Finite sample properties for closed loop system identification

¹WANG JIANHONG, ¹CHEN PENG, ²RICARDO A. RAMIREZ-MENDOZA

¹School of Electronic Engineering and Automation, Jiangxi University of Science and Technology, Ganzhou, 343100, CHINA ²School of Engineering and Sciences, Tecnologico de Monterrey, Monterrey, Ave. Eugenio Garza Sada

2501, Monterrey, N.L., 64849, MEXICO

Abstract: In this paper, after closed loop system identification is reviewed, asymptotic analysis and finite sample analysis for closed loop system identification are studied respectively, corresponding to the infinite data and finite data. More specifically, within the framework of infinite data, the cost function is modified to its simplified form, and one optimal feedback controller is obtained based on our own derivations. The simplified cost function and optimal feedback controller are benefit for practical application. Furthermore, the asymptotic variance of that optimal feedback controller is also yielded from the point of asymptotic analysis. In the case of finite data, finite sample properties are constructed for closed loop system identification, then one difference between the sampled identification criterion and its corresponding expected criterion is derived as an explicit form, which can bound one guaranteed interval for the sampled identification criterion. Finally, one simulation example is used to prove the efficiency of our proposed theories.

Key-Words: Asymptotic analysis; Finite sample analysis; Closed loop identification; Optimal feedback controller

Received: February 3, 2021. Revised: July 4, 2021. Accepted: July 13, 2021. Published: July 21, 2021.

1 Introduction

Many practical systems operate under feedback control, which is due to required safety of operation or to unstable behavior of the plant, as occurs in many industrial production processes, such as paper production, glass production etc. During the feedback structure, i.e. closed loop system, the output variable is turned back to the input through a feedback controller. The most merit for closed loop system or feedback controller is to achieve the perfect tracking, it means the output variable for the whole closed loop system can track our designed output trajectory. As a consequence experiments can only be performed under presence of a stabilizing controller. Even in situations where plants are stable, production restrictions can be strong reasons for not allowing experiments in open loop. As plant and feedback controller exist in closed loop system simultaneously, and the design for feedback controller is dependent of this considered plant, so plant must be identified firstly before feedback controller design. This detailed identification

process for plant corresponds to closed loop system identification.

Theoretical research on system identification has been matured, for example, the detailed statistical identification methods and the asymptotic or convergence properties are analyzed from the point of time domain perspective in reference [1], where some stochastic probability inequalities corresponding to the guaranteed interval estimations are proposed. Furthermore, these various statistical identification methods are the theoretical basis for our new identification strategies. In order to be different from the point of time domain, reference [2] studies the identification methods in frequency domain, and points out how to avoid the aliasing effect in carrying out the Fourier transformation. The recursive identification methods on line are proposed in [3], where the computational complexity and real time algorithms are also given for linear system. It means that the next parameter estimation can be updated recursively through iterating computation. Now the identification of nonlinear system is still under investigation, as there

is no universal identification method that can be applied to any form of nonlinear systems. So far some special nonlinear systems are studied, such as Wiener system, Hammerstein system and linear parameter varying system etc. The instrumental variable method is studied for linear parameter varying system [4], and the norm regularization identification is proposed in identifying the switched linear regression model [5]. Reference [6] gives the state space identification for nonlinear system with dual linear blocks. For two special classical nonlinear systems-Wiener system and Hammerstein system, the frequency weighted maximum likelihood identification is put out for identifying Wiener system, and a real-time robust set estimation is given for Hammerstein system. Then based on these two special cases, many researchers start to study another identification strategies and the identifiability for different series combination [7]. Due to white noise is an ideal case, it does not exist in engineering and in additional, deriving statistical properties of noise is often very difficult in practice, as it is usually not possible to measure the external noise directly. As linear system identification is a mature field, then the most common classical approaches can be used to perform the prediction of the output value, such as prediction error method, maximum likelihood method, etc, in case of the linear dynamic system. But the above classical approaches are not used in nonlinear system identification, as it is not easy to expand the nonlinear system and the nonlinear system identification can not reformulated as one parametric optimization problem easily. It means that the subject of ongoing research is concentrated nonlinear system identification, i.e. on the prediction of the output value coincides with the considered nonlinear dynamic system. The use of orthonormal basis functions is proposed to be a systematic construction mechanism of stable time domain kernels for impulse response estimation of linear time invariant systems [8], further two proposed weights of the orthonormal basis functions are regarded as decay terms, guaranteeing stability in the associated hypothesis space. A Bayesian nonparametric approach is proposed to estimate multi-input multi-output linear parameter varying models under the general noise model structure [9], where this Bayesian nonparametric approach is based on the estimation of the one step ahead predictor. Due to linear parameter varying models are the special nonlinear systems, efficient maximum likelihood estimation is considered on the basis conditional probability density and total probability theorem [10]. In [11], adopting maximum entropy arguments, a new L2 penalty is derived and the Hankel matrix is exploited to achieve the maximum entropy vector kernels for multi-input multi-output system identification. Alternatively, a new kernel based regulation approach, that combines ideas from machine learning, statistics and dynamical systems, has been introduced recently [12], and further extended in [13], where a new sparse multiple kernel based regularization is provided to alleviate the latter computational cost. Moreover prediction error method framework is well suited to identify a large variety of noise and plant models, see [14] for an overview. Due to linear parameter varying models cover a large of processes and noise representations, where the Box-Jenkins model is the most general form, prediction error identification of Box-Jenkins models leads to a nonlinear optimization problem [15], which is solved by a scaled gradient projection method. Another important issue in identifying the linear parameter varying models is capturing the structural dependency on the scheduling signal. In the parametric case, the structural dependency is generally characterized by using a pre-specified set of basis functions [16], which require significant prior knowledge of the underlying system. The nonparametric methods for the identification of linear parameter varying models offer attractive alternative approaches to capture the underlying dependencies directly from measured data without covering any parameterization [17]. A Bayesian nonparametric approach on the basis of the estimation of the one step ahead predictor is

proposed to estimate the linear parameter varying models [18], and some predictors associated with the input and output signals are modeled as the stable impulse response models. The problem of controlling the linear parameter varying model is considered in [19], and the control strategy can be applied in a wafer stage. During the system identification process, many identification methods are proposed to identify these unknown parameters, for example the classical least squares method, instrumental variable method, maximum likelihood estimation method, prediction error method, Bayesian method, etc [20]. As the goal of model predictive control is to control the closed loop system in order to track a desired output reference and reject disturbances. Moreover, the considered controller may enforce input and output constraints. So after introducing control input and output value as decision variables simultaneously, such a control objective can be formulated as a quadratic programming problem with inequality constrained condition. It is well known that the dual of the quadratic programming problem is an unconstrained optimization problem [21]. The most important element in model predictive control is the prediction of output value. After deriving the prediction of output value by prediction error method and then substituting it into one considered cost function, we take the derivative of the cost function with respect to input value to obtain one optimal input [22]. But the problem of deriving the prediction of output value is dependent of external noise, which is always assumed to be independent and identically distributed white noise. Due to white noise is an ideal case, it does not exist in engineering and in additional, deriving statistical properties of noise is often very difficult in practice as it is usually not possible to measure noise directly [23]. Although model predictive control has been found to be quite a robust type of control in most reported applications, some new and very promising results allow one to think that this control technique will experience greater expansion within this community in near recent years [24].

