Journal of Computer Engineering
ISSN 10003428

Original Research Article
Vol 12, Issue 4, 2023 | ISSN: 1000-3428

A Novel Optimization Framework for Identifying Hammerstein–Wiener Nonlinear Systems Under Noisy Conditions

Dr. Fatima Rizvi, Dr. Kaveh Mirzaei, Dr. Amir Hossein Ghasemi

Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, Canada; Department of Mechanical Engineering, Sharif University of Technology, Tehran, Iran

Abstract—Hammerstein–Wiener model is a block-oriented model where a linear dynamic system is surrounded by two static nonlinearities at its input and output and could be used to model various processes. This paper contains an optimization approach method for analysing the problem of Hammerstein–Wiener systems identification. The method relies on reformulate the identification problem; solve it as constraint quadratic problem and analysing its solutions. During the formulation of the problem, effects of adding noise to both input and output signals of nonlinear blocks and disturbance to linear block, in the emerged equations are discussed. Additionally, the possible parametric form of matrix operations to reduce the equation size is presented. To analyse the possible solutions to the mentioned system of equations, a method to reduce the difference between the number of equations and number of unknown variables by formulate and importing existing knowledge about nonlinear functions is presented. Obtained equations are applied to an instance H–W system to validate the results and illustrate the proposed method.

I. Introduction

ONE of the famous categories of nonlinear systems is named 'block oriented systems', in which the system is assumed to consist of a combination of linear dynamic blocks and nonlinear static blocks. This category is further divided into different subcategories. Two of these typical subcategories are known as Hammerstein systems and Wiener systems, which are a serial combination of a static nonlinear block and a linear dynamic one. By serializing a Hammerstein and a Wiener system, a new subcategory named 'Hammerstein–Wiener' emerged, which is also the focus of this study [1].

The above mentioned categories could be used to model different systems in various applications such as continuous stirred tank reactors [2], Quality of Service (QoS) performance and resource management of software systems [3], pH neutralisation processes [4], or Brushless DC (BLDC) motors [5]. Moreover, static nonlinearity can be seen in sensors and actuators, so a linear system having these sensors and actuators can be modelled as a Hammerstein–Wiener system [1]. Various methods have been used in previous researches to identify Hammerstein, Wiener, or Hammerstein–Wiener systems from 1998 till date.

Examples of a wide range methods for the identification of Hammerstein systems is the algorithm uses the Singular Values Decomposition (SVD) technique and triangular basis functions and an approach for Wiener systems identification is the semiparametric Bayesian [6], [7]. Regarding the identification of Hammerstein–Wiener systems, one of the first methods used was a two stage identification algorithm; the stages included using the recursive least squares method and calculation of singular value decomposition of two matrices, whose dimensions are fixed and do not depend on the number of the data points [8]. Four years later, the researcher presented a blind approach to solve the problem [9]. In the first mentioned method, nonlinear blocks were assumed to be approximated by a polynomial or series of orthogonal functions while in the second one there was no particular shape. Some researchers deal with the identification of the H–W by iterative solutions, such as the method in which nonlinear functions are approximated by cubic splines [10]. Further, there are frequency based methods which consider a combination of sinusoids with a random phase as the system input [11]. Apart from these methods, the identification problem is also solved with using refined instrumental variable method [12] and maximum likelihood method [13].

The objective of this study is to propose a new evaluation method for the H–W system identification problem and study its requirements, useful approximations and assumptions. First, by considering the state space form of the linear block, the identification problem's the relationships among the sample data and unknown variables are formulated in the form of a system of equations. Then, several unknown variables and equations are reduced by means of matrix operations. Following this, by considering the number of equations, unknown variables and their coefficients, the number of possible solutions is evaluated. Then, the matrix sizes are reduced by applying elementary row operation on them, and a solution to find values of nonlinear functions will be proposed. Next, the effect of adding noise to input and output signals of nonlinear blocks and disturbance to the linear block on the mentioned equations is considered, and values of unknown functions are found using the constrained least–squares method. Finally, formulated knowledge about nonlinear functions are described for reducing the number of unknown variables and also proper assumptions that could increase the known parameters and limit the set of solutions are considered. All proposed methods are examined by applying them on a sample H-W system to validate the equations. The approach presented

How to cite: Dr. Fatima Rizvi, Dr. Kaveh Mirzaei, Dr. Amir Hossein Ghasemi (2023). A Novel Optimization Framework for Identifying Hammerstein–Wiener Nonlinear Systems Under Noisy Conditions. Journal of Computer Engineering, 12(4), 47–55. DOI: https://doi.org/10.5281/zenodo.19348490

here differs from existing literature in followings; first, the focus is on arriving at an analytical solution to the problem rather than a numerical one, second, it considered multivariate systems and third, it analysied the effect of adding noise to both input and output of nonlinear blocks.

This paper is organized as follows. In Section II the problem formulation is presented, and it is simplified and discussed in Section III. Section IV provides details the enhancement of the method in the presence of noise and disturbance; the solution to the problem using a constrained least–squares algorithm is presented. Section V presents simulation results, and lastly, the conclusions are discussed in Section VI.

II. Problem Statement

For a single input single output (SISO) system with an SISO linear block, the H–W model with an injective output function, which is illustrated in Fig. 1, can be written as shown below by considering the state space form of the linear system.

