Journal of Industrial and Management Optimization

https://www.aimspress.com/journal/jimo

JIMO, 22(6): 3023–3049.
DOI:10.3934/jimo.2026111
Received: 10 February 2026
Revised: 12 May 2026
Accepted: 15 May 2026
Published: 28 May 2026

Research article

An adaptive Levenberg-Marquardt identification algorithm for time-delay linear discrete periodic systems

Zhuo Chen[1], Yingying Guo[2], Zhuocheng Zou[3] and Zhen Zhang[4,*]

[1] School of Management and Economics, North China University of Water Resources and Electric Power, Zhengzhou 450046, China

[2] School of Materials Science and Engineering, North China University of Water Resources and Electric Power, Zhengzhou 450046, China

[3] Zhejiang Shenjia Technology Co., Ltd. No. 133 Wuchang Road, Hangzhou, Zhejiang 310023, China

[4] Taizhou Yinxin Zhishu Technology Co., Ltd. No. 383, Shide Road, Fuxi Street, Tiantai, Zhejiang 317200, China

* **Correspondence:** Email: zhangzhen@yinlun.cn.

Abstract: This paper proposes an adaptive auxiliary-model-based Levenberg-Marquardt (AM-ILM) algorithm for parameter identification of linear discrete-time periodic systems with time delays. Unlike existing recursive least squares (AM-RLS) based methods that assume slowly varying parameters, the proposed LM framework dynamically balances gradient descent and Gauss-Newton steps via an adaptive regularization strategy, enabling accurate tracking of periodic variations. An auxiliary model reconstructs unmeasurable intermediate variables from input-output data, addressing the non-causal structure caused by time delays. The algorithm introduces two innovations: an adaptive regularization that adjusts the damping parameter, and an auxiliary model that reconstructs latent variables. By replacing unknown variables with auxiliary models and designing an adaptive updating strategy, an improved LM algorithm is developed. Numerical examples demonstrate its effectiveness, with comparisons to LM, AM-RLS, RLS, fully connected neural network, model-agnostic meta-learning, and Bayesian algorithms. Identification results under varying parameters and initial conditions, along with an application study, are provided. The results show that the proposed algorithm effectively handles the coupling between time delay and periodic variations, achieving superior accuracy and stability.

Keywords: periodic system identification; time-delay periodic systems; auxiliary models; adaptive LM algorithm

Mathematics Subject Classification: 93B30, 93C55, 65K10, 39Axx

1. Introduction

In modern control theory, modeling and identification of complex dynamical systems continue to pose fundamental challenges. The unique dynamic behaviors of systems exhibiting both time delays and periodic variations [1–3] further complicate the identification process. Time delay phenomena, prevalent in industrial systems [4, 5], often lead to performance degradation or even system instability. When time-delay characteristics are combined with periodic dynamics, the resulting complex behaviors frequently exceed the capabilities of traditional analytical methods [6, 7]. The identification of such systems is particularly challenging for two main reasons. On one hand, time delay introduces a non-causal dependence in the input-output relationship, as the current output depends on past inputs shifted by an unknown or time-varying delay. This non-causal structure complicates the formulation of the information matrix and makes conventional identification methods, which assume causal relationships, difficult to apply. On the other hand, periodic variations in system parameters imply that the underlying dynamics change periodically over time, requiring identification algorithms to track time-varying parameters while maintaining convergence stability. When these two features coexist, they interact in complex ways—for instance, the delay can induce periodic solutions or stability switches—making the identification problem substantially more difficult than the case where only one of these features is present. Regarding the identification of time-delay systems, a relatively mature theoretical framework has been established in the academic community [8–10]. For systems with constant time delays, well-developed theories and applications exist for both open-loop and closed-loop identification, with methods broadly categorized into time-domain and frequency-domain approaches. Time-domain methods are typically based on the prediction error minimization criterion, achieving joint estimation by constructing an augmented information matrix that incorporates the delay parameter. Frequency-domain methods, on the other hand, exploit the linear phase characteristic introduced by the time delay in the transfer function, extracting delay information through frequency response analysis. However, these methods generally assume the system parameters to be time-invariant, making them difficult to generalize directly to periodically time-varying scenarios.

Among existing approaches, the Levenberg-Marquardt (LM) algorithm has emerged as a particularly effective framework for solving nonlinear least squares problems in system identification [11]. Originally proposed as a damped least-squares method, the LM algorithm uniquely combines gradient descent and Gauss-Newton iteration, exhibiting superior convergence properties when dealing with ill-conditioned Jacobian matrices—a common challenge in time delay periodic system identification. As a classical optimization technique, the LM algorithm has found widespread application in numerous scientific and engineering fields. In terms of solving nonlinear problems, Li et al. enhanced the LM algorithm by incorporating normalized cross-correlation, thereby improving breast image registration accuracy for early cancer screening [12]. Tong et al. refined the LM-Kaczmarz method through nonsmooth convex constraints and adaptive parameter selection, enabling efficient solution of multiple nonlinear ill-posed equations [13]. Tang and Zhou further proposed a modified LM algorithm with Broyden-like updates, ensuring both global convergence and local quadratic convergence under certain conditions [14]. In the field of neural network training, Mirza et al. applied Levenberg-Marquardt backpropagation neural networks (LMB-NN) to analyze activation energy in tri-hybrid nanofluid two-phase flow [15]. Qureshi et al. employed Levenberg-Marquardt neural networks (LMA-NN) to study reactive radiative nanofluid flow, demonstrating the LM algorithm's enhanced convergence and predictive accuracy [16]. Zahoor

Shah et al. analyzed chemically reactive Carreau nanoflow using stochastic analysis on Levenberg Marquardt backpropagation neural networks (SALMBNNs) [17], while Albasheir et al. studied magneto-thermal effects on Carreau stagnation flow [18]. Both studies leveraged LM optimization to reformulate fluid dynamics problems as neural network tasks, thereby advancing computational methodologies. In parameter estimation, Nautiyal et al. employed the LM method for photovoltaic module parameter identification, simplifying the five-parameter single-diode model to enhance maximum power point tracking (MPPT) accuracy [19]. Liu and Chen developed an LM-based simultaneous estimator for time delays and system parameters, utilizing finite element discretization and error mitigation strategies [20].These studies demonstrate that the LM algorithm can significantly enhance the precision and efficiency of parameter estimation, providing robust support for system modeling and analysis. To further improve the performance of the LM algorithm, researchers have introduced various modifications and innovations. Zhang et al. developed a frozen LM-Kaczmarz method with Nesterov acceleration and convex penalties, maintaining initial Fréchet derivatives to enhance computational efficiency and accuracy for ill-posed problems [21]. Boos et al. proposed an LM variant with a singular scaling matrix, achieving local quadratic convergence and ensuring the existence of stationary points through globalization techniques [22]. These advancements have expanded the applicability of the LM algorithm to a broader range of fields.

For the identification of periodic systems, traditional strategies typically employ lifting techniques to transform periodically time-varying systems into high-dimensional time-invariant systems for processing [23, 24]. While this approach preserves the periodic structure information of the system, the computational complexity increases exponentially with the period length, and it handles time delays in a rather unnatural manner. In recent years, auxiliary model-based methods have enabled the transformation of complex problems into more tractable forms, thereby enhancing the feasibility of numerical solution and analysis of models [25]. Hierarchical identification methods [26] further decompose large-scale systems into lower-dimensional subsystems, an idea that shares similarity with the lifting technique for periodic systems. Moreover, two-stage parameter estimation methods for continuous-time systems have also been developed [27]. Nevertheless, research that simultaneously incorporates both time delays and periodicity into a unified identification framework remains relatively limited. Additionally, when the system is multivariable with partially-coupled information vectors, existing recursive identification methods face further difficulties [28]. Recently, a recursive identification algorithm for discrete time-delay periodic linear systems was proposed using auxiliary models and recursive least squares (RLS) [29]. While the RLS-based method provides a baseline for this class of systems, it assumes the system parameters are slowly time-varying and cannot efficiently handle rapid periodic fluctuations or ill-conditioned Jacobians that arise due to the coupling between delay and periodicity. The RLS algorithm minimizes a quadratic cost using a constant forgetting factor, which fails to adjust the step size according to local curvature. Like other existing works, it either focuses on time-delay identification while neglecting time-varying parameters, or addresses periodicity while avoiding the non-causality introduced by time delays. The complex dynamic behaviors arising from the coupling of these two factors—such as delay-induced periodic solutions, stability switches, and resonance phenomena—impose more stringent requirements on the design of identification algorithms. Therefore, developing identification methods capable of handling both time delays and periodicity is an urgent need for engineering applications.

Despite the widespread application of the Levenberg-Marquardt algorithm in nonlinear system identification, several fundamental limitations hinder its direct application to time-delay periodic

systems. First, the presence of time delay introduces a non-causal structure in the information vector, making the conventional Jacobian formulation ill-posed when the delay is unknown or varies periodically. Second, existing LM implementations assume time-invariant parameters and rely on fixed regularization strategies, which cannot capture the intrinsic time-varying nature of periodic systems and often lead to slow convergence or divergence when parameter variations occur over time. Third, the auxiliary variables required for constructing the information matrix are unmeasurable in practical scenarios, yet existing LM-based methods lack an effective mechanism to recover these latent states in a recursively consistent manner. These limitations motivate the development of an enhanced LM framework that jointly addresses time delay, periodicity, and latent variable estimation within a unified identification structure. In this context, this paper proposes an improved Levenberg-Marquardt algorithm based on auxiliary models to address the parameter identification problem of linear discrete-time periodic systems with time delays under the least squares framework. The proposed algorithm incorporates an adaptive updating strategy that not only improves the accuracy of parameter identification but also enhances the algorithm's robustness when dealing with complex dynamic systems. This approach provides a novel methodology for handling periodic systems with time delays, offering both theoretical and practical advancements in the field. The main contributions of this paper are twofold. First, we devise a damping parameter update rule (Eq. (3.13)) that depends on the prediction error norm, allowing the algorithm to automatically shift between gradient descent (when error is large or curvature changes rapidly) and Gauss-Newton (when near a minimum). This is in sharp contrast to fixed or heuristically tuned damping in conventional LM. Second, we develop an auxiliary model that generates surrogate signals for unmeasurable intermediate variables without requiring state measurements. This integration resolves the non-causality caused by time delays and enables recursive estimation. While existing works introduce various modifications to the LM algorithm, none simultaneously address the coupling between time delay and periodic parameter variations. The combination of these two innovations enables a unified treatment of these challenging features—a capability that existing LM methods lack. Numerical comparisons demonstrate that the proposed LM approach achieves faster convergence and lower steady-state error than the existing RLS+auxiliary model method.

The remainder of this paper is structured as follows. In Section 2, we present a detailed formulation of the state-space model for linear discrete-time periodic systems with time delays, analyzing their distinctive characteristics and inherent complexities. This section also elucidates our identification methodology based on measurable input-output data. Section 3 introduces the proposed auxiliary model-based improved Levenberg-Marquardt algorithm, with comprehensive derivation of its theoretical framework. Section 4 validates the correctness and effectiveness of the proposed algorithm through carefully designed numerical case studies, demonstrating its practical performance in real-world applications. Finally, we conclude the paper with a summary of key findings and contributions.

Notation 1 In the upcoming discussion, T represents the matrix transpose, I represents the unit matrix with appropriate dimensions, t indicates the time index, σ is the transfer operator of $\nu(t)$ with actions: $\sigma\nu(t) = \nu(t+1)$, $\sigma^{-1}\nu(t) = \nu(t-1)$, and $\|\bullet\|$ represents the L_2 norm.