Based on above formulations on system identification or closed loop identification, after the excitation input is used to excite the closed loop system, then output variable or signal is measured by some sensors. The unknown plant is identified from these input and output sequence, i.e. the knowledge about the unknown plant are extracted from the measured data sequences. This identification process is similar to recent machine learning and data mining. This similarity between system identification and data mining motivates lots of research on applying data mining into system identification. But in the studying for system identification, lots of considerations are put on optimal input design and identification algorithm, few on identification accuracy, i.e. how about the quality about the identification model and can we use this identification model to replace the original plant? To reply this question, asymptotic theory, proposed in [1], is proposed to measure the quality of the identification model. All the derived results hold in case of infinite data, i.e. the number of measured data N approaches to infinity $(N \rightarrow \infty)$ in classical reference [1]. For example, asymptotic theory is applied to analyze asymptotic properties for one special set membership identification algorithm [25]. But in practice the measured data is finite, it means that the number of measured data approaches to infinity is an ideal case. This case of finite data causes all the derived asymptotical results in [1] and [25] not hold in reality. The quality analysis about the identification model on the condition of finite data corresponds to finite sample property for system identification. For this problem of finite sample analysis, the author considers finite sample properties of virtual reference feedback tuning control and derive one probability bound to quantify the difference between the expected identification cost and its sampled cost [26]. Furthermore, in [27] the author gives the number of measured data by using generalization of independent block sequence, which corresponds to the sample complexity in system identification.

From our previously published work [28] about stealth identification for closed loop system structure, the identification problem for closed loop system with unknown controller and nonlinear controller is tackled, and the new prediction error and inverse covariance matrix are found to be independent of the unknown or nonlinear controller. In this paper, here we do not concentrate on the identification algorithm for closed loop system again, but turn to identification accuracy analysis for closed loop system. As there are two kinds of accuracy analysis, i.e. asymptotic analysis and finite sample analysis, so we consider the identification accuracy analysis for closed loop system from the point of asymptotic analysis and finite sample analysis respectively. More specifically, after reviewing some existed results about closed loop system identification, the cost function, used to identify the unknown parameter vector in case of parameterized plant, is reformulated to its reduced form. The spectral estimation for unknown plant is obtained as a nonparametric estimation, which is widely used in flutter identification. Moreover, based on our derived reduced cost function, one optimal feedback controller is yielded to be dependent of the identified plant. These reduced cost function and optimal feedback controller are benefit for later practical application, such as closed loop model validation, optimal input design etc. Then asymptotical analysis is applied to derive the variance for the optimal feedback controller, and the derivation of the reduced cost function and the optimal feedback controller are full of some asymptotic results. On the contrary, finite sample properties for closed loop system are also studied to quantify the difference between identification criterion and expected criterion. Furthermore, the VC dimension, coming from machine learning, is used to bound this difference. Generally, this paper considers the closed loop system from our published word [28]., and extends the asymptotic analysis and finite sample analysis for closed loop system identification. Our derived new results in this paper are completed than our published work

[26] and [27], as asymptotic analysis and finite sample analysis are all considered here.

The paper is organized as follows. In section 2, closed loop system is described and the classical identification algorithm for closed loop system is reviewed on the basis of the parametrized plant. In section 3, the cost function or identification criterion is adjusted to its other form, then spectral estimation for the unknown plant is derived to be one nonparametric estimation. Based on this adjusted cost function, one optimal feedback control is given, and asymptotic analysis are studied during these mathematical derivations, such as the variance analysis for the optimal feedback control. This section 3 corresponds to asymptotic analysis for closed loop system identification. In section 4, finite sample properties for closed loop system identification are considered in case of finite data. Section 5 ends the paper with final conclusion.

2. Closed loop system identification

Some different forms exists for closed loop system structure in practice, here consider one typical closed loop system with one linear time invariant feedback controller in Figure 1.





where in Figure 1, r(t) is external excitation input, chosen by the designer, $P_0(z)$ is the true or nominal plant, $H_0(z)$ is one noise filter, C(z) is the feedback controller, whatever it is unknown or nonlinear, our proposed stealth identification strategy can be used to design it. u(t) and y(t)are input and output for plant $P_0(z)$. e(t) is a white noise with zero mean white Gaussian and mutually uncorrelated, v(t) is the filtered noise by passing white noise e(t) through that noise filter $H_0(z)$, z^{-1} is the backward shift operator.

Observing Figure 1 again, the process about identifying the plant $P_0(z)$ and noise filter $H_0(z)$ corresponds to closed loop system identification, and we have $v(t) = H_0(z)e(t)$. From Figure 1, we have the following equities easily.

$$\begin{cases} y(t) = P_0(z)u(t) + H_0(z)e(t) \\ u(t) = r(t) - C(z)y(t) \end{cases}$$
(1)

After simple computations, we have

$$\begin{cases} y(t) = \frac{P_0(z)}{1 + P_0(z)C(z)}r(t) + \frac{H_0(z)}{1 + P_0(z)C(z)}e(t) \\ u(t) = \frac{1}{1 + P_0(z)C(z)}r(t) - \frac{C(z)H_0(z)}{1 + P_0(z)C(z)}e(t) \end{cases}$$
(2)

To simplify notation, one sensitivity function $S_0(z)$ is defined as.

$$S_0(z) = \frac{1}{1 + P_0(z)C(z)}$$
(3)

Then equation (2) can be simplified as.

$$\begin{cases} y(t) = P_0(z)S_0(z)r(t) + H_0(z)S_0(z)e(t) \\ u(t) = S_0(z)r(t) - C(z)H_0(z)S_0(z)e(t) \end{cases}$$
(4)

Here $(P_0(z), H_0(z))$ are unknown in closed loop system and they are regarded as true values or nominated values. The closed loop identification is to use input-output variables $\{r(t), y(t)\}$ to estimate $(P_0(z), H_0(z))$.

