Original research paper

Data-driven parameter tuning for rational feedforward controller: Achieving optimal estimation via instrumental variable

Weicai Huang1,2 | Kaiming Yang1,2 | Yu Zhu1,2 | Sen Lu1,2

1 State Key Laboratory of Tribology, Department of Mechanical Engineering, Tsinghua University, Beijing, P.R. China
2 Beijing Key Laboratory of Precision/Ultra-Precision Manufacturing Equipments and Control, Tsinghua University, Beijing, P.R. China

Correspondence
Kaiming Yang, State Key Laboratory of Tribology, Department of Mechanical Engineering, Tsinghua University, Beijing 100084, P.R. China.
Email: yangkm@tsinghua.edu.cn

Funding information
the National Science and Technology Major Project of China, Grant/Award Number: 2017ZX02102004

Abstract
Feedforward control has been widely used to improve the tracking performance of precision motion systems. This paper develops a new data-driven feedforward tuning approach associated with rational basis functions. The aim is to obtain the global optimum with optimal estimation accuracy. First, the instrumental variable is employed to ensure the unbiased estimation of the global optimum. Then, the optimal instrumental variable which leads to the highest estimation accuracy is derived, and a new refined instrumental variable method is exploited to estimate the optimal instrumental variable. Moreover, the estimation accuracy of the optimal parameter is further improved through the proposed parameter updating law. Simulations are conducted to test the parameter estimation accuracy of the proposed approach, and it is demonstrated that the global optimum is unbiasedly estimated with optimal parameter estimation accuracy in terms of variance with the proposed approach. Experiments are performed and the results validate the excellent performance of the proposed approach for varying tasks.

1 | INTRODUCTION

The general trend of precision motion systems is that the requirements related to throughput and processing quality are ever increasing [1, 2]. Feedforward control is widely used to meet these requirements since feedforward can compensate for the known disturbances before they affect the systems [3, 4]. Traditionally, the feedforward controller is designed to approximate the inverse of the parametric model of the controlled system, examples include zero-phase error-tracking control [5, 6], high-order feedforward control [7] and inversion-based control [8]. The performances of these methods are highly dependent on the quality of the parametric model of the plant. The achievable performance is restricted by the inevitable model uncertainties and modelling error, especially for precision motion systems which are generally with complex dynamics [9].

Compared to traditional methods, data-driven methods are carried out using the measured data directly, and no accurate parametric model is required [10, 11]. Consequently, the restrictions resulted from the usage of parametric model are greatly mitigated. In data-driven methods, the feedforward tuning problem is converted into the feedforward signal or feedforward controller optimisation problem. According to the object to be optimised, the data-driven methods can be divided into signal-based methods and filter-based methods, which optimise the feedforward signal and feedforward controller, respectively.

A typical signal-based method is iterative learning control (ILC) [12–14], which exploits signals measured from the last trial to generate the feedforward signal for the next trial. Many applications have been reported where the application of ILC leads to significant performance improvement [15–17]. Moreover, many signal-based methods are developed based on ILC, such as segment ILC [18], projection-based ILC [19] and nonlinear ILC [20]. Another widely used signal-based method is signal decomposition [21, 22], in which the feedforward signal is linearly parameterised and the optimal coefficients are pursued. Superior performance can be achieved using signal-based methods. However, the feedforward signal obtained from signal-based methods is optimal only for a specific reference. Therefore, performance deterioration is inevitable while executing non-repeating tasks, that is, the extrapolation properties of signal-based methods are poor.

Compared to signal-based methods, filter-based methods achieve better extrapolation properties while maintaining the
high tracking performance. Specifically, the feedforward controller is parameterised in filter-based methods and the optimal parameter is determined through parameter optimisation algorithms [23, 24]. A widely used strategy is to parameterise the feedforward controller linearly using polynomial basis functions (PBFs) [25–28]. This linear parameterisation guarantees the stability of the feedforward controller inherently and makes the parameter optimisation problem convex. These merits facilitate the application of PBFs and this approach has been extended towards input shaping [29] and multivariable systems [11, 30]. The main drawback of PBFs is that the resulting feedforward controller can only approximate the inverse of the plant with unit denominator, which severely limits the achievable performance and extrapolation properties. A more promising strategy is to parameterise the feedforward controller using rational basis functions (RBFs) [31, 32]. Compared to PBFs, the feedforward controller parameterised using RBFs is able to approximate the inverse plants of general rational motion systems. Hence, higher tracking performance and extrapolation properties are available. However, the parameter optimisation problem associated with RBFs is non-convex. Pre-existing approaches seek the optimal parameter through non-linear optimisation approach [31], which cannot ensure the global optimisation in general.

Besides, instrumental variable is an important component in many filter-based methods. The application of the instrumental variable generally leads to unbiased estimation of the optimal parameter [3, 27], and the parameter estimation accuracy in terms of variance can be optimised [28, 33]. However, in pre-existing filter-based approaches associated with the instrumental variable, the noise in the estimated regressor is ignored while calculating the covariance matrix. Therefore, the achieved estimation accuracy after optimisation is non-optimal actually.

Based on the instrumental variable, this paper proposes a novel data-driven parameter tuning approach for feedforward controller parameterised using RBFs. The proposed approach aims at obtaining the global optimum with optimal estimation accuracy. Specifically, an analysis framework based on the instrumental variable is proposed, which leads to the unbiased estimation of the optimal parameter inherently. The estimation accuracy in terms of variance is analysed, in which the noise in the estimated regressor will be taken into consideration to ensure optimal parameter estimation accuracy after optimisation. The optimal instrumental variable is determined such that the covariance matrix is minimised. A refined instrumental variable (RIV) approach is proposed to estimate the optimal instrumental variable. Moreover, a novel parameter updating law is proposed to further improve the estimation accuracy of the global optimum based on the above optimisation. The performance of the proposed approach is tested through simulations and experiments. The results indicate that the proposed approach not only estimates the global optimum unbiasedly with optimal accuracy but also achieves high performance for varying tasks.

The preliminary results have been published in [34], in which the least square method is exploited to estimate the optimal parameter unbiasedly for feedforward controller parameterised using RBFs. This paper greatly improves the parameter estimation accuracy, compared to [34], through the introduction of an instrumental variable. As shown here in the simulations and experiments, the improved parameter estimation accuracy of the global optimum with the proposed approach contributes to higher reference-tracking performance and extrapolation properties.

This paper is organised as follows. In Section 2, the notations used here are provided. In Section 3, the problem to be solved is thoroughly stated. In Section 4, the optimal parameter is obtained with an optimal accuracy using the instrumental variable. In Section 5, the proposed parameter updating law which further improves the parameter estimation accuracy is introduced in detail. In Section 6, simulations are conducted to test the parameter estimation accuracy of the proposed approach. In Section 7, the reference-tracking performance and extrapolation properties of the proposed approach are compared with pre-existing approaches via experiments. In Section 8, the overall conclusions of this paper are given.