2. The system model formulation

Consider the following state-space model for linear discrete periodic systems with d-unit time delays:

$$\begin{cases} x(t+1) = A(t)x(t) + B(t)u(t-d), \\ y(t) = C(t)x(t) + v(t), \end{cases} \tag{2.1}$$

$$A(t) := \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 1 \\ -a_n(t) & -a_{n-1}(t) & -a_{n-2}(t) & \cdots & -a_1(t) \end{bmatrix},$$

$$B(t) := \begin{bmatrix} b_n(t) & b_{n-1}(t) & b_{n-2}(t) & \cdots & b_1(t) \end{bmatrix}^{\mathrm{T}},$$

$$C(t) := \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \end{bmatrix},$$

where $x(t) \in \mathbb{R}^n$ is the state vector, $u(t) \in \mathbb{R}^r$ is the input vector, $u(t-d) \in \mathbb{R}^r$ represents the d-unit time delay of system input, $y(t) \in \mathbb{R}^m$ is the output vector, $v(t) \in \mathbb{R}$ stands for the uncertain random noise, and $A(t) \in \mathbb{R}^{n\times n}$, $B(t) \in \mathbb{R}^{n\times r}$, and $C(t) \in \mathbb{R}^{1\times n}$ are periodic system parameter matrices with period T.

Remark 2.1. Although the time delay d is assumed known, simply shifting the input by d does not trivialize the identification problem in the presence of periodic parameter variations. The shifted input remains coupled with the unknown and periodically time-varying matrix $B(t)$ in the state equation. Moreover, the delay affects the order of the numerator polynomial in the input-output model, thereby influencing the structure of the information matrix.

Assume that n is known and the input satisfies $u(t) = 0$ for all $t \le 0$. The identification of time delay linear discrete periodic systems with uncertain parameters and unmeasurable states in (2.1) poses significant challenges. To circumvent the state measurement limitation, we adopt an input-output data-driven identification framework. This approach leverages measurable excitation inputs and corresponding system outputs, effectively bypassing the need for direct state observation while preserving identifiability conditions.

Lemma 2.1. *For the linear discrete periodic system* (2.1) *with period T and time delay d, there exist constant coefficients $\varsigma_{i,j}$ and time-varying coefficients $\varpi_{i,k}(t)$ such that its transfer function can be decomposed into T time-invariant sampled transfer functions $H_i(z,t)$, i.e.,*

$$G(\sigma, t) = \sum_{i=0}^{T-1} H_i(\sigma^T, t)\sigma^{-i}, \tag{2.2}$$

where $i \in \{0, \ldots, T-1\}$, and each $H_i(z,t)$ is a rational function in z, i.e.,

$$H_i(z,t) = \frac{\sum_{k=0}^{n+d-1} \varpi_{i,k}(t) z^{-k}}{1 + \sum_{j=1}^{n} \varsigma_{i,j} z^{-j}}. \tag{2.3}$$

Proof. Based on the periodic system theory in [30], the input-output relationship of system (2.1) can be expressed as

$$y(t) = \sum_{j=0}^{\infty} M_j(t-d)u(t-j-d) + v(t), \tag{2.4}$$

where the Markov coefficients $M_j(t)$ satisfy the periodicity condition: $M_j(t+T) = M_j(t)$.

Introducing the shift operator σ, the periodic transfer operator for system (2.1) is given by

$$G(\sigma, t) = \sum_{j=0}^{\infty} M_j(t-d)\sigma^{-j}. \tag{2.5}$$

By exploiting the periodicity, we decompose the summation index modulo T as $j = mT + i$, where $m \geq 0$ and $i \in \{0, \ldots, T-1\}$. Substituting into (2.5) yields

$$G(\sigma, t) = \sum_{i=0}^{T-1}\sum_{m=0}^{\infty} M_{mT+i}(t-d)\sigma^{-(mT+i)} = \sum_{i=0}^{T-1}\Big[\sum_{m=0}^{\infty} M_{mT+i}(t-d)(\sigma^T)^{-m}\Big]\sigma^{-i}. \tag{2.6}$$

Define the sampled transfer functions as

$$H_i(z, t) = \sum_{m=0}^{\infty} M_{mT+i}(t-d)z^{-m}, \quad i = 0, \ldots, T-1. \tag{2.7}$$

Then we obtain the compact decomposition

$$G(\sigma, t) = \sum_{i=0}^{T-1} H_i(\sigma^T, t)\sigma^{-i}. \tag{2.8}$$

This is (2.2).

For rational periodic systems, the Markov coefficients $M_{mT+i}(t)$ decay exponentially with m, hence each $H_i(z, t)$ can be expressed as a ratio of two polynomials. That is, there exist constants $\varsigma_{i,j}$ and periodic functions $\varpi_{i,k}(t)$ such that

$$H_i(z, t) = \frac{\sum_{k=0}^{n+d-1} \varpi_{i,k}(t)z^{-k}}{1 + \sum_{j=1}^{n} \varsigma_{i,j}z^{-j}}. \tag{2.9}$$

This completes the proof of Lemma 2.1. □

Lemma 2.2. *For the linear discrete periodic system* (2.1) *with period T and time delay d, there exist time-invariant coefficients $\varsigma_1, \ldots, \varsigma_p$ and periodically time-varying coefficients $\varpi_0(t), \ldots, \varpi_q(t)$ such that its input-output relation can be written as*

$$y(t) = \sum_{k=1}^{p} \varsigma_k\, y(t-k\tau) + \sum_{j=0}^{q} \varpi_j(t)\, u(t-j-d) + v(t), \tag{2.10}$$

where τ is a multiple of the period T.

Proof. Since (2.4) is realized in a finite-dimensional state space realization (i.e., (2.1)), its transfer operator can be written as

$$G(\sigma, t) = [d(\sigma, t)]^{-1} n(\sigma, t), \tag{2.11}$$

where $d(\sigma, t)$ and $n(\sigma, t)$ are polynomials in σ with T-periodic coefficients. Also,

$$\begin{aligned} d(\sigma, t) &= \sigma^{n+d-1} + a_1(t)\sigma^{n+d-2} + \cdots + a_{n+d-1}(t), \\ n(\sigma, t) &= b_0(t)\sigma^{n+d-1} + b_1(t)\sigma^{n+d-2} + \cdots + b_{n+d-1}(t), \end{aligned} \tag{2.12}$$

with $a_i(t + T) = a_i(t)$, $b_j(t + T) = b_j(t)$, and the delay d appearing as the excess of the numerator degree over the denominator degree.

Multiplying both sides of $y(t) = G(\sigma, t)u(t)$ by $d(\sigma, t)$ yields

$$d(\sigma, t)\, y(t) = n(\sigma, t)\, u(t). \tag{2.13}$$

Expanding the polynomials and rearranging terms gives a difference equation with periodic coefficients:

$$y(t) + \sum_{k=1}^{n+d-1} a_k(t)\, y(t - k) = \sum_{j=0}^{n+d-1} b_j(t)\, u(t - j - d). \tag{2.14}$$

Because the coefficients $a_k(t)$ are periodic with period T, the explicit time-dependence on the left-hand side can be removed by constructing a T-step recurrence.

Define the extended state vector as

$$\mathbf{Y}(t) = \big[y(t), y(t - 1), \ldots, y(t - T + 1)\big]^{\mathrm{T}}. \tag{2.15}$$

Writing equation (2.14) at times $t, t - 1, \ldots, t - T + 1$ and exploiting the periodicity $a_k(t + T) = a_k(t)$, algebraic elimination leads to a relation that involves only delays that are multiples of the period:

$$y(t) = \sum_{\ell=1}^{L} \varsigma_\ell\, y(t - \ell T) + \sum_{j=0}^{Q} \varpi_j(t)\, u(t - j - d) + \bar{v}(t), \tag{2.16}$$

where the coefficients ς_ℓ are time-invariant, obtained from products of the original periodic coefficients $a_k(\cdot)$; the coefficients $\varpi_j(t)$ are periodic with period T, inherited from the original periodic coefficients $b_j(t)$ together with the elimination procedure; and $\bar{v}(t)$ collects the noise term $v(t)$ and its delayed versions.

Taking $\tau = T$ (or any integer multiple thereof), equation (2.16) can be written in the standard form

$$y(t) = \sum_{k=1}^{p} \varsigma_k\, y(t - k\tau) + \sum_{j=0}^{q} \varpi_j(t)\, u(t - j - d) + v(t) \tag{2.17}$$

after renaming $p = L$, $q = Q$, and absorbing the noise term into $v(t)$.

Since every operation in the derivation (shifting, multiplication by periodic coefficients, and algebraic elimination) preserves periodicity, each coefficient $\varpi_j(t)$ inherits the period T from the original coefficients $b_j(t)$; i.e., $\varpi_j(t + T) = \varpi_j(t)$, $j = 0, 1, \ldots, q$. Therefore, the proof of Lemma 2.2 is completed. □

The representation $y(t) = \sum_{k=1}^{p} \zeta_k y(t-k\tau) + \sum_{j=0}^{q} \pi_j(t) u(t-j-d) + v(t)$ uses only output samples at multiples of τ, where τ is an integer multiple of the period T. This downsampling potentially discards inter-sample information, which may raise concerns about identifiability in time-varying periodic systems. Nevertheless, the following considerations justify its use under appropriate conditions.

First, the parameter τ is typically chosen as the period T (the smallest integer multiple). By the lifting technique, the original T-periodic system can be transformed into a T-step time-invariant system whose state and output are sampled every T steps; the downsampled output sequence $y(kT)$ then exactly corresponds to the output of this lifted system, and the parameters ζ_k become time-invariant. Second, regarding persistency of excitation, assume that the original system satisfies the persistent excitation condition: There exist constants $\alpha, \beta > 0$ such that for sufficiently large N,

$$\alpha I \leq \frac{1}{N} \sum_{t=1}^{N} \psi(t)\psi^{\mathrm{T}}(t) \leq \beta I,$$

where $\psi(t)$ is the full information vector using all available data. Under this condition, the downsampled information matrix (constructed from samples at times $t, t-\tau, t-2\tau, \ldots$) also satisfies a similar inequality, provided the system is observable from the downsampled output. Hence, no essential information is lost for parameter identification. Third, sufficiency for identifiability: For linear periodic systems with period T, the lifted T-step ahead model is known to be identifiable if the original system is observable and the input is sufficiently rich. The representation (2.10) with $\tau = T$ preserves all information necessary to estimate the time-invariant parameters ζ_k and the periodic coefficients $\pi_j(t)$ up to a periodic transformation. Finally, in practical implementation, the proposed algorithm does not actually discard intermediate outputs; it directly uses the full input-output sequence $\{u(t), y(t)\}$ to construct the auxiliary model and update the estimates. The downsampled form (2.10) serves mainly as a theoretical bridge to obtain a time-invariant denominator polynomial, simplifying the derivation of the identification model. The final algorithm (Algorithm 1) operates on every sample, thereby avoiding any actual loss of information.

Therefore, while the representation (2.10) omits inter-sample outputs, it remains a valid and sufficient model for parameter identification under the standard assumptions of persistency of excitation and observability of the lifted system. For $\tau > T$, the same conclusions hold but require a proportionally stronger excitation condition.