To show $(P_0(z), H_0(z))$ are unknown, one unknown parameter vector θ is introduced to parametrize the closed loop system, i.e. $(P(z,\theta), H(z,\theta))$. It means closed loop system is parametrized by one unknown parameter vector θ , then the goal of closed loop identification is to estimate that unknown parameter vector θ .

substituting the parametrized form $(P(z,\theta),H(z,\theta))$ into equation (2), the parametrized relations are given as.

$$\begin{cases} y(t) = \frac{P(z,\theta)}{1+P(z,\theta)C(z)}r(t) + \frac{H(z,\theta)}{1+P(z,\theta)C(z)}e(t) \\ u(t) = \frac{1}{1+P(z,\theta)C(z)}r(t) - \frac{C(z)H(z,\theta)}{1+P(z,\theta)C(z)}e(t) \end{cases}$$
(5)

Combing equation (2) and (5), $\{u(t), y(t)\}_{t=1}^{N}$ are input-output variables for plant $P_0(z)$, and $\{u(t,\theta), y(t,\theta)\}_{t=1}^{N}$ are parametrized input-output variables for plant $P(z,\theta)$. To identify that unknown parameter vector θ in $(P(z,\theta), H(z,\theta))$, classical prediction error strategy is applied here, where *N* denotes the number of input-output variables.

$$\hat{y}(t,\theta) = \frac{1+P(z,\theta)C(z)}{H(z,\theta)} \frac{P(z,\theta)}{1+P(z,\theta)C(z)} r(t) + \left[1-\frac{1+P(z,\theta)C(z)}{H(z,\theta)}\right] y(t) = \frac{P(z,\theta)}{H(z,\theta)} r(t) + \frac{H(z,\theta)-1-P(z,\theta)C(z)}{H(z,\theta)} y(t)$$
(6)

Define the residual $\varepsilon(t,\theta)$ as that.

$$\varepsilon(t,\theta) = y(t) - \hat{y}(t,\theta)$$

= $\frac{1 + P(z,\theta)C(z)}{H(z,\theta)} \left[y(t) - \frac{P(z,\theta)}{1 + P(z,\theta)C(z)} r(t) \right]$ (7)

Based on the measured input-output variables $Z_N = \{r(t), y(t)\}_{t=1}^N$, that unknown parameter vector θ is given as the following numerical optimization problem.

$$\hat{\theta}_{N} = \arg\min_{\theta} V_{N}(t,\theta,Z_{N}) = \arg\min_{\theta} \frac{1}{N} \sum_{t=1}^{N} \varepsilon^{2}(t,\theta)$$
(8)

where $\hat{\theta}_N$ denoted the parameter estimation based on measured input-output $Z_N = \{r(t), y(t)\}_{t=1}^N$.

Substituting residual $\varepsilon(t,\theta)$ into numerical optimization (8), then lots of optimization methods can be applied to minimize that cost function or identification criterion $\frac{1}{N}\sum_{t=1}^{N}\varepsilon^{2}(t,\theta)$, for example least squares method, Newton method, and gradient method etc.

The above description on closed loop system identification is existed and mature, referring to reference [1].

3. Asymptotic analysis

As this paper does not concern on the optimal input design and identification algorithm for closed loop system , but on finite sample analysis on closed loop system identification. Before starting our new results about finite sample analysis, we firstly give our derived asymptotic analysis result, as some results in this section 3 will be used in further study. From some knowledge about statistical signal processing, assume external excitation input r(t) be quasi-stationary with power spectrum $\phi_r(w)$, where *w* is one spectrum variable. The variance for white noise e(t) is λ_0 , then power spectrum $\phi_v(w)$ for the filter noise v(t) is

$$\phi_{v}(w) = H_{0}(e^{jw})H_{0}^{*}(e^{jw})\lambda_{0}$$

where notation * means complex conjugate.

3.1 Asymptotic analysis for cost function

Using the uncorrelated condition between r(t) and e(t), the output spectrum and input spectrum in equation (4) are yielded respectively as.

$$\begin{cases} \phi_{y}(w) = |P_{0}|^{2} |S_{0}|^{2} \phi_{r}(w) + |H_{0}|^{2} |S_{0}|^{2} \lambda_{0} \\ \phi_{u}(w) = |S_{0}|^{2} \phi_{r}(w) + |C|^{2} |H_{0}|^{2} |S_{0}|^{2} \lambda_{0} \\ = \phi_{u}^{r}(w) + \phi_{u}^{e}(w) \end{cases}$$
(9)

where in equation (9), variance z is neglected, $\phi_u^r(w)$ and $\phi_u^e(w)$ correspond to components, coming from the external excitation input and external noise respectively.

Similarly the cross spectrums are given as.

$$\begin{cases} \phi_{yu}(w) = P_0 |S_0|^2 \phi_r(w) - C |H_0|^2 |S_0|^2 \lambda_0 \\ \phi_{ue}(w) = -CH_0 S_0 \lambda_0 \end{cases}$$
(10)

Observing equation (9) and (10), we see that

$$\begin{aligned} \lambda_{0}\phi_{u}(w) - |\phi_{ue}(w)|^{2} &= \lambda_{0} |S_{0}|^{2} \phi_{r}(w) \\ &+ |C|^{2} |H_{0}|^{2} |S_{0}|^{2} \lambda_{0}^{2} - |C|^{2} |H_{0}|^{2} |S_{0}|^{2} \lambda_{0}^{2} \qquad (11) \\ &= \lambda_{0} |S_{0}|^{2} \phi_{r}(w) = \lambda_{0}\phi_{u}^{r}(w) \end{aligned}$$

The above spectrum relations are benefit for spectral estimation of the transfer function $P_0(z)$. From a practical point of view, the spectral estimate is defined as.

$$\hat{P}(e^{jw}) = \frac{\phi_{yu}(w)}{\phi_{u}(w)} = \frac{P_{0}|S_{0}|^{2} \phi_{r}(w) - C|H_{0}|^{2}|S_{0}|^{2} \lambda_{0}}{|S_{0}|^{2} \phi_{r}(w) + |C|^{2}|H_{0}|^{2}|S_{0}|^{2} \lambda_{0}}$$

$$= \frac{P_{0}(e^{jw})\phi_{r}(w) - C(e^{jw})|H_{0}(e^{jw})|^{2} \lambda_{0}}{\phi_{r}(w) + |C(e^{jw})|^{2}|H_{0}(e^{jw})|^{2} \lambda_{0}} \qquad (12)$$

$$= \frac{P_{0}(e^{jw})\phi_{r}(w) - C(e^{jw})\phi_{v}(w)}{\phi_{r}(w) + |C(e^{jw})|^{2} \phi_{v}(w)}$$

When the number of measured data approaches to infinity $(N \rightarrow \infty)$, the above relations hold, i.e.

$$\begin{cases} P_0\left(e^{jw}\right)\phi_r\left(w\right) - C\left(e^{jw}\right)\phi_v\left(w\right) \to \phi_{yu}\left(w\right) \\ \phi_r\left(w\right) + \left|C\left(e^{jw}\right)\right|^2\phi_v\left(w\right) \to \phi_u\left(w\right) \end{cases}$$
(13)

