Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005

The Size of the Membership Set in the Presence of Disturbance and Parameter Uncertainty

Wataru Kitamura, Yasumasa Fujisaki, and Er-Wei Bai

***Abstract*— The size of the membership set is analyzed probabilistically in the presence of not only bounded disturbance but also l_2 bounded parameter uncertainty. Especially, the diameter of the membership set is estimated with a probabilistic confidence for a finite number of samples, where the regressor is assumed to be persistently exciting and the disturbance and the parameter uncertainty are assumed to be random variables. It is also shown that, under the same assumptions, the diameter converges to zero as the number of samples tends to infinity.**

I. Introduction

The membership set represents the entire set of system parameters which are consistent with measured data and a priori information. Nowadays, many algorithms to estimate this membership set are available. See for example [1], [2], [3], [4]. On the other hand, there are a few studies [5], [6], [7] on fundamental properties of the membership set, where the relationship between the number of input-output data and the size of the membership set are investigated. In particular, it is shown that, in a stochastic setting, the membership set converges to the true parameter of the system as the number of samples tends to infinity.

However, it should be noticed that, in the literature [5], [6], [7], the properties have been explored for a certain systems with additive bounded disturbance, i.e., no parameter uncertainty of the system is considered. Since we can not neglect parameter uncertainty in practical situation and its influence appears in output signals of the system not as an additive bounded noise but as an unbounded noise whose amplitude depends on the magnitude of input signals, we need a theory which clarifies the fundamental properties of the membership set in the presence of both disturbance and parameter uncertainty.

From this point of view, the authors have investigated the size of the membership set in the presence of not only disturbance but also parameter uncertainty [8], [9], where the diameter of the membership set is estimated with a probabilistic confidence for a finite number of samples in a stochastic setting. Then, it is shown that the diameter converges to zero as the number of samples tends to infinity. However, in these papers, the regressors are assumed to be periodic due to a technicality of the proofs, which is restrictive for practical application.

W. Kitamura was with Graduate School of Science and Technology, Kobe University. He is currently with Murata Machinery, Ltd..

Y. Fujisaki is with Department of Computer and Systems Engineering, Kobe University, Nada, Kobe 657-8501, Japan. E-mail: fujisaki@cs.kobe-u.ac.jp

E.-W. Bai is with Department of Electrical and Computer Engineering, University of Iowa, Iowa City, IA 52242, U.S.A. E-mail: erwei@icaen.uiowa.edu

In this paper, we investigate the size of the membership set in the presence of bounded disturbance and l_2 bounded parameter uncertainty without periodicity of the regressors. To this end, we employ a useful lemma in [7], which gives a set of non-periodic regressors that can play a role of periodic regressors. For a finite number of samples, we first derive the lower and the upper bounds of the probability so that the diameter of the membership set is greater than or equal to a given value, where we assume that the regressor is persistently exciting and the disturbance and the parameter uncertainty are random variables which take their extreme values with nonzero probability. Then, we show that, under the same assumptions, the diameter of the membership set converges to zero as the number of samples tends to infinity.

Note that both lower and upper bounds of the probability regarding the diameter of the membership set are obtained in this paper. That is, apart from the assumption of periodicity of the regressors, the results of this paper are in sharp contrast to the previous results [8], [9], where only an upper bound of the probability is derived. As we will see in the subsequent sections, the lower and the upper bounds of the probability are exponential functions with respect to the number of samples. This implies that the probability itself is also an exponential function with respect to the number of samples.

The outline of this paper is as follows. In Section II, we define the membership set in the presence of disturbance and parameter uncertainty and give its characterization. In Section III, we state the main results of the paper. Introducing a few definitions on regressors, disturbance, and parameter uncertainty, we give an estimation of the size of the membership set. Section IV illustrates the main results through numerical examples. In Section V, we make some concluding remarks.

II. The Membership Set

Let us describe a single-input single-output (SISO) discrete time system as

$$y_i = {\phi_i}^{\mathrm{T}}(\theta + \eta_i) + v_i \tag{1}$$

where $y_i \in \mathbb{R}$ is the system output, $\phi_i \in \mathbb{R}^m$ is the measurable regression vector, $\theta \in \mathbb{R}^m$ is a nominal parameter of the system to be identified, $v_i \in \mathbb{R}$ is an unknown disturbance, and $\eta_i \in \mathbb{R}^m$ is an unknown parameter uncertainty. The disturbance and the parameter uncertainty are bounded, that is,

$$|v_i| \leq \epsilon, \quad \|\eta_i\|_2 \leq \delta \tag{2}$$

0-7803-9568-9/05/$20.00 ©2005 IEEE

where the bounds ϵ and δ are given, and $\|\bullet\|_2$ is the l_2 norm

$$\|\eta_i\|_2 \doteq ({\eta_i}^{\mathrm{T}}\eta_i)^{1/2}.$$

Then, the membership set is defined as all estimates $\hat{\theta}$ that are consistent with the given data $\{y_i, \phi_i\}, i = 1, 2, \ldots, N$ and the given bounds of the disturbance and the parameter uncertainty, that is,

$$\Omega^N(\epsilon, \delta) \doteq \bigcap_{i=1}^{N} \Big\{ \hat{\theta} \in \mathbb{R}^m : \exists \eta_i, \exists v_i \text{ s.t.} \\ y_i = {\phi_i}^{\mathrm{T}}(\hat{\theta} + \eta_i) + v_i,\ |v_i| \leq \epsilon,\ \|\eta_i\|_2 \leq \delta \Big\}.$$