The periodic matrix equations arising from such models have been studied numerically in [31]. As a natural progression of the input-output paradigm, we give the following equivalent reformulation of (2.1) based on Lemma 2.1 and Lemma 2.2:

$$y(t) = \frac{\varpi(\sigma, t)}{\varsigma(\sigma)} u(t) + v(t), \tag{2.18}$$

where $\varsigma(\sigma)$ and $\varpi(\sigma, t)$ are determined by (2.19) and (2.20), respectively. This parametric decomposition explicitly separates the time-invariant denominator $\varsigma(\sigma)$ from the time-varying numerator $\varpi(\sigma, t)$, enabling systematic parameter estimation.

$$\varsigma(\sigma) = 1 + \varsigma_1 \sigma^{-1} + \varsigma_2 \sigma^{-2} + \cdots + \varsigma_p \sigma^{-p}, \tag{2.19}$$

$$\varpi(\sigma, t) = \varpi_0(t) + \varpi_1(t)\sigma^{-1} + \varpi_2(t)\sigma^{-2} + \cdots + \varpi_q(t)\sigma^{-q}. \tag{2.20}$$

Therefore, the identification of system (2.1) can be transformed into the identification of ς_i and $\varpi_j(t)$.

To address the challenge of unknown intermediate variables in the system, we propose an auxiliary model-based improved Levenberg-Marquardt algorithm for parameter estimation. The key innovation lies in the construction of an intermediate variable that bridges measurable inputs/outputs with unknown parameters. Specifically, we define

$$i(t) = \frac{\varpi(\sigma, t)}{\varsigma(\sigma)} u(t). \tag{2.21}$$

Through an algebraic shifting operation, the relationship in (2.21) can be equivalently expressed as

$$\varsigma(\sigma) i(t) = \varpi(\sigma, t) u(t). \tag{2.22}$$

The complete parameter vector θ is then constructed by

$$\theta := [\varsigma_1, \varsigma_2, \cdots, \varsigma_p, \varpi_0(t), \varpi_1(t), \varpi_2(t), \cdots, \varpi_q(t)]^{\mathrm{T}}. \tag{2.23}$$

Correspondingly, the information matrix $\psi(t)$ is defined as

$$\psi(t) = [-i(t-1), -i(t-2), \ldots, -i(t-p), u(t), u(t-1), \ldots, u(t-q)]^{\mathrm{T}}. \tag{2.24}$$

Substituting (2.19) and (2.20) into (2.22) and applying algebraic rearrangement, we derive the expanded expression for the intermediate variable:

$$\begin{aligned} i(t) &= [1 - \varsigma(\sigma)] i(t) + \varpi(\sigma, t) u(t) \\ &= -\varsigma_1 i(t-1) - \varsigma_2 i(t-2) - \cdots - \varsigma_p i(t-p) + \varpi_0(t) u(t) + \varpi_1(t) u(t-1) + \cdots + \varpi_q(t) u(t-q) \\ &= \psi^{\mathrm{T}}(t) \theta. \end{aligned} \tag{2.25}$$

Combining (2.18), (2.21), and (2.25) yields the final identification model for system (2.1):

$$y(t) = \psi^{\mathrm{T}}(t) \theta + v(t). \tag{2.26}$$

3. The improved LM identification algorithm

The parameter identification of linear discrete periodic systems with time delays presents a fundamental challenge due to the presence of unmeasurable intermediate variables. These latent variables not only introduce additional nonlinearity to the system dynamics but also necessitate robust identification algorithms capable of handling increased parametric uncertainty. To overcome these limitations, we develop an improved Levenberg-Marquardt algorithm that incorporates auxiliary modeling techniques. The proposed algorithm distinguishes itself from existing LM variants through two key features: an adaptive regularization strategy that adjusts the damping parameter in response to local curvature, and an auxiliary model that reconstructs latent variables without requiring state measurements. These features collectively enable the algorithm to handle the coupling between time delay and periodic variations—a capability not present in prior LM-based methods.

Conventional LM methods rely on measurable state information, which is unavailable in time-delay periodic systems. To address this, the auxiliary model is constructed to recursively generate surrogate signals $\hat{i}(t)$ from input-output data, explicitly incorporating the time delay d and synchronizing with the

system period T. Unlike standard LM algorithms that assume time-invariant parameters, the proposed update law tracks periodic variations by aligning the information matrix $\hat{\psi}(t)$ with the system period T. Moreover, unlike state estimation techniques such as Kalman filtering—which require prior knowledge of system matrices and are not designed for non-causal time-delay structures—the auxiliary model operates directly within the identification framework, enabling latent variable reconstruction without a separate state-space model while naturally accommodating both time delays and periodic variations.

The identification process begins with the definition of the criterion function:

$$\mathbb{V}(\theta) = \sum_{t=1}^{l} [e(t)]^2 = \sum_{t=1}^{l} \left[y(t) - \psi^{\mathrm{T}}(t)\theta \right]^2, \tag{3.1}$$

where θ denotes the parameter vector to be estimated and $\psi(t)$ represents the information matrix at time t. This quadratic form is chosen to facilitate the derivation of closed-form solutions while maintaining convexity properties essential for convergence.

The Jacobian matrix $J(\theta)$ quantifies the sensitivity of the criterion function to parameter variations. For the current sample at time t, the Jacobian row vector is constructed as

$$J(t) = \left[\frac{\partial e(t)}{\partial \theta_1}, \ldots, \frac{\partial e(t)}{\partial \theta_n} \right] = -\psi^{\mathrm{T}}(t), \tag{3.2}$$

i.e.,

$$J_j(t) = -\psi_j(t). \tag{3.3}$$

Building upon this Jacobian formulation, we define

$$\begin{aligned} \tilde{\gamma}(t) &= \left(J(t)^{\mathrm{T}} J(t) + \lambda(t) I \right)^{-1} J^{\mathrm{T}}(t) \\ &= -\left(\psi(t)\psi^{\mathrm{T}}(t) + \lambda(t) I \right)^{-1} \psi(t). \end{aligned} \tag{3.4}$$

Let the effective gain matrix $\gamma(t)$ be

$$\gamma(t) = -\tilde{\gamma}(t). \tag{3.5}$$

Then the parameter update using the true information matrix is

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \gamma(t) \left[y(t) - \psi^{\mathrm{T}}(t)\hat{\theta}(t-1) \right]. \tag{3.6}$$

The regularization parameter $\lambda(t)$ is adaptively updated as

$$\lambda(t) = \frac{\beta \left\| \mathbb{V}[\hat{\theta}(t)] \right\|^2}{1 + \alpha \left\| \mathbb{V}[\hat{\theta}(t)] \right\|^2}, \quad \alpha + \beta = 1. \tag{3.7}$$

Notably, the implementation of (3.4)–(3.7) requires knowledge of intermediate variables $i(t-i)$ $(i = 1, 2, \cdots, 2n)$, which are inherently unmeasurable. To resolve this practical limitation, we develop an auxiliary model that generates surrogate signals $\hat{i}(t)$ through the following procedure.

First, the information matrix $\hat{\psi}(t)$ is constructed using only previously estimated auxiliary variables:

$$\hat{\psi}(t) = [-\hat{i}(t-1), \ldots, -\hat{i}(t-p),\ u(t), \ldots, u(t-q)]^{\mathrm{T}}. \tag{3.8}$$

Second, the parameter vector is updated using the current $\hat{\psi}(t)$:

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \gamma(t)[y(t) - \hat{\psi}^{\mathrm{T}}(t)\hat{\theta}(t-1)]. \tag{3.9}$$

Third, the auxiliary variable is updated as

$$\hat{i}(t) = \hat{\psi}^{\mathrm{T}}(t)\hat{\theta}(t). \tag{3.10}$$

This ordering ensures that no implicit equation appears; each variable is computed from previously available quantities. Consequently, existence and uniqueness of the solution at each time step are trivial.

For completeness, the algorithmic implementation uses the estimated information matrix $\hat{\psi}(t)$ in place of the true $\psi(t)$. Hence, we rewrite the key steps as

$$\gamma(t) = (\hat{\psi}(t)\hat{\psi}^{\mathrm{T}}(t) + \lambda(t)I)^{-1}\hat{\psi}(t), \tag{3.11}$$

$$\hat{\theta}(t) = \hat{\theta}(t-1) + \gamma(t)\Big[y(t) - \hat{\psi}^{\mathrm{T}}(t)\hat{\theta}(t-1)\Big], \tag{3.12}$$

$$\lambda(t) = \frac{\beta\left\|\mathbb{V}[\hat{\theta}(t)]\right\|^2}{1+\alpha\left\|\mathbb{V}[\hat{\theta}(t)]\right\|^2}, \quad \alpha+\beta = 1. \tag{3.13}$$

Thus, the complete auxiliary-model-based improved Levenberg-Marquardt (AM-ILM) algorithm is given by (3.11)–(3.13) together with the adaptive regularization (3.7). The overall procedure is summarized in Algorithm 1.

Algorithm 1 AM-ILM identification algorithm for time delay linear discrete periodic systems.

1. Initialize the length of the data to L and the tolerance ε with $\varepsilon > 0$. At $t = 1$, set the initial value $\hat{\theta}(0) = 1/\rho_0$, $\hat{i}(t-i) = 0$, $u(t-i) = 0$, $y(t-i) = 0$ $(i = 1, 2, \cdots, 2n)$, $\lambda(0) = 0.02$, $\alpha = 0.1$, $\beta = 0.9$, $\varepsilon = 10^{-6}$, $\gamma(0) = 0$.
2. For $t = 1, 2, 3, \cdots, L$, collect input $u(t)$ and output $y(t)$.
3. Construct information matrix $\hat{\psi}(t)$ using equation (3.8).
4. Compute Jacobian matrix using equations (3.2) and (3.3).
5. Calculate gain $\gamma(t)$ using equation (3.11).
6. Update $\hat{\theta}(t)$ using equation (3.12).
7. Estimate auxiliary vector $\hat{i}(t)$ using equation (3.10).
8. **if** $\frac{\|\hat{\theta}(t)-\hat{\theta}(t-1)\|^2}{\|\hat{\theta}(t)\|^2+\delta} < \varepsilon$, (with $\delta = 10^{-12}$ to avoid division by zero)
   go to step 9.
   else
   Update $\lambda(t)$ by equation (3.13),
   $t := t + 1$, go to step 2.
   end if
9. **Output:** Parameter estimate $\hat{\theta}$.

Remark 3.1. The persistent excitation condition requires that $\exists\, \alpha, \beta > 0$ such that for all t

$$\alpha I \le \frac{1}{T}\sum_{k=t}^{t+T-1}\hat{\psi}(k)\hat{\psi}^{\mathrm{T}}(k) \le \beta I. \tag{3.14}$$

The convergence properties of Algorithm 1 are established through the following two fundamental lemmas. First, we present a convexity result for the criterion function.

