Received: 21 February 2020 | Revised: 25 August 2020 | Accepted: 6 October 2020

DOI: 10.1049/cth2.12052

IET Control Theory & Applications

The Institution of Engineering and Technology WILEY

ORIGINAL RESEARCH PAPER

Hammerstein–Wiener identification of industrial plants: A pressure control valve case study

Leili Esmaeilani[1] | **Jafar Ghaisari**[1] | **Mohammad Ali Bagherzadeh**[2]

[1] Department of Electrical and Computer Engineering, Isfahan University of Technology, Isfahan, Iran

[2] Department of Electrical and Computer Engineering, Graduate University of Advanced Technology, Kerman, Iran

Correspondence
Mohammad Ali Bagherzadeh, Department of Electrical and Computer Engineering, Graduate University of Advanced Technology, Kerman 7631885356, Iran.
Email: ma.bagherzadeh@kgut.ac.ir

Abstract

Hammerstein–Wiener (H–W) identification problem is investigated under practical considerations for industrial plants. The main consideration in these cases is inability to apply arbitrary input signals. An algorithm is proposed to identify H–W model for industrial systems, in the presence of noise and disturbance. The identification problem is changed into a constrained optimisation problem and is solved by minimising the non-linear functions modelling error. To improve the efficiency of the proposed method, the known information about non-linear functions are used. To make the proposed method applicable for general systems with hard non-linearities, non-linear functions are described by some ordered pairs. A pressure control valve that is used for anti-surge protection of an air compressor in the second refinery of South Pars Gas Complex is considered as the case study to demonstrate the effectiveness of the proposed algorithm. Seventy-five percent of the gathered data during the time is used for identification and the remaining 25% is used for verification of the resulted model. Verification of the obtained H–W model showed that accuracy of the identified model is more than 95%.

1 | INTRODUCTION

Accurate modelling and identification of industrial processes is essential for optimising the performance control. System identification of such cases could not be done just by conventional linear identification methods and neglecting the non-linear effects. Non-linearities are inevitable in most practical engineering systems which mostly arise from dead-zone and/or saturation of sensors or actuators. For example, in a Programmable Logic Controller (PLC), analog Input/Output (I/O) modules limit the passing signals between their minimum and maximum values. Control valve actuator is another industrial example with non-linearities including saturation, dead-zone, or hysteresis.

A famous category of non-linear mathematical models is the class of block oriented models, which are constructed by interconnections of linear dynamic models with static (or memory-less) non-linearities. 'Hammerstein' and 'Wiener' models are two well known subsets of this category that are formed by series interconnection of a linear dynamic block and a static non-linear one [1]. By serialising a Hammerstein and a Wiener model, the new subset 'Hammerstein–Wiener (H–W)' is emerged, which is the focus of this study. H–W modelling is appropriate for describing the behaviour of many practical applications such as continuous stirred tank reactors [2], Quality of Service (QoS) performance and resource management of software systems [3], PH neutralisation processes [4], Brushless DC (BLDC) motors [5], wideband RF transmitters [6], and lead-zinc flotation plant [7]. Moreover, sensors and actuators with saturation, dead-zone, and quantisation are modelled by static non-linearities. So each linear plant combined with these non-linear sensors and actuators can be represented as an H–W model [8]. Control valve is also an industrial equipment which can be described by block oriented models [9, 10].

So far, various methods have been used in previous research articles to identify Hammerstein, Wiener, or H–W models. Examples of wide ranging methods for the identification of Hammerstein and Wiener systems are [11–14]. Regarding the identification of H–W models, the two stage identification algorithm of [15] is one of the earliest methods. The stages included using the recursive least sum of squares method and calculation of singular value decomposition of two matrices whose dimensions are fixed and do not depend on the number of data points. The non-linear blocks were assumed to be approximated by a series of orthogonal functions. Four years later, a blind approach was presented in [16] to solve the problem, considering the non-linear blocks as a polynomial. Some researchers also proposed iterative methods for identification of H–W models. For example in [17, 18], non-linear functions are approximated by cubic splines and also the least-squares-based iterative technique for special case with two-segment polynomial input block and backlash output block characteristics. Furthermore, there are frequency based methods which consider a combination of sinusoids with a random phase as the system input [19], while some methods use pseudorandom binary sequence (PRBS) signals [20]. Apart from these methods, the identification problem is also solved by using refined instrumental variable method [21], particle swarm optimisation [22], the data filtering-based recursive generalised extended least squares algorithm [23] and maximum likelihood method [24].

The objective of this paper is to propose a novel method to identify H–W model for real word industrial plants. First, by considering the observable canonical state space realisation of the linear block, the relationship between the sampled data and unknown variables are presented in the form of a non-linear system of equations. Then, the problem of estimating the parameters of H–W model is reformulated as an optimisation problem. To investigate the existence and uniqueness of the solution, number of equations and unknown variables are discussed about. Moreover, known information about non-linear functions is taken into account to obtain a unique solution. Therefore, some equality and inequality constraints are derived based on the known behaviours of non-linear functions by formulating the non-linear functions and non-equalities related to used parameters or hyper-parameters. In this regard, different formulating methods, such as parametric and quantised methods and non-parametric methods (like Gaussian Process (GP)) were discussed. The constrained optimisation problem should be solved in the next step to obtain a unique solution for the identification problem. For this purpose, an iterative Newton approach is employed to solve the optimisation problem by minimising a constrained quadratic function that is conducted at each iteration of the algorithm. Furthermore, the influence of noise and disturbances on the accuracy of the identified H–W model is discussed. The proposed identification method is applied on a case study to verify the effectiveness of the results.

A pressure control valve (PCV) that is used for anti-surge protection of an air compressor in the second refinery of South Pars Gas Complex is considered as the case study. The compressor compresses the dehydrated atmospheric pressure air to be used as utility and instrument air or source of nitrogen generation package in refinery. The anti-surge valve is an important element of the system because it protects the compressor from working in the surge area. The actuating element of the PCV applies the control command received from the analog output module of control system. The actual position of valve is also sent to the control system through the analog input module. An appropriate identification method should be used to model the relation of these signals and verify if the anti-surge valve is properly protecting the compressor from working in the surge area [25]. A set of input-output data, measured from an industrial PCV, is investigated to model its closing behaviour and approve the proposed identification methods.

The main contributions of this paper are summarised as follows.

- The identification method is proposed for the general case with all types of noise and disturbances applied to linear and non-linear blocks.
- In real-world applications, the known information about the non-linear functions is used to improve the accuracy of the identified model.
- Invertibility of non-linear functions is relaxed and any memoryless function, even hard non-linearity, can be considered for non-linear blocks.
- The proposed method can be generalised to multi-variable plants.
- The gathered input-output signals may belong to different time intervals with various initial conditions. Therefore, the proposed method can be used in applications with different modes.

This paper was organised as follows. Section 2 illustrates the structure of the H–W model and the procedure of formulating the identification problem. In this section, preliminaries and assumptions for identifiability and uniqueness of the result are presented. Moreover, H–W identification problem is first formulated for SISO systems, which is also generalised to MIMO systems. In Section 3, a method is proposed to identify the H–W model for real-world applications. The proposed method is an iterative algorithm which utilises the Newton method in combination with Karush-Kuhn-Tucker (KKT) conditions for solving the emerged optimisation problem from the H–W model equations. In Section 4, a PCV is considered as the case study. Valve actuation system is explained and possible sources of non-linearities in the H–W model are explained. Moreover, required steps of the proposed method for identifying the H–W model for PCV are discussed and the identification results are presented. Finally, the paper is concluded in Section 5.

Notations. Superscript ^{T} stands for matrix transposition. For each matrix A with independent columns, superscript † denotes the left pseudo-inverse of matrix A; so that $A^\dagger$ is equal to $(A^TA)^{-1}A^T$. Kronecker product of matrices A and B is shown by $A \otimes B$. Moreover, $vec(A)$ indicates the vectorisation function of matrix A, which converts matrix $A_{m\times n}$ to a column vector x by stacking the columns of matrix A below one another. In contrast, for vector x with mn entries, $[x]_m$ shows a matrix with m rows, which satisfies $vec([x]_m) = x$. An $m \times n$ Matrix where every entry is equal to one is also called "all-ones matrix" and is shown by $\mathbb{1}_{m\times n}$; while $\mathbb{0}_{m\times n}$ is used to represent an $m \times n$ null matrix (i.e. A matrix, whose all entries are zero). I_n also denotes the "Identity Matrix" of size n. For the MIMO function f, notation $\nabla_x f(x)$ denotes the gradient of function f with respect to its input, which is the extension of partial derivation in SISO functions f, showing by $\frac{\partial f}{\partial u}$.

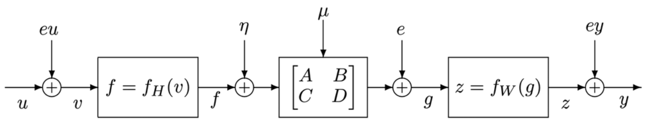

FIGURE 1 Block diagram of H–W model

Definition ([26]). "Extended Identity Matrix" $\mathcal{I}^{(m,n)}(x, y, z)$ is an $m \times n$ sparse matrix with z non-zero elements (equal to 1), where $0 \leq x \leq m$, $0 \leq y \leq n$, and $0 \leq z \leq \min\{(m-x), (n-y)\}$. The structure of $\mathcal{I}$ which simplifies the formulation of identification problem is shown in (1).

$$\mathcal{I}^{(m,n)}(x, y, z) = \begin{bmatrix} \mathbb{0}_{x\times y} & \mathbb{0}_{x\times z} & \mathbb{0}_{x\times n-z-y} \\ \mathbb{0}_{z\times y} & I_z & \mathbb{0}_{z\times n-z-y} \\ \mathbb{0}_{(m-z-x)\times y} & \mathbb{0}_{(m-z-x)\times z} & \mathbb{0}_{(m-z-x)\times(n-z-y)} \end{bmatrix}. \quad (1)$$

2 | PROBLEM FORMULATION

In this section, a method is proposed to formulate the H–W identification problem for the system shown in Figure 1. The model shown in Figure 1 is the most general form of H–W model in which all types of noise and disturbances are considered [26]. As shown in Figure 1, u, y, and x represent the input, output and state vectors respectively. η and e represent the noise and/or disturbance signals added to the intermediate signals f and g. Moreover, μ indicates the external disturbance applied to the linear dynamic block. The input (u) and output (y) signals can be measured through the time, while the intermediate signals f and g can not be measured. The shown signals in Figure 1 during the time $\{1, 2, \ldots, N\}$ are the vectors given in (2) and assumed to be independent and limited.

$$u := \begin{bmatrix} u_1 \\ \vdots \\ u_N \end{bmatrix}, \quad y := \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}, \quad x := \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix}, \quad (2a)$$

$$f := \begin{bmatrix} f_1 \\ \vdots \\ f_N \end{bmatrix}, \quad g := \begin{bmatrix} g_1 \\ \vdots \\ g_N \end{bmatrix}.$$

$$eu := \begin{bmatrix} eu_1 \\ \vdots \\ eu_N \end{bmatrix}, \quad ey := \begin{bmatrix} ey_1 \\ \vdots \\ ey_N \end{bmatrix}, \quad \mu := \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_N \end{bmatrix}, \quad (2b)$$

$$\eta := \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_N \end{bmatrix}, \quad e := \begin{bmatrix} e_1 \\ \vdots \\ e_N \end{bmatrix}.$$

At sampling time t, the state-space equations of the linear dynamic block in Figure 1 are expressed as (3).

$$\begin{bmatrix} x_{t+1} \\ g^{-1}\left(y_t - ey_t\right) \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} x_t \\ f(u_t + eu_t) + \eta_t \end{bmatrix} + \begin{bmatrix} \mu_t \\ e_t \end{bmatrix}. \quad (3)$$

It should be noted that the considered system may be either SISO or MIMO. At sampling time t, dimension of x_t is n and dimensions of f_t and g_t are respectively n_f and n_g, which are known constants. If the investigated system is SISO, n_f and n_g are equal to 1 (i.e. f and g are $N \times 1$ vectors.).