We first rewrite the system (1) as

$$y_i - {\phi_i}^{\mathrm{T}}\theta = v_i + {\phi_i}^{\mathrm{T}}\eta_i.$$

Notice that the disturbance and the parameter uncertainty are bounded as

$$|v_i + {\phi_i}^{\mathrm{T}}\eta_i| \leq \epsilon + \delta\|\phi_i\|_2 \tag{3}$$

where the equality in the above can be attained with

$$v_i = \pm\epsilon, \qquad \eta_i = \pm\frac{\delta}{\|\phi_i\|_2}\phi_i$$

which meet (2) [9]. Thus, we can represent the membership set as

$$\Omega^N(\epsilon, \delta) = \bigcap_{i=1}^{N} \Big\{ \hat{\theta} \in \mathbb{R}^m : \\ -(\epsilon + \delta\|\phi_i\|_2) \leq y_i - {\phi_i}^{\mathrm{T}}\hat{\theta} \leq \epsilon + \delta\|\phi_i\|_2 \Big\}. \tag{4}$$

That is, the entire region of system parameter $\hat{\theta}$ can be characterized as an intersection of inequalities.

We remark that the parameter uncertainty behaves as an additional disturbance. However, notice here that the resultant total bound $\epsilon + \delta\|\phi_i\|_2$ depends on the regressor and is time varying. In the rest of the paper, we use this convenient characterization (4) of the membership set.

III. Probabilistic Analysis of the Size of the Membership Set

In this section, we investigate the size of the membership set in a stochastic setting, where the disturbance and the parameter uncertainty are regarded as random variables.

Let us introduce some definitions. The following [10] is a deterministic version of the definition of persistently exciting.

Definition 1: The regressor $\phi_i,\ i = 1, 2, \ldots, N$ is said to be persistently exciting (PE) if there exist some $\alpha > 0$, $\beta > 0$, $n_0 > 0$ such that

$$\alpha^2 I \leq \sum_{i=i_0}^{i_0+n_0-1} \phi_i {\phi_i}^{\mathrm{T}} \leq \beta^2 I$$

for $1 \leq i_0 \leq N - n_0 + 1$. Here, the matrix inequality $A \leq B$ means $x^T(B - A)x \geq 0$ for all x.

Let $\mathbb{N}$ denote the set of positive integers and $\mathcal{N}(\bullet)$ denote the number of elements of a given set $\bullet$. The following lemma is taken from [7].

Lemma 1: Assume that the regressor ϕ_i, $i = 1, 2, \ldots, N$ is PE. Suppose that $N = 2rn_0$, where $r \in \mathbb{N}$. In addition, let $\mathcal{P}$ be an integer defined by

$$\mathcal{P} \doteq 2 + 2q(q-1)^{m-2}$$

where

$$q \doteq \min_p \left\{ p \in \mathbb{N} : p\alpha \geq \pi\beta\sqrt{n_0(m-1)} \right\}.$$

Then, there exists a set $\mathcal{K}$ of subsequences $\{i_k\}$ of time series having all of the following properties (i) – (iii):

(i) $\mathcal{N}(\{i_k\}) \geq r, \quad \forall \{i_k\} \in \mathcal{K}$;
(ii) $\mathcal{N}(\mathcal{K}) \leq \mathcal{P}$;
(iii) For each $x \in \mathbb{R}^m$, there is a $\{i_k\} \in \mathcal{K}$ such that either one of the following holds:

$$x^{\mathrm{T}}\phi_{i_k} \leq -\frac{\alpha}{2\sqrt{n_0}}\|x\|_2, \quad \forall k; \tag{5}$$

$$x^{\mathrm{T}}\phi_{i_k} \geq \frac{\alpha}{2\sqrt{n_0}}\|x\|_2, \quad \forall k. \tag{6}$$

This lemma means the following: Assume that the regressor is PE. If we take at most $\mathcal{P}$ kinds of subsequences $\{i_k\}$ whose length is of r at least, then there exists a subsequence $\{i_k\}$ such that each of $\{\phi_{i_k}\}$ has the same direction for any fixed $x \in \mathbb{R}^m$ in the sense of (5) or (6). We also remark that r is proportional to N, while $\mathcal{P}$ is an integer depending on m, α, β, and n_0.

The following definitions are introduced in [5] and [9], which are probabilistic properties on the disturbance and the parameter uncertainty.

Definition 2: Suppose that the disturbance v_i is a random variable satisfying $|v_i| \leq \epsilon$. The bound ϵ is said to be tight if for any $\rho > 0$ and each i

$$p_{vL}(\rho) \leq \mathrm{Prob}\{-\epsilon \leq v_i \leq -(\epsilon - \rho)\} \leq p_{vU}(\rho)$$
$$p_{vL}(\rho) \leq \mathrm{Prob}\{\epsilon - \rho \leq v_i \leq \epsilon\} \leq p_{vU}(\rho)$$

for some $0 < p_{vL}(\rho)$, $p_{vU}(\rho) \leq 1$. Here, $\mathrm{Prob}\{\bullet\}$ denotes the probability that the event $\bullet$ occurs.