2 | PRELIMINARIES

This paper considers only discrete-time, linear time-invariant and single-input single-output (SISO) systems. Finite-time task is performed and signals are assumed to be of length \( N \in \mathbb{Z}^+ \). A positive definite matrix \( A \) is denoted as \( A > 0 \). For a vector \( u \in \mathbb{R}^N \), the weighted two-norm is given by \( \| u \|^2_w = u^T W u \), with \( (\cdot)^T \) the transpose of a matrix or vector. Besides, the \( k \)th element of \( u \) is expressed as \( u[k] \) when zero initial condition is assumed. As finite-time task with length \( N \) is performed, it follows that \( u[k] = 0 \) when \( k < 0 \) and \( k > N-1 \). Therefore, the relation between \( u \) and \( y \) can be rewritten as

\[
\begin{bmatrix}
y[0] \\
y[1] \\
\vdots \\
y[N-1]
\end{bmatrix} = \begin{bmatrix}
g[0] & 0 & \cdots & 0 \\
g[1] & g[0] & \ddots & \vdots \\
\vdots & \vdots & \ddots & 0 \\
g[N-1] & \cdots & g[1] & g[0]
\end{bmatrix} \begin{bmatrix}
u[0] \\
u[1] \\
\vdots \\
u[N-1]
\end{bmatrix},
\]

where \( G \in \mathbb{R}^{N \times N} \) is the lower triangular Toeplitz matrix corresponding to \( G(q^{-1}) \).

3 | PROBLEM FORMULATION

3.1 | System description

This paper employs the two degrees-of-freedom (2-DOFs) control system, as shown in Figure 1. The plant \( P \) is generally...
unknown. The feedback controller $C$ is applied for robust stability and noise rejection. Meanwhile, the feedforward controller $F$ is exploited to reduce the reference-tracking error. In Figure 1, $r$ is a finite-time reference with length $N_r$, $y$ is the measured output, and $e = r - y$ is the tracking error. Besides, the noise $v$ is given by $v = H(q^{-1})w$, where $w$ is normally distributed white noise with zero mean and variance $\lambda^2$, and $H(q^{-1})$ is a monic and stable transfer function.

From Figure 1, the tracking error $e$ is given by
\[ e = Sr - F \cdot S_r r - S_v \]  
where $S = 1/(1 + RC)$ is the sensitivity function and $S_r = P/(1 + RC)$ is the process sensitivity function. It is observed from (2) that $e$ can be divided into two parts: the reference-induced error $e_r = Sr - F \cdot S_r r$ and the noise-induced error $e_v = -S_v$. Similar to Assumption 2.1 in [28], it is assumed that $C$ is designed such that $S(q^{-1})H(q^{-1}) = 1$, which makes $e_r = -w$.

### 3.2 Feedforward parameterisation

In filter-based feedforward tuning approaches, $F$ is parameterised and the feedforward tuning problem is converted into parameter optimisation problem. This paper parameterises $F$ using RBFs as follows:
\[ F(\theta, q^{-1}) = \frac{A(\theta^A, q^{-1})}{B(\theta^B, q^{-1})} \]  
and
\[
\begin{align*}
A(\theta^A, q^{-1}) &= \Psi^A \cdot \theta^A = \sum_{i=1}^{n_A} \Psi^A_i(q^{-1}) \cdot \theta^A[i], \\
B(\theta^B, q^{-1}) &= 1 + \Psi^B \cdot \theta^B = 1 + \sum_{i=1}^{n_B} \Psi^B_i(q^{-1}) \cdot \theta^B[i],
\end{align*}
\]
where $\Psi = [\Psi^A, \Psi^B]$ is the selected basis function vector with $\Psi^A = [\Psi^A_1, \Psi^A_2, \ldots, \Psi^A_{n_A}]$ and $\Psi^B = [\Psi^B_1, \Psi^B_2, \ldots, \Psi^B_{n_B}]$, and $\theta = [(\theta^A)^T, (\theta^B)^T]^T$ is the parameter vector to be optimised.

### 3.3 Problem statements

Generally, $F$ is designed to eliminate $e_r$ to enhance the reference-tracking performance. Ideally, $e_r = 0$ holds for arbitrary references when $F = P^{-1}$. Hence, good tracking performance and extrapolation properties can be achieved simultaneously by making $F$ as close as possible to $P^{-1}$. With the parameterisation in (3) and (4), $F$ is able to approximate the inverse plants of general rational motion systems. Therefore, the key to achieve high performance is to find the global optimum accurately. Biasedness and variance are two critical indicators of the parameter estimation accuracy, which are the main focuses of this paper. Therefore, the aim of this paper is to propose a novel data-driven parameter tuning approach for feedforward controller parameterised using (3) and (4), and the detailed requirements for the proposed approach are as follows:

(R1) achieve unbiased estimation of the global optimum;
(R2) achieve optimal estimation accuracy in terms of variance.

**Remark 1.** Indeed, the selection of the basis function $\Psi$ has great influence on the convergence performance of filter-based methods. However, similar to pre-existing literatures [23–34], the detailed selection method of $\Psi$ is out of the scope of this paper and will be an interesting topic for the future research. Hence, the focus of this paper is to accurately obtain the global optimum with the given basis function $\Psi$, and the widely used differentiators [25–28] and delay units [3, 11, 23, 24] are selected as the basis functions for the simulations and experiments, respectively.

### 4 Estimation of the optimal parameter based on instrumental variable

This paper introduces the proposed approach in detail in Sections 4 and 5, which constitutes the main contribution of this paper. For convenience of reading, the framework of Sections 4 and 5 is presented in Figure 2.
As shown in Figure 2, the expression of the global optimum will be given in (10) and (11) in Section 4.1 based on the feedforward parameterisation in (3) and (4). In Sections 4.2–4.4, the variables required in (10) and (11), including \( S_r \), \( S_f \) and the instrumental variable, are derived to make the global optimum calculable. Specifically, the estimation of \( S_r \) and \( S_f \) is realised in Section 4.2. The covariance matrix originated from the noise components in the estimated \( S_r \) and \( S_f \) is computed and minimised in Section 4.3. An RIV method is proposed in Section 4.4 to estimate the optimal instrumental variable obtained in Section 4.2. Substitute the \( S_r \) and \( S_f \) estimated in Section 4.2 and the optimal instrumental variable in Section 4.4 into (10) and (11), the global optimum can be obtained.

### 4.1 Determine the optimal parameter via instrumental variable

Ideally, the global optimum \( \theta_{opt} \) is the parameter such that \( e_r(\theta_{opt}) = 0 \). Combining the feedforward parameterisation in (3) and (4), we have that

\[
B^A(\theta^A_{opt}) S_r - A^T(\theta^A_{opt}) S_f = 0.
\]

(5)

In order to estimate \( \theta_{opt} \) unbiasedly from (5), the objective criterion based on the instrumental variable is given by

\[
V(\theta) = \left\| Z^T L(q^{-1}) \eta(\theta) \right\|_W^2,
\]

(6)

where \( Z \in \mathbb{R}^{N \times t_e} \) is the instrumental variable, \( L(q^{-1}) \) is the pre-filter, \( W > 0 \) is the weighting matrix and \( \eta(\theta) \) is given by

\[
\eta(\theta) = B(\theta^B) S_r - A(\theta^A) S_f.
\]

(7)

Substituting (4) into (7), we have that

\[
\eta(\theta) = S_r - \phi \theta,
\]