For a real-world application, lots of equivalent H–W models with identical input-output behaviour exist. For instance, if the input non-linear function is multiplied to a constant number and the state space matrices B and D of linear dynamic block are divided to the same number, the models are equivalent. Hence, identifiability of real-world applications to distinguish static non-linearities f and g from the linear function is an important issue in the identification of H–W models. The identifiability conditions depend on both input-output signals (u and y) and characteristics of linear dynamic block and non-linear static blocks. Therefore, every identification method for H–W model should contain some uniqueness techniques and identifiability assumptions. In some research articles such as [16], minimality assumption is considered to avoid pole–zero cancellations which makes the parametrisation non-unique and guarantee the identifiability.

In this work, to avoid equivalent models with different parameters, observable canonical state space realisation is considered for the linear block (A, B, C, D) as well as the minimality assumption. Therefore, as it is shown in (4), the linear block could be uniquely identified by parameter vectors a, b and d.

$$A = \begin{bmatrix} 0 & 0 & \ldots & 0 & -a_n \\ 1 & 0 & \ldots & 0 & -a_{n-1} \\ 0 & 1 & \ldots & 0 & -a_{n-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & -a_1 \end{bmatrix} \rightarrow a = \begin{bmatrix} a_n \\ a_{n-1} \\ a_{n-2} \\ \vdots \\ a_1 \end{bmatrix}, \quad (4a)$$

$$b = vec(B), \quad (4b)$$

$$d = vec(D). \quad (4c)$$

The relation between matrix A and vector a of (4a) is given in (5), considering the definition of extended identity matrix.

$$A = \mathcal{I}^{(n,n)}(1, 0, n-1) - \mathcal{I}^{(1,n)}(0, n-1, 1) \otimes a. \quad (5)$$

Matrix C can also be rewritten as

$$C = \begin{bmatrix} 0 & \ldots & 0 & 1 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \ldots & 0 & 1 \end{bmatrix} = \mathcal{I}^{(1,n)}(0, n-1, 1) \otimes \mathbb{1}_{n_g \times 1}. \quad (6)$$

Remark 1. For SISO systems, B is a single-column vector and D is a scalar. Therefore, $b = B$ and $d = D$. Moreover, $C = [0 \ \dots \ 0 \ 1] = \mathcal{I}^{(1,n)}(0, n-1, 1)$.

In the formulation of this paper, some pre-known information about the practical aspects of the investigated plant should be available to guarantee the uniqueness of the result. The following assumptions are also considered for the identifiability of the proposed method.

Assumption 1. The input and output noise signals (eu and ey) are considered to be bounded with sufficiently small amplitude.

A non-linear function can be represented either by a parametric function, or by a set of ordered pairs of the function. Due to the ability of ordered pairs in representing complicated non-linear functions, even functions with hard non-linearities, this paper considers sets of ordered pairs for representation of non-linear functions f and g. It should be noted that using more ordered pairs, represents the function in more details.

Assumption 2. The values or range of variations of some ordered pairs of non-linear functions are assumed to be known.

Assumption 3. Some information about the range of non-linear function or the maximum and minimum slopes of the non-linear functions in the whole domain (or an interval) is assumed to be known.

It is worth noting that the information mentioned in Assumption 2 and Assumption 3 are used to avoid unacceptable solutions in the presence of noise and disturbance. The more information about non-linear functions is used, the more efficient model is identified. In fact, the system of equations describing the H–W identification is under-determined since number of equations is less than unknowns. Number of equations is equal to the number of samples (N), while number of unknowns in a SISO system is $2N + 2n + 1$. If the input signal (u) is rich in harmonics, the values of u will be repeated in many order pairs. Therefore, if the input noise signal (eu) is sufficiently small, non-linear function f corresponding to these values of u will be almost similar and the number of unknowns will be decreased. But in industrial plants that input signal u is not usually rich of harmonics, identifiability is much more complicated. A novel method is presented herein to solve the identification problem of H–W models for industrial plants.

In this method, an auxiliary vector u' is considered. u' is a known vector with the elements distributed in the domain of u. For instance, if the distribution is uniform, u' will be evenly spaced values within the domain of u. f' is also defined as the function output that is related to u' (i.e. $f' = f_H(u')$). Using the defined vectors f' and g' as unknowns, a system of non-linear equations is determined which will be solved by the method presented in the next section. When the auxiliary unknowns f' and g' are determined, unknown values f can be estimated using interpolation methods. Interpolation is a method of constructing new data points within the range of a discrete set of known data points [27]. In this work, ordered pairs of the non-linear functions are written and interpolated in the following form.

$$f = \Omega_f f' + \epsilon_f, \tag{7a}$$

$$g = \Omega_g g' + \epsilon_g. \tag{7b}$$

To consider the input range and also interpolation or modelling errors, ϵ_f and ϵ_g are added to the results. For instance, if the function $f(u)$ is represented by GP, f and f' are functions output related to u and u', respectively. Therefore, ϵ_f has Gaussian distribution (i.e. $\epsilon_f \sim N(0, K''_f - K'_f K_f^{-1} {K'_f}^T)$) and matrix $\Omega_f = K'_f K_f^{-1}$, where K_f, K'_f and K''_f represents kernel matrices $k(u', u')$, $k(u', u)$ and $k(u, u)$, respectively [28]. The flexibility of this method is that f' should not be only $f_H(u')$ and can also be considered as Gaussian Process (GP) non-parametric functions or even it can be coefficients of a limited order polynomial which models the function. Moreover, the input space of the functions may be quantised and the outputs are considered to be piecewise constant signals (i.e. outputs in each quantisation level is assumed to be equal).

Remark 2. Choosing the length of vector u', there is a trade off between number of solutions and accuracy of H–W identification. In fact, for smaller number of elements, identification results will be converged to a unique solution; but the accuracy will be degraded. So, length of vector u' will be selected large enough (fewer than N) to maintain the sufficient accuracy of the H–W model and additional information about the behaviour of the plant will be applied in order to guarantee the uniqueness of the solution as well.

The effect of the input and output measurement noise signals are also considered in the terms ϵ_f and ϵ_g, respectively. If the measurement noise has sufficiently small amplitude, first order Taylor approximation will be acceptable and distributions of ϵ_f and ϵ_g are multiples of the distribution of the measurement noise plus interpolation error.

As it is described in [26], the function derivatives could also be approximated as a constant matrix multiplied by f' or g'. For each input sample, the derivative can be approximated according to its distances to the nearest element of u' in each side as

$$\frac{\partial f}{\partial u}(u) \approx \frac{f(u + \triangle u_1) - f(u - \triangle u_2)}{\triangle u_1 + \triangle u_2}.$$

In the rest of this section, the problem formulation is divided in two parts. First, equations that describe the relationship between signals and parameters of the H–W model will be given for SISO systems. Then the problem formulation will be extended to MIMO systems.

2.1 | Problem formulation of SISO systems

If the vector of unknown variables is defined as $\alpha^{[1]} = \begin{bmatrix} x \\ f' \\ g' \\ \eta + \epsilon_f \end{bmatrix}$, the following formulation given in Theorem 1 will be obtained.

Theorem 1. *For the SISO H–W model shown in Figure 1, system of linear equations* $\Gamma^{[1]}\alpha^{[1]} - \beta^{[1]} = \vartheta$ *with*

$$\beta^{[1]} = \begin{bmatrix} m_\mu \\ m_e + m_{\epsilon_g} \\ m_\eta + m_{\epsilon_f} \end{bmatrix}, \tag{8a}$$

$$\vartheta = \begin{bmatrix} \mu - m_\mu \\ e - m_e + \epsilon_g - m_{\epsilon_g} \\ \eta - m_\eta + \epsilon_f - m_{\epsilon_f} \end{bmatrix}, \tag{8b}$$

$$\Gamma^{[1]} = \begin{bmatrix} \Gamma_A^{[1]} & \Gamma_B^{[1]}\Omega_f & \mathbb{0}_{n(N-1)\times N} & \Gamma_B^{[1]} \\ \Gamma_C^{[1]} & \Gamma_D^{[1]}\Omega_f & \Omega_g & \Gamma_D^{[1]} \\ \mathbb{0}_{N\times nN} & \mathbb{0}_{N\times N} & \mathbb{0}_{N\times N} & I_N \end{bmatrix}, \tag{8c}$$

describes the relationship between the signals and parameters, where m_η*,* m_e*,* m_μ*,* m_{ϵ_f}*, and* m_{ϵ_g} *correspondingly denote the mean values of Gaussian signals* η*,* e*,* μ*,* ϵ_f*, and* ϵ_g*, and block matrices* $\Gamma_A^{[1]}$*,* $\Gamma_B^{[1]}$*,* $\Gamma_C^{[1]}$*, and* $\Gamma_D^{[1]}$ *are expressed as*

$$\Gamma_A^{[1]} = \mathcal{I}^{(N-1,N)}(0,1,N-1) \otimes I_n - \mathcal{I}^{(N-1,N)}(0,0,N-1) \otimes A, \tag{9a}$$

$$\Gamma_B^{[1]} = -\mathcal{I}^{(N-1,N)}(0,0,N-1) \otimes b, \tag{9b}$$

$$\Gamma_C^{[1]} = -I_N \otimes \mathcal{I}^{(1,n)}(0,n-1,1), \tag{9c}$$

$$\Gamma_D^{[1]} = -d\, I_N. \tag{9d}$$

Proof. Considering vector definitions (2), one can rewrite Equation (3) at time t as

$$x_{t+1} - Ax_t - Bf_t - B\eta_t = \mu_t, \tag{10a}$$

$$g_t - Cx_t - Df_t - D\eta_t = e_t. \tag{10b}$$

Using the definition "extended identity matrix", it can be verified that considering the block matrices $\Gamma_A^{[1]}, \Gamma_B^{[1]}, \Gamma_C^{[1]}$, and $\Gamma_D^{[1]}$ in (9), Equations (10) will be rewritten as

$$\begin{bmatrix} \Gamma_A^{[1]} & \Gamma_B^{[1]} & \mathbb{0}_{(N-1)n\times N} & \Gamma_B^{[1]} \\ \Gamma_C^{[1]} & \Gamma_D^{[1]} & I_N & \Gamma_D^{[1]} \end{bmatrix} \begin{bmatrix} x \\ f \\ g \\ \eta \end{bmatrix} = \begin{bmatrix} \mu \\ e \end{bmatrix}. \tag{11}$$

By replacing f and g by the right hand side of Equations (7) and also adding $\eta + \epsilon_f$ to both sides of (11) to consider the noise and disturbance signals in ϑ, system of linear equations $\Gamma^{[1]}\alpha^{[1]} - \beta^{[1]} = \vartheta$ with the parameters given in (8) and (9) will be determined. □

Remark 3. In the system of linear equations given in Theorem 1, the block matrices (9) and hence the matrix $\Gamma^{[1]}$ is independent from $\alpha^{[1]}$. Moreover, the values of measured inputs and outputs are used in calculating the elements of Ω_f and Ω_g.

One may consider a different representations and define unknown variables as $\begin{bmatrix} a^T & b^T & d \end{bmatrix}^T$ to drive another formulation, given in Theorem 2. It is important to mention that both systems of equations in Theorem 1 and Theorem 2 are the same (since ϑ is identical in both theorems), but with different representations and unknown variables definition.

