

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Adaptive Estimation and Control for Systems with Parametric and Nonparametric Uncertainties

Hongbin Ma* and Kai-Yew Lum†

Temasek Laboratories, National University of Singapore

tslmh@nus.edu.sg*

tsllumky@nus.edu.sg†

Abstract

Adaptive control has been developed for decades, and now it has become a rigorous and mature discipline which mainly focuses on dealing parametric uncertainties in control systems, especially linear parametric systems. Nonparametric uncertainties were seldom studied or addressed in the literature of adaptive control until new areas on exploring limitations and capability of feedback control emerged in recent years. Comparing with the approach of robust control to deal with parametric or nonparametric uncertainties, the approach of adaptive control can deal with relatively larger uncertainties and gain more flexibility to fit the unknown plant because adaptive control usually involves adaptive estimation algorithms which play role of “learning” in some sense.

This chapter will introduce a new challenging topic on dealing with both parametric and nonparametric internal uncertainties in the same system. The existence of both two kinds of uncertainties makes it very difficult or even impossible to apply the traditional recursive identification algorithms which are designed for parametric systems. We will discuss by examples why conventional adaptive estimation and hence conventional adaptive control cannot be applied directly to deal with combination of parametric and nonparametric uncertainties. And we will also introduce basic ideas to handle the difficulties involved in the adaptive estimation problem for the system with combination of parametric and nonparametric uncertainties. Especially, we will propose and discuss a novel class of adaptive estimators, i.e. *information-concentration (IC) estimators*. This area is still in its infant stage, and more efforts are expected in the future for gaining comprehensive understanding to resolve challenging difficulties.

Furthermore, we will give two concrete examples of semi-parametric adaptive control to demonstrate the ideas and the principles to deal with both parametric and nonparametric uncertainties in the plant. (1) In the first example, a simple first-order discrete-time nonlinear system with both kinds of internal uncertainties is investigated, where the uncertainty of non-parametric part is characterized by a Lipschitz constant L , and the nonlinearity of parametric part is characterized by an exponent index b . In this example, based on the idea of the IC estimator, we construct a unified adaptive controller in both cases of $b = 1$ and

$b > 1$, and its closed-loop stability is established under some conditions. When the parametric part is bilinear ($b = 1$), the conditions given reveal the magic number $\frac{3}{2} + \sqrt{2}$ which appeared in previous study on capability and limitations of the feedback

mechanism. (2) In the second example with both parametric uncertainties and non-parametric uncertainties, the controller gain is also supposed to be unknown besides the unknown parameter in the parametric part, and we only consider the noise-free case. For this model, according to some *a priori* knowledge on the non-parametric part and the unknown controller gain, we design another type of adaptive controller based on a gradient-like adaptation law with time-varying deadzone so as to deal with both kinds of uncertainties. And in this example we can establish the asymptotic convergence of tracking error under some mild conditions, although these conditions required are not as perfect as in the first

example in sense that $L < 0.5$ is far away from the best possible bound $\frac{3}{2} + \sqrt{2}$.

These two examples illustrate different methods of designing adaptive estimation and control algorithms. However, their essential ideas and principles are all based on the *a priori* knowledge on the system model, especially on the parametric part and the non-parametric part. From these examples, we can see that the closed-loop stability analysis is rather nontrivial. These examples demonstrate new adaptive control ideas to deal with two kinds of internal uncertainties simultaneously and illustrates our elementary theoretical attempts in establishing closed-loop stability.

1. Introduction

This chapter will focus on a special topic on adaptive estimation and control for systems with parametric and nonparametric uncertainties. Our discussion on this topic starts with a very brief introduction to adaptive control.

1.1 Adaptive Control

As stated in [SB89], "Research in adaptive control has a long and vigorous history" since the initial study in 1950s on adaptive control which was motivated by the problem of designing autopilots for air-craft operating at a wide range of speeds and altitudes. With decades of efforts, adaptive control has become a rigorous and mature discipline which mainly focuses on dealing parametric uncertainties in control systems, especially linear parametric systems.

From the initial stage of adaptive control, this area has been aiming at study how to deal with *large uncertainties* in control systems. This goal of adaptive control essentially means that one adaptive control law cannot be a *fixed* controller with fixed structure and fixed parameters because any fixed controller usually can only deal with *small uncertainties* in control systems. The fact that most fixed controllers with certain structure (e.g. linear feedback control) designed for an exact system model (called *nominal model*) can also work for a small range of changes in the system parameter is often referred to as *robustness*, which is the kernel concept of another area, *robust control*. While robust control focuses on studying the stability margin of fixed controllers (mainly linear feedback controller), whose

design essentially relies on priori knowledge on exact nominal system model and bounds of uncertain parameters, adaptive control generally does not need a priori information about the bounds on the uncertain or (slow) time-varying parameters. Briefly speaking, comparing with the approach of robust control to deal with parametric or nonparametric uncertainties, the approach of adaptive control can deal with relatively larger uncertainties and gain more flexibility to fit the unknown plant because adaptive control usually involves adaptive estimation algorithms which play role of “learning” in some sense.

The advantages of adaptive control come from the fact that adaptive controllers can adapt themselves to modify the control law based on estimation of unknown parameters by recursive identification algorithms. Hence the area of adaptive control has close connections with system identification, which is an area aiming at providing and investigating mathematical tools and algorithms that build dynamical models from measured data. Typically, in system identification, a certain model structure is chosen by the user which contains unknown parameters and then some recursive algorithms are put forward based on the structural features of the model and statistical properties of the data or noise. The methods or algorithms developed in system identification are borrowed in adaptive control in order to estimate the unknown parameters in the closed loop. For convenience, the parameter estimation methods or algorithms adopted in adaptive control are often referred to as *adaptive estimation* methods. Adaptive estimation and system identification share many similar characteristics, for example, both of them originate and benefit from the development of statistics. One typical example is the frequently used least-squares (LS) algorithm, which gives parameter estimation by minimizing the sum of squared errors (or residuals), and we know that LS algorithm plays important role in many areas including statistics, system identification and adaptive control. We shall also remark that, in spite of the significant similarities and the same origin, adaptive estimation is different from system identification in sense that adaptive estimation serves for adaptive control and deals with dynamic data generated in the closed loop of adaptive controller, which means that statistical properties generally cannot be guaranteed or verified in the analysis of adaptive estimation. This unique feature of adaptive estimation and control brings many difficulties in mathematical analysis, and we will show such difficulties in later examples given in this paper.

1.2 Linear Regression Model and Least Square Algorithm

Major parts in existing study on regression analysis (a branch of statistics) [DS98, Ber04, Wik08j], time series analysis [BJR08, Tsa05], system identification [Lju98, VV07] and adaptive control [GS84, AW89, SB89, CG91, FL99] center on the following linear regression model

$$z_k = \theta^T \phi_k + v_k \quad (1)$$

where $\{z_k\}$, ϕ_k , v_k represent observation data, regression vector and noise disturbance (or external uncertainties), respectively. Here θ is the unknown parameter to be estimated. Linear regression models have many applications in many disciplines of science and engineering [Wik08g, web08, DS98, Hel63, Wei05, MPV07, Fox97, BDB95]. For example, as

stated in [web08], *Linear regression is probably the most widely used, and useful, statistical technique for solving environmental problems. Linear regression models are extremely powerful, and have the power to empirically tease out very complicated relationships between variables.* Due to the importance of model (1.1), we list several simple examples for illustration:

- Assume that a series of (stationary) data (x_k, y_k) ($k = 1, 2, \dots, N$) are generated from the following model

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

where β_0, β_1 are unknown parameters, $\{x_k\}$ are i. i. d. taken from a certain probability distribution, and $\varepsilon_k \approx N(0, \sigma^2)$ is random noise independent of X . For this model, let $\theta = [\beta_0, \beta_1]^T$, $\phi_k = [1, x_k]^T$, then we have $y_k = \theta^T \phi_k + \varepsilon_k$. This example is a classic topic in statistics to study the statistical properties of parameter estimates $\hat{\theta}_N$ as the data size N grows to infinity. The statistical properties of interests may include $E(\hat{\theta} - \theta)$, $\text{Var}(\hat{\theta})$, and so on.

- Unlike the above example, in this example we assume that x_k and x_{k+1} have close relationship modeled by

$$x_{k+1} = \beta_0 + \beta_1 x_k + \varepsilon_k$$

where β_0, β_1 are unknown parameters, and $\varepsilon_k \approx N(0, \sigma^2)$ are i. i. d. random noise independent of $\{x_1, x_2, \dots, x_k\}$.

This model is an example of linear time series analysis, which aims to study asymptotic statistical properties of parameter estimates $\hat{\theta}_N$ under certain assumptions on statistical properties of ε_k . Note that for this example, it is possible to deduce an explicit expression of x_k in terms of ε_j ($j = 0, 1, \dots, k-1$).

- In this example, we consider a simple control system

$$x_{k+1} = \beta_0 + \beta_1 x_k + b u_k + \varepsilon_k$$

where $b \neq 0$ is the controller gain, ε_k is the noise disturbance at time step k . For this model, in case where b is known *a priori*, we can take; $\theta = [\beta_0, \beta_1]^T$, $\phi_k = [1, x_{k-1}]^T$, $z_k = x_k - b u_{k-1}$; otherwise, we can take $\theta = [\beta_0, \beta_1, b]^T$, $\phi_k = [1, x_{k-1}]^T$, $z_k = x_k - b u_{k-1}$. In both cases, the system can be rewritten as

$$z_k = \theta^T \phi_k + \varepsilon_k$$

which implies that intuitively, θ can be estimated by using the identification algorithm since both data z_k and ϕ_k are available at time step k . Let $\hat{\theta}_k$ denote the parameter estimates at time step $\hat{\theta}_k$, then we can design the control signal u_k by regarding $\hat{\theta}_k$ as the real parameter θ :

$$u_k = \frac{1}{\hat{b}} [r_{k+1} - \hat{\beta}_0 - \hat{\beta}_1 x_k]$$

where $\{r_k\}$ is the known reference signal to be tracked, and \hat{b} , $\hat{\beta}_0$, $\hat{\beta}_1$ are estimates of b , β_0 , β_1 , respectively. Note that for this example, the closed-loop system will be very complex because the data generated in the closed loop essentially depend on all history signals. In the closed-loop system of an adaptive controller, generally it is difficult to analyze or verify statistical properties of signals, and this fact makes that adaptive estimation and control cannot directly employ techniques or results from system identification. Now we briefly introduce the frequently-used LS algorithm for model (1.1) due to its importance and wide applications [LH74, Gio85, Wik08e, Wik08f, Wik08d]. The idea of LS algorithm is simply to minimize the sum of squared errors, that is to say,

$$\hat{\theta}_n^{LS} \triangleq \arg \min_{\zeta} \sum_{k=1}^n [z_k - \zeta^T \phi_k]^T [z_k - \zeta^T \phi_k] \quad (1.2)$$

This idea has a long history rooted from great mathematician Carl Friedrich Gauss in 1795 and published first by Legendre in 1805. In 1809, Gauss published this method in volume two of his classical work on celestial mechanics, *heoria Motus Corporum Coelestium in sectionibus conicis solem ambientium* [Gau09], and later in 1829, Gauss was able to state that the LS estimator is optimal in the sense that in a linear model where the errors have a mean of zero, are uncorrelated, and have equal variances, the best linear unbiased estimators of the coefficients is the least-squares estimators. This result is known as the Gauss-Markov theorem [Wik08a].

By Eq. (1.2), at every time step, we need to minimize the sum of squared errors, which requires much computation cost. To improve the computational efficiency, in practice we often use the recursive form of LS algorithm, often referred to as *recursive LS algorithm*, which will be derived in the following. First, introducing the following notations

$$Z_n = \begin{bmatrix} z_1^T \\ \vdots \\ z_n^T \end{bmatrix}, \quad \Phi_n = \begin{bmatrix} \phi_1^T \\ \vdots \\ \phi_n^T \end{bmatrix}, \quad V_n = \begin{bmatrix} v_1^T \\ \vdots \\ v_n^T \end{bmatrix}, \quad (1.3)$$

and using Eq. (1.1), we obtain that

$$Z_n = \Phi_n \theta + V_n.$$

Noting that

$$\begin{aligned}
 & \sum_{k=1}^n [z_k - \zeta^\tau \phi_k]^\tau [z_k - \zeta^\tau \phi_k] \\
 &= [Z_n - \Phi_n \zeta]^\tau [Z_n - \Phi_n \zeta] \\
 &= Z_n^\tau Z_n - 2\Phi_n^\tau \zeta + \zeta^\tau \Phi_n^\tau \Phi_n \zeta \\
 &= [\zeta - (\Phi_n^\tau \Phi_n)^+ \Phi_n^\tau Z_n]^\tau \Phi_n^\tau \Phi_n [\zeta - (\Phi_n^\tau \Phi_n)^+ \Phi_n^\tau Z_n] \\
 &\quad + Z_n^\tau [I - \Phi_n (\Phi_n^\tau \Phi_n)^+ \Phi_n^\tau] Z_n
 \end{aligned}$$

where the last equation is derived from properties of Moore-Penrose pseudoinverse [Wik08h]

$$\Phi_n^\tau = \Phi_n^\tau \Phi_n \Phi_n^+ = \Phi_n^\tau \Phi_n (\Phi_n^\tau \Phi_n)^+ \Phi_n^\tau$$

we know that the minimum of $[Z_n - \Phi_n \zeta]^\tau [Z_n - \Phi_n \zeta]$ can be achieved at

$$\hat{\theta}_n^{LS} = (\Phi_n^\tau \Phi_n)^+ \Phi_n^\tau Z_n \quad (1.4)$$

which is the LS estimate of θ . Let

$$P_n \triangleq (\Phi_n^\tau \Phi_n)^+$$

and then, by Eq. (1.3), with the help of matrix inverse identity

$$[A^{-1} + A^{-1} B^\tau C^{-1} B A^{-1}]^{-1} = A - B^\tau (C + B A^{-1} B^\tau)^{-1} B$$

we can obtain that

$$\begin{aligned}
 P_n &= (P_{n-1}^{-1} + \phi_n \phi_n^\tau)^{-1} \\
 &= [A^{-1} + A^{-1} B^\tau C^{-1} B A^{-1}]^{-1} \\
 &= P_{n-1} - (P_{n-1} \phi_n) [1 + (\phi_n^\tau P_{n-1}) P_{n-1}^{-1} (P_{n-1} \phi_n)]^{-1} (\phi_n^\tau P_{n-1}) \\
 &= P_{n-1} - a_n P_{n-1} \phi_n \phi_n^\tau P_{n-1}
 \end{aligned}$$

where

$$a_n = (1 + \phi_n^\tau P_{n-1} \phi_n)^{-1}$$

Further,

$$\begin{aligned}
\hat{\theta}_n &= P_n \sum_{k=1}^n \phi_k z_k^\tau \\
&= [P_{n-1} - a_n P_{n-1} \phi_n \phi_n^\tau P_{n-1}] \left[\sum_{k=1}^{n-1} \phi_k z_k^\tau + \phi_n z_n^\tau \right] \\
&= \hat{\theta}_{n-1} - a_n P_{n-1} \phi_n \phi_n^\tau P_{n-1} \hat{\theta}_{n-1} + P_{n-1} \phi_n z_n^\tau - a_n P_{n-1} \phi_n \phi_n^\tau P_n \\
&= \hat{\theta}_{n-1} - a_n P_{n-1} \phi_n \phi_n^\tau P_{n-1} \hat{\theta}_{n-1} + P_{n-1} \phi_n [1 - a_n \phi_n^\tau P_{n-1} \phi_n] z_n^\tau \\
&= \hat{\theta}_{n-1} - a_n P_{n-1} \phi_n \phi_n^\tau P_{n-1} \hat{\theta}_{n-1} + a_n P_{n-1} \phi_n z_n^\tau \\
&= \hat{\theta}_{n-1} + a_n P_{n-1} \phi_n (z_n^\tau - \phi_n^\tau \theta_{n-1})
\end{aligned}$$

Thus, we can obtain the following recursive LS algorithm

$$\hat{\theta}_n = \hat{\theta}_{n-1} + a_n P_{n-1} \phi_n (z_n^\tau - \phi_n^\tau \theta_{n-1})$$

where P_{n-1} and θ_{n-1} reflect only information up to step $n-1$, while a_n , ϕ_n and $z_n^\tau - \phi_n^\tau \theta_{n-1}$ reflect information up to step n .

In statistics, besides linear parametric regression, there also exist generalized linear models [Wik08b] and non-parametric regression methods [Wik08i], such as kernel regression [Wik08c]. Interested readers can refer to the wiki pages mentioned above and the references therein.

1.3 Uncertainties and Feedback Mechanism

By the discussions above, we shall emphasize that, in a certain sense, linear regression models are kernel of classical (discrete-time) adaptive control theory, which focuses to cope with the parametric uncertainties in linear plants. In recent years, parametric uncertainties in nonlinear plants have also gained much attention in the literature [MT95, Bos95, Guo97, ASL98, GHZ99, LQF03]. Reviewing the development of adaptive control, we find that parametric uncertainties were of primary interests in the study of adaptive control, no matter whether the considered plants are linear or nonlinear. Nonparametric uncertainties were seldom studied or addressed in the literature of adaptive control until some new areas on understanding limitations and capability of feedback control emerged in recent years. Here we mainly introduce the work initiated by Guo, who also motivated the authors' exploration in the direction which will be discussed in later parts.

Guo's work started from trying to understand fundamental relationship between the uncertainties and the feedback control. Unlike traditional adaptive theory, which focuses on investigating closed-loop stability of certain types of adaptive controllers, Guo began to think over a general set of adaptive controllers, called *feedback mechanism*, i.e., all possible feedback control laws. Here the feedback control laws need not be restricted in a certain class of controllers, and any series of mappings from the space of history data to the space of control signals is regarded as a feedback control law. With this concept in mind, since the most fundamental concept in automatic control, *feedback*, aims to reduce the effects of the

plant uncertainty on the desired control performance, by introducing the set F of internal uncertainties in the plant and the whole feedback mechanism U , we wonder the following basic problems:

1. Given an uncertainty set F , does there exist any feedback control law in U which can stabilize the plant? This question leads to the problem of how to characterize the maximum capability of feedback mechanism.
2. If the uncertainty set F is too large, is it possible that any feedback control law in U cannot stabilize the plant? This question leads to the problem of how to characterize the limitations of feedback mechanism.