(8)

with the regressor \( \phi = [\phi^A, \phi^B] \) and

\[
\begin{align*}
\phi^A &= \begin{bmatrix}
\Psi^A S_f r_1, \Psi^A S_f r_2, \ldots, \Psi^A S_f r_{t_e}
\end{bmatrix},
\phi^B &= \begin{bmatrix}
-\Psi^B S_r, -\Psi^B S_f r_1, -\Psi^B S_f r_2, \ldots, -\Psi^B S_f r_{t_e}
\end{bmatrix}.
\end{align*}
\]

(9)

It is observed from (8) that \( \eta(\theta) \) is linear with respect to \( \theta \), which implies that \( V(\theta) \) is a quadratic function of \( \theta \). Therefore, \( \theta_{opt} \) can be directly solved as the analytical solution to \( \partial V(\theta)/\partial \theta = 0 \), which yields

\[
\theta_{opt} = \left( R_{\phi \phi}^T W R_{\phi \phi} \right)^{-1} R_{\phi \phi}^T W R_{\phi \phi}.
\]

(10)

where

\[
\begin{align*}
R_{\phi \phi} &= Z^T L(q^{-1}) \phi,
R_{\phi \phi} &= Z^T L(q^{-1}) S_f.
\end{align*}
\]

(11)

**Remark 2.** The motivations of selecting \( V(\theta) \) in (6) as the objective criterion are twofold. First, the resulting optimisation problem is convex and the global optimum \( \theta_{opt} \) can be directly obtained from (10) and (11), which is different from the non-convex optimisation problem suffered when the norm of the reference-induced error is optimised directly. Second, the introduction of the instrumental variable in (6) is crucial for the optimisation of parameter estimation accuracy in Sections 4 and 5.

### 4.2 Estimation of \( S_r \) and \( S_f \)

It is observed from (9) to (11) that the calculation of \( \theta_{opt} \) requires the knowledge of \( Z, L(q^{-1}), W, \Psi, S_r \) and \( S_f \). Of all these variables, \( Z, L(q^{-1}), W \) will be determined by minimising the covariance matrix in the next subsection, the basis function \( \Psi \) is selected by the user in advance, but \( S_r \) and \( S_f \) cannot be given directly due to the unknown plant. This paper estimates \( S_r \) and \( S_f \) using the measured data from a reference-tracking trial marked as \( E_0 \). Specifically, \( E_0 \) is carried out without feedforward control, that is, \( F = 0 \). Define \( e_0 \) and \( y_0 \) as the tracking error and measured output of \( E_0 \), they are given by

\[
\begin{align*}
e_0 &= S_r - S_0,
y_0 &= C \cdot S_f + S_0.
\end{align*}
\]

(12)

with \( n_0 = H(q^{-1}) w_0 \) the measurement noise in \( E_0 \). Define \( \hat{S}_r \) and \( \hat{S}_f \) as the estimated \( S_r \) and \( S_f \), respectively. Ignore the noise term \( \tilde{S}_0 \) in (12), \( \hat{S}_r \) and \( \hat{S}_f \) can be directly solved as follows:

\[
\begin{align*}
\hat{S}_r &= e_0,
\hat{S}_f &= C^{-1} y_0.
\end{align*}
\]

(13)

Combining (12) and (13), the relation between the estimated values and real values is given by

\[
\begin{align*}
\hat{S}_r &= S_r - S_0,
\hat{S}_f &= S_f + C^{-1} S_0.
\end{align*}
\]

(14)

Considering that \( e_0 \) is with zero mean, \( \hat{S}_r \) and \( \hat{S}_f \) obtained from (13) are the unbiased estimations of \( S_r \) and \( S_f \), respectively. Combining the preselected \( \psi \) and the instrumental variable determined in the next subsection, the optimal parameter can be calculated through (10).
4.3 Estimation accuracy of the optimal parameter

In the following description, \( \hat{\phi}, R_{\phi \phi}, R_{\phi}, \) and \( \hat{\theta}_{\text{opt}} \) denote the estimated variables corresponding to the noise-free variables \( \phi, R_{\phi \phi}, R_{\phi}, \) and \( \theta_{\text{opt}} \). Specifically, noise-free variables are the real values calculated based on real variables \( \hat{s}_r \) and \( \hat{s}_p \), while the estimated variables are calculated based on the estimated variables \( \hat{s}_r \) and \( \hat{s}_p \) in (13). Therefore, noise components induced by \( n_0 \) in (14) are included in all the estimated variables, which is the main difference between the estimated variables and their corresponding noise-free variables.

The difference between \( \hat{\theta}_{\text{opt}} \) and \( \theta_{\text{opt}} \) is given by

\[
\hat{\theta}_{\text{opt}} - \theta_{\text{opt}} = \left( R_{\phi \phi}^T W R_{\phi \phi} \right)^{-1} R_{\phi \phi}^T W ( R_{\phi} - R_{\phi \phi} \theta_{\text{opt}} ),
\]

(15)

Define \( \hat{\phi} = \phi - \phi \) as the noise-induced component in the estimated regressor \( \hat{\phi} \). Combining (9) and (14), we have that \( \hat{\phi} = [\hat{\phi}^A, \hat{\phi}^B] \) and

\[
\begin{cases}
\hat{\phi}^A = [\psi_1^A, \psi_2^A, \ldots, \psi_n^A] C^{-1} \hat{s}_q, \\
\hat{\phi}^B = [\psi_1^B, \psi_2^B, \ldots, \psi_n^B] \hat{s}_q.
\end{cases}
\]

(16)

It is observed from (5) that \( s_r = \phi \hat{\phi}_{\text{opt}} \). Combining with (11) and (16), it follows that

\[
R_{\phi} - R_{\phi \phi} \theta_{\text{opt}} = -Z^T ( L(q^{-1}) K(q^{-1}) w_0 ),
\]

(17)

where

\[
K(q^{-1}) = A ( \theta_{\text{opt}}^A ) : C^{-1} + B ( \theta_{\text{opt}}^B ).
\]

(18)

The asymptotic distribution of \( \hat{\theta}_{\text{opt}} \) is given by

\[
\sqrt{N} \left( \hat{\theta}_{\text{opt}} - \theta_{\text{opt}} \right) \xrightarrow{\text{dist}} N(0, \Sigma_{\theta}),
\]

(19)

where \( N(0, \Sigma_{\theta}) \) denotes the normal distribution with zero mean, \( \Sigma_{\theta} \) is the covariance matrix, which is typically used as the performance index in terms of parameter estimation accuracy and can be minimised to improve the parameter estimation accuracy [28]. Based on (15) and (17), we have that

\[
P_{\theta} = E \left( \hat{\theta}_{\text{opt}} - \theta_{\text{opt}} \right) \left( \hat{\theta}_{\text{opt}} - \theta_{\text{opt}} \right)^T
= \lambda^2 \left( R_{\phi \phi}^T W R_{\phi \phi} \right)^{-1} R_{\phi \phi}^T W Z^T L K^T L^T Z W^T R_{\phi \phi} \left( R_{\phi \phi}^T W R_{\phi \phi} \right)^{-1},
\]

(20)

where \( E \) denotes the expectation with respect to the noise \( r \), and \( L \) and \( K \) are the lower triangular Toeplitz matrices of \( L(q^{-1}) \) and \( K(q^{-1}) \), respectively.

