An Approach to Nonlinear Stochastic System Identification: Simulation Experiments

*Akira Ohsumi (Prof. Emeritus of Kyoto Institute of Technology**)
Kentaro Kameyama (Fukui National College of Technology ***)

Abstract—In this paper, the usefulness of the nonlinear identification method proposed by one of the authors is demonstrated by simulation experiments.

Key Words: System identification, Nonlinear system

1 Introduction and Problem Statement

Recently, the modeling of nonlinear black-box system from a couple of input and output data is recognized as one of the most important problems in modern engineering, even in non-engineering field.

One of the authors1) has proposed an identification method for a class of MIMO (multi-input multi-output) stochastic systems which involve unknown nonlinearities in system state and, possibly, input, incorporated with the standard subspace-based identification method developed for linear systems.

Before reviewing our method, let us state our identification problem. Assume that we are given a couple of input and output data, \( \{u(k), y(k)\} \) \( u(k) \in \mathbb{R}^n \), \( y(k) \in \mathbb{R}^m \) which are acquired from the relevant black-box system. To fix the idea, the black-box system is supposed to be an MIMO stochastic system involving some nonlinearities in the system state and possibly input; and assume that the input-output relationship is described by the following state-space model:

\[
\begin{align*}
\mathbf{x}(k+1) &= A\mathbf{x}(k) + B\mathbf{u}(k) + f[\mathbf{x}(k), \mathbf{u}(k)] + w(k) \quad (1) \\
y(k) &= C\mathbf{x}(k) + D\mathbf{u}(k) + v(k), 
\end{align*}
\]

where \( \mathbf{x}(k) \in \mathbb{R}^n \) is the system state; \( w(k) \in \mathbb{R}^n \) and \( v(k) \in \mathbb{R}^m \) are mutually independent white Gaussian noise sequences with zero-means and covariance matrices \( Q \) and \( R \), i.e., \( \mathcal{E}\{w(k)w^T(j)\} = Q\delta_{kj} \), \( \mathcal{E}\{v(k)v^T(j)\} = R\delta_{kj} \) (\( \delta_{kj} \): Kronecker delta; the superscript \( T \) denotes the transpose); and \( f(\cdot, \cdot) \) is the \( n \)-vector-valued nonlinear function. The system order \( n \) is unknown, while rank \( C = m \) is assumed.

Then, the identification problem is to identify the system quadruplet \((A, B, C, D)\) and the covariances \((Q, R)\), the system order \( n \) and the nonlinear term \( f(\cdot, \cdot) \) (within a similarity transformation) from the acquired input and output data \( \{u(k), y(k)\} \).

The purpose of this paper is to verify the usefulness of the identification method proposed in Ref. 1) by simulation experiments.

2 Review of the Method

In this section, we review briefly our nonlinear identification method.

The mathematical model of the relevant nonlinear black-box system is established, incorporating the subspace identification method, by the following three steps.

(i) Extraction of Nonlinear Effect from Observation Data: Having premised that the input-output relationship is expressed by (1) and (2), it will be natural to consider that the output \( y(k) \) consists of “linear” and “nonlinear” portions. So, in order to extract the nonlinear portion from the observation data \( \{y(k)\} \), regarding temporarily the nonlinear term \( f[x(k), u(k)] \) in (1) as a random disturbance \( f(k) \) with zero-mean and independent of \( w(k) \), we consider the model:

\[
\begin{align*}
x(k+1) &= Ax(k) + Bu(k) + f(k) + w(k) \\
y(k) &= Cx(k) + Du(k) + v(k). 
\end{align*}
\]

Here, in order to identify the system quadruplet \((A, B, C, D)\) and noise covariances \((Q, R)\), we lump random terms together as

\[
\zeta(k) = f(k) + w(k)
\]

whose covariance is

\[
Q_\zeta = \mathcal{E}\{\zeta(k)\zeta^T(k)\} = \mathcal{E}\{f(k)f^T(k)\} + Q.
\]

