A New Robust Identification Algorithm for Hammerstein-Like System Using Identification Error Structure

ZHENGBIN LI1, LIJUN MA2, AND YONGQIANG WANG3

1School of Electronic Information and Electrical Engineering, Anyang Institute of Technology, Anyang 455000, China
2Power Plant, Anyang Iron and Steel Company, Ltd., Anyang 455000, China
3Iron-Making Plant, Anyang Iron and Steel Company, Ltd., Anyang 455000, China

Corresponding author: Zhengbin Li (ligzwq08052008@163.com)

This work was supported by the Key Specialized Research and Development Projects of Henan Province under Grant 202102210337.

ABSTRACT A new robust identification algorithm is introduced in this study for a Hammerstein-like system based on identification error information. With the help of the half-substitution idea, the identification model is converted into a compact model where the coupling parameters are avoided. To reduce the effect of noise signals, a filter gain is proposed to obtain helpful system data. Then, on the basis of the filtered variables and developed forcing variables, the identification error information is extracted from the helpful system data. By using the identification error data, a new parameter estimation adaptive law that differs from the classic prediction error method is derived. Therefore, a new identification scheme framework is proposed and the weakness of the prediction error method is improved. Simulations and a real-life plant are presented to test the validity and practicality of the presented identification approach. The results of parameter estimation and estimation error qualitatively demonstrated the advantages of the proposed algorithm. The computational complexity and performance evaluation indicators results of the developed algorithm quantitatively indicated that the proposed algorithm produces higher estimation performance compared with the other algorithms.

INDEX TERMS Hammerstein-like system, memory nonlinearity, parameter estimation, robust algorithm, identification error.

I. INTRODUCTION

Block-oriented models provide a flexible and comprehensible nonlinear mathematical model for nonlinear engineering processes due to separable characteristics [1], [2], in which block-oriented models are composed of some linear time invariant subsystems and memoryless nonlinear elements. Hence, block-oriented models have been used to model a variety of practical systems, including the pH process [3], servo motor system [4], battery systems [5], health management system [6], and other processes. However, block-oriented models with memoryless nonlinear elements exhibit limited modeling performance and can only effectively establish a mathematical model for processes without memory characteristic. A Hammerstein system is a well-known block-oriented model that describes an actuator in series with a linear time invariant subsystem. Therefore, research on the identification of a Hammerstein system is one of the most popular topics in block-oriented system identification.

For the past decades, various identification schemes have been developed and used in theoretical framework and practical applications, such as recursive estimation type [7], [8], iteration identification type [9], [10], expectation maximization identification type [3], [12], maximum likelihood identification [14] and subspace identification [15], [16] and fractional order identification [17], [18], etc. Mao et al. [9] studied a parameter separation recursive least squares identification scheme for a Hammerstein model by using a decomposition technique and a filter, wherein the product items of the estimated parameters are avoided. Sobhani Tehrani and Kearney [13] used a novel method for non-parametric identification method to achieve the parameter information recovery of a parameter-varying Hammerstein systems, in which the large colored is considered to validate the robustness of the developed method. The nonlinearities in the aforementioned Hammerstein systems are memoryless, and thus cannot meet the needs of actual processes, particularly nonlinear
real-life systems with memory characteristic. To extend the application range and prospect of a Hammerstein system, scholars and engineers have used a memory nonlinear element to replace the memoryless nonlinear element, producing a Hammerstein-like system [19]–[21]. A Hammerstein-like system exhibits a stronger modelling performance than a Hammerstein system due to the memory nonlinear block, in which the identification of such system becomes a challenging problem. Inspired by the advantages and difficulties of a Hammerstein-like system, the current study focuses on the system identification of Hammerstein-like system with memory hysteresis nonlinearity. The considered system is shown in Fig. 1. Hysteresis nonlinearity is denoted by $f$, and the linear time invariant subsystem is described by $L$. The system input, output and internal variable are denoted by $u(t)$, $x(t)$ and $y(t)$, respectively.

Some effective and useful identification schemes have been published [3], [22]–[24] in recent years. Krikelis et al. [20] designed a subspace identification algorithm for a Hammerstein-like system with hysteretic nonlinearity, in which the memory hysteresis is represented by using recurrent artificial neural networks. A least-squares like estimator was presented on the basis of the periodic input signals by Brouri et al. [25] to achieve the parameter estimation for the Hammerstein-Wiener-like system. In [26], Ai et al. used a Hammerstein-like system with hysteresis to model a pneumatic muscle actuator based on a back propagation (BP) neural network, in which model parameters were estimated using a nonlinear least squares estimator. Li et al. [27] proposed an adaptive estimation method to recover the information of the Wiener-Hammerstein-like system, in which the convergence rate can be lifted by using attenuation coefficient. An artificial bee colony algorithm was developed by Yang et al. [28] to estimate model information of Hammerstein-like system, and experiment verified the effectiveness of the proposed algorithm. From the preceding reports on Hammerstein-like system identification, we can observe that most identification schemes have been developed using the prediction or observation error data [29]. Unsatisfactory estimation performance will emerge when the strong noise and low-quality initial values exist [32]–[34]. To address the aforementioned problems, some intelligence algorithms are used. In [30], Liu et al. published a neural network-based estimator for identifying vehicle sideslip angle, in which better identification performance can be achieved compared with kinematics and dynamics estimators. She [31] et al. proposed an incremental capacity (IC) analysis to estimate battery state of health by considering an equivalent IC-value calculation. In accordance with the previous work, we use identification error data in the current study to develop an identification algorithm that is directly related to parameter estimation to implement parameter update instead of other error data. Nevertheless, the identification error data are difficult to obtain in parameter estimation due to the unknown real value of estimated parameters, especially in the identification of a discrete system [35], [36]. Hence, we need to use available system data to derive information of identification error, which is also the original purpose of this current study.