Lemma 3.1. *$\forall \theta_1, \theta_2$ and $\eta \in [0, 1]$,*

$$\mathbb{E}[\mathbb{V}(\eta\theta_1 + (1-\eta)\theta_2)] \le \eta\,\mathbb{E}[\mathbb{V}(\theta_1)] + (1-\eta)\,\mathbb{E}[\mathbb{V}(\theta_2)]. \tag{3.15}$$

Proof. First of all, we have

$$\begin{aligned}\mathbb{V}(\eta\theta_1 + (1-\eta)\theta_2) &= \sum_{t=1}^{l}\big[y(t) - \hat{\psi}^{\mathrm{T}}(t)(\eta\theta_1 + (1-\eta)\theta_2)\big]^2 \\ &= \eta\mathbb{V}(\theta_1) + (1-\eta)\mathbb{V}(\theta_2) - \eta(1-\eta)(\theta_1-\theta_2)^{\mathrm{T}}\Big(\sum_{t=1}^{l}\hat{\psi}(t)\hat{\psi}^{\mathrm{T}}(t)\Big)(\theta_1-\theta_2).\end{aligned} \tag{3.16}$$

Since $\sum_{t=1}^{l}\hat{\psi}(t)\hat{\psi}^{\mathrm{T}}(t) \ge 0$ (positive semidefinite), the last term is nonpositive, yielding

$$\mathbb{V}(\eta\theta_1 + (1-\eta)\theta_2) \le \eta\mathbb{V}(\theta_1) + (1-\eta)\mathbb{V}(\theta_2). \tag{3.17}$$

Thus $\mathbb{V}$ is convex.

Now take expectation on both sides with respect to the noise distribution. Expectation is linear and preserves the inequality direction, hence

$$\mathbb{E}[\mathbb{V}(\eta\theta_1 + (1-\eta)\theta_2)] \le \eta\,\mathbb{E}[\mathbb{V}(\theta_1)] + (1-\eta)\,\mathbb{E}[\mathbb{V}(\theta_2)].$$

Thus, the Lemma 3.1 is proved. □

We now proceed to demonstrate the monotonic error decrease property and establish mean-square convergence.

Lemma 3.2. *For the linear discrete periodic system* (2.1) *with time delays and period T, assume that the persistent excitation condition* (3.14) *holds, the noise $v(t)$ is zero-mean white noise with finite variance σ_v^2, and the gain matrix $\gamma(t)$ is defined by* (3.11)*. Then there exists a constant $C > 0$ such that*

$$\limsup_{t\to\infty}\mathbb{E}[\|\tilde{\theta}(t)\|^2] \le C\sigma_v^2, \tag{3.18}$$

where $\tilde{\theta}(t) = \hat{\theta}(t) - \theta$. In particular, if $\sigma_v^2 = 0$, then

$$\lim_{t\to\infty}\mathbb{E}\left[\left\|\hat{\theta}(t) - \theta\right\|^2\right] = 0. \tag{3.19}$$

Proof. First, define the parameter estimation error as

$$\tilde{\theta}(t) = \hat{\theta}(t) - \theta. \tag{3.20}$$

Substituting (3.12) into (3.20) gives

$$\tilde{\theta}(t+1) = \hat{\theta}(t) + \gamma(t+1)\left[y(t+1) - \hat{\psi}^{\mathrm{T}}(t+1)\hat{\theta}(t)\right] - \theta. \tag{3.21}$$

According to (3.20), (3.21) can be rewritten as

$$\begin{aligned}\tilde{\theta}(t+1) &= \hat{\theta}(t+1) - \theta \\ &= \hat{\theta}(t) + \gamma(t+1)\left[y(t+1) - \hat{\psi}^{\mathrm{T}}(t+1)\hat{\theta}(t)\right] - \theta \\ &= \tilde{\theta}(t) + \gamma(t+1)\left[y(t+1) - \hat{\psi}^{\mathrm{T}}(t+1)\hat{\theta}(t)\right].\end{aligned} \tag{3.22}$$

Substituting (2.26) into (3.22) yields

$$\begin{aligned}\tilde{\theta}(t+1) &= \tilde{\theta}(t) + \gamma(t+1)\left[y(t+1) - \hat{\psi}^{\mathrm{T}}(t+1)\hat{\theta}(t)\right] \\ &= \tilde{\theta}(t) + \gamma(t+1)\left[\hat{\psi}^{\mathrm{T}}(t+1)\theta + v(t+1) - \hat{\psi}^{\mathrm{T}}(t+1)\hat{\theta}(t)\right] \\ &= \tilde{\theta}(t) + \gamma(t+1)\left[\psi^{\mathrm{T}}(t+1)\theta + v(t+1) - \hat{\psi}^{\mathrm{T}}(t+1)(\theta + \tilde{\theta}(t))\right] \\ &= (I - \gamma(t+1)\hat{\psi}^{\mathrm{T}}(t+1))\tilde{\theta}(t) - \gamma(t+1)\Delta\psi^{\mathrm{T}}(t+1)\theta + \gamma(t+1)v(t+1),\end{aligned} \tag{3.23}$$

where $\Delta\psi(t) := \hat{\psi}(t) - \psi(t)$.

By (3.11), $\gamma(t+1) = (\hat{\psi}(t+1)\hat{\psi}^{\mathrm{T}}(t+1) + \lambda(t+1)I)^{-1}\hat{\psi}(t+1)$. Hence,

$$\gamma(t+1)\hat{\psi}^{\mathrm{T}}(t+1) = [\hat{\psi}(t+1)\hat{\psi}(t+1)^{\mathrm{T}} + \lambda(t+1)I]^{-1}\hat{\psi}(t+1)\hat{\psi}^{\mathrm{T}}(t+1). \tag{3.24}$$

Let $\sigma_i(t+1) \geq 0$ be the eigenvalues of $\hat{\psi}(t+1)\hat{\psi}^{\mathrm{T}}(t+1)$. Then the eigenvalues of $\gamma(t+1)\hat{\psi}^{\mathrm{T}}(t+1)$ are

$$\mu_i(t+1) = \frac{\sigma_i(t+1)}{\sigma_i(t+1) + \lambda(t+1)} \in (0,1). \tag{3.25}$$

Thus, $0 < \gamma(t+1)\hat{\psi}^{\mathrm{T}}(t+1) < I$.

According to (3.24), define $\rho_0 = \max_i(1 - \mu_i(t+1))^2$, which satisfies $0 < \rho_0 < 1$. Since $\mu_i(t+1) \in (0,1)$, the contraction factor $\rho_0 = \max_i(1 - \mu_i(t+1))^2$ satisfies $\rho_0 \in (0,1)$ for each t, and moreover $\bar{\rho}_0 = \sup_t \rho_0(t) < 1$ holds uniformly. Then for any vector x,

$$\|(I - \gamma(t+1)\hat{\psi}^{\mathrm{T}}(t+1))x\|^2 \leq \rho_0\|x\|^2. \tag{3.26}$$

Persistent excitation (3.14) implies $\|\gamma(t+1)\|^2 \leq g_{\max}$.

Let $\Delta i(t) = \hat{i}(t) - i(t)$. From (3.10) and (2.25), we have

$$\Delta i(t) = -\sum_{k=1}^{p} \zeta_k \Delta i(t-k) + \hat{\psi}^{\mathrm{T}}(t)\tilde{\theta}(t). \tag{3.27}$$

The homogeneous part of (3.27) has the characteristic polynomial $z^p + \sum_{k=1}^{p} \zeta_k z^{p-k}$, which is stable [32], so there exist constants $c_0 > 0$, $\lambda \in (0,1)$ such that for any initial condition,

$$|\Delta i(t)| \leq c_0\lambda^t \max_{1\leq k\leq p} |\Delta i(0)| + c_0 \sum_{k=0}^{t-1} \lambda^{t-1-k}\left|\hat{\psi}^{\mathrm{T}}(k)\tilde{\theta}(k)\right|. \tag{3.28}$$

Persistent excitation yields $\|\hat{\psi}(k)\| \leq M$ almost surely, so $|\hat{\psi}^{\mathrm{T}}(k)\tilde{\theta}(k)| \leq M\|\tilde{\theta}(k)\|$. Squaring (3.28), taking expectations, and using the Cauchy-Schwarz inequality gives

$$\mathbb{E}[\Delta i(t)^2] \leq c_1\lambda^{2t} + c_2 \sum_{k=0}^{t-1} \lambda^{2(t-1-k)}\mathbb{E}[\|\tilde{\theta}(k)\|^2], \tag{3.29}$$

with $c_1, c_2 > 0$. Since $\|\hat{\psi}(t) - \psi(t)\|^2 = \sum_{k=1}^{p} \Delta i(t-k)^2$, there exists $c_3 > 0$ such that

$$\mathbb{E}[\|\Delta\psi(t+1)\|^2] \le c_3\lambda^{2t} + c_3 \sum_{k=0}^{t} \lambda^{2(t-k)}\mathbb{E}[\|\tilde{\theta}(k)\|^2]. \tag{3.30}$$

Let $\Delta I(t) = [\Delta i(t-1), \dots, \Delta i(t-p)]^{\mathrm{T}}$, and define the Lyapunov function

$$V(t) = \|\tilde{\theta}(t)\|^2 + \alpha_0\|\Delta I(t)\|^2, \tag{3.31}$$

where $\alpha_0 > 0$ will be chosen later. Using the dynamics of $\Delta I(t)$ derived from (3.27) and the stability of the homogeneous part, there exist constants $0 < \rho_i < 1$ and $c_i > 0$ such that

$$\mathbb{E}[\|\Delta I(t+1)\|^2 \mid \mathcal{F}_t] \le \rho_i\|\Delta I(t)\|^2 + c_i\|\tilde{\theta}(t)\|^2, \tag{3.32}$$

where $\mathcal{F}_t$ is the σ-algebra generated by all data up to time t.

Take conditional expectation in (3.23). Because $v(t+1)$ is independent of $\mathcal{F}_t$ and zero-mean, and $\Delta\psi(t+1)$ is $\mathcal{F}_t$-measurable, we obtain from (3.26), (3.28), and (3.30):

$$\begin{aligned}\mathbb{E}[V(t+1) \mid \mathcal{F}_t] &\le \rho_0\|\tilde{\theta}(t)\|^2 + 2g_{\max}^2\|\theta\|^2\mathbb{E}[\|\Delta\psi(t+1)\|^2 \mid \mathcal{F}_t] \\ &\quad + 2g_{\max}^2\sigma_v^2 + \alpha_0(c_i\|\tilde{\theta}(t)\|^2 + \rho_i\|\Delta I(t)\|^2) \\ &\le (\rho_0 + 2g_{\max}^2\|\theta\|^2c_3 + \alpha_0c_i)\|\tilde{\theta}(t)\|^2 \\ &\quad + (2g_{\max}^2\|\theta\|^2c_3 + \alpha_0\rho_i)\|\Delta I(t)\|^2 + 2g_{\max}^2\sigma_v^2 + \delta(t),\end{aligned} \tag{3.33}$$

where $\delta(t) \le c_4\lambda^{2t}$ is an exponentially decaying term.