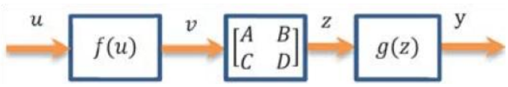


Fig. 1 Hammerstein–Wiener system block diagram

$$\begin{bmatrix} x_{t+1} \\ z_t \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} x_t \\ y_t \end{bmatrix}$$
$$\rightarrow \begin{bmatrix} x_{t+1} \\ g^{-1}(y_t) \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} x_t \\ f(u_t) \end{bmatrix} \quad (1)$$

If the samples from input u and output y are gathered during t=1 to T times of an arbitrary time unit, all gathered samples can be formulated, as seen in (3) in the matrix form, after defining the matrices shown in (2).

x_1^2 ⎤⎡ uu^{12} ⎤, F ⎡ $f f(u^u u_{..T}{}^{12})$ ⎤ x ( )

$X=$ ⎡ $x^{\cdots}{}_{T01}$ ⎤, $U=$ ⎡ $u_{\cdot T1}$ ⎤ $y_1=$

x ( ) y g- ( ) X

$y_{.2}$ ⎡ $g^{-1}(y_2)$ ⎤, φ ⎡ $GF-$ ⎤ (2)

$Y =$, $G=$ ⎡ g 1..y_T =

⎣ y_T ⎦ -( )

$$\Psi = \begin{bmatrix} \Psi_{(1,1)} & \Psi_{(1,2)} & \Psi_{(1,3)} \\ \Psi_{(2,1)} & \Psi_{(2,2)} & \Psi_{(2,3)} \end{bmatrix}$$

$$\Psi\varphi = 0$$

$$\Psi_{(1,1)} = \begin{bmatrix} I & 0 & \cdots & 0 & -0A \\ -0A & I & \cdots & 0 & \\ & -A & \cdots & 0 & \\ & \cdots & \cdots & \cdots & \\ & \cdots & -A & I & 00.. \end{bmatrix},$$

$$\Psi_{(2,1)} = \begin{bmatrix} -C & 0 & \cdots & 0 & 0 \\ 0 & -C & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & -C & 0 \end{bmatrix} \quad (3)$$

$\Psi_{(1,2)} = -I_{T*T} \otimes B$, $\Psi_{(1,3)} = -0_{T*T} \otimes B$ $\Psi_{(2,2)} = -I_{T*T} \otimes D$, $\Psi_{(2,3)} = -I_{T*T}$

As it can be seen, if the linear system has an order of n, the number of rows in Ψ will be T(n+1), while the number of columns is (T+1)n+2T which is equal to unknown variables that are collected in the Φ vector. Therefore, the problem can have several solutions. It is worth mentioning that the above homogenous system always has the trivial solution Φ=0 which means $f(u)=g^{-1}(y)=0$ for every u and y. In the following sections, methods of evaluating and limiting these solutions will be considered.

III. Proposed Method for Noise Free System

The static behaviour of f and g^{-1} implies the same outputs in the presence of the same inputs. By defining F^* and G^* and changing the above equation to the following form, the static behaviour of the first and last blocks will be considered. It is worth mentioning that a unique function returns the same data as in its input but with no repetitions.

(4)

$-$

$\lceil\ \Psi_{(1,1)} \quad | \ \Psi_{*(1,2)} \quad | \quad 0 \quad \rceil$

$\Psi_{*} = || \lfloor - + - + - \quad ||$

$\Psi_{(2,1)} \rightarrow | \ \Psi\Psi_{*}\varphi^{*}{}_{*(2=}\ 0_{,2)} | \ \Psi_{*(2,3)}$

Owing to the special form of (1,1) and (2,1) matrix blocks in Ψ, matrix elementary row operations can be used to reduce the number of equations and unknown variables. The effect of elementary row operation on Ψ is shown in (5) and (6).

$$\Psi'^{(1,1)} = \begin{bmatrix} I & 0 & \cdots & 0 & -A \\ 0 & I & \ldots & 0 & -A^2 \\ 0 & 0 & \ldots & 0 & -A^3 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \ldots & 0 & I & -A^T \end{bmatrix}$$

$$\Psi'^{(2,1)} = \begin{bmatrix} 0 & 0 & \cdots & 0 & -\sum_{i=1}^{k} c_i A_{(i,:)} \\ 0 & 0 & \ldots & 0 & -\sum_{i=1}^{k} c_i A^2{}_{(i,:)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 0 & -\sum_{i=1}^{k} c_i A^T{}_{(i,:)} \end{bmatrix} \qquad (5)$$

$$\Psi'^{(1,2)}_{t,i} = \sum_{l=0}^{t-1} \left(A^l \Psi^{(1,2)}_{t-l,i} \right) = -\sum_{l=0}^{t-1} \left(A^l B \Gamma^{(1)}_{t-l,i} \right)$$

$$\Psi'^{(2,2)}_{t,i} = \left(\sum_{p=1}^{k} c_p \left(\Psi^{(1,2)'}_{t,i} \right)_p \right) + \Psi^{(2,2)}_{t,i} = -D*\Gamma^{(1)}_{t,i} - \sum_{l=0}^{t-1} \left(\left(\sum_{p=1}^{k} c_p \left(A^l\right)_p \right) *B*\Gamma^{(1)}_{t-l,i} \right)$$

$$\Psi'^{(2,3)} = -\Gamma^{(2)} \otimes I_T$$

$$\Psi = \begin{bmatrix} - & + & - & + & - \\ \Psi'^{(2,1)} & | & \Psi'^{(2,2)} & | & \Psi'^{(2,3)} \end{bmatrix} \rightarrow \Psi'\phi^* = 0 \qquad (6)$$

$, \ \lceil\ \Psi'^{(1,1)} \quad | \quad \Psi'^{(1,2)} \quad | \quad 0 \quad \rceil$

If the state–space realization of the linear system has an observable canonical form, it can be reduced even further by applying elementary row operation, as observed in (7) and (8).