Theorem 2. *For the SISO H–W model shown in Figure 1, system of linear equations* $\Gamma^{[2]}\alpha^{[2]} - \beta^{[2]} = \vartheta$ *describes the relationship between the H–W model signals and parameters, considering* ϑ *defined by (8b) and* $\alpha^{[2]} = \begin{bmatrix} a^T & b^T & d \end{bmatrix}^T$*, where* $\beta^{[2]}$ *is defined as (13) and* $\Gamma^{[2]}$ *is*

$$\Gamma^{[2]} = \begin{bmatrix} \Gamma_a^{[2]} & \Gamma_b^{[2]} & \mathbb{0}_{n(N-1)\times 1} \\ \mathbb{0}_{N\times n} & \mathbb{0}_{N\times n} & -\left[\Omega_f f' + \epsilon_f + \eta\right] \\ \mathbb{0}_{N\times n} & \mathbb{0}_{N\times n} & \mathbb{0}_{N\times 1} \end{bmatrix}, \tag{12a}$$

where

$$\Gamma_a^{[2]} = \left(\mathcal{I}^{(N-1,N)}(0,0,N-1) \otimes \mathcal{I}^{(1,n)}(0,n-1,1)\right)x \otimes I_n, \tag{12b}$$

$$\Gamma_b^{[2]} = -\left(\mathcal{I}^{(N-1,N)}(0,0,N-1)\left(\Omega_f f' + \epsilon_f + \eta\right)\right) \otimes I_n. \tag{12c}$$

$$\beta^{[2]} = \begin{bmatrix} \left(\mathcal{I}^{(N-1,N)}(0,1,N-1) \otimes I_n - \mathcal{I}^{(N-1,N)}(0,0,N-1) \otimes \mathcal{I}^{(n,n)}(1,0,n-1)\right)x + m_\mu \\ -\Omega_g g' + \left(I_N \otimes \mathcal{I}^{(1,n)}(0,n-1,1)\right)x + m_e + m_{\epsilon_g} \\ m_\eta + m_{\epsilon_f} - \eta - \epsilon_f \end{bmatrix} \tag{13}$$

Proof. Let us rewrite the first equation of Theorem 1 as $\Gamma_A^{[1]} x + \Gamma_B^{[1]} \Omega_f f' + \Gamma_B^{[1]} (\eta + \epsilon_f) - \mu = 0$. This equation can be rewritten as

$$\begin{aligned}\Gamma_A^{[1]} x + \Gamma_B^{[1]} \Omega_f f' + \Gamma_B^{[1]} (\eta + \epsilon_f) - \mu \\ &= \Gamma_A^{[1]} x + \Gamma_B^{[1]} (\Omega_f f' + \epsilon_f + \eta) - \mu \\ &= \Gamma_A^{[1]} x + \Gamma_B^{[1]} (f + \eta) - \mu = 0.\end{aligned}$$

Substituting (5) in (9a), the following equation will be resulted for $\Gamma_A^{[1]} x$.

$$\begin{aligned}\Gamma_A^{[1]} x = \Big[&\mathcal{I}^{(N-1,N)}(0,1,N-1) \otimes I_n \\ &- \mathcal{I}^{(N-1,N)}(0,0,N-1) \otimes \mathcal{I}^{(n,n)}(1,0,n-1) \Big] x \\ &+ \left(\mathcal{I}^{(N-1,N)}(0,0,N-1) \otimes \mathcal{I}^{(1,n)}(0,n-1,1) \otimes a \right) x.\end{aligned}$$

The right hand side of this equation contains two parts. The first part is not related to $\alpha^{[2]} = \begin{bmatrix} a^T & b^T & d \end{bmatrix}^T$ and can be considered in $\beta^{[2]}$; while vector a is seen in the second part. Using the properties of Kronecker product and vectorisation operator ($vec(.)$), it can be verified that for any three matrices P, Q, and R with compatible dimensions that matrix multiplication PQR is defined, $vec(PQR) = (R^T \otimes P) vec(Q)$ is satisfied [29]. Therefore, defining matrices $P := a$, $Q := x^T$, and $R := (\mathcal{I}^{(N-1,N)}(0,0,N-1) \otimes \mathcal{I}^{(1,n)}(n-1,0,1))^T$, the following equations are satisfied.

$$\begin{aligned}&\left(\mathcal{I}^{(N-1,N)}(0,0,N-1) \otimes \mathcal{I}^{(1,n)}(n-1,0,1) \otimes a \right) x \\ &\quad = vec\left(a\, x^T \left(\mathcal{I}^{(N,N-1)}(0,0,N-1) \otimes \mathcal{I}^{(n,1)}(0,n-1,1) \right) \right) \\ &\quad = \left(\left(\mathcal{I}^{(N-1,N)}(0,0,N-1) \otimes \mathcal{I}^{(1,n)}(0,n-1,1) \right) x \otimes I_n \right) a.\end{aligned}$$

Moreover, for $\Gamma_B^{[1]} (f + \eta)$, the following equations hold.

$$\begin{aligned}\Gamma_B^{[1]} (f + \eta) &= -\left(\mathcal{I}^{(N-1,N)}(0,0,N-1) \otimes b \right) (f + \eta) \\ &= -\, vec\left(b\, (f + \eta)^T \mathcal{I}^{(N,N-1)}(0,0,N-1) \right) \\ &= -\left(\mathcal{I}^{(N-1,N)}(0,0,N-1)(f + \eta) \otimes I_n \right) b.\end{aligned}$$

Similarly, other coefficients from the second equation of Theorem 1 can be calculated as $\Gamma_D^{[1]} (f + \eta) = -d\ (f + \eta)$. Other remaining parts of the equations could be seen in $\beta^{[2]}$. □

2.2 | Problem formulation of MIMO systems

One of the main contributions of this paper is the ability of the proposed method to formulate multi-variable plants in which number of signals passing between linear and non-linear blocks are not necessarily equal. Here, the problem formulation is adapted for MIMO systems.

Proposition 1. *System of linear equations $\Gamma^{[1]} \alpha^{[1]} - \beta^{[1]} = \vartheta$ describes the relationship between signals and parameters of the H–W model shown in Figure 1, with*

$$\alpha^{[1]} = \begin{bmatrix} x \\ f' \\ g' \\ \eta + \epsilon_f \end{bmatrix}, \quad \beta^{[1]} = \begin{bmatrix} m_\mu \\ m_e + m_{\epsilon_g} \\ m_\eta + m_{\epsilon_f} \end{bmatrix}, \tag{15a}$$

$$\vartheta = \begin{bmatrix} \mu - m_\mu \\ e - m_e + \epsilon_g - m_{\epsilon_g} \\ \eta - m_\eta + \epsilon_f - m_{\epsilon_f} \end{bmatrix}, \tag{15b}$$

$$\Gamma^{[1]} = \begin{bmatrix} \Gamma_A^{[1]} & \Gamma_B^{[1]} \Omega_f & \mathbb{0}_{n(N-1) \times n_g N} & \Gamma_B^{[1]} \\ \Gamma_C^{[1]} & \Gamma_D^{[1]} \Omega_f & \Omega_g & \Gamma_D^{[1]} \\ \mathbb{0}_{n_f N \times nN} & \mathbb{0}_{n_f N \times n_{f'} N} & \mathbb{0}_{n_f N \times n_{g'} N} & I_{n_f N} \end{bmatrix}, \tag{15c}$$

where m_η, m_e, m_μ, m_{ϵ_f}, and m_{ϵ_g} correspondingly denote the mean values of Gaussian signals η, e, μ, ϵ_f, and ϵ_g, and block matrices $\Gamma_A^{[1]}$, $\Gamma_B^{[1]}$, $\Gamma_C^{[1]}$, and $\Gamma_D^{[1]}$ are expressed as

$$\Gamma_A^{[1]} = \mathcal{I}^{(N-1,N)}(0,1,N-1) \otimes I_n \\ - \mathcal{I}^{(N-1,N)}(0,0,N-1) \otimes A, \tag{16a}$$

$$\Gamma_B^{[1]} = -\mathcal{I}^{(N-1,N)}(0,0,N-1) \otimes B, \tag{16b}$$

$$\Gamma_C^{[1]} = -I_N \otimes C = -I_N \otimes \mathcal{I}^{(1,n)}(0,n-1,1) \otimes \mathbb{1}_{n_g \times 1}, \tag{16c}$$

$$\Gamma_D^{[1]} = -I_N \otimes D. \tag{16d}$$

Proof. Similar to the proof of Theorem 1 and using and properties of extended identity matrix given in Appendix 1 of [26]. □

Moreover, if one considers unknown variables as $\begin{bmatrix} a^T & b^T & d^T \end{bmatrix}^T$, the formulation given in Proposition 2 will be obtained.

Proposition 2. *Considering $\alpha^{[2]} = \begin{bmatrix} a^T & b^T & d^T \end{bmatrix}^T$ and ϑ defined by (15b), system of linear equations $\Gamma^{[2]} \alpha^{[2]} - \beta^{[2]} = \vartheta$ describes the relationship between the H–W model signals and parameters shown in*

Figure 1, where $\Gamma^{[2]}$, an independent matrix from $\alpha^{[2]}$, is defined as

$$\Gamma^{[2]} = \begin{bmatrix} \Gamma_a^{[2]} & \Gamma_b^{[2]} & \mathbb{0}_{n(N-1)\times n_g n_f} \\ \mathbb{0}_{n_g N\times n} & \mathbb{0}_{n_g N\times n_f n} & \Gamma_d^{[2]} \\ \mathbb{0}_{n_f N\times n} & \mathbb{0}_{n_f N\times n_f n} & \mathbb{0}_{n_f N\times n_f n_g} \end{bmatrix}, \tag{17a}$$

where

$$\Gamma_a^{[2]} = \left(\mathcal{I}^{(N-1,N)}(0,0,N-1) \otimes \mathcal{I}^{(1,n)}(0,n-1,1)\right)x \otimes I_n, \tag{17b}$$

$$\Gamma_b^{[2]} = -\left(\mathcal{I}^{(N-1,N)}(0,0,N-1)\left[\Omega_f f' + \epsilon_f + \eta\right]_{n_f}^T\right) \otimes I_n, \tag{17c}$$

$$\Gamma_d^{[2]} = -\left[\Omega_f f' + \epsilon_f + \eta\right]_{n_f}^T \otimes I_{n_g}. \tag{17d}$$

and $\beta^{[2]}$ is defined as (18).

$$\beta^{[2]} = \begin{bmatrix} \left(\mathcal{I}^{(N-1,N)}(0,0,N-1) \otimes I_n - \mathcal{I}^{(N-1,T)}(0,1,N-1)\right. \\ \left.\otimes \mathcal{I}^{(n,n)}(1,0,n-1)\right) x + m_\mu \\ -\Omega_g g' + \left(I_N \otimes \mathcal{I}^{(1,n)}(0,n-1,1) \otimes \mathbb{1}_{n_g\times 1}\right) x + m_e + m_{\epsilon_g} \\ m_\eta + m_{\epsilon_f} - \eta - \epsilon_f \end{bmatrix} \tag{18}$$

Proof. Similar to the proof of Theorem 2. □

2.3 | Pre-known information about non-linear functions

As it was previously discussed, lots of equivalent H–W models exist for a real-world application. Herein, the pre-known information about the non-linear functions are used to limit the solution of H–W identification problem to the desired one. The following shows how this information can be used in the formulation of non-linear functions.