Since $\rho_0 < 1$ and $\rho_i < 1$, we may select α_0 sufficiently small so that

$$\rho_\theta := \rho_0 + 2g_{\max}^2\|\theta\|^2c_3 + \alpha_0c_i < 1, \qquad \rho_I := 2g_{\max}^2\|\theta\|^2c_3 + \alpha_0\rho_i < 1. \tag{3.34}$$

Let $\rho = \max\{\rho_\theta, \rho_I\}$; then $0 < \rho < 1$. Notice that $\|\tilde{\theta}(t)\|^2 + \|\Delta I(t)\|^2 \le \max(1, 1/\alpha_0)V(t)$. Consequently, from (3.33) we obtain for some $0 < \rho' < 1$ and $C' > 0$,

$$\mathbb{E}[V(t+1) \mid \mathcal{F}_t] \le \rho' V(t) + C'\sigma_v^2 + \delta(t). \tag{3.35}$$

Take full expectation in (3.35) and iterate:

$$\mathbb{E}[V(t)] \le (\rho')^t\mathbb{E}[V(0)] + \frac{C'}{1-\rho'}\sigma_v^2 + \sum_{k=0}^{t-1}(\rho')^{t-1-k}\delta(k). \tag{3.36}$$

Because $\delta(k) \le c_4\lambda^{2k}$ and we can ensure $\lambda^2 < \rho'$ by a suitable choice of α, the last sum is bounded and tends to zero as $t \to \infty$. Hence,

$$\limsup_{t\to\infty} \mathbb{E}[V(t)] \le \frac{C'}{1-\rho'}\sigma_v^2. \tag{3.37}$$

Since $\|\tilde{\theta}(t)\|^2 \le V(t)$, we finally obtain

$$\limsup_{t\to\infty} \mathbb{E}[\|\tilde{\theta}(t)\|^2] \le \frac{C'}{1-\rho'}\sigma_v^2 =: C\sigma_v^2. \tag{3.38}$$

If $\sigma_v^2 = 0$, then $C' = 0$, and

$$\lim_{t\to\infty} \mathbb{E}[\|\tilde{\theta}(t)\|^2] = 0. \tag{3.39}$$

Therefore, Lemma 3.2 is proved. □

Remark 3.2. From (3.11), we have $\gamma(t)\hat{\psi}^{\mathrm{T}}(t) = (\hat{\psi}\hat{\psi}^{\mathrm{T}} + \lambda I)^{-1}\hat{\psi}\hat{\psi}^{\mathrm{T}}$, whose eigenvalues are $\sigma_i/(\sigma_i + \lambda) \in (0, 1)$. Hence $0 < \gamma\hat{\psi}^{\mathrm{T}} < I$, which guarantees the contraction property used in Lemma 3.2.

Lemma 3.3. *Under the same conditions as Lemma 3.2 (zero-mean white noise with finite variance σ_v^2, persistent excitation* (3.14)*, and asymptotic stability of the denominator polynomial $1 + \sum_{k=1}^{p} \zeta_k z^{-k}$), let $\hat{\psi}(t)$ be generated by the auxiliary model* (3.8)–(3.10)*, and let $\psi(t)$ be the true information matrix* (2.24)*. Then there exists a constant $D > 0$ such that*

$$\limsup_{t\to\infty} \mathbb{E}[\|\hat{\psi}(t) - \psi(t)\|^2] \le D\sigma_v^2. \tag{3.40}$$

In particular, if $\sigma_v^2 = 0$, then

$$\lim_{t\to\infty} \mathbb{E}[\|\hat{\psi}(t) - \psi(t)\|^2] = 0, \tag{3.41}$$

and $\mathbb{E}[\hat{\psi}(t)] \to \psi(t)$.

Proof. Set $\Delta i(t) = \hat{i}(t) - i(t)$. From (2.24) and (3.8), we have

$$\hat{\psi}(t) - \psi(t) = [-\Delta i(t-1), \dots, -\Delta i(t-p), 0, \dots, 0]^{\mathrm{T}}, \tag{3.42}$$

hence $\|\hat{\psi}(t) - \psi(t)\|^2 = \sum_{k=1}^{p} \Delta i(t-k)^2$. Thus it suffices to analyse $\Delta i(t)$.

From the true model (2.25) and the auxiliary model (3.9)–(3.10), we obtain

$$i(t) = -\sum_{k=1}^{p} \zeta_k i(t-k) + \sum_{j=0}^{q} \pi_j(t) u(t-j-d), \tag{3.43}$$

$$\hat{i}(t) = -\sum_{k=1}^{p} \hat{\zeta}_k(t)\hat{i}(t-k) + \sum_{j=0}^{q} \hat{\pi}_j(t) u(t-j-d). \tag{3.44}$$

Let $\tilde{\theta}(t) = \hat{\theta}(t) - \theta$ with components $\tilde{\zeta}_k(t) = \hat{\zeta}_k(t) - \zeta_k$, $\tilde{\pi}_j(t) = \hat{\pi}_j(t) - \pi_j(t)$. Subtracting (3.43) from (3.44) and using

$$\hat{\psi}^{\mathrm{T}}(t)\tilde{\theta}(t) = -\sum_{k=1}^{p} \tilde{\zeta}_k(t)\hat{i}(t-k) + \sum_{j=0}^{q} \tilde{\pi}_j(t) u(t-j-d), \tag{3.45}$$

we obtain

$$\Delta i(t) = -\sum_{k=1}^{p} \zeta_k \Delta i(t-k) + \hat{\psi}^{\mathrm{T}}(t)\tilde{\theta}(t). \tag{3.46}$$

The homogeneous part of (3.46) is stable. As shown in the proof of Lemma 3.2 (specifically, inequalities (3.28)–(3.30) and the subsequent derivation), there exist constants $c_1, c_2 > 0$ and $\lambda \in (0, 1)$ such that

$$\mathbb{E}[\Delta i(t)^2] \le c_1\lambda^{2t} + c_2 \sum_{k=0}^{t-1} \lambda^{2(t-1-k)} \mathbb{E}[\|\tilde{\theta}(k)\|^2]. \tag{3.47}$$

By Lemma 3.2, $\mathbb{E}[\|\tilde{\theta}(k)\|^2]$ is bounded and $\limsup_{k\to\infty} \mathbb{E}[\|\tilde{\theta}(k)\|^2] \le C\sigma_v^2$. Hence, there exists C_1 such that

$$\sum_{k=0}^{t-1} \lambda^{2(t-1-k)} \mathbb{E}[\|\tilde{\theta}(k)\|^2] \le C_1\sigma_v^2. \tag{3.48}$$

Therefore,

$$\limsup_{t\to\infty} \mathbb{E}[\Delta i(t)^2] \le \frac{c_2 C_1}{1-\lambda^2}\sigma_v^2. \tag{3.49}$$

If $\sigma_v^2 = 0$, Lemma 3.2 yields $\mathbb{E}[\|\tilde{\theta}(k)\|^2] \to 0$, so the sum in (3.47) vanishes and $\mathbb{E}[\Delta i(t)^2] \to 0$. Since $\|\hat{\psi}(t) - \psi(t)\|^2 = \sum_{k=1}^{p} \Delta i(t-k)^2$, we obtain

$$\limsup_{t\to\infty} \mathbb{E}[\|\hat{\psi}(t) - \psi(t)\|^2] \le p \cdot \frac{c_2 C_1}{1-\lambda^2}\sigma_v^2 \triangleq D\sigma_v^2, \tag{3.50}$$

and the upper bound is zero when $\sigma_v^2 = 0$. Finally, $\|\mathbb{E}[X]\|^2 \le \mathbb{E}[\|X\|^2]$ implies $\mathbb{E}[\hat{\psi}(t)] \to \psi(t)$. □

Based on the above analysis, we now establish the convergence of Algorithm 1 and present the relevant conclusions in the form of Theorem 3.1.

Theorem 3.1. *Consider the linear discrete periodic system* (2.1) *with time delay d and period T. Assume that the noise sequence v(t) is zero-mean Gaussian white noise with finite variance σ_v^2, independent of the input u(t) (u(t) is deterministic) and the information matrix ψ(t), and the persistent excitation condition* (3.14) *holds.* $\mathbb{E}[\Delta\psi(t) \mid \mathcal{F}_{t-1}] = 0$*, where $\mathcal{F}_{t-1}$ is the σ-algebra generated by all observations up to time t − 1 (i.e.,*$\mathcal{F}_{t-1} = \sigma\{u(1), \dots, u(t-1),\ y(1), \dots, y(t-1)\}$*). Let $\hat{\theta}(t)$ be the estimate obtained by Algorithm 1 using the estimated information matrix $\hat{\psi}(t)$. Then $\hat{\theta}(t)$ is an asymptotically unbiased estimate of θ.*

Proof. By (3.11) and the persistent excitation condition, there exists a constant $0 < \rho < 1$ such that $\|I - \gamma(t)\hat{\psi}^{\mathrm{T}}(t)\|_2 \le \rho$ a.s. for all sufficiently large t, and $\|\gamma(t)\|_2$ is uniformly bounded.

Taking expectation of (3.23) and using $\mathbb{E}[v(t+1) \mid \mathcal{F}_t] = 0$ together with the $\mathcal{F}_t$-measurability of $\gamma(t+1)$ and $\hat{\psi}(t+1)$, we obtain

$$\mathbb{E}[\tilde{\theta}(t+1)] = \mathbb{E}\left[(I - \gamma(t+1)\hat{\psi}^{\mathrm{T}}(t+1))\mathbb{E}[\tilde{\theta}(t) \mid \mathcal{F}_t]\right] - \mathbb{E}\left[\gamma(t+1)\Delta\psi^{\mathrm{T}}\right]\theta. \tag{3.51}$$

Denote $\beta(t) = \mathbb{E}[\tilde{\theta}(t)]$, $\Phi(t) = \mathbb{E}[I - \gamma(t+1)\hat{\psi}^{\mathrm{T}}(t+1)]$, and $\Gamma(t) = \mathbb{E}[\gamma(t+1)\Delta\psi^{\mathrm{T}}(t+1)]$. Then,

$$\beta(t+1) = \Phi(t)\beta(t) - \Gamma(t)\theta. \tag{3.52}$$

Since $\mathbb{E}[\Delta\psi(t) \mid \mathcal{F}_{t-1}] = 0$ and $\gamma(t)$ is $\mathcal{F}_{t-1}$-measurable, hence

$$\mathbb{E}[\gamma(t)\Delta\psi^{\mathrm{T}}(t)] = \mathbb{E}\left[\gamma(t)\,\mathbb{E}[\Delta\psi^{\mathrm{T}}(t) \mid \mathcal{F}_{t-1}]\right] = 0, \tag{3.53}$$

i.e., $\Gamma(t) = 0$. Therefore, (3.52) reduces to $\beta(t+1) = \Phi(t)\beta(t)$.

From Lemma 3.2 ((3.25) and (3.30)), we have $0 < \gamma(t)\hat{\psi}^{\mathrm{T}}(t) < I$ almost surely, and $\|I - \gamma(t)\hat{\psi}^{\mathrm{T}}(t)\|^2 \le \sqrt{\rho_0} < 1$. Hence,

$$\|I - \mathbb{E}[\gamma(t)\hat{\psi}^{\mathrm{T}}(t)]\|^2 \le \mathbb{E}[\|I - \gamma(t)\hat{\psi}^{\mathrm{T}}(t)\|^2] \le \sqrt{\rho_0} < 1, \tag{3.54}$$

which implies that $I - \mathbb{E}[\gamma(t)\hat{\psi}^{\mathrm{T}}(t)]$ is a contraction (its spectral radius is less than 1).

Consequently, there exists a constant $0 < \rho < 1$ such that

$$\|\beta(t)\| \le \rho\|\beta(t-1)\|. \tag{3.55}$$

Iterating this inequality gives $\beta(t) \to 0$, i.e.,

$$\lim_{t\to\infty} \mathbb{E}[\hat{\theta}(t)] = \theta.$$

Thus, $\hat{\theta}(t)$ is an asymptotically unbiased estimate of θ. This completes the proof of Theorem 3.1. □

4. Simulation and analysis

In this section, we employ Algorithm 1 to identify the system parameters of the unit-time delay system (2.1) with period $T = 3$ and $T = 4$, respectively. We evaluate the performance of Algorithm 1 through Example 4.1, and compare it with the AM-RLS, LM, fully connected neural network (FCN), model-agnostic meta-learning (MAML), Bayesian and RLS algorithms. Example 4.2 further examines the robustness of Algorithm 1 under different initial conditions and varying system parameters ($T = 4$, $d = 2$). Example 4.3 presents an application. In the numerical simulations, the input signal $u(t)$ is a sequence of independent and identically distributed (i.i.d.) continuous random variables following a Bernoulli 0-1 distribution, and the variances of the noise terms $v(t)$ in these examples are $\delta^2 = 0.5$, $\delta^2 = 0.1$, and $\delta^2 = 0.01$ respectively.