Definition 3: Suppose that the parameter uncertainty η_i is a random vector satisfying $\|\eta_i\|_2 \leq \delta$. The bound δ is said to be tight if for any $\mu > 0$, each i, and any ϕ_i

$$p_{\eta L}(\mu) \leq \mathrm{Prob}\{-\delta\|\phi_i\|_2 \leq {\phi_i}^{\mathrm{T}}\eta_i \leq -(\delta - \mu)\|\phi_i\|_2\} \\ \leq p_{\eta U}(\mu)$$
$$p_{\eta L}(\mu) \leq \mathrm{Prob}\{(\delta - \mu)\|\phi_i\|_2 \leq {\phi_i}^{\mathrm{T}}\eta_i \leq \delta\|\phi_i\|_2\} \\ \leq p_{\eta U}(\mu)$$

for some $0 < p_{\eta L}(\mu)$, $p_{\eta U}(\mu) \leq 1$.

The tightness of the bounds of v_i and η_i means that v_i and η_i take around their extreme values with nonzero probability. This is obvious for v_i. To see this fact for η_i, we rewrite the formulae in Definition 3 as

$$p_{\eta L}(\mu) \leq \mathrm{Prob}\{-\delta \leq \frac{\phi_i^{\mathrm{T}}}{\|\phi_i\|_2}\eta_i \leq -(\delta - \mu)\} \leq p_{\eta U}(\mu)$$

$$p_{\eta L}(\mu) \leq \mathrm{Prob}\{\delta - \mu \leq \frac{\phi_i^{\mathrm{T}}}{\|\phi_i\|_2}\eta_i \leq \delta\} \leq p_{\eta U}(\mu).$$

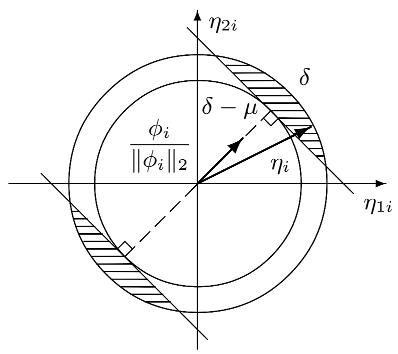

Fig. 1. Tightness in the case $m = 2$

When $m = 2$, we can illustrate this probabilistic statement with Fig. 1. That is, for a fixed ϕ_i, η_i takes a value in each shaded portion with a probability specified by $p_{\eta L}(\mu)$ and $p_{\eta U}(\mu)$. Since this holds for any ϕ_i, we see that, with a nonzero probability, η_i takes a value in the ring determined by $\delta - \mu$ and δ, i.e., around the surface of the possible values.

Then, the next lemma is obtained.

Lemma 2: Assume that the disturbance v_i and the parameter uncertainty η_i are independent random variables, and their bounds ϵ and δ are tight. Then, for any $\rho > 0$, $\mu > 0$, each i, and any ϕ_i

$$\begin{aligned} p_{eL}(\rho,\mu) \le \mathrm{Prob}\big\{&-(\epsilon+\delta\|\phi_i\|_2) \le v_i + {\phi_i}^{\mathrm{T}}\eta_i \\ &\le -(\epsilon-\rho) - (\delta-\mu)\|\phi_i\|_2\big\} \\ \le p_{eU}&(\rho,\mu) \end{aligned} \tag{7}$$

$$\begin{aligned} p_{eL}(\rho,\mu) \le \mathrm{Prob}\big\{&\epsilon-\rho+(\delta-\mu)\|\phi_i\|_2 \\ &\le v_i + {\phi_i}^{\mathrm{T}}\eta_i \le \epsilon+\delta\|\phi_i\|_2\big\} \\ \le p_{eU}&(\rho,\mu) \end{aligned} \tag{8}$$

for some $0 < p_{eL}(\rho,\mu)$, $p_{eU}(\rho,\mu) \le 1$, where

$$\begin{aligned} p_{eL}(\rho,\mu) &\doteq p_{vL}(\rho)p_{\eta L}(\mu) \\ p_{eU}(\rho,\mu) &\doteq p_{vU}(\rho) + p_{\eta U}(\mu) - p_{vU}(\rho)p_{\eta U}(\mu). \end{aligned}$$

Proof: We first show the inequality (8). The lower bound of the probability in (8) is obtained if we evaluate the probability as

$$\begin{aligned} &\mathrm{Prob}\big\{\epsilon-\rho+(\delta-\mu)\|\phi_i\|_2 \le v_i + {\phi_i}^{\mathrm{T}}\eta_i \le \epsilon+\delta\|\phi_i\|_2\big\} \\ &\ge \mathrm{Prob}\big\{\epsilon-\rho \le v_i \le \epsilon\big\} \\ &\quad \cdot \mathrm{Prob}\big\{(\delta-\mu)\|\phi_i\|_2 \le {\phi_i}^{\mathrm{T}}\eta_i \le \delta\|\phi_i\|_2\big\} \\ &\ge p_{vL}(\rho)p_{\eta L}(\mu). \end{aligned}$$