The identification of \((A, B, C, D)\) and \((Q_\zeta, R)\) may be done based on (3) with (4) and (5) using the standard subspace-based identification method to obtain their estimates \((\hat{A}, \hat{B}, \hat{C}, \hat{D})\), \((\hat{Q}_\zeta, \hat{R})\) as well as the system order \( \hat{n} \). Although these estimates are rough ones, we use the estimate of \( Q \) in such a way that \( \hat{Q} = \rho \hat{Q}_\zeta \) where \( \rho \) is a design parameter such that \( 0 < \rho < 1 \).
Then, in order to make the nonlinear portion hidden in observation data conspicuous, we diminish the “linear” portion in \( \{ y(k) \} \) as thoroughly as possible. To do this, consider the linear system,

\[
\begin{align*}
x_L(k+1) &= \hat{A}x_L(k) + \hat{B}u(k) \\
y_L(k) &= \hat{C}x_L(k) + \hat{D}u(k)
\end{align*}
\]

with a proper initial value \( x_L(0) \), where \( x_L(k) \in R^n \), \( y_L(k) \in R_m \), \( u(k) \) is the same input as that in (1) and (2), and \( (\hat{A}, \hat{B}, \hat{C}, \hat{D}) \) are rough and ready estimates as mentioned above. Let us define

\[ y_F(k) = y(k) - y_L(k). \tag{7} \]

This is considered to be due to the nonlinearity involved in the original observation data and the random disturbances \( w(k) \) and \( v(k) \).

(ii) Estimation of Nonlinearity: Consider the following linear system whose output is \( y_F(k) \) defined by (7):

\[
\begin{align*}
x_F(k+1) &= \hat{A}x_F(k) + \hat{\bar{f}}(k) + \tilde{w}(k) \\
y_F(k) &= \hat{\bar{C}}x_F(k) + \tilde{v}(k)
\end{align*}
\]

in which \( x_F(k) \in R^\hat{n} \); \( \tilde{w}(k) \in R^\hat{n} \) and \( \tilde{v}(k) \in R^m \) are newly introduced (computer-generated) random sequences having zero-means and covariances \( Q \) and \( \tilde{R} \), respectively, which are identified in Step (i); and \( \hat{\bar{f}}(k) \) denotes the similarity-transformed \( f(k) \).

The input \( \{ \hat{f}(k) \} \) should be determined such that the output of (8) is exactly the data \( \{ y_F(k) \} \); so that the inverse system of (8) will be constructed to produce \( \{ \hat{f}(k) \} \). Assume that \( \{ \hat{f}(k) \} \) obeys the dynamics,

\[ \hat{f}(k+1) = \hat{f}(k) + \eta(k), \tag{9} \]

where \( \eta(k) \) is the (computer-generated) white Gaussian sequence with zero-mean and known covariance \( Q_\eta(\geq 0) \).

Then, augmenting (9) with \( x_F(k) \)-process and constructing the Kalman filter, we get the estimate process for \( \hat{f}(k) \) as follows:

\[
\begin{align*}
\hat{f}(k+1|k) &= \hat{f}(k|k-1) + \hat{P}_{21}\hat{C}^T(\hat{C}\hat{P}_{11}\hat{C}^T + \hat{R})^{-1}\cdot[y_F(k) - \hat{\bar{C}}x_F(k|k-1)] \\
\hat{\bar{x}}_F(k+1|k) &= \hat{\bar{x}}_F(k|k-1) + \hat{\bar{f}}(k|k-1) \\
+ (\hat{A}\hat{P}_{11} + \hat{P}_{21})\hat{C}^T(\hat{C}\hat{P}_{11}\hat{C}^T + \hat{R})^{-1}\cdot[y_F(k) - \hat{\bar{C}}x_F(k|k-1)],
\end{align*}
\]

(10)

where \( \hat{f}(k|k-1) \) and \( \hat{\bar{x}}_F(k|k-1) \) are Kalman estimates based on the data \( \{ y_F(k) \} \) of \( \hat{f}(k) \) and \( x_F(k) \), respectively; and \( \hat{P}_{ij} (i, j = 1, 2; \hat{P}_{11} \in R^{\hat{n} \times \hat{n}}; \hat{P}_{21} = \hat{P}_{12}^T) \) is the \((i,j)\)-component of the matrix \( \hat{P} \in R^{2\hat{n} \times 2\hat{n}} \) which satisfies algebraic Riccati equation:

\[
\hat{P} = A_0[\hat{P} - \hat{PC}_0^T(\hat{C}_0\hat{PC}_0^T + \hat{R})^{-1}\hat{C}_0\hat{P}]A_0^T + Q_0, \tag{11}
\]

where

\[
A_0 = \begin{bmatrix} \hat{A} & I_n \\ 0 & I_{\hat{n}} \end{bmatrix} \in R^{2\hat{n} \times 2\hat{n}}, \ C_0 = [\hat{C} & 0] \in R^{m \times 2\hat{n}}
\]

\[
Q_0 = \text{block diag} \{ \hat{Q}, \hat{Q}_\eta \}.
\]

The obtained \( \hat{f}(k|k-1) \) is used as the estimate of \( \hat{f}(k) \), viz.,

\[ \hat{f}(k) = \hat{f}(k|k-1). \tag{12} \]

Note that (10) is the inverse system of (8) in the sense that the input and output, \( \hat{f}(k) \) and \( y_F(k) \), in (8) are converted inversely in (10) to \( \hat{f}(k|k-1) \) (= \( \hat{f}(k) \)) as output and input \( y_F(k) \).