However, some uncertainties are inevitable when the identification data are recorded. These disturbances, modelling errors, and various uncertainties may influence system stability, parameter estimation accuracy, and may even produce instability and oscillation phenomenon [37]. Thus, to avoid these problems, several necessary measures must be taken in the design of identification methods and control algorithms. In particular, a filter is commonly used to improve the aforementioned issues. All types of developed filters to achieve the parameter estimation and the effective control are presented, including Kalman filter [38], linear filter [39], transversal recursive filter [40], particle filter [41], adaptive filter [42], and dissipative filter [43], etc. Li and Liu [44] discussed a hierarchical multi-innovation gradient scheme for the Hammerstein system by using a linear filter. Estimation performance was enhanced using a long innovation length. In [45], Subrahmanya and Shin proposed an adaptive all-pass filter for delay systems by using a least mean square structure, in which the accurate and capable performance of tracking time-varying delays is enhanced compared with other schemes. Filipovic et al. [46] a robust filter for a stochastic systems with parameter uncertainties, in which an upper bound of the filtering error is guaranteed. Wang [47] et al. reported an adaptive extended Kalman filter for identifying the roll angle and rate of vehicles. They used recursive least squares with a forgetting factor to estimate the height center of gravity. A $H_{\infty}$ Kalman filter was presented by Zhao and Mili [48] for estimating a two-axis generator model, wherein the unknown state and model uncertainties are estimated. Although the filter can implement effective parameter identification, common filters impose strict conditions, such as a strictly real condition [44], redundant adjustment parameters [49], and less modelling uncertainties [50], etc. To simplify the filter design, some filters with less restrictive conditions have been proposed and used. One good solution was put forward by Na [36], [51], in which strict conditions of the system model are not required and has only one tuning parameter. Inspired by the work in [36], [51], filter gain is applied to conduct data preprocessing in the current paper.

In this study, we will provide a novel methodology for developing an identifier for a Hammerstein-like system with memory hysteresis nonlinearity. This methodology differs from the traditional prediction error method. In the algorithm design process, a filter and some forcing variables with adaptive attenuation coefficient are used to derive identification.
error information. Based on the obtained error information, a novel adaptive law is developed. The contributions of the current work are summarized as follows:

1. A compact identification model without coupling parameters is constructed on the basis of the half-substitution technology, such that being highly time-consuming is avoided.

2. A filter gain is proposed to obtain useful identification data from the collected system data, in which the strict restrictions are not required compared with published filters [38], [39], [42].

3. An identification error information is derived on the basis of several developed imposed variables and filtered data, providing a new method for obtaining identification error data.

4. A novel parameter estimation adaptive law is developed using identification error data compared to the traditional estimators [3], [20], [23], so that a new identification scheme framework is reported. The verification results based on simulation and experiment show that the designed scheme can achieve a better identification performance compared with existing schemes.

This study is organized as follows. Section 2 offers the problem considered. In Section 3, the design of the identification algorithm is constructed on the basis of the half-substitution technology, such that being highly time-consuming is avoided.

II. PROBLEM DESCRIPTION

Let’s give some symbols, \( x(t) \) describes the value of \( x \). \( \hat{x}(t) \) be the estimation value of \( x \), \( x^T(t) \) is the transpose of \( x(t) \). \( x^{-1}(t) \) is the inverse matrix of \( x(t) \). \( \hat{x}(t) \) represents the estimation error of \( x(t) \). \( \lambda_{max} \{x\} \) is the maximum eigenvalue of \( x \), \( E[x|y] \) denotes conditional expectation. I describes the unit vector. \( x(0) \) denotes the initial value of \( x(t) \). \( \max \{a, b\} \) is the maximum value of \( a, b \). \( \min \{a, b\} \) is the minimum value of \( a, b \).

In this study, we consider a Hammerstein-like system with memory hysteresis nonlinearity, whose dynamics are expressed by

\[
\begin{align*}
 f : \dot{x}(t) &= f_1(x(t), u(t))\dot{u}+(t) + f_2(x(t), u(t))\dot{u}-(t), \\
 x(0) &= x_0, \\
 L : y(t) &= \sum_{i=1}^{n_a} a_i x(t-i) - \sum_{j=1}^{n_b} b_j y(t-j) + \omega(t),
\end{align*}
\]

where \( f \) is Duhem hysteresis model, \( f_1(x(t), u(t)) \) and \( f_2(x(t), u(t)) \) are two functions, \( \dot{u}-(t) := \min\{0, \dot{u}(t)\}, \dot{u}+(t) := \max\{0, \dot{u}(t)\} \). \( L \) denotes the linear system. \( n_a \) and \( n_b \) are orders. \( \omega(t) \) is the additional noise. \( u(t), y(t) \) describes the input, output of the system. \( a_i, b_j \) denote coefficients.

This study aims at estimating parameters of the considered system in (1)-(2) by proposing a new identification approach, analysing the convergence nature of the developed scheme, testing the practicality and efficiency of the algorithm in Section 3 using the examples.