The philosophical thoughts to these problems result in fruitful study [Guo97, XG00, ZG02, XG01, LX06, Ma08a, Ma08b].

The first step towards this direction was made in [Guo97], where Guo attempted to answer the following question for a nontrivial example of discrete-time nonlinear polynomial plant model with parametric uncertainty: What is the largest nonlinearity that can be dealt with by feedback? More specifically, in [Guo97], for the following nonlinear uncertain system

$$y_{t+1} = \theta \phi_t + u_t + w_{t+1}, \quad \phi_t = O(y_t^b), \quad b > 0 \quad (1.5)$$

where θ is the unknown parameter, b characterizes the nonlinear growth rate of the system, and $\{w_t\}$ is the Gaussian noise sequence, a critical stability result is found – system (1.5) is not a.s. globally stabilizable if and only if $b \geq 4$. This result indicates that there exist limitations of the feedback mechanism in controlling the discrete-time nonlinear adaptive systems, which is not seen in the corresponding continuous-time nonlinear systems (see [Guo97, Kan94]). The “impossibility” result has been extended to some classes of uncertain nonlinear systems with unknown vector parameters in [XG99, Ma08a] and a similar result for system (1.5) with bounded noise is obtained in [LX06].

Stimulated by the pioneering work in [Guo97], a series of efforts ([XG00, ZG02, XG01, MG05]) have been made to explore the maximum capability and limitations of feedback mechanism. Among these work, a breakthrough for non-parametric uncertain systems was made by Xie and Guo in [XG00], where a class of first-order discrete-time dynamical control systems

$$y_{t+1} = f(y_t) + u_t + w_{t+1}, \quad f(\cdot) \in \mathcal{F}(L) \quad (1.6)$$

is studied and another interesting critical stability phenomenon is proved by using new techniques which are totally different from those in [Guo97]. More specifically, in [XG00], $\mathcal{F}(L)$ is a class of nonlinear functions satisfying Lipschitz condition, hence the Lipschitz constant L can characterize the size of the uncertainty set $\mathcal{F}(L)$. Xie and Guo obtained the following results: if $L \geq \frac{3}{2} + \sqrt{2}$, then there exists a feedback control law such that for any

$f \in \mathcal{F}(L)$, the corresponding closed-loop control system is globally stable; and if $L < \frac{3}{2} + \sqrt{2}$, then for any feedback control law and any $y_0 \in \mathbb{R}^1$, there always exists

some $f \in F(L)$ such that the corresponding closed-loop system is unstable. So for system (1.6), the “magic” number $\frac{3}{2} + \sqrt{2}$ characterizes the capability and limits of the whole

feedback mechanism. The impossibility part of the above results has been generalized to similar high-order discrete-time nonlinear systems with single Lipschitz constant [ZG02] and multiple Lipschitz constants [Ma08a]. From the work mentioned above, we can see two different threads: one is focused on parametric nonlinear systems and the other one is focused on non-parametric nonlinear systems. By examining the techniques in these threads, we find that different difficulties exist in the two threads, different controllers are designed to deal with the uncertainties and completely different methods are used to explore the capability and limitations of the feedback mechanism.

1.4 Motivation of Our Work

From the above introduction, we know that only parametric uncertainties were considered in traditional adaptive control and non-parametric uncertainties were only addressed in recent study on the whole feedback mechanism. This motivates us to explore the following problems: When both parametric and non-parametric uncertainties are present in the system, what is the maximum capability of feedback mechanism in dealing with these uncertainties? And how to design feedback control laws to deal with both kinds of internal uncertainties? Obviously, in most practical systems, there exist parametric uncertainties (unknown model parameters) as well as non-parametric uncertainties (e.g. unmodeled dynamics). Hence, it is valuable to explore answers to these fundamental yet novel problems. Noting that parametric uncertainties and non-parametric uncertainties essentially have different nature and require completely different techniques to deal with, generally it is difficult to deal with them in the same loop. Therefore, adaptive estimation and control in systems with parametric and non-parametric uncertainties is a new challenging direction. In this chapter, as a preliminary study, we shall discuss some basic ideas and principles of adaptive estimation in systems with both parametric and non-parametric uncertainties; as to the most difficult adaptive control problem in systems with both parametric and non-parametric uncertainties, we shall discuss two concrete examples involving both kinds of uncertainties, which will illustrate some proposed ideas of adaptive estimation and special techniques to overcome the difficulties in the analysis closed-loop system. Because of significant difficulties in this new direction, it is not possible to give systematic and comprehensive discussions here for this topic, however, our study may shed light on the aforementioned problems, which deserve further investigation.

The remainder of this chapter is organized as follows. In Section 2, a simple semi-parametric model with parametric part and non-parametric part will be introduced first and then we will discuss some basic ideas and principles of adaptive estimation for this model. Later in Section 3 and Section 4, we will apply the proposed ideas of adaptive estimation and investigate two concrete examples of discrete-time adaptive control: in the first example, a discrete-time first-order nonlinear semi-parametric model with bounded external noise disturbance is discussed with an adaptive controller based on information-contraction estimator, and we give rigorous proof of closed-loop stability in case where the uncertain parametric part is of linear growth rate, and our results reveal again the magic number

$\frac{3}{2} + \sqrt{2}$; in the second example, another noise-free semi-parametric model with parametric uncertainties and non-parametric uncertainties is discussed, where a new adaptive controller based on a novel type of update law with deadzone will be adopted to stabilize the system, which provides yet another view point for the adaptive estimation and control problem for the semi-parametric model. Finally, we give some concluding remarks in Section 5.

2. Semi-parametric Adaptive Estimation: Principles and Examples

2.1 One Semi-parametric System Model

Consider the following semi-parametric model

$$z_k = \theta^T \phi_k + f(\phi_k) + \varepsilon_k \quad (2.1)$$

where $\theta \in \Theta$ denotes unknown parameter vector, $f(\cdot) \in F$ denotes unknown function and $\varepsilon_k \in \Delta_k$ denote external noise disturbance. Here Θ , F and Δ_k represent *a priori* knowledge on possible θ , $f(\phi_k)$ and ε_k , respectively. In this model, let

$$v_k = f(\phi_k) + \epsilon_k$$

then Eq. (2.1) becomes Eq. (1.1). Because each term of right hand side of Eq. (2.1) involves uncertainty, it is difficult to estimate θ , $f(\phi_k)$ and ε_k simultaneously.

Adaptive estimation problem can be formulated as follows: Given *a priori* knowledge on θ , $f(\cdot)$ and ε_k , how to estimate θ and $f(\cdot)$ according to a series of data $\{\phi_k, z_k; k = 1, 2, \dots, n\}$. Or in other words, given *a priori* knowledge on θ and v_k , how to estimate θ and v_k according to a series of data $\{\phi_k, z_k; k = 1, 2, \dots, n\}$.

Now we list some examples of *a priori* knowledge to show various forms of adaptive estimation problem.

Example 2.1 As to the unknown parameter θ , here are some commonly-seen examples of *a priori* knowledge:

- There is no any *a priori* knowledge on θ except for its dimension. This means that θ can be arbitrary and we do not know its upper bound or lower bound.
- The upper and lower bounds of θ are known, i.e. $\underline{\theta} \leq \theta \leq \bar{\theta}$, where $\underline{\theta}$ and $\bar{\theta}$ are constant vector and the relationship " \leq " means element-wise "less or equal".
- The distance between θ and a nominal θ_0 is bounded by a known constant, i.e. $\|\theta - \theta_0\| \leq r_\theta$, where $r_\theta \geq 0$ is a known constant and θ_0 is the center of set Θ .
- The unknown parameter lies in a known countable or finite set of values, that is to say, $\theta \in \{\theta_1, \theta_2, \theta_3, \dots\}$.

Example 2.2 As to the unknown function $f(\cdot)$, here are some possible examples of *a priori* knowledge:

- $f(x) = 0$ for all x . This case means that there is no unmodeled dynamics.

- Function f is bounded by other known functions, that is to say, $\underline{f}(x) \leq f(x) \leq \overline{f}(x)$ for any x .
- The distance between f and a nominal f_0 is bounded by a known constant, i.e. $\|f - f_0\| \leq r_f$, where $r_f \geq 0$ is a known constant and f_0 can be regarded as the center of a ball F in a metric functional space with norm $\|\cdot\|$.
- The unknown function lies in a known countable or finite set of functions, that is to say, $f \in \{f_1, f_2, f_3, \dots\}$.
- Function f is Lipschitz, i.e. $f(x_1) - f(x_2) \leq L|x_1 - x_2|$ for some constant $L > 0$.
- Function f is monotone (increasing or decreasing) with respect to its arguments.
- Function f is convex (or concave).
- Function f is even (or odd).

Example 2.3 As to the unknown noise term ε_k , here are some possible examples of a priori knowledge:

- Sequence $\varepsilon_k = 0$. This case means that no noise/disturbance exists.
- Sequence ε_k is bounded in a known range, that is to say, $\underline{\varepsilon} \leq \varepsilon_k \leq \overline{\varepsilon}$ for any k . One special case is $\underline{\varepsilon} = -\overline{\varepsilon}$.
- Sequence ε_k is bounded by a diminishing sequence, e.g. $|\varepsilon_k| \leq \frac{1}{k}$ for any k . This case means that the noise disturbance converges to zero with a certain rate. Other typical rate sequences include $\{\frac{1}{k^2}\}$, $\{\delta^k\}$ ($0 < \delta < 1$), and so on.
- Sequence ε_k is bounded by other known sequences, that is to say, $\underline{\varepsilon}_k \leq \varepsilon_k \leq \overline{\varepsilon}_k$ for any k . This case generalizes the above cases.
- Sequence ε_k is in a known finite set of values, that is to say, $\varepsilon_k \in \{e_1, e_2, \dots, e_N\}$. This case may happen in digital systems where all signals can only take values in a finite set.
- Sequence ε_k is oscillatory with specific patterns, e.g. $\varepsilon_k > 0$ if k is even and $\varepsilon_k < 0$ if k is odd.
- Sequence ε_k has some statistical properties, for example, $Ee_k = 0$, $Ee_k^2 = \sigma^2$; for another example, sequence $\{\varepsilon_k\}$ is i.i.d. taken from a probability distribution e.g. $\varepsilon_k \approx U(0,1)$.

Parameter estimation problems (without non-parametric part) involving statistical properties of noise disturbance are studied extensively in statistics, system identification and traditional adaptive control. However, we shall remark that other non-statistic descriptions on a priori knowledge is more useful in practice yet seldom addressed in existing literature. In fact, in practical problems, usually the probability distribution of the noise/disturbance (if any) is not known and many cases cannot be described by any probability distribution since noise/disturbance in practical systems may come from many different types of sources. Without any a priori knowledge in mind, one frequently-used way to handle the noise is to simply assume the noise is Gaussian white noise, which is

reasonable in a certain sense. But in practice, from the point of view of engineering, we can usually conclude the noise/disturbance is bounded in a certain range. This chapter will focus on uncertainties with non-statistical *a priori* knowledge. Without loss of generality, in this section we often regard $v_k = f(\phi_k) + \varepsilon_k$ as a whole part, and correspondingly, *a priori* knowledge on v_k , (e.g. $\underline{v}_k \leq v_k \leq \bar{v}_k$), should be provided for the study.

2.2 An Example Problem

Now we take a simple example to show that it may not be appropriate to apply traditional identification algorithms blindly so as to get the estimate of unknown parameter. Consider the following system

$$z_k = \theta \phi_k + f(\phi_k, k) + \varepsilon_k \quad (2.2)$$

where θ , $f(\cdot)$ and ε_k are unknown parameter, unknown function and unmeasurable noise, respectively. For this model, suppose that we have the following *a priori* knowledge on the system:

- No *a priori* knowledge on θ is known.
- At any step k , the term $f(\phi_k, k)$ is of form $f(\phi_k, k) = \exp(\xi_k \phi_k)$. Here $\{\xi_k\}$ is an unknown sequence satisfying $0 \leq \xi_k \leq 1$.
- Noise ε_k is diminishing with $|\varepsilon_k| \leq \frac{1}{k}$.

And in this example, our problem is how to use the data generated from model (2.2) so as to get a good estimate of true value of parameter θ . In our experiment, the data is generated by the following settings ($k = 1, 2, \dots, 50$):

$$\theta = 5, \phi_k = \frac{k}{10}, f(\phi_k, k) = \exp(|\sin k| \phi_k), \varepsilon_k = \frac{1}{k}(\alpha_k - 0.5)$$

where $\{\alpha_k\}$ are i.i.d. taken from uniform distribution $U(0, 1)$. Here we have $N = 50$ groups of data (ϕ_k, z_k) .

Since model (2.2) involves various uncertainties, we rewrite it into the following form of linear regression

$$z_k = \theta \phi_k + v_k \quad (2.3)$$

by letting

$$v_k = f(\phi_k, k) + \varepsilon_k.$$

From the *a priori* knowledge for model (2.2), we can obtain the following *a priori* knowledge for the term v_k

$$\underline{v}_k \leq v_k \leq \bar{v}_k$$

where

$$\begin{aligned}\bar{v}_k &= \begin{cases} \exp(\phi_k) + \frac{1}{k} & \text{if } \phi_k \geq 0 \\ 1 + \frac{1}{k} & \text{if } \phi_k < 0 \end{cases} \\ \underline{v}_k &= \begin{cases} 1 + \frac{1}{k} & \text{if } \phi_k \geq 0 \\ \exp(\phi_k) + \frac{1}{k} & \text{if } \phi_k < 0 \end{cases}\end{aligned}$$

Since model (2.3) has the form of linear regression, we can use try traditional identification algorithms to estimate θ . Fig. 1 illustrates the parameter estimates for this problem by using standard LS algorithm, which clearly show that LS algorithm cannot give good parameter estimate in this example because the final parameter estimation error $\tilde{\theta}_k = \hat{\theta} - \theta \approx 5.68284$ is very large.

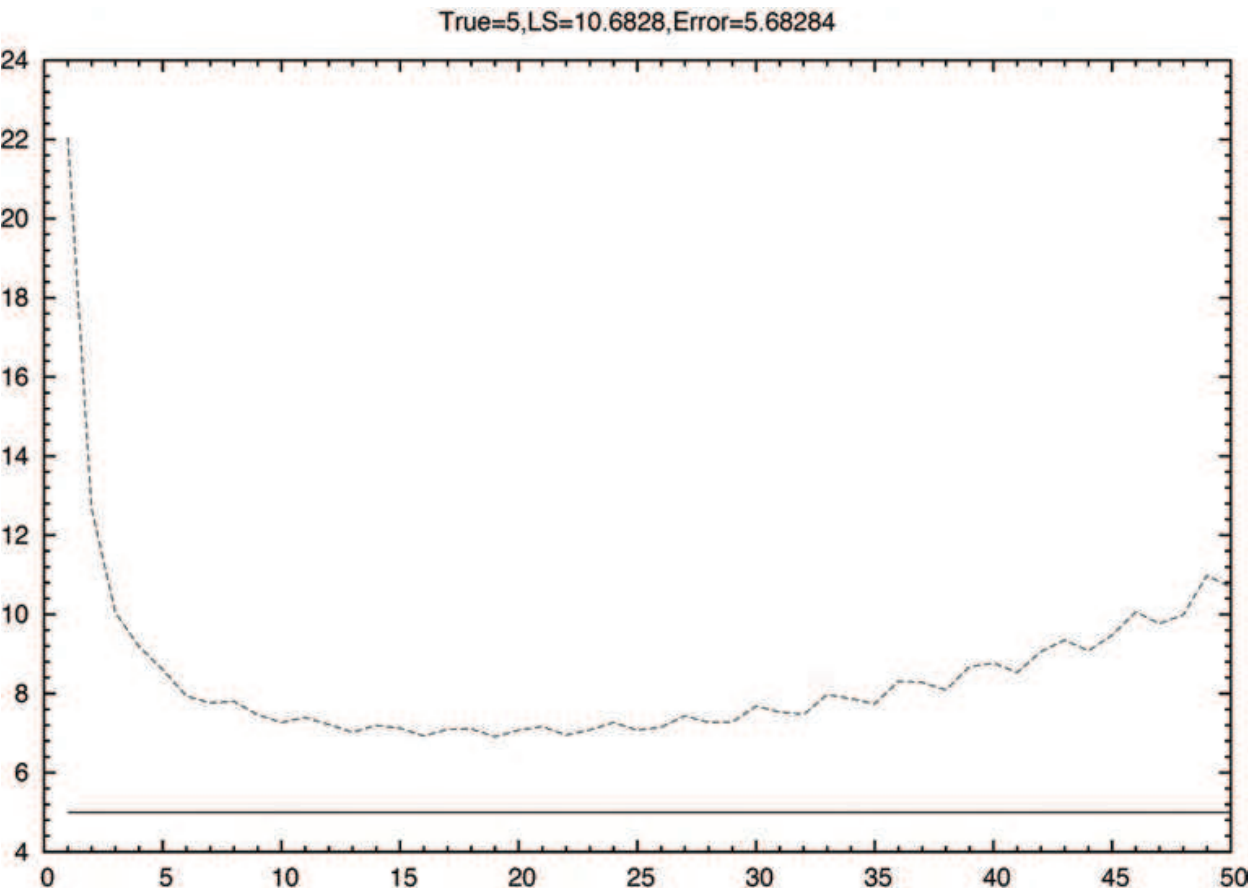


Fig. 1. The dotted line illustrates the parameter estimates obtained by standard least-squares algorithm. The straight line denotes the true parameter.