One of the pre-known information about non-linear functions f and g may be their output ranges in all or some of their points. These points may be not involved in the measured set but they also could be formulated by the linear transformation of function parameters (f' and g') similar to equations (7). Therefore, these types of pre-known information may be written in the form of inequality constraints (19).

$$\Gamma_{\mathrm{ineq}f}^{[1]} f' - \beta_{\mathrm{ineq}f}^{[1]} \le 0 \tag{19a}$$

$$\Gamma_{\mathrm{ineq}g}^{[1]} g' - \beta_{\mathrm{ineq}g}^{[1]} \le 0 \tag{19b}$$

Limitation in the slope of non-linear functions, that may be different in various ranges, may be another pre-known information about non-linear functions. For SISO systems, the key note in formulation of these information is to approximate the slope in each point with

$$\frac{\partial f}{\partial u}(u) \approx \frac{f(u+\triangle u) - f(u-\triangle u)}{2\triangle u}.$$

It should be noted that Δu in the above approximation refers to the distance between elements of the auxiliary vector u'. Obviously, the smaller Δu is (corresponded to larger number of elements of u'), the more precise approximation of df/du is resulted. However, the uniqueness of the result may be lost if there is not enough information about the behaviour of the plant.

By formulating the values of functions in some points with parameters and approximating the slope, another set of constraints in the form of (20) will be obtained which assures the non-linear functions perform the desired behaviour.

$$\Gamma_{\mathrm{ineq}f}^{[2]} f' - \beta_{\mathrm{ineq}f}^{[2]} \le 0 \tag{20a}$$

$$\Gamma_{\mathrm{ineq}g}^{[2]} g' - \beta_{\mathrm{ineq}g}^{[2]} \le 0 \tag{20b}$$

It should be noted that matrices $\Gamma_{\mathrm{ineq}f}^{[.]}$ and $\Gamma_{\mathrm{ineq}g}^{[.]}$ and vectors $\beta_{\mathrm{ineq}f}^{[.]}$ and $\beta_{\mathrm{ineq}g}^{[.]}$ in constraints (19) and (20) do not depend on f' and g'. Moreover, the acceptable range of set of parameters in f' and g' should satisfy the above linear inequalities. It should also be noticed that the terms ϵ_f and ϵ_g in (7) are inserted to consider the effect of measurement noise (*eu* and *ey*) on non-linear functions f and g. The range and distribution of these terms are related to both measurement noise signals distribution and also the behaviour of non-linear functions. The relationship may be approximated as (21), using the first order approximation of Taylor expansion, where the matrices $\triangle_f(f')$ and $\triangle_g(g')$ contain the approximations of function slope in all points.

$$\epsilon_f \approx \triangle_f(f')e_f \approx \Gamma_f^*(e_u)f' \tag{21a}$$

$$\epsilon_g \approx \triangle_g(g')e_g \approx \Gamma_g^*(e_y)g' \tag{21b}$$

To ensure the noise and disturbance signal, absolute levels may be limited according to assumptions and by considering (21), following equations could be written.

$$\triangle_\vartheta^{[1]} \vartheta - \beta_{\mathrm{ineq}\vartheta}^{[3]} \le 0 \tag{22a}$$

$$\triangle_\vartheta^{[2]} \vartheta + \Gamma_{\mathrm{ineq}f}^{[3]} f' - \beta_{\mathrm{ineq}f}^{[3]} \le 0 \tag{22b}$$

$$\triangle_\vartheta^{[3]} \vartheta + \Gamma_{\mathrm{ineq}g}^{[3]} g' - \beta_{\mathrm{ineq}g}^{[3]} \le 0 \tag{22c}$$

One can combine the constraints presented in (19), (20) and (22) and rewrite them in the form of $\triangle_{\vartheta}\vartheta + \triangle_{f} f' + \triangle_{g} g' - \beta_{\text{ineq}} \leq 0$.

3 | PROPOSED IDENTIFICATION METHOD

In this section, a constrained optimisation method is proposed for solving the identification problem. The goal of optimisation methods may be minimising or maximising a real function, and the optimisation problem may be solved subject to inequality or equality constrains. One of the subcategories of optimisation problem is finding the least square of residuals. The residuals may be the difference between an observed value, and the fitted value provided by a model and may be caused by entered noise to the model. Least-squares approximate the solution of overdetermined equations and is a standard approach in regression analysis and identification. If the optimum solution of optimisation problem is an optimum solution within a neighbouring set of candidate solutions, it will be called a local optimum. While if this is the optimal solution among all possible solutions, not just those in a particular neighbourhood of values, it will be called the global optimum.

In this section, we first propose a lemma which will be used to solve the H–W identification problem. Then an optimisation procedure for the identification problem will be offered in a theorem.

Lemma 1. *Consider the vectors* $\alpha = \begin{bmatrix} \alpha_n^{[1]^T} & \alpha_n^{[2]^T} \end{bmatrix}^T$, $\beta = \begin{bmatrix} \beta_n^{[1]^T} & \beta_n^{[2]^T} \end{bmatrix}^T$ *and* $\vartheta(\alpha) = \Gamma^{[1]}\alpha^{[1]} - \beta^{[1]} = \Gamma^{[2]}\alpha^{[2]} - \beta^{[2]}$, *where elements of* $\Gamma^{[1]}$ *and* $\beta^{[1]}$ *are functions of* $\alpha^{[2]}$ *and elements of* $\Gamma^{[2]}$ *and* $\beta^{[2]}$ *are functions of* $\alpha^{[1]}$. *The obtained* α^* *by the following iterative equation will be a local minimum of the scaler function* $\vartheta(\alpha)^T\vartheta(\alpha)$ *if* $\begin{bmatrix} \Gamma^{[1]} & \Gamma^{[2]} \end{bmatrix}$ *is a column independent matrix.*

$$\Gamma_n^{[1]^T}\left(I_{n_\vartheta} - \Gamma_n^{[2]}\Gamma_n^{[2]^\dagger}\right)\left(\Gamma_n^{[1]}\alpha_{n+1}^{[1]} - \beta_n^{[1]}\right) = \mathbb{0}_{n_{\alpha^{[1]}}\times 1}, \tag{23a}$$

$$\Gamma_n^{[2]^T}\left(I_{n_\vartheta} - \Gamma_n^{[1]}\Gamma_n^{[1]^\dagger}\right)\left(\Gamma_n^{[2]}\alpha_{n+1}^{[2]} - \beta_n^{[2]}\right) = \mathbb{0}_{n_{\alpha^{[2]}}\times 1}. \tag{23b}$$

Proof. First, it needs to be reminded that $\vartheta^T(\alpha)\vartheta(\alpha)$ is a scaler function and the size of vectors and matrices are

$$n_{\vartheta(\alpha)} = n_{\beta^{[1]}} = n_{\beta^{[2]}} = n_\beta,$$

$$n_\alpha \leq n_\beta,$$

$$\Gamma^{[1]} \in \mathbb{R}^{n_\beta \times n_{\alpha^{[1]}}},$$

$$\Gamma^{[2]} \in \mathbb{R}^{n_\beta \times n_{\alpha^{[2]}}}.$$

To apply the multivariate Gauss-Newton method, Jacobian of $\vartheta(\alpha)$ should be first calculated as follows.

$$J_\vartheta(\alpha) = \left[J_\vartheta(\alpha^{[1]})\ J_\vartheta(\alpha^{[2]})\right] = \left[\Gamma^{[1]}\ \Gamma^{[2]}\right]. \tag{24}$$

Then the local minimum of $\vartheta^T(\alpha)\vartheta(\alpha)$ could be calculated by the iterative equation

$$\alpha_{n+1} = \alpha_n - \left(J_\vartheta(\alpha_n)^T J_\vartheta(\alpha_n)\right)^{-1} J_\vartheta(\alpha_n)^T \vartheta(\alpha_n), \tag{25}$$

where $\vartheta(\alpha_n)$ is the left pseudo-inverse of the Jacobian matrix. The left pseudo-inverse of a matrix can be defined if it is full column rank, that is the matrix columns are linearly independent. In this case, $J_\vartheta(\alpha_n)^T J_\vartheta(\alpha_n)$ is a full rank matrix and its inverse is existed. The dimension of the Jacobian matrix is equal to $n_\beta \times n_\alpha$, so $n_\alpha \leq n_\beta$ and $\text{rank}(J_\vartheta(\alpha_n)) = n_\alpha$ are the necessary conditions to have the left pseudo-inverse. To avoid calculating the matrix inversion $(J_\vartheta(\alpha_n)^T J_\vartheta(\alpha_n))^{-1}$ in (25), both sides of the equation are multiplied by $(J_\vartheta(\alpha_n)^T J_\vartheta(\alpha_n))$ to obtain the following equation.

$$J_\vartheta(\alpha_n)^T J_\vartheta(\alpha_n)\alpha_{n+1} = J_\vartheta(\alpha_n)^T J_\vartheta(\alpha_n)\alpha_n - J_\vartheta(\alpha_n)^T \vartheta(\alpha_n).$$

By replacing the Jacobian matrix from (24), the equation will be changed to

$$\begin{bmatrix} \Gamma_n^{[1]^T}\Gamma_n^{[1]} & \Gamma_n^{[1]^T}\Gamma_n^{[2]} \\ \Gamma_n^{[2]^T}\Gamma_n^{[1]} & \Gamma_n^{[2]^T}\Gamma_n^{[2]} \end{bmatrix}\begin{bmatrix} \alpha_{n+1}^{[1]} \\ \alpha_{n+1}^{[2]} \end{bmatrix} = \begin{bmatrix} \Gamma_n^{[1]^T}\Gamma_n^{[1]} & \Gamma_n^{[1]^T}\Gamma_n^{[2]} \\ \Gamma_n^{[2]^T}\Gamma_n^{[1]} & \Gamma_n^{[2]^T}\Gamma_n^{[2]} \end{bmatrix}\begin{bmatrix} \alpha_n^{[1]} \\ \alpha_n^{[2]} \end{bmatrix} - \begin{bmatrix} \Gamma_n^{[1]^T} \\ \Gamma_n^{[2]^T} \end{bmatrix}\vartheta(\alpha_n).$$

Moreover, the following equation will be calculated, if $\vartheta(\alpha_n)$ is replaced by its equivalent value.

$$\begin{bmatrix} \Gamma_n^{[1]^T}\Gamma_n^{[1]} & \Gamma_n^{[1]^T}\Gamma_n^{[2]} \\ \Gamma_n^{[2]^T}\Gamma_n^{[1]} & \Gamma_n^{[2]^T}\Gamma_n^{[2]} \end{bmatrix}\alpha_{n+1} = \begin{bmatrix} \Gamma_n^{[1]^T}\Gamma_n^{[1]} & 0 \\ 0 & \Gamma_n^{[2]^T}\Gamma_n^{[2]} \end{bmatrix}\alpha_n + \begin{bmatrix} 0 & \Gamma_n^{[1]^T} \\ \Gamma_n^{[2]^T} & 0 \end{bmatrix}\beta_n. \tag{26}$$

Equation (26) will be rewritten in the following form, by multiplying both sides of (26) to $\begin{bmatrix} ({\Gamma_n^{[1]}}^T \Gamma_n^{[1]})^{-1} & \mathbb{0}_{n_{\alpha[1]}\times n_{\alpha[2]}} \\ \mathbb{0}_{n_{\alpha[2]}\times n_{\alpha[1]}} & ({\Gamma_n^{[2]}}^T \Gamma_n^{[2]})^{-1} \end{bmatrix}$.

$$\begin{bmatrix} I_{n_{\alpha[1]}} & {\Gamma_n^{[1]}}^\dagger \Gamma_n^{[2]} \\ {\Gamma_n^{[2]}}^\dagger \Gamma_n^{[1]} & I_{n_{\alpha[2]}} \end{bmatrix} \alpha_{n+1} = \begin{bmatrix} I_{n_{\alpha[1]}} & \mathbb{0}_{n_{\alpha[1]}\times n_{\alpha[2]}} \\ \mathbb{0}_{n_{\alpha[2]}\times n_{\alpha[1]}} & I_{n_{\alpha[2]}} \end{bmatrix} \alpha_n + \begin{bmatrix} \mathbb{0}_{n_{\alpha[1]}\times n_\vartheta} & {\Gamma_n^{[1]}}^\dagger \\ {\Gamma_n^{[2]}}^\dagger & \mathbb{0}_{n_{\alpha[2]}\times n_\vartheta} \end{bmatrix} \beta_n.$$

It should be noted that $\begin{bmatrix} I_{n_{\alpha[1]}} & \mathbb{0}_{n_{\alpha[1]}\times n_{\alpha[2]}} \\ \mathbb{0}_{n_{\alpha[2]}\times n_{\alpha[1]}} & I_{n_{\alpha[2]}} \end{bmatrix}$ in the above equation is identical to $I_{(n_{\alpha[1]}+n_{\alpha[2]})}$.