On the other hand, the upper bound in (8) is obtained as

$$\begin{aligned} &\mathrm{Prob}\big\{\epsilon-\rho+(\delta-\mu)\|\phi_i\|_2 \le v_i + {\phi_i}^{\mathrm{T}}\eta_i \le \epsilon+\delta\|\phi_i\|_2\big\} \\ &\le 1 - \mathrm{Prob}\big\{-\epsilon \le v_i \le \epsilon-\rho\big\} \\ &\qquad \cdot \mathrm{Prob}\big\{-\delta\|\phi_i\|_2 \le {\phi_i}^{\mathrm{T}}\eta_i \le (\delta-\mu)\|\phi_i\|_2\big\} \\ &\le 1 - (1-p_{vU}(\rho))(1-p_{\eta U}(\mu)) \\ &= p_{vU}(\rho) + p_{\eta U}(\mu) - p_{vU}(\rho)p_{\eta U}(\mu). \end{aligned}$$

Similarly, we can derive the lower and upper bounds of the probability in (7). ■

This lemma shows that, if the bound ϵ of v_i and the bound δ of η_i are tight, the bound $\epsilon + \delta\|\phi_i\|_2$ of $v_i + {\phi_i}^{\mathrm{T}}\eta_i$ is also tight, that is, $v_i + {\phi_i}^{\mathrm{T}}\eta_i$ takes around its extreme value with nonzero probability.

Then, we obtain the following, which is the main result of this paper.

Theorem 1: Consider the system (1). Assume that the regressor ϕ_i is deterministic, PE, and $\underline{\beta} \le \|\phi_i\|_2 \le \overline{\beta}$, $\forall i$. In addition, v_i and η_i are independent random variables, and their bounds are tight. Then, the size of the membership set is estimated as

$$\mathrm{Prob}\left\{\mathrm{dia}\,\Omega^N(\epsilon,\delta) \ge \frac{4\sqrt{n_0}}{\alpha}(\rho+\mu\underline{\beta})\right\} \ge p_{rL}(\rho,\mu) \tag{9}$$

$$\mathrm{Prob}\left\{\mathrm{dia}\,\Omega^N(\epsilon,\delta) \ge \frac{4\sqrt{n_0}}{\alpha}(\rho+\mu\overline{\beta})\right\} \le p_{rU}(\rho,\mu) \tag{10}$$

where $\mathrm{dia}\,\Omega^N(\epsilon,\delta)$ is the diameter of the membership set

$$\mathrm{dia}\,\Omega^N(\epsilon,\delta) \doteq \sup_{\hat{\theta}_1,\hat{\theta}_2 \in \Omega^N(\epsilon,\delta)} \|\hat{\theta}_1 - \hat{\theta}_2\|_2$$

and $p_{rL}(\rho,\mu)$, $p_{rU}(\rho,\mu)$ are positive constants defined by

$$\begin{aligned} &p_{rL}(\rho,\mu) \\ &\quad \doteq \left[\max\left\{0, 1 - 2p_{eU}\left(\frac{4\overline{\beta}\sqrt{n_0}}{\alpha}\rho, \frac{4\overline{\beta}\sqrt{n_0}}{\alpha}\mu\right)\right\}\right]^N \\ &p_{rU}(\rho,\mu) \doteq \min\left\{1, \mathcal{P}\left(1-p_{eL}(\rho,\mu)\right)^r\right\}. \end{aligned}$$

Proof: We first show (9). Let us define

$$e_i \doteq v_i + {\phi_i}^{\mathrm{T}}\eta_i, \quad \nu_i \doteq \epsilon + \delta\|\phi_i\|_2.$$

Note that $\nu_i \pm e_i \ge 0$, which follows (3). Then, we rewrite the membership set (4) as

$$\Omega^N(\epsilon,\delta) = \bigcap_{i=1}^{N}\Big\{\hat{\theta} \in \mathbb{R}^m : \\ -(\nu_i + e_i) \le {\phi_i}^{\mathrm{T}}(\theta - \hat{\theta}) \le \nu_i - e_i\Big\}. \tag{11}$$

Let us consider the set $\mathcal{S}$ as

$$\mathcal{S} \doteq \bigcap_{i=1}^{N}\Big\{\hat{\theta} \in \mathbb{R}^m : \\ |{\phi_i}^{\mathrm{T}}(\theta - \hat{\theta})| < \min\left[\min_j(\nu_j + e_j), \min_j(\nu_j - e_j)\right]\Big\}$$

where $j = 1, 2, \ldots, N$. Since we see that

$$\Omega^N(\epsilon,\delta) \supset \mathcal{S},$$

there exists at least an $i^* \in \{1, 2, \ldots, N\}$ satisfying

$$\min\left[\min_j(\nu_j + e_j), \min_j(\nu_j - e_j)\right] \le |{\phi_{i^*}}^{\mathrm{T}}(\theta - \widetilde{\theta})| \tag{12}$$

for any $\widetilde{\theta} \in \Omega^N(\epsilon,\delta)\backslash\mathcal{S}$, where $\Omega^N(\epsilon,\delta)\backslash\mathcal{S}$ denotes the complement of $\mathcal{S}$ with respect to $\Omega^N(\epsilon,\delta)$. Using $\|\phi_i\|_2 \le$