As shown in (20), \( P_{\theta} \) is a function of \( Z \), \( L(q^{-1}) \) and \( W^\prime \). In order to achieve optimal estimation accuracy (i.e. requirement \( R_2 \)), \( Z \), \( L(q^{-1}) \), and \( W^\prime \) should be determined such that \( P_{\theta} \) is minimised. Next, the lower bound of \( P_{\theta} \) is derived and the optimal instrumental variable is given in Theorem 1.

Theorem 1. The lower bound of \( P_{\theta} \), which is denoted as \( P_{\theta}^{\text{opt}} \), is given by

\[
P_{\theta}^{\text{opt}} = \lambda^2 \left( K^{-1}(q^{-1}) \phi \right)^T \left( K^{-1}(q^{-1}) \phi \right)^{-1}.
\]

(21)

The equivalence between \( P_{\theta} \) and \( P_{\theta}^{\text{opt}} \) holds by selecting \( Z, L(q^{-1}), W' \) as follows:

\[
Z^{\text{opt}} = K^{-1}(q^{-1}) \phi, \quad L^{\text{opt}}(q^{-1}) = K^{-1}(q^{-1}), \quad W'^{\text{opt}} = I_{N \times N}.
\]

(22)

Proof. Define two matrices \( \alpha \) and \( \beta \) as follows:

\[
\alpha = R_{\phi \phi}^T W Z^T L K, \quad \beta = ( K^{-1}(q^{-1}) \phi )^T.
\]

(23)

Therefore, \( P_{\theta} \) in (20) can be rewritten using \( \alpha \) and \( \beta \) as follows:

\[
P_{\theta} = \lambda^2 ( \alpha \beta^T )^{-1} \alpha \beta^T ( \beta \alpha^T )^{-1}.
\]

(24)

Using Lemma A.4 in [35], we have that

\[
P_{\theta} \geq \lambda^2 ( \beta \beta^T )^{-1} = P_{\theta}^{\text{opt}}.
\]

(25)

Besides, it is obvious that \( P_{\theta} \) is equal to \( P_{\theta}^{\text{opt}} \) while selecting the variables as (22). Hence, \( P_{\theta}^{\text{opt}} \) is the optimal covariance matrix and the optimal variables are presented in (22), which completes the proof.

Remark 3. According to the theory about the instrumental variable [35], unbiased estimation of \( \theta_{\text{opt}} \) (i.e. requirement \( R_1 \)) can be achieved from (10) if \( Z \) is correlated with the reference \( r \) and uncorrelated with the noise \( r \). Obviously, these two conditions are both satisfied by the \( Z^{\text{opt}} \) in (22). Therefore, \( \hat{\theta}_{\text{opt}} \) is an unbiased estimation of \( \theta_{\text{opt}} \) inherently.

Remark 4. From (22), it is observed that the dimension of \( Z^{\text{opt}} \) satisfies \( n_\phi = n_\phi \). Therefore, \( R_{\phi \phi} \in \mathbb{R}^{n_\phi \times n_\phi} \) is square and (10) can be simplified as follows:

\[
\hat{\theta}_{\text{opt}} = R_{\phi \phi}^{-1} R_{\phi}.
\]

(26)

Remark 5. In pre-existing approaches [28], \( \phi \) is ignored while calculating and minimising \( P_{\theta} \), which yields

\[
Z^{\text{opt}} = \phi, \quad L^{\text{opt}}(q^{-1}) = 1, \quad W'^{\text{opt}} = I_{N \times N}.
\]

(27)

In Section 6, simulations are conducted and it is shown that the variables in (27) are non-optimal actually.
The proposed parameter updating law

It is observed from (22) that the calculation of \( Z^{\text{opt}} \) and \( L^{\text{opt}}(q^{-1}) \) requires the accurate knowledge of the unknown plant \( P \) and the real optimal parameter \( \theta_{\text{opt}} \), which are generally unavailable in practical implementations. Hence, an RIV method is developed in this subsection to obtain \( Z^{\text{opt}} \) and \( L^{\text{opt}}(q^{-1}) \) iteratively based on \( \hat{S}_r \) and \( \hat{S}_{pr} \) obtained in (13) without exploiting any information of \( P \) and \( \theta_{\text{opt}} \).

In this subsection, an auxiliary index \( j \) is introduced to mark the iteration number in the proposed RIV method. Specifically, the parameter \( \theta^{<j>} \), the instrumental variable \( Z^{<j>} \) and the filter \( L^{<j>}(q^{-1}) \) are updated after each iteration, and \( \theta_{\text{opt}} \), \( Z^{\text{opt}} \) and \( L^{\text{opt}}(q^{-1}) \) can be obtained after convergence.

In the proposed RIV method, the initial parameter \( \theta^{<0>} \) is selected and known by the user in advance. The steps of the proposed RIV method in the \( j \)th iteration are as follows.

1. Obtain the estimation of the noise-free \( S_r \) and \( S_{pr} \) as follows:
   \[
   \begin{align*}
   \hat{S}_r^{<j>} &= \left( C + F(\theta^{<j-1>}) \right)^{-1} \tilde{r}, \\
   \hat{L}_{pr}^{<j>} &= F(\theta^{<j-1>}) \hat{S}_r^{<j>}.
   \end{align*}
   \tag{28}
   \]

2. Based on (28), we have that \( \hat{\phi}^{<j>} = \left[ (\hat{\phi}^A)^{<j>}, (\hat{\phi}^B)^{<j>} \right] \) and
   \[
   \begin{align*}
   (\hat{\phi}^A)^{<j>} &= \left[ \Psi_1^{A, <j>, \hat{S}_r}, \Psi_2^{A, <j>, \hat{S}_r}, \ldots, \Psi_n^{A, <j>, \hat{S}_r} \right], \\
   (\hat{\phi}^B)^{<j>} &= \left[ -\Psi_1^{B, \hat{S}_r}, -\Psi_2^{B, \hat{S}_r}, \ldots, -\Psi_n^{B, \hat{S}_r} \right].
   \end{align*}
   \tag{29}
   \]

3. Calculate the estimation of \( K(q^{-1}) \) as follows:
   \[
   K^{<j>} = A \left( (\theta^A)^{<j-1>} \right) C^{-1} + B(\theta^B)^{<j-1>}).
   \tag{30}
   \]

4. Combining (29) and (30), the approximation of \( Z^{\text{opt}} \) and \( L^{\text{opt}}(q^{-1}) \) are derived as follows:
   \[
   Z^{<j>} = (K^{<j>})^{-1} \hat{\phi}^{<j>}, \\
   L^{<j>}(q^{-1}) = (K^{<j>})^{-1}.
   \tag{31}
   \]

5. Substituting (31) into (11), the estimations of \( R^{\phi} \) and \( R^{\psi} \) are given by
   \[
   \begin{align*}
   R^{\phi}^{<j>} &= \left( Z^{<j>} \right)^T \left( L^{<j>}(q^{-1}) \hat{\phi} \right), \\
   R^{\psi}^{<j>} &= \left( Z^{<j>} \right)^T \left( L^{<j>}(q^{-1}) \hat{S}_r \right).
   \end{align*}
   \tag{32}
   \]

6. The estimation of \( \theta_{\text{opt}} \) is determined as follows:
   \[
   \hat{\theta}^{<j>} = \left( R^{\phi}^{<j>} \right)^{-1} R^{\psi}^{<j>}.
   \tag{33}
   \]

(7) Check if the parameter \( \theta^{<j>} \) is convergent: if convergence is achieved, let \( \hat{\theta}_{\text{opt}} = \theta^{<j>} \), \( Z^{\text{opt}} = Z^{<j>} \) and \( L^{\text{opt}}(q^{-1}) = L^{<j>}(q^{-1}) \); if not, let \( j = j + 1 \) and return to the step (1).