Example 4.1. *Consider a periodic system of the form* (2.1) *with period* $T = 3$ *and time delay* $d = 1$. *The system matrices are given by*

$$a(t) = \begin{cases} 0.2, & t = 3k \\ 0.26, & t = 3k+1 \\ -1, & t = 3k+2 \end{cases}, \quad b(t) = \begin{cases} 0.1, & t = 3k \\ 0.58, & t = 3k+1 \\ -2, & t = 3k+2 \end{cases}. \tag{4.1}$$

According to Lemma 2.2, the periodic transfer operator of system (2.1) can be derived as

$$G(\sigma, t) = \frac{b(t+1)\sigma^2 + b(t+2)\sigma + b(t)}{\sigma^3 - a(t)}. \tag{4.2}$$

Based on Lemma 2.1, the periodic transfer operator (4.2) can be decomposed into three time-invariant rational functions in σ^3. For $i = 0, 1, 2$, the sampled transfer functions are

$$\begin{aligned} H_0(\sigma^3, t) &= \frac{b(t)}{\sigma^3 - a(t)}, \\ H_1(\sigma^3, t) &= \frac{b(t+1)\sigma^3}{\sigma^3 - a(t+1)}, \\ H_2(\sigma^3, t) &= \frac{b(t+2)\sigma^3}{\sigma^3 - a(t+2)}. \end{aligned} \tag{4.3}$$

To obtain a representation with a time-invariant denominator, we define

$$d(\sigma, t) = \sigma^3 - a(t), \quad n(\sigma, t) = b(t+1)\sigma^2 + b(t+2)\sigma + b(t). \tag{4.4}$$

Pre-multiply both polynomials by

$$d_*(\sigma, t) = (\sigma^3 - a(t+1))(\sigma^3 - a(t+2)). \tag{4.5}$$

Then, the transfer operator becomes

$$G(\sigma, t) = [d_*(\sigma, t)d(\sigma, t)]^{-1} d_*(\sigma, t)n(\sigma, t), \tag{4.6}$$

and the new denominator is

$$d_*(\sigma,t)d(\sigma,t) = \sigma^9 - \varsigma_1\sigma^6 - \varsigma_2\sigma^3 - \varsigma_3, \tag{4.7}$$

with constant coefficients

$$\begin{aligned}
\varsigma_1 &= a(t) + a(t+1) + a(t+2),\\
\varsigma_2 &= -a(t)a(t+1) - a(t)a(t+2) - a(t+1)a(t+2),\\
\varsigma_3 &= a(t)a(t+1)a(t+2).
\end{aligned} \tag{4.8}$$

The numerator is

$$d_*(\sigma,t)n(\sigma,t) = \varpi_1(t)\sigma^8 + \cdots + \varpi_8(t)\sigma + \varpi_9(t), \tag{4.9}$$

with

$$\begin{aligned}
\varpi_1(t) &= b(t+1),\\
\varpi_2(t) &= b(t+2),\\
\varpi_3(t) &= b(t),\\
\varpi_4(t) &= -[a(t+1) + a(t+2)]b(t+1),\\
\varpi_5(t) &= -[a(t+1) + a(t+2)]b(t+2),\\
\varpi_6(t) &= -[a(t+1) + a(t+2)]b(t),\\
\varpi_7(t) &= a(t+1)a(t+2)b(t+1),\\
\varpi_8(t) &= a(t+1)a(t+2)b(t+2),\\
\varpi_9(t) &= a(t+1)a(t+2)b(t).
\end{aligned} \tag{4.10}$$

Then, we obtain

$$\theta = [-0.54, 0.408, -0.052, 0, 0.58, -2, 0.1, 0.4292, -1.48, 0.074, -0.1508, -0.52, -0.026].$$

The identification performance of Algorithm 1 (AM-ILM) is quantitatively evaluated in Figure 1, with precision metrics defined as

$$\varepsilon = \frac{||\theta(t) - \theta||^2}{||\theta||^2}. \tag{4.11}$$

To comprehensively evaluate the proposed Algorithm 1, we select several representative baseline methods: RLS and AM-RLS represent conventional recursive least-squares approaches; classical LM represents fixed-damping nonlinear least-squares; FCN and MAML serve as representatives of black-box modeling and rapid adaptation, respectively; and Bayesian methods represent probabilistic inference frameworks. The comparative results are listed in Table 1, and the performance metrics of the algorithms are illustrated in Figure 2 and tabulated in Table 2.

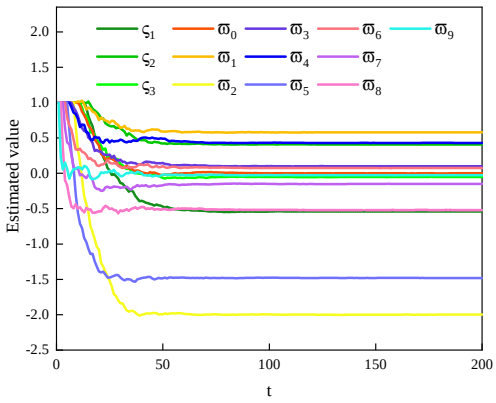

Figure 1. Evolution of parameter estimates ς_i and $\varpi_j(t)$ over time.

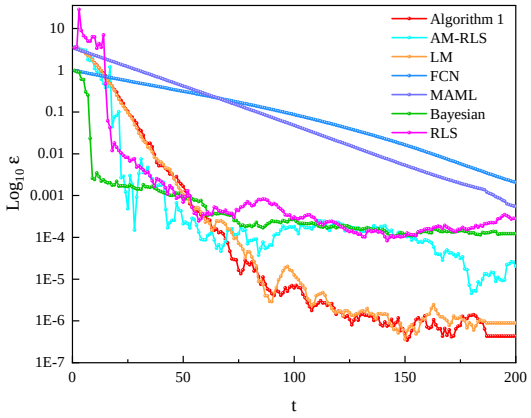

Figure 2. Comparison of identification accuracy ε for different algorithms.

Table 1. Estimated value of ς_i and $\varpi_j(t)$ by the different algorithms.

Algorithm	*m*	ς_1	ς_2	ς_3	$\varpi_0(t)$	$\varpi_1(t)$	$\varpi_2(t)$	$\varpi_3(t)$	$\varpi_4(t)$	$\varpi_5(t)$	$\varpi_6(t)$	$\varpi_7(t)$	$\varpi_8(t)$	$\varpi_9(t)$
AM-ILM	5	1	1	1	1	1	1	1	1	1	0.78781	0.43779	-0.16857	0.06242
	20	0.39764	0.82910	0.43663	0.39635	0.75221	-1.07983	0.30675	0.50551	-1.38721	0.10251	-0.22111	-0.53750	-0.03741
	50	-0.47089	0.42016	-0.06477	-0.02326	0.59076	-1.99629	0.14643	0.48942	-1.48514	0.10261	-0.17072	-0.50760	-0.02004
	100	-0.54075	0.40911	-0.05160	-0.00023	0.57918	-1.99627	0.10023	0.43239	-1.47614	0.07412	-0.15007	-0.51820	-0.02517
	150	-0.54047	0.40708	-0.05198	-0.00071	0.57966	-1.99978	0.10007	0.42964	-1.48039	0.07369	-0.15096	-0.51979	-0.02738
	200	-0.54024	0.40795	-0.05187	-0.00005	0.58012	-1.99998	0.09975	0.42975	-1.48112	0.07393	-0.15082	-0.52003	-0.02613
AM-RLS	200	-0.54139	0.40877	-0.0524	-0.00005	0.57995	-2.00091	0.10273	0.42909	-1.48093	0.07598	-0.15079	-0.51988	-0.02534
LM	200	-0.53954	0.40684	-0.05292	-0.00099	0.57940	-2.00090	0.10069	0.42993	-1.47912	0.07358	-0.15074	-0.52052	-0.02692
FCN	200	-0.54368	0.40309	-0.05755	0.00712	0.57825	-1.87956	0.10734	0.43353	-1.45258	0.07537	-0.1566	-0.51401	-0.02865
MAML	200	-0.53501	0.39881	-0.04865	0.00917	0.57765	-1.95109	0.10526	0.41849	-1.47036	0.10678	-0.14682	-0.50573	-0.02752
Bayesian	200	-0.54532	0.41197	-0.05437	-0.00334	0.58406	-2.00745	0.09626	0.42029	-1.49294	0.08214	-0.17057	-0.52180	-0.01731
RLS	200	-0.52656	0.40217	-0.05698	0.00023	0.58253	-1.98895	0.07258	0.41940	-1.45440	0.06303	-0.15748	-0.51288	-0.02519
	true value	-0.54	0.408	-0.052	0	0.58	-2	0.1	0.4292	-1.48	0.074	-0.1508	-0.52	-0.026

Table 2. Comparison of metrics for algorithms.

Algorithm	*MAE*	*MSE*	*RMSE*	*MAPE*
AM-ILM	0.000193	5.98E-07	0.000773	0.000193
AM-RLS	0.000674	7.9E-07	0.000887	0.000365
LM	0.000235	1.27E-06	0.001129	0.000235
FCN	0.000277	7.21E-07	0.000848	0.00027688
MAML	0.000797	1.09E-05	0.003299	0.000797
Bayesian	0.003486	5.05E-04	0.022472	0.003486
RLS	0.029492	0.001497	0.038398	0.014099

Figures 1 and 2 and Tables 1 and 2 jointly demonstrate that Algorithm 1 outperforms the AM-RLS, LM, FCN, MAML, Bayesian, and RLS algorithms in terms of convergence speed, final accuracy, and multiple error metrics (mean absolute error (MAE), mean squared error (MSE), root mean squared error (RMSE), and mean absolute percentage error (MAPE)), and has the smallest fluctuations.

Next, we examine the applicability of Algorithm 1 with respect to different initial conditions and varying system parameters.

Example 4.2. *Consider a periodic system of the form* (2.1) *with period* $T = 4$ *and time delay* $d = 2$. *The system matrices are given by*

$$a(t) = \begin{cases} 0.15, & t = 4k \\ 0.32, & t = 4k+1 \\ -0.8, & t = 4k+2 \\ 0.45, & t = 4k+3 \end{cases}, \quad b(t) = \begin{cases} 0.08, & t = 4k \\ 0.62, & t = 4k+1 \\ -1.5, & t = 4k+2 \\ 0.23, & t = 4k+3 \end{cases}. \tag{4.12}$$

Subsequently, we can obtain

$$d(\sigma, t) = \sigma^4 - a(t), \tag{4.13}$$

$$n(\sigma, t) = b(t+2)\sigma^3 + b(t+3)\sigma^2 + b(t)\sigma + b(t+1), \tag{4.14}$$

$$d_*(\sigma, t) = (\sigma^4 - a(t+1))(\sigma^4 - a(t+2))(\sigma^4 - a(t+3)). \tag{4.15}$$

Hence, θ is given by

$$\theta = [0.12, 0.4765, -0.186, 0.01728, 0, -1.5, 0.23, 0.08, 0.62, -0.045, 0.069,$$
$$0.0024, 0.0186, 0.708, -0.1085, -0.03776, -0.2926, -0.1728, 0.0265, 0.009, 0.07].$$

During the numerical simulation process, a higher noise level (for example, $\delta^2 = 0.5$) would reduce the stability and accuracy of Algorithm 1. This indicates the expected trade-off relationship between noise intensity and recognition performance. Therefore, in this case, a lower noise condition ($\delta^2 = 0.1$) was chosen for the simulation. Figure 3 and Table 3 present the identification results of Algorithm 1.