At this stage, it should be referred that there is another possibility to determine the nonlinearity \( \hat{f}(k) \) as a control problem. Instead of (8), consider the linear system,

\[
\begin{align*}
x_C(k+1) &= \hat{A}\hat{x}_C(k) + \hat{\bar{f}}(k) + \tilde{w}(k) \\
y_C(k) &= \hat{\bar{C}}x_C(k) + \tilde{v}(k)
\end{align*}
\]

Then, \( \{ \hat{\bar{f}}(k) \} \) is generated in such a way that the output \( y_C(k) \) tracks \( y_F(k) \) given by (7). This leads us to the stochastic tracking control problem to determine \( \{ \hat{\bar{f}}(k) \} \) as a control sequence to follow the target sequence \( \{ y_F(k) \} \). However, this approach is a little complicated compared with the inverse system (i.e., Kalman filter) because the (time-)backward computation is required to determine \( \hat{f}(k) \) as a feedback control sequence. So, we do not mention this approach any further\(^2\).

(iii) Modeling of Nonlinearity Using Basis Functions: Once we have obtained the estimate of nonlinearity \( \{ \hat{\bar{f}}(k) \} \) like (12), we model it using some basis functions. To do this consider the system:

\[
\begin{align*}
\Sigma_D: \quad &\begin{cases} x_D(k+1) = \hat{A}x_D(k) + \hat{B}u(k) + \hat{\bar{f}}(k) \\ y_D(k) = \hat{\bar{C}}x_D(k) + \hat{D}u(k) \end{cases} \tag{14}
\end{align*}
\]

and, as opposed to this, let

\[
\begin{align*}
\Sigma_M: \quad &\begin{cases} x_M(k+1) = \hat{A}x_M(k) + \hat{B}u(k) \\ y_M(k) = \hat{\bar{C}}x_M(k) + \hat{D}u(k) \end{cases} + \mu(k, \phi; \theta) \tag{15}
\end{align*}
\]

where \( x_D(k), x_M(k) \in R^n \); \( y_D(k), y_M(k) \in R^m \); \( u(k) \in R^d \) is the same input as that in (1) and (2); \( \{ \hat{f}(k) \} \) is given numerically as stated in Step (ii);
\( \mu(k, \varphi; \theta) \) in (15) is the \( n \)-vector-valued function to model \( f(k) \), \( \varphi = \varphi(x_M, u) \) is the user-defined basis functions; and \( \theta = \theta(k) \) is the parameter vector to be determined.

There are many possibilities to set the vector function \( \mu(k, \varphi; \theta) \). For instance, let

\[
\mu(k, \varphi; \theta) = \sum_{i=1}^{p} \theta_i(k) \varphi_i[x_M(k), u(k)],
\]

where \( \varphi_i(\cdot, \cdot) \) is \( n \)-vector-valued basis function and \( \{\theta_i(k)\} \) are scalar parameters to be determined.

Then, the vector-valued function \( \mu(k, \varphi; \theta) \) can be expressed as

\[
\begin{aligned}
\mu(k, \varphi; \theta) &= \Phi[x_M(k), u(k)] \theta(k), \\
\theta(k) &= [\theta_1(k) \cdots \theta_p(k)]^T \in R^p \\
\Phi(\cdot, \cdot) &= [\varphi_1(\cdot, \cdot) \cdots \varphi_p(\cdot, \cdot)] \in R^{n \times p};
\end{aligned}
\]

and hence, (15) is rewritten as

\[
\Sigma_M: \begin{cases} x_M(k + 1) = \hat{A}x_M(k) + \hat{B}u(k) \quad + \Phi[x_M(k), u(k)] \theta(k) \quad y_M(k) = \hat{C}x_M(k) + \hat{D}u(k). \end{cases}
\]

The parameter vector \( \theta(k) \) is determined sequentially in such a way that the functional defined by

\[
J_k(\theta) = \|y_D(k) - y_M(k)\|_W^2 \quad (k = 1, 2, \cdots)
\]

becomes minimal, where \( W \) is a symmetric and positive-definite matrix and \( \|y\|_W^2 \) is the Euclidean norm. Under the assumption that \( x_D(0) = x_M(0) \), the result is given by

\[
\theta(k) = \{\Phi^T[x_M(k), u(k)] \hat{C}^T W \hat{C} \Phi[x_M(k), u(k)]\}^{-1} \Phi^T[x_M(k), u(k)] \hat{C}^T W \hat{C} \lambda_k
\]

where \( \lambda_k \) will be given below (for the derivation of (20), see Ref. 1).