\textbf{Remark 6.} Note that the above iterations are performed offline and no actual experiment is involved, which indicates that only one experiment, that is \( E_{01} \), is required to obtain \( \hat{\theta}_{\text{opt}} \).

\textbf{Remark 7.} Approach to deal with the potential instability of \( C^{-1} \) in (13), \( (C + F(\theta^{<j-1>})^{-1} \) in (28) and \( (K^{<j>})^{-1} \) in (31) is presented in the Appendix of [27].

\section{5 | FURTHER IMPROVEMENT OF PARAMETER ESTIMATION ACCURACY}

As shown in Figure 2, the proposed parameter updating law which aims at improving the parameter estimation accuracy is given in Section 5.1. The unbiasedness and variance of the parameters obtained with this parameter updating law are analysed in Sections 5.2 and 5.3, respectively.

\subsection{5.1 | The proposed parameter updating law}

In order to further improve the parameter estimation accuracy based on the above analysis, a novel parameter updating law is proposed as follows:

\[
\hat{\theta}_k = \hat{\theta}_{k-1} + \frac{1}{k} \hat{\Gamma}^{-1} \hat{\Omega}_k,
\tag{34}
\]

where \( k \) is the iteration number, and

\[
\begin{align*}
\hat{\Gamma} &= (Z^{\text{opt}})^T \left( L^{\text{opt}}(q^{-1}) \right), \\
\hat{\Omega}_k &= (Z^{\text{opt}})^T \left( L^{\text{opt}}(q^{-1}) \eta(\hat{\theta}_{k-1}) \right),
\end{align*}
\tag{35}
\]

where \( \eta(\hat{\theta}_{k-1}) \) is the estimated \( \eta(\hat{\theta}_{k-1}) \) in noisy condition, which is given by

\[
\eta(\hat{\theta}_{k-1}) = B(\hat{\theta}_{k-1}) \hat{r}_{k-1} = \tilde{r} - \hat{\phi} \hat{\theta}_{k-1} - B(\hat{\theta}_{k-1}) \hat{\eta}_{k-1}.
\tag{36}
\]

With the parameter updating law in (34), it is obvious that the \( \hat{\theta}_{\text{opt}} \) obtained in Section 4 is \( \hat{\theta}_1 \) when the initial parameter \( \hat{\theta}_0 \) is zero. The estimation accuracy of \( \hat{\theta}_1 \) has been analysed in Section 4. In the following description, the unbiasedness and variance of \( \hat{\theta}_k(k \geq 2) \) will be investigated.

\textbf{Remark 8.} Different from the iterations in Section 4.4, the iterations in (34) is not offline, and one reference-tracking trial is required in every iteration to update \( \eta(\hat{\theta}_{k-1}) \).
5.2 Unbiasedness

According to Remark 3, \(\theta_1\) is the unbiased estimation of \(\theta_{\text{opt}}\), that is,

\[
E(\theta_1 - \theta_{\text{opt}}) = 0. \tag{37}
\]

Combining (34)–(36), the difference between \(\theta_k(k \geq 2)\) and \(\theta_{\text{opt}}\) is given by

\[
\theta_k - \theta_{\text{opt}} = \left[ I_{nq} - \frac{1}{k} \Gamma^{-1}(Z^{\text{opt}})^T (L^{\text{opt}} \Phi) \right] (\theta_k - \theta_{\text{opt}}) \nonumber \\
- \frac{1}{k} \Gamma^{-1}(Z^{\text{opt}})^T (L^{\text{opt}} B(\theta_k - \theta_{\text{opt}}) X_{k-1}^1 ). \tag{38}
\]

Therefore, we have that

\[
E(\theta_k - \theta_{\text{opt}}) = \frac{k-1}{k} E(\theta_{k-1} - \theta_{\text{opt}}) \nonumber \\
= \frac{1}{k} E(\theta_1 - \theta_{\text{opt}}) = 0. \tag{39}
\]

Equation (39) indicates that \(\theta_k\) is still the unbiased estimation of \(\theta_{\text{opt}}\) when \(k \geq 2\).

5.3 Estimation accuracy in terms of variance

Define \(P_k\) as the covariance matrix of \(\theta_k\), that is,

\[
P_k = E \left[ (\theta_k - E(\theta_k))(\theta_k - E(\theta_k))^T \right] \nonumber \\
= E \left[ (\theta_k - \theta_{\text{opt}})(\theta_k - \theta_{\text{opt}})^T \right]. \tag{40}
\]

Obviously, we have that \(P_1 = P_{\theta_{\text{opt}}}^{\text{opt}}\). Combining (38) and (40), yields

\[
P_k = \left( \frac{k-1}{k} \right)^2 P_{k-1} + \frac{\lambda^2}{k^2} \Gamma^{-1}(Z^{\text{opt}})^T I_{\text{opt}} B_1 B_1^T I_{\text{opt}}^T Z^{\text{opt}} T^{-1} \nonumber \\
= \frac{1}{k^2} P_1 + \sum_{s=2}^{k} \frac{\lambda^2}{k^2} \Gamma^{-1}(Z^{\text{opt}})^T I_{\text{opt}} B_1 B_1^T I_{\text{opt}}^T Z^{\text{opt}} T^{-1}, \tag{41}
\]

where \(I_{\text{opt}}\) and \(B_1\) are the Toeplitz matrices of \(L^{\text{opt}}(q^{-1})\) and \(B(\theta_{\text{opt}})\), respectively, and \(\Gamma\) is the noise-free variable corresponding to \(\tilde{\Gamma}\) in (35).

As \(\theta_s(s \geq 1)\) is the unbiased estimation of \(\theta_{\text{opt}}\), it can be assumed that \(B(\theta_s) \approx B(\theta_{\text{opt}})\). Based on the above assumption, (41) can be rewritten as follows:

\[
P_k = \frac{1}{k^2} P_1 + \frac{(k-1)\lambda^2}{k^2} \Gamma^{-1}(Z^{\text{opt}})^T I_{\text{opt}} B_1 B_1^T I_{\text{opt}}^T Z^{\text{opt}} T^{-1}. \tag{42}
\]

Table 1 Procedure of the proposed approach

| Step | Description |
|------|-------------|
| (a)  | Determine the initial parameter \(\theta^{(0)}\) |
| (b)  | Conduct the reference-tracking trial \(E_{0i}\) and record the data \(e_0\) and \(j_0\) |
| (c)  | Estimate \(\hat{\lambda}_m\) and \(\hat{\lambda}_s\) using (33) |
| (d)  | Estimate \(Z^{\text{opt}}\) and \(L^{\text{opt}}(q^{-1})\) with the proposed RIV method in Section 4.4 |
| (e)  | Optimizing \(\hat{\lambda}_{1}\) using (26) |
| (f)  | For \(k \geq 2\), conduct a reference-tracking trial with the parameter \(\theta_{k-1}\) and record the signal \(y(\theta_{k-1})\) |
| (g)  | Calculate \(\theta_{k}\) directly using (34) |
| (h)  | Return to Step (i) to start new iteration until the parameter estimation accuracy is satisfactory |

where \(B_{\text{opt}}\) is the Toeplitz matrix of \(B(\theta_{\text{opt}})\). Note that \(P_1\) and \(\Gamma^{-1}(Z^{\text{opt}})^T I_{\text{opt}} B_{\text{opt}} B_{\text{opt}}^T I_{\text{opt}}^T Z^{\text{opt}} T^{-1}\) in (42) are both constant, which yields

\[
\lim_{k \to +\infty} P_k = 0. \tag{43}
\]

Therefore, the covariance matrix can be arbitrarily small after enough iterations, which implies that the parameter estimation accuracy can be greatly improved.