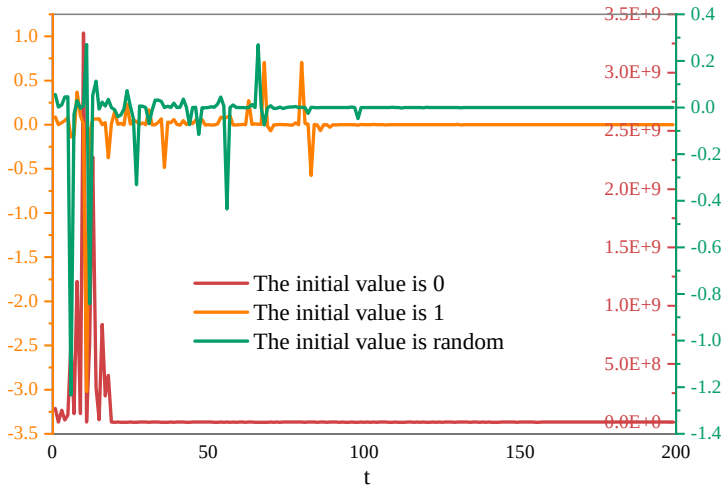

Figure 3. Comparison of identification accuracy ε for different initial conditions.

Table 3. Comparison of metrics for different initial conditions.

Initial value	*MAE*	*MSE*	*RMSE*
0	6.60635e-03	3.82230E-03	6.18248e-02
1	1.95454e-02	9.65976e-3	9.82841e-02
Random	6.70325e-03	1.8355e-3	4.28432e-02

Figure 3 and Table 3 present the identification results of Algorithm 1 under different initial conditions. When the initial value is 0, Algorithm 1 diverges severely in the first 20 steps and then converges to a level similar to other cases.

Besides, the computational complexity of Algorithm 1 is dominated by the matrix inversion in the gain computation step (Eq. (3.11)), requiring $O(n^3)$ operations per iteration, where n is the number of unknown parameters. AM-RLS and classical RLS achieves $O(n^2)$ complexity due to its recursive structure, while standard LM shares the same $O(n^3)$ complexity as AM-ILM but requires more iterations to converge. Although AM-ILM has higher per-iteration cost than RLS, it converges in significantly fewer iterations (typically 50-100 vs. 300-600 for RLS), resulting in comparable total runtime (approximately 50 ms on a standard CPU), which is sufficient for real-time applications in periodic manufacturing systems.

Based on the above content, the following key observations are made:

- The estimated parameters asymptotically converge to their true values with increasing time steps t, despite significant initial fluctuations. All of the above algorithms achieve stabilization of identification errors after a finite number of iterations.
- While Algorithm 1 and AM-RLS, LM, FCN, MAML, Bayesian, and RLS algorithms successfully identify parameters in linear discrete periodic systems with time delays, Algorithm 1 exhibits consistently higher accuracy.
- The MAPE values for initial condition 1 (9.69%) and random initialization (6.55%) are both below 10%, demonstrating high identification accuracy regardless of initialization. The slightly better performance of random initialization suggests stochastic perturbations help escape poor local optima in early iterations.

These results indicate that Algorithm 1 exhibits superior accuracy for parameter identification in linear discrete periodic time-delay systems. Moreover, Algorithm 1, LM, AM-RLS, and RLS demonstrate significantly better stability than other compared methods. In contrast, the remaining algorithms operate near critical stability, requiring careful evaluation in practical applications. Thus, the proposed auxiliary model-based improved Levenberg-Marquardt algorithm (Algorithm 1) effectively addresses parameter identification challenges in such systems.

In the following, in order to verify the applicability of Algorithm 1 in practical engineering, it is applied to the parameter identification of the excitation system. The main object we are considering is the excitation system in large-scale hydropower stations. The stable operation of the excitation system can enhance the convergence and stability of the equipment. The mathematical model of a typical AC excitation system is shown in Figure 4.

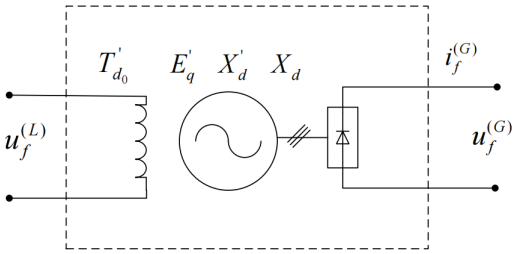

Figure 4. Mathematical model of the AC excitation system.

In this section, the initial reference value of the considered system is set to the calibrated terminal voltage of the generator. Starting from the electromagnetic power deviation ΔP, the voltage deviation ΔU_t, and the angular velocity $\Delta\omega$ of the excitation system, these three parameters are taken as the state variables of the system state-space equation. Based on the discussions and analyses in the existing literature, the state equation of the linearized state-space representation of the system can be written as

$$\dot{X} = AX + BU \tag{4.16}$$

with

$$A = \begin{bmatrix} \frac{S_E - S_V}{T_d S_V} & S'_E & -\frac{R_V S_E}{T_d S_V} \\ -\frac{\omega_0}{H} & -\frac{D}{H} & 0 \\ \frac{S_E - S_V}{T_d R_V S_V} & \frac{S'_E - S_V}{R_V} & -\frac{S_E}{T_d S_V} \end{bmatrix},$$

$$B = \begin{bmatrix} \frac{R'_E}{T_{d0}} & 0 & \frac{R'_E}{T_{d0} R_V} \end{bmatrix}^{\mathrm{T}},$$

$$X = [\Delta P \quad \Delta\omega \quad \Delta V_t]^{\mathrm{T}},$$

$$U = \Delta E_t.$$

In equation (4.16), H represents the moment of inertia of the system in seconds; T_{d0} is the time constant of the excitation winding when the stator winding is open; ΔV_t denotes the generator terminal voltage deviation; ΔE_t denotes the main generator excitation winding voltage; S_E, S'_E, R_E, and R'_E are constants related to the electromagnetic power deviation ΔP, as shown in equations (4.17) and (4.18).

$$\Delta P_e = S_E \Delta\delta + R_E \Delta E_q, \tag{4.17}$$

$$\Delta P_e = S'_E \Delta\delta + R'_E \Delta E'_q. \tag{4.18}$$

When analyzing and controlling the system, it is often necessary to discretize the system model. Here, the sampling period of the excitation system is set to T, the number of sampling periods in the system is set to L, and the angular frequency of the motor is ω_n. Then the following equation clearly holds:

$$T = \frac{2\pi}{L\omega_n}. \tag{4.19}$$

Meanwhile, let t denote the number of sampling times during discretization of the system. Then the following equations hold:

$$\begin{cases} X(t+1) = A(t)X(t) + B(t)U(t), \\ Y(t) = C(t)X(t). \end{cases} \tag{4.20}$$

The state-space equation shown above is the linear discrete periodic model of the excitation system after discretization. However, during the actual operation of the system, various uncertain factors often cause interference, leading to time delays in the system. Therefore, based on the periodic linear discrete model of the excitation system without disturbance, the time delay and noise interference of the system are further considered. The state-space model of the excitation system is then given as

$$\begin{cases} X(t+1) = A(t)X(t) + B(t)U(t-1), \\ Y(t) = C(t)X(t) + v(t), \end{cases} \tag{4.21}$$

where $X(t)$ is the state vector of the considered excitation system, $Y(t)$ is the output vector of the system, and $U(t)$ is the input vector of the system. The system matrices $A(t)$, $B(t)$, and $C(t)$ appearing in the state equation and output equation all have period T. Suppose $v(t)$ is a set of random white noise sequences with zero mean, and $U(t-1)$ represents the existence of time delay in the system, i.e., the state equation of the excitation system state-space model contains one unit of input time delay.

The simulation parameters of this system are

$$a(t) = \begin{cases} 0, & t = 2k \\ 2.7, & t = 2k+1 \end{cases}, \quad b(t) = \begin{cases} 1.4, & t = 2k \\ -0.7, & t = 2k+1 \end{cases}. \tag{4.22}$$

By the same token,

$$\theta = [2.7, 0, 0, 1.4, -0.7, -3.78, 1.89].$$

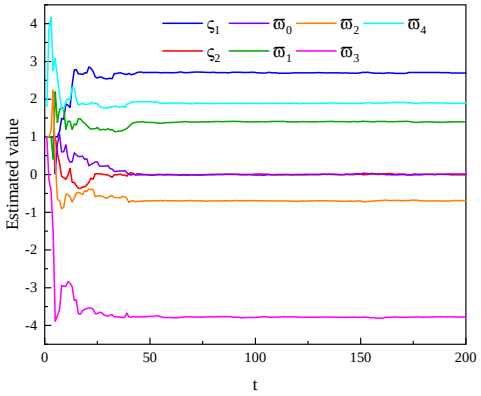

Figure 5. Evolution of parameter estimates ς_i and $\varpi_j(t)$ over time.

The identification results are shown in Figures 5 and 6. As illustrated in the figures, with the recursion of iteration t, the identification error decreases continuously, and the estimated values gradually approach their true values, which demonstrates the effectiveness of the proposed algorithm for parameter identification of the time-delay linear discrete periodic system in practical hydropower excitation

applications. It should be noted that this example is a simplified conceptual demonstration; practical application would require a more accurate excitation system model (e.g., IEEE standard models) and calibration with real operational data.

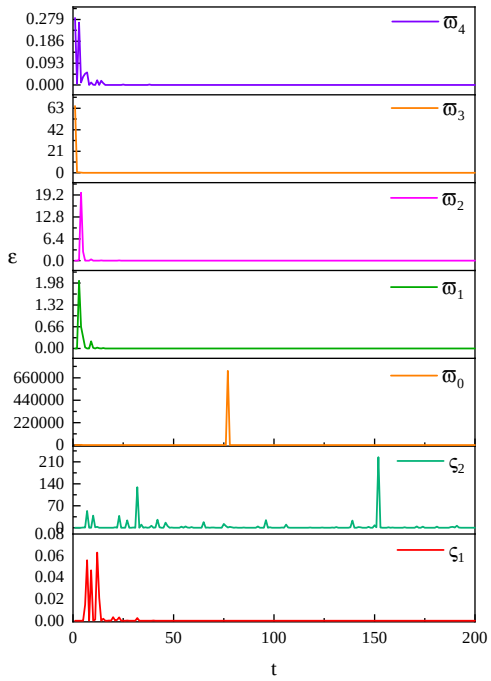

Figure 6. Comparison of identification accuracy ε for Algorithm 1.