Moreover, by multiplying both equation sides of previous equation to

$$\begin{bmatrix} I_{n_{\alpha[1]}} & -{\Gamma_n^{[1]}}^\dagger \Gamma_n^{[2]} \\ -{\Gamma_n^{[2]}}^\dagger \Gamma_n^{[1]} & I_{n_{\alpha[2]}} \end{bmatrix},$$

the set of Equations (27) will be obtained which describes two sets of equations for calculating each of $\alpha_{n+1}^{[1]}$ and $\alpha_{n+1}^{[2]}$.

$$\begin{bmatrix} I_{n_{\alpha[1]}} - {\Gamma_n^{[1]}}^\dagger \Gamma_n^{[2]} {\Gamma_n^{[2]}}^\dagger \Gamma_n^{[1]} & \mathbb{0}_{n_{\alpha[1]}\times n_{\alpha[2]}} \\ \mathbb{0}_{n_{\alpha[2]}\times n_{\alpha[1]}} & I_{n_{\alpha[2]}} - {\Gamma_n^{[2]}}^\dagger \Gamma_n^{[1]} {\Gamma_n^{[1]}}^\dagger \Gamma_n^{[2]} \end{bmatrix} \alpha_{n+1} = \begin{bmatrix} I_{n_{\alpha[1]}} & -{\Gamma_n^{[1]}}^\dagger \Gamma_n^{[2]} \\ -{\Gamma_n^{[2]}}^\dagger \Gamma_n^{[1]} & I_{n_{\alpha[2]}} \end{bmatrix} \alpha_n + \begin{bmatrix} -{\Gamma_n^{[1]}}^\dagger \Gamma_n^{[2]} {\Gamma_n^{[2]}}^\dagger & {\Gamma_n^{[1]}}^\dagger \\ {\Gamma_n^{[2]}}^\dagger & -{\Gamma_n^{[2]}}^\dagger \Gamma_n^{[1]} {\Gamma_n^{[1]}}^\dagger \end{bmatrix} \beta_n. \tag{27}$$

The first set of Equations in (27) is

$$\left(I_{n_{\alpha[1]}} - {\Gamma_n^{[1]}}^\dagger \Gamma_n^{[2]} {\Gamma_n^{[2]}}^\dagger \Gamma_n^{[1]}\right) \alpha_{n+1}^{[1]} = \left[I_{n_{\alpha[1]}} \quad -{\Gamma_n^{[1]}}^\dagger \Gamma_n^{[2]}\right] \alpha_n + \left[-{\Gamma_n^{[1]}}^\dagger \Gamma_n^{[2]} {\Gamma_n^{[2]}}^\dagger \quad {\Gamma_n^{[1]}}^\dagger\right] \beta_n,$$

which can be rewritten as the following equation, by replacing α_n and β_n with $\left[\alpha_n^{[1]T} \; \alpha_n^{[2]T}\right]^T$ and $\left[\beta_n^{[1]T} \; \beta_n^{[2]T}\right]^T$.

$$\left(I_{n_{\alpha[1]}} - {\Gamma_n^{[1]}}^\dagger \Gamma_n^{[2]} {\Gamma_n^{[2]}}^\dagger \Gamma_n^{[1]}\right) \alpha_{n+1}^{[1]} = \alpha_n^{[1]} - {\Gamma_n^{[1]}}^\dagger \Gamma_n^{[2]} {\Gamma_n^{[2]}}^\dagger \beta_n^{[1]} - {\Gamma_n^{[1]}}^\dagger \left(\Gamma_n^{[2]} \alpha_n^{[2]} - \beta_n^{[2]}\right).$$

By replacing the $\Gamma_n^{[2]}\alpha_n^{[2]} - \beta_n^{[2]}$ with $\Gamma_n^{[1]}\alpha_n^{[1]} - \beta_n^{[1]}$, the following equation is calculated, since $(I_{n_{\alpha[1]}} - {\Gamma_n^{[1]}}^\dagger \Gamma_n^{[2]} {\Gamma_n^{[2]}}^\dagger \Gamma_n^{[1]})$ is equal to ${\Gamma_n^{[1]}}^\dagger (I_{n_\vartheta} - \Gamma_n^{[2]} {\Gamma_n^{[2]}}^\dagger) \Gamma_n^{[1]}$.

$$\left(I_{n_{\alpha[1]}} - {\Gamma_n^{[1]}}^\dagger \Gamma_n^{[2]} {\Gamma_n^{[2]}}^\dagger \Gamma_n^{[1]}\right) \alpha_{n+1}^{[1]} = \left({\Gamma_n^{[1]}}^\dagger - {\Gamma_n^{[1]}}^\dagger \Gamma_n^{[2]} {\Gamma_n^{[2]}}^\dagger\right) \beta_n^{[1]}.$$

So, the first set of equations in (27) can be simplified and rewritten as (23a).

Similar steps can be applied to the second set of equations in (27) and (23b) would be resulted. □

The Gauss-Newton algorithm is a modification of Newton's method for finding a minimum of a function which can be used to minimise a sum of squared function values. The Gauss-Newton algorithm will be derived from Newton's method for function optimisation via an approximation. In the Newton method, the Hessian of the function should be calculated, while the optimised value in the Gauss-Newton method is obtained by ignoring the second-order derivative terms. Appendix A will discuss about how to use the Newton method for the optimisation problem defined in Lemma 1.

By considering the constraints mentioned in (19) and (20), Equation (23a) should be solved as a constrained quadratic problem. Considering $\vartheta^T\vartheta$ as the goal function in the minimisation problem, in many practical cases, the elements of vector ϑ become less than the constraints mentioned in (21). Elsewhere, the constraint optimisation problem should be solved.

Remark 4. The necessary conditions to guarantee the existence of an optimal solution for a non-linear programming with both equality and inequality constraints are known as Karush-Kuhn-Tucker (KKT) conditions.

KKT conditions [30]: KKT conditions are first derivative necessary conditions for a solution of non-linear programming to be optimal, provided that some regularity conditions are satisfied. Suppose that the objective function $f : \mathbb{R}^n \to \mathbb{R}$ and the constraint functions $g_i : \mathbb{R}^n \to \mathbb{R} \quad (i = 1, \dots, m)$ and $h_j : \mathbb{R}^n \to \mathbb{R} \quad (j = 1, \dots, \ell)$ are continuously differentiable at a point x^*. If x^* is a local minimum of $f(x)$ subject to primal feasibility conditions

$$g_i(x^*) \le 0 \qquad (i = 1, \dots, m),$$
$$h_j(x^*) = 0 \qquad (j = 1, \dots, \ell),$$

then constants μ_i $(i = 1, \dots, m)$ and β_j $(j = 1, \dots, \ell)$, exist such that

$$\nabla f(x^*) + \sum_{i=1}^{m} \mu_i \quad \nabla g_i(x^*) + \sum_{j=1}^{l} \beta_j \nabla h_j(x^*) = 0, \tag{28a}$$

$$\mu_i g_i(x^*) = 0 \qquad (i = 1, \dots, m), \tag{28b}$$

$$\mu_i \geq 0 \qquad (i = 1, \dots, m). \tag{28c}$$

The constants μ_i $(i = 1, \dots, m)$ and β_j $(j = 1, \dots, \ell)$ are called the KKT multipliers.

Remark 5. In the particular case that no inequality constraints exist in the constrained non-linear programming ($m = 0$), the KKT conditions are identical to the Lagrange conditions, and the KKT multipliers are also called the Lagrange multipliers.

Combining the Newton method and KKT conditions, an iterative method for solving constrained optimisation problems will be resulted which is called the Sequential Quadratic Programming (SQP). In each iteration of SQP, a constrained quadratic program should be solved. The solving procedure for identification problem is described in the following theorem.

Theorem 3. *Consider the vectors* $\alpha = \left[\alpha_n^{[1]\,T} \ \alpha_n^{[2]\,T}\right]^T$, $\beta = \left[\beta_n^{[1]\,T} \ \beta_n^{[2]\,T}\right]^T$ *and* $\vartheta(\alpha) = \Gamma^{[1]}\alpha^{[1]} - \beta^{[1]} = \Gamma^{[2]}\alpha^{[2]} - \beta^{[2]}$ *where elements of* $\Gamma^{[1]}$ *and* $\beta^{[1]}$ *are functions of* $\alpha^{[2]}$ *and on the other hand elements of* $\Gamma^{[2]}$ *and* $\beta^{[2]}$ *are functions of* $\alpha^{[1]}$. *The obtained* α^* *by solving the following iterative constrained quadratic programming problem will be a local minimum of the scaler function* $\vartheta(\alpha)^T\vartheta(\alpha)$ *subject to* $\Gamma_{eq}\alpha - \beta_{eq} = 0$ *and* $\Gamma_{ineq}\alpha - \beta_{ineq} \leq 0$ *if* $\left[\Gamma^{[1]} \ \Gamma^{[2]}\right]$ *is a column independent matrix.*

$$\min\left\{ \alpha_{k+1}^T \begin{bmatrix} {\Gamma^{[1]}_k}^T\Gamma^{[1]}_k & {\Gamma^{[1]}_k}^T\Gamma^{[2]}_k \\ {\Gamma^{[2]}_k}^T\Gamma^{[1]}_k & {\Gamma^{[2]}_k}^T\Gamma^{[2]}_k \end{bmatrix} \alpha_{k+1} - \left(\begin{bmatrix} 2{\Gamma^{[1]}_k}^T\Gamma^{[1]}_k & {\Gamma^{[1]}_k}^T\Gamma^{[2]}_k \\ {\Gamma^{[2]}_k}^T\Gamma^{[1]}_k & 2{\Gamma^{[2]}_k}^T\Gamma^{[2]}_k \end{bmatrix} \alpha_k + \begin{bmatrix} 0 & {\Gamma^{[1]}_k}^T \\ {\Gamma^{[2]}_k}^T & 0 \end{bmatrix} \beta_k \right)^T \alpha_{k+1} \right\},$$

$$\textit{subject to} \begin{cases} \Gamma_{eq}\alpha_{k+1} - \beta_{eq} = 0 \\ \Gamma_{neq}\alpha_{k+1} - \beta_{neq} \leq 0 \end{cases}.$$

Proof. Using the results of [31], solution of the SQP iterative problem for the constrained problem $\min\limits_x f(x)$ subject to $h(x) = 0$ and $g(x) \leq 0$ will be

$$\min\left\{ \nabla_x f(x_k)^T (x_{k+1} - x_k) + \frac{1}{2}(x_{k+1} - x_k)^T \mathcal{HL}_x(x_k, \mu_k, \lambda_k)(x_{k+1} - x_k) \right\},$$

$$\text{subject to} \begin{cases} h(x_k) + \nabla_x h(x_k)(x_{k+1} - x_k) = 0 \\ g(x_k) + \nabla_x g(x_k)(x_{k+1} - x_k) \leq 0 \end{cases},$$

where ∇_x and $\mathcal{HL}_x$ correspondingly denote the gradient and Hessian functions.