5.4 Procedure of the proposed approach

The overall procedure of the proposed approach is presented in Table 1.

6 SIMULATIONS

6.1 Set-ups

In the following simulations, the controlled plant and the feedback controller are given by

\[
P(q^{-1}) = \frac{0.04(1 - 1.953q^{-1} + 0.9629q^{-2})}{(1-q^{-1})^2(1-1.995q^{-1} + 0.9987q^{-2})}, \tag{44}
\]

\[
C(q^{-1}) = \frac{3.879q^{-1}(1 - 0.9882q^{-1})(1 - 1.990q^{-1} + 0.9938q^{-2})}{(1 - 1.951q^{-1} + 0.9605q^{-2})(1 - 0.9551q^{-1} + 0.2816q^{-2})}. \tag{45}
\]

The noise \(w\) is normally distributed white noise with zero mean and standard deviation \(\bar{\lambda} = 5\) nm. \(H(q^{-1})\) is derived such that \(H(q^{-1})S(q^{-1}) = 1\) is satisfied. The references to be tracked in the simulations are the fourth-order trajectory \(r_s\), as shown in Figure 3. The parameters of \(r_s\) are provided in Table 2.
In $Q_1$, the coefficient $k$ in experiments $E_2$ and $E_4$ in [34] will be high enough to ensure the optimal estimation accuracy of $\theta_{opt}$. In $Q_2$ and $Q_3$, $Z^{opt}$ and $L^{opt}(q^{-1})$ are assumed to be unknown and the RIV method presented in Section 4.4 is applied to determine $Z^{opt}$ and $L^{opt}(q^{-1})$. Five computational iterations are performed in the RIV method for both $Q_2$ and $Q_3$, which are enough for them to be convergent. By contrast, $Z^{opt}$ and $L^{opt}(q^{-1})$ are assumed to be known in advance in $Q_4$ and can be directly used without computational iterations.

In order to test the parameter estimation accuracy of $Q_1$–$Q_4$, a Monte Carlo simulation is conducted with 1000 realisations to track the reference $r_t$ for each of the four methods. To enable fair comparison, the initial parameters for $Q_1$–$Q_4$ are all set as $\theta_0 = [0, 0, 0, 0, 0]$. In each realisation, 10 iterations are performed for $Q_2$–$Q_4$, and one iteration is performed for $Q_1$ as $Q_1$ is non-iterative.

### 6.2 Simulation results

#### 6.2.1 Unbiasedness and variance

The means and standard deviations of the parameters obtained from $Q_1$ to $Q_4$ after the first iteration are presented in Table 3. It is observed that the means of the parameters obtained from $Q_1$ to $Q_4$ are all unbiased. For $Q_1$, this is owing to the conduction of repeated experiments $E_3$ and $E_4$ in [34]. For $Q_2$–$Q_4$, the unbiased estimation is guaranteed as the two conditions in Remark 3 are both met by the instrumental variables applied in all these three methods. In $Q_1$, four experiments are required to achieve the unbiased estimation of $\theta_{opt}$. By contrast, only one experiment, that is $E_0$ in Section 4.2, is needed for $Q_2$–$Q_4$ in the first iteration. Therefore, the proposed approach achieves unbiased estimation of $\theta_{opt}$ more efficiently than $Q_1$.

From Table 3, it can also be observed that the standard deviations of parameters obtained by $Q_3$ is much smaller than $Q_1$ and $Q_2$ and very close to $Q_4$. Two conclusions can be made based on the above results: (1) the parameter estimation accuracy of the proposed approach is greatly improved compared to $Q_1$ after the application of the instrumental variable and the minimisation of the covariance matrix in Section 4.3; (2) optimal estimation accuracy is achieved by the instrumental variable proposed in (22).

The evolutions of standard deviations of $Q_2$–$Q_4$ are illustrated in Figure 4. It is shown that the curves of $Q_1$ are always much lower than $Q_2$ and almost overlap with $Q_4$. This validates the optimal estimation accuracy of the proposed instrumental variable again. Moreover, it can be observed from Figure 4 that the standard deviations of $Q_2$–$Q_4$ are all reduced iteration by iteration. Hence, the improvement of parameter estimation accuracy brought in by the proposed parameter updating law in Section 5 is verified.
6.2.2 | Worst-case performance

In order to illustrate the relation between the parameter estimation accuracy and reference-tracking performance, the worst-case performance is investigated. Specifically, the worst-case tracking error means that the reference-tracking error signal whose 2-norm is the largest among the 1000 realisations. Define $\theta_{wc}^i$ as the parameter in the $i$th iteration which leads to the worst performance among the 1000 realisations, that is,

$$\theta_{wc}^i = \max_{\theta_{k(i)}} \left\| \varepsilon(\theta_{k(i)}) \right\|_2, \quad (i = 1, 2, \ldots, 1000),$$  \hspace{1cm} (47)

where the auxiliary index $i$ means the $i$th realisation. Hence, the worst-case tracking error is given by $\varepsilon_{wc} = \varepsilon(\theta_{wc}^i)$. The evolutions of the 2-norm of $\varepsilon_{wc}$ of $Q_1$-$Q_4$ are presented in Figure 5. Several worst-case tracking errors of $Q_1$-$Q_4$ are given in Figure 6. It can be observed that the noise-induced error is dominant in the worst-case tracking errors of $Q_2$ and $Q_4$, and the reference-induced error is almost eliminated with only one iteration. By contrast, significant reference-induced error still exists in the worst-case tracking errors of $Q_1$ and $Q_3$ after the first iteration. In the subsequent iterations, with the parameter updated using (34), the reference-induced error in the worst-case tracking error of $Q_2$ is reduced iteration by iteration, as shown in Figure 5. It can be observed from Figure 6 that...
the reference-induced error nearly disappears in the worst-case tracking error of $Q_2$ after the 10th iteration.

Compare the above results with Figure 4 and Table 3, it can be concluded that the magnitude of worst-case tracking error is highly correlated with the parameter estimation accuracy. Specifically, the improvement of parameter estimation accuracy will lead to the reduction of worst-case tracking error. Therefore, the robustness against the noise is improved after the optimisation of the parameter estimation accuracy with the proposed approach.