5. Conclusions

This paper has proposed an auxiliary model-based improved Levenberg–Marquardt (AM-ILM) algorithm for parameter identification of linear discrete-time periodic systems with time delays. The AM-ILM algorithm integrates an auxiliary model that directly reconstructs unmeasurable latent variables from input–output data. These two innovations—adaptive regularization and auxiliary-model-based reconstruction—are not simultaneously available in any previous Levenberg–Marquardt variant. To comprehensively evaluate the proposed algorithm, we selected several representative baselines covering different paradigms: RLS and AM-RLS for conventional recursive least-squares, classical LM for fixed-damping nonlinear least-squares, FCN and MAML for black-box and meta-learning approaches, and Bayesian methods for probabilistic inference. Numerical simulations demonstrate that the proposed algorithm achieves superior accuracy, faster convergence, and improved stability compared to these benchmark methods. In particular, compared with the existing AM-RLS method [29], the proposed AM-ILM achieves faster convergence and lower steady-state error under the same conditions. Future work may integrate Bayesian optimization or neural networks [33] for multi-model fusion, and extend the framework to Wiener nonlinear systems using metaheuristic optimization (e.g., zebra optimization algorithm [34]). Furthermore, we have included validation on real-field data as a priority future work.

Author contributions

Zhuo Chen: Conceptualization, Resources, Supervision, Methodology, Validation, Writing-original draft, Software. Yingying Guo: Conceptualization, Methodology, Data curation, Visualization, Formal analysis, Writing-original draft, Investigation, Validation. Zhuocheng Zou: Conceptualization, Methodology, Investigation, Writing-review & editing. Zhen Zhang: Supervision, Validation, Writing-review & editing.

Conflict of interest

The authors declare that there exists no conffict of interest that could potentially compromise the neutrality of the reported research.

Use of Generative-AI tools declaration

The authors declare that they did not utilize any artificial intelligence (AI) tools in the creation of this article.

References

1. J. H. Li, T. C. Zong, G. P. Lu, Parameter identification of Hammerstein-Wiener nonlinear systems with unknown time delay based on the linear variable weight particle swarm optimization, *ISA Trans.,* **120** (2022), 89–98. http://dx.doi.org/10.1016/j.isatra.2021.03.021
2. J. Yuan, Y. C. Ding, S. M. Fei, Y. Q. Chen, Identification and parameter sensitivity analyses of time-delay with single-fractional-pole systems under actuator rate limit effect, *Mech. Syst. Signal Process.,* **163** (2022), 108111. https://doi.org/10.1016/j.ymssp.2021.108111
3. Z. S. Wang, C. Y. Wang, L. H. Ding, Z. Wang, S. N. Liang, Parameter identification of fractional-order time delay system based on Legendre wavelet, *Mech. Syst. Signal Process.,* **163** (2022), 108141. https://doi.org/10.1016/j.ymssp.2021.108141
4. C. Y. Liu, C. Sun, Robust parameter identification of a nonlinear impulsive time-delay system in microbial fed-batch process, *Appl. Math. Model.,* **111** (2022), 160–175. https://doi.org/10.1016/j.apm.2022.06.032
5. L. H. Zhu, Y. X. Tang, S. L. Shen, Pattern study and parameter identification of a reaction-diffusion rumor propagation system with time delay, *Chaos Sol. Fract.,* **166** (2023), 112970. https://doi.org/10.1016/j.chaos.2022.112970
6. D. W. Hu, Z. Q. Zhang, Four positive periodic solutions of a discrete time delayed predator–prey system with nonmonotonic functional response and harvesting, *Comput. Math. Appl.,* **56** (2008), 3015–3022. https://doi.org/10.1016/j.camwa.2008.09.009
7. C. E. D. Souza, D. Coutinho, Robust stability and control of uncertain linear discrete-time periodic systems with time-delay, *Automatica,* **50** (2014), 431–441. https://doi.org/10.1016/j.automatica.2013.11.038

8. L. N. Wei, W. H. Chen, S. X. Luo, Stabilization of discrete-time switched linear systems with time-varying delays via nearly-periodic impulsive control, *J. Franklin Inst.,* **356** (2019), 8996–9022. https://doi.org/10.1016/j.jfranklin.2019.07.003
9. S. Gao, H. H. Guo, T. R. Chen, The existence of periodic solutions for discrete-time coupled systems on networks with time-varying delay, *Physica A.,* **526** (2019), 120876. https://doi.org/10.1016/j.physa.2019.04.112
10. E. Aranda-Escolástico, L. J. Colombo, M. Guinaldo, Periodic event-triggered targeted shape control of Lagrangian systems with discrete-time delays, *ISA Trans.,* **117** (2021), 139–149. https://doi.org/10.1016/j.isatra.2021.01.048
11. J. J. Moré , The Levenberg-Marquardt algorithm: Implementation and theory, *Numer. Anal.*, **105** (1978), 105–116. https://doi.org/10.1007/BFb0067700
12. G. Li, N. S. S. Win, M. L. Fan, J. T. Li, L. Lin, Enhance registration precision of transmission breast images utilizing improved Levenberg-Marquardt optimization algorithm with normalized cross-correlation, *Comput. Biol. Med.*, **186** (2025), 109654. https://doi.org/10.1016/j.compbiomed.2025.109654
13. S. S. Tong, M. Y. Liu, H. Y. Shu, P. P. Wang, Adaptive Levenberg–Marquardt-Kaczmarz method for nonlinear ill-posed problems with non-smooth constraints, *J. Comput. Appl. Math.*, **463** (2025), 116504. https://doi.org/10.1016/j.cam.2025.116504
14. J. Y. Tang, J. C. Zhou, A Levenberg–Marquardt type algorithm with a Broyden-like update technique for solving nonlinear equations, *J. Comput. Appl. Math.*, **460** (2025), 116401. https://doi.org/10.1016/j.cam.2024.116401
15. C. R. Mirza, M. Abbas, S. A. Idris, Y. Khan, A. Alameer, A. B. Rajab, et al., Intelligent computing technique to analyze the two-phase flow of dusty trihybrid nanofluid with Cattaneo-Christov heat flux model using Levenberg-Marquardt Neural-Networks, *Case Stud. Therm. Eng.*, **68** (2025), 105891. https://doi.org/10.1016/j.csite.2025.105891
16. H. Qureshi, U. Khaliq, Z. Shah, H. Abutuqayqah, M. Waqas, S. Saleem, et al., Artificial intelligence-based analysis employing Levenberg Marquardt neural networks to study chemically reactive thermally radiative tangent hyperbolic nanofluid flow considering Darcy-Forchheimer theory, *J. Radiat. Res. Appl. Sci.*, **18** (2024), 101253. https://doi.org/10.1016/j.jrras.2024.101253
17. Z. Shah, S. Alzhrani, M. A. Z. Raja, A. A. Pasha, F. Shahzad, W. A. Khan, Stochastic analysis through Levenberg Marquardt backpropagation neural networks for radiative Carreau nanofluid flow subject to chemical reaction, *Ain Shams Eng. J.*, **15** (2024), 103100. https://doi.org/10.1016/j.asej.2024.103100
18. N. A. Albasheir, Z. Shah, M. A. Z. Raja, A. A. Touati, M. M. A. Almazah, M. Jawaid, etal., Ai-powered Levenberg-Marquardt neural networks for implementation of novel flux conditions in Carreau nanomaterial stagnation-point flow magnetic considering field and heat source effects, *Measurement*, **243** (2025), 116304. https://doi.org/10.1016/j.measurement.2024.116304
19. D. C. Nautiyal, S. Tripathi, H. S. Sahu, Parameter estimation of solar PV module using levenberg-marquardt method for maximum power point calculation, *IFAC-Pap.*, **57** (2024), 373–378. https://doi.org/10.1016/j.ifacol.2024.05.064
20. B. Y. Liu, W. F. Chen, Time delay and model parameter estimation for nonlinear system with simultaneous approach, *J. Process Control*, **139** (2024), 103234. https://doi.org/10.1016/j.jprocont.2024.103234

21. X. Y. Zhang, G. Y. Gao, Z. W. Fu, Y. Li, B. Han, A frozen Levenberg-Marquardt-Kaczmarz method with convex penalty terms and two-point gradient strategy for ill-posed problems, *Appl. Numer. Math.*, **209** (2025), 187–207. https://doi.org/10.1016/j.apnum.2024.11.014
22. E. Boos, D. S. Gonçalves, F. S. V. Bazán, Levenberg-Marquardt method with singular scaling and applications, *Appl. Math. Comput.*, **474** (2024), 128688. https://doi.org/10.1016/j.amc.2024.128688
23. N. Zhao, S. Gorbachev, D. Yue, C. X. Dou, X. P. Xie, The effect of communication delays on the frequency stability of power systems integrated with inverter air conditioners, *Sustain. Energy Grids Netw.*, **32** (2022), 100920. https://doi.org/10.1016/j.segan.2022.100920
24. M. Breakspear, Dynamic models of large-scale brain activity, *Nat. Neurosci.*, **20** (2017), 340–352. https://doi.org/10.1038/nn.4497
25. Y. S. Zhao, Y. Ji, Separable synchronous auxiliary model adaptive momentum estimation strategy for a time-varying system with colored noise from on-line measurements, *ISA Trans.*, **157** (2024), 213–223. https://doi.org/10.1016/j.isatra.2024.11.048
26. F. Ding, L. Xu, X. Zhang, Hierarchical generalized extended parameter identification for multivariable equation-error ARMA-like systems by using the filtering identification idea, *Annu. Rev. Control.*, **60** (2025), 100993. https://doi.org/10.1016/j.arcontrol.2025.100993
27. F. Ding, L. Xu, P. Liu, X. F. Wang, Two-stage parameter estimation methods for linear time-invariant continuous-time systems, *Syst. Control Lett.*, **204** (2025), 106166. https://doi.org/10.1016/j.sysconle.2025.106166
28. F. Ding, X. L. Luan, L. Xu, Hierarchical recursive gradient parameter identification for multi-input ARX systems with partially-coupled information vectors, *Int. J. Adapt. Control Signal Process.*, **39** (2025), 1978–1995. https://doi.org/10.1002/acs.4036
29. L. L. Lv, J. L. Zhao, B. Q. Zheng, J. W. Shen, H. C. Yan, A recursive identification algorithm for discrete time-delay periodic linear systems, *J. Comput. Appl. Math.*, **461** (2025), 116447. https://doi.org/10.1016/j.cam.2024.116447
30. S. Bittanti, P. Colaneri, *Periodic systems: Filtering and control,* Springer, **1** (2009), 61–80. https://dl.acm.org/doi/book/10.5555/1521564
31. L. L. Lv, J. B. Chen, Z. Zhang, B. W. Wang, L. Zhang, A numerical solution of a class of periodic coupled matrix equations, *J. Franklin Inst.*, **358** (2021), 2039–2059. https://doi.org/10.1016/j.jfranklin.2020.11.022
32. Y. F. Tian, Y. Yang, X. Y. Ma, X. J. Su, Stability of discrete-time delayed systems via convex function-based summation inequality, *Appl. Math. Lett.*, **145** (2023), 108764. https://doi.org/10.1016/j.aml.2023.108764
33. F. Ding, Y. S. Xiao, L. Xu, Z. M. Fang, Hierarchical stochastic gradient and hierarchical multi-innovation stochastic gradient identification for multivarible ARX models, *Int. J. Adapt. Control Signal Process.*, **40** (2026), 79–102. https://doi.org/10.1002/acs.4081
34. L. L. Lv, Y. L. Zhang, Y. Wu, C. B. Xu, Z. Y. Li, H. C. Yan, Parameter estimation of wiener nonlinear systems based on gradient iteration theory and zebra optimization algorithm, *J. Franklin Inst.*, **363** (2026), 108499. https://doi.org/10.1016/j.jfranklin.2026.108499

AIMS Press