_q(\theta_k))^{-1} G_q^\top C_q G_q(\theta_k) \right)} $$

where $\theta_{k+1} = \arg \min_{\theta} \theta^\top T^\top \left( (G_q^\top (\theta_k) C_q G_q(\theta_k))^{-1} \otimes D_q \right) T \theta$

where $T$ (a sparse matrix of zeros and ones) is chosen such that $\text{vec}(G_q) = T \hat{\theta}$. In each iteration, minimization w.r.t. $\theta$ is attained by solving a generalized eigenvector decomposition problem on the pair $(Q, R)$ with

$$ Q = T^\top \left( (G_q^\top (\theta_k) C_q G_q(\theta_k))^{-1} \otimes D_q \right) T $$
$$ R = T^\top \left( (G_q^\top (\theta_k) C_q G_q(\theta_k))^{-1} \otimes C_q \right) T $$

where $\theta$ is the eigenvector that belongs to the smallest eigenvalue $\mu$.

### III. A NONLINEAR EXTENSION

In order to further generalize the Koopmans–Levin method to nonlinear systems, linear components $y_i$ and $u_i$ give way to the nonlinearities that occur in the model. For instance, for the components $y_1, u_1, u_2^2$ and $y_1 u_i$, the $(N - q + 1) \times nq$ matrix $Z_q$ takes the form

$$ Z_q = \begin{bmatrix} y_1 & \cdots & y_q & u_1 & \cdots & u_q \\ y_2 & \cdots & y_{q+1} & u_2 & \cdots & u_{q+1} \\ \vdots & & \vdots & & \vdots \\ u_1^2 & \cdots & u_2^2 & y_1 u_1 & \cdots & y_q u_q \\ u_2^2 & \cdots & u_{q+1}^2 & y_2 u_2 & \cdots & y_{q+1} u_{q+1} \\ \vdots & & \vdots & & \vdots \end{bmatrix} $$

and the general form for the matrix $G_q$ becomes

$$ G_q = \begin{bmatrix} G_1 \\ G_2 \\ \vdots \\ G_n \end{bmatrix} $$

$n$ being the number of nonlinear components (in our case, $n = 4$) and $G_k$ encapsulating the parameters for the respective nonlinearity, $0 < k \leq n$. Notice that the matrix product $Z_q G_q$ entails that the system is still linear in parameters. However, the covariance matrix structure $C_q$ is no longer a single diagonal matrix but is replaced by a matrix polynomial $C_q(\mu) = \mu C_q^{(1)} + \mu^2 C_q^{(2)} + \ldots + \mu^n C_q^{(n)}$

in which $C_q^{(k)}$ is the $k$th coefficient of the matrix polynomial $C_q(\mu)$. One can use the following identities in deriving $C_q(\mu)$:

$$ E(x_i^p) = E(x_{0, i} + n_i)^p $$
$$ E(n_i^2) = (2p - 1)(2p - 3) \ldots 1 \sigma^2p $$
$$ E(n_i^{2p-1}) = 0 $$

$E(x_i) \approx \bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i$

$E(x_i, x_{i-\tau}) = E(x_i) E(x_{i-\tau})$

in which $n_i$ is a zero-mean $\sigma^2$-variance normally distributed random value at time instant $i$. [12] features a detailed description as well as an introductory example on how to derive these terms. Figure 2 shows a sample covariance polynomial for $q = 2$. Observe that the matrix entries usually depend not only on noise parameters $\sigma_\epsilon$ and $\sigma_y$ but also on (means of) the observations themselves.

The use of covariance matrix polynomials instead of regular covariance matrices necessitates some modifications to the objective function (5) as well as the iteration scheme (7). By including the noise magnitude within the trace operator in (5) yielding

$$ \text{trace} \left( (G_q^\top \mu C_q G_q)^{-1} G_q^\top D_q G_q \right) $$

it is apparent that the matrix product approaches the unit matrix should the best possible model parameters and noise covariance matrix be used. In this spirit, (5) can be reformulated as

$$ J = \left( \text{trace} \left( (G_q^\top \mu C_q G_q)^{-1} G_q^\top D_q G_q \right) - (q-m) \right)^2 $$

with $m$ being the the order of the dynamic model to estimate where the minimum of $J$ is attained when both model and noise magnitude estimates best match observations. Substituting the noise covariance polynomial $C_q(\mu)$ into (8), we get a parameter estimation scheme for nonlinear systems. Thus, we propose the following differentiable objective function:

$$ J = \left( \text{trace} \left( \gamma^{-1} \delta \right) - (q-m) \right)^2 $$

where

$$ \delta = G_q^\top(\theta) D_q G_q(\theta) $$
$$ \gamma = G_q^\top(\theta) C_q(\mu) G_q(\theta) $$

As the function (9) is differentiable, a direct search utilizing the Levenberg-Marquardt method yields model parameter...
and noise magnitude estimates. However, the Levenberg-Marquardt method finds a single local minimum in a single run, making the scheme sensitive to initial values.