Remark 1: Hysteresis is a common memory nonlinearity in practical systems [52]–[54]. However, hysteresis is a destructive nonlinearity that leads to oscillation, lag and even instability in practical systems [55]. We need to identify the unknown hysteresis parameters, and offers parameter information for the subsequent control scheme design. The Duhem model in Eq.(1) cannot represent frequently-used hysteresis models based on different functions \( f(x(t), u(t)) \). The hysteresis model can be represented by

\[
\begin{align*}
 h_1(x(t), u(t)) &= P(x(t), u(t)) + Q(x(t), u(t)) \\
 h_2(x(t), u(t)) &= -P(x(t), u(t)) + Q(x(t), u(t))
\end{align*}
\]

According to [61], \( P(x(t), u(t)) \) and \( Q(x(t), u(t)) \) are defined by

\[
\begin{align*}
 P(x(t), u(t)) &= -\gamma x(t) + \gamma p_0 \text{sgn}(u(t)) \sum_{i=1}^{n} p_i(u(t))^{2i-2}, \\
 Q(x(t), u(t)) &= \sum_{i=1}^{n} q_i u(t-i)^{2i-2},
\end{align*}
\]

where \( \text{sgn}(\cdot) \) denotes the sign function. \( \gamma, p_0, p_i \) and \( q_{i-1} \) represent the constants.

From (3)-(6), the discrete-time form of Duhem hysteresis is provided by

\[
\begin{align*}
 x(t) &= x(t-1) + [\gamma \max\{0, \dot{u}(t-1)\} \\
 &\quad + \gamma \sum_{i=1}^{n} p_i(u(t-1))^{2i-1} - \gamma u(t-1)] \\
 &\quad \times [u(t) - u(t-1)] + \sum_{i=1}^{n} q_i (u(t-1))^{2i-2} \\
 &\quad \times (u(t) - u(t-1)).
\end{align*}
\]

Based on the half-substitution idea [62], by inserting (7) into (2), we have the following identification model of this paper:

\[
\begin{align*}
 y(t) &= \theta^T \varphi(t) + \omega(t),
\end{align*}
\]

the observation vector is described as

\[
\varphi(t) = [x(t-2), \text{sgn}(u(t-2))\Delta u(t), u(t-2)\Delta u(t), \ldots, u(t-2)^{2n-1}\Delta u(t), -u(t-2)\Delta u(t)],
\]

\[
\Delta u(t) = u(t) - u(t-1).
\]
\[ \Delta u(t), \ldots, u(t-2)\Delta u(t), x(t-2), \ldots, x(t-n_a), -y(t-1), \ldots, -y(t-n_b) \]  
\[ \theta = [a_1, a_1 y p_0, \ldots, a_1 y p_n, a_1 y, a_1 q_0, \ldots, a_1 q_{n-1}, a_2, \ldots, a_2 n_0, b_1, \ldots, b_2 n_0]^T, \]  
where \( \Delta u(t) = u(t-1) - u(t-2) \).

To complete the goal of this paper, some common assumptions are offered as follows [63], [64]:

**Assumption 1**: The signal \( u(t) \) is a persistent excitation input signal.

**Assumption 2**: The linear submodel \( L \) is stable.

**Assumption 3**: The system initial values are chosen as zero.

**Assumption 4**: The orders \( n_a \) and \( n_b \), \( n \) are known and the corresponding coefficients \( a_i, b_j, p_i, p_0 \) and \( q_i \) are unknown.

**Assumption 5**: The additional noise \( \omega(t) \) and input signal \( u(t) \) are independent of each other.

Assumption 1 shows the condition that the modality of the system is fully excited by using the input signal. The work condition for the linear subsystem \( L \) is described by Assumption 2, the physically realizable of the system is displayed in Assumption 3. In Assumption 4, the purpose of this paper is given. The noise assumption is indicated by Assumption 5.

### III. ESTIMATION ALGORITHM

In this section, we develop an adaptive identifier for achieving the parameter identification of the considered system. The designed estimator exhibits a better identification performance than the traditional estimator via a new parameter update law framework.

Data preprocessing is typically implemented using the developed filter, but the complex filter structure is the obstacle for designing an effective filter. Therefore, the study of simple and effective filters is a popular topic. On this basis, filter gain is introduced to filter noise data from contaminated system data.

The output and observation data (8)-(9), noise data is included in \( y(t) \) and \( \phi(t) \). To extract usefulness data, the variables \( y(t) \) and \( \phi(t) \) are filtered. For achieving this aim, \( y_f(t) \) and \( \phi_f(t) \) are given as follows:

\[ y_f(t) = \frac{\gamma}{\gamma+1} y(t-1) + \frac{1}{\gamma+1} y(t), \quad y_f(0) = 0.001, \]  
\[ \phi_f(t) = \frac{\gamma}{\gamma+1} \phi_f(t-1) + \frac{1}{\gamma+1} \phi_f(t), \quad \phi_f(0) = 0.001, \]  
where the gain of filter is denoted by \( \gamma \).

**Remark 2**: The gain \( \gamma \) of the filter provides a simple structure, and can effectively achieve the data preprocessing by choosing an appropriate parameter.

The identification error is unknown in the process of parameter estimation. The identification error also makes designing the identification algorithm difficult. In order to solve this problem, this paper deduces the identification error expression only according to the available data with batch data. Consequently, the imposed variables \( J(t) \) and \( K(t) \) are defined by

\[ J(t) = \frac{1}{1+\rho(t)} J(t-1) + \frac{1}{1+\rho(t)} \phi_f(t) \phi_f^T(t), \]  
\[ K(t) = \frac{1}{1+\rho(t)} U(t) V^{-1}(t) J(t-1) + \frac{y_f(t) \phi_f^T(t)}{1+\rho(t)}, \]  
\[ U(t) = \phi_f(t) \phi_f(t)^T, \]  
\[ V(t) = \phi_f(t) \phi_f(t)^T, \]  
where \( J(0) = 0.001, K(0) = 0.001, \rho(t) = e^{-\tau t}/(1 + e^{-\tau t^2}), \tau > 0 \) describes an adaptive attenuation coefficient. \( y_f(t) = [y_f(1), \ldots, y_f(g)], \phi_f(t) = [\phi_f(1), \ldots, \phi_f(g)] \), the batch data length \( g \) is \( 0 < g < N \). \( N \) is data length.