From the asymptotic theory, we see that if *N* tends to infinity, spectral estimate $\hat{P}(e^{jw})$ converges to expression

$$\frac{P_{0}\left(e^{jw}\right)\phi_{r}\left(w\right)-C\left(e^{jw}\right)\phi_{v}\left(w\right)}{\phi_{r}\left(w\right)+\left|C\left(e^{jw}\right)\right|^{2}\phi_{v}\left(w\right)}$$

but it is one biased estimate. As residual $\varepsilon(t,\theta)$ is implicit with unknown parameter vector θ , we rewrite it as that.

$$\varepsilon(t,\theta) = H^{-1}(z,\theta) \Big[y(t) - P(z,\theta)u(t) \Big]$$
(14)

Further suppose one true parameter vector θ_0 exist and satisfy

$$P(z,\theta_0) = P_0(z); H(z,\theta_0) = H_0(z)$$

After substituting equation (2) into (14), it holds that

$$y(t) - P(z,\theta)u(t) = \frac{P_0(z)}{1 + P_0(z)C(z)}r(t) + \frac{H_0(z)}{1 + P_0(z)C(z)}e(t) - \frac{P(z,\theta)}{1 + P_0(z)C(z)}r(t) - \frac{P(z,\theta)C(z)H_0(z)}{1 + P_0(z)C(z)}e(t) (15) = \frac{P_0(z) - P(z,\theta)}{1 + P_0(z)C(z)}r(t) + \frac{H_0(z)[1 - P(z,\theta)C(z)]}{1 + P_0(z)C(z)}e(t)$$

Then residual $\varepsilon(t,\theta)$ is rewritten as again.

$$\varepsilon(t,\theta) = \frac{P_{0}(z) - P(z,\theta)}{H(z,\theta)} \frac{1}{1 + P_{0}(z)C(z)} r(t) + \frac{H_{0}(z)}{H(z,\theta)} \frac{\left[1 - P(z,\theta)C(z)\right]}{1 + P_{0}(z)C(z)} e(t) = \frac{P_{0}(z) - P(z,\theta)}{H(z,\theta)} S_{0}(z)r(t) + \frac{H_{0}(z)}{H(z,\theta)} \frac{\left[1 - P(z,\theta)C(z)\right]}{1 + P_{0}(z)C(z)} e(t)$$
(16)

On the condition of this above new residual, then the cost function $V_N(\theta)$ in equation (8) is also rewritten as that.

$$V_{N}(\theta) = \frac{1}{N} \sum_{t=1}^{N} \varepsilon^{2}(t,\theta)$$

= $\int_{-\pi}^{\pi} \frac{\left|P_{0}(e^{jw}) - P(e^{jw},\theta)\right|^{2}}{H(e^{jw},\theta)} \phi_{u}^{r}(w) dw$ (17)
+ $\int_{-\pi}^{\pi} \frac{1}{\left|H(e^{jw},\theta)^{2}\right|} \left|\frac{1 + P(e^{jw},\theta)C(e^{jw})}{1 + P_{0}(e^{jw})C(e^{jw})}\right|^{2} \phi_{v}(w) dw$

where in equation (17), Parseval's relations is used. From equation (17), the minimum value is yielded in case of $P(e^{jw}, \theta) \rightarrow P_0(e^{jw})$ or $\hat{\theta} \rightarrow \theta_0$, but they hold when *N* tends to infinity. This condition corresponds to the asymptotic analysis.

3.2 Asymptotic analysis for optimal feedback controller

In section 3.1, asymptotic theory is applied to rewrite the cost function $V_N(\theta)$, based on our derived cost function (17), we seek its optimal feedback controller C(z).

Adding one weight matrix N(w) into the derived cost function $V_N(\theta)$, we denote the cost function $V_N(\theta)$ as $V(\phi_u(w), \phi_{ue}(w))$, i.e.

$$V(\phi_{u}(w),\phi_{ue}(w)) = \int_{-\pi}^{\pi} \left[\frac{\lambda_{0}N_{12}(w) - 2\operatorname{Re}[N_{12}(w)\phi_{ue}(-w)]}{\lambda_{0}\phi_{u}(w) - |\phi_{ue}(w)|^{2}} + \frac{N_{22}(w)\phi_{u}(w)}{\lambda_{0}\phi_{u}(w) - |\phi_{ue}(w)|^{2}} \right] \phi_{v}(w) dw$$
(18)

where the weight matrix N(w) is one 2×2 matrix, i.e.

$$N(w) = \begin{bmatrix} N_{11}(w) & N_{12}(w) \\ N_{21}(w) & N_{22}(w) \end{bmatrix}$$

Rewrite the weight matrix N(w) as one special form, i.e.

$$N(w) = S(w) \begin{bmatrix} \left| M\left(e^{jw}\right) \right|^2 & M\left(e^{jw}\right) \\ M\left(e^{-jw}\right) & 1 \end{bmatrix}$$
(19)

Substituting the above chosen weight matrix into cost function, it holds that.

$$\frac{\lambda_{0}S(w)|M(e^{jw})|^{2} - 2[S(w)M(e^{jw})\phi_{ue}(w)]}{\lambda_{0}\phi_{u}(w) - |\phi_{ue}(w)|^{2}} + \frac{S(w)\phi_{u}(w)}{\lambda_{0}\phi_{u}(w) - |\phi_{ue}(w)|^{2}}\phi_{v}(w) \\
= S(w) \\
\times \frac{\lambda_{0}|M(e^{jw})|^{2} - 2[M(e^{jw})\phi_{ue}(w)] + \phi_{u}(w)}{\lambda_{0}\phi_{u}(w) - |\phi_{ue}(w)|^{2}}\phi_{v}(w) \\
= \frac{S(w)\phi_{v}(w)}{\lambda_{0}} \\
\times \frac{\lambda_{0}^{2}|M(e^{jw})|^{2} - 2\lambda_{0}M(e^{jw})\phi_{ue}(w) + \phi_{ue}^{2}(w)}{\lambda_{0}\phi_{u}(w) - |\phi_{ue}(w)|^{2}}\phi_{v}(w) \\
+ \frac{\lambda_{0}\phi_{u}(w) - \phi_{ue}^{2}(w)}{\lambda_{0}\phi_{u}(w) - |\phi_{ue}(w)|^{2}}\phi_{v}(w) \\
= \frac{S(w)\phi_{v}(w)}{\lambda_{0}\phi_{u}(w) - |\phi_{ue}(w)|^{2}}\phi_{v}(w)$$
(20)

Then the cost function is that.