Therefore, (dropping the subscript \( M \) on \( x_M(k) \) and \( y_M(k) \) to write), we obtain the following models for the couple of (1) and (2):

\[
\begin{aligned}
x(k + 1) &= \hat{A}x(k) + \hat{B}u(k) \\
&\quad + \Phi[x(k), u(k)] \theta(k) + \hat{w}(k) \\
y(k) &= \hat{C}x(k) + \hat{D}u(k) + \hat{v}(k)
\end{aligned}
\]

in which \( \theta(k) \) is given by

\[
\theta(k) = \{\Phi^T[x(k), u(k)] \hat{C}^T W \hat{C} \Phi[x(k), u(k)]\}^{-1} \Phi^T[x(k), u(k)] \hat{C}^T W \hat{C} \lambda_k
\]

and \( \lambda_k \) is computed recursively by\(^2\)

\[
\begin{aligned}
\lambda_{k+1} &= \hat{A}\{\lambda_k - \Phi[x(k), u(k)] \theta(k)\} + \hat{f}(k + 1), \\
\lambda_0 &= \hat{f}(0) \quad (k = 1, 2, \cdots).
\end{aligned}
\]

Consequently, the identified mathematical model is given by

\[
\begin{aligned}
x(k + 1) &= \hat{A}x(k) + \hat{B}u(k) \\
&+ \Phi[x(k), u(k)] \theta(k) + \hat{w}(k) \\
y(k) &= \hat{C}x(k) + \hat{D}u(k) + \hat{v}(k)
\end{aligned}
\]

with

\[
\theta(k) = \{\Phi^T[x(k), u(k)] \hat{C}^T W \hat{C} \Phi[x(k), u(k)]\}^{-1} \Phi^T[x(k), u(k)] \hat{C}^T W \hat{C} \lambda_k
\]

\[
\begin{aligned}
\lambda_{k+1} &= \hat{A}\{\lambda_k - \Phi[x(k), u(k)] \theta(k)\} + \hat{f}(k + 1) \\
\text{I.C.:} &\quad x(0) = x_0, \quad y(0) = y_0, \quad \lambda_0 = \hat{f}(0)
\end{aligned}
\]

The initial condition \( x(0) \) should be estimated beforehand; otherwise it will be set arbitrarily.

\[\text{3} \quad \text{Simulation Experiment—A Quarter-Car}\]

(i) \textbf{Studied Model:}

Consider a quarter-car model of automotive suspension (see, e.g. Ref. 3). It is often used to demonstrate the application of the identification algorithm on a manageable example\(^4\). Its model is represented by a two-mass model (2DOF) as follows:

\[
\begin{aligned}
m_c \ddot{x}_c(t) &+ d_c \dot{x}_c(t) - \ddot{x}_w(t) \\
&+ k_0(x_c - x_u) = 0 \\
m_w \ddot{x}_w(t) &+ d_w \dot{x}_w(t) - \ddot{x}_c(t) + k_w x_w(t) \\
&- k_0(x_c - x_u) = k_w x_w(t),
\end{aligned}
\]

where \( x_c(t) \) and \( x_w(t) \) are vertical displacements of the chassis (body) of the car and the wheel, respectively, from their equilibrium states; \( x_u(t) \) the vertical road displacement input; \( m_c \) and \( m_w \) are masses of the body and the wheel of the car; \( d_c \) is the damper (shock absorber) coefficient between body and wheel; and \( k_w \) is spring constant of the tire (tire stiffness). The nonlinearity appears in the characteristics of the suspension spring. Its reaction force exhibits the nonlinearity\(^4\):

\[
k_0(\xi) = k_c \left( \xi + \alpha \xi^3 + \beta \tan \frac{\pi \xi}{2 \xi_{\max}} \right),
\]

where \( k_c \) is the spring constant; \( \alpha, \beta \) are positive constants; and \( \xi_{\max} \geq |x_c(t) - x_w(t)| \) is the suspension travel limit. For derivation of (25), see Appendix A.
Now, write the nonlinear reaction force \( k_0(\xi) \) as
\[
k_0(\xi) = k_c \xi + f_s(\xi),
\] (27)
where
\[
f_s(\xi) = k_c \left( \alpha \xi^3 + \beta \tan \frac{\pi \xi}{2 \xi_{\text{max}}} \right),
\] (28)
and write the road input as \( u(t) \), i.e.,
\[
u(t) = x_g(t).
\] (29)