$\overline{\beta}$, $\forall i$, we have

$$\begin{aligned}\mathrm{dia}\,\Omega^N(\epsilon,\delta) &\geq \sup_{\hat\theta\in\Omega^N(\epsilon,\delta)} \|\theta-\hat\theta\|_2 \\ &\geq (\overline{\beta})^{-1} \sup_{i,\ \hat\theta\in\Omega^N(\epsilon,\delta)} \|\phi_i\|_2\|\theta-\hat\theta\|_2 \\ &\geq (\overline{\beta})^{-1} \sup_{i,\ \hat\theta\in\Omega^N(\epsilon,\delta)} |{\phi_i}^{\mathrm{T}}(\theta-\hat\theta)|. \qquad (13)\end{aligned}$$

From inequality (12), we obtain

$$\begin{aligned}&\sup_{i,\ \hat\theta\in\Omega^N(\epsilon,\delta)} |{\phi_i}^{\mathrm{T}}(\theta-\hat\theta)| \\ &\geq \min\left[\min_j(\nu_j+e_j), \min_j(\nu_j-e_j)\right] \\ &= \min_j\left(\min\left[\nu_j+e_j, \nu_j-e_j\right]\right). \qquad (14)\end{aligned}$$

Substituting this inequality (14) into (13), we see that

$$\mathrm{dia}\,\Omega^N(\epsilon,\delta) \geq (\overline{\beta})^{-1}\min_j\left(\min\left[\nu_j+e_j, \nu_j-e_j\right]\right).$$

Hence, we have

$$\begin{aligned}&\mathrm{Prob}\left\{\mathrm{dia}\,\Omega^N(\epsilon,\delta) \geq \frac{4\sqrt{n_0}}{\alpha}(\rho+\mu\underline{\beta})\right\} \\ &\geq \mathrm{Prob}\left\{(\overline{\beta})^{-1}\min_j\left(\min\left[\nu_j+e_j, \nu_j-e_j\right]\right.\right. \\ &\qquad\qquad \left.\geq \frac{4\sqrt{n_0}}{\alpha}(\rho+\mu\underline{\beta})\right\} \\ &= \mathrm{Prob}\left\{-\nu_i+\frac{4\overline{\beta}\sqrt{n_0}}{\alpha}(\rho+\mu\underline{\beta}) \leq e_i\right. \\ &\qquad\qquad \left.\leq \nu_i-\frac{4\overline{\beta}\sqrt{n_0}}{\alpha}(\rho+\mu\underline{\beta}),\quad \forall i\right\}. \qquad (15)\end{aligned}$$

Now, using Lemma 2 with $\|\phi_i\|_2 \geq \underline{\beta}$, $\forall i$, we see that

$$\begin{aligned}&\mathrm{Prob}\left\{-\nu_i+(\widehat\rho+\widehat\mu\underline{\beta}) \leq e_i \leq \nu_i-(\widehat\rho+\widehat\mu\underline{\beta})\right\} \\ &\geq \mathrm{Prob}\{-\nu_i+(\widehat\rho+\widehat\mu\|\phi_i\|_2) \leq e_i \leq \nu_i-(\widehat\rho+\widehat\mu\|\phi_i\|_2)\} \\ &\geq 1-2p_{eU}(\widehat\rho,\widehat\mu) \qquad (16)\end{aligned}$$

holds for any $\widehat\rho$, $\widehat\mu$, i, and ϕ_i. Here we set $\widehat\rho = 4\overline{\beta}\sqrt{n_0}\rho/\alpha$, $\widehat\mu = 4\overline{\beta}\sqrt{n_0}\mu/\alpha$ in (16). We also substitute these $\widehat\rho$ and $\widehat\mu$ into (15), and evaluate it with (16). Then, we obtain the conclusion (9). We remark that the lower bound of (16) is 0 since probability is a nonnegative number.

Next, we show (10). Let us introduce $\hat\theta$ which denotes a member of $\Omega^N(\epsilon,\delta)$. We first fix $\hat\theta$. Let $x=\theta-\hat\theta$. From Lemma 1–(iii), there exists a subsequence $\{i_k\}\in\mathcal{K}$ such that either (5) or (6) holds. If (5) holds, then from (11)

$$-\nu_{i_k}-e_{i_k} \leq -\frac{\alpha}{2\sqrt{n_0}}\|\theta-\hat\theta\|_2,\quad \forall k.$$

It turns out that

$$\frac{\alpha}{2\sqrt{n_0}}\|\theta-\hat\theta\|_2 \leq \min_k(\nu_{i_k}+e_{i_k}). \qquad (17)$$