**Remark 3**: The adaptive attenuation coefficient \( \rho(t) \) is presented to improve new data utilisation, avoiding the data inundation problem. \( \theta^0 = U(t)V^{-1}(t) \) is developed based on batch data and the variables \( y_f(t) \) and \( \phi_f(t) \) which are used to derive identification error data.

Go a step further, \( H(t) \) is provided as follows:

\[ H(t) = \hat{\theta}^T(t-1) J(t) - K(t), \]  
where the parameter estimation value is represented by \( \hat{\theta}(t) \).

By inserting (13)-(14) into (17), we have

\[ H(t) = \hat{\theta}^T(t-1) J(t) - \theta^T J(t) + \beta(t) \]  
\[ = -\hat{\theta}^T(t-1) J(t) + \beta(t) \]  
where \( \beta(t) = -\omega_f(t) \phi_f^T(t)/(1+\rho(t)) \), the identification error \( \hat{\theta}(t-1) \) is defined by \( \theta(t-1) = \theta - \hat{\theta}(t-1) \).

Based on the derived identification error expression \( H(t) \), a new estimator is given as follows:

\[ \hat{\theta}(t) = \hat{\theta}(t-1) - \Gamma(t) J(t) H(t)^T, \]  
\[ \Gamma(t) = \Gamma(t-1) - \frac{\Gamma(t-1)(t) J(t) \Gamma(t-1)}{I + J^T(t) \Gamma(t-1) J(t)}. \]  

**Remark 4**: \( H(t) \) includes identification error data \( \hat{\theta} \). Thus, we use the developed variables \( J(t) \), \( K(t) \) and \( H(t) \) to obtain the expression of the identification error by using \( y_f(t) \) and \( \phi_f(t) \). This construction method provides a new perspective for using the estimation error to design the identification algorithm. In Eq.(19), a parameter adaptive law is developed using \( H(t) \) rather than the popular predictive error \( \gamma(t) - \hat{\theta}^T \phi_f(t) \). The parameter update law is adjusted in real time based on the identification error itself, which is directly related to parameter estimation.

**Remark 5**: \( \Gamma(t) \) is a recursive variable gain based on system data \( J(t) \) instead of constant gain. It improves correction performance and is easy to run online.

For variable \( x(t) \) that cannot be directly measured, we apply the reference model identification idea [65] to obtain the indirectly measured \( x_{aux}(t) \). Then, \( x(t) \) is replaced using \( x_{aux}(t) \). Thus, unmeasured problem of \( x(t) \) is addressed.
In accordance with the information of the origin system (7), the reference model of $x_{aux}(t)$ is constructed by

$$x_{aux}(t) = x_{aux}(t - 1) + [\hat{\gamma} \rho_0 \text{sgn}(u(t - 1))]$$

$$+ \hat{\gamma} \sum_{i=1}^{n} \hat{\beta}_i (u(t - 1))^{2i-1} - \hat{\gamma} x_{aux}(t - 1)$$

$$\times |u(t) - u(t - 1)| + \sum_{i=1}^{n} \hat{\xi}_{i-1}(u(t - 1))^{2i-2}$$

$$\times (u(t) - u(t - 1)), \quad x_{aux}(0) = 0.001. \quad (21)$$

To increase the readability of the design scheme, the pseudocode of the proposed method is illustrated in Algorithm 1.

**Algorithm 1 Proposed Algorithm**

**Input:** Let $t = 1$, provide the initial values $y(t) = 0.001$ and $\varphi(y) = 0.001$, $\hat{\theta}(0) = I/p0$, $p0 = 10^3$, $x_{aux}(0) = 0.001$, $K(0) = 0.001$. Give a value for the length of total data $N$.

**Result:** $y(t)$, $\varphi(y)$, $\hat{\beta}(t)$, $x_{aux}(t)$, $J(t)$, $K(t)$.

1. **for** $t = 1 : N$ **do**
2. **Sample** the data $u(t)$ and $y(t)$
3. **Construct** the filter data $y(t)$ and $\varphi(y)$ by (11), (12)
4. **Construct** the variable $\hat{\beta}(t)$ by (13), $K(t)$ by (14), and $\hat{\beta}(t)$ by (17)
5. **Construct** the variable $\hat{\theta}(t)$ by (20)
6. **Construct** the variable $x_{aux}(t)$ by (21)
7. **Generate** the parameter estimation $\hat{\theta}(t)$ by (19)
8. Read $\hat{\theta}(t), \hat{\beta}(t), \cdots$ from $\hat{\theta}(t)$
9. **end for**

**IV. CONVERGENCE OF THE PROPOSED SCHEME**

This section introduces the convergence of the estimator based on the persistent excitation condition.

**Theorem 1:** Suppose that $\mathcal{F}_t$ is an algebra sequence which consists of $\beta(t)$, the following noise conditions are satisfied: [66]:

(A1) $E[\beta(t)|\mathcal{F}_{t-1}] = 0$, a.s.

(A2) $E[\beta^2(t)|\mathcal{F}_{t-1}] \leq \sigma_\beta^2 < \infty$ and there exist $\mu, \nu > 0$, such that the following persistent excitation (A3) holds.

(A3) $\mu I_n \leq 1/t \sum_{i=1}^{T} J(i) J^T(i) \leq \nu I_n$

Then, the parameter identification error can converge to zero, i.e.,

$$\lim_{t \to \infty} ||\hat{\theta}(t) - \theta||^2 = 0 \quad (22)$$

**Proof:** According to Eq.(19), it yields

$$\theta - \hat{\theta}(t) = \theta - \hat{\theta}(t - 1) + \Gamma(t) J(t) H^T(t)$$

$$\hat{\theta}(t) = \hat{\theta}(t - 1) + \Gamma(t) J(t) [\hat{J}(t) + \beta(t)], \quad (23)$$

where $\hat{J}(t) = \hat{\beta}^T(t - 1) J(t)$.