Iterative schemes based on matrix decomposition are more robust against getting stuck in local minima as they deliver one way to solve a polynomial eigenvector decomposition of the polynomial eigenvalue problem reduces to a general-poly.

\[ C(\mu) = \begin{bmatrix} \sigma_u^2 & 0 & 0 & 0 & \bar{y}u \sigma_u^2 & \bar{y}u \sigma_u^2 & 0 \\ 0 & \sigma_u^2 & 0 & 0 & \bar{y}u \sigma_u^2 & \bar{y}u \sigma_u^2 & 0 \\ 0 & 0 & \sigma_u^2 & 3\bar{u}u^2 & \bar{y}u \sigma_u^2 & 0 & 0 \\ \bar{y}u \sigma_u^2 & \bar{y}u \sigma_u^2 & 3\bar{u}u^2 & 6\bar{u}u^2 & 2\bar{y}u \sigma_u^2 & 3\bar{y}u \sigma_u^2 & 0 \\ \bar{u}u \sigma_u^2 & 0 & \bar{y}u \sigma_u^2 & 3\bar{y}u \sigma_u^2 & \bar{y}^2 \sigma_u^2 + \sigma_p^2 u^2 & 0 & 0 \\ 0 & \bar{y}u \sigma_u^2 & 0 & \bar{y}u \sigma_u^2 & 3\bar{y}u \sigma_u^2 & \bar{y}^2 \sigma_u^2 + \sigma_p^2 u^2 & 0 \end{bmatrix} \begin{bmatrix} \ldots & 0 & 0 & 0 & 0 \end{bmatrix} - \mu^2 \begin{bmatrix} 3\sigma_u^2 & \sigma_u^2 & 0 & 0 \\ \sigma_u^2 & 3\sigma_u^2 & 0 & 0 \\ 0 & 0 & \sigma_u^2 & 0 \\ \sigma_u^2 & 0 & \sigma_u^2 & 0 \end{bmatrix} \begin{bmatrix} \ldots & 0 & 0 & 0 & 0 \end{bmatrix}. \]

As the linearized problem has eigenvectors of dimension \( np \) rather than \( m \), the “best” polynomial eigenvector that belongs to the eigenvalue \( \mu \) becomes the portion \( v_k \) of the linearized eigenvector \( \psi(\mu)x = 0 \) that gives the smallest normalized residual, i.e.

\[ v_k = \arg \min_{v_k} \frac{\sum_k |\langle \Psi(\mu)v_k \rangle|}{\sum_k |v_k|}. \]

With the iterative scheme (10) at hand, a few initial iterations can be used to seed the Levenberg-Marquardt search with appropriate initial values reducing the likelihood of (9) getting stuck in a local minimum.

IV. MODEL AND NOISE ESTIMATION

Contrary to the hidden assumption in the previous section, in a real-world scenario, the true noise structure \( C_\rho \) (or equivalently, a noise direction \( \rho \) that determines the ratio of input and output noise variances for a unit magnitude noise) is seldom at our disposal. As the final step of the parameter estimation method, we propose means to estimate \( C_\rho \) for white noise.

One way to parametrize noise variances, as in (4), is by writing \( \sigma_u^2 = \mu \cos^2 \rho \) and \( \sigma_p^2 = \mu \sin^2 \rho \) such that \( C_q = C_q(\mu, \rho) \). Let \( \hat{\theta} \) denote (unit-normalized) estimates obtained with a particular assumption of \( \rho \) using (10). Introduce the notations

\[ \hat{\delta} = G_q^\dagger(\hat{\theta})D_qG_q(\hat{\theta}) \]
\[ \hat{\gamma} = G_q^\dagger(\hat{\theta})C_q(\hat{\theta}) \]

Varying \( \rho \) in the range from 0 to \( \frac{\pi}{2} \), one can discover the “true” value by minimizing the loss function

\[ J(\rho) = \| \hat{\delta} - \hat{\gamma} \|_F \]

where \( \| \cdot \|_F \) denotes the Frobenius norm (a technique called covariance matching in [13]) or the “inverted” loss function

\[ J(\rho) = \text{trace}(\hat{\delta}^{-1}\hat{\gamma}) \]

or the so-called Itakura–Saito matrix divergence

\[ J(\rho) = \text{trace}(\gamma \hat{\delta}^{-1} - \log(\det(\gamma \hat{\delta}^{-1})) - n \]

where \( n \) is the dimension of the square matrices involved. The minimum value for \( J \) in the above equations yields the estimated value for \( \rho \).