One may then argue that why LS algorithm fails here is just because the term v_k is in fact biased and we indeed do not utilize the *a priori* knowledge on v_k . Therefore, we may try a modified LS algorithm for this problem: let

$$\begin{aligned}
 c_k &= \frac{1}{2}(\underline{v}_k + \bar{v}_k) \\
 d_k &= \frac{1}{2}(\bar{v}_k - \underline{v}_k) \\
 w_k &= v_k - c_k \\
 y_k &= z_k - c_k
 \end{aligned}$$

then we can conclude that $y_k = \theta^\tau \phi_k + w_k$ and $w_k \in [-d_k, d_k]$, where $[-d_k, d_k]$ is a symmetric interval for every k . Then, intuitively, we can apply LS algorithm to data $\{(\phi_k, z_k), k = 1, 2, \dots, N\}$. The curve of parameter estimates obtained by this modified LS algorithm is plotted in Fig. 2. Since the modified LS algorithm has removed the bias in the *a priori* knowledge, one may expect the modified LS algorithm may give better parameter estimates, which can be verified from Fig. 2 since the final parameter estimation error $\tilde{\theta}_N = \hat{\theta}_N - \theta \approx -1.83314$. In this example, although the modified LS algorithm can work better than the standard LS algorithm, the modified LS algorithm in fact does not help much in solving our problem since the estimation error is still very large comparing with the true value of the unknown parameter.

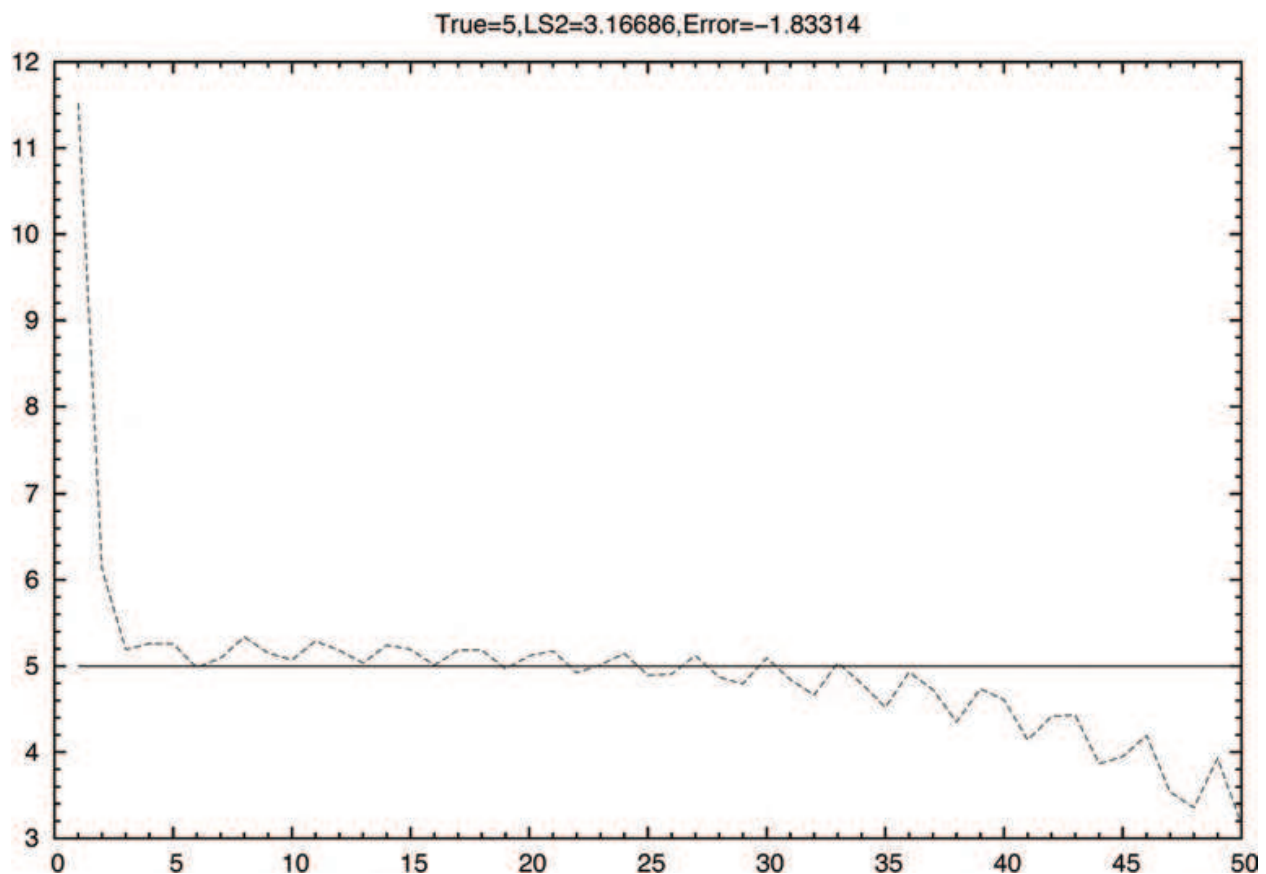


Fig. 2. The dotted line illustrates the parameter estimates obtained by modified least-squares algorithm. The straight line denotes the true parameter.

From this example, we do not aim to conclude that traditional identification algorithms developed in linear regression are not good, however, we want to emphasize the following particular point: *Although traditional identification algorithms (such as LS algorithm) are very powerful and useful in practice, generally it is not wise to apply them blindly when the matching conditions, which guarantee the convergence of those algorithms, cannot be verified or asserted a priori.* This particular point is in fact one main reason why the so-called *minimum-variance self tuning regulator*, developed in the area of adaptive control based on the LS algorithm, attracted several leading scholars to analyze its closed-loop stability throughout past decades from the early stage of adaptive control.

To solve this example and many similar examples with *a priori* knowledge, we will propose new ideas to estimate the parametric uncertainties and the non-parametric uncertainties.

2.3 Information-Concentration Estimator

We have seen that there exist various forms of *a priori* knowledge on system model. With the *a priori* knowledge, how can we estimate the parametric part and the non-parametric part? Now we introduce the so-called information-concentration estimator. The basic idea of this estimator is, the *a priori* knowledge at each time step can be regarded as some constraints of the unknown parameter or function, hence the growing data can provide more and more information (constraints) on the true parameter or function, which enable us to reduce the uncertainties step by step. We explain this general idea by the simple model

$$z_k = \theta^T \phi_k + v_k \quad (2.4)$$

with *a priori* knowledge that $\theta \in \Theta \subseteq R^d, v_k \in V_k$. Then, at k -th step ($k \geq 1$), with the current data k, ϕ_k, z_k we can define the so-called *information set* I_k at step k :

$$I_k \triangleq \{\theta \in \Theta : z_k - \theta^T \phi_k \in V_k\}. \quad (2.5)$$

For convenience, let $I_0 = \Theta$. Then we can define the so-called *concentrated information set* C_k at step k as follows

$$C_k = \bigcap_{j=0}^k I_j \quad (2.6)$$

which can be recursively written as

$$C_k = C_{k-1} \cap I_k \quad (2.7)$$

with initial set $C_0 = \Theta$. Eq. (2.7) with Eq. (2.5) is called *information-concentration estimator* (short for *IC estimator*) throughout this chapter, and any value in the set C_k can be taken as one possible estimate of unknown parameter θ at time step k . The IC estimator differs from existing parameter identification in the sense that the IC estimator is in fact a set-

valued estimator rather than a real-valued estimator. In practical applications, generally C_k is a domain in R^d , and naturally we can take the center point of C_k as $\hat{\theta}_k$.

Remark 2.1 *The definition of information set varies with system model. In general cases, it can be extended to the set of possible instances of θ (and/or f) which do not contradict with the data at step k . We will see an example involving unknown f in next section.*

From the definition of the IC estimator, the following proposition can be obtained without difficulty:

Proposition 2.1 *Information-concentration estimator has the following properties:*

(i) *Monotonicity:* $C_0 \supseteq C_1 \supseteq C_2 \supseteq \dots$

(ii) *Convergence:* Sequence $\{C_k\}$ has a limit set $C_\infty = \bigcap_{k=1}^{\infty} C_k$;

(iii) *If the system model and the a priori knowledge are correct, then C_∞ must be a non-empty set with property $\theta \in C_\infty$ and any element of C_∞ can match the data and the model;*

(iv) *If $C_\infty = \emptyset$, then the data $\{\phi_k, z_k\}$ cannot be generated by the system model used by the IC estimator under the specified a priori knowledge.*

Proposition 2.1 tells us the following particular points of the IC estimator: property (i) implies that the IC estimator will provide more and more exact estimation; property (ii) means that there exists a limitation in the accuracy of estimation; property (iii) means that true parameter lies in every C_k if the system model and a priori knowledge are correct; and property (iv) means that the IC estimator provides also a method to validate the system model and the a priori knowledge. Now we discuss the IC estimator for model (2.4) in more details. In the following discussions, we only consider a typical a priori knowledge on $\underline{v}_k \leq v_k \leq \bar{v}_k$ are two known sequences of vectors (or scalars).

2.3.1 Scalar case: $d = 1$

By Eq. (2.5), we have

$$I_k = \{\theta \in \Theta : \underline{v}_k \leq z_k - \theta \phi_k \leq \bar{v}_k\}$$

Solving the inequality in I_k , we obtain that

$$\theta \phi_k \in [z_k - \underline{v}_k, z_k - \bar{v}_k]$$

and consequently, if $\phi_k \neq 0$, then we have

$$\theta \in [\underline{b}_k, \bar{b}_k]$$

where

$$\begin{aligned}\underline{b}_k &= \frac{\min(\text{sign}(\phi_k)(z_k - \underline{v}_k), \text{sign}(\phi_k)(z_k - \bar{v}_k))}{|\phi_k|}, \\ \bar{b}_k &= \frac{\max(\text{sign}(\phi_k)(z_k - \underline{v}_k), \text{sign}(\phi_k)(z_k - \bar{v}_k))}{|\phi_k|},\end{aligned}$$

Here $\text{sign}(x)$ denotes the sign of x : $\text{sign}(x) = 1, 0, -1$ for positive number, zero, and negative number, respectively. Then, by Eq. (2.7), we can explicitly obtain that

$$C_k = [\underline{\beta}_k, \bar{\beta}_k]$$

where $\underline{\beta}_k$ and $\bar{\beta}_k$ can be recursively obtained by

$$\begin{aligned}\underline{\beta}_k &= \max(\underline{\beta}_{k-1}, \underline{b}_k) \\ \bar{\beta}_k &= \min(\bar{\beta}_{k-1}, \bar{b}_k)\end{aligned}$$

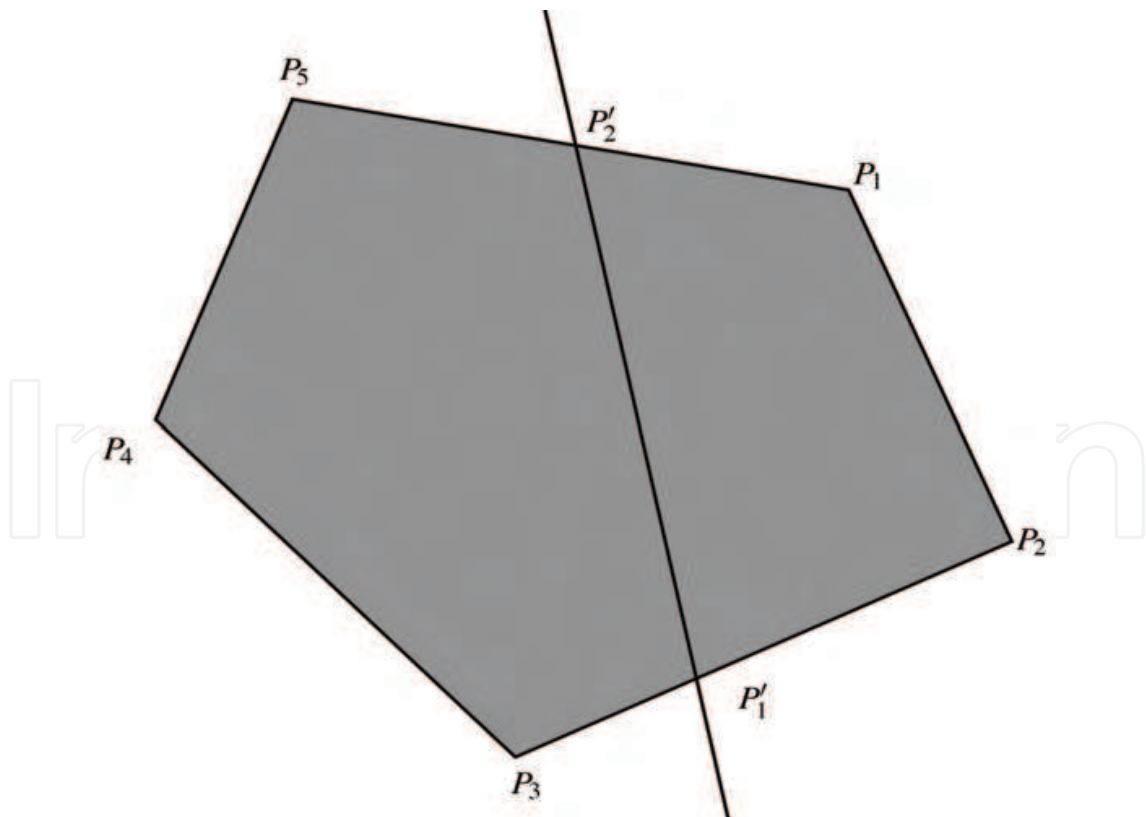


Fig. 3. The straight line may intersect the polygon V and split it into two sub-polygons, one of which will become new polygon V' . The polygon V' can be efficiently calculated from the polygon V .

2.3.2 Vector case: $d > 1$

In case of $d > 1$, since θ and ϕ_k are vectors, we cannot directly obtain explicit solution of inequality

$$\theta^T \phi_k \in [z_k - \underline{v}_k, z_k - \bar{v}_k] \quad (2.8)$$

Notice that Eq. (2.8) can be rewritten into two separate inequalities:

$$\phi_k^T \cdot \theta \leq z_k - \bar{v}_k, \quad (-\phi_k)^T \cdot \theta \leq -(z_k - \underline{v}_k)$$

we need only study linear equalities of the form $\phi^T \theta \leq c$. Generally speaking, the solution to a system of inequalities represents a polyhedral (or polygonal) domain in R^d , hence we need only determine the vertices of the polyhedral (or polygonal) domain. In case of $d = 2$, it is easy to graph linear equalities since every inequality $\phi^T \theta \leq c$ represents a half-plane. In general case, let $v_k = \{v_i, i = 1, 2, \dots, p_k\}$ denote the distinct vertices of the domain C_k and p_k denote the number of vertices of domain C_k , then we discuss how to deduce V_k from V_{k-1} . The domain C_k has two more linear constraints than the domain C_{k-1}

$$\phi_{k,j}^T \cdot \theta \leq c_{k,j}, j = 1, 2$$

with

$$\begin{aligned} \phi_{k,1} &= \phi_k, & c_{k,1} &= z_k - \bar{v}_k \\ \phi_{k,2} &= -\phi_k, & c_{k,2} &= -(z_k - \underline{v}_k) \end{aligned}$$

We need only add these two constraints one by one, that is to say,

$$\begin{aligned} \mathcal{V}'_{k-1} &= \text{AddLinearConstraint}(\mathcal{V}_{k-1}, \phi_{k,1}, c_{k,1}) \\ \mathcal{V}_k &= \text{AddLinearConstraint}(\mathcal{V}'_{k-1}, \phi_{k,2}, c_{k,2}) \end{aligned}$$

where $\text{AddLinearConstraint}(\mathcal{V}, \phi, c)$ is an algorithm whose function is to add linear constraint $\phi^T \theta \leq c$ to the polygon represented by vertex set V and to return the vertex set of the new polygon with added constraint.

Now we discuss how to implement the algorithm **AddLinearConstraint**.

2D Case: In case of $d = 2$, $\phi^T \theta \leq c$ represents a straight line which splits the plane into two half-planes (see Fig. 3). In this case, we can use an efficient algorithm **AddLinearConstraint2D** which is listed in Algorithm 1. Its basic idea is to simply test each vertex of V to see whether to keep original vertex or generate new vertex. The time

complexity of Algorithm 1 is $O(s)$, where s is the number of vertices of domain V . Note that it is possible that $V' = \emptyset$ if the straight line $L: \phi^T \theta \leq c$ does not intersect with the polygon V and any vertex P_i of polygon V does not satisfy $\phi^T P_i > c$. And the vertex number of polygon V' can in fact vary within the range from 0 to s according to the geometric relationship between the straight line L and the polygon V .

Algorithm 1
AddLinearConstraint2D
(\mathcal{V}, ϕ, c): Add linear constraint $\phi^T \theta \leq c$ ($\phi \in \mathbb{R}^2$) to a polygon \mathcal{V}

Input: \mathcal{V} : represented by clock-wise arranged vertices P_1, P_2, \dots, P_s
Output: \mathcal{V}' : clock-wise arranged vertices

1: Let $P_0 \leftarrow P_s$
2: Let Q be an empty queue
3: **for** $i = 1$ to s **do**
4: Let $\delta_i \leftarrow \phi^T P_i - c$
5: **end for**
6: Let $\delta_0 \leftarrow \delta_s$
7: **for** $i = 1$ to s **do**
8: **if** $(\delta_i = 0)$ or $(\delta_i < 0 \text{ and } \delta_{i-1} \leq 0)$ **then**
9: Append vertex P_i to queue Q
10: **else if** $(\delta_i > 0 \text{ and } \delta_{i-1} < 0)$ or $(\delta_i < 0 \text{ and } \delta_{i-1} > 0)$ **then**
11: Let $w \leftarrow \frac{\delta_{i-1}}{\delta_{i-1} - \delta_i}$
12: Generate new vertex $P' \leftarrow w \cdot P_i + (1 - w) \cdot P_{i-1}$
13: Append vertex P' to queue Q
14: **end if**
15: **end for**
16: Return vertices in queue Q as the vertices of new domain \mathcal{V}'

High-dimensional Case: In case of $d > 2$, $\phi^T \theta \leq c$ represents a hyperplane which splits the whole space into two half-hyperplanes. Unlike in case of $d = 2$, the vertices in this case generally cannot be arranged in a certain natural order (such as clock-wise order). In this case, we can use an algorithm AddLinearConstraintND which is listed in Algorithm 2. The idea of this algorithm is to classify the vertices of V first according to their relationship with the hyperplane determined by hyperplane $\phi^T \theta \leq c$.