By defining $W(t) = \hat{\beta}^T(t) \Gamma^{-1}(t) \hat{\theta}(t)$, by inserting (23) into $W(t)$, it has

$$W(t) = [\hat{\theta}(t - 1) + \Gamma(t) J(t) [\hat{J}(t) + \beta(t)]]^T \Gamma^{-1}(t)$$

$$\times \hat{\theta}(t - 1) + \Gamma(t) J(t) [\hat{J}(t) + \beta(t))] \Gamma^{-1}(t)$$

$$= \hat{\beta}^T(t - 1) \Gamma^{-1}(t) \hat{\theta}(t - 1) + 2 \hat{\beta}^T(t - 1) J(t)$$

$$\times - (J(t) + \beta(t)) + \Gamma^T(t) \Gamma(t) [\hat{J}(t) + \beta(t)]^2$$

$$= W(t - 1) - \hat{J}(t)^2 + 2 \hat{J}(t) \beta(t) + \hat{J}(t) \Gamma(t)$$

$$\times J(t) \beta(t)^2$$

$$- 2 J(t) \Gamma(t) \beta(t)$$

$$= W(t - 1) - [1 - J(t) \Gamma(t) \beta(t)] \hat{J}(t)^2$$

$$+ 2 [1 - J(t) \Gamma(t) \beta(t)] \hat{J}(t) \beta(t) + J(t) \Gamma(t)$$

$$\times J(t) \beta(t)^2$$.

(24)

Based on matrix inversion theory, it obtains

$$1 - J(t) \Gamma(t) \beta(t) = \frac{1}{1 + J(t) \Gamma(t) \beta(t)} > 0. \quad (25)$$

Then, (24) can be written as

$$W(t) \leq W(t - 1) + J(t) \Gamma(t) \beta(t)^2$$

$$+ \hat{J}(t)^2 + 2 \hat{J}(t) \beta(t) + J(t) \Gamma(t)$$

$$\times J(t) \beta(t)^2$$.

(26)

Since $\hat{J}(t)$ and $J(t) \Gamma(t) \beta(t)$ are uncorrelated with $\beta(t)$, and using the theorem of martingale convergence, (A1) and (A2) to (26), we obtain

$$E[W(t)|\mathcal{F}_{t-1}] \leq W(t - 1) + J(t) \Gamma(t) \beta(t)^2$$

$$\leq E[S(t)|\mathcal{F}_{t-1}] \leq S(t - 1) + J(t) \Gamma(t) \beta(t)^2$$

$$\leq E[S(t)|\mathcal{F}_{t-1}] \leq S(t - 1) + J(t) \Gamma(t) \beta(t)^2$$

$$\leq E[S(t)|\mathcal{F}_{t-1}] \leq S(t - 1) + J(t) \Gamma(t) \beta(t)^2$$

(28)

Applying the martingale theorem to (28), we known that $S(t)$ converges to $S_0$, i.e.,

$$S(t) = \frac{W(t)}{[\ln \Gamma(t)]^p} \to S_0 < \infty, \quad a.s. \quad (29)$$

where $S_0$ is a finite random variable, or

$$W(t) \leq \eta [\ln \Gamma(t - 1)]^p, \quad a.s., \quad t \to \infty, \quad (30)$$

where $\eta$ is a large random variable. According to $W(t)$, $\hat{\theta}(t)$ yields

$$\hat{\theta}(t)^2 \leq \frac{\eta [\ln \Gamma(t - 1)]^p}{\lambda_{\min} \Gamma(t - 1)} \leq \frac{\eta [\ln \Gamma(t - 1)]^p}{\lambda_{\min} \Gamma(t - 1)}$$

(31)

Based on (A3), we have

$$\lambda_{\min} \Gamma(t - 1) \leq \nu t + n \Gamma(t - 1) \beta(t)$$

$$\lambda_{\min} \Gamma(t - 1) \geq \mu t$$

(32)

where $\Gamma(0)$ represents the initial state of $\Gamma(0)$. According to (32)-(33), (31) gives

$$\lim_{t \to \infty} \frac{\hat{\theta}(t) - \theta}{\mu t} = \frac{\varphi(n \ln(nvt + n \Gamma^{-1}(0))}{\mu t} = 0, \quad a.s.$$
V. SIMULATION AND EXPERIMENT

This section provides the simulation and experiment results with the developed method and some existing identification methods. Before the simulation and experiment, we give some guidelines on key parameters.

1. The initial value of the estimated value \( \hat{\theta}(t) \) is set to small value, e.g., 0,1,0.01,0.001 based on the literature [67].
2. The filter gain can affect the outcome, since the noise frequency is unknown, we use the trial-and-error manner to chosen the filter gain values.
3. The initial value of the adaptive attenuation coefficient is chosen as 0 ∼ 1 according to the selection criteria of the forgetting factor method.
4. The selection criterion of parameter \( g \) is greater than the number of estimated parameters.

A. SIMULATION EXAMPLE

Example 1: Consider a Hammerstein system with Duhem hysteresis as follows:

Duhem Model: \( x(t) = 0.1(u(t) - u(t - 1)) + x(t - 1) + 0.1x(t - 1)[u(t) - u(t - 1)] - 0.2x(t - 1)[u(t) - u(t - 1)], L : y(t) = x(t - 1) + 0.8x(t - 2) - 0.5y(t - 1) - 0.35y(t - 2) + \omega(t). \)

The input signal is a random sequence with a zero mean and a unit variable. Additional noise is a white noise.