Fig. 2. Example covariance matrix for the polynomial components \( y_i, u_i, u_i^2 \) and \( y_iu_i \) with \( q = 2 \). Entries not indicated take a value of zero.
As an alternative to the two-stage estimation procedure outlined above, a single-stage strategy might theoretically also be employed. Let
\[
\delta = G_q^T(\theta)D_qG_q(\theta) \\
\gamma = G_q^T(\theta)C_q(\mu, \rho)G_q(\theta)
\]
and introduce the objective functions
\[
J_1 = (\text{trace} (\gamma^{-1}\delta) - (q - m))^2 \\
J_2 = (\text{trace} (\delta^{-1}\gamma) - (q - m))^2 \\
J_3 = \|\delta - \gamma\|^2_F = \text{trace}\left\{ (\delta - \gamma)^T (\delta - \gamma) \right\}
\]
Notice that all are functions of \(\theta, \mu\) as well as \(\rho\) simultaneously. Due to their nonlinearity, they are likely to exhibit convergence to local minima if started with the wrong initial values.

V. SIMULATION RESULTS

As an introductory example, consider the nonlinear Mackey–Glass chaotic process that models white blood cell production. The process is described by the following state equation
\[
x_{0,i+1} = -ax_{0,i} + \frac{b, x_{0,i-\tau}}{1 + x_{0,i-\tau}}^T \tag{15}
\]
where there exists no external excitation,
\[
x_{0,i} = a \quad \text{if} \quad 0 \leq i \leq \tau \\
a = 0.9 \quad b = 0.2 \quad \tau = 17 \quad p = 10
\]
The relationship (15) given in explicit form can be reformulated in implicit form
\[
x_{0,i+1} = ax_{0,i}^p - bx_{0,i-\tau} + ax_{0,i-\tau} - bx_{0,i-\tau} = 0
\]
A sequence of \(N = 1000\) data are observed with a noise of \(\sigma^2 = 0.05\). Notice the equality constraint on certain parameter coefficients expressible with the structural constraint matrix
\[
S = \begin{bmatrix}
1 & 0 & 1 & 0 & 0 \\
0 & 1 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1
\end{bmatrix}
\]
with the original model parameter vector \(\theta\) (without repeated coefficients) relating to the extended vector \(\tilde{\theta}\) (with repeated coefficients) via \(S\tilde{\theta} = \theta\). As a result of a simulation with \(M = 100\) runs, \(\tilde{\theta} = 0.9147 \pm 0.0124\) and \(\tilde{b} = -0.1653 \pm 0.0334\), that reproduce the original process fairly accurately.

Let us now draw our attention to an artificial yet relatively complex process described by the nonlinear relationship
\[
y_{0,i} = p_1 y_{0,i-1} + p_2 y_{0,i-2} + p_3 u_{0,i-1} + \\
+ p_4 u_{0,i-1} + p_5 y_{0,i-1} u_{0,i-2} \\
+ p_6 u_{0,i-1} y_{0,i-1}
\]

comprising of both linear and polynomial terms as well as cross-correlating terms. The true parameter values are set to
\[
p_1 = 1.5 \quad p_2 = -0.7 \quad p_3 = 1 \quad p_4 = -0.3 \quad p_5 = -0.05 \quad p_6 = 0.1
\]
and a configuration of \(N = 500\) samples, \(q = 6, \sigma_u = 0.01\) and \(\sigma_y = 0.01\) is set up.\(^1\) Next, a Monte-Carlo simulation of \(M = 100\) runs of the outlined nonlinear extension to the generalized Koopmans–Levin method has been carried out with a known noise structure. Table I analyzes the consistency of the estimation scheme.

Figure 3 shows how to discover noise covariance structure (i.e. the noise “direction” \(\rho\)) by minimizing the distance or divergence measures (12) and (13) over an interval to arrive at estimates for all parameters. In accordance, a different series of Monte-Carlo experiments have been conducted, in which for each independent simulation run a noise structure discovery step shown in Figure 3 has been carried out, with model parameters and noise magnitude estimated for each potential value of \(\rho\), and the distance measure evaluated. The comprehensive results, in which all model and noise parameters are estimated, are shown in Table II.