Then, (25) is expressed as follows:
\[
\begin{pmatrix}
m_c \ddot{x}_c(t) + d_c \dot{x}_c(t) - \dot{x}_w(t) \\
+k_c \{ x_c(t) - x_w(t) \} + f_s(x_c - x_w) = 0 \\
m_w \ddot{x}_w(t) + d_c \{ \dot{x}_w(t) - \dot{x}_c(t) \} + k_w x_w(t) \\
+k_c \{ \dot{x}_w(t) - \dot{x}_c(t) \} \\
-f_s(x_c - x_w) = k_w u(t)
\end{pmatrix}
\] (30)

(ii) **Generation of Solution Process**:

The system model (30) can be expressed in the second-order vector form:
\[
M \ddot{z}(t) + C \dot{z}(t) + K z(t) + c_0 f_s [z(t)] = g u(t),
\] (31)
where \( z(t) = [x_c(t), x_w(t)]^T = [z_1(t), z_2(t)]^T \); and
\[
M = \begin{bmatrix} m_c & 0 \\
0 & m_w \end{bmatrix}, C = \begin{bmatrix} d_c & -d_c \\
-d_c & d_c \end{bmatrix}, \quad K = \begin{bmatrix} k_c & -k_c \\
-k_c & k_c + k_w \end{bmatrix}, \quad c_0 = \begin{bmatrix} 1 \\
-1 \end{bmatrix}, \quad g = \begin{bmatrix} 0 \\
k_w \end{bmatrix}
\]
f_s(z) = f_s(x_c - x_w) = f_s(z_1 - z_2).

The solutions \((x_c(t), x_w(t)) = (z_1(t), z_2(t))\) can be obtained by solving (31) numerically with small time step \( h > 0 \) using some method such as Runge-Kutta algorithm. However, it is doubtful whether the conventional numerical methods are suitable or not whenever the input \( u(t) \) is random process. The main reason is that the derivatives \( \dot{x}_c(t), \dot{x}_w(t), \dot{x}_c(t) \) and \( \ddot{x}_w(t) \) are merely formal, to the contrary, do not exist any longer in the ordinary sense\(^5\) because both \( x_c(t) \) and \( x_w(t) \) are random processes.

Keeping in mind, let us discretize the continuous-time differential equation (31) according to the idea of the mean forward and backward derivatives in stochastic mechanics\(^8\)-\(^8\). The derivation of discrete-time versions for (30) or (31) is given in Appendix B.

(iii) **Generation of Observation Data \( \{y(k)\} \)**:

The observations are made on the velocities of masses, \( \dot{x}_c(t) \) and \( \dot{x}_w(t) \), and the suspension length \( x_c(t) - x_w(t) \). Noting that \( \ddot{x}_c(t) = \dot{z}_1(t), \ddot{x}_w(t) = \dot{z}_2(t) \) and \( x_c(t) - x_w(t) = (z_1(t) - z_2(t) \), define the observations \( y_1(t), y_2(t) \) and \( y_3(t) \) (with each additive noise) as follows:
\[
\begin{aligned}
y_1(t) &= \dot{z}_1(t) + \text{noise} \\
y_2(t) &= \dot{z}_2(t) + \text{noise} \\
y_3(t) &= (z_1(t) - z_2(t)) + \text{noise}.
\end{aligned}
\] (32)

Let \( \{t_k\} \ (t_0 = 0, k = 0, 1, 2, \cdots) \) be sampling times for observation. The sampling interval \( \Delta = t_{k+1} - t_k \) is set several times of \( h \). According to Appendix B, the derivation \( \dot{z}_i(t) \) (\( i = 1, 2 \)) is discretized as
\[
\dot{z}_i(k) = \frac{1}{2h} \{ z_i(t_k + h) - z_i(t_k - h) \},
\]
so we have for \( i = 1, 2 \)
\[
y_i(t_k) = \frac{1}{2h} \{ z_i(t_k + h) - z_i(t_k - h) \} + v_i(t_k), \] (33)

where \( v_i(t_k) \) is modeled by a white Gaussian noise.