The common least-squares based reference model (LS-RM) in [20] and stochastic gradient scheme based auxiliary model (SG-AM) [68] are selected as two comparison algorithms that are designed based on popular prediction error data. The initial values of the three identification methods are listed as follows:

1. Based on guidances, the parameters applied of the presented estimator in the numerical example are \( \gamma = 4, \rho(0) = 0.85, \tau = 2, \Gamma(0) = 100 \ast diag([0.01, 0.01, 0.01, 0.01, 0.01, 0.01, 0.01, 0.01])^T, \hat{\theta}(0) = I/p0, p0 = 10^3, x_{aux}(0) = 0.001. \)
2. The initial values of the LS-RM are \( \hat{\theta}(0) = I/p0, p0 = 10^3, x_{aux}(0) = 0.001, P = eye(8) \ast p0 \) based on the selection criteria of the least-squares method.
3. The initial values of the SG-AM are \( \hat{\theta}(0) = I/p0, p0 = 10^3, x_{aux}(0) = 0.001, r = 1 \) based on the selection criteria of the stochastic gradient method.

Figs.2-3 plots the evolutionary curves of parameter identification with the proposed parameter adaptive identifier (11)-(20), LS-RM and SG-AM methods. One can be found that during early parameter estimation time, the estimation curves show a certain number of oscillation, and the estimated parameters finally tend to be near the expected values. The identification error results shown in Fig.4 indicate that as data are added, faster convergence speed and higher identification precision can be provided by the proposed adaptive estimator compared with LS-RM and SG-AM because identification error data are used to modify the parameter adaptive law instead of prediction error data.

The identification nature of the proposed parameter adaptive law is then assessed using model verification. In Fig.5, the model prediction output together with the real output is provided. It can be observed that the model output with the developed adaptive law can describe system dynamic characteristic with a small tracking error. The constancy results of the proposed scheme based on 100 monte carlo experiments are displayed in Fig.6. In the implementation process, the parameter identification error curves exhibits a small fluctuation. The model verification and monte carlo experiment results demonstrate the effectiveness and stability of the estimator presented in this study.

Based on the multiplication and addition float numbers, we can evaluate the computational efficiency of the parameter.
estimation algorithm [69]. Table 1 provides the calculation efficiency of the three estimation algorithms with \( n = 6 \), \( g = 6 \). It can be observed from Table 1 that LS-RM produces higher computational burden than SG-AM due to the computing covariance matrix. The developed algorithm gives slightly higher floating numbers than LS-RM owing to the computation problem of \( K(t) \) and \( T(t) \). In accordance with the estimation profiles shown in Figs. 2-3 and the identification error profiles plotted in Fig. 4, it can be found that the developed method sacrifices computational burden to achieve higher identification performance.

Example 2: To test the effectiveness of the developed estimator further, a Wiener-like system with hysteresis is considered [70].

\[
L: y(t) = x(t - 1) + 1.2x(t - 2) - 0.32y(t - 1) - 0.7y(t - 2) + \omega(t).
\]

Duhem Model: \( x(t) = 0.001(u(t) - u(t - 1)) + x(t - 1) + 0.002x(t - 1)|u(t) - u(t - 1)| - 0.0001u(t - 1)|u(t) - u(t - 1)| \),

The parameter estimation curves are shown in Figs. 7-8. It can be found that the parameter estimation can be effectively achieved by the three estimators, but the developed method provides faster convergence speed due to the addition of the adaptive attenuation coefficient. It also offers higher estimation accuracy because the adaptive law is updated using the estimation error itself. Fig. 9 plots the parameter identification error profiles obtained by the considered estimator. It can be seen that the estimation error decreases with an increase in data, indicating that the three methods are effective. The proposed algorithm achieves fast convergence performance and high estimation accuracy, suggesting that exhibits better performance than the two other methods. In Fig. 10, the model validation results are described. One can be observed that the developed scheme can achieve better
tracking output results with the smallest tracking output error compared to the SG-AM and LS-RM methods, demonstrating the advantage of the developed method.

**B. EXPERIMENT**

A servo drive system is chosen as a verification platform because the Hammerstein system with hysteresis can be used to represent the servo drive system, as plotted in Fig.11. Analog/digital (A/D) and digital/analog (D/A) converters are used to connect the personal computer with visual C++ compiler and digital signal processor (DSP) (TMS320F28335), the drive motor (HC-PQ43BK) is applied to drive the mechanical arm platform.

The servo drive system is expressed as the following mathematical equation [71]

$$\begin{align*}
\dot{x}_1 &= x_2 \\
\dot{x}_2 &= \frac{1}{J} (-K_2 x_2 + K_1 u - T_f - T_l) \\
y &= x_1
\end{align*}$$

where $\theta_i = [K_2/J, K_1/J, T_c/J, B/J]^T$, $i = 1, 2, 3, 4$. $x = [x_1, x_2]^T = [q, \dot{q}]^T$. The angular position and velocity are described by $q$ and $\dot{q}$, respectively. The inductance, armature current and resistance are defined as $L_a$, $I_a$, $R_a$. The payload load, torque and friction force are denoted by $T_l$, $T_m$, $T_f$, separately. The motor-inertia and system input are $J$ and $u$, separately. $K_T$ describes electro-mechanical constant; $K_E$ represents the backelectromotive constant.