For the sake of comparison, an instrumental variable scheme, based on the bias-compensating least-squares technique for nonlinear polynomial systems (NBCLS) has been included in the table. The scheme minimizes the objective function [5]
\[
J = \| d_{IV} - c_{IV} - (D_{IV} - C_{IV})\tilde{\theta} \| \tag{16}
\]
where \(\tilde{\theta} = (D_{IV} - C_{IV})^{-1}(d_{IV} - c_{IV})\) with \(D_{IV}\) and \(d_{IV}\) being the (rectangular) covariance matrices of the regressor vector and the output vector, respectively, w.r.t. so-called instruments \(C_{IV}(\mu, \rho)\) and \(c_{IV}(\mu, \rho)\)

\(^1\)Simulation examples are available as part of [4].
TABLE I

Consistency analysis of parameter estimates with \( q = 6 \) and known noise structure.

<table>
<thead>
<tr>
<th>nonlinearity</th>
<th>NBCLS (16)</th>
<th>Frobenius norm (12)</th>
<th>“inverted” (13)</th>
<th>Itakura–Saito (14)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( y_{0,1-1} )</td>
<td>-1.4996 ± 0.0045</td>
<td>-1.5002 ± 0.0039</td>
<td>-1.4997 ± 0.0044</td>
<td>-1.4998 ± 0.0036</td>
</tr>
<tr>
<td>( y_{0,1-2} )</td>
<td>0.6998 ± 0.0038</td>
<td>0.7002 ± 0.0032</td>
<td>0.6997 ± 0.0035</td>
<td>0.7000 ± 0.0028</td>
</tr>
<tr>
<td>( u_{0,1-1} )</td>
<td>-1.0015 ± 0.0096</td>
<td>-0.9998 ± 0.0182</td>
<td>-1.0021 ± 0.0221</td>
<td>0.9960 ± 0.0198</td>
</tr>
<tr>
<td>( u_{0,1-2} )</td>
<td>0.3258 ± 0.3173</td>
<td>0.3101 ± 0.0953</td>
<td>0.2928 ± 0.1053</td>
<td>0.3004 ± 0.0898</td>
</tr>
<tr>
<td>( y_{0,1-2} + y_{0,1-2} )</td>
<td>0.0050 ± 0.0040</td>
<td>0.0492 ± 0.0114</td>
<td>0.0496 ± 0.0104</td>
<td>0.0505 ± 0.0102</td>
</tr>
<tr>
<td>( u_{0,1-2} + u_{0,1-2} )</td>
<td>-0.1038 ± 0.0477</td>
<td>-0.1048 ± 0.0394</td>
<td>-0.1017 ± 0.0307</td>
<td>0.1010 ± 0.0338</td>
</tr>
<tr>
<td>( \sigma^2 ) ( 10^{-3} )</td>
<td>0.0064 ± 0.0153</td>
<td>0.1017 ± 0.0134</td>
<td>0.0999 ± 0.0099</td>
<td>0.1011 ± 0.0102</td>
</tr>
<tr>
<td>( \sigma^2 ) ( 10^{-3} )</td>
<td>0.1021 ± 0.0406</td>
<td>0.0915 ± 0.0138</td>
<td>0.0944 ± 0.0126</td>
<td>0.0911 ± 0.0102</td>
</tr>
</tbody>
</table>

TABLE II

Simulation results with unknown noise structure and various methods to match covariance matrices.

VI. Conclusion

After a brief description of the Koopmans–Levin method and its generalization for linear systems, we introduced a nonlinear extension by modifying the objective function and the iteration scheme of the linear generalization. As a result, the original dependency on the model parameters \( \theta \) has been extended with dependency on the noise magnitude \( \mu \). Consequently, the iterative problem is tackled efficiently by solving a polynomial rather than a generalized eigenvalue problem. Next, an optimization scheme of minimizing an error term over an angle range has been shown to give a noise structure estimate thus yielding estimates for all model and noise parameters. Finally, objective functions that encompass all model and noise parameters have been proposed, which can be fed to direct search methods to produce parameter estimates.

ACKNOWLEDGMENTS

This work has been supported by the fund of the Hungarian Academy of Sciences for control research and the Hungarian National Research Fund (grant number T68370).

REFERENCES