$$\min_{\phi_{u}(w),\phi_{ue}(w)} \int_{-\pi}^{\pi} \frac{S(w)\phi_{v}(w)}{\lambda_{0}} \left[1 + \frac{\left|\lambda_{0}^{2}M\left(e^{jw}\right) - \phi_{ue}\left(w\right)\right|^{2}}{\lambda_{0}\phi_{u}\left(w\right) - \left|\phi_{ue}\left(w\right)\right|^{2}} \right] dw$$
(21)

Applying our existed relations in equation (21).

$$\begin{cases} \phi_{ue}(w) = -CH_{0}S_{0}\lambda_{0} \\ \phi_{u}(w) = |S_{0}|^{2} \phi_{r}(w) + |C|^{2} |H_{0}|^{2} |S_{0}|^{2} \lambda_{0} \\ = \phi_{u}^{r}(w) + \phi_{u}^{e}(w) \end{cases}$$

i.e. it holds that

$$\frac{\left|\lambda_{0}^{2}M\left(e^{jw}\right)-\phi_{ue}\left(w\right)\right|^{2}}{\lambda_{0}\phi_{u}\left(w\right)-\left|\phi_{ue}\left(w\right)\right|^{2}}=\frac{\lambda_{0}^{2}M\left(e^{jw}\right)+CH_{0}S_{0}\lambda_{0}}{\lambda_{0}S_{0}\phi_{r}\left(w\right)}$$
(22)

substituting equation (22) into (21), the cost function is reduced to that.

$$\min_{\phi_{u}(w),\phi_{uc}(w)} \int_{-\pi}^{\pi} \frac{\lambda_{0}^{2} M\left(e^{jw}\right) + C H_{0} S_{0} \lambda_{0}}{\lambda_{0} S_{0} \phi_{r}\left(w\right)} dw \qquad (23)$$

It is easy to see that the optimal feedback controller C(z) must satisfy that.

$$C(z)H_0(z)S_0(z) = -M(z)$$

i.e.

$$\frac{C(z)H_0(z)}{1+P_0(z)C(z)} = -M(z)$$

$$C(z)H_0(z) = -M(z)\left[1+P_0(z)C(z)\right]$$

It yields that

$$C(z) = -\frac{M(z)}{P_0(z)M(z) + H_0(z)}$$
(24)

where in equation (24), M(z) is given, and $(P_0(z), H_0(z))$ can be replaced by their parametrized form $(P(z, \theta), H(z, \theta))$, then optimal feedback controller C(z) is chosen as.

$$C(z) = -\frac{M(z)}{P(z,\theta)M(z) + H(z,\theta)}$$
(25)

During above mathematical derivations, asymptotic theory is used too, for example, that Parseval's relation and some spectral relations.

Observing equation (25) again, optimal feedback controller C(z) is dependent of the parametrized form $(P(z,\theta), H(z,\theta))$, so we need to analyze this closed dependence through the variance expression for our derived optimal feedback controller.

Consider two disturbances or perturbs exist in parametrized forms $(P(z,\theta), H(z,\theta))$,, without loss of generality set

$$(P(z,\theta),H(z,\theta)) = (P_0(z),H_0(z)) + (\Delta P,\Delta H),$$

i.e.

$$\begin{pmatrix} P(z,\theta), H(z,\theta) \end{pmatrix} = \begin{pmatrix} P_0(z), H_0(z) \end{pmatrix} + (\Delta P, \Delta H)$$

$$\begin{cases} \Delta P = P(z,\theta) - P_0(z) \\ \Delta H = H(z,\theta) - H_0(z) \end{cases}$$
 (26)

substituting these two perturbs into optimal feedback controller, we have

$$C(z) = \frac{M(z)}{[P_{0}(z) + \Delta P]M(z) + [H_{0}(z) + \Delta H]}$$

$$= \frac{M(z)}{P_{0}(z)M(z) + \Delta PM(z) + [H_{0}(z) + \Delta H]}$$

$$= M(z) \begin{bmatrix} \frac{1}{P_{0}(z)M(z) + H_{0}(z)} \\ -\frac{1}{[P_{0}(z)M(z) + H_{0}(z)]^{2}} \\ \times [\Delta PM(z) + \Delta H] \end{bmatrix}$$

$$= \frac{M(z)}{P_{0}(z)M(z) + H_{0}(z)}$$

$$- \frac{M(z)}{[P_{0}(z)M(z) + H_{0}(z)]^{2}} [\Delta PM(z) + \Delta H]$$

$$= C_{0}(z)$$

$$- \frac{M(z)}{[P_{0}(z)M(z) + H_{0}(z)]^{2}} [\Delta PM(z) + \Delta H]$$

where for the latter analysis, $C_0(z)$ means the nominate controller.

Define the perturb, existing in optimal feedback controller as

$$\Delta C = C(z) - C_0(z)$$

$$= -\frac{M(z)}{\left[P_0(z)M(z) + H_0(z)\right]^2} \left[M(z) \quad 1\right] \left[\begin{array}{c} \Delta P\\ \Delta H\end{array}\right] (28)$$

$$= -\left[C_0^2(z) \quad \frac{C_0(z)}{P_0(z)M(z) + H_0(z)}\right] \left[\begin{array}{c} \Delta P\\ \Delta H\end{array}\right]$$

where the following relations are used.

$$C_{0}(z) - \frac{M(z)}{\left[P_{0}(z)M(z) + H_{0}(z)\right]^{2}} \left[\Delta PM(z) + \Delta H\right]$$

= $C_{0}(z) - \frac{M^{2}(z)}{\left[P_{0}(z)M(z) + H_{0}(z)\right]^{2}} \Delta P$
 $- \frac{M(z)}{\left[P_{0}(z)M(z) + H_{0}(z)\right]^{2}} \Delta H$
= $C_{0}(z) - C_{0}^{2}(z)\Delta P - \frac{C_{0}(z)}{P_{0}(z)M(z) + H_{0}(z)} \Delta H$

Based on the derived perturb ΔC , variance of the optimal feedback controller C(z) is given as.

$$P_{c} = E \Big[C(z) - C_{0}(z) \Big]^{2}$$

$$= \Big[C_{0}^{2}(z) \frac{C_{0}(z)}{P_{0}(z)M(z) + H_{0}(z)} \Big] \Big[\Delta P \\ \Delta H \Big]$$

$$\times [\Delta P \quad \Delta H] \Big[\frac{C_{0}^{2}(z)}{P_{0}(z)M(z) + H_{0}(z)} \Big]$$
(29)
$$= \Big[C_{0}^{2}(z) \frac{C_{0}(z)}{P_{0}(z)M(z) + H_{0}(z)} \Big]$$

$$\times Cov_{(P,H)} \Big[\frac{C_{0}^{2}(z)}{P_{0}(z)M(z) + H_{0}(z)} \Big]$$

where $Cov_{(P,H)}$ is variance for parametrized forms $(P(z,\theta), H(z,\theta))$ and its explicit form for $Cov_{(P,H)}$ is seen in our previous work [28]. Substituting our existed result from [28] in equation (29), then variance P_c is yielded as that.