$\lceil 0 \quad 0 \quad \ldots\ldots$

$0 \quad -a a_{nn}]|||||||\ 1$

$0 \quad \ldots\ldots \quad 0$

$-a_{n-1}\ |||||||]$

$A= \quad 0 \quad 1 \quad 0 \quad - \quad {}_{-2}$

$\ldots \quad \ldots \quad \ldots \quad \ldots \quad \ldots$

$\lfloor$

$b 0_n\ 0 \quad 1 \quad -a_1$

$[|||| \ b_n \ ]||||| \ C [||| \ 0 \ \ldots\ldots \ 0 \ 1 \ ]||]$

$\lfloor$

$B= \quad \ldots{}_{-1} \quad = \quad \ldots \quad \ldots \quad \ldots \quad \ldots \qquad (7)$

$0 \quad 0 \quad 1$

b_1

$$A^t{}_{n,:} := n^{th}\ row\ of\ A^t \quad , \quad a_0 := 1$$

$$\begin{cases} A^t{}_{n,:} = A_{n-t+1,:} - \sum_{i=1}^{t-1} a_i A^{t-i}{}_{n,:} & t \le n \\ \sum_{i=0}^{n} a_i A^{t-i}{}_{n,:} = 0 & t > n \end{cases}$$

$$\forall n<t\leq T:\ \begin{cases} \Psi"^{(1)}_{t-n,i}=\sum_{r=0}^{n} a_r\, \Psi'^{(2,2)'}_{t-r,i} \\ \Psi"^{(2)}_{t-n,j}=\sum_{r=0}^{n} a_r\, \Psi^{(2,3)}_{t-r,j} \end{cases}$$

$$\rightarrow \begin{cases} \Psi"^{(1)}_{t-n,i}=-\sum_{r=0}^{n} a_r X_r \\ X_r = D\Gamma^{(1)}_{t-r,i}+\sum_{h=0}^{t-r-1} X_h \\ X_h = \left(A^h\right)_{n,:} B\Gamma^{(1)}_{t-r-h,i} \\ \Psi"^{(2)}_{t-n,j}=-\sum_{r=0}^{n} a_r \Gamma^{(2)}_{t-r,j} \end{cases}$$

$$\Psi"=\left[\ \Psi"^{(1)}\ \ \middle|\ \ \Psi"^{(2)}\ \right],\ \phi"^{*}=\begin{bmatrix} F^{*} \\ - \\ G^{*} \end{bmatrix}$$

$$\rightarrow \Psi"\,\phi"^{*}=0 \tag{8}$$

The equation can be easily simplified to what is shown in (9).

$$\Psi"^{*}=\left[\ \Psi"^{*(1)}\ \ \middle|\ \ \Psi"^{*(2)}\ \right]$$

$$\Psi"^{*(1)}_{t-n,i}=D*\Gamma^{(1)}_{t-n,i}+\sum_{h=0}^{t-n-1}\left(\left(A^h\right)_{n,:}B\Gamma^{(1)}_{t-n-h,i}\right) \tag{9}$$

$$\Psi"^{*(2)}_{t-n,j}=\Gamma^{(2)}_{t-n,j}$$

$$\Psi"^{*}\varphi"^{*}=0$$

The number solutions to the above problem can be infinite if the rank of Ψ is less than the number of elements in Φ. By considering (9), number of free variables can be determined
using (10).

$$\begin{cases} rank\left(\Psi"^{*}\right)\leq min(T-n,l+k) \\ \qquad T \\ \phi"^{*}\ elements=l+k\leq 2 \\ free\ \text{variables}=l+k-rank\left(\Psi"^{*}\right) \end{cases} \tag{10}$$

$$\rightarrow\begin{cases} free\ \text{variables}\geq l+k+\ - \\ free\ \text{variables}\geq 0 \qquad n \end{cases}$$

$$T$$

According to definition of $\Gamma^{(1)}$ and $\Gamma^{(2)}$, the number of free variables is equal to values of F* and G* that should be known to find unique solutions to other elements of them. To have at least one non–zero solution, there should be more than one free variables. So l+k+n should be less than T, and the number of non–unique elements in U and Y should be more than T+n. Hence, the number of known value points of functions should be equal to the number of free variables; the values for these points should be known to find the values at other points. According to the samples' quantities and their similarities, if the repetition is sufficiently high, linear system coefficients can be determined by setting the rank of Ψ"* equal to l+k-1. In this case, only one known point in either the input or output function values is sufficient to determine the values of functions at all other points. By separation of the known points and unknown points in Ψ"* and sending the known points plus their coefficients to the other side of the equality sign in equation 9, another system of linear equations will emerge that has only one solution which could be obtained as explained below.

There are a variety of methods to solve a system of linear equations. The following equation shows the simplest form of solving the equation in which Ψ^{+} is called the pseudo–inverse, left inverse, or generalized inverse of Ψ [14], [15].

$$\Psi\phi=\Upsilon\rightarrow\phi=\left(\Psi^T\Psi\right)^{-1}\Psi^T\Upsilon\rightarrow\phi=\Psi^{+}\Upsilon \tag{11}$$

By considering the singular value decomposition of Ψ, its pseudo–inverse could be calculated as follows. In this equation, Σ is a diagonal matrix with singular values of Ψ on the diagonal [15].

$$\Psi=U\Sigma V^{T}\rightarrow\Psi^{+}=V\Sigma^{+}U^{T} \tag{12}$$

However, it should be considered that linking the H–W identification problem to the above solution for a system of equations makes it very sensitive to even small errors in calculations and measurements.