Here, let \( y(k) = y(t_k) = y(k\Delta) = [y_1(t_k), y_2(t_k), y_3(t_k)]^T \), \( Z_k = [z(t_k + h), z(t_k), z(t_k - h)]^T \in \mathbb{R}^3 \), and \( v_3(t_k) \) be also a white Gaussian noise. Then, the observation model is given by
\[
y(k) = H Z_k + v(k), \] (34)
in which \( v(k) = v(t_k) = [v_1(t_k), v_2(t_k), v_3(t_k)]^T \), and
\[
H = \begin{bmatrix} \frac{1}{2h} & 0 & 0 & -\frac{1}{2h} \\
0 & \frac{1}{2h} & 0 & 0 \\
0 & 0 & -1 & 0 \end{bmatrix}.
\]

(iii) **Selection of Basis Functions**:

Several basis functions in modeling nonlinear systems are introduced in the literature (e.g., see Refs. 9, 10). In our identified model (24), it should be noted that the computation of the matrix inverse is necessary for obtaining \( \theta(k) \). A crucial point is how we can guarantee the existence of the inverse of \( \Phi^T(x, u(k)) \). The observation model \( \Phi^T(x, u(k)) \) becomes a diagonal matrix,
\[
\Phi(x, u) = \begin{bmatrix} \varphi_{11}(x, u) & 0 & \cdots \\
0 & \ddots & \cdots \\
\varphi_{\hat{n}n}(x, u) & 0 & \cdots \end{bmatrix}.
\] (36)

Recalling the properties of the determinant for square matrices \( A, B \), \( \det (AB) = \det A \cdot \det B \) and \( \det A^T = \det A \), we have
\[
\det [\Phi^T(x, u) (\hat{C}^T W \hat{C}) \Phi(x, u)] = [\det \Phi(x, u)]^2 \cdot \det (\hat{C}^T W \hat{C}).
\]
From this we know that the inverse of \( \Phi^T(x, u) \cdot (\hat{C}^T \hat{W} \hat{C}) \Phi(x, u) \) exists as far as \( \det (\hat{C}^T \hat{W} \hat{C}) \neq 0 \) and \( \det \Phi(x, u) \neq 0 \) for any \((x(k), u(k))\) hold. Furthermore, since \( \Phi(x, u) \) is the diagonal matrix, its column vectors are linearly independent, so that we may set simply as
\[
\varphi_{11}(x, u) = \varphi_{22}(x, u) = \cdots = \varphi_{\hat{n} \hat{n}}(x, u).
\]

In the simulation studies, let \( \varphi_{ii}(x, u) \) be
\[
\varphi_{ii}[x(k), u(k)] = \sum_{\nu=1}^q \{\sin \nu x_i(k) + \cos \nu x_i(k)\}
\]
\( (i = 1, 2, \cdots, \hat{n}). \) (37)

(iv) Simulation Results:

In the simulation studies, as an input \( u(t) \) (road input) a white Gaussian noise with zero-mean and variance parameter \( Q_u = 10^2 \) was used, while as observation noise \( v(k) (= [v_1(k), v_2(k), v_3(k)]^T) \) three independent white noises with zero-means and the same variance were used such that \( R = 10^{-6} I \) (\( I \): unit matrix). Regarding physical parameters were set as: \( m_c = 600 \) [kg], \( m_w = 15 \) [kg], \( d_s = 5000 \) [N·m\(^{-1}\)·s], \( k_c = k_w = 35000 \) [N·m\(^{-1}\)], \( \xi_{\text{max}} = 60 \times 10^{-3} \) [m].

System matrices were identified using the N4SID algorithm and the system order was determined as \( \hat{n} = 4 \). As for the basis nonlinear function \( \varphi_i(\cdot, \cdot) \), we set it as (37) with \( q = 100 \), while for the user-defined parameters, \( \rho = 0.5 \), \( W = I \). Other parameters for simulations were: \( h = 10^{-2} \) [s], \( \Delta = 3 \times 10^{-4} (= 3h) \) [s], \( x_0(t_0) = x_w(t_0) = 0 \), \( \dot{x}_0(t_0) = \ddot{x}_w(t_0) = 0 \).

Several experiments were performed with various set of parameters \((\alpha, \beta)\) which influence the nonlinearity on the reaction force. A typical one of the simulations is depicted in Figs. 1 and 2 for \((\alpha, \beta) = (0.1, 0.1)\).