In this theorem, the Lagrangian function is

$$\mathcal{L}(\alpha, \mu, \lambda) = \vartheta^T(\alpha)\vartheta(\alpha) + \lambda^T(\Gamma_{eq}\alpha - \beta_{eq}) + \mu^T(\Gamma_{neq}\alpha - \beta_{neq}).$$

Therefore, gradient and Hessian functions of it will be

$$\nabla_\alpha \mathcal{L}(\alpha, \mu, \lambda) = 2\left[\Gamma^{[1]} \ \Gamma^{[2]}\right]^T \vartheta + \Gamma_{eq}^T\lambda + \Gamma_{neq}^T\mu, \tag{29a}$$

$$\mathcal{HL}_\alpha(\alpha, \mu, \lambda) = 2\begin{bmatrix} \Gamma^{[1]\,T}\Gamma^{[1]} & \frac{\partial \Gamma^{[1]\,T}\vartheta}{\partial \alpha^{[2]}} \\ \frac{\partial \Gamma^{[2]\,T}\vartheta}{\partial \alpha^{[1]}} & \Gamma^{[2]\,T}\Gamma^{[2]} \end{bmatrix} \approx 2\begin{bmatrix} \Gamma^{[1]\,T}\Gamma^{[1]} & \Gamma^{[1]\,T}\Gamma^{[2]} \\ \Gamma^{[2]\,T}\Gamma^{[1]} & \Gamma^{[2]\,T}\Gamma^{[2]} \end{bmatrix}. \tag{29b}$$

So, by substituting (29a) and (29b) in SQP algorithm, the theorem will be proved. □

Applying this theorem on the system of linear equations given in Theorem 1 and Theorem 2 (or Proposition 1 and Proposition 2 for MIMO systems) and also the constraints in (19) and (20) will result the identified H–W model. In the following section, the proposed algorithm for identification of H–W model is applied on a PCV as a case study.

4 | CASE STUDY: PRESSURE CONTROL VALVE

4.1 | Valve actuation system

Control valves in industrial processes possibly present various non-linearities, which may degrade control performance or even cause oscillations arisen in control loops. The control loops oscillate due to valve non-linearities, such as stiction, hysteresis, and dead-band. Compressed air supply is an essential utility in every pneumatic based company like Oil & Gas refineries. Air compressor receives the atmospheric air as its input stream and increases its pressure to a desired level. A PCV is installed parallel to the compressor discharge line, in the exhaust line to

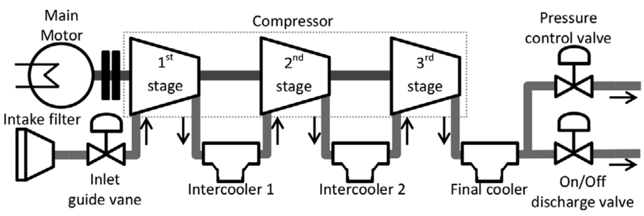

FIGURE 2 Schematic view of main parts in the air compressor and location of the PCV

protect the compressor against the surge phenomena. It is opened to ensure minimum air flow through the compressor. PCV controls the exhaust air flow rate by varying the size of the flow passage as directed by a signal from a controller. The location of PCV in a three-stage air compressor is shown in Figure 2.

The control command of PCV is calculated by a PLC, based on the compressor discharge pressure. The logic has three modes, called full open, auto, and manual. The full open mode of PCV is the safe mode of the compressor in which surge probability is minimal. This mode is selected by logic when the compressor is either in the stop mode or during the start-up phase. It is also used in case of any emergency situation when the full opening command is suddenly sent to the PCV. The logic suddenly switches to emergency situation mode in case of either very low discharge flow, very high discharge pressure or very high absolute slope of the discharge pressure. In other cases and if the manual mode has not been selected by the operator, the valve position is controlled by the maximum output of two PID controllers that both use the discharge pressure as their manipulated values. The logic also limits the output value between 0 and 100 percent.

PCV actuates by a diaphragm that is connected to its stem which is moved by air pressure (pneumatic force). Nowadays, The pneumatic diaphragm actuator is the most commonly used type of PCVs in industrial processes due to its high strength, its size and weight, and its simple and inexpensive construction. Diaphragm actuators usually use pressured air between 3∼15 Psig to operate the valve. In addition to the air pressure, a return spring is mounted under the diaphragm which applies a force in the opposite direction of the air force. In compressors, the safe mode following failure of the air signal would be when the PCV is driven in full open mode. To deal with this eventuality, the spring would drive the valve stem upward, when an air failure arises.

The calculated control command by PLC is transferred to the actuation system through an analog output module, in the form of 4∼20 mA current signal. In case of any wire break or signal failure the valve should switch to its safe mode. The analog output module of PLC is calibrated to linearly convert the logic output value to 4∼20mA current signal. In the ideal case, it is assumed that the calibration is perfect; but in practice, non-linearities affect the precise calibration. To consider this issue in the modelling, Hammerstein block (the input static non-linear function) is added to the input of the linear dynamic block representing the ideal model of PCV.

The electrical signal 4∼20 mA is sent to an I2P module to convert it to 3∼15 Psig pneumatic signal which makes the valve stem move between 0 to 100 percent of its travelling range. 4 mA and 20 mA signals are corresponding to 3 Psig (fully opens the valve) and 15 Psig (fully closes the valve) respectively. Increasing the air pressure makes the diaphragm move the valve stem downward and hence shutting the vent. For valves where the accurate and rapid control without error or hysteresis is required, valve positioners are used. The stem position is transferred to the positioner by a rod and is compared with the input signal to increase or decrease the air load pressure driving the actuator and to ensure the accurate position. In modern positioners, the I2P is also embedded in positioner and has an input connection for the electrical signal. The I2P module is also calibrated to linearly scale the mid-range values between these extreme values and actuate the valve between 0 to 100 percent of its travelling range. However, small calibration errors and non-linearities may affect the performance of PCV. It should also be noted that the position of PCV is measured using LVDTs or potentiometers to be transmitted to analog input module of PLC as a 4∼20mA electric signal. The measuring instruments usually have saturation, dead-band or any other type of non-linearities in its response. Therefore, Wiener block (the output static non-linear function) is added to the output of linear dynamic block to take into account the calibration errors and non-linearities of I2P as well as non-linearities of the measuring devices. So, a PCV is modelled as a linear block which represents the time evolution of the ideal system and the delays, along with non-linear blocks which are used to model the discussed sources of non-linearities in the positioner, transmitter, and valve structure. Due to having both non-linear and linear blocks in the model describing a PCV, H–W model which is illustrated in Figure 1 is an appropriate choice for modelling its behaviour. Non-linear blocks may contain hard non-linearities such as saturation or dead-zone.

In case of any emergency situation, the PCV should open the vent rapidly to guarantee the safety of air compressor. Therefore a quick-exhaust valve and a silencer is installed in the air exhaust line from diaphragm. Existence of the quick-exhaust valve and the effect of return spring cause the valve to have different models for opening and closing operations. This study focused on the valve modelling during its closing operation, because the opening operation is so quick that can not be sampled. It is worth noting that the PLC variables are sent to a monitoring system via a fast industrial network with an ignorable delay. The trend of data during the time is saved in a database in the monitoring system with an adjustable sampling time which is imported to MATLAB software for verification of the proposed method. The PCV actuation and monitoring system of this study is shown in Figure 3 [32].

4.2 | Identification of pressure control valve

Since the opening command is issued only in emergency situations, the trend of data during twelve opening conditions are gathered. The acquired data during nine closing operations are

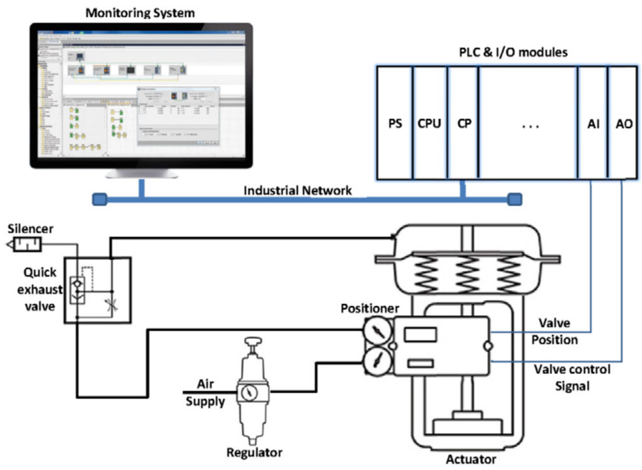

FIGURE 3 Actuation and monitoring system of the PCV

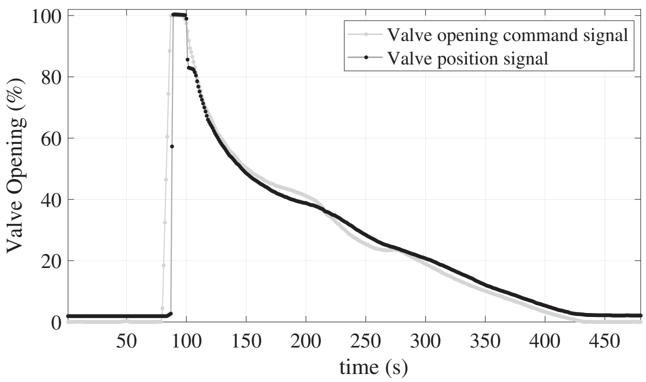

FIGURE 4 An instance of the valve opening-closing command and the valve position responded to it

used for H–W identification and the data of the rest three operations are used for verification of the obtained model. The command signal sent to the valve and its response is shown in Figure 4 at one of the opening-closing operations.

The following steps are used to identify this system through the proposed algorithm:

i) **Data preparation**: The trend of data during a three month period is gathered from the compressor. But only 100 minute of it is useful for identification purpose and the valve was fully closed in other times. The data was Exported from compressor monitoring system to excel and also imported in Matlab. To apply the proposed algorithm on the investigated PCV, the time periods of data which are rich of information should be separated from the other parts.
As it is mentioned, because of using quick exhaust valve, the model of PCV actuation system during opening and closing are different and only measured data in the closing operation is used for identification. It is noteworthy that even if a fixed sampling time is defined for the data acquisition system, the time intervals between each two consecutive samples are not exactly the same.

ii) **Interpolation of data**: Since the sampling time of the gathered data is considered to be fixed in the proposed algorithm (and most of others), interpolation at equal intervals is applied in order to uniform the time intervals of the data. By this step, the time can be neglected in the proposed identification algorithm.

iii) **Quantisation of non-linear functions**: Partitioning of measured input and output non-linear functions should be done based on the known information about distinct behaviours of non-linear functions in different regions. In this method, hard non-linearities, such as dead-band and saturation are allowed. The number of partitions in measured input and output range should be small enough to make the system overdetermined. On the other hand, using more partitions make it possible to specify the functions in more details. In this case study, a limited slope is expected in all regions of the Hammerstein non-linear function. So, a same limitation on the slope is assumed for all regions of the non-linear function.
Each region is partitioned to several sub-regions. The function value at the centre of each sub-region is considered as the elements of f' and g' vectors. After specifying partitions, Ω_f and Ω_g can be calculated.

iv) **Specifying the constraints**: The constrains of this case study are formulated in the form of equalities or inequalities which are linear in parameters. The constraints include the limitations of the slopes in the non-linear functions, in addition to the constraints shown in (19) and (20). For this purpose, the noise and disturbance amplitude are limited first. Then, the non-linear input function output range and slope and also inverse of non-linear output function input range and slope are limited. Moreover, the measurement noise considered to have limited range; so its effect on the functions outputs are limited according to their approximated slopes. In this case study, the functions outputs are limited between 0 to 100 and the absolute function slopes between 5% to 95% of output function range and whole input function range are limited between 0 to 5.

v) **Determine initial functions and linear block parameters to specify α_0**: Using the known information, based of the physical considerations, an initial guess for the non-linear functions. However, a common choice for the starting point can be a linear shape with slope of one and y-intercept of zero. In this case study, a linear function with slope one is considered for the starting point.

vi) **Solving the constrained quadratic problem**: The constrained quadratic problem should be solved, using the iterative method given in Theorem 3. Solving the constrained problem, parameters of the linear block and the ordered pairs of non-linear functions will be determined. The iteration of the algorithm in Theorem 3 continues while the sum of squares of its decreasing rate is less than a threshold.