IV. Robustness Enhancement

First, to make the solution less sensitive to noise, disturbance and errors, the effect of adding Gaussian noise and disturbance to the H–W model in the emerged equations will be considered. The error, noise, and disturbance will be added to the system input or output as linear block inputs, outputs, or states. Then, solutions of constrained linear least–squares problems will be used to solve the system of linear equations in a more robust manner.

A. Formulate Linear Block Noise Effect

Noise and disturbance could be added to linear block inputs, states, and outputs as shown in Fig. 2. The addition of noise to the system changes (1) into the following form:

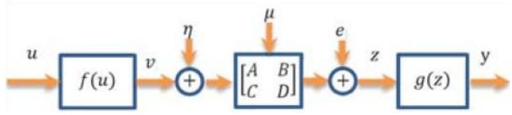


Fig. 2 H–W system with additional noise to linear block

$$\begin{bmatrix} x_{t+1} \\ g^{-1}(y_t) \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} x_t \\ f(u_t)+\eta_t \end{bmatrix} + \begin{bmatrix} \mu_t \\ e_t \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} x_t \\ f(u_t) \end{bmatrix} + \begin{bmatrix} \mu_t + B\eta_t \\ e_t + D\eta_t \end{bmatrix} \quad (13)$$

If the mentioned noises are independent of each other and spread through the time with Gaussian distribution, (3) will then be change to the following form.

$$\begin{cases} \mu_t \sim N(\mu, \Sigma_\mu^T \Sigma_\mu) \\ \eta_t \sim N(\eta, \Sigma_\eta^T \Sigma_\eta) \quad \rightarrow \Psi\phi = \Upsilon \\ e_t \sim N(e, \Sigma_e^T \Sigma_e) \end{cases}$$

$$\Upsilon \sim N\left(\begin{bmatrix} \Upsilon_1 \\ \Upsilon_2 \end{bmatrix}, \begin{bmatrix} \Sigma_1^T \Sigma_1 & 0 \\ 0 & \Sigma_2^T \Sigma_1 \end{bmatrix} \right) \quad (14)$$

$$\Upsilon_1 = (\mu + B\eta) \otimes I_{T+1}$$
$$\Upsilon_2 = (e + D\eta) \otimes I_T$$
$$\Sigma_1^T \Sigma_1 = \left(\Sigma_\mu^T \Sigma_\mu + B \Sigma_\eta^T \Sigma_\eta B^T \right) \otimes I_{T+1}$$
$$\Sigma_2^T \Sigma_2 = \left(\Sigma_e^T \Sigma_e + D \Sigma_\eta^T \Sigma_\eta D^T \right) \otimes I_T$$

Multiplying the matrix Ψ by Ξ, calculated from Cholesky factorization of the variance matrix, the variance of the resulted Y is normalised as shown below.

$$\Xi \Xi^T = \begin{bmatrix} \Sigma_1^T \Sigma_1 & 0 \\ 0 & \Sigma_2^T \Sigma_2 \end{bmatrix} \rightarrow$$
$$\Xi^{-1} \Psi \phi \sim N\left(\Xi^{-1} \begin{bmatrix} \Upsilon_1 \\ \Upsilon_2 \end{bmatrix}, I_{(n(T+1)+size_{z_t} T)} \right) \quad (15)$$

B. Formulate Measurement Noise Effect

When the Gaussian independent measurement noise is added to the input and output of the system as illustrated in Fig. 3, (14) changes as follows.

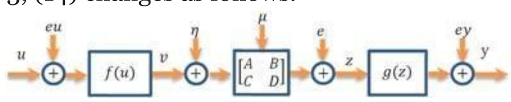


Fig. 3 H–W system with additional noise to system input, system output, and linear block

$$\Psi\varphi + \Psi_e \varphi_e = Y$$

$$F_e = \begin{bmatrix} f(u_1+eu_1) - f(u_1) \\ f(u_2+eu_2) - f(u_2) \\ \vdots \\ f(u_T+eu_T) - f(u_T) \end{bmatrix},$$

$$G_e = \begin{bmatrix} g^{-1}(y_1+ey_1) - g^{-1}(y_1) \\ g^{-1}(y_2+ey_2) - g^{-1}(y_2) \\ \vdots \\ g^{-1}(y_T+ey_T) - g^{-1}(y_T) \end{bmatrix}, \quad (16)$$

$$\phi_e = \begin{bmatrix} F_e \\ - \\ G_e \end{bmatrix}, \quad \Psi_e = \begin{bmatrix} \Psi^{(1,2)} & | & \Psi^{(1,3)} \\ - & + & - \\ \Psi^{(2,2)} & | & \Psi^{(2,3)} \end{bmatrix}$$

By considering the first order approximation and Taylor's theorem shown in (17), (16) is simplified to the form shown in (18), if the noise value is small enough. It is worth mentioning that $D_f(u)$ represents the matrix containing partial derivatives of the function f.