$$P_{C} = \begin{bmatrix} C_{0}^{2}(z) & \frac{C_{0}(z)}{P_{0}(z)M(z) + H_{0}(z)} \end{bmatrix}$$

$$\times \begin{bmatrix} \frac{n}{N} \frac{|H_{0}|^{2} \lambda_{0}}{|S_{0}|\phi_{r}(w)} & 0\\ 0 & \frac{n}{N} |H_{0}|^{2} \end{bmatrix} \begin{bmatrix} C_{0}^{2}(z) \\ C_{0}(z) \\ \frac{P_{0}(z)M(z) + H_{0}(z)}{|P_{0}(z)M(z) + H_{0}(z)} \end{bmatrix}$$
(30)

where *n* is the number of parameter vector θ . Further observing equation (28) again, the biased error is that.

$$E[\Delta C] = -C_0^2(z)E[\Delta P] + \frac{C_0(z)}{P_0(z)M(z) + H_0(z)}E[\Delta H]$$
(31)

So if all parametrized forms $(P(z,\theta), H(z,\theta))$ are unbiased estimates, i.e. $E[\Delta P] = E[\Delta H] = 0$, then our derived optimal feedback controller C(z) is also one unbiased estimate, i.e.

$$E[C(z)] = C_0(z), N \to \infty$$

Generally asymptotic theory is applied to analyze that cost function, optimal feedback controller and its corresponding variance.

4. Finite sample analysis

During above mathematical derivations, some results hold on the basis of infinite data, i.e. $N \rightarrow \infty$, for example that asymptotic variance (30) for the optimal feedback controller C(z). We

regard those results in section 3 as asymptotic analysis, but in practice, that finite data is an ideal case, and the number of measured data is truly finite. Then within this framework of finite sample data, how can we use the sampled identification criterion to replace its corresponding limit or expected criterion? This section is to measure the difference between the sampled identification criterion and expected criterion. Based on this explicit difference, the number of measured data is obtained to guarantee the difference not great one given positive scalar. This necessary number of sampled data

structure. For the sake of completeness, consider that residual $\varepsilon(t,\theta)$ (7) again.

corresponds to sample complexity in closed loop

$$\varepsilon(t,\theta) = y(t) - \hat{y}(t,\theta)$$

= $\frac{1 + P(z,\theta)C(z)}{H(z,\theta)} \left[y(t) - \frac{P(z,\theta)}{1 + P(z,\theta)C(z)} r(t) \right]$

Substituting that optimal feedback controller C(z) (25) into above residual, it holds that.

$$P(z,\theta)C(z) = -\frac{P(z,\theta)M(z)}{P(z,\theta)M(z) + H(z,\theta)}$$

$$1 + P(z,\theta)C(z) = 1 - \frac{P(z,\theta)M(z)}{P(z,\theta)M(z) + H(z,\theta)}$$

$$= \frac{H(z,\theta)}{P(z,\theta)M(z) + H(z,\theta)}$$

$$\frac{P(z,\theta)}{1 + P(z,\theta)C(z)} = \frac{P(z,\theta)[P(z,\theta)M(z) + H(z,\theta)]}{H(z,\theta)}$$
(32)

Combing above equities, we get

$$\varepsilon(t,\theta) = \frac{1}{P(z,\theta)M(z) + H(z,\theta)}$$

$$\times \left[y(t) - \frac{P(z,\theta)[P(z,\theta)M(z) + H(z,\theta)]}{H(z,\theta)} r(t) \right] \quad (33)$$

$$= \frac{1}{P(z,\theta)M(z) + H(z,\theta)} y(t) - \frac{P(z,\theta)}{H(z,\theta)} r(t)$$

Observing that identification criterion (8) again, as only N terms exist in this identification criterion, we call it as the empirical value or sampled value of identification criterion, and denote it as

$$V_{N}(\theta) = \frac{1}{N} \sum_{t=1}^{N} \varepsilon^{2}(t,\theta)$$
(34)

To measure this sampled identification criterion, we

need one ideal value to compare, i.e. one expected criterion, which is defined as follows.

$$V(\theta) = \frac{1}{N} \sum_{t=1}^{N} E\left[\varepsilon^{2}(t,\theta)\right]$$
(35)

where expectation operation E[] is introduced in

above equation (35).

To compare the difference between $V_{N}(\theta)$ and $V(\theta)$,

i.e. $|V_N(\theta) - V(\theta)|$, the square term is needed. From equation (1), we have

$$\begin{cases} y^{2}(t) = P_{0}^{2}(z)S_{0}^{2}(z)r^{2}(t) + H_{0}^{2}(z)S_{0}^{2}(z)e^{2}(t) \\ + 2P_{0}(z)H_{0}(z)S_{0}^{2}(z)r(t)e(t) \\ y(t)r(t) = P_{0}(z)S_{0}(z)r^{2}(t) + H_{0}(z)S_{0}(z)r(t)e(t) \end{cases}$$
(36)

Using equation (36), we compute the square term $\varepsilon^2(t,\theta)$ as follows.

$$\begin{aligned} \varepsilon^{2}(t,\theta) &= \left[\frac{1}{P(z,\theta)M(z) + H(z,\theta)}\right]^{2} \\ \times \left[P_{0}^{2}(z)S_{0}^{2}(z)r^{2}(t) + H_{0}^{2}(z)S_{0}^{2}(z)e^{2}(t)\right] \\ &+ 2P_{0}(z)H_{0}(z)S_{0}^{2}(z)r(t)e(t) \end{bmatrix} \\ + \left[\frac{P(z,\theta)}{H(z,\theta)}\right]^{2}r^{2}(t) \\ - \frac{2P(z,\theta)}{\left[P(z,\theta)M(z) + H(z,\theta)\right]H(z,\theta)} \\ \times \left[P_{0}(z)S_{0}(z)r^{2}(t) + H_{0}(z)S_{0}(z)r(t)e(t)\right] \\ &= \left[\frac{P_{0}^{2}(z)S_{0}^{2}(z)}{\left[P(z,\theta)M(z) + H(z,\theta)\right]^{2}} \\ + \left[\frac{P(z,\theta)}{H(z,\theta)}\right]^{2} \\ - \frac{2P(z,\theta)P_{0}(z)S_{0}(z)}{\left[P(z,\theta)M(z) + H(z,\theta)\right]^{2}H(z,\theta)} \right]r^{2}(t) \\ &+ \left[\frac{2P_{0}(z)H_{0}(z)S_{0}^{2}(z)}{\left[P(z,\theta)M(z) + H(z,\theta)\right]^{2}} \\ + \left[\frac{2P(z,\theta)H_{0}(z)S_{0}(z)}{\left[P(z,\theta)M(z) + H(z,\theta)\right]^{2}} \\ + \left[\frac{2P(z,\theta)H_{0}(z)S_{0}(z)}{\left[P(z,\theta)M(z) + H(z,\theta)\right]H(z,\theta)}\right]r(t)e(t) (37) \\ &+ \frac{H_{0}^{2}(z)S_{0}^{2}(z)}{\left[P(z,\theta)M(z) + H(z,\theta)\right]^{2}}e^{2}(t) \end{aligned}$$