7 | EXPERIMENTS

7.1 | System description

Experiments are carried out on the ultraprecision reticle stage developed in our lab. As shown in Figure 7, the reticle stage is driven by moving-magnet planar motors. Specifically, the coils and the permanent-magnet array are fixed on the stator and mover of the reticle stage, respectively. The mover is magnetically suspended above the stator while the reticle stage moves. Therefore, the motion of the reticle stage is frictionless. Moreover, with this moving-magnet structure, the cooling system is installed on the stator and no cable is connected to the mover. Hence, the disturbances resulted from the cable is avoided compared to moving-coil planar motors. The position of the mover is measured by the laser interferometers with subnanometer resolution. The reticle stage is controlled in all six DOFs, including three translations ($x$, $y$, and $z$) and three rotations ($t_x$, $t_y$, and $t_z$). Static decoupling matrix is applied to make the plant diagonally dominant, which enables the application of SISO feedback and feedforward control to each DOF of the reticle stage. The sampling period of the control system is $T_s = 0.0002$ s.

This paper considers the scanning direction (i.e., $y$ direction) for the experiments. The references to be applied include $r_1$, $r_2$, and $r_3$, which are presented in Figure 4 and Table 2. In the experiments, the reference-tracking performance and extrapolation properties of the following three methods are compared.

$M_1$: standard ILC in [12];
$M_2$: the pre-existing approach associated with RBFs in [34];
$M_3$: the proposed approach mentioned here.

In $M_1$, the optimal feedforward signal is pursued and no feedforward controller parameterisation is required. Based on the results of preliminary experiments, the feedforward controller is parameterised as follows for $M_2$ and $M_3$:

$$F(\theta, q^{-1}) = \left(1 - \frac{g^{-1}}{\bar{g}}\right)^2 \cdot \frac{\theta^A[1] + \theta^A[2]q^{-1} + \theta^A[3]q^{-2}}{1 + \theta^B[1]q^{-1} + \theta^B[2]q^{-2}},$$  \hfill (48)$$

where the term $\left(1 - \frac{g^{-1}}{\bar{g}}\right)^2$ is applied to compensate for the dominant rigid-body dynamics.

7.2 | Experimental results

7.2.1 | Reference-tracking performance

The reference-tracking performance of $M_1$–$M_3$ towards the reference $r_1$ is investigated. In this case study, no feedforward control is applied in the initial task for all the three methods. Ten iterations are performed for $M_1$ and $M_3$, and $M_2$ is executed non-iteratively.

The performance of these three methods are presented in Figures 8 and 9. It is shown that great performance enhancement is achieved with all three methods, and $M_1$ achieves the best performance. This result validates the superior performance of signal-based methods towards repeating tasks. On the other hand, it can be observed that $M_3$ achieves higher performance than $M_2$ with the same feedforward parameterisation. This can be explained by the conclusion of the simulations that higher parameter estimation accuracy contributes to smaller worst-case error.

Actually, the optimality of the parameter obtained by $M_3$ can be reflected by Figure 10. In the proposed RIV method in
FIGURE 9 The achieved reference-tracking errors of $M_1$ (after 10 iterations), $M_2$ (after 1 iteration) and $M_3$ (after 10 iterations) while tracking $r_I$.

FIGURE 10 The estimated $\hat{S}_r$ and $\hat{S}_{pr}$ from (13) and the noise-free estimation $S_r^{<10^2}$ and $S_{pr}^{<10^2}$ from (28).

Note: $\hat{S}_r$ in (a) is also the initial reference-tracking error for $M_1$–$M_3$.

In order to test the extrapolation properties of $M_1$–$M_3$, the optimal feedforward signal ($M_1$) or feedforward controller ($M_2$ and $M_3$) obtained while tracking $r_I$ are directly applied to track the other two references, that is, $r_{II}$ and $r_{III}$.

The reference-tracking errors of $M_1$–$M_3$ are depicted in Figure 11. It is obvious that great performance deterioration is encountered with $M_1$ after the reference changes, which reveals the poor extrapolation properties of $M_1$. Comparatively, the performance of $M_2$ and $M_3$ is maintained even if the reference is changed. Hence, the great advantage of filter-based method over signal-based method on extrapolation properties is validated.

On the other hand, the performance of $M_3$ is better than $M_2$ for all the three references. Fundamentally, the extrapolation properties of filter-based method are determined by the degree of approximation of the feedforward controller $F$ to the inverse plant $P^{-1}$. It has been shown in the simulations that the parameter estimation accuracy of $M_3$ is much better than $M_2$. Therefore, the feedforward controller $F$ obtained by $M_3$ is probably much closer to $P^{-1}$ than that of $M_2$. This explains the better performance of $M_3$ compared to $M_2$ while tracking $r_{II}$ and $r_{III}$. The high performance resulted from the accurate feedforward parameter estimation of $M_3$ is validated.

8 | CONCLUSION

This paper proposed a novel data-driven rational feedforward tuning approach. In the proposed approach, the instrumental variable was exploited, and the unbiased estimation of the optimal parameter was achieved inherently. In order to achieve the optimal estimation accuracy, the variance of the estimated value of the optimal parameter was analysed. The optimal instrumental variable which leads to the minimisation of the covariance matrix was derived, and an RIV approach was proposed to estimate the optimal instrumental variable. Based on the above optimal instrumental variable, a novel parameter updating law was proposed. It is proved that the proposed parameter updating law was able to improve the parameter estimation accuracy while maintaining the unbiasedness. Throughout the whole parameter tuning process, only data measured from reference-tracking trials was required. Hence, the proposed approach was totally data-driven. In the simulations and experiments, the advantages of the proposed...
approach were demonstrated: (1) the unbiasedness and optimal estimation accuracy in terms of variance; (2) high performance for repeating tasks and good extrapolation properties. The research is ongoing towards extending the proposed approach to multivariable and parameter-varying systems.

ACKNOWLEDGEMENT

This work was supported by the National Science and Technology Major Project of China under grant 2017ZX02102004.