Algorithm 2 AddLinearConstraintND(V, ϕ, c): Add linear constraint $\phi^T \theta \leq c$ ($\phi \in \mathbb{R}^d$) to a polyhedron V

2.3.3 Implementation issues

In the IC estimator, the key problem is to calculate the information set I_k or the concentrated information set C_k at every step. From the discussions above, we can see that it is easy to solve this basic problem in case of $d = 1$. However, in case of $d > 1$, generally the vertex

number of domain C_k may grow as $k \rightarrow \infty$. Therefore, it may be impractical to implement the IC estimator in case of $d > 1$ since it may require growing memory as $k \rightarrow \infty$. To overcome this problem, noticing the fact that the domain C_k will shrink gradually as $k \rightarrow \infty$ in order to get a feasible IC estimate of the unknown parameter vector, generally we need not use too many vertices to represent the exact concentrated information set C_k . That is to say, in practical implementation of IC estimator in high-dimensional case, we can use a domain \hat{C}_k with only a small number (say up to M) of vertices to approximate the exact concentrated information set C_k . With such an idea of *approximate IC estimator*, the issue of computational complexity will not hinder the applications of IC estimator.

We consider two typical cases of *approximate IC estimator*. One typical case is that $\hat{C}_k \supseteq C_k$ for any k , and the other case is that $\hat{C}_k \supseteq C_k$ for any k . Let $\hat{C}_\infty = \bigcap_{k=1}^{\infty} \hat{C}_k$, then in the former case (called *loose IC estimator*, see Fig. 4), we must have

$$C_\infty = \bigcap_{k=1}^{\infty} C_k \subseteq \hat{C}_\infty$$

which means that we will never mistakenly exclude the true parameter from the concentrated approximate information sets; while in the latter case (called *tight IC estimator*, see Fig. 5), we must have

$$C_\infty = \bigcap_{k=1}^{\infty} C_k \supseteq \hat{C}_\infty$$

which means that the true parameter may be outside of \hat{C}_∞ however any value in \hat{C}_∞ can be served as good estimate of true parameter.

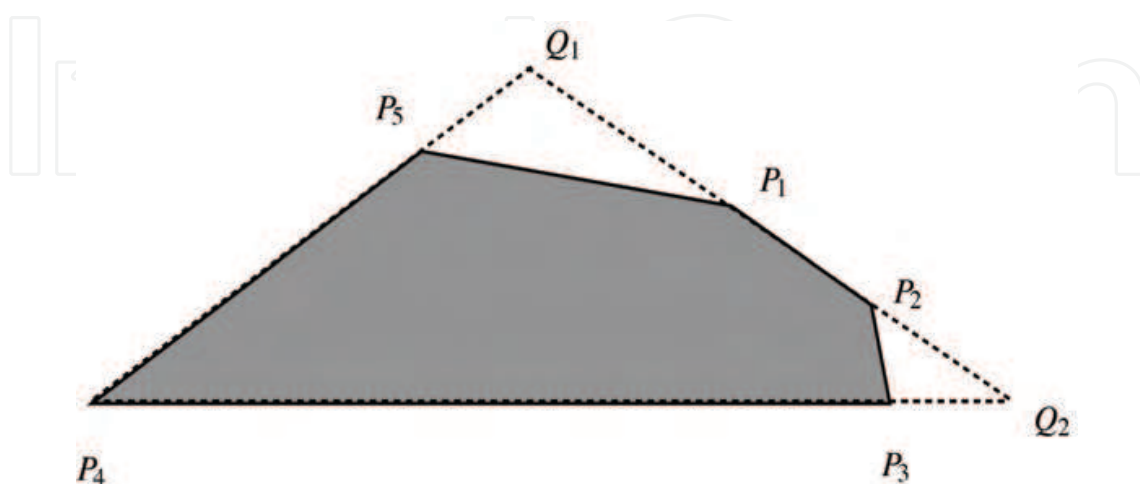


Fig. 4. Idea of loose IC estimator: The polygon $P_1P_2P_3P_4P_5$ can be approximated by a triangle $Q_1P_4Q_2$. Here $M = 3$.

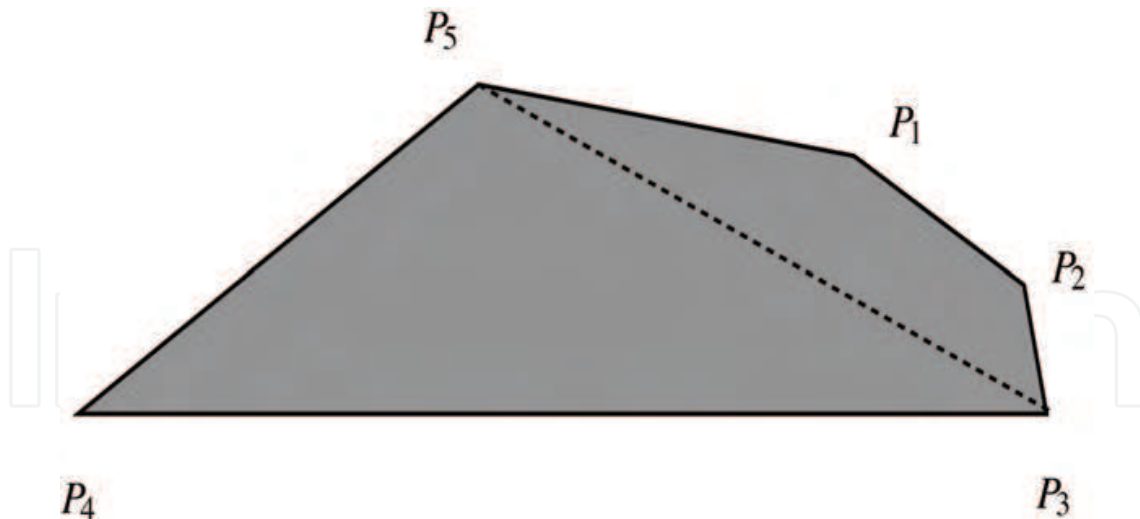


Fig. 5. Idea of tight IC estimator: The polygon $P_1P_2P_3P_4P_5$ can be approximated by a triangle $P_3P_4P_5$. Here $M = 3$.

Now we discuss implementation details of *tight IC estimator* and *loose IC estimator*. Without loss of generality, we only explain the ideas in case of $d = 2$. Similar ideas can be applied in cases of $d > 2$ without difficulty.

Tight IC estimator: To implement a tight IC estimator, one simple approach is to modify Algorithm 1 so as it just keeps up to M vertices in the queue Q . To get good approximation, in the loop of Algorithm 1, it is suggested to abandon the generated vertex P' (in Line 12 of Algorithm 1) which is very close to existing vertex P_j (let $j = i$ if $\delta_i < 0$ and $\delta_{i-1} > 0$ or $j = i - 1$ if $\delta_i > 0$ and $\delta_{i-1} < 0$). The closeness between P' and existing vertex P_j can be measured by checking the corresponding weight w .

Loose IC estimator: To implement a loose IC estimator, one simple approach is to modify Algorithm 1 so as it can generate M vertices which surround all vertices in the queue Q . To this end, in the loop of Algorithm 1, if the generated vertex P' (in Line 12 of Algorithm 1) is very close to existing vertex P_j (let $j = i$ if $\delta_i < 0$ and $\delta_{i-1} > 0$ or $j = i - 1$ if $\delta_i > 0$ and $\delta_{i-1} < 0$), we can simply append vertex P_j instead of P' to queue Q . In this way, we can avoid increasing the vertex number by generating new vertices. The closeness between P' and existing vertex P_j can be measured by checking the corresponding weight w .

Besides the ideas of tight or loose IC estimator, to reduce the complexity of IC estimator, we can also use other flexible approaches. For example, to avoid growth in the vertex number of V_k as $k \rightarrow \infty$, we can approximate V_k by using a simple outline rectangle (see Fig. 6) every certain steps. For a polygon V_k with vertices P_1, P_2, \dots, P_s , we can easily obtain its outline rectangle by algorithm FindPolygonBounds listed in Algorithm 3. Here for convenience, the operators max and min for vectors are defined element-wisely, i.e.

$$\begin{aligned} \max(A, B) &\triangleq [\max(A_1, B_1), \max(A_2, B_2), \dots, \max(A_d, B_d)]^\tau \\ \min(A, B) &\triangleq [\min(A_1, B_1), \min(A_2, B_2), \dots, \min(A_d, B_d)]^\tau \end{aligned}$$

where $A = [A_1, A_2, \dots, A_d]^\tau, B = [B_1, B_2, \dots, B_d]^\tau$ are two vectors in R_n .

Algorithm 3 FindPolygonBounds(\mathcal{V}): Find the outline rectangle of polygon \mathcal{V}

Input: \mathcal{V} : represented by vertices P_1, P_2, \dots, P_s
Output: \mathcal{V}' : the outline rectangle represented by upper and lower bounds of each component of vertices

- 1: Let $\bar{B} \leftarrow P_1$
 - 2: Let $\underline{B} \leftarrow P_1$
 - 3: **for** $i = 2$ to s **do**
 - 4: Let $\bar{B} \leftarrow \max(\bar{B}, P_i)$
 - 5: Let $\underline{B} \leftarrow \min(\underline{B}, P_i)$
 - 6: **end for**
 - 7: Return vertices of rectangle represented by coordinate bounds \bar{B} and \underline{B} as the vertices of new domain \mathcal{V}'
-

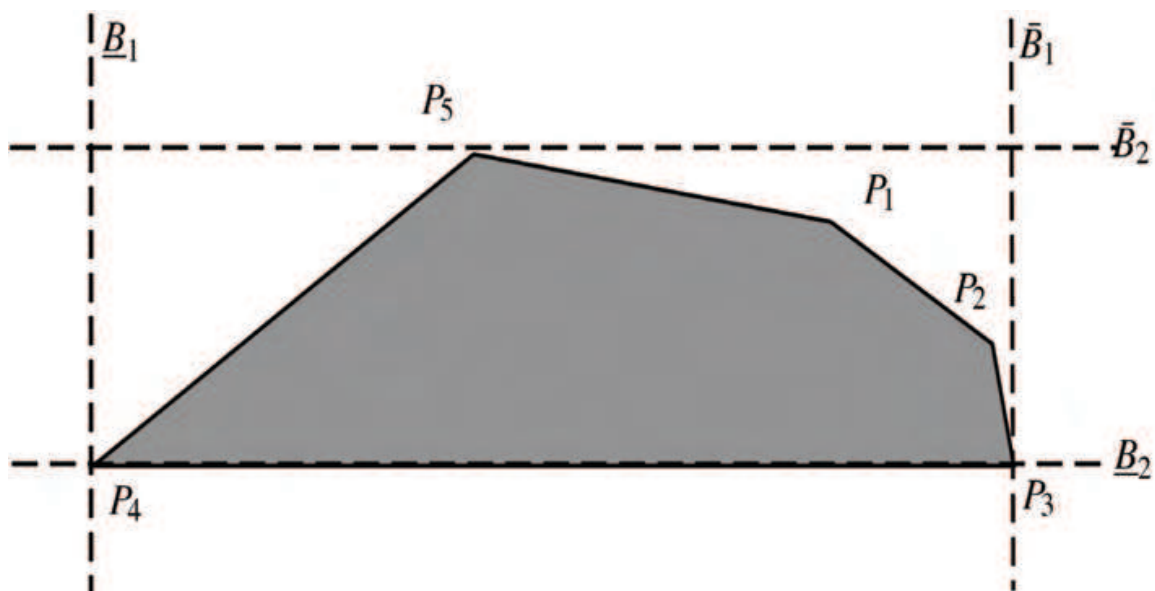


Fig. 6. Idea of outline rectangle: The polygon $P_1P_2P_3P_4P_5$ can be approximated by an outline rectangle. In this case, $\underline{B}_1, \bar{B}_1$ denote the lower bound and upper bound in the x -axis (1st component of each vertex), and $\underline{B}_2, \bar{B}_2$ denote the lower bound and upper bound in the y -axis (2nd component of each vertex)

2.4 IC Estimator vs. LS Estimator

2.4.1 Illustration of IC Estimator

Now we go back to the example problem discussed before. For this example, ϕ_k and z_k are scalars, hence we need only apply the IC estimator introduced in Section 2.3.1. Since IC estimator yields concentrated information set C_k at every step, we can take any value in

C_k as parameter estimate of true parameter. In this example, C_k is an interval at every step. For comparison with other parameter estimation methods, we simply take $\hat{\theta}_k = \frac{1}{2}(\bar{b}_k + \underline{b}_k)$, i.e. the center of interval C_k , as the parameter estimate at step k .

In Fig. 7, we plot three curves \bar{b}_k , \underline{b}_k and $\hat{\theta}_k$. From this figure, we can see that, for this particular example, with the help of *a priori* knowledge, the upper estimates \bar{b}_k and lower estimates \underline{b}_k given by the IC estimator converge to true parameter $\theta = 5$ quickly, and consequently $\hat{\theta}_k$ also converges to true parameter.

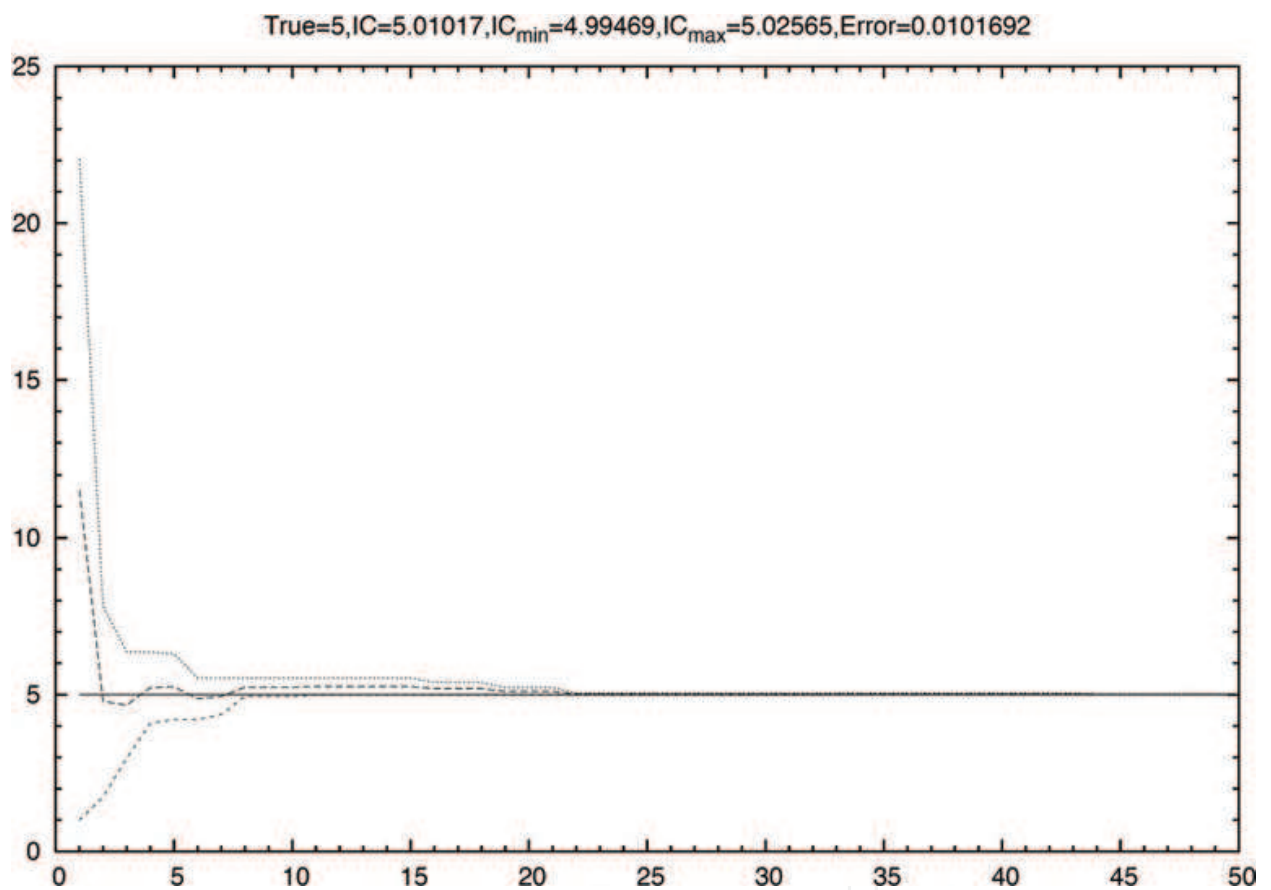


Fig. 7. This figure illustrates the parameter estimates obtained by the proposed information-concentration estimator. The upper curve and lower curve represent the upper bounds \bar{b}_k and lower bounds \underline{b}_k for the parameter estimates. We use the center curve $\hat{\theta}_k = \frac{1}{2}(\bar{b}_k + \underline{b}_k)$ to yield the parameter estimates.

We should also remark that the parameter estimates given by the IC estimator are not necessarily convergent as in this example. Whether the IC parameter estimates converge

largely depend on the accuracy of *a priori* knowledge and the richness of the practical data. Note that the IC estimator generally does not require classical richness concepts (like *persistent excitation*) which are useful in the analysis of traditional recursive identification algorithms.