Parameter identification is conducted using the developed scheme, LS-RM and SG-AM on the basis of the given signal $y_r = 0.8 \sin(1/5\pi t)$. The experimental processes are described in Fig.12, which presents the results of the parameter identification. From these results, we can observe that the smooth curve and faster convergence rate can be obtained using the developed method based on identification error data. For LS-RM and SG-AM, however, large oscillation exists, during in the initial stage of parameter estimation. Such oscillation originates from the prediction error data. Fig.13 plots the track output and error based three estimators. As shown in Fig.13, it is obvious that a smaller tracking error and smooth tracking output can be achieved by using the presented new identification framework. Although, LS-RM
and SG-AM can also represent the dynamic output responses of the system, large track errors and the vibration phenomena occur.

Table 2 presents the identification performance evaluation by using the following indicators. (1) Mean Squared Error (MSE): \( MSE = \frac{1}{Var^2} \sum_{t=1}^{L} e^2(t) \), where \( e(t) = y(t) - \hat{y}(t) \); (2) Prediction Error Mean (PEM): \( PEM = \frac{1}{L} \sum_{t=1}^{L} e(t) \), where \( L \) represents the sample length, \( Var^2 \) denotes the output variance. A small indicator shows that the three identification schemes can estimate servo drive system parameters, demonstrating the usefulness of the considered methods. Among these indicators, the indicator of the proposed method is the smallest, signifying that the estimation model established by the proposed scheme is the best. Therefore, the developed method exhibits better estimation performance than the SG-AM and LS-RM schemes.

VI. CONCLUSION
A new robust identification algorithm for the Hammerstein-like system has been reported in this study, providing a new identification scheme framework for estimator design. The primary idea is to use identification error data to develop a parameter estimation adaptive law, in which a filter technique is proposed to filter noise data. Then, some forcing variables are introduced to obtain identification error data from system data. A novel parameter estimation adaptive law is derived on the basis of the identification error data. The modified gain is improved in this law by using a recursive gain form. Hence, this novel method is distinct from traditional identification methods based on prediction errors. Quantitative and qualitative analysis results demonstrated the advantage of the proposed method by using the simulation and experiment. In the experiment, we can choose a small filter gain for the high-frequency disturbance signal to reduce noise interference. For the model error, we select the estimation model with the smallest output error as the final model to remedy the adverse effect of model error. The limitation of the proposed method in practical applications is adjusting the parameters of filter gain and the adaptive attenuation coefficient via trial-and-error manner. Future work will use to other nonlinear systems to test the availability of the proposed method.
and to adjust the aforementioned parameters by using the relationship between estimation error value and parameters. In the further, more complex nonlinear systems (e.g., time-varying nonlinear systems, multivariable systems) are used to verify the effectiveness of the proposed algorithm, and fractional operator theory is used to improve the performance of the design algorithm.

VII. DISCUSSION

In simulation examples, parameter estimation and identification result displayed that the faster convergence rate and high estimation accuracy. While the parameter estimation result displayed that the faster convergence rate can be obtained by the proposed method. Meanwhile, the good estimation accuracy is produced by using the tracking output and performance indicators.

REFERENCES

[1] F. Giri and E.-W. Bai, Block-Oriented Nonlinear System Identification, Berlin, Germany: Springer, 2010.

[2] G. Savaia, G. Panzani, M. Corno, J. Cecconi, and S. M. Savaresi, “Hammerstein–Wiener modelling of a magneto-rheological dampers considering the magnetization dynamics,” Control Eng. Pract., vol. 112, Jul. 2021, Art. no. 104829.

[3] D. Wang, S. Zhang, M. Gan, and J. Qiu, “A novel EM identification method for Hammerstein systems with missing output data,” IEEE Trans. Ind. Informat., vol. 16, no. 4, pp. 2500–2508, Apr. 2020.

[4] V. Prasad and U. Mehta, “Modeling and parametric identification of Hammerstein systems with time delay and asymmetric dead-zones using fractional differential equations,” Mech. Syst. Signal Process., vol. 167, Mar. 2022, Art. no. 108568.

[5] K.-K. Xu, H.-D. Yang, and C.-J. Zhu, “A novel extreme learning machine-based Hammerstein-Wiener model for complex nonlinear industrial processes,” Neurocomputing, vol. 358, pp. 246–254, Sep. 2019.

[6] J. Lin, G. Liao, M. Chen, and H. Yin, “Two-phase degradation modeling and remaining useful life prediction using nonlinear Wiener process,” Comput. Ind. Eng., vol. 160, Oct. 2021, Art. no. 107533.

[7] S. Dong, L. Yu, W.-A. Zhang, and B. Chen, “Robust extended recursive least squares identification algorithm for Hammerstein systems with dynamic disturbances,” Digit. Signal Process., vol. 101, Jun. 2020, Art. no. 102716.

[8] B. Balasangiam and K. R. Pattipati, “On the identification of electrical equivalent circuit models based on noisy measurements,” IEEE Trans. Instrum. Meas., vol. 70, pp. 1–16, 2021.

[9] Y. Mao, F. Ding, L. Xu, and T. Hayat, “Highly efficient parameter estimation algorithms for Hammerstein non-linear systems,” IET Control Theory Appl., vol. 13, no. 4, pp. 477–485, Mar. 2019.

[10] Z. Lu, N. Wang, and C. Yang, “A novel iterative identification based on the optimised topology for common state monitoring in wireless sensor networks,” Int. J. Syst. Sci., vol. 53, no. 1, pp. 25–39, Jan. 2022.

[11] Y. Pu, Y. Yang, and J. Chen, “Maximum likelihood iterative algorithm for Hammerstein systems with hard nonlinearities,” Int. J. Control, Autom. Syst., vol. 18, no. 11, pp. 2879–2899, Nov. 2020.

[12] N. Sammaknejad, Y. Zhao, and B. Huang, “A review of the expectation maximization algorithm in data-driven process identification,” J. Process Control, vol. 73, pp. 123–136, Jan. 2019.