REFERENCES

1. Zhu, Y., et al.: Convergence performance oriented data-driven tuning method for parameterised controller design with cases investigation. IET Control Theory Appl. 10(12), 1322–1330 (2016)
2. Li, M., et al.: State/model-free variable-gain discrete sliding mode control for an ultraprecision wafer stage. IEEE Trans. Ind. Electron. 64(8), 6695–6705 (2017)
3. Li, M., et al.: An integrated data-driven inverse discrete sliding mode control for an ultraprecision wafer stage. IEEE Trans. Ind. Electron. 64(5), 4139–4149 (2017)
4. Li, M., et al.: Data-driven iterative feedforward controller with rational parametrization: achieving optimality for varying tasks. J. Frankl. Inst. 356(12), 6352–6372 (2019)
5. Butterworth, J.A., Pao, L.Y., Abramovich, D.Y.: Analysis and comparison of three discrete-time feedforward model-inverse control techniques for nonminimum-phase systems. Mechatronics 22(5), 577–587 (2012)
6. Tomizuka, M.: Zero phase error tracking algorithm for digital control. J. Dyn. Syst., Meas. Control 190(1), 65–68 (1987)
7. Lambrechts, P., Boerlage, M., Steinbuch, M.: Trajectory planning and feedforward control for electromechanical motion systems. Control Eng. Pract. 13(2), 145–157 (2005)
8. van Zundert, J., Oomen, T.: On inversion-based approaches for feedforward and ILC. Mechatronics 50, 282–291 (2018)
9. Devasia, S.: Should model-based inverse inputs be used as feedforward controller under plant uncertainty? IEEE Trans. Autom. Control 47(11), 1865–1871 (2002)
10. Heertjes, M.: Data-based motion control of wafer scanners. IFAC-PapersOnLine 49(13), 1–12 (2016)
11. Jiang, Y., et al.: A data-driven iterative decoupling feedforward control strategy with application to an ultraprecision motion stage. IEEE Trans. Ind. Electron. 62(1), 620–627 (2015)
12. Bristow, D.A., Tharayil, M., Alleyne, A.G.: A survey of iterative learning control. IEEE Control System Mag. 27(4), 20–25 (2007)
13. Chen, W., Tomizuka, M.: Dual-stage iterative learning control for MIMO mismatched system with application to robots with joint elasticity. IEEE Trans. Ind. Electron. 22(4), 1350–1361 (2014)
14. Huang, D., et al.: High-performance tracking of piezoelectric positioning stage using current-cycle iterative learning control with gain scheduling. IEEE Trans. Ind. Electron. 61(2), 1085–1098 (2014)
15. Roovers, D., Bosgra, O.: Synthesis of robust multivariable iterative learning controllers with application to a wafer stage motion system. Int. J. Control 73(10), 968–979 (2000)
16. Mishra, S., Coaplen, J., Tomizuka, M.: Precision positioning of wafer scanners segmented iterative learning control for nonrepetitive disturbances. IEEE Control Syst. Mag. 27(4), 20–25 (2007)
17. Mishra, S., Tomizuka, M.: Projection-based iterative learning control for wafer scanner systems. IEEE/ASME Trans. Mechatronics 14(3), 388–393 (2009)
18. Heerjes, M., Tso, T.: Nonlinear iterative learning control with applications to lithographic machinery. Control Eng. Pract. 15(12), 1545–1555 (2007)
19. Kaseminsup, Y., et al.: Reference-tracking feedforward control design for linear dynamical systems through signal decomposition. In: Proceedings of the American Control Conference, Seattle, WA, pp. 2387–2392, May 2017
20. Benidirhoum, T., Bouakrif, F., Zasadzinski, M.: Iterative learning radial basis function neural networks control for unknown multi-input multi-output nonlinear systems with unknown control direction. Trans. Inst. Meas. Control 41(12), 3452–3467 (2019)
21. Jiang, Y., et al.: Optimal feedforward control with a parametric structure applied to a wafer stage. Proc. Inst. Mech. Eng. I J. Syst. Control Eng. 228(2), 97–106 (2014)
22. Huang, D., et al.: Data-based switching feedforward control for repeating and varying tasks: with application to an ultraprecision wafer stage. IEEE Trans. Ind. Electron. 66(11), 8670–8680 (2019)
23. van der Meulen, S.H., Tousain, R.L., Bosgra, O.H.: Fixed structure feedforward controller design exploiting iterative trials: application to a wafer stage and a desktop printer. J. Dyn. Syst. Meas. Control 130(5), 0510061–05100616 (2008)
24. Song, F., et al.: Data-driven iterative feedforward tuning for a wafer stage: a high-order approach based on instrumental variables. IEEE Trans. Ind. Electron. 66(4), 3106–3116 (2018)
25. Boeren, F., Oomen, T., Steinbuch, M.: Iterative motion feedforward tuning: a data-driven approach based on instrumental variable identification. Control Eng. Pract. 37, 11–19 (2015)
26. Boeren, F., Brauwen, D., Oomen, T.: Enhancing feedforward controller tuning via instrumental variables: with application to nanopositioning. Int. J. Control 90, 746–764 (2017)
27. Boeren, F., et al.: Joint input shaping and feedforward for point-to-point motion: automated tuning for an industrial nanopositioning system. Mechatronics 24(6), 572–581 (2014)
28. Heerjes, M., Hennekens, D., Steinbuch, M.: MIMO feedforward design in wafer scanners using a gradient approximation-based algorithm. Control Eng. Pract. 18, 495–506 (2010)
29. Bolder, J., Oomen, T.: Rational basis functions in iterative learning control – with experimental verification on a motion system. IEEE Trans. Control Syst. Technol. 23(2), 722–729 (2015)
30. Blanken, I., et al.: Batch-to-batch rational feedforward control: from iterative learning to identification approaches, with application to a wafer stage. IEEE/ASME Trans. Mechatron. 22(2), 826–837 (2017)
31. Boeren, F., et al.: Optimal estimation of rational feedforward control via instrumental variables: with application to a wafer stage. Asia. J. Control 2021;15:937–948. https://doi.org/10.1049/cth2.12093
32. Kaseminsup, Y., et al.: Reference-tracking feedforward control design for linear dynamical systems through signal decomposition. In: Proceedings of the American Control Conference, Seattle, WA, pp. 2387–2392, May 2017
33. Benidirhoum, T., Bouakrif, F., Zasadzinski, M.: Iterative learning radial basis function neural networks control for unknown multi-input multi-output nonlinear systems with unknown control direction. Trans. Inst. Meas. Control 41(12), 3452–3467 (2019)
34. Jiang, Y., et al.: Optimal feedforward control with a parametric structure applied to a wafer stage. Proc. Inst. Mech. Eng. I J. Syst. Control Eng. 228(2), 97–106 (2014)
35. Li, M., et al.: Data-based switching feedforward control for repeating and varying tasks: with application to an ultraprecision wafer stage. IEEE Trans. Ind. Electron. 66(11), 8670–8680 (2019)
36. van der Meulen, S.H., Tousain, R.L., Bosgra, O.H.: Fixed structure feedforward controller design exploiting iterative trials: application to a wafer stage and a desktop printer. J. Dyn. Syst. Meas. Control 130(5), 0510061–05100616 (2008)
37. Song, F., et al.: Data-driven iterative feedforward tuning for a wafer stage: a high-order approach based on instrumental variables. IEEE Trans. Ind. Electron. 66(4), 3106–3116 (2018)
38. Boeren, F., Oomen, T., Steinbuch, M.: Iterative motion feedforward tuning: a data-driven approach based on instrumental variable identification. Control Eng. Pract. 37, 11–19 (2015)
39. Boeren, F., Brauwen, D., Oomen, T.: Enhancing feedforward controller tuning via instrumental variables: with application to nanopositioning. Int. J. Control 90, 746–764 (2017)
40. Boeren, F., et al.: Joint input shaping and feedforward for point-to-point motion: automated tuning for an industrial nanopositioning system. Mechatronics 24(6), 572–581 (2014)
41. Heerjes, M., Hennekens, D., Steinbuch, M.: MIMO feedforward design in wafer scanners using a gradient approximation-based algorithm. Control Eng. Pract. 18, 495–506 (2010)
42. Bolder, J., Oomen, T.: Rational basis functions in iterative learning control – with experimental verification on a motion system. IEEE Trans. Control Syst. Technol. 23(2), 722–729 (2015)
43. Blanken, I., et al.: Batch-to-batch rational feedforward control: from iterative learning to identification approaches, with application to a wafer stage. IEEE/ASME Trans. Mechatron. 22(2), 826–837 (2017)
44. Boeren, F., et al.: Optimal estimation of rational feedforward control via instrumental variables: with application to a wafer stage. Asia. J. Control 20(3), 975–992 (2018)
45. Huang, W., et al.: Data-driven rational feedforward tuning with application to an ultraprecision wafer stage. Proc. Inst. Mech. Eng. I J. Syst. Control Eng. 234(6), 748–758 (2020)
46. Söderström, T., Stoica, P.: System Identification. Prentice-Hall, Hmpestead (1989)