On the other hand, if (6) holds, in the same way, we have

$$\frac{\alpha}{2\sqrt{n_0}}\|\theta-\hat\theta\|_2 \leq \min_k(\nu_{i_k}-e_{i_k}) \qquad (18)$$

from (11). Hence, from (17) and (18), we obtain

$$\frac{\alpha}{2\sqrt{n_0}}\|\theta-\hat\theta\|_2 \leq \psi(\{i_k\})$$

where

$$\psi(\{i_k\}) \doteq \begin{cases}\min_k(\nu_{i_k}+e_{i_k}) & \text{if (5) holds} \\ \min_k(\nu_{i_k}-e_{i_k}) & \text{if (6) holds.}\end{cases}$$

Lemma 1–(iii) says that there exists a subsequences $\{i_k\}\in\mathcal{K}$ for any $\hat\theta\in\Omega^N(\epsilon,\delta)$. Thus, we see

$$\frac{\alpha}{2\sqrt{n_0}}\sup_{\hat\theta\in\Omega^N(\epsilon,\delta)}\|\theta-\hat\theta\|_2 \leq \max_{\{i_k\}\in\mathcal{K}}\psi(\{i_k\}).$$

Since the right-hand side of the above inequality can be evaluated as

$$\begin{aligned}&\sup_{\hat\theta_1,\hat\theta_2\in\Omega^N(\epsilon,\delta)}\|\hat\theta_1-\hat\theta_2\|_2 \\ &= \sup_{\hat\theta_1,\hat\theta_2\in\Omega^N(\epsilon,\delta)}\|(\hat\theta_1-\theta)-(\hat\theta_2-\theta)\|_2 \\ &\leq \sup_{\hat\theta\in\Omega^N(\epsilon,\delta)} 2\|\theta-\hat\theta\|_2,\end{aligned}$$

we see that

$$\begin{aligned}\mathrm{dia}\,\Omega^N(\epsilon,\delta) &= \sup_{\hat\theta_1,\hat\theta_2\in\Omega^N(\epsilon,\delta)}\|\hat\theta_1-\hat\theta_2\|_2 \\ &\leq \frac{4\sqrt{n_0}}{\alpha}\max_{\{i_k\}\in\mathcal{K}}\psi(\{i_k\}). \qquad (19)\end{aligned}$$

We can therefore estimate the diameter of the membership set as

$$\begin{aligned}&\mathrm{Prob}\left\{\mathrm{dia}\,\Omega^N(\epsilon,\delta) \geq \frac{4\sqrt{n_0}}{\alpha}(\rho+\mu\overline{\beta})\right\} \\ &\leq \mathrm{Prob}\left\{\max_{\{i_k\}\in\mathcal{K}}\psi(\{i_k\}) \geq \rho+\mu\overline{\beta}\right\}. \qquad (20)\end{aligned}$$

Suppose that $\psi(\{i_k\})$ takes a maximum value when

$$\psi(\{i_k\}) = \min_k(\nu_{i_k}+e_{i_k})$$

holds. From Lemma 1–(ii) and $\|\phi_i\|_2 \leq \overline{\beta}$, $\forall i$, we have

$$\begin{aligned}&\mathrm{Prob}\left\{\max_{\{i_k\}\in\mathcal{K}}\min_k(\nu_{i_k}+e_{i_k}) \geq \rho+\mu\overline{\beta}\right\} \\ &= \mathrm{Prob}\left\{\bigcup_{\{i_k\}\in\mathcal{K}}\left\{\min_k(\nu_{i_k}+e_{i_k}) \geq \rho+\mu\overline{\beta}\right\}\right\} \\ &\leq \mathcal{N}(\mathcal{K})\cdot\mathrm{Prob}\left\{\min_k(\nu_{i_k}+e_{i_k}) \geq \rho+\mu\overline{\beta}\right\} \\ &\leq \mathcal{P}\cdot\mathrm{Prob}\left\{\nu_{i_k}+e_{i_k} \geq \rho+\mu\|\phi_{i_k}\|_2,\quad \forall k\right\}. \qquad (21)\end{aligned}$$

Here, recall that Lemma 2 says that

$$\mathrm{Prob}\left\{\nu_i+e_i \geq \rho+\mu\|\phi_i\|_2\right\} \leq 1-p_{eL}(\rho,\mu) \qquad (22)$$

holds for any ρ, μ, i, and ϕ_i. Substituting this inequality and Lemma 1–(i) into (21), we obtain

$$\begin{aligned}&\mathcal{P}\cdot\mathrm{Prob}\left\{\nu_{i_k}+e_{i_k} \geq \rho+\mu\|\phi_{i_k}\|_2,\ \forall k\right\} \\ &\leq \mathcal{P}\left(1-p_{eL}(\rho,\mu)\right)^r. \qquad (23)\end{aligned}$$

On the other hand, suppose that $\psi(\{i_k\})$ takes a maximum value when

$$\psi(\{i_k\}) = \min_k(\nu_{i_k} - e_{i_k})$$

holds. In the same way, we obtain

$$\begin{aligned} &\mathrm{Prob}\left\{\max_{\{i_k\}\in\mathcal{K}}\min_k(\nu_{i_k} - e_{i_k}) \geq \rho + \mu\overline{\beta}\right\} \\ &\leq \mathcal{P}\cdot \mathrm{Prob}\left\{\nu_{i_k} - e_{i_k} \geq \rho + \mu\|\phi_{i_k}\|_2, \quad \forall k\right\} \\ &\leq \mathcal{P}\left(1 - p_{eL}(\rho,\mu)\right)^r. \end{aligned} \tag{24}$$