Taking expectation operation E[] on both sides of equation (37), it yields that difference between $|V_N(\theta) - V(\theta)|$.

$$\frac{1}{N}\sum_{t=1}^{N}\varepsilon^{2}(t,\theta) - \frac{1}{N}\sum_{t=1}^{N}E\varepsilon^{2}(t,\theta)$$

$$= \frac{1}{N}\sum_{t=1}^{N}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} \left[\frac{H_{0}^{2}(z)S_{0}^{2}(z)}{\left[P(z,\theta)M(z) + H(z,\theta)\right]^{2}} \right]_{k,l}$$

$$\times \left(e(t-k)e(t-l) - E\left[e(t-k)e(t-l)\right]\right)$$

$$+ \frac{1}{N}\sum_{t=1}^{N}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} \left[\frac{2P_{0}(z)H_{0}(z)S_{0}^{2}(z)}{\left[P(z,\theta)M(z) + H(z,\theta)\right]^{2}} - \frac{2P(z,\theta)H_{0}(z)S_{0}(z)}{\left[P(z,\theta)M(z) + H(z,\theta)\right]H(z,\theta)} \right]_{k,l}$$

$$\left(r(t-k)e(t-l) - E\left[r(t-k)e(t-l)\right]\right)$$

$$= \frac{1}{N}\sum_{t=1}^{N}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} \left[N_{1}(z)\right]_{k,l,\theta} \left(e(t-k)e(t-l) - \lambda_{0}\right)$$

$$+ \frac{1}{N}\sum_{t=1}^{N}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty} \left[N_{2}(z)\right]_{k,l,\theta} \left(r(t-k)e(t-l) - \lambda_{0}\right)$$
(38)

where in equation (38), we use the following known equities.

$$r(t-k)r(t-l) = E(r(t-k)r(t-l))$$
$$E(e(t-k)e(t-l)) = \begin{cases} \lambda_0 & k = l \\ 0 & k \neq l \end{cases}$$
$$E(r(t-k)e(t-l)) = 0$$

To give a bound for the difference (38), set

$$n_{1} = \sup_{\theta} \left[N_{1}(z) \right]_{k,l,\theta}; n_{2} = \sup_{\theta} \left[N_{2}(z) \right]_{k,l,\theta}$$

Then equation (38) is simplified as that.

$$\frac{1}{N}\sum_{t=1}^{N}\varepsilon^{2}(t,\theta) - \frac{1}{N}\sum_{t=1}^{N}E\varepsilon^{2}(t,\theta)$$
$$\leq \frac{1}{N}\sum_{t=1}^{N}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}n_{1}(e(t-k)e(t-l)-\lambda_{0})$$
$$+ \frac{1}{N}\sum_{t=1}^{N}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}n_{2}\binom{r(t-k)e(t-l)}{-E[r(t-k)e(t-l)]}$$

From the statistical signal processing, the bound for the first subtract is given as.

$$\frac{1}{N}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}n_{1}\left(e(t-k)e(t-l)-\lambda_{0}\right)$$

$$\leq\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}n_{1}\left(k+l+1\right)\varepsilon_{0}\leq\varepsilon_{1}$$
(39)

where ε_0 and ε_1 are two positive scalars, similarly the similar result holds for the second subtract.

$$\frac{1}{N}\sum_{t=1}^{N}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}n_{2}\binom{r(t-k)e(t-l)}{-E[r(t-k)e(t-l)]} \leq \varepsilon_{2}$$

$$(40)$$

where ε_2 is also one positive scalar.'

Observing equation (39) and (40), and set $\varepsilon = \varepsilon_1 + \varepsilon_2$, then we have

$$\left|\frac{1}{N}\sum_{t=1}^{N}\varepsilon^{2}(t,\theta)-\frac{1}{N}\sum_{t=1}^{N}E\varepsilon^{2}(t,\theta)\right|\leq\varepsilon=\varepsilon_{1}+\varepsilon_{2}$$
(41)

Equation (41) means after chosen one suitable sample complexity N, the difference

$$\left|\frac{1}{N}\sum_{t=1}^{N}\varepsilon^{2}(t,\theta)-\frac{1}{N}\sum_{t=1}^{N}E\varepsilon^{2}(t,\theta)\right|$$

will at least one given positive scalar ε . If ε is small enough, then *N* will be increased, i.e. *N* is increased with ε reducing.

5. Simulation example

In this section we apply our derived results on one single input and single output system, controlled by one feedback controller. The true data generating system is given as follows.

$$P_0(z) = \frac{0.25z^{-1} + 0.12z^{-2}}{1 - 1.6z^{-1} + 0.8z^{-2} - 0.64z^{-3} + 0.65z^{-4}}$$
$$= \frac{0.25z^3 + 0.12z^2}{z^4 - 1.6z^3 + 0.8z^2 - 0.64z + 0.65}$$
$$H_0(z) = \frac{1 + 0.2z^{-1}}{1 + 0.5z^{-1}} = \frac{z + 0.2}{z + 0.5} = 1 - \frac{0.3}{z + 0.5}$$

Their corresponding parametrized forms are denoted as

$$P(z,\theta) = \frac{a_5 z^{-1} + a_6 z^{-2}}{1 + a_1 z^{-1} + a_2 z^{-2} + a_3 z^{-3} + a_4 z^{-4}}$$
$$= \frac{a_5 z^3 + a_6 z^2}{z^4 + a_1 z^3 + a_2 z^2 + a_3 z + a_4}$$
$$H(z,\theta) = \frac{z + b_2}{z + b_1}$$

A Gaussian white noise e(t) with variance $\lambda_0 = 0.5$ is added through the noise filter $H_0(z)$. The sampled time is $T_s = 1$ second, the true parameter vector θ_0 is defined as.