[13] E. Sobhani Tehrani and R. E. Keanney, “A non-parametric approach for identification of parameter varying Hammerstein systems,” IEEE Access, vol. 10, pp. 6348–6362, 2022.

[14] J. Li and J. Zhang, “Maximum likelihood identification of dual-rate Hammerstein output-error moving average system,” IET Control Theory Appl., vol. 14, no. 8, pp. 1089–1101, May 2020.

[15] J. Hou, F. Chen, P. Li, and Z. Zhu, “Gray-box parsimonious subspace identification of Hammerstein-type systems,” IEEE Trans. Ind. Electron., vol. 68, no. 10, pp. 9941–9951, Oct. 2021.
[38] A. Louis, G. Ledwich, G. Walker, and Y. Mishra, “Measurement sensitivity and estimation error in distribution system state estimation using augmented complex Kalman filter,” J. Modern Power Syst. Clean Energy, vol. 8, no. 4, pp. 657–668, 2020.

[39] H. Xu, F. Ma, F. Ding, L. Xu, A. Alsaedi, and T. Hayat, “Data filtering-based recursive identification for an exponential autoregressive moving average model by using the multi-innovation theory,” IET Control Theory Appl., vol. 14, no. 17, pp. 2526–2534, Nov. 2020.

[40] M. Setareh, M. Parniani, and F. Aminifar, “Non-stationary stabilized fast transversal RLS filter for online power system modal estimation,” IEEE Trans. Power Syst., vol. 34, no. 4, pp. 2744–2754, Jul. 2019.

[41] B. Liu, G. Gui, S.-Y. Matsushita, and L. Xu, “Adaptive filtering algorithm for direction-of-arrival (DOA) estimation with small snapshots,” Digit. Signal Process., vol. 94, pp. 84–95, Nov. 2019.

[42] X. Zhang, S. He, V. Stoianovic, X. Luan, and F. Liu, “Finite-time asynchronous dissipative filtering of conic-type nonlinear Markov jump systems,” Sci. China Inf. Sci., vol. 64, no. 5, May 2021, Art. no. 152206.

[43] M. Li and X. Liu, “Maximum likelihood least squares based iterative estimation for a class of bilinear systems using the data filtering technique,” Int. J. Control, Autom. Syst., vol. 18, no. 6, pp. 1581–1592, Jun. 2020.

[44] N. Subrahmanya and Y. C. Shin, “Adaptive divided difference filtering for simultaneous state and parameter estimation,” Automatica, vol. 45, no. 7, pp. 1686–1693, 2009.

[45] V. Filipovic, N. Nedic, and V. Stoianovic, “Robust identification of pneumatic servo actuators in the real situations,” Forschung Ingenieurwesen, vol. 75, no. 4, pp. 183–196, Dec. 2011.

[46] C. Wang, Z. Wang, L. Zhang, D. Cao, and D. G. Dorrell, “A vehicle rollover evaluation system based on enabling state and parameter estimation,” IEEE Trans. Ind. Informat., vol. 17, no. 6, pp. 4003–4013, Jun. 2021.

[47] J. Zhao and L. Mili, “A decentralized H-infinity unscented Kalman filter for dynamic state estimation against uncertainties,” IEEE Trans. Smart Grid, vol. 10, no. 5, pp. 4870–4880, Sep. 2018.

[48] H. Ben Sassi, F. Errahimi, and N. ES-Sbai, “State of charge estimation of lithium-ion batteries with potential and SOC estimation,” IEEE Trans. Energy Storage, vol. 3, no. 1, pp. 1427–1434, Jun. 2017.

[49] J. Zambrano, J. Sanchis, J. M. Herrero, and M. Martinez, “WH-MOEA: A multi-objective evolutionary algorithm for wiener-hammerstein system identification. A novel approach for trade-off analysis between complexity and accuracy,” IEEE Access, vol. 8, pp. 228655–228674, 2020.

[50] L. Xu, F. Ding, X. Lu, L. Wan, and J. Sheng, “Hierarchical multi-innovation generalised extended stochastic gradient methods for multivariable equation-error autoregressive moving average systems,” IET Control Theory Appl., vol. 14, no. 10, pp. 1276–1286, Jul. 2020.

[51] F. Ding, P. X. Liu, and G. Liu, “Multiinnovation least-squares identification for system modeling,” IEEE Trans. Syst., Man, Cybern., B, Cybern., vol. 40, no. 3, pp. 767–778, Jun. 2010.

[52] G. C. Goodwin and K. S. Sin, “Adaptive Filtering Prediction and Control,” Upper Saddle River, NJ, USA: Prentice-Hall, 1984.

[53] L. Ljung, “System Identification: Theory for the User,” 2nd ed., Upper Saddle River, NJ, USA: Prentice-Hall, 1999.

[54] Y. Pu, Y. Yang, and J. Chen, “Some stochastic gradient algorithms for Hammerstein systems with piecewise linearity,” Circuits, Syst., Signal Process., vol. 40, no. 4, pp. 1635–1651, Apr. 2021.

[55] F. Ding and T. Chen, “Hierarchical gradient-based identification of multi-variable discrete-time systems,” Automatica, vol. 41, no. 2, pp. 315–325, Feb. 2005.

[56] S. Gupta, A. K. Sahoo, and U. K. Sahoo, “Wireless sensor network-based distributed approach to identify spatio-temporal Volterra model for industrial distributed parameter systems,” IEEE Trans. Ind. Informat., vol. 16, no. 12, pp. 7671–7681, Dec. 2020.

[57] G. Zhang, J. Chen, and Z. Li, “Identifier-based adaptive robust control for servomechanisms with improved transient performance,” IEEE Trans. Ind. Electron., vol. 57, no. 7, pp. 2536–2547, Jul. 2010. ** **