$$f(u+e) - f(u) \approx D_f(u)e$$
$$if\ SISO\ case:\ D_f(u) = \frac{\partial f}{\partial u}(u)$$
$$\approx \frac{f(u+\Delta u_1) - f(u)}{2\Delta u_1} - \frac{f(u-\Delta u_2) - f(u)}{2\Delta u_2} \rightarrow$$
$$f(u+e) - f(u) \approx e\Delta \begin{bmatrix} f(u-\Delta u_2) \\ f(u) \\ f(u+\Delta u_1) \end{bmatrix}$$
$$\Delta = \begin{bmatrix} \left(\frac{-1}{2\Delta u_2}\right) & \left(\frac{1}{2\Delta u_2} - \frac{1}{2\Delta u_1}\right) & \left(\frac{1}{2\Delta u_1}\right) \end{bmatrix} \quad (17)$$

$$\begin{cases} \Psi_e^{(1,2)} = \Psi^{(1,2)} + \Psi^{(1,2)} E_u \Delta_u \\ \Psi_e^{(2,2)} = \Psi^{(2,2)} + \Psi^{(2,2)} E_u \Delta_u \\ \Psi_e^{(1,3)} = \Psi^{(1,3)} + \Psi^{(1,3)} E_y \Delta_y \\ \Psi_e^{(2,3)} = \Psi^{(2,3)} + \Psi^{(2,3)} E_y \Delta_y \end{cases}$$

$$\Psi' = \begin{bmatrix} \Psi^{(1,1)} & | & \Psi_{e(1,2)} & | & \Psi_{e(1,3)} \\ - & + & - & + & - \\ \Psi_{(2,1)} & | & \Psi_{e(2,2)} & | & \Psi_{e(2,3)} \end{bmatrix}$$

(18)

e

$$\Psi_e' \phi \approx \Upsilon$$
$$\Psi_e' = \Psi + \Delta_\Psi$$
$$mean(\Delta_\Psi) = 0$$

C. Optimal Solution

To find an optimal solution for (15), instead of a solution to (11), algorithms to solve a constrained linear system of equations, which are shown in (19), could be used. As it can be seen in (19), these algorithms are linear least–squares' solvers with bounds or linear constraints [16], [17].

$$\min_{\phi} \frac{1}{2}\| \Psi\phi - \Upsilon\|_2^2 \ \ such\ that \begin{cases} A\phi \le b \\ C\phi = d \\ \phi_{lb} \le \phi \le \phi_{ub} \end{cases} \tag{19}$$

Apart from adding some known elements of Φ to the equations, using the above algorithm makes it possible to limit the solution and enhance the problem by considering other assumptions and approximations. These assumptions could be the determining limits for elements of Φ; they could even be determining limits for the derivatives of input or output functions obtained from the approximations in (16) by changing them to linear constraints. The gradients in this case are calculated by subtracting the function values of two consecutive outputs and dividing the result by their difference. In case of a limited range assumption for function values, the lower bounds and upper bounds for unknown variables which are shown by Φ_{lb} and Φ_{ub} are known. Additionally, the gradient of input and output functions can be limited by considering the first order approximation in calculating the gradient and the construction of matrix A and vector b.

Another useful approximation could be the quantization of the input and output range. From the point of view of a linear system of equations, this approximation will reduce the number of unknown variables by increasing the repetitions of measured samples.

One of the trivial solutions for Φ is the unique value of X, Y and Z elements, which obey the following equation. To avoid this situation, if possible, the minimum difference between two points in f or g^{-1} functions can be determined as another assumption.

$$\forall t : 1..T : \begin{cases} x_t = x_0 \\ f(u_t) = f_0 \\ g^{-1}(y_t) = g_0^{-1} \end{cases} , \quad \begin{cases} x_0 = (I-A)^{-1} B f_0 \\ g_0^{-1} = \left(C(I-A)^{-1}B + D \right) f_0 \end{cases} \tag{20}$$

The above examples present only a few of the various constraints that could be applied to the system for obtaining the desired solution and limit the number of possible solutions to the mentioned system of linear equations. In other words, by increasing the assumption set of possible solutions, number of solutions will be decreased.

V. Simulation Results

An example of an H–W system can be seen in Fig. 4.

The first subplot shows the measured input while the last one shows the output of the system. Nonlinear functions are illustrated by the second and fifth subplots. The third and fourth subplots belong to the immeasurable input and output of the linear block, respectively. The states are also shown at the bottom of this figure.

Fig. 5 shows the result of solving (4) by calculating the pseudo–inverse of matrix Ψ, illustrated in (11), to find the values of nonlinear functions in sample points. In this system T, n, l and k are equal to 201, 2, 197, and 201, respectively. The red stars are the 203 known points of the functions while the blue stars are the 603 calculated ones when solving the problem. As it can be seen in most of the points, the blue stars are superimposed on the real values which are shown by green stars.

It is important to remember that determining the solution using this method could generate many imperfect results because of the existence of very small errors in calculation or noise. This is because of the nature of the solutions for the problem involving a system of linear equations. To reduce the errors in this case, apart from using iterative algorithms for the problem in (19), some assumptions based on the known properties of input and output functions, such as a limited value range, could be considered.