 $\theta_0 = [-1.6, 0.8, -0.64, 0.65, 0.25, 0.12, 0.5, 0.2]^T$

The data generating system is operated in one closed loop system with one unit feedback controller. In solving that numerical optimization problem to identify the unknown parameter vector, the initial value for unknown parameter vector θ_{init} is chosen as.

$$\theta_{init} = [-1.7, 0.7, -0.4, 0.8, 0.15, 0.1, 0.4, 0.1]^T$$



The input variable u(t) has a weight of 1 and the excitation signal r(t) has also one bound, i.e. $-1 \le r(t) \le 1$. The applied input signal is shown in Figure 2, and the output signal y(t) is measured through some physical devices. Figure 3 shows the observed output signal. When using predictor error identification method to identify the unknown parameters, the condition for termination the recursive method is to guarantee the estimation error

$$\frac{\left\|\hat{\theta}(t) - \theta\right\|}{\left\|\theta\right\|}$$

be less than one small constant value, for example, 0.05.





As the goal of finite sample analysis is to drive one suitable sample complexity, which can give one reference about using the sample sum to replace the expectation operation, it means this suitable sample complexity can guarantee the error between the sample sum and the expectation operation be less than one given positive scalar. More specifically, the above description about finite sample analysis is proven through comparing the estimations and their own true values respectively. Figure 4 shows the comparisons for estimations $\{a_4, b_1\}$ and their true values. The black straight lines denote the true values $\{0.65, 0.5\}$, and the red curves correspond to the estimation values. From Figure 4, we see although the deviations between them are obvious, but after 60 iteration steps, the estimations are closed to their true values. It tells that when the sample complexity satisfies $N \ge 60$, then the difference between the sample identification criterion and expected criterion will be zero, i.e. all estimations converge to their true values.

6. Conclusion

In this paper, we study the asymptotic analysis and finite sample analysis for closed loop system identification, which correspond to infinite data and finite data. Firstly in the case of infinite data, the cost function and optimal feedback controller are obtained by our own mathematical derivations in order to simplify the identification process for closed loop system, then the variance for the optimal feedback controller is derived too. Secondly for the finite sample data, the bound for the difference between the sampled identification criterion and expected criterion is constructed. The research about applying our derived results into more complex closed loop system will be the main idea of the next paper.

References

- [1]. Ljung, L. System identification: Theory for the user: Prentice Hall. 1999.
- [2]. Pintelon R, Schoukens J. System identification: a frequency domain approach. New York: IEEE Press, 2001.
- [3]. Boyd S, L Vandenberghe. Convex optimization: UK: Cambridge university press,2008.
- [4]. Er Wei Bai. Kernal based approaches to local nonlinear nonparametric variable selection. Automatica, 2014, 50(1): 100-113.
- [5]. Henrik Ohlsson. Identification of switched linear regression models using sum of norms regularization. Automatica, 2013, 49(4): 1045-1050.
- [6]. Anne Van Mulders. Identification of systems with localised nonlinearity: from state space to block structured models. Automatica, 2013, 49(5):

1392-1396.

- [7]. G. Pillonetto. Kernel methods in system identification, machine learning and function estimation: a survey. Automatica, 2014, 50(3): 657-682.
- [8]. Paul M J Vanden Hof. Identification of dynamic models in complex networks with prediction error methods. Automatica, 2013, 49(10): 2994-3006.
- [9]. Johan Paduart. Identification of nonlinear systems using polynomial nonlinear state space models. Automatica, 2010, 46(4): 647-656.
- [10]. Anne Van Mulders. Two nonlinear optimization methods for black box identification compared. Automatica, 2010, 46(10): 1675-1681.
- [11]. Melanie Zeilinger. Real time suboptimal model predictive control using a combination of explicit MPC and online optimization. IEEE Transactions of Automatic Control, 2011, 56(7): 1524-1534.
- [12]. Christian Feller. An improved algorithm for combinatorial multi-parameteric quadratic programming. Automatica, 2013, 49(5): 1370-1376.
- [13]. G. Pillonetto. A new kernel based approach for linear system identification. Automatica, 2010, 46(1): 81-93.
- [14]. Torston soderstorn. Accuracy analysis of a covariance matching approach for identify errors in variables systems. Automatica, 2011, 47(1): 272-282.
- [15]. Maarten Schoukens. Identification of Wiener Hammerstein systems by a nonparametric separation of the best linear approximation. Automatica, 2014, 50(2): 628-634.
- [16]. Carlo Novara, Simone Formentin, Sergio M Savaresi. Data driven design of two degrees of freedom nonlinear controllers, Automatica, 2016, 72(10): 19-27.
- [17]. J M Bravo, T Alamo, M Vasallo. A general framework for predictions based on bounding techniques and local approximation, IEEE Transaction on Automatic Control, 2017, 62(7): 3430-3435.
- [18]. Marko Tanaskovic, Lorenzo Fagiano, Roy Smith. Adaptive receding horizon control for constrained MIMO systems, Automatica, 2014, 50(12):

3019-3029.

- [19]. Marko Tanaskovic, Lorenzo Fagiano, Carlo Novara. Data driven control of nonlinear systems: an on line direct approach, Automatica, 2017, 75(1):11-10.
- [20]. M Casini, A Garulli, A Vicino. A linear programming approach to online set membership parameter estimation for linear regression models, International Journal of Adaptive Control and Signal Processing, 2017, 31(3): 360-378.
- [21]. Wang Jianhong. Application of bounded error identification into model predictive control, accepted in ISA Transactions, 2017.
- [22]. X Zhang, M Kamgarpour, A Georghiou. Robust optimal control with adjustable uncertainty sets, Automatica, 2017, 75(1): 249-259.
- [23]. J M Bravo, T Alamo, E F Camacho, Bounded error identification of systems with time varying parameters, IEEE Transactions on Automatic control, 2016, 51(7):1144-1150.
- [24]. Algo Care, Balazs Csorji, M C Campi, Finite sample system identification: an overview and a new correlation method, IEEE Control Systems Letters, 2018, 2(1): 61-66.
- [25]. Erik Weyer, M C Campi, Asymptotic properties of SPS confidence regions, Automatica, 2017, 82(8): 287-294.
- [26]. Wang Jianhong. Finite sample properties of virtual reference feedback tuning control, Journal of Systems Science and Complexity, 2018, 31(3), 664-676.
- [27]. Wang Jianhong, Ricardo A. Ramirez-Mendoza. Finite sample properties of virtual reference feedback tuning with two degrees of freedom controllers, ISA Transactions, 2020, 99(4), 37-49.
- [28]. Wang Jianhong, Ricardo A. Ramirez-Mendoza. Stealth identification strategy for closed loop system structure, International Journal of Systems Science, 2020, 51(6), 1084-1101.

Creative Commons Attribution License 4.0 (Attribution 4.0 International, CC BY 4.0)

This article is published under the terms of the Creative Commons Attribution License 4.0 https://creativecommons.org/licenses/by/4.0/deed.en_US