2.4.2 Advantages of IC Estimator

We have seen practical effects of IC estimator for the simple example given above. Why can it perform better than the LS estimator? Roughly speaking, comparing with traditional identification algorithm like LS algorithm, the proposed IC estimator has the following advantages:

1. It can make full use of *a priori* information and posterior information. And in the ideal case, no information is wasted in the iteration process of the IC estimator. This property is not seen in traditional identification algorithms since only partial information and certain stochastic *a priori* knowledge can be utilized in those algorithms.
2. It does not give single parameter estimate at every step; instead, it gives a (finite or infinite) set of parameter estimates at every step. This property is also unique since traditional identification algorithms always give parameter estimates directly.
3. It can gradually find out all (or most) possible values of true parameters; and this property can even help people to check the consistence between the practical data and the system model with *a priori* knowledge. This property distinguishes traditional identification algorithms in sense that traditional identification algorithms generally have no mechanism to validate the correctness of the system model.
4. The *a priori* knowledge can vary from case to case, not necessarily described in the language of probability theory or statistics. This property enables the IC estimator to handle various kinds of non-statistic *a priori* knowledge, which cannot be dealt with by traditional identification algorithms.
5. It has great flexibilities in its implementation, and its design is largely determined by the characteristics of *a priori* knowledge. *The IC estimator has only one basic principle – information concentration!* Any practical implementation approach using such a principle can be regarded as an IC estimator. We have discussed some implementation details for a certain type of IC estimator in last subsection, which have shown by examples how to design the IC estimator according the known *a priori* knowledge and how to reduce computational complexity in practical implementation.
6. Its accuracy will never degrade as time goes by. Generally speaking, the more steps calculated, the more data involved, and the more accurate the estimates are. Generally speaking, traditional identification algorithms can only have similar property (called *strong consistency*) under certain matching conditions.
7. The IC estimator can not only provide reasonably good parameter estimates but also tell people how accurate these estimates are. In our previous example, when we use

$\hat{\theta}_k = \frac{1}{2}(\bar{b}_k + \underline{b}_k)$ as the parameter estimate, we know also that the absolute parameter

estimation error $|\tilde{\theta}| = |\hat{\theta} - \theta|$ will not exceed $\frac{1}{2}(\bar{b}_k + \underline{b}_k)$. In some sense, such a property may be conceptually similar to the so-called *confidence level* in statistics.

2.4.3 Disadvantages of IC Estimator

Although the IC estimator has many advantages over traditional identification algorithms, it may have the following disadvantages:

1. The proposed IC estimator is relatively difficult to incorporate stochastic *a priori* knowledge on noise term, especially unbounded random noise. In fact, in such cases without non-parametric uncertainties, traditional identification algorithms like LS algorithm may be more suitable and efficient to estimate the unknown parameter.
2. The efficiency of IC estimator largely depends on its implementation via the characteristics of the *a priori* knowledge. Generally speaking, the IC estimator may involve a little more computation operations than recursive identification algorithms like LS algorithm. We shall remark also that this point is not always true since the numerical operations involved in the IC estimator are relatively simple (see algorithms listed before), while many traditional identification algorithms may involve costly numerical operations like matrix product, matrix inversion, etc.
3. Although the IC estimator has simple and elegant properties such as monotonicity and convergence, *due to its nature of set-valued estimator*, no explicit and recursive expressions can be given directly for the IC parameter estimates, which may bring mathematical difficulties in the applications of the IC estimator. However, generally speaking, we also know that closed-loop analysis for adaptive control using traditional identification algorithms is not easy, too.

Summarizing the above, we can conclude that the IC estimator provides a new approach or principle to estimate parametric and even non-parametric uncertainties, and we have shown that it is possible to design efficient IC estimator according to characteristics of *a priori* knowledge.

3. Semi-parametric Adaptive Control: Example 1

In this section, we will give a first example of semi-parametric adaptive control, whose design is essentially based on the IC estimator introduced in last section.

3.1 Problem Formulation

Consider the following system

$$y_{t+1} = \theta \phi_t + f(y_t) + u_t + w_{t+1} \quad (3.1)$$

where y_t , u_t and w_t are the output, input and noise, respectively; $f(\cdot) \in F(L)$ is an unknown function (the set $F(L)$ will be defined later) and θ is an unknown parameter. To make further study, the following assumptions are used throughout this section:

Assumption 3.1 The unknown function $f : R \rightarrow R$ belongs to the following uncertainty set

$$\mathcal{F}(L) = \{f : |f(x) - f(y)| \leq L|x - y| + c\} \quad (3.2)$$

where c is an arbitrary non-negative constant.

Assumption 3.2 The noise sequence $\{w_t\}$ is bounded, i.e.

$$|w_t| \leq w$$

where w is an arbitrary positive constant.

Assumption 3.3 The tracking signal $\{y_t^*\}$ is bounded, i.e.

$$|y_t^*| \leq S, \forall t \geq 0 \quad (3.3)$$

where S is a positive constant.

Assumption 3.4 In the parametric part $\theta\phi_t$, we have no any a priori information of the unknown parameter θ , but $\phi_t = g(y_t)$ is measurable and satisfies

$$M' \leq \left| \frac{g(x_1) - g(x_2)}{x_1^b - x_2^b} \right| \leq M \quad (3.4)$$

for any $x_1 \neq x_2$, where $M' \leq M$ are two positive constants and $b \geq 1$ is a constant.

Remark 3.1 Assumption 3.4 implies that function $g(\cdot)$ has linear growth rate when $b = 1$. Especially when $g(x) = x$, we can take $M = M' = 1$. Condition (3.4) need only hold for sufficiently large x_1 and x_2 , however we require it holds for all $x_1 \neq x_2$ to simplify the proof. We shall also remark that Sokolov [Sok03] has ever studied the adaptive estimation and control problem for a special case of model (3.1), where ϕ_t is simply taken as αy_t .

Remark 3.2 Assumption 3.4 excludes the case where $g(\cdot)$ is a bounded function, which can be handled easily by previous research. In fact, in that case $w'_{t+1} = \theta\phi_t + w_{t+1}$ must be bounded, hence by the result of [XG00], system (3.1) is stabilizable if and only if $L < \frac{3}{2} + \sqrt{2}$.

3.2 Adaptive Controller Design

In the sequel, we shall construct a unified adaptive controller for both cases of $b = 1$ and $b > 1$. For convenience, we introduce some notations which are used in later parts. Let $I = [a, b]$ be

an interval, then $m(I) = \frac{1}{2}(a + b)$ ($a + b$) denotes the center point of interval I , and

$r(I) = \frac{1}{2}|b - a|$ denotes the radius of interval I . And correspondingly, we let

$I(x, \delta) = [x - \delta, x + \delta]$ denote a closed interval centered at $x \in R$ with radius $\delta \geq 0$.

Estimate of Parametric Part: At time t , we can use the following information: $y_0, y_1, \dots, y_t, u_0, u_1, \dots, u_{t-1}$ and $\phi_1, \phi_2, \dots, \phi_t$. Define

$$z_j \triangleq y_{j+1} - u_j \quad (3.5)$$

and

$$I_t \triangleq \bigcap_{(i,j) \in J_t} I\left(\frac{z_j - z_i}{\phi_j - \phi_i}, \frac{L|y_j - y_i|}{|\phi_j - \phi_i|} + \frac{2w+c}{|\phi_j - \phi_i|}\right) \quad (3.6)$$

where

$$J_t \triangleq \{(i,j) \in \mathcal{N} : i < j < t, \phi_i \neq \phi_j\} \quad (3.7)$$

then, we can take

$$\hat{\theta}_t = m(I_t), \quad \delta_t = r(I_t) \quad (3.8)$$

as the estimate of parameter θ at time t and corresponding estimate error bound, respectively. With $\hat{\theta}_t$ and δ_t defined above, $\bar{\theta}_t = \hat{\theta}_t + \delta_t$ and $\underline{\theta}_t = \hat{\theta}_t - \delta_t$ are the estimates of the upper and lower bounds of the unknown parameter θ , respectively.

According to Eq. (3.6), obviously we can see that $\{\bar{\theta}_t\}$ is a non-increasing sequence and $\{\underline{\theta}_t\}$ is non-decreasing.

Remark 3.3 Note that Eq. (3.6) makes use of a priori information on nonlinear function $f(\cdot)$. This estimator is another example of the IC estimator which demonstrates how to design the IC estimator according to the Lipschitz property of function $f(\cdot)$. With similar ideas, the IC estimator can be designed based on other forms of a priori information of function $f(\cdot)$.

Estimate of Non-parametric Part: Since the non-parametric part $f(y_t)$ may be unbounded and the parametric part is also unknown, generally speaking it is not easy to estimate the non-parametric part directly. To resolve this problem, we choose to estimate

$$g_t \triangleq \theta \phi_t + f(y_t)$$

as a whole part rather than to estimate $f(y_t)$ directly. In this way, consequently, we can obtain the estimate of $f(y_t)$ by removing the estimate of parametric part from the estimate of g_t .

Define

$$i_t = \arg \min_{i < t} |y_t - y_i| \quad (3.9)$$

then, we get

$$\begin{aligned}
 g_t &= g_t - z_{i_t} + z_{i_t} \\
 &= [\theta \phi_t + f(y_t)] - [\theta \phi_{i_t} + f(y_{i_t}) + w_{i_t+1}] + z_{i_t} \\
 &= [\theta(\phi_t - \phi_{i_t}) + z_{i_t}] + [f(y_t) - f(y_{i_t}) - w_{i_t+1}]
 \end{aligned} \tag{3.10}$$

Thus, intuitively, we can take

$$\hat{g}_t \triangleq \hat{\theta}_t(\phi_t - \phi_{i_t}) + z_{i_t} = \hat{\theta}_t(\phi_t - \phi_{i_t}) + (y_{i_t+1} - u_{i_t}) \tag{3.11}$$

as the estimate of g_t at time t .

Design of Control u_t : Let

$$\begin{aligned}
 \bar{b}_t &\triangleq \max_{i \leq t} y_i = \max(\bar{b}_{t-1}, y_t) \\
 \underline{b}_t &= \min_{i \leq t} y_i = \min(\underline{b}_{t-1}, y_t).
 \end{aligned} \tag{3.12}$$

Under Assumptions 3.1-3.4, we can design the following control law

$$u_t = \begin{cases} -\hat{g}_t + y_{t+1}^* & \text{if } |y_t - y_{i_t}| \leq D \\ -\hat{g}_t + \frac{1}{2}(\bar{b}_t + \underline{b}_t) & \text{if } |y_t - y_{i_t}| > D \end{cases} \tag{3.13}$$

where D is an appropriately large constant, which will be addressed in the proof later.

Remark 3.4 The controller designed above is different from most traditional adaptive controllers in its special form, information utilization and computational complexity. To reduce its computational complexity, the interval I_t given by Eq. (3.6) can be calculated recursively based on the idea in Eq. (3.12).

3.3 Stability of Closed-loop System

In this section, we shall investigate the closed-loop stability of system (3.1) using the adaptive controller given above. We only discuss the case that the parametric part is of linear growth rate, i.e. $b = 1$. For the case where the parametric part is of nonlinear growth rate, i.e. $b > 1$, though simulations show that the constructed adaptive controller can stabilize the system under some conditions, we have not rigorously established corresponding theoretical results; further investigation is needed in the future to yield deeper understanding.

3.3.1 Main Results

The adaptive controller constructed in last section has the following property:

Theorem 3.1 When $b = 1$, $\frac{ML}{M'} < \frac{3}{2} + \sqrt{2}$, the controller defined by Eqs. (3.5) – (3.13) can guarantee that the output $\{y_t\}$ of the closed-loop system is bounded. More precisely, we have

$$\limsup_{t \rightarrow \infty} |y_t - y_t^*| \leq 2w + c. \quad (3.14)$$

Based on Theorem 3.1, we can classify the capability and limitations of feedback mechanism for the system (3.1) in case of $b = 1$ as follows:

Corollary 3.1 For the system (3.1) with both parametric and non-parametric uncertainties, the following results can be obtained in case of $b = 1$:

- (i) If $b = 1$, $\frac{ML}{M'} < \frac{3}{2} + \sqrt{2}$, then there exists a feedback control law guaranteeing that the closed-loop system is stabilized.
- (ii) When $\phi_t = y_t$ (i.e. $g(x) = x$), the presence of uncertain parametric part $\theta\phi_t$ does not reduce the critical value $\frac{3}{2} + \sqrt{2}$ of the feedback mechanism which is determined by the uncertainties of non-parametric part.

Proof of Corollary 3.1: (i) This result follows from Theorem 3.1 directly. (ii) When $g(x) = x$, we can take $M = M' = 1$. In this case, the sufficiency can be immediately obtained via Theorem 3.1; on the other hand, the necessity can be obtained by the “impossibility” part of Theorem

1 in [XG00]. In fact, if $L \geq \frac{3}{2} + \sqrt{2}$, for any given control law $\{u_t\}$, we need only take the parameter $\theta = 0$, then by [XG00, Theorem 2.1], there exists a function f such that system (3.1) cannot be stabilized by the given control law.

Remark 3.5 As we have mentioned in the introduction part, system (1.6), a special case of system (3.1), has been studied in [XG00]. Comparing system (3.1) and system (1.6), we can see that system (3.1) has also parametric uncertainty besides nonparametric uncertainty and noise disturbance. Hence intuitively speaking, it will be more difficult for the feedback mechanism to deal with uncertainties in system (3.1) than those in system (1.6). Noting that $M' \leq M$, we know this fact has been partially verified by Theorem 3.1. And Corollary 3.1 (ii) indicates that in the special case of $\phi_t = y_t$, since the structure of parametric part is completely determined, the uncertainty in non-parametric part becomes the main difficulty in designing controller, and the parametric uncertainty has no influence on the capability of the feedback mechanism, that is to say, the feedback mechanism can still deal with the non-parametric uncertainty characterized by the set $F(L)$ with $L < \frac{3}{2} + \sqrt{2}$.

Remark 3.6 Theorem 3.1 is also consistent with classic results on adaptive control for linear systems. In fact, when $L = 0$, the non-parametric part $f(y_t)$ vanishes, consequently system (3.1) becomes a linear-in-parameter system

$$y_{t+1} = \theta\phi_t + u_t + w_{t+1} \quad (3.15)$$

where θ is the unknown parameter, and $\phi_t = g(y_t)$ can have arbitrary linear growth rate because by Theorem 3.1, we can see that no restrictions are imposed on the values of M and M' when $L = 0$. Based on the knowledge from existing adaptive control theory [CG91], system (3.15) can be always stabilized by algorithms such as minimum-variance adaptive controller no matter how large the θ is. Thus the special case of Theorem 3.1 reveals again the well-known result in a new way, where the adaptive controller is defined by Eq. (3.13) together with Eqs. (3.5) – (3.12).

Corollary 3.2 If $b = 1$, $\frac{ML}{M'} < \frac{3}{2} + \sqrt{2}$, $c = w = 0$, then the adaptive controller defined by Eqs. (3.5) – (3.13) can asymptotically stabilize the corresponding noise-free system, i.e.

$$\lim_{t \rightarrow \infty} |y_t - y_t^*| = 0. \quad (3.16)$$

3.3.2 Preliminary Lemmas

To prove Theorem 3.1, we need the following Lemmas:

Lemma 3.1 Assume $\{x_n\}$ is a bounded sequence of real numbers, then we must have

$$\lim_{n \rightarrow \infty} \min_{i < n} |x_n - x_i| = 0. \quad (3.17)$$

Proof: It is a direct conclusion of [XG00, Lemma 3.4]. It can be proved by argument of contradiction.

Lemma 3.2 Assume that $L \in (0, \frac{3}{2} + \sqrt{2})$, $d \geq 0$, $n_0 \geq 0$. If non-negative sequence $\{h_n, n \geq 0\}$ satisfies

$$h_{n+1} \leq \left(L \max_{i \leq n} h_i - \frac{1}{2} \sum_{i=0}^n h_i + d \right)^+, \forall n \geq n_0 \quad (3.18)$$

where $x^+ = \max(x, 0)$, $\forall x \in R$, then we must have

$$\lim_{n \rightarrow \infty} \sum_{i=0}^n h_i < \infty. \quad (3.19)$$

Proof: See [XG00, Lemma 3.3].

3.3.3 Proof of Theorem 3.1

Proof of Theorem 3.1: We divide the proof into four steps. In Step 1, we deduce the basic relation between y_{t+1} and $\tilde{\theta}_t$, and then a key inequality describing the upper bound of $|y_t - y_{i_t}|$ is established in Step 2. Consequently, in Step 3, we prove that $|y_t - y_{i_t}| \rightarrow 0$

as $t \rightarrow \infty$ if y_t is not bounded, and hence the boundedness of output sequence $\{y_t\}$ can be guaranteed. Finally, in the last step, the bound of tracking error can be further estimated based on the stability result obtained in Step 3.

Step 1: Let

$$\begin{aligned}\tilde{\theta}_t &\triangleq \theta - \hat{\theta}_t \\ y_{t+1}^\# &\triangleq \theta\phi_t + f(y_t) + w_{t+1} - \hat{g}_t\end{aligned}\quad (3.20)$$

then, by definition of u_t and Eq. (3.13), obviously we get

$$y_{t+1} = \begin{cases} y_{t+1}^\# + y_{t+1}^* & \text{if } |y_t - y_{i_t}| \leq D \\ y_{t+1}^\# + \frac{1}{2}(\bar{b}_t + \underline{b}_t) & \text{if } |y_t - y_{i_t}| > D \end{cases}\quad (3.21)$$

Now we discuss $y_{t+1}^\#$. By Eq. (3.11) and Eq. (3.1), we get

$$\begin{aligned}y_{t+1}^\# &= \theta\phi_t + f(y_t) + w_{t+1} - \hat{g}_t \\ &= \theta\phi_t + f(y_t) + w_{t+1} - \hat{\theta}_t(\phi_t - \phi_{i_t}) \\ &\quad - (\theta\phi_{i_t} + f(y_{i_t}) + w_{i_t+1}) \\ &= (\theta - \hat{\theta}_t)(\phi_t - \phi_{i_t}) \\ &\quad + [f(y_t) - f(y_{i_t})] + (w_{t+1} - w_{i_t+1}) \\ &= \tilde{\theta}_t(\phi_t - \phi_{i_t}) + [f(y_t) - f(y_{i_t})] + (w_{t+1} - w_{i_t+1})\end{aligned}\quad (3.22)$$

In case of $\phi_t = \phi_{i_t}$, i.e. $y_t = y_{i_t}$, obviously we get

$$|y_{t+1}^\#| = |w_{t+1} - w_{i_t+1}| \leq 2w;\quad (3.23)$$

otherwise, we get

$$y_{t+1}^\# = (\tilde{\theta}_t + D_{t,i_t})(\phi_t - \phi_{i_t}) + (w_{t+1} - w_{i_t+1})\quad (3.24)$$

where

$$D_{i,j} \triangleq \frac{f(y_i) - f(y_j)}{\phi_i - \phi_j}.$$

Obviously $D_{ij} = D_{ji}$. In the latter case, i.e. when $\phi_t \neq \phi_{i_t}$, for any $(i, j) \in J_t$, noting that

$$\begin{aligned}
 z_j - z_i &= (y_{j+1} - u_j) - (y_{i+1} - u_i) \\
 &= \theta(\phi_j - \phi_i) + [f(y_j) - f(y_i)] \\
 &\quad + [w_{j+1} - w_{i+1}]
 \end{aligned} \tag{3.25}$$

we obtain that

$$\theta - \frac{z_j - z_i}{\phi_j - \phi_i} = -D_{i,j} - \frac{w_{j+1} - w_{i+1}}{\phi_j - \phi_i}. \tag{3.26}$$

Therefore

$$\begin{aligned}
 \tilde{\theta}_t + D_{t,i_t} &= \left(\theta - \frac{z_j - z_i}{\phi_j - \phi_i}\right) - \left(\hat{\theta}_t - \frac{z_j - z_i}{\phi_j - \phi_i}\right) + D_{t,i_t} \\
 &= D_{t,i_t} - D_{i,j} - \frac{w_{j+1} - w_{i+1}}{\phi_j - \phi_i} + \Delta_{i,j}(t)
 \end{aligned} \tag{3.27}$$

where

$$\Delta_{i,j}(t) = \frac{z_j - z_i}{\phi_j - \phi_i} - \hat{\theta}_t. \tag{3.28}$$

Step 2: Since $\frac{ML}{M'} < \frac{3}{2} + \sqrt{2}$, there exists a constant $\varepsilon > 0$ such that $\frac{ML}{M'} + \varepsilon < \frac{3}{2} + \sqrt{2}$.

Let

$$B_t \triangleq [\underline{b}_t, \bar{b}_t], \Delta B_t = B_t - B_{t-1} \tag{3.29}$$

and consequently

$$|B_t| = \bar{b}_t - \underline{b}_t, |\Delta B_t| = |B_t| - |B_{t-1}| \tag{3.30}$$

By the definitions of \underline{b}_t , \bar{b}_t and B_t , we obtain that

$$|B_{t+1}| = \begin{cases} |B_t| & \text{if } y_{t+1} \in B_t \\ \frac{1}{2}|B_t| + |y_{t+1} - \frac{1}{2}(\underline{b}_t + \bar{b}_t)| & \text{if } y_{t+1} \notin B_t \end{cases} \tag{3.31}$$

By the definition of i_t , obviously we get

$$|y_t - y_{i_t}| \begin{cases} = |\Delta B_t| & \text{if } y_t \notin B_{t-1} \\ \leq |\Delta B_i| & \text{if } y_t \in \Delta B_i \subset B_{t-1} \end{cases} \quad (3.32)$$

Step 3: Based on Assumption 3.4, for any fixed $\varepsilon > 0$, we can take constants D and D' such that $|\phi_i - \phi_j| > D' > \frac{4M(2w+c)}{\varepsilon}$ when $|y_t - y_{i_t}| > D$. Now we are ready to show that for any $s > 0$, there always exists $t > s$ such that $|y_t - y_{i_t}| > D$.

In fact, suppose that it is not true, then there must exist $s > 0$ such that $|y_t - y_{i_t}| > D$ for any $t > s$, correspondingly $|\phi_t - \phi_{i_t}| > D'$. Consequently, by the definition of D , for sufficiently large t and $j < t$, we obtain that

$$\left| \frac{w_{j+1} - w_{i_j+1}}{\phi_j - \phi_{i_j}} \right| \leq \left| \frac{2w}{D'} \right| < \frac{1}{4M}\varepsilon \quad (3.33)$$

together with the definition of $\hat{\theta}_t$, we know that for any $s < i < j < t$,

$$|\Delta_{i,j}(t)| = \left| \frac{z_j - z_i}{\phi_j - \phi_i} - \hat{\theta}_t \right| \leq \frac{L}{M'} + \frac{2w+c}{|\phi_j - \phi_i|}. \quad (3.34)$$

hence for $s < j < t, i = i_j$, we get

$$|\Delta_{j,i_j}(t)| = |\Delta_{i,j}(t)| \leq \frac{L}{M'} + \frac{2w+c}{D'} \leq \frac{L}{M'} + \frac{1}{4M}\varepsilon. \quad (3.35)$$

Now we consider $D_{t,i_t} - D_{j,i_j}$.

Let $d_n = D_{n,i_n}$, then, by the definition of $D_{i,j}$, noting that $|y_j - y_i| \geq |y_j - y_{i_j}| > D$ for any $j > s$, we obtain that

$$|D_{i,j}| = \left| \frac{f(y_i) - f(y_j)}{y_i - y_j} \right| \cdot \left| \frac{y_i - y_j}{\phi_i - \phi_j} \right| \leq \frac{L}{M'}, \quad (3.36)$$

so we can conclude that $\{d_n, n > s\}$ is bounded. Then, by Lemma 3.1, we conclude that

$$\lim_{t \rightarrow \infty} \min_{s < j < t} |d_t - d_j| = 0. \quad (3.37)$$

Consequently there exists $s' > s$ such that for any $t > s'$, we can always find a corresponding $j=j(t)$ satisfying

$$|D_{t,i_t} - D_{j,i_j}| = |d_t - d_j| < \frac{1}{4M}\epsilon. \quad (3.38)$$

Summarizing the above, for any $t > s'$, by taking $j = j(t)$, we get

$$\begin{aligned} |\tilde{\theta}_t + D_{t,i_t}| &= |D_{t,i_t} - D_{j,i_j} - \frac{w_{j+1} - w_{i+1}}{\phi_j - \phi_{i_j}} + \Delta_{i_j,j}(t)| \\ &\leq |D_{t,i_t} - D_{j,i_j}| + |\frac{w_{j+1} - w_{i+1}}{\phi_j - \phi_{i_j}}| + |\Delta_{i_j,j}(t)| \\ &\leq \frac{1}{4M}\epsilon + \frac{1}{4M}\epsilon + (\frac{L}{M'} + \frac{1}{4M}\epsilon) \\ &= \frac{L}{M'} + \frac{3}{4M}\epsilon \\ &\triangleq L_\epsilon. \end{aligned} \quad (3.39)$$

Therefore

$$\begin{aligned} |y_{t+1}^\#| &= |(\tilde{\theta}_t + D_{t,i_t}) \frac{\phi_t - \phi_{i_t}}{y_t - y_{i_t}} (y_t - y_{i_t}) + (w_{t+1} - w_{i_t+1})| \\ &\leq L_\epsilon M |y_t - y_{i_t}| + 2w. \end{aligned} \quad (3.40)$$

Since $|y_t - y_{i_t}| > D$, we know that

$$y_{t+1} = y_{t+1}^\# + \frac{1}{2}(\bar{b}_t + \underline{b}_t). \quad (3.41)$$

From Eq. (3.39) together with the result in Step 2, we obtain that

$$\begin{aligned} |B_t| \leq |B_{t+1}| &\leq \max\{|B_t|, \frac{1}{2}|B_t| + |y_{t+1} - \frac{1}{2}(\bar{b}_t + \underline{b}_t)|\} \\ &= \max\{|B_t|, \frac{1}{2}|B_t| + |y_{t+1}^\#|\} \end{aligned} \quad (3.42)$$

Thus noting (3.40), we obtain the following key inequality:

$$|\Delta B_t| \leq (L_\epsilon M |y_t - y_{i_t}| + 2w - \frac{1}{2}|B_t|)^+ \quad (3.43)$$

where

$$L_\epsilon M = (\frac{L}{M'} + \frac{3}{4M}\epsilon)M = \frac{ML}{M'} + \frac{3}{4}\epsilon < \frac{3}{2} + \sqrt{2}. \quad (3.44)$$

Considering the arbitrariness of $t > s'$, together with Lemma 3.2, we obtain that

$$\sum_{j>s'} |\Delta B_j| < \infty, \quad (3.45)$$

and consequently $\{|B_t|\}$ must be bounded. By applying Lemma 3.1 again, we conclude that

$$|y_t - y_{i_t}| \leq \min_{i<t} |y_t - y_i| \rightarrow 0 \quad (3.46)$$

which contradicts the former assumption!

Step 4: According to the results in Step 3, for any $s > 0$, there always exists $t > s$ such that

$|y_t - y_{i_t}| \leq D$. Then, we can easily obtain that $\{|\tilde{\theta}_t|\}$ is bounded, say $|\tilde{\theta}_t| \leq L'$.

Considering that

$$y_{t+1}^\# = \tilde{\theta}_t(\phi_t - \phi_{i_t}) + [f(y_t) - f(y_{i_t})] + (w_{t+1} - w_{i_t+1}) \quad (3.47)$$

we can conclude that

$$\begin{aligned} |y_{t+1}| &\leq |y_{t+1}^\# + y_{t+1}^*| \\ &\leq L'|\phi_t - \phi_{i_t}| + (L|y_t - y_{i_t}| + c) + 2w \\ &\leq Y \end{aligned} \quad (3.48)$$

where $Y = L'MD + LD + c + 2w + S$.

The proof below is similar to that in [XG00]. Let

$$t_0 = \inf_{t>0} \{t : |y_t| \leq Y\}, \quad t_n = \inf_{t>t_{n-1}} \{t : |y_t| \leq Y\}. \quad (3.49)$$

Because of the result obtained above, we conclude that for any $n \geq 1$, t_n is well-defined and $t_n < \infty$. Let $v_n = y_{t_n}$, then obviously $\{v_n\}$ is bounded. Then, by applying Lemma 3.1, we get

$$\min_{i<n} |v_n - v_i| \rightarrow 0 \quad (3.50)$$

as $n \rightarrow \infty$. Thus for any $\varepsilon > 0$, there exists an integer n_0 such that for any $n > n_0$,

$$\min_{i<n} |v_n - v_i| < \varepsilon. \quad (3.51)$$

So

$$|y_{t_n} - y_{i_{t_n}}| = \min_{i<t_n} |y_{t_n} - y_i| \leq \min_{i<n} |y_{t_n} - y_{t_i}| < \varepsilon. \quad (3.52)$$

By taking ε sufficiently small, we obtain that

$$|y_{t_n+1}| \leq L'M\varepsilon + L\varepsilon + c + 2w + S \leq Y \quad (3.53)$$

for any $n > n_0$.

Thus based on definition of t_n , we conclude that $t_{n+1} = t_n + 1$! Therefore for any $t \geq t_{n_0}$,

$$|y_t| \leq Y \quad (3.54)$$

which means that the sequence $\{y_t\}$ is bounded.

Finally, by applying Lemma 3.1 again, for sufficiently large t , $|y_t - y_i^*| \leq \varepsilon$ consequently

$$|y_{t+1} - y_{t+1}^*| = |y_{t+1}^\#| \leq L'M\varepsilon + L\varepsilon + c + 2w. \quad (3.55)$$

Because of arbitrariness of ε , Theorem 3.1 is true.

3.4 Simulation Study

In this section, two simulation examples will be given to illustrate the effects of the adaptive controller designed above. In both simulations, the tracking signal is taken as

$y_t^* = 10 \sin \frac{t}{10}$ and the noise sequence is i.i.d. randomly taken from uniform distribution

$U(0, 1)$. The simulation results for two examples are depicted in Figure 8 and Figure 9, respectively. In each figure, the output sequence y_t and the reference sequence y_t^* are

plotted in the top-left subfigure; the tracking error sequence $e_t = y_t - y_t^*$ is plotted in the

bottom-left subfigure; the control sequence u_t is plotted in the top-right subfigure; and the

parameter θ together with its upper and lower estimated bounds is plotted in the bottom-right subfigure.

Simulation Example 1: This example is for case of $b = 1$, and the unknown plant is

$$y_{t+1} = f(y_t) + \theta g(y_t) + w_{t+1}, \quad f(\cdot) \in \mathcal{F}(L) \quad (3.56)$$

with $L = 2.9 < \frac{3}{2} + \sqrt{2}$, $g(x) = x$ (i.e. $b = 1, M = M' = 1$)

and

$$f(x) = 1.4x \sin \log(|x| + 1). \quad (3.57)$$

For this example, we can verify that

$$|f'(x)| = 1.4|\sin \log(|x| + 1) \pm \frac{x \cos \log(|x| + 1)}{|x| + 1}| \leq 2.8 < L \tag{3.58}$$

consequently $|f(x) - f(y)| < L|x - y|$, i.e. $f(\cdot) \in F(L)$..

Simulation Example 2: This example is for case of $b > 1$, and the unknown plant is

$$y_{t+1} = f(y_t) + \theta g(y_t) + w_{t+1}, \quad f(\cdot) \in \mathcal{F}(L) \tag{3.59}$$

with $L = 2.9$, $g(x) = x^2$ (i.e. $b = 2$, $M = M' = 1$), and

$$f(x) = 2x + \sin x^2. \tag{3.60}$$

For this example, we can verify that $|f(x) - f(y)| < L|x - y| + 2$, i.e. $f(\cdot) \in F(L)$.

From the simulation results, we can see that in both examples, the adaptive controller can track the reference signal successfully. The simulation study verified our theoretical result and indicate that under some conditions, the adaptive control law constructed in this paper can deal with both parametric and non-parametric uncertainties, even in some cases when the parametric part is of nonlinear growth rate. In case of $b = 1$, the stabilizability criteria have been completely characterized by a simple algebraic condition; however, in case of $b > 1$, it is very difficult to give complete theoretical characterization. Note that usually more accurate estimate of parameter can be obtained in case of $b > 1$ than in case of $b = 1$, however, worse transient performance may be encountered.

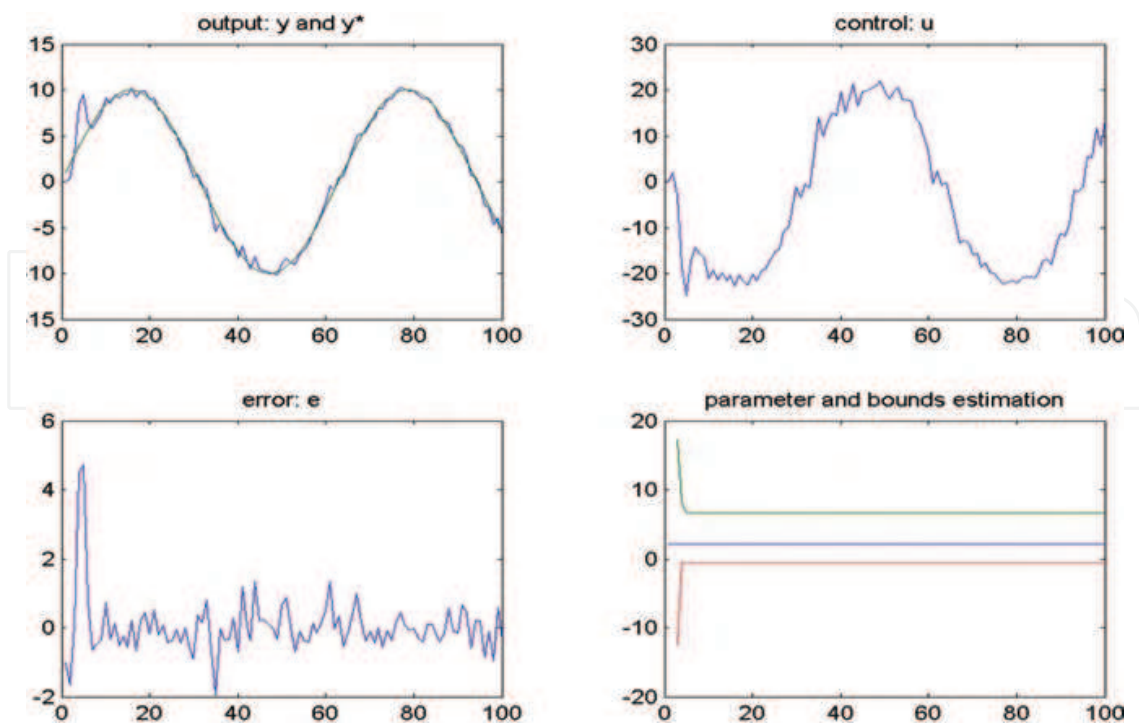


Fig. 8. Simulation example 1: ($g(x) = x$, $b = 1, M = M' = 1$)

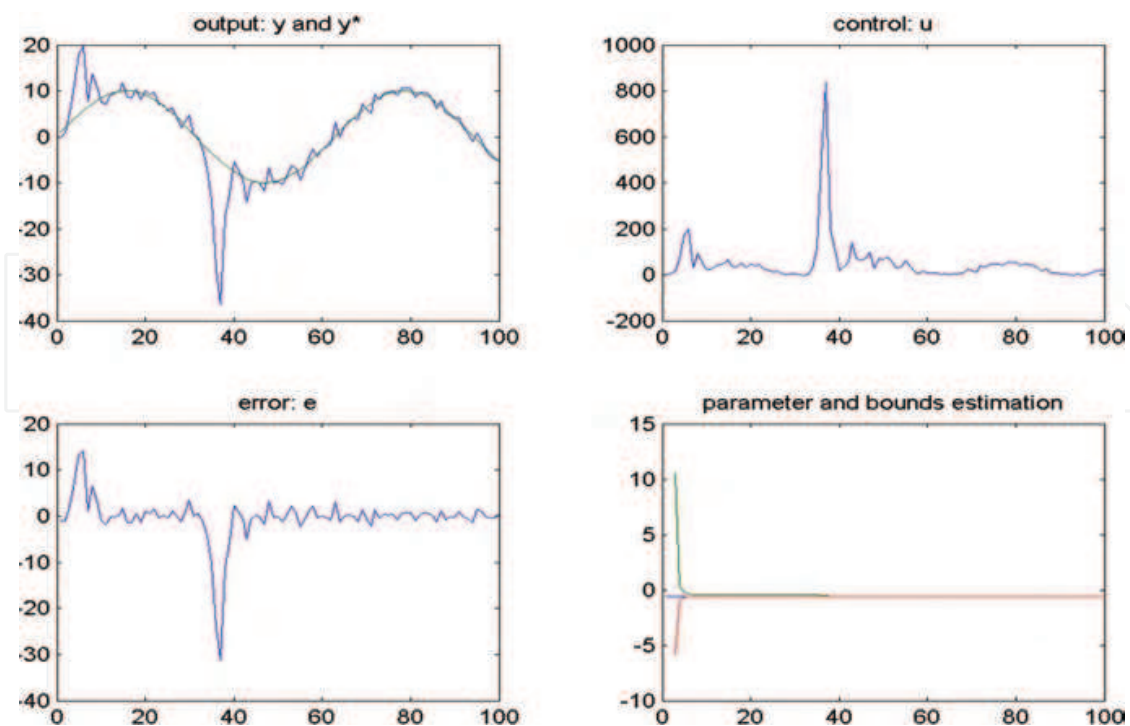


Fig. 9. Simulation example 2: ($g(x) = x_2$, $b = 2$, $M = M' = 1$)

4. Semi-parametric Adaptive Control: Example 2

In this section, we shall give another example of adaptive estimation and control for a semi-parametric model. Although the system considered in this section is similar to the model considered in last section, there are several particular points in this example:

- The controller gain in this model is also unknown with *a priori* knowledge on its sign and its lower bound.
- The system is noise-free, and correspondingly the asymptotic tracking is rigorously established in this example.
- The algorithm in this example has a form of gradient algorithm, however, it partially makes use of *a priori* knowledge on the non-parametric part.
- Due to the limitation of this algorithm and technical difficulties, unlike the algorithm in last section, we can only establish stability of the closed-loop system under condition $0 < L < 0.5$ for the parametric part, which is much stronger than the condition $0 \leq L < \frac{3}{2} + \sqrt{2}$.

This example is given here only for the purpose of demonstrating that there exist other possible ways to make use of *a priori* knowledge on the parametric uncertainties and non-parametric uncertainties. By comparing the examples in this section and last section, the readers may get a deeper understanding to adaptive estimation and control problems for semi-parametric models.

4.1 Problem Formulation

We consider the following system model

$$y_{k+1} = \theta \Phi(y(k)) + f(y_k) + bu_k \quad (4.1)$$

where $y_k \in R^1$ and $u_k \in R^1$ are output and control signals, respectively. Here $\theta \in R^1$ is the unknown parameter, $b \in R^1$ is the unknown controller gain, $\Phi(\cdot)$ is a known function, and $f(\cdot)$ is the unknown function. We have the following *a priori* knowledge on the real system:

Assumption 4.1 The nonparametric uncertain function $f(\cdot)$ is Lipschitz, i.e., $\|f(x_1) - f(x_2)\| \leq L \|x_1 - x_2\|$, $\forall x_1, x_2 \in R$, where $L < 0.5$. The known function $\Phi(\cdot)$ is also a Lipschitz function with Lipschitz constant L .

Assumption 4.2 The sign of unknown controller gain b is known. Without loss of generality, we assume that $b \geq \underline{b} > 0$ where \underline{b} is a known constant.

Assumption 4.3 The reference signal y_k^* is a known bounded deterministic signal.

The control objective is to design the control law u_k such that the output signal y_k asymptotically tracks a bounded reference trajectory y_k^* and all the closed-loop signals are guaranteed to be bounded.

4.2 Adaptive Control Design

To design the adaptive controller, the following notations will be used throughout this section:

$$\begin{aligned} l_k &= \arg \min_{l \leq k-1} |y_k - y_l| \\ e_k &= y_k - y_k^* \end{aligned} \quad (4.2)$$

Obviously, at time step k , with the history information $\{y_j, j \leq k\}$ and the *a priori* knowledge, the index l_k and the tracking error e_k are available. Later we will see important roles of l_k and e_k in the controller design.

Estimation of parametric part: The estimates of the parameter θ and the controller gain b at time step k are denoted by $\hat{\theta}_k$ and \hat{b}_k , respectively. We design the following adaptive update law to update the parameter estimates recursively:

$$\begin{aligned}
\hat{\theta}_k &= \hat{\theta}_{k-1} + \gamma \frac{a_k e_k [\Phi(y_{k-1}) - \Phi(y_{l_{k-1}})]}{D_{k-1}} \\
\hat{b}_k &= \begin{cases} \hat{b}'_k & \text{if } \hat{b}'_k > \underline{b} \\ \underline{b} & \text{otherwise} \end{cases} \\
\hat{b}'_k &= \hat{b}_{k-1} + \text{Proj}_{\underline{b}} \left\{ \frac{\gamma a_k e_k}{D_{k-1}} [u_{k-1} - u_{l_{k-1}}] \right\} \\
D_{k-1} &= \frac{1}{2} \|\Phi(y_{k-1}) - \Phi(y_{l_{k-1}})\|^2 + \frac{1}{2} [u_{k-1} - u_{l_{k-1}}]^2
\end{aligned}$$

where $0 < \gamma < 1$ and the coefficient a_k is defined by a time-varying deadzone:

$$a_k = \begin{cases} 1 - \frac{\lambda |y_{k-1} - y_{l_{k-1}}|}{|e_k|} & \text{if } |e_k| > \lambda |y_{k-1} - y_{l_{k-1}}| \\ 0 & \text{otherwise} \end{cases} \quad (4.4)$$

Estimation of non-parametric part: As in last section, we do not estimate the non-parametric part directly. Instead, we try to estimate the parametric part and non-parametric part as a whole part

$$y_k^\# \triangleq \theta \Phi(y_k) + f(y_k) \quad (4.5)$$

Noticing of the system model (4.1), we know that

$$y_{k+1} = y_k^\# + b u_k \quad (4.6)$$

consequently, from Eqs. (4.5) and (4.6), it is easy to derive

$$\begin{aligned}
y_k^\# &= y_k^\# - y_{l_k}^\# + y_{l_k}^\# \\
&= \theta [\Phi(y_k) - \Phi(y_{l_k})] + f(y_k) - f(y_{l_k}) \\
&\quad + y_{l_k+1} - b u_{l_k}
\end{aligned} \quad (4.7)$$

Since function $f(\cdot)$ is unknown and parameters θ and b are unknown, we simply estimate $y_k^\#$ by the following equation

$$\hat{y}_k^\# = \hat{\theta}_k [\Phi(y_k) - \Phi(y_{l_k})] + y_{l_k+1} - \hat{b}_k u_{l_k} \quad (4.8)$$

where $\hat{\theta}_k$ and \hat{b}_k are regarded as true parameters, and the unknown term $f(y_k) - f(y_{l_k})$ in Eq. (4.7) is simply dropped off.

Adaptive control law: By Eq. (4.6), according to the certainty equivalence principle, we can design the following adaptive control law

$$u_k = -\frac{1}{\hat{b}_k}(\hat{y}_k^\# - y_{k+1}^*) \quad (4.9)$$

Where \hat{b}_k and $\hat{y}_k^\#$ are given by Eqs. (4.3) and (4.7). The closed-loop stability will be given later.

4.3 Asymptotic Tracking Performance

4.3.1 Main Results

Theorem 4.1 *In the closed-loop system (4.1) with control law (4.9) and parameters adaptation law (4.3), under Assumptions 4.1 – 4.3, all the signals in the closed-loop system are bounded and further the tracking error e_k will converge to zero.*

4.3.2 Preliminaries

Definition 4.1 Let x_k and y_k ($k \geq 0$) be two discrete-time scalar or vector signals.

- We denote $x_k = O[y_k]$, if there exist positive constants m_1 and m_2 such that $\|x_k\| \leq m_1 \max_{j \leq k} \|y_j\| + m_2$, $\forall k > k_0$ and k_0 is the initial time step.
- We denote $x_k = o[y_k]$, if there exists a sequence α_k satisfying $\lim_{k \rightarrow \infty} \alpha_k \rightarrow 0$ such that $\|x_k\| \leq m_1 \max_{j \leq k} \|y_j\| + m_2$, $\forall k > k_0$.
- We denote $x_k \sim y_k$ if they satisfy $x_k = O[y_k]$ and $y_k = O[x_k]$.

Lemma 4.1 Consider the following parameter update law

$$\hat{\theta}_k = \begin{cases} \hat{\theta}'_k & \text{if } \hat{\theta}'_k > \mu \\ \mu & \text{otherwise} \end{cases} \quad (4.10)$$

$$\hat{\theta}'_k = \hat{\theta}_{k-1} + \text{Proj}_{\hat{\theta}}(\eta_k) \quad (4.11)$$

$$\text{Proj}_{\hat{\theta}}(\eta_k) = \begin{cases} -\eta_k, & \text{if } \hat{\theta}_{k-1} = \mu \text{ and } \eta_k < 0 \\ \eta_k, & \text{otherwise} \end{cases} \quad (4.12)$$

where $\theta \in R$ is an unknown scalar, $\hat{\theta}_k$ is its estimate at time step k , μ is the lower bound of θ , and $\eta_k \in R$ is any sequence. Then, $\hat{\theta}_k \geq \mu$ is guaranteed and the following properties hold:

$$\tilde{\theta}_k'^2 \geq \tilde{\theta}_k^2, \text{Proj}_{\hat{\theta}}^2(\eta_k) = \eta_k^2, \tilde{\theta}_k^T (\text{Proj}_{\hat{\theta}}(\eta_k) - \eta_k) \leq 0$$

where $\tilde{\theta}_k' = \hat{\theta}_k' - \theta$ and $\tilde{\theta}_k = \hat{\theta}_k - \theta$.

Proof: According to Eqs. (4.10) and (4.11), it is obvious that $\hat{\theta}_k \geq \mu$ always hold. From Eq. (4.12), we see that $|\text{Proj}_{\hat{\theta}}(\eta_k)| = |\eta_k|$, hence $\text{Proj}_{\hat{\theta}}^2(\eta_k) = \eta_k^2$. Further, we have

$$\begin{aligned} & \tilde{\theta}_{k-1} (\text{Proj}_{\hat{\theta}}(\eta_k) - \eta_k) \\ = & \begin{cases} (\hat{\theta}_{k-1} - \theta)(-\eta_k - \eta_k) \leq 0, & \text{if } \hat{\theta}_{k-1} = \mu \text{ means } (\hat{\theta}_{k-1} - \theta) \leq 0 \text{ and } \eta_k < 0 \\ (\hat{\theta}_{k-1} - \theta)(\eta_k - \eta_k) = 0 & \text{otherwise} \end{cases} \end{aligned}$$

From (4.10), we see that $\hat{\theta}_k' = \hat{\theta}_k$ if $\hat{\theta}_k' > \mu$ such that $\tilde{\theta}_k'^2 = \tilde{\theta}_k^2$ when $\hat{\theta}_k' > \mu$. Noticing that when $\hat{\theta}_k' \leq \mu$, we have $\mu \leq \theta$, so that

$$\begin{aligned} \tilde{\theta}_k'^2 &= (\hat{\theta}_k' - \theta)^2 = [(\hat{\theta}_k' - \mu) + (\mu - \theta)]^2 \\ &= (\hat{\theta}_k' - \mu)^2 + (\mu - \theta)^2 + 2(\hat{\theta}_k' - \mu)(\mu - \theta) \\ &\geq (\mu - \theta)^2 = (\hat{\theta}_k - \theta)^2 = \tilde{\theta}_k^2 \end{aligned} \quad (4.13)$$

Therefore, we always have $\tilde{\theta}_k'^2 \geq \tilde{\theta}_k^2$. This completes the proof.

Lemma 4.2 Given a bounded sequence $X_k \in R^m$. Define

$$l_k = \arg \min_{l \leq k-1} \|X_k - X_l\|$$

Then, we have

$$\lim_{k \rightarrow \infty} \|X_k - X_{l_k}\| = 0$$

Proof: This lemma is an extension of Lemma 3.1. Its proof can be found in [Ma06].

Lemma 4.3 (Key Technical Lemma) Let $\{S_t\}$ be a sequence of real numbers and $\{\sigma_t\}$ be a sequence of vectors such that

$$\|\sigma_t\| \leq c_1 + c_2 \max_{0 \leq k \leq t} |s_k|$$

Assume that

$$z_t = \frac{s_t^2}{\alpha_1 + \alpha_2 \sigma_t^T \sigma_t} \rightarrow 0 \text{ as } t \rightarrow \infty$$

where $\alpha_1 > 0, \alpha_2 > 0$. Then $\|\sigma_t\|$ is bounded.

Proof: This lemma can be found in [AW89, GS84].

4.3.3 Proof of Theorem 4.1

Define parameter estimate errors $\tilde{b}_k = \hat{b}_k - b$ and $\tilde{\theta}_k = \hat{\theta}_k - \theta$. From Eqs. (4.7) and (4.8), we have

$$\begin{aligned} \tilde{y}_k^\# &\triangleq \hat{y}_k^\# - y_k^\# \\ &= \tilde{\theta}_k [\Phi(y_k) - \Phi(y_{l_k})] - f(y_k) + f(y_{l_k}) - \tilde{b}_k u_{l_k} \end{aligned} \quad (4.14)$$

Then, we can derive the following error dynamics:

$$\begin{aligned} e_{k+1} &\triangleq y_{k+1} - y_{k+1}^* \\ &= y_k^\# + \hat{b}_k u_k - \tilde{b}_k u_k - y_{k+1}^* \\ &= -\tilde{y}_k^\# - \tilde{b}_k u_k \\ &= -\tilde{\theta}_k [\Phi(y_k) - \Phi(y_{l_k})] + f(y_k) - f(y_{l_k}) - \tilde{b}_k [u_k - u_{l_k}] \end{aligned} \quad (4.15)$$

According to Assumption 4.1, we have

$$|f(y_k) - f(y_{l_k})| \leq \lambda |y_k - y_{l_k}| \quad (4.16)$$

where λ can be any constant satisfying $L \leq \lambda < 0.5$.

From the error dynamics Eq. (4.15), we have

$$\begin{aligned} e_k &= -\tilde{\theta}_{k-1} [\Phi(y_{k-1}) - \Phi(y_{l_{k-1}})] + f(y_{k-1}) - f(y_{l_{k-1}}) \\ &\quad - \tilde{b}_{k-1} [u_{k-1} - u_{l_{k-1}}] \end{aligned} \quad (4.17)$$

Choose Lyapunov function candidate as

$$V_k = \tilde{\theta}_k^2 + \tilde{b}_k^2 \quad (4.18)$$

From the adaptation laws (4.3), we obtain that

$$\tilde{\theta}_k^2 - \tilde{\theta}_{k-1}^2 = \gamma^2 \frac{a_k^2 e_k^2}{D_{k-1}^2} [\Phi(y_{k-1}) - \Phi(y_{l_{k-1}})]^2 + 2\gamma \frac{a_k e_k}{D_{k-1}} [\Phi(y_{k-1}) - \Phi(y_{l_{k-1}})] \quad (4.19)$$

$$\tilde{g}_k^2 - \tilde{g}_{k-1}^2 \leq \tilde{g}_k'^2 - \tilde{g}_{k-1}^2 \quad (4.20)$$

$$= \gamma^2 \frac{a_k^2 e_k^2}{D_{k-1}^2} [u_{k-1} - u_{l_{k-1}}]^2 + 2\gamma \frac{a_k e_k}{D_{k-1}} [u_{k-1} - u_{l_{k-1}}] \quad (4.21)$$

Together with the error dynamics Eq. (4.17), we can derive that the difference of V_k

$$\Delta V_k = V_k - V_{k-1} \leq -\frac{2\gamma(1-\gamma)a_k^2 e_k^2}{D_{k-1}} \quad (4.22)$$

Noting that $0 < \gamma < 1$ and taking summation on both hand sides of Eq. (4.22), we obtain

$$\sum_{k=0}^{\infty} 2\gamma(1-\gamma) \frac{a_k^2 e_k^2}{D_{k-1}} \leq V(0) - V(\infty)$$

Which implies

$$\lim_{k \rightarrow \infty} \frac{a_k^2 e_k^2}{D_{k-1}} = 0 \quad (4.23)$$

and the boundedness of $\hat{\theta}_k$ and \hat{b}_k . Considering $y_k \sim e_k$, we have

$$|y_{k-1}| \leq \max_{j \leq k} \{|e_j|\} + C_2, \quad k > k_0$$

where C_2 are some constants. From the definition of deadzone in Eq. (4.4), we have $|e_k| - \lambda|y_{k-1} - y_{l_{k-1}}| \leq a_k|e_k|$.

Therefore, we have

$$\begin{aligned}
|y_{k-1}| &\leq \max_{j \leq k} \{|e_j|\} + C_2 \\
&= \max_{j \leq k} \{|e_j| - \lambda|y_{j-1} - y_{l_{j-1}}| + \lambda|y_{j-1} - y_{l_{j-1}}|\} + C_2 \\
&\leq \max_{j \leq k} \{a_j|e_j|\} + \lambda \max_{j \leq k-1} \{|y_j - y_{l_j}|\} + C_2, \quad k > k_0
\end{aligned} \tag{4.24}$$

Therefore, we have

$$\max_{j \leq k-1} \{|y_j|\} \leq \max_{j \leq k} \{a_j|e_j|\} + 2\lambda \max_{j \leq k-1} \{|y_j|\} + C_3, \quad k > k_0 \tag{4.25}$$

Note that $\lambda < 0.5$, we have

$$\max_{j \leq k-1} \{|y_j|\} \leq \frac{1}{1-2\lambda} \max_{j \leq k} \{a_j|e_j|\} + \frac{C_3}{1-2\lambda}, \quad k > k_0 \tag{4.26}$$

holds for all $\lambda < \lambda^*$, where C_3 is some finite number. Note that inequality Eq. (4.26) means that $y_{k-1} = O[a_k e_k]$. Further we have

$$D_{k-1}^{\frac{1}{2}} \leq |\Phi(y_{k-1}) - \Phi(y_{l_{k-1}})| + |u_{k-1} - u_{l_{k-1}}| = O[y_{k-1}] = O[a_k e_k]$$

Therefore, we can apply the Key Technical Lemma (Lemma 4.3) to Eq. (4.23) and obtain that

$$\lim_{k \rightarrow \infty} a_k e_k = 0 \tag{4.27}$$

which guarantees the boundedness of y_k from Eq. (4.26) and thus, the boundedness of output y_k , tracking error e_k . Therefore, applying Lemma 4.2 yields

$$\lim_{k \rightarrow \infty} |y_k - y_{l_k}| = 0 \tag{4.28}$$

Next, we will show that $\lim_{k \rightarrow \infty} \alpha_k e_k \rightarrow 0$ leads to $\lim_{k \rightarrow \infty} e_k \rightarrow 0$. From the definition of deadzone in Eq. (4.4), we have $a_k \in [0, 1]$. Let us define the following sets:

$$\begin{aligned}
Z_1^+ &= \{k | a_k = 0, k \in Z^+\} \\
Z_2^+ &= \{k | a_k \neq 0, k \in Z^+\}
\end{aligned} \tag{4.29}$$

which results in $Z_1^+ \cap Z_2^+ = \emptyset$ and $Z_1^+ \cup Z_2^+ = Z^+$. The following three cases need to be considered. In every case, we only need to discuss the case where k belongs to an infinite set.

Case i). Z_1^+ is an infinite set and Z_2^+ is a finite set. Let us discuss $k \in Z_1^+$. From the definition in Eq. (4.29), it follows that $a_k = 0$. Hence it is clear from the definition of deadzone (4.4) that $0 \leq |e_k| \leq \lambda |y_{k-1} - y_{l_{k-1}}|$ which means $\lim_{k \rightarrow \infty} e_k \rightarrow 0$ according to Eq. (4.28).

Case ii). Z_1^+ is a finite set and Z_2^+ is an infinite set. Let us discuss $k \in Z_2^+$. From the definition in (4.29), it follows that $a_k \neq 0$. Hence it is clear from deadzone (4.4) that $|e_k| = a_k |e_k| + \lambda |y_{k-1} - y_{l_{k-1}}|$ which means $\lim_{k \rightarrow \infty} e_k = 0$ due to Eqs. (4.27) and (4.28).

Case iii). Z_1^+ and Z_2^+ are infinite sets. If $k \in Z_1^+$ then $a_k = 0$. Following Case i) gives $\lim_{k \rightarrow \infty} e_k = 0$. Otherwise, $a_k \neq 0$, it follows from Case ii) that $\lim_{k \rightarrow \infty} e_k = 0$.

Based on the discussions for the above three cases, we can conclude that $\lim_{k \rightarrow \infty} a_k e_k = 0$ implies that $\lim_{k \rightarrow \infty} e_k = 0$. This completes the proof.

5. Conclusion

In this chapter, we have formulated and discussed the adaptive estimation and control problems for a class of semi-parametric models with both parametric uncertainty and non-parametric uncertainty. For a typical semi-parametric system model, we have discussed new ideas and principles in how to estimate the unknown parameters and non-parametric part by making full use of *a priori* knowledge, and for a typical type of *a priori* knowledge on the non-parametric part, we have proposed novel information-concentration estimator so as to deal with both kinds of uncertainties in the system, and some implementation issues in various cases have been discussed with applicable algorithm descriptions. Furthermore, we have applied the ideas of adaptive estimation for semi-parametric model into two examples of adaptive control problem for two typical semi-parametric control systems, and discussed in details how to establish the closed-loop stability of the whole system with semi-parametric adaptive estimator and controller. Our discussions have demonstrated that the topic in this chapter is very challenging yet important due to its wide background. Especially, for the closed-loop analysis problem of semi-parametric adaptive control, the examples given in this chapter illustrate different methods to overcome the difficulties.

In the first example of semi-parametric adaptive control, we have investigated a simple first-order nonlinear system with both non-parametric uncertainties and parametric uncertainties, which is largely motivated by the recent-year exploration of the capability and limitations of the feedback mechanism. For this model, based on the principle of the proposed IC estimator, we have constructed a unified adaptive controller which can be used in both cases of $b = 1$ and $b > 1$. When the parametric part is of linear growth rate ($b = 1$), we have proved the closed-loop stability under some assumptions and a simple algebraic

condition $\frac{ML}{M'} < \frac{3}{2} + \sqrt{2}$, which reveals essential connections with the known magic

number $L = \frac{3}{2} + \sqrt{2}$ discovered in recent work [XG00] on the study of feedback mechanism capability.

In the second example of semi-parametric adaptive control, we further assume that the control gain is also unknown, yet the system is noise-free, and we have designed an adaptive controller based on gradient-like estimation algorithm with time-varying deadzone according to the *a priori* knowledge on the non-parametric part and the unknown controller gain. In this example, although we cannot establish perfect results revealing the magic

number $\frac{3}{2} + \sqrt{2}$ as in the first example, we can still establish good results of asymptotic

tracking performance under some mild conditions. This example has demonstrated yet another method to deal with uncertainties in semi-parametric model.

Finally, we shall remark that the discussed topic in this chapter is still in its infant stage, and many more challenging problems can be investigated in the future. These problems may root in wide practical background where the system model is only partially known *a priori*, that is to say, the major part of the system can be parameterized and the other part is unknown and non-parameterized with only limited *a priori* knowledge. Solving such problems can definitely improve our understanding to the whole feedback mechanism and help us gain more insights on the capability of adaptive control, especially non-traditional adaptive control methods which were not extensively addressed and studied in previous study. Therefore, we expect more theoretical study in this new topic, i.e. semi-parametric adaptive estimation and control.

6. Acknowledgement

The authors would like to thank Chenguang Yang for his helpful discussion and valuable comments.

7. References

- [ASL98] A. M. Annaswamy, F. P. Skantzes, and A.-P. Loh, "Adaptive control of continuous time systems with convex/concave parametrization," *Automatica*, vol. 34, no. 1, pp. 33–49, 1998.
- [AW89] K. Aström and B. Wittenmark, *Adaptive Control*. Addison-Wesley Publ. Comp., 1989.
- [BDB95] R. M. Bethea, B. S. Duran, and T. L. Boullion, *Statistical Methods for Engineers and Scientists*. CRC Press, 1995.
- [Ber04] R. A. Berk, *Regression Analysis: A Constructive Critique*. SAGE, 2004.
- [BJR08] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, *Time Series Analysis: Forecasting and Control*, 4th ed. Wiley, Jun 2008.
- [Bos95] J. D. Boskovic, "Stable adaptive control of a class of first-order nonlinearly parameterized plants," *IEEE Transactions on Automatic Control*, vol. 40, pp. 347–350, 1995.
- [CG91] H. F. Chen and L. Guo, *Identification and Stochastic Adaptive Control*. Boston, MA: Birkhäuser, 1991.

- [DS98] N. R. Draper and H. Smith, *Applied Regression Analysis*, 3rd ed. Wiley-Interscience, Apr 1998.
- [FL99] G. Feng and R. Lozano, *Adaptive Control Systems*. Butterworth Heinemann, 1999.
- [Fox97] J. Fox, *Applied Regression Analysis, Linear Models, and Related Methods*. SAGE, 1997.
- [Gau09] C. F. Gauss, *Theoria motus corporum coelestium in sectionibus conicis solem ambientium: in sectionibus conicis solem ambientium* (Theorie der Bewegung der Himmelskörper, die die Sonne in Kegelschnitten umkreisen). Dover, New York, 1809, english translation by C. H. Davis, reprinted 1963.
- [GHZ99] S. S. Ge, C. C. Hang, and T. Zhang, "A direct adaptive controller for dynamic systems with a class of nonlinear parameterizations," *Automatica*, vol. 35, pp. 741–747, (1999).
- [Gio85] A. A. Giordano, *Least Square Estimation With Applications to Digital Signal Processing*. Wiley-Interscience, Mar 1985.
- [GS84] G. Goodwin and K. Sin, *Adaptive Filtering, Prediction and Control*. Englewood Cliffs, NJ: Prentice-Hall, 1984.
- [Guo97] L. Guo, "On critical stability of discrete-time adaptive nonlinear control," *IEEE Transactions on Automatic Control*, vol. 42, no. 11, pp. 1488–1499, 1997.
- [Hel63] Z. Hellwig, *Linear regression and its application to economics*. Pergamon Press, 1963.
- [Kan94] I. Kanellakopoulos, "A discrete-time adaptive nonlinear system," *IEEE Transactions on Automatic Control*, vol. 39, Nov. 1994.
- [LH74] C. Lawson and R. Hanson, *Solving Least Squares Problems*. Prentice-Hall, 1974.
- [Lju98] L. Ljung, *System Identification – Theory for the User*, 2nd ed. Upper Saddle River, NJ 07458: Prentice Hall, 1998.
- [LQF03] A. P. Loh, C. Y. Qu, and K. F. Fong, "Adaptive control of discrete time systems with concave/convex parametrizations," *IEEE Transactions on Automatic Control*, vol. 48, no. 6, 2003.
- [LX06] C. Y. Li and L. L. Xie, "On robust stability of discrete-time adaptive nonlinear control," *Systems and Control Letters*, vol. 55, no. 6, pp. 452–458, June 2006.
- [Ma06] H. B. Ma, "Capability and limitation of feedback mechanism in dealing with uncertainties of some discrete-time control systems," Phd Thesis, Graduate School of Chinese Academy of Sciences, Beijing, China, June 2006.
- [Ma08a] —, "Further results on limitations to the capability of feedback," *International Journal of Control*, vol. 81, no. 1, pp. 21 – 42, 2008. [Online]. Available: <http://www.informaworld.com/10.1080/00207170701218333>
- [Ma08b] —, "An "impossibility" theorem on a class of high-order discrete-time nonlinear control systems," *Systems & Control Letters*, vol. 57, no. 6, pp. 497–504, 2008. [Online]. Available: <http://dx.doi.org/10.1016/j.sysconle.2007.11.008>
- [MG05] H. B. Ma and L. Guo, "An "impossibility" theorem on second-order discrete-time nonlinear control systems," in *Proceedings of the 24th Chinese Control Conference*. South China University of Technology Press: Guangzhou, July. 2005, pp. 57–61.
- [MPV07] D. C. Montgomery, E. A. Peck, and G. G. Vining, *Introduction to Linear Regression Analysis, Solutions Manual*, 4th ed., ser. Wiley Series in Probability and Statistics. Wiley, John Sons, Incorporated, Feb 2007.
- [MT95] R. Marino and P. Tomei, *Nonlinear adaptive design: Geometric, adaptive, and robust*. London: Printice-Hall, 1995.

- [SB89] S. Sastry and M. Bodson, *Adaptive Control: Stability, Convergence, and Robustness*. Englewood, New Jersey: Prentice-Hall, 1989.
- [Sok03] V. F. Sokolov, "Adaptive suboptimal tracking for the first-order plant with lipschitz uncertainty," *IEEE Transactions on Automatic Control*, vol. 48, no. 4, Apr 2003.
- [Tsa05] R. S. Tsay, *Analysis of Financial Time Series*, 2nd ed., ser. Wiley Series in Probability and Statistics. Wiley, John Sons, Incorporated, Aug 2005, no. 543.
- [VV07] M. Verhaegen and V. Verdult, *Filtering and System Identification: A Least Squares Approach*, 1st ed. Cambridge University Press, May 2007.
- [web08] "Statistics & data analysis for environmental science and management," 2008, available online from [http://www.esm.ucsb.edu/academics/courses/206/Readings/reader Ch8.pdf](http://www.esm.ucsb.edu/academics/courses/206/Readings/reader%20Ch8.pdf).
- [Wei05] S. Weisberg, *Applied Linear Regression*, 3rd ed. Wiley-Interscience, Jan 2005.
- [Wik08a] Wikipedia, "Gauss-Markov theorem – wikipedia, the free encyclopedia," 2008. [Online]. Available: [http://en.wikipedia.org/w/index.php?title=Gauss%E2%80%93Markov theorem&oldid=232973932](http://en.wikipedia.org/w/index.php?title=Gauss%E2%80%93Markov%20theorem&oldid=232973932)
- [Wik08b] —, "Generalized linear model – wikipedia, the free encyclopedia," 2008. [Online]. Available: [http://en.wikipedia.org/w/index.php?title=Generalized linear model&oldid=237425065](http://en.wikipedia.org/w/index.php?title=Generalized%20linear%20model&oldid=237425065)
- [Wik08c] —, "Kernel regression – wikipedia, the free encyclopedia," 2008. [Online]. Available: [http://en.wikipedia.org/w/index.php?title=Kernel regression&oldid=234955359](http://en.wikipedia.org/w/index.php?title=Kernel%20regression&oldid=234955359)
- [Wik08d] —, "Kriging – wikipedia, the free encyclopedia," 2008. [Online]. Available: <http://en.wikipedia.org/w/index.php?title=Kriging&oldid=234876679>
- [Wik08e] —, "Least squares – wikipedia, the free encyclopedia," 2008. [Online]. Available: [http://en.wikipedia.org/w/index.php?title=Least squares&oldid=233230305](http://en.wikipedia.org/w/index.php?title=Least%20squares&oldid=233230305)
- [Wik08f] —, "Linear least squares – wikipedia, the free encyclopedia," 2008. [Online]. Available: [http://en.wikipedia.org/w/index.php?title=Linearleast squares&oldid=233230344](http://en.wikipedia.org/w/index.php?title=Linear%20least%20squares&oldid=233230344)
- [Wik08g] —, "Linear regression – wikipedia, the free encyclopedia," 2008. [Online]. Available: [http://en.wikipedia.org/w/index.php?title=Linear regression&oldid=237064808](http://en.wikipedia.org/w/index.php?title=Linear%20regression&oldid=237064808)
- [Wik08h] —, "Moore-Penrose pseudoinverse – wikipedia, the free encyclopedia," 2008. [Online]. Available: [http://en.wikipedia.org/w/index.php?title=Moore-Penrose pseudoinverse&oldid=237863190](http://en.wikipedia.org/w/index.php?title=Moore-Penrose%20pseudoinverse&oldid=237863190)
- [Wik08i] —, "Nonparametric regression – wikipedia, the free encyclopedia," 2008. [Online]. Available: [http://en.wikipedia.org/w/index.php?title=Nonparametric regression&oldid=235577025](http://en.wikipedia.org/w/index.php?title=Nonparametric%20regression&oldid=235577025)
- [Wik08j] —, "Regression analysis – wikipedia, the free encyclopedia," 2008. [Online]. Available: [http://en.wikipedia.org/w/index.php?title=Regression analysis&oldid=232221082](http://en.wikipedia.org/w/index.php?title=Regression%20analysis&oldid=232221082)
- [XG99] L. L. Xie and L. Guo, "Fundamental limitations of discrete-time adaptive nonlinear control," *IEEE Transactions on Automatic Control*, vol. 44, no. 9, pp. 1777–1782, 1999.
- [XG00] —, "How much uncertainty can be dealt with by feedback?" *IEEE Transactions on Automatic Control*, vol. 45, no. 12, pp. 2203–2217, 2000.

- [XG01] F. Xue and L. Guo, "Necessary and sufficient conditions for adaptive stabilizability of jump linear systems," *Communications in Information and Systems*, vol. 1, no. 2, pp. 205–224, Apr. 2001.
- [YDG+08] C. G. Yang, S. L. Dai, S. S. Ge, Z. X. Jiao, and T. H. Lee, "Adaptive asymptotic tracking control of a class of strict-feedback discrete-time nonlinear systems," *IEEE Transactions on Automatic Control*, 2008, to be submitted.
- [ZG02] Y. X. Zhang and L. Guo, "A limit to the capability of feedback," *IEEE Transactions on Automatic Control*, vol. 47, no. 4, pp. 687–692, 2002.



Adaptive Control

Edited by Kwanho You

ISBN 978-953-7619-47-3

Hard cover, 372 pages

Publisher InTech

Published online 01, January, 2009

Published in print edition January, 2009

Adaptive control has been a remarkable field for industrial and academic research since 1950s. Since more and more adaptive algorithms are applied in various control applications, it is becoming very important for practical implementation. As it can be confirmed from the increasing number of conferences and journals on adaptive control topics, it is certain that the adaptive control is a significant guidance for technology development. The authors of the chapters in this book are professionals in their areas and their recent research results are presented in this book which will also provide new ideas for improved performance of various control application problems.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Hongbin Ma and Kai-Yew Lum (2009). Adaptive Estimation and Control for Systems with Parametric and Nonparametric Uncertainties, Adaptive Control, Kwanho You (Ed.), ISBN: 978-953-7619-47-3, InTech, Available from:

http://www.intechopen.com/books/adaptive_control/adaptive_estimation_and_control_for_systems_with_parametric_and_nonparametric_uncertainties

INTECH
open science | open minds

InTech Europe

University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2009 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the [Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License](https://creativecommons.org/licenses/by-nc-sa/3.0/), which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.

IntechOpen

IntechOpen