To solve the problem in the above example, an approximation of equality is considered. The approximation involves determining each two consequent input points in

Journal of Computer Engineering
ISSN 10003428

Original Research Article
Vol 12, Issue 4, 2023 | ISSN: 1000-3428

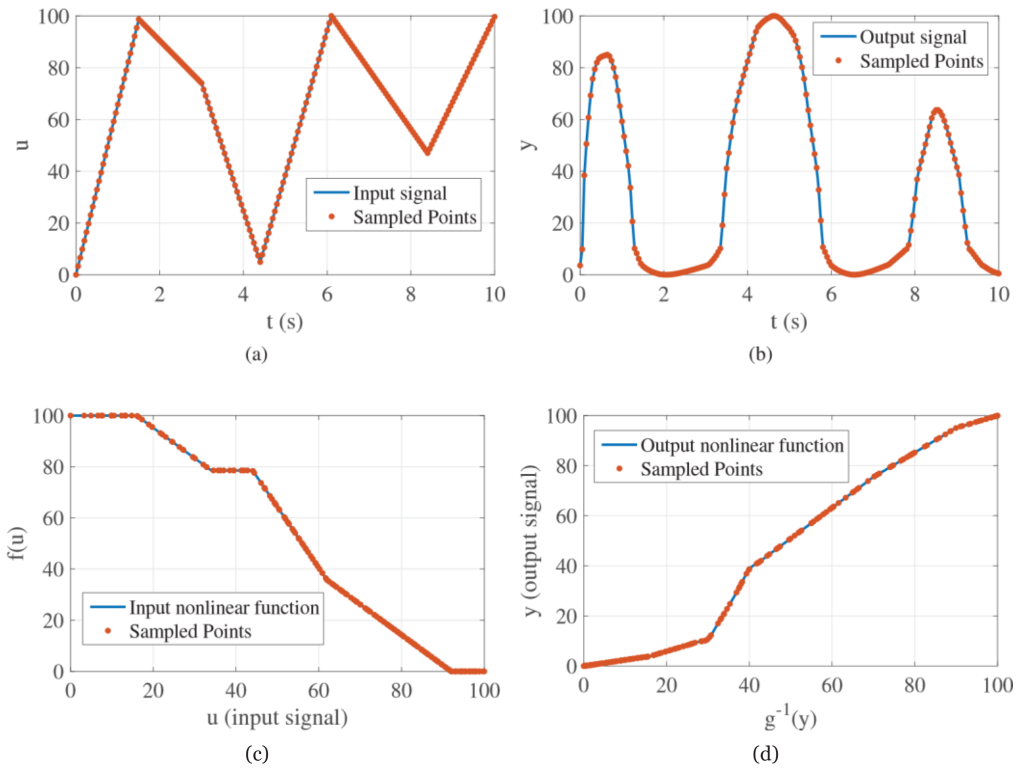


Fig. 4 An H–W system example, sample points in each subplot shown by red dots. (a), (b) Input & Output Signals. (c), (d) nonlinear functions

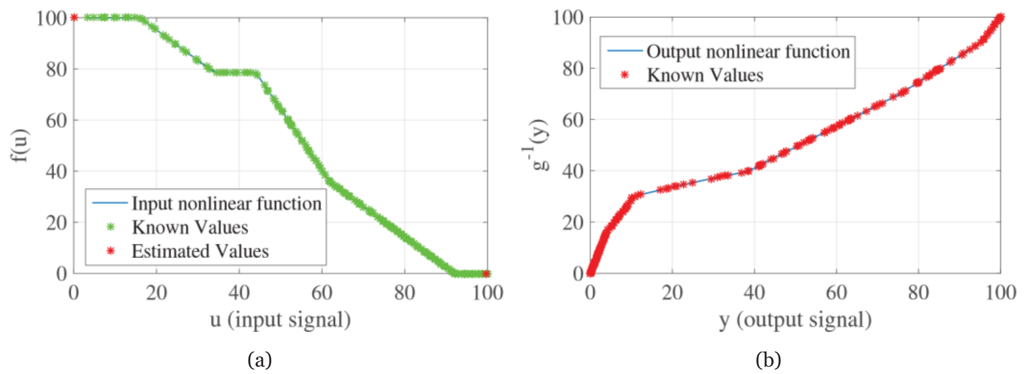


Fig. 5 Identification result of nonlinear static functions of example H–W system while some of the values are known

the functions that have the same output values. Using this approach for all points reduce the number of unknowns in (9) by half. Another similar approximation is to quantize the input and output range, and assume equal function values for each section. Each of these quantized discrete values should contain at least one sample point, and the number of intervals should be selected to have only one free variable according to (9). Fig. 6 shows the result of input quantization in f and g^{-1} to 12 parts in the u and y axes, in the example. As seen in this example, only one known point is needed. The results for input and output functions are limited between -10% and 110%.

As seen in Fig. 6, by applying this method to the sample system shown in Fig. 4, the number of variables that should be known before identification of nonlinear functions reduces to one, and the result seems to be acceptable despite quantization errors in input and output functions.

A. Simulation Results in Presence of Noise

To have an intuitive view of the effect of noise and disturbance on the presented methods and compare them to systems without noise, the simulations were repeated. Fig. 7 shows the signals in Fig. 3's system in the presence of 3% of signal range noise in each signal to which noise was added. The noise is large enough to be seen in the picture shown in Fig. 7.

How to cite: Dr. Fatima Rizvi, Dr. Kaveh Mirzaei, Dr. Amir Hossein Ghasemi (2023). A Novel Optimization Framework for Identifying Hammerstein–Wiener Nonlinear Systems Under Noisy Conditions. Journal of Computer Engineering, 12(4), 47–55. DOI: https://doi.org/10.5281/zenodo.19348490

The resulting output of the constrained iterative algorithm for the above system in both quantized and non–quantized input and output functions is shown in Figs. 8 and 9, respectively. As seen, the existence of Gaussian noise makes the solution even better in case of quantization. The non–Gaussian quantization error, which exists in this example as the slope of the functions, affected the results and distanced them from desired values.