Figure 1 depicts the comparison of the outputs of the original observation data \( y(k) \) (shown by red solid line), the linear model \( y_L(k) \) (green dashed line) and the nonlinear model \( y_{NL}(k) \) (blue dash-dotted line). From the top to the bottom, each component concerning \( y_1 \), \( y_2 \), \( y_3 \) are shown. Top figure shows the output of the chassis (body) velocity \( \dot{x}_c(k) \), middele the wheel velocity \( \dot{x}_w(k) \) and the bottom the suspension length \( x_c(k) - x_w(k) \). From the figure, we see that our identification method tracks considerably well the original data.

Figure 2 illustrates the behaviors of the adjust parameters \( \{\theta_i(k)\} \ (i = 1, 2, 3, 4) \). Comparing Fig. 2 with Fig. 1, we recognize that when the nonlinearity outbreaks in the system dynamics, the adjust parameters work well.

In the experiments, we mention that a term \( \varepsilon I \) has added such that \( \{\varepsilon I + \Phi^T(x, u) \cdot (\hat{C}^T \hat{W} \hat{C}) \cdot \Phi(x, u)\}^{-1} \) in (24) to compute \( \theta(k) \) with \( \varepsilon = 10^3 \) to make sure the prevention of its explosion.

4 Conclusion

The usefulness of the nonlinear stochastic system identification algorithm proposed earlier by one of the authors has been demonstrated using a quarter-car model as a studied example. The crucial problem is how to select the user-defined basis functions. A
typical one is proposed in the numerical simulations, showing the good fitness to the original data.

References

APPENDIX A. Derivation of Quarter-Car Model
The Lagrangian of the quarter-car is
\[ L(t) = T(t) - V(t) \]  
(A.1)
in which \( T(t) \) and \( V(t) \) are kinetic and potential energies given by
\[ T(t) = \frac{1}{2} m_c \dot{x}_c^2(t) + \frac{1}{2} m_w \dot{x}_w^2(t), \]  
(A.2)
\[ V(t) = \frac{1}{2} k_w (x_w(t) - x_g(t))^2 + \int_0^{x(t)-x_w(t)} k_o(\xi) \, d\xi, \]  
(A.3)
where \( \xi = x_c - x_w \). Furthermore, the dissipation energy is
\[ F(t) = \frac{1}{2} d_c (\dot{x}_c(t) - \dot{x}_w(t))^2. \]  
(A.4)
Taking \( x_c(t) \) and \( x_w(t) \) as generalized coordinates, we have the Lagrange's equations of motion (11):
\[
\begin{align*}
\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{x}_c} \right) - \frac{\partial L}{\partial x_c} + \frac{\partial F}{\partial x_c} &= 0, \\
\frac{d}{dt} \left( \frac{\partial L}{\partial \dot{x}_w} \right) - \frac{\partial L}{\partial x_w} + \frac{\partial F}{\partial x_w} &= 0.
\end{align*}
\]
(A.5)
Substituting (A.1) and (A.4) into (A.5) and noting that
\[ \frac{\partial L}{\partial x_c} = -k_0(\xi) \int_0^{x_c-x_w} k_o(\xi) \, d\xi \frac{\partial \xi}{\partial x_c} - k_0(\xi) (x_c(t) - x_w(t)) \]
\[ \frac{\partial L}{\partial x_w} = -k_0(\xi) \int_0^{x_c-x_w} k_o(\xi) \, d\xi \frac{\partial \xi}{\partial x_w} + k_0(\xi) \int_0^{x_c-x_w} k_o(\xi) \, d\xi.
\]
(Q.E.D.)

APPENDIX B. Discretization of Differential Equations with Additive Random Disturbance
As mentioned in Section 3, the derivatives of \( x_c(t) \) (\( \equiv z_1(t) \)) and \( x_w(t) \) (\( \equiv z_2(t) \)) have no their mathematical meanings as derivatives if the additive input is random. So, in such a case, any formal numerical method introduced in the literature are inadequate. To remedy the situation, we use the following two mean derivatives for a stochastic process \( z(t) \), according to the Nelson's stochastic mechanics (12), (13):
\[ D_z(t) = \lim_{h \downarrow 0} E \left\{ \frac{z(t+h) - z(t)}{h} \right\} \left| P_t \right\} \]  
(B.1)
\[ D_{\theta} z(t) = \lim_{h \downarrow 0} E \left\{ \frac{z(t) - z(t-h)}{h} \right\} \left| F_t \right\}, \]  
(B.2)