Figure 5 summarises the proposed identification method for the PCV.

Applying proposed identification method on the gathered data and considering the linear block as a third order SISO

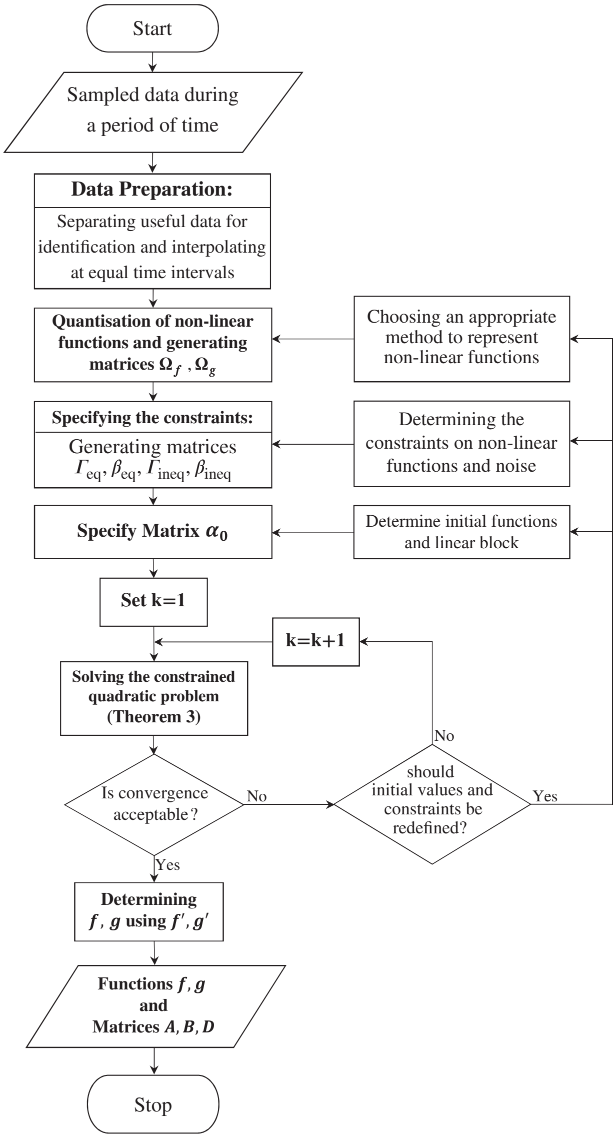

FIGURE 5 Flowchart of the proposed identification method

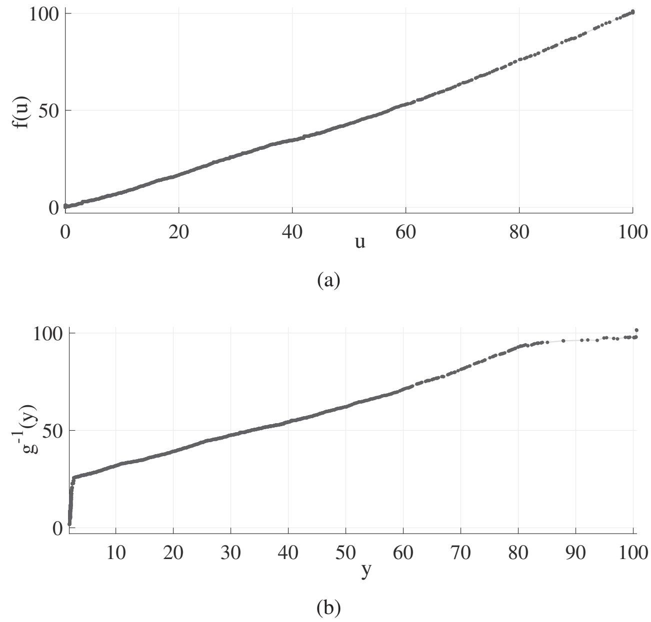

FIGURE 6 The identified non-linear function of PCV model: (a) Hammerstein function, (b) Wiener function

system with the observable canonical state space realisation, $a = [-0.0009\ 0.7650\ -1.7484]^T$, $b = [0.7959\ -1.7600\ 0.9884]^T$, and $d = [0]$ will be resulted. In other words, state space matrices (A, B, C, D) of the linear block in (3) can be written as

$$A = \begin{bmatrix} 0 & 0 & 0.0009 \\ 1 & 0 & -0.7650 \\ 0 & 1 & 1.7484 \end{bmatrix}, B = \begin{bmatrix} 0.7959 \\ -1.7600 \\ 0.9884 \end{bmatrix}, C = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}, D = [0].$$

Non-linear functions f and g^{-1} are also shown in Figure 6. It should be noted that in Figure 6(b), instead of the Wiener function, its inverse is illustrated, due to the effect of hard non-linearities (such as saturation) on it. Moreover, Figure 6(b) illustrates hard non-linearities and non-linear function $g^{-1}(y)$ is not one-to-one. Hard non-linearities are seen near $y = 0\%$ and $y = 100\%$; while the function is one-to-one and almost linear in the middle-range of y. Shape of the Wiener function in Figure 6(b) is due to the calibration of the PCV and/or I/O cards. As it is illustrated in this figure, decreasing the control command from 100% to 80%, no significant changes occur in the operation of the PCV. Continuing decreasing the control command less than 80%, the PCV starts to follow the control signal linearly. This is mainly due to the dead-band in the PCV. Since the PCV is a normally open pneumatic valve, there is an inverse relationship between the percentage of the valve opening and the air pressure applied to the valve stem (i.e. valve is closed by the air pressure, and is opened by a return spring). The non-linear behaviour of $g^{-1}(y)$ near $y = 0\%$ is also intentionally set in the calibration process to guarantee the valve being closed in the normal performance of system.

Verification of the obtained H–W model with the data during three closing operations which are not used in the identification process is illustrated in Figure 7. It can be seen that accuracy of the identified model is more than 95%. It is worth noting that the main reason of larger error (almost 5%) at the beginning of the closing operations is that fewer number of ordered pairs are sampled near $y = 100\%$. For $y < 80\%$ with more sampled data, the identification error is less that 0.5%. So, if a more accurate model is required, data acquisition with higher sampling rates should be used. Time evaluation of states for the identified linear block is also shown in Figure 8 for a closing operation used in the identification process.

5 | CONCLUSION

This paper has considered the problem of identifying industrial plants. An algorithm is proposed for identification of H–W model, based on constrained optimisation. To convert the

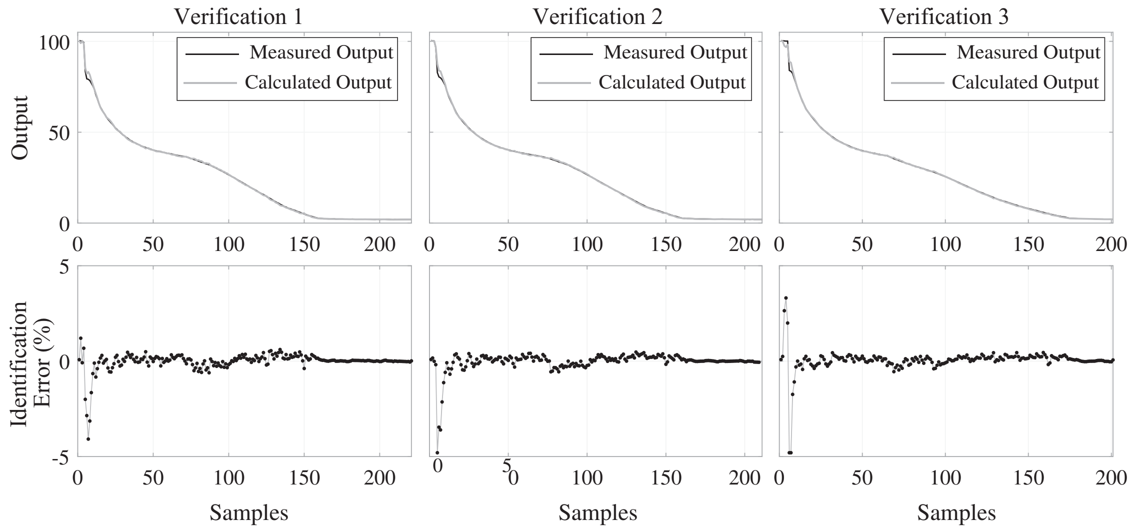

FIGURE 7 Verification of the identified H–W model in three closing operations of the PCV

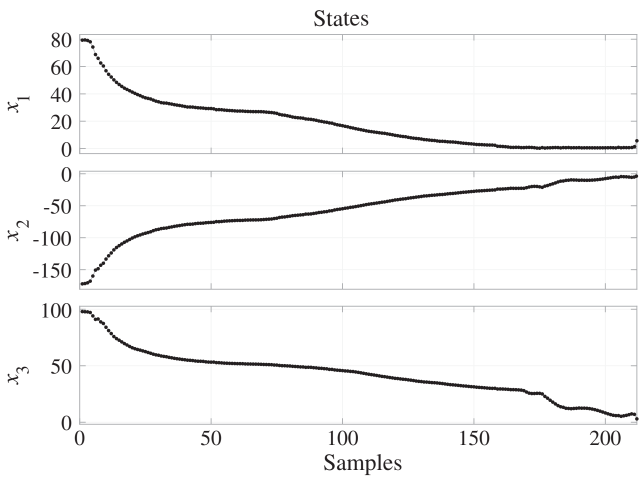

FIGURE 8 Time evaluation of states, during a closing operation

identification problem into an optimisation problem, noise and disturbance terms are sent to one side of equalities and the least square approach is used. Since the least square method is a solution for overdetermined systems, one method to solve the problem is reducing the number of unknown variables to have an overdetermined systems. So, known behaviour of non-linear functions are taken into account to describe them. To ensure that the non-linear functions have desired behaviour and prevent the out of range errors, or fast changing functions, some constraints are defined. The constraints consist of limiting function outputs ranges, slopes and known values of functions. In the proposed method which can be used for MIMO systems, iterative multivariate SQP method is applied to address the non-linear optimisation problem.

The main advantage of the proposed algorithm over the existing ones is that the most general form of practical systems with all types of noise and disturbances applied to linear and non-linear blocks are considered. Moreover, some of the restrictive assumptions are relaxed. For example, invertibility of non-linear functions is relaxed, and even hard non-linearities can be considered as non-linear blocks. In addition, the known information about the practical interpretation of non-linear functions is used to improve the accuracy of the identified model. It should also be noted that sampled data for this method may belong to different time intervals with various initial conditions. Therefore, the method can be used in applications with different modes, such as the switched systems. For example in our case study, the dynamics of PCV are different in opening and closing operation modes.

To evaluate the effectiveness of the proposed method in identification of real-world applications, a valve actuation system is investigated. A pressure control valve is identified using the real data gathered from the industrial monitoring system of a gas refinery. Verification of the identified H–W model by data which is not used in the identification shows 95% accuracy of the model.

ACKNOWLEDGEMENTS

The authors would like to thank the South Pars Gas Complex for issuance of data collection permission, and hope the research results would be useful and encourage the company for funding in the future.