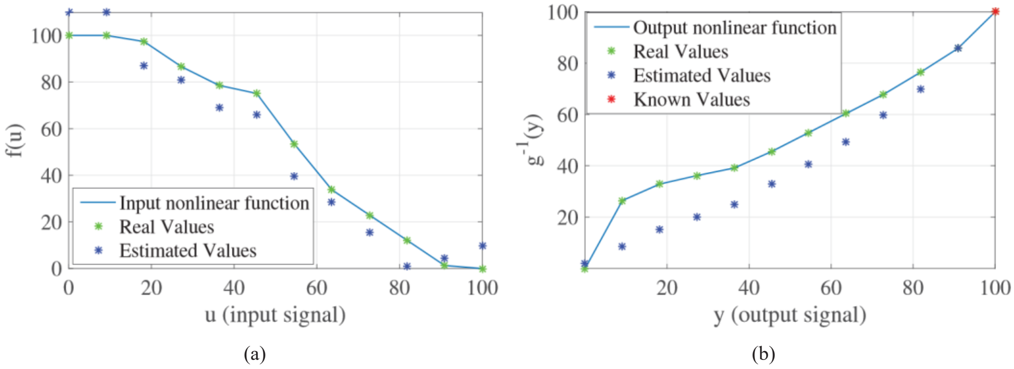


Fig. 6 Estimating input and output functions in instant H–W system by using constrained linear least–squares algorithm, after quantization input of f and g^{-1} to 12 parts in u and y axes and considering only one known point

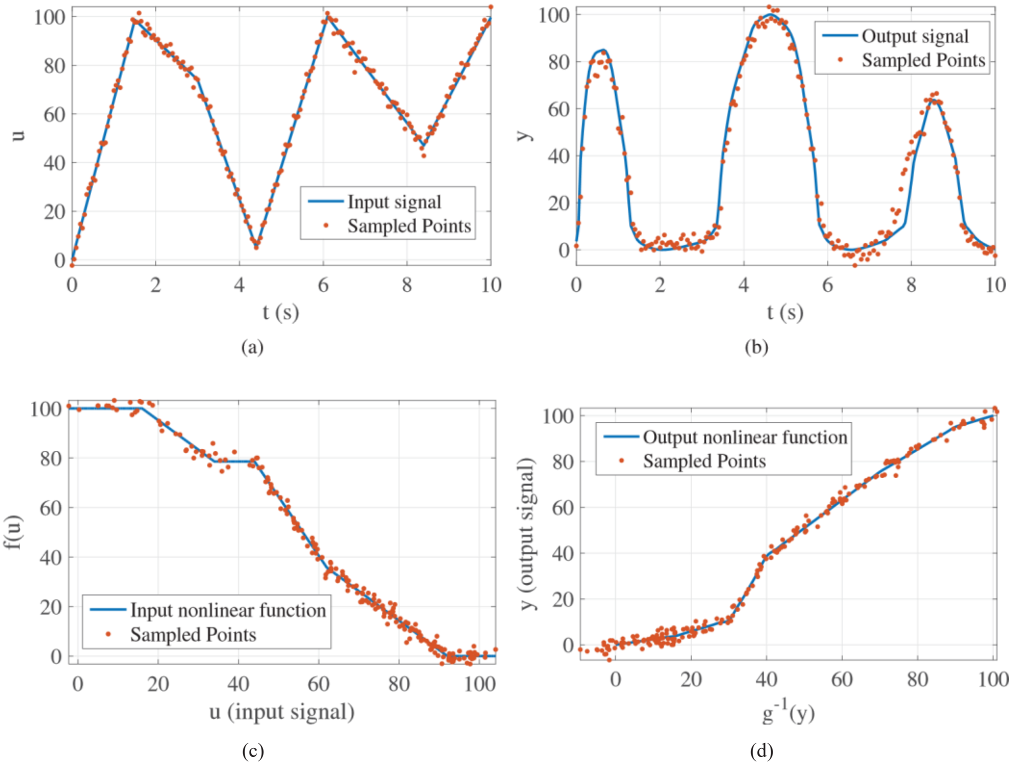
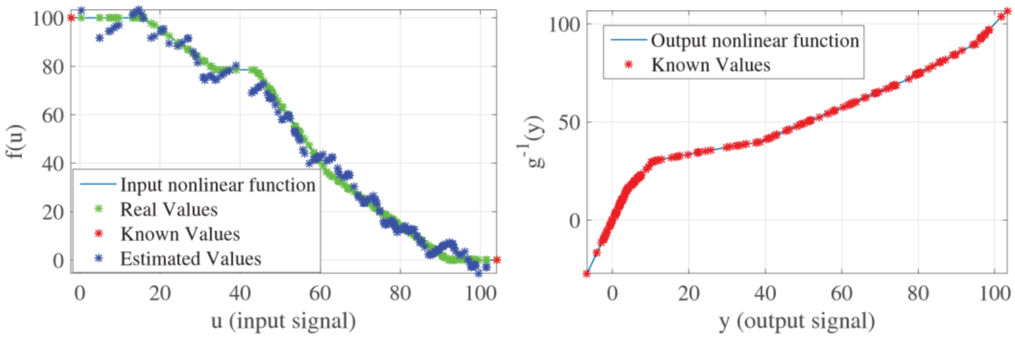