1425
where $\mathcal{E}\{\cdot\}$ denotes the expectation, and $\mathcal{P}_t$, $\mathcal{F}_t$ are increasing and decreasing families of $\sigma$-algebra of the process $z(t)$ such that $\mathcal{P}_s \subset \mathcal{P}_t$ and $\mathcal{F}_s \supset \mathcal{F}_t$ for $s \leq t$. These are called mean forward and backward derivatives, respectively. The conditions in (B.1) and (B.2) may be replaced by the realization $z(t) = \hat{z}$ at the present time $t$. If $z(t)$ is smooth, then $Dz(t) = D\hat{z}(t) = dz(t)/dt$

Based on these definitions, let us derive the discrete version for the second-order stochastic equation (31), assuming that the input $u(t)$ is a random process. Then, the formal derivatives $\hat{z}(t)$ and $\ddot{z}(t)$ in the Langevin-type equation (31) should be interpreted as

$$\begin{align*}
\hat{z}(t) &\sim \frac{1}{2} \{ Dz(t) + D_s z(t) \} \\
\ddot{z}(t) &\sim \frac{1}{4} \{ D^2 D_s z(t) + D_s Dz(t) \}.
\end{align*}$$

Therefore, (31) should be replaced by

$$M \frac{1}{2} \{ D^2 D_s z(t) + D_s Dz(t) \} + C \frac{1}{2} \{ Dz(t) + D_s z(t) \} + K z(t) + c_0 f_s[z(t)] = g u(t). \tag{B.4}$$

For the numerical computation of (B.4), let $h(t > 0)$ be the (sufficiently small) step length and denote as $t_j = jh$ ($t_0 = 0$, $j = 0, 1, 2, \cdots$). Approximating as

$$\begin{align*}
Dz(t_j) &\approx \frac{z(t_{j+1}) - z(t_j)}{h} \\
D_s z(t_j) &\approx \frac{z(t_j) - z(t_{j-1})}{h},
\end{align*}$$

we have

$$\begin{align*}
&\frac{1}{2} \{ Dz(t_j) + D_s z(t_j) \} = \frac{1}{2h} \{ z(t_{j+1}) - z(t_{j-1}) \} \\
&\frac{1}{2} \{ D^2 D_s z(t_j) + D_s Dz(t_j) \} \\
&= \frac{1}{h^2} \{ z(t_{j+1}) - 2z(t_j) + z(t_{j-1}) \},
\end{align*}$$

so that (B.4) is discretized as

$$\frac{M}{h^2} \{ z(t_{j+1}) - 2z(t_j) + z(t_{j-1}) \} + \frac{C}{2h} \{ z(t_{j+1}) - z(t_{j-1}) \} + K z(t_j) + c_0 f_s[z(t_j)] = g u(t_j). \tag{B.7}$$

Rearranging terms in (B.7) yields

$$M_d z(t_{j+1}) + C_d z(t_j) + K_d z(t_{j-1})$$

$$+ h^2 c_0 f_s[z(t_j)] = h g u(t_j) \tag{B.8}$$

in which

$$M_d = \frac{M}{h} + \frac{C}{2}, \quad C_d = -\frac{2M}{h} + hK,$$

$$K_d = \frac{M}{h} - \frac{C}{2}. \tag{B.9}$$

It is noted that $M_d, C_d$ and $K_d$ have no their meanings of mass, damping and stiffness in (B.8). Such equation as (B.8) is called the discrete structural dynamics.$^{77}$

Initial conditions for (B.8) are given as follows. For (30), let the initial conditions be

$$\begin{align*}
x_e(0) &= x_{e0}, \quad \dot{x}_e(0) = \dot{x}_{e0}; \\
x_w(0) &= x_{w0}, \quad \dot{x}_w(0) = \dot{x}_{w0}.
\end{align*}$$

Then, $z(0) = [x_{e0}; x_{w0}]^T \equiv \tilde{z}_0$ and $\dot{z}(0) = [\dot{x}_{e0}; \dot{x}_{w0}]^T \equiv \dot{\tilde{z}}_0$ are given. Since $\tilde{z}_0 = \frac{1}{2h} \{ z(t_1) - z(t_{-1}) \}$, we get $z(t_0) = 0$ for convenience to get the initial conditions for (B.8) as

$$\begin{align*}
z(t_1) &= z(h) = 2h\tilde{z}_0, \quad z(t_0) = z(0) = \tilde{z}_0, \\
z(t_1) &= z(-h) = 0.
\end{align*}$$

Solving (B.8) under these initial conditions with random input $u(t_j)$, we can obtain the solution sequence $\{z(t_j)\}$, or $\{x_e(t_j), x_w(t_j)\}$ ($j = 0, 1, 2, \cdots$).

(Q.E.F.)