REFERENCES

1. Giri, F., Bai, E.-W.: Block-Oriented Nonlinear System Identification. Springer, New York (2010)
2. Hong, M., Cheng, S.: Hammerstein-Wiener model predictive control of continuous stirred tank reactor. In: Lecture Notes in Electrical Engineering, vol. 97, pp. 235–242. Springer, New York (2011)
3. Patikirikorala, T., et al.: Hammerstein-Wiener nonlinear model based predictive control for relative QoS performance and resource management of software systems. Control Eng. Pract. 20 (1), 49–61 (2012)
4. Gomez, J.C., et al.: Wiener model identification and predictive control of a pH neutralisation process. IEE Proc., Control Theory Appl. 151 (3), 329–338 (2004)

5. Nemati, A., Faieghi, M.: The performance comparison of ANFIS and Hammerstein-Wiener models for BLDC motors. In: Lecture Notes in Electrical Engineering, vol. 97, pp. 29–37. Springer, New York (2011)
6. Taringou, F., et al.: Behaviour modelling of wideband RF transmitters using hammerstein–wiener models. IET Circuits, Devices Syst. 4 (4), 282 (2010)
7. Zhang, J., et al.: Disturbance-encoding-based neural hammerstein-wiener model for industrial process predictive control. IEEE Trans. Syst. Man Cybern. Syst. 1–12 (2020), https://doi.org/10.1109/TSMC.2020.3004382
8. Alsaedi, A., et al.: Combined state and multi-innovation parameter estimation for an input non-linear state-space system using the key term separation. IET Control Theory Appl. 10 (13), 1503–1512 (2016)
9. Bauer, M., et al.: The current state of control loop performance monitoring - A survey of application in industry. J. Process Control 38, 1–10 (2016)
10. Bacci di Capaci, R., et al.: System identification applied to stiction quantification in industrial control loops: A comparative study. J. Process Control 46, 11–23 (2016)
11. Ma, L., Liu, X.: A nonlinear recursive instrumental variables identification method of Hammerstein ARMAX system. Nonlinear Dyn. 79 (2), 1601–1613 (2014)
12. Lindsten, F., et al.: Bayesian semiparametric Wiener system identification. Automatica 49 (7), 2053–2063 (2013)
13. Guo, J., Liu, H.: Hammerstein system identification with quantised inputs and quantised output observations. IET Control Theory Appl. 11, 593–599 (2017)
14. Dong, S., et al.: Identification of hammerstein systems with time delay under load disturbance. IET Control Theory Appl. 12 (7), 942–952 (2018)
15. Bai, E.W.: An optimal two stage identification algorithm for Hammerstein-Wiener nonlinear systems. Proc. Am. Control Conf. 5 (3), 2756–2760 (1998)
16. Bai, E.W., Fu, M.: A blind approach to Hammerstein model identification. IEEE Trans. Signal Process. 50 (7), 1610–1619 (2002)
17. Zhu, Y.: Estimation of an N-L-N Hammerstein-Wiener model. 15th Triennial World Congress of the International Federation of Automatic Control 38(9), 1607–1614 (2002), https://doi.org/10.1016/S0005-1098(02)00062-6
18. Vörös, J.: Iterative identification of nonlinear dynamic systems with output backlash using three-block cascade models. Nonlinear Dyn. 79 (3), 2187–2195 (2015)
19. Crama, P., Schoukens, J.: Hammerstein-Wiener system estimator initialization. Automatica 40 (9), 1543–1550 (2004)
20. MacArthur, J.W.: A new approach for nonlinear process identification using orthonormal bases and ordinal splines. J. Process Control 22 (2), 375–389 (2012)
21. Ni, B., et al.: A refined instrumental variable method for hammerstein-wiener continuous-time model identification. IFAC Proc. Vol. 16 (PART 1), 1061–1066 (2012)
22. Boubaker, S.: Identification of nonlinear Hammerstein system using mixed integer-real coded particle swarm optimization: application to the electric daily peak-load forecasting. Nonlinear Dyn. 90 (2), 797–814 (2017)
23. Wang, Y., Ding, F.: Recursive least squares algorithm and gradient algorithm for Hammerstein–Wiener systems using the data filtering. Nonlinear Dyn. 84 (2), 1045–1053 (2016)
24. Wills, A., et al.: Identification of Hammerstein–Wiener models. Automatica 49 (1), 70–81 (2013)
25. Esmaeilani, L., Ghaisari, J.: Surge cause study in three-stages air compressors. In: 2016 24th Iranian Conference on Electrical Engineering, ICEE 2016, Shiraz (2016)
26. Esmaeilani, L., et al.: Bayesian approach to identify Hammerstein–Wiener non-linear model in presence of noise and disturbance. IET Control Theory Appl. 13 (3), 367–376 (2018)
27. Steffensen, J.F.: Interpolation: Second Edition. Dover Publications, New York (2006)
28. Seeger, M.: Gaussian Processes for Machine Learning. vol. 14. MIT Press, Cambridge (2004)
29. Laub, A.J.: Matrix Analysis for Scientists and Engineers. SIAM, Philadelphia (2005)
30. Boyd, S., Vandenberghe, L.: Convex Optimization. Cambridge University Press, Cambridge (2004)
31. Kress, B.C.: Numerical Optimization. Springer Science & Business Media, Heidelberg (2015)
32. Esmaeilani, L., et al.: Control valve position response identification by Matlab. In: Conference Record - IEEE Instrumentation and Measurement Technology Conference, Pisa (2015)

How to cite this article: Esmaeilani L, Ghaisari J, Bagherzadeh MA. Hammerstein–Wiener identification of industrial plants: A pressure control valve case study. *IET Control Theory Appl.* 2020;1–16. https://doi.org/10.1049/cth2.12052

APPENDIX A

A.1 | Newton method for optimisation

In an especial case of the least square problem defined in Lemma 1, the matrices $\Gamma^{[1]}$ and $\Gamma^{[2]}$ and vectors $\beta^{[1]}$ and $\beta^{[2]}$ could be written in the form of

$$\Gamma^{[1]}(\alpha^{[2]}) = \left[\beta_\Gamma^{[1]}\alpha^{[2]} + \gamma_\Gamma^{[1]}\right]_{n_\vartheta}^T,$$

$$\Gamma^{[2]}(\alpha^{[1]}) = \left[\beta_\Gamma^{[2]}\alpha^{[1]} + \gamma_\Gamma^{[2]}\right]_{n_\vartheta}^T,$$

$$\beta^{[1]}(\alpha^{[2]}) = \beta_\beta^{[1]}\alpha^{[2]} + \gamma_\beta^{[1]},$$

$$\beta^{[2]}(\alpha^{[1]}) = \beta_\beta^{[2]}\alpha^{[1]} + \gamma_\beta^{[2]},$$

where $\beta_\Gamma^{[1]}, \beta_\Gamma^{[2]}, \beta_\beta^{[1]}, \beta_\beta^{[2]}, \gamma_\Gamma^{[1]}, \gamma_\Gamma^{[2]}, \gamma_\beta^{[1]}$ and $\gamma_\beta^{[2]}$ do not depend on α.

Moreover, Hessian matrix of the scaler number $\vartheta^T\vartheta$ with respect to α is calculated as

$$H_\alpha(\vartheta^T\vartheta) = \begin{bmatrix} \frac{\partial^2(\vartheta^T\vartheta)}{\partial {\alpha^{[1]}}^2} & \frac{\partial^2(\vartheta^T\vartheta)}{\partial\alpha^{[1]}\partial\alpha^{[2]}} \\ \frac{\partial^2(\vartheta^T\vartheta)}{\partial\alpha^{[2]}\partial\alpha^{[1]}} & \frac{\partial^2(\vartheta^T\vartheta)}{\partial {\alpha^{[2]}}^2} \end{bmatrix},$$

where

$$\frac{\partial^2(\vartheta^T\vartheta)}{\partial {\alpha^{[1]}}^2} = {\Gamma^{[1]}}^T\Gamma^{[1]},$$

$$\frac{\partial^2(\vartheta^T\vartheta)}{\partial\alpha^{[1]}\partial\alpha^{[2]}} = \frac{\partial\left({\Gamma^{[2]}}^T\vartheta\right)}{\partial\alpha^{[1]}},$$

$$\frac{\partial^2(\vartheta^T\vartheta)}{\partial\alpha^{[2]}\partial\alpha^{[1]}} = \frac{\partial\left({\Gamma^{[1]}}^T\vartheta\right)}{\partial\alpha^{[2]}},$$

$$\frac{\partial^2(\vartheta^T\vartheta)}{\partial {\alpha^{[2]}}^2} = {\Gamma^{[2]}}^T\Gamma^{[2]}.$$

Moreover, the following equations hold for $\Gamma^{[2]^T}\vartheta(\alpha)$.

$$\begin{aligned}\Gamma^{[2]^T}\vartheta(\alpha) &= \Gamma^{[2]^T}\left(\Gamma^{[1]}\alpha^{[1]} + \beta^{[1]}\right)\\ &= \Gamma^{[2]^T}\Gamma^{[1]}\alpha^{[1]} + \Gamma^{[2]^T}\beta^{[1]}\\ &= vec\left(\Gamma^{[2]^T}\Gamma^{[1]}\alpha^{[1]}\right) + vec\left(\Gamma^{[2]^T}\beta^{[1]}\right)\\ &= vec\left(I_{n_\vartheta}\Gamma^{[2]^T}\Gamma^{[1]}\alpha^{[1]}\right) + vec\left(I_{n_\vartheta}\Gamma^{[2]^T}\beta^{[1]}\right)\\ &= (\alpha^{[1]^T}\Gamma^{[1]^T} \otimes I_{n_\vartheta})\ vec\left(\Gamma^{[2]^T}\right)\\ &\quad + (\beta^{[1]^T} \otimes I_{n_\vartheta})\ vec\left(\Gamma^{[2]^T}\right)\\ &= (\alpha^{[1]^T}\Gamma^{[1]^T} \otimes I_{n_\vartheta})(\beta_\Gamma^{[2]}\alpha^{[1]} + \gamma_\Gamma^{[2]})\\ &\quad + (\beta^{[1]^T} \otimes I_{n_\vartheta})(\beta_\Gamma^{[2]}\alpha^{[1]} + \gamma_\Gamma^{[2]})\end{aligned}$$

Therefore, the partial derivative of $\Gamma^{[2]^T}\vartheta(\alpha)$ with respect to $\alpha^{[1]}$ will be written as

$$\begin{aligned}\frac{\partial\left(\Gamma^{[2]^T}\vartheta(\alpha)\right)}{\partial\alpha^{[1]}} &= \Gamma^{[2]^T}\Gamma^{[1]} + (\alpha^{[1]^T}\Gamma^{[1]^T} \otimes I_{n_\vartheta})\beta_\Gamma^{[2]}\\ &\quad + (\beta^{[1]^T} \otimes I_{n_\vartheta})\beta_\Gamma^{[2]}\\ &= \Gamma^{[2]^T}\Gamma^{[1]} + (\vartheta(\alpha)^T \otimes I_{n_\vartheta})\beta_\Gamma^{[2]}.\end{aligned}$$

And finally, the Hessian matrix will be as follows.

$$H_\alpha(\vartheta^T\vartheta) = \begin{bmatrix}\Gamma^{[1]T}\Gamma^{[1]} & \Gamma^{[2]T}\Gamma^{[1]}\\ \Gamma^{[1]T}\Gamma^{[2]} & \Gamma^{[2]T}\Gamma^{[2]}\end{bmatrix} + \begin{bmatrix}\mathbb{0}_{n_{\alpha^{[1]}}\times n_{\alpha^{[1]}}} & (\vartheta(\alpha)^T \otimes I_{n_\vartheta})\beta_\Gamma^{[2]}\\ (\vartheta(\alpha)^T \otimes I_{n_\vartheta})\beta_\Gamma^{[1]} & \mathbb{0}_{n_{\alpha^{[2]}}\times n_{\alpha^{[2]}}}\end{bmatrix}$$

The Newton method for optimisation uses the Hessian matrix instead of $2J^TJ$ matrix in Gauss-Newton method.