Fig. 7 An H–W system example with noise and disturbance. (a), (b) Input & Output Signals. (c), (d) nonlinear functions

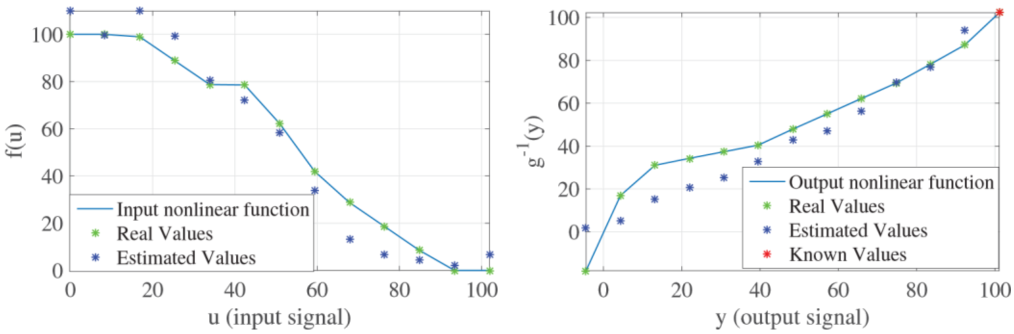
VI. Conclusion

In this paper, an algorithm to identify the input and output nonlinear functions of Hammerstein–Wiener systems is presented. Also, the effects of adding noise and disturbance to the system were evaluated. The mentioned problem is formulated as a least-square optimization problem. As seen in the presentation of the results, this problem could not be solved in the absence of sufficient repetition in measured data. So as a part of this method, it is illustrated how to add some constrained approximations and iterative algorithms to minimize errors. By formulating existing knowledge about nonlinear functions and techniques like quantization, the problem is solved and, the lack of repetition in measured data has been compensated. Finally, the proposed method is also validated by an example both in the absence and presence of noise and disturbance. Additionally, the quantization assumption is tested on this example to minimize the required known parts of nonlinear functions for solving the problem.



(a) (b)

Fig. 8 Estimating input and output functions in noisy H–W system by using constrained solutions to linear least–squares problem and putting limits on functions rates and values



(a) (b)

Fig. 9 Estimating input and output functions in noisy H–W system by using constrained solutions to linear least–squares problem, after quantization and assuming limits for functions rates and values

REFERENCES

[1] F. Giri and E.-W. Bai, "Block-oriented Nonlinear System Identification," London: Springer-Verlag, 2010.

[2] M. Hong and S. Cheng, "Hammerstein–Wiener Model Predictive Control of Continuous Stirred Tank Reactor," Electronics and Signal Processing, ed: Springer, 2011, pp. 235-242.

[3] T. Patikirikorala, L. Wang, A. Colman, and J. Han, "Hammerstein–Wiener nonlinear model based predictive control for relative QoS performance and resource management of software systems," Control Engineering Practice, vol. 20, pp. 49-61, 2012.

[4] J. C. Gómez, A. Jutan, and E. Baeyens. "Wiener model identification and predictive control of a pH neutralisation process," IEE Proceedings Control Theory and Applications 151(3), 329-338, 2004.

[5] A. Nemati and M. Faieghi, "The Performance Comparison of ANFIS and Hammerstein–Wiener Models for BLDC Motors," Electronics and Signal Processing, ed: Springer, 2011, pp. 29-37.

[6] K. Elleuch and A. Chaari, "Modeling and identification of hammerstein system by using triangular basis functions," International Journal of Electrical, Computer, Energetic, Electronic and Communication Engineering, vol. 1, p. 1, 2011.

[7] F. Lindsten, T. B. Schön, and M. I. Jordan, "Bayesian semiparametric Wiener system identification," Automatica, vol. 49, pp. 2053-2063, 2013.

[8] E.W. Bai, "An optimal two-stage identification algorithm for Hammerstein–Wiener nonlinear systems," Automatica, vol. 34, pp. 333-338, 1998.

[9] E.W. Bai, "A blind approach to the Hammerstein–Wiener model identification," Automatica, vol. 38, pp. 967-979, 2002.

[10] Y. Zhu, "Estimation of an N–L–N Hammerstein–Wiener model," Automatica, vol. 38, pp. 1607-1614, 9// 2002.

[11] P. Crama and J. Schoukens, "Hammerstein–Wiener system estimator initialization," Automatica, vol. 40, pp. 1543-1550, 2004.

[12] B. Ni, M. Gilson, and H. Garnier, "Refined instrumental variable method for Hammerstein–Wiener continuous-time model identification," IET Control Theory & Applications, vol. 7, pp. 1276-1286, 2013.

[13] A. Wills, T. B. Schön, L. Ljung, and B. Ninness, "Identification of Hammerstein–Wiener models," Automatica, vol. 49, pp. 70-81, 2013.

[14] G. Golub and W. Kahan, "Calculating the singular values and pseudo-inverse of a matrix," Journal of the Society for Industrial and Applied Mathematics, Series B: Numerical Analysis, vol. 2, pp. 205-224, 1965.

[15] R. Bouldin, "The pseudo-inverse of a product," SIAM Journal on Applied Mathematics, vol. 24, pp. 489-495, 1973.

[16] P. E. Gill, W. Murray, and M. H. Wright, "Practical optimization," New York: Academic Press, 1981.

[17] T. F. Coleman and Y. Li, "A reflective Newton method for minimizing a quadratic function subject to bounds on some of the variables," SIAM Journal on Optimization, vol. 6, pp. 1040-1058, 1996.