Using the inequalities (20), (21), (23), and (24), we conclude (10). ■

Theorem 1 implies convergence property of the diameter of $\Omega^N(\epsilon,\delta)$. In fact, we see that

$$p_{rL}(\rho,\mu) \to 0 \qquad \text{as} \quad N \to \infty$$

and

$$p_{rU}(\rho,\mu) \to 0 \qquad \text{as} \quad N \to \infty.$$

That is, the probability with which the diameter of the membership set is greater than a given bound converges to zero when the number of samples N tends to infinity. It should be also noticed that the lower and upper bounds in Theorem 1 clarify that this convergence rate of the probability is $\mathcal{O}(N)$.

Note that Theorem 1 holds for any ρ and μ, which leads to the following corollary.

Corollary 1: Suppose that all assumptions of Theorem 1 are satisfied. Then, $\mathrm{dia}\,\Omega^N(\epsilon,\delta)$ converges to 0 in probability when $N \to \infty$.

That is, the membership set converges to the nominal parameter of the system in probability. We remark that this result includes the existing result [9] as a special case, where the regressor is assumed to be periodic.

We also remark that Theorem 1 shows a quantitative relationship between the diameter of the membership set and the number of samples. Let us consider the case that $p_s > 0$ is the probability that we want to ensure and $(4\sqrt{n_0}/\alpha)(\rho + \mu\overline{\beta})$ is the bound of the diameter of the membership set that we want to specify. Since the inequality

$$p_s \geq \mathcal{P}\left(1 - p_{eL}(\rho,\mu)\right)^r$$

can be solved with respect to r as

$$r \geq \frac{\ln\left(p_s/\mathcal{P}\right)}{\ln(1 - p_{eL}(\rho,\mu))},$$

we can know the necessary number of samples explicitly if the probability of $p_{vL}(\rho)$, $p_{\eta L}(\mu)$ are available.

IV. Numerical Example

Let us consider a system (1) whose nominal parameter is

$$\theta = \begin{bmatrix} 5 \\ -5 \end{bmatrix}.$$

We generated the regressor ϕ_i randomly according to uniform distribution, where the magnitude of the regressor ϕ_i was set as $\|\phi_i\|_2 = 1$, i.e., $\underline{\beta} = \overline{\beta} = 1$. For the generated sequences of ϕ_i, we confirmed that ϕ_i was PE with $\alpha = 1.06$, $\beta = 2.44$, and $n_0 = 7$. Then, $\mathcal{P} = 40$.

We also generated the disturbance v_i and the parameter uncertainty η_i randomly, where independent random sequences were used. Here we set $\epsilon = 1$ and $\delta = 1$, and v_i and η_i are generated randomly according to uniform distribution over $|v_i| \leq 1$ and $\|\eta_i\|_2 \leq 1$ respectively. In this case, the probabilities of the tightness $p_{vL}(\rho)$, $p_{vU}(\rho)$, $p_{\eta L}(\mu)$, and $p_{\eta U}(\mu)$ are calculated exactly as

$$\begin{aligned} p_{vL}(\rho) &= p_{vU}(\rho) = \frac{\rho}{2\epsilon} \\ p_{\eta L}(\mu) &= p_{\eta U}(\mu) \\ &= \frac{1}{\pi}\left\{\cos^{-1}\left(1-\frac{\mu}{\delta}\right) - \left(1-\frac{\mu}{\delta}\right)\sqrt{\frac{2\mu}{\delta} - \left(\frac{\mu}{\delta}\right)^2}\right\} \end{aligned}$$

where $p_{\eta L}(\mu)$ and $p_{\eta U}(\mu)$ are obtained by calculating the volume of the shaded potions in Fig. 1.

We set $\rho = 0.1$ and $\mu = 0.1$, which means that

$$\frac{4\sqrt{n_0}}{\alpha}(\rho + \mu\underline{\beta}) = \frac{4\sqrt{n_0}}{\alpha}(\rho + \mu\overline{\beta}) = 1.9922.$$

Then, for several numbers of repetitions r, we computed the theoretical probabilities $p_{rL}(\rho,\mu)$, $p_{rU}(\rho,\mu)$. We also estimated the probability p_{r_M} based on the randomly generated sequences, each of which was determined by the average of $1,000$ trials. Note that p_{r_M} was computed in detail for small r, i.e., at each 10 steps from $r = 1$ to 51 and at each 50 steps from $r = 101$ to 2001. The results are shown in Fig. 2. In this figure, a broken line represents $p_{rL}(\rho,\mu)$, a solid line represents $p_{rU}(\rho,\mu)$, and • stands for p_{r_M}, where these probabilities are drawn with respect to the number of samples.

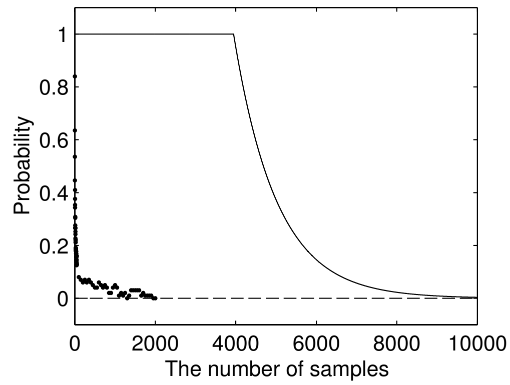

Fig. 2. Relation between the number of samples and the risk ($\rho = 0.1$, $\mu = 0.1$)

We also performed the same procedure for the case that $\rho = 0.005$ and $\mu = 0.005$, i.e.,

$$\frac{4\sqrt{n_0}}{\alpha}(\rho + \mu\underline{\beta}) = \frac{4\sqrt{n_0}}{\alpha}(\rho + \mu\overline{\beta}) = 0.0966.$$

The probabilities $p_{rL}(\rho,\mu)$, $p_{rU}(\rho,\mu)$, and p_{r_M} are shown in Fig. 3, where a broken line represents $p_{rL}(\rho,\mu)$ a solid

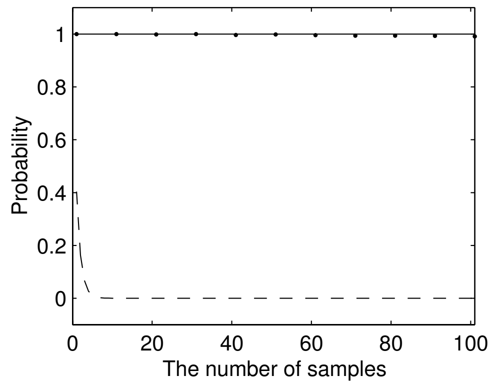

Fig. 3. Relation between the number of samples and the risk ($\rho = 0.005$, $\mu = 0.005$)

line represents $p_{rU}(\rho, \mu)$, and • denotes p_{r_M}. For this case, we used an equidistant step size for r regarding p_{r_M}, i.e., we compute p_{r_M} at each 10 steps from $r = 1$ to 101.

Fig. 2 and 3 show that $p_{rL}(\rho, \mu) \leq p_{rM} \leq p_{rU}(\rho, \mu)$ for any numbers of samples, which is consistent with Theorem 1. Furthermore, Fig. 2 shows that the probability that the diameter of the membership set is less than or equal to 1.9922 converges to zero as the number of samples tends to infinity, which is also a consequence of Theorem 1.

V. Concluding Remarks

In this paper, we have analyzed the size of the membership set probabilistically in the presence of disturbance and parameter uncertainty. In particular, we have estimated the diameter of the membership set for a finite number of samples, where we have assumed that the regressor is PE and the disturbance and the parameter uncertainty are tight, while we have not assumed that the regressor is periodic. We have also shown that, under the same conditions, the estimated diameter converges to zero as the number of samples tends to infinity. The result means that, even in the context of the membership set in the presence of both disturbance and parameter uncertainty, the PE property and the tightness are crucial in order to obtain a good estimate of the system parameter.

References

[1] G. Belforte, B. Bona and V. Cerone: Parameter Estimation Algorithms for a Set-Membership Description of Uncertainty; *Automatica*, Vol. 26, No. 5, pp. 887–898 (1990)

[2] M. Milanese and A. Vicino: Optimal Estimation Theory for Dynamic Systems with Set Membership Uncertainty — An Overview; *Automatica*, Vol. 27, No. 6, pp. 997–1009 (1991)

[3] Special Issue on Bounded-Error Estimation; *International Journal of Adaptive Control and Signal Processing*, Vol. 8, No. 1 (1994)

[4] Special Issue on Bounded-Error Estimation; *International Journal of Adaptive Control and Signal Processing*, Vol. 9, No. 1 (1995)

[5] E.-W. Bai, R. Tempo and H. Cho: Membership Set Estimators — Size, Optimal Input, Complexity and Relations with Least Squares; *IEEE Transactions on Circuits and Systems*, Part I, Vol. 42, No. 5, pp. 266–277 (1995)

[6] E.-W. Bai, H. Cho and R. Tempo: Convergence Properties of the Membership Set; *Automatica*, Vol. 34, No. 10, pp. 1245–1249 (1998)

[7] H. Akçay: The Size of the Membership Set in a probabilistic framework; *Automatica*, Vol. 42, No. 2, pp. 253–260 (2004)

[8] W. Kitamura, Y. Fujisaki and E.-W. Bai: Infinite and Finite Sample Properties of Set Membership Identification in a Stochastic Setting; *Proceedings of 41st IEEE Conference on Decision and Control*, Las Vegas, USA, pp. 269–274 (2002)

[9] W. Kitamura and Y. Fujisaki: Convergence Properties of the Membership Set in the Presence of Disturbance and Parameter Uncertainty; *Transaction of the Society of Instrument and Control Engineers*, Vol. 39, No. 4, pp. 382–387 (2003)

[10] S. Dasgupta and Y.-F. Huang: Asymptotically Convergent Modified Recursive Least Squares with Data Dependent Updating and Forgetting Factor for System with Bounded Noise; *IEEE Transaction on Information Theory*, Vol. 33, No. 3, pp. 